Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 12–22

ROBUST COMBINATORIAL OPTIMIZATION WITH VARIABLE BUDGETED UNCERTAINTY

MICHAEL POSS

ABSTRACT: We introduce a new model for robust combinatorial optimization where the uncertain parameters belong to the image of multifunctions of the problem variables. In particular, we study the variable budgeted uncertainty, an extension of the budgeted uncertainty introduced by Bertsimas and Sim. Variable budgeted uncertainty can provide the same probabilistic guarantee as the budgeted uncertainty while being less conservative for vectors with few non-zero components. The feasibility set of the resulting optimization problem is in general non-convex so that we propose a mixed-integer programming reformulation for the problem, based on the dualization technique often used in robust linear programming. We seek how to extend these results to non-binary variables and to more general multifunctions involving uncertainty set defined by conic constraints that are affine in the problem variables. We present a computational comparison of the budgeted uncertainty and the variable budgeted uncertainty on the robust knapsack problem. The experiments show a reduction of the price of robustness by an average factor of 18%.

KEYWORDS: Robust optimization; Multifunction; Combinatorial optimization; Variable uncertainty; Chance constraint.

1. Introduction

For a large class of mathematical programs it is very hard or even impossible to compute exactly the parameters. Two main frameworks have been introduced to address this difficulty: stochastic programming and robust programming. Stochastic programming supposes that the unknown parameters are described by known random variables and replaces the deterministic constraints by chance constraints that must be satisfied with a given probability. Robust programming supposes that the unknown parameters belong to known uncertainty sets and imposes that the constraints be feasible for all values of the parameters in the uncertainty sets. The two frameworks have also been considered in a dynamic context, where subsets of the decision variables must be fixed only after part of the uncertainty has been revealed. This

Received June 14, 2012.
Michael Poss is supported by CMUC and FCT (Portugal), through European program COMPETE/FEDER within project PEst-C/MAT/UI0324/2011 and under the postdoctoral scholarship SFRH/BPD/76331/2011.

is usually known as multi-stage stochastic programming [7] and adjustable robust optimization [2], respectively. In this paper, we are interested by the static situation where all decisions must be taken before the uncertainty is revealed.

Chance constraints suffer from two main drawbacks: (i) the probability distributions of the random parameters are often impossible to describe with precision and (ii) the resulting optimization problems are very hard to solve exactly. In contrast, computing uncertainty sets for robust constraints requires less information on the parameters and, as long as these sets can be described by a conic system of constraints, the resulting optimization problems are essentially of the same difficulty as their deterministic counterparts. The tractability of robust optimization has motivated the large interest on the topic for the last ten years, see Ben-Tal et al. [2]. Between these two extremes, ambiguous chance constraints models constraints where the parameters are described by random variables which are not known exactly. Namely, the constraint must be satisfied with a given probability for all probability distributions in a given set [9].

Among the large literature on robust optimization, researchers have proposed uncertainty sets that allow a robust constraint to approximate a chance constraint or an ambiguous chance constraints in the following sense: any solution to the robust constraint will be feasible for the original chance constraint or ambiguous chance constraint. For instance, Ben-Tal and Nemirovski [4] and Berstimas and Sim [6] have introduced uncertainty sets for which a robust constraint approximates an ambiguous chance constraint where the coefficients are described by bounded random perturbations that are only assumed to be symmetrically and independently distributed around their means.

In this paper, we introduce a novel model for combinatorial optimization under uncertainty. Given a linear constraint $\sum_{i=1}^n a_i x_i \leq b$, its robust counterpart is defined as

$$\sum_{i=1}^{n} a_i x_i \leq b, \quad a \in \mathbf{U}, \tag{1}$$

where $\mathbf{U} \subset \mathbb{R}^n$ is the uncertainty set. Herein, we extend inequality (1) by considering the constraint

$$\sum_{i=1}^{n} a_i x_i \leq b, \quad a \in \mathcal{U}(x), \tag{2}$$

where $\mathcal{U} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a multifunction of x. We recall in Section 2.1 the uncertainty polytope $\mathbf{U}^\Gamma$ introduced in [6] and introduce in Section 2.2 a multifunction $\mathcal{U}^\gamma$ that generalizes $\mathbf{U}^\Gamma$. Model $\mathcal{U}^\gamma$ is motivated by the probabilistic bounds discussed in Section 3. We see that $\mathcal{U}^\gamma$ is less conservative than $\mathbf{U}^\Gamma$ while protecting the associated ambiguous chance constraint with the same probability. We provide an example showing that the feasibility set delimited by (2) is in general non-convex, which contrasts with the convexity of robust linear programs. We show in Section 4 how the classical dualization technique can be generalized to $\mathcal{U}^\gamma$. Section 5 extends our results to non-binary variables and considers the case of multifunctions described by conic constraints that are affine in the problem variables. Section 6 provides a numerical evaluation of model $\mathcal{U}^\gamma$ on the robust knapsack problem. We conclude the paper in Section 7.

In the rest of this paper, $\|x\| = \sum_{i=1}^n |x_i|$ denotes the L^1 norm and $\sum$ refers to the summation over the set $\{1, \dots, n\}$ unless stated otherwise.

2. Variable budgeted uncertainty

2.1. Static model. Combinatorial optimization problems that feature robust constraints of the type (1) have witnessed an increasing attention in recent years, see for instance [1, 13]. This success can be explained by two main reasons. First, Ben-Tal and Nemirovski [3] have shown how the infinite set of constraints (1) can be reformulated as a finite set of constraints by introducing a new set of real variables. Given a linear description of $\mathbf{U}$, this reformulation adds n linear constraints to the problem as well as a number of variables equal to the number of linear constraints that define $\mathbf{U}$; we come back to this technique in Section 4. Second, Berstimas and Sim [6] have introduced a rich class of uncertainty polytopes with a conservatism that can be regulated by parameter Γ:

$$\mathbf{U}^\Gamma := \left\{ a \in \mathbb{R}^n : a_i = \overline{a}_i + \delta_i \hat{a}_i, 0 \leq \delta_i \leq 1, \sum \delta_i \leq \Gamma \right\}. \tag{3}$$

Clearly, increasing Γ increases the size of $\mathbf{U}^\Gamma$ and thus, the conservatism of the approach. For instance, if $\Gamma \geq n$, then all components of a can take simultaneously their peak values and the robust constraint become $\sum(\overline{a}_i + \hat{a}_i)x_i \leq b$. On the opposite side, if $\Gamma = 0$, all components of a are equal to their non-peak values and the robust constraint becomes $\sum \overline{a}_i x_i \leq b$. Varying Γ between 0 and n enables Berstimas and Sim [6] to define a wide variety of uncertainty polytopes.

