The Price of Robustness

Dimitris Bertsimas [*] Melvyn Sim [†]

August 2001; revised August, 2002

Abstract

A robust approach to solving linear optimization problems with uncertain data has been proposed in the early 1970s, and has recently been extensively studied and extended. Under this approach, we are willing to accept a suboptimal solution for the nominal values of the data, in order to ensure that the solution remains feasible and near optimal when the data changes. A concern with such an approach is that it might be too conservative. In this paper we propose an approach that attempts to make this tradeoff more attractive, that is we investigate ways to decrease what we call the price of robustness. In particular, we flexibly adjust the level of conservatism of the robust solutions in terms of probabilistic bounds of constraint violations. An attractive aspect of our method is that the new robust formulation is also a linear optimization problem. We thus naturally extend our methods to discrete optimization problems in a tractable way. We report numerical results for a portfolio optimization problem, a knapsack problem, and a problem from the Net Lib library.

[*]Boeing Professor of Operations Research, Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, E53-363, Cambridge, MA 02139, dbertsim@mit.edu. The research of the author was partially supported by the Singapore-MIT alliance.

[†]Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, melvyn@mit.edu. The research of the author is supported by a graduate scholarship from the National University of Singapore

1 Introduction

The classical paradigm in mathematical programming is to develop a model that assumes that the input data is precisely known and equal to some nominal values. This approach, however, does not take into account the influence of data uncertainties on the quality and feasibility of the model. It is therefore conceivable that as the data takes values different than the nominal ones several constraints may be violated, and the optimal solution found using the nominal data may be no longer optimal or even feasible. This observation raises the natural question of designing solution approaches that are immune to data uncertainty, that is they are "robust."

To illustrate the importance of robustness in practical applications, we quote from the case study by Ben-Tal and Nemirovski [1] on linear optimization problems from the Net Lib library:

> In real-world applications of Linear Programming, one cannot ignore the possibility that a small uncertainty in the data can make the usual optimal solution completely meaningless from a practical viewpoint.

The need naturally arises to develop models that are immune, as far as possible, to data uncertainty. The first step in this direction was taken by Soyster [9] who proposes a linear optimization model to construct a solution that is feasible for all data that belong in a convex set. The resulting model produces solutions that are too conservative in the sense that we give up too much of optimality for the nominal problem in order to ensure robustness (see the comments of Ben-Tal and Nemirovski [1]). Soyster considers the linear optimization problem

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c'x} \\ \text{subject to} & \sum_{j=1}^{n} \boldsymbol{A_j} x_j \leq \boldsymbol{b}, \qquad \forall \boldsymbol{A_j} \in K_j,\ j = 1, \ldots, n \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{array}$$

where the uncertainty sets K_j are convex. Note that the case considered is "column-wise" uncertainty, i.e., the columns $\boldsymbol{A_j}$ of the constraint matrix are known to belong to a given convex set K_j. Soyster shows that the problem is equivalent to

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c'x} \\ \text{subject to} & \sum_{j=1}^{n} \overline{\boldsymbol{A}}_j x_j \leq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{array} \tag{1}$$

where $\overline{a}_{ij} = \sup_{\boldsymbol{A_j} \in K_j}(A_{ij})$.

A significant step forward for developing a theory for robust optimization was taken independently by Ben-Tal and Nemirovski [1, 2, 3] and El-Ghaoui et al. [5, 6]. To address the issue of over-conservatism, these papers proposed less conservative models by considering uncertain linear problems with ellipsoidal uncertainties, which involve solving the *robust counterparts* of the nominal problem in the form of conic quadratic problems (see [2]). With properly chosen ellipsoids, such a formulation can be used as a reasonable approximation to more complicated uncertainty sets. However, a practical drawback of such an approach, is that it leads to nonlinear, although convex, models, which are more demanding computationally than the earlier linear models by Soyster [9] (see also the discussion in [1]).

In this research we propose a new approach for robust linear optimization that retains the advantages of the linear framework of Soyster [9]. More importantly, our approach offers full control on the degree of conservatism for every constraint. We protect against violation of constraint i deterministically, when only a prespecified number Γ_i of the coefficients changes, that is we guarantee that the solution is feasible if less than Γ_i uncertain coefficients change. Moreover, we provide a probabilistic guarantee that even if more than Γ_i change, then the robust solution will be feasible with high probability. In the process we prove a new, to the best of our knowledge, tight bound on sums of symmetrically distributed random variables. In this way, the proposed framework is at least as flexible as the one proposed by Ben-Tal and Nemirovski [1, 2, 3] and El-Ghaoui et al. [5, 6] and possibly more. Unlike these approaches, the robust counterparts we propose are linear optimization problems, and thus our approach readily generalizes to discrete optimization problems. To the best of our knowledge, there was no similar work done in the robust discrete optimization domain that involves deterministic and probabilistic guarantees of constraints against violation.

Structure of the paper. In Section 2, we present the different approaches for robust linear optimization from the literature and discuss their merits. In Section 3 we propose the new approach and show that it can be solved as a linear optimization problem. In Section 4, we show that the proposed robust LP has attractive probabilistic and deterministic guarantees. Moreover, we perform sensitivity analysis of the degree of protection the proposed method offers. We provide extensions to our basic framework dealing with correlated uncertain data in Section 5. In Section 6, we apply the proposed approach to a portfolio problem, a knapsack problem and a problem from the Net Lib library. Finally, Section 7 contains some concluding remarks.

2 Robust Formulation of Linear Programming Problems

2.1 Data uncertainty in linear optimization

We consider the following nominal linear optimization problem:

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u}. \end{aligned}$$

In the above formulation, we assume that data uncertainty only affects the elements in matrix $\boldsymbol{A}$. We assume without loss of generality that the objective function $\boldsymbol{c}$ is not subject to uncertainty, since we can use the objective maximize z, add the constraint $z - \boldsymbol{c}'\boldsymbol{x} \leq 0$, and thus include this constraint into $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$.

Model of Data Uncertainty U:

Consider a particular row i of the matrix $\boldsymbol{A}$ and let J_i the set of coefficients in row i that are subject to uncertainty. Each entry $a_{ij},\ j \in J_i$ is modeled as a symmetric and bounded random variable $\tilde{a}_{ij}, j \in J_i$ (see [1]) that takes values in $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. Associated with the uncertain data $\tilde{a}_{ij}$, we define the random variable $\eta_{ij} = (\tilde{a}_{ij} - a_{ij})/\hat{a}_{ij}$, which obeys an unknown, but symmetric distribution, and takes values in $[-1, 1]$.

2.2 The robust formulation of Soyster

As we have mentioned in the Introduction Soyster [9] considers column-wise uncertainty. Under the model of data uncertainty U, the robust Formulation (1) is as follows:

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{subject to} \quad & \sum_j a_{ij} x_j + \sum_{j \in J_i} \hat{a}_{ij} y_j \leq b_i \quad \forall i \\ & -y_j \leq x_j \leq y_j \quad\quad\quad\quad \forall j \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} \\ & \boldsymbol{y} \geq \boldsymbol{0}. \end{aligned} \tag{2}$$

Let $\boldsymbol{x}^*$ be the optimal solution of Formulation (2). At optimality clearly, $y_j = |x_j^*|$, and thus

$$\sum_j a_{ij} x_j^* + \sum_{j \in J_i} \hat{a}_{ij} |x_j^*| \leq b_i \quad \forall i.$$

We next show that for every possible realization $\tilde{a}_{ij}$ of the uncertain data, the solution remains feasible, that is the solution is "robust." We have

$$\sum_j \tilde{a}_{ij} x_j^* = \sum_j a_{ij} x_j^* + \sum_{j \in J_i} \eta_{ij} \hat{a}_{ij} x_j^* \leq \sum_j a_{ij} x_j^* + \sum_{j \in J_i} \hat{a}_{ij} |x_j^*| \leq b_i \quad \forall i$$

For every ith constraint, the term, $\sum_{j \in J_i} \hat{a}_{ij} |x_j|$ gives the necessary "protection" of the constraint by maintaining a gap between $\sum_j a_{ij} x_j^*$ and b_i.

2.3 The robust formulation of Ben-Tal and Nemirovski

Although the Soyster's method admits the highest protection, it is also the most conservative in practice in the sense that the robust solution has an objective function value much worse than the objective function value of the solution of the nominal linear optimization problem. To address this conservatism, Ben-Tal and Nemirovski [1] propose the following robust problem:

$$\begin{array}{lll} \text{maximize} & \boldsymbol{c}'\boldsymbol{x} & \\ \text{subject to} & \sum_j a_{ij} x_j + \sum_{j \in J_i} \hat{a}_{ij} y_{ij} + \Omega_i \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 z_{ij}^2} \leq b_i & \forall i \\ & -y_{ij} \leq x_j - z_{ij} \leq y_{ij} & \forall i, j \in J_i \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} & \\ & \boldsymbol{y} \geq \boldsymbol{0}. & \end{array} \tag{3}$$

Under the model of data uncertainty U, the authors have shown that the probability that the i constraint is violated is at most $\exp(-\Omega_i^2/2)$. The robust Model (3) is less conservative than Model (2) as every feasible solution of the latter problem is a feasible solution to the former problem.

We next examine the sizes of Formulations (2) and (3). We assume that there are k coefficients of the $m \times n$ nominal matrix $\boldsymbol{A}$ that are subject to uncertainty. Given that the original nominal problem has n variables and m constraints (not counting the bound constraints), Model (2) is a linear optimization problem with $2n$ variables, and $m + 2n$ constraints. In contrast, Model (3) is a second order cone problem, with $n + 2k$ variables and $m + 2k$ constraints. Since Model (3) is a nonlinear one, it is particularly not attractive for solving robust discrete optimization models.

3 The New Robust Approach

In this section, we propose a robust formulation that is linear, is able to withstand parameter uncertainty under the model of data uncertainty U without excessively affecting the objective function, and readily extends to discrete optimization problems.

We motivate the formulation as follows. Consider the ith constraint of the nominal problem $\boldsymbol{a_i'x} \leq b_i$. Let J_i be the set of coefficients a_{ij}, $j \in J_i$ that are subject to parameter uncertainty, i.e., $\tilde{a}_{ij}$, $j \in J_i$ takes values according to a symmetric distribution with mean equal to the nominal value a_{ij} in the interval $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. For every i, we introduce a parameter Γ_i, not necessarily integer, that takes values in the interval $[0, |J_i|]$. As it would become clear below, the role of the parameter Γ_i is to adjust the robustness of the proposed method against the level of conservatism of the solution. Speaking intuitively, it is unlikely that all of the a_{ij}, $j \in J_i$ will change. Our goal is to be protected against all cases that up to $\lfloor \Gamma_i \rfloor$ of these coefficients are allowed to change, and one coefficient a_{it} changes by $(\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it}$. In other words, we stipulate that nature will be restricted in its behavior, in that only a subset of the coefficients will change in order to adversely affect the solution. We will develop an approach, that has the property that if nature behaves like this, then the robust solution will be feasible **deterministically**, and moreover, even if more than $\lfloor \Gamma_i \rfloor$ change, then the robust solution will be feasible **with very high probability**.

