MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Probabilistic design of a robust controller using a parameter-dependent Lyapunov function

Yasuaki Oishi

METR 2004–35 June 2004

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

Probabilistic design of a robust controller using a parameter-dependent Lyapunov function*

Yasuaki Oishi†

An extension of a randomized algorithm is considered for the use of a parameter-dependent Lyapunov function. The proposed algorithm is considered to be useful for a less conservative design of a robust state-feedback controller against nonlinear parametric uncertainty. Indeed, one can avoid two sources of conservatism with the proposed algorithm: covering the uncertainty by a polytope and using a common Lyapunov function for all parameter values. After a bounded number of iterations, the proposed algorithm either gives a probabilistic solution to the provided control problem with high confidence or detects infeasibility of the problem in an approximated sense. Convergence to a non-strict deterministic solution is discussed. Usefulness of the proposed algorithm is illustrated by a numerical example.

Keywords. randomized algorithm, parameter-dependent Lyapunov function, conservatism, parameter-dependent linear matrix inequality, computational complexity, semidefinite programming.

1. Introduction

Although a robust-controller design is a classical control problem, it still remains difficult against nonlinear parametric uncertainty. One existing approach to this problem consists of two simplifications: (i) to cover the uncertainty by a polytope and consider the plants corresponding to the vertices of the polytope, and (ii) to assume a common Lyapunov function for all of those plants. These two simplifications enable us to formulate and solve the above problem in terms of a linear matrix inequality (LMI). On the other hand, these simplifications produce conservatism, that is, even if the original problem is solvable, the simplified one may not.

In order to reduce conservatism, we propose in this paper a randomized algorithm that allows the use of a parameter-dependent Lyapunov function. This is based on a randomized algorithm, which was proposed in [13] as an extension of [17, 4, 16, 8], and also on a technique to use a parameter-dependent Lyapunov function, which was developed by de Oliveira *et al.* [5] with further extensions by [1, 2, 18, 20, 6, 7]. In particular, the proposed algorithm does not need to cover the uncertainty and allows a Lyapunov function having any parameter dependence.

*Released on June 14, 2004

†Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan; phone: +81-3-5841-6921; fax: +81-3-5841-6886; email: oishi@mist.i.u-tokyo.ac.jp

A problem to be considered is formulated in terms of a parameter-dependent LMI having two types of variables: parameter-independent variables, which describe a controller, and parameter-dependent variables, which correspond to a Lyapunov function. The proposed algorithm iterates the following three steps: (i) random selection of uncertain parameter values; (ii) optimization of the parameter-dependent variables; (iii) update of the parameter-independent variables based on the sensitivity of the preceding optimization problem. This algorithm terminates after a bounded number of iterations; At its termination, it gives a probabilistic solution of the provided problem with high confidence or detects that the problem has no deterministic solution in an approximated sense. The above sensitivity is efficiently evaluated based on a theory of semidefinite programming. In [17, 4, 16, 8], it is shown that their algorithms finds a non-strict deterministic solution instead of a probabilistic one after a finite number of iterations with probability one. We can prove a corresponding result on our algorithm. However, this result is not emphasized in this paper because it has some practical difficulties as is discussed in [13] for a special case.

The contents of this paper have been presented in earlier forms at the conferences [11, 12].

In Section 2, the problem to be considered is presented with an example. An algorithm is proposed in Section 3 with a theoretical guarantee on its performance. Section 4 describes how one can compute an infimum and a subgradient, which are required in the algorithm. In Section 5, convergence to a non-strict deterministic solution is discussed. Section 6 provides a numerical example and Section 7 concludes this paper.

Throughout this paper, ln stands for the natural logarithm. For a real number a, the symbol $\lceil a \rceil$ denotes the minimum integer that is larger than or equal to a. The symbol $\mathbb{R}^n$ designates the n-dimensional Euclidean space and vol the volume in $\mathbb{R}^n$. The symbol $\|\cdot\|$ denotes the Euclidean norm. The transpose of a matrix or a vector is expressed by $^{\mathrm{T}}$. The maximum (most positive) eigenvalue of a symmetric matrix A is written as $\overline{\lambda}[A]$. Negative definiteness and negative semidefiniteness of a symmetric matrix A are denoted by $A \prec O$ and $A \preceq O$, respectively.

2. Problem

The problem to be considered is stated in a general form. Let $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})$ be a symmetric-matrix-valued function in $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^m$, and $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$. At this moment, the parameter set Θ can be any set. The function $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})$ is affine in $\boldsymbol{x}$ and $\boldsymbol{y}$ and is written as

$$V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}) = V_0(\boldsymbol{\theta}) + \sum_{i=1}^{n} x_i V_i(\boldsymbol{\theta}) + \sum_{i=1}^{m} y_i V_{n+i}(\boldsymbol{\theta}),$$

where x_i and y_i are the ith elements of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively.

Problem 1. Find $\boldsymbol{x} \in \mathbb{R}^n$ such that for any $\boldsymbol{\theta} \in \Theta$ there exists $\boldsymbol{y} \in \mathbb{R}^m$ satisfying $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}) \prec O$. □

The choice of $\boldsymbol{y}$ can depend on the parameter $\boldsymbol{\theta}$. The inequality $V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}) \prec O$ is called a *parameter-dependent LMI.* When Θ consists of infinitely many components, our problem is equivalent to solving infinitely many LMIs with infinitely many variables. Note also that our problem reduces to the one considered in the existing papers [17, 4, 16, 8, 13] in the special case that V does not depend on $\boldsymbol{y}$.

Here is an example that can be formulated in the above general form.

Example 1. Consider a discrete-time plant $\boldsymbol{\xi}[k+1] = A(\boldsymbol{\theta})\boldsymbol{\xi}[k] + B(\boldsymbol{\theta})\boldsymbol{u}[k]$, which depends nonlinearly on a time-invariant but uncertain parameter $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$. Suppose that we want a state-feedback controller $\boldsymbol{u}[k] = K\boldsymbol{\xi}[k]$ that robustly stabilizes this plant for all $\boldsymbol{\theta} \in \Theta$. This problem can be stated in terms of a parameter-dependent LMI based on the result of de Oliveira *et al.* [5]. Namely, consider to find two matrices G and L such that for any $\boldsymbol{\theta} \in \Theta$ there exists a symmetric matrix Y satisfying

$$-\begin{bmatrix} Y & A(\boldsymbol{\theta})G + B(\boldsymbol{\theta})L \\ G^{\mathrm{T}}A(\boldsymbol{\theta})^{\mathrm{T}} + L^{\mathrm{T}}B(\boldsymbol{\theta})^{\mathrm{T}} & G + G^{\mathrm{T}} - Y \end{bmatrix} \prec O. \tag{1}$$

