Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Stochastic Ellipsoid Methods with Multiple Cuts

Takayuki Wada [*,**] and **Yasumasa Fujisaki** [***]

[*] *Graduate School of Science and Technology, Kobe University, Nada, Kobe 657–8501, Japan (Tel: +81-78-803-6251; e-mail: takayuki@cs23.cs.kobe-u.ac.jp).*
[**] *Research Fellow of the Japan Society for the Promotion of Science (DC).*
[***] *Graduate School of Engineering, Kobe University, Nada, Kobe 657–8501, Japan (Tel: +81-78-803-6246; e-mail: fujisaki@cs.kobe-u.ac.jp).*

Abstract: Robust control systems synthesis is generally recast as a class of robust feasibility problems which is to find a solution satisfying a set of parameter-dependent convex constraints for all possible parameter values. For this class of the problems, a stochastic ellipsoid method with multiple cuts each of which corresponds to each of the constraint is proposed, where a new update rule is presented for constructing a smaller ellipsoid which contains the intersection of a previous ellipsoid and half spaces determined by given multiple subgradients. Moreover, we show an explicit relation between the volume of the ellipsoid updated by the original method and that of the proposed method. A quantitative analysis of the volume of the updated ellipsoid is also provided, which leads to a further modification of the algorithm for achieving fast convergence.

Keywords: Robust control; Parameter-dependent linear matrix inequalities; Probabilistic robustness; Convex optimization.

1. INTRODUCTION

Robust control synthesis such as guaranteed cost regulator or LPV control can generally be recast as a class of robust feasibility problems, which is to find a solution satisfying a parameter-dependent constraint. Even though this problem has a convex constraint in design variable, finding a solution is deterministically intractable due to its parameter-dependency, that is, we have to handle infinite number of constraints depending on uncertain parameters in principle.

In order to cope with this difficulty, the stochastic ellipsoid method (Kanev, De Schutter, and Verhaegen, 2003; Oishi, 2007) has been proposed. This algorithm employs random sampling of the uncertain parameter and uses the ellipsoid method (Bland, Goldfarb, and Todd, 1981) for updating the candidate of the solution. Then, a critical issue for improving this type of the algorithms is to construct a smaller ellipsoid which contains the intersection of the previous ellipsoid and the possible solution set determined by a subgradient of the constraint with the sampled parameter.

In the previous work (Wada and Fujisaki, 2006), the authors utilized a fact that the robust control synthesis is originally formulated with several matrix inequalities corresponding to, e.g., state feedback gain, observer gain, coupling condition, and so on, and proposed a modified stochastic ellipsoid method. This idea was further explored in the context of switched system design (Wada and Fujisaki, 2007). Note here that the algorithm (Wada and Fujisaki, 2006) handles these multiple constraints separately in each iteration, which can lead to more updates of the candidate of the solution for one random sample, while the original algorithm can lead to only one update for one random sample. As a result, with this algorithm, we can reduce the total number of random samples that is necessary for convergence, which was extensively demonstrated through numerical examples.

In contrast to the previous work, in this paper, we propose a new update rule of an ellipsoid which directly utilizes multiple subgradients given by the multiple constraints. In other words, we here consider *one-shot* update with given multiple cuts, while the previous approach employs a set of updates corresponding to the given cuts. Then, we focus our attention on *theoretical aspects* of the proposed algorithm. In fact, we prove that the volume of the ellipsoid updated by the proposed update rule with multiple subgradients is always smaller than that of the ellipsoid updated by the original update rule with one subgradient. Moreover, we derive an explicit relation between the volume of the ellipsoid updated by the original method and that of the proposed method. We also provide a quantitative analysis of the volume of the updated ellipsoid, which leads to a further modification of the algorithm for achieving fast convergence.

We hasten to note that there are a few other update rules of an ellipsoid which use multiple subgradients. For example, in *surrogate cuts* (Goldfarb and Todd, 1982), multiple subgradients are combined into a particular cut which is deeper than any cut derived by the given subgradients. That is, the ellipsoid is updated *indirectly* with the cut. On the other hand, in the literature (Shor and Gershovich, 1979; Ech-Cherif and Ecker, 1984), the updated ellipsoid is constructed directly with multiple cuts if its center and its directions of all principal axes can be determined in advance. Unfortunately, this property is not satisfied in general except when the number of the given subgradients is *two*. In contrast to these existing update rules, the proposed method enables us to update the ellipsoid *directly* with multiple cuts in general even when the number of the subgradients is *greater than two*.

This paper is organized as follows. In Section 2, we state the problem formulation of the robust feasibility problem with multiple constraints. In section 3, we propose a stochastic ellipsoid method with multiple cuts to solve this problem. Section 4 presents numerical examples. Finally, we make some concluding remarks in Section 5.

2. PROBLEM FORMULATION

Let us consider a robust feasibility problem with multiple constraints:

$$\text{Find } z \in \mathbb{R}^n \quad \text{s.t. } v_i(z,\theta) \le 0 \qquad \forall\theta \in \Theta \in \mathbb{R}^p, \ \forall i \in \{1,2,\ldots,m\} \tag{1}$$

where Θ is a measurable set and $v_i : \mathbb{R}^n \times \Theta \mapsto \mathbb{R}$ is a measurable function and convex in z. In this paper, we assume that $n \ge 2$, which has been used in the literature of standard ellipsoid methods (Bland, Goldfarb, and Todd, 1981).

Since we do not assume that each constraint v_i of this problem is convex in θ, it is generally difficult to find an element of the solution set

$$\mathcal{S} \doteq \{z \in \mathbb{R}^n : v_i(z,\theta) \le 0 \quad \forall\theta \in \Theta, \forall i \in \{1,2,\ldots,m\}\}$$

if we take a deterministic approach.

