On the Polynomial Solvability of Distributionally Robust k-Sum Optimization

Anulekha Dhara[*] Karthik Natarajan[‡]

Submitted: December 2015, First Revision: April 2016, Second Revision: July 2016

Abstract

In this paper, we define a distributionally robust k-sum optimization problem as the problem of finding a solution that minimizes the worst-case expected sum of up to the k largest costs of the elements in the solution. The costs are random with a joint probability distribution that is not completely specified but rather assumed to be known to lie in a set of probability distributions. For $k = 1$, this reduces to a distributionally robust bottleneck optimization problem while for $k = n$, this reduces to distributionally robust minimum sum optimization problem. Our main result is that for a Fréchet class of discrete marginal distributions with finite support, the distributionally robust k-sum combinatorial optimization problem is solvable in polynomial time if the deterministic minimum sum problem is solvable in polynomial time. This extends the result of Punnen and Aneja (Operations Research Letters, 18(5), 1996) from the deterministic to the robust case. We show that this choice of the set of distributions helps preserve the submodularity of the k-sum objective function which is an useful structural property for optimization problems.

1 Introduction

Consider a generic combinatorial optimization problem of the following form: Let $\mathcal{N} = \{1, 2, \ldots, n\}$ be a finite ground set and $\mathcal{F}$ be a family of subsets of $\mathcal{N}$. Associated with each $F \in \mathcal{F}$ is a n-dimensional

[*]Engineering System and Design, Singapore University of Technology and Design, Singapore 487372. Email: anulekha@sutd.edu.sg

[†]Engineering System and Design, Singapore University of Technology and Design, Singapore 487372. Email: karthik_natarajan@sutd.edu.sg

[‡]This project was partly funded by the SUTD-MIT International Design Center grant number IDG31300105 on 'Optimization for Complex Discrete Choice' and the MOE Tier 2 grant number MOE2013-T2-2-168 on 'Distributional Robust Optimization for Consumer Choice in Transportation Systems'.

characteristic vector $\mathbf{x}$ where $x_i = 1$ if $i \in F$ and 0 otherwise. Let $\mathcal{X}$ denote the set of characteristic vectors associated with $\mathcal{F}$. Each element $i \in \mathcal{N}$ is associated with a nonnegative cost $c_i \in \mathbb{R}_+$. For a given cost vector $\mathbf{c} \in \mathbb{R}^n_+$ and a feasible solution $\mathbf{x} \in \mathcal{X}$, an ordering of the elements in $\mathcal{N}$ in term of non-decreasing costs is denoted by $\sigma(\cdot) = \{\sigma(1), \sigma(2), \ldots, \sigma(n)\}$ where $c_{\sigma(1)}x_{\sigma(1)} \geq c_{\sigma(2)}x_{\sigma(2)} \geq \ldots \geq c_{\sigma(n)}x_{\sigma(n)}$. For a given integer $k \in \{1, 2, \ldots, n\}$, the k-sum combinatorial optimization is defined as follows:

$$\min_{\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n} \sum_{i=1}^{k} c_{\sigma(i)} x_{\sigma(i)}, \tag{1.1}$$

where the sum of up to the k largest costs is minimized. Equivalently, this problem can be formulated as the following minimax optimization problem:

$$\min_{\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n} \max_{\mathbf{y} \in \mathcal{U}_k \subseteq \{0,1\}^n} \sum_{i=1}^{n} c_i x_i y_i, \tag{1.2}$$

where the inner maximization is over the set $\mathcal{U}_k$ which is a uniform matroid of rank k and is defined as follows:

$$\mathcal{U}_k = \left\{ \mathbf{y} \;\middle|\; \sum_{i=1}^{n} y_i \leq k,\; y_i \in \{0, 1\}, \forall i \in \mathcal{N} \right\}. \tag{1.3}$$

In particular, for $k = n$, the objective function reduces to the sum objective $(\sum_i c_i x_i)$ and for $k = 1$, the objective function reduces to the bottleneck objective $(\max_i c_i x_i)$.

The k-sum problem and in particular the minimum sum and bottleneck optimization problems have been extensively studied in literature. Clearly, the general problem in (1.1) is NP-hard since the minimum sum problem is already NP-hard. Gupta and Punnen [17] and Punnen and Aneja [25] showed that the solution to the k-sum problem can be obtained by solving $O(n)$ minimum sum problems. Thus the k-sum optimization problem can be solved in polynomial time whenever the associated minimum sum problem can be solved in polynomial time. Furthermore, Punnen and Aneja [25] showed that if the minimum sum problem has a polynomial time ε-approximation scheme, the k-sum problem also has a polynomial time ε-approximation scheme. As a special case of the k-sum optimization problem, the bottleneck optimization problem has been extensively studied in the literature. Gross [15] studied the problem of assigning jobs to parallel machines so as to minimize the largest completion time and modeled it as a bottleneck assignment problem. The problem is formulated as follows:

$$\min_{\mathbf{x} \in \mathcal{X}_{LAP}} \max_{i,j} c_{ij} x_{ij}, \tag{1.4}$$

where c_{ij} is the time to perform job i on machine j and the feasible set of the linear assignment problem

(LAP) is given as:

$$\mathcal{X}_{LAP} = \left\{ \boldsymbol{x} \;\middle|\; \sum_{i=1}^{n} x_{ij} = 1, \forall j \in \mathcal{N},\ \sum_{j=1}^{n} x_{ij} = 1, \forall i \in \mathcal{N},\ x_{ij} \in \{0,1\}, \forall i,j \in \mathcal{N} \right\}. \tag{1.5}$$

Gross [15] developed a threshold algorithm to solve this problem in polynomial time. Edmonds and Fulkerson [11] generalized this algorithm and the duality theory of blocking systems to a larger class of bottleneck combinatorial optimization problems. Other results in this area include an algorithm to solve the bottleneck location problem proposed by Hsu and Nemhauser [19] and a branch and bound algorithm to solve the bottleneck traveling salesperson problem by Carpaneto, Martello and Toth [9]. Gabow and Tarjan [12] developed polynomial time algorithms for the bottleneck spanning tree problem in a directed graph and the bottleneck maximum cardinality matching problem. The k-sum objective has been studied in a facility location problem by Slater [27] and Tamir [28]. In the p-facility, k-centrum problem, the objective is to locate p service facilities to minimize the sum of the k farthest nodes (customers) from the service facilities. Tamir [28] developed polynomial time algorithms to solve the p-facility k-centrum problem for path and tree graphs. Grygiel [16] studied the k-sum assignment problem and developed a polynomial algorithm to solve the problem.

In Section 2, we review the literature most relevant to our work from stochastic, robust and distributionally robust combinatorial optimization problems. In Section 3, we formulate the distributionally robust k-sum optimization problem. For the Fréchet class of distributions with discrete marginal distributions with finite support, we formulate the problem as a polynomial sized mixed integer linear program. As an implication of the formulation, we show that the distributionally robust k-sum optimization problem is solvable in polynomial time if the deterministic minimum sum problem is solvable in polynomial time. We also provide extensions of this result to develop a polynomial time ε-approximation guarantee for the distributionally robust k-sum optimization problem given a polynomial time ε-approximation guarantee of the deterministic minimum sum problem. Extension to instances where other distributional information is available is provided. We show that for simple uncertainty sets such as ellipsoids, the robust k-sum optimization problem can already be NP-hard even if the deterministic minimum sum problem is solvable in polynomial time. In Section 4, we discuss the submodularity structure of the k-sum objective function and show that this important combinatorial property is preserved under our choice of the class of distributions. We conclude in Section 5 with our results illustrating that the Fréchet class of distributions is important from a computational tractability perspective for this class of problems.