Then $K := LG^{-1}$ gives a desired controller. Here, $\boldsymbol{\eta}[k]^{\mathrm{T}}Y\boldsymbol{\eta}[k]$ works as a Lyapunov function for the transposed version of the closed-loop system $\boldsymbol{\eta}[k+1] = [A(\boldsymbol{\theta}) + B(\boldsymbol{\theta})K]^{\mathrm{T}}\boldsymbol{\eta}[k]$. Since Y can be chosen depending on $\boldsymbol{\theta}$, the Lyapunov function is allowed to be parameter-dependent. This means that a less conservative result can be expected than in the case that a common Lyapunov function is used for all $\boldsymbol{\theta}$. In order to express this problem in our general form, construct $\boldsymbol{x}$ by the components of G and L, construct $\boldsymbol{y}$ by the independent components of Y, and write the left-hand side of (1) as V. □

The idea in this example has been extended to robust performance problems [1, 6] and to problems on continuous-time systems [2, 18, 20, 7]. These can be stated in our general form, too.

The $\boldsymbol{x}$ desired in Problem 1 is called a *deterministic solution.* The set of all deterministic solutions is referred to as the *solution set* $S \subseteq \mathbb{R}^n$. Rather than a deterministic solution itself, we consider to obtain an approximate solution in this paper. For a given probability measure P on Θ and a given $0 < \epsilon < 1$, we mean by a *probabilistic solution* an $\boldsymbol{x} \in \mathbb{R}^n$ that satisfies

$$P\big\{\boldsymbol{\theta} \in \Theta : \exists \boldsymbol{y} \in \mathbb{R}^m \text{ s.t. } V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}) \prec O\big\} > 1 - \epsilon.$$

By a *non-strict deterministic solution*, we mean an $\boldsymbol{x} \in \mathbb{R}^n$ such that for any $\boldsymbol{\theta} \in \Theta$ there exists $\boldsymbol{y} \in \mathbb{R}^m$ satisfying $V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}) \preceq O$. Note that the strict inequality in Problem 1 is replaced by the non-strict one.

3. Algorithm

This is the main section of this paper and proposes a randomized algorithm for finding a probabilistic solution of Problem 1. The basic idea is as follows. Note that $V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}) \prec O$ is equivalent to $\overline{\lambda}[V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})] < 0$. Hence, for provided $\boldsymbol{x}$ and $\boldsymbol{\theta}$, there exists $\boldsymbol{y} \in \mathbb{R}^m$ satisfying $V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}) \prec O$ if and only if $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})] < 0$. This $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})]$ is convex in $\boldsymbol{x}$ because $\overline{\lambda}[V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})]$ is convex in $\boldsymbol{x}$ and $\boldsymbol{y}$. By these facts, we can extend the randomized algorithms in [13], which are for the special case that V does not depend on $\boldsymbol{y}$. Although the extended algorithm requires computation of the infimum value, $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta})]$, and its subgradient with respect to $\boldsymbol{x}$, their computation is efficiently carried out as we will see in the next section.

In [13], two randomized algorithms are presented: the gradient-based algorithm, which was originally proposed by Polyak and Tempo [17] and Calafiore and Polyak [4], and the ellipsoid-based algorithm, which was by Kanev *et al.* [8]. We present only an extension of the ellipsoid-based algorithm because an extension of the gradient-based one is similar.

The algorithm to be proposed iteratively updates an ellipsoid, which is expected to contain a deterministic solution. The ellipsoid at the kth iteration, $E^{(k)}$, is specified by its center $\boldsymbol{x}^{(k)}$ and a positive definite matrix $Q^{(k)}$ in the form of $\{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{x} - \boldsymbol{x}^{(k)})^{\mathrm{T}}(Q^{(k)})^{-1}(\boldsymbol{x} - \boldsymbol{x}^{(k)}) < 1\}$.

Before executing the algorithm, we choose an initial ellipsoid $E^{(0)} = (\boldsymbol{x}^{(0)}, Q^{(0)})$ and three numbers $0 < \mu$, $0 < \epsilon < 1$, and $0 < \delta < 1$. We usually choose these three numbers close to zero. We use two counters in the algorithm: k counts the number of iterations and ℓ the number of updates. We define an integer

$$\overline{\ell} := \left\lceil 2(n+1)\ \ln \frac{\operatorname{vol} E^{(0)}}{\mu} \right\rceil$$

and a function

$$\kappa(\ell) := \left\lceil \left(\ln \frac{\pi^2(\ell+1)^2}{6\delta}\right) \Big/ \ln \frac{1}{1-\epsilon} \right\rceil.$$

Algorithm 1.

0. Set $k := 0$ and $\ell := 0$.

1. If ℓ reaches $\overline{\ell}$, stop with no output.

2. If $\lambda^{(k-1)}, \lambda^{(k-2)}, \dots, \lambda^{(k-\kappa(\ell))}$ are all well-defined and negative, stop and give $\boldsymbol{x}^{(k)}$ as an output.

3. Randomly sample $\boldsymbol{\theta}^{(k)} \in \Theta$ according to the given probability measure P.

4. Check $\lambda^{(k)} := \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})]$ for negativity.

Case 1 The value $\lambda^{(k)}$ is nonnegative.

Compute a subgradient, say $\boldsymbol{d}^{(k)}$, of $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})]$ as a convex function in $\boldsymbol{x}$. Update the current ellipsoid by setting

$$\boldsymbol{x}^{(k+1)} := \boldsymbol{x}^{(k)} - \frac{Q^{(k)}\boldsymbol{d}^{(k)}}{(n+1)\sqrt{(\boldsymbol{d}^{(k)})^{\mathrm{T}}Q^{(k)}\boldsymbol{d}^{(k)}}},$$

$$Q^{(k+1)} := \frac{n^2}{n^2-1}\left[Q^{(k)} - \frac{2Q^{(k)}\boldsymbol{d}^{(k)}(\boldsymbol{d}^{(k)})^{\mathrm{T}}Q^{(k)}}{(n+1)(\boldsymbol{d}^{(k)})^{\mathrm{T}}Q^{(k)}\boldsymbol{d}^{(k)}}\right].$$

Case 2 The value $\lambda^{(k)}$ is negative.

Keep the current ellipsoid by setting $\boldsymbol{x}^{(k+1)} := \boldsymbol{x}^{(k)}$ and $Q^{(k+1)} := Q^{(k)}$.