We consider the following (still nonlinear) formulation:

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{c'x} \\ \text{subject to} \quad & \sum_j a_{ij} x_j + \max_{\{S_i \cup \{t_i\} |\ S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} y_j + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} y_t \right\} \leq b_i \quad \forall i \\ & -y_j \leq x_j \leq y_j \quad \forall j \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} \\ & \boldsymbol{y} \geq \boldsymbol{0}. \end{aligned} \tag{4}$$

If Γ_i is chosen as an integer, the ith constraint is protected by $\beta_i(\boldsymbol{x}, \Gamma_i) = \max_{\{S_i |\ S_i \subseteq J_i, |S_i| = \Gamma_i\}} \{\sum_{j \in S_i} \hat{a}_{ij} |x_j|\}$. Note that when $\Gamma_i = 0$, $\beta_i(\boldsymbol{x}, \Gamma_i) = 0$ the constraints are equivalent to that of the nominal problem. Likewise, if $\Gamma_i = |J_i|$, we have Soyster's method. Therefore, by varying $\Gamma_i \in [0, |J_i|]$, we have the flexibility of adjusting the robustness of the method against the level of conservatism of the solution.

In order to reformulate Model (4) as a linear optimization model we need the following proposition.

Proposition 1 *Given a vector* $\boldsymbol{x}^*$, *the protection function of the* i*th constraint,*

$$\beta_i(\boldsymbol{x}^*, \Gamma_i) = \max_{\{S_i \cup \{t_i\} |\ S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} |x_j^*| + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} |x_j^*| \right\} \tag{5}$$

equals to the objective function of the following linear optimization problem:

$$\begin{aligned} \beta_i(\boldsymbol{x^*}, \Gamma_i) = \text{maximize} \quad & \sum_{j \in J_i} \hat{a}_{ij} |x_j^*| z_{ij} \\ \text{subject to} \quad & \sum_{j \in J_i} z_{ij} \le \Gamma_i \\ & 0 \le z_{ij} \le 1 \qquad \forall j \in J_i. \end{aligned} \tag{6}$$

Proof : Clearly the optimal solution value of Problem (6) consists of $\lfloor \Gamma_i \rfloor$ variables at 1 and one variable at $\Gamma_i - \lfloor \Gamma_i \rfloor$. This is equivalent to the selection of subset $\{S_i \cup \{t_i\} | \ S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \backslash S_i\}$ with corresponding cost function $\sum_{j \in S_i} \hat{a}_{ij} |x_j^*| + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} |x_j^*|$. ■

We next reformulate Model (4) as a linear optimization model.

Theorem 1 *Model (4) has an equivalent linear formulation as follows:*

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{c'x} \\ \text{subject to} \quad & \sum_j a_{ij} x_j + z_i \Gamma_i + \sum_{j \in J_i} p_{ij} \le b_i && \forall i \\ & z_i + p_{ij} \ge \hat{a}_{ij} y_j && \forall i, j \in J_i \\ & -y_j \le x_j \le y_j && \forall j \\ & l_j \le x_j \le u_j && \forall j \\ & p_{ij} \ge 0 && \forall i, j \in J_i \\ & y_j \ge 0 && \forall j \\ & z_i \ge 0 && \forall i. \end{aligned} \tag{7}$$

Proof : We first consider the dual of Problem (6):

$$\begin{aligned} \text{minimize} \quad & \sum_{j \in J_i} p_{ij} + \Gamma_i z_i \\ \text{subject to} \quad & z_i + p_{ij} \ge \hat{a}_{ij} |x_j^*| && \forall i, j \in J_i \\ & p_{ij} \ge 0 && \forall j \in J_i \\ & z_i \ge 0 && \forall i. \end{aligned} \tag{8}$$

By strong duality, since Problem (6) is feasible and bounded for all $\Gamma_i \in [0, |J_i|]$, then the dual problem (8) is also feasible and bounded and their objective values coincide. Using Proposition 1, we have that $\beta_i(\boldsymbol{x^*}, \Gamma_i)$ is equal to the objective function value of Problem 8. Substituting to Problem (4) we obtain that Problem (4) is equivalent to the linear optimization problem (7). ■

Remark : The robust linear optimization Model (7) has $n + k + 1$ variables and $m + k + n$ constraints, where $k = \sum_i |J_i|$ the number of uncertain data, contrasted with $n + 2k$ variables and $m + 2k$ constraints for the nonlinear Formulation (3). In most real-world applications, the matrix $\boldsymbol{A}$ is sparse. An attractive characteristic of Formulation (7) is that it preserves the sparsity of the matrix $\boldsymbol{A}$.

4 Probability Bounds of Constraint Violation

It is clear by the construction of the robust formulation that if up to $\lfloor \Gamma_i \rfloor$ of the J_i coefficients a_{ij} change within their bounds, and up to one coefficient a_{it_i} changes by $(\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it}$, then the solution of Problem (7) will remain feasible. In this section, we show that under the Model of Data Uncertainty U, the robust solution is feasible with high probability. The parameter Γ_i controls the tradeoff between the probability of violation and the effect to the objective function of the nominal problem, which is what we call **the price of robustness**.

In preparation for our main result in this section, we prove the following proposition.

Proposition 2 *Let $\boldsymbol{x^*}$ be an optimal solution of Problem (7). Let S_i^* and t_i^* be the set and the index respectively that achieve the maximum for $\beta_i(\boldsymbol{x^*}, \Gamma_i)$ in Eq. (5). Suppose that the data in matrix $\boldsymbol{A}$ are subjected to the model of data uncertainty U.*

(a) *The probability that the ith constraint is violated satisfies:*

$$\Pr\left(\sum_j \tilde{a}_{ij} x_j^* > b_i\right) \leq \Pr\left(\sum_{j \in J_i} \gamma_{ij}\eta_{ij} \geq \Gamma_i\right)$$

where

$$\gamma_{ij} = \begin{cases} 1, & \text{if } j \in S_i^* \\ \dfrac{\hat{a}_{ij}|x_j^*|}{\hat{a}_{ir^*}|x_{r^*}^*|}, & \text{if } j \in J_i \backslash S_i^* \end{cases}$$

and

$$r^* = \arg \min_{r \in S_i^* \cup \{t_i^*\}} \hat{a}_{ir}|x_r^*|.$$

(b) *The quantities γ_{ij} satisfy $\gamma_{ij} \leq 1$ for all $j \in J_i \backslash S_i^*$.*

Proof : **(a)** Let x^*, S_i^* and t_i^* be the solution of Model (4). Then the probability of violation of the ith constraint is as follows:

$$\begin{aligned} \Pr\left(\sum_j \tilde{a}_{ij} x_j^* > b_i\right) &= \Pr\left(\sum_j a_{ij} x_j^* + \sum_{j \in J_i} \eta_{ij}\hat{a}_{ij} x_j^* > b_i\right) \\ &\leq \Pr\left(\sum_{j \in J_i} \eta_{ij}\hat{a}_{ij}|x_j^*| > \sum_{j \in S_i^*} \hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*|\right) \end{aligned} \quad (9)$$

$$\begin{aligned}
&= \Pr\left(\sum_{j\in J_i\backslash S_i^*} \eta_{ij}\hat{a}_{ij}|x_j^*| > \sum_{j\in S_i^*} \hat{a}_{ij}|x_j^*|(1-\eta_{ij}) + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*|\right) \\
&\leq \Pr\left(\sum_{j\in J_i\backslash S_i^*} \eta_{ij}\hat{a}_{ij}|x_j^*| > \hat{a}_{ir^*}|x_{r^*}^*|\left(\sum_{j\in S_i^*}(1-\eta_{ij}) + (\Gamma_i - \lfloor\Gamma_i\rfloor)\right)\right) \qquad (10) \\
&= \Pr\left(\sum_{j\in S_i^*}\eta_{ij} + \sum_{j\in J_i\backslash S_i^*}\frac{\hat{a}_{ij}|x_j^*|}{\hat{a}_{ir^*}|x_{r^*}^*|}\eta_{ij} > \Gamma_i\right) \\
&= \Pr\left(\sum_{j\in J_i}\gamma_{ij}\eta_{ij} > \Gamma_i\right) \\
&\leq \Pr\left(\sum_{j\in J_i}\gamma_{ij}\eta_{ij} \geq \Gamma_i\right).
\end{aligned}$$

Inequality (9) follows from Inequality (4), since $\boldsymbol{x^*}$ satisfies

$$\sum_j a_{ij}x_j^* + \sum_{j\in S_i^*}\hat{a}_{ij}y_j + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}y_{t_i^*} \leq b_i.$$

Inequality (10) follows from $1-\eta_{ij}\geq 0$ and $r^* = \arg\min_{r\in S_i^*\cup\{t_i^*\}}\hat{a}_{ir}|x_r^*|$.

(**b**) Suppose there exist $l \in J_i\backslash S_i^*$ such that $\hat{a}_{il}|x_l^*| > \hat{a}_{ir^*}|x_{r^*}^*|$. If $l \neq t_i^*$, then, since $r^* \in S_i^* \cup \{t^*\}$, we could increase the objective value of $\sum_{j\in S_i^*}\hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it^*}|x_{t^*}^*|$ by exchanging l with r^* from the set $S_i^*\cup\{t_i^*\}$. Likewise, if $l = t_i^*$, $r^*\in S_i^*$, we could exchange t_i^* with r^* in the set S_i^* to increase the same objective function. In both cases, we arrive at a contradiction that $S_i^*\cup\{t_i^*\}$ is an optimum solution to this objective function. ■

We are naturally led to bound the probability $\Pr(\sum_{j\in J_i}\gamma_{ij}\eta_{ij} \geq \Gamma_i)$. The next result provides a bound that is independent of the solution $\boldsymbol{x^*}$.

Theorem 2 *If $\eta_{ij}, j\in J_i$ are independent and symmetrically distributed random variables in $[-1,1]$, then*

$$\Pr\left(\sum_{j\in J_i}\gamma_{ij}\eta_{ij} \geq \Gamma_i\right) \leq \exp\left(-\frac{\Gamma_i^2}{2|J_i|}\right). \qquad (11)$$

Proof : Let $\theta > 0$. Then

$$\begin{aligned}
\Pr\left(\sum_{j\in J_i}\gamma_{ij}\eta_{ij} \geq \Gamma_i\right) &\leq \frac{E[\exp(\theta\sum_{j\in J_i}\gamma_{ij}\eta_{ij})]}{\exp(\theta\Gamma_i)} \qquad (12) \\
&= \frac{\prod_{j\in J_i} E[\exp(\theta\ \gamma_{ij}\eta_{ij})]}{\exp(\theta\Gamma_i)} \qquad (13)
\end{aligned}$$

$$= \frac{\prod_{j\in J_i} 2\int_0^1 \sum_{k=0}^{\infty} \frac{(\theta\gamma_{ij}\eta)^{2k}}{(2k)!} dF_{\eta_{ij}}(\eta)}{\exp(\theta\Gamma_i)} \tag{14}$$

$$\leq \frac{\prod_{j\in J_i} \sum_{k=0}^{\infty} \frac{(\theta\gamma_{ij})^{2k}}{(2k)!}}{\exp(\theta\Gamma_i)}$$

$$\leq \frac{\prod_{j\in J_i} \exp\left(\frac{\theta^2\gamma_{ij}^2}{2}\right)}{\exp(\theta\Gamma_i)}$$

$$\leq \exp\left(|J_i|\frac{\theta^2}{2} - \theta\Gamma_i\right). \tag{15}$$

Inequality (12) follows from Markov's inequality, Eqs. (13) and (14) follow from the independence and symmetric distribution assumption of the random variables η_{ij}. Inequality (15) follows from $\gamma_{ij} \leq 1$. Selecting $\theta = \Gamma_i/|J_i|$, we obtain (11). ■

Remark : While the bound we established has the attractive feature that is independent of the solution $\boldsymbol{x}^*$, it is not particularly attractive especially when $\frac{\Gamma_i^2}{2|J_i|}$ is small. We next derive the best possible bound, i.e., a bound that is achievable. We assume that $\Gamma_i \geq 1$.

