OJMO

Open Journal of Mathematical Optimization

Hande Yaman

Short Paper - A Note on Robust Combinatorial Optimization with Generalized Interval Uncertainty

Volume 4 (2023), article no. 4 (7 pages)

https://doi.org/10.5802/ojmo.23

Article submitted on October 14, 2022, revised on March 12, 2023,
accepted on March 7, 2023.

CENTRE MERSENNE

Open Journal of Mathematical Optimization is a member of the
Centre Mersenne for Open Scientific Publishing
http://www.centre-mersenne.org/
e-ISSN: 2777-5860

Short Paper - A Note on Robust Combinatorial Optimization with Generalized Interval Uncertainty

Hande Yaman
ORSTAT, Faculty of Economics and Business
KU Leuven, 3000 Leuven, Belgium
hande.yaman@kuleuven.be

Abstract

In this paper, we consider a robust combinatorial optimization problem with uncertain weights and propose an uncertainty set that generalizes interval uncertainty by imposing lower and upper bounds on deviations of subsets of items. We prove that if the number of such subsets is fixed and the family of these subsets is laminar, then the robust combinatorial optimization problem can be solved by solving a fixed number of nominal problems. This result generalizes a previous similar result for the case where the family of these subsets is a partition of the set of items.

Digital Object Identifier 10.5802/ojmo.23

Keywords robust combinatorial optimization, interval uncertainty, budgeted uncertainty, complexity.

1 Introduction

We consider a combinatorial optimization problem

$$(CO) \qquad \min_{x\in X} cx,$$

where $N = \{1,\dots,n\}$ is the ground set of items, $X \subset \{0,1\}^n$ is the set of feasible solutions x with x_i equal to 1 if item $i \in N$ is part of the solution and 0 otherwise and c is the vector of weights (a row vector of size n). We are interested in the case where the weights are not known with certainty; but an uncertainty set U, which is the set of all possible realizations of weights (with no associated probability distribution), is available. The aim is to find a robust solution, i.e., a feasible solution whose worst-case weight is minimum. So we would like to solve the robust combinatorial optimization problem with uncertainty set U, which is defined as:

$$RCO(U) \qquad \min_{x\in X}\max_{c\in U} cx.$$

A common way of modeling uncertainty in robust optimization is defining a nonempty interval for each unknown parameter. This is rather natural and easy compared to defining a set of discrete scenarios. In interval uncertainty, we consider scenarios where the weight of item $i \in N$ takes a value in a nonempty interval $[l_i, u_i]$, independently of the weights of other items. Then, a worst-case scenario for any feasible solution is the scenario in which all weights are equal to their upper bounds. However, this is an extreme scenario and making a decision based solely on it is considered to be very conservative. Different uncertainty sets have been proposed with the aim of limiting this conservatism. Our aim in this study is to propose a polyhedral uncertainty set, that we refer to as "generalized deviation budgeted uncertainty set" and denote by U_{G-dev}. This set is the set of all c that satisfy

$$c_i = l_i + \delta_i \qquad i \in N, \tag{1}$$

$$\sum_{i\in N_k} \delta_i \le \Lambda_k \qquad k \in K, \tag{2}$$

$$\sum_{i\in N_k} \delta_i \ge \lambda_k \qquad k \in K, \tag{3}$$

$$\delta_i \le u_i - l_i \qquad i \in N, \tag{4}$$

$$\delta_i \ge 0 \qquad i \in N, \tag{5}$$

where $K = \{1, \ldots, m\}$, m is a nonnegative integer, $N_k \subseteq N$, $\Lambda_k \geq \lambda_k \geq 0$ for all $k \in K$.

Set U_{G-dev} generalizes the locally budgeted uncertainty model, U_{L-dev}, proposed by Goerigk and Lendl [6]. More precisely, U_{L-dev} corresponds to the special case of U_{G-dev} with no lower bounds, i.e., $\lambda_k = 0$ for all $k \in K$. Goerigk and Lendl study the case where $N_1, \ldots, N_m$ is a partition of N. Set U_{L-dev} is also used by Gounaris et al. [8] to model the demand uncertainty in the robust capacitated vehicle routing problem.