2 Literature Review

Consider the stochastic version of the combinatorial optimization problem where the costs are random. Denote the random cost vector as $\tilde{\mathbf{c}}$ with a probability distribution P. Problem (1.2) is then replaced by the stochastic optimization problem:

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n} \mathbb{E}_P\left(\max_{\mathbf{y}\in\mathcal{U}_k}\sum_{i=1}^{n}\tilde{c}_i x_i y_i\right), \tag{2.1}$$

where the expected sum of up to the k largest costs is minimized. Yechiali [31] studied the maximization version of the stochastic bottleneck assignment problem which is formulated as:

$$\max_{\mathbf{x}\in\mathcal{X}_{LAP}} \mathbb{E}_P\left(\min_{i,j:x_{ij}=1}\tilde{c}_{ij}\right). \tag{2.2}$$

For the cases where: (a) $\tilde{c}_{ij}$ are independent exponentially distributed random variables with means $1/\mu_{ij}$ and (b) $\tilde{c}_{ij}$ are independent Weibull distributed random variables with scale parameter μ_{ij} and shape parameter α, the assignment that maximizes the expected minimum cost is found by solving the deterministic minimum sum cost assignment problem with costs set to μ_{ij}. Other problems that have been studied include the stochastic bottleneck transportation problem (see Geetha and Nair [14]) and the stochastic bottleneck spanning tree problem (see Ishii and Nishida [20]).

An alternate approach that has received a lot of attention in the recent years to deal with optimization problems under uncertainty is robust optimization (see Ben-Tal, El Ghaoui and Nemirovskii [3]). In the standard robust optimization approach, the uncertain data $\mathbf{c}$ is assumed to belong to an uncertainty set Ω. However no additional distributional information is provided. The robust approach is to then minimize the worst-case cost as follows:

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n} \max_{\mathbf{c}\in\Omega}\left(\max_{\mathbf{y}\in\mathcal{U}_k}\sum_{i=1}^{n}c_i x_i y_i\right). \tag{2.3}$$

Two of the popular uncertainty sets that have been studied in combinatorial optimization problems are: (a) the interval uncertainty set where $\boldsymbol{c} \in [\underline{\boldsymbol{c}}, \overline{\boldsymbol{c}}]$ and (b) the scenario uncertainty set where $\boldsymbol{c} \in \left\{\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(m)}\right\}$. The reader is referred to Kouvelis and Yu [21] for results on robust discrete optimization. The interval uncertainty case reduces to solving the deterministic k-sum optimization at $\boldsymbol{c} = \overline{\boldsymbol{c}}$ since the worst-case is attained at this value. Thus solving the robust k-sum optimization is easy under interval uncertainty if solving the deterministic minimum sum problem is easy. The results for the scenario uncertainty set are however very different. Even the robust minimum sum shortest path problem with $m = 2$ scenarios is NP-hard (see Kouvelis and Yu [21]). However for the bottleneck optimization problem, Aissi, Bazgan and Vanderpooten [2] showed that the robust version with scenario uncertainty is easy if the deterministic bottleneck optimization problem is easy. Bertsimas and Sim [7] extended the interval uncertainty set by allowing for only a subset of the costs to deviate from the lower bound and proposed a mixed integer program to solve the robust optimization problem. Under this model, they showed that the robust counterpart for the minimum sum problem is solvable in polynomial time if the deterministic combinatorial optimization problem is solvable in polynomial time.

More recently, there has been a growing interest in an intermediate approach that lies between the stochastic and robust optimization approaches. In this approach, which is popularly referred to as *distributionally robust optimization*, the goal is to find the solution that minimizes the expected cost for the worst-case distribution P belonging to the set $\mathbb{P}$ of probability distributions. The distributionally robust optimization problem associated with (1.2) is formulated as:

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n} \max_{P\in\mathbb{P}} \mathbb{E}_P \left(\max_{\mathbf{y}\in\mathcal{U}_k} \sum_{i=1}^{n} \tilde{c}_i x_i y_i \right). \tag{2.4}$$

The main focus of research in distributionally robust optimization is to identify sets of distributions which are practical from a modeling perspective and yet provide computational tractability from an optimization perspective. The reader is referred to Bertsimas et. al [5], Delage and Ye [10], Zymler, Kuhn and Rustem [32] and, Wiesemann, Kuhn and Sim [29] for some of the recent results in this area. While most of the results focus on distributionally robust linear and convex optimization problems, fewer results are available for distributionally robust integer programs.

Of particular relevance to this paper is the Fréchet class of distributions where information on the marginal distributions of the random costs is provided but no information on the joint distribution is assumed. This provides a modeling approach to capture arbitrary dependence structure among the random variables without specifying them explicitly. Bertsimas, Natarajan and Teo [6] studied the problem of computing the worst-case expected cost $\max_{P\in\mathbb{P}} \mathbb{E}_P \left(\max_{\mathbf{y}\in\mathcal{U}} \sum_i \tilde{c}_i y_i\right)$ for general sets $\mathcal{U} \subseteq \{0,1\}^n$. Building on the results of Meilijson and Nadas [23], they reformulated the problem of computing the tight upper bound on the expected optimal objective value as a concave separable maximization problem over the convex hull of the set $\mathcal{U}$. As an implication of this result, they showed that given marginal distributions or a few marginal moments, the tight upper bound is computable in polynomial time if the corresponding deterministic combinatorial optimization problem is solvable in polynomial time. Natarajan, Shi and Toh [24] extended this model to minimizing the worst-case expected regret in combinatorial optimization problems under the assumption that the uncertainty set is in the form of a interval set along with information on marginal distributions. The distributionally robust counterpart was formulated in [24] as polynomial sized mixed-integer linear and conic programs and for the subset selection problem, polynomial time solvability of the problem was shown. Another relevant paper in this area is the work of Agrawal et. al. [1] who studied a general class of distributionally robust optimization problem of the form $\min_{\mathbf{x}\in\mathcal{X}} \max_{P\in\mathbb{P}} \mathbb{E}_P(h(\mathbf{x},\tilde{\mathbf{c}}))$ where $\mathbb{P}$ is the Fréchet class of distributions. For a given $\boldsymbol{x}$, they showed that for binary random variables the problem of computing the expected value under the worst-case distribution is NP-hard, even when restricted to functions $h(\mathbf{x},\cdot)$ that are monotone and submodular in the random variables. Their proposed approach to tackle the problem is to use an independent distribution to approximate the worst-case distribution and solve the corresponding stochastic problem using sampling based techniques or other algorithmic techniques. For problems such as the uncapacitated facility location, Steiner trees, and submodular function minimization such as in the bottleneck assignment problem, they showed that this approach approximates the robust model to a constant factor while it can be arbitrarily bad for supermodular functions. Another recent approach to tackle two-stage distributionally robust integer programs is the work of Hanasusanto, Wiesemann and Kuhn [18] who applied the notion of K-adaptability to develop mixed integer linear programming formulations which are shown to be tight under certain conditions on the set of distributions and the feasible region.