Theorem 3 **(a)** *If $\eta_{ij}, j \in J_i$ are independent and symmetrically distrubuted random variables in $[-1,1]$, then*

$$\Pr\left(\sum_{j\in J_i} \gamma_{ij}\eta_{ij} \geq \Gamma_i\right) \leq B(n,\Gamma_i), \tag{16}$$

where

$$B(n,\Gamma_i) = \frac{1}{2^n}\left\{(1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{n}{l} + \mu\sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right\} = \frac{1}{2^n}\left\{(1-\mu)\binom{n}{\lfloor\nu\rfloor} + \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right\}, \tag{17}$$

where $n = |J_i|$, $\nu = \frac{\Gamma_i+n}{2}$ and $\mu = \nu - \lfloor\nu\rfloor$.

(b) *The bound (16) is tight for η_{ij} having a discrete probability distribution:* $\Pr(\eta_{ij} = 1) = 1/2$ *and* $\Pr(\eta_{ij} = -1) = 1/2$, *$\gamma_{ij} = 1$, an integral value of $\Gamma_i \geq 1$ and $\Gamma_i + n$ being even.*

(c) *The bound (16) satisfies*

$$B(n,\Gamma_i) \leq (1-\mu)C(n,\lfloor\nu\rfloor) + \sum_{l=\lfloor\nu\rfloor+1}^{n} C(n,l), \tag{18}$$

where

$$C(n,l) = \begin{cases} \frac{1}{2^n}, & \text{if } l\text{=0 or } l\text{=}n, \\ \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(n-l)l}}\exp\left(n\log\left(\frac{n}{2(n-l)}\right) + l\log\left(\frac{n-l}{l}\right)\right), & \text{otherwise.} \end{cases} \tag{19}$$

(d) *For* $\Gamma_i = \theta\sqrt{n}$,

$$\lim_{n\to\infty} B(n, \Gamma_i) = 1 - \Phi(\theta), \tag{20}$$

where

$$\Phi(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta} \exp\left(-\frac{y^2}{2}\right) dy$$

is the cumulative distribution function of a standard normal.

Proof : See appendix (Section 8).

Remarks:

(a) While Bound (16) is best possible (Theorem 3(b)), it poses computational difficulties in evaluating the sum of combination functions for large n. For this reason, we have calculated Bound (18), which is simple to compute and, as we will see, it is also very tight.

(b) Eq. (20) is a formal asymptotic theorem that applies when $\Gamma_i = \theta\sqrt{n}$. We can use the De Moivre-Laplace approximation of the Binomial distribution to obtain the approximation

$$B(n, \Gamma_i) \approx 1 - \Phi\left(\frac{\Gamma_i - 1}{\sqrt{n}}\right), \tag{21}$$

that applies, even when Γ_i does not scale as $\theta\sqrt{n}$.

(c) We next compare the bounds: (11) (Bound 1), (16) (Bound 2), (18) (Bound 3) and the approximate bound (21) for $n = |J_i| = 10, 100, 2000$. In Figure 1 we compare Bounds 1 and 2 for $n = 10$ that clearly show that Bound 2 dominates Bound 1 (in this case there is no need to calculate Bounds 3 and the approximate bound as n is small). In Figure 2 we compare all bounds for $n = 100$. It is clear that Bound 3, which is simple to compute, is identical to Bound 2, and both Bounds 2 and 3 dominate Bound 1 by an order of magnitude. The approximate bound provides a reasonable approximation to Bound 2. In Figure 3 we compare Bounds 1 and 3 and the approximate bound for $n = 2000$. Bound 3 is identical to the approximate bound, and both dominate Bound 1 by an order of magnitude. In summary, in the remainder of the paper, we will use Bound 3, as it is simple to compute, it is a true bound (as opposed to the approximate bound), and dominates Bound 1. To amplify this point, Table 1 illustrates the choice of Γ_i as a function of $n = |J_i|$ so that the probability that a constraint is violated is less than 1%, where we used Bounds 1, 2, 3 and the approximate bound to evaluate the probability. It is clear that using Bounds 2,3 or the approximate bound gives essentially identical values of Γ_i, while using Bound 1 leads to unnecessarily higher values of Γ_i. For $|J_i| = 200$, we need to use $\Gamma = 33.9$, i.e., only 17% of the number of uncertain data, to guarantee violation probability of less than 1%. For constraints with fewer number of uncertain data such as $|J_i| = 5$, it is necessary to ensure full protection, which is equivalent to the Soyster's method. Clearly, for constraints with large number of uncertain data, the proposed approach is capable of delivering less conservative solutions compared to the Soyster's method.

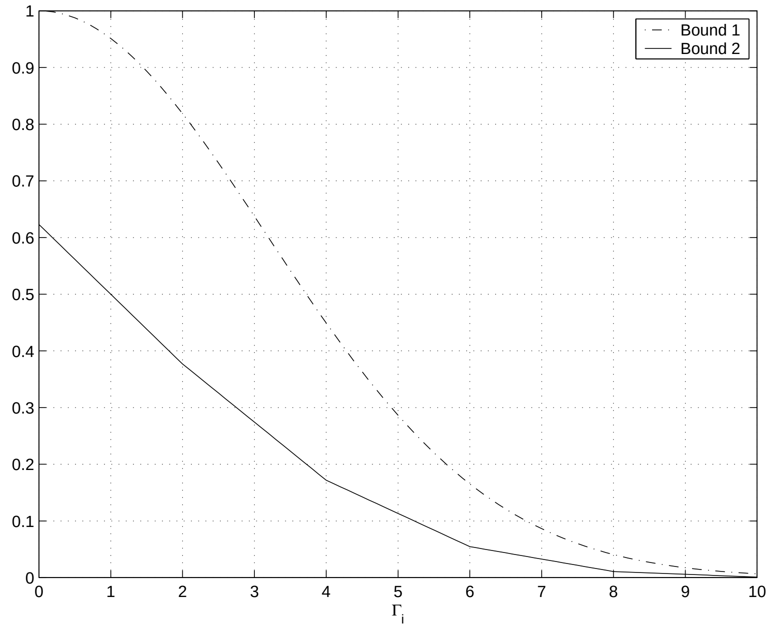

Figure 1: Comparison of probability bounds for $n = |J_i| = 10$.

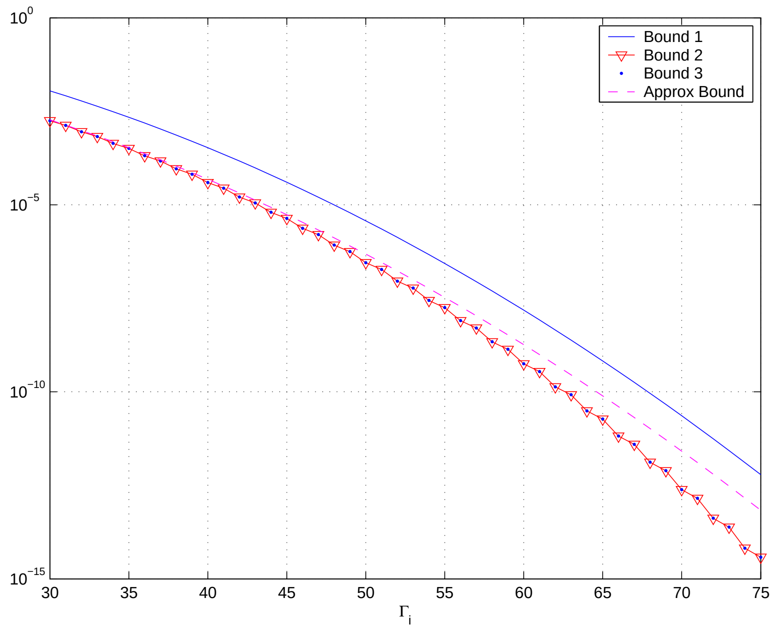

Figure 2: Comparison of probability bounds for $n = |J_i| = 100$.

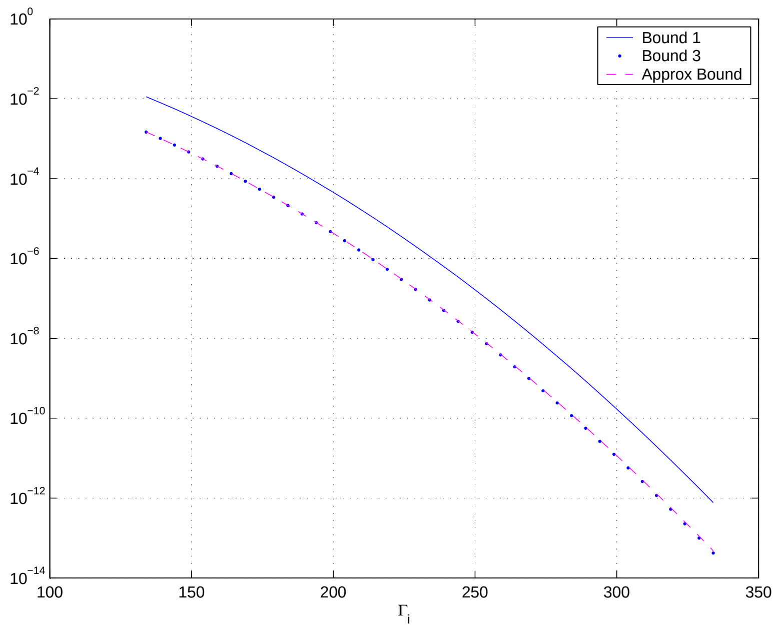

Figure 3: Comparison of probability bounds for $n = |J_i| = 2000$.

$\|J_i\|$	Γ_i **from Bound 1**	Γ_i **from Bounds 2, 3**	Γ_i **from Approx.**
5	5	5	5
10	9.6	8.2	8.4
100	30.3	24.3	24.3
200	42.9	33.9	33.9
2000	135.7	105	105

Table 1: Choice of Γ_i as a function of $n = |J_i|$ so that the probability of constraint violation is less than 1%.

4.1 On the conservatism of robust solutions

We have argued so far that the linear optimization framework of our approach has some computational advantages over the conic quadratic framework of Ben-Tal and Nemirovski [1, 2, 3] and El-Ghaoui et al. [5, 6] especially with respect to discrete optimization problems. Our objective in this section is to provide some insight, but not conclusive evidence, on the degree of conservatism for both approaches.

