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

Distribution-Dependent Robust Linear Optimization with Asymmetric Uncertainty and Application to Optimal Control ⋆

Ioannis Ch. Paschalidis * **Seong-Cheol Kang** ** **Keyong Li** **

* *Center for Information & Systems Eng., Boston University, Brookline, MA 02446, USA (yannisp@bu.edu, http://ionia.bu.edu/)*
** *Boston University, USA (jsckang@bu.edu; likeyong@ieee.org)*

Abstract: We consider a linear programming problem in which the constraint matrix is uncertain. Each element of the constraint matrix is modeled as a random variable whose range is asymmetrically bounded around its mean. We construct a formulation that yields a solution with a better objective value, compared to the classical robust optimization approach, while taking the risk that the solution may become infeasible to the original problem. We address the risk by establishing upper bounds on the probability that it violates the constraints of the problem. These bounds exploit full distributional information on the random elements or limited distributional information such as the true means or sample means of the random elements. We explore the application of our methodology to the optimal control of linear uncertain systems with constraints.

1. INTRODUCTION

When real-world applications are formulated as mathematical optimization problems, sometimes problem data are subject to uncertainty due to their random nature, measurement errors, or other reasons. In this case, ignoring data uncertainty and solving an optimization problem with fixed (nominal) values for the uncertain data often results in a non-robust solution, in the sense that even small changes in the problem data can easily render the solution infeasible. For some applications, such a solution could be totally useless.

Robust optimization is a methodology for dealing with optimization problems with data uncertainty. In the classical sense, the goal of robust optimization is to find a "safe" solution that is immune to data uncertainty (i.e., a solution with guaranteed feasibility). Soyster [1973] is an early work on this topic, and more recent results include El Ghaoui et al. [1998], Ben-Tal and Nemirovski [1998, 1999, 2002].

Clearly the classical robust optimization approach is appropriate for applications where infeasibility of a solution cannot be accepted at all (e.g., design of engineering structures like bridges, dams, tunnels, etc.). When an application can tolerate a small chance of infeasibility, however, a solution from this approach tends to be too conservative. For the latter class of applications, methods that produce a less conservative solution with a certain probabilistic guarantee of feasibility could prove to be useful.

Toward this end, a relaxed robust optimization approach has recently emerged. In essence, this approach produces a solution with an improved objective value by taking into consideration only partial realizations of the uncertain data in the optimization process, while taking a certain risk of the solution becoming infeasible. Bertsimas and Sim [2004] considered a linear programming (LP) problem in which each element of the constraint matrix is modeled as a random variable. The range of each random element was assumed to be *symmetrically bounded* around its mean; in other words, the mean is equal to the midpoint of the range. They assumed that the probability distributions of the random elements are unknown except that they are symmetric. They proposed a formulation that excludes certain realizations of the random elements. As such, the solution of the formulation is not guaranteed to be feasible to the original LP problem. Instead, they endowed the solution with a probabilistic guarantee of feasibility by establishing an upper bound on the probability that it violates the constraints of the problem.

Paschalidis and Kang [2005, 2006] improved the results of Bertsimas and Sim [2004] by exploiting distributional information on the random elements. They developed several new upper bounds on the constraint violation probability, which make use of full or limited distributional information. They showed that these bounds are stronger than the one given by Bertsimas and Sim [2004]. This result is significant for real-world applications because stronger bounds leads to improved solutions, under the same probabilistic guarantee of feasibility. This was numerically demonstrated through an inventory control problem with quality of service constraints in Paschalidis and Kang [2005].

The work of Bertsimas and Sim [2004] and Paschalidis and Kang [2005, 2006] both assumed *symmetric data uncertainty*: the ranges of the random elements are symmetrically bounded around their means. This assumption is natural and satisfactory when the probability distributions of the random elements are symmetric over their ranges (e.g., uniform distributions or truncated normal distributions). Moreover, it renders convenience when constructing the formulation and establishing bounds on the constraint vi-

⋆ Research partially supported by the NSF under grants DMI-0330171, CNS-0435312, ECS-0426453, EFRI-0735974, and by the DOE under grant DE-FG52-06NA27490.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC

10.3182/20080706-5-KR-1001.3759

olation probability. In many real-world applications, however, it is often found that the probability distributions of random data are asymmetric over asymmetrically bounded ranges. This paper deals with robust optimization of LP problems under this *asymmetric data uncertainty.*

In Section 2, we consider an LP problem in which each element of the constraint matrix is a random variable whose range is asymmetrically bounded around its mean. We construct a formulation that yields a solution with a better objective value compared to the classical robust optimization approach. Section 3 is concerned with the probability that the solution of the formulation may become infeasible to the original LP problem by violating its constraints. We establish an upper bound on the constraint violation probability, which uses full distributional information on the random elements. We also derive a bound that requires only limited distributional information, namely the means of the random elements and their range information. Sometimes the true means of the random elements are not known, but sample means are available instead. For this case, we develop another bound that utilizes the sample means. Unlike other bounds, this one holds with a certain probability. After we conduct brief numerical tests in Section 4, we explore the application of our methodology to the optimal control of linear uncertain systems with constraints in Section 5.

2. THE ROBUST FORMULATION FOR ASYMMETRIC DATA UNCERTAINTY

Consider the LP problem with data uncertainty

$$\begin{aligned} \max\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \mathbf{A}\mathbf{x} \le \mathbf{b} \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}, \end{aligned} \tag{1}$$

where $\mathbf{c}, \mathbf{l}, \mathbf{u} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} = (a_{ij})$ is an $m \times n$ matrix, $\mathbf{x} \in \mathbb{R}^n$ is the vector of decision variables, and $\mathbf{c}'$ denotes the transpose of the vector $\mathbf{c}$. We assume, without loss of generality, that only the elements of the matrix $\mathbf{A}$ are subject to uncertainty. It is assumed that $a_{ij} \in [a_{ij}^L, a_{ij}^U]$, where $a_{ij}^L \le a_{ij}^U$. Let $\overline{a}_{ij} \triangleq E[a_{ij}]$, and define the *forward deviation* as $d_{ij}^F = a_{ij}^U - \overline{a}_{ij}$ and the *backward deviation* as $d_{ij}^B = \overline{a}_{ij} - a_{ij}^L$. Using these deviations, we can rewrite $a_{ij} \in [\overline{a}_{ij} - d_{ij}^B, \overline{a}_{ij} + d_{ij}^F]$. For each row i of $\mathbf{A}$, we define $J_i \triangleq \{j \mid a_{ij}^L < a_{ij}^U\}$, i.e., $J_i \triangleq \{j \mid a_{ij} \text{ is random}\}$. We assume that a_{ij}, for all i and $j \in J_i$, are independent random variables.