2.2. Variable model. The uncertainty sets proposed in [6] suffer from a practical drawback: they are independent from the value of x. Because of this, binary vectors with few non-zero components are more protected than binary vectors with larger numbers of non-zero components. For instance, consider two binary vectors x^1 and x^2 feasible for constraint (1) and suppose that $\|x^1\| = \Gamma$ while $\|x^2\| = 2\Gamma$. The robust constraints associated to x^1 and x^2 are

$$\sum_{i:x_i^1=1} a_i \leq b, \quad a \in \mathbf{U}^\Gamma, \tag{4}$$

and

$$\sum_{i:x_i^2=1} a_i \leq b, \quad a \in \mathbf{U}^\Gamma, \tag{5}$$

respectively. In a relative point of view, vector x^1 is more protected than vector x^2 since it is ensured that constraint (4) is feasible against the simultaneous perturbation of all of its terms while constraint (5) is only protected against the simultaneous perturbation of half of its terms. This relative point of view has a natural probabilistic interpretation. If $\tilde{a}_i$ are random variables distributed between $\overline{a}_i - \hat{a}_i$ and $\overline{a}_i + \hat{a}_i$, then the probability that constraint $\sum_{i:x_i^1=1} \tilde{a}_i \leq b$ be violated is always zero while the probability that $\sum_{i:x_i^2=1} \tilde{a}_i \leq b$ be violated can be strictly positive for particular choices of b and probability distributions.

To avoid this conservatism, we introduce in this paper a novel model of uncertainty. Instead of considering an uncertainty set $\mathbf{U} \subseteq \mathbb{R}^n$ independent of x, we introduce a multifunction of x (point-to-set mapping) $\mathcal{U} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. For each value of x, uncertainty set $\mathcal{U}(x) \subseteq \mathbb{R}^n$ contains all feasible values for the uncertain parameters a. We consider in particular multifunctions that are generalizations of the budgeted uncertainty. Given a non-negative function $\gamma : \mathbb{R}^n \to \mathbb{R}_+$, we define the variable budgeted uncertainty as

$$\mathcal{U}^\gamma(x) := \left\{ a \in \mathbb{R}^n : a_i = \overline{a}_i + \delta_i \hat{a}_i, 0 \leq \delta_i \leq 1, \sum \delta_i \leq \gamma(x) \right\}. \tag{6}$$

If γ is constantly equal to Γ, then $\mathcal{U}^\gamma(x)$ coincide with $\mathbf{U}^\Gamma$ for any x. In general however, $\mathcal{U}^\gamma$ enables us to avoid to overprotect vectors with few components, yielding a less conservative model than $\mathbf{U}^\Gamma$. The pendant of (1) for the variable budgeted uncertainty is

$$\sum a_i x_i \leq b, \quad a \in \mathcal{U}^\gamma(x). \tag{7}$$

3. Probabilistic bounds

Using the probabilistic bounds derived in [6], we show in this section how $\mathcal{U}^\gamma$ enables us to guarantee exactly the same protection level for every $x \in \{0,1\}^n$. The following notation is used throughout: $\tilde{a}_i = \overline{a}_i + \eta_i \hat{a}_i$ denotes the random variable associated with parameter a_i and η_i, $i = 1, \dots, n$, are arbitrary random variables independently and symmetrically distributed in $[-1,1]$. Given any vector x^* that satisfies the robust constraint (1) for $\mathbf{U}^\Gamma$, Berstimas and Sim [6] prove that

$$P\left(\sum \tilde{a}_i x_i^* > b\right) \leq \exp\left(-\frac{\Gamma^2}{2n}\right). \tag{8}$$

We adapt below their bound to variable budgeted uncertainty. First we need to make bound (8) dependent on x.

Lemma 1. *Let x^* be a binary vector that satisfies the robust constraint* (1) *for $\mathbf{U}^\Gamma$. It holds that*

$$P\left(\sum \tilde{a}_i x_i^* > b\right) \leq \exp\left(-\frac{\Gamma^2}{2\|x^*\|}\right).$$

Proof: If $\|x^*\| \leq \Gamma$ then $P\left(\sum \tilde{a}_i x_i^* > b\right) = 0$. Hence, suppose that $\|x^*\| > \Gamma$, so that x^* satisfies the robust constraint

$$\sum_{i:x_1^*=1} a_i x_i \leq b, \quad a \in \mathbf{U}^\Gamma. \tag{9}$$

Then,

$$\begin{aligned} P\left(\sum_{i=1}^n \tilde{a}_i x_i^* > b\right) &= P\left(\sum_{i:x_1^*=1} \tilde{a}_i x_i^* > b\right) \\ &\leq \exp\left(-\frac{\Gamma^2}{2\|x^*\|}\right), \end{aligned} \tag{10}$$

where (10) follows from Proposition 2 and Theorem 2 from [6] applied to vector x^* that satisfies robust constraint (9), which contains with $\|x^*\|$ terms. ■

Define the non-negative function $\alpha_\epsilon(x) = (-2\ln(\epsilon)\|x\|)^{1/2}$. The bound from Lemma 1 implies that any binary vector x^* that satisfies a robust constraint for uncertainty set $\mathbf{U}^{\alpha_\epsilon(x^*)}$ will also satisfy the chance constraint with probability $1-\epsilon$. Because $\mathbf{U}^{\alpha_\epsilon(x^*)} = \mathcal{U}^{\alpha_\epsilon}(x^*)$, this result can be applied to variable budgeted uncertainty.

Corollary 1. *Let x^* be a binary vector and consider $\epsilon \in (0,1)$. If x^* satisfies constraint* (7) *with $\gamma(x^*) = \alpha_\epsilon(x^*)$, then $P\left(\sum \tilde{a}_i x_i^* > b\right) \le \epsilon$.*

The interest of Corollary 1 lies in the fact that $\alpha_\epsilon(x)$ is an increasing function of $\|x\|$. Then, taking x^1 and x^2 such that $\|x^1\| \le \|x^2\|$, we have that $\mathcal{U}^{\alpha_\epsilon}(x^1) \subseteq \mathcal{U}^{\alpha_\epsilon}(x^2)$. Hence, $\mathcal{U}^{\alpha_\epsilon}$ enables us to impose that x^1 be protected against a smaller uncertainty set than x^2, and the uncertainty sets are tailored in such a way that the probability of violating the constraint is at most ϵ for both x^1 and x^2. In contrast, the classical budgeted uncertainty would have to protect both x^1 and x^2 against the variations of a in the uncertainty set $\mathbf{U}^{\alpha_\epsilon(n)}$.