3 Distributionally Robust k-Sum Optimization

Associated with each random cost $\tilde{c}_i$ for $i \in \mathcal{N}$ is a marginal distribution P_i with support contained in the set $\Omega_i \subseteq [\underline{c}_i, \overline{c}_i]$. Denote the set of all possible joint distributions with support in $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ and marginal distributions P_i for $i \in \mathcal{N}$ by $\mathbb{P}(P_i,\ i \in \mathcal{N})$. This set of joint distributions is commonly known as the Fréchet class of distributions. One feasible distribution in this set is the independent (product) distribution given by $P_1 \times P_2 \times \ldots \times P_n \in \mathbb{P}(P_i,\ i \in \mathcal{N})$. The distributionally robust counterpart associated with (1.2) for the Fréchet class of distributions is formulated as:

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n} \max_{P\in\mathbb{P}(P_i,\, i\in\mathcal{N})} \mathbb{E}_P\left(\max_{\mathbf{y}\in\mathcal{U}_k} \sum_{i=1}^{n} \tilde{c}_i x_i y_i\right). \tag{3.1}$$

Below we state a key result that we use in solving (3.1), the proof of which can be found in Meilijson and Nadas [23], Bertsimas, Natarajan and Teo [6] and Natarajan, Shi and Toh [24].

Theorem 1. *For each $i \in \mathcal{N}$, assume that the marginal distribution P_i of the random variable $\tilde{c}_i$ with* *support* $\Omega_i \subseteq [\underline{c}_i, \bar{c}_i]$ *is given. For* $\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n$*, define the worst-case expected cost:*

$$Z(\mathbf{x}) = \max_{P \in \mathbb{P}(P_i, i \in \mathcal{N})} \mathbb{E}_P \left(\max_{\mathbf{y} \in \mathcal{U}_k \subseteq \{0,1\}^n} \sum_{i=1}^n \tilde{c}_i x_i y_i \right). \tag{3.2}$$

Let:

$$\bar{Z}(\mathbf{x}) = \min_{\mathbf{w} \in \mathbb{R}^n} \left(\max_{\mathbf{y} \in \mathcal{U}_k \subseteq \{0,1\}^n} \sum_{i=1}^n w_i y_i + \sum_{i=1}^n \mathbb{E}_{P_i}[\tilde{c}_i x_i - w_i]^+ \right), \tag{3.3}$$

where $[z]^+ = \max(z, 0)$ *for* $z \in \mathbb{R}$*. Then* $Z(\mathbf{x}) = \bar{Z}(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n$*.*

Since the constraint matrix in the $\mathcal{U}_k = \left\{ \mathbf{y} \;\middle|\; \sum_i y_i \leq k,\ y_i \in \{0,1\}, \forall i \right\}$ is totally unimodular, the inner maximization problem in (3.3) is solvable as a compact linear program as follows (see Schrijver [26]):

$$\max_{\mathbf{y} \in \mathcal{U}_k} \mathbf{w}^T \mathbf{y} = \max \left\{ \mathbf{w}^T \mathbf{y} \;\middle|\; \sum_{i=1}^n y_i \leq k,\ 0 \leq y_i \leq 1, \forall i \in \mathcal{N} \right\}. \tag{3.4}$$

Corresponding to the linear program (3.4), the associated dual linear program is:

$$\max_{\mathbf{y} \in \mathcal{U}_k} \mathbf{w}^T \mathbf{y} \quad = \quad \min \left\{ k\lambda_0 + \sum_{i=1}^n \lambda_i \;\middle|\; \lambda_0 + \lambda_i \geq w_i, \forall i \in \mathcal{N},\ \lambda_0 \geq 0,\ \lambda_i \geq 0, \forall i \in \mathcal{N} \right\}. \tag{3.5}$$

Combining Theorem 1 and the dual formulation (3.5), we can compute $Z(\mathbf{x})$ as:

$$\begin{aligned} Z(\mathbf{x}) \quad = \quad \min \quad & k\lambda_0 + \sum_{i=1}^n \lambda_i + \sum_{i=1}^n \mathbb{E}_{P_i}[\tilde{c}_i x_i - w_i]^+ \\ \text{s.t.} \quad & \lambda_0 + \lambda_i \geq w_i, && \forall i \in \mathcal{N}, \\ & \lambda_0 \geq 0, \\ & \lambda_i \geq 0, && \forall i \in \mathcal{N}, \\ & \mathbf{w} \in \mathbb{R}^n. \end{aligned} \tag{3.6}$$

It is easy to check that at optimality, the value λ_i is equal to $[w_i - \lambda_0]^+$. The distributionally robust k-sum optimization problem is then solvable as:

$$\min_{\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n, \lambda_0 \geq 0, \boldsymbol{w} \in \mathbb{R}^n} \left(k\lambda_0 + \sum_{i=1}^n [w_i - \lambda_0]^+ + \sum_{i=1}^n \mathbb{E}_{P_i}[\tilde{c}_i x_i - w_i]^+ \right). \tag{3.7}$$

For a fixed $\boldsymbol{x} \in \mathcal{X} \subseteq \{0,1\}^n$ and $\lambda_0 \geq 0$, the objective function is separable in the w_i variables. The optimal solution for each i is obtained at $w_i = \lambda_0$ as follows:

$$\min_{w_i \in \mathbb{R}} ([w_i - \lambda_0]^+ + \mathbb{E}_{P_i}[\tilde{c}_i x_i - w_i]^+) = \mathbb{E}_{P_i}[\tilde{c}_i x_i - \lambda_0]^+. \tag{3.8}$$

This can be verified by observing that for $w_i \leq \lambda_0$, we can increase the w_i to the value λ_0 keeping the first term unchanged while possibly reducing the second, while for $w_i \geq \lambda_0$, we can decrease w_i to the value λ_0 decreasing the first term at a rate of 1 while possibly increasing the second term at most a rate of 1. Thus, we can reformulate the distributionally robust k-sum optimization problem as:

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n,\lambda_0\geq 0}\left(k\lambda_0+\sum_{i=1}^{n}\mathbb{E}_{P_i}[\tilde{c}_i x_i-\lambda_0]^+\right). \tag{3.9}$$

Equivalently, this problem can be formulated as:

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n,\lambda_0\geq 0}\left(k\lambda_0+\sum_{i=1}^{n}\mathbb{E}_{P_i}[\tilde{c}_i-\lambda_0]^+ x_i\right), \tag{3.10}$$

where the equivalence comes from $x_i \in \{0,1\}$ and $\lambda_0 \geq 0$. For $k = n$, $\lambda_0 = 0$ is optimal in (3.10) and the distributionally robust k-sum optimization problem reduces to $\min_{\mathbf{x}\in\mathcal{X}} \boldsymbol{\mu}^T\mathbf{x}$ where $\boldsymbol{\mu} = \mathbb{E}(\tilde{\mathbf{c}})$. This corresponds to solving the deterministic minimum sum optimization problem with the costs set to mean.