Given a constraint $\boldsymbol{a}'\boldsymbol{x} \leq b$, with $\boldsymbol{a} \in [\overline{\boldsymbol{a}} - \hat{\boldsymbol{a}}, \overline{\boldsymbol{a}} + \hat{\boldsymbol{a}}]$, the robust counterpart of Ben-Tal and Nemirovski [1, 2, 3] and El-Ghaoui et al. [5, 6] in its simplest form of ellipsoidal uncertainty (Formulation (3) includes combined interval and ellipsoidal uncertainty) is:

$$\overline{\boldsymbol{a}}'\boldsymbol{x} + \Omega||\hat{\boldsymbol{A}}\boldsymbol{x}|| \leq b,$$

where $\hat{\boldsymbol{A}}$ is a diagonal matrix with elements $\hat{a}_i$ in the diagonal. Ben-Tal and Nemirovski [1] show that under the model of data uncertainty U for $\boldsymbol{a}$, the probability that the constraint is violated is bounded above by $\exp(-\Omega^2/2)$.

The robust counterpart of the current approach is

$$\overline{\boldsymbol{a}}'\boldsymbol{x} + \beta(\boldsymbol{x}, \Gamma) \leq b,$$

where we assumed that Γ is integral and

$$\beta(\boldsymbol{x}, \Gamma) = \max_{S,\ |S|=\Gamma} \sum_{i \in S} \hat{a}_i |x_i|.$$

From Eq. (11), the probability that the constraint is violated under the model of data uncertainty U for $\boldsymbol{a}$ is bounded above by $\exp(-\Gamma^2/(2n))$. Note that we do not use the stronger bound (16) for simplicity.

Let us select $\Gamma = \Omega\sqrt{n}$ so that the bounds for the probability of violation are the same for both approaches. The protection levels are $\Omega||\hat{\boldsymbol{A}}\boldsymbol{x}||$ and $\beta(\boldsymbol{x}, \Gamma)$. We will compare the protection levels both from a worst and an average case point of view in order to obtain some insight on the degree of conservatism. To simplify the exposition we define $y_i = \hat{a}_i|x_i|$. We also assume without loss of generality that $y_1 \geq y_2 \geq \ldots \geq y_n \geq 0$. Then the two protection levels become $\Omega||\boldsymbol{y}||$ and $\sum_{i=1}^{\Gamma} y_i$.

For $\Gamma = \theta\sqrt{n}$, and $y_1^0 = \ldots = y_\Gamma^0 = 1$, $y_k^0 = 0$ for $k \geq \Gamma + 1$, we have $\sum_{i=1}^{\Gamma} y_i^0 = \Gamma = \theta\sqrt{n}$, while $\Omega||\boldsymbol{y}|| = \theta^{3/2} n^{1/4}$, i.e., in this example the protection level of the conic quadratic framework is asymptotically smaller than our framework by a multiplicative factor of $n^{1/4}$. This order of the magnitude is in fact worst possible, since

$$\sum_{i=1}^{\Gamma} y_i \leq \sqrt{\Gamma}||\boldsymbol{y}|| = \sqrt{\frac{n}{\Gamma}}(\Omega||\boldsymbol{y}||),$$

which for $\Gamma = \theta\sqrt{n}$ leads to

$$\sum_{i=1}^{\Gamma} y_i \leq \frac{n^{1/4}}{\theta^{1/2}}(\Omega||\boldsymbol{y}||).$$

Moreover, we have

$$\Omega||\boldsymbol{y}||_2 \;\leq\; \Omega\sqrt{\sum_{i=1}^{\Gamma} y_i^2 + y_\Gamma^2(n - \Gamma)}$$

$$
\begin{aligned}
&\leq \Omega \sqrt{\sum_{i=1}^{\Gamma} y_i^2 + \left(\frac{\sum_{i=1}^{\Gamma} y_i}{\Gamma} \right)^2 (n - \Gamma)} \\
&\leq \Omega \sqrt{\left(\sum_{i=1}^{\Gamma} y_i \right)^2 + \left(\sum_{i=1}^{\Gamma} y_i \right)^2 \left(\frac{n - \Gamma}{\Gamma^2} \right)} \\
&= \frac{\Gamma}{\sqrt{n}} \sum_{i=1}^{\Gamma} y_i \sqrt{1 + \frac{n - \Gamma}{\Gamma^2}} \\
&= \sqrt{\frac{\Gamma^2 + n - \Gamma}{n}} \sum_{i=1}^{\Gamma} y_i.
\end{aligned}
$$

If we select $\Gamma = \theta\sqrt{n}$, which makes the probability of violation $\exp(-\theta^2/2)$, we obtain that

$$
\Omega||\boldsymbol{y}|| \leq \sqrt{1 + \theta^2} \sum_{i=1}^{\Gamma} y_i.
$$

Thus, in the worst case the protection level of our framework can only be smaller than the conic quadratic framework by a multiplicative factor of a constant. We conclude that in the worst case, the protection level for the conic quadratic framework can be smaller than our framework by a factor of $n^{1/4}$, while the protection of our framework can be smaller than the conic quadratic framework by at most a constant.

Let us compare the protection levels on average, however. In order to obtain some insight let us assume that y_i are independently and uniformly distributed in $[0, 1]$. Simple calculations show that for the case in question ($\Omega = \Gamma/\sqrt{n}$, $\Gamma = \theta\sqrt{n}$)

$$
E[\Omega||\boldsymbol{y}||] = \Theta(\sqrt{n}), \qquad E\left[\max_{S, |S| = \Gamma} \sum_{i \in S} y_i \right] = \Theta(\sqrt{n}),
$$

which implies that on average the two protection levels are of the same order of magnitude.

It is admittedly unclear whether it is the worst or the average case we presented which is more relevant, and thus the previous discussion is inconclusive. It is fair to say, however, that both approaches allow control of the degree of conservatism by adjusting the parameters Γ and Ω. Moreover, we think that the ultimate criterion for comparing the degree of conservatism of these methods will be computation in real problems.

4.2 Local Sensitivity Analysis of the Protection Level

Given the solution of Problem (7), it is desirable to estimate the change in the objective function value with respect to the change of the protection level Γ_i. In this way we can assess the price of increasing or decreasing the protection level Γ_i of any constraint. Note that when Γ_i changes, only one parameter in the coefficient matrix in Problem (7) changes. Thus, we can use results from sensitivity analysis (see Freund [7] for a comprehensive analysis) to understand the effect of changing the protection level under nondegeneracy assumptions.

Theorem 4 *Let $\boldsymbol{z^*}$ and $\boldsymbol{q^*}$ be the optimal nondegenerate primal and dual solutions for the linear optimization problem (7) (under nondegeneracy, the primal and dual optimal solutions are unique). Then, the derivative of the objective function value with respect to protection level Γ_i of the ith constraint is*

$$-z_i^* q_i^* \tag{22}$$

where z_i^* is the optimal primal variable corresponding to the protection level Γ_i and q_i^* is the optimal dual variable of the ith constraint.

Proof : We transform Problem (7) in standard form,

$$\begin{aligned} G(\Gamma_i) = \text{maximize} \quad & \boldsymbol{c'x} \\ \text{subject to} \quad & \boldsymbol{Ax} + \Gamma_i z_i \boldsymbol{e_i} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned}$$

where $\boldsymbol{e_i}$ is a unit vector with an one in the ith position. Let $\boldsymbol{B}$ be the optimal basis, which is unique under primal and dual nondegeneracy. If the column $\Gamma_i \boldsymbol{e_i}$ corresponding to the variable z_i is not in the basis, then $z_i^* = 0$. In this case, under dual nondegeneracy all reduced costs associated with the nonbasic variables are strictly negative, and thus a marginal change in the protection level does not affect the objective function value. Eq. (22) correctly indicates the zero variation.

If the column $\Gamma_i \boldsymbol{e_i}$ corresponding to the variable z_i is in the basis, and the protection level Γ_i changes by $\Delta\Gamma_i$, then $\boldsymbol{B}$ becomes $\boldsymbol{B} + \Delta\Gamma_i \boldsymbol{e_i e_i'}$. By the *Matrix Inversion Lemma* we have:

$$(\boldsymbol{B} + \Delta\Gamma_i \boldsymbol{e_i e_i'})^{-1} = \boldsymbol{B}^{-1} - \frac{\Delta\Gamma_i \boldsymbol{B}^{-1}\boldsymbol{e_i e_i'}\boldsymbol{B}^{-1}}{1 + \Delta\Gamma_i \boldsymbol{e_i'}\boldsymbol{B}^{-1}\boldsymbol{e_i}}.$$

Under primal and dual nondegeneracy, for small changes $\Delta\Gamma_i$, the new solutions preserve primal and dual feasibility. Therefore, the corresponding change in the objective function value is,

$$G(\Gamma_i + \Delta\Gamma_i) - G(\Gamma_i) = -\frac{\Delta\Gamma_i \boldsymbol{c_B'}\boldsymbol{B}^{-1}\boldsymbol{e_i e_i'}\boldsymbol{B}^{-1}\boldsymbol{b}}{1 + \Delta\Gamma_i \boldsymbol{e_i'}\boldsymbol{B}^{-1}\boldsymbol{e_i}} = -\frac{\Delta\Gamma_i z_i^* q_i^*}{1 + \Delta\Gamma_i \boldsymbol{e_i'}\boldsymbol{B}^{-1}\boldsymbol{e_i}},$$

where $\boldsymbol{c_B}$ is the part of the vector $\boldsymbol{c}$ corresponding to the columns in $\boldsymbol{B}$. Thus,

$$G'(\Gamma_i) = \lim_{\Delta\Gamma_i \to 0} \frac{G(\Gamma_i + \Delta\Gamma_i) - G(\Gamma_i)}{\Delta\Gamma_i} = -z_i^* q_i^*.$$

■

Remark : An attractive aspect of Eq. (22) is its simplicity as it only involves only the primal optimal solution corresponding to the protection level Γ_i and the dual optimal solution corresponding to the ith constraint.

5 Correlated Data

So far we assumed that the data are independently uncertain. It is possible, however, that the data are correlated. In particular, we envision that there are few sources of data uncertainty that affect all the data. More precisely, we assume that the model of data uncertainty is a follows.

Correlated Model of Data Uncertainty C:

Consider a particular row i of the matrix $\boldsymbol{A}$ and let J_i the set of coefficients in row i that are subject to uncertainty. Each entry $a_{ij},\ j \in J_i$ is modeled as

$$\tilde{a}_{ij} = a_{ij} + \sum_{k \in K_i} \tilde{\eta}_{ik} g_{kj}$$

and $\tilde{\eta}_{ik}$ are independent and symmetrically distributed random variables in $[-1, 1]$.

Note that under this model, there are only $|K_i|$ sources of data uncertainty that affect the data in row i. Note that these sources of uncertainty affect all the entries $a_{ij},\ j \in J_i$. For example if $|K_i| = 1$, then all data in a row are affected by a single random variable. For a concrete example, consider a portfolio construction problem, in which returns of various assets are predicted from a regression model. In this case, there are a few sources of uncertainty that affect globally all the assets classes.

