Robust Linear Optimization With Recourse

Aurélie Thiele
Lehigh University

Report No. 05T-007

Robust Linear Optimization With Recourse

Aurélie Thiele*

May 2005

Abstract

We propose an approach to linear optimization with recourse that does not involve a probabilistic description of the uncertainty, and allows the decision-maker to adjust the degree of robustness of the model while preserving its linear properties. We model random variables as uncertain parameters belonging to a polyhedral uncertainty set and minimize the sum of the first-stage costs and the worst-case second-stage costs over that set. The decision-maker's conservatism is taken into account through a budget of uncertainty, which determines the size of the uncertainty set around the mean of the random variables. We establish that the robust problem is a linear programming problem with a potentially very large number of constraints, and describe how the classical Benders decomposition algorithm can be adapted to the robust approach. Furthermore, in the case of simple recourse, we show that the robust problem can be formulated as a series of m linear programming problems of size similar to the original deterministic problem, where m is the number of random variables. Numerical results are encouraging.

1 Introduction

Linear optimization with recourse was first introduced by Dantzig in [14] as a mathematical framework for sequential decision-making under uncertainty. In that setting, we must make some decisions before discovering the actual value taken by the random variables but have the opportunity to take further action once uncertainty has been revealed, with the objective of minimizing total expected cost. This framework later became known as stochastic programming and is described in detail in the monographs by Birge and Louveaux [11] and Kall and Wallace [22]. However, as early as the mid-1960s, researchers such as Zácková [30] recognized the practical limitations of the expected-value paradigm, which requires the exact knowledge of the underlying probability distributions. The fact that such probabilities are very hard to estimate in practice

*Department of Industrial and Systems Engineering, Lehigh University, Harold S. Mohler Laboratory Room 329, Bethlehem, PA 18015, aurelie.thiele@lehigh.edu.

motivated the development of a minimax approach, where the decision-maker minimizes the maximal expected cost over a family of probability distributions. It has received significant attention in the stochastic programming literature, for instance from Dupacova [15, 16, 17], whose work along with Záčková's findings [30] laid the foundation for subsequent research efforts. Other early references include Jagannathan [21], who studied stochastic linear programming with simple recourse when the first two moments of the distributions are known, and Birge and Wets [12], who focused on bounding and approximating stochastic problems. More recently, Shapiro and Ahmed [25] and Shapiro and Kleywegt [26] have investigated further the theoretical properties of minimax stochastic optimization, while Takriti and Ahmed describe in [29] an application to electricity contracts. The main drawback of the stochastic minimax approach is that the solution methods proposed in the literature (a stochastic gradient technique in Ermoliev et. al. [20], a bundle method in Breton and El Hachem [13], a cutting plane algorithm in Riis and Andersen [23], to name a few), all require finding explicitly the worst-case probability for the current candidate solution at each step of the algorithm, and hence suffer from dimensionality problems. These are particularly acute here as stochastic programming often yields large-scale formulations. Although Shapiro and Ahmed [25] and Shapiro and Kleywegt [26] have studied specific classes of problems for which the minimax framework leads to traditional stochastic formulations, no such approach has been developed to date for the general case. Furthermore, while the field of stochastic programming has seen in recent years a number of algorithmic advances, e.g., sampling methods (Shapiro [24]), the problem with recourse remains significantly more difficult to solve than its deterministic linear counterpart, and does not allow for easy insights into the impact of randomness on the optimal decision variables.

Therefore, the need arises to develop an approach to linear optimization with recourse that does not involve a probabilistic description of the uncertainty, remains tractable in a wide range of settings, and yields theoretical insights into the way randomness affects the optimal solution. The purpose of this paper is to present such an approach, based on robust optimization. While robust optimization has been previously used in stochastic programming as a method to incorporate cost variability in the objective function (Takriti and Ahmed [28]), we consider here a different methodology, which was developed independently under the same name. What we refer to as robust optimization addresses data uncertainty in mathematical programming problems by finding the optimal solution for the worst-case instances of unknown but bounded parameters. This approach was pioneered in 1973 by Soyster [27], who proposed a model that guarantees feasibility for all instances of the parameters within a convex set. However, the resulting solution is very conservative, in the sense that it is too far from optimality in the nominal model to be of practical interest for real-life implementation. This issue of overconservatism hindered the adoption of robust techniques in optimization problems until the mid-1990s, when Ben-Tal and Nemirovski [3, 4, 5], El Ghaoui and Lebret [18] and El Ghaoui et. al. [19] started investigating models where feasibility of a linear programming problem is guaranteed with high probability. They focus on ellipsoidal uncertainty sets, which allow for important insights into the robust framework but increase the complexity of the problem considered, e.g., yield second-order cone problems as the robust counterpart of linear models. In contrast, Bertsimas and Sim [8] study polyhedral uncertainty sets, which do not change the class of the problem at hand, and explicitly quantify the trade-off between performance and conservatism in terms of probabilistic bounds of constraint violation. An advantage of their approach is that it can be easily extended to integer and mixed-integer programming problems (Bertsimas and Sim [9]). While robust optimization has been applied in the references above as a way to address parameter uncertainty, Bertsimas and Thiele [10] use this framework to model random variables and address uncertainty on the underlying distributions in a multi-period inventory problem. Their approach highlights the potential of robust optimization for dynamic decision-making in presence of randomness. A first step towards implementing robust techniques in stochastic programming with recourse was taken by Ben-Tal et. al. [6], who coined the term "adjustable decision variables" as a synonym for second-stage solutions. Unfortunately, the robust counterpart in their approach is computationally intractable, which leads them to restrict the second-stage variables to affine functions of the uncertain data. Atamtürk and Zhang propose in [1] a model for two-stage optimization that does not involve affinely adjustable decision variables, in the context of network design under uncertain demand. Similarly, we do not impose any limitations on the structure of the recourse. The model presented here is very broad in scope, in the sense that we develop a robust approach for generic two-stage stochastic problems with uncertainty on the right-hand side. We believe that the framework proposed in this paper offers a new perspective on linear programming with recourse that combines the decision-maker's degree of conservatism and the uncertainty on the probability distributions in a tractable manner.

Specifically, we make the following contributions:

1. We address right-hand-side uncertainty in linear programming problems with recourse by modelling random variables as uncertain parameters in a polyhedral uncertainty set. The level of conservatism of the optimal solution is flexibly adjusted by setting a parameter called the "budget of uncertainty" to an appropriate value.

2. We propose a solution technique based on Benders' decomposition that is computationally less demanding than its stochastic counterpart, and yields the worst-case realization of the uncertain parameters within the uncertainty set. In other words, the robust approach is equivalent to the deterministic problem solved for a specific value of the parameters.

3. We formulate the robust problem with simple recourse as a series of m linear programming problems similar in size to the model without uncertainty, where m is the number of uncertain random variables.

The structure of the paper is as follows. In Section 2, we define the model of uncertainty and present the main ideas underlying the robust approach. We focus on the case of simple recourse in Section 3 and extend our results to the general case in Section 4, with an emphasis on tractability, the relationship to the deterministic models and insights into the optimal policy. We present computational results in Section 5. Finally, Section 6 contains some concluding remarks.

2 Problem Overview

2.1 Optimization With Recourse

The focus of this paper is on two-stage linear optimization with right-hand side uncertainty, which was first described by Dantzig in [14]. The deterministic problem can be formulated as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \mathbf{d}'\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\mathbf{y} = \mathbf{b}, \\ & \mathbf{x}, \mathbf{y} \geq \mathbf{0}, \end{aligned} \tag{1}$$

with the following notations:

$\mathbf{x}$: the first-stage decision variables,
$\mathbf{y}$: the second-stage decision variables,
$\mathbf{c}$: the first-stage costs,
$\mathbf{d}$: the second-stage costs,
$\mathbf{A}$: the first-stage coefficient matrix,
$\mathbf{B}$: the second-stage coefficient matrix,
$\mathbf{b}$: the requirement vector.

For clarity in the exposition, we do not include constraints that are only on the first-stage variables. They can however be incorporated to the model without changing any of the structural results presented in Sections 3 and 4.

In many applications, the requirement vector is random and the decision-maker implements the first-stage variables without knowing the actual requirements ("here-and-now"), but chooses the second-stage variables only after the uncertainty has been revealed ("wait-and-see"). This has traditionally been modelled using stochastic programming techniques, i.e., by assuming that the requirements obey a known probability distribution and minimizing the expected cost of the problem. In mathematical terms, if we define the recourse function, once the first-stage decisions have been implemented and the realization of the uncertainty has become known, as:

$$\begin{aligned} \mathbf{Q}(\mathbf{x}, \mathbf{b}) = \quad \min \quad & \mathbf{d}'\mathbf{y} \\ \text{s.t.} \quad & \mathbf{B}\,\mathbf{y} = \mathbf{b} - \mathbf{A}\,\mathbf{x}, \\ & \mathbf{y} \geq \mathbf{0}, \end{aligned} \tag{2}$$

the stochastic counterpart of Problem (1) can be formulated as a nonlinear problem:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \mathbf{E}_{\mathbf{b}}[\mathbf{Q}(\mathbf{x}, \mathbf{b})] \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{3}$$

If the uncertainty is discrete, consisting of Ω possible requirement vectors each occurring with probability π_ω, $\omega = 1, \ldots, \Omega$, Problem (3) becomes a linear programming problem:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \sum_{\omega=1}^{\Omega} \pi_\omega \cdot \mathbf{d}'\mathbf{y}_\omega \\ \text{s.t.} \quad & \mathbf{A}\,\mathbf{x} + \mathbf{B}\,\mathbf{y}_\omega = \mathbf{b}_\omega, \quad \forall \omega, \\ & \mathbf{x},\, \mathbf{y}_\omega \geq \mathbf{0}, \qquad\qquad \forall \omega. \end{aligned} \tag{4}$$

However, a realistic description of the uncertainty generally requires a high number of scenarios. Therefore, the deterministic equivalent (4) is often a large-scale problem, which necessitates the use of special-structure algorithms such as decomposition methods or Monte-Carlo simulations (see Birge and Louveaux [11] and Kall and Wallace [22] for an introduction to these techniques). Problem (4) can thus be considerably harder to solve than Problem (1), although both are linear. The difficulty in estimating probability distributions accurately also hinders the practical implementation of these techniques.

