A Robust Approach to a Class of Uncertain Optimization Problems with Imprecise Probabilities

Adam Kasperski
Faculty of Computer Science and Management
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
Email: adam.kasperski@pwr.edu.pl

Paweł Zieliński
Faculty of Fundamental Problems of Technology
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
Email: pawel.zielinski@pwr.edu.pl

Abstract—In this paper a class of discrete optimization problems with uncertain costs is discussed. The uncertainty is modeled by providing a discrete scenario set, in which each scenario represents a possible realization of the element costs (the problem parameters). It is assumed that a partial information about scenario occurrence probabilities is also available. Namely, each such a probability is known to belong to a given closed interval. Several criteria for choosing a solution, such as the expected value, the value at risk, the conditional value at risk, and the tail \( \alpha \)-mean are considered. A solution minimizing one of these criteria for the worst possible probability distribution in scenario set is computed. The computational complexity of the problems under consideration is explored. Some exact and approximation algorithms for them are proposed.

I. INTRODUCTION

In this paper wish to investigate a class of discrete optimization problems, consisting in determining an object (a feasible solution) composed of elements of a given finite set \( E \). Each element in \( E \) has a nonnegative cost and we are interested in finding an object with the minimum total cost. Several discrete optimization problems of practical importance belong to this class. For instance, the shortest path problem, the minimum spanning tree problem, or generally - the network problems (see, e.g., [1]), in which \( E \) consists of all edges of a given graph. A common assumption in this class of problems is that the element costs are precisely known. However, in practice this is rarely the case. Typically, the uncertainty is modeled by specifying a set, denoted by \( \mathcal{U} \), of all possible realizations of the element costs, called scenarios. There are two popular methods of defining the set of scenarios, namely the interval and discrete uncertainty representations [2].

In the interval uncertainty representation, a closed interval is assigned to each element, which means that its cost will take some value within the interval, but it is not possible to predict at present which one. Thus \( \mathcal{U} \) is the Cartesian product of these intervals. A natural criterion to choose a robust solution under the interval uncertainty is the maximum regret, used by risk averse decision makers, which is the maximum difference between the cost of a solution and the optimal cost over all scenarios. This leads to the class of problems, where the maximum regret is minimized. This criterion has been applied to many network optimization problems (see [3] for a survey). A more elaborate approach is collecting both the intervals and the plausible element cost values, and modeling the uncertain costs by fuzzy intervals in the setting possibility theory [4]. The fuzzy intervals are regarded as possibility distributions, describing the sets of more or less plausible values of element costs. Consequently, we obtain a joint possibility distribution, induced by these fuzzy intervals, on the set of scenarios \( \mathcal{U} \). For finding a “good” solution one can adopt a possibility based criterion, called necessary soft optimality [5], which has been originally proposed in [6], [7] for linear programming problem with a fuzzy objective function. Choosing a best necessarily soft optimal solution is a direct generalization of minimizing the maximum regret to the fuzzy case. We refer the reader to [8] for a deeper discussion of the minmax regret and fuzzy approaches, and to [9] for a review of different concepts in fuzzy optimization.

In the discrete uncertainty representation, scenario set \( \mathcal{U} \) is defined by explicitly listing all possible scenarios. Hence, \( \mathcal{U} \) is a finite set. In order to choose a solution the maximum and the maximum regret criteria, used in robust optimization, are widely adopted. If the maximum criterion is applied, then we seek a solution with the smallest maximum total cost over all scenarios. Similarly as in the interval case, the criteria have been applied to many discrete optimization problems (see [3] for a survey). However, this robust approach to decision making is often regarded as too conservative and assumes that decision makers are very risk averse. In particular, the maximum criterion takes into account only one, the worst-case scenario, ignoring the information connected with the remaining scenarios. Hence, there is a need to soften the very conservative maximum (regret) criterion.