Analogously to (4), we propose the following robust formulation:

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{c'x} \\ \text{subject to} \quad & \sum_j a_{ij}x_j + \max_{\{S_i \cup \{t_i\} | S_i \subseteq K_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in K_i \setminus S_i\}} \left\{ \sum_{k \in S_i} |\sum_{j \in J_i} g_{kj} x_j| + (\Gamma_i - \lfloor \Gamma_i \rfloor) |\sum_{j \in J_i} g_{t_i j} x_j| \right\} \leq b_i \quad \forall i \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u}, \end{aligned} \tag{23}$$

which can be written as a linear optimization problem as follows:

$$
\begin{array}{lll}
\text{maximize} & \boldsymbol{c}'\boldsymbol{x} & \\
\text{subject to} & \displaystyle\sum_j a_{ij}x_j + z_i\Gamma_i + \sum_{k\in K_i} p_{ik} \le b_i & \forall i \\
& z_i + p_{ik} \ge y_{ik} & \forall i,\ k \in K_i \\
& -y_{ik} \le \sum_{j\in J_i} g_{kj}x_j \le y_{ik} & \forall i,\ k \in K_i \\
& l_j \le x_j \le u_j & \forall j \\
& p_{ik}, y_{ik} \ge 0 & \forall i,\ k \in K_i \\
& z_i \ge 0 & \forall i.
\end{array}
\tag{24}
$$

Analogously to Theorem 3, we can show that the probability that the ith constraint is violated is at most $B(|K_i|, \Gamma_i)$ defined in Eq. (16).

6 Experimental Results

In this section, we present three experiments illustrating our robust solution to problems with data uncertainty. In the first experiment, we show that our methods extend to discrete optimization problems with data uncertainty, by solving a knapsack problem with uncertain weights. The second example is a simple portfolio optimization problem from Ben-Tal and Nemirovski [2], which has data uncertainty in the objective function. In the last experiment we apply our method to a problem PILOT4 from the well known Net Lib collection to examine the effectiveness of our approach to real world problems.

6.1 The Robust Knapsack Problem

The proposed robust formulation in Theorem 1 in the case in which the nominal problem is a mixed integer programming (MIP) model, i.e., some of the variables in the vector $\boldsymbol{x}$ take integer values, is still a MIP formulation, and thus can be solved in the same way that the nominal problem can be solved. Moreover, both the deterministic guarantee as well as the probabilistic guarantee (Theorem 3) that our approach provides is still valid. As a result, our approach applies for addressing data uncertainty for MIPs. To the best of our knowledge, there is no prior research in robust discrete optimization that is both tractable computationally and involves a probabilistic guarantee for constraint violation.

We apply in this section our approach to zero-one knapsack problems that are subject to data uncertainty. Our objective in this section is to examine whether our approach is computationally tractable, and whether it succeeds in reducing the price of robustness.

The zero-one knapsack problem is the following discrete optimization problem:

$$\begin{aligned} \text{maximize} \quad & \sum_{i \in N} c_i x_i \\ \text{subject to} \quad & \sum_{i \in N} w_i x_i \leq b \\ & x_i \in \{0, 1\}. \end{aligned}$$

Although the knapsack problem is *NP-hard*, for problems of moderate size, it is often solved to optimality using state-of-the-art MIP solvers. For this experiment, we use CPLEX 6.0 to solve to optimality a random knapsack problem of size, $|N| = 200$.

Regarding the uncertainty model for data, we assume the weights $\tilde{w}_i$ are uncertain, independently distributed and follow symmetric distributions in $[w_i - \delta_i, w_i + \delta_i]$. An application of this problem is to maximize the total value of goods to be loaded on a cargo that has strict weight restrictions. The weight of the individual item is assumed to be uncertain, independent of other weights and follows a symmetric distribution. In our robust model, we want to maximize the total value of the goods but allowing a maximum of 1% chance of constraint violation. The robust model of Theorem 1 is as follows:

$$\begin{aligned} \text{maximize} \quad & \sum_{i \in N} c_i x_i \\ \text{subject to} \quad & \sum_{i \in N} w_i x_i + \zeta(x, \Gamma) \leq b \\ & x_i \in \{0, 1\}. \end{aligned}$$

where

$$\zeta(x, \Gamma) = \max_{\{S \cup \{t\} |\ S \subseteq N, |S| = \lfloor \Gamma \rfloor, t \in N \setminus S\}} \left\{ \sum_{j \in S} \delta_j x_j + (\Gamma - \lfloor \Gamma \rfloor) \delta_t x_t \right\}.$$

For the random knapsack example, we set the capacity limit, b to 4000, the nominal weight, w_i being randomly chosen from the set $\{20, 21, \ldots, 29\}$ and the cost c_i randomly chosen from the set $\{16, 17, \ldots, 77\}$. We set the weight uncertainty δ_i to equal 10% of the nominal weight. The time to solve the robust discrete problems to optimality using CPLEX 6.0 on a Pentium II 400 PC ranges from 0.05 to 50 seconds.

Figure 4 illustrates the effect of the protection level on the objective function value. In the absence of protection to the capacity constraint, the optimal value is $5,592$. However, with maximum protection, that is admitting the Soyster's method, the optimal value is reduced by 5.5% to $5,283$. In Figure 5, we plot the optimal value with respect to the approximate probability bound of constraint violation. In Table 2, we present a sample of the objective function value and the probability bound of constraint violation.

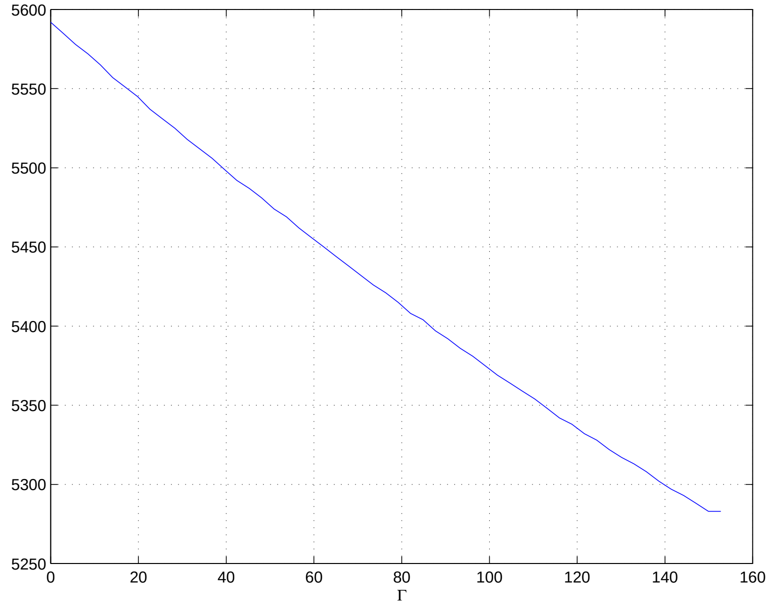

Figure 4: Optimal value of the robust knapsack formulation as a function of Γ.

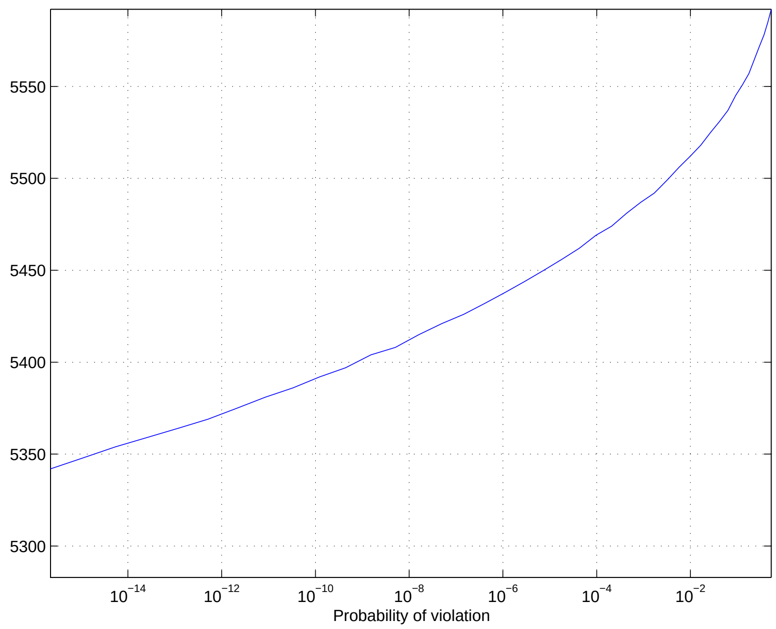

Figure 5: Optimal value of the robust knapsack formulation as a function of the probability bound of constraint violation given in Eq. (18).

It is interesting to note that the optimal value is marginally affected when we increase the protection level. For instance, to have a probability guarantee of at most 0.57% chance of constraint violation, we only reduce the objective by 1.54%. We can summarize the key insights in this example:

Γ	Probability bound	Optimal Value	Reduction
2.8	4.49×10^{-1}	5585	0.13%
14.1	1.76×10^{-1}	5557	0.63%
25.5	4.19×10^{-2}	5531	1.09%
36.8	5.71×10^{-3}	5506	1.54%
48.1	4.35×10^{-4}	5481	1.98%
59.4	1.82×10^{-5}	5456	2.43%
70.7	4.13×10^{-7}	5432	2.86%
82.0	5.04×10^{-9}	5408	3.29%
93.3	3.30×10^{-11}	5386	3.68%
104.7	1.16×10^{-13}	5364	4.08%
116.0	2.22×10^{-16}	5342	4.47%

Table 2: Results of Robust Knapsack Solutions.

1. Our approach succeeds in reducing the price of robustness, that is we do not heavily penalize the objective function value in order to protect ourselves against constraint violation.

2. The proposed robust approach is computationally tractable in that the problem can be solved in reasonable computational times.

6.2 A Simple Portfolio Problem

In this section we consider a portfolio construction problem consisting of a set of N stocks ($|N| = n$). Stock i has return $\tilde{p}_i$ which is of course uncertain. The objective is to determine the fraction x_i of wealth invested in stock i, so as to maximize the portfolio value $\sum_{i=1}^{n} \tilde{p}_i x_i$. We model the uncertain return $\tilde{p}_i$ as a random variable that has an arbitrary symmetric distribution in the interval $[p_i - \sigma_i, p_i + \sigma_i]$, where p_i is the expected return and σ_i is a measure of the uncertainty of the return of stock i. We further assume that the returns $\tilde{p}_i$ are independent.

The classical approach in portfolio construction is to use quadratic optimization and solve the following problem:

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} p_i x_i - \phi \sum_{i=1}^{n} \sigma_i^2 x_i^2 \\ \text{subject to} \quad & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0, \end{aligned}$$

where we interpret σ_i as the standard deviation of the return for stock i, and ϕ is a parameter that controls the tradeoff between risk and return. Applying our approach, we will solve instead the following problem (which can be reformulated as a linear optimization problem as in Theorem 7):

$$\begin{aligned} \text{maximize} \quad & z \\ \text{subject to} \quad & z \leq \sum_{i=1}^{n} p_i x_i - \beta(\boldsymbol{x}, \Gamma) \\ & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0. \end{aligned} \tag{25}$$

where

$$\beta(\boldsymbol{x}, \Gamma) = \max_{\{S \cup \{t\} \mid \; S \subseteq N, |S| = \lfloor \Gamma \rfloor, t \in N \setminus S\}} \left\{ \sum_{j \in S} \sigma_j x_j + (\Gamma - \lfloor \Gamma \rfloor) \sigma_t x_t \right\}$$

In this setting Γ is the protection level of the actual portfolio return in the following sense. Let $\boldsymbol{x}^*$ be an optimal solution of Problem (25) and let z^* be the optimal solution value of Problem (25). Then, $\boldsymbol{x}^*$ satisfies that $\Pr(\tilde{\boldsymbol{p}}' \boldsymbol{x}^* < z^*)$ is less than or equal to than the bound in Eq. (18). Ben-Tal and Nemirovski [2] consider the same portfolio problem using $n = 150$,

$$p_i = 1.15 + i \frac{0.05}{150}, \qquad \sigma_i = \frac{0.05}{450} \sqrt{2in(n+1)}.$$

Note that in this experiment, stocks with higher returns are also more risky.

