Robust Semidefinite Programming Problems with General Nonlinear Parameter Dependence: Application of the DC-Representations

Yasuki Oishi * Teodor Alamo **

* Department of Systems Design and Engineering, Nanzan University, Seireicho 27, Seto 489-0863, Japan (email: oishi@nanzan-u.ac.jp).
** Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Camino de los Descubrimientos s/n. 41092 Sevilla, Spain (email: alamo@cartuja.us.es).

Abstract: Two conservative approaches are proposed to a semidefinite programming problem nonlinearly dependent on uncertain parameters. These approaches are applicable to general nonlinear parameter dependence not necessarily polynomial or rational. They are based on a mild assumption that the parameter dependence is expressed as the difference of two convex functions. The first approach uses constant bounds on the parameter dependence. Optimization of the bounds is reduced to convex nonsmooth minimization in general. The second approach uses parameter-dependent bounds for a less conservative result. Optimization of the bounds is immediate when the gravity center is computable for the parameter set. Numerical examples are presented for illustration of the approaches.

Keywords: robust semidefinite programming, nonlinear parameter dependence, DC-representation, convexity, conjugate functions, linear matrix inequalities, conservatism.

1. INTRODUCTION

A robust semidefinite programming problem (robust SDP problem) is an important optimization problem, to which many control problems are reduced (Ben-Tal & Nemirovski, 2001; Scherer, 2006; Ben-Tal, El Ghaoui, & Nemirovski, 2009). There, a linear objective function is to be optimized under a linear matrix inequality (LMI) constraint dependent on uncertain parameters. When the parameter dependence is affine, this problem is exactly solvable by noticing the constraints corresponding to the vertices of the parameter set. Note however that the problem is NP-hard even in this simple case (Nemirovski, 1993). When the parameter dependence is not affine, a conservative approach is usually taken: the parameter-dependent constraint is replaced by its sufficient condition expressed by a finite number of parameter-independent LMI constraints. Considerable efforts have been paid in the case of polynomial or rational parameter dependence (e.g., Ben-Tal & Nemirovski, 1998; El Ghaoui, Oustry, & Lebret, 1998; Bliman, 2004; Chesi, Garulli, Tesi, & Vicino, 2005; Scherer, 2005; Scherer & Hol, 2006; Peaucelle & Sato, 2009; Oishi, 2009). There, conservatism can be reduced to any desired degree. On the other hand, the case of general nonlinear parameter dependence has not been considered so much except for a few works (Papachristodoulou & Prajna, 2005; Chesi & Hung, 2008; Oishi & Fujioka, 2010).

In this paper, we consider a robust SDP problem with general nonlinear parameter dependence and show the usefulness of the, so-called, DC-representations, which will be explained later. Specifically, we consider minimization of a linear objective function subject to an LMI constraint affinely dependent on q continuous nonlinear functions \( a_j(\theta) \), \( j = 1, 2, \ldots, q \), where \( \theta \) is a p-dimensional uncertain parameter. This constraint has to be satisfied for all \( \theta \) in the set \( \Theta \), which is a p-dimensional closed convex polytope (i.e., bounded polyhedron) having a nonempty interior. (See Section 2.1 for the explicit form of the problem.) One possible approach to this problem is to obtain upper and lower bounds, \( a_{\underline{p}} \leq a_j(\theta) \leq \overline{a}_j (\theta \in \Theta) \), for each \( j = 1, 2, \ldots, q \) and to replace \( a_j(\theta) \) in the original LMI constraint either by \( a_{\underline{p}} \) or \( \overline{a}_j \). The resulting \( 2^q \) LMI constraints constitute a sufficient condition for the original LMI constraint to be satisfied. Hence, we can minimize the objective function subject to these LMI constraints and regard the obtained solution as an approximate solution of the original problem. Now, we have two issues in this approach. First, the upper and lower bounds on \( a_j(\theta) \) can be difficult to find. Second, this approach can be considerably conservative because it uses constant bounds and neglects how \( a_j(\theta) \) depends on \( \theta \). (See Section 5 for an example.)

In this paper, we would like to address these two issues using the DC-representations. Here, the letters “DC” mean the difference of two convex functions. We assume that each of the nonlinear functions \( a_j(\theta) \) is expressed as the
difference of two convex functions. Then, we can compute upper and lower bounds on \( a_j(\theta) \) with this representation. Furthermore, we can obtain parameter-dependent bounds incorporating the parameter dependence of \( a_j(\theta) \). The constant bounds as well as the parameter-dependent bounds form parameterized families and they can be optimized in some specific sense. Between these two types of bounds, the parameter-dependent bounds look better because they can incorporate the parameter dependence of \( a_j(\theta) \) and they have computational advantage in their optimization. The notion of the DC-representation has been utilized in the field of optimization (Tuy, 1998; Horst & Thoai, 1999). Its application in control was considered by Tuan, Apkarian, Hosoe, and Tuy (2000), Tuan, Hosoe, and Tuy (2000), Bravo, Alamo, Fiacchini, and Camacho (2007), and Alamo, Bravo, Redondo, and Camacho (2008).

