Randomized algorithms to solve parameter-dependent linear matrix inequalities*

Yasuaki Oishi† and Hidenori Kimura‡

Abstract

The randomized algorithm of Polyak and Tempo (2000), which consists of random sampling and subgradient descent, is generalized in order to solve parameter-dependent linear matrix inequalities and its computational complexity is analyzed. This paper first examines an algorithm obtained by direct generalization of Polyak and Tempo's and shows that its expected time to achieve a solution is infinite. Then this paper improves this algorithm so that its expected achievement time becomes finite. An explicit upper bound of the expected achievement time is given in a special case. A numerical example is provided.

Keywords: randomized algorithms, parameter-dependent linear matrix inequalities, computational complexity, curse of dimensionality, linear parameter-varying systems.

1. Introduction

Parameter-dependent linear matrix inequalities (LMIs) often appear when we consider analysis and synthesis of linear parameter-varying (LPV) systems, which are useful to model time-varying or nonlinear systems. Parameter-dependent LMIs are discussed in this context by the authors such as Becker, Packard, Philbrick, and Balas (1993), Becker and Packard (1994), and Apkarian, Gahinet, and Becker (1995), and their advanced use taking account of the rate of parameter change is considered by the authors including Watanabe, Uchida, Fujita, and Shimemura (1994), Gahinet, Apkarian, and Chilali (1996), Feron, Apkarian, and Gahinet (1996), Wu, Yang, Packard, and Becker (1996), Yu and Sideris (1997), and Apkarian and Adams (1998).

Since a parameter-dependent LMI is equivalent to a set of infinitely many LMIs, it is difficult to solve in general. One conventional approach is to make its parameter dependence affine by approximation (Ohara, Ide, Yamaguchi, & Ohno, 1999) or by introduction of new parameters (Masubuchi & Shimemura, 1999). However, this approach often causes conservatism or increase of the problem size. Another approach is to grid the parameter set and to consider the LMI only at the grid points (Wu, Yang, Packard, & Becker, 1996; Apkarian & Adams, 1998). In this approach, there is a concern that the LMI may not be satisfied between the grid points. Moreover, it has a problem called the "curse of

*Submitted to *Automatica* as a brief paper on July 17, 2001

†Corresponding author. Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan; phone: +81-3-5841-6906; fax: +81-3-5841-6886; email: oishi@simplex.t.u-tokyo.ac.jp

‡Department of Complexity Science and Engineering, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan; email: kimura@crux.t.u-tokyo.ac.jp

dimensionality," that is, the number of the grid points increases exponentially fast as the parameter dimension grows.

We try in this paper a probabilistic approach to parameter-dependent LMIs. The algorithms to be proposed are obtained by generalizing a recent work of Polyak and Tempo (2000), where they considered design of a robust linear quadratic regulator. Since our approach does not require the number of LMIs to be finite, it does not suffer from the mentioned problems of the conventional approaches. Although there are many works on a probabilistic approach or randomized algorithms (for example, Ray & Stengel, 1993; Khargonekar & Tikku, 1996; Barmish & Lagoa, 1997; Tempo, Bai, & Dabbene, 1997; Vidyasagar, 1999; Chen & Zhou, 2000), most of them utilize random sampling only. The proposed algorithms are more systematic than the existing ones because they use a combination of random sampling and subgradient descent.

By utilizing the result of Polyak and Tempo (2000), it is possible to show that the proposed algorithms arrive at a solution after a finite time with a probability one. However, it is not obvious how long this time is. This paper spends its main efforts in evaluating the expected time to achieve a solution. Particularly, it first shows that the expected achievement time is *infinite* when we use the algorithm obtained by direct generalization of Polyak and Tempo's one. This implies that this basic algorithm requires long computational time and is not practical. This paper then shows that one can improve this algorithm so that its expected achievement time is *finite*. An explicit upper bound of this time is given in a special case. Furthermore, it is shown that the improved algorithm overcomes the curse of dimensionality in the case that the LMI has to be satisfied not for all the parameters but for *almost* all of them. The contents of this paper are to be presented in the conference (Oishi & Kimura, 2001).

After finishing the main body of the present research, the author is notified that Fujisaki, Dabbene, and Tempo (2001) independently consider generalization of the algorithm of Polyak and Tempo. However, they concentrate on a special type of parameter-dependent LMIs and do not perform so detailed analysis of computational complexity as is found in the present paper.

The following notation is used. Let $\mathbb{R}^n$ mean the n-dimensional Euclidean space. The symbol $\|\cdot\|$ stands for the 2-norm of a vector and the symbol $\|\cdot\|_{\mathrm{F}}$ for the Frobenius norm of a matrix. Define $\overline{\lambda}(A)$ as the maximum (*i.e.*, the most positive) eigenvalue of a symmetric matrix A. For two symmetric matrices A and B, the inequalities $A < B$ and $A \leq B$ imply that $B - A$ is positive definite and is positive semidefinite, respectively. For a symmetric matrix A, the symbol A^+ means the projection onto the cone of positive semidefinite matrices. That is, with an orthogonal matrix U that makes $U^{\mathrm{T}}AU = \varXi = (\xi_{ij})$ diagonal, the matrix A^+ is defined as $U\varXi^+U^{\mathrm{T}}$, where the (i,j)-element of $\varXi^+$ is ξ_{ij} if $\xi_{ij} > 0$ and is zero otherwise.

2. Basic algorithm

A problem to be solved is presented first.

Problem. The parameter set P is the n-dimensional hypercube $[-1,1]^n$ and the variable domain Q is a bounded closed convex subset of $\mathbb{R}^d$ having its interior. Moreover, $V(\boldsymbol{q},\boldsymbol{p})$ is a symmetric-matrix-valued function of $\boldsymbol{q} \in Q$ and $\boldsymbol{p} \in P$, which is affine with respect to $\boldsymbol{q}$ and is continuously differentiable with respect to $\boldsymbol{p}$. Find a $\boldsymbol{q}$ satisfying $V(\boldsymbol{q},\boldsymbol{p}) \le O$ for any $\boldsymbol{p} \in P$. □

One can formulate analysis and synthesis of LPV systems into this form possibly using information on the rate of parameter change. Details are found in the references cited in the previous section. See Section 5 for an example.