The classical robust optimization approach seeks a maximizing $\mathbf{x}$ that is also guaranteed to satisfy the constraints for all realizations of $\mathbf{A}$. Such an $\mathbf{x}$ is obtained by solving the formulation

$$\begin{aligned} z_F = \max\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \max\nolimits_{\mathbf{a}_i \in \mathscr{U}_i} \{\mathbf{a}_i'\mathbf{x}\} \le b_i, \quad \forall i \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}, \end{aligned} \tag{2}$$

where the *uncertainty set* for the ith row, $\mathscr{U}_i$, is given by

$$\mathscr{U}_i \triangleq \{\mathbf{a}_i \mid a_{ij} \in [\overline{a}_{ij} - d_{ij}^B, \overline{a}_{ij} + d_{ij}^F], \forall j\}.$$

We refer to the formulation (2) as the *classical robust formulation* or *"fat" formulation.* It is not difficult to show that (2) can be rewritten as the LP formulation

$$\begin{aligned} z_F = \max\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \textstyle\sum_j \overline{a}_{ij} x_j + \sum_{j \in J_i} y_{ij} \le b_i, \quad \forall i \\ & y_{ij} \ge d_{ij}^F x_j, \quad \forall i, \forall j \in J_i \\ & y_{ij} \ge -d_{ij}^B x_j, \quad \forall i, \forall j \in J_i \\ & y_{ij} \ge 0, \quad \forall i, \forall j \in J_i \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}. \end{aligned} \tag{3}$$

An optimal solution (and other feasible solutions as well) of (2) never violates the constraints $\mathbf{A}\mathbf{x} \le \mathbf{b}$ of (1). This guaranteed feasibility, however, comes with a degradation of objective value. We are interested in another formulation that yields a solution with a better objective value, in exchange for a possibility of the solution becoming infeasible. To that end, for each row i of $\mathbf{A}$, we introduce a parameter, called the *uncertainty budget*, $\Gamma_i \in [0, |J_i|]$. We then define the *restricted uncertainty set* $\mathscr{R}_i(\Gamma_i) \subseteq \mathscr{U}_i$ as

$$\begin{aligned} \mathscr{R}_i(\Gamma_i) \triangleq \{\mathbf{a}_i \mid a_{ij} \in [\overline{a}_{ij} - \beta_{ij} d_{ij}^B, \overline{a}_{ij} + \beta_{ij} d_{ij}^F], \forall j; \\ 0 \le \beta_{ij} \le 1, \forall j; \textstyle\sum_{j \in J_i} \beta_{ij} \le \Gamma_i\}. \end{aligned} \tag{4}$$

The Γ_i restricts the "variability" of the random elements in row i. One may envision that it excludes those cases where every random element simultaneously takes a value far from its mean, which can be presumed rare from a practical perspective.

Our goal is to find a maximizing $\mathbf{x}$ that is guaranteed to satisfy the constraints $\mathbf{A}\mathbf{x} \le \mathbf{b}$ as long as the uncertain vector $\mathbf{a}_i$ belongs to the set $\mathscr{R}_i(\Gamma_i)$ for all i. For that purpose, we construct the *robust formulation*

$$\begin{aligned} z_R(\mathbf{\Gamma}) = \max\ & \mathbf{c}'\mathbf{x} \\ \text{s.t.}\ & \max\nolimits_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)} \{\mathbf{a}_i'\mathbf{x}\} \le b_i, \quad \forall i \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}. \end{aligned} \tag{5}$$

Since the robust formulation (5) can be viewed as a relaxation of the fat formulation (2), $z_R(\mathbf{\Gamma}) \ge z_F$. However, an optimal solution of (5) may violate the constraint $\mathbf{A}\mathbf{x} \le \mathbf{b}$. This constraint violation probability will be addressed in Section 3.

We now show that (5) can be recast as an equivalent LP formulation. For any $\mathbf{x}$, the maximization problem in the ith constraint of (5) is written as

$$\max\ \mathbf{a}_i'\mathbf{x} \tag{6}$$

$$\text{s.t.}\ a_{ij} \le \overline{a}_{ij} + \beta_{ij} d_{ij}^F, \quad \forall j \tag{6a}$$

$$a_{ij} \ge \overline{a}_{ij} - \beta_{ij} d_{ij}^B, \quad \forall j \tag{6b}$$

$$\textstyle\sum_{j \in J_i} \beta_{ij} \le \Gamma_i \tag{6c}$$

$$\beta_{ij} \le 1, \quad \forall j \tag{6d}$$

$$\beta_{ij} \ge 0, \quad \forall j,$$

where a_{ij} and β_{ij} are the decision variables. Let λ_{ij}, μ_{ij}, z_i, and p_{ij} be the dual variables for the constraints (6a)–(6d), respectively. Then the dual of (6) is given by (after some simplifications)

$$\begin{aligned} \min\ & \textstyle\sum_j \overline{a}_{ij} x_j + \Gamma_i z_i + \sum_j p_{ij} \\ \text{s.t.}\ & \lambda_{ij} - \mu_{ij} = x_j, \quad \forall j \\ & z_i + p_{ij} \ge d_{ij}^F \lambda_{ij} + d_{ij}^B \mu_{ij}, \quad \forall j \\ & \lambda_{ij}, \mu_{ij}, p_{ij} \ge 0, \quad \forall j \\ & z_i \ge 0. \end{aligned} \tag{7}$$

Theorem 1. The robust formulation (5) is equivalent to the LP formulation

$$z_R(\mathbf{\Gamma}) = \max\ \mathbf{c}'\mathbf{x} \tag{8}$$

$$\begin{aligned} \text{s.t.}\ & \textstyle\sum_j \overline{a}_{ij}x_j + \Gamma_i z_i + \sum_j p_{ij} \le b_i && \forall i \\ & \lambda_{ij} - \mu_{ij} = x_j, && \forall i, j \\ & z_i + p_{ij} \ge d_{ij}^F \lambda_{ij} + d_{ij}^B \mu_{ij}, && \forall i, j \\ & \lambda_{ij}, \mu_{ij}, p_{ij} \ge 0, && \forall i, j \\ & z_i \ge 0, && \forall i \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}. \end{aligned}$$

