Annals of Operations Research manuscript No.
(will be inserted by the editor)

Distribution-dependent robust linear optimization with applications to inventory control

**Seong-Cheol Kang · Theodora S. Brisimi ·
Ioannis Ch. Paschalidis**

the date of receipt and acceptance should be inserted later

Abstract This paper tackles linear programming problems with data uncertainty and applies it to an important inventory control problem. Each element of the constraint matrix is subject to uncertainty and is modeled as a random variable with a bounded support. The classical robust optimization approach to this problem yields a solution with guaranteed feasibility. As this approach tends to be too conservative when applications can tolerate a small chance of infeasibility, one would be interested in obtaining a less conservative solution with a certain probabilistic guarantee of feasibility. A robust formulation in the literature produces such a solution, but it does not use any distributional information on the uncertain data. In this work, we show that the use of distributional information leads to an equally robust solution (i.e., under the same probabilistic guarantee of feasibility) but with a better objective value. In particular, by exploiting distributional information, we establish stronger upper bounds on the constraint violation probability of a solution. These bounds enable us to "inject" less conservatism into the formulation, which in turn yields a more cost-effective solution (by 50% or more in some numerical instances). To illustrate the effectiveness of our methodology, we consider a discrete-time stochastic inventory control problem with certain quality of service constraints. Numerical tests demonstrate that the use of distributional information in the robust optimization of the inventory control problem results in 36%–54% cost savings, compared to the case where such information is not used.

Research partially supported by the NSF under grants EFRI-0735974, CNS-1239021, IIS-1237022, by the DOE under grant DE-FG52-06NA27490, by the ARO under grants W911NF-11-1-0227 and W911NF-12-1-0390, and by the ONR under grant N00014-10-1-0952.

Seong-Cheol Kang
The Korea Transport Institute, Goyang-si, Korea, e-mail: jsckang@koti.re.kr

Theodora S. Brisimi
Department of Electrical & Computer Eng., Boston University, Boston, MA 02215, USA, e-mail: tbrisimi@bu.edu

Ioannis Ch. Paschalidis
Corresponding author. Department of Electrical & Computer Eng. and Division of Systems Eng., Boston University, Boston, MA 02215, USA, e-mail: yannisp@bu.edu, url: http://ionia.bu.edu/

Keywords Robust optimization, Linear programming, Data uncertainty, Inventory Control, Quality-of-Service.

1 Introduction

Linear programming (LP) is a ubiquitous tool in manufacturing systems and service operations. Theory and methods are very mature, there are many excellent solvers to choose from, and LPs can be solved in polynomial time with interior-point methods. Alas, the world is neither (always) linear nor certain. In this paper we focus on the latter shortcoming of LP-based modeling, that is, the presence of uncertainty in the problem data.

Assuming *certainty equivalence* offers a way to deal with uncertainty: for every uncertain data element use a nominal value (usually its mean) and form a *nominal formulation* which remains an LP formulation. A solution obtained in this manner, however, is non-robust 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 useless.

For this reason, the classical robust linear optimization approach focused on a formulation, which we will refer to as the *fat formulation*, whose solution is immune to data uncertainty (i.e., protected against infeasibility). In the early 1970's, Soyster [Soy73] considered a convex optimization problem with a linear objective function and a set-inclusive constraint that requires nonnegative combinations of convex sets to be contained in another convex set. As a special case, Soyster noted that the problem can be viewed as an "inexact LP problem," where each column vector of the constraint matrix $\mathbf{A}$ is only known to belong to a convex set. In our terminology, it is the fat formulation of an LP problem where the constraint matrix has *column-wise uncertainty*. Soyster showed that the fat formulation can be recast as an LP formulation.

Ben-Tal and Nemirovski [BTN99] pointed out that the case of column-wise uncertainty used in [Soy73] is extremely conservative. They instead considered *row-wise uncertainty* where the rows of the constraint matrix $\mathbf{A}$ are known to belong to given convex sets. In this case, they argued that the fat formulation does not necessarily lead to an LP formulation; for example, when the uncertainty sets for the rows of $\mathbf{A}$ are ellipsoids, the fat formulation turns out to be a conic quadratic problem.

Clearly, the guaranteed feasibility makes the fat formulations (i.e., the worst-case approach) of [Soy73] and [BTN99] suitable for applications in which infeasibility cannot be accepted at all (e.g., design of engineering structures like bridges considered in Ben-Tal and Nemirovski [BTN99]). When applications can tolerate a small chance of infeasibility, however, solutions from those formulations tend to be too conservative. This is especially the case if the worst-case occurs very rarely. Hence for the latter class of applications, methods that produce less conservative solutions with certain probabilistic guarantees of feasibility could prove to be useful.

To that end, Bertsimas and Sim [BS04] considered a "relaxed" robust linear optimization. In contrast to the column-wise and row-wise uncertainty used in [Soy73] and in [BTN99], respectively, they considered *element-wise uncertainty* where each uncertain element of the constraint matrix $\mathbf{A}$ is modeled as an independent bounded random variable. They assumed that the probability distributions of the uncertain elements are unknown except that they are symmetric. To construct a less conservative formulation than the fat formulation, which we will refer to as the *robust formulation*, they used a set of parameters that "restrict" the variability of the uncertain elements. They showed that their robust formulation can be recast as an LP formulation, which is favorable from a computational point of view. The difference between the objective values of the nominal formulation and their robust formulation was termed the *price of robustness*. For a probabilistic guarantee of feasibility of an optimal solution of their robust formulation, they established upper bounds on the constraint violation probability of the solution.

As in [BS04], the robust model used in this paper is a special case of the robust counterpart of a linear optimization model when uncertain coefficients are assumed to belong to a D-norm induced uncertainty set (see [BPS04]). Unlike, however, the earlier work, we exploit distributional information on the uncertain elements in robust linear optimization. We will show that by using full or limited distributional information, one can obtain a better solution than earlier approaches, but with the same probabilistic guarantee of feasibility. The crux of the matter is that by exploiting distributional information, we can establish stronger bounds on the constraint violation probability. This enables us to "inject" less conservatism into the formulation, which in turn yields a more cost-effective solution. In particular, we establish three types of bounds on the constraint violation probability: (*i*) bounds that use the full distribution of the uncertain problem data, (*ii*) bounds that use the full distribution of data and the optimal solution to the robust problem, and (*iii*) bounds that use up to the first two moments of the problem data.

Our motivation comes from the emerging abundance of data in many real-world applications. By mining these data suitably, one can obtain highly reliable distributional information and, as we will show, put it to good use. When data are not available and distributional information can not be obtained, then the approach of [BS04] is appropriate. However, our work can help quantify the benefits that can result from data collection and implementation of estimation techniques for obtaining distributional information. We expect that in many settings these benefits can exceed the associated costs. In this spirit, one can think of the gain in objective value that stems from our robust optimization approach as the *estimation discount on the price of robustness*.

The work in this paper has parallels with ideas developed in the context of *robust control*. Using the terminology in [DDB95], we are interested in parametric uncertainty and consider a stochastic model of uncertainty rather than a deterministic model which would necessitate a worst case analysis. To illustrate the effectiveness of our robust optimization approach, we apply it to a stochastic inventory control problem.[1] Inventory control has, of course, a long history going back to the seminal results of Clark and Scarf [CS60]. We are interested in inventory control problems that enforce explicit *quality of service (QoS)* or *service level* constraints. Under stochastic demand and production models, such problems have been analyzed in Paschalidis and Liu [PL03], Bertsimas and Paschalidis [BP01], Paschalidis et al. [PLCP04], and Del Vecchio and Paschalidis [DVP06] using large

[1] For other applications of robust optimization, refer to Bertsimas et al. [BBC11] and Ben-Tal et al. [BTGN09].

deviations techniques. Yet, the analysis is complex and uses the special structure of the models to solve the corresponding large deviations problem. Formulating the inventory control problems as static optimization problems ignores the inherent uncertainty but has the advantage that many modeling complexities (e.g., lead times, ordering costs, etc.) can be more easily incorporated.

Bertsimas and Thiele [BT06] addressed inventory control problems to minimize total ordering, holding, and shortage costs from the robust optimization perspective of [BS04]. A different robust model for inventory management was considered by See and Sim [SS10]. Bienstock and Ozbay [BO08] propose an algorithm to iteratively solve a robust min-max problem for setting the base-stock level in a single buffer under uncertain demand. Here, we consider an inventory control problem similar to one in [BT06], but our approach differs from theirs in two aspects. First, we incorporate certain QoS constraints into the problem instead of using shortage costs. QoS constraints are often used in managing supply chains, partly because shortage costs are hard to quantify. Our second, and more subtle, difference with [BT06] is the way we construct the robust formulation. This will be elaborated on in Section 3.2.

The paper is organized as follows. In Section 2, and in the form of background, we deal with robust optimization for an LP problem with element-wise uncertainty. We follow [BS04] to construct the robust formulation for the LP problem and derive its equivalent LP formulation. Using distributional information on the uncertain elements, we develop new stronger bounds on the constraint violation probability. We explain that stronger bounds lead to an equally robust solution with a better objective value (i.e., a more cost-effective solution under the same probabilistic guarantee of feasibility). The case of using limited distributional information in the form of the first and second moments as well as full distributional information is investigated. In Section 3, we consider a discrete-time stochastic inventory control problem with QoS constraints. We form the robust formulation of the problem and show that the optimal ordering quantities of this formulation correspond to a base-stock policy. Using the results of Section 2, bounds on the probability that the optimal ordering quantities violate the QoS constraints are developed. Through numerical tests, we demonstrate the cost-effectiveness of our approach. Finally, concluding remarks are given in Section 4.

Notational Conventions: Throughout the paper, we use boldface lowercase letters to denote vectors and boldface uppercase letters to represent matrices. Occasionally we use boldface uppercase Greek letters to denote vectors. All vectors are assumed to be column vectors. When we specify the components of a vector, however, we write $\mathbf{x} = (x_1, \ldots, x_n)$ for the column vector $\mathbf{x}$ to save space. $\mathbf{x}'$ represents the transpose of $\mathbf{x}$; however sometimes it also means a vector different from $\mathbf{x}$. It will be clear from the context which one is the case. The vector of all zeros is denoted by $\mathbf{0}$. $\mathbf{A} = (a_{ij})$ denotes a matrix whose (i, j)th element is a_{ij}. We will use $\mathbf{a}_i$ to denote the ith row of $\mathbf{A}$ which, using our convention, we will assume to be a column vector. $\mathbf{P}[A]$ means the probability of the event A. For a random variable X, we write $\mathbf{E}[X]$ and $\mathbf{Var}(X)$ for its mean and variance, respectively. The floor function of a real number x, denoted by $\lfloor x \rfloor$, returns the largest integer less than or equal to x. The ceiling function, denoted by $\lceil x \rceil$, gives the smallest integer not less than x.

2 Robust Linear Optimization

2.1 A Linear Programming Problem with Data Uncertainty

Let us consider the LP problem[2]

$$\begin{array}{rl} \text{maximize} & \mathbf{c}'\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \end{array} \tag{1}$$

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, and $\mathbf{x} \in \mathbb{R}^n$ is the vector of decision variables. We assume, without loss of generality, that only the elements of the matrix $\mathbf{A}$ are subject to uncertainty. Indeed, if $\mathbf{c}$ and $\mathbf{b}$ are also uncertain, the problem can be reformulated so that uncertainty exists only in the entries of $\mathbf{A}$ ([BT06], [BTGN09]).

Each uncertain element of $\mathbf{A}$ is modeled as a random variable whose range is *symmetrically bounded* around its mean (or the nominal value). In particular, we assume that $a_{ij} \in [\overline{a}_{ij} - \hat{a}_{ij}, \overline{a}_{ij} + \hat{a}_{ij}]$, where $\overline{a}_{ij} = \mathbf{E}[a_{ij}]$ and $\hat{a}_{ij} \geq 0$. Note that if $\hat{a}_{ij} = 0$, the corresponding a_{ij} is deterministic and takes the value $\overline{a}_{ij}$. For each row i of $\mathbf{A}$, we define $J_i \triangleq \{j \mid \hat{a}_{ij} > 0\}$, i.e., $J_i \triangleq \{j \mid a_{ij} \text{ is uncertain}\}$. We assume that a_{ij}, for all i and $j \in J_i$, are independent of each other. It will be generally assumed that the probability distribution of each uncertain a_{ij} is known. However, we will also consider the case that only limited distributional information on a_{ij} is available, such as the first moment, or the first and second moments of a_{ij}. The following symmetry assumption on the distribution will be in effect for some of the results we will present.

Assumption A
For all $i, j \in J_i$, the probability distribution of a_{ij} is symmetric *over $[\overline{a}_{ij} - \hat{a}_{ij}, \overline{a}_{ij} + \hat{a}_{ij}]$, that is,*

$$F_{a_{ij}}(\overline{a}_{ij} - a) = 1 - F_{a_{ij}}(\overline{a}_{ij} + a), \quad 0 \leq a \leq \hat{a}_{ij},$$

where $F_{a_{ij}}$ is the cumulative distribution function of a_{ij}.

Given the data uncertainty structure for $\mathbf{A}$, one may elect to solve the *nominal formulation*, where each uncertain a_{ij} is replaced by its mean value:

$$\begin{array}{rl} z_N = \text{maximize} & \mathbf{c}'\mathbf{x} \\ \text{subject to} & \sum_j \overline{a}_{ij} x_j \leq b_i, \quad i = 1, \ldots, m, \\ & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}. \end{array} \tag{2}$$

One disadvantage of the nominal formulation is that its optimal solution is likely to violate the constraints $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ of (1). This leads one to consider a solution that is guaranteed to satisfy $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ for all realizations of the uncertain a_{ij}'s, while maximizing the objective value. To obtain such a solution, one needs to solve the *fat formulation*

$$\begin{array}{rl} z_F = \text{maximize} & \mathbf{c}'\mathbf{x} \\ \text{subject to} & \max_{\mathbf{a}_i \in \mathscr{U}_i} \{\mathbf{a}_i'\mathbf{x}\} \leq b_i, \quad i = 1, \ldots, m, \\ & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \end{array} \tag{3}$$

[2] The main role of the constraints $\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$ is to make the feasible region bounded, so that the case of unbounded objective value can be avoided. Other than that, they do not play any other role in the analysis that follows.

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

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

It is not difficult to see that the formulation (3) can be written as the LP formulation

$$\begin{aligned} z_F = \text{maximize} \quad & \mathbf{c}'\mathbf{x} && (4)\\ \text{subject to} \quad & \sum_j \overline{a}_{ij}x_j + \sum_{j\in J_i} \hat{a}_{ij}y_j \le b_i, \quad i = 1,\ldots,m, \\ & -y_j \le x_j \le y_j, \quad \forall\, j \in \textstyle\bigcup_1^m J_i, && (4a)\\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}. \end{aligned}$$

We note that the constraints (4a) can be replaced with $-\mathbf{y} \le \mathbf{x} \le \mathbf{y}$. This possibly adds to the formulation the additional constraints $-y_j \le x_j \le y_j$ for all $j \notin \bigcup_1^m J_i$. Their presence, however, does not alter the optimal solution of (4). The following Lemma is pretty standard in robust optimization, hence, we skip the proof.

Lemma 1 *Let $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ be an optimal solution of (4). Then $\hat{\mathbf{x}}$ is a feasible solution of (1) for every possible realization of $\mathbf{A}$. Moreover, $z_F \le z_N$.*

As Lemma 1 implies, by solving the fat formulation, one may obtain a solution with an inferior objective value in exchange for its *full robustness* to (i.e., immunity against) data uncertainty.