Goerigk and Lendl prove that $RCO(U_{L-dev})$ can be solved by solving 2^m nominal problems when $N_1, \ldots, N_m$ is a partition of N. Consequently, if m is fixed and if CO can be solved in polynomial time, then $RCO(U_{L-dev})$ can be solved in polynomial time. The authors conclude their paper by noting that their positive results do not translate directly to settings where $N_1, \ldots, N_m$ is not necessarily a partition of N. In this study, we give another positive result for the case where $N_1, \ldots, N_m$ is a laminar set family. The family $N_1, \ldots, N_m$ is called laminar if for any $k_1 < k_2$ in K, either N_{k_1} and N_{k_2} are disjoint or one is a subset of the other, i.e., $N_{k_1} \cap N_{k_2} = \emptyset$ or $N_{k_1} \subset N_{k_2}$ or $N_{k_2} \subset N_{k_1}$. Laminar families have already been studied in the context of uncertainty sets in robust optimization. Indeed, Gounaris et al. [7] extend their study on the robust capacitated vehicle routing problem to the case of inclusion-constrained budgeted uncertainty sets, i.e., the case in which the family of sets $N_1, \ldots, N_m$ is laminar. Wiesemann et al. [10] call the laminarity conditions as nesting conditions for the confidence sets in their ambiguity set.

The uncertainty model U_{G-dev} is interesting in modeling parameters correlated in a particular way. We give two examples here. Suppose that we would like to locate blood collection points in such a way that the population that can reach at least one of these points within a threshold walking distance is maximized. Let N be the set of streets in the area and c_i denote the population on street $i \in N$. It is difficult to come up with exact populations as people move to different parts of a city depending on the hour of the day. Hence, it would be natural to propose intervals rather than point estimates. But in addition to defining an interval for each c_i, we may also estimate lower and upper bounds on the population in neighborhoods composed of several streets and in regions composed of several neighborhoods and this additional information can be used to limit the conservatism. Consider another setting where N is a set of products and c_i denotes the demand for product $i \in N$. Products may be grouped into subsets based on their properties, such as color, model and year of release, in such a way that products in the same group may be substitutes for each other. Then, the uncertainty set can be improved by including lower and upper bounds on the aggregate demands.

To motivate the use of both lower and upper bounds on deviations, we provide a small example.

► **Example.** Let $N = \{1, \ldots, 8\}$, $l = (1,1,1,1,1,1,1,1)$ and $u = (4,5,4,5,4,5,4,5)$. Suppose that we would like to solve the selection problem where we choose four items with minimum weight.

First, we consider the simple interval uncertainty model and let $U^1 = \{c : c_i = l_i + \delta_i, \delta_i \in [0, u_i - l_i] \text{ for } i \in N\}$. Problem $RCO(U^1)$ has a unique optimal solution $(1,0,1,0,1,0,1,0)$ with worst-case weight equal to 16. Next, we add a budget constraint to U^1 and obtain $U^2 = \{c : c_i = l_i + \delta_i, \delta_i \in [0, u_i - l_i] \text{ for } i \in N, \sum_{i \in N} \delta_i \leq 10\}$. Now any solution with four items is optimal for $RCO(U^2)$. Note that, under both uncertainty sets, for any feasible solution, there exists a worst-case scenario in which the weights of the items that are not part of the solution are at their lower bounds. For instance, solution $(1,0,1,0,1,0,1,0)$ is optimal for $RCO(U^2)$ and its worst-case weight is 14 with the weights of items 2, 4, 6 and 8 equal to 1.

Now suppose that we have additional information about the deviations. Let $N_k = \{2k-1, 2k\}$ for $k = 1,2,3,4$ and consider the uncertainty set

$$U^3 = \left\{ c : c_i = l_i + \delta_i, \delta_i \in [0, u_i - l_i] \text{ for } i \in N, \sum_{i \in N} \delta_i \leq 10, \sum_{i \in N_k} \delta_i \geq 2 \text{ for } k = 1,2,3,4 \right\}.$$

Under U^3, solution $(1,0,1,0,1,0,1,0)$ has worst-case weight 14 whereas solution $(1,1,1,1,0,0,0,0)$ has worst-case weight 10. Indeed the latter solution is optimal for $RCO(U^3)$. With U^3, in the worst-case scenario for solution $(1,1,1,1,0,0,0,0)$, items 5, 6, 7 and 8, the items that are not part of the solution, do not have their weights all at their lower bounds; the weights of these four items have a deviation of at least four units in total.

Our main result in this paper is the following:

► **Theorem 1.** *Problem $RCO(U_{G-dev})$ can be solved by solving 2^m nominal problems if the family of sets $N_1, \ldots, N_m$ is laminar.*

Theorem 1 generalizes the result of Goerigk and Lendl for a partition to a laminar family of sets. We note later that imposing lower bounds on deviations would play no role when $N_1, \ldots, N_m$ is a partition (see Corollary 8).

In the next section, we review other uncertainty sets and similar results obtained for them. Section 3 gives a proof of Theorem 1. We conclude with some research directions in Section 4.