Proof. Let $(\mathbf{x}^*, \lambda_{ij}^*, \mu_{ij}^*, z_i^*, p_{ij}^*)$ be an optimal solution of (8). Let $\tilde{\mathbf{x}}$ be an optimal solution of (5). The equivalence will be established by showing that $\mathbf{x}^*$ is feasible to (5) and $\mathbf{c}'\mathbf{x}^* = \mathbf{c}'\tilde{\mathbf{x}}$. Fix $\mathbf{x} = \mathbf{x}^*$ in (6) and (7), and let $(\mathbf{a}_i^*, \beta_{ij}^*)$ be an optimal solution of (6). Since $(\lambda_{ij}^*, \mu_{ij}^*, z_i^*, p_{ij}^*)$ is a feasible solution of (7), the weak duality between (6) and (7) yields

$$\mathbf{a}_i^{*\prime}\mathbf{x}^* \le \textstyle\sum_j \overline{a}_{ij}x_j^* + \Gamma_i z_i^* + \sum_j p_{ij}^*.$$

Since $\mathbf{a}_i^{*\prime}\mathbf{x}^* = \max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)}\{\mathbf{a}_i'\mathbf{x}^*\}$ and $(\mathbf{x}^*, \lambda_{ij}^*, \mu_{ij}^*, z_i^*, p_{ij}^*)$ is feasible to (8), we have, for all i,

$$\max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)}\{\mathbf{a}_i'\mathbf{x}^*\} \le \textstyle\sum_j \overline{a}_{ij}x_j^* + \Gamma_i z_i^* + \sum_j p_{ij}^* \le b_i.$$

This shows that $\mathbf{x}^*$ is feasible to (5), implying that $\mathbf{c}'\mathbf{x}^* \le \mathbf{c}'\tilde{\mathbf{x}}$. Next, set $\mathbf{x} = \tilde{\mathbf{x}}$ in (6) and (7), and let $(\tilde{\mathbf{a}}_i, \tilde{\beta}_{ij})$ be an optimal solution of (6). By the strong duality, there exists a feasible $(\tilde{\lambda}_{ij}, \tilde{\mu}_{ij}, \tilde{z}_i, \tilde{p}_{ij})$ to (7) such that

$$\max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)}\{\mathbf{a}_i'\tilde{\mathbf{x}}\} = \tilde{\mathbf{a}}_i'\tilde{\mathbf{x}} = \textstyle\sum_j \overline{a}_{ij}\tilde{x}_j + \Gamma_i \tilde{z}_i + \sum_j \tilde{p}_{ij}.$$

Since $\tilde{\mathbf{x}}$ is feasible to (5), we have, for all i,

$$b_i \ge \max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)}\{\mathbf{a}_i'\tilde{\mathbf{x}}\} = \textstyle\sum_j \overline{a}_{ij}\tilde{x}_j + \Gamma_i \tilde{z}_i + \sum_j \tilde{p}_{ij}.$$

This shows that $(\tilde{\mathbf{x}}, \tilde{\lambda}_{ij}, \tilde{\mu}_{ij}, \tilde{z}_i, \tilde{p}_{ij})$ satisfies the first set of the constraints of (8). Since the other constraints of (8) are also satisfied by $(\tilde{\mathbf{x}}, \tilde{\lambda}_{ij}, \tilde{\mu}_{ij}, \tilde{z}_i, \tilde{p}_{ij})$, it is a feasible solution of (8), from which we have $\mathbf{c}'\tilde{\mathbf{x}} \le \mathbf{c}'\mathbf{x}^*$. ■

Since the robust formulation (5) is parameterized by Γ_i, $i = 1, \ldots, m$, one can expect that adjusting values for Γ_i's gives a flexibility for the formulation. Indeed, if $\Gamma_i = |J_i|$ for all i, $\mathscr{R}_i(\Gamma_i) = \mathscr{U}_i$ for all i. Consequently, (5) becomes the fat formulation (2). On the other hand, if $\Gamma_i = 0$ for all i, $\mathscr{R}_i(\Gamma_i) = \{\overline{\mathbf{a}}_i\}$ for all i. In this case, (5) is simply reduced to the *nominal formulation*

$$\begin{aligned} & \max\ \mathbf{c}'\mathbf{x} \\ & \text{s.t.}\ \textstyle\sum_j \overline{a}_{ij}x_j \le b_i \quad \forall i \\ & \qquad \mathbf{l} \le \mathbf{x} \le \mathbf{u}. \end{aligned} \tag{9}$$

3. BOUNDS ON THE CONSTRAINT VIOLATION PROBABILITY

Unless $\Gamma_i = |J_i|$ for all i, an optimal solution $\mathbf{x}^*$ of the robust formulation (5) (which is obtained by solving the equivalent LP formulation (8)) may violate the constraints $\mathbf{Ax} \le \mathbf{b}$ of (1). Let us consider the ith constraint violation probability $P[\mathbf{a}_i'\mathbf{x}^* > b_i]$. Computing this probability exactly is often a challenging task. Therefore we are interested in establishing (easily computable) upper bounds on this probability.

3.1 A Distribution-Dependent Bound

One of such bounds can be obtained as follows.

$$\begin{aligned} P\big[\mathbf{a}_i'\mathbf{x}^* > b_i\big] &= P\big[\textstyle\sum_{j\in J_i} a_{ij}x_j^* + \sum_{j\notin J_i} \overline{a}_{ij}x_j^* > b_i\big] \\ &= P\big[\textstyle\sum_{j\in J_i} x_j^* a_{ij} > B_i(\mathbf{x}^*)\big] \\ &\le P\big[\textstyle\sum_{j\in J_i} x_j^* a_{ij} \ge B_i(\mathbf{x}^*)\big], \end{aligned}$$

where $B_i(\mathbf{x}^*) = b_i - \sum_{j\notin J_i} \overline{a}_{ij}x_j^*$. Using Markov's inequality and the independence of a_{ij}'s, $\forall j \in J_i$, we have, for any $\theta \ge 0$,

$$\begin{aligned} P\big[\textstyle\sum_{j\in J_i} x_j^* a_{ij} \ge B_i(\mathbf{x}^*)\big] &\le e^{-\theta B_i(\mathbf{x}^*)} E\big[e^{\theta \sum_{j\in J_i} x_j^* a_{ij}}\big] \\ &= e^{-\theta B_i(\mathbf{x}^*)} \textstyle\prod_{j\in J_i} E\big[e^{\theta x_j^* a_{ij}}\big] \\ &= \exp\big[-\theta B_i(\mathbf{x}^*) + \textstyle\sum_{j\in J_i} \Lambda_{a_{ij}}(\theta x_j^*)\big], \end{aligned}$$

where $\Lambda_{a_{ij}}(\theta x_j^*) \triangleq \log E[e^{\theta x_j^* a_{ij}}]$. Optimizing over $\theta \ge 0$, we obtain the following result.