Berstimas and Sim [6] mention that bound (8) is not very tight. For this reason, they introduce more complex bounds that provide tighter approximations of the probability $P\left(\sum \tilde{a}_i x_i^* > b\right)$. The strongest of these bounds states that any vector x^* that satisfies the robust constraint (1) for $\mathbf{U}^\Gamma$ satisfies

$$P\left(\sum \tilde{a}_i x_i^* > b\right) \le B(n,\Gamma) = \frac{1}{2^n}\left((1-\mu)\binom{n}{\lfloor \nu \rfloor} + \sum_{l=\lfloor \nu \rfloor+1}^{n} \binom{n}{l}\right), \quad (11)$$

where $\nu = (\Gamma+n)/2$, $\mu = \nu - \lfloor \nu \rfloor$. Their experiments show that the bound provided by $B(n,\Gamma)$ is one order of magnitude smaller than bound (8) for $n = 100$ and $n = 2000$.

Using the same reasoning as the one used for Lemma 1, we can make this bound dependent of $\|x^*\|$ by considering the subconstraint that contains only terms where x^* is different from zero.

Lemma 2. *Let x^* be a binary vector that satisfies the robust constraint* (1) *for $\mathbf{U}^\Gamma$. It holds that $P\left(\sum \tilde{a}_i x_i^* > b\right) \le B(\|x^*\|,\Gamma)$.*

To use this bound in the context of $\mathcal{U}^\gamma$, we need to solve the following equation in variable Γ

$$B(\|x^*\|, \Gamma) - \epsilon = 0. \tag{12}$$

The unicity of solutions to equation (12) follows from the following property of $B(n, \Gamma)$.

Lemma 3. *Function $B(n, \Gamma)$ is strictly decreasing in Γ.*

Proof: Let $\delta > 0$ be small enough so that $\left\lfloor \frac{n+\Gamma+\delta}{2} \right\rfloor = \left\lfloor \frac{n+\Gamma}{2} \right\rfloor$. Then,

$$B(n, \Gamma + \delta) - B(n, \Gamma) = -\frac{\delta}{2^n} \binom{n}{\lfloor \nu \rfloor} < 0.$$

■

Because $B(n, \Gamma)$ is strictly decreasing in Γ, equation (12) has at most one solution for all $B(n, \Gamma)$, which we denote $\beta_\epsilon(x)$. However, the equation does not always have a solution, that is, $\beta_\epsilon(x)$ is not defined for all $x \in \{0, 1\}^n$ and $\epsilon > 0$. Taking for instance x^* with only one component equal to one (i.e. $\|x\| = 1$), the minimum of $B(\|x^*\|, \Gamma)$ is equal to 0.5 and is obtained for $\Gamma = 1$. Hence, $\beta_\epsilon(x^*)$ is not defined for $\epsilon < 0.5$.

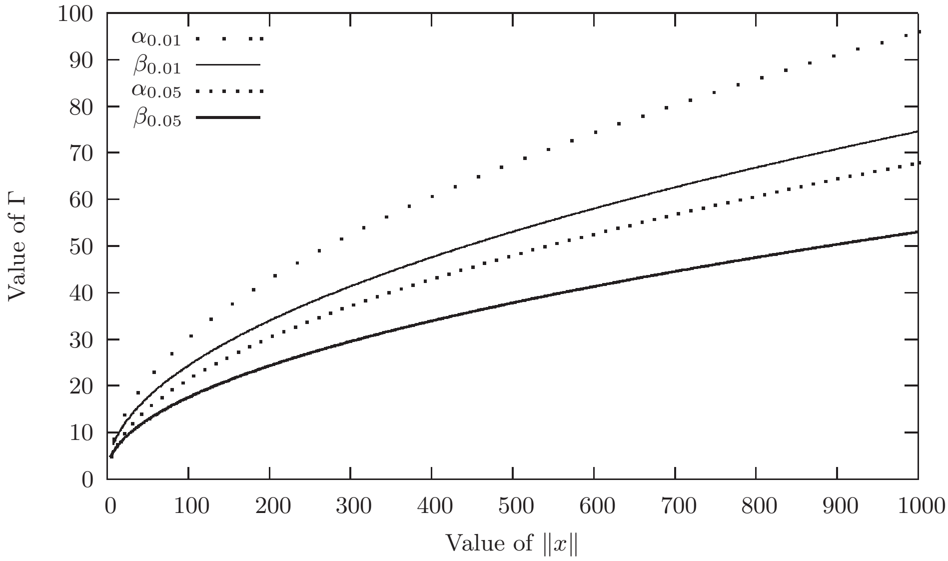
In practice, $\beta_\epsilon(x)$ is defined for relatively small values of $\|x\|$. Because $B(n, \Gamma)$ is decreasing in Γ, we can solve equation (12) to the required precision by using a dichotomic search. Namely, we first evaluate $B(n, n/2)$. If $B(n, n/2) > \epsilon$, then we know that $\beta_\epsilon(x) \in [0, n/2]$. Otherwise, this means that $\beta_\epsilon(x) \in [n/2, n]$. We can proceed this way up to the required precision on Γ. We have computed $\beta_{0.01}$ and $\beta_{0.05}$ numerically up to a precision of 0.01. These computations have shown that equation (12) has a solution for ϵ equal to 0.01 and 0.05 when $\|x^*\|$ is greater than or equal to 8 and 5, respectively.

Whenever $\beta_\epsilon(x)$ is defined, it can be used to ensure that the probabilistic constraint is satisfied.

Corollary 2. *Let x^* be a binary vector and consider $\epsilon \in (0, 1)$ so that $\beta_\epsilon(x^*)$ is well-defined. If x^* satisfies the robust constraint (7) with $\gamma(x^*) = \beta_\epsilon(x^*)$ then $P\left(\sum \tilde{a}_i x_i^* > b\right) \leq \epsilon$.*

We compare in Figure 1 the values of the two bounds discussed previously for ϵ equal to 0.01 and 0.05. We see that $\beta_{0.01} \leq \alpha_{0.01}$ and $\beta_{0.05} \leq \alpha_{0.05}$ for all values of $\|x\|$ depicted in the figure. This was expected because in [6], it was observed that $B(n, \Gamma)$ is tighter than bound (8). For this reason, we focus on β_ϵ in the rest of the paper.

Figure 1. Comparison of the two bounds.



We present in Section 6 numerical results for the knapsack problem showing that the use of $\mathcal{U}^{\beta_\epsilon}$ reduces the cost of protecting the solution with probability 0.99 or 0.95 by 18% on average. More complex problems may witness more important cost reductions. For instance, the protection cost is reduced to zero for the problem described in Example 1.