5. If $\lambda^{(k)} \geq 0$, set $\ell := \ell + 1$.

6. Set $k := k + 1$. Go back to Step 1. □

The performance of this algorithm is guaranteed by the following theorem, which is a direct generalization of Theorem 21 in [13], a result on the special case that V does not depend on $\boldsymbol{y}$. In particular, the statement (a) evaluates the computational complexity and the statements (b) and (c) give properties of the output. Although the proof is almost the same as in the special case, it is presented in Appendix A for convenience of the readers. Notice that the probabilistic behavior of the algorithm is determined by the sequence $\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \ldots$. We hence analyze the behavior of the algorithm according to the probability measure on such sequences, which can be derived from P and is denoted by P^∞. See for example [21, Section II.3.4].

Theorem 1. *The following statements hold on Algorithm 1.*

(a) The number of iterations k is bounded as

$$k \leq \overline{\ell}\kappa(\overline{\ell}-1) = \overline{\ell}\left\lceil \left(\ln\frac{\pi^2\overline{\ell}^2}{6\delta}\right) \Big/ \ln\frac{1}{1-\epsilon}\right\rceil =: \overline{k}.$$

(b) If the algorithm terminates at Step 1, the volume of the set $E^{(0)} \cap S$ is less than μ.

(c) The probability that the algorithm stops at Step 2 but still the corresponding output $\boldsymbol{x}^{(k)}$ fails to satisfy $P\{\boldsymbol{\theta} \in \Theta : \exists \boldsymbol{y} \in \mathbb{R}^m \text{ s.t. } V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}) \prec O\} > 1 - \epsilon$ is less than or equal to δ, where the probability is measured with respect to P^∞.

The statement (a) of Theorem 1 means that our algorithm stops in $\overline{k}$ iterations. The number $\overline{k}$ is of order $\mathrm{O}((\overline{\ell}/\epsilon)\ln(\overline{\ell}^2/\delta))$, which is a polynomial in n, $\ln(\mathrm{vol}\, E^{(0)}/\mu)$, $1/\epsilon$, and $\ln(1/\delta)$. Note that this number does not depend on the dimension of the parameter p. This forms a sharp contrast with deterministic algorithms whose complexity is usually of exponential order of p.

When the algorithm stops at Step 1, we see $\mathrm{vol}\,(E^{(0)} \cap S) < \mu$ by the statement (b). This means that our choice of an initial ellipsoid $E^{(0)}$ is not good or the solution set S itself is too small. On the other hand, termination at Step 2 implies that the associated output is a probabilistic solution with high confidence. Note that the algorithm always stops at Step 2 if $\mathrm{vol}\,(E^{(0)} \cap S) \geq \mu$.

As is considered in Section 6 of [13] for the special case, we can adapt the algorithm for an optimization problem. Indeed, since the algorithm gives a negative or a positive result, its bisectional use leads us to finding a solution that approximately optimizes a given objective function. For the explicit algorithm and its properties, see Section 6 of [13].

4. Computation of the infima and the subgradients

In order to carry out Algorithm 1 in the previous section, we need to compute the infimum value, $\lambda^{(k)} = \inf_{\boldsymbol{y} \in \mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})]$, and its subgradient $\boldsymbol{d}^{(k)}$. This section provides a method for this task.

Let us notice that the infimum value $\lambda^{(k)}$ is the optimal value of the following semidefinite programming problem:

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)}) - \lambda I \preceq O, \end{array} \tag{2}$$

where the optimization variables are $\boldsymbol{y}$ and λ. By following a general theory on semidefinite programming (see, *e.g.*, [19]), its dual problem is

$$\begin{array}{ll} \text{maximize} & \mathrm{tr}\,[UV(\boldsymbol{x}^{(k)}, \boldsymbol{0}, \boldsymbol{\theta}^{(k)})] \\ \text{subject to} & U \succeq O, \quad \mathrm{tr}\, U = 1, \\ & \mathrm{tr}\,[UV_{n+i}(\boldsymbol{\theta}^{(k)})] = 0 \ \text{ for } i = 1, \ldots, m, \end{array} \tag{3}$$

with the optimization variable being U, where tr designates the trace of a matrix. The following properties hold on this dual problem.

Theorem 2. *If the optimal value of the primal problem (2) is finite, the dual problem (3) has an optimal solution and the optimal value is equal to the infimum value $\lambda^{(k)}$. In this case, the set of all subgradients of the infimum value is equal to the set of vectors $\begin{bmatrix} \mathrm{tr}\,[UV_1(\boldsymbol{\theta}^{(k)})] & \cdots & \mathrm{tr}\,[UV_n(\boldsymbol{\theta}^{(k)})] \end{bmatrix}^{\mathrm{T}}$, where U is an optimal solution of (3).*

Proof. The first statement follows from the duality theorem on semidefinite programming [19, Theorem 4.1.3]. Notice that the Slater condition is satisfied in our primal problem (2), that is, there exist $\boldsymbol{y}$ and λ that satisfy $V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)}) - \lambda I \prec O$. The second statement is derived from sensitivity analysis on semidefinite programming [19, Theorem 4.1.2]. □

Theorem 2 reduces computation of the infimum value $\lambda^{(k)}$ and the subgradient $\boldsymbol{d}^{(k)}$ to solving the dual semidefinite programming problem (3). Since one can efficiently solve a semidefinite programming problem, computation of $\lambda^{(k)}$ and $\boldsymbol{d}^{(k)}$ is now possible.

Remark 1. In some situations, one can check $\lambda^{(k)}$ for negativity without computing its value explicitly. In the case of Example 1, for example, $\lambda^{(k)}$ is negative if and only if the controller corresponding to $\boldsymbol{x}^{(k)}$ stabilizes the plant at $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$. Stability of the system can be checked by location of its poles. Using such an alternative is preferable when it is faster and/or numerically more robust than semidefinite programming. □

Remark 2. In the algorithms of [17, 4, 16, 8], the norm $\|V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})^+\|_{\mathrm{F}}$ is used in place of our $\overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$, where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm and $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})^+$ is the projection of $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})$ onto the cone of positive semidefinite matrices. There are two reasons why we adopt the maximum eigenvalue. First, the infimum value and its subgradient can be efficiently computed as we have seen. Second, $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}) \prec O$ can be differentiated from $V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}) \preceq O$ with $\overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$, while this is not possible with $\|V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})^+\|_{\mathrm{F}}$. The latter point is important because a strict inequality is often required in control applications. The use of $\overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$ is due to [10]. □

5. Convergence to a non-strict deterministic solution

In the papers [17, 4, 16, 8], they considered a non-strict deterministic solution in our terminology and showed that their randomized algorithms obtain this solution after a finite number of iterations with probability one. We prove in this section a corresponding result for our algorithm and investigate the expected number of necessary iterations. For this purpose, we slightly modify Algorithm 1 in Section 3 by removing Steps 1 and 2. Similarly to [17, 4, 16, 8], the following assumptions are made.