Optimization Results

Let $\boldsymbol{x}^*(\Gamma)$ be an optimal solution to Problem (25) corresponding to the protection level Γ. A classical measure of risk is the *Standard Deviation*,

$$w(\Gamma) = \sqrt{\sum_{i \in N} \sigma_i^2 (x_i^*(\Gamma))^2}.$$

We first solved Problem (25) for various levels of Γ. Figure 6 illustrates the performance of the robust solution as a function of the protection level Γ, while Figure 7 shows the solution itself for various levels of the protection level. The solution exhibits some interesting “phase transitions” as the protection level increases:

1. For $\Gamma \leq 17$, both the expected return as well as the risk adjusted return (the objective function value) gradually decrease. Starting with $\Gamma = 0$, for which the solution consists of the stock 150 that has the highest expected return, the portfolio becomes gradually more diversified putting more weight on stocks with higher ordinal numbers. This can be seen for example for $\Gamma = 10$ in Figure 7.

2. For $17 < \Gamma \leq 41$, the risk adjusted return continues to gradually decrease as the protection level increases, while the expected return is insensitive to the protection level. In this range, $x_i* = \frac{\sum_i (1/\sigma_i)}{\sigma_i}$, i.e., the portfolio is fully diversified.

3. For $\Gamma \geq 41$, there is a sudden phase transition (see Figure 6). The portfolio consists of only stock 1, which is the one that has the largest risk adjusted return $p_i - \sigma_i$. This is exactly the solution given by the Soyster method as well. In this range both the expected and the risk adjusted returns are insensitive to Γ.

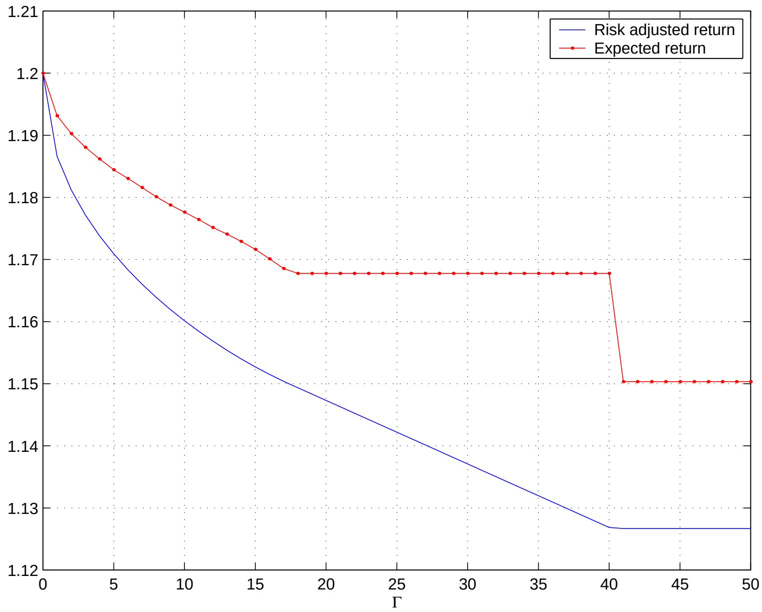

Figure 6: The return and the objective function value (risk adjusted return) as a function of the protection level Γ.

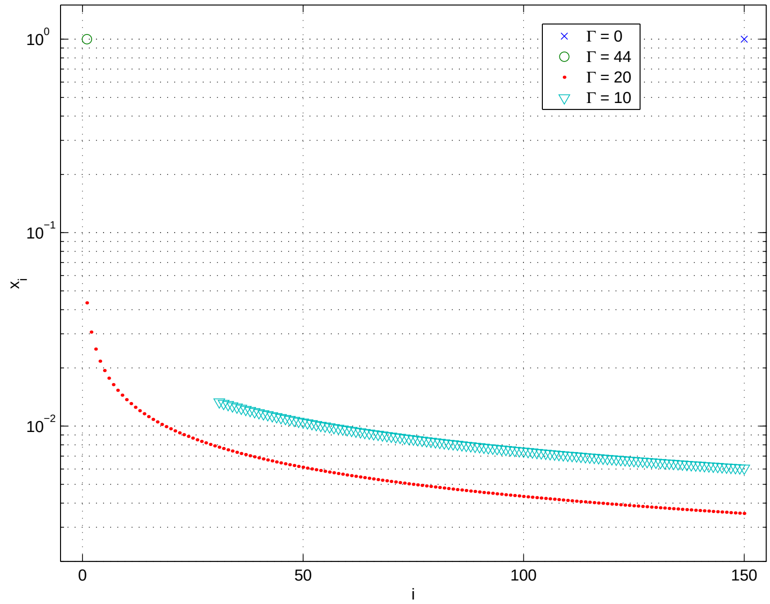

Figure 7: The solution of the portfolio for various protection levels.

Simulation Results

To examine the quality of the robust solution, we run 10,000 simulations of random yields and compare robust solutions generated by varying the protection level Γ. As we have discussed, for the worst case simulation, we consider the distribution with $\tilde{p}_i$ taking with probability $1/2$ the values at $p_i \pm \sigma_i$. In Figure 8, we compare the theoretical bound in Eq. (18) with the fraction of the simulated portfolio returns falling below the optimal solution, z^*. The empirical results suggest that the theoretical bound is close to the empirically observed values.

In Table 3, we present the results of the simulation indicating the tradeoff between risk and return. The corresponding plots are also presented in Figures 9 and 10. As expected as the protection level increases, the expected and maximum returns decrease, while the minimum returns increase. For instance, with $\Gamma \geq 15$, the minimum return is maintained above 12% for all simulated portfolios.

This example suggests that our approach captures the tradeoff between risk and return, very much like the mean variance approach, but does so in a linear framework. Additionally the robust approach provides both a deterministic guarantee about the return of the portfolio, as well as a probabilistic guarantee that is valid for all symmetric distributions.

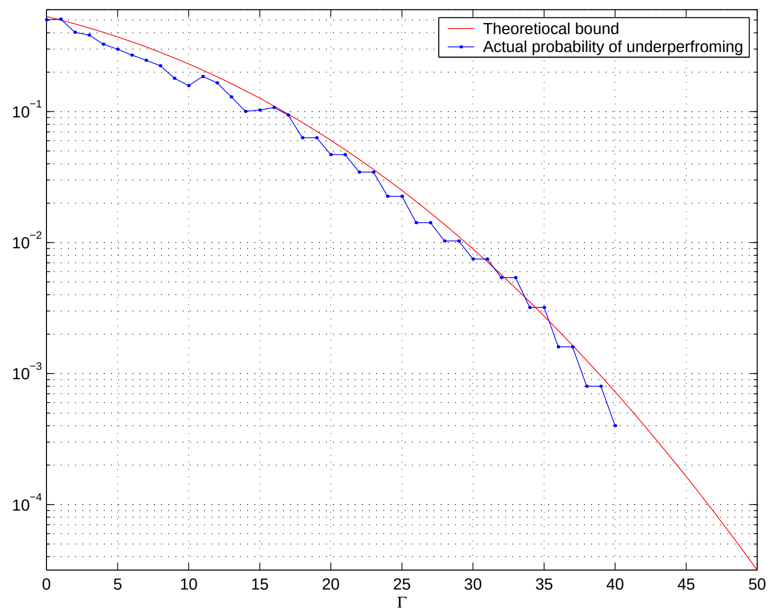

Figure 8: Simulation study of the probability of underperforming the nominal return as a function of Γ.

Γ	Prob. violation	Exp. Return	Min. Return	Max. Return	$w(\Gamma)$
0	0.5325	1.200	0.911	1.489	0.289
5	0.3720	1.184	1.093	1.287	0.025
10	0.2312	1.178	1.108	1.262	0.019
15	0.1265	1.172	1.121	1.238	0.015
20	0.0604	1.168	1.125	1.223	0.013
25	0.0250	1.168	1.125	1.223	0.013
30	0.0089	1.168	1.125	1.223	0.013
35	0.0028	1.168	1.125	1.223	0.013
40	0.0007	1.168	1.125	1.223	0.013
45	0.0002	1.150	1.127	1.174	0.024

Table 3: Simulation results given by the robust solution.

6.3 Robust Solutions of a Real-World Linear Optimization Problem

As noted in [1], optimal solutions of linear optimization problems may become severely infeasible if the nominal data are slightly perturbed. In this experiment, we applied our method to the problem PILOT4 from the Net Lib library of problems. Problem PILOT4 is a linear optimization problem

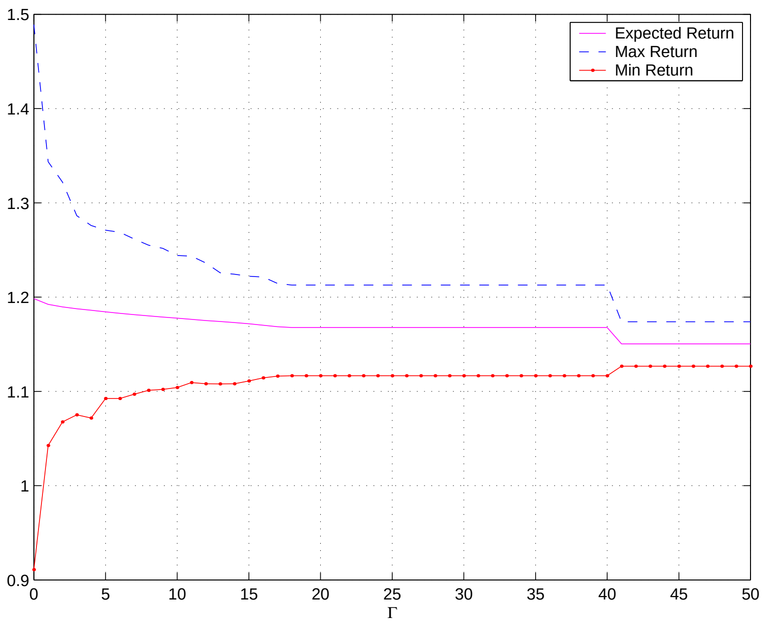

Figure 9: Empirical Result of expected, maximum and minimum yield.