2.2 The Robust Approach

In contrast with the stochastic programming framework, robust optimization models random variables using uncertainty sets rather than probability distributions. The objective is then to minimize the worst-case cost in that set. Specifically, let $\mathcal{B}$ be the uncertainty set of the requirement vector, of known mean $\overline{\mathbf{b}}$. The robust problem with recourse is formulated as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \max_{\mathbf{b} \in \mathcal{B}} \mathbf{Q}(\mathbf{x}, \mathbf{b}) \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{5}$$

If $\mathcal{B} = \{\overline{\mathbf{b}}\}$, Problem (5) is the nominal problem. As $\mathcal{B}$ expands around $\overline{\mathbf{b}}$, the decision-maker protects the system against more realizations of the random variables and the solution becomes more robust, but also more conservative. If the decision-maker does not take uncertainty into account, he might incur very large costs once the uncertainty has been revealed. On the other hand, if he includes every possible outcome in his model, he will protect the system against realizations that would indeed be detrimental to his profit, but are also very unlikely to happen. The question of choosing uncertainty sets that yield a good trade-off between performance and conservatism is central to robust optimization.

Following the approach developed by Bertsimas and Sim [8, 9] and Bertsimas and Thiele [10], we focus on polyhedral uncertainty sets and model the random variable b_i, $i = 1, \dots, m$, as a parameter of known mean $\overline{b}_i$ and belonging to the interval $[\overline{b}_i - \widehat{b}_i, \overline{b}_i + \widehat{b}_i]$. Equivalently:

$$b_i = \overline{b}_i + \widehat{b}_i\, z_i,\ \ |z_i| \leq 1,\ \forall i. \tag{6}$$

To avoid overprotecting the system, we impose the constraint:

$$\sum_{i=1}^{m} |z_i| \leq \Gamma, \tag{7}$$

which bounds the total scaled deviation of the parameters from their mean. Such a constraint was first proposed by Bertsimas and Sim [8] in the context of linear programming with uncertain coefficients. The parameter Γ, which we assume to be integer, is called the budget of uncertainty. $\Gamma = 0$ yields the nominal problem and hence does not incorporate uncertainty at all, while $\Gamma = m$ corresponds to interval-based uncertainty sets and leads to the most conservative case. In summary, we will consider the following uncertainty set:

$$\mathcal{B} = \left\{ b_i = \overline{b}_i + \widehat{b}_i\, z_i,\ \forall i,\ \mathbf{z} \in \mathcal{Z} \right\}, \tag{8}$$

with:

$$\mathcal{Z} = \left\{ \sum_{i=1}^{m} |z_i| \leq \Gamma,\ |z_i| \leq 1, \forall i \right\}. \tag{9}$$

In the remainder of the paper, we investigate how Problem (5) can be solved efficiently for the polyhedral set defined in Equation (8), with an emphasis on the link with deterministic linear models, and how the robust approach can help us gain insights into the impact of the uncertainty on the optimal solution.

3 The Case of Simple Recourse

3.1 The Setup

In problems with simple recourse, the decision-maker is able to address excess or shortage for each of the requirements independently. For instance, he might pay a unit penalty d_i^-, resp.

d_i^+, for falling short of, resp. exceeding, the (random) target b_i for each i. We describe an application of this setting to multi-item newsvendor problems in Section 5.1.

The deterministic model can be formulated as:

$$\begin{aligned} \min\quad & \mathbf{c}'\mathbf{x} + (\mathbf{d}^-)'\mathbf{y}^- + (\mathbf{d}^+)'\mathbf{y}^+ \\ \text{s.t.}\quad & \mathbf{A}\,\mathbf{x} + \mathbf{y}^- - \mathbf{y}^+ = \mathbf{b}, \\ & \mathbf{x},\, \mathbf{y}^-,\, \mathbf{y}^+ \geq \mathbf{0}. \end{aligned} \tag{10}$$

The recourse function defined in Equation (2) becomes:

$$\begin{aligned} \mathbf{Q}(\mathbf{x}, \mathbf{b}) = \quad \min\quad & (\mathbf{d}^-)'\mathbf{y}^- + (\mathbf{d}^+)'\mathbf{y}^+ \\ \text{s.t.}\quad & \mathbf{y}^- - \mathbf{y}^+ = \mathbf{b} - \mathbf{A}\,\mathbf{x}, \\ & \mathbf{y}^-,\, \mathbf{y}^+ \geq \mathbf{0}, \end{aligned} \tag{11}$$

It is straightforward to see that $\mathbf{Q}(\mathbf{x}, \mathbf{b})$ is available in closed form:

$$\mathbf{Q}(\mathbf{x}, \mathbf{b}) = \sum_{i=1}^{m} \left[d_i^- \cdot \max\{0, b_i - (\mathbf{A}\mathbf{x})_i\} + d_i^+ \cdot \max\{0, (\mathbf{A}\mathbf{x})_i - b_i\} \right]. \tag{12}$$

However, we will focus on Problem (11) to build a linear, tractable robust model. As Problem (11) is always feasible and its optimal value is finite, we obtain an equivalent characterization of the recourse function by invoking strong duality:

$$\begin{aligned} \mathbf{Q}(\mathbf{x}, \mathbf{b}) = \quad \max\quad & (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p} \\ \text{s.t.}\quad & -\mathbf{d}^+ \leq \mathbf{p} \leq \mathbf{d}^-. \end{aligned} \tag{13}$$

Therefore, the robust approach is concerned with developing efficient ways to solve:

$$\min_{\mathbf{x} \geq \mathbf{0}} \left[\mathbf{c}'\mathbf{x} + \max_{\mathbf{b} \in \mathcal{B},\, -\mathbf{d}^+ \leq \mathbf{p} \leq \mathbf{d}^-} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p} \right], \tag{14}$$

where $\mathcal{B}$ has been defined in Equation (8).

3.2 The Robust Problem as a Large-Scale Linear Problem

In this section, we show that the robust problem (14) can be formulated as a large-scale linear programming problem and discuss solution techniques. We will need the following lemmas:

Lemma 3.1 *The function:*

$$\mathbf{R}(\mathbf{x}, \mathbf{p}) = \max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p} \tag{15}$$

is convex in $\mathbf{p}$.

Proof: For any $\mathbf{x}$, $\mathbf{p}_1$, $\mathbf{p}_2$, for any $\alpha \in [0, 1]$, we have:

$$\max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'(\alpha\, \mathbf{p}_1 + (1 - \alpha)\, \mathbf{p}_2) \leq \alpha \cdot \max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p}_1 + (1 - \alpha) \cdot \max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p}_2, \tag{16}$$

i.e., $\mathbf{R}(\mathbf{x}, \alpha\,\mathbf{p_1} + (1-\alpha)\,\mathbf{p_2}) \le \alpha\,\mathbf{R}(\mathbf{x}, \mathbf{p_1}) + (1-\alpha)\,\mathbf{R}(\mathbf{x}, \mathbf{p_2})$. □

Lemma 3.2 *Let* $\mathbf{x}$, $\mathbf{p}$ *be given. Then:*

$$\max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p} = (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \begin{array}{ll} \min & \lambda\,\Gamma + \sum_{i=1}^{m} \mu_i \\ s.t. & \lambda + \mu_i \ge \widehat{b}_i\,|p_i|, \quad \forall i, \\ & \lambda,\ \mu_i \ge 0, \qquad \forall i. \end{array} \tag{17}$$

Proof: We have:

$$\max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p} = (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \max_{\mathbf{z}\in\mathcal{Z}} \sum_{i=1}^{m} p_i\,\widehat{b}_i\,z_i, \tag{18}$$

where $\mathcal{Z}$ has been defined in Equation (9), or equivalently:

$$\max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p} = (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \max_{\mathbf{z}'\in\mathcal{Z}'} \sum_{i=1}^{m} |p_i|\,\widehat{b}_i\,z_i', \tag{19}$$

with $\mathcal{Z}' = \left\{ \sum_{i=1}^{m} z_i' \le \Gamma,\ 0 \le z_i' \le 1,\ \forall i \right\}$. Since $\mathcal{Z}'$ is nonempty and bounded, strong duality holds, which yields Problem (17). □

Let $\mathbf{p_k}$, $k = 1, \ldots, 2^m$ be the extreme points of $[-\mathbf{d}^+, \mathbf{d}^-]$.

Theorem 3.3 (The Robust Problem) *The robust problem (14) is equivalent to the linear programming problem:*

$$\begin{array}{lll} \min & \mathbf{c}'\mathbf{x} + Z & \\ s.t. & Z \ge (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p_k} + \lambda_k\,\Gamma + \sum_{i=1}^{m} \mu_{ik}, & \forall k, \\ & \lambda_k + \mu_{ik} \ge \widehat{b}_i\,|p_{ik}|, & \forall i, k, \\ & \lambda,\ \mu,\ \mathbf{x} \ge \mathbf{0}. & \end{array} \tag{20}$$

Problem (20) has exponentially many constraints.

Proof: From Lemma 3.1, $\mathbf{R}(\mathbf{x}, \mathbf{p}) = \max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p}$ is convex in $\mathbf{p}$, and therefore reaches its maximum over $\mathbf{p} \in [-\mathbf{d}^+, \mathbf{d}^-]$ at an extreme point $\mathbf{p_k}$ of the polyhedron. It follows that the robust problem (14) is equivalent to:

$$\begin{array}{ll} \min & \mathbf{c}'\mathbf{x} + Z \\ \text{s.t.} & Z \ge \max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p_k},\ \forall k, \\ & \mathbf{x} \ge \mathbf{0}. \end{array} \tag{21}$$

We then apply Lemma 3.2 to $\max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p_k}$ and conclude by noting that, for any set S and function f, $Z \geq \min_{s\in S} f(s)$ if and only if $Z \geq f(s)$ for some $s \in S$. □

Large-scale linear programming problems such as Problem (20) can be solved by applying techniques based on delayed constraint generation, also known as cutting plane methods. Bertsimas and Tsitsiklis provide an introduction to these techniques in [7]. (The reader is also referred to Birge and Louveaux [11] and Kall and Wallace [22] for an extensive treatment of these methods in the context of stochastic optimization.) We focus here on Benders decomposition, which was developed by Benders in [2], and show below how it can be adapted to the robust methodology.