We therefore take a probabilistic approach (Tempo, Calafiore, and Dabbene, 2004). That is, we first introduce a probability measure $\mathcal{P}$ into the set Θ. Then, we consider a $z \in \mathbb{R}^n$ which satisfies

$$\mathcal{P}\{\theta \in \Theta : v_i(z,\theta) \le 0 \quad \forall i \in \{1,2,\ldots,m\}\} \ge 1-\varepsilon$$

for a given accuracy $\varepsilon \in (0,1)$. That is, we consider an approximate solution which satisfies the parameter-dependent constraints for *almost* all parameter values. We call such a $z \in \mathbb{R}^n$ a *probabilistic solution* with an accuracy ε.

Our problem in this paper is to develop a randomized algorithm which finds a probabilistic solution of the problem (1) within a given risk $\delta \in (0,1)$ or says that the problem (1) is infeasible in a deterministic sense.

3. RANDOMIZED ALGORITHM

In this section, we propose a stochastic ellipsoid method with multiple cuts for seeking a probabilistic solution within a specified risk $\delta \in (0,1)$. We first choose an initial candidate $z_0 \in \mathbb{R}^n$ of the solution, a positive definite matrix $Q_0 > 0$, and a parameter $\mu > 0$. With z_0, Q_0, we set the initial ellipsoid $\mathcal{E}_0$ which is described as

$$\mathcal{E}_0 \doteq \{z \in \mathbb{R}^n : (z-z_0)^{\mathrm{T}} Q_0^{-1}(z-z_0) \le 1\}.$$

From these parameters, we compute two integers

$$\bar{\ell} \doteq \left\lceil 2(n+1)\ln\frac{\mathrm{Vol}(\mathcal{E}_0)}{\mu}\right\rceil, \qquad \bar{\kappa} \doteq \left\lceil \ln\frac{\bar{\ell}}{\delta}\Big/\ln\frac{1}{1-\varepsilon}\right\rceil \tag{2}$$

where $\mathrm{Vol}(\mathcal{E}_0)$ denotes the volume of the initial ellipsoid $\mathcal{E}_0$.

In the following algorithm, k denotes the k-th iteration and ℓ corresponds to the number of updates, and κ indicates that the ellipsoid has not been updated for consecutive κ random samples. Then, z_k and Q_k are sequentially updated according to the update rule of the ellipsoid method if z_k does not satisfy $v_i(z_k,\theta_k) \le 0$ for all $i \in \{1,2,\ldots,m\}$ for a random sample θ_k. When the ellipsoid has not been updated consecutive $\bar{\kappa}$ times, this algorithm stops with z_k as output. If ℓ reaches $\bar{\ell}$, this algorithm stops indicating infeasibility.

Algorithm 1.

1. Set $k := 0$ and $\ell := 0$.
2. If $\ell \ge \bar{\ell}$, stop the algorithm indicating infeasibility.
3. Set $\kappa := 0$.
4. If $\kappa = \bar{\kappa}$, stop the algorithm with z_k as output.
5. Draw $\theta_k \in \Theta$ according to $\mathcal{P}$.
6. If $v_i(z_k,\theta_k) \le 0$ for all $i \in \{1,2,\ldots,m\}$, set $\kappa := \kappa + 1$.
7. If there exists $i \in \{1,2,\ldots,m\}$ such that $v_i(z_k,\theta_k) > 0$, compute subgradients of v_i with respect to z at z_k and θ_k such that $v_i(z_k,\theta_k) > 0$, select subgradients g_i, $i = 1,2,\ldots,q$ such that $g_i^{\mathrm{T}} Q_k g_j \le 0$ for all $i \ne j$, and define matrix G composed of g_i, $i = 1,2,\ldots,q$ as
$$G \doteq \left[g_1\ g_2 \ldots g_q\right], \qquad 1 \le q < n.$$
Update the candidate of the solution as
$$z_{k+1} := z_k - \gamma Q_k \widetilde{G} e \tag{3}$$
$$Q_{k+1} := \eta\left(Q_k - \sigma Q_k \widetilde{G}\widetilde{G}^{\mathrm{T}} Q_k\right) \tag{4}$$
where
$$\widetilde{G} \doteq G(G^{\mathrm{T}} Q_k G)^{-1/2} \tag{5}$$
$$\alpha \doteq \frac{-(n-2)+\sqrt{(n-2)^2+4(n-q)}}{2(n-q)} \tag{6}$$
$$\gamma \doteq \frac{\alpha}{1+2\alpha} \tag{7}$$
$$\sigma \doteq \frac{2\alpha}{1+2\alpha} \tag{8}$$
$$\eta \doteq \frac{1+2\alpha+q\alpha^2}{1+2\alpha} \tag{9}$$
and e denotes the q dimensional vector all elements of which are equal to 1. Set $\ell := \ell + q$.
8. Set $k := k+1$ and go to Step 2.

Then, a significant feature of the proposed algorithm appears in the procedure at Step 7, where the ellipsoid is updated via a set of given multiple subgradients. On the other hand, in the original stochastic ellipsoid method (Kanev, De Schutter, and Verhaegen, 2003; Oishi, 2007), the update rule uses only one subgradient per one update. Another important difference is to use $\ell := \ell + q$ in the stopping rule, while Oishi (2007) uses $\ell := \ell + 1$.

Remark 1. When $q = 1$, the update rule at Step 7 of Algorithm 1 reduces to the update rule of the original stochastic ellipsoid method (Kanev, De Schutter, and Verhaegen, 2003; Oishi, 2007).

Now, we define the k-th ellipsoid $\mathcal{E}_k$ and half spaces $\mathcal{H}_i$, $i = 1,2,\ldots q$ as

$$\mathcal{E}_k \doteq \{z \in \mathbb{R}^n : (z-z_k)^{\mathrm{T}} Q_k^{-1}(z-z_k) - 1 \le 0\}$$
$$\mathcal{H}_i \doteq \{z \in \mathbb{R}^n : g_i^{\mathrm{T}}(z-z_k) \le 0\}, \qquad i = 1,2,\ldots,q.$$

Then, we obtain the following result.