Proposition 2. The ith constraint violation probability is upper bounded as follows:

$$P\big[\mathbf{a}_i'\mathbf{x}^* > b_i\big] \le \exp\Big[-\sup_{\theta\ge 0}\Big\{\theta B_i(\mathbf{x}^*) - \sum_{j\in J_i} \Lambda_{a_{ij}}(\theta x_j^*)\Big\}\Big]. \tag{10}$$

Because the log moment generating function $\Lambda_{a_{ij}}(\theta x_j^*)$ is convex in θ, $\sup_{\theta\ge 0}\{\cdot\}$ in (10) is a convex optimization problem, which is efficiently solvable from the computational complexity standpoint.

3.2 A Bound with Limited Distributional Information

The bound in (10) requires full distributional information on a_{ij}'s in order to compute $E[e^{\theta x_j^* a_{ij}}]$. In some applications, however, only limited distributional information, such as the means of a_{ij}'s, might be available. To derive a bound for this case, we employ the idea of upper bounding $E[e^{\theta x_j^* a_{ij}}]$ using the means of a_{ij}'s and their range information (instead of computing it exactly). To that end, we will use the following inequality due to Bennett [1962] (also see Dembo and Zeitouni [1998]): Let $X \le b$ is a random variable with $\overline{x} = E[X]$ and $E[(X - \overline{x})^2] \le c^2$ for some $c > 0$. Then for any $\theta \ge 0$,

$$E\big[e^{\theta X}\big] \le e^{\theta\overline{x}}\Big[\tfrac{(b-\overline{x})^2}{(b-\overline{x})^2+c^2} e^{-\frac{\theta c^2}{b-\overline{x}}} + \tfrac{c^2}{(b-\overline{x})^2+c^2} e^{\theta(b-\overline{x})}\Big]. \tag{11}$$

Now suppose that we know $a_{ij} \in [\overline{a}_{ij} - d_{ij}^B, \overline{a}_{ij} + d_{ij}^F]$, but do not know the probability distribution of a_{ij}. Consider the following probability distribution

$$f_{a_{ij}}(a) = \begin{cases} d_{ij}^F/(d_{ij}^F + d_{ij}^B) & \text{if } a = \overline{a}_{ij} - d_{ij}^B, \\ d_{ij}^B/(d_{ij}^F + d_{ij}^B) & \text{if } a = \overline{a}_{ij} + d_{ij}^F. \end{cases} \tag{12}$$

This probability distribution yields $E[a_{ij}] = \overline{a}_{ij}$. Notice that among all the probability distributions that yield $E[a_{ij}] = \overline{a}_{ij}$, (12) has the maximum variance. Hence for any such probability distribution for a_{ij}, we have $E[(a_{ij} - \overline{a}_{ij})^2] \le d_{ij}^F d_{ij}^B$, where $d_{ij}^F d_{ij}^B$ is the variance of (12).

Let $X = x_j^* a_{ij}$. Then $X \le \overline{a}_{ij}x_j^* + \max\{-d_{ij}^B x_j^*, d_{ij}^F x_j^*\}$ and $E[(X - \overline{x})^2] \le x_j^{*2} d_{ij}^F d_{ij}^B$. Using the inequality (11), it can be shown that

$$E\big[e^{\theta x_j^* a_{ij}}\big] \le \tfrac{d_{ij}^F}{d_{ij}^F + d_{ij}^B} e^{\theta x_j^* a_{ij}^L} + \tfrac{d_{ij}^B}{d_{ij}^F + d_{ij}^B} e^{\theta x_j^* a_{ij}^U}.$$

Then following the same steps used for the bound in (10), we obtain, for any $\theta \ge 0$,

$$\begin{aligned} P\big[\mathbf{a}_i'\mathbf{x}^* > b_i\big] &\le e^{-\theta B_i(\mathbf{x}^*)} \textstyle\prod_{j\in J_i} E\big[e^{\theta x_j^* a_{ij}}\big] \\ &\le e^{-\theta B_i(\mathbf{x}^*)} \textstyle\prod_{j\in J_i} \Big\{\tfrac{d_{ij}^F}{d_{ij}^F + d_{ij}^B} e^{\theta x_j^* a_{ij}^L} + \tfrac{d_{ij}^B}{d_{ij}^F + d_{ij}^B} e^{\theta x_j^* a_{ij}^U}\Big\} \\ &= \exp\big[-\theta B_i(\mathbf{x}^*) + \textstyle\sum_{j\in J_i} \log\big(g_{ij} e^{\theta x_j^* a_{ij}^L} + h_{ij} e^{\theta x_j^* a_{ij}^U}\big)\big], \end{aligned}$$

where $g_{ij} \triangleq d_{ij}^F/(d_{ij}^F + d_{ij}^B)$ and $h_{ij} \triangleq d_{ij}^B/(d_{ij}^F + d_{ij}^B)$. Again, optimizing over $\theta \ge 0$, the following result is attained.

Proposition 3. The ith constraint violation probability is upper bounded as follows:

$$P\big[\mathbf{a}_i'\mathbf{x}^* > b_i\big] \le \exp\Big[-\sup_{\theta\ge 0}\Big\{\theta B_i(\mathbf{x}^*) - \sum_{j\in J_i}\log\big(g_{ij}e^{\theta x_j^* a_{ij}^L} + h_{ij}e^{\theta x_j^* a_{ij}^U}\big)\Big\}\Big]. \quad (13)$$

From the way the bound in (13) was established, one can infer that it will be weaker than the bound in (10). Nonetheless, we believe that the former can be useful in many cases.