2 Related results and a compact formulation

The best-known uncertainty set with limited conservatism is the budgeted uncertainty set proposed by Bertsimas and Sim [2],

$$U_{card} = \{c : c_i = l_i + \xi_i(u_i - l_i), \xi_i \in [0,1] \text{ for } i \in N, \sum_{i\in N} \xi_i \leq \Gamma\},$$

where $\Gamma \leq n$ is a nonnegative integer. The extreme points of this uncertainty set correspond to the scenarios in which Γ parameters take their highest values while the other $n - \Gamma$ parameters take their lowest (estimated) values. Bertsimas and Sim show that $RCO(U_{card})$ can be solved by solving at most $n+1$ nominal problems. Note that there is a difference in the definition of the deviation compared to the one of Goerigk and Lendl [6]. In the cardinality budgeted uncertainty of Bertsimas and Sim, the bound Γ is on the fraction deviations (of $u_i - l_i$) whereas Goerigk and Lendl impose the bound Λ on the actual deviations. This difference in the definition of deviations is quite significant when one looks at the complexity of the resulting problems: for $m = 1$, solving $RCO(U_{L-dev})$ requires solving two nominal problems, whereas one may need to solve up to $n+1$ nominal problems to solve $RCO(U_{card})$.

Poss [9] generalizes the cardinality budgeted uncertainty set by considering m knapsack constraints. The knapsack uncertainty set is

$$U_{knap} = \{c : c_i = l_i + \xi_i(u_i - l_i) \text{ for } i \in N, \sum_{i\in N} a_{ki}\xi_i \leq b_k \text{ for } k = 1, \ldots, m, 0 \leq \xi \leq \bar{\xi}\},$$

where $a \in \mathbb{R}_+^{m\times n}$, $b \in \mathbb{R}_+^m$ and $\bar{\xi} \in \mathbb{R}_+^n$. He proves that $RCO(U_{knap})$ can be solved in polynomial time if m is fixed and CO can be solved in polynomial time. The set U_{G-dev} does not generalize U_{knap} since a is arbitrary in U_{knap} and U_{knap} does not generalize U_{G-dev} since $a \geq 0$ and constraints (3) cannot be modeled.

Similar uncertainty models have been previously used to model the uncertainty in the traffic demand between origin-destination pairs in the context of a network design problem. Duffield et al. [4] and Fingerhut et al. [5] propose the hose model that specifies aggregate bounds on the traffic adjacent at nodes rather than bounds on pairwise traffic demands. The hose model is extended to a hybrid model by adding box constraints on the individual traffic demands by Altın et al. [1]. Ambiguity sets used in distributionally robust optimization have also been extended in similar ways, see, e.g., Wiesemann et al. [10] and Chan et al. [3].

Since U_{G-dev} is polyhedral, we can easily obtain a compact mixed integer programming formulation for $RCO(U_{G-dev})$. With U_{G-dev}, the worst-case weight for a solution $x \in X$ can be computed by solving the linear program

$$\sum_{i\in N} l_i x_i + \max \sum_{i\in N} x_i \delta_i \quad \text{s.t. (2)–(5).}$$

If U_{G-dev} is not empty, then this linear program is feasible and bounded and hence has the same optimal value as its dual. Consequently, the robust counterpart, denoted by $RCO(U_{G-dev})$, can be stated as follows.

$$\min \sum_{i\in N} l_i x_i + \sum_{k\in K} (\Lambda_k \sigma_k - \lambda_k \upsilon_k) + \sum_{i\in N} (u_i - l_i)\rho_i \tag{6}$$

$$\text{s.t. } x \in X, \tag{7}$$

$$\sum_{k\in K} a_{ki}(\sigma_k - \upsilon_k) + \rho_i \geq x_i \qquad i \in N, \tag{8}$$

$$\sigma, \upsilon, \rho \geq 0, \tag{9}$$

where, for $i \in N$ and $k \in K$, $a_{ki} = 1$ if $i \in N_k$ and 0 otherwise.

If $K = \emptyset$, then we get the interval uncertainty model. It is easy to see that in this case there is an optimal solution in which $\rho_i = x_i$ for all $i \in N$, i.e., the robust problem is equivalent to the nominal problem with $c_i = u_i$ for all $i \in N$.

3 Proof of Theorem 1

We prove a series of lemmas to prove Theorem 1. For $k' \in K$, let $K_{k'} = \{k \in K : N_k \subset N_{k'}\}$ and for $T \subseteq K$, let $N_T = \cup_{k \in T} N_k$. We first give a lemma that is used repeatedly in the proofs of the subsequent lemmas.