2.2 The Robust Formulation

In order to construct a formulation that is less conservative than the fat formulation (3), we follow Bertsimas and Sim [BS04] and introduce the *uncertainty budget* $\Gamma_i \in [0, |J_i|]$ for each row $i = 1,\ldots,m$ of the matrix $\mathbf{A}$. The role of the uncertainty budget Γ_i is to impose the *uncertainty budget constraint*

$$\sum_{j\in J_i} \frac{|a_{ij} - \overline{a}_{ij}|}{\hat{a}_{ij}} \le \Gamma_i.$$

That is, Γ_i limits the deviations of a_{ij}, $\forall\, j \in J_i$, from their respective mean values by imposing an ℓ_1-norm constraint on the vector of $\left(\frac{a_{ij}-\overline{a}_{ij}}{\hat{a}_{ij}}\right)$ for $j \in J_i$. The motivation for this constraint comes from the expectation that not all uncertain elements of the ith row can take their extreme values at the same time. We define the *restricted uncertainty set* $\mathscr{R}_i(\Gamma_i)$ as

$$\mathscr{R}_i(\Gamma_i) \triangleq \left\{\mathbf{a}_i \mid a_{ij} \in [\overline{a}_{ij} - \hat{a}_{ij}, \overline{a}_{ij} + \hat{a}_{ij}], \forall\, j;\ \sum_{j\in J_i} \frac{|a_{ij} - \overline{a}_{ij}|}{\hat{a}_{ij}} \le \Gamma_i\right\}.$$

In other words, $\mathscr{R}_i(\Gamma_i)$ is the set of all realizations of $\mathbf{a}_i$ that satisfy the uncertainty budget constraint.

We define the *robust formulation* as

$$\begin{aligned} z_R(\mathbf{\Gamma}) = \text{maximize} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)}\{\mathbf{a}_i'\mathbf{x}\} \le b_i, \quad i = 1,\ldots,m, && (5)\\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}, \end{aligned}$$

where $z_R(\mathbf{\Gamma})$ denotes the optimal objective value for a given $\mathbf{\Gamma} = (\Gamma_1, \ldots, \Gamma_m)$. The following Lemma establishes a useful monotonicity property; we include a proof in Appendix A for completeness.

Lemma 2 *$z_R(\mathbf{\Gamma})$ is nonincreasing as $\mathbf{\Gamma}$ increases componentwise, and $z_F \leq z_R(\mathbf{\Gamma}) \leq z_N$.*

Based on Lemma 2, one is able to strike a balance between the robustness of a solution and its objective value by varying Γ_i between 0 and $|J_i|$ for all i. If $\Gamma_i = |J_i|$, the ith constraint is *fully protected* against violation. If $\Gamma_i = 0$, on the other hand, the ith constraint is *not protected*. In this respect, Γ_i can also be viewed as the *level of protection* for the ith constraint. The difference $z_N - z_R(\mathbf{\Gamma})$ is the degradation of the objective value that results from improving the level of protection of the constraints by selecting $\mathbf{\Gamma}$.

We now show that the robust formulation (5) can be recast as an LP formulation. The formulation we obtain is equivalent to the one in Bertsimas and Sim [BS04]; we provide a more concise proof for completeness and because the derivation is useful in deriving our bounds and in generalizing to asymmetrically bounded random variables as well as different types of the uncertainty budget constraint.

For any $\mathbf{x}$, the maximization problem in the ith constraint of (5) can be written as

$$\begin{aligned}
\text{maximize} \quad & \mathbf{a}_i'\mathbf{x} && (6)\\
\text{subject to} \quad & a_{ij} \leq \overline{a}_{ij} + \hat{a}_{ij}, \quad \forall j, && (6a)\\
& a_{ij} \geq \overline{a}_{ij} - \hat{a}_{ij}, \quad \forall j, && (6b)\\
& \sum_{j \in J_i} w_{ij} \leq \Gamma_i, && (6c)\\
& w_{ij} \geq (a_{ij} - \overline{a}_{ij})/\hat{a}_{ij}, \quad \forall j \in J_i, && (6d)\\
& w_{ij} \geq -(a_{ij} - \overline{a}_{ij})/\hat{a}_{ij}, \quad \forall j \in J_i, && (6e)\\
& w_{ij} \geq 0, \quad \forall j \in J_i,
\end{aligned}$$

where a_{ij}, $\forall j$ and w_{ij}, $\forall j \in J_i$ are the decision variables. Let $\lambda_{ij}, \mu_{ij}, z_i, \nu_{ij}, \tau_{ij}$ be the dual variables for the constraints (6a)–(6e), respectively. Then the dual of (6) is given by (after some simplifications)

$$\begin{aligned}
\text{minimize} \quad & \sum_j \overline{a}_{ij} x_j + \sum_{j \in J_i} \hat{a}_{ij}(\lambda_{ij} + \mu_{ij}) + \Gamma_i z_i && (7)\\
\text{subject to} \quad & \lambda_{ij} - \mu_{ij} + \nu_{ij} - \tau_{ij} = x_j, \quad \forall j \in J_i,\\
& z_i - \hat{a}_{ij}(\nu_{ij} + \tau_{ij}) \geq 0, \quad \forall j \in J_i,\\
& \lambda_{ij}, \mu_{ij}, \nu_{ij}, \tau_{ij} \geq 0, \quad \forall j \in J_i,\\
& z_i \geq 0.
\end{aligned}$$

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

$$\begin{aligned}
z_R(\mathbf{\Gamma}) = \text{maximize} \quad & \mathbf{c}'\mathbf{x} && (8)\\
\text{subject to} \quad & \sum_j \overline{a}_{ij} x_j + \sum_{j \in J_i} \hat{a}_{ij}(\lambda_{ij} + \mu_{ij}) + \Gamma_i z_i \leq b_i \quad i = 1, \ldots, m,
\end{aligned}$$

$$\lambda_{ij} - \mu_{ij} + \nu_{ij} - \tau_{ij} = x_j, \quad i = 1, \ldots, m, \quad \forall\, j \in J_i, \tag{8a}$$
$$z_i - \hat{a}_{ij}(\nu_{ij} + \tau_{ij}) \geq 0, \quad i = 1, \ldots, m, \quad \forall\, j \in J_i, \tag{8b}$$
$$\lambda_{ij}, \mu_{ij}, \nu_{ij}, \tau_{ij} \geq 0, \quad i = 1, \ldots, m, \quad \forall\, j \in J_i,$$
$$z_i \geq 0, \quad i = 1, \ldots, m,$$
$$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}.$$

Proof: Let $(\mathbf{x}^*, \lambda^*_{ij}, \mu^*_{ij}, z^*_i, \nu^*_{ij}, \tau^*_{ij})$ be an optimal solution of (8). Let $\tilde{\mathbf{x}}$ be an optimal solution of (5). We will establish the equivalence by showing that $\mathbf{x}^*$ is a feasible solution of (5) with $\mathbf{c}'\mathbf{x}^* = \mathbf{c}'\tilde{\mathbf{x}}$.

Fix $\mathbf{x} = \mathbf{x}^*$ in (6) and (7), and let $(\mathbf{a}^*_i, w^*_{ij})$ be an optimal solution of (6). Since $(\lambda^*_{ij}, \mu^*_{ij}, z^*_i, \nu^*_{ij}, \tau^*_{ij})$ is a feasible solution of (7), the weak duality between (6) and (7) yields

$$\mathbf{a}^{*\prime}_i \mathbf{x}^* \leq \sum_j \overline{a}_{ij} x^*_j + \sum_{j \in J_i} \hat{a}_{ij}(\lambda^*_{ij} + \mu^*_{ij}) + \Gamma_i z^*_i.$$

From $\mathbf{a}^{*\prime}_i \mathbf{x}^* = \max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)} \{\mathbf{a}'_i \mathbf{x}^*\}$ and the feasibility of $(\mathbf{x}^*, \lambda^*_{ij}, \mu^*_{ij}, z^*_i, \nu^*_{ij}, \tau^*_{ij})$ to (8), we have for all i

$$\max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)} \{\mathbf{a}'_i \mathbf{x}^*\} \leq \sum_j \overline{a}_{ij} x^*_j + \sum_{j \in J_i} \hat{a}_{ij}(\lambda^*_{ij} + \mu^*_{ij}) + \Gamma_i z^*_i \leq b_i.$$

This shows that $\mathbf{x}^*$ is feasible to (5), implying that $\mathbf{c}'\mathbf{x}^* \leq \mathbf{c}'\tilde{\mathbf{x}}$.

Next, set $\mathbf{x} = \tilde{\mathbf{x}}$ in (6) and (7), and let $(\tilde{\mathbf{a}}_i, \tilde{w}_{ij})$ be an optimal solution of (6). By the strong duality, there exists a feasible $(\tilde{\lambda}_{ij}, \tilde{\mu}_{ij}, \tilde{z}_i, \tilde{\nu}_{ij}, \tilde{\tau}_{ij})$ to (7) such that

$$\max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)} \{\mathbf{a}'_i \tilde{\mathbf{x}}\} = \tilde{\mathbf{a}}'_i \tilde{\mathbf{x}} = \sum_j \overline{a}_{ij} \tilde{x}_j + \sum_{j \in J_i} \hat{a}_{ij}(\tilde{\lambda}_{ij} + \tilde{\mu}_{ij}) + \Gamma_i \tilde{z}_i.$$

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

$$b_i \geq \max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)} \{\mathbf{a}'_i \tilde{\mathbf{x}}\} = \sum_j \overline{a}_{ij} \tilde{x}_j + \sum_{j \in J_i} \hat{a}_{ij}(\tilde{\lambda}_{ij} + \tilde{\mu}_{ij}) + \Gamma_i \tilde{z}_i.$$

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

It can be shown that the LP formulation (8) can be reshaped into the following LP formulation derived in Bertsimas and Sim [BS04]; see Appendix B for an explanation.

$$
\begin{aligned}
z_R(\mathbf{\Gamma}) = \text{maximize} \quad & \mathbf{c}'\mathbf{x} && \quad (9)\\
\text{subject to} \quad & \sum_j \overline{a}_{ij} x_j + \Gamma_i z_i + \sum_{j \in J_i} p_{ij} \leq b_i, && i = 1, \ldots, m,\\
& z_i + p_{ij} \geq \hat{a}_{ij} y_j, && i = 1, \ldots, m, \quad \forall\, j \in J_i,\\
& p_{ij} \geq 0, && i = 1, \ldots, m, \quad \forall\, j \in J_i,\\
& z_i \geq 0, && i = 1, \ldots, m,\\
& -y_j \leq x_j \leq y_j, && \forall\, j \in \textstyle\bigcup_1^m J_i,\\
& y_j \geq 0, && \forall\, j \in \textstyle\bigcup_1^m J_i,\\
& \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}.
\end{aligned}
$$

2.3 Bounds on the Constraint Violation Probability

For a given $\mathbf{\Gamma} > \mathbf{0}$, let $\mathbf{x}^*$ be an optimal solution of the robust formulation (5), which is obtained by solving the LP formulation (8) (or (9)). Let us call $\mathbf{x}^*$ the *robust solution*. To avoid degenerate cases and without loss of generality, we assume $|\mathbf{x}^*| > \mathbf{0}$.[3] Unless $\Gamma_i = |J_i|$ for all i, $\mathbf{x}^*$ may well violate the constraints $\mathbf{Ax} \leq \mathbf{b}$ of (1) due to the uncertainty in $\mathbf{A}$.

Computing the constraint violation probability $\mathbf{P}\left[\sum_j a_{ij}x_j^* > b_i\right]$, $i = 1, \ldots, m$, exactly is often a challenging task. Therefore one would be interested in upper bounds on this probability. Bertsimas and Sim [BS04] assumed that the probability distributions of the random a_{ij}'s are unknown except that they are symmetric. Under this assumption, they derived the following bound on the ith constraint violation probability:

$$\mathbf{P}\left[\sum_j a_{ij}x_j^* > b_i\right] \leq \exp\left(-\frac{\Gamma_i^2}{2|J_i|}\right). \tag{10}$$

Henceforth, we will use the terminology "Bound (10)" to refer to the right hand side of (10) and similar terminologies for such bounds. Bound (10) is an *a priori* bound in the sense that it is not a function of the robust solution $\mathbf{x}^*$. Moreover, it does not use any distributional information on a_{ij}, $\forall\, j \in J_i$, other than the symmetry of the probability distributions. For these reasons, Bound (10) can be weaker than other bounds that exploit the distributional information and/or the robust solution $\mathbf{x}^*$.

In this section, we develop new bounds on the constraint violation probability, which are stronger than Bound (10). This will be accomplished by making use of full or limited distributional information on a_{ij}'s. Obtaining stronger bounds is important because, as we shall show later, they lead to a more cost-effective solution under the same probabilistic guarantee of feasibility. The following lemma will be used in the results that we will present.

Lemma 3 *Let S_i be a subset of J_i such that $|S_i| = \lfloor \Gamma_i \rfloor$, and let $t_i \in J_i \setminus S_i$. Then*

$$\max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)} \{\mathbf{a}_i'\mathbf{x}^*\} = \sum_j \overline{a}_{ij}x_j^* + \max_{S_i \cup \{t_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it_i}|x_{t_i}^*| \right\}, \tag{11}$$

where it is understood that $\{t_i\} = \emptyset$ if Γ_i is an integer.

Proof: For all $j \notin J_i$, the left hand side of (11) takes the constant value $\sum_{j \notin J_i} \overline{a}_{ij}x_j^*$. Sort x_j^*, $\forall\, j \in J_i$, in the nonincreasing order of $|x_j^*|$. Let $x_{s_1}^*, \ldots, x_{s_{\lfloor \Gamma_i \rfloor}}^*$ be the first $\lfloor \Gamma_i \rfloor$ elements in that order. To maximize the left hand side of (11), for $k = 1, \ldots, \lfloor \Gamma_i \rfloor$, we choose $\overline{a}_{is_k} + \hat{a}_{is_k}$ if $x_{s_k}^* > 0$ and $\overline{a}_{is_k} - \hat{a}_{is_k}$ otherwise. When Γ_i is not an integer, we can consider one more element, $x_{s_{\lfloor \Gamma_i \rfloor + 1}}^*$. If it is positive, we choose $\overline{a}_{is_{\lfloor \Gamma_i \rfloor + 1}} + (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{is_{\lfloor \Gamma_i \rfloor + 1}}$; otherwise, we choose $\overline{a}_{is_{\lfloor \Gamma_i \rfloor + 1}} - (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{is_{\lfloor \Gamma_i \rfloor + 1}}$. Note the scaling factor $(\Gamma_i - \lfloor \Gamma_i \rfloor)$ in the last step, which ensures $\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i)$. The right hand side of (11) achieves the same goal. □

[3] If not, we can always fix to zero all zero components of $\mathbf{x}^*$ and recast the LP problem in a lower-dimensional space.

2.3.1 A Distribution-Dependent Bound

We first derive an *a priori* bound that utilizes the probability distributions of the random a_{ij}'s. Let $\eta_{ij} = (a_{ij} - \overline{a}_{ij})/\hat{a}_{ij}$ for all i and $j \in J_i$. Define the logarithmic moment generating function of η_{ij} as $\Lambda_{\eta_{ij}}(\theta) \triangleq \log \mathbf{E}[e^{\theta\eta_{ij}}]$.