3.3 A Bound Based on Sample Data

The underlying assumption for the bound in (13) is that the true means of the random a_{ij}'s are known. One may argue that this assumption is problematic because in some applications only estimates of the true means could be available from observed data. We now present a bound that is suitable for this situation.

Let $a_{ij}^1, \ldots, a_{ij}^{N_{ij}}$ be N_{ij} samples of a random a_{ij}. Let $\widetilde{a}_{ij} = \frac{1}{N_{ij}}\sum_{k=1}^{N_{ij}} a_{ij}^k$ be the sample mean of a_{ij}^k's. Hoeffding's inequality (Hoeffding [1963]) yields that for $t > 0$,

$$P\big[x_j^*\widetilde{a}_{ij} - E[x_j^* a_{ij}] \le -t\big] \le \exp\Big[-\frac{2N_{ij}t^2}{x_j^{*2}(d_{ij}^F + d_{ij}^B)^2}\Big].$$

Letting $\epsilon_{ij} = \exp\big[-\frac{2N_{ij}t^2}{x_j^{*2}(d_{ij}^F + d_{ij}^B)^2}\big]$ and then solving for t, we obtain

$$P\Big[x_j^*\widetilde{a}_{ij} - E[x_j^* a_{ij}] \le -\sqrt{\frac{x_j^{*2}(d_{ij}^F + d_{ij}^B)^2}{2N_{ij}}\log\frac{1}{\epsilon_{ij}}}\Big] \le \epsilon_{ij}.$$

In other words, the unknown true mean $E[x_j^* a_{ij}]$ is upper bounded by

$$E[x_j^* a_{ij}] \le x_j^*\widetilde{a}_{ij} + \sqrt{\frac{x_j^{*2}(d_{ij}^F + d_{ij}^B)^2}{2N_{ij}}\log\frac{1}{\epsilon_{ij}}} \quad (14)$$

with probability at least $1 - \epsilon_{ij}$.

Now consider the following inequality that follows from the convexity of the exponential function e^x: If X is a random variable such that $a \le X \le b$, then for any real number θ,

$$E\big[e^{\theta X}\big] \le \frac{b - E[X]}{b-a}e^{\theta a} + \frac{E[X]-a}{b-a}e^{\theta b}. \quad (15)$$

Applying the inequality (15) to $X = x_j^* a_{ij}$, it can be seen that

$$E\big[e^{\theta x_j^* a_{ij}}\big] \le E[x_j^* a_{ij}]\frac{e^{\theta s_{ij}} - e^{\theta r_{ij}}}{|x_j^*|(d_{ij}^F + d_{ij}^B)} + \frac{s_{ij}e^{\theta r_{ij}} - r_{ij}e^{\theta s_{ij}}}{|x_j^*|(d_{ij}^F + d_{ij}^B)}, \quad (16)$$

where $r_{ij} \triangleq \min\{x_j^* a_{ij}^L, x_j^* a_{ij}^U\}$ and $s_{ij} \triangleq \max\{x_j^* a_{ij}^L, x_j^* a_{ij}^U\}$. Since $e^{\theta s_{ij}} - e^{\theta r_{ij}} \ge 0$ for all $\theta \ge 0$, we replace $E[x_j^* a_{ij}]$ in (16) with the upper bound in (14) and conclude that for any $\theta \ge 0$

$$E\big[e^{\theta x_j^* a_{ij}}\big] \le \Big\{x_j^*\widetilde{a}_{ij} + \sqrt{\frac{x_j^{*2}(d_{ij}^F + d_{ij}^B)^2}{2N_{ij}}\log\frac{1}{\epsilon_{ij}}}\Big\}\frac{e^{\theta s_{ij}} - e^{\theta r_{ij}}}{|x_j^*|(d_{ij}^F + d_{ij}^B)} + \frac{s_{ij}e^{\theta r_{ij}} - r_{ij}e^{\theta s_{ij}}}{|x_j^*|(d_{ij}^F + d_{ij}^B)} \quad (17)$$

with probability at least $1 - \epsilon_{ij}$. Denote the right hand side of the inequality (17) by $f_{ij}(\theta)$. Following the same steps used for the bound in (13), we obtain the following result.

Proposition 4. Let $0 < \epsilon_{ij} < 1$ for all $j \in J_i$. The following bound on the ith constraint violation probability holds with probability at least $\prod_{j\in J_i}(1 - \epsilon_{ij})$:

$$P\big[\mathbf{a}_i'\mathbf{x}^* > b_i\big] \le \exp\Big[-\sup_{\theta\ge 0}\Big\{\theta B_i(\mathbf{x}^*) - \sum_{j\in J_i}\log f_{ij}(\theta)\Big\}\Big]. \quad (18)$$

We emphasize that unlike the bounds in (10) and (13) that hold deterministically, the bound in (18) is valid with a certain probability.

4. NUMERICAL RESULTS

We consider a 10×10 matrix $\mathbf{A}$, where all elements are assumed to be random, i.e., $|J_i| = 10$ for all i. A data set $(\mathbf{c}, \mathbf{b}, \mathbf{l}, \mathbf{u}, a_{ij}^U, a_{ij}^L)$, which constitutes a problem instance, is specified as follows: c_j is randomly selected from $[-30, 30]$ for all j; b_i is randomly selected from $[5, 15]$ for all i; l_j and u_j are set to -10 and 10, respectively, for all j; a_{ij}^U is randomly chosen from $[-10, 30]$ for all i and j; a_{ij}^L is set to $a_{ij}^U - \delta_{ij}$, where δ_{ij} is randomly drawn from $[1, 10]$ for all i and j. In this way, we generate 5 data sets (i.e., 5 problem instances). We assume that each random a_{ij} takes the following linearly decreasing density function over its range:

$$f_{a_{ij}}(a) = \frac{2}{(a_{ij}^U - a_{ij}^L)^2}(a_{ij}^U - a), \quad a_{ij}^L \le a \le a_{ij}^U.$$

Given a problem instance, let z_F be the optimal objective value of the fat formulation (2). Let $z_R(7.5)$ denote the optimal objective value of the robust formulation (5) when $\Gamma_i = 7.5$, $i = 1, \ldots, 10$. We define $z_R(5.0)$ and $z_R(2.5)$ similarly. Table 1 shows z_F, $z_R(7.5)$, $z_R(5.0)$, and $z_R(2.5)$ for each problem instance. As Γ_i's take a smaller value, the robust formulation becomes less conservative, yielding a higher objective value.