▶ **Lemma 2.** *Let $k' \in K$ and $T \subseteq K_{k'}$. If the family of sets $N_1, \ldots, N_m$ is laminar, there exists a subset $S \subseteq T$ such that $\sum_{k \in S} a_{ki} = 1$ for all $i \in N_T$ and $\sum_{k \in S} a_{ki} = 0$ for all $i \in N_{k'} \setminus N_T$.*

Proof. Let $k' \in K$ and $T \subseteq K_{k'}$ and suppose that the family of sets $N_1, \ldots, N_m$ is laminar. Start with an empty set S. Let k be an index in T with largest cardinality N_k (break ties arbitrarily). Add k to S and remove k and all l with $N_l \subset N_k$ from T. Continue until T is empty. It is easy to see that N_{k_1} and N_{k_2} are disjoint for any distinct k_1 and k_2 in S. Moreover, since $S \subseteq T$, $N_S \subseteq N_T$. Let $i \in N_{k_1}$ with $k_1 \in T$. Either $k_1 \in S$ and $i \in N_S$ or $k_1 \notin S$ and there exists $k_2 \in S$ with $N_{k_1} \subset N_{k_2}$. In the latter case, $i \in N_{k_2} \subseteq N_S$. Hence $N_T \subseteq N_S$. Overall, $N_T = N_S$. This together with the fact that N_{k_1} and N_{k_2} are disjoint for any distinct k_1 and k_2 in S gives the result. ◀

▶ **Lemma 3.** *If the family of sets $N_1, \ldots, N_m$ is laminar, then for $k' \in K$ and $i \in N_{k'}$, constraint (8) is the same as*

$$\sum_{k \in K_{k'}} a_{ki}(\sigma_k - \upsilon_k) + \sum_{k \in K : N_{k'} \subseteq N_k} (\sigma_k - \upsilon_k) + \rho_i \geq x_i. \tag{10}$$

Proof. Let $k' \in K$ and $i \in N_{k'}$ and suppose that the family of sets $N_1, \ldots, N_m$ is laminar. Then, for any $k \in K \setminus \{k'\}$, either $N_k \subset N_{k'}$ or $N_k' \subset N_k$ or $N_k \cap N_{k'} = \emptyset$. If k is such that $N_k' \subset N_k$, then $a_{ki} = 1$ and if k is such that $N_k \cap N_{k'} = \emptyset$, then $a_{ki} = 0$ since $i \in N_{k'}$. ◀

Let P be the convex hull of the feasible set of $RCO(U_{G-dev})$ as modeled above, i.e., $P = conv(\{(x, \sigma, \upsilon, \rho) : (7)–(9)\})$. Let e_i^n be the i-th unit vector of size n and e_k^m be the k-th unit vector of size m. We will use the properties of extreme points of P in proving our result. The first property is given without proof as it is easy to demonstrate.

▶ **Lemma 4.** *An extreme point $(x, \sigma, \upsilon, \rho)$ of P satisfies $\sigma_k \upsilon_k = 0$ for all $k \in K$.*

Different from the above property, the following ones rely on the family of sets $N_1, \ldots, N_m$ being laminar.

▶ **Lemma 5.** *If the family of sets $N_1, \ldots, N_m$ is laminar, an extreme point $(x, \sigma, \upsilon, \rho)$ of P satisfies*

$$\sum_{k \in K : N_{k'} \subseteq N_k} (\sigma_k - \upsilon_k) \geq 0$$

for all $k' \in K$.

Proof. Suppose that the family of sets $N_1, \ldots, N_m$ is laminar and let $p = (x, \sigma, \upsilon, \rho)$ be an extreme point of P and $k' \in K$. We look at three cases:

Case 1: $\upsilon_{k'} > 0$. Assume, for the sake of deriving a contradiction, that $\sum_{k \in K : N_{k'} \subseteq N_k} (\sigma_k - \upsilon_k) < 0$. From Lemma 4, we know that $\sigma_{k'} = 0$. For $i \in N_{k'}$, by Lemma 3, we know that p satisfies constraint (10). This constraint can be rewritten as

$$\sum_{k \in K_{k'}} a_{ki}\sigma_k + \rho_i \geq x_i + \sum_{k \in K_{k'}} a_{ki}\upsilon_k - \sum_{k \in K : N_{k'} \subseteq N_k} (\sigma_k - \upsilon_k).$$

Using $x, \upsilon \geq 0$ and $\sum_{k \in K : N_{k'} \subseteq N_k} (\sigma_k - \upsilon_k) < 0$, we see that the right hand side is positive. Hence for $i \in N_{k'}$, we have $\sum_{k \in K_{k'}} a_{ki}\sigma_k + \rho_i > 0$.