Example 1. *Let a_i be random variables independently and symmetrically distributed in $[-1, 1]$ and let m and n be any integers such that (i) $m \leq \beta_{0.01}(n)$ and (ii) $\beta_{0.01}(m) \leq m-1$. Consider then the following combinatorial optimization problem:*

$$\begin{aligned} \max \quad & c\,x_m \\ s.t. \quad & x_{i+1} \leq x_i, \qquad\qquad i = 1, \ldots, n-1 \\ & \sum_{i=1}^{n} a_i x_i \leq m-1, \\ & x \in \{0,1\}^n. \end{aligned}$$

The deterministic version of the problem replaces the random variables a_i by their mean values, 0, yielding the solution $x_i^* = 1$ for $1 \leq i \leq m$ and $x_i^* = 0$ otherwise, with objective value c. Using uncertainty models $\mathbf{U}^{\beta_{0.01}(n)}$ and $\mathcal{U}^{\beta_{0.01}}$, we can ensure that constraint $\sum_{i=1}^{n} \tilde{a}_i x_i \leq m-1$ is satisfied with probability 0.99 by any solution to the robust models. The optimal solution *costs to models* $\mathbf{U}^{\beta_{0.01}(n)}$ *and* $\mathcal{U}^{\beta_{0.01}}$ *are 0 and c, respectively. Therefore, the protection cost for model* $\mathbf{U}^{\beta_{0.01}(n)}$ *is equal to c while there is no protection cost for model* $\mathcal{U}^{\beta_{0.01}}$.

In this section, we have motivated the introduction of our new model, $\mathcal{U}^\gamma$. Given a function γ properly chosen, the model can be less conservative than $\mathbf{U}^\Gamma$ while ensuring the required level of protection. In spite of this good news, using model $\mathcal{U}^\gamma$ is more complex than using model $\mathbf{U}^\Gamma$. This arises from the following observation. A robust constraint subject to a non-empty uncertainty polytope, such as $\mathbf{U}^\Gamma$, can always be rewritten as a finite set of linear constraints. This reformulation makes possible to solve a large class of robust combinatorial optimization problems efficiently. The situation is more complex in the case of budgeted variable uncertainty $\mathcal{U}^\gamma$. We show in the example below that the feasible region of the vectors that satisfy (7) is in general non-convex.

Example 2. *Consider the feasibility set of a linear constraint in two variables*

$$\mathbf{X} := \{x \in \mathbb{R}^2 \ s.t. \ a_1x_1 + a_2x_2 \leq 1 \ for\ all\ a \in \mathcal{U}^\gamma(x)\}.$$

where

$$\mathcal{U}^\gamma(x) := \left\{a \in \mathbb{R}^2 : a = (1,1) + (2,0)\delta, 0 \leq \delta \leq 1, \delta_1 + \delta_2 \leq x_1 + x_2\right\},$$

Set $\mathbf{X}$ *is non-convex because* $x^1 = (0.5, 0) \in \mathbf{X}$, $x^2 = (0,1) \in \mathbf{X}$ *and* $0.5x^1 + 0.5x^2 \notin \mathbf{X}$.

In view of Example 2, combinatorial optimization problems that present robust variable constraints (7) belong to the general class of non-convex Mixed-Integer Non-Linear Programming. Although some progress has been made for that type of problems [8], they remain very hard to solve to optimality in general, especially when the constraints do not define a known structure. Nevertheless, we show in the next section how the dualization technique can be applied to $\mathcal{U}^\gamma$. To simplify the notations, we often omit the index ϵ in the rest of the paper.

4. Dualization

We recall in this section the classical dualization technique used in robust linear programming and show how it extends to the case of variable uncertainty. The method described below requires that function γ involved in the definition of $\mathcal{U}^\gamma$ be an affine function of x. Recall, however, that neither α

nor β are affine functions so that we can not apply our method directly to $\mathcal{U}^\beta$. Therefore, we define below another class of multifunctions that enables us to approximate $\mathcal{U}^\beta$ by using only affine functions of x.

Let $\gamma^1, \dots, \gamma^m$ be affine functions of x defined by $\gamma^j(x) = \gamma_0^j + \sum \gamma_i^j x_i$. For each $x \in \{0,1\}^n$, the set $\mathcal{U}^{\gamma^1 \dots \gamma^m}(x)$ contains vectors $a \in \mathbb{R}^n$ such that $a_i = \overline{a}_i + \delta_i \hat{a}_i$ and

$$\begin{aligned} 0 \ \leq \ \delta_i \ &\leq \ 1, && i = 1, \dots, n && (13)\\ \sum \delta_i \ &\leq \ \gamma^j(x), && j = 1 \dots, m. && (14) \end{aligned}$$

In what follows, we will use multifunction $\mathcal{U}^{\gamma^1 \dots \gamma^m}$ as an approximation of $\mathcal{U}^\beta$. To ensure that $\mathcal{U}^{\gamma^1 \dots \gamma^m}$ yields the same probabilistic guarantee as $\mathcal{U}^\beta$, functions γ^j, $j = 1, \dots, m$, must be greater than or equal to β for all $x \in \{0,1\}^n$.

Lemma 4. *Let $\gamma^1, \dots, \gamma^m$ be affine functions of x such that $\gamma^j(x) \geq \beta(x)$ for all $x \in \{0,1\}^n$. If $x^* \in \{0,1\}^n$ satisfies the robust constraint $\sum a_i x_i^* \leq b$, for all $a \in \mathcal{U}^{\gamma^1 \dots \gamma^m}(x)$, then $P(\sum \tilde{a}_i x_i^* > b) \leq \epsilon$.*

Proof: Since $\gamma^j(x^*) \geq \beta(x^*)$ for each $j = 1, \dots, m$, it holds that $\mathcal{U}^{\gamma^1 \dots \gamma^m}(x^*) \subseteq \mathcal{U}^\beta(x^*)$. Hence, the constraint $\sum a_i x_i^* \leq b$ is satisfied for all $a \in \mathcal{U}^\beta(x^*)$ and Corollary 2 implies the result. ■

The next result shows to handle the upper approximation provided by $\gamma^1, \dots, \gamma^m$.