3.1 Construction of Worst-Case Distribution

In this section, we provide a construction of the worst-case joint distribution and illustrate with a simple example that the optimal solution from the distributionally robust optimization problem can be different from the optimal solution with independent distributions. For a given $\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n$, the worst-case expected cost $Z(\mathbf{x})$ is given as:

$$Z(\mathbf{x}) = \min_{\lambda_0\geq 0}\left(k\lambda_0+\sum_{i=1}^{n}\mathbb{E}_{P_i}[\tilde{c}_i-\lambda_0]^+ x_i\right). \tag{3.11}$$

We now provide an explicit construction of the joint distribution that attains the bound. Consider the case where the marginal distributions of $\tilde{\mathbf{c}}$ are discrete with finite support. Let $\{c_i^{(j)}, j \in \mathcal{J}_i\}$ denote the set of values taken by the random variable $\tilde{c}_i$ where $J_i = |\mathcal{J}_i|$ is the size of the support set. Furthermore, let p_{ij} denote the probability that $\tilde{c}_i$ takes the value $c_i^{(j)}$ for $j \in \mathcal{J}_i$. Note that $p_{ij} \geq 0$ and $\sum_j p_{ij} = 1$. Then the worst-case expected cost in (3.11) is computed as the optimal objective value to the linear program:

$$\begin{aligned} Z(\mathbf{x}) \quad = \quad \min \quad & k\lambda_0+\sum_{i=1}^{n}\sum_{j=1}^{J_i} t_{ij}p_{ij} \\ \text{s.t.} \quad & t_{ij} \geq c_i^{(j)}x_i - \lambda_0 x_i, \quad \forall j \in \mathcal{J}_i, \forall i \in \mathcal{N}, \\ & t_{ij} \geq 0, \quad \forall j \in \mathcal{J}_i, \forall i \in \mathcal{N}, \\ & \lambda_0 \geq 0. \end{aligned} \tag{3.12}$$

The dual linear program is formulated as:

$$
\begin{aligned}
Z(\mathbf{x}) \quad = \quad \max \quad & \sum_{i=1}^{n} \sum_{j=1}^{J_i} c_i^{(j)} \alpha_{ij} x_i \\
\text{s.t.} \quad & \sum_{i=1}^{n} \sum_{j=1}^{J_i} \alpha_{ij} x_i \leq k, \\
& \alpha_{ij} \leq p_{ij}, \qquad\qquad \forall j \in \mathcal{J}_i, \forall i \in \mathcal{N}, \\
& \alpha_{ij} \geq 0, \qquad\qquad \forall j \in \mathcal{J}_i, \forall i \in \mathcal{N},
\end{aligned}
\tag{3.13}
$$

where strong duality holds since the primal problem is a linear program. Let α^* denote the optimal vector for the dual linear program (3.13). Define $y_i^* = \sum_j \alpha_{ij}^* x_i$ for each i. Clearly, $\mathbf{y}^*$ lies in the the convex hull of the set $\mathcal{U}_k$ which is defined by the polytope $conv(\mathcal{U}_k) = \left\{ \mathbf{y} \;\middle|\; \sum_i y_i \leq k,\ y_i \in [0,1], \forall i \in \mathcal{N} \right\}$. This implies, that we can express $\mathbf{y}^*$ as a convex combination of the extreme points of the polytope as follows:

$$
\begin{aligned}
\sum_{\mathbf{y} \in ext(conv(\mathcal{U}_k))} \beta_{\mathbf{y}} \mathbf{y} \quad &= \quad \mathbf{y}^*, \\
\sum_{\mathbf{y} \in ext(conv(\mathcal{U}_k))} \beta_{\mathbf{y}} \quad &= \quad 1, \\
\beta_{\mathbf{y}} \quad &\geq \quad 0,
\end{aligned}
\tag{3.14}
$$

where $ext(\mathcal{S})$ is the set of extreme points of the convex set $\mathcal{S}$. We now construct the extremal distribution using a two-step procedure as follows:

(a) Choose randomly the extreme points $\mathbf{y} \in ext(conv(\mathcal{U}_k))$ with probabilities $\beta_{\mathbf{y}}$.

(b) Generate the random variables for each $\tilde{c}_i$ independently as follows: If $y_i = 1$, generate the random variable $\tilde{c}_i$ as follows:

$$\tilde{c}_i = c_i^{(j)} \text{ with probability } \frac{\alpha_{ij}}{\sum_j \alpha_{ij}}, \ \forall j \in \mathcal{J}_i,$$

and if $y_i = 0$, generate the random variable $\tilde{c}_i$ as follows:

$$\tilde{c}_i = c_i^{(j)} \text{ with probability } \frac{p_{ij} - \alpha_{ij}}{\sum_j (p_{ij} - \alpha_{ij})}, \ \forall j \in \mathcal{J}_i.$$

The proof of tightness of this distribution can be found in Bertsimas, Natarajan and Teo [6] and Natarajan, Shi and Toh [24]. We next provide a simple example to illustrate the construction of the distribution and to compare the optimal robust solution with the solution with independent distributions.

Example 1. *Consider a project selection problem where a set of projects is defined as* $\mathcal{N} = \{1, 2, 3\}$*. The constraints are (a) projects 1 and 2 must be jointly selected if chosen and (b) either project 1 or project 3 must be chosen. The feasible region is given as:*

$$\mathcal{X} = \left\{ \boldsymbol{x} \;\middle|\; x_1 - x_2 = 0,\ x_1 + x_3 = 1,\ x_1, x_2, x_3 \in \{0, 1\} \right\}.$$

There are two feasible solutions in this case with $(x_1, x_2, x_3) = (1, 1, 0)$ *(select projects 1 and 2) and* $(0, 0, 1)$ *(select project 3). Assume that the duration of projects 1 and 2 are random while the duration of project 3 is deterministic with the marginal distribution of the duration given as follows:*

$$\tilde{c}_1 = \begin{cases} 0, & \textit{with probability } \frac{1}{2}, \\ 1, & \textit{with probability } \frac{1}{2}, \end{cases} \quad \tilde{c}_2 = \begin{cases} 0, & \textit{with probability } \frac{1}{2}, \\ 1, & \textit{with probability } \frac{1}{2}. \end{cases} \quad \tilde{c}_3 = \begin{cases} c_3, & \textit{with probability } 1. \end{cases}$$

The objective is to choose a set of projects that satisfies the constraints so as to minimize the maximum duration of the project in the chosen set (bottleneck objective).

(a) For independent project durations, the joint distribution is given as:

$$(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3) = \begin{cases} (0, 0, c_3), & \textit{with probability } \frac{1}{4}, \\ (0, 1, c_3), & \textit{with probability } \frac{1}{4}, \\ (1, 0, c_3), & \textit{with probability } \frac{1}{4}, \\ (1, 1, c_3), & \textit{with probability } \frac{1}{4}. \end{cases}$$

If $\mathbf{x} = (1, 1, 0)$*, the project completion time is given as:*

$$\max(\tilde{c}_1, \tilde{c}_2) = \begin{cases} 0, & \textit{with probability } \frac{1}{4}, \\ 1, & \textit{with probability } \frac{3}{4}, \end{cases}$$

with an expected project completion time of $3/4$*. If* $\mathbf{x} = (0, 0, 1)$*, the project completion time is deterministic and equal to* c_3*. The optimal solution is to choose projects 1 and 2 if* $c_3 > 3/4$ *and to choose project 3 if* $c_3 < 3/4$*.*

(b) We now consider the worst-case expected cost using the proposed model. If $\mathbf{x} = (1, 1, 0)$ *from Theorem 1 and (3.13), the worst-case expected project completion time is computed by solving:*

$$\begin{aligned} Z(\{1,1,0\}) \quad = \quad \max \quad & \alpha_{12} + \alpha_{22} \\ \textit{s.t.} \quad & \alpha_{11} + \alpha_{12} + \alpha_{21} + \alpha_{22} \le 1, \\ & \alpha_{11} \le \frac{1}{2},\ \alpha_{12} \le \frac{1}{2},\ \alpha_{21} \le \frac{1}{2},\ \alpha_{22} \le \frac{1}{2},\ \alpha_{31} \le 1 \\ & \alpha_{11} \ge 0,\ \alpha_{12} \ge 0,\ \alpha_{21} \ge 0,\ \alpha_{22} \ge 0,\ \alpha_{31} \ge 0. \end{aligned} \tag{3.15}$$