Let $T = \{k \in K_{k'} : \sigma_k > 0\}$. Let $i \in N_{k'} \setminus N_T$. By definition of T, $\sum_{k \in K_{k'}} a_{ki}\sigma_k = 0$. Since $\sum_{k \in K_{k'}} a_{ki}\sigma_k + \rho_i > 0$, we have $\rho_i > 0$. By Lemma 2, there exists a subset $S \subseteq T$ such that $\sum_{k \in S} a_{ki} = 1$ for all $i \in N_T$ and $\sum_{k \in S} a_{ki} = 0$ for all $i \in N_{k'} \setminus N_T$. Using this together with $\rho_i > 0$ for all $i \in N_{k'} \setminus N_T$, we see that there exists a small positive ϵ such that the two points $p^1 = (x, \sigma + \sum_{k \in S} \epsilon e_k^m, \upsilon + \epsilon e_{k'}^m, \rho + \sum_{i \in N_{k'} \setminus N_T} \epsilon e_i^n)$ and $p^2 = (x, \sigma - \sum_{k \in S} \epsilon e_k^m, \upsilon - \epsilon e_{k'}^m, \rho - \sum_{i \in N_{k'} \setminus N_T} \epsilon e_i^n)$ are both in P. In addition, $p = \frac{1}{2}p^1 + \frac{1}{2}p^2$. This is in contradiction with p being an extreme point of P. Hence if $\upsilon_{k'} > 0$, then $\sum_{k \in K : N_{k'} \subseteq N_k} (\sigma_k - \upsilon_k) \geq 0$.

Case 2: $\upsilon_{k'} = 0$ and there exists k with $N_{k'} \subset N_k$ and $\upsilon_k > 0$. Let k'' be the index of the smallest cardinality set containing set $N_{k'}$ and $\upsilon_{k''} > 0$. Since the family of sets is laminar, any $k \in K$ with $N_{k'} \subset N_k$ is such that $N_k \subset N_{k''}$ or $N_{k''} \subseteq N_k$. By the choice of $N_{k''}$, we know that $\upsilon_k = 0$ for all $k \in K$ with $N_{k'} \subseteq N_k \subset N_{k''}$. Hence $\sum_{k\in K:N_{k'}\subseteq N_k}(\sigma_k - \upsilon_k) = \sum_{k\in K:N_{k'}\subseteq N_k\subset N_{k''}} \sigma_k + \sum_{k\in K:N_{k''}\subseteq N_k}(\sigma_k - \upsilon_k)$. From case 1, since $\upsilon_{k''} > 0$, we know that $\sum_{k\in K:N_{k''}\subseteq N_k}(\sigma_k - \upsilon_k) \geq 0$. Using this together with $\sigma \geq 0$, we have $\sum_{k\in K:N_{k'}\subseteq N_k\subset N_{k''}} \sigma_k + \sum_{k\in K:N_{k''}\subseteq N_k}(\sigma_k - \upsilon_k) \geq 0$ and so $\sum_{k\in K:N_{k'}\subseteq N_k}(\sigma_k - \upsilon_k) \geq 0$.

Case 3: $\upsilon_{k'} = 0$ and $\upsilon_k = 0$ for all k with $N_{k'} \subset N_k$. Then,

$$\sum_{k\in K:N_{k'}\subseteq N_k}(\sigma_k - \upsilon_k) = \sum_{k\in K:N_{k'}\subseteq N_k} \sigma_k \geq 0. \qquad ◀$$

▶ **Lemma 6.** *Suppose that the family of sets $N_1, \ldots, N_m$ is laminar and $(x, \sigma, \upsilon, \rho)$ is an extreme point of P. For $k' \in K$, if $\sigma_{k'} > 0$ and $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) = 0$, then $\sigma_{k'} = 1$. If there exists k'' with $N_{k''} \subset N_{k'}$ and $\sigma_{k''} + \upsilon_{k''} > 0$ and $\sigma_k = \upsilon_k = 0$ for all k with $N_{k''} \subset N_k \subset N_{k'}$, then $\upsilon_{k''} = 1$ and $\sigma_{k''} = 0$.*

Proof. Suppose that the family of sets $N_1, \ldots, N_m$ is laminar and $(x, \sigma, \upsilon, \rho)$ is an extreme point of P. Let $k' \in K$ be such that $\sigma_{k'} > 0$ and $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) = 0$. By Lemma 4, we know that $\upsilon_{k'} = 0$. Let $K^1 = \{k \in K_{k'} : \sigma_k > 0\}$, $K^2 = \{k \in K_{k'} : \upsilon_k > 0\}$ and $T = K^1 \cup K^2$. We know that $\sum_{k\in K_{k'}} a_{ki}\sigma_k = 0$ and $\sum_{k\in K_{k'}} a_{ki}\upsilon_k = 0$ for all $i \in N_{k'} \setminus N_T$. By Lemma 2, there exists a subset $S \subseteq T$ such that $\sum_{k\in S} a_{ki} = 1$ for all $i \in N_T$ and $\sum_{k\in S} a_{ki} = 0$ for all $i \in N_{k'} \setminus N_T$.