Remark. Although the parameter set P is a general hyper rectangle in many LPV systems, it can be transformed to the above hypercube $[-1,1]^n$ without loss of generality. It is not practically difficult to limit the variable domain to a bounded Q by performing some preliminary computation. One may notice that a strict inequality $V(\boldsymbol{q},\boldsymbol{p}) < O$ is considered in many applications instead of $V(\boldsymbol{q},\boldsymbol{p}) \le O$ above. However, this difference is not essential in numerical computation and the improved algorithm to be introduced gives a solution $\boldsymbol{q}$ that satisfies the strict inequality. □

It is assumed throughout this paper that there exists a $\boldsymbol{q}$ satisfying the strict inequality $V(\boldsymbol{q},\boldsymbol{p}) < O$ for any $\boldsymbol{p} \in P$. Let us call the region $\{\boldsymbol{q} \in Q : V(\boldsymbol{q},\boldsymbol{p}) \le O$ for any $\boldsymbol{p} \in P\}$ the *solution region*. The solution region contains by assumption a d-dimensional closed ball whose center $\boldsymbol{q}^*$ attains $V(\boldsymbol{q}^*,\boldsymbol{p}) < O$ for any $\boldsymbol{p} \in P$. Fix one such ball and write its radius as r^*.

We now present an algorithm to solve the problem above. It is a generalization of the algorithm of Polyak and Tempo (2000), which is for design of a robust quadratic regulator. In order to state the algorithm, some preparation is needed. Choose an initial point $\boldsymbol{q}^0$ in the variable domain Q. We introduce into the parameter set P a probability density function possessing a finite upper bound $\overline{\mu}_P$ and a positive lower bound $\underline{\mu}_P$. Note that the inequality $V(\boldsymbol{q},\boldsymbol{p}) \not\le O$ is equivalent to $\|V(\boldsymbol{q},\boldsymbol{p})^+\|_{\mathrm{F}} > 0$. When $\|V(\boldsymbol{q},\boldsymbol{p})^+\|_{\mathrm{F}} > 0$ holds, we define $\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q},\boldsymbol{p})^+\|_{\mathrm{F}}$ to be the vector whose i-th element is

$$\frac{1}{\|V(\boldsymbol{q},\boldsymbol{p})^+\|_{\mathrm{F}}}\ \mathrm{tr}\left[V(\boldsymbol{q},\boldsymbol{p})^+\,\frac{\partial V(\boldsymbol{q},\boldsymbol{p})}{\partial q_i}\right],$$

where $\boldsymbol{q} = [q_1 \quad \ldots \quad q_d]^{\mathrm{T}}$. It is possible to show that $\|V(\boldsymbol{q},\boldsymbol{p})^+\|_{\mathrm{F}}$ is a convex function of $\boldsymbol{q}$ and $\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q},\boldsymbol{p})^+\|_{\mathrm{F}}$ is its subgradient. Furthermore, let $\mathit{\Pi}$ be the projection onto the variable domain Q, that is, $\mathit{\Pi}\boldsymbol{q} := \arg\min_{\widetilde{\boldsymbol{q}}\in Q}\|\widetilde{\boldsymbol{q}} - \boldsymbol{q}\|$ for $\boldsymbol{q} \in \mathbb{R}^d$. Finally, define $r^0, r^1, \ldots$ to be a sequence of positive numbers that monotonically decreases to zero and satisfies $\sum_{\ell=0}^{\infty}(r^\ell)^2 = \infty$.

Algorithm.

0. $k := 0$. $\ell := 0$.
1. Randomly sample $\boldsymbol{p}^k$ from P according to the probability density function on P.
2. If $V(\boldsymbol{q}^k, \boldsymbol{p}^k) \not\leq O$, put

$$\boldsymbol{q}^{k+1} := \Pi\big[\boldsymbol{q}^k - \gamma^k \nabla_{\boldsymbol{q}} \|V(\boldsymbol{q}^k, \boldsymbol{p}^k)^+\|_{\mathrm{F}}\big],$$
$$\ell := \ell + 1,$$

where

$$\gamma^k := \frac{\|V(\boldsymbol{q}^k, \boldsymbol{p}^k)^+\|_{\mathrm{F}} + r^\ell \big\|\nabla_{\boldsymbol{q}} \|V(\boldsymbol{q}^k, \boldsymbol{p}^k)^+\|_{\mathrm{F}}\big\|}{\big\|\nabla_{\boldsymbol{q}} \|V(\boldsymbol{q}^k, \boldsymbol{p}^k)^+\|_{\mathrm{F}}\big\|^2};$$

Otherwise, just put $\boldsymbol{q}^{k+1} := \boldsymbol{q}^k$.

3. $k := k + 1$. Return to Step 1. □

This algorithm produces a sequence of points $\boldsymbol{q}^k$. One can prove the next property on this sequence as is found in Appendix A.

Proposition 1. *Let $\ell(k)$ be the value of ℓ at the k-th iteration of the preceding algorithm. If $r^{\ell(k)} \leq r^*$ and $V(\boldsymbol{q}^k, \boldsymbol{p}^k) \not\leq O$, there holds*

$$\|\boldsymbol{q}^{k+1} - \boldsymbol{q}^*\|^2 \leq \|\boldsymbol{q}^k - \boldsymbol{q}^*\|^2 - [r^{\ell(k)}]^2.$$

When the point $\boldsymbol{q}^k$ is not in the solution region, the probability that a correction step is made in Step 2 is positive by assumption. This implies that correction of $\boldsymbol{q}^k$ is made after a finite number of iterations with a probability one. Combining this observation with the fact that the squared sum of $\{r^\ell\}$ is infinite, we have the following corollary. See Polyak and Tempo (2000) for more details.

Corollary 1. *With a probability one, there exists a finite number k_1 such that the point $\boldsymbol{q}^k$ produced by the preceding algorithm is contained in the solution region for any $k \geq k_1$.*

This property still holds in fact even if Q is unbounded. Among the numbers that can be k_1, we refer to the smallest one as the *achievement time* k_{A}. The above corollary claims that there exists a finite achievement time with a probability one.

A problem here is that it is not obvious how large the achievement time k_{A} is. Although Polyak and Tempo (2000) provide an answer to this question, it is not applicable in our generalized setting where the set P is continuous and the number r^ℓ is not constant.

3. Infiniteness of the expected achievement time

Since the achievement time k_{A} is a random number, it is easier to evaluate its expectation $\mathrm{E}_P(k_{\mathrm{A}})$, which is called the *expected achievement time* henceforth. The value of $\mathrm{E}_P(k_{\mathrm{A}})$ depends on the choice of the initial point $\boldsymbol{q}^0$. In order to evaluate it in an averaged case, we suppose that the initial point $\boldsymbol{q}^0$ is chosen according to some probability density function fixed on Q and consider the expectation $\mathrm{E}_Q[\mathrm{E}_P(k_{\mathrm{A}})]$. The claim of this section is that this expectation is infinite. It means that the preceding algorithm often requires a long time to achieve a solution and is not appropriate for a practical use. Improvement of the algorithm is certainly necessary, which is a task of the next section.

We add the following assumptions in order to obtain the mentioned result. Suppose that there exists a pair of $\boldsymbol{q} \in Q$ and $\boldsymbol{p} \in P$ that satisfies $V(\boldsymbol{q}, \boldsymbol{p}) \not\leq O$. The probability density function introduced into Q is assumed to have a positive lower bound $\underline{\mu}_Q$.