Theorem 1. *Consider robust constraint*

$$\begin{aligned} a^T x \ &\leq \ b, && a \in \mathcal{U}^{\gamma^1 \dots \gamma^m}(x), \\ x \ &\in \ \{0,1\}^n, && \end{aligned} \qquad (15)$$

and suppose that $\gamma^1, \dots, \gamma^m$ *are affine functions of* x*, non-negative for* $x \in \{0,1\}^n$*. Then,* (15) *is equivalent to*

$$\sum_{i=1}^{n} \overline{a}_i x_i + \sum_{j=1}^{m} \left(\gamma_0^j z_j + \sum_{i=1}^{n} \gamma_i^j w_{ji} \right) + \sum_{i=1}^{n} p_i \le b \tag{16}$$

$$\sum_{j=1}^{m} z_j + p_i \ge \hat{a}_i x_i, \qquad i = 1, \dots, n, \tag{17}$$

$$w_{ji} - z_j \ge -\max_j(\hat{a}_j)(1 - x_i), \quad i = 1, \dots, n, \quad j = 1, \dots, m, \tag{18}$$

$$p, w, z \ge 0. \tag{19}$$

$$x \in \{0,1\}^n. \tag{20}$$

Proof: The proof works in two steps. The first step applies the dualization technique from [3]. Let p and z be the dual multipliers associated to constraints (13) and (14), respectively. For any $x \in \{0,1\}^n$, $\gamma^j(x) \ge 0$ so that $\mathcal{U}^{\gamma^1 \dots \gamma^m}(x)$ is non-empty and bounded. Hence, strong duality in linear programming implies that (15) is equivalent to

$$\begin{aligned}
\sum_{i=1}^{n} \overline{a}_i x_i + \sum_{j=1}^{m} \gamma^j(x) z_j + \sum_{i=1}^{n} p_i \ &\le \ b && (21)\\
\sum_{j=1}^{m} z_j + p_i \ &\ge \ \hat{a}_i x_i, \quad i = 1, \dots, n, && (22)\\
p, z \ &\ge \ 0, && (23)\\
x \ &\in \ \{0,1\}^n. && (24)
\end{aligned}$$

We are left to linearize the bilinear terms in (21): $\gamma(x)^j z_j = (\gamma_0^j + \sum \gamma_i^j x_i) z_j$. Introducing nm real variables w_{ji} to represent products $x_i z_j$, constraint (21) for x binary can be rewritten as

$$\sum_{i=1}^{n} \overline{a}_i x_i + \sum_{j=1}^{m} \left(\gamma_0 z_j + \sum_{i=1}^{n} \gamma_i w_{ji} \right) + \sum_{i=1}^{n} p_i \leq b \tag{25}$$

$$w_{ji} - z_j \geq -M(1 - x_i), \quad i = 1, \dots, n, \quad j = 1, \dots, m, \tag{26}$$

$$w \geq 0, \tag{27}$$

where M is a constant large enough. Because each z_j must satisfies constraints (22), M may be as large as $\max_i \hat{a}_i$. Constraint (25) does not impose additional restriction on the minimal value of z_j, so that we can choose M equal to $\max_i \hat{a}_i$. Regrouping constraints (22)–(27) yields the result. ■

Some care must be taken when choosing functions $\gamma^1, \dots, \gamma^m$. If β were a concave and differentiable function defined for all $x \in [0,1]^n$, its differentials would provide the best over-approximating affine functions. Unfortunately, β is defined only for $x \in \{0,1\}^n$ so that no differential is available. Moreover, the finite differences of step equal to 1 may not be sufficient because there exists $x^*, x' \in \{0,1\}^n$ with $\|x^* - x'\| = 1$ such that

$$\beta(x') > \beta(x^*) + (x' - x^*)(\beta(x') - \beta(x^*)).$$

Therefore, we define the angular coefficient of each function γ^j by using $x^*, x' \in \{0,1\}^n$ with $\|x^* - x'\| > s$ for some step s. In our experiments, we use s equal to 5. Notice that because $\beta(x)$ only depends $\|x\|$, we only use affine functions where the coefficients of all variables are equal. Then, to ensure that for each $x \in \{0,1\}^n$, $\gamma^j(x)$ is greater than or equal to $\beta(x)$, we add the constant term

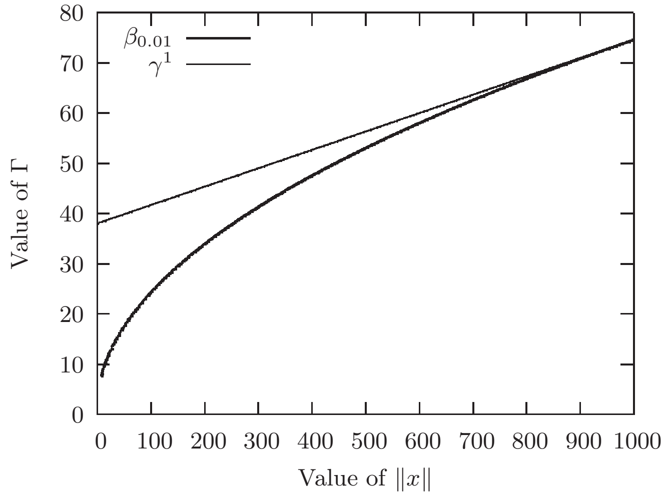
$$\max_{x \in \{0,1\}^n} |\beta(x) - \gamma^j(x)|. \tag{28}$$

Problem (28) is solved by enumeration over $\|x\| \in \{1, \dots, n\}$. We present in Figure 2 examples of piece-wise affine over-approximations of β for $\epsilon = 0.01$.

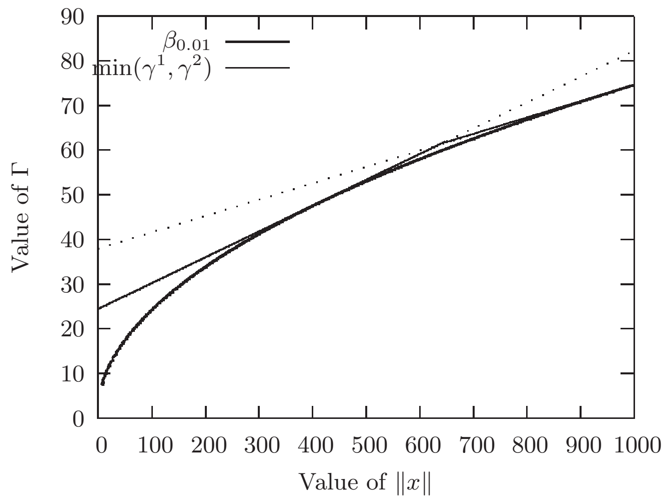
5. Extensions

5.1. Non-binary variables. We show in this subsection how to extends the results developed in Sections 3 and 4 to a robust constraint that contains

FIGURE 2. Approximating β.



(a) one piece-wise linear function



(b) two piece-wise linear functions

bounded real or integer variables, in addition to the binary variables considered so far. Recall that our motivation for introducing multifunctions $\mathcal{U}^\alpha$ and $\mathcal{U}^\beta$ arise from the probabilistic bounds computed in [6]: these bounds depend on the number of non-zero elements in a robust constraint, which we adapt to variable uncertainty with binary variables by using the L^1 norm $\| * \|$.