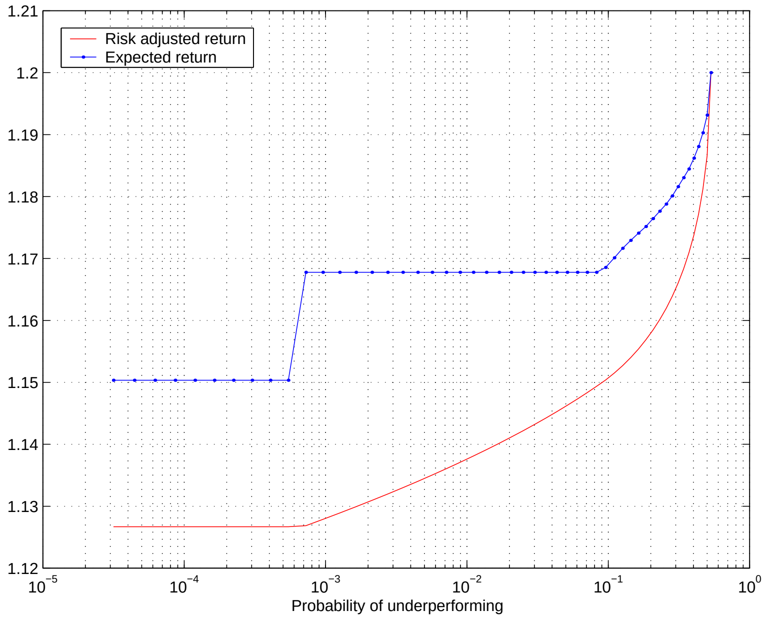

Figure 10: Tradeoffs between probability of underperforming and returns.

with 411 rows, 1000 columns, 5145 nonzero elements and optimum objective value, −2581.1392613. It contains coefficients such as 717.562256, -1.078783, -3.053161, -.549569, -22.634094, -39.874283, which seem unnecessarily precise. In our study, we assume that the coefficients of this type that participate

$\lvert J_i \rvert$	# constraints	$\lvert J_i \rvert$	# constraints
1	21	24	12
11	4	25	4
14	4	26	8
15	4	27	8
16	4	28	4
17	8	29	8
22	4		

Table 4: Distributions of $|J_i|$ in PILOT4.

in the inequalities of the formulation have a maximum 2% deviation from the corresponding nominal values. Table 4 presents the distributions of the number of uncertain data in the problem. We highlight that each of the constraints has at most 29 uncertain data.

We solve the robust problem (7) and report the results in Table 5. In Figure 11, we present the efficient frontier of the probability of constraint violation and cost.

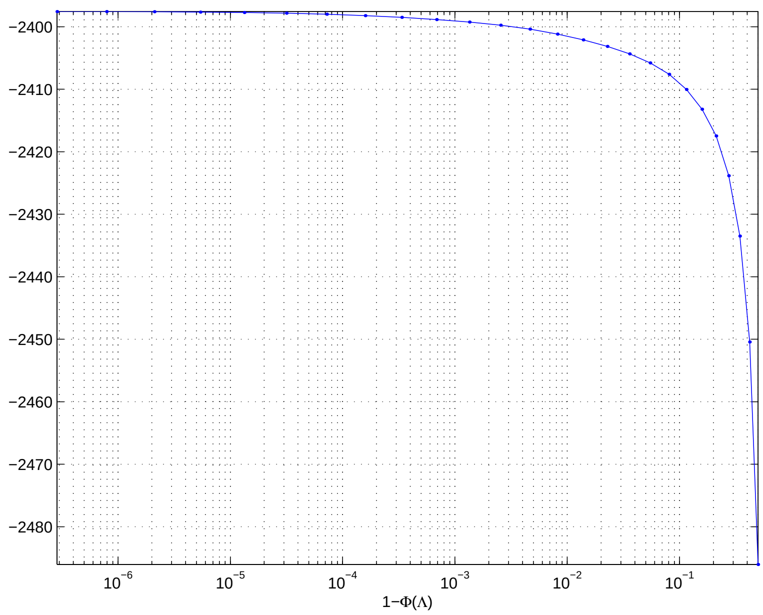

Figure 11: The tradeoff between cost and robustness.

We note that the cost of full protection (Soyster's method) is equal to -2397.5798763. In this example, we observe that relaxing the need of full protection, still leads to a high increase in the cost

Optimal Value	% Change	Prob. violation
-2486.0334864	3.68%	0.5
-2450.4323768	5.06%	0.421
-2433.4959374	5.72%	0.345
-2413.2013332	6.51%	0.159
-2403.1494574	6.90%	0.0228
-2399.2592198	7.05%	0.00135
-2397.8404967	7.10%	3.17×10^{-5}
-2397.5798996	7.11%	2.87×10^{-7}
-2397.5798763	7.11%	9.96×10^{-8}

Table 5: The tradeoff between optimal cost and robustness.

unless one is willing to accept unrealistically high probabilities for constraint violation. We attribute this to the fact that there are very few uncertain coefficients in each constraint (Table 4), and thus probabilistic protection is quite close to deterministic protection.

7 Conclusions

The major insights from our analysis are:

1. Our proposed robust methodology provides solutions that ensure deterministic and probabilistic guarantees that constraints will be satisfied as data change.

2. Under the proposed method, the protection level determines probability bounds of constraint violation, which do not depend on the solution of the robust model.

3. The method naturally applies to discrete optimization problems.

4. Comparing Figures 5 for a problem with $n = 200$ uncertain coefficients and 11 for a problem in which the maximum number of uncertain coefficients per row is 29, we observe that in order to have probability of violation of the constraints 10^{-4} we have an objective function change of 2% in the former case and 7% in the latter case. We feel that this is indicative of the fact that the attractiveness of the method increases as the number of uncertain data increases.

8 Appendix: Proof of Theorem 3

In this appendix we present the proof of Theorem 3.

(a) The proof follows from Proposition 2 parts (a) and (b). To simplify the exposition we will drop the subscript i, which represents the index of the constraint. We prove the bound in (16) by induction on n. We define the auxiliary quantities:

$$\nu(\Gamma, n) = \frac{\Gamma + n}{2}, \qquad \mu(\Gamma, n) = \nu(\Gamma, n) - \lfloor \nu(\Gamma, n) \rfloor, \qquad \Upsilon(s, n) = \frac{1}{2^n} \sum_{l=s}^{n} \binom{n}{l}.$$

The induction hypothesis is formulated as follows:

$$\Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma\right) \leq \begin{cases} (1 - \mu(\Gamma, n))\Upsilon(\lfloor \nu(\Gamma, n) \rfloor, n) + \mu(\Gamma, n)\Upsilon(\lfloor \nu(\Gamma, n) \rfloor + 1, n) & \text{if } \Gamma \in [1, n] \\ 0 & \text{if } \Gamma > n. \end{cases}$$

For $n = 1$, then $\Gamma = 1$, and so $\nu(1,1) = 1, \mu(1,1) = 0$, $\Upsilon(1,1) = 1/2$ leading to:

$$\begin{aligned} \Pr(\eta_1 \geq \Gamma) &\leq \Pr(\eta_1 \geq 0) \\ &\leq \frac{1}{2} \\ &= (1 - \mu(1,1))\Upsilon(\lfloor \nu(1,1) \rfloor, 1) + \mu(1,1)\Upsilon(\lfloor \nu(1,1) \rfloor + 1, 1). \end{aligned}$$

Assuming the induction hypothesis holds for n, we have

$$\begin{aligned} \Pr\left(\sum_{j=1}^{n+1} \gamma_j \eta_j \geq \Gamma\right) &= \int_{-1}^{1} \Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma - \gamma_{n+1}\eta_{n+1} | \eta_{n+1} = \eta\right) dF_{\eta_{n+1}}(\eta) \\ &= \int_{-1}^{1} \Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma - \gamma_{n+1}\eta\right) dF_{\eta_{n+1}}(\eta) \qquad (26) \\ &= \int_{0}^{1} \left[\Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma - \gamma_{n+1}\eta\right) + \Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma + \gamma_{n+1}\eta\right)\right] dF_{\eta_{n+1}}(\eta) \qquad (27) \\ &\leq \max_{\phi \in [0, \gamma_{n+1}]} \left\{\Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma - \phi\right) + \Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma + \phi\right)\right\} \int_{0}^{1} dF_{\eta_{n+1}}(\eta) \\ &= \frac{1}{2} \max_{\phi \in [0, \gamma_{n+1}]} \left\{\Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma - \phi\right) + \Pr\left(\sum_{j=1}^{n} \gamma_j \eta_j \geq \Gamma + \phi\right)\right\} \\ &\leq \frac{1}{2} \max_{\phi \in [0, \gamma_{n+1}]} \Psi_n(\phi) \qquad (28) \\ &\leq \frac{1}{2}\Psi_n(1) \qquad (29) \\ &= (1 - \mu(\Gamma, n+1))\Upsilon(\lfloor \nu(\Gamma, n+1) \rfloor, n+1) + \\ &\quad \mu(\Gamma, n+1)\Upsilon(\lfloor \nu(\Gamma, n+1) \rfloor + 1, n+1), \qquad (30) \end{aligned}$$

where

$$\Psi_n(\phi) = (1-\mu(\Gamma-\phi,n)\Upsilon(\lfloor\nu(\Gamma-\phi,n)\rfloor,n) + \mu(\Gamma-\phi,n)\Upsilon(\lfloor\nu(\Gamma-\phi,n)\rfloor+1,n)$$
$$+(1-\mu(\Gamma+\phi,n))\Upsilon(\lfloor\nu(\Gamma+\phi,n)\rfloor,n) + \mu(\Gamma+\phi,n)\Upsilon(\lfloor\nu(\Gamma+\phi,n)\rfloor+1,n).$$

Eqs. (26) and (27) follow from the assumption that η_j's are independent, symmetrically distributed random variables in $[-1,1]$. Inequality (28) represents the induction hypothesis. Eq. (30) follows from:

$$\begin{aligned}
\Psi_n(1) &= (1-\mu(\Gamma-1,n))\Upsilon(\lfloor\nu(\Gamma-1,n)\rfloor,n) + \mu(\Gamma-1,n)\Upsilon(\lfloor\nu(\Gamma-1,n)\rfloor+1,n) \\
&\quad +(1-\mu(\Gamma+1,n))\Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor,n) + \mu(\Gamma+1,n)\Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor+1,n) \qquad (31)\\
&= (1-\mu(\Gamma+1,n))(\Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor-1,n) + \Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor,n)) + \\
&\quad \mu(\Gamma+1,n)(\Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor,n) + \Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor+1,n)) \qquad (32)\\
&= 2\{(1-\mu(\Gamma+1,n))\Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor,n+1) + \mu(\Gamma+1,n)\Upsilon(\lfloor\nu(\Gamma+1,n)\rfloor+1,n+1)\} \qquad (33)\\
&= 2\{(1-\mu(\Gamma,n+1))\Upsilon(\lfloor\nu(\Gamma,n+1)\rfloor,n+1) + \mu(\Gamma,n+1)\Upsilon(\lfloor\nu(\Gamma,n+1)\rfloor+1,n+1)\} \qquad (34)
\end{aligned}$$

Eqs. (31) and (32) follow from noting that $\mu(\Gamma-1,n) = \mu(\Gamma+1,n)$ and $\lfloor\nu(\Gamma-1,n)\rfloor = \lfloor\nu(\Gamma+1,n)\rfloor-1$. Eq. (33) follows from the claim that $\Upsilon(s,n)+\Upsilon(s+1,n) = 2\Upsilon(s+1,n+1)$, which is presented next:

$$\begin{aligned}
\Upsilon(s,n)+\Upsilon(s+1,n) &= \frac{1}{2^n}\left\{\sum_{l=s}^{n}\binom{n}{l} + \sum_{l=s+1}^{n}\binom{n}{l}\right\} \\
&= \frac{1}{2^n}\left(\sum_{l=s}^{n-1}\left[\binom{n}{l}+\binom{n}{l+1}\right]+1\right) \\
&= \frac{1}{2^n}\left(\sum_{l=s}^{n-1}\binom{n+1}{l+1}+1\right) \\
&= \frac{1}{2^n}\sum_{l=s+1}^{n+1}\binom{n+1}{l} \\
&= 2\Upsilon(s+1,n+1),
\end{aligned}$$

and Eq. (34) follows from $\mu(\Gamma+1,n) = \mu(\Gamma,n+1) = (\Gamma+n+1)/2$.