For $i \in N_{k'}$, constraint (10) becomes $\sigma_{k'} + \sum_{k\in K_{k'}} a_{ki}\sigma_k - \sum_{k\in K_{k'}} a_{ki}\upsilon_k + \rho_i \geq x_i$. If $i \in N_{k'} \setminus N_T$, then this constraint further becomes $\sigma_{k'} + \rho_i \geq x_i$. If $x_i = 0$, then the constraint cannot be tight since $\sigma_{k'} > 0$ and $\rho_i \geq 0$. Let $N' = \{i \in N_{k'} \setminus N_T : x_i = 1\}$. Suppose that $N' = \emptyset$ or $\rho_i > 0$ for all $i \in N'$. Then, there exists a small positive ϵ such that the two points $p^1 = (x, \sigma + \sum_{k\in S\cap K^1} \epsilon e^m_k - \epsilon e^m_{k'}, \upsilon - \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho + \sum_{i\in N'} \epsilon e^n_i)$ and $p^2 = (x, \sigma - \sum_{k\in S\cap K^1} \epsilon e^m_k + \epsilon e^m_{k'}, \upsilon + \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho - \sum_{i\in N'} \epsilon e^n_i)$ are both in P and $p = \frac{1}{2}p^1 + \frac{1}{2}p^2$. This is in contradiction with p being an extreme point of P. Hence there exists $i \in N'$ with $\rho_i = 0$. Then, $\sigma_{k'} + \rho_i \geq x_i$, $x_i = 1$ and $\rho_i = 0$ give $\sigma_{k'} \geq 1$. If $\sigma_{k'} > 1$, then there exists a small positive ϵ such that the two points $p^1 = (x, \sigma + \sum_{k\in S\cap K^1} \epsilon e^m_k - \epsilon e^m_{k'}, \upsilon - \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho)$ and $p^2 = (x, \sigma - \sum_{k\in S\cap K^1} \epsilon e^m_k + \epsilon e^m_{k'}, \upsilon + \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho)$ are both in P. Since $p = \frac{1}{2}p^1 + \frac{1}{2}p^2$, p cannot an extreme point of P. Thus we can conclude that if p is an extreme point of P, then $\sigma_{k'} = 1$.

Consider the same k' and suppose that there exists k'' with $N_{k''} \subset N_{k'}$ and $\upsilon_{k''} > 0$, $\upsilon_{k''} \neq 1$ and $\sigma_k = \upsilon_k = 0$ for all k with $N_{k''} \subset N_k \subset N_{k'}$. Let $K^1 = \{k \in K_{k''} : \sigma_k > 0\}$, $K^2 = \{k \in K_{k''} : \upsilon_k > 0\}$ and $T = K^1 \cup K^2$. By Lemma 2, there exists a subset $S \subseteq T$ such that $\sum_{k\in S} a_{ki} = 1$ for all $i \in N_T$ and $\sum_{k\in S} a_{ki} = 0$ for all $i \in N_{k''} \setminus N_T$. For $i \in N_{k''} \setminus N_T$, constraint (10) becomes $\sigma_{k'} - \upsilon_{k''} + \rho_i \geq x_i$ and since $\upsilon_{k''} \neq \sigma_{k'}$, either $\rho_i > 0$ or $\sigma_{k'} - \upsilon_{k''} + \rho_i > x_i$. Let $N'' = \{i \in N_{k''} \setminus N_T : \rho_i > 0\}$. There exists a small positive ϵ such that the two points $p^1 = (x, \sigma + \sum_{k\in S\cap K^1} \epsilon e^m_k, \upsilon + \epsilon e^m_{k''} - \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho + \sum_{i\in N''} \epsilon e^n_i)$ and $p^2 = (x, \sigma - \sum_{k\in S\cap K^1} \epsilon e^m_k, \upsilon - \epsilon e^m_{k''} + \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho - \sum_{i\in N''} \epsilon e^n_i)$ are both in P and $p = \frac{1}{2}p^1 + \frac{1}{2}p^2$. Hence for p to be an extreme point, we need $\upsilon_{k''} = 1$.