Assumption 1. For $\boldsymbol{x}^* \in \mathbb{R}^n$, the probability $P\{\boldsymbol{\theta} \in \Theta : \inf_{\boldsymbol{y} \in \mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^*, \boldsymbol{y}, \boldsymbol{\theta})] > 0\}$ is positive whenever $\inf_{\boldsymbol{y} \in \mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^*, \boldsymbol{y}, \boldsymbol{\theta})] > 0$ holds for some $\boldsymbol{\theta} \in \Theta$. □

Assumption 2. The intersection between the initial ellipsoid $E^{(0)}$ and the solution set S has a positive volume. □

Theorem 3. *Under Assumptions 1 and 2, there exists with probability one a finite k such that $\boldsymbol{x}^{(k)}$ produced by the modified algorithm is a non-strict deterministic solution, where the probability is measured with P^∞.*

Proof. The proof is basically the same as that of Lemma 2 in [8]. Since $E^{(k)} \supseteq E^{(0)} \cap S$ as in the proof of Theorem 1 (b), the volume of the ellipsoid $E^{(k)}$ has to be greater than or equal to that of $E^{(0)} \cap S$, which is positive due to Assumption 2. This means that the number of updates is finite by the same reasoning as in the proof of Theorem 1 (b). In the modified algorithm, it is allowed that no update is made for an infinite number of consecutive iterations. However, the probability that this occurs is zero by Assumption 1. □

For each sequence $\{\boldsymbol{\theta}^{(k)}\}$, there may exist multiple k's having the property of the theorem. We write the minimum such k as k_{N}, which stands for the number of iterations necessary to find a non-strict deterministic solution. Note that k_{N} is a random number.

Though the above result is theoretically attractive, it is less practical than we seek for a probabilistic solution as in the preceding sections. The reason is as follows. First, it is difficult to choose an initial ellipsoid $E^{(0)}$ so that Assumption 2 holds because we do not know the location of the solution set S in many cases. If Assumption 2 fails to hold, the finite-time convergence is not guaranteed. Second, it is difficult to detect when $\boldsymbol{x}^{(k)}$ becomes a non-strict deterministic solution. Indeed, in order to detect it, we need to check $\inf_{\boldsymbol{y}\in\mathbb{R}^m} V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}) \preceq O$ for all $\boldsymbol{\theta} \in \Theta$, which is impossible in a typical case that Θ consists of infinitely many parameter values. Finally, the number of necessary iterations k_{N} is distributed with a heavy tail. In fact, the expectation of this number is infinite as in the following theorem. The corresponding result has been reported in the special case that V does not depend on $\boldsymbol{y}$ [15, 13].

Theorem 4. *Suppose that an initial ellipsoid $E^{(0)}$ is provided. Under Assumptions 3–6 below, the number of necessary iterations k_{N} in the modified algorithm has infinite expectation, where the expectation is taken according to P^{∞}.*

The proof is similar to that in the special case. See Appendix B.

Due to the dependence on $\boldsymbol{y}$, the required assumptions are stronger than in the special case.

Assumption 3. The parameter set Θ is a bounded closed set having a nonempty interior and a finite-measure boundary. There exists an open set that includes Θ and has $V_0(\boldsymbol{\theta}), \ldots, V_{n+m}(\boldsymbol{\theta})$ continuously differentiable there. The probability measure P has a density function possessing a finite upper bound and a positive lower bound. □

Assumption 4. There exists a finite sequence of parameter values $\{\widehat{\boldsymbol{\theta}}^{(k)}\}_{k=0}^{k^*}$ having the following properties, where $\widehat{\boldsymbol{\theta}}^{(0)}, \ldots, \widehat{\boldsymbol{\theta}}^{(k^*-1)}$ are in the interior of Θ and $\widehat{\boldsymbol{\theta}}^{(k^*)}$ is in Θ. If we choose $\boldsymbol{\theta}^{(0)} = \widehat{\boldsymbol{\theta}}^{(0)}$, $\ldots$, $\boldsymbol{\theta}^{(k^*)} = \widehat{\boldsymbol{\theta}}^{(k^*)}$ in the modified algorithm, we obtain $\boldsymbol{x}^{(0)} = \widehat{\boldsymbol{x}}^{(0)}, \ldots, \boldsymbol{x}^{(k^*)} = \widehat{\boldsymbol{x}}^{(k^*)}$, with which the following statements hold:

(a) For each of $k = 0, \ldots, k^* - 1$, there holds $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\widehat{\boldsymbol{x}}^{(k)}, \boldsymbol{y}, \widehat{\boldsymbol{\theta}}^{(k)})] > 0$;

(b) $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\widehat{\boldsymbol{x}}^{(k^*)}, \boldsymbol{y}, \widehat{\boldsymbol{\theta}}^{(k^*)})] = \max_{\boldsymbol{\theta}\in\Theta} \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\widehat{\boldsymbol{x}}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta})] = 0$;

(c) For each of $k = 0, \dots, k^*$, there exists $\widehat{\boldsymbol{y}}^{(k)} \in \mathbb{R}^m$ that attains the infimum value $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\widehat{\boldsymbol{x}}^{(k)}, \boldsymbol{y}, \widehat{\boldsymbol{\theta}}^{(k)})]$;

(d) For each of $k = 0, \dots, k^*$, the maximum eigenvalue $\overline{\lambda}[V(\widehat{\boldsymbol{x}}^{(k)}, \widehat{\boldsymbol{y}}^{(k)}, \widehat{\boldsymbol{\theta}}^{(k)})]$ is simple;

(e) For each of $k = 0, \dots, k^*$, the Hessian of $\overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$ with respect to $\boldsymbol{y}$ is positive definite at $(\widehat{\boldsymbol{x}}^{(k)}, \widehat{\boldsymbol{y}}^{(k)}, \widehat{\boldsymbol{\theta}}^{(k)})$;

(f) The gradient of $\overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$ with respect to $\boldsymbol{\theta}$ is nonzero at $(\widehat{\boldsymbol{x}}^{(k^*)}, \widehat{\boldsymbol{y}}^{(k^*)}, \widehat{\boldsymbol{\theta}}^{(k^*)})$. □

The assumption (d) guarantees the existence of the Hessian and the gradient considered in the assumptions (e) and (f). By the assumptions (c), (d), and (e), the infimum value, $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$, is continuously differentiable in neighborhoods of $(\widehat{\boldsymbol{x}}^{(k)}, \widehat{\boldsymbol{\theta}}^{(k)})$, $k = 0, \dots, k^*$. In particular, the gradient of $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta})]$ with respect to $\boldsymbol{\theta}$ is nonzero at $(\widehat{\boldsymbol{x}}^{(k^*)}, \widehat{\boldsymbol{\theta}}^{(k^*)})$ by the assumption (f). Due to this fact and the assumption (b), the point $\widehat{\boldsymbol{\theta}}^{(k^*)}$ has to lie on the boundary of Θ.