Theorem 2 *Let Assumption A be in effect.*
(a) For the ith constraint

$$\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] \leq \exp\Big(-\sup_{\theta\geq 0}\Big\{\theta\Gamma_i - \sum_{j\in J_i}\Lambda_{\eta_{ij}}(\theta)\Big\}\Big). \tag{12}$$

(b) Bound (12) < Bound (10).

Proof: (a) The proof of this part proceeds similarly to the proofs of Proposition 2 and Theorem 2 of Bertsimas and Sim [BS04]. Under Assumption A, the random variables η_{ij} are symmetrically distributed over $[-1, 1]$. Let $S_i^* \cup \{t_i^*\} = \arg\max_{S_i\cup\{t_i\}}\{\sum_{j\in S_i}\hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i}|x_{t_i}^*|\}$ in (11), where again it is understood that $\{t_i^*\} = \emptyset$ if Γ_i is an integer. Let $r = \mathrm{argmin}_{j\in S_i^*\cup\{t_i^*\}}\hat{a}_{ij}|x_j^*|$. (Notice that if Γ_i is an integer, $r = \mathrm{argmin}_{j\in S_i^*}\hat{a}_{ij}|x_j^*|$; otherwise, $r = t_i^*$.) We have

$$\begin{aligned}
\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] &= \mathbf{P}\Big[\sum_j \overline{a}_{ij}x_j^* + \sum_{j\in J_i}\hat{a}_{ij}x_j^*\eta_{ij} > b_i\Big] \\
&= \mathbf{P}\Big[\sum_{j\in J_i}\hat{a}_{ij}|x_j^*|\eta_{ij} > b_i - \sum_j \overline{a}_{ij}x_j^*\Big] \\
&\leq \mathbf{P}\Big[\sum_{j\in J_i}\hat{a}_{ij}|x_j^*|\eta_{ij} > \sum_{j\in S_i^*}\hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*|\Big] \\
&= \mathbf{P}\Big[\sum_{j\in J_i\setminus S_i^*}\hat{a}_{ij}|x_j^*|\eta_{ij} > \sum_{j\in S_i^*}\hat{a}_{ij}|x_j^*|(1-\eta_{ij}) + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*|\Big] \\
&\leq \mathbf{P}\Big[\sum_{j\in J_i\setminus S_i^*}\hat{a}_{ij}|x_j^*|\eta_{ij} > \hat{a}_{ir}|x_r^*|\Big\{\sum_{j\in S_i^*}(1-\eta_{ij}) + \Gamma_i - \lfloor\Gamma_i\rfloor\Big\}\Big] \\
&= \mathbf{P}\Big[\sum_{j\in J_i\setminus S_i^*}\frac{\hat{a}_{ij}|x_j^*|}{\hat{a}_{ir}|x_r^*|}\eta_{ij} + \sum_{j\in S_i^*}\eta_{ij} > \Gamma_i\Big] \\
&= \mathbf{P}\Big[\sum_{j\in J_i}\gamma_{ij}\eta_{ij} > \Gamma_i\Big] \leq \mathbf{P}\Big[\sum_{j\in J_i}\gamma_{ij}\eta_{ij} \geq \Gamma_i\Big],
\end{aligned}$$

where

$$\gamma_{ij} = \begin{cases} \hat{a}_{ij}|x_j^*|/\hat{a}_{ir}|x_r^*| & \text{if } j \in J_i \setminus S_i^*, \\ 1 & \text{if } j \in S_i^*. \end{cases}$$

The first equality follows from the definitions of η_{ij}'s, and the second equality from the fact that $\hat{a}_{ij}x_j^*\eta_{ij}$ is equal to $\hat{a}_{ij}|x_j^*|\eta_{ij}$ in distribution. The first inequality is due to the feasibility of $\mathbf{x}^*$ in (5), namely (cf. Lemma 3),

$$\max_{\mathbf{a}_i\in\mathscr{R}_i(\Gamma_i)}\{\mathbf{a}_i'\mathbf{x}^*\} = \sum_j \overline{a}_{ij}x_j^* + \sum_{j\in S_i^*}\hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*| \leq b_i.$$

The second inequality follows from $\hat{a}_{ir}|x_r^*| = \min_{j\in S_i^*\cup\{t_i^*\}} \hat{a}_{ij}|x_j^*|$. The fourth equality holds because $\sum_{j\in S_i^*} 1 = \lfloor \Gamma_i \rfloor$. Note that $0 \le \gamma_{ij} \le 1$, $\forall\, j \in J_i$, because $\hat{a}_{ir}|x_r^*| \ge \hat{a}_{ij}|x_j^*|$, $\forall\, j \in J_i \setminus S_i^*$. (If $\hat{a}_{ir}|x_r^*| < \hat{a}_{ij}|x_j^*|$ for some $j \in J_i \setminus S_i^*$, r cannot belong to $S_i^* \cup \{t_i^*\}$.)

Using Markov's inequality, for all $\theta \ge 0$ we obtain

$$\begin{aligned}
\mathbf{P}\Big[\sum_{j\in J_i} \gamma_{ij}\eta_{ij} \ge \Gamma_i\Big] &\le e^{-\theta\Gamma_i}\mathbf{E}\big[e^{\theta\sum_{j\in J_i}\gamma_{ij}\eta_{ij}}\big] \\
&= e^{-\theta\Gamma_i}\prod_{j\in J_i}\mathbf{E}\big[e^{\theta\gamma_{ij}\eta_{ij}}\big] \\
&= e^{-\theta\Gamma_i}\prod_{j\in J_i}\int_{-1}^{1}\sum_{k=0}^{\infty}\frac{(\theta\gamma_{ij}\eta)^k}{k!}dF_{\eta_{ij}}(\eta) \\
&\le e^{-\theta\Gamma_i}\prod_{j\in J_i}\int_{-1}^{1}\sum_{k=0}^{\infty}\frac{(\theta\eta)^k}{k!}dF_{\eta_{ij}}(\eta) \qquad (13)\\
&= e^{-\theta\Gamma_i}\prod_{j\in J_i}\mathbf{E}\big[e^{\theta\eta_{ij}}\big] \\
&= \exp\Big(-\theta\Gamma_i + \sum_{j\in J_i}\Lambda_{\eta_{ij}}(\theta)\Big).
\end{aligned}$$

The first equality follows from the independence of the random variables η_{ij}, and the second equality from the Maclaurin series for $e^{\theta\gamma_{ij}\eta_{ij}}$. The second inequality is due to the following two properties that result from the symmetry of the probability distribution of η_{ij} and $0 \le \gamma_{ij} \le 1$: for all $k = 0, 1, \ldots,$

$$\begin{aligned}
\int_{-1}^{1}\frac{(\theta\gamma_{ij}\eta)^{2k+1}}{(2k+1)!}dF_{\eta_{ij}}(\eta) &= \int_{-1}^{1}\frac{(\theta\eta)^{2k+1}}{(2k+1)!}dF_{\eta_{ij}}(\eta) = 0, \\
\int_{-1}^{1}\frac{(\theta\gamma_{ij}\eta)^{2k}}{(2k)!}dF_{\eta_{ij}}(\eta) &\le \int_{-1}^{1}\frac{(\theta\eta)^{2k}}{(2k)!}dF_{\eta_{ij}}(\eta).
\end{aligned}$$

Optimizing over θ, we obtain

$$\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] \le \exp\Big(-\sup_{\theta\ge 0}\Big\{\theta\Gamma_i - \sum_{j\in J_i}\Lambda_{\eta_{ij}}(\theta)\Big\}\Big).$$

(b) For any $\theta \ge 0$ and symmetric probability distribution for a_{ij}, $\forall\, j \in J_i$,

$$\begin{aligned}
\theta\Gamma_i - \sum_{j\in J_i}\Lambda_{\eta_{ij}}(\theta) &= \theta\Gamma_i - \sum_{j\in J_i}\log\int_{-1}^{1}\sum_{k=0}^{\infty}\frac{(\theta\eta)^k}{k!}dF_{\eta_{ij}}(\eta) \\
&= \theta\Gamma_i - \sum_{j\in J_i}\log\int_{-1}^{1}\sum_{k=0}^{\infty}\frac{(\theta\eta)^{2k}}{(2k)!}dF_{\eta_{ij}}(\eta) \\
&\ge \theta\Gamma_i - \sum_{j\in J_i}\log\sum_{k=0}^{\infty}\frac{\theta^{2k}}{(2k)!}\int_{-1}^{1}dF_{\eta_{ij}}(\eta) \\
&= \theta\Gamma_i - \sum_{j\in J_i}\log\sum_{k=0}^{\infty}\frac{\theta^{2k}}{(2k)!}
\end{aligned}$$

$$\geq \theta\Gamma_i - \sum_{j\in J_i} \log \sum_{k=0}^{\infty} \frac{\theta^{2k}}{2^k k!}$$
$$= \theta\Gamma_i - \sum_{j\in J_i} \log e^{\theta^2/2},$$

where the second equality follows from the symmetry of the probability distributions, and the first inequality from $\eta_{ij} \in [-1, 1]$. Note that the equality holds throughout only when $\mathbf{P}[\eta_{ij} = -1] = \mathbf{P}[\eta_{ij} = 1] = 1/2$ and $\theta = 0$. However since

$$\frac{\partial}{\partial\theta}\left\{\theta\Gamma_i - \sum_{j\in J_i} \Lambda_{\eta_{ij}}(\theta)\right\}_{\theta=0} = \Gamma_i - \sum_{j\in J_i} \mathbf{E}[\eta_{ij}] = \Gamma_i > 0,$$

$\theta = 0$ cannot maximize $\theta\Gamma_i - \sum_{j\in J_i} \Lambda_{\eta_{ij}}(\theta)$. Therefore

$$\exp\left(-\sup_{\theta\geq 0}\left\{\theta\Gamma_i - \sum_{j\in J_i} \Lambda_{\eta_{ij}}(\theta)\right\}\right) < \exp\left(-\sup_{\theta\geq 0}\left\{\theta\Gamma_i - \sum_{j\in J_i} \log e^{\theta^2/2}\right\}\right)$$
$$= \exp\left(-\sup_{\theta\geq 0}\left\{\theta\Gamma_i - |J_i|\frac{\theta^2}{2}\right\}\right)$$
$$= \exp\left(-\frac{\Gamma_i^2}{2|J_i|}\right).$$

□

Since $\Lambda_{\eta_{ij}}(\theta)$ is a convex function of θ, $\sup_{\theta\geq 0}\{\cdot\}$ in Bound (12) is a convex optimization problem, which can be efficiently solved. Like Bound (10), Bound (12) is nonincreasing in Γ_i. The monotonicity of Bounds (10) and (12) and Theorem 2(b) lead to the following important implication: to ensure the same constraint violation probability, Bound (12) requires smaller Γ_i than Bound (10) does. Since $z_R(\mathbf{\Gamma})$ is a nonincreasing function of $\mathbf{\Gamma}$, one can achieve a higher $z_R(\mathbf{\Gamma})$ by using Γ_i required by Bound (12), while maintaining the same constraint violation probability. The following corollary formalizes this deduction.

Corollary 1 *Let ϵ_i be the maximum allowable violation probability for the ith constraint. Let $\mathbf{\Gamma}^B = (\Gamma_1^B, \ldots, \Gamma_m^B)$ (respectively, $\mathbf{\Gamma}^D$) be such that Γ_i^B (respectively, Γ_i^D) is the smallest Γ_i satisfying Bound (10) $\leq \epsilon_i$ (respectively, Bound (12) $\leq \epsilon_i$) for $i = 1, \ldots, m$. Then $z_R(\mathbf{\Gamma}^B) \leq z_R(\mathbf{\Gamma}^D)$.*

In other words, Bound (12) enables one to obtain an equally robust solution with a better objective value. The difference $z_R(\mathbf{\Gamma}^D) - z_R(\mathbf{\Gamma}^B)$ is the gain in the objective value that results from exploiting distributional information on the uncertain problem data, which we refer to as the *estimation discount on the price of robustness*. We remark that Γ_i^B and Γ_i^D in Corollary 1 can be determined by binary search on $[0, |J_i|]$ since Bounds (10) and (12) are monotone in Γ_i.

2.3.2 Solution-Dependent Bounds

In the derivation of Bound (12), there are a few steps that can potentially cause the bound to be loose. For instance, the inequality (13), which was used to make Bound (12) independent of the robust solution $\mathbf{x}^*$, is such a step. Tightening the bound leads us to obtain the bound in Theorem 3, which relies on the probability distributions of the random a_{ij}'s as well as the robust solution $\mathbf{x}^*$. (This bound can be called an *a posteriori* bound in the sense that it is a function of $\mathbf{x}^*$.) The derivation is actually simpler than that of Bound (12) and the resulting bound is stronger than Bound (12). A host of numerical examples we present later demonstrate that the difference can be dramatic.

Theorem 3 *Let Assumption A be in effect.*
(a) Let $C_i(\mathbf{x}^*) = b_i - \sum_j \overline{a}_{ij} x_j^*$ *for all* i *and* $\beta_{ij} = \hat{a}_{ij}|x_j^*|$ *for all* i *and* $j \in J_i$*. Then for the* i*th constraint*

$$\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] \le \exp\Big(-\sup_{\theta\ge 0}\Big\{\theta C_i(\mathbf{x}^*) - \sum_{j\in J_i} \Lambda_{\eta_{ij}}(\theta\beta_{ij})\Big\}\Big). \tag{14}$$

(b) Bound (14) $\le$ *Bound (12).*

Proof: (a) Using the random variables η_{ij} introduced earlier and following the steps of the proof of Thm. 2(a), we have

$$\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] \le \mathbf{P}\Big[\sum_{j\in J_i} \beta_{ij}\eta_{ij} \ge C_i(\mathbf{x}^*)\Big]. \tag{15}$$

From Markov's inequality, for all $\theta \ge 0$

$$\begin{aligned}\mathbf{P}\Big[\sum_{j\in J_i} \beta_{ij}\eta_{ij} \ge C_i(\mathbf{x}^*)\Big] &\le e^{-\theta C_i(\mathbf{x}^*)}\mathbf{E}\big[e^{\theta\sum_{j\in J_i}\beta_{ij}\eta_{ij}}\big] \\ &= e^{-\theta C_i(\mathbf{x}^*)}\prod_{j\in J_i}\mathbf{E}\big[e^{\theta\beta_{ij}\eta_{ij}}\big] \\ &= \exp\Big(-\theta C_i(\mathbf{x}^*) + \sum_{j\in J_i}\Lambda_{\eta_{ij}}(\theta\beta_{ij})\Big).\end{aligned}$$

Optimizing over θ, we obtain

$$\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] \le \exp\Big(-\sup_{\theta\ge 0}\Big\{\theta C_i(\mathbf{x}^*) - \sum_{j\in J_i} \Lambda_{\eta_{ij}}(\theta\beta_{ij})\Big\}\Big).$$

(b) As in the proof of Theorem 2(a), let $S_i^* \cup \{t_i^*\} = \arg\max_{S_i\cup\{t_i\}}\big\{\sum_{j\in S_i}\hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i}|x_{t_i}^*|\big\}$ and $r = \text{argmin}_{j\in S_i^*\cup\{t_i^*\}}\hat{a}_{ij}|x_j^*|$. Consider the probability in (15) and scale both sides of $\sum_{j\in J_i}\beta_{ij}\eta_{ij} \ge C_i(\mathbf{x}^*)$ by $1/\beta_{ir}$. Then following the remaining steps in (a), we obtain

$$\mathbf{P}\Big[\sum_j a_{ij}x_j^* > b_i\Big] \le \exp\Big(-\sup_{\theta\ge 0}\Big\{\theta\frac{C_i(\mathbf{x}^*)}{\beta_{ir}} - \sum_{j\in J_i} \Lambda_{\eta_{ij}}\Big(\theta\frac{\beta_{ij}}{\beta_{ir}}\Big)\Big\}\Big). \tag{16}$$

This bound is equivalent to Bound (14). We will show that Bound (16) $\le$ Bound (12). Let $\gamma_{ij} \triangleq \frac{\beta_{ij}}{\beta_{ir}}$ for $j \in J_i \setminus S_i^*$ and $\delta_{ij} \triangleq \frac{\beta_{ij}}{\beta_{ir}}$ for $j \in S_i^*$. Note that $0 \le \gamma_{ij} \le 1$ and $\delta_{ij} \ge 1$.