To highlight the advantages of implementing the method in the robust framework, as opposed to the stochastic model, we first provide a brief summary of Benders decomposition in stochastic programming. It is well known that the master problem can be formulated as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \sum_{\omega=1}^{\Omega} \pi_\omega Z_\omega \\ \text{s.t.} \quad & Z_\omega \geq \mathbf{p}_\mathbf{k}'(\mathbf{b}_\omega - \mathbf{A}\,\mathbf{x}) \quad \forall k, \omega, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{22}$$

where the $\mathbf{p_k}$ are the extreme points of $[-\mathbf{d}^+, \mathbf{d}^-]$. Enumerating all these extreme points obviously leads to a very large number of constraints (here, $\Omega \cdot 2^m$, excluding nonnegativity), and the main idea underlying Benders decomposition is to generate constraints only when they prove to be necessary to the formulation, i.e., when they are violated by the current solution in the master problem. At each iteration, the decision-maker solves a relaxed master problem, which has only a few of the original constraints, and obtains an optimal solution $\tilde{\mathbf{x}}$ and the corresponding value of the recourse function $\widetilde{Z}_\omega$ when scenario $\omega = 1, \ldots, \Omega$ is realized. Then he must check if that solution is optimal for the original problem. Therefore, he solves the recourse problem in each scenario with the first-stage decision variables set to $\tilde{\mathbf{x}}$. If $\widetilde{Z}_\omega \leq \mathbf{Q}(\tilde{\mathbf{x}}, \mathbf{b}_\omega)$ for each ω, the problem has been solved. Otherwise, the decision-maker adds the cut: $Z_\omega \geq \mathbf{p}_{\mathbf{k}(\omega)}'(\mathbf{b}_\omega - \mathbf{A}\,\mathbf{x})$ for each ω such that $\widetilde{Z}_\omega > \mathbf{Q}(\tilde{\mathbf{x}}, \mathbf{b}_\omega)$, where $\mathbf{p}_{\mathbf{k}(\omega)}$ is the optimal solution of Problem (13) when $\mathbf{b}$ is set to $\mathbf{b}_\omega$, and reiterates. Although some variants of Benders decomposition apply cuts to the whole recourse function $\sum_{\omega=1}^{\Omega} \pi_\omega \mathbf{Q}(\mathbf{x}, \mathbf{b}_\omega)$, these methods always require solving to optimality the recourse problems for each scenario. While these problems are similar to each other and can each be solved efficiently by applying for instance the dual simplex method, the large number of subproblems is a drawback in implementing the approach in many real-life settings. In contrast, Theorem 3.4 shows that Benders decomposition applied to robust problems with simple recourse involves only one subproblem, which can be solved in closed form. This plays a key role in the tractability of the robust approach.

Theorem 3.4 (Benders decomposition for robust problems) *Problem (20) can be solved as follows. At iteration S, $S \geq 1$, we consider the relaxed master problem:*

$$
\begin{aligned}
\min \quad & \mathbf{c}'\mathbf{x} + Z \\
s.t. \quad & Z \geq (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p}_{\mathbf{k(s)}} + \lambda_s \Gamma + \sum_{i=1}^{m} \mu_{is}, && \forall s = 1, \ldots, S-1, \\
& \lambda_s + \mu_{is} \geq \widehat{b}_i \, |p_{ik(s)}|, && \forall i = 1, \ldots, m, \forall s = 1, \ldots, S-1, \\
& \lambda_s, \, \mu_{\mathbf{s}} \geq \mathbf{0}, && \forall s = 1, \ldots, S-1, \\
& \mathbf{x} \geq \mathbf{0},
\end{aligned}
\tag{23}
$$

with the convention that $Z \geq L$ in iteration 1 where L is a lower bound for the second-stage cost. Let $(\tilde{\mathbf{x}}, \tilde{\lambda}, \tilde{\mu}, \widetilde{Z})$ be the optimal solution of Problem (23), and let $\mathcal{I}$ be the set of indices corresponding to the Γ greatest Δ_i, with Δ_i given by:

$$
\Delta_i = \max\left\{ (\overline{b}_i + \widehat{b}_i - (\mathbf{A}\tilde{\mathbf{x}})_i)\, d_i^-, ((\mathbf{A}\tilde{\mathbf{x}})_i - \overline{b}_i + \widehat{b}_i)\, d_i^+ \right\} - \max\left\{ (\overline{b}_i - (\mathbf{A}\tilde{\mathbf{x}})_i)\, d_i^-, ((\mathbf{A}\tilde{\mathbf{x}})_i - \overline{b}_i)\, d_i^+ \right\}. \tag{24}
$$

The recourse function $\mathcal{Q}(\tilde{\mathbf{x}}) = \max_{\mathbf{b} \in \mathcal{B}} \mathbf{Q}(\tilde{\mathbf{x}}, \mathbf{b})$ verifies:

$$
\begin{aligned}
\mathcal{Q}(\tilde{\mathbf{x}}) = \sum_{i \in \mathcal{I}} \max & \left\{ (\overline{b}_i + \widehat{b}_i - (\mathbf{A}\tilde{\mathbf{x}})_i)\, d_i^-, ((\mathbf{A}\tilde{\mathbf{x}})_i - \overline{b}_i + \widehat{b}_i)\, d_i^+ \right\} \\
& + \sum_{i \notin \mathcal{I}} \max \left\{ (\overline{b}_i - (\mathbf{A}\tilde{\mathbf{x}})_i)\, d_i^-, ((\mathbf{A}\tilde{\mathbf{x}})_i - \overline{b}_i)\, d_i^+ \right\}.
\end{aligned}
\tag{25}
$$

If $\widetilde{Z} \geq \mathcal{Q}(\tilde{\mathbf{x}})$, Problem (20) has been solved.
If $\widetilde{Z} < \mathcal{Q}(\tilde{\mathbf{x}})$, the solution $\tilde{\mathbf{x}}$ of the relaxed problem is not optimal for Problem (20). Let $p_{ik(S)}$ be equal to d_i^- if $\overline{b}_i \geq (\mathbf{A}\tilde{\mathbf{x}})_i + \dfrac{d_i^+ - d_i^-}{d_i^+ + d_i^-} \widehat{b}_i \cdot 1_{\{i \in \mathcal{I}\}}$ and $-d_i^+$ otherwise. We add the constraints:

$$
\begin{aligned}
& Z \geq (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p}_{\mathbf{k(S)}} + \lambda_S \Gamma + \sum_{i=1}^{m} \mu_{iS}, \\
& \lambda_S + \mu_{iS} \geq \widehat{b}_i \, |p_{i\,k(S)}|, \ \forall i, \\
& \lambda_S, \, \mu_{\mathbf{S}} \geq \mathbf{0}.
\end{aligned}
\tag{26}
$$

to the relaxed problem (23) and reiterate.

Proof: This is a straightforward application of the Benders decomposition technique to Problem (20). The only point specific to the robust framework is the computation of the recourse function $\mathcal{Q}(\tilde{\mathbf{x}})$. We note that, for any first-stage decision vector $\mathbf{x}$:

$$
\mathcal{Q}(\mathbf{x}) = \max_{\mathbf{b} \in \mathcal{B}} \max_{-\mathbf{d}^+ \leq \mathbf{p} \leq \mathbf{d}^-} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p}, \tag{27}
$$

$$
= \max_{\mathbf{b} \in \mathcal{B}} \sum_{i=1}^{m} \max\left\{ (b_i - (\mathbf{A}\,\mathbf{x})_i)\, d_i^-, ((\mathbf{A}\,\mathbf{x})_i - b_i)\, d_i^+ \right\}, \tag{28}
$$

$$= \sum_{i=1}^{m} \max\left\{(\overline{b}_i - (\mathbf{A}\,\mathbf{x})_i)\, d_i^-, ((\mathbf{A}\,\mathbf{x})_i - \overline{b}_i)\, d_i^+\right\} + \max_{\mathbf{z}' \in \mathcal{Z}'} \sum_{i=1}^{m} \Delta_i\, z_i', \tag{29}$$

where $\mathcal{Z}' = \left\{\sum_{i=1}^{m} z_i' \leq \Gamma,\ 0 \leq z_i' \leq 1\ \forall i\right\}$.

Whether the worst case is reached when b_i deviates up or down (to its lowest or highest value) is captured by the value of Δ_i. It is then obvious that $\max_{\mathbf{z}' \in \mathcal{Z}'} \sum_{i=1}^{m} \Delta_i\, z_i'$ is equal to $\sum_{i \in \mathcal{I}} \Delta_i$. We compute $\mathbf{p_{k(S)}}$ by studying where (in the left- or right-hand side) each of the inner maximizations in Equation (25) reaches its optimum. □

Remark: Instead of considering the set of equations (26), it is possible to add one single cut at each step of the algorithm, specifically:

$$Z \geq (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p_{k(S)}} + \nu_{k(S)}, \tag{30}$$

where $\nu_{k(S)}$ is the optimal objective value of:

$$\begin{aligned} \min \quad & \lambda_S\, \Gamma + \sum_{i=1}^{m} \mu_{iS} \\ \text{s.t.} \quad & \lambda_S + \mu_{iS} \geq \widehat{b}_i\, |p_{ik(S)}|, \quad \forall i, \\ & \lambda_S,\ \mu_{iS} \geq 0, \qquad\qquad\quad \forall i. \end{aligned} \tag{31}$$

However, for tractability purposes, Formulation (26) has the advantage of not requiring any auxiliary optimization.

Corollary 3.5 (Worst-case uncertainty) *Upon completion of the algorithm, a worst-case vector* $\mathbf{b}$ *is obtained by selecting a cut* $Z \geq (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p_k} + \lambda_k\, \Gamma + \sum_{i=1}^{m} \mu_{ik}$ *which is tight at optimality, and solving* $\max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p_k}$ *for that vector* $\mathbf{p_k}$. *Specifically,* $\mathbf{b}$ *verifies for all i:*

$$b_i = \overline{b}_i + \widehat{b}_i \cdot sign(p_{ik}) \cdot 1_{\{i \in \mathcal{S}\}}, \tag{32}$$

where $sign(x) = 1$ *if* $x \geq 0$ *and* -1 *otherwise,* $\mathcal{S}$ *is the set of the* Γ *greatest* $\widehat{b}_i\, |p_{ik}|$ *with* $|\mathcal{S}| = \Gamma$ *(ties can be broken arbitrarily) and* $1_{\{i \in \mathcal{S}\}} = 1$ *if* $i \in \mathcal{S}$ *and* 0 *otherwise.*

Proof: Follows immediately from Theorem 3.3 and using the fact that $\max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p_k} = \max_{\mathbf{b} \in \mathcal{B}} \mathbf{b}'\mathbf{p_k} = \max_{\mathbf{z} \in \mathcal{Z}} \sum_{i=1}^{m} \widehat{b}_i\, p_{ik}\, z_i$. □

Benders decomposition is hence faster in the robust framework than in its stochastic counterpart, in the sense that each iteration after solving the relaxed master problem only involves computing a quantity available in closed form, as opposed to solving Ω optimization problems. In particular, since we have m random variables, which will at least take 2 values each (but often more to yield a finer description of uncertainty), we have replaced the optimization of at least 2^m linear problems by the evaluation of a mathematical expression that can be computed without using any optimization software.