To count the number of non-zeroes for real or integer variables, we need to use a function different from $\| * \|$. Given an arbitrary vector $x^* \in \mathbb{R}^n$, let $\mathbf{1}^{x^*}$ be the binary vector whose i-ith coordinate is equal to 0 if x^*_i is equal to

0 and equal to 1 otherwise. We define $\|x^*\|_0 := \|\mathbf{1}^{x^*}\|$, sometimes abusively called the L^0 norm. This function enables us to extend α_ϵ and β_ϵ to values of x^* different from zero and one: $\alpha^0_\epsilon(x^*) = (-2\ln(\epsilon)\|x^*\|_0)^{1/2}$ and $\beta^0_\epsilon(x^*)$ is the unique solution to equation $B(\|x^*\|_0, \Gamma) - \epsilon = 0$ in variable Γ (if the solution exists). The values of $\|x\|_0$ and $\|x\|$ coincide for any binary vector x so that functions α_ϵ and α^0_ϵ (resp. β_ϵ and β^0_ϵ) coincide on $\{0,1\}^n$. We see immediately that the results from Section 3 extend to α^0_ϵ and β^0_ϵ. Namely, Corollary 2 becomes the proposition below, and we can similarly adapt Corollary 1.

Proposition 1. *Let x^* be a vector in $\mathbb{R}^n$ and consider $\epsilon \in (0,1)$ so that $\beta^0_\epsilon(x^*)$ is well-defined. If x^* satisfies the robust constraint* (7) *with $\gamma(x^*) = \beta^0_\epsilon(x^*)$ then $P\left(\sum \tilde{a}_i x^*_i > b\right) \leq \epsilon$.*

The reformulation from Section 4 can also be extended to bounded real or integer variables. By assumption, we know that there exists a positive real M large enough so that any vector x^* feasible for our problem belongs to $B_M(0)$, the ball centered at the origin of radius M. Hence, $\|x\|_0$ can be expressed by introducing an auxiliary binary vector y, equal to $\mathbf{1}^{x^*}$:

$$x_i \leq My_i, \qquad i = 1, \ldots, n, \tag{29}$$

$$y \in \{0,1\}^n. \tag{30}$$

To extend Theorem 1 to non-binary variables and avoid products of real variables, we need to restrict ourselves to affine functions γ^j that depend only on $r := \|x\|_0$: $\gamma^j(x) = \gamma^j_0 + \gamma^j_1\|x\|_0$. We present below an extension of Theorem 1 where it is supposed that the affine functions depend on a unique variable r. This restricted version of the result does not prevent us from approximating α^0 and β^0 with affine functions because α^0 and β^0 can be expressed as functions of $r := \|x\|_0$. Similarly, we mentioned already that α and β can be expressed as functions of $r := \|x\|$. The proof of the next result is essentially the same as the proof of Theorem 1.

Theorem 2. *Consider robust constraint*

$$\begin{aligned} a^T x &\leq b, \qquad a \in \mathcal{U}^{\gamma^1 \ldots \gamma^m}(x), \\ x &\in B_M(0), \end{aligned} \tag{31}$$

and suppose that $\gamma^1, \dots, \gamma^m$ *are affine functions of* $r := \|x\|_0$*, non-negative for all* $x \in B_M(0)$*. Then,* (31) *is equivalent to*

$$\begin{aligned}
\sum_{i=1}^n \overline{a}_i x_i + \sum_{j=1}^m \left(\gamma_0^j z_j + \gamma_1^j \sum_{i=1}^n w_{ji} \right) + \sum_{i=1}^n p_i \le b \qquad\qquad & \\
w_{ji} - z_j \ge - \max_j(\hat{a}_j)(1 - y_i), \quad & i = 1, \dots, n, \\
& j = 1, \dots, m, \\
x \in B_M(0), \quad & \\
(17), (19), (29), (30). \qquad\qquad &
\end{aligned}$$

5.2. Conic uncertainty sets. So far we have only considered a very particular multifunction, $\mathcal{U}^\gamma$, which can be seen as a generalization of the budgeted uncertainty set from [6], $\mathbf{U}^\Gamma$. This is motivated by the probabilistic bounds that Berstimas and Sim [6] have provided for $\mathbf{U}^\Gamma$, and the fact that these bounds extend to $\mathcal{U}^\gamma$ for proper choices of γ. However, the concept of variable uncertainty could by applied to more general multifunctions $\mathcal{U}$. For instance, the specific knowledge of a practical problem could motivate the use of a multifunction different from $\mathcal{U}^\gamma$, justified by the application.

The aim of this subsection is to show that Theorem 1 can be generalized to multifunctions whose images are sets defined by conic inequalities that depend affinely on x. More precisely, we consider again a robust constraint

$$a^T x \le b, \quad (a, b) \in \mathcal{U}(x), \tag{32}$$

and we suppose that the uncertainty set is defined by

$$\mathcal{U}(x) := \{(a, b) = (a^0, b^0) + \sum_{\ell=1}^L (a^\ell, b^\ell)\zeta_\ell : \zeta \in \mathcal{Z}(x)\},$$

where the perturbation set $\mathcal{Z}(x)$ is given by the conic representation

$$\mathcal{Z}(x) := \{\zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^K : P(x)\zeta + Q(x)u + p(x) \in \mathbf{K}\}, \tag{33}$$

where $\mathbf{K}$ is closed convex pointed cone in $\mathbb{R}^N$ with nonempty interior, and for all x, $P(x), Q(x)$ are given matrices and $p(x)$ is a given vector. In the case where $\mathbf{K}$ is not a polyhedral cone, we assume that the Slater's condition holds for all $x \in \{0,1\}^n$, see for instance [2]. As in the classical case, where $P(x), Q(x)$ and $p(x)$ are constant, we can apply the strong duality of conic programming and replace (32) by a finite set of conic constraints, see for instance [5] for a proof.

Proposition 2. *The robust constraint* (32) *can be represented by the following system of conic inequalities in variables* $x \in \mathbb{R}^n, y \in \mathbb{R}^N$*:*

$$p^T(x)y + (a^0)^T \le b^0 \tag{34}$$

$$Q^T(x)y = 0 \tag{35}$$

$$(P^T(x)y)_\ell + (a^\ell)^T x = b^\ell, \ \ell = 1, \dots, L, \tag{36}$$

$$y \in \mathbf{K}_*,$$

where $\mathbf{K}_* = \{y : y^T z \ge 0 \; \forall \; z \in \mathbf{K}\}$ *is the cone dual to* $\mathbf{K}$*.*

If $p(x), P(x)$, and $Q(x)$ are affine functions of x, the products of variables that appear in constraints (34)–(36) can be linearized with the help of big-M coefficients, yielding a mixed-integer conic reformulation for constraint (32) that generalizes the reformulation of Theorem 1.