It is not difficult to see that the function $\sup_{\boldsymbol{p} \in P} \overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p})]$ is continuous and is convex with respect to $\boldsymbol{q}$. By the assumptions so far, there holds $\sup_{\boldsymbol{p} \in P} \overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p})] > 0$ at some $\boldsymbol{q} \in Q$ while $\sup_{\boldsymbol{p} \in P} \overline{\lambda}[V(\boldsymbol{q}^*, \boldsymbol{p})] < 0$ at $\boldsymbol{q}^*$. Accordingly, on the line segment connecting these two points, there exists one and only one $\boldsymbol{q}$ that attains $\sup_{\boldsymbol{p} \in P} \overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p})] = 0$. Let us write it as $\boldsymbol{q}^\sharp$ from now on. This $\boldsymbol{q}^\sharp$ belongs to the interior of Q. It is possible to construct a $(d-1)$-dimensional manifold in a neighborhood of $\boldsymbol{q}^\sharp$ so that it contains $\boldsymbol{q}^\sharp$ and $\sup_{\boldsymbol{p} \in P} \overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p})] = 0$ holds anywhere on this manifold. We make here our final assumption, that is, there is no $\boldsymbol{p} = [p_1 \quad \ldots \quad p_n]^{\mathrm{T}} \in P$ that satisfies

$$\det V(\boldsymbol{q}^\sharp, \boldsymbol{p}) = 0, \quad \frac{\partial \det V(\boldsymbol{q}^\sharp, \boldsymbol{p})}{\partial p_1} = 0, \quad \ldots, \quad \frac{\partial \det V(\boldsymbol{q}^\sharp, \boldsymbol{p})}{\partial p_n} = 0$$

simultaneously. This assumption is considered to be mild enough because $n+1$ equalities are not simultaneously satisfied in an n-dimensional space in general.

Theorem 1. *Under the assumptions so far, there holds*

$$\mathrm{E}_Q\big[\mathrm{E}_P(k_{\mathrm{A}})\big] = \infty.$$

The proof of this theorem requires the following lemma, which is proved in Appendix B.

Lemma 1. *Define* $\widehat{\boldsymbol{q}}(\alpha) := \boldsymbol{q}^\sharp + \alpha(\boldsymbol{q}^\sharp - \boldsymbol{q}^*)$. *Under the assumptions so far, there exists a positive* C *such that*

$$\mathrm{Prob}\{\boldsymbol{p} \in P : V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}) \not\leq O\} \leq C\alpha$$

holds for any small enough nonnegative α.

Proof of Theorem 1. Suppose that the initial point $\boldsymbol{q}^0$ is chosen at $\widehat{\boldsymbol{q}}(\alpha)$, $\alpha > 0$. Since this $\boldsymbol{q}^0$ is not in the solution region, at least one correction step is required to arrive at a solution. The expected number of steps required to make one correction step is

$$\frac{1}{\mathrm{Prob}\{\boldsymbol{p} \in P : V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}) \not\leq O\}},$$

which is greater than or equal to $1/C\alpha$ by Lemma 1. Since the expected achievement time $\mathrm{E}_P(k_\mathrm{A})$ is greater than or equal to this value, we have $\mathrm{E}_P(k_\mathrm{A}) \geq 1/C\alpha$.

Consider a $(d-1)$-dimensional manifold S on which $\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q},\boldsymbol{p})] = 0$ and the point $\boldsymbol{q}^\sharp$ is located. If we choose this manifold S small enough, there exists a positive α^0 such that the inequality

$$\mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q} + \alpha(\boldsymbol{q}^\sharp - \boldsymbol{q}^*), \boldsymbol{p}) \not\leq O\} \leq 2C\alpha$$

holds for any $\boldsymbol{q} \in S$ and any $0 < \alpha < \alpha^0$. Since the expectation $\mathrm{E}_Q[\mathrm{E}_P(k_\mathrm{A})]$ is the integration over the whole Q, the integration over its subset $\{\boldsymbol{q} + \alpha(\boldsymbol{q}^\sharp - \boldsymbol{q}^*) : \boldsymbol{q} \in S \text{ and } 0 < \alpha < \alpha_0\}$ is less than or equal to this. We consequently have

$$\mathrm{E}_Q[\mathrm{E}_P(k_\mathrm{A})] \geq \underline{\mu}_Q D \int_0^{\alpha_0} \mathrm{E}_P(k_\mathrm{A})\, \mathrm{d}\alpha \geq \underline{\mu}_Q D \int_0^{\alpha_0} \frac{1}{2C\alpha}\, \mathrm{d}\alpha = \infty$$

with some positive number D. □

4. Improved algorithm

As we can see in the proof of Theorem 1, the reason of the infinite expected achievement time is that the probability to make a correction step is close to zero near the solution region. We hence increase this probability hoping to have the finite expected achievement time. Let $a^0, a^1, \ldots$ be a sequence of positive numbers that monotonically decreases to zero and replace $V(\boldsymbol{q}^k, \boldsymbol{p}^k)$ in the preceding algorithm by $V(\boldsymbol{q}^k, \boldsymbol{p}^k) + a^\ell I$. In this improved algorithm, a correction step is made if $V(\boldsymbol{q}^k, \boldsymbol{p}^k) + a^\ell I \not\leq O$, which holds with more probability than $V(\boldsymbol{q}^k, \boldsymbol{p}^k) \not\leq O$.

Before investigating the expected achievement time, we show that the achievement time remains finite with a probability one after the improvement above. Define $a^* := -\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q}^*, \boldsymbol{p})]$, which is positive by the definition of $\boldsymbol{q}^*$. Let $\ell(k)$ be the value of ℓ at the k-th iteration of the improved algorithm and suppose $a^{\ell(k)} < a^*$. Then, since the function $\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q},\boldsymbol{p})]$ is convex, the ball with the center $\boldsymbol{q}^*$ and the radius $[1 - a^{\ell(k)}/a^*]r^*$ is included by the region $\{\boldsymbol{q} \in Q : \sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q},\boldsymbol{p}) + a^{\ell(k)} I] \leq 0\}$. This enables us to use Proposition 1 with replacing $r^{\ell(k)} \leq r^*$ by $r^{\ell(k)} \leq [1 - a^{\ell(k)}/a^*]r^*$ and $V(\boldsymbol{q}^k, \boldsymbol{p}^k)$ by $V(\boldsymbol{q}^k, \boldsymbol{p}^k) + a^{\ell(k)} I$. Following the reasoning to obtain Corollary 1, we arrive at the result below.