Since $\mathbf{x}^*$ is a feasible solution of (5) (cf. Lemma 3),

$$b_i - \sum_j \overline{a}_{ij} x_j^* \geq \sum_{j \in S_i^*} \hat{a}_{ij} |x_j^*| + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i^*} |x_{t_i^*}^*|. \tag{17}$$

Multiplying both sides of (17) by $1/\beta_{ir}$, we obtain

$$\frac{C_i(\mathbf{x}^*)}{\beta_{ir}} \geq \sum_{j \in S_i^*} \delta_{ij} + (\Gamma_i - \lfloor \Gamma_i \rfloor) \gamma_{it_i^*} = \sum_{j \in S_i^*} \delta_{ij} + (\Gamma_i - \lfloor \Gamma_i \rfloor),$$

where the equality follows from the fact that if Γ_i is not an integer, then $r = t_i^*$. Therefore

$$\begin{aligned} &\exp\Big(-\sup_{\theta \geq 0}\Big\{\theta \frac{C_i(\mathbf{x}^*)}{\beta_{ir}} - \sum_{j \in J_i} \Lambda_{\eta_{ij}}\Big(\theta \frac{\beta_{ij}}{\beta_{ir}}\Big)\Big\}\Big) \\ = \;& \exp\Big(-\sup_{\theta \geq 0}\Big\{\theta \frac{C_i(\mathbf{x}^*)}{\beta_{ir}} - \sum_{j \in J_i \setminus S_i^*} \Lambda_{\eta_{ij}}(\theta \gamma_{ij}) - \sum_{j \in S_i^*} \Lambda_{\eta_{ij}}(\theta \delta_{ij})\Big\}\Big) \\ \leq \;& \exp\Big(-\sup_{\theta \geq 0}\Big\{\sum_{j \in S_i^*} \theta \delta_{ij} + \theta(\Gamma_i - \lfloor \Gamma_i \rfloor) - \sum_{j \in J_i \setminus S_i^*} \Lambda_{\eta_{ij}}(\theta \gamma_{ij}) - \sum_{j \in S_i^*} \Lambda_{\eta_{ij}}(\theta \delta_{ij})\Big\}\Big). \end{aligned} \tag{18}$$

Because $0 \leq \gamma_{ij} \leq 1$, we also have

$$\begin{aligned} &\exp\Big(-\sup_{\theta \geq 0}\Big\{\theta \Gamma_i - \sum_{j \in J_i} \Lambda_{\eta_{ij}}(\theta)\Big\}\Big) \\ = \;& \exp\Big(-\sup_{\theta \geq 0}\Big\{\theta \lfloor \Gamma_i \rfloor + \theta(\Gamma_i - \lfloor \Gamma_i \rfloor) - \sum_{j \in J_i \setminus S_i^*} \Lambda_{\eta_{ij}}(\theta) - \sum_{j \in S_i^*} \Lambda_{\eta_{ij}}(\theta)\Big\}\Big) \\ \geq \;& \exp\Big(-\sup_{\theta \geq 0}\Big\{\theta \lfloor \Gamma_i \rfloor + \theta(\Gamma_i - \lfloor \Gamma_i \rfloor) - \sum_{j \in J_i \setminus S_i^*} \Lambda_{\eta_{ij}}(\theta \gamma_{ij}) - \sum_{j \in S_i^*} \Lambda_{\eta_{ij}}(\theta)\Big\}\Big), \end{aligned} \tag{19}$$

where the inequality follows from $\Lambda_{\eta_{ij}}(\theta) \geq \Lambda_{\eta_{ij}}(\theta \gamma_{ij})$ due to the symmetry of the probability distribution of η_{ij} over $[-1, 1]$ (cf. the proof of Theorem 2(a)).

Now our goal is to show that $\exp(\cdot)$ in (18) is no greater than $\exp(\cdot)$ in (19). This is true if for all $\theta \geq 0$

$$\sum_{j \in S_i^*} \theta \delta_{ij} - \sum_{j \in S_i^*} \Lambda_{\eta_{ij}}(\theta \delta_{ij}) \geq \theta \lfloor \Gamma_i \rfloor - \sum_{j \in S_i^*} \Lambda_{\eta_{ij}}(\theta). \tag{20}$$

The inequality (20) holds if for each $j \in S_i^*$

$$\theta \delta_{ij} - \Lambda_{\eta_{ij}}(\theta \delta_{ij}) \geq \theta - \Lambda_{\eta_{ij}}(\theta). \tag{21}$$

Taking the exponential function on both sides of (21) and multiplying both sides by $\mathbf{E}[e^{\theta \delta_{ij} \eta_{ij}}] \mathbf{E}[e^{\theta \eta_{ij}}]$, we obtain

$$\mathbf{E}\big[e^{\theta(\delta_{ij} + \eta_{ij})}\big] \geq \mathbf{E}\big[e^{\theta(1 + \delta_{ij} \eta_{ij})}\big]. \tag{22}$$

The inequality (22) holds if for all realizations of η_{ij}

$$\delta_{ij} + \eta_{ij} \geq 1 + \delta_{ij}\eta_{ij}. \tag{23}$$

Since $\delta_{ij} \geq 1$ and $-1 \leq \eta_{ij} \leq 1$, the inequality (23), which can be rewritten as $(\delta_{ij} - 1)(1 - \eta_{ij}) \geq 0$, holds. □

Recall that Bounds (10) and (12) require only Γ_i (and not Γ_j's, $j \neq i$) for the ith constraint violation probability. In contrast, Bound (14) depends on *all* Γ_i's because $\mathbf{x}^*$ is a function of $\mathbf{\Gamma}$.

If Assumption A is not in effect, a slightly different bound is obtained as shown in Corollary 2. This bound can be useful because in many real-world applications the symmetry assumption could be restrictive.

Corollary 2 *Let* $C_i(\mathbf{x}^*) = b_i - \sum_j \overline{a}_{ij} x_j^*$ *for all* i *and* $\kappa_{ij} = \hat{a}_{ij} x_j^*$ *for all* i *and* $j \in J_i$. *Then for the* i*th constraint*

$$\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] \leq \exp\Big(-\sup_{\theta \geq 0}\Big\{\theta C_i(\mathbf{x}^*) - \sum_{j \in J_i} \Lambda_{\eta_{ij}}(\theta \kappa_{ij})\Big\}\Big). \tag{24}$$

Proof:

$$\begin{aligned}\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] &= \mathbf{P}\Big[\sum_j \overline{a}_{ij} x_j^* + \sum_{j \in J_i} \hat{a}_{ij} x_j^* \eta_{ij} > b_i\Big] \\ &\leq \mathbf{P}\Big[\sum_{j \in J_i} \kappa_{ij}\eta_{ij} \geq C_i(\mathbf{x}^*)\Big],\end{aligned}$$

and the remaining steps are identical to those of Theorem 3(a). □

2.3.3 Moment-Dependent Bounds

Bounds (12), (14), and (24) on the constraint violation probability made full use of the probability distributions of the random a_{ij}'s to compute the moment generating functions of η_{ij}'s. (Recall that $\eta_{ij} = (a_{ij} - \overline{a}_{ij})/\hat{a}_{ij}$.) Therefore, if the probability distributions are not known, but instead some limited distributional information is available, we need a different strategy to establish bounds on the constraint violation probability. We employ the idea of upper bounding the moment generating functions of η_{ij}'s (in lieu of computing them exactly) using the first and second moments of a_{ij}'s and their range information. To that end, we use the following result due to Bennett [Ben62] (also see Dembo and Zeitouni [DZ98]): Let $X \leq b$ be a random variable with $\overline{x} = \mathbf{E}[X]$ and $\mathbf{E}[(X - \overline{x})^2] \leq c^2$ for some $c > 0$. Then for any $\theta \geq 0$,

$$\mathbf{E}\big[e^{\theta X}\big] \leq e^{\theta \overline{x}}\Big\{\frac{(b-\overline{x})^2}{(b-\overline{x})^2 + c^2} e^{-\theta \frac{c^2}{b-\overline{x}}} + \frac{c^2}{(b-\overline{x})^2 + c^2} e^{\theta(b-\overline{x})}\Big\}. \tag{25}$$

Theorem 4 below features two bounds. The first bound requires the robust solution $\mathbf{x}^*$ and the first and second moments (or equivalently, the mean and variance) of a_{ij} for all i and $j \in J_i$. The second bound, on the other hand, does not need the second moments. As shown in the theorem, both bounds are stronger than Bound (10). Another interesting observation is that both bounds do not require the symmetry assumption, Assumption A.

Theorem 4 *Let $C_i(\mathbf{x}^*) = b_i - \sum_j \overline{a}_{ij} x_j^*$ for all i and $\kappa_{ij} = \hat{a}_{ij} x_j^*$ for all i and $j \in J_i$.*
(a) Let $\sigma_{ij}^2 = \mathbf{Var}(\eta_{ij})$ for all i and $j \in J_i$. Then for the ith constraint

$$\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] \le \exp\bigg(-\sup_{\theta \ge 0}\bigg\{\theta C_i(\mathbf{x}^*) - \sum_{j \in J_i} \log\Big(\frac{1}{1+\sigma_{ij}^2} e^{-\theta|\kappa_{ij}|\sigma_{ij}^2} + \frac{\sigma_{ij}^2}{1+\sigma_{ij}^2} e^{\theta|\kappa_{ij}|}\Big)\bigg\}\bigg). \quad (26)$$

(b) For the ith constraint

$$\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] \le \exp\bigg(-\sup_{\theta \ge 0}\bigg\{\theta C_i(\mathbf{x}^*) - \sum_{j \in J_i} \log\Big(\frac{1}{2} e^{-\theta\kappa_{ij}} + \frac{1}{2} e^{\theta\kappa_{ij}}\Big)\bigg\}\bigg). \quad (27)$$

(c) Bound (26) $\le$ Bound (27) $<$ Bound (10).

Proof: (a) Using $\eta_{ij} = (a_{ij} - \overline{a}_{ij})/\hat{a}_{ij}$, we have

$$\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] = \mathbf{P}\Big[\sum_j \overline{a}_{ij} x_j^* + \sum_{j \in J_i} \hat{a}_{ij} x_j^* \eta_{ij} > b_i\Big] \le \mathbf{P}\Big[\sum_{j \in J_i} \kappa_{ij}\eta_{ij} \ge C_i(\mathbf{x}^*)\Big].$$

From Markov's inequality, for all $\theta \ge 0$ we obtain

$$\mathbf{P}\Big[\sum_{j \in J_i} \kappa_{ij}\eta_{ij} \ge C_i(\mathbf{x}^*)\Big] \le e^{-\theta C_i(\mathbf{x}^*)} \mathbf{E}\big[e^{\theta \sum_{j \in J_i} \kappa_{ij}\eta_{ij}}\big] = e^{-\theta C_i(\mathbf{x}^*)} \prod_{j \in J_i} \mathbf{E}\big[e^{\theta\kappa_{ij}\eta_{ij}}\big].$$

Let $X = \kappa_{ij}\eta_{ij}$. Since $\eta_{ij} \in [-1, 1]$, $X \le |\kappa_{ij}|$. Moreover, $\mathbf{E}[X] = \kappa_{ij}\mathbf{E}[\eta_{ij}] = 0$ and $\mathbf{Var}(X) = \kappa_{ij}^2\sigma_{ij}^2$. Then using the inequality (25), we have

$$\mathbf{E}\big[e^{\theta\kappa_{ij}\eta_{ij}}\big] \le \frac{\kappa_{ij}^2}{\kappa_{ij}^2 + \kappa_{ij}^2\sigma_{ij}^2} e^{-\theta\frac{\kappa_{ij}^2\sigma_{ij}^2}{|\kappa_{ij}|}} + \frac{\kappa_{ij}^2\sigma_{ij}^2}{\kappa_{ij}^2 + \kappa_{ij}^2\sigma_{ij}^2} e^{\theta|\kappa_{ij}|}.$$

Combining the results,

$$\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] \le e^{-\theta C_i(\mathbf{x}^*)} \prod_{j \in J_i} \bigg\{\frac{\kappa_{ij}^2}{\kappa_{ij}^2 + \kappa_{ij}^2\sigma_{ij}^2} e^{-\theta\frac{\kappa_{ij}^2\sigma_{ij}^2}{|\kappa_{ij}|}} + \frac{\kappa_{ij}^2\sigma_{ij}^2}{\kappa_{ij}^2 + \kappa_{ij}^2\sigma_{ij}^2} e^{\theta|\kappa_{ij}|}\bigg\}$$
$$= \exp\bigg(-\theta C_i(\mathbf{x}^*) + \sum_{j \in J_i} \log\Big(\frac{1}{1+\sigma_{ij}^2} e^{-\theta|\kappa_{ij}|\sigma_{ij}^2} + \frac{\sigma_{ij}^2}{1+\sigma_{ij}^2} e^{\theta|\kappa_{ij}|}\Big)\bigg).$$

The bound is obtained by optimizing over θ.
(b) Since $\eta_{ij} \in [-1, 1]$, $\mathbf{Var}(\kappa_{ij}\eta_{ij}) \le \kappa_{ij}^2$. Following the same steps as in part (a), we have for all $\theta \ge 0$

$$\mathbf{P}\Big[\sum_j a_{ij} x_j^* > b_i\Big] \le e^{-\theta C_i(\mathbf{x}^*)} \prod_{j \in J_i} \mathbf{E}\big[e^{\theta\kappa_{ij}\eta_{ij}}\big]$$

$$\le e^{-\theta C_i(\mathbf{x}^*)} \prod_{j \in J_i} \left\{ \frac{\kappa_{ij}^2}{2\kappa_{ij}^2} e^{-\theta \frac{\kappa_{ij}^2}{|\kappa_{ij}|}} + \frac{\kappa_{ij}^2}{2\kappa_{ij}^2} e^{\theta |\kappa_{ij}|} \right\}$$
$$= \exp\Big(-\theta C_i(\mathbf{x}^*) + \sum_{j \in J_i} \log\Big(\frac{1}{2} e^{-\theta \kappa_{ij}} + \frac{1}{2} e^{\theta \kappa_{ij}}\Big)\Big),$$

where the second inequality follows from the inequality (25) with $X = \kappa_{ij}\eta_{ij}$. Optimizing over θ, we obtain the bound.
(c) To prove Bound (26) $\le$ Bound (27), we will show for any fixed θ and κ_{ij}

$$\frac{1}{1+\sigma_{ij}^2} e^{-\theta|\kappa_{ij}|\sigma_{ij}^2} + \frac{\sigma_{ij}^2}{1+\sigma_{ij}^2} e^{\theta|\kappa_{ij}|} \le \frac{1}{2} e^{-\theta \kappa_{ij}} + \frac{1}{2} e^{\theta \kappa_{ij}}. \tag{28}$$

To simplify notation, let $x \triangleq \sigma_{ij}^2$ and denote the left hand side of (28) by $f(x)$. Note that $0 \le x \le 1$. We will first show that $f(x)$ is a nondecreasing function. Since

$$\frac{df(x)}{dx} = -\frac{1+\theta|\kappa_{ij}| + \theta|\kappa_{ij}|x}{(1+x)^2} e^{-\theta|\kappa_{ij}|x} + \frac{1}{(1+x)^2} e^{\theta|\kappa_{ij}|},$$

showing $df(x)/dx \ge 0$ is equivalent to showing

$$e^{\theta|\kappa_{ij}| + \theta|\kappa_{ij}|x} \ge 1 + \theta|\kappa_{ij}| + \theta|\kappa_{ij}|x.$$

This inequality holds due to the fact that $e^y \ge 1 + y$ for any y. Therefore the maximum of $f(x)$ is attained at $x = 1$, from which the inequality (28) follows.