The optimal solution to the dual linear program (3.15) is clearly $\alpha_{12}^* = \alpha_{22}^* = 1$*,* $\alpha_{11}^* = \alpha_{21}^* = 0$ *with* $\alpha_{31}^* \in [0,1]$*. This corresponds to* $\mathbf{y}^* = (1/2, 1/2, 0)$ *with* $\beta_{(0,1,0)} = \beta_{(1,0,0)} = 1/2$*. We then construct the worst-case joint distribution from the previous construction as follows:*

$$(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3) = \begin{cases} (0, 1, c_3), & \textit{with probability } \frac{1}{2}, \\ (1, 0, c_3), & \textit{with probability } \frac{1}{2}. \end{cases}$$

The distribution of the durations of projects 1 and 2 is perfectly negatively correlated in this case. The project completion time in this instance is deterministic and equal to 1*. The optimal solution is to choose projects 1 and 2 if* $c_3 > 1$ *and to choose project 3 if* $c_3 < 1$*. This simple example already illustrates the difference in the optimal solutions for* $c_3 \in [3/4, 1]$ *under the independent distribution and the worst-case distribution.*

3.2 Polynomial Time Solvable Instances and ε-Approximation Guarantees

In this section, we start by formulating the distributionally robust k-sum optimization problem as a mixed integer linear program for discrete marginal distributions. From formulation (3.9), we can rewrite the problem as follows:

$$\begin{aligned} \min \quad & k\lambda_0 + \sum_{i=1}^{n} \sum_{j=1}^{J_i} t_{ij} p_{ij} \\ \text{s.t.} \quad & \lambda_0 \geq 0, \\ & t_{ij} \geq c_i^{(j)} x_i - \lambda_0, \qquad \forall j \in \mathcal{J}_i, \forall i \in \mathcal{N}, \\ & t_{ij} \geq 0, \qquad\qquad\quad\;\; \forall j \in \mathcal{J}_i, \forall i \in \mathcal{N}, \\ & \mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n. \end{aligned} \tag{3.16}$$

The size of the mixed integer linear program is polynomial in the total number of support points for each marginal distribution and the total number of random variables. We next identify a class of distributionally robust k-sum optimization problems that is solvable in polynomial time. We focus on $1 \leq k < n$ and show that the optimal λ_0 can be restricted to a finite set of points.

Proposition 1. *Assume that the marginal distribution of* $\tilde{c}_i$ *is discrete and* $\tilde{c}_i = c_i^{(j)}$ *with probability* $p_{ij}, j \in \mathcal{J}_i, i \in \mathcal{N}$*. Then the objective function for (3.10) attains its minimum for the* λ_0 *variable in the finite set:*

$$\lambda_0 \in \{0\} \cup \{c_i^{(j)} \;\big|\; j \in \mathcal{J}_i, i \in \mathcal{N}\}.$$

Proof. For the discrete marginal distributions, problem (3.10) can be written as

$$\min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n} Z(\mathbf{x}) \quad = \quad \min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n} \min_{\lambda_0\geq 0}\left(k\lambda_0 + \sum_{i=1}^{n}\sum_{j=1}^{J_i}[c_i^{(j)}-\lambda_0]^+ p_{ij}x_i\right).$$

For fixed $\mathbf{x}\in\mathcal{X}$, the objective function in (3.10) is a piecewise linear convex function in λ_0. Since the objective value goes to infinity as λ_0 goes to $+\infty$, hence its minimum value occurs at one of the break points $\{0\}\cup\{c_i^{(j)} : j\in\mathcal{J}_i,\ i\in\mathcal{N}\}$. □

Define $\Lambda_0 = \{0\}\cup\{c_i^{(j)} \mid j\in\mathcal{J}_i, i\in[N]\}$, and

$$h_i(\lambda_0) \quad = \quad \sum_{j=1}^{J_i}[c_i^{(j)}-\lambda_0]^+ p_{ij}, \forall i\in\mathcal{N}, \tag{3.17}$$

$$h_0(\lambda_0) \quad = \quad k\lambda_0. \tag{3.18}$$

Then problem (3.17) can be rewritten as

$$\min_{\lambda_0\in\Lambda_0}\ \min_{\mathbf{x}\in\mathcal{X}\subseteq\{0,1\}^n}\sum_{i=1}^{n} h_i(\lambda_0)x_i + h_0(\lambda_0). \tag{3.19}$$

Hence we can solve the distributionally robust k-sum optimization problem as follows: For each $\lambda_0\in\Lambda_0$, solve the deterministic minimum sum optimization problem $g(\lambda_0) = \min_{\mathbf{x}\in\mathcal{X}}\sum_i h_i(\lambda_0)x_i$ where $h_i(\lambda_0)$ is defined as in (3.17). Then select the smallest value from $\{g(\lambda_0)+h_0(\lambda_0) : \lambda_0\in\Lambda_0\}$, where $h_0(\lambda_0)$ is defined as in (3.18). Therefore, we have the following theorem:

Theorem 2. *If the deterministic minimum sum optimization problem is polynomial time solvable, then under the Fréchet class of discrete distributions with finite support, the distributionally robust k-sum optimization problem is solvable in polynomial time.*

Remark 1. *As an implication of Theorem 2, consider the distributionally robust bottleneck matching problem studied in Agrawal et. al. [1]. In this problem the objective is to find a perfect matching of minimum worst-case expected cost where the cost of a matching is determined by the most expensive edge in the matching. Agrawal et. al. [1] showed that if the dependency structure is unknown, the random variables can be assumed to be independent to get $e/(e-1)$ approximation for the distributionally robust problem. Our result shows that the distributionally robust problem in fact can be solved in polynomial time in this case.*

We now extend the result to instances where the deterministic minimum sum problem has a polynomial time ε-approximation guarantee and show that the distributionally robust k-sum optimization problem preserves the polynomial time ε-approximation guarantee. For any $\lambda_0 \in \Lambda_0$, let $\bar{\boldsymbol{x}}(\lambda_0) \in \mathcal{X}$ be the ε-approximate solution for the minimum sum optimization problem $g(\lambda_0) = \min_{\mathbf{x}\in\mathcal{X}} \sum_i h_i(\lambda_0)x_i$, namely:

$$g(\lambda_0) \leq \sum_{i=1}^{n} h_i(\lambda_0)\bar{x}_i(\lambda_0) \leq (1+\varepsilon)g(\lambda_0). \tag{3.20}$$

Let $\lambda_0^* \in \Lambda_0$ be the minimizer of $\sum_{i=1}^{n} h_i(\lambda_0)\bar{x}_i(\lambda_0) + h_0(\lambda_0)$ over Λ_0 which along with inequality (3.20) implies that:

$$\sum_{i=1}^{n} h_i(\lambda_0^*)\bar{x}_i(\lambda_0^*) + h_0(\lambda_0^*) \leq (1+\varepsilon) \min_{\lambda_0\in\Lambda_0} \left(g(\lambda_0) + h_0(\lambda_0)\right). \tag{3.21}$$

This leads to the following theorem.

Theorem 3. *If the deterministic minimum sum optimization problem has a polynomial time ε-approximation, then under the Fréchet class of discrete distributions with finite support, the distributionally robust k-sum optimization problem has a polynomial time ε-approximation.*

3.3 Extensions to Other Distribution Sets

In the previous section, the polynomial solvability of the distribution robust k-sum optimization problem was discussed given the marginal distribution P_i of the random variable $\tilde{c}_i,\ i \in \mathcal{N}$ with finite, discrete support. In this section, we extend the polynomial solvability result to other distribution sets involving limited information on marginal moments. We also consider joint information on the parameters captured through an ellipsoid set and show that the problem becomes NP-hard in this case. Specifically, we show polynomial time solvability for two cases: (a) the range, mean and (b) the range, mean and mean absolute deviation. In these cases, the univariate moment bounds $\sup_{P_i\in\mathbb{P}_i} \mathbb{E}_{P_i}[\tilde{c}_i - \lambda_0]^+$ are known in closed form (see Madansky [22], and Ben-Tal and Hochman [4]) wherein the extremal univariate distribution reduces to a two-point or three-point discrete distribution respectively.