The following notation is used. The symbol \( \mathbb{R}^n \) stands for the set of \( n \)-dimensional real column vectors. The symbol \( \mathbf{T} \) expresses the transpose of a vector. The symbol \( I \) denotes the identity matrix of appropriate size. For a symmetric matrix \( A \), the inequality \( A \succeq O \) means that \( A \) is positive semidefinite. For a closed convex polytope \( \Theta \), the symbol \( \text{ver} \Theta \) indicates the set of the vertices of \( \Theta \). The number of the vertices is expressed by \( |\text{ver} \Theta| \).

2. PREPARATIONS

2.1 Problem to be considered

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

\[
\begin{align*}
P & : \quad \text{minimize} \quad c^T x \\
& \quad \text{subject to} \\
F_0(x) + \sum_{i=1}^{p} \theta_i F_i(x) + \sum_{j=1}^{q} a_j(\theta) F_{p+j}(x) & \geq O \quad (\forall \theta \in \Theta) (1) \end{align*}
\]

Here, \( x \in \mathbb{R}^n \) is a design variable; \( c \in \mathbb{R}^n \) is a given vector; \( F_i(x), i = 0, \ldots, p + q \), are affine functions of \( x \) whose values are \( m \times m \) symmetric matrices; \( \Theta = (\theta_1, \theta_2, \ldots, \theta_p)^\mathbf{T} \) is an uncertain parameter taking the value in the parameter set \( \Theta \); \( \Theta \) is a closed convex polytope in \( \mathbb{R}^p \) having a nonempty interior; \( a_j(\theta), j = 1, 2, \ldots, q \), are continuous nonlinear functions of \( \theta \in \Theta \). This problem is difficult to solve because the constraint (1) has to be satisfied for all \( \theta \) in \( \Theta \).

2.2 DC-representation

Let \( \Theta \) be a closed convex polytope in \( \mathbb{R}^p \) having a nonempty interior. A continuous function \( a(\theta) \) on \( \Theta \) is said to have a DC-representation if it is expressed as the difference of two convex functions, that is,

\[
a(\theta) = b(\theta) - c(\theta) \quad (\theta \in \Theta),
\]

where \( b(\theta) \) and \( c(\theta) \) are convex functions on \( \Theta \).

\textbf{Example 1.} Suppose that \( a(\theta) = \sin \theta \) and \( \Theta = [-1, 1] \). Its DC-representation is given by \( b(\theta) = e^{\theta}/2 \) and \( c(\theta) = e^{-\theta}/2 \).

Suppose that \( a(\theta) = \sin \theta \) and \( \Theta = [-2, 2] \). The second derivative of \( \sin \theta \) is larger than or equal to \( -1 \) on \( \Theta \), which implies the convexity of \( \sin \theta + \theta^2/2 \) on \( \Theta \). Hence, \( b(\theta) = \sin \theta + \theta^2/2 \) and \( c(\theta) = \theta^2/2 \) give the DC-representation.

Polynomials and rational functions have DC-representations if they are continuous on \( \Theta \). Any function continuous on \( \Theta \) can be approximated to any precision by a function possessing a DC-representation. Moreover, the second example above shows that an explicit DC-representation can be obtained for a given function if a lower bound is known on its second derivative (or the Hessian matrix in the case of \( p > 1 \)). Computation of such a bound is discussed by Adjiman and Floudas (1996). It is hence reasonably mild to assume the nonlinear functions \( a_j(\theta) \) in the problem \( P \) to have known DC-representations.

3. APPROACH WITH CONSTANT BOUNDS

We assume in the problem \( P \) that the nonlinear function \( a_j(\theta) \) has a known DC-representation

\[
a_j(\theta) = b_j(\theta) - c_j(\theta) \quad (\theta \in \Theta)
\]

for each \( j = 1, 2, \ldots, q \), where \( b_j(\theta) \) and \( c_j(\theta) \) are convex functions on \( \Theta \). We consider in this section how to obtain upper and lower bounds on \( a_j(\theta) \) through maximization of the concave functions \( \lambda^T \hat{\theta} - b_j(\hat{\theta}) \) and \( \mu^T \hat{\theta} - c_j(\hat{\theta}) \), though it requires a numerical method.
technique\textsuperscript{1} in general (e.g., Bertsekas, 1999). Then, we carry out minimization and maximization in (4) and (5), respectively, to obtain the bounds. This step requires only comparison of a finite number of values.