Theorem 2. The ellipsoids $\mathcal{E}_k$, $\mathcal{E}_{k+1}$ determined by the update rule at Step 7 satisfy

$$\mathcal{E}_k \cap \left(\bigcap_{i=1}^{q} \mathcal{H}_i\right) \subseteq \mathcal{E}_{k+1}. \tag{10}$$

Proof. We first define $\tilde{g}_1,\ldots,\tilde{g}_q$ and W as

$$\left[\tilde{g}_1\ \ldots\ \tilde{g}_q\right] \doteq \left[g_1\ \ldots\ g_q\right] W, \qquad W \doteq (G^{\mathrm{T}} Q_k G)^{-1/2}.$$

Note that $(G^{\mathrm{T}}Q_kG)^{1/2} \succ 0$ is positive definite and thus non-singular. Furthermore, all off diagonal elements of $(G^{\mathrm{T}}Q_kG)^{1/2}$ are non-positive since we select the vectors $g_1, g_2, \ldots, g_q$ such that all off diagonal elements of $G^{\mathrm{T}}Q_kG$ are non-positive (Alefeld and Schneider, 1982). Then, all elements of the inverse matrix W of $(G^{\mathrm{T}}Q_kG)^{1/2}$ are non-negative (Graham (1987), Lemma 5.1).

From this property, we see that for any $z \in \bigcap_{i=1}^{q} \mathcal{H}_i$,

$$\tilde{g}_j^{\mathrm{T}}(z - z_k) = \sum_{i=1}^{q} w_{ij} g_i^{\mathrm{T}}(z - z_k) \le 0. \tag{11}$$

We further define half spaces $\widetilde{\mathcal{H}}_j$ with $\tilde{g}_j$, $j = 1, 2, \ldots, q$ as

$$\widetilde{\mathcal{H}}_j \doteq \{z \in \mathbb{R}^n : \tilde{g}_j^{\mathrm{T}}(z - z_k) \le 0\}.$$

Since $\bigcap_{i=1}^{q} \mathcal{H}_i \subseteq \widetilde{\mathcal{H}}_j$ for all $j = 1, 2, \ldots, q$, we obtain

$$\bigcap_{i=1}^{q} \mathcal{H}_i \subseteq \bigcap_{j=1}^{q} \widetilde{\mathcal{H}}_j. \tag{12}$$

This is the first step of this proof.

Then, we go to the next step. We notice that , for any $z \in \mathcal{E}_k$,

$$\begin{aligned}
0 &\ge (z - z_k)^{\mathrm{T}} Q_k^{-1}(z - z_k) - 1 \\
&= (z - z_k + Q_k\tilde{g}_j - Q_k\tilde{g}_j)^{\mathrm{T}} Q_k^{-1}(z - z_k + Q_k\tilde{g}_j - Q_k\tilde{g}_j) - 1 \\
&= (z - z_k + Q_k\tilde{g}_j)^{\mathrm{T}} Q_k^{-1}(z - z_k + Q_k\tilde{g}_j) \\
&\quad - (z - z_k + Q_k\tilde{g}_j)^{\mathrm{T}}\tilde{g}_j - \tilde{g}_j^{\mathrm{T}}(z - z_k + Q_k\tilde{g}_j) \\
&= (z - z_k + Q_k\tilde{g}_j)^{\mathrm{T}} Q_k^{-1}(z - z_k + Q_k\tilde{g}_j) \\
&\quad - (z - z_k)^{\mathrm{T}}\tilde{g}_j - \tilde{g}_j^{\mathrm{T}}(z - z_k) - 2 \\
&\ge -2\left(\tilde{g}_j^{\mathrm{T}}(z - z_k) + 1\right)
\end{aligned} \tag{13}$$

holds. From (11) and (13), we obtain

$$\begin{aligned}
&(z - z_k)^{\mathrm{T}}\tilde{g}_j(\tilde{g}_j^{\mathrm{T}}(z - z_k) + 1) + ((z - z_k)^{\mathrm{T}}\tilde{g}_j + 1)\tilde{g}_j^{\mathrm{T}}(z - z_k) \\
&\quad = 2(z - z_k)^{\mathrm{T}}\tilde{g}_j\tilde{g}_j^{\mathrm{T}}(z - z_k) + (z - z_k)^{\mathrm{T}}\tilde{g}_j + \tilde{g}_j^{\mathrm{T}}(z - z_k) \\
&\quad \le 0.
\end{aligned} \tag{14}$$

Now, we define strips $\widetilde{\mathcal{L}}_j$, $j = 1, 2, \ldots, q$ as

$$\begin{aligned}
\widetilde{\mathcal{L}}_j \doteq \{z \in \mathbb{R}^n :\ & 2(z - z_k)^{\mathrm{T}}\tilde{g}_j\tilde{g}_j^{\mathrm{T}}(z - z_k) \\
& + (z - z_k)^{\mathrm{T}}\tilde{g}_j + \tilde{g}_j^{\mathrm{T}}(z - z_k) \le 0\}.
\end{aligned}$$

Then we see that

$$\mathcal{E}_k \cap \left(\bigcap_{i=1}^{q} \widetilde{\mathcal{H}}_i\right) = \mathcal{E}_k \cap \left(\bigcap_{i=1}^{q} \widetilde{\mathcal{L}}_i\right). \tag{15}$$