6. Computational experiments

This section studies two numerical aspects of the robust knapsack problem under uncertainty model $\mathcal{U}^\gamma$. First, we compare the prices of robustness of $\mathcal{U}^\gamma$ and $\mathbf{U}^\Gamma$: we show that the cost of protecting the capacity constraint with a probability of 0.99 or 0.95 is strictly less for $\mathcal{U}^\gamma$ than for $\mathbf{U}^\Gamma$. Second, we study the computational complexity of $\mathcal{U}^\gamma$. We perform this comparison for the binary knapsack problem as well as for its linear relaxation. We performed our experiments on a computer equiped with a processor Intel Core i5 at 2.53 GHz and 4 GB of RAM memory and calling CPLEX 12.1 in JAVA via Concert Technology [11].

Given a set of n items, each with profit p_i and weight a_i, the knapsack problem aims at choosing a subset of these items not exceeding the available capacity b and maximizing the profit:

$$\max \quad \sum_{i=1}^n p_i x_i$$

$$\text{s.t.} \quad \sum_{i=1}^n a_i x_i \le b, \tag{37}$$

$$x \in \{0,1\}^n. \tag{38}$$

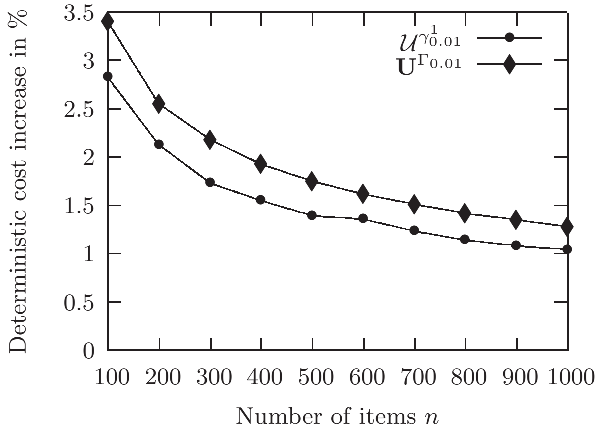
Although $\mathcal{NP}$-hard to solve exactly, state of the art MIP solvers can easily solve instances of the knapsack problem with thousands of variables. This problem is central to integer programming because many complicate integer programs feature capacity constraints like (37). For this reason, the problem is often used to benchmark new approaches for integer linear programs in the uncertain context. In stochastic programming, [12] use the knapsack problem to test their sample average approximation algorithm and [10] use the problem to test their simple-recourse reformulations and evaluate its complexity. In robust programming, [6] use the problem to evaluate the cost of protecting the capacity constraint for various probability guarantees.

To evaluate our new model $\mathcal{U}^{\beta}$, we generate our instances similarly to [6]. We consider different item numbers $n \in \{100, 200, \ldots, 1000\}$ and set the capacity to $b = 20n$ for each value of n. For each value of n, we generate randomly five instances as follows. For each $i = 1, \ldots, n$, the average weight $\overline{a}_i$ is chosen uniformly from the set $\{20, 21, \ldots, 29\}$, the deviation $\hat{a}_i$ is equal to 10% of $\overline{a}_i$, and the profit p_i is chosen uniformly from the set $\{16, 17, \ldots, 77\}$.

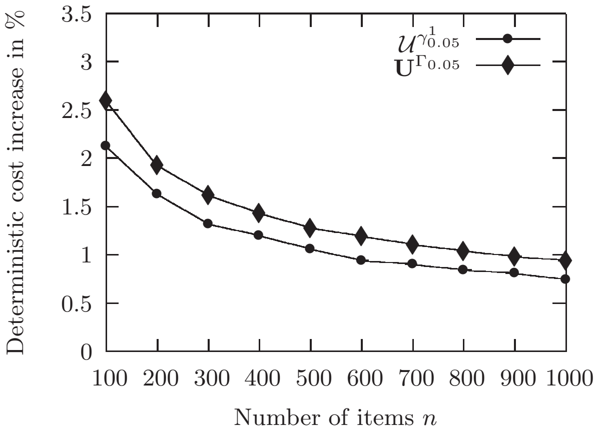
We compare in Figure 3 the optimal protection costs of the different models with binary variables. For each value of n, we compute $\Gamma_\epsilon = \beta_\epsilon(n)$. Then, we compute the affine function γ_ϵ that overestimate β_ϵ as described in Section 4. Let $p(det)$, $p(\mathbf{U}^{\Gamma_\epsilon})$, and $p(\mathcal{U}^{\gamma_\epsilon})$ denote the optimal solution costs to, respectively, the deterministic model and the robust models where constraint (37) must be satisfied for all values of a in $\mathbf{U}^{\Gamma_\epsilon}$ or $\mathcal{U}^{\gamma_\epsilon}$. We compute the cost $c(*)$ of protecting a solution with a given probability for model $*$ as $c(*) = \frac{p(*)-p(det)}{p(det)}$. We present in Figure 3 the geometric means of these protection costs for each value of n. On average, $c(\mathcal{U}^{\gamma_\epsilon})$ is 18% less than $c(\mathbf{U}^{\Gamma_\epsilon})$.

We compare then the computational complexity of models $\mathcal{U}^{\gamma_\epsilon}$ and $\mathbf{U}^{\Gamma_\epsilon}$. Let $t(*)$ be the solution time in seconds to solve model $*$ to optimality. The solution time was less than 10 second for any of our instances. For $\epsilon = 0.01$, the geometric mean of the ratios $t(\mathcal{U}^{\gamma_{0.01}})/t(\mathbf{U}^{\Gamma_{0.01}})$ is equal to 1.7, with a maximum value of 7.68. For $\epsilon = 0.05$, these values increase to 2.5 and 10, respectively. The ratios do not increase with the problem size. In addition to the solution times, we investigate the bound provided by the linear relaxation of the problem. In particular, we want to understand whether the bound proposed in Theorem 1 for the big-M coefficients is tight. Our results show that the gap between the linear relaxation and the solution of the problem are very close for both models. In fact, the gap of model $\mathcal{U}^{\gamma_\epsilon}$ is 10% better in average than the one of $\mathbf{U}^{\Gamma_\epsilon}$. However, replacing $\max_j(\hat{a}_j)$ by very large

FIGURE 3. Cost of protecting the solution for the knapsack problem.