Table 1. Optimal objective values

Instance	z_F	$z_R(7.5)$	$z_R(5.0)$	$z_R(2.5)$
1	1548.88	1558.85	1595.91	1665.22
2	1263.90	1278.91	1310.90	1362.31
3	969.39	979.63	1040.84	1146.12
4	257.25	262.48	291.98	380.81
5	1149.24	1181.40	1273.51	1421.42
Average	1037.73	1052.25	1102.63	1195.18

Given an optimal solution $\mathbf{x}^*$ of the robust formulation (5), let P_i be the value of the bound in (10) for the ith constraint. Then the probability that $\mathbf{x}^*$ remains feasible to (1) is at least $\prod_{i=1}^m(1 - P_i)$. In Table 2, we report this *probability of feasibility* of $\mathbf{x}^*$ for each problem instance when $\Gamma_i = 5.0$ for all i. We also use the bound in (13) instead of the bound in (10) and report the results. As one might expect, the bound in (10), which is stronger than the bound in (13), results in a higher probability of feasibility. When the bound (10) is used, the probability of feasibility of $\mathbf{x}^*$ is almost 1.0 for each problem instance; in other words, it is extremely unlikely that $\mathbf{x}^*$ violates any of the constraints $\mathbf{Ax} \le \mathbf{b}$ of (1). Meanwhile, as can be seen in Table 1, the robust formulation (5) improves the objective value by 6.3% on average, compared to the fat formulation (2). This demonstrates that when an application can tolerate a small chance of infeasibility of a solution, our robust optimization approach would be preferred to the classical robust optimization approach.

5. ROBUST CONTROL OF LINEAR UNCERTAIN SYSTEMS WITH CONSTRAINTS

In this section, we apply the robust optimization approach of Section 2 to linear uncertain systems with constraints.

Table 2. Probability of feasibility when $\Gamma_i = 5.0$ for all i

Instance	From (10)	From (13)
1	0.999971	0.804643
2	0.999665	0.597228
3	0.999970	0.713272
4	0.999981	0.652228
5	0.999980	0.761439

Consider the discrete-time linear uncertain system

$$\mathbf{x}^{k+1} = \mathbf{A}\mathbf{x}^k + \mathbf{B}\mathbf{u}^k, \quad k = 0, \ldots, N-1$$

subject to the constraints

$$\begin{aligned} &\mathbf{u}^L \le \mathbf{u}^k \le \mathbf{u}^U, \quad && k = 0, \ldots, N-1, \\ &\mathbf{x}^k \in \mathcal{X}^k, && k = 0, \ldots, N, \end{aligned}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a given matrix, $\mathbf{u}^U, \mathbf{u}^L \in \mathbb{R}^m$ are given vectors, and $\mathbf{x}^k \in \mathbb{R}^n$ and $\mathbf{u}^k \in \mathbb{R}^m$ are the state and control vectors at time k, respectively. We assume that the matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$ is uncertain within the range

$$\overline{\mathbf{B}} - \mathbf{D}^B \le \mathbf{B} \le \overline{\mathbf{B}} + \mathbf{D}^F, \tag{19}$$

where $\overline{\mathbf{B}}$ is a nominal matrix and $\mathbf{D}^F, \mathbf{D}^B > \mathbf{0}$. The sets $\mathcal{X}^k$ are defined as

$$\mathcal{X}^k = \begin{cases} \{\mathbf{x}^k \mid \mathbf{x}^k \ge \mathbf{x}^L\} & \text{if } k = N, \\ \{\mathbf{x}^k \mid \mathbf{x}^k \ge \mathbf{x}^L \text{ and} \\ \quad \exists\, \mathbf{u}^k, \mathbf{A}\mathbf{x}^k + \mathbf{B}\mathbf{u}^k \in \mathcal{X}^{k+1}, \forall \mathbf{B}\} & \text{otherwise,} \end{cases}$$

where $\mathbf{x}^L$ is a given vector.

For given matrices $\mathbf{Q}, \mathbf{P} \in \mathbb{R}^{n\times n}$ and $\mathbf{R} \in \mathbb{R}^{n\times m}$, we consider the control synthesis based on the *closed-loop optimal control problem* (cf. Bemporad et al. [2003])

$$\begin{aligned} J_k^*(\mathbf{x}^k) = \min_{\mathbf{u}^k} \; & \left\|\mathbf{Q}\mathbf{x}^k\right\|_1 + \left\|\mathbf{R}\mathbf{u}^k\right\|_1 + J_{k+1}^*(\mathbf{A}\mathbf{x}^k + \overline{\mathbf{B}}\mathbf{u}^k) \quad (20) \\ \text{s.t.}\; & \mathbf{u}^L \le \mathbf{u}^k \le \mathbf{u}^U \\ & \mathbf{x}^k \ge \mathbf{x}^L \\ & \mathbf{A}\mathbf{x}^k + \mathbf{B}\mathbf{u}^k \in \mathcal{X}^{k+1} \end{aligned}$$

for $k = 0, \ldots, N-1$ and with the boundary conditions $J_N^*(\mathbf{x}^N) = \left\|\mathbf{P}\mathbf{x}^N\right\|_1$ and $\mathbf{x}^N \ge \mathbf{x}^L$. As can be seen from the objective function, the goal of this problem is to regulate the system to the origin, given the initial state vector $\mathbf{x}^0$. To solve (20), a multiparametric LP solver (Kvasnica et al. [2004]) is used. The constraint $\mathbf{A}\mathbf{x}^k + \mathbf{B}\mathbf{u}^k \in \mathcal{X}^{k+1}$ is then explicitly written as

$$\mathbf{L}^{k+1}(\mathbf{A}\mathbf{x}^k + \mathbf{B}\mathbf{u}^k) \ge \mathbf{f}^{k+1}, \tag{21}$$

where the matrix $\mathbf{L}^{k+1}$ and the vector $\mathbf{f}^{k+1}$ are returned from the multiparametric LP solver after it solves (20) for time $k+1$. It can be shown that $\mathcal{X}^k = \{\mathbf{x}^k \mid \mathbf{L}^k\mathbf{x}^k \ge \mathbf{f}^k\}$ (Bemporad et al. [2003]).