As constraints in cutting plane methods are only generated when needed, the optimal solution will in general be reached after enumerating only a few of the extreme points of the dual feasible set $[-\mathbf{d}^+, \mathbf{d}^-]$. However, there is no guarantee that this will always happen, which raises the following question: can we use the special structure of the robust optimization problem (14) to devise a more efficient algorithm? This will be the purpose of Section 3.3.

3.3 The Robust Problem as a Series of Linear Problems

In what follows, we show that the robust problem with simple recourse can be solved as a series of linear problems of moderate size and identify the worst-case vector $\mathbf{b} \in \mathcal{B}$ in this framework.

Theorem 3.6 (The Robust Problem) *The optimal solution* $\mathbf{x}$ *to the robust problem (14) can be found by solving the following* m *linear problems and keeping the solution corresponding to the problem with the smallest optimal value:*

Problem j, j=1,...,m:

$$
\begin{aligned}
\min \quad & \mathbf{c}'\mathbf{x} + Z && (33)\\
s.t. \quad & Z \geq \sum_{i \neq j} u_i, && (34)\\
& Z \geq \left[(\mathbf{A}\mathbf{x})_j - \overline{b}_j + \Gamma\,\widehat{b}_j\right] d_j^+ + \sum_{i \neq j} v_i, && (35)\\
& Z \geq \left[\overline{b}_j + \Gamma\,\widehat{b}_j - (\mathbf{A}\,\mathbf{x})_j\right] d_j^- + \sum_{i \neq j} w_i, && (36)\\
& \mathbf{A}\mathbf{x} - (1/\mathbf{d}^+) \cdot \mathbf{u} \leq \overline{\mathbf{b}} - \widehat{\mathbf{b}}, && (37)\\
& \mathbf{A}\mathbf{x} + (1/\mathbf{d}^-) \cdot \mathbf{u} \geq \overline{\mathbf{b}} + \widehat{\mathbf{b}}, && (38)\\
& \mathbf{A}\mathbf{x} - (1/\mathbf{d}^+) \cdot \mathbf{v} \leq \overline{\mathbf{b}} - \widehat{\mathbf{f}}_\mathbf{j}^+, && (39)\\
& \mathbf{A}\mathbf{x} + (1/\mathbf{d}^-) \cdot \mathbf{v} \geq \overline{\mathbf{b}} + \widehat{\mathbf{f}}_\mathbf{j}^-, && (40)\\
& \mathbf{A}\mathbf{x} - (1/\mathbf{d}^+) \cdot \mathbf{w} \leq \overline{\mathbf{b}} - \widehat{\mathbf{g}}_\mathbf{j}^+, && (41)\\
& \mathbf{A}\mathbf{x} + (1/\mathbf{d}^-) \cdot \mathbf{w} \geq \overline{\mathbf{b}} + \widehat{\mathbf{g}}_\mathbf{j}^-, && (42)\\
& \mathbf{x} \geq \mathbf{0}, && (43)
\end{aligned}
$$

with the following notations for all i:

$$\begin{cases} f_{ij}^{+} = \max\left(0, \widehat{b}_i - \dfrac{d_j^+}{d_i^+}\,\widehat{b}_j\right), \; f_{ij}^{-} = \max\left(0, \widehat{b}_i - \dfrac{d_j^+}{d_i^-}\,\widehat{b}_j\right), \\ g_{ij}^{+} = \max\left(0, \widehat{b}_i - \dfrac{d_j^-}{d_i^+}\,\widehat{b}_j\right), \; g_{ij}^{-} = \max\left(0, \widehat{b}_i - \dfrac{d_j^-}{d_i^-}\,\widehat{b}_j\right), \\ (1/d^+)_i = 1/d_i^+, \; (1/d^-)_i = 1/d_i^-. \end{cases}$$

Each problem has $n+3\,m+1$ *decision variables and* $6\,m+3$ *constraints, excluding nonnegativity. Therefore, the robust problem can be solved efficiently by standard linear optimization packages.*

Proof: Using the definition of $\mathcal{B}$ in Equation (8) and $\mathcal{Z}$ in Equation (9), we have:

$$\max_{\mathbf{b}\in\mathcal{B},\, -\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} (\mathbf{b}-\mathbf{A}\mathbf{x})'\mathbf{p} = \max_{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} \left\{ (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \max_{\mathbf{z}\in\mathcal{Z}} \sum_{i=1}^{m} p_i \widehat{b}_i\, z_i \right\}. \tag{44}$$

From Lemma 3.2, $\max_{\mathbf{z}\in\mathcal{Z}} \sum_{i=1}^{m} p_i\,\widehat{b}_i\,z_i$ is equivalent to:

$$\begin{array}{lll} \min & \lambda\,\Gamma + \sum_{i=1}^{m} \mu_i & \\ \text{s.t.} & \lambda + \mu_i \ge \widehat{b}_i\,|p_i|, & \forall i, \\ & \lambda,\, \mu_i \ge 0, & \forall i. \end{array} \tag{45}$$

It is straightforward to see that the optimal solution of Problem (45) verifies:

$$\mu_i = \max(0, \widehat{b}_i|p_i| - \lambda),\;\; \forall i. \tag{46}$$

As a result, $\max_{\mathbf{b}\in\mathcal{B},\, -\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} (\mathbf{b}-\mathbf{A}\mathbf{x})'\mathbf{p}$ can be rewritten as:

$$\max_{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} \left\{ (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \min_{\lambda\ge 0} \left[\lambda\,\Gamma + \sum_{i=1}^{m} \max(0, \widehat{b}_i\,|p_i| - \lambda) \right] \right\}. \tag{47}$$

For a given $\mathbf{p}$, the function:

$$F(\lambda) = \lambda\,\Gamma + \sum_{i=1}^{m} \max(0, \widehat{b}_i\,|p_i| - \lambda) \tag{48}$$

is piecewise linear, convex in λ, with breakpoints at $\widehat{b}_i\,|p_i|$, $i = 1,\ldots,m$, and its optimum is reached at the Γ-th greatest $\widehat{b}_i\,|p_i|$. Consequently, solving Problem (14) amounts to solving:

$$\min_{\mathbf{x}\ge\mathbf{0}} \left[\mathbf{c}'\mathbf{x} + \max_{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} \left\{ (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \lambda\,\Gamma + \sum_{i=1}^{m} \max(0, \widehat{b}_i\,|p_i| - \lambda) \right\} \right], \tag{49}$$

for $\lambda = \widehat{b}_j\,|p_j|$, $j = 1, \ldots, m$, and keeping the problem that yields the smallest objective at optimality. (This is because $\min_{\lambda \geq 0}\left[\lambda\,\Gamma + \sum_{i=1}^{m} \max(0, \widehat{b}_i\,|p_i| - \lambda)\right] \leq \lambda\,\Gamma + \sum_{i=1}^{m} \max(0, \widehat{b}_i\,|p_i| - \lambda)$ for all $\lambda \geq 0$, so that the optimal solution to Problem (47) will be less than or equal to the optimal solution to Problem (49) for the m possible values of λ considered, and the inequality will be tight for the optimal λ.)

We now focus on rewriting Problem (49) for $\lambda = \widehat{b}_j\,|p_j|$ as a linear programming problem. Our goal is thus to solve:

$$\min_{\mathbf{x} \geq \mathbf{0}} \left[\mathbf{c}'\mathbf{x} + \max_{-\mathbf{d}^+ \leq \mathbf{p} \leq \mathbf{d}^-} \left\{ (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p} + \widehat{b}_j\,|p_j|\,\Gamma + \sum_{i=1}^{m} \max\left(0, \widehat{b}_i\,|p_i| - \widehat{b}_j\,|p_j|\right) \right\} \right] \tag{50}$$

For a given p_j, the functions $(\overline{b}_i - (\mathbf{A}\,\mathbf{x})_i)\,p_i + \max(0, \widehat{b}_i\,|p_i| - \widehat{b}_j\,|p_j|)$ with $i \neq j$ are convex in p_i and therefore reach their maximum over $[-d_i^+, d_i^-]$ at an extremity of the feasible set. Hence, the inner maximization problem in Problem (50) can be rewritten as:

$$\begin{aligned} &\max_{-d_j^+ \leq p_j \leq d_j^-} \Big[(\overline{b}_j - (\mathbf{A}\,\mathbf{x})_j)\,p_j + \widehat{b}_j\,|p_j|\,\Gamma \\ &+ \sum_{i \neq j} \max\left\{ -(\overline{b}_i - (\mathbf{A}\,\mathbf{x})_i)\,d_i^+ + \max(0, \widehat{b}_i\,d_i^+ - \widehat{b}_j\,|p_j|), (\overline{b}_i - (\mathbf{A}\,\mathbf{x})_i)\,d_i^- + \max(0, \widehat{b}_i\,d_i^- - \widehat{b}_j\,|p_j|) \right\} \Big]. \end{aligned} \tag{51}$$

The function to be maximized in Problem (51) is convex in p_j over $[-d_j^+, 0]$ and $[0, d_j^-]$, and therefore reaches its maximum over $[-d_j^+, d_j^-]$ at either $-d_j^+$, 0 or d_j^-. As a result, Problem (50) is equivalent to:

$$\begin{aligned} \min_{\mathbf{x} \geq \mathbf{0}} \mathbf{c}'\mathbf{x} + \max \Bigg\{ & \sum_{i \neq j} \max\left\{ [(\mathbf{A}\mathbf{x})_i - \overline{b}_i + \widehat{b}_i]\,d_i^+, [\overline{b}_i + \widehat{b}_i - (\mathbf{A}\mathbf{x})_i]\,d_i^- \right\}, \\ & [(\mathbf{A}\mathbf{x})_j - \overline{b}_j + \widehat{b}_j\,\Gamma]\,d_j^+ + \sum_{i \neq j} \max\left\{ [(\mathbf{A}\mathbf{x})_i - \overline{b}_i + f_{ij}^+]\,d_i^+, [\overline{b}_i + f_{ij}^- - (\mathbf{A}\mathbf{x})_i]\,d_i^- \right\}, \\ & [\overline{b}_j + \widehat{b}_j\,\Gamma - (\mathbf{A}\mathbf{x})_j]\,d_j^- + \sum_{i \neq j} \max\left\{ [(\mathbf{A}\mathbf{x})_i - \overline{b}_i + g_{ij}^+]\,d_i^+, [\overline{b}_i + g_{ij}^- - (\mathbf{A}\mathbf{x})_i]\,d_i^- \right\} \Bigg\}, \end{aligned} \tag{52}$$

with $f_{ij}^+ = \max\left(0, \widehat{b}_i - \frac{d_j^+}{d_i^+}\,\widehat{b}_j\right)$, $f_{ij}^- = \max\left(0, \widehat{b}_i - \frac{d_j^+}{d_i^-}\,\widehat{b}_j\right)$, $g_{ij}^+ = \max\left(0, \widehat{b}_i - \frac{d_j^-}{d_i^+}\,\widehat{b}_j\right)$ and $g_{ij}^- = \max\left(0, \widehat{b}_i - \frac{d_j^-}{d_i^-}\,\widehat{b}_j\right)$ for all i. Linearizing the convex piecewise linear terms concludes the proof. □

We now compute the worst-case value of the uncertainty. For this we need the optimal dual vector $\mathbf{p}^*$.