Proposition 2 allows us to construct a sufficient condition for the constraint (1) of the problem P. Let us choose the vectors $\lambda(j)$ and $\mu(j)$ arbitrarily from $\mathbb{R}^p$ for each $j = 1, 2, \ldots, q$ and consider the following condition on $x \in \mathbb{R}^n$:

$$F_0(x) + \sum_{i=1}^{p} \theta_i F_i(x) + \sum_{j=1}^{q} \alpha_j F_{p+j}(x) \geq 0 \quad (\forall \alpha_j \in \{\pi_{j,\lambda(j)}, \pi_{j,\mu(j)}\}, \; j = 1, 2, \ldots, q; \; \forall \theta \in \text{ver} \Theta). \quad (6)$$

Here, all the $2^n$ assignments are considered for $\alpha_j$, $j = 1, 2, \ldots, q$, and all the vertices of $\Theta$ are considered for $\theta$.

Proposition 3. Let $\lambda(j)$ and $\mu(j)$ be some vectors in $\mathbb{R}^p$ for each $j = 1, 2, \ldots, q$. Then, for a given $x \in \mathbb{R}^n$, the condition (1) is satisfied if the condition (6) is satisfied. \hfill \Box

Proof. Proposition 2 assures $\tilde{a}_j(\lambda(j)) \leq a_j(\theta) \leq \tilde{a}_j(\mu(j)) (\theta \in \theta)$ for each $j = 1, 2, \ldots, q$. Since the condition (6) is convex in $\theta$ and $\alpha_j$, it implies the condition (1). \hfill \Box

We now have the following approach to the problem $P$. We replace the constraint (1) with its sufficient condition (6) to have a standard SDP problem solvable with the interior-point method. The optimal solution of this modified problem can be used as an approximate solution of the original problem $P$. The condition (6) consists of $2^q |\text{ver} \Theta|$ LMIs with $|\text{ver} \Theta|$ meaning the number of vertices of $\Theta$. Although this number becomes large for a large $q$, it may be reasonable for the NP-hard problem $P$.

The choice of the vectors $\lambda(j)$ and $\mu(j)$ is arbitrary. It would be nice, however, if we can choose optimal vectors. Specifically, the desired $\lambda$ and $\mu$ would be such that the lower bound $\tilde{a}_j(\lambda)$ and minimize the upper bound $\tilde{a}_j(\mu)$, respectively. The result of Carrizosa (2001) is useful for this purpose. (A related result can be found in Bravo, et al. (2007).) He showed that these maximization and minimization can be reduced to minimization of some convex functions. This minimization is however not easy in general because the minimized functions do not have analytic expressions and may not be differentiable at some points. On the other hand, in the special case that $\Theta$ is a simplex, we can obtain the optimal $\lambda$ and $\mu$ by solving linear equations. Indeed, $\lambda$ is optimal when the value of $\lambda^T \theta - c_j(\theta)$ is identical for all $\theta \in \text{ver} \Theta$. A similar property holds on the optimal $\mu$ but with $\mu^T \theta - b_j(\theta)$ this time.

Example 4. We continue the second example of Example 1, where $a(\theta) = \sin \theta$ and $\Theta = \{-2, 2\}$. The functions $b(\theta) = \sin \theta + \theta^2/2$ and $c(\theta) = \theta^2/2$ give the DC-representation.

Let us compute the optimal lower bound $\tilde{a}_\lambda$. Since the parameter set $\Theta$ is a simplex, the optimal $\lambda$ can be obtained by solving $\lambda^T (-2) - c(-2) = 2 \lambda^T - c(2), \text{ i.e., } \lambda = 0$. Since $b_j(\theta) = -\max_{\theta \in \Theta}[-\sin \theta - \theta^2/2]$, we carry out the

\textsuperscript{1}Maximization is not necessary to have $b_j(\theta)$ if $\lambda$ is known to be a subgradient of $b_j(\theta)$ at some point $\hat{\theta} \in \Theta$. Indeed, in this case, we have $b_j(\lambda) = \lambda^T (\theta - \lambda) + b_j(\theta)$. The same comment applies to $c_{j,\mu}(\theta)$, too. This property will be used in Proposition 8.

numerical maximization to have the explicit expression $b_j(\theta) \approx -0.400$. To compute $\tilde{a}_\lambda = \min_{\theta \in \Theta} [b_j(\theta) - c(\theta)]$, we evaluate $b_j(\theta) - c(\theta)$ at $\theta = -2$ and $\theta = 2$. The two values are equal and give the desired optimal lower bound $\tilde{a}_\lambda \approx -2.400$. Computation of the optimal upper bound $\tilde{a}_\mu$ is similar. The result is $\bar{a}_\mu \approx 2.103$ for $\mu \approx 0.455$.