Given an uncertainty budget $\Gamma \in [0, nm]$, we define

$$\begin{aligned} \mathscr{R}(\Gamma) \triangleq \big\{\mathbf{B} \mid\, & b_{ij} \in [\bar{b}_{ij} - \beta_{ij}d_{ij}^B, \bar{b}_{ij} + \beta_{ij}d_{ij}^F], \forall\, i, j; \\ & 0 \le \beta_{ij} \le 1, \forall\, i, j;\; \textstyle\sum_i \sum_j \beta_{ij} \le \Gamma\big\}, \end{aligned} \tag{22}$$

where $\bar{b}_{ij}$, d_{ij}^B, and d_{ij}^F are the (i,j)th components of $\overline{\mathbf{B}}$, $\mathbf{D}^B$, and $\mathbf{D}^F$, respectively. Recall that in Section 2 we introduced Γ_i for each row i of the uncertain constraint matrix $\mathbf{A}$ (see (4)), because each constraint of $\mathbf{A}\mathbf{x} \le \mathbf{b}$ has a distinct set of random elements. That is not the case for the constraints (21): The presence of $\mathbf{L}^{k+1}$ may cause different constraints to have common random elements. For instance, the uncertain b_{ij} may appear not only in the ith constraint of (21), but also in other constraints. Therefore, it is more appropriate to use a single (global) uncertainty budget.

We construct the *robust formulation* of the closed-loop optimal control problem (20) as follows (cf. (5)):

$$\begin{aligned} J_k^*(\mathbf{x}^k) = \min_{\mathbf{u}^k} \; & \left\|\mathbf{Q}\mathbf{x}^k\right\|_1 + \left\|\mathbf{R}\mathbf{u}^k\right\|_1 + J_{k+1}^*(\mathbf{A}\mathbf{x}^k + \overline{\mathbf{B}}\mathbf{u}^k) \quad (23) \\ \text{s.t.}\; & \mathbf{u}^L \le \mathbf{u}^k \le \mathbf{u}^U \\ & \mathbf{x}^k \ge \mathbf{x}^L \\ & \min_{\mathbf{B}\in\mathscr{R}(\Gamma)} \big\{(\mathbf{L}^{k+1}\mathbf{B}\mathbf{u}^k)_r\big\} \ge (\mathbf{f}^{k+1} - \mathbf{L}^{k+1}\mathbf{A}\mathbf{x}^k)_r, \forall\, r, \quad (23a) \end{aligned}$$

where $(\mathbf{y})_r$ denotes the rth element of the vector $\mathbf{y}$. If $\Gamma = 0$, then $\mathscr{R}(\Gamma) = \{\overline{\mathbf{B}}\}$ and consequently (23) becomes the *nominal formulation* (cf. (9)). If, on the other hand, $\Gamma = nm$, then (23) becomes the *fat formulation* (cf. (2)) which represents the classical worst-case approach.

For any $\mathbf{u}^k$, the minimization problem in the rth constraint of (23a) can be written as

$$\begin{aligned} \min \; & \sum_{j=1}^m \sum_{i=1}^n \ell_{ri}^{k+1} b_{ij} u_j^k && (24) \\ \text{s.t.}\; & b_{ij} \le \bar{b}_{ij} + \beta_{ij}d_{ij}^F, \;\; i = 1, \ldots, n,\; j = 1, \ldots, m && (24a) \\ & b_{ij} \ge \bar{b}_{ij} - \beta_{ij}d_{ij}^B, \;\; i = 1, \ldots, n,\; j = 1, \ldots, m && (24b) \\ & \sum_{i=1}^n \sum_{j=1}^m \beta_{ij} \le \Gamma && (24c) \\ & \beta_{ij} \le 1, \;\; i = 1, \ldots, n,\; j = 1, \ldots, m && (24d) \\ & \beta_{ij} \ge 0, \;\; i = 1, \ldots, n,\; j = 1, \ldots, m, \end{aligned}$$

where ℓ_{ri}^{k+1} is the (r,i)th element of $\mathbf{L}^{k+1}$. Let λ_{rij}^k, μ_{rij}^k, z_r^k, and p_{rij}^k be the dual variables for the constraints (24a)–(24d). Then the dual of (24) is (after some simplifications)

$$\begin{aligned} \max \; & \sum_{i=1}^n \sum_{j=1}^m \bar{b}_{ij} \ell_{ri}^{k+1} u_j^k - \Gamma z_r^k - \sum_{i=1}^n \sum_{j=1}^m p_{rij}^k \\ \text{s.t.}\; & -\lambda_{rij}^k + \mu_{rij}^k = \ell_{ri}^{k+1} u_j^k, && \forall\, i, j \\ & z_r^k + p_{rij}^k \ge d_{ij}^F \lambda_{rij}^k + d_{ij}^B \mu_{rij}^k, && \forall\, i, j \\ & \lambda_{rij}^k, \mu_{rij}^k, p_{rij}^k \ge 0, && \forall\, i, j \\ & z_r^k \ge 0. \end{aligned}$$

The following proposition shows that the robust formulation (23) can be recast as a formulation with linear constraints. The proof technique is similar to that of Theorem 1 and is omitted due to space limitations.

Proposition 5. The robust formulation (23) is equivalent to the formulation

$$\begin{aligned} J_k^*(\mathbf{x}^k) = \min_{\mathbf{u}^k} \; & \left\|\mathbf{Q}\mathbf{x}^k\right\|_1 + \left\|\mathbf{R}\mathbf{u}^k\right\|_1 + J_{k+1}^*(\mathbf{A}\mathbf{x}^k + \overline{\mathbf{B}}\mathbf{u}^k) \quad (25) \\ \text{s.t.}\; & \mathbf{u}^L \le \mathbf{u}^k \le \mathbf{u}^U \\ & \mathbf{x}^k \ge \mathbf{x}^L \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^m \bar{b}_{ij} \ell_{ri}^{k+1} u_j^k - \Gamma z_r^k - \sum_{i=1}^n \sum_{j=1}^m p_{rij}^k \ge (\mathbf{f}^{k+1} - \mathbf{L}^{k+1}\mathbf{A}\mathbf{x}^k)_r, \forall\, r \\ -\lambda_{rij}^k + \mu_{rij}^k = \ell_{ri}^{k+1} u_j^k, \quad & \forall\, r, i, j \\ z_r^k + p_{rij}^k \ge d_{ij}^F \lambda_{rij}^k + d_{ij}^B \mu_{rij}^k, \quad & \forall\, r, i, j \\ \lambda_{rij}^k, \mu_{rij}^k, p_{rij}^k \ge 0, \quad & \forall\, r, i, j \\ z_r^k \ge 0, \quad & \forall\, r. \end{aligned}$$

Our robust formulation (23) of the closed-loop optimal control problem (20) distinguishes itself from the work of Bemporad et al. [2003] in two aspects: First, the uncertainty budget Γ provides the formulation with flexibility so that it can produce different control laws without changing its underlying structure. Second, its equivalent formulation (25) offers computational advantages over the formulation of Bemporad et al. [2003] which requires the enumeration of all the vertices of the uncertainty set.