Finally, we move on to the last step of the proof. We notice that, for any $z \in \mathcal{E}_k \cap \left(\bigcap_{j=1}^{q} \widetilde{\mathcal{L}}_j\right)$,

$$\begin{aligned}
&(z - z_k)^{\mathrm{T}} Q_k^{-1}(z - z_k) - 1 \\
&\quad + r\sum_{j=1}^{q}\left(2(z - z_k)^{\mathrm{T}}\tilde{g}_j\tilde{g}_j^{\mathrm{T}}(z - z_k) + (z - z_k)^{\mathrm{T}}\tilde{g}_j + \tilde{g}_j^{\mathrm{T}}(z - z_k)\right) \\
&\le 0,
\end{aligned}$$

where r is any positive number. Thus, we set $r = \alpha > 0$. Applying (3), and (4) to the above inequality, we can derive the updated ellipsoid $\mathcal{E}_{k+1}$ as

$$\begin{aligned}
&(z - z_k)^{\mathrm{T}} Q_k^{-1}(z - z_k) - 1 \\
&\quad + \alpha\sum_{j=1}^{q}\left(2(z - z_k)^{\mathrm{T}}\tilde{g}_j\tilde{g}_j^{\mathrm{T}}(z - z_k) + (z - z_k)^{\mathrm{T}}\tilde{g}_j + \tilde{g}_j^{\mathrm{T}}(z - z_k)\right) \\
&= (z - z_k)^{\mathrm{T}} Q_k^{-1}(z - z_k) - 1 \\
&\quad + \alpha\left(2(z - z_k)^{\mathrm{T}}\widetilde{G}\widetilde{G}^{\mathrm{T}}(z - z_k) + (z - z_k)^{\mathrm{T}}\widetilde{G}e + e^{\mathrm{T}}\widetilde{G}^{\mathrm{T}}(z - z_k)\right) \\
&= (z - z_k)^{\mathrm{T}}\ Q_k^{-1}(z - z_k) + \gamma^2\,(1 + 2\alpha)\, e^{\mathrm{T}}e - \eta \\
&\quad + 2\alpha(z - z_k)^{\mathrm{T}}\widetilde{G}\widetilde{G}^{\mathrm{T}}(z - z_k) \\
&\quad + \gamma\,(1 + 2\alpha)\,(z - z_k)^{\mathrm{T}}\widetilde{G}e + \gamma\,(1 + 2\alpha)\, e^{\mathrm{T}}\widetilde{G}^{\mathrm{T}}(z - z_k) \\
&= \left(z - z_k + \gamma Q_k\widetilde{G}e\right)^{\mathrm{T}}\left(Q_k^{-1} + 2\alpha\widetilde{G}\widetilde{G}^{\mathrm{T}}\right)\left(z - z_k + \gamma Q_k\widetilde{G}e\right) - \eta \\
&= \left(z - z_k + \gamma Q_k\widetilde{G}e\right)^{\mathrm{T}}\left(Q_k - \sigma Q_k\widetilde{G}\widetilde{G}^{\mathrm{T}}Q_k\right)^{-1}\left(z - z_k + \gamma Q_k\widetilde{G}e\right) \\
&\quad - \eta \\
&= \eta\left((z - z_{k+1})^{\mathrm{T}} Q_{k+1}^{-1}(z - z_{k+1}) - 1\right).
\end{aligned} \tag{16}$$

With (12), (15), (16), and $\eta > 0$, we conclude that

$$\mathcal{E}_k \cap \left(\bigcap_{i=1}^{q} \mathcal{H}_i\right) \subseteq \mathcal{E}_k \cap \left(\bigcap_{j=1}^{q} \widetilde{\mathcal{H}}_j\right) = \mathcal{E}_k \cap \left(\bigcap_{j=1}^{q} \widetilde{\mathcal{L}}_j\right) \subseteq \mathcal{E}_{k+1}.$$

□

This theorem shows that the proposed update rule appropriately works. That is, the updated ellipsoid $\mathcal{E}_{k+1}$ contains the intersection of the previous ellipsoid $\mathcal{E}_k$ and the half spaces $\mathcal{H}_i$, $i = 1, 2, \ldots, q$ which involve the solution set $\mathcal{S}$. This property is necessary for ellipsoid method.

Another key issue is that the volume of $\mathcal{E}_{k+1}$ should be less than that of $\mathcal{E}_k$. Here, we prove that this property actually holds showing a relation between the volume of the updated ellipsoid and the number of given subgradients.

Theorem 3. For given $n \in \{1, 2, \ldots\}$, $q \in \{1, 2, \ldots, n-1\}$,

$$\frac{\mathrm{Vol}(\mathcal{E}_{k+1})}{\mathrm{Vol}(\mathcal{E}_k)} = f(n, q) < 1 \tag{17}$$

holds, where $\mathcal{E}_k$ and $\mathcal{E}_{k+1}$ denote the original ellipsoid and the updated ellipsoid respectively,

$$f(n, q) \doteq \sqrt{\frac{(1 + 2\alpha + q\alpha^2)^n}{(1 + 2\alpha)^{q+n}}} \tag{18}$$

and α is given by (6).

Moreover, a given $q \in \{1, 2, \ldots, n-2\}$,

$$f(n, q) > f(n, q+1) \tag{19}$$

holds. For a given $q \in \{2, 3, \ldots, n-1\}$,

$$f(n, q) < (f(n, 1))^q \tag{20}$$

holds.