4. APPROACH WITH PARAMETER-DEPENDENT BOUNDS

The bounds in the previous section are constant in $\theta$ and neglect how the functions $a_j(\theta)$ depend on $\theta$. They can give a considerably conservative result when incorporation of this dependence is important for solving $P$. In this section, we introduce parameter-dependent bounds using the DC-representations of $a_j(\theta)$. These bounds give another sufficient condition for the constraint (1) of the problem $P$.

The considered bounds are the following.

Proposition 5. For each $j = 1, 2, \ldots, q$, define the functions on $\Theta$ as

$$\Lambda_{j,\lambda}(\theta) := b_j(\lambda) - c_j(\theta), \quad \Gamma_{j,\mu}(\theta) := b_j(\theta) - c_j(\mu)(\theta) \quad (7)$$

with $\lambda$ and $\mu$ being any vectors in $\mathbb{R}^p$. Then, $\Lambda_{j,\lambda}(\theta)$ is concave and $\Gamma_{j,\mu}(\theta)$ is convex. Moreover, they satisfy

$$\Lambda_{j,\lambda}(\theta) \leq a_j(\theta) \leq \Gamma_{j,\mu}(\theta) \quad (\theta \in \Theta). \quad \hfill \Box$$

Proof. The concavity of $\Lambda_{j,\lambda}(\theta)$ follows from the affinity of $b_j(\lambda)$ and the convexity of $c_j(\theta)$. The convexity of $\Gamma_{j,\mu}(\theta)$ follows similarly. The desired inequality is a consequence of $b_j(\lambda) \leq b_j(\theta)$ and $c_j(\mu)(\theta) \leq c_j(\theta)$. \hfill \Box

With the parameter-dependent bounds $\Lambda_{j,\lambda}(\theta)$ and $\Gamma_{j,\mu}(\theta)$, we have the following sufficient condition for (1). Let $\lambda(j)$
and $\mu(j)$ be some vectors in $\mathbb{R}^p$ for each $j = 1, 2, \ldots, q$. Then, for a given $x \in \mathbb{R}^n$, we consider the condition

$$F_0(x) + \sum_{j=1}^{q} \rho_j F_j(x) + \sum_{j=1}^{q} \rho_j F_{p+j}(x) \geq 0$$

(8)

Here, $\rho_j(\theta)$ is a function chosen from the two candidate functions $\xi_{j,\lambda}(\theta)$ and $\tau_{j,\mu}(\theta)$ for each $j$. All the $2^q$ choices are considered for $\rho_j(\theta)$, $j = 1, 2, \ldots, q$, and all the vertices in $\Theta$ are considered for $\theta$.

Proposition 6. For a given $x \in \mathbb{R}^n$, the condition (1) is satisfied if the condition (8) is satisfied.

Proof. Let $v$ be any nonzero vector in $\mathbb{R}^m$. The proof is complete if we can show

$$v^TF_0(x)v + \sum_{j=1}^{q} \rho_j v^TF_j(x)v + \sum_{j=1}^{q} \rho_j v^TF_{p+j}(x)v \geq 0$$

for any $\theta \in \Theta$, which is the inequality (1) multiplied by $v^T$ and $v$. Toward this end, multiply $v^T$ and $v$ to the assumption (8). Defining

$$f_j(\theta) := \min\{F_{j,\lambda}(\theta)v^TF_{p+j}(x)v, \tau_{j,\mu}(\theta)v^TF_{p+j}(x)v\}$$

for $j = 1, 2, \ldots, q$, we have

$$v^TF_0(x)v + \sum_{j=1}^{q} \rho_j v^TF_j(x)v + \sum_{j=1}^{q} f_j(\theta) \geq 0 \quad (\theta \in \Theta).$$

(9)

Pick up any $j$ and suppose $v^TF_{p+j}(x)v \geq 0$. In this case, the function $f_j(\theta)$ is equal to $F_{j,\lambda}(\theta)v^TF_{p+j}(x)v$, which is concave in $\theta$ and less than or equal to $a_j(\theta)v^TF_{p+j}(x)v$ by Proposition 5. Repeating a similar discussion in the case of $v^TF_{p+j}(x)v < 0$, we see that, in any case, $f_j(\theta)$ is concave and satisfies

$$a_j(\theta)v^TF_{p+j}(x)v \geq f_j(\theta) \quad (\theta \in \Theta)$$

(10)

for each $j$. The concavity of $f_j(\theta)$ implies that the inequality (10) holds for all $\theta \in \Theta$ and the inequality (11) implies that the desired inequality (9) holds for all $\theta \in \Theta$. The proof is complete.