(a) **Range and mean:** For known range $\Omega_i = [\underline{c}_i, \bar{c}_i]$ and mean μ_i, the marginal moment model under consideration is:

$$\mathbb{P}_i = \left\{P_i : \mathbb{E}_{P_i}(\tilde{c}_i) = \mu_i,\ \mathbb{E}_{P_i}[\mathbb{I}_{\Omega_i}(\tilde{c})] = 1\right\}. \tag{3.22}$$

The extremal distribution for the problem $\sup_{P_i\in\mathbb{P}_i}[\tilde{c}_i - \lambda_0]^+$ is given by the following two-point discrete distribution:

$$\tilde{c}_i = \begin{cases} \underline{c}_i, & \text{with } P_i(\tilde{c}_i = \underline{c}_i) = \dfrac{\bar{c}_i - \mu_i}{\bar{c}_i - \underline{c}_i} \\ \bar{c}_i, & \text{with } P_i(\tilde{c}_i = \bar{c}_i) = \dfrac{\mu_i - \underline{c}_i}{\bar{c}_i - \underline{c}_i} \end{cases}$$

(b) **Range, mean and mean absolute deviation:** For known range $\Omega_i = [\underline{c}_i, \bar{c}_i]$, mean μ_i and mean absolute deviation δ_i, the marginal moment model under consideration is

$$\mathbb{P}_i = \{P_i : \mathbb{E}_{P_i}(\tilde{c}_i) = \mu_i,\ \mathbb{E}_{P_i}(|\tilde{c}_i - \mu_i|) = \delta_i,\ \mathbb{E}_{P_i}[\mathbb{I}_{\Omega_i}(\tilde{c})] = 1\}\,. \tag{3.23}$$

The extremal distribution for the problem $\sup_{P_i \in \mathbb{P}_i}[\tilde{c}_i - \lambda_0]^+$ is given by the following three-point discrete distribution:

$$\tilde{c}_i = \begin{cases} \underline{c}_i, & \text{with } P_i(\tilde{c}_i = \underline{c}_i) = \dfrac{\delta_i}{2(\mu_i - \underline{c}_i)} =: p_i \\ \bar{c}_i, & \text{with } P_i(\tilde{c}_i = \bar{c}_i) = \dfrac{\delta_i}{2(\bar{c}_i - \mu_i} =: q_i \\ \mu_i, & \text{with } P_i(\tilde{c}_i = \mu_i) = 1 - p_i - q_i \end{cases}$$

In the case, when the marginal information available is the range $\Omega_i = [\underline{c}_i, \bar{c}_i]$, mean μ_i and standard deviation σ_i, the extremal distribution for the univariate problem depends on the parameter λ_0 (see Birge and Maddox [8]). Characterizing the complexity of the distributionally robust k-sum optimization problem in this case remains an open question.

The distributionally robust k-sum optimization problem for models (a) and (b) is given as:

$$\min_{\mathbf{x} \in \mathcal{X} \subseteq \{0,1\}^n, \lambda_0 \geq 0} \left(k\lambda_0 + \sum_{i=1}^{n} \sup_{P_i \in \mathbb{P}_i} \mathbb{E}_{P_i}[\tilde{c}_i x_i - \lambda_0]^+ \right), \tag{3.24}$$

where the inner supremum is over the set of univariate distributions with the given information (see [24]). Since the worst-case univariate distribution in (3.22) and (3.23) is independent of the λ_0 variable, we can directly apply Theorem 2 using the two and three point marginal distributions to obtain the following result.

Theorem 4. *Assume that for each random variable $\tilde{c}_i$, $i \in \mathcal{N}$ with support $\Omega_i = [\underline{c}_i, \bar{c}_i]$, the available information is either: (a) the range, mean or (b) the range, mean and mean absolute deviation is given. If the deterministic minimum sum optimization problem is solvable in polynomial time, then the distributionally robust k-sum problem (3.24) is solvable in polynomial time for both cases (a) and (b).*

We end this section by showing that with even a simple uncertainty set such as an ellipsoid, the robust minimum sum problem is already NP-hard. Consider a set of distributions defined as:

$$\mathbb{P} = \{P : P(\tilde{\mathbf{c}} \in \mathscr{C}) = 1\} \qquad \text{where} \quad \mathscr{C} = \left\{\mathbf{c} : \|\mathbf{\Sigma}^{1/2}\mathbf{c}\| \leq 1\right\}, \tag{3.25}$$

where the matrix $\mathbf{\Sigma} \succ 0$. For $k = n$ with the set of distribution $\mathbb{P}$ in (3.25) and the set $\mathcal{X} = \{0,1\}^n$, the distributionally robust minimum sum optimization problem (3.1) reduces to:

$$\min_{\mathbf{x} \in \{0,1\}^n} \max_{\mathbf{c} \in \mathscr{C}} \mathbf{c}^T \mathbf{x}. \tag{3.26}$$

The inner maximization problem is solvable in closed form with the optimal cost vector given by:

$$\mathbf{c}^* = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}{\sqrt{\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}}. \tag{3.27}$$

The robust minimum sum in (3.26) then reduces to:

$$\min_{\mathbf{x}\in\{0,1\}^n} \sqrt{\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}, \tag{3.28}$$

which is equivalent to solving the optimization problem $\min_{\mathbf{x}\in\{0,1\}^n} \boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}$. This is the quadratic unconstrained binary optimization problem which is known to be NP-hard even for positive definite matrices $\boldsymbol{\Sigma}$ (see Garey and Johnson [13]).

4 Submodularity of the Distributionally Robust k-Sum Objective

In this section, we study the submodularity property of the k-sum objective function and show that this extends to the distributionally robust objective. For a given cost vector $\mathbf{c} \in \mathbb{R}^n_+$ and a set $\mathcal{S} \subseteq \mathcal{N}$, define:

$$Z(\mathcal{S};\mathbf{c}) = \max\left\{\sum_{i\in\mathcal{S}} c_i y_i \;\middle|\; \sum_{i\in\mathcal{S}} y_i \le k,\ y_i \in \{0,1\}, \forall i \in \mathcal{S}\right\},$$

where $Z(\emptyset;\mathbf{c}) = 0$. The function $Z(\mathcal{S};\mathbf{c})$ is the weighted rank function on an uniform matroid and is known to satisfy the following two conditions:

(a) Monotone: For every $\mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{N}$, we have that $Z(\mathcal{T};\mathbf{c}) \le Z(\mathcal{S};\mathbf{c})$.

(b) Submodular: For every $\mathcal{S},\mathcal{T} \subseteq \mathcal{N}$, we have that $Z(\mathcal{S};\mathbf{c}) + Z(\mathcal{T};\mathbf{c}) \ge Z(\mathcal{S}\cap\mathcal{T};\mathbf{c}) + Z(\mathcal{S}\cup\mathcal{T};\mathbf{c})$.