To show Bound (27) < Bound (10), notice that Bound (27) can be viewed as a special case of Bound (14), where the probability distribution of η_{ij} is defined as

$$f_{\eta_{ij}}(\eta) = \begin{cases} \frac{1}{2} & \text{if } \eta = -1, 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then by invoking Theorem 3(b) and Theorem 2(b), it follows that Bound (27) < Bound (10). $\square$

The function $\log\big(\frac{1}{1+\sigma_{ij}^2} e^{-\theta|\kappa_{ij}|\sigma_{ij}^2} + \frac{\sigma_{ij}^2}{1+\sigma_{ij}^2} e^{\theta|\kappa_{ij}|}\big)$ in Bound (26) can be viewed as the logarithmic moment generating function of the random variable Y_{ij} whose probability distribution is

$$f_{Y_{ij}}(y) = \begin{cases} \frac{1}{1+\sigma_{ij}^2} & \text{if } y = -|\kappa_{ij}|\sigma_{ij}^2, \\ \frac{\sigma_{ij}^2}{1+\sigma_{ij}^2} & \text{if } y = |\kappa_{ij}|. \end{cases}$$

This observation implies that $\sup_{\theta \ge 0}\{\cdot\}$ in the bound is a convex optimization problem. Consequently, Bound (26) can be computed efficiently. A similar argument can be made for Bound (27).

2.4 The Inverse Problem

We have discussed so far how to obtain "good" bounds on the constraint violation probability for a given $\mathbf{\Gamma}$. We now change the perspective and consider the *inverse problem*: how to find a $\mathbf{\Gamma}$ that maximizes $z_R(\mathbf{\Gamma})$ when the maximum allowable violation probability for the ith constraint, ϵ_i, is given for all i. Using Bound (12), the inverse problem is posed as follows: find a $\mathbf{\Gamma}$ such that

$$\begin{aligned} \text{maximize} \quad & z_R(\mathbf{\Gamma}) \\ \text{subject to} \quad & \text{Bound (12)} \le \epsilon_i, \quad i = 1, \ldots, m. \end{aligned} \tag{29}$$

Since $z_R(\mathbf{\Gamma})$ is a nonincreasing function of $\mathbf{\Gamma}$ (cf. Lemma 2), one needs to seek as small a $\mathbf{\Gamma}$ as possible, which ensures that the constraints of (29) are satisfied. Such a $\mathbf{\Gamma}$ can be found through binary search because Bound (12) is monotone in Γ_i (cf. Corollary 1).

As Bound (14) is stronger than Bound (12), the use of the former in place of the latter in (29), in principle, would result in a better objective value. However, finding an optimal $\mathbf{\Gamma}$ in this case appears to be hard unless the LP problem (1) has some special structure (see, for instance, Section 3.3 where this is the case). A suboptimal approach based on binary search is as follows: First determine the smallest $\mathbf{\Gamma}$, denoted by $\mathbf{\Gamma}^D$, for which Bound (12) $\le \epsilon_i$ for all i. (cf. Corollary 1). Set $\mathbf{\Gamma} = \mathbf{\Gamma}^D/2$, solve (8) with $\mathbf{\Gamma}$, and compute Bound (14) for all i. If Bound (14) yields a value smaller than ϵ_i for all i, set $\mathbf{\Gamma} = \mathbf{\Gamma}/2$. Otherwise, set $\mathbf{\Gamma} = (\mathbf{\Gamma} + \mathbf{\Gamma}^D)/2$. This process is repeated until the interval for $\mathbf{\Gamma}$ becomes reasonably small. The suboptimality of this approach stems from simultaneously increasing or decreasing all the components of $\mathbf{\Gamma}$ in the same proportion. Thus, one might be tempted to amend the approach with the "fine-tuning" step that performs binary search only on some Γ_i's. However, this does not guarantee that an optimal $\mathbf{\Gamma}$ is found because of two reasons: First, Bound (14) for the ith constraint is a function of $\mathbf{\Gamma}$, not of only Γ_i. Second, in general, Bound (14) is not monotone in Γ_i.

We note a close relation between the inverse problem (29) and the chance constrained problem[4]

$$\begin{aligned} \text{maximize} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{P}\Big[\sum_j a_{ij} x_j > b_i\Big] \le \epsilon_i, \quad i = 1, \ldots, m \\ & \mathbf{l} \le \mathbf{x} \le \mathbf{u}. \end{aligned} \tag{30}$$

Under the given constraint violation probabilities, the chance constrained problem (30) produces a better optimal objective value than the inverse problem (29). However, the chance constrained problem is hard to solve because its feasible set is nonconvex and/or because the deterministic analytical expression for $\mathbf{P}\big[\sum_j a_{ij} x_j > b_i\big] \le \epsilon_i$ is difficult to obtain. This intractability has been observed in several works; see Ben-Tal et al. [BTGN09]. As a result, one then resorts to either convex relaxations as in Nemirovski and Shapiro [NS06], Ben-Tal et al. [BTGN09, Chap. 2], Chen et al. [CSS07], or Monte-Carlo sampling as in Calafiore and El

[4] For more on chance constrained problems, see Kall and Wallace [KW94], Birge and Louveaux [BL97], and Nemirovski and Shapiro [NS06].

Ghaoui [CG06]. More recent work on chance constraints has also looked at "joint" chance constraints when one seeks to bound the probability of violating a set of constraints rather than an individual one; see Chen et al. [CSST10] and Zymler et al. [ZKR13].

The inverse problem we presented can serve as an approximation to the chance constrained problem. There are advantages in using the inverse problem: Once $\mathbf{\Gamma}$ is determined, the optimization problem one needs to solve remains an LP, whereas other convex relaxations of (30) often "lift" the class of the problem from an LP to a more complex optimization problem (e.g., an SOCP). Another benefit of the inverse problem is that once $\mathbf{\Gamma}$ is determined, multiple problem instances arising from different $\mathbf{c}$, $\mathbf{b}$, $\mathbf{l}$, and $\mathbf{u}$ can be solved using the same $\mathbf{\Gamma}$, because the changes in those data do not affect Bound (12).

2.5 Numerical Results

We first compare the two *a priori* bounds, Bound (10) and Bound (12), for $\Gamma_i = 5$ and varying $|J_i|$. We consider three symmetric probability distributions for a random a_{ij}: triangle, uniform, and reverse-triangle. (The density functions of the triangle and reverse-triangle distributions are depicted in Fig. 1.)

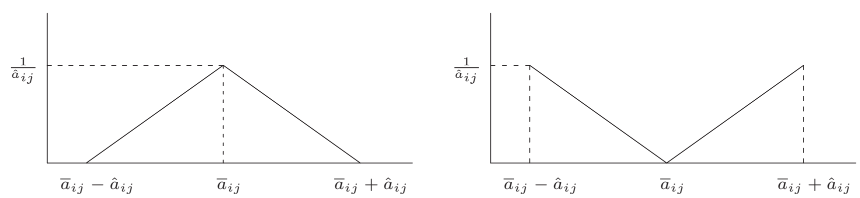

Fig. 1 The triangle (left) and reverse-triangle (right) distributions.

In Table 1, Bound (12)-T (Bound (12)-U, Bound (12)-R, respectively) denotes the value of Bound (12) when all the random a_{ij}'s have the triangle (uniform, reverse-triangle, respectively) distribution. The results show that Bound (12) is tighter than Bound (10); in particular, when all the random a_{ij}'s have the triangle distribution, the differences between the two bounds are most significant. We also observe that as the variance of the probability distribution increases (the triangle distribution having the lowest variance and the reverse-triangle having the highest), Bound (12) increases as well.

We next compare the minimum Γ_i's for which Bound (10) and Bound (12) are less than or equal to $\epsilon_i = 0.05$, respectively. Recall that such Γ_i's can be determined by binary search because both bounds are monotone in Γ_i. Table 2 shows such Γ_i's for various $|J_i|$ and the three symmetric probability distributions. Γ_i from Bound (12) is always smaller than that from Bound (10), regardless of probability distributions. It is also seen that Γ_i from Bound (12) becomes larger as the variance of the probability distribution increases.

We now assess the estimation discount on the price of robustness for randomly generated LP problems; i.e., how much the objective value of the robust solution

Table 1 *A priori* bounds on the constraint violation probability when $\Gamma_i = 5$.

$\lvert J_i \rvert$	Bound (10)	Bound (12)-T	Bound (12)-U	Bound (12)-R
7	0.17	0.000002	0.0013	0.013
10	0.29	0.00028	0.017	0.067
20	0.54	0.022	0.15	0.28
30	0.66	0.08	0.28	0.43
40	0.73	0.15	0.39	0.53
50	0.78	0.22	0.47	0.61

Table 2 Minimum Γ_i's for which $\mathbf{P}\left[\sum_j a_{ij} x_j^* > b_i\right] \leq 0.05$.

$\lvert J_i \rvert$	From Bound (10)	From Bound (12)-T	From Bound (12)-U	From Bound (12)-R
10	7.76	3.12	4.34	5.25
50	17.34	7.06	9.97	12.16
100	24.52	10.01	14.12	17.29
200	34.67	14.17	20.02	24.47
500	54.81	22.34	31.62	38.70
1000	77.64	31.62	44.68	54.81

improves when Bound (12) is used, compared to Bound (10) being used (cf. Corollary 1). Consider a 10×10 matrix $\mathbf{A}$, where all elements are assumed to be random, i.e., $|J_i| = 10$ for all i. We further assume that all a_{ij} are uniform random variables and that $\mathbf{P}\left[\sum_j a_{ij} x_j^* > b_i\right] \leq 0.05$ is required for all i.

A data set $(\mathbf{c}, \mathbf{b}, \mathbf{l}, \mathbf{u}, \overline{a}_{ij}, \hat{a}_{ij})$, which constitutes a problem instance, is specified as follows: c_j is randomly selected from $[-50, 50]$ for all j; b_i is randomly selected from $[0, 100]$ for all i; l_j is randomly chosen from $[-20, 0]$ for all j; u_j is randomly chosen from $[0, 20]$ for all j; $\overline{a}_{ij}$ is randomly drawn from $[-100, 100]$ for all i and j; $\hat{a}_{ij}$ is randomly drawn from $[1, 50]$ for all i and j. In this way, we generate 20 data sets (i.e., 20 problem instances).

In order to guarantee at most 0.05 violation probability for each constraint, Bound (10) and Bound (12) require $\Gamma_i^B = 7.76$ and $\Gamma_i^D = 4.34$ respectively for each i (see Table 2). Table 3 reports $z_R(\mathbf{\Gamma}^B)$ and $z_R(\mathbf{\Gamma}^D)$ for each problem instance as well as the averages over all the problem instances. The results demonstrate that by using the *distribution-dependent* bound (12), one can improve the objective value by more than 50% in some instance and 23% on average compared to using the *distribution-independent* bound (10), without compromising the level of robustness of a solution.

For each problem instance, we also compute Bound (14) for each i using the robust solution $\mathbf{x}^*$ of (8) where $\mathbf{\Gamma}$ is set to $\mathbf{\Gamma}^D$. Among these bound values, the maximum is reported in Table 3. It is notable that the *solution-dependent* bound (14) yields significantly smaller values than 0.05.

3 An Inventory Control Problem with Quality of Service Constraints

3.1 Problem Setting

Consider a single-station single-item stochastic inventory control problem over N discrete time periods. We use the following notation: Let x_k denote the inventory

Table 3 Robust objective values and solution-dependent bound.

Instance	$z_R(\mathbf{\Gamma}^B)$	$z_R(\mathbf{\Gamma}^D)$	$\frac{z_R(\mathbf{\Gamma}^D)-z_R(\mathbf{\Gamma}^B)}{z_R(\mathbf{\Gamma}^B)} \times 100\%$	Bound (14)
1	25.71	34.58	34.5%	0.001609
2	31.86	36.69	15.2%	0.001087
3	308.16	353.55	14.7%	0.000088
4	31.36	45.58	45.4%	0.000600
5	22.13	24.96	12.8%	0.002661
6	95.83	107.49	12.2%	0.000563
7	53.58	61.60	15.0%	0.002647
8	182.74	229.15	25.4%	0.000730
9	84.90	122.26	44.0%	0.000553
10	387.36	464.09	19.8%	0.000081
11	35.37	49.86	41.0%	0.006664
12	103.20	113.74	10.2%	0.000174
13	10.00	11.54	15.4%	0.003834
14	66.37	78.61	18.4%	0.001264
15	52.79	61.72	16.9%	0.000521
16	48.71	59.85	22.9%	0.001783
17	18.53	23.77	28.3%	0.002384
18	71.07	105.99	49.1%	0.002786
19	66.32	68.07	2.6%	0.000099
20	96.72	148.01	53.0%	0.000223
Average	89.64	110.06	22.8%	-

at the beginning of the kth period; w_k the demand during the kth period; u_k the stock ordered at the beginning of the kth period; c the ordering cost per unit stock; and h the holding cost per unit stock per period.

We assume that each demand w_k is an independent random variable taking values from $[\overline{w}_k - \hat{w}_k, \overline{w}_k + \hat{w}_k]$, where $\overline{w}_k > \hat{w}_k > 0$ and $\overline{w}_k = \mathbf{E}[w_k]$. The following symmetry assumption on the probability distribution of w_k is analogous to Assumption A and will be needed for some of the results later on.

Assumption B
For all k, the probability distribution of w_k is symmetric *over $[\overline{w}_k - \hat{w}_k, \overline{w}_k + \hat{w}_k]$, that is,*

$$F_{w_k}(\overline{w}_k - w) = 1 - F_{w_k}(\overline{w}_k + w), \quad 0 \le w \le \hat{w}_k,$$

where F_{w_k} is the cumulative distribution function of w_k.

We also assume that the stock ordered at the beginning of the kth period is delivered instantly, i.e., the lead time is zero. If we further assume that excess demand is backlogged and filled as soon as additional stock becomes available, inventory evolves according to

$$x_{k+1} = x_k + u_k - w_k, \quad k = 1, \ldots, N. \tag{31}$$

The following *quality of service (QoS)* constraints will be incorporated into the inventory control problem:

$$x_k + u_k \ge w_k, \quad k = 1, \ldots, N.$$

We consider an inventory control problem at the *strategic level*, where the ordering decisions should be made before any of demands are observed. A few examples that fit this case are when it takes considerable amount of time to produce (or ship) items or when the supplier allows discounts for the orders placed in advance. We will mention how to adjust our model to the *operational* inventory control problem at the end of Section 3.3.