Proposition 2. *With a probability one, there exists a finite number k_1 such that the point $\boldsymbol{q}^k$ produced by the improved algorithm is contained in the solution region for any $k \geq k_1$.*

Again we refer to the smallest k_1 as the achievement time k_A.

We now consider the expected achievement time $\mathrm{E}_P(k_\mathrm{A})$ in the improved algorithm. As in the previous section, it is assumed that there exists a pair of $\boldsymbol{q} \in Q$ and $\boldsymbol{p} \in P$ satisfying $V(\boldsymbol{q},\boldsymbol{p}) \not\leq O$. We do not need here the assumption on the density function in Q or on the point $\boldsymbol{q}^\sharp$. Let G be a positive number with which

$$\left\| e_1 \frac{\partial V(\boldsymbol{q},\boldsymbol{p})}{\partial p_1} + \cdots + e_n \frac{\partial V(\boldsymbol{q},\boldsymbol{p})}{\partial p_n} \right\|_\mathrm{F} < G$$

holds for any $\boldsymbol{q} \in Q$, any $\boldsymbol{p} \in P$, and any unit vector $[e_1 \ \ \dots \ \ e_n]^{\mathrm{T}}$. Define $R := \sup_{\boldsymbol{q}\in Q} \|\boldsymbol{q} - \boldsymbol{q}^*\|$. The lemma below is important for our purpose.

Lemma 2. *Let a be any nonnegative number and $\boldsymbol{q}^\sharp$ be any vector satisfying $\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p})] = 0$. There then holds*

$$\mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^\sharp, \boldsymbol{p}) + aI \not\leq O\} \geq \min\Big\{1,\ \underline{\mu}_P\Big(\frac{a}{\sqrt{n}G}\Big)^n\Big\}.$$

This lemma is proved in Appendix C. From this lemma the following theorems are derived.

Theorem 2. *For any initial point $\boldsymbol{q}^0$, the expected achievement time $\mathrm{E}_P(k_{\mathrm{A}})$ defined for the improved algorithm is finite.*

Theorem 3. *Choose $r^\ell := r^0/\sqrt{\ell+1}$ and $a^\ell := a^0/\sqrt{\ell+1}$. Then, there holds*

$$\mathrm{E}_P(k_{\mathrm{A}}) < \max\left\{\mathrm{e}^{(R/r^0)^2}\Big(\frac{r^0}{r^*} + \frac{a^0}{a^*} + 1\Big)^2,\ \ \frac{\mathrm{e}^{(R/r^0)^2}[\mathrm{e}^{(R/r^0)^2} + 1]^{n/2}(\sqrt{n}G)^n}{\underline{\mu}_P(a^0)^n}\Big(\frac{r^0}{r^*} + \frac{a^0}{a^*} + 1\Big)^{n+2}\right\}. \quad (1)$$

The right-hand side of (1) increases faster than the exponential order of n. This suggests that our algorithm suffers from the curse of dimensionality like the gridding algorithm, which was discussed in Section 1. However, the curse of dimensionality is removed if it is sufficient to obtain an "approximate solution" that satisfies the provided LMI not for all $\boldsymbol{p} \in P$ but for almost all of them. This is shown next.

Theorem 4. *Let ϵ and δ be any positive numbers less than unity and let $\{\boldsymbol{q}^k\}$ be a sequence produced by the improved algorithm. With a probability greater than $1-\delta$, there exists a nonnegative integer k less than or equal to*

$$\frac{2\mathrm{e}^{(R/r^0)^2}}{\epsilon}\Big(\frac{r^0}{r^*} + \frac{a^0}{a^*} + 1\Big)^2 + \frac{1}{2\epsilon^2}\ln\frac{1}{\delta} \quad (2)$$

that satisfies $\mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^k, \boldsymbol{p}) \leq O\} \geq 1 - \epsilon$. Here $\ln$ denotes the natural logarithm.

Choose ϵ and δ to be small positive numbers. Then, with a high probability, the time to obtain an approximate solution has the upper bound (2), which is independent of the parameter dimension n. The present algorithm is therefore advantageous to the gridding algorithm and other deterministic algorithms when the parameter dimension n is high.

Proof of Theorem 2. If $\boldsymbol{q}^0$ is in the solution region, the achievement time is zero. We suppose in the rest of the proof that $\boldsymbol{q}^0$ is not in the solution region.

Let ℓ_0 be the smallest integer ℓ that attains $r^\ell \leq (1 - a^\ell/a^*)r^*$. Note that the same inequality remains valid for any $\ell \geq \ell_0$. Let ℓ_1 be the smallest integer satisfying $\sum_{\ell=\ell_0}^{\ell_1-1}(r^\ell)^2 \geq R^2$. The number of correction steps required to achieve the solution region must be less than or equal to ℓ_1 and, thus, is finite. It is left for us to show that the number of steps required to make one correction step is finite on the average.

Suppose that $\boldsymbol{q}^k$ is not in the solution region. Since the corresponding ℓ has to be less than or equal to ℓ_1, there holds $a^\ell \geq a^{\ell_1}$. Furthermore, there exists a point $\boldsymbol{q}^\sharp$ satisfying $\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p})] = 0$ on the line segment connecting $\boldsymbol{q}^k$ and $\boldsymbol{q}^*$. There holds by definition that $r^{\ell_1} \leq (1 - a^{\ell_1}/a^*)r^*$, which implies $a^{\ell_1} < a^*$ and then $\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q}^*, \boldsymbol{p}) + a^{\ell_1} I] < 0$. Since $\overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p}) + a^{\ell_1} I]$ is a convex function of $\boldsymbol{q}$, if $\overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p}) + a^{\ell_1} I] > 0$, it is inferred that $\overline{\lambda}[V(\boldsymbol{q}^k, \boldsymbol{p}) + a^{\ell_1} I] > 0$. By these observations, we can see that

$$\begin{aligned}\mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^k, \boldsymbol{p}) + a^\ell I \not\leq O\} &\geq \mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^k, \boldsymbol{p}) + a^{\ell_1} I \not\leq O\} \\ &\geq \mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^\sharp, \boldsymbol{p}) + a^{\ell_1} I \not\leq O\} \\ &\geq \min\left\{1,\ \underline{\mu}_P \left(\frac{a^{\ell_1}}{\sqrt{n}G}\right)^n\right\}. \end{aligned} \tag{3}$$

The last inequality follows from Lemma 2.

The expected number of steps required to make one correction step is the reciprocal of the leftmost expression in (3), which is shown to be finite. The proof is hence complete. □

Proof of Theorem 3. We evaluate ℓ_0 and ℓ_1 in the previous proof explicitly.