Proof. We first show the statement (17). The ratio of the volume of the updated ellipsoid $\mathcal{E}_{k+1}$ to that of the original ellipsoid $\mathcal{E}_k$ is given by

$$\begin{aligned}
\frac{\mathrm{Vol}(\mathcal{E}_{k+1})}{\mathrm{Vol}(\mathcal{E}_k)} &= \sqrt{\frac{\det\left[\eta\left(Q_k - \sigma Q_k\widetilde{G}\widetilde{G}^{\mathrm{T}}Q_k\right)\right]}{\det[Q_k]}} \\
&= \sqrt{\det[\eta I]\det[I - \sigma\widetilde{G}^{\mathrm{T}}Q_k\widetilde{G}]} \\
&= \sqrt{\frac{(1 + 2\alpha + q\alpha^2)^n}{(1 + 2\alpha)^{q+n}}} \\
&= f(n, q).
\end{aligned}$$

Now, we regard n as a constant and define the function $\tilde{f}(q,\alpha)$ of $q \in \mathbb{R}$, $q \geq 0$ and $\alpha \in \mathbb{R}$, $\alpha \geq 0$ as

$$\tilde{f}(q,\alpha) \doteq n\log(1+2\alpha+q\alpha^2) - (q+n)\log(1+2\alpha). \quad (21)$$

Since $f(n,q)$ and $\tilde{f}(q,\alpha)$ have the same order relation for α determined by $q \in \{1,2,\ldots,n-1\}$ and (6), we will use $\tilde{f}(q,\alpha)$ instead of $f(n,q)$ to show statements (17), (19), and (20) in the following discussion.

Notice first that the partial derivative of $\tilde{f}(q,\alpha)$ with respect to α is given by

$$\frac{\partial}{\partial\alpha}\tilde{f}(q,\alpha) = 2q\frac{(n-q)\alpha^2+(n-2)\alpha-1}{(1+2\alpha+q\alpha^2)(1+2\alpha)}.$$

Thus, the positive stationary point of $\tilde{f}(q,\alpha)$ is

$$\alpha = \frac{-(n-2)+\sqrt{(n-2)^2+4(n-q)}}{2(n-q)}.$$

Since $n-q>0$, $\tilde{f}(q,\alpha)$ is minimized by this stationary point. Noting that $f(n,q)=1$ at $\alpha=0$ and this stationary point is equal to α given by (6), we see that the statement (17) holds.

To show the statement (19), let

$$\alpha_q \doteq \frac{-(n-2)+\sqrt{(n-2)^2+4(n-q)}}{2(n-q)},$$

where $n>q>0$.

We first show that

$$\tilde{f}(q,\alpha_q) \geq \tilde{f}(q+\tilde{q},\alpha_q) \quad (22)$$

for any $\tilde{q} \in (0,n-q)$. We differentiate $\tilde{f}(q+\tilde{q},\alpha_q)$ with respect to $\tilde{q}$ and obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tilde{q}}\tilde{f}(q+\tilde{q},\alpha_q) &= \left(\frac{n\alpha_q^2}{1+2\alpha_q+(q+\tilde{q})\alpha_q^2} - \log(1+2\alpha_q)\right)\\ &\leq \left(\frac{n\alpha_q^2}{1+2\alpha_q+(q+\tilde{q})\alpha_q^2} - \frac{2\alpha_q}{1+2\alpha_q}\right)\\ &= \frac{-\tilde{q}\alpha_q^3-n\alpha_q^2}{(1+2\alpha_q+(q+\tilde{q})\alpha_q^2)(1+2\alpha_q)}\\ &< 0,\end{aligned}$$

where we use Lemma 6 (See Appendix A) and the identity

$$(n-q)\alpha_q^2+(n-2)\alpha_q-1=0.$$

This inequality implies (22).

Following the proof of the statement (17), we can see that, for any fixed $q<n$, $\tilde{f}(q,\alpha)$ takes the minimum at $\alpha=\alpha_q$. Thus, if $\alpha_{q+1}\neq\alpha_q$, we have

$$\tilde{f}(q+1,\alpha_q) > \tilde{f}(q+1,\alpha_{q+1}).$$

From this inequality and (22), we see

$$\tilde{f}(q,\alpha_q) \geq \tilde{f}(q+1,\alpha_q) > \tilde{f}(q+1,\alpha_{q+1}),$$

which meets (19).

To show the statement (20), we first show that

$$\tilde{f}(q,\alpha_1) - q\tilde{f}(1,\alpha_1) < 0, \qquad q \in (1,n) \quad (23)$$

holds. We differentiate the right hand side of (23) with respect to q and substitute $\alpha_1 = 1/(n-1)$. Then, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\tilde{f}(q,\alpha_1)-q\tilde{f}(q,\alpha_1)\right) = \frac{n}{n^2+q-1} + n\log\left(1-\frac{1}{n^2}\right).$$

Furthermore, we can derive

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\tilde{f}(q,\alpha_1)-q\tilde{f}(1,\alpha_1)\right) < \frac{n}{n^2} + n\left(-\frac{1}{n^2}\right) = 0,$$

where we use $n^2+q-1>n^2$ and Lemma 6 (See Appendix A). Since $\tilde{f}(q,\alpha_1)-q\tilde{f}(q,\alpha_1)=0$ at $q=1$, this inequality implies (23).

Following the proof of the statement (17), we can see that $\tilde{f}(q,\alpha)$ takes the minimum at $\alpha=\alpha_q$ for any fixed $q<n$. Thus, if $\alpha_1\neq\alpha_q$ and $q\neq 1$, we have

$$\tilde{f}(q,\alpha_q) < \tilde{f}(q,\alpha_1).$$

From this inequality and (23), we obtain

$$\tilde{f}(q,\alpha_q) < \tilde{f}(q,\alpha_1) < q\tilde{f}(1,\alpha_1)$$

which meets (20). □

The statement (17) of this theorem means that we can always obtain the updated ellipsoid of which volume is smaller than that of the original ellipsoid. Furthermore, the statement (19) shows that if we use many subgradients , the updated ellipsoid has a smaller volume. We also remark that, from the statement (20), the volume of the ellipsoid updated with q subgradients is smaller than the volume of the ellipsoid updated q times with one subgradient. This property enables us to replace $\ell := \ell+1$ with $\ell := \ell+q$ at Step 7 in Algorithm 1.