From the definition of the expectation of a random variable (as a non-negative weighted combination), it is easy to see that these two properties extend to the expected weighted rank function on an uniform matroid with random $\tilde{\mathbf{c}}$ as follows:

(a) Monotone: For every $\mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{N}$, we have that $\mathbb{E}[Z(\mathcal{T};\tilde{\mathbf{c}})] \le \mathbb{E}[Z(\mathcal{S};\tilde{\mathbf{c}})]$.

(b) Submodular: For every $\mathcal{S},\mathcal{T} \subseteq \mathcal{N}$, we have that $\mathbb{E}[Z(\mathcal{S};\mathbf{c})]+\mathbb{E}[Z(\mathcal{T};\mathbf{c})] \ge \mathbb{E}[Z(\mathcal{S}\cap\mathcal{T};\mathbf{c})]+\mathbb{E}[Z(\mathcal{S}\cup\mathcal{T};\mathbf{c})]$.

We now show that these properties extends to the distributionally robust objective function:

$$Z(\mathcal{S}) = \max_{P\in\mathbb{P}_{\mathcal{S}}} \mathbb{E}_P\left[\max\left\{\sum_{i\in\mathcal{S}} \tilde{c}_i y_i \;\middle|\; \sum_{i\in\mathcal{S}} y_i \le k,\ y_i \in \{0,1\}, \forall i \in \mathcal{S}\right\}\right]. \tag{4.1}$$

Note that while submodularity is preserved under expectation, it is generally not preserved under maximization. We next show that in the special case of Fréchet class of distributions, this important property still holds.

Theorem 5. *The function $Z(\mathcal{S})$ in (4.1) is monotone and submodular.*

Proof. From the results in Section 3, we can compute the distributionally robust objective in (4.1) as the optimal objective value to the problem:

$$Z(\mathcal{S}) = \min_{\lambda_0 \geq 0} \left(k\lambda_0 + \sum_{i \in \mathcal{S}} \mathbb{E}_{P_i}[\tilde{c}_i - \lambda_0]^+ \right).$$

Let $f_i(\lambda_0) = \mathbb{E}_{P_i}[\tilde{c}_i - \lambda_0]^+$. Then clearly, $f_i(\lambda_0)$ is a non-negative, convex and non-increasing function in λ_0. We next show the the two properties:

(a) To show that $Z(\cdot)$ is monotone:

Consider $\mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{N}$. Let $\lambda_{\mathcal{S}} \in \text{argmin}\left\{k\lambda_0 + \sum_{i \in \mathcal{S}} f_i(\lambda_0) \mid \lambda_0 \geq 0\right\}$. Then, we have:

$$\begin{aligned}
Z(\mathcal{T}) &= \min_{\lambda_0 \geq 0} \left(k\lambda_0 + \sum_{i \in \mathcal{T}} f_i(\lambda_0) \right) \\
&\leq k\lambda_{\mathcal{S}} + \sum_{i \in \mathcal{T}} f_i(\lambda_{\mathcal{S}}) && \text{[Since } \lambda_{\mathcal{S}} \text{ is feasible]} \\
&= k\lambda_{\mathcal{S}} + \sum_{i \in \mathcal{S}} f_i(\lambda_{\mathcal{S}}) - \sum_{i \in \mathcal{S} \setminus \mathcal{T}} f_i(\lambda_{\mathcal{S}}) \\
&= Z(\mathcal{S}) - \sum_{i \in \mathcal{S} \setminus \mathcal{T}} f_i(\lambda_{\mathcal{S}}) && \text{[Since } Z(\mathcal{S}) = k\lambda_{\mathcal{S}} + \textstyle\sum_{i \in \mathcal{S}} f_i(\lambda_{\mathcal{S}})\text{]} \\
&\leq Z(\mathcal{S}) && \text{[Since } f_i(\cdot) \text{ is non-negative].}
\end{aligned}$$

(b) To show that $Z(\cdot)$ is submodular:

Let $\lambda_{\mathcal{S}} \in \text{argmin}\left\{k\lambda_0 + \sum_{i \in \mathcal{S}} f_i(\lambda_0) \mid \lambda_0 \geq 0\right\}$ and $\lambda_{\mathcal{T}} \in \text{argmin}\left\{k\lambda_0 + \sum_{i \in \mathcal{T}} f_i(\lambda_0) \mid \lambda_0 \geq 0\right\}$. Set $\lambda_{\mathcal{S} \cap \mathcal{T}} = \alpha\lambda_{\mathcal{S}} + (1-\alpha)\lambda_{\mathcal{T}}$ and $\lambda_{\mathcal{S} \cup \mathcal{T}} = (1-\alpha)\lambda_{\mathcal{S}} + \alpha\lambda_{\mathcal{T}}$ for some $\alpha \in [0, 1]$. Then, we have:

$$\begin{aligned}
Z(\mathcal{S} \cap \mathcal{T}) + Z(\mathcal{S} \cup \mathcal{T}) &\leq k\lambda_{\mathcal{S} \cap \mathcal{T}} + k\lambda_{\mathcal{S} \cup \mathcal{T}} + \sum_{i \in \mathcal{S} \cap \mathcal{T}} f_i(\lambda_{\mathcal{S} \cap \mathcal{T}}) + \sum_{i \in \mathcal{S} \cup \mathcal{T}} f_i(\lambda_{\mathcal{S} \cup \mathcal{T}}) \\
&\quad \text{[Since } \lambda_{\mathcal{S} \cap \mathcal{T}} \text{ and } \lambda_{\mathcal{S} \cup \mathcal{T}} \text{ are feasible]} \\
&= k\lambda_{\mathcal{S}} + k\lambda_{\mathcal{T}} + \sum_{i \in \mathcal{S} \cap \mathcal{T}} f_i(\lambda_{\mathcal{S} \cap \mathcal{T}}) + \sum_{i \in \mathcal{S} \cup \mathcal{T}} f_i(\lambda_{\mathcal{S} \cup \mathcal{T}}) \\
&\quad \text{[Since } \lambda_{\mathcal{S} \cap \mathcal{T}} + \lambda_{\mathcal{S} \cup \mathcal{T}} = \lambda_{\mathcal{S}} + \lambda_{\mathcal{T}}\text{]} \\
&\leq k\lambda_{\mathcal{S}} + k\lambda_{\mathcal{T}} + \sum_{i \in \mathcal{S} \cap \mathcal{T}} (\alpha f_i(\lambda_{\mathcal{S}}) + (1-\alpha) f_i(\lambda_{\mathcal{T}})) \\
&\quad + \sum_{i \in \mathcal{S} \cup \mathcal{T}} ((1-\alpha) f_i(\lambda_{\mathcal{S}}) + \alpha f_i(\lambda_{\mathcal{T}})) \\
&\quad \text{[From the convexity of } f_i(\cdot)\text{]}
\end{aligned}$$