We are left to show that $\Psi_n(\phi)$ is a monotonically non-decreasing function in the domain $\phi \in [0,1]$, which implies that for any $\phi_1,\phi_2 \in [0,1]$ such that $\phi_1 > \phi_2$, $\Psi_n(\phi_1)-\Psi_n(\phi_2) \geq 0$. We fix Γ and n. To simplify the notation we use: $\mu(\phi) = \mu(\Gamma+\phi,n) = (\Gamma+\phi+n)/2$, $\nu(\phi) = \nu(\Gamma+\phi,n)$. For any choice of ϕ_1 and ϕ_2, we have $\rho = \lfloor\nu_{-\phi_1}\rfloor \leq \lfloor\nu_{-\phi_2}\rfloor \leq \lfloor\nu_{\phi_2}\rfloor \leq \lfloor\nu_{\phi_1}\rfloor \leq \rho+1$. Therefore, we consider the following cases:

For $\rho = \lfloor\nu_{-\phi_1}\rfloor = \lfloor\nu_{-\phi_2}\rfloor = \lfloor\nu_{\phi_2}\rfloor = \lfloor\nu_{\phi_1}\rfloor$,

$$\mu_{-\phi_1}-\mu_{-\phi_2} = -\frac{\phi_1-\phi_2}{2}$$

$$\begin{aligned}
\mu_{\phi_1} - \mu_{\phi_2} &= \frac{\phi_1 - \phi_2}{2} \\
\Psi_n(\phi_1) - \Psi_n(\phi_2) &= \frac{\phi_1 - \phi_2}{2}\{\Upsilon(\rho, n) - \Upsilon(\rho+1, n) - \Upsilon(\rho, n) + \Upsilon(\rho+1, n)\} \\
&= 0.
\end{aligned}$$

For $\rho = \lfloor \nu_{-\phi_1} \rfloor = \lfloor \nu_{-\phi_2} \rfloor = \lfloor \nu_{\phi_2} \rfloor - 1 = \lfloor \nu_{\phi_1} \rfloor - 1$,

$$\begin{aligned}
\mu_{-\phi_1} - \mu_{-\phi_2} &= -\frac{\phi_1 - \phi_2}{2} \\
\mu_{\phi_1} - \mu_{\phi_2} &= \frac{\phi_1 - \phi_2}{2} \\
\Psi_n(\phi_1) - \Psi_n(\phi_2) &= \frac{\phi_1 - \phi_2}{2}\{\Upsilon(\rho, n) - 2\Upsilon(\rho+1, n) + \Upsilon(\rho+2, n)\} \\
&= \frac{\phi_1 - \phi_2}{2^{n+1}(n+1)}\binom{n+1}{\rho+1}(1 + 2\rho - n) \\
&\geq \frac{\phi_1 - \phi_2}{2^{n+1}(n+1)}\binom{n+1}{\rho+1}\left(1 + 2\left\lfloor \frac{1 + n - \phi_1}{2} \right\rfloor - n\right) \\
&\geq 0.
\end{aligned}$$

For $\rho = \lfloor \nu_{-\phi_1} \rfloor = \lfloor \nu_{-\phi_2} \rfloor - 1 = \lfloor \nu_{\phi_2} \rfloor - 1 = \lfloor \nu_{\phi_1} \rfloor - 1$,

$$\begin{aligned}
\mu_{-\phi_1} - \mu_{-\phi_2} &= -\frac{\phi_1 - \phi_2}{2} + 1 \\
\mu_{\phi_1} - \mu_{\phi_2} &= \frac{\phi_1 - \phi_2}{2} \\
\Psi_n(\phi_1) - \Psi_n(\phi_2) &= (1 - \mu_{-\phi_1})\{\Upsilon(\rho, n) - 2\Upsilon(\rho+1, n) + \Upsilon(\rho+2, n)\} \\
&\geq 0.
\end{aligned}$$

For $\rho = \lfloor \nu_{-\phi_1} \rfloor = \lfloor \nu_{-\phi_2} \rfloor = \lfloor \nu_{\phi_2} \rfloor = \lfloor \nu_{\phi_1} \rfloor - 1$,

$$\begin{aligned}
\mu_{-\phi_1} - \mu_{-\phi_2} &= -\frac{\phi_1 - \phi_2}{2} \\
\mu_{\phi_1} - \mu_{\phi_2} &= \frac{\phi_1 - \phi_2}{2} - 1 \\
\Psi_n(\phi_1) - \Psi_n(\phi_2) &= \mu_{\phi_1}\{\Upsilon(\rho, n) - 2\Upsilon(\rho+1, n) + \Upsilon(\rho+2, n)\} \\
&\geq 0.
\end{aligned}$$

(b) Let η_j obey a discrete probability distribution $\Pr(\eta_j = 1) = 1/2$ and $\Pr(\eta_j = -1) = 1/2$, $\gamma_j = 1$, $\Gamma \geq 1$ is integral and $\Gamma + n$ is even. Let S_n obeys a Binomial distribution with parameters n and $1/2$. Then,

$$\Pr\left(\sum_{j=1}^{n} \eta_j \geq \Gamma\right) = \Pr(S_n - (n - S_n) \geq \Gamma)$$

$$
\begin{aligned}
&= \Pr(2S_n - n \geq \Gamma) \\
&= \Pr(S_n \geq \frac{n+\Gamma}{2}) \\
&= \frac{1}{2^n} \sum_{l=\frac{n+\Gamma}{2}}^{n} \binom{n}{l},
\end{aligned} \tag{35}
$$

which implies that the bound (16) is indeed tight.

(c) From Eq. (17), we need to find an upper bound for the function $\frac{1}{2^n}\binom{n}{l}$. From Robbins [8] we obtain for $n \geq 1$,

$$
\sqrt{2\pi} n^{n+1/2} \exp(-n + 1/(12n+1)) \leq n! \leq \sqrt{2\pi} n^{n+1/2} \exp(-n + 1/(12n)),
$$

we can establish for $l \in \{1, \ldots, n-1\}$,

$$
\begin{aligned}
\frac{1}{2^n}\binom{n}{l} &= \frac{n!}{2^n (n-l)! l!} \\
&\leq \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{(n-l)l}} \exp\left(\frac{1}{12n} - \frac{1}{12(n-l)+1} - \frac{1}{12l+1}\right) \left(\frac{n}{2(n-l)}\right)^n \left(\frac{n-l}{l}\right)^l \\
&\leq \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{(n-l)l}} \left(\frac{n}{2(n-l)}\right)^n \left(\frac{n-l}{l}\right)^l \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{(n-l)l}} \exp\left(n \log\left(\frac{n}{2(n-l)}\right) + l \log\left(\frac{n-l}{l}\right)\right),
\end{aligned} \tag{36}
$$

where Eq. (36) follows from

$$
\frac{1}{12(n-l)+1} + \frac{1}{12l+1} \geq \frac{2}{(12(n-l)+1) + (12l+1)} = \frac{2}{12n+2} > \frac{1}{12n}.
$$

For $l = 0$ and $l = n$ $\frac{1}{2^n}\binom{n}{l} = \frac{1}{2^n}$.

(d) Bound (16) can be written as

$$
B(n, \Gamma_i) = (1-\mu)\Pr(S_n \geq \lfloor \nu \rfloor) + \mu \Pr(S_n \geq \lfloor \nu \rfloor + 1),
$$

where S_n represents a Binomial distribution with parameters n and $1/2$. Since $\Pr(S_n \geq \lfloor \nu \rfloor + 1) \leq \Pr(S_n \geq \lfloor \nu \rfloor)$, we have

$$
\Pr(S_n \geq \nu + 1) = \Pr(S_n \geq \lfloor \nu \rfloor + 1) \leq B(n, \Gamma_i) \leq \Pr(S_n \geq \lfloor \nu \rfloor) = \Pr(S_n \geq \nu),
$$

since S_n is a discrete distribution. For $\Gamma_i = \theta\sqrt{n}$, where θ is a constant, we have

$$
\Pr\left(\frac{S_n - n/2}{\sqrt{n}/2} \geq \theta + \frac{2}{\sqrt{n}}\right) \leq B(n, \Gamma_i) \leq \Pr\left(\frac{S_n - n/2}{\sqrt{n}/2} \geq \theta\right).
$$

By the central limit theorem, we obtain that

$$\lim_{n\to\infty} \Pr\left(\frac{S_n - n/2}{\sqrt{n}/2} \geq \theta + \frac{2}{\sqrt{n}}\right) = \lim_{n\to\infty} \Pr\left(\frac{S_n - n/2}{\sqrt{n}/2} \geq \theta\right) = 1 - \Phi(\theta),$$

where $\Phi(\theta)$ is the cumulative distribution function of a standard normal. Thus, for $\Gamma_i = \theta\sqrt{n}$, we have

$$\lim_{n\to\infty} B(n, \Gamma_i) = 1 - \Phi(\theta).$$

■

Acknowledgements

We would like to thank the reviewers of the paper for several insightful comments.

References

[1] Ben-Tal, A., Nemirovski, A. (2000): Robust solutions of Linear Programming problems contaminated with uncertain data, *Math. Program.*, 88, 411-424.

[2] Ben-Tal, A., Nemirovski, A. (1999): Robust solutions to uncertain programs, *Oper. Res. Letters*, 25, 1-13.

[3] Ben-Tal, A., Nemirovski, A. (1998): Robust convex optimization, *Math. Oper. Res.*, 23, 769-805.

[4] Ben-Tal, A., El-Ghaoui, L., Nemirovski, A. (2000): Robust semidefinite programming, in Saigal, R., Vandenberghe, L., Wolkowicz, H., eds., *Semidefinite programming and applications*, Kluwer Academic Publishers.

[5] El-Ghaoui, Lebret, H. (1997): Robust solutions to least-square problems to uncertain data matrices, *SIAM J. Matrix Anal. Appl.*, 18, 1035-1064.

[6] El-Ghaoui, L., Oustry, F., Lebret, H. (1998): Robust solutions to uncertain semidefinite programs, *SIAM J. Optim.*, 9, 33-52.

[7] Freund, R. M. (1985): Postoptimal Analysis of a Linear Program Under Simultaneous Changes in Matrix Coefficients, *Math. Prog. Study*, 24, 1-13.

[8] Robbins, H. (1955): A Remark of Stirling's Formula. *Amer. Math. Monthly*, 62, 26-29.

[9] Soyster, A.L. (1973): Convex programming with set-inclusive constraints and applications to inexact linear programming, *Oper. Res.*, 21, 1154-1157.