Lemma 3.7 (Optimal dual variables) *Let Problem j defined by Equations (33)-(43) for some j yield the smallest objective among the m problems considered in Theorem 3.6. An optimal dual* *vector $\mathbf{p}^*$ is obtained as follows.*

(i) *If Constraint (34) is tight at optimality, then $p_j^* = 0$ and for all $i \neq j$, $p_i^* = -d_i^+$ if row i of Constraint (37) is tight at optimality and $p_i^* = d_i^-$ if it is not.*

(ii) *If Constraint (35) is tight at optimality, then $p_j^* = -d_j^+$ and for all $i \neq j$, $p_i^* = -d_i^+$ if row i of Constraint (39) is tight at optimality and $p_i^* = d_i^-$ if it is not.*

(iii) *If Constraint (36) is tight at optimality, then $p_j^* = d_j^-$ and for all $i \neq j$, $p_i^* = -d_i^+$ if row i of Constraint (41) is tight at optimality and $p_i^* = d_i^-$ if it is not.*

If several of the constraints (34)-(36) are tight at the optimal solution, any of the corresponding cases (i)-(iii) can be chosen to define $\mathbf{p}^*$.

Proof: Follows directly from the proof of Theorem 3.6. It is obvious from the definition of Problem j that at least one constraint among Equations (34)-(36) is tight at optimality, which yields cases (i) to (iii). Furthermore, the tight constraints among Equations (34)-(42) enable us to identify where the convex piecewise linear functions in Problem (52) reach their maximum. □

Finally, we derive the worst-case value of the uncertain vector $\mathbf{b}$ in Corollary 3.8.

Corollary 3.8 (Worst-case uncertainty) *Let $\mathbf{p}^*$ be the optimal dual vector obtained in Lemma 3.7 and let $\mathcal{S}$ be a set of the Γ greatest $\widehat{b}_i|p_i^*|$ with $|\mathcal{S}| = \Gamma$. (Ties can be broken arbitrarily.) Then a worst-case vector $\mathbf{b}$ in the recourse problem (14) is given for all $i = 1, \ldots, m$ by:*

$$b_i = \overline{b}_i + \widehat{b}_i \cdot sign(p_i^*) \cdot 1_{\{i \in \mathcal{S}\}}, \tag{53}$$

where $sign(x) = 1$ if $x \geq 0$ and -1 otherwise, and $1_{\{i\in\mathcal{S}\}} = 1$ if $i \in \mathcal{S}$ and 0 otherwise.

Proof: Follows from solving $\max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p}^*$, that is, $\max_{\mathbf{z}\in\mathcal{Z}} \sum_{i=1}^{m} \widehat{b}_i\, p_i^*\, z_i$. □

Remark: As a consequence of Corollary 3.8, the worst case is to have a higher requirement than average for item i ($b_i = \overline{b}_i + \widehat{b}_i$) if the total shortage cost $d_i^-\widehat{b}_i$ is above the threshold λ^*, less than average ($b_i = \overline{b}_i - \widehat{b}_i$) if the total surplus cost $d_i^+\widehat{b}_i$ is above the same threshold λ^*, and equal to its average ($b_i = \overline{b}_i$) otherwise. There might be more than one worst-case vector $\mathbf{b}$. λ^* is set so that at most Γ requirements differ from their nominal value.

4 The Case of General Recourse

4.1 The Setup

We now extend the framework developed in Section 3 for simple recourse to the general case. From Equation (5), the robust problem can be formulated as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \max_{\mathbf{b}\in\mathcal{B}} \mathbf{Q}(\mathbf{x},\mathbf{b}) \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{54}$$

where the recourse function $\mathbf{Q}(\mathbf{x},\mathbf{b})$ has been defined in Equation (2):

$$\begin{aligned} \mathbf{Q}(\mathbf{x},\mathbf{b}) = \quad \min \quad & \mathbf{d}'\mathbf{y} \\ \text{s.t.} \quad & \mathbf{B}\,\mathbf{y} = \mathbf{b} - \mathbf{A}\,\mathbf{x}, \\ & \mathbf{y} \geq \mathbf{0}, \end{aligned} \tag{55}$$

We assume that relative complete recourse holds, so that Problem (55) always has a finite optimal value. Therefore, by strong duality, Problem (55) is equivalent to:

$$\begin{aligned} \mathbf{Q}(\mathbf{x},\mathbf{b}) = \quad \max \quad & (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p} \\ \text{s.t.} \quad & \mathbf{B}'\,\mathbf{p} \leq \mathbf{d}. \end{aligned} \tag{56}$$

Hence, the focus of this section will be to solve:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \max_{\mathbf{b}\in\mathcal{B},\ \mathbf{B}'\mathbf{p}\leq\mathbf{d}} (\mathbf{b} - \mathbf{A}\,\mathbf{x})'\mathbf{p} \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{57}$$

4.2 Large-Scale Formulation

In this section, we present a large-scale linear formulation of Problem (57). Let $\mathbf{p_k}$, $k = 1, \ldots, K$, be the extreme points of $\{\mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$.

Theorem 4.1 (The Robust Problem) *The robust problem (57) is equivalent to the linear programming problem:*

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + Z \\ s.t. \quad & Z \geq (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p_k} + \lambda_k\,\Gamma + \sum_{i=1}^{m} \mu_{ik}, && \forall k, \\ & \lambda_k + \mu_{ik} \geq \widehat{b}_i\,|p_{ik}|, && \forall i,k, \\ & \lambda_k, \mu_{ik} \geq 0, && \forall i,k, \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{58}$$

Proof: See the proof of Theorem 3.3. □

Problem (58) has potentially a very large number of constraints, which motivates the use of an algorithm based on delayed constraint generation as in Section 3.2. In order to decide which constraints to incorporate in the relaxed master problem, we need to evaluate the recourse function at the candidate solution. This is the purpose of Lemma 4.2. While the methods we provide are widely applicable, the most efficient algorithms will take advantage of the structure of the set $\{\mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$. For instance, in the case of simple recourse described in Section 3.2, we have used that this set was separable in the p_i, $i = 1, \ldots, m$.

Lemma 4.2 (Computing the recourse function) *Let* $\mathbf{x}$ *be a first-stage decision vector. The value* $\mathcal{Q}(\mathbf{x})$ *of the recourse function at* $\mathbf{x}$*, which is the optimal solution of:*

$$\max_{\mathbf{B}'\mathbf{p}\leq\mathbf{d}} \max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p}, \tag{59}$$

can be computed using either one of the following two methods.

(a) Let $\mathbf{p_1}, \ldots, \mathbf{p_K}$ *be the extreme points of* $\{\mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$*. Then Problem (59) is equivalent to the linear programming problem:*

$$\begin{array}{lll} \max & Q & \\ s.t. & Q \geq (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'\mathbf{p_k} + \lambda_k \Gamma + \sum_{i=1}^{m} \mu_{ik}, & \forall k, \\ & \lambda_k + \mu_{ik} \geq \widehat{b}_i\, |p_{ki}|, & \forall i, k, \\ & \lambda_k,\, \mu_{ik} \geq 0, & \forall i, k. \end{array} \tag{60}$$

(b) Problem (59) is also equivalent to the mixed-integer programming problem:

$$\begin{array}{ll} \max & (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'(\mathbf{p}^+ - \mathbf{p}^-) + \widehat{\mathbf{b}}'(\mathbf{q}^+ + \mathbf{q}^-) \\ s.t. & \mathbf{B}'(\mathbf{p}^+ - \mathbf{p}^-) \leq \mathbf{d}, \\ & \mathbf{e}'(\mathbf{r}^+ + \mathbf{r}^-) \leq \Gamma, \\ & \mathbf{q}^+ \leq M\,\mathbf{r}^+, \\ & \mathbf{q}^- \leq M\,\mathbf{r}^-, \\ & \mathbf{0} \leq \mathbf{q}^+ \leq \mathbf{p}^+, \\ & \mathbf{0} \leq \mathbf{q}^- \leq \mathbf{p}^-, \\ & \mathbf{r}^+, \mathbf{r}^- \in \{0,1\}^m, \\ & \mathbf{p}^+,\, \mathbf{p}^- \geq \mathbf{0}, \end{array} \tag{61}$$

where $\mathbf{e}$ *is the vector of all ones and* M *is a large positive number.*

Proof: (a) We know from Lemma 3.1 that $\mathbf{R}(\mathbf{x}, \mathbf{p}) = \max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})'\mathbf{p}$ is convex in $\mathbf{p}$, and hence reaches its maximum over $\{\mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$ at an extreme point of the set. It follows that Problem (59) is equivalent to:

$$
\begin{aligned}
\max \quad & Q \\
\text{s.t.} \quad & Q \geq \max_{\mathbf{b} \in \mathcal{B}} (\mathbf{b} - \mathbf{A}\mathbf{x})' \mathbf{p_k}, \ \forall k,
\end{aligned} \tag{62}
$$

where the $\mathbf{p_k}$, $k = 1, \ldots, K$, are the extreme points of $\{\mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$. Applying Lemma 3.2 to each constraint in Problem (62) yields Problem (60).