In the present approach, the utilized sufficient condition (8) consists of $2^q$ ver $\Theta$ LMs. This number is the same as the approach in the previous section. On the other hand, the present approach is less conservative than the previous one when the two approaches use the same $\lambda(j)$ and $\mu(j)$ for each $j$. This is shown by the next result.

Proposition 7. Let $\lambda(j)$ and $\mu(j)$ be some vectors in $\mathbb{R}^p$ for each $j = 1, 2, \ldots, q$. Then, the condition (8) is satisfied if the condition (6) is satisfied.

Proof. The claim follows if we have

$$\underline{\xi}_{j,\lambda}(\theta) \leq \xi_{j,\lambda}(\theta) \quad (\theta \in \Theta),$$

(12)

$$\tau_{j,\mu}(\theta) \leq \tau_{j,\mu}(\theta) \quad (\theta \in \Theta)$$

(13)

for each $j = 1, 2, \ldots, q$. Recall that

$$\underline{\xi}_{j,\lambda}(\theta) = \min_{\theta \in \Theta} [b_j(\theta) - c_j(\theta)],$$

$$\xi_{j,\lambda}(\theta) = b_j(\theta) - c_j(\theta).$$

Hence, the inequality (12) follows immediately. The inequality (13) is similarly derived.

We next consider how the vectors $\lambda(j)$ and $\mu(j)$ should be chosen for the parameter-dependent bounds. One possibility is to use the $\lambda$ and $\mu$ that optimize the constant bounds $\underline{\xi}_{j,\lambda}$ and $\tau_{j,\mu}$. Proposition 7 then guarantees a less conservative result for the parameter-dependent bounds. A problem here is that computation of such $\lambda$ and $\mu$ is not easy except for the special case that $\Theta$ is a simplex. Even if such $\lambda$ and $\mu$ are computed, we still need concave maximization to have the explicit expressions of $b_{j,\lambda}(\theta)$ and $c_{j,\mu}(\theta)$. It is also questionable if optimization of the constant bounds really leads to optimization of the parameter-dependent bounds.

We here consider another criterion for the choice of $\lambda$ and $\mu$. It does not guarantee a less conservative result for the present approach in the sense of Proposition 7. It is however reasonable to use and convenient to compute the corresponding bounds. In particular, we employ the following quantity as a measure of conservatism:

$$V_j(\lambda, \mu) := \int_{\Theta} \tau_{j,\mu}(\theta) - \xi_{j,\lambda}(\theta) \, d\theta$$

(14)

and choose $\lambda$ and $\mu$ so as to minimize this measure for each $j$. The measure $V_j(\lambda, \mu)$ stands for the volume of the region sandwiched by the parameter-dependent bounds $\xi_{j,\lambda}(\theta)$ and $\tau_{j,\mu}(\theta)$. Since $\xi_{j,\lambda}(\theta)$ and $\tau_{j,\mu}(\theta)$ are used in place of $a_j(\theta)$, the smaller $V_j(\lambda, \mu)$ means the less conservative approximation in the present approach. It is thus reasonable to choose $\lambda$ and $\mu$ so as to minimize $V_j(\lambda, \mu)$. Such $\lambda$ and $\mu$ have the following simple characterization.

Proposition 8. For each $j = 1, 2, \ldots, q$, the measure of conservatism, $V_j(\lambda, \mu)$ in (14), is minimized if and only if $\lambda$ and $\mu$ are subgradients of $b_{j,\lambda}(\theta)$ and $c_{j,\mu}(\theta)$, respectively, at the gravity center of $\Theta$, that is, $\theta^c := \int_{\Theta} \theta \, d\theta / \int_{\Theta} d\theta$. Moreover, for these $\lambda$ and $\mu$, the following relationships hold:

$$b_{j,\lambda}(\theta) = \lambda^T(\theta - \theta^c) + b_j(\theta^c), \quad c_{j,\mu}(\theta) = \mu^T(\theta - \theta^c) + c_j(\theta^c).$$

(15)