The strategic inventory control problem with the QoS constraints, which aims to minimize total ordering and holding costs, is formulated as the following LP problem with data uncertainty:

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{N} (cu_k + hx_{k+1}) \\ \text{subject to} \quad & x_{k+1} = x_k + u_k - w_k, \quad k = 1, \dots, N, \\ & x_k + u_k \geq w_k, \quad k = 1, \dots, N, \\ & u_k \geq 0, \quad k = 1, \dots, N, \end{aligned} \tag{32}$$

where x_1 is given. From the inventory evolution equations (31), we obtain $x_{k+1} = x_1 + \sum_{j=1}^{k}(u_j - w_j)$. Thus we can eliminate x_k, $k = 2, \dots, N$, from (32), yielding

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{N} \big(c + h(N-k+1)\big) u_k + Nhx_1 - h \sum_{k=1}^{N} (N-k+1) w_k \\ \text{subject to} \quad & \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k} w_j, \quad k = 1, \dots, N, \\ & u_k \geq 0, \quad k = 1, \dots, N. \end{aligned} \tag{33}$$

By defining $a_k \triangleq N - k + 1$, $k = 1, \dots, N$, and introducing the auxiliary variable z for the objective function, we can rewrite (33) as

$$\begin{aligned} \text{minimize} \quad & z \\ \text{subject to} \quad & \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k} w_j, \quad k = 1, \dots, N, \\ & z - \sum_{k=1}^{N} (c + ha_k) u_k \geq Nhx_1 - h \sum_{k=1}^{N} a_k w_k, \\ & u_k \geq 0, \quad k = 1, \dots, N. \end{aligned} \tag{34}$$

If we view (34) in terms of the generic LP problem (1), the vector $\mathbf{b}$ is uncertain because the demands w_k are random elements. The coefficients of the decision variables, $\mathbf{u} = (u_1, \dots, u_N)$ and z, are not subject to uncertainty, i.e., the matrix $\mathbf{A}$ is deterministic. As discussed in Section 2.1, we may transform (34) in such a way that only the elements of the matrix $\tilde{\mathbf{A}}$ are random. In this particular problem, however, it will be more convenient not to do so.

3.2 The Robust Inventory Control Problem

When we presented the robust formulation for the LP problem (1) in Section 2.2, the uncertainty budget Γ_i was introduced for each row i of the matrix $\mathbf{A}$, i.e., for each constraint of the LP problem. It was because each constraint has a distinct set of random elements. Put differently, the random elements in row i affect only the ith constraint. This is not the case for the inventory control problem (34). Each random element w_k influences more than one constraint. For instance, w_1 is involved in all of the constraints. Therefore, we believe that it is more appropriate to use a single uncertainty budget Γ globally for the problem. (We point out that Bertsimas and Thiele [BT06] used a different Γ_k for each period k.)

The uncertainty budget $\Gamma \in [0, N]$ imposes the uncertainty budget constraint

$$\sum_{k=1}^{N} \frac{|w_k - \overline{w}_k|}{\hat{w}_k} \leq \Gamma.$$

Define the restricted uncertainty set $\mathscr{R}(\Gamma)$ as

$$\mathscr{R}(\Gamma) \triangleq \left\{ \mathbf{w} \mid w_k \in [\overline{w}_k - \hat{w}_k, \overline{w}_k + \hat{w}_k],\, \forall\, k;\ \sum_{k} \frac{|w_k - \overline{w}_k|}{\hat{w}_k} \leq \Gamma \right\}.$$

The robust formulation is then given by

$$\begin{aligned} z_R(\Gamma) = \text{minimize} \quad & z \\ \text{subject to} \quad & \sum_{j=1}^{k} u_j \geq \max_{\mathbf{w} \in \mathscr{R}(\Gamma)} \left\{ -x_1 + \sum_{j=1}^{k} w_j \right\}, \quad k = 1, \ldots, N, \\ & z - \sum_{k=1}^{N} (c + h a_k) u_k \geq \max_{\mathbf{w} \in \mathscr{R}(\Gamma)} \left\{ N h x_1 - h \sum_{k=1}^{N} a_k w_k \right\}, \\ & u_k \geq 0, \quad k = 1, \ldots, N. \end{aligned} \tag{35}$$

For $k = 1, \ldots, \lfloor \Gamma \rfloor$, we have

$$\max_{\mathbf{w} \in \mathscr{R}(\Gamma)} \left\{ -x_1 + \sum_{j=1}^{k} w_j \right\} = -x_1 + \sum_{j=1}^{k} (\overline{w}_j + \hat{w}_j).$$

For $k = \lfloor \Gamma \rfloor + 1, \ldots, N$, let S_k be a subset of the index set $\{1, \ldots, k\}$ such that $|S_k| = \lfloor \Gamma \rfloor$. In addition, let t_k be an index such that $t_k \in \{1, \ldots, k\} \setminus S_k$. Then we have

$$\max_{\mathbf{w} \in \mathscr{R}(\Gamma)} \left\{ -x_1 + \sum_{j=1}^{k} w_j \right\} = -x_1 + \sum_{j=1}^{k} \overline{w}_j + \max_{S_k \cup \{t_k\}} \left\{ \sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) \hat{w}_{t_k} \right\}.$$

We also have

$$\max_{\mathbf{w} \in \mathscr{R}(\Gamma)} \left\{ N h x_1 - h \sum_{k=1}^{N} a_k w_k \right\} = \\ N h x_1 - h \sum_{k=1}^{N} a_k \overline{w}_k + h \max_{S_N \cup \{t_N\}} \left\{ \sum_{j \in S_N} a_j \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) a_{t_N} \hat{w}_{t_N} \right\}.$$

Hence (35) becomes

$$z_R(\Gamma) = \text{minimize} \quad z \tag{36}$$

$$
\begin{aligned}
\text{subject to}\quad & \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k} (\overline{w}_j + \hat{w}_j), \quad k = 1, \ldots, \lfloor \Gamma \rfloor, \\
& \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k} \overline{w}_j + \max_{S_k \cup \{t_k\}} \Big\{ \sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) \hat{w}_{t_k} \Big\}, \\
& \qquad\qquad\qquad\qquad\qquad\qquad k = \lfloor \Gamma \rfloor + 1, \ldots, N, \\
& z - \sum_{k=1}^{N} (c + h a_k) u_k \geq N h x_1 - h \sum_{k=1}^{N} a_k \overline{w}_k, \\
& \qquad\qquad + h \max_{S_N \cup \{t_N\}} \Big\{ \sum_{j \in S_N} a_j \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) a_{t_N} \hat{w}_{t_N} \Big\}, \\
& u_k \geq 0, \quad k = 1, \ldots, N.
\end{aligned}
$$

We refer to (36) as the *robust inventory control problem.*

Let us characterize the optimal ordering quantities of the robust inventory control problem. To make the exposition simple, define

$$
M(k) \triangleq \max_{S_k \cup \{t_k\}} \{ \sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) \hat{w}_{t_k} \},\ k = \lfloor \Gamma \rfloor + 1, \ldots, N
$$

,

$$
A(N) \triangleq \max_{S_N \cup \{t_N\}} \{ \sum_{j \in S_N} a_j \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) a_{t_N} \hat{w}_{t_N} \}.
$$

Eliminating the auxiliary variable z from (36), we obtain

$$
\begin{aligned}
\text{minimize}\quad & \sum_{k=1}^{N} (c + h a_k) u_k + N h x_1 - h \sum_{k=1}^{N} a_k \overline{w}_k + h A(N) \\
\text{subject to}\quad & \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k} (\overline{w}_j + \hat{w}_j), \quad k = 1, \ldots, \lfloor \Gamma \rfloor, \\
& \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k} \overline{w}_j + M(k), \quad k = \lfloor \Gamma \rfloor + 1, \ldots, N, \\
& u_k \geq 0, \quad k = 1, \ldots, N.
\end{aligned} \tag{37}
$$

Define the constant demands $\widetilde{w}_k$ such that

$$
\sum_{j=1}^{k} \widetilde{w}_j = \begin{cases} \sum_{j=1}^{k} (\overline{w}_j + \hat{w}_j), & k = 1, \ldots, \lfloor \Gamma \rfloor, \\ \sum_{j=1}^{k} \overline{w}_j + M(k), & k = \lfloor \Gamma \rfloor + 1, \ldots, N, \end{cases}
$$

from which we obtain

$$
\widetilde{w}_k = \begin{cases} \overline{w}_k + \hat{w}_k & \text{if } k = 1, \ldots, \lfloor \Gamma \rfloor, \\ \overline{w}_{\lfloor \Gamma \rfloor + 1} + M(\lfloor \Gamma \rfloor + 1) - \sum_{j=1}^{\lfloor \Gamma \rfloor} \hat{w}_j & \text{if } k = \lfloor \Gamma \rfloor + 1, \\ \overline{w}_k + M(k) - M(k-1) & \text{if } k = \lfloor \Gamma \rfloor + 2, \ldots, N. \end{cases} \tag{38}
$$

Rewriting (37) in term of the demands $\widetilde{w}_k$, we have

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{N}(c+ha_k)u_k + Nhx_1 - h\sum_{k=1}^{N} a_k\widetilde{w}_k + C \\ \text{subject to} \quad & \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k}\widetilde{w}_j, \quad k=1,\dots,N, \\ & u_k \geq 0, \quad k=1,\dots,N, \end{aligned}$$

where

$$C = h\sum_{k=1}^{\lfloor \Gamma \rfloor}(\lfloor \Gamma \rfloor - k + 1)\hat{w}_k + h\sum_{k=\lfloor \Gamma \rfloor+1}^{N-1} M(k) + a_N hM(N) + hA(N).$$

Since C is a constant, it can be ignored in minimization. Introducing the auxiliary variable z for the objective function minus C, we have

$$\begin{aligned} \text{minimize} \quad & z \\ \text{subject to} \quad & \sum_{j=1}^{k} u_j \geq -x_1 + \sum_{j=1}^{k}\widetilde{w}_j, \quad k=1,\dots,N, \\ & z - \sum_{k=1}^{N}(c+ha_k)u_k \geq Nhx_1 - h\sum_{k=1}^{N} a_k\widetilde{w}_k, \\ & u_k \geq 0, \quad k=1,\dots,N. \end{aligned} \tag{39}$$

Theorem 5 *Assume that $x_1 < \sum_{k=1}^{N}\widetilde{w}_k$, where $\widetilde{w}_k$ are given in (38). The optimal ordering quantities u_k^* of the robust inventory control problem (36) correspond to the base-stock policy*

$$u_k^* = \begin{cases} \widetilde{w}_k - x_k & \text{if } x_k < \widetilde{w}_k, \\ 0 & \text{otherwise.} \end{cases} \tag{40}$$

The cost of this optimal policy is $c\left(\sum_{k=1}^{N}\widetilde{w}_k - x_1\right) + h\sum_{k=1}^{L}\left(x_1 - \sum_{j=1}^{k}\widetilde{w}_j\right) + C$, where $L = \max\left\{k \mid x_1 - \sum_{j=1}^{k}\widetilde{w}_j \geq 0\right\}$ and C is given above.

Proof: From (40), u_k^* are recursively determined as

$$u_k^* = \begin{cases} 0, & \text{if } k = 1,\dots,L, \\ -x_1 + \sum_{j=1}^{L+1}\widetilde{w}_j, & \text{if } k = L+1, \\ \widetilde{w}_k, & \text{if } k = L+2,\dots,N. \end{cases} \tag{41}$$

Let $z^* = \sum_{k=1}^{N}(c+ha_k)u_k^* + Nhx_1 - h\sum_{k=1}^{N} a_k\widetilde{w}_k$. Then $(\mathbf{u}^*, z^*)$ is a feasible solution of (39). The dual of (39) is

$$\text{maximize} \quad \sum_{k=1}^{N}\lambda_k\left(-x_1 + \sum_{j=1}^{k}\widetilde{w}_j\right) + \mu\left(Nhx_1 - h\sum_{k=1}^{N} a_k\widetilde{w}_k\right)$$

$$\text{subject to} \quad \sum_{j=k}^{N} \lambda_j \le \mu(c + ha_k), \quad k = 1, \dots, N,$$
$$\mu = 1, \quad \lambda_k \ge 0, \quad k = 1, \dots, N,$$

where λ_k are the dual variables for the first N constraints of (39) and μ for the next constraint. Consider a dual feasible solution $(\boldsymbol{\lambda}^*, \mu^*)$ that satisfies the following set of equations:

$$\mu^* = 1, \quad \lambda_k^* = 0, \quad k = 1, \dots, L,$$
$$\sum_{j=k}^{N} \lambda_j^* = \mu^*(c + ha_k), \quad k = L+1, \dots, N,$$

from which we obtain

$$\lambda_k^* = \begin{cases} 0, & \text{if } k = 1, \dots, L, \\ h, & \text{if } k = L+1, \dots, N-1, \\ c+h, & \text{if } k = N. \end{cases}$$

It is easy to verify that the primal-dual feasible solution pair, $(\mathbf{u}^*, z^*)$ and $(\boldsymbol{\lambda}^*, \mu^*)$, satisfies the complementary slackness conditions. Therefore $(\mathbf{u}^*, z^*)$ is an optimal solution of (39), which establishes the optimality of the policy (40). The cost of the optimal policy is equal to $z^* + C = \sum_{k=1}^{N}(c + ha_k)u_k^* + Nhx_1 - h\sum_{k=1}^{N} a_k \widetilde{w}_k + C$. Using (41) and after some algebra, we obtain the desired result. □

We now derive an LP formulation equivalent to the robust inventory control problem (36). First, observe that for each $k = \lfloor \Gamma \rfloor + 1, \dots, N$, $\max_{S_k \cup \{t_k\}} \{ \sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor)\hat{w}_{t_k} \}$ equals to the optimal objective value of the following LP problem whose feasible set is non-empty and bounded:

$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{k} \hat{w}_j y_{kj} \\ \text{subject to} \quad & \sum_{j=1}^{k} y_{kj} \le \Gamma, \\ & 0 \le y_{kj} \le 1, \quad j = 1, \dots, k. \end{aligned} \tag{42}$$

The dual of (42) is

$$\begin{aligned} \text{minimize} \quad & \Gamma p_k + \sum_{j=1}^{k} q_{kj} \\ \text{subject to} \quad & p_k + q_{kj} \ge \hat{w}_j, \quad j = 1, \dots, k \\ & p_k, q_{kj} \ge 0, \quad j = 1, \dots, k, \end{aligned} \tag{43}$$

where p_k is the dual variable for the constraint $\sum_{j=1}^{k} y_{kj} \le \Gamma$ and q_{kj} the dual variables for the constraints $y_{kj} \le 1$, $j = 1, \dots, k$. It can also be seen that $\max_{S_N \cup \{t_N\}} \{ \sum_{j \in S_N} a_j \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor) a_{t_N} \hat{w}_{t_N} \}$ is equal to the optimal objective value of the following LP problem whose feasible set is also non-empty and bounded:

$$\begin{aligned} \text{maximize} \quad & \sum_{k=1}^{N} a_k \hat{w}_k v_k \\ \text{subject to} \quad & \sum_{k=1}^{N} v_k \le \Gamma, \\ & 0 \le v_k \le 1, \quad k = 1, \dots, N. \end{aligned} \tag{44}$$

The dual of (44) is

$$\begin{aligned} \text{minimize} \quad & \Gamma r + \sum_{k=1}^{N} s_k \\ \text{subject to} \quad & r + s_k \ge a_k \hat{w}_k, \quad k = 1, \dots, N \\ & r, s_k \ge 0, \quad k = 1, \dots, N, \end{aligned} \tag{45}$$

where r and s_k are the dual variables. The next Proposition provides the LP formulation of the robust inventory control problem (36). The proof is similar to that of Theorem 1 and it is given in Appendix C.