Remark 4. We can replace $(G^{\mathrm{T}}Q_kG)^{1/2}$ with $\tilde{W}$ at Step 7 in Algorithm 1, where $\tilde{W}$ is a triangular matrix which satisfies $\tilde{W}^{\mathrm{T}}\tilde{W} = G^{\mathrm{T}}Q_kG$. This is because the statement (10) of Theorem 2 still holds since all off diagonal elements of $\tilde{W}$ are nonnegative and all diagonal elements of $\tilde{W}$ are positive (Graham (1987), Lemma 5.7).

Finally, we summarize the properties of Algorithm 1, where we see that several useful properties of the stochastic ellipsoid method (Oishi, 2007) are preserved.

Theorem 5.

(1) The number of random samples is less than or equal to $\bar{\kappa}\bar{\ell}$.
(2) The number of updates of ellipsoid is less than or equal to $\bar{\ell}$.
(3) When Algorithm 1 stops at Step 4 with z_k as output, the probability that satisfies

$$\mathcal{P}\{\theta\in\Theta : v_i(z_k,\theta)\leq 0, \forall i\in\{1,2,\ldots,m\}\} \leq 1-\varepsilon$$

is less than or equal to δ.
(4) When Algorithm 1 stops at Step 2 indicating infeasibility,

$$\mathrm{Vol}(\mathcal{S}\cap\mathcal{E}_0) < \mu$$

holds.

Proof. The statements (1) and (2) follow from the construction of Algorithm 1.

The proof of the statement (3) is similar to the literature (Oishi, 2007). We first define two events, F_ℓ and B_ℓ:

F_ℓ : The number of updates of ellipsoid reaches ℓ, then the algorithm stops at Step 5 with z_ℓ as output.

B_ℓ : The random sample z_ℓ satisfies

$$\mathcal{P}\{\theta\in\Theta : v_i(z_\ell,\theta)\leq 0, i\in\{1,2,\ldots,m\}\} \leq 1-\varepsilon.$$

That is, z_ℓ is not a probabilistic solution with a given accuracy ε.

Our aim is to show the probability that the output z_ℓ of the algorithm is not a probabilistic solution is less than or equal to δ for any $\ell = 1,2,\ldots,\bar{\ell}$. That is, we show that

$$\mathcal{P}^{\infty}\left\{\bigcup_{\ell=1}^{\bar{\ell}}(F_\ell \cap B_\ell)\right\} \leq \delta$$

holds.

Notice that, for each $\ell = 1, 2, \ldots, \bar{\ell}$,

$$\mathcal{P}^{\infty}\{F_\ell \cap B_\ell\} \leq \mathcal{P}^{\bar{\kappa}}\{F_\ell | B_\ell\}\mathcal{P}^{\bar{\kappa}}\{B_\ell\} \leq \mathcal{P}^{\bar{\kappa}}\{F_\ell | B_\ell\} \leq (1-\varepsilon)^{\bar{\kappa}}$$

holds. Since the selection (2) implies

$$\bar{\ell}(1-\varepsilon)^{\bar{\kappa}} \leq \delta,$$

we obtain

$$\mathcal{P}^{\infty}\left\{\bigcup_{\ell=1}^{\bar{\ell}}(F_\ell \cap B_\ell)\right\} \leq \bar{\ell}(1-\varepsilon)^{\bar{\kappa}} \leq \delta$$

which shows the statement (3) of Theorem 5.

Finally, we show the statement (4). The equation (20) shows that

$$\mathrm{Vol}(\mathcal{E}_{\ell+1}) < \exp\left(-\frac{q}{2(n+1)}\right)\mathrm{Vol}(\mathcal{E}_\ell)$$

when we update the ellipsoid with q subgradients. Therefore, if $\ell \geq \bar{\ell}$, the above inequality and Theorem 2 imply

$$\mathrm{Vol}(\mathcal{S} \cap \mathcal{E}_0) < \mathrm{Vol}(\mathcal{E}_{\bar{\ell}}) < \exp\left(-\frac{\bar{\ell}}{2(n+1)}\right)\mathrm{Vol}(\mathcal{E}_0) \leq \mu.$$

□

This theorem shows that Algorithm 1 always stops in a finite number of iterations. When the algorithm stops at Step 4 with z_k as output, the probability that the obtained z_k is a probabilistic solution with ε is greater than $1-\delta$. When the algorithm stops at Step 2 indicating infeasibility, the volume of intersection between the solution set $\mathcal{S}$ and the given initial ellipsoid $\mathcal{E}_0$ is less than or equal to μ. That is, we see that the solution set is too small to find a solution.

4. NUMERICAL EXAMPLES

In this section, we illustrate the effectiveness of the proposed algorithm. We here consider $\mathcal{H}_\infty$ control problem for linear parameter-varying (LPV) systems, which is taken from Fujisaki, Dabbene, and Tempo (2003).

Let us consider an LPV system

$$\begin{aligned}\dot{x}(t) &= A(\theta(t))x(t) + B_1(\theta(t))d(t) + B_2(\theta(t))u(t)\\ e(t) &= C_1(\theta(t))x(t) + D_{12}(\theta(t))u(t)\\ y(t) &= C_2(\theta(t))x(t) + D_{21}(\theta(t))d(t)\end{aligned}$$

and an LPV controller

$$\begin{aligned}\dot{x}_c(t) &= A_c(\theta(t))x_c(t) + B_c(\theta(t))y(t)\\ u(t) &= C_c(\theta(t))x_c(t)\end{aligned}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the state, $d(t) \in \mathbb{R}^{n_d}$ is the disturbance, $u(t) \in \mathbb{R}^{n_u}$ is the control input, $e(t) \in \mathbb{R}^{n_e}$ is the controlled output, $y(t) \in \mathbb{R}^{n_y}$ is the measurement output, and $\theta(t) \in \Theta$ is the scheduling parameter. We assume that

$$D_{12}^{\mathrm{T}}(\theta)\begin{bmatrix} C_1(\theta) & D_{12}(\theta)\end{bmatrix} = \begin{bmatrix} 0 & I\end{bmatrix}$$
$$\begin{bmatrix} B_1(\theta) \\ D_{21}(\theta)\end{bmatrix} D_{21}^{\mathrm{T}}(\theta) = \begin{bmatrix} 0 \\ I\end{bmatrix} \qquad \forall\theta \in \Theta.$$