Proof. As noticed in the previous section, the functions $b_{j,\lambda}(\theta)$ and $c_{j,\mu}(\theta)$, which are used in the definitions of $\xi_{j,\lambda}(\theta)$ and $\tau_{j,\mu}(\theta)$, are related to the notion of conjugate functions. Consider in general a convex function $f(\theta)$ on $\Theta$. Then, its conjugate $f^*(\lambda)$ is defined as

$$f^*(\lambda) := \max_{\theta \in \Theta} [\lambda^T \theta - f(\theta)]$$

for $\lambda \in \mathbb{R}^p$ (e.g., Hiriart-Urruty & Lemaréchal, 1993; Bertsekas, Nedić, & Ozdaglar, 2003). The conjugate function $f^*(\lambda)$ is convex because it is the maximum of the family of affine functions. With this notion, we can write

$$\xi_{j,\lambda}(\theta) = \lambda^T \theta - b_j^*(\lambda) - c_j(\theta), \quad \tau_{j,\mu}(\theta) = b_j(\theta) - \mu^T \theta + c_j^*(\mu)$$

for any $j, \lambda$, and $\mu$. The measure of conservatism can be rewritten as

$$V_j(\lambda, \mu) = \int_{\Theta} [b_j(\theta) - \mu^T \theta + c_j^*(\mu) - \lambda^T \theta + b_j^*(\lambda) + c_j(\theta)] \, d\theta$$

$$= \int_{\Theta} b_j(\theta) + c_j(\theta) \, d\theta - \lambda^T \int_{\Theta} \theta \, d\theta + b_j^*(\lambda) \int_{\Theta} \, d\theta$$

$$- \mu^T \int_{\Theta} \theta \, d\theta + c_j^*(\mu)$$

(16)

for each $j$. This is minimized when both $-\lambda^T \theta^c + b_j^*(\lambda)$ and $-\mu^T \theta^c + c_j^*(\mu)$ are minimized. Let us consider when.
the first function, \(-\lambda^T \theta c + b^*(\lambda)\), is minimized. Due to the convexity of this function, it is minimized if and only if \(\lambda\) is a subgradient of \(b^*(\lambda)\). This occurs if and only if \(\lambda\) is a subgradient of \(b_j(\theta)\) by a property of conjugate functions (Corollary 1.1 of Hiriart-Urruty and Lemaréchal (1993)). Similar equivalence holds on the second function.

Suppose that \(\lambda\) is a subgradient of \(b_j(\theta)\) at \(\theta = \theta^e\). Then, the graph of \(\lambda^T (\theta - \theta^e) + b^*(\theta^e)\) gives a hyperplane supporting the graph of \(b_j(\theta)\) from below. This means that this function is identical with \(b_j,\lambda(\theta)\). Similar discussion is possible on \(c_{j,\mu}(\theta)\).

Example 9. This is a continuation of Example 4, where \(a(\theta) = \sin \theta\) and \(\Theta = [-\pi, 2\pi]\). With the DC-representation \(b(\theta) = \sin \theta + \theta^2/2\) and \(c(\theta) = \theta^2/2\), we compute the optimal parameter-dependent bounds \(\Gamma_{\lambda}(\theta)\) and \(\pi_{\mu}(\theta)\).

The gravity center of \(\Theta = [-\pi, 2\pi]\) is \(\theta^e = 0\). The function \(b(\theta) = \sin \theta + \theta^2/2\) is differentiable there and has the gradient 1. Hence, by Proposition 8, \(\lambda = 1\) is optimal and the corresponding \(b_j(\theta) = \lambda (\theta - \theta^e) + b^*(\theta^e)\). We thus obtain the optimal lower bound \(\Gamma_{\lambda}(\theta) = \theta - \theta^e/2\). On the other hand, the function \(c(\theta) = \theta^2/2\) has the gradient 0 at \(\theta = 0\). Hence, \(\mu = 0\) is optimal and the corresponding \(c_{j,\mu}(\theta) = \mu (\theta - \theta^e) + c(\theta^e) = 0\). The obtained optimal upper bound is \(\pi_{\mu}(\theta) = \sin \theta + \theta^2/2\). These bounds are shown in Figure 1.

In a large part of \(\Theta\), the parameter-dependent bounds are closer to the original function \(a(\theta)\) than the constant bounds \(\overline{a}_\lambda\) and \(\underline{pi}_\mu\) considered in Example 4. It is seen however that \(\overline{a}_\lambda\) is closer to \(a(\theta)\) than \(\underline{pi}_\mu(\theta)\) around \(\theta = -\pi\) and that \(\underline{pi}_\mu\) is closer than \(\Gamma_{\lambda}(\theta)\) around \(\theta = 2\). This is not surprising. Indeed, the uniform superiority of the parameter-dependent bounds is not guaranteed in the sense of Proposition 7 because different values are used for \(\lambda\) and \(\mu\) between the two types of bounds.