Since $\boldsymbol{x}^{(k^*)}$ depends on the choice of $\boldsymbol{\theta}^{(0)}, \dots, \boldsymbol{\theta}^{(k^*-1)}$, it is possible to regard $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k^*)})]$ as a function of $\boldsymbol{\theta}^{(0)}, \dots, \boldsymbol{\theta}^{(k^*)}$. This function is continuously differentiable in a neighborhood of $(\widehat{\boldsymbol{\theta}}^{(0)}, \dots, \widehat{\boldsymbol{\theta}}^{(k^*)})$ by the above discussion and the update rule of the algorithm. It is also possible to regard $\max_{\boldsymbol{\theta}^{(k^*)}\in\Theta} \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k^*)})]$ as a function of $\boldsymbol{\theta}^{(0)}, \dots, \boldsymbol{\theta}^{(k^*-1)}$. The next assumption is made on this function.

Assumption 5. The maximum value, $\max_{\boldsymbol{\theta}^{(k^*)}\in\Theta} \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k^*)})]$, is attained at a unique $\boldsymbol{\theta}^{(k^*)}$ for each $(\boldsymbol{\theta}^{(0)}, \dots, \boldsymbol{\theta}^{(k^*-1)})$ in a neighborhood of $(\widehat{\boldsymbol{\theta}}^{(0)}, \dots, \widehat{\boldsymbol{\theta}}^{(k^*-1)})$. □

It follows from this assumption that the maximum value, $\max_{\boldsymbol{\theta}^{(k^*)}\in\Theta} \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k^*)})]$, is continuously differentiable in a neighborhood of $(\widehat{\boldsymbol{\theta}}^{(0)}, \dots, \widehat{\boldsymbol{\theta}}^{(k^*-1)})$ and that the unique maximizing $\boldsymbol{\theta}^{(k^*)}$ is continuous there. The final assumption is now made.

Assumption 6. There exists at least one $k = 0, \dots, k^* - 1$ such that the gradient of $\max_{\boldsymbol{\theta}^{(k^*)}\in\Theta} \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k^*)})]$ with respect to this $\boldsymbol{\theta}^{(k)}$ is nonzero at $(\widehat{\boldsymbol{\theta}}^{(0)}, \dots, \widehat{\boldsymbol{\theta}}^{(k^*-1)})$. □

6. Example

In order to illustrate our approach, we consider robust stabilization of a tower crane model, which is taken from [22] with some simplification. Application of Algorithm 1 successfully gave a probabilistic solution as we will see below. On the other hand, we were not able to find a solution due to the conservatism with the existing approach based on a polytopic cover and a common Lyapunov function. These results show practical usefulness of our approach.

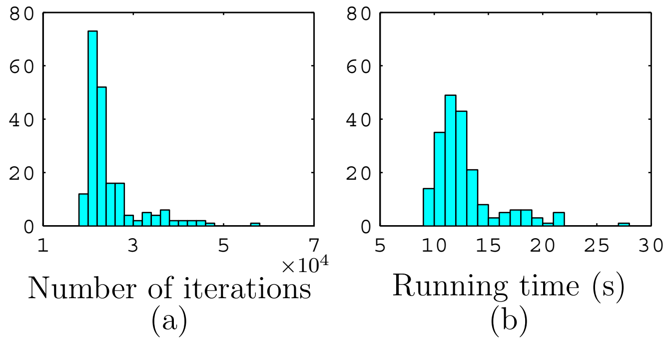

Figure 1. Histograms for 200 executions of the algorithm

By discretizing the model with the sampling period 0.01 s, we obtain the plant $\boldsymbol{\xi}[k+1] = A\boldsymbol{\xi}[k] + B\boldsymbol{u}[k]$, where the dimension of $\boldsymbol{\xi}[k]$ is four. The coefficients A and B nonlinearly depend on the three-dimensional parameter $\boldsymbol{\theta}$, which expresses the rope length, the boom angle at the equilibrium point, and the load weight. More details of the plant is found in [14]. It is supposed that each component of $\boldsymbol{\theta}$ can take any value in some interval, which means that our parameter set Θ is a hyper rectangle. We consider to design a robustly stabilizing controller by using the formulation of Example 1 and applying Algorithm 1, which results in $n = 20$ and $m = 10$. We use the uniform distribution as P and choose the initial ellipsoid $E^{(0)}$ by letting $\boldsymbol{x}^{(0)}$ be the zero vector and $Q^{(0)}$ be the identity. The parameters are set as $\mu = 10^{-9}\mathrm{vol}\, E^{(0)}$, $\epsilon = 0.001$, and $\delta = 0.0001$.

We executed the algorithm 200 times with Pentium 4 of 2.4 GHz and memory of 2.0 GByte. We used SDPA-M[1] of the version 2.00 to solve the semidefinite programming problem (3). In every execution, the algorithm stopped at Step 2 and gave an output. Figure 1 (a) shows the histogram of the number of iterations and (b) shows that of the running time. We can see from Figure 1 (b) that the running time is less than 20 seconds in many cases, which is practical enough. By Theorem 1 (c), on the other hand, the probability that the algorithm stops at Step 2 but the obtained output is not a probabilistic solution of $\epsilon = 0.001$ is less than or equal to $\delta = 0.0001$. This means that the obtained output is a probabilistic solution with high confidence.

7. Conclusion

A randomized algorithm that allows the use of a parameter-dependent Lyapunov function is proposed. It is considered to be useful for less conservative robust-controller design against nonlinear parametric uncertainty. This algorithm provides a probabilistic solution after a bounded number of iterations, which is of polynomial order in the problem size. Convergence to a non-

[1] http://grid.r.dendai.ac.jp/sdpa/index.html

strict deterministic solution is considered and its practical difficulties are pointed out. Since our problem is described in a general form, the application of the proposed algorithm is not limited to robust-controller design.