that is,

$$
\begin{aligned}
Z(\mathcal{S}\cap\mathcal{T})+Z(\mathcal{S}\cup\mathcal{T}) &= k\lambda_{\mathcal{S}}+k\lambda_{\mathcal{T}}+\sum_{i\in\mathcal{S}\cap\mathcal{T}}(f_i(\lambda_{\mathcal{S}})+f_i(\lambda_{\mathcal{T}})) \\
&\quad+\sum_{i\in\mathcal{S}\setminus\mathcal{T}}((1-\alpha)f_i(\lambda_{\mathcal{S}})+\alpha f_i(\lambda_{\mathcal{T}}))+\sum_{i\in\mathcal{T}\setminus\mathcal{S}}((1-\alpha)f_i(\lambda_{\mathcal{S}})+\alpha f_i(\lambda_{\mathcal{T}})) \\
&= Z(\mathcal{S})+Z(\mathcal{T})-\alpha\sum_{i\in\mathcal{S}\setminus\mathcal{T}}(f_i(\lambda_{\mathcal{S}})-f_i(\lambda_{\mathcal{T}})) \\
&\quad+(1-\alpha)\sum_{i\in\mathcal{T}\setminus\mathcal{S}}(f_i(\lambda_{\mathcal{S}})-f_i(\lambda_{\mathcal{T}})) \\
&\quad\text{[Since } Z(\mathcal{S})=k\lambda_{\mathcal{S}}+\textstyle\sum_{i\in\mathcal{S}}f_i(\lambda_{\mathcal{S}}) \text{ and } Z(\mathcal{T})=k\lambda_{\mathcal{T}}+\textstyle\sum_{i\in\mathcal{T}}f_i(\lambda_{\mathcal{T}})\text{]} \\
&\leq Z(\mathcal{S})+Z(\mathcal{T}) \quad \text{[By choosing } \alpha=0 \text{ if } \lambda_{\mathcal{S}}>\lambda_{\mathcal{T}} \text{ and } \alpha=1 \text{ if } \lambda_{\mathcal{T}}>\lambda_{\mathcal{S}}\text{]}. \qquad \square
\end{aligned}
$$

A natural implication of this result is that the problem of maximizing the nonnegative monotone submodular function in (4.1) subject to cardinality constraints:

$$\max_{|\mathcal{S}|\leq r} Z(\mathcal{S}),$$

admits a $1-1/e$ approximation algorithm (see Nemhauser, Wolsey and Fisher [30]).

5 Conclusions

In this paper, we have introduced an extension of the k-sum combinatorial optimization to a distributionally robust version where the distribution of the costs is only partly specified. This includes the bottleneck optimization and the minimum sum problem as special cases. Our main results show that for the Fréchet class of discrete marginal distributions with finite support, many of the nice properties established for the deterministic k-sum combinatorial optimization extends to the distributionally robust case. This includes polynomial time solvability when the deterministic minimum sum problem is solvable in polynomial time, ε-approximation guarantees when the deterministic minimum sum problem has a ε-approximation guarantee and submodularity of the objective function. Our results also indicate that the choice of the distributional uncertainty set plays an important role in terms of computational tractability, given that even for simple uncertainty sets such as ellipsoids, the k-sum combinatorial optimization problem can become NP-hard when passing from the deterministic to the distributionally robust formulation. Our results extend the results of Punnen and Aneja [25] to the distributionally robust setting and the results of Agrawal et. al. [1] to identify instances when the distributionally robust integer program is solvable in polynomial time. One possible future research question is to identify other sets of distributions beyond the ones considered in this paper when the problem is solvable in polynomial time. Furthermore, developing algorithms and characterizing the complexity of the distributionally robust k-sum problem with continuous marginal distributions and with second and higher order moments would also be useful.

Acknowledgements

The authors would like to thank the Editor, the Associate Editor and two anonymous reviewers for their valuable comments that significantly helped improve the manuscript. The authors would also like to thank Professor Teo Chung-Piaw for originally posing the question on the submodularity of the robust bottleneck objective to us.

References

[1] S. Agrawal, Y. Ding, A. Saberi, and Y. Ye. Price of correlations in stochastic optimization. *Operations Research*, 60(1):150–162, 2012.

[2] H. Aissi, C. Bazgan, and D. Vanderpooten. Min-max and min-max regret versions of combinatorial optimization problems: A survey. *European Journal of Operational Research*, 197(2):427–438, 2009.

[3] A. Ben-Tal, L. El Ghaoui, and A. S. Nemirovskii. *Robust Optimization*. Princeton University Press, 2009.

[4] A. Ben-Tal and E. Hochman. More bounds on the expectation of a convex function of a random variable. *Journal of Applied Probability*, 9(4):803–812, 1972.

[5] D. Bertsimas, X. V. Doan, K. Natarajan, and C. P. Teo. Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research*, 35(3):580–602, 2010.

[6] D. Bertsimas, K. Natarajan, and C. P. Teo. Probabilistic combinatorial optimization: moments, semidefinite programming and asymptotic bounds. *SIAM Journal on Optimization*, 15(1):185–209, 2004.

[7] D. Bertsimas and M. Sim. Robust discrete optimization and network flows. *Mathematical Programming*, 98(1–3):49–71, 2003.

[8] J. R. Birge and M. J. Maddox. Bounds on expected project tardiness. *Operations Research*, 43(5):838–850, 1995.

[9] G. Carpaneto, S. Martello, and P. Toth. An algorithm for the bottleneck traveling salesman problem. *Operations Research*, 32(2):380–389, 1984.

[10] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.

[11] J. Edmonds and D. R. Fulkerson. Bottleneck extrema. *Journal of Combinatorial Theory*, 8:299–306, 1970.

[12] H. N. Gabow and R. E. Tarjan. Algorithms for two bottleneck optimization problems. *Journal of Algorithms*, 9(3):411–417, 1988.

[13] M. R. Garey and D. S. Johnson. Computers and intractability: A guide to the theory of npcompleteness. 109, 1979.

[14] S. Geetha and K. P. K. Nair. A stochastic bottleneck transportation problem. *Journal of the Operational Research Society*, 45:583–588, 1994.

[15] O. Gross. The bottleneck assignment problem. Technical report, 1959.

[16] G. Grygiel. Algebraic $\sum_k$-assignment problems. *Control and Cybernetics*, 10(3-4):155–165, 1981.

[17] S. K. Gupta and A. P. Punnen. k-sum optimization problems. *Operations Research Letters*, 9(2):121–126, 1990.

[18] G. A. Hanasusanto, D. Kuhn, and W. Wiesemann. K-adaptability in two-stage distributionally robust binary programming. *Operations Research Letters*, 44(1):6–11, 2016.

[19] W. L. Hsu and G. L. Nemhauser. Easy and hard bottleneck location problems. *Discrete Applied Mathematics*, 1(3):209–215, 1979.

[20] H. Ishii and T. Nishida. Stochastic bottleneck spanning tree problem. *Networks*, 13(3):443–449, 1983.

[21] P. Kouvelis and G. Yu. *Robust discrete optimization and its applications*. Kluwer Academic Publishers, 1997.

[22] A. Madansky. Bounds on the expectation of a convex function of a multivariate random variable. *The Annals of Mathematical Statistics*, 30(3):743–746, 1959.

[23] I. Meilijson and A. Nadas. Convex majorization with an application to the length of critical paths. *Journal of Applied Probability*, 16(3):671–677, 1979.

[24] K. Natarajan, D. Shi, and K. C. Toh. A probabilistic model for minmax regret in combinatorial optimization. *Operations Research*, 62(1):160–181, 2014.

[25] A. P. Punnen and Y. P. Aneja. On k-sum optimization. *Operations Research Letters*, 18(5):233–236, 1996.

[26] A. Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1986.

[27] P. J. Slater. Centers to centroids in graphs. *Journal of Graph Theory*, 2(3):209–222, 1978.

[28] A. Tamir. The k-centrum multi-facility location problem. *Discrete Applied Mathematics*, 109(3):293–307, 2001.

[29] W. Wiesemann, D. Kuhn, and M. Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014.

[30] L. A. Wolsey, M. L. Fisher, and G. L. Nemhauser. An analysis of approximations for maximizing submodular set functions. *Mathematical Programming*, 14(1):265–294, 1978.

[31] U. Yechiali. A stochastic bottleneck assignment problem. *Management Science*, 14(11):732–734, 1968.

[32] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2013.