Solve $r^\ell \leq (1 - a^\ell/a^*)r^*$ to have $\ell \geq (r^0/r^* + a^0/a^*)^2 - 1$, which implies $\ell_0 < (r^0/r^* + a^0/a^*)^2$. Since $\sum_{\ell=\ell_0}^{\ell-1} (r^\ell)^2 > \int_{\ell_0}^{\ell} (r^0)^2/(x+1)\,\mathrm{d}x = (r^0)^2 \ln[(\ell+1)/(\ell_0+1)]$, the inequality $\ell \geq \mathrm{e}^{(R/r^0)^2}(\ell_0 + 1) - 1$ guarantees $\sum_{\ell=\ell_0}^{\ell-1} (r^\ell)^2 > R^2$. Therefore, we have

$$\ell_1 < \mathrm{e}^{(R/r^0)^2}(\ell_0 + 1) < \mathrm{e}^{(R/r^0)^2}\left[\left(\frac{r^0}{r^*} + \frac{a^0}{a^*}\right)^2 + 1\right] < \mathrm{e}^{(R/r^0)^2}\left(\frac{r^0}{r^*} + \frac{a^0}{a^*} + 1\right)^2$$

and

$$\frac{1}{a^{\ell_1}} = \frac{\sqrt{\ell_1 + 1}}{a^0} < \frac{1}{a^0}\sqrt{\mathrm{e}^{(R/r^0)^2}\left[\left(\frac{r^0}{r^*} + \frac{a^0}{a^*}\right)^2 + 1\right] + 1} < \frac{1}{a^0}\sqrt{\left[\mathrm{e}^{(R/r^0)^2} + 1\right]\left(\frac{r^0}{r^*} + \frac{a^0}{a^*} + 1\right)^2}.$$

The expected achievement time is no greater than the product of ℓ_1 and the reciprocal of the rightmost expression in (3). Substitution of the above inequalities gives the desired result. □

Proof of Theorem 4. Let $\widetilde{k}$ be the smallest integer greater than (2). Suppose that the inequality

$$\mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^k, \boldsymbol{p}) \not\leq O\} > \epsilon \tag{4}$$

holds for any $0 \leq k \leq \widetilde{k} - 1$. This implies that the point $\boldsymbol{q}^k$ does not arrive in the solution region at least until $k = \widetilde{k} - 1$, which ensures $\widetilde{\ell} < \ell_1$, where $\widetilde{\ell}$ is the value of ℓ at the time $\widetilde{k} - 1$ and ℓ_1 is the number defined in the proof of Theorem 3. Note that the inequality $\widetilde{\ell} < \ell_1$ is equivalent to $\epsilon - \widetilde{\ell}/\widetilde{k} > \epsilon - \ell_1/\widetilde{k}$. Consider in general that one carries out $\widetilde{k}$ Bernoulli trials whose probability of success is s. It is known as Chernoff's inequality that the probability to have $s - \widetilde{\ell}/\widetilde{k} \geq A$, where $\widetilde{\ell}$ is the number of success and A is any nonnegative number, is less than or equal to $\exp(-2\widetilde{k}A^2)$ (see, *e.g.*, p. 22, Vidyasagar, 1997).

Since we have (4) for any $0 \le k \le \widetilde{k}-1$ and the number $\epsilon - \ell_1/\widetilde{k}$ is nonnegative by assumption, the probability to have $\epsilon - \widetilde{\ell}/\widetilde{k} > \epsilon - \ell_1/\widetilde{k}$ is less than or equal to $\exp[-2(\widetilde{k}\epsilon - \ell_1)^2/\widetilde{k}]$. This quantity is less than δ again by assumption. Now it is seen that the inequality (4) fails to hold at least one $0 \le k \le \widetilde{k}-1$ with a probability greater than $1-\delta$. □

5. Example

The proposed two algorithms, the basic one and the improved one, are applied to an example problem.

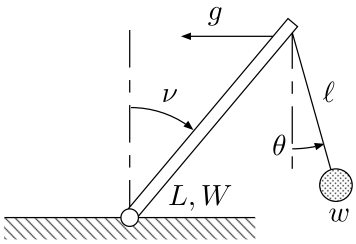


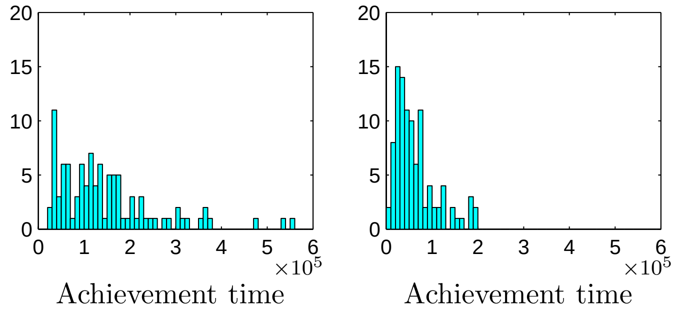
Figure 1. A considered plant

We consider a plant given in Figure 1, which is a simplification of the tower crane investigated by Takagi and Nishimura (1999). A uniform boom of the length L[m] and the weight W[kg] is connected to the ground by a free joint. From the top of the boom, a load of the weight w[kg] is hung down by a rope of the length ℓ[m]. The angles from the vertical line to the boom and to the rope are ν[rad] and θ[rad], respectively. One can add a force g[N] to the top of the boom in the horizontal direction. Especially, the force that makes the equilibrium boom angle equal to some given value ν_0[rad] is denoted by g_0[N].

Linearize the plant dynamics around the equilibrium point and choose $\boldsymbol{x} = [\nu - \nu_0 \quad \theta \quad \dot{\nu} \quad \dot{\theta}]^{\mathrm{T}}$ as a state vector and $u = g - g_0$ as an input. Then we have an LPV system $\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + \boldsymbol{b}u(t)$. The matrix A and the column vector $\boldsymbol{b}$ are functions of the parameter $\boldsymbol{p} := [\ell \quad \nu_0 \quad w]^{\mathrm{T}}$. Suppose that $L = 0.71\,\mathrm{m}$ and $W = 0.205\,\mathrm{kg}$ here and that the parameter $\boldsymbol{p}$ can take any value in the set $P = \{\boldsymbol{p} : 0.5\,\mathrm{m} \le \ell \le 1\,\mathrm{m},\quad (40\pi/180)\mathrm{rad} \le \nu_0 \le (50\pi/180)\mathrm{rad},\quad 0.1\,\mathrm{kg} \le w \le 0.2\,\mathrm{kg}\}$. We want to see whether the state feedback $u = \boldsymbol{f}^{\mathrm{T}}\boldsymbol{x}$, $\boldsymbol{f}^{\mathrm{T}} = [12.25 \quad -12.21 \quad 4.91 \quad 1.16]$, stabilizes this system for any parameter $\boldsymbol{p}$ in P. We suppose here for simplicity that the time derivative of the parameter $\boldsymbol{p}$ is zero. Let us consider a Lyapunov function having the form $\boldsymbol{x}^{\mathrm{T}}L(\boldsymbol{p})\boldsymbol{x}$ with $L(\boldsymbol{p}) = L_0 + \ell L_1 + (\cos\nu_0)L_2 + wL_3$. Then, a sufficient condition for stability is the existence of symmetric matrices $L_0, \dots, L_3$ that satisfy