(a) $\epsilon = 0.01$



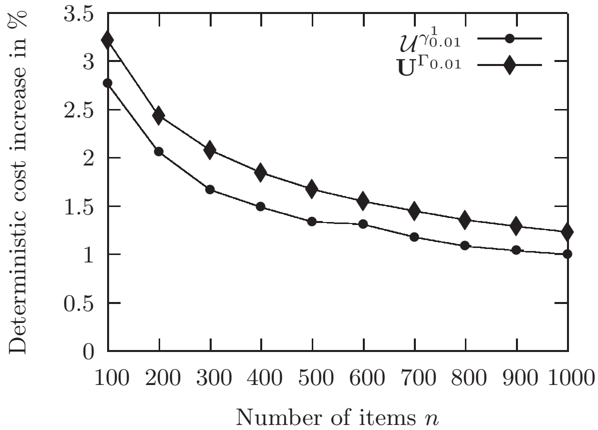
(b) $\epsilon = 0.05$

numbers reduce significantly the bound provided by the linear relaxation. For instance, setting M to 10000 multiplies the gap by an average factor of four.

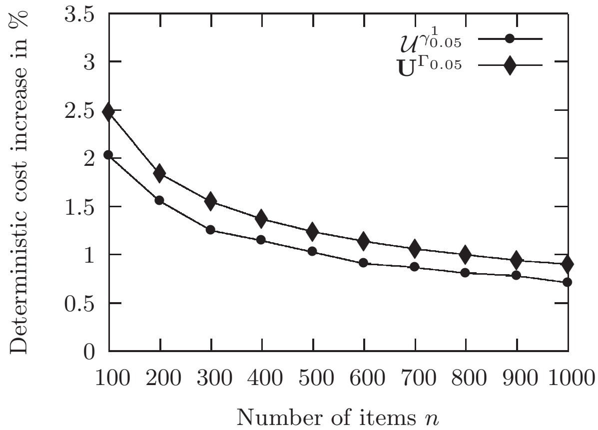
We have also tested more refined linearizations, using two and three linear over-approximations. Unreported results have shown that the solution times tend to increase more than linearly with the number of linear functions used while decreasing the protecting cost by less than 1%.

We turn then to the linear relaxation of the knapsack problem, where the binary restriction (37) is replaced by $x \in [0,1]^n$. For this model, the robust counterpart using model $\mathbf{U}^{\Gamma_\epsilon}$ is a linear program, while the robust counterpart of model $\mathcal{U}^{\gamma_\epsilon}$ requires the introduction of binary variables to describe $\| * \|_0$, see Theorem 2. Similarly to the case of the binary knapsack

FIGURE 4. Cost of protecting the solution for the fractional knapsack problem.



(a) $\epsilon = 0.01$



(b) $\epsilon = 0.05$

problem, we compute the optimal protection costs of the two models and compare the geometric means of their ratios in Figure 4. On average, model $\mathcal{U}^{\gamma_\epsilon}$ is 18% cheaper than $\mathbf{U}^{\Gamma_\epsilon}$, as in the case of the binary knapsack problem. Without surprise, the solution times of $\mathcal{U}^{\gamma_\epsilon}$ are higher than those of $\mathbf{U}^{\Gamma_\epsilon}$, but the average ratio $t(\mathcal{U}^{\gamma_{0.01}})/t(\mathbf{U}^{\Gamma_{0.01}})$ are hard to compute exactly because a large part of the solutions times for $\mathcal{U}^{\gamma_{0.01}}$ are to small to be measured. For both models, most instances are solved in less than 0.05 seconds, and the ratio $t(\mathcal{U}^{\gamma_{0.01}})/t(\mathbf{U}^{\Gamma_{0.01}})$ can be as large as 60 for the instances with 1000 items.

7. Conclusion

We have presented a new model for combinatorial optimization under uncertainty where the uncertain parameters are allowed to vary in uncertainty sets defined by multifunctions of the problem variables. We have motivated this general model by proposing a particular example of multifunction $\mathcal{U}^\gamma$ that extends the budgeted uncertainty set $\mathbf{U}^\Gamma$ studied by Berstimas and Sim [6]. Using multifunction $\mathcal{U}^\gamma$, one can obtain less conservative solutions than using $\mathbf{U}^\Gamma$, while ensuring the same probabilistic satisfaction of the constraint.

We have then shown how the robust counterpart of linear constraints can be computed in the new model, assuming that the uncertainty sets are delimited by conic inequalities that depend affinely on the variables of the problem. If the variables of the original problem are all binary, the dualization of the new model introduces a polynomial number of real variables and constraints, in addition to those added by the classical model. In that case, the new model can be almost as easy to solve as the classical robust model. However, if the original problem contains fractional or integer variables, the dualization requires to introduce additional binary variables and big-M coefficients.

We have provided computational experiments for the robust knapsack problem. The experiments show that using model $\mathcal{U}^\gamma$ reduces the cost of protecting the constraint of model $\mathbf{U}^\Gamma$ by 18% on average both for the binary and the fractional versions of the problem.

References

[1] A. Agra, M. Christiansen, R. Figueiredo, L. Magnus Hvattum, M. Poss, and C. Requejo. Layered formulation for the robust vehicle routing problem with time windows. In *Proceedings of ISCO 2012, Athens, Greece*, Lecture Notes in Computer Science 7422. Springer, 2012.

[2] A. Ben-Tal, L. El Ghaoui, and A.S. Nemirovski. *Robust optimization.* Princeton University Press, 2009.

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

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

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

[6] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52:35–53, 2004.

[7] J. R. Birge and F. V. Louveaux. *Introduction to Stochastic programming (2nd edition).* Springer Verlag, New-York, 2011.

[8] S. Burer and A.N. Letchford. Non-convex mixed-integer nonlinear programming: a survey, 2012. Submitted to Surveys in Operational Research and Management Science.

[9] E. Erdogan and G. Iyengar. Ambiguous chance constrained problems and robust optimization. *Mathematical Programming*, 107:37–61, 2006.
[10] B. Fortz, M. Labbé, F. V. Louveaux, and M. Poss. Stochastic binary problems with simple penalties for capacity constraints violations. *Mathematical Programming*, 2012. In press.
[11] IBM. IBM-ILOG Cplex Optimizer, 2012.
[12] A. J. Kleywegt, A. Shapiro, and T. Homem de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.*, 12(2):479–502, 2002.
[13] O. Klopfenstein and D. Nace. Cover inequalities for robust knapsack sets—application to the robust bandwidth packing problem. *Networks*, 59(1):59–72, 2012.

Michael Poss
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal
E-mail address: mposs@gmail.com