Proposition 1 *The robust inventory control problem (36) is equivalent to the LP formulation*

$$\begin{aligned} z_R(\Gamma) = \text{minimize} \quad & z \\ \text{subject to} \quad & \sum_{j=1}^{k} u_j \ge -x_1 + \sum_{j=1}^{k} (\overline{w}_j + \hat{w}_j), \quad k = 1, \dots, \lfloor \Gamma \rfloor, \\ & \sum_{j=1}^{k} u_j \ge -x_1 + \sum_{j=1}^{k} \overline{w}_j + \Gamma p_k + \sum_{j=1}^{k} q_{kj}, \quad k = \lfloor \Gamma \rfloor + 1, \dots, N, \\ & p_k + q_{kj} \ge \hat{w}_j, \quad k = \lfloor \Gamma \rfloor + 1, \dots, N, \quad j = 1, \dots, k, \\ & z - \sum_{k=1}^{N} (c + h a_k) u_k \ge N h x_1 - h \sum_{k=1}^{N} a_k \overline{w}_k + h \Big(\Gamma r + \sum_{k=1}^{N} s_k \Big), \\ & r + s_k \ge a_k \hat{w}_k, \quad k = 1, \dots, N, \\ & p_k,\ q_{kj} \ge 0, \qquad k = \lfloor \Gamma \rfloor + 1, \dots, N, \quad j = 1, \dots, k, \\ & r,\ s_k \ge 0, \qquad k = 1, \dots, N, \\ & u_k \ge 0, \qquad k = 1, \dots, N. \end{aligned} \tag{46}$$

3.3 Bounds on the QoS Constraint Violation Probability

Assume $\Gamma > 0$. Let $(\mathbf{u}^*, z^*)$ be an optimal solution of the robust inventory control problem (36), which is obtained by solving (46). We examine the probability that $\mathbf{u}^*$ violates the QoS constraints $\sum_{j=1}^{k} u_j \ge -x_1 + \sum_{j=1}^{k} w_j$, $k = 1, \dots, N$.

Clearly, none of the constraints for $k = 1, \ldots, \lfloor \Gamma \rfloor$ are violated by $\mathbf{u}^*$ because Γ provides full protection for those constraints. For $k = \lfloor \Gamma \rfloor + 1, \ldots, N$, the following theorem provides upper bounds on the violation probability of the kth period QoS constraint. The proofs of the three parts are similar to those of Theorem 2(a), Corollary 2, and Theorem 3(b), respectively; we omit the details in the interest of space. Let $z_k = (w_k - \overline{w}_k)/\hat{w}_k$, $k = 1, \ldots, N$, and let $\Lambda_{z_k}(\theta) \triangleq \log \mathbf{E}[e^{\theta z_k}]$ denote the logarithmic moment generating function of z_k.

Proposition 2 *(a) Let Assumption B be in effect. Then for* $k = \lfloor \Gamma \rfloor + 1, \ldots, N$

$$\mathbf{P}\Big[\sum_{j=1}^{k} u_j^* < -x_1 + \sum_{j=1}^{k} w_j\Big] \leq \exp\Big(-\sup_{\theta \geq 0}\Big\{\theta\Gamma - \sum_{j=1}^{k} \Lambda_{z_j}(\theta)\Big\}\Big). \tag{47}$$

(b) Let $C_k(\mathbf{u}^*) = \sum_{j=1}^{k} u_j^* + x_1 - \sum_{j=1}^{k} \overline{w}_j$. *Then for* $k = \lfloor \Gamma \rfloor + 1, \ldots, N$

$$\mathbf{P}\Big[\sum_{j=1}^{k} u_j^* < -x_1 + \sum_{j=1}^{k} w_j\Big] \leq \exp\Big(-\sup_{\theta \geq 0}\Big\{\theta C_k(\mathbf{u}^*) - \sum_{j=1}^{k} \Lambda_{z_j}(\theta\hat{w}_j)\Big\}\Big). \tag{48}$$

(c) Under Assumption B, Bound (48) $\leq$ Bound (47).

Under Assumption B, Bound (10) developed in [BS04] can be tailored to the inventory control problem as follows: for $k = \lfloor \Gamma \rfloor + 1, \ldots, N$

$$\mathbf{P}\Big[\sum_{j=1}^{k} u_j^* < -x_1 + \sum_{j=1}^{k} w_j\Big] \leq \exp\Big(-\frac{\Gamma^2}{2k}\Big). \tag{49}$$

By applying Theorem 2(b), one can show that Bound (47) < Bound (49).

Bound (47) depends on the probability distributions of w_k's, but not on $\mathbf{u}^*$. Bound (48), on the other hand, relies on both the probability distributions and $\mathbf{u}^*$. Moreover, Bound (48) does not require the symmetry assumption, Assumption B. The following two lemmas establish some properties of these bounds.

Lemma 4 *Given Γ, Bound (47) is a nondecreasing function of k.*

Proof: Since $\mathbf{E}[z_j] = 0$, Jensen's inequality yields $\mathbf{E}[e^{\theta z_j}] \geq \exp\big(\mathbf{E}[\theta z_j]\big) = 1$, from which we have $\Lambda_{z_j}(\theta) \geq 0$. Let $k' > k$. Then

$$\theta\Gamma - \sum_{j=1}^{k'} \Lambda_{z_j}(\theta) = \theta\Gamma - \sum_{j=1}^{k} \Lambda_{z_j}(\theta) - \sum_{j=k+1}^{k'} \Lambda_{z_j}(\theta) \leq \theta\Gamma - \sum_{j=1}^{k} \Lambda_{z_j}(\theta).$$

Hence

$$\sup_{\theta \geq 0}\Big\{\theta\Gamma - \sum_{j=1}^{k'} \Lambda_{z_j}(\theta)\Big\} \leq \sup_{\theta \geq 0}\Big\{\theta\Gamma - \sum_{j=1}^{k} \Lambda_{z_j}(\theta)\Big\}$$

and the lemma follows. □

Lemma 5 *Given k, Bounds (47) and (48) are nonincreasing functions of Γ.*

Proof: It is immediate that Bound (47) is a nonincreasing function of Γ. To show the monotonicity of Bound (48), consider Γ and Γ', where $\Gamma < \Gamma' < k$. Let $(\mathbf{u}^*, z^*)$ and $(\mathbf{u}', z')$ be optimal solutions of the robust inventory control problem (36) associated with Γ and Γ', respectively. Let $S_k^* \cup \{t_k^*\} = \arg\max_{S_k \cup \{t_k\}} \{\sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor)\hat{w}_{t_k}\}$, and define S_k', t_k' similarly. We need to consider three cases.

First, suppose $x_1 \le \sum_{j=1}^k \overline{w}_j + \sum_{j \in S_k^*} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor)\hat{w}_{t_k^*}$. Then at optimality the kth period QoS constraint of (36) will be binding for both Γ and Γ', i.e.,

$$\sum_{j=1}^k u_j^* = -x_1 + \sum_{j=1}^k \overline{w}_j + \sum_{j \in S_k^*} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor)\hat{w}_{t_k^*},$$
$$\sum_{j=1}^k u_j' = -x_1 + \sum_{j=1}^k \overline{w}_j + \sum_{j \in S_k'} \hat{w}_j + (\Gamma' - \lfloor \Gamma' \rfloor)\hat{w}_{t_k'}.$$

Since $\sum_{j \in S_k^*} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor)\hat{w}_{t_k^*} < \sum_{j \in S_k'} \hat{w}_j + (\Gamma' - \lfloor \Gamma' \rfloor)\hat{w}_{t_k'}$, we have $C_k(\mathbf{u}^*) < C_k(\mathbf{u}')$.

Second, suppose $\sum_{j=1}^k \overline{w}_j + \sum_{j \in S_k^*} \hat{w}_j + (\Gamma - \lfloor \Gamma \rfloor)\hat{w}_{t_k^*} < x_1 \le \sum_{j=1}^k \overline{w}_j + \sum_{j \in S_k'} \hat{w}_j + (\Gamma' - \lfloor \Gamma' \rfloor)\hat{w}_{t_k'}$. Then the kth period QoS constraint will be binding at optimality for Γ'. However, the constraint will not be binding at optimality for Γ, resulting in $\sum_{j=1}^k u_j^* = 0$. This implies $C_k(\mathbf{u}^*) \le C_k(\mathbf{u}')$.

Third, suppose $x_1 > \sum_{j=1}^k \overline{w}_j + \sum_{j \in S_k'} \hat{w}_j + (\Gamma' - \lfloor \Gamma' \rfloor)\hat{w}_{t_k'}$. In this case, the kth period QoS constraint will not be binding at optimality for Γ as well as for Γ'. Hence $\sum_{j=1}^k u_j^* = \sum_{j=1}^k u_j' = 0$, and we have $C_k(\mathbf{u}^*) = C_k(\mathbf{u}')$.

Therefore $C_k(\mathbf{u}^*)$ is not greater than $C_k(\mathbf{u}')$, from which the lemma follows. □

We next consider the inverse problem (cf. Section 2.4) for the inventory control problem: find a Γ such that

$$\begin{aligned} &\text{maximize} \quad z_R(\Gamma) \\ &\text{subject to} \quad \text{Bound (48)} \le \epsilon_k, \quad k = 1, \ldots, N, \end{aligned} \tag{50}$$

where ϵ_k is the maximum allowable violation probability for the kth period QoS constraint. (Note that we are using the solution-dependent bound (48) in the inverse problem. Refer to the related discussion in Section 2.4.) An optimal (i.e., smallest possible) Γ, denoted by Γ^S, can be found as follows: Assume that the probability distributions of w_k's are symmetric. (We need this assumption in order to use Bound (47) during the process.) First for each $k = 1, \ldots, N$, we compute the minimum Γ, denoted by Γ_k, for which Bound (47) $\le \epsilon_k$. Note that Γ_k can be obtained through binary search over the interval $[0, k]$, because Bound (47) is monotone in Γ. Let $\Gamma^D = \max_k \Gamma_k$. Clearly, Γ^S lies in the interval $[0, \Gamma^D]$ due to Proposition 2(c). Suppose we choose some $\Gamma' \in [0, \Gamma^D]$. We solve (46) with Γ' and then compute Bound (48) for each k, denoted by p_k. If $p_k \le \epsilon_k$ for all k, Γ^S lies in $[0, \Gamma']$ because Bound (48) is also monotone in Γ. Otherwise, Γ^S belongs to $[\Gamma', \Gamma^D]$. This implies that Γ^S can be found by binary search.

The inverse problem is formulated at the *strategic level* in the sense that once the optimal Γ^S is found, the robust ordering quantities u_k^*, $k = 1, \ldots, N$, are determined by (46) before any of the demands are observed. The framework, however, can be easily extended to handle *operational level* inventory control decisions; namely, the ordering decision u_k^* can be made after the previous demands $w_1, \ldots, w_{k-1}$ have been observed. To that end, one can use the inverse problem in a *receding horizon* manner: First solve the inverse problem and order u_1^* only. Observe w_1 and compute x_2. With x_2 being the initial inventory, one now has an inventory control problem of $N-1$ periods. Solve the inverse problem for this $N-1$ period inventory control problem and order u_2^* only. Observe w_2 and compute x_3. And repeat the process. The ordering decisions of the operational inventory control problem are certainly less conservative than those of the strategic counterpart since the former problem allows the decision maker to use more information. Other than the receding horizon approach, an alternative can be to consider adaptive policies but with a simple linear structure; see Ben-Tal et al. [BTGGN04], See and Sim [SS10], Bertimas et al. [BIP11], and Bertsimas and Goyal [BG12].

3.4 Numerical Results

We set $N = 30$, $c = 2$, $h = 1$, and $x_1 = 0$. Let $\overline{\mathbf{w}} = (\overline{w}_1, \ldots, \overline{w}_N)$ and $\hat{\mathbf{w}} = (\hat{w}_1, \ldots, \hat{w}_N)$. For each k, $\overline{w}_k$ is randomly drawn from the range $[30, 300]$; for each k, $\hat{w}_k$ is randomly chosen from the range $[1, \overline{w}_k - 1]$. In this manner, we generate 20 instances of $(\overline{\mathbf{w}}, \hat{\mathbf{w}})$. We consider three symmetric probability distributions for w_k: triangle, uniform, and reverse-triangle. (For the triangle and reverse-triangle distributions, see Section 2.5.) For simplicity, we assume that all w_k have the same distribution. We further assume that the kth period QoS constraint violation probability is required to be no more than 0.05 for all k.

Let Γ^D and Γ^B denote the minimum values of Γ for which Bound (47) and Bound (49) are less than or equal to 0.05 for all k, respectively. Since both bounds are nondecreasing functions of k for a given Γ, it suffices to consider the last period to determine Γ^D and Γ^B. Γ^D depends on the probability distributions of w_k's; computations (binary search) yield that Γ^D is equal to 5.45, 7.68, and 9.38 for the triangle, uniform, and reverse-triangle distributions, respectively. Γ^B equals to 13.42 regardless of probability distributions. Let Γ^S be the optimal Γ of the inverse problem (50).

For each instance of $(\overline{\mathbf{w}}, \hat{\mathbf{w}})$ and probability distribution, we obtain the optimal costs of the robust inventory control problem (36), $z_R(\Gamma^B)$ and $z_R(\Gamma^D)$. We also determine Γ^S and compute $z_R(\Gamma^S)$. In Table 4, we list the average $z_R(\Gamma^B)$, $z_R(\Gamma^D)$, Γ^S, and $z_R(\Gamma^S)$ over the 20 instances of $(\overline{\mathbf{w}}, \hat{\mathbf{w}})$. It can be seen that the stronger bounds we have derived using distributional information lead to significant cost savings (up to 54% in these examples) over the *distribution-independent* robust approach of [BS04]. For comparison purposes, we also computed the optimal cost and QoS constraint violation probability if one solves the nominal (rather than the robust) problem. It turns out, that the corresponding violation probability is quite high, namely, about 0.5 with either distribution. Of course, the optimal cost in that case is substantially lower than the robust cost.

To assess numerically the tightness of our best bound we computed the value of Γ, say Γ^*, that is required for the actual QoS constraint violation probability to remain below 0.05. To that end, we used simulation to estimate the violation probability. We found that the cost $z_R(\Gamma^*)$ corresponding to Γ^* is no more than

16% lower than $z_R(\Gamma^S)$, suggesting that our bound is reasonably tight. We also note that we performed computations for other values of the problem parameters as well (e.g., for a holding cost of $h = 0.5$) and the qualitative conclusions remain the same.

Table 4 Robust inventory control cost $z_R(\Gamma)$ and Γ^S.

Distribution	$z_R(\Gamma^B)$	$z_R(\Gamma^D)$	Γ^S	$z_R(\Gamma^S)$	$\frac{z_R(\Gamma^D)-z_R(\Gamma^B)}{z_R(\Gamma^B)} \times 100\%$	$\frac{z_R(\Gamma^S)-z_R(\Gamma^B)}{z_R(\Gamma^B)} \times 100\%$
Triangle	75847.83	51867.60	2.55	35192.38	−31.6%	−53.6%
Uniform	75847.83	60975.82	3.81	43321.65	−19.6%	−42.9%
Reverse-triangle	75847.83	66466.91	4.85	49024.66	−12.4%	−35.4%

Let $\mathbf{u}_B^*$, $\mathbf{u}_D^*$, and $\mathbf{u}_S^*$ be optimal ordering solutions of the robust inventory control problem associated with Γ^B, Γ^D, and Γ^S, respectively. For each instance of $(\overline{\mathbf{w}}, \hat{\mathbf{w}})$ and probability distribution, we simulate the inventory system with $\mathbf{u}_B^*$, $\mathbf{u}_D^*$, and $\mathbf{u}_S^*$. Each simulation consists of 100,000 runs, and the cost is averaged over the 100,000 runs. In Table 5, we report the costs further averaged over the 20 instances of $(\overline{\mathbf{w}}, \hat{\mathbf{w}})$.