$$\begin{bmatrix} [A(\boldsymbol{p}) + \boldsymbol{b}(\boldsymbol{p})\boldsymbol{f}^{\mathrm{T}}]^{\mathrm{T}}L(\boldsymbol{p}) + L(\boldsymbol{p})[A(\boldsymbol{p}) + \boldsymbol{b}(\boldsymbol{p})\boldsymbol{f}^{\mathrm{T}}] & \\ & -L(\boldsymbol{p}) \end{bmatrix} < O$$

for any $\boldsymbol{p}$ in P. The problem to find $L_0, \dots, L_3$ is formulated into the general form described in Section 2. Define $\boldsymbol{q}$ to be a vector consisting of the elements of $L_0, \dots, L_3$ and write the left-hand side of the above inequality as $V(\boldsymbol{q}, \boldsymbol{p})$. The matrices $L_0, \dots, L_3$ are seeked for in the region where each element of them belongs to the interval $[-2, 2]$. The variable domain Q is defined as such. We choose the probability density functions in P and Q to be those corresponding to the uniform distribution. The sequences $\{r^\ell\}$ and $\{a^\ell\}$ are defined by $r^\ell = a^\ell = 1/\sqrt{\ell+1}$.



(a) Basic algorithm (b) Improved algorithm

Figure 2. Distribution of the achievement time for 100 trials

In Figure 2, (a) and (b) are histograms of the achievement time when we execute the basic and the improved algorithms 100 times for each. While the achievement time of the basic algorithm distributes in a rather wide range, that of the improved algorithm concentrates in the range of $k \leq 2 \times 10^5$. This shows the effect of the improvement. Strictly speaking, the word "achievement time" is not used here in its original sense. Since it is difficult to judge whether a provided $\boldsymbol{q}^k$ satisfies $V(\boldsymbol{q}^k, \boldsymbol{p}) \leq O$ for all $\boldsymbol{p} \in P$, we regard $\boldsymbol{q}^k$ to be a solution if this $\boldsymbol{q}^k$ satisfies the inequality for all of 10000 $\boldsymbol{p}$'s randomly sampled from P. Hence, this is an achievement time to an "approximate solution" considered in Theorem 4. By the result of Khargonekar and Tikku (1996) and Tempo, Bai, and Dabbene (1997), the fact that $V(\boldsymbol{q}^k, \boldsymbol{p}) \leq O$ holds for all of randomly sampled 10000 $\boldsymbol{p}$'s implies that the inequality $\mathrm{Prob}\{\boldsymbol{p} \in P : V(\boldsymbol{q}^k, \boldsymbol{p}) \leq O\} \geq 0.999$ holds with a probability greater than 0.9999. From this we expect that the achievement time measured here sufficiently reflects the properties of the achievement time in the original sense.

The time to execute 2×10^5 iterations in the improved algorithm is measured on Pentium III 1.0 GHz with 256 MBytes memory. On the average of 100 executions, the time is 106 s without the achievement judgment described above.

6. Conclusion

Randomized algorithms for parameter-dependent LMIs are proposed as a generalization of the algorithm of Polyak and Tempo (2000) and their computational complexity is analyzed.

The following problems are left unsolved. The proposed algorithms do not contain an explicit termination rule. Since it is difficult to judge whether the considered LMI is satisfied for all the parameters $\boldsymbol{p}$ in P, it is considered to be practical to terminate the algorithms when the LMI is satisfied for all of many $\boldsymbol{p}$'s randomly sampled from P. If one adds this termination rule to our algorithms, their computational complexity is not sufficiently characterized by the iteration number k. One has to reexamine their complexity using a more appropriate index. The computational time presented in the previous section is not very long but not very short either. More speedup is preferable. It would be useful if one can explicitly evaluate the upper bounds given in Theorems 3 and 4. However, this is difficult because these upper bounds include unknown numbers r^* and a^*. Explicitly computable upper bounds are desired. Finally, it is assumed throughout this paper that the considered LMI has a solution. It is also important to investigate the case that the LMI has no solution.

Acknowledgment

Thanks are due to Professors Roberto Tempo and Yasumasa Fujisaki. They provided the authors the full-paper version of Polyak and Tempo (2000). Their comments on an earlier version of this paper motivated improvement of the example in Section 5 and derivation of Theorem 4.

A. Proof of Proposition 1

The proof is a complete parallel to that of Theorem 1 of Polyak and Tempo (2000).

We write r instead of $r^{\ell(k)}$ for simplicity. Since $r \le r^*$, the point

$$\widetilde{\boldsymbol{q}} := \boldsymbol{q}^* + r\,\frac{\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}}{\big\|\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big\|}$$

is included in the solution region and thus $\|V(\widetilde{\boldsymbol{q}},\boldsymbol{p}^k)^+\|_{\mathrm{F}} = 0$. From $V(\boldsymbol{q}^k,\boldsymbol{p}^k) \not\leq O$ it follows that

$$\begin{aligned}
\|\boldsymbol{q}^{k+1}-\boldsymbol{q}^*\|^2 \le& \big\|\boldsymbol{q}^k - \gamma^k\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}} - \boldsymbol{q}^*\big\|^2 \\
=& \|\boldsymbol{q}^k-\boldsymbol{q}^*\|^2 + (\gamma^k)^2\big\|\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big\|^2 - 2\gamma^k\big[\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big]^{\mathrm{T}}(\boldsymbol{q}^k-\widetilde{\boldsymbol{q}}) \\
&- 2\gamma^k\big[\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big]^{\mathrm{T}}(\widetilde{\boldsymbol{q}}-\boldsymbol{q}^*).
\end{aligned} \tag{5}$$

Since $\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}$ is a subgradient, there holds on the second last term of (5) $0 = \|V(\widetilde{\boldsymbol{q}},\boldsymbol{p}^k)^+\|_{\mathrm{F}} \ge \|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}} + \big[\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big]^{\mathrm{T}}(\widetilde{\boldsymbol{q}}-\boldsymbol{q}^k)$. On the last term of (5) we have $\big[\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big]^{\mathrm{T}}(\widetilde{\boldsymbol{q}}-\boldsymbol{q}^*) = r\big\|\nabla_{\boldsymbol{q}}\|V(\boldsymbol{q}^k,\boldsymbol{p}^k)^+\|_{\mathrm{F}}\big\|$. Substitution of these relations into (5) gives the desired inequality.