Acknowledgments

The author is thankful to the following people for their comments: Stephen Boyd, Marco Campi, Fabrizio Dabbene, Shinji Hara, Satoru Iwata, Masakazu Kojima, Kazuo Murota, Maria Prandini, Roberto Tempo, and Takashi Tsuchiya. This work is supported by the 21st Century COE Program on Information Science and Technology Strategic Core and a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

A. Proof of Theorem 1

(a) By the construction of the algorithm, at most $\kappa(\ell_0)$ iterations are made during the period that the counter ℓ has the value ℓ_0. Since the algorithm stops when ℓ reaches $\overline{\ell}$, the number of iterations k has to be less than or equal to $\kappa(0) + \cdots + \kappa(\overline{\ell} - 1) \leq \overline{\ell}\kappa(\overline{\ell} - 1)$.

(b) Suppose that the ellipsoid $E^{(k)}$ is updated. Then, the ellipsoid $E^{(k+1)}$ is in fact the minimum-volume ellipsoid that contains the intersection between the original ellipsoid $E^{(k)}$ and the half space $\{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{d}^{(k)})^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^{(k)}) < 0\}$ [3]. As we will see below, this half space includes the solution set S. Therefore $E^{(k+1)}$ includes the set $E^{(k)} \cap S$. It can also be shown that $\operatorname{vol} E^{(k+1)} < (\operatorname{vol} E^{(k)})\mathrm{e}^{-1/2(n+1)}$ [3]. Hence, if the ellipsoid $E^{(k)}$ has experienced $\overline{\ell}$ updates, it satisfies $E^{(k)} \supseteq E^{(0)} \cap S$ and $\operatorname{vol} E^{(k)} < (\operatorname{vol} E^{(0)})\mathrm{e}^{-\overline{\ell}/2(n+1)} \leq \mu$, which shows the claim.

We show that the half space $\{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{d}^{(k)})^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^{(k)}) < 0\}$ includes the solution set S. To this end, note that $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})] \geq \inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})] + (\boldsymbol{d}^{(k)})^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^{(k)}) \geq (\boldsymbol{d}^{(k)})^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^{(k)})$ because $\boldsymbol{d}^{(k)}$ is the subgradient of $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})]$ and the value $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})]$ is nonnegative. By this inequality, if $\boldsymbol{x}$ belongs to S, it satisfies $\inf_{\boldsymbol{y}\in\mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}^{(k)})] < 0$ and thus $(\boldsymbol{d}^{(k)})^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^{(k)}) < 0$.

(c) This is an extension of a result of [9] and [23]. Write the solution candidate after ℓ updates as $\boldsymbol{x}^{[\ell]}$ noting that the candidate remains unchanged until the next update. We define two events for each of $\ell = 0, 1, \ldots$:

- M_ℓ : The number of updates reaches ℓ and the candidate $\boldsymbol{x}^{[\ell]}$ is not updated for consecutive $\kappa(\ell)$ iterations;
- B_ℓ : The number of updates reaches ℓ and $\boldsymbol{x}^{[\ell]}$ satisfies $P\{\boldsymbol{\theta} \in \Theta : \exists \boldsymbol{y} \in \mathbb{R}^m \text{ s.t. } V(\boldsymbol{x}^{[\ell]}, \boldsymbol{y}, \boldsymbol{\theta}) \prec O\} \leq 1 - \epsilon$.

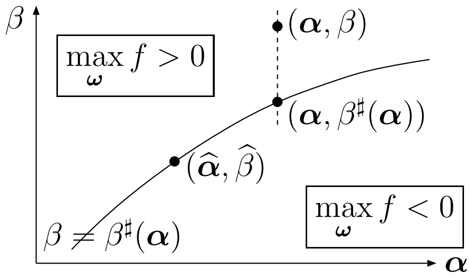

Figure 2. The function $\beta^\sharp(\boldsymbol{\alpha})$

In the following, we will show

$$P^\infty\big[(M_0 \cap B_0) \cup (M_1 \cap B_1) \cup \cdots\big] \le \delta, \tag{4}$$

which establishes the claim.

We have, for any ℓ,

$$P^\infty(M_\ell \cap B_\ell) = P^\infty(M_\ell | B_\ell) P^\infty(B_\ell) \le P^\infty(M_\ell | B_\ell) \le (1-\epsilon)^{\kappa(\ell)},$$

where $P^\infty(M_\ell | B_\ell)$ expresses the conditional probability of M_ℓ with B_ℓ being assumed. Therefore, we have

$$\begin{aligned} P^\infty\big[(M_0 \cap B_0) \cup (M_1 \cap B_1) \cup \cdots\big] &\le (1-\epsilon)^{\kappa(0)} + (1-\epsilon)^{\kappa(1)} + \cdots \\ &\le \frac{6\delta}{\pi^2}\Big(1 + \frac{1}{2^2} + \cdots\Big) = \delta, \end{aligned}$$

which shows (4).

B. Proof of Theorem 4

We prepare some notation. As was noticed after Assumption 4, the infimum value, $\inf_{\boldsymbol{y} \in \mathbb{R}^m} \overline{\lambda}[V(\boldsymbol{x}^{(k^*)}, \boldsymbol{y}, \boldsymbol{\theta}^{(k^*)})]$, is regarded as a function of $(\boldsymbol{\theta}^{(0)}, \ldots, \boldsymbol{\theta}^{(k^*)})$, which is continuously differentiable in a neighborhood of $(\widehat{\boldsymbol{\theta}}^{(0)}, \ldots, \widehat{\boldsymbol{\theta}}^{(k^*)})$. Let us write this function as $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega})$, where $[\boldsymbol{\alpha}^{\mathrm{T}} \quad \beta]^{\mathrm{T}}$ stands for $[(\boldsymbol{\theta}^{(0)})^{\mathrm{T}} \quad \ldots \quad (\boldsymbol{\theta}^{(k^*-1)})^{\mathrm{T}}]^{\mathrm{T}}$ with β being the last element of the vector and $\boldsymbol{\omega}$ stands for $\boldsymbol{\theta}^{(k^*)}$. We also use the notation $[\widehat{\boldsymbol{\alpha}}^{\mathrm{T}} \quad \widehat{\beta}]^{\mathrm{T}} = [(\widehat{\boldsymbol{\theta}}^{(0)})^{\mathrm{T}} \quad \ldots \quad (\widehat{\boldsymbol{\theta}}^{(k^*-1)})^{\mathrm{T}}]^{\mathrm{T}}$ and $\widehat{\boldsymbol{\omega}} = \widehat{\boldsymbol{\theta}}^{(k^*)}$. By rearranging the order of the elements if necessary, we can assume due to Assumption 6 that $(\partial/\partial\beta) \max_{\boldsymbol{\omega} \in \Theta} f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) \neq 0$ at $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$. Using the implicit function theorem, we can define a continuously differentiable function $\beta^\sharp(\boldsymbol{\alpha})$ so that $\beta^\sharp(\widehat{\boldsymbol{\alpha}})$ is equal to $\widehat{\beta}$ and $\max_{\boldsymbol{\omega} \in \Theta} f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) = 0$ is equivalent to $\beta = \beta^\sharp(\boldsymbol{\alpha})$ in a neighborhood of $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$. The situation is illustrated in Figure 2.