Now suppose that there exists k'' with $N_{k''} \subset N_{k'}$ and $\sigma_{k''} > 0$ and $\sigma_k = \upsilon_k = 0$ for all k with $N_{k''} \subset N_k \subset N_{k'}$. As we have done above, define $K^1 = \{k \in K_{k''} : \sigma_k > 0\}$, $K^2 = \{k \in K_{k''} : \upsilon_k > 0\}$ and $T = K^1 \cup K^2$. For $i \in N_{k''} \setminus N_T$, constraint (10) is $\sigma_{k'} + \sigma_{k''} + \rho_i \geq x_i$ and is never tight since the left hand side is larger than 1. Since there exists a subset $S \subseteq T$ such that $\sum_{k\in S} a_{ki} = 1$ for all $i \in N_T$ and $\sum_{k\in S} a_{ki} = 0$ for all $i \in N_{k''} \setminus N_T$ by Lemma 2, one can find a small positive ϵ such that the two points $p^1 = (x, \sigma + \sum_{k\in S\cap K^1} \epsilon e^m_k - \epsilon e^m_{k''}, \upsilon - \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho)$ and $p^2 = (x, \sigma - \sum_{k\in S\cap K^1} \epsilon e^m_k + \epsilon e^m_{k''}, \upsilon + \sum_{k\in S\cap K^2} \epsilon e^m_k, \rho)$ are both in P. Since $p = \frac{1}{2}p^1 + \frac{1}{2}p^2$, p cannot be an extreme point, proving that such k'' does not exist. ◀

▶ **Lemma 7.** *Suppose that the family of sets $N_1, \ldots, N_m$ is laminar and $(x, \sigma, \upsilon, \rho)$ is an extreme point of P. For $k' \in K$, if $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) = 0$, then $\upsilon_{k'} = 0$ and $\sigma_{k'}$ is 0 or 1 and if $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) > 0$, then $\sigma_{k'} = 0$ and $\upsilon_{k'}$ is 0 or 1. Also $\sum_{k\in K} a_{ki}(\sigma_k - \upsilon_k) \in \{0, 1\}$ for all $i \in N$.*

Proof. Let $k' \in K$. If $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) = 0$, then, as $\sum_{k\in K:N_{k'}\subseteq N_k}(\sigma_k - \upsilon_k) \geq 0$ by Lemma 5 and $\sigma_{k'}\upsilon_{k'} = 0$ by Lemma 4, we know that $\upsilon_{k'} = 0$. Also, by Lemma 6, we have $\sigma_{k'}$ is 0 or 1. If $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) > 0$, then let $k_1, \ldots, k_t$ be such that $N_{k'} \subset N_{k_t} \subset \ldots, N_{k_1}$ with $\sigma_{k_j} + \upsilon_{k_j} > 0$ for all $j = 1, \ldots, t$ and $\sigma_k = \upsilon_k = 0$ for all k with $N_{k'} \subset N_k$ and $k \neq k_j$ for all $j \in \{1, \ldots, t\}$. Since $\sum_{k\in K:N_{k_1}\subset N_k}(\sigma_k - \upsilon_k) = 0$, we have $\sigma_{k_1} = 1$ and $\upsilon_{k_1} = 0$. Then, by Lemma 6, $\sigma_{k_2} = 0$ and $\upsilon_{k_2} = 1$. Repeating this argument, we can show that if $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) = 0$, then $\upsilon_{k'} = 0$ and $\sigma_{k'}$ is 0 or 1 and if $\sum_{k\in K:N_{k'}\subset N_k}(\sigma_k - \upsilon_k) > 0$, then $\sigma_{k'} = 0$ and $\upsilon_{k'}$ is 0 or 1. Finally $\sum_{k\in K} a_{ki}(\sigma_k - \upsilon_k) \in \{0, 1\}$ for all $i \in N$ follows as a consequence. ◀

Before proceeding to the proof of Theorem 1, we have a corollary that states that if $N_1, \ldots, N_m$ is a partition of N in U_{G-dev}, then we obtain the locally budgeted uncertainty set proposed by Goerigk and Lendl [6].