Table 5 Simulated inventory control costs.

Distribution	$c(\mathbf{u}_B^*)$	$c(\mathbf{u}_D^*)$	$c(\mathbf{u}_S^*)$	$\frac{c(\mathbf{u}_D^*)-c(\mathbf{u}_B^*)}{c(\mathbf{u}_B^*)} \times 100\%$	$\frac{c(\mathbf{u}_S^*)-c(\mathbf{u}_B^*)}{c(\mathbf{u}_B^*)} \times 100\%$
Triangle	45638.19	33058.74	23736.86	−27.6%	−48.0%
Uniform	45871.10	38228.70	28570.15	−16.7%	−37.7%
Reverse-triangle	46106.14	41360.53	31971.06	−10.3%	−30.7%

During the 100,000 simulation runs, we also count for each period the number of times that shortage occurs. Dividing the frequency of shortage of period k by 100,000, we obtain the empirical stockout probability for period k, i.e., the empirical kth period QoS constraint violation probability. Let $p_k(\mathbf{u}^*)$ denote the empirical QoS constraint violation probability for period k. In the interest of space, we report in Table 6 $p_N(\mathbf{u}^*)$ only. We note that we computationally confirmed that the QoS constraint violation probabilities are smaller for smaller values of k. This is due to the fact that the effects of the uncertain demand are compounded as k increases. In the table, "avg" denotes the average value of $p_N(\mathbf{u}^*)$ over the 20 instances of $(\overline{\mathbf{w}}, \hat{\mathbf{w}})$. We interpret "min" and "max" accordingly. The insight from Table 6 is that even $\mathbf{u}_S^*$ is an overly conservative ordering solution, let alone $\mathbf{u}_B^*$ and $\mathbf{u}_D^*$; the empirical QoS constraint violation probabilities are well below 0.05. This raises the question of how to further improve Bound (48), but in the meantime it demonstrates the necessity of using Bound (48) rather than Bound (47) and Bound (49) whenever possible.

4 Conclusion

In this paper we considered an LP problem in which each element of the constraint matrix is subject to uncertainty. To obtain a less conservative solution with a certain probabilistic guarantee of feasibility, we constructed the robust formulation,

Table 6 Empirical QoS constraint violation probabilities.

Distribution	$p_N(\mathbf{u}_B^*)$			$p_N(\mathbf{u}_D^*)$			$p_N(\mathbf{u}_S^*)$		
	avg	min	max	avg	min	max	avg	min	max
Triangle	0	0	0	.0000015	0	.00002	.00712	.00454	.00787
Uniform	0	0	0	.000007	0	.00004	.00658	.00456	.00730
Reverse-triangle	0	0	0	.0000125	0	.00004	.00633	.00536	.00723

which can be recast as an LP formulation. Using the probability distributions of the uncertain elements, we showed that we can improve the quality of a solution without compromising its robustness (quantified as the constraint violation probability of the solution). To that end, we derived a new bound on the constraint violation probability. This bound is distribution-dependent, but is independent of the solution. We showed that the bound is stronger than a distribution-independent bound given in [BS04]. We also derived another bound that utilizes the solution as well as the probability distributions. We proved that this solution-dependent bound is stronger (and in our numerical tests significantly so) than all other bounds we presented. The case where only limited distributional information on the uncertain elements is available was also considered. We derived two bounds that exploit the first moments, and the first and second moments of the uncertain elements, respectively, and the solution. We showed that these bounds are also stronger than the bound given in [BS04].

As an application, we considered a discrete-time stochastic inventory control problem with QoS constraints. We constructed a robust formulation and showed that its optimal ordering quantities form a base-stock policy. We derived two bounds on the probability that the QoS constraints are violated. We explained how these two bounds can be used together to obtain a better solution of the inventory control problem via the so-called inverse problem. In some of the examples we provided, cost savings amount to 36%–54%.

In closing, distributional information can be very valuable in a robust optimization context. Thus, when it is not readily available, one should consider estimating it because the benefits can significantly outweigh the estimation costs.

Appendix

A Proof of Lemma 2

Proof: Consider $\mathbf{\Gamma}^1$ and $\mathbf{\Gamma}^2$, where $\mathbf{\Gamma}^1 \leq \mathbf{\Gamma}^2$. Since for any $\mathbf{x}$

$$\max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i^1)} \{\mathbf{a}_i'\mathbf{x}\} \leq \max_{\mathbf{a}_i \in \mathscr{R}_i(\Gamma_i^2)} \{\mathbf{a}_i'\mathbf{x}\}, \quad i = 1, \ldots, m,$$

the feasible region of (5) with $\mathbf{\Gamma}^1$ is no smaller than that of (5) with $\mathbf{\Gamma}^2$. Therefore, $z_R(\mathbf{\Gamma})$ is nonincreasing as $\mathbf{\Gamma}$ increases componentwise. If $\Gamma_i = 0$ for all i, $\mathscr{R}_i(\Gamma_i) = \{\bar{\mathbf{a}}_i\}$ and consequently (5) is reduced to (2). On the other hand, if $\Gamma_i = |J_i|$ for all i, the uncertainty budget constraint becomes redundant, making $\mathscr{R}_i(\Gamma_i) = \mathscr{U}_i$. Hence, it follows that $z_F \leq z_R(\mathbf{\Gamma}) \leq z_N$. □

B Reshaping the LP formulation (8) into another LP formulation (9)

Notice that for any i and $j \in J_i$, at most one of the constraints (6a) and (6b) is binding at optimality. Therefore, the corresponding dual variables satisfy $\lambda_{ij}\mu_{ij} = 0$. Since $\lambda_{ij}, \mu_{ij} \geq 0$, this implies that $\lambda_{ij} + \mu_{ij} = |\lambda_{ij} - \mu_{ij}|$. The similar observation regarding the constraints (6d) and (6e) leads to $\nu_{ij}\tau_{ij} = 0$ and $\nu_{ij} + \tau_{ij} = |\nu_{ij} - \tau_{ij}|$. It can also be noted that $\lambda_{ij} - \mu_{ij}$ and $\nu_{ij} - \tau_{ij}$ have the same sign: if $\lambda_{ij} > 0$, then $\nu_{ij} > 0$. (if $\mu_{ij} > 0$, then $\tau_{ij} > 0$.) Using these observations, we obtain from the constraints (8a)

$$|\lambda_{ij} - \mu_{ij} + \nu_{ij} - \tau_{ij}| = |x_j| \quad \Leftrightarrow \quad |\lambda_{ij} - \mu_{ij}| + |\nu_{ij} - \tau_{ij}| = |x_j| \tag{51}$$

$$\Leftrightarrow \quad \lambda_{ij} + \mu_{ij} + \nu_{ij} + \tau_{ij} = |x_j|, \tag{52}$$

where the relation in (51) follows from the fact that $\lambda_{ij} - \mu_{ij}$ and $\nu_{ij} - \tau_{ij}$ have the same sign. Using (52), the constraints (8b) can be rewritten as

$$z_i - \hat{a}_{ij}(\nu_{ij} + \tau_{ij}) \geq 0 \quad \Leftrightarrow \quad z_i + \hat{a}_{ij}(\lambda_{ij} + \nu_{ij}) \geq \hat{a}_{ij}|x_j|.$$

By defining $p_{ij} \triangleq \hat{a}_{ij}(\lambda_{ij} + \mu_{ij})$ and substituting y_j for $|x_j|$ with the addition of the constraints $-y_j \leq x_j \leq y_j$ and $y_j \geq 0$, the LP formulation (8) becomes the LP formulation (9).

C Proof of Proposition 1

Proof: Let $(\tilde{\mathbf{u}}, \tilde{z}, \tilde{p}_k, \tilde{q}_{kj}, \tilde{r}, \tilde{s}_k)$ be an optimal solution of (46). Let $(\mathbf{u}^*, z^*)$ be an optimal solution of (36). We will show that $(\tilde{\mathbf{u}}, \tilde{z})$ is feasible to (36) with $\tilde{z} = z^*$. For $k = \lfloor\Gamma\rfloor+1, \ldots, N$, we have

$$\begin{aligned}\sum_{j=1}^{k} \tilde{u}_j &\geq -x_1 + \sum_{j=1}^{k} \overline{w}_j + \Gamma\tilde{p}_k + \sum_{j=1}^{k} \tilde{q}_{kj} \\ &\geq -x_1 + \sum_{j=1}^{k} \overline{w}_j + \max_{S_k \cup \{t_k\}} \Big\{ \sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor\Gamma\rfloor)\hat{w}_{t_k} \Big\},\end{aligned}$$

where the second inequality follows from $(\tilde{p}_k, \tilde{q}_{kj})$ being feasible to (43) and the weak duality between (43) and (42). We also have

$$\begin{aligned}\tilde{z} - \sum_{k=1}^{N}(c + ha_k)\tilde{u}_k &\geq Nhx_1 - h\sum_{k=1}^{N} a_k\overline{w}_k + h\Big(\Gamma\tilde{r} + \sum_{k=1}^{N}\tilde{s}_k\Big) \\ &\geq Nhx_1 - h\sum_{k=1}^{N} a_k\overline{w}_k + h\max_{S_N \cup \{t_N\}}\Big\{\sum_{j \in S_N} a_j\hat{w}_j + (\Gamma - \lfloor\Gamma\rfloor)a_{t_N}\hat{w}_{t_N}\Big\},\end{aligned}$$

where the second inequality follows from $(\tilde{r}, \tilde{s}_k)$ being feasible to (45) and the weak duality between (45) and (44). These inequalities show that $(\tilde{\mathbf{u}}, \tilde{z})$ is a feasible solution of (36). Hence $\tilde{z} \geq z^*$.

For $k = \lfloor\Gamma\rfloor+1, \ldots, N$, let (p_k^*, q_{kj}^*) be an optimal solution of (43). Then by the strong duality between (43) and (42), we have $\Gamma p_k^* + \sum_{j=1}^{k} q_{kj}^* = \max_{S_k \cup \{t_k\}}\{\sum_{j \in S_k} \hat{w}_j + (\Gamma - \lfloor\Gamma\rfloor)\hat{w}_{t_k}\}$. Let (r^*, s_k^*) be an optimal solution of (45). Again the strong duality between (45) and (44) implies $\Gamma r^* + \sum_{k=1}^{N} s_k^* = \max_{S_N \cup \{t_N\}}\{\sum_{j \in S_N} a_j\hat{w}_j + (\Gamma - \lfloor\Gamma\rfloor)a_{t_N}\hat{w}_{t_N}\}$. From these equalities, it can be seen that $(\mathbf{u}^*, z^*, p_k^*, q_{kj}^*, r^*, s_k^*)$ is a feasible solution of (46), from which we have $z^* \geq \tilde{z}$. □

References

[BBC11] D. Bertsimas, D. Brown, and C. Caramanis, *Theory and Applications of Robust Optimization*, SIAM review **53** (2011), no. 3, 464–501.

[Ben62] G. Bennett, *Probability inequalities for the sum of independent random variables*, Journal of the American Statistical Association **57** (1962), 33–45.

[BG12] D. Bertsimas and V. Goyal, *On the power and limitations of affine policies in two-stage adaptive optimization*, Mathematical programming **134** (2012), no. 2, 491–531.

[BIP11] D. Bertsimas, D. A. Iancu, and P. A. Parrilo, *A hierarchy of near-optimal policies for multistage adaptive optimization*, IEEE Trans. Automat. Contr. **56** (2011), no. 12, 2809–2824.

[BL97] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*, Springer, New York, 1997.

[BO08] D. Bienstock and N. Özbay, *Computing robust basestock levels*, Discrete Optimization **5** (2008), no. 2, 389–414.

[BP01] D. Bertsimas and I. Ch. Paschalidis, *Probabilistic service level guarantees in make-to-stock manufacturing systems*, Operations Research **49** (2001), no. 1, 119–133.

[BPS04] D. Bertsimas, D. Pachamanova, and M. Sim, *Robust linear optimization under general norms*, Operations Research Letters **32** (2004), no. 6, 510–516.

[BS04] D. Bertsimas and M. Sim, *The price of robustness*, Operations Research **52** (2004), no. 1, 35–53.

[BT06] D. Bertsimas and A. Thiele, *A robust optimization approach to inventory theory*, Operations Research **54** (2006), no. 1, 150–168.

[BTGGN04] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, *Adjustable robust solutions of uncertain linear programs*, Mathematical Programming **99** (2004), no. 2, 351–376.

[BTGN09] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*, Princeton, 2009.

[BTN99] A. Ben-Tal and A. Nemirovski, *Robust solutions of uncertain linear programs*, Operations Research Letters **25** (1999), no. 1, 1–13.

[CG06] G. Calafiore and L. El Ghaoui, *Distributionally robust chance-constrained linear programs with applications*, J. Optimization Th. Applicat. (2006).

[CS60] A. J. Clark and H. E. Scarf, *Optimal policies for a multi-echelon inventory problem*, Management Science **6** (1960), 475–490.

[CSS07] X. Chen, M. Sim, and P. Sun, *A robust optimization perspective on stochastic programming*, Operations Research **55** (2007), no. 6, 1058–1071.

[CSST10] W. Chen, M. Sim, J. Sun, and C.-P. Teo, *From CVaR to uncertainty set: Implications in joint chance-constrained optimization*, Operations research **58** (2010), no. 2, 470–485.

[DDB95] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of uncertain systems: A linear programming approach*, Prentice Hall, 1995.

[DVP06] C. Del Vecchio and I. Ch. Paschalidis, *Enforcing service-level constraints in supply chains with assembly operations*, IEEE Trans. Automat. Contr. **51** (2006), no. 12, 2000–2005.

[DZ98] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd ed., Springer-Verlag, NY, 1998.

[KW94] P. Kall and S. W. Wallace, *Stochastic programming*, John Wiley & Sons, Chichester, UK, 1994.

[NS06] A. Nemirovski and A. Shapiro, *Convex approximations of chance constrained programs*, SIAM Journal on Optimization **17** (2006), no. 4, 969–996.

[PL03] I. Ch. Paschalidis and Y. Liu, *Large deviations-based asymptotics for inventory control in supply chains*, Operations Research **51** (2003), no. 3, 437–460.

[PLCP04] I. Ch. Paschalidis, Y. Liu, C. G. Cassandras, and C. Panayiotou, *Inventory control for supply chains with service level constraints: A synergy between large deviations and perturbation analysis*, The Annals of Operations Research **126** (2004), 231–258, (Special Volume on Stochastic Models of Production-Inventory Systems).

[Soy73] A. L. Soyster, *Convex programming with set-inclusive constraints and applications to inexact linear programming*, Operations Research **21** (1973), 1154–1157.

[SS10] C.-T. See and M. Sim, *Robust approximation to multiperiod inventory management*, Operations research **58** (2010), no. 3, 583–594.

[ZKR13] S. Zymler, D. Kuhn, and B. Rustem, *Distributionally robust joint chance constraints with second-order moment information*, Mathematical Programming **137** (2013), no. 1-2, 167–198.