It is known that if $X = X^{\mathrm{T}}$ and $Y = Y^{\mathrm{T}}$ satisfy the matrix inequalities

$$V_1(X,\theta) \leq 0, \quad V_2(Y,\theta) \leq 0, \quad V_3(X,Y,\theta) \leq 0 \quad \forall\theta \in \Theta,$$

there exists an LPV controller such that the closed loop system is quadratically stable and the upper bound of $\mathcal{L}_2$ gain from $d(t)$ to $e(t)$ is less than or equal to γ, where

$$\begin{aligned}V_1(X,\theta) &\doteq A(\theta)X + XA^{\mathrm{T}}(\theta) + XC_1^{\mathrm{T}}(\theta)C_1(\theta)X\\ &\quad + \gamma^{-2}B_1(\theta)B_1^{\mathrm{T}}(\theta) - B_2(\theta)B_2^{\mathrm{T}}(\theta) + \epsilon I\\ V_2(Y,\theta) &\doteq A^{\mathrm{T}}(\theta)Y + YA(\theta) + YB_1(\theta)B_1^{\mathrm{T}}(\theta)Y\\ &\quad + \gamma^{-2}C_1^{\mathrm{T}}(\theta)C_1(\theta) - C_2^{\mathrm{T}}(\theta)C_2(\theta) + \epsilon I\\ V_3(X,Y) &\doteq -\begin{bmatrix} X & \gamma^{-1}I \\ \gamma^{-1}I & Y\end{bmatrix}\end{aligned}$$

and $\varepsilon > 0$ is a given small constant.

For this problem, if we introduce

$$v(z,\theta) \doteq \left\|\begin{bmatrix} [V_1(X,\theta)]^+ & 0 & 0 \\ 0 & [V_2(Y,\theta)]^+ & 0 \\ 0 & 0 & [V_3(X,Y)]^+\end{bmatrix}\right\| \tag{24}$$

we have a robust feasibility problem (1) with $m = 1$, where z is a vector composed of the variables in X and Y, $\|\cdot\|$ denotes Frobenius norm, and $[\cdot]^+$ denotes projection onto the cone of positive semidefinite matrix. On the other hand, if we introduce

$$\begin{aligned}v_1(z,\theta) &\doteq \|[V_1(X,\theta)]^+\|\\ v_2(z,\theta) &\doteq \|[V_2(Y,\theta)]^+\|\\ v_3(z) &\doteq \|[V_3(X,Y)]^+\|,\end{aligned} \tag{25}$$

the LPV control problem reduces to a robust feasibility problem (1) with $m = 3$. In fact, we see that z satisfies the condition

$$v(z,\theta) \leq 0 \qquad \forall\theta \in \Theta$$

if and only if z satisfies the multiple conditions

$$v_1(z,\theta) \leq 0, \quad v_2(z,\theta) \leq 0, \quad v_3(z) \leq 0 \quad \forall\theta \in \Theta.$$

To illustrate the effectiveness of the proposed algorithm, we solved the robust feasibility problem with three constraints (25) and that with one constraint (24), then compared two results. The former problem was solved by the randomized algorithm proposed by Oishi (2007).

In this numerical example, we set the coefficients of the state space equation as

$$A(\theta) = \begin{bmatrix} 0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{bmatrix} + \begin{bmatrix} \theta_1&\theta_5&\theta_9&\theta_{13}\\ \theta_2&\theta_6&\theta_{10}&\theta_{14}\\ \theta_3&\theta_7&\theta_{11}&\theta_{15}\\ \theta_4&\theta_8&\theta_{12}&\theta_{16}\end{bmatrix} = A_0 + \Delta$$

$$B_1 = \begin{bmatrix} B_2 & 0\end{bmatrix}, \quad B_2 = \begin{bmatrix} 0&1&0&0\\0&0&0&1\end{bmatrix}^{\mathrm{T}}$$

$$C_1 = \begin{bmatrix} C_2 \\ 0\end{bmatrix}, \quad C_2 = \begin{bmatrix} 1&0&0&0\\0&0&1&0\end{bmatrix}$$

$$D_{12} = \begin{bmatrix} 0 & I\end{bmatrix}^{\mathrm{T}}, \quad D_{21} = \begin{bmatrix} 0 & I\end{bmatrix}$$

$$\Theta = \{\Delta^{\mathrm{T}}\Delta \leq 10^{-2}I\}$$

$$n_x = 4, \quad n_d = n_e = 4, \quad n_u = n_y = 2.$$

Thus, we see $z \in \mathbb{R}^{20}$ from $X \in \mathbb{R}^{5\times5}$ and $Y \in \mathbb{R}^{5\times5}$. The above setting gives $f(20,1) = 0.975$, $f(20,2) = 0.951$ and $f(20,3) = 0.927$. From these values, we can expect this algorithm fast convergence and reduce the computational time if updates with multiple cuts occur many times.