When the gravity center \(\theta^e\) is computable for \(\Theta\), the following approach is possible to the problem \(P\). For each \(j = 1, 2, \ldots, q\), we compute subgradients of \(b_j(\theta)\) and \(c_{j,\mu}(\theta)\) at \(\theta = \theta^e\) and choose them for \(\lambda(j)\) and \(\mu(j)\), respectively. Proposition 8 then guarantees their optimality with respect to \(V_j(\lambda, \mu)\) and provides the explicit expressions of \(b_j,\lambda(j)(\theta)\) and \(c_{j,\mu(j)}(\theta)\). We obtain the optimal bounds \(\Gamma_{\lambda,j}(\theta)\) and \(\pi_{\mu,j}(\theta)\) with (7) and substitute them into the condition (8). Using this condition in place of (1), we obtain an approximate problem of \(P\), which is solvable with the interior-point method. If we want to reduce the conservatism of the present approach, we can divide the parameter set \(\Theta\) as in Section 3.

Computation of the gravity center \(\theta^e\) is immediate when \(\Theta\) is a simplex or a parallelepiped. In many control problems, \(\Theta\) is a hyperrectangle, which is a special parallelepiped. On the other hand, the exact computation of \(\theta^e\) is difficult when \(\Theta\) is a general polytope (Rademacher, 2007). A solution would be approximate computation of \(\theta^e\) with a randomized technique (e.g., Bertsimas & Vempala, 2004).

5. EXAMPLES

We applied the proposed two approaches to two example problems. The proposed approaches worked well for both problems. In particular, they gave approximate solutions of the problems, whose precision became better as the division of the parameter region became finer.

Example 10. Set \(a(\theta) = \sin \theta\) and \(\Theta = [-\pi, 2\pi]\). Let \(F\) be a \(3 \times 3\) symmetric matrix whose elements are randomly generated according to the standard normal distribution. We consider the problem:

\[
\minimize z \quad \text{subject to} \quad zI - \theta \text{diag}(x_1, x_2, x_3) + a(\Theta)F \succeq 0 \quad (\forall \theta \in \Theta),
\]

\[
zI + \theta \text{diag}(x_1, x_2, x_3) - a(\Theta)F \succeq 0 \quad (\forall \theta \in \Theta),
\]

with the real scalar design variables \(z, x_1, x_2,\) and \(x_3\). Here, \(\text{diag}(x_1, x_2, x_3)\) stands for the \(3 \times 3\) diagonal matrix with the diagonal elements \(x_1, x_2,\) and \(x_3\). In this problem, uniform approximation of \(a(\theta)F\) is considered with \(\theta \text{diag}(x_1, x_2, x_3)\). Hence, it is important to take account of the parameter dependence of \(a(\theta)\).

For 100 randomly generated \(F\), we applied our two approaches and obtained the approximate solutions. Let \(z_{\text{cons}}\) stand for the minimum given by the approach with the constant bounds; let \(z_{\text{dep}}\) stand for the minimum given by the approach with the parameter-dependent bounds. These are upper bounds on the true minimum of the problem. For reference, we also computed its lower bound \(z_{\text{zamp}}\) by minimizing \(z\) subject to the inequalities (15) and (16) but only for the 101 values of \(\theta\) equally spaced in \(\Theta\). The normalized approximation error for the approach with the constant bounds, i.e., \((z_{\text{cons}} - z_{\text{zamp}}))/z_{\text{zamp}}\), had the mean 2.66 and the standard deviation 1.14. On the other hand, the normalized approximation error for the approach with the parameter-dependent bounds, i.e., \((z_{\text{dep}} - z_{\text{zamp}}))/z_{\text{zamp}}\), had the mean 3.33 and the standard deviation 0.98. Although the result of the parameter-dependent approach was worse, the situation changed when \(\Theta\) was divided into \([-2, 0]\) and \([0, 2]\). Namely, the mean of the normalized error became 0.71 for the approach with the constant bounds and 0.35 for the approach with the parameter-dependent bounds.

Example 11. We apply our approaches to design of a sampled-data control system with an uncertain sampling interval. Consider a linear continuous-time plant described by \(\dot{x}(t) = Ax(t) + Bu(t)\) with \(A = \begin{pmatrix} 0 & 1 \\ 0 & -0.1 \end{pmatrix}\), \(B = \begin{pmatrix} 0 \\ 0.1 \end{pmatrix}\).