B. Proof of Lemma 1

The proof requires two lemmas. Let $\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp,\boldsymbol{p})$ denote the vector whose i-th element is $(\partial/\partial p_i)\det V(\boldsymbol{q}^\sharp,\boldsymbol{p})$.

Lemma 3. *There exists a (possibly multiple)* $\boldsymbol{p}^\sharp \in P$ *that satisfies* $\overline{\lambda}[V(\boldsymbol{q}^\sharp,\boldsymbol{p}^\sharp)] = 0$. *This* $\boldsymbol{p}^\sharp$ *is on the surface of* P *and the value of* $\|\nabla_{\boldsymbol{p}}\det V(\boldsymbol{q}^\sharp,\boldsymbol{p}^\sharp)\|$ *has a positive lower bound common to all of such* $\boldsymbol{p}^\sharp$. *If the parameter dimension* n *is even, the angle between the two vectors* $\nabla_{\boldsymbol{p}}\det V(\boldsymbol{q}^\sharp,\boldsymbol{p}^\sharp)$ *and* $\boldsymbol{x}-\boldsymbol{p}^\sharp$ *is no greater than* $\pi/2$ *for any* $\boldsymbol{x}\in P$ *other than* $\boldsymbol{p}^\sharp$. *If* n *is odd, the same is true with* $-\nabla_{\boldsymbol{p}}\det V(\boldsymbol{q}^\sharp,\boldsymbol{p}^\sharp)$ *instead of* $\nabla_{\boldsymbol{p}}\det V(\boldsymbol{q}^\sharp,\boldsymbol{p}^\sharp)$.

Proof. The existence of $\boldsymbol{p}^\sharp$ is ensured by the definition of $\boldsymbol{q}^\sharp$ and the compactness of P. Since all the eigenvalues of $V(\boldsymbol{q}^\sharp, \boldsymbol{p})$ are negative at any $\boldsymbol{p}$ that cannot be $\boldsymbol{p}^\sharp$, the function $\det V(\boldsymbol{q}^\sharp, \boldsymbol{p})$ takes the minimum value at $\boldsymbol{p}^\sharp$ if n is even and the maximum value otherwise. By assumption, the gradient $\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)$ is not the zero vector, which means that $\boldsymbol{p}^\sharp$ has to be on the surface of P. Because the set $\{\boldsymbol{p} \in P : \overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p})] = 0\}$ is closed and bounded, there exists a positive minimum value for $\|\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)\|$. The last property is a consequence of the Karush-Kuhn-Tucker necessary condition for optimality. □

Lemma 4. *There exists a positive number g such that $\det V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}) = 0$ implies $\max_{i=1,\dots,n} |p_i| \geq 1 - g\alpha$ for any small enough nonnegative α and any $\boldsymbol{p} \in P$.*

Proof. Let $\boldsymbol{p}^\sharp$ be a vector considered in Lemma 3. With a guarantee of Lemma 3, one can choose g so that the inequality $g\|\nabla_{\boldsymbol{p}} V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)\| > \left|(\partial/\partial\alpha) \det V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}^\sharp)|_{\alpha=0}\right|$ holds for any $\boldsymbol{p}^\sharp$. We negate the claim of the lemma with this g and derive contradiction.

Assume that there hold $\det V(\widehat{\boldsymbol{q}}(\alpha^j), \boldsymbol{x}^j) = 0$ and $\max_{i=1,\dots,n} |x_i^j| < 1 - g\alpha^j$ for some sequence of positive numbers $\alpha^1, \alpha^2, \dots$ that monotonically decreases to zero and for some sequence of parameters $\boldsymbol{x}^1, \boldsymbol{x}^2, \dots$ in P. Here, x_i^j denotes the i-th element of $\boldsymbol{x}^j$. The sequence $\{\boldsymbol{x}^j\}$ has to have an accumulation point, which has to be one of the considered $\boldsymbol{p}^\sharp$'s. We can assume that the sequence $\{\boldsymbol{x}^j\}$ converges to this $\boldsymbol{p}^\sharp$ by taking its subsequence if necessary. Expand $\det V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p})$ around $\alpha = 0$ and $\boldsymbol{p} = \boldsymbol{p}^\sharp$ to have

$$0 = \det V(\widehat{\boldsymbol{q}}(\alpha^j), \boldsymbol{x}^j) = \frac{\partial \det V(\widehat{\boldsymbol{q}}(\widetilde{\alpha}^j), \widetilde{\boldsymbol{x}}^j)}{\partial \alpha} \alpha^j + [\nabla_{\boldsymbol{p}} \det V(\widehat{\boldsymbol{q}}(\widetilde{\alpha}^j), \widetilde{\boldsymbol{x}}^j)]^{\mathrm{T}} (\boldsymbol{x}^j - \boldsymbol{p}^\sharp), \tag{6}$$

where $\widetilde{\alpha}^j \to 0$ and $\widetilde{\boldsymbol{x}}^j \to \boldsymbol{p}^\sharp$ as $j \to \infty$. From the assumption $\max_{i=1,\dots,n} |x_i^j| < 1 - g\alpha^j$, it follows that

$$g\alpha^j < \left| \frac{[\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)]^{\mathrm{T}}}{\|\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)\|} (\boldsymbol{x}^j - \boldsymbol{p}^\sharp) \right|.$$

Take the limit $j \to \infty$ in this inequality and use (6) to conclude

$$\begin{aligned} g\|\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)\| &\leq \lim_{j\to\infty} \left| \frac{1}{\alpha^j} [\nabla_{\boldsymbol{p}} \det V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp)]^{\mathrm{T}} (\boldsymbol{x}^j - \boldsymbol{p}^\sharp) \right| \\ &= \lim_{j\to\infty} \left| \frac{1}{\alpha^j} [\nabla_{\boldsymbol{p}} \det V(\widehat{\boldsymbol{q}}(\widetilde{\alpha}^j), \widetilde{\boldsymbol{x}}^j)]^{\mathrm{T}} (\boldsymbol{x}^j - \boldsymbol{p}^\sharp) \right| = \left| \frac{\partial \det V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}^\sharp)}{\partial \alpha} \Big|_{\alpha=0} \right|. \end{aligned}$$

This contradicts with the definition of g. □

We now prove Lemma 1. Suppose that $V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}) \not\leq O$ for some small enough nonnegative α and some $\boldsymbol{p} \in P$. Since $V(\widehat{\boldsymbol{q}}(0), \boldsymbol{p}) \leq O$, there exists $0 \leq \widetilde{\alpha} < \alpha$ with which $\det V(\widehat{\boldsymbol{q}}(\widetilde{\alpha}), \boldsymbol{p}) = 0$. Lemma 4 is invoked here to give $\max_{i=1,\dots,n} |p_i| \geq 1 - g\widetilde{\alpha} > 1 - g\alpha$.