For numerical tests, we use the following data: $N = 4$,

$$\mathbf{A} = \begin{bmatrix} 0.8 & 0.4 \\ -0.4 & 0.8 \end{bmatrix},\ \overline{\mathbf{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \mathbf{D}^B = \begin{bmatrix} 0.4 \\ 0.4 \end{bmatrix},\ \mathbf{D}^F = \begin{bmatrix} 0.7 \\ 0.7 \end{bmatrix},$$

$$\mathbf{u}^U = 3,\ \mathbf{u}^L = -3,\ \mathbf{x}^L = \begin{bmatrix} -1.5 \\ -1.5 \end{bmatrix},\ \mathbf{Q} = \mathbf{P} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},\ \mathbf{R} = 0.4.$$

With four different values for Γ (0.8, 1.0, 1.5, and 2.0), we solve (25) recursively from $k = N-1$ to $k = 0$. As a result, for each value of Γ, a piecewise affine control law of the following form is obtained:

$$\mathbf{u}^k = \mathbf{h}_i'\mathbf{x}^k + g_i \ \text{ for } \mathbf{x}^k \in \mathscr{P}_i, \quad i = 1, \ldots, s, \quad k = 0, \ldots, N-1,$$

where s is the number of partitions of the admissible set for $\mathbf{x}^k$, $\mathscr{P}_i$'s are polytopes such that $\bigcup_{i=1}^{s} \mathscr{P}_i = \mathcal{X}^k$, and $\mathbf{h}_i$ and g_i are a vector and a scalar determined by the multiparametric LP solver for partition i.

We then simulate the system with the control law associated with each value of Γ, where the control is applied in a receding horizon manner: namely, only $\mathbf{u}^0$ is used at each time. The time horizon of a simulation is set to 20. For each control law, we perform 200 simulations by creating 200 random pairs of $(\mathbf{x}^0, \{\mathbf{B}^k\}_{k=1}^{20})$, where $\mathbf{B}^k$ is a realization of $\mathbf{B}$ at time k. The components of $\mathbf{x}^0$ are uniformly distributed in the ranges $[-2, 4]$ and $[0, 6]$, respectively. The $\mathbf{B}^k$ are generated assuming that each component has a triangle distribution over the corresponding range (cf. (19)). Table 3 lists the average cost of the 200 simulations for each control law. It is demonstrated that as the value of Γ gets smaller, the corresponding control law leads to cost savings.

Table 3. Simulated costs

Γ	0.8	1.0	1.5	2.0
Average cost	8.3960	8.5989	8.6525	9.1385

Fig. 1 shows two sample trajectories of the state vector $\mathbf{x}^k$ and the control $\mathbf{u}^k$ when $\Gamma = 0.8$ and $\Gamma = 2.0$, respectively.

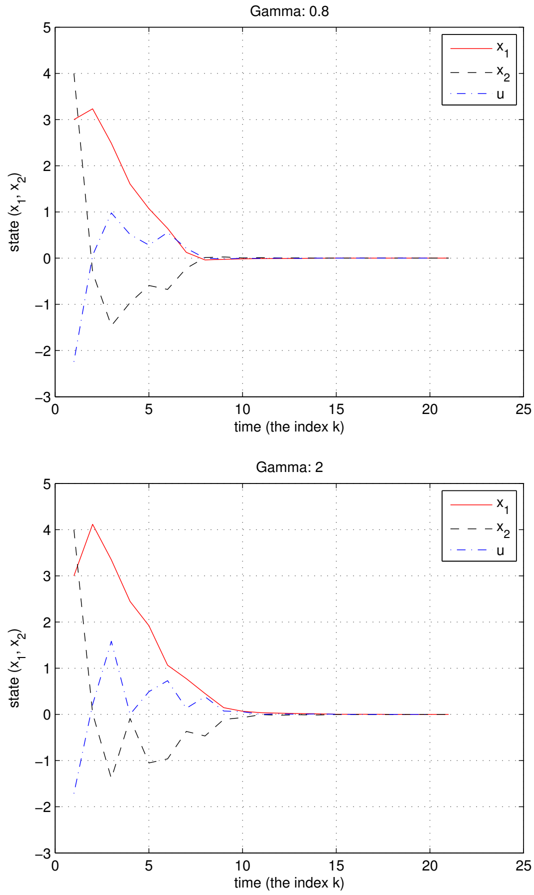

Fig. 1. Sample trajectories of the state vector $\mathbf{x}^k$ and the control $\mathbf{u}^k$

REFERENCES

A. Bemporad, F. Borrelli, and M. Morari. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(9):1600–1606, 2003.

A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25:1–13, 1999.

A. Ben-Tal and A. Nemirovski. Robust optimization – methodology and applications. *Mathematical Programming, Series B*, 92:453–480, 2002.

G. Bennett. Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association*, 57:33–45, 1962.

D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, New York, 2nd edition, 1998.

L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1):33–52, 1998.

W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

M. Kvasnica, P. Grieder, and M. Baotić. Multi-Parametric Toolbox (MPT), 2004. URL `http://control.ee.ethz.ch/ mpt/`.

I. Ch. Paschalidis and S.-C. Kang. Robust linear optimization: On the benefits of distributional information and applications in inventory control. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 4416–4421, Seville, Spain, 2005.

I. Ch. Paschalidis and S.-C. Kang. On the benefits of distributional information in robust linear optimization. In *Proceedings of the 5th IFAC Symposium on Robust Control Design*, Toulouse, France, 2006.

A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, 1973.