We would like to design a matrix \(K\) so that the sampled-data state-feedback control

\[
\dot{u}(t) = K x(t_k) \quad (t_k \leq t < t_{k+1}; \ k = 0, 1, 2, \ldots)
\]

stabilizes the plant, where \(t_0, t_1, t_2, \ldots\) are sampling instants that satisfy \(0 = t_0 < t_1 < t_2 < \cdots\) and \(\lim_{k \to \infty} t_k = \infty\). We assume here that the sampling interval \(t_{k+1} - t_k\) is uncertain and can vary with \(k\), motivated by the need in the networked and embedded control. In particular, we consider the (rather extreme) situation that \(t_{k+1} - t_k\) can take any value in \((0, 200)\) and look for a stabilizing \(K\) even in this situation.

This type of control problems has been considered by several authors (e.g., Fridman, Seuret, & Richard, 2004; Hetel, Daafouz, & Iung, 2006; Mirkin, 2007; Naghshtabrizi, Hespanha, & Teel, 2008). We here follow the approach
of Oishi and Fujioka (2010) and reduce the problem to a robust SDP problem:

\[
\begin{align*}
\text{minimize} & \quad z \\
\text{subject to} & \quad Q - I \succeq O,
\end{align*}
\]

(the formula (17) at the bottom of this page) with \( \Theta := [0, 200] \) and

\[
E(\theta) := \frac{1}{\theta} \int_0^\theta e^{At} dt = \begin{pmatrix}
\frac{1}{\theta} & -10 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & -e^{0.10 - 0.10}\theta
\end{pmatrix}
\begin{pmatrix}
10 \\
0 1
\end{pmatrix}.
\]

We define \( E(\theta) = I \) at \( \theta = 0 \). In this problem, the design variables are a scalar \( z \), a 2 \times 2 symmetric matrix \( Q \), and a 1 \times 2 matrix \( L \). If a feasible solution \( (z, Q, L) \) such that \( z < 0 \) is found for this problem, a stabilizing state-feedback gain \( K \) is obtained by \( LQ^{-1} \).

Let us set \( a_1(\theta) = \sqrt{\theta} \) and \( a_2(\theta) = (e^{-0.1\theta} - 1)/(-0.1\theta) \).

We define \( a_2(\theta) = 1 \) at \( \theta = 0 \). We first notice that the constraint of the problem includes the product \( a_1(\theta)a_2(\theta) \).

To remove this product, we apply the technique of Oishi (2009) and Oishi and Isaka (2009). Then, we can obtain a sufficient condition, which is affine in \( a_1(\theta) \) and \( a_2(\theta) \). Replacing the constraint by this sufficient condition, we can apply our approaches. We choose \( b_1(\theta) \) and \( c_2(\theta) \) to be zero noting the concavity of \( a_1(\theta) \) and the convexity of \( a_2(\theta) \), respectively.

The obtained minimum was \( 1.75 \times 10^{-5} \) in the approach with the constant bounds and \( 3.87 \times 10^{-4} \) in the approach with the parameter-dependent bounds. This means that both approaches failed to give a stabilizing state-feedback gain. The situation changed when \( \Theta \) was divided into two subintervals \([0, 100]\) and \([100, 200]\). Indeed, with both approaches, a feasible solution with a negative \( z \) was found, which gave a stabilizing state-feedback gain.

Oishi and Fujioka (2010) gave an approach specialized for analysis and design of this type of sampled-data control systems. The performance of our present approaches were comparable to this specialized approach. Indeed, the specialized approach failed to give a state-feedback gain when \( \Theta \) was not divided. It failed again with the division consisting of \([0, 100]\) and \([100, 200]\). After some trials and errors, it succeeded with the division consisting of \([0, 25]\) and \([25, 200]\).

6. CONCLUSION

Two conservative approaches are proposed for robust SDP problems with general nonlinear parameter dependence. The first approach is based on the computation of constant bounds for each nonlinear function. It is conceptually simple but has a computational issue in the optimization of the bounds. The second approach is based on the computation of parameter-dependent bounds. It has a potential to give a less conservative result by exploiting the parameter dependence of the problem. It also has a computational advantage when the gravity center is computable for the parameter set. Both approaches rely on the assumption that the nonlinear parameter dependence is expressed as the difference of two convex functions. Since this assumption is mild enough, the proposed approaches can be applied to a large class of problems.

\[
\begin{pmatrix}
-E(\theta)(AQ + BL) - (AQ + BL)^T E(\theta)^T + z I \\
\sqrt{\theta}(AQ + BL)^T E(\theta)^T
\end{pmatrix} \succeq Q \quad (\theta \in \Theta)
\]

(17)

ACKNOWLEDGMENTS

The authors are thankful to Fabrizio Dabbene, Yoshio Ebihara, Dimitri Peaucelle, and Roberto Tempo for their valuable comments.

REFERENCES