(b) Using the notation $\mathbf{p} = \mathbf{p}^+ - \mathbf{p}^-$ with $\mathbf{p}^+$, $\mathbf{p}^- \geq \mathbf{0}$, Problem (59) can be rewritten as:

$$
\begin{aligned}
\max \quad & (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'(\mathbf{p}^+ - \mathbf{p}^-) + \max_{\mathbf{z} \in \mathcal{Z}} \sum_{i=1}^{m} \widehat{b}_i \, (p_i^+ - p_i^-) \, z_i \\
\text{s.t.} \quad & \mathbf{B}'(\mathbf{p}^+ - \mathbf{p}^-) \leq \mathbf{d}, \\
& \mathbf{p}^+, \mathbf{p}^- \geq \mathbf{0},
\end{aligned} \tag{63}
$$

where $\mathcal{Z}$ was defined in Equation (9). We note that $\max_{\mathbf{z} \in \mathcal{Z}} \sum_{i=1}^{m} \widehat{b}_i \, (p_i^+ - p_i^-) \, z_i$ is equivalent to:

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{m} \widehat{b}_i \, (p_i^+ - p_i^-) \, (z_i^+ - z_i^-) \\
\text{s.t.} \quad & \sum_{i=1}^{m} (z_i^+ + z_i^-) \leq \Gamma, \\
& 0 \leq z_i^-, z_i^+ \leq 1, \ \forall i,
\end{aligned} \tag{64}
$$

where $z_i = z_i^+ - z_i^-$ and $|z_i| = z_i^+ + z_i^-$ for all i. Without loss of generality, we can assume that $p_i^+ \, z_i^- = p_i^- \, z_i^+ = 0$ for all i, as it is suboptimal to select $z_i < 0$ when $p_i > 0$, and $z_i > 0$ when $p_i < 0$. Therefore, Problem (63) becomes:

$$
\begin{aligned}
\max \quad & (\overline{\mathbf{b}} - \mathbf{A}\mathbf{x})'(\mathbf{p}^+ - \mathbf{p}^-) + \sum_{i=1}^{m} \widehat{b}_i \, (p_i^+ z_i^+ + p_i^- z_i^-) \\
\text{s.t.} \quad & \mathbf{B}'(\mathbf{p}^+ - \mathbf{p}^-) \leq \mathbf{d}, \\
& \mathbf{e}'(\mathbf{z}^+ + \mathbf{z}^-) \leq \Gamma, \\
& \mathbf{0} \leq \mathbf{z}^-, \mathbf{z}^+ \leq \mathbf{e}, \\
& \mathbf{p}^+, \mathbf{p}^- \geq \mathbf{0},
\end{aligned} \tag{65}
$$

where $\mathbf{e}$ is the vector of all ones. Formulation (61) follows by introducing new nonnegative variables $\mathbf{q}^+$, $\mathbf{q}^-$ and binary variables $\mathbf{r}^+$, $\mathbf{r}^-$ to obtain a linear objective function and enforce, through additional constraints, that $q_i^+ = p_i^+ z_i^+$ and $q_i^- = p_i^- z_i^-$ for all i at optimality. □

Theorem 4.3, which adapts the Benders decomposition algorithm to the robust framework, is the main result of this section.

Theorem 4.3 (Benders decomposition for robust problems: the general case) *Problem (58) can be solved as follows. At iteration S, $S \geq 1$, we consider the relaxed master problem:*

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + Z \\ s.t. \quad & Z \geq (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'(\mathbf{p}_{\mathbf{k}_s}) + \nu_{k_s}, \quad \forall s = 1, \ldots, S-1, \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{66}$$

with the convention that $Z \geq L$ on iteration 1 where L is a lower bound for the second-stage cost. Let $(\tilde{\mathbf{x}}, \widetilde{Z})$ be the optimal solution of Problem (66), and let $\mathcal{Q}(\tilde{\mathbf{x}})$ be the value of the recourse function at $\tilde{\mathbf{x}}$ obtained in Lemma 4.2.
If $\widetilde{Z} \geq \mathcal{Q}(\tilde{\mathbf{x}})$, Problem (58) has been solved and $\tilde{\mathbf{x}}$ is the optimal solution.
If $\widetilde{Z} < \mathcal{Q}(\tilde{\mathbf{x}})$, the solution $\tilde{\mathbf{x}}$ is not optimal for Problem (58). Let $\mathbf{p}_{\mathbf{k}\mathbf{S}}$ and $\nu_{\mathbf{k}\mathbf{S}}$ be such that $\mathcal{Q}(\tilde{\mathbf{x}}) = (\overline{\mathbf{b}} - \mathbf{A}\tilde{\mathbf{x}})'(\mathbf{p}_{\mathbf{k}\mathbf{S}}) + \nu_{k_S}$. We add the cut:

$$Z \geq (\overline{\mathbf{b}} - \mathbf{A}\,\mathbf{x})'\mathbf{p}_{\mathbf{k}\mathbf{S}} + \nu_{k_S}, \tag{67}$$

to the relaxed problem (66) and reiterate.

Proof: See the proof of Theorem 3.4. □

4.3 A Suboptimal Solution

We have presented in Section 4.2 an exact algorithm to find the optimal solution to the robust problem. However, since this method requires an auxiliary optimization procedure at each step of the algorithm to compute the recourse function, we are interested in speeding it up by initializing the first-stage variables appropriately. We do so by solving a relaxation of the two-stage problem, which allows us to use the techniques developed in Section 3. This method also provides us with an upper bound to the optimal cost in the robust framework.

Let $d_i^+ = -\min\{p_i | \mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$ and $d_i^- = \max\{p_i | \mathbf{B}'\mathbf{p} \leq \mathbf{d}\}$ for all $i = 1, \ldots, m$.

Theorem 4.4 (Relaxed Problem) *An upper bound to the optimal cost of Problem (57) and a candidate first-stage solution are obtained by solving the following two-stage linear problem with simple recourse:*

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \mathbf{d}'\mathbf{y} + \max_{\mathbf{b} \in \mathcal{B}} \mathbf{Q}(\mathbf{x}, \mathbf{y}, \mathbf{b}) \\ s.t. \quad & \mathbf{x}, \mathbf{y} \geq \mathbf{0}, \end{aligned} \tag{68}$$

where the recourse function $\mathbf{Q}(\mathbf{x}, \mathbf{y}, \mathbf{b})$ is defined by:

$$\begin{aligned} \mathbf{Q}(\mathbf{x}, \mathbf{y}, \mathbf{b}) = \quad \min \quad & (\mathbf{d}^-)'\mathbf{y}^- + (\mathbf{d}^+)'\mathbf{y}^+ \\ s.t. \quad & \mathbf{y}^- - \mathbf{y}^+ = \mathbf{b} - (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}), \\ & \mathbf{y}^-, \mathbf{y}^+ \geq \mathbf{0}. \end{aligned} \tag{69}$$

Proof: For any $\mathbf{p}$ such that $\mathbf{B}'\mathbf{p} \leq \mathbf{d}$ and for any $\mathbf{y} \geq \mathbf{0}$, we have:

$$(\mathbf{d} - \mathbf{B}'\mathbf{p})'\mathbf{y} \geq 0. \tag{70}$$

Let $\mathbf{R}(\mathbf{x},\mathbf{p}) = \max_{\mathbf{b}\in\mathcal{B}} (\mathbf{b}-\mathbf{A}\mathbf{x})'\mathbf{p}$. It follows from Equation (70) that:

$$\max_{\mathbf{B}'\mathbf{p}\le\mathbf{d}} \mathbf{R}(\mathbf{x},\mathbf{p}) \le \max_{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} \left[\mathbf{R}(\mathbf{x},\mathbf{p}) + (\mathbf{d}-\mathbf{B}'\mathbf{p})'\mathbf{y}\right], \tag{71}$$

for any $\mathbf{y}\ge\mathbf{0}$, where we have used that the set $\{\mathbf{B}'\mathbf{p}\le\mathbf{d}\}$ is included in $\{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-\}$. In particular, Equation (71) holds for the value of $\mathbf{y}$ that yields the lowest right-hand side. This yields:

$$\min_{\mathbf{x}\ge\mathbf{0}} \left\{\mathbf{c}'\mathbf{x} + \max_{\mathbf{B}'\mathbf{p}\le\mathbf{d}} \mathbf{R}(\mathbf{x},\mathbf{p})\right\} \le \min_{\mathbf{x},\mathbf{y}\ge\mathbf{0}} \left\{\mathbf{c}'\mathbf{x} + \mathbf{d}'\mathbf{y} + \max_{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-} \left[\mathbf{R}(\mathbf{x},\mathbf{p}) - (\mathbf{B}\,\mathbf{y})'\mathbf{p}\right]\right\}, \tag{72}$$

or equivalently:

$$\min_{\mathbf{x}\ge\mathbf{0}} \left\{\mathbf{c}'\mathbf{x} + \max_{\mathbf{B}'\mathbf{p}\le\mathbf{d},\mathbf{b}\in\mathcal{B}} (\mathbf{b}-\mathbf{A}\mathbf{x})'\mathbf{p}\right\} \le \min_{\mathbf{x},\mathbf{y}\ge\mathbf{0}} \left\{\mathbf{c}'\mathbf{x} + \mathbf{d}'\mathbf{y} + \max_{-\mathbf{d}^+\le\mathbf{p}\le\mathbf{d}^-,\mathbf{b}\in\mathcal{B}} \left[(\mathbf{b}-\mathbf{A}\mathbf{x})'\mathbf{p} - (\mathbf{B}\,\mathbf{y})'\mathbf{p}\right]\right\}. \tag{73}$$

The left-hand side is the robust problem (57) and the right-hand side can be rewritten as Problem (68) by invoking strong duality for the inner maximization over $\mathbf{p}$. □

Remark: In many practical cases, $\mathbf{y}=\mathbf{0}$ at optimality, for instance if $\mathbf{B}=\mathbf{A}$ and $\mathbf{d}>\mathbf{c}$, where the second stage consists in ordering raw materials at a higher unit cost for the same production process as in the first stage. This amounts to optimizing $\mathbf{R}(\mathbf{x},\mathbf{p})$ over a box uncertainty set that contains $\{\mathbf{B}'\mathbf{p}\le\mathbf{d}\}$ but is more convenient to manipulate. However, if the decision-maker uses, say, a different production process in the second stage, the additional ordering cost might be offset by the increase in productivity, yielding a nonzero $\mathbf{y}$.

5 Computational Results

In this section, we present two numerical experiments to illustrate the robust methodology. The first example is a multi-item newsvendor problem with penalties on surplus and shortage, which falls within the framework developed in Section 3. The second example is a production planning problem, where the recourse is to buy additional raw material at a higher cost. This is a case of general recourse, for which we apply the approach described in Section 4.