We chose several values of γ and executed Algorithms 1 100 times, where we chose the probability measure $\mathcal{P}$ as uniform distribution on the set Θ (Calafiore et al., 2000), $\delta = 0.01$, $\varepsilon = 0.01$, $\mu = 1$, and the initial ellipsoid $\mathcal{E}_0$ by letting $z_0 = 0$,

Table 1. Average numbers of random samples and average computational times with each γ

γ	1.5	2.5	3.5	4.5	5.5
$k^{\mathrm{m}}_{\mathrm{ave}}$	1,100.4	1,140.9	1,228.8	1,975.2	1,465.1
$r^{\mathrm{m}}_{\mathrm{feas}}$	0.00	0.00	0.00	1.00	1.00
$t^{\mathrm{m}}_{\mathrm{ave}}$ [s]	6.91	7.16	7.61	9.64	7.06
$k^{\mathrm{s}}_{\mathrm{ave}}$	1,782.0	1,782.0	1,825.2	2,753.6	1,714.9
$r^{\mathrm{s}}_{\mathrm{feas}}$	0.00	0.00	0.00	1.00	1.00
$t^{\mathrm{s}}_{\mathrm{ave}}$ [s]	10.19	10.14	10.36	13.17	8.06

$Q_0 = 100I$. The result is shown in Table 1, where k_{ave} denotes the average number of random samples,

$$r_{\mathrm{feas}} \doteq \frac{\text{The number of trials terminating at Step 4}}{\text{The number of trials } (= 100)},$$

and t_{ave} denotes the average computational time measured on a PC with Pentium 4 3.0 GHz and 2.0 GByte memory. The superscripts 's' and 'm' to k_{ave}, r_{feas} and t_{ave} correspond to the problem (24) and the problem (25) respectively.

Table 1 says that the proposed algorithm generally performed better than the original algorithm did. Especially, this result is consistent with the following observation. If the feasibility set is very small, a candidate z fails to satisfy $v(z,\theta)$ for almost all $\theta \in \Theta$. This could be also true even if we replace $v(z,\theta)$ with $v_i(z,\theta)$, $i = 1,2,\dots,m$. Thus, updates with multiple cuts could occur many times, which leads to fast convergence and reduction of the computational time.

5. CONCLUSION

In this paper, we have proposed a new update rule of ellipsoid with multiple cuts in the context of randomized algorithms, and have investigated its performance. In particular, we have shown that we can always obtain a smaller ellipsoid by the proposed algorithm. Then, we have studied the performance of the proposed algorithm, which enables us to modify the stopping rule of the stochastic ellipsoid method. We have demonstrated this effect through a numerical example.

Recently, Calafiore and Dabbene (2007) proposed the probabilistic analytic center cutting plane method which is a randomized algorithm similar to the stochastic ellipsoid method discussed in this paper. We remark that a "multiple cuts" technique could be developed in that context. In this regard, a deterministic analytic center cutting plane method employing multiple subgradients has been developed by Ye (1997).

ACKNOWLEDGEMENTS

This work was supported in part by the Grant-in-Aid for Scientific Research (C), 17560395, Japan Society for the Promotion of Science.

REFERENCES

G. Alefeld and N. Schneider. On square roots of M-matrices. *Linear Algebra and Its Applications*, 42:119–132, 1982.

R. G. Bland, D. Goldfarb, and M. J. Todd. The ellipsoid method: A survey. *Operations Research*, 29(6):1039–1091, 1981.

G. Calafiore and F. Dabbene. A probabilistic analytic center cutting plane method for feasibility of uncertain LMIs. *Automatica*, 43(12):2022–2033, 2007.

G. Calafiore, F. Dabbene, and R. Tempo. Randomized algorithms for probabilistic robustness with real and complex structured uncertainty. *IEEE Transactions on Automatic Control*, 45(12):2218–2235, 2000.

A. Ech-Cherif and J. Ecker. A class of rank-two ellipsoid algorithms for convex programming. *Mathematical Programing*, 29:187–202, 1984.

Y. Fujisaki, F. Dabbene, and R. Tempo. Probabilistic design of LPV control systems. *Automatica*, 39(8):1323–1337, 2003.

D. Goldfarb and M. J. Todd. Modifications and implemantation of the ellipsoid algorithm for linear programming. *Mathematical Programing*, 23(1):1–19, 1982.

A. Graham. *Nonnegative matrices and applicable topics in linear algebra*. Halsted Press, New York, 1987.

S. Kanev, B. De Schutter, and M. Verhaegen. An ellipsoid algorithm for probabilistic robust controller design. *Systems & Control Letters*, 49(5):365–375, 2003.

Y. Oishi. Polynomial-time algorithms for probabilistic solutions of parameter-dependent linear matrix inequalities. *Automatica*, 43(3):538–545, 2007.

N. Shor and V. Gershovich. Family of algorithms for solving convex programming problems. *Kibernetika*, 15(4):62–67, 1979. translated in *Cybernetics*, 15(4):502–507 1979.

R. Tempo, G. Calafiore, and F. Dabbene. *Randomized Algorithms for Analysis and Control of Uncertain Systems*. Springer-Verlag, London, 2004.

T. Wada and Y. Fujisaki. Efficient randomized algorithms for robust feasibility problems. *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, 2251–2254, 2006.

T. Wada and Y. Fujisaki. Probabilistic design of switched systems. *Proceedings of European Control Conference 2007*, 2611–2618, 2007.

Y. Ye. Complexity analysis of the analytic center cutting plane method that uses multiple cuts. *Mathematical Programming*, 78(1):85–104, 1997.

Appendix A. COMPLEMENTS TO THEOREM 2

Lemma 6.

(1) For any $x \in (0,1)$,

$$\log \frac{1}{1-x} - x \geq 0 \tag{A.1}$$

holds.

(2) For any $x \in (-1,\infty)$,

$$\log(1+x) - x \leq 0 \tag{A.2}$$

holds.

Proof. Since $\exp(-x)$ is convex in x, we have

$$\exp(-x) \geq 1 - x.$$

Then, for any $x \in (0,1)$, we see

$$x \geq \log \frac{1}{1-x},$$

which implies (A.1).

Similarly, since $\exp(x)$ is convex in x, we have

$$\exp(x) \geq 1 + x.$$

Then, for any $x \in (-1,\infty)$, we see

$$x \geq \log(1+x),$$

which implies (A.2). □