For each pair of $(\boldsymbol{\alpha}, \beta)$, a point $\boldsymbol{x}^{(k^*)}$ is determined. A pair $(\boldsymbol{\alpha}, \beta)$ satisfies $\max_{\boldsymbol{\omega} \in \Theta} f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) > 0$ if and only if the corresponding $\boldsymbol{x}^{(k^*)}$ is not a non-strict deterministic solution, in which case this $\boldsymbol{x}^{(k^*)}$ needs to be further updated. We will evaluate the probability of this update, which is equal to $P^\infty\{\boldsymbol{\omega} \in \Theta : f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) \geq 0\}$. Indeed, the two lemmas below show that this probability is bounded by a linear function of $|\beta - \beta^\sharp(\boldsymbol{\alpha})|$. Once such evaluation is obtained, the theorem is easily proved.

We need more preparation to present the first lemma. By Assumption 5, the maximum value, $\max_{\boldsymbol{\omega}\in\Theta} f(\boldsymbol{\alpha}, \beta^\sharp(\boldsymbol{\alpha}), \boldsymbol{\omega})$, is attained at a unique $\boldsymbol{\omega}$, which is written as $\boldsymbol{\omega}^\sharp(\boldsymbol{\alpha})$. This $\boldsymbol{\omega}^\sharp(\boldsymbol{\alpha})$ is continuous in $\boldsymbol{\alpha}$ as was noted after Assumption 5. Distinguishing the first element of $\boldsymbol{\omega}$ from the others, we write $\boldsymbol{\omega} = [\omega_1 \;\; \boldsymbol{\omega}_2^{\mathrm{T}}]^{\mathrm{T}}$, $\widehat{\boldsymbol{\omega}} = [\widehat{\omega}_1 \;\; \widehat{\boldsymbol{\omega}}_2^{\mathrm{T}}]^{\mathrm{T}}$, and so on. We can assume without loss of generality that $(\partial/\partial\omega_1) f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega})$ is nonzero at $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta}, \widehat{\boldsymbol{\omega}})$ since the gradient is nonzero as was discussed after Assumption 4.

Lemma 1. *There exist a neighborhood of $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$ and a positive number c such that, if $(\boldsymbol{\alpha}, \beta)$ belongs to that neighborhood and $\boldsymbol{\omega} \in \Theta$ satisfies $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) = 0$, there exists a real number t such that $[\omega_1 + t \;\; \boldsymbol{\omega}_2^{\mathrm{T}}]^{\mathrm{T}}$ is on the boundary of Θ and $|t| \leq c|\beta - \beta^\sharp(\boldsymbol{\alpha})|$.*

Proof. Since $(\partial/\partial\omega_1) f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega})$ is nonzero at $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta}, \widehat{\boldsymbol{\omega}})$, we can apply the implicit function theorem and define a continuously differentiable function $\omega_1(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}_2)$ so that $\omega_1(\widehat{\boldsymbol{\alpha}}, \widehat{\beta}, \widehat{\boldsymbol{\omega}}_2)$ is equal to $\widehat{\omega}_1$ and $f(\boldsymbol{\alpha}, \beta, [\omega_1, \boldsymbol{\omega}_2^{\mathrm{T}}]^{\mathrm{T}}) = 0$ is equivalent to $\omega_1 = \omega_1(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}_2)$ in a neighborhood of $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta}, \widehat{\boldsymbol{\omega}})$. Moreover, it is possible to assume the existence of a positive number c such that

$$\left|\frac{\partial\omega_1(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}_2)}{\partial\beta}\right| = \left|\frac{(\partial/\partial\beta) f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega})}{(\partial/\partial\omega_1) f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega})}\right| \leq c. \tag{5}$$

At $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$, the relationship $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) = 0$ holds only at $\widehat{\boldsymbol{\omega}}$; Hence, it is possible to choose a neighborhood of $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$, in which $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) = 0$ holds only at $\boldsymbol{\omega}$ obtained as $[\omega_1(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}_2) \;\; \boldsymbol{\omega}_2^{\mathrm{T}}]^{\mathrm{T}}$. Now, suppose that $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) = 0$ holds with $(\boldsymbol{\alpha}, \beta)$ in this neighborhood. By integrating (5) from β to $\beta^\sharp(\boldsymbol{\alpha})$, we have the lemma because no $\boldsymbol{\omega} \in \Theta$ satisfies $f(\boldsymbol{\alpha}, \beta^\sharp(\boldsymbol{\alpha}), \boldsymbol{\omega}) = 0$ except for $\boldsymbol{\omega}^\sharp(\boldsymbol{\alpha})$. □

Lemma 2. *There exist a neighborhood of $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$ and a positive number c' such that all $(\boldsymbol{\alpha}, \beta)$ in that neighborhood satisfy $P\{\boldsymbol{\omega} \in \Theta : f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) \geq 0\} \leq c'|\beta - \beta^\sharp(\boldsymbol{\alpha})|$.*

Proof. In Lemma 1, a neighborhood of $(\widehat{\boldsymbol{\alpha}}, \widehat{\beta})$ can be chosen to be convex without loss of generality. Suppose that $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega}) \geq 0$ holds for some $(\boldsymbol{\alpha}, \beta)$ in that neighborhood and $\boldsymbol{\omega} \in \Theta$. Since $f(\boldsymbol{\alpha}, \beta^\sharp(\boldsymbol{\alpha}), \boldsymbol{\omega}) \leq 0$, continuity of $f(\boldsymbol{\alpha}, \beta, \boldsymbol{\omega})$ implies that there exists $\widetilde{\beta}$ satisfying $f(\boldsymbol{\alpha}, \widetilde{\beta}, \boldsymbol{\omega}) = 0$ on the line segment connecting β and $\beta^\sharp(\boldsymbol{\alpha})$. Lemma 1 guarantees the existence of t such that $[\omega_1 + t \;\; \boldsymbol{\omega}_2^{\mathrm{T}}]^{\mathrm{T}}$ is on the boundary of Θ and $|t| \leq c|\widetilde{\beta} - \beta^\sharp(\boldsymbol{\alpha})| \leq c|\beta - \beta^\sharp(\boldsymbol{\alpha})|$. Due to Assumption 3, the probability of $\boldsymbol{\omega} \in \Theta$ having this property is bounded by $c'|\beta - \beta^\sharp(\boldsymbol{\alpha})|$ for some $c' > 0$. Note that we can choose the same c' for all $\boldsymbol{\alpha}$ in a neighborhood of $\widehat{\boldsymbol{\alpha}}$. □