5.1 Newsvendor Problem

We first implement the robust methodology in a multi-item newsvendor problem. The decision-maker orders perishable items subject to uncertain demand, given a budget constraint, and incurs surplus and shortage costs for each item at the end of the time period. His goal is to minimize total cost. We use the following notations:

n : the number of items,

c_i : the unit ordering cost of item i,

d_i^+ : the unit holding cost of item i,

d_i^- : the unit shortage cost of item i,

b_i : the demand for item i,

A : the budget available.

The deterministic problem can be formulated as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \sum_{i=1}^{n} \max\left\{ d_i^- \left(b_i - x_i\right), d_i^+ \left(x_i - b_i\right) \right\} \\ \text{s.t.} \quad & \mathbf{c}'\mathbf{x} \le A, \\ & x_i \ge 0, \forall i, \end{aligned} \tag{74}$$

or equivalently, as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + (\mathbf{d}^-)'\mathbf{y}^- + (\mathbf{d}^+)'\mathbf{y}^+ \\ \text{s.t.} \quad & \mathbf{x} + \mathbf{y}^- - \mathbf{y}^+ = \mathbf{b}, \\ & \mathbf{c}'\mathbf{x} \le A, \\ & \mathbf{x} \ge \mathbf{0}. \end{aligned} \tag{75}$$

Problem (75) is an example of linear programming problem with simple recourse and therefore can be analyzed using the techniques described in Section 3. We consider a case with 50 items and capacity 5,000 units, with ordering cost $c_i = 1$, nominal demand $\overline{b}_i = 8 + 2\,i$, maximum deviation of the demand from its nominal value $\widehat{b}_i = 0.5 \cdot \overline{b}_i$, surplus penalty $d_i^+ = i$, and shortage penalty $d_i^- = 2\,i$ for each $i = 1, \ldots, 50$. We apply the modified Benders decomposition algorithm presented in Section 3.2 using AMPL/CPLEX v.8.1 on a Pentium IV.

Figure 1 illustrates the effect of the budget of uncertainty Γ on the expected newsvendor cost. Sample averages have been computed on a sample of 500 realizations of the demands, which we generated using independent normal random variables with mean $\overline{b}_i$ and standard deviation $0.2 \cdot \overline{b}_i$ for all i. We observe that the average cost first decreases with Γ, as incorporating a small amount of uncertainty in the model yields to more robust solutions, then reaches its minimum and starts increasing with Γ, as the solution becomes overly conservative. The optimal trade-off is reached for $\Gamma = 5$, which is consistent with the guidelines provided by Bertsimas and Sim in [8], namely, that the budget of uncertainty should be of the order of $\sqrt{n}$ (here, $\sqrt{50} \approx 7.1$). The worst case when $\Gamma = 5$ corresponds to the case where the demand for the last 5 items (items 46 to 50) is equal to its highest value, and demand for the other items is equal to its mean.

Figure 2 shows the number of iterations needed to reach optimality in the delayed constraint generation algorithm, as a function of the budget of uncertainty. While the feasible set $[-\mathbf{d}^+, \mathbf{d}^-]$

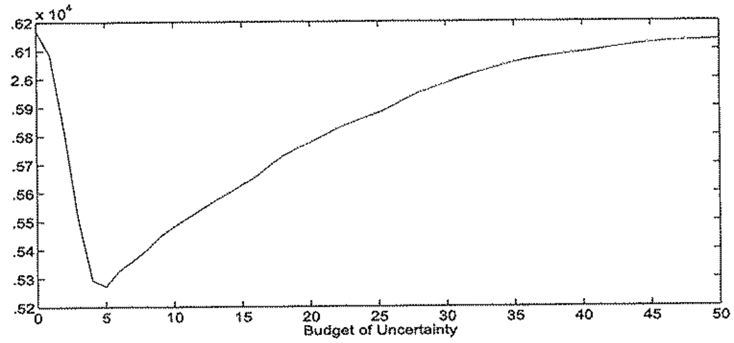

Figure 1: Impact of the budget of uncertainty on the expected cost.

has 2^{50} extreme points, i.e., Benders decomposition could in theory generate $2^{50} \approx 1.1 \cdot 10^{15}$ constraints, the algorithm converges in at most 334 iterations. We note that the number of iterations increases with the protection level Γ, so that if we take $\Gamma = 5$ as suggested by Figure 1, we only need 196 iterations to reach optimality.

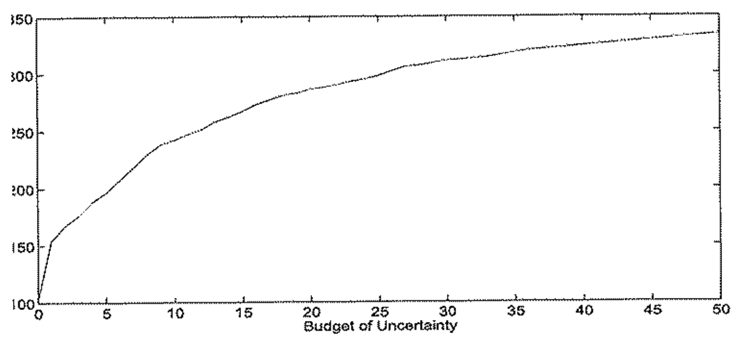

Figure 2: Number of iterations as a function of the budget of uncertainty.

Hence, in this example with simple recourse, the robust optimization approach incorporates uncertainty in a tractable manner while preserving performance.

5.2 Production Planning

Here, we consider a production planning example where the demand is uncertain but must be met. Once demand has been revealed, the decision-maker has the option to buy additional raw material at a more expensive cost and re-run the production process, so that demand for all products is satisfied. His goal is to minimize the ordering cost of raw materials in both stages. We define the following notations:

m : the number of raw materials,

n : the number of finished products,

$\mathbf{c}$: the first-stage unit ordering cost of the raw materials,

$\mathbf{d}$: the second-stage unit ordering cost of the raw materials,

$\mathbf{x}$: the raw materials ordered in the first stage,

$\mathbf{y}$: the raw materials ordered in the second stage,

$\mathbf{u}$: the products manufactured in the first stage,

$\mathbf{v}$: the products manufactured in the second stage,

$\mathbf{A}$: the first-stage productivity matrix,

$\mathbf{B}$: the second-stage productivity matrix,

$\mathbf{b}$: the demand for the finished products.

We assume that all coefficients of the matrices $\mathbf{A}$ and $\mathbf{B}$ are nonnegative. The deterministic problem can be formulated as:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} + \mathbf{d}'\mathbf{y} \\ \text{s.t.} \quad & \mathbf{Au} + \mathbf{Bv} \leq \mathbf{x} + \mathbf{y}, \\ & \mathbf{Au} \leq \mathbf{x}, \\ & \mathbf{u} + \mathbf{v} \geq \mathbf{b}, \\ & \mathbf{x},\ \mathbf{y},\ \mathbf{u},\ \mathbf{v} \geq \mathbf{0}. \end{aligned} \tag{76}$$

The recourse function once the demand is known is solution of the following linear programming problem:

$$\begin{aligned} \mathbf{Q}(\mathbf{x}, \mathbf{u}, \mathbf{b}) = \min \quad & \mathbf{d}'\mathbf{y} \\ \text{s.t.} \quad & \mathbf{Bv} - \mathbf{y} \leq \mathbf{x} - \mathbf{Au}, \\ & \mathbf{v} \geq \mathbf{b} - \mathbf{u}, \\ & \mathbf{y},\ \mathbf{u},\ \mathbf{v} \geq \mathbf{0}. \end{aligned} \tag{77}$$

Once we have identified the worst-case demand $\tilde{\mathbf{b}}$, we solve the dual of Problem (77):

$$\begin{aligned} \max \quad & (\tilde{\mathbf{b}} - \mathbf{u})'\mathbf{p} + (\mathbf{Au} - \mathbf{x})'\mathbf{q} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{p} \leq \mathbf{B}'\mathbf{q}, \\ & \mathbf{0} \leq \mathbf{q} \leq \mathbf{d}, \end{aligned} \tag{78}$$

obtaining the optimal dual variables $\mathbf{p}^*$, $\mathbf{q}^*$, and add a cut of the type:

$$Z \geq (\mathbf{A}'\mathbf{q}^* - \mathbf{p}^*)'\mathbf{u} - (\mathbf{q}^*)'\mathbf{x} + \tilde{\mathbf{b}}'\mathbf{p}^*, \tag{79}$$

to the Benders decomposition algorithm described in Section 4.2.

We now focus on finding the worst-case demand $\tilde{\mathbf{b}}$. We obtain easily the optimal solution to Problem (77) in closed form:

$$\mathbf{Q}(\mathbf{x},\mathbf{u},\mathbf{b}) = \sum_{i=1}^{m} d_i \max\left(0, \sum_{j=1}^{n} A_{ij}\, u_j + \sum_{j=1}^{n} B_{ij}\, \max(0, b_j - u_j) - x_i\right). \tag{80}$$

Since this function is convex in $\mathbf{b}$, its maximum over $\mathcal{B}$ will be reached at an extreme point of the feasible set. Therefore, with the notations:

$$\overline{A}_i(\mathbf{x},\mathbf{u}) = \sum_{j=1}^{n} A_{ij}\, u_j + \sum_{j=1}^{n} B_{ij}\, \max(0, \overline{b}_j - u_j) - x_i,\ \forall i, \tag{81}$$

and:

$$\Delta_j(\mathbf{u}) = \max(\overline{b}_j + \widehat{b}_j, u_j) - \max(\overline{b}_j, u_j),\ \forall j, \tag{82}$$

the robust second-stage problem becomes:

$$\max_{\mathbf{z}\in\mathcal{Z}} \sum_{i=1}^{m} d_i \max\left(0, \overline{A}_i(\mathbf{x},\mathbf{u}) + \sum_{j=1}^{n} B_{ij}\Delta_j(\mathbf{u})\, z_j\right). \tag{83}$$

The key is to find which raw materials might be in shortage after the first stage, i.e., verify $\overline{A}_i(\mathbf{x},\mathbf{u}) + \sum_{j=1}^{n} B_{ij}\Delta_j(\mathbf{u})\, z_j \geq 0$ at optimality. For any $\mathbf{z}$ in $\mathcal{Z}$, we have:

$$0 \leq \sum_{j=1}^{n} B_{ij}\Delta_j(\mathbf{u})\, z_j \leq \max_{\mathbf{z}\in\mathcal{Z}} \sum_{j=1}^{n} B_{ij}\Delta_j(\mathbf{u})\, z_j. \tag{84}$$

Therefore, if $\overline{A}_i(\mathbf{x},\mathbf{u}) \geq 0$, we know that $i \in \mathcal{I}$, where $\mathcal{I}$ is the set of raw materials in shortage. Similarly, if $\overline{A}_i(\mathbf{x},\mathbf{u}) + \max_{\mathbf{z}\in\mathcal{Z}} \sum_{j=1}^{n} B_{ij}\Delta_j(\mathbf{u})\, z_j \leq 0$, then $i \notin \mathcal{I}$, in which case the index i does not affect Problem (83) (for the current first-stage solution). The remaining raw materials will be classified either by enumeration with the help of upper and lower bounds on Problem (83) to decide which cases are not worth exploring:

$$\sum_{i=1}^{m} d_i \max\{0, \overline{A}_i(\mathbf{u})\} + \max_{\mathbf{z}\in\mathcal{Z}} \sum_{i=1}^{m}\sum_{j=1}^{n} d_i B_{ij}\Delta_j(\mathbf{u}) z_j \leq \mathcal{Q}(\mathbf{x},\mathbf{u}) \leq \sum_{i=1}^{m} d_i \left(\max\{0, \overline{A}_i(\mathbf{u})\} + \max_{\mathbf{z}\in\mathcal{Z}} \sum_{j=1}^{n} B_{ij}\Delta_j(\mathbf{u}) z_j\right). \tag{85}$$

Alternatively, the lower bound can be used as an approximation to the recourse function in the master problem. It can speed up this part of the algorithm as we know that a cut will be required if Z is less than the lower bound. As a trade-off, this implies generating suboptimal cuts in the master problem.