By the above reasoning, the set $\{\boldsymbol{p} \in P : 1 \geq \max_{i=1,\dots,n} |p_i| \geq 1 - g\alpha\}$ includes the set $\{\boldsymbol{p} \in P : V(\widehat{\boldsymbol{q}}(\alpha), \boldsymbol{p}) \not\leq O\}$. The probability measure of the former set is at most $\overline{\mu}_P\left[2^n - (2 - 2g\alpha)^n\right]$, which is less than or equal to $\overline{\mu}_P 2^n n g\alpha$ for a small enough nonnegative α.

C. Proof of Lemma 2

If $a = 0$ the considered inequality is trivial. Suppose that $0 < a < 2\sqrt{n}G$. Since $\sup_{\boldsymbol{p}\in P} \overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p})] = 0$, there exists $\boldsymbol{p} \in P$ satisfying $\overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p})] = 0$. Write it as $\boldsymbol{p}^\sharp$ to have $\overline{\lambda}[V(\boldsymbol{q}^\sharp, \boldsymbol{p}^\sharp) + aI] = a$. From the definition of G, it follows that $G\|\boldsymbol{p} - \boldsymbol{p}'\| > \|V(\boldsymbol{q}, \boldsymbol{p}) - V(\boldsymbol{q}, \boldsymbol{p}')\|_{\mathrm{F}} \geq \left|\overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p}) + aI] - \overline{\lambda}[V(\boldsymbol{q}, \boldsymbol{p}') + aI]\right|$ for any $\boldsymbol{p}$ and $\boldsymbol{p}'$ in P. This implies that there holds $V(\boldsymbol{q}^\sharp, \boldsymbol{p}) + aI \not\leq O$ in the set $\{\boldsymbol{p} : \|\boldsymbol{p} - \boldsymbol{p}^\sharp\| \leq a/G\}$. Irrespective of the location of $\boldsymbol{p}^\sharp$, the intersection of this set and P includes the n-dimensional hypercube each edge of which has the length $a/\sqrt{n}G$. Since this hypercube ensures $V(\boldsymbol{q}^\sharp, \boldsymbol{p}) + aI \not\leq O$ in its inside and has the volume $(a/\sqrt{n}G)^n$, the desired inequality holds. If $a \geq 2\sqrt{n}G$, again the desired inequality is trivial because the set $\{\boldsymbol{p} : \|\boldsymbol{p} - \boldsymbol{p}^\sharp\| \leq a/G\}$ includes the whole P.

References

Apkarian, P., & Adams, R. J. (1998). Advanced gain-scheduling techniques for uncertain systems. *IEEE Transactions on Control Systems Technology*, *6* (1), 21–32.

Apkarian, P., Gahinet, P., & Becker, G. (1995). Self-scheduled H^∞ control of linear parameter-varying systems: a design example. *Automatica*, *31* (9), 1251–1261.

Barmish, B. R., & Lagoa, C. M. (1997). The uniform distribution: a rigorous justification for its use in robustness analysis. *Mathematics of Control, Signals, and Systems*, *10* (3), 203–222.

Becker, G., & Packard, A. (1994). Robust performance of linear parametrically varying systems using parametrically-dependent linear feedback. *Systems & Control Letters*, *23* (3), 205–215.

Becker, G., Packard, A., Philbrick, D., & Balas, G. (1993). Control of parametrically-dependent linear systems: a single quadratic Lyapunov approach. *Proceedings of the 1993 American Control Conference*, San Francisco, CA (pp. 2795–2799).

Chen, X. & Zhou, K. (2000). Constrained robustness analysis by randomized algorithms. *IEEE Transactions on Automatic Control*, *45* (6), 1180–1186.

Feron, E., Apkarian, P., & Gahinet, P. (1996). Analysis and synthesis of robust control systems via parameter-dependent Lyapunov functions. *IEEE Transactions on Automatic Control*, *41* (7), 1041–1046.

Fujisaki, Y., Dabbene, F., & Tempo, R. (2001). Probabilistic robust design of LPV control systems. To be presented at *the 40th IEEE Conference on Decision and Control*, Orlando, FL.

Gahinet, P., Apkarian, P., & Chilali, M. (1996). Affine parameter-dependent Lyapunov functions and real parametric uncertainty. *IEEE Transactions on Automatic Control*, *41* (3), 436–442.

Khargonekar, P., & Tikku, A. (1996). Randomized algorithms for robust control analysis and synthesis have polynomial complexity. *Proceedings of the 35th IEEE Conference on Decision and Control*, Kobe, Japan (pp. 3470–3475).

Masubuchi, I., & Shimemura, E. (1999). On application of the descriptor form to design of gain scheduling systems (in Japanese). *Transactions of the Institute of Systems, Control and Information Engineers*, *12* (7), 390–394.

Ohara, A., Ide, M., Yamaguchi, Y., & Ohno, M. (1999). LPV modeling and gain scheduled control of ALFLEX vehicle flight systems (in Japanese). *Transactions of the Institute of Systems, Control and Information Engineers*, *12* (11), 655–663.

Oishi, Y., & Kimura, H. (2001). Randomized algorithms to solve parameter-dependent linear matrix inequalities and their computational complexity. To be presented at *the 40th IEEE Conference on Decision and Control*, Orlando, FL.

Polyak, B. T., & Tempo, R. (2000). Probabilistic robust design with linear quadratic regulators. *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia (pp. 1037–1042). The full-paper version: *IRITI-CNR Report 7-00*.

Ray, L. R., & Stengel, R. F. (1993). A Monte Carlo approach to the analysis of control system robustness. *Automatica*, *29* (1), 229–236.

Takagi, K., & Nishimura, H. (1999). Gain-scheduled control of a tower crane considering varying load-rope length. *JSME International Journal*, *Ser. C*, *42* (4), 914–921.

Tempo, R., Bai, E. W., & Dabbene, F. (1997). Probabilistic robustness analysis: explicit bounds for the minimum number of samples. *Systems & Control Letters*, *30* (5), 237–242.

Vidyasagar, M. (1997). *A Theory of Learning and Generalization: With Applications to Neural Networks and Control Systems*. London, UK: Springer.

Vidyasagar, M. (1999). Randomized algorithms for robust controller synthesis using statistical learning theory. In Y. Yamamoto and S. Hara, *Learning, Control and Hybrid Systems* (pp. 3–24). London, UK: Springer.

Watanabe, R., Uchida, K., Fujita, M., & Shimemura, E. (1994). L^2 gain and H^∞ control of linear systems with scheduling parameter. *Proceedings of the 33rd IEEE Conference on Decision and Control*, Lake Buena Vista, FL (pp. 1412–1414).

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

Yu, J., & Sideris, A. (1997). H^∞ control with parametric Lyapunov functions. *Systems & Control Letters*, *30* (2–3), 57–69.