▶ **Corollary 8.** *If $N_1, \ldots, N_m$ is a partition of N, then the lower bounds λ_k for $k \in K$ are redundant in $RCO(U_{G-dev})$.*

Proof. Suppose that $N_1, \ldots, N_m$ is a partition of N and let $(x, \sigma, \upsilon, \rho)$ be an extreme point of P. For any $k' \in K$ we have $\sum_{k \in K: N_{k'} \subset N_k} (\sigma_k - \upsilon_k) = 0$ since no subset is contained in another one. Then, by Lemma 7, $\upsilon_{k'} = 0$ and $\sigma_{k'}$ is 0 or 1. Hence $RCO(U_{G-dev})$ has an optimal solution with $\upsilon = 0$. This implies that the lower bounds λ_k for $k \in K$ play no role and can be dropped. ◀

Proof of Theorem 1. Suppose that the family of sets $N_1, \ldots, N_m$ is laminar and let $p = (x, \sigma, \upsilon, \rho)$ be an extreme point of P. Let $K^+ = \{k \in K : \sigma_k + \upsilon_k > 0\}$. Let $k' \in K^+$ be such that there exists no $k \in K^+$ with $N_{k'} \subset N_k$. Then, by Lemma 7, we know that $\sigma_{k'} = 1$ and $\upsilon_{k'} = 0$. Let $k'' \in K^+$ be such that $N_{k''} \subset N_{k'}$ and there exists no $k \in K^+$ with $N_{k''} \subset N_k \subset N_{k'}$. Then, again by Lemma 7, we know that $\sigma_{k''} = 0$ and $\upsilon_{k''} = 1$. Using Lemma 7 in this way, we can identify the values of all σ and υ. Hence, the number of possible choices for σ and υ at an extreme point is 2^m.

When σ and υ are fixed, model (6)–(9) reduces to

$$\sum_{k \in K} (\Lambda_k \sigma_k - \lambda_k \upsilon_k) + \min \sum_{i \in N} l_i x_i + \sum_{i \in N} (u_i - l_i)(x_i - \sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k))^+$$
$$\text{s.t. } x \in X,$$

since there exists an optimal solution in which $\rho_i = (x_i - \sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k))^+$ for all $i \in N$. Let $i \in N$. From Lemma 7, we know that $\sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k) \in \{0, 1\}$. If $\sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k) = 1$, then $(x_i - \sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k))^+ = 0$ since $x_i \leq 1$. If $\sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k) = 0$, then $(x_i - \sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k))^+ = x_i$. Then, the remaining problem is the nominal problem where, for $i \in N$, $c_i = l_i$ if $\sum_{k \in K} a_{ki}(\sigma_k - \upsilon_k) = 1$ and $c_i = u_i$ otherwise. ◀

4 Concluding Remarks

We conclude with two remarks: First, the structure of the constraints that relate the dual variables σ, υ and ρ and the original variables x is much easier to study when we impose bounds on the actual deviations rather than imposing them on the cardinality of items whose weights deviate. Second, imposing lower bounds on the deviations in addition to the upper bounds can play a more important role when one wants to minimize the worst-case regret rather than an absolute measure as we have done here. However the robust counterpart becomes much more challenging in that case.

References

1 Ayşegül Altın, Hande Yaman, and Mustafa Ç. Pınar. A hybrid polyhedral uncertainty model for the robust network loading problem. In *Performance models and risk management in communications systems*, volume 46 of *Springer Optimization and Its Applications*, pages 157–172. Springer, 2011.

2 Dimitris Bertsimas and Melvyn Sim. Robust discrete optimization and network flows. *Math. Program.*, 98(1):49–71, 2003.

3 Timothy C. Y. Chan, Zuo-Jun Max Shen, and Auyon Siddiq. Robust defibrillator deployment under cardiac arrest location uncertainty via row-and-column generation. *Oper. Res.*, 66(2):358–379, 2018.

4 Nick G. Duffield, Pawan Goyal, Albert Greenberg, Partho Mishra, Kadangode K Ramakrishnan, and Jacobus E. van der Merive. A flexible model for resource management in virtual private networks. In *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, pages 95–108. ACM Press, 1999.

5 J. Andrew Fingerhut, Subhash Suri, and Jonathan S. Turner. Designing least-cost nonblocking broadband networks. *J. Algorithms*, 24(2):287–309, 1997.

6 Marc Goerigk and Stefan Lendl. Robust combinatorial optimization with locally budgeted uncertainty. *Open J. Math. Optim.*, 2(3):18, 2021.

7 Chrysanthos E. Gounaris, Panagiotis P. Repoussis, Christos D. Tarantilis, Wolfram Wiesemann, and Christodoulos A. Floudas. An adaptive memory programming framework for the robust capacitated vehicle routing problem. *Transp. Sci.*, 50(4):1239–1260, 2016.

8 Chrysanthos E. Gounaris, Wolfram Wiesemann, and Christodoulos A. Floudas. The robust capacitated vehicle routing problem under demand uncertainty. *Oper. Res.*, 61(3):677–693, 2013.

9 Michael Poss. Robust combinatorial optimization with knapsack uncertainty. *Discrete Optim.*, 27:88–102, 2018.
10 Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Oper. Res.*, 62(6):1358–1376, 2014.