In the numerical implementation, we consider 2 raw materials and 30 products. Demands for the products are i.i.d., with mean $\overline{b} = 10$ and maximum deviation is $\widehat{b} = 5$. All raw materials have a first-stage, resp. second-stage, ordering cost of 10, resp. 20 per unit. The productivity coefficients in matrix $\mathbf{A}$ are generated using uniform distributions in $[5, 15)$. The productivity coefficients in matrix $\mathbf{B}$ are equal to: $B_{ij} = 2\, A_{ij}$, which models the fact that producing in the second stage uses more resources than in the first one. We apply the modified Benders decomposition algorithm presented in Section 4.2 using AMPL/CPLEX v.8.1 on a Pentium IV.

Figure 3 illustrates the effect of the budget of uncertainty Γ on the expected production cost. Sample averages have been computed on a sample of 500 realizations of the demands, which we generated using independent normal random variables with mean $\bar{b} = 10$ and standard deviation $\sigma = 2$. Again, we observe a trade-off between performance and conservatism, as the expected cost first decreases when the decision-maker plans for a limited amount of uncertainty, but ultimately starts increasing as the system becomes overprotected. The optimal choice of the budget of uncertainty is $\Gamma = 4$, which matches the guidelines provided by Bertsimas and Sim in [8], as they suggest to take $\Gamma \approx \sqrt{n}$ with n the number of random variables (here, $\sqrt{30} \approx 5.5$.)

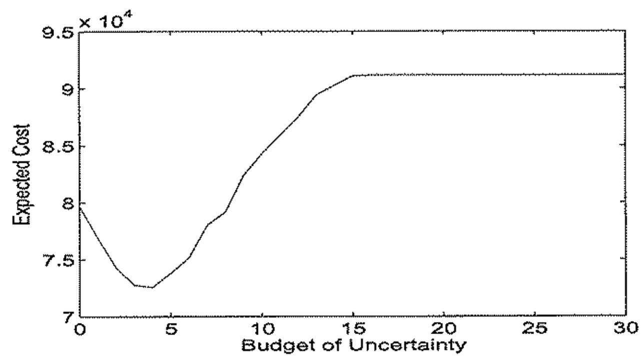

Figure 3: Impact of the budget of uncertainty on the expected cost.

Figure 4 shows the number of iterations as the budget of uncertainty increases. This number increases steadily up to 2,472 iterations, while the robust problem solved with $\Gamma = 4$ requires 947 iterations to reach optimality. A reason for the added computational burden in this problem with general recourse is that the dual feasible set of the recourse function has a large number of extreme points, which are relatively close from each other, and therefore slow the convergence of the algorithm when each cut only results in a small change in the cost.

5.3 Summary of Results

The numerical results in this section suggest that robust optimization is a promising methodology to address sequential decision-making with simple recourse, illustrated in Section 5.1. The method also performs well in the case of general recourse, presented in Section 5.2. In particular, we have observed in both experiments that protecting the system against a small amount of uncertainty will yield a smaller expected cost than not protecting the system at all, i.e., solving the deterministic problem, or protecting the system against any possible outcome, i.e., solving

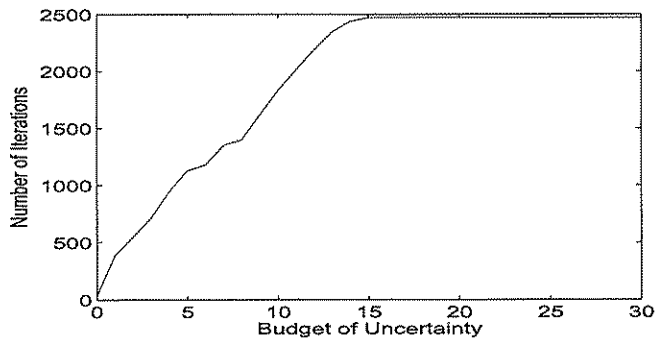

Figure 4: Number of iterations as a function of the budget of uncertainty.

the worst-case problem. This matches the intuition. The robust approach also has advantages over traditional stochastic optimization, as it does not require to solve the recourse problem for every possible scenario.

6 Conclusions

We have proposed an approach to linear optimization with recourse that is robust with respect to the underlying probabilities and can be solved efficiently using standard linear optimization software. Specifically, instead of relying on the actual distribution, which would be difficult to estimate accurately, or a family of distributions, which would significantly increase the complexity of the problem at hand, we have modelled random variables as uncertain parameters in a polyhedral uncertainty set and analyzed the problem for the worst-case instance within that set. We have shown that this robust formulation can be solved efficiently (a) by adapting Benders' decomposition, with computational advantages over the traditional stochastic framework, and (b) in the case of simple recourse, by considering a series of linear programming problems of size similar to the original deterministic problem. We have implemented this methodology to the multi-item newsvendor problem and a production planning example, with encouraging computational results.

References

[1] Atamtürk, A., M. Zhang: Two-stage robust network flow and design under demand uncertainty. Working paper, Department of Industrial Engineering and Operations Research, University of California at Berkeley, Berkeley, CA (2004)

[2] Benders, J.F.: Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* **4** 238-252 (1962)

[3] Ben-Tal, A., A. Nemirovski: Robust convex optimization. *Mathematics of Operations Research* **23** 769-805 (1998)

[4] Ben-Tal, A., A. Nemirovski: Robust solutions to uncertain programs. *Operations Research Letters* **25** 1-13 (1999)

[5] Ben-Tal, A., A. Nemirovski: Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming* **88** 411-424 (2000)

[6] Ben-Tal, A., A. Goryashko, E. Gulitzer, A. Nemirovski: Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99** 351-376 (2004)

[7] Bertsimas, D., J. Tsitsiklis: *Introduction to Linear Optimization*, Athena Scientific (1997)

[8] Bertsimas, D., M. Sim: The price of robustness. *Operations Research* **52** 35-53 (2004)

[9] Bertsimas, D., M. Sim: Robust discrete optimization and network flows. *Mathematical Programming* **98** 43-71 (2004)

[10] Bertsimas, D., A. Thiele: A robust optimization approach to inventory theory. *Operations Research*, to appear (2005)

[11] Birge, J., F. Louveaux: *Introduction to Stochastic Programming*, Springer Series in Operations Research (1997)

[12] Birge, J., R. Wets: Computing bounds for stochastic programming problems by means of a generalized moment problem. *Mathematics of Operations Research* **12** 149-162 (1987)

[13] Breton, M., S. El Hachem: Algorithms for the solution of stochastic dynamic minimax problems. *Computational Optimization and Applications* **4** 317-345 (1995)

[14] Dantzig, G.B.: Linear programming under uncertainty. *Management Science* **1** 197-206 (1955)

[15] Dupacova, J.: On minimax decision rule in stochastic linear programming, in: A. Prekopa, ed., *Studies in Mathematical Programming*, pp.47-60, Akademiai Kiado, Budapest (1980)

[16] Dupacova, J.: The minimax approach to stochastic programming and an illustrative application. *Stochastics* **20** 73-88 (1987)

[17] Dupacova, J.: Stochastic programming: minimax approach, *Encyclopedia of Optimization*, Ch.A.Floudas and P.M. Pardalos, eds., Vol.V, pp. 327-330, Kluwer (2001)

[18] El Ghaoui, L., H. Lebret: Robust solutions to least-square problems with uncertain data matrices. *SIAM Journal on Matrix Analysis and Applications* **18** 1035-1064 (1997)

[19] El Ghaoui, L., F. Oustry, H. Lebret: Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization* **9** 33-52 (1998)

[20] Ermoliev, Y., A. Gaivoronsky, C. Nedeva: Stochastic optimization problem with partially known distribution functions. *SIAM Journal on Control and Optimization* **23** 697-716 (1985)

[21] Jagannathan, R.: Minimax procedure for a class of linear programs under uncertainty. *Operations Research* **25** 173-177 (1977)

[22] Kall, P., S. Wallace: *Stochastic Programming*, Wiley (1994)

[23] Riis, M., K. A. Andersen: Applying the minimax criterion in stochastic recourse programs, Technical Report, University of Aarhus, Department of Operations Research, Aarhus, Denmark (2002)

[24] Shapiro A.: Monte-Carlo Sampling Methods, in *Stochastic Programming*, A. Ruszczynski, A., A. Shapiro (eds), volume 10 in Handbooks in Operations Research and Management Science, Elsevier (2003)

[25] Shapiro, A., S. Ahmed: On a class of minimax stochastic programs. *SIAM Journal on Optimization* **14** 1237-1249 (2004)

[26] Shapiro, A., A. Kleywegt: Minimax analysis of stochastic programs. *Optimization Methods and Software* **17** 523-542 (2002)

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

[28] Takriti, S., S. Ahmed: On robust optimization of two-stage systems. *Mathematical Programming* **99** 109-126 (2004)

[29] Takriti, S., S. Ahmed: Managing short-term electricity contracts under uncertainty: a minimax approach. Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA (2002)

[30] Zácková, J.: On minimax solution of stochastic linear programming problems. *Cas. Pest. Mat.* **91** 423-430 (1966)