We now prove the theorem. We can assume without loss of generality that $(\partial/\partial\beta)\max_{\boldsymbol{\omega}\in\Theta} f(\widehat{\boldsymbol{\alpha}}, \widehat{\beta}, \boldsymbol{\omega})$ is not only nonzero but also positive. We choose positive numbers a and b small enough that the set $A = \{(\boldsymbol{\alpha}, \beta) : \|\boldsymbol{\alpha}-\widehat{\boldsymbol{\alpha}}\| \leq a$ and $\beta^\sharp(\boldsymbol{\alpha}) < \beta \leq \beta^\sharp(\boldsymbol{\alpha})+b\}$ is contained in the neighborhood of Lemma 2. For a pair $(\boldsymbol{\alpha}, \beta) \in A$, the corresponding $\boldsymbol{x}^{(k^*)}$ is not a non-strict deterministic solution and thus needs to be updated. Because the probability of update is bounded by $c'|\beta - \beta^\sharp(\boldsymbol{\alpha})|$ due to Lemma 2, the expected number of iterations required to make one update is larger than or equal to $1/c'|\beta - \beta^\sharp(\boldsymbol{\alpha})|$. If we integrate this number in A with P^∞, it gives a lower bound of the expected number of iterations necessary to find a non-strict deterministic solution. The result is infinity, which completes the proof.

References

[1] P. Apkarian, P. C. Pellanda, and H. D. Tuan, "Mixed H^2/H^∞ multi-channel linear parameter-varying control in discrete time," *Systems & Control Letters*, vol. 41, no. 5, pp. 333–346, 2000.

[2] P. Apkarian, H. D. Tuan, and J. Bernussou, "Continuous-time analysis, eigenstructure assignment, and H^2 synthesis with enhanced linear matrix inequalities (LMI) characterizations," *IEEE Transactions on Automatic Control*, vol. 46, no. 12, pp. 1941–1946, 2001.

[3] R. G. Bland, D. Goldfarb, and M. J. Todd, "The ellipsoid method: A survey," *Operations Research*, vol. 29, no. 6, pp. 1039–1091, 1981.

[4] G. Calafiore and B. Polyak, "Stochastic algorithms for exact and approximate feasibility of robust LMIs," *IEEE Transactions on Automatic Control*, vol. 46, no. 11, pp. 1755–1759, 2001.

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

[6] M. C. de Oliveira, J. C. Geromel, and J. Bernussou, "Extended H^2 and H^∞ norm characterizations and controller parametrizations for discrete-time systems," *International Journal of Control*, vol. 75, no. 9, pp. 666–679, 2002.

[7] Y. Ebihara and T. Hagiwara, "New dilated LMI characterizations for continuous-time control design and robust multiobjective control," in *Proceedings of the 2002 American Control Conference*, Anchorage, USA, May 2002, pp. 47–52.

[8] S. Kanev, B. De Schutter, and M. Verhaegen, "An ellipsoid algorithm for probabilistic robust controller design," *Systems & Control Letters*, vol. 49, no. 5, pp. 365–375, 2003.

[9] P. Khargonekar and A. Tikku, "Randomized algorithms for robust control analysis and synthesis have polynomial complexity," in *Proceedings of the 35th IEEE Conference on Decision and Control*, Kobe, Japan, December 1996, pp. 3470–3475.

[10] D. Liberzon and R. Tempo, "Gradient algorithms for finding common Lyapunov functions," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, USA, December 2003, pp. 4782–4787.

[11] Y. Oishi, "An ellipsoid randomized algorithm for the use of non-common Lyapunov functions," in *Proceedings of the Third SICE Annual Conference on Control Systems*, Kobe, Japan, May 2003, pp. 473–476.

[12] Y. Oishi, "Probabilistic design of a robust state-feedback controller based on parameter-dependent Lyapunov functions," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, USA, December 2003, pp. 1920–1925.

[13] Y. Oishi, "Polynomial-time algorithms for probabilistic solutions of parameter-dependent linear matrix inequalities," *Technical Report METR 2004-23*, Department of Mathematical Informatics, The University of Tokyo, Japan, May 2004, http://www.keisu.t.u-tokyo.ac.jp/Research/techrep.0.html

[14] Y. Oishi and H. Kimura, "Randomized algorithms to solve parameter-dependent linear matrix inequalities," *Technical Report METR 2001-04*, Department of Mathematical Informatics, The University of Tokyo, Japan, September 2001, http://www.keisu.t.u-tokyo.ac.jp/Research/techrep.0.html

[15] Y. Oishi and H. Kimura, "Computational complexity of randomized algorithms for solving parameter-dependent linear matrix inequalities," *Automatica*, vol. 39, no. 12, pp. 2149–2156, 2003.

[16] B. T. Polyak, "Random algorithms for solving convex inequalities," in *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, D. Butnariu, Y. Censor, and S. Reich, Eds. Amsterdam, The Netherlands: Elsevier, 2001, pp. 409–422.

[17] B. T. Polyak and R. Tempo, "Probabilistic robust design with linear quadratic regulators," *Systems & Control Letters*, vol. 43, no. 5, pp. 343–353, 2001.

[18] U. Shaked, "Improved LMI representation for the analysis and the design of continuous-time systems with polytopic type uncertainty," *IEEE Transactions on Automatic Control*, vol. 46, no. 4, pp. 652–656, 2001.

[19] A. Shapiro, "Duality, optimality conditions, and perturbation analysis," in *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds. Boston, USA: Kluwer, 2000, pp. 68–92.

[20] T. Shimomura, M. Takahashi, and T. Fujii, "Extended-space control design with parameter-dependent Lyapunov functions," in *Proceedings of the 40th IEEE Conference on Decision and Control*, Orlando, USA, December 2001, pp. 2157–2162.

[21] A. N. Shiryaev, *Probability*, second ed. New York, USA: Springer, 1996.

[22] K. Takagi and H. Nishimura, "Gain-scheduled control of a tower crane considering varying load-rope length," *JSME International Journal*, ser. C, vol. 42, no. 4, pp. 914–921, 1999.

[23] R. Tempo, E. W. Bai, and F. Dabbene, "Probabilistic robustness analysis: Explicit bounds for the minimum number of samples," *Systems & Control Letters*, vol. 30, no. 5, pp. 237–242, 1997.