In real life decision makers often have some additional information provided with \( \mathcal{U} \). For instance, an information about an importance of each scenario. So, to improve the situation one can express the importance of scenarios by assigning some weights to them and apply the Ordered Weighted Averaging aggregation operator (OWA for short) [10] that uses this information while computing a solution [11]. Hence, the weights allow decision makers to take their attitude towards a risk into account. It is worth pointing out that the OWA criterion generalizes the classical criteria such as the
maximum, minimum, average, Hurwicz etc. Another example of additional information provided with $\mathcal{U}$ is a probability distribution in $\mathcal{U}$, where the probabilities are known precisely or can be estimated. A frequent goal, in this case, is to minimize the expected value criterion (the expected cost) of a solution built (see, e.g., [12], [13]). This criterion assumes that decision maker is risk neutral and sometimes a solution found may be questionable, especially when it is implemented only once. A criterion that exploits the information about both an importance and probability of occurrence of each scenario is the Weighted Ordered Weighted Averaging aggregation operator [14] (WOWA for short). It is worthwhile to mention that the WOWA criterion is a generalization of the expected value and OWA. In consequence, it also generalizes the maximum, minimum, average, etc. The WOWA operator has been recently applied in [15] to discrete optimization problems under the discrete uncertainty representation. In order to represent decision maker’s attitude towards a risk, the tail mean quantity based criteria have been recently proposed. Among them, there are the value at risk, the conditional value at risk, and the tail $\alpha$-mean [17].

In this paper we deal with the class of discrete optimization problems with uncertain element costs under the discrete scenario uncertainty representation. We assume that the occurrence probabilities of scenarios in $\mathcal{U}$ are imprecise. We only known that they belong to closed intervals. In order to choose a solution we can adopt the robust approach (the maximum criterion) to the expected value criterion with imprecise scenario probabilities, i.e. we seek a solution that minimizes the worst case expected value of the cost with respected to the probabilities modeled by the intervals. Such model has been originally proposed in [18] for stochastic linear programming problems. Moreover, we propose further extensions of this model to the discrete optimization problems with the worst case versions of the risk criteria, namely the value at risk, the conditional value at risk, and the tail $\alpha$-mean. We give an analysis of computational complexity of the problems under consideration, namely we show when the problems are polynomially solvable, NP-hard, approximable or inapproximable. For tractable cases, we propose some approximation algorithms and exact methods based on mixed integer programming (MIP) formulations.

The paper is organized as follows. In Section II we formally state the class of discrete optimization problems with uncertain element costs under the discrete uncertainty representation and imprecise probabilities in scenario set $\mathcal{U}$. We also give a brief exposition of the tail mean quantity based criteria adopted to the problems under consideration. In Section III we show exact and approximation methods for solving the problems with the expected cost criterion. In Sections IV, V and VI we examine the problems with the tail mean quantity based criteria and show polynomially solvable, NP-hard and approximable cases of the problems. We finish the paper with some conclusions.

![Fig. 1. A sample SHORTEST PATH problem with 4 scenarios $S_1 = (6, 5, 6, 8, 0), S_2 = (2, 3, 0, 1, 0), S_3 = (1, 1, 1, 0, 0), S_4 = (1, 4, 1, 0, 0)$ and $p = (0, 1, 0, 4, 0, 5, 2)$. The costs of all three paths under scenarios are shown in the table.](image)

II. PRELIMINARIES

We now give a formal definition of a class of discrete optimization problems with uncertain costs under the discrete uncertainty representation. Let $E = \{e_1, \ldots, e_n\}$ be a finite set of elements and let $\Phi \subseteq 2^E$ be a set of feasible solutions. We are given a finite scenario set $\mathcal{U} = \{S_1, \ldots, S_K\}$. Let $c_{ij} \geq 0$ be a cost of element $e_i$ under scenario $S_j$, $j \in [K]$ denotes the set $\{1, \ldots, K\}$. We will use $f(X, S_j) = \sum_{i \in X} c_{ij}$ to denote the cost of solution $X$ under scenario $S_j$. Let $p_j$ be the probability of the event that scenario $S_j$ will occur. It thus holds $p_j \in [0, 1]$ for each $j \in [K]$ and $p_1 + \cdots + p_K = 1$. Let us denote $p = (p_1, \ldots, p_K)$, so $p$ induces a probability distribution in scenario set $\mathcal{U}$. A random cost of a given solution $X \in \Phi$ is a discrete random variable, denoted by $F(X, p)$, with parameters $X$ and $p$ and the probability distribution defined as follows: $\Pr[F(X, p) = f(X, S_j)] = p_j$, for $j \in [K]$.

We now discuss several criteria of choosing a solution in the problem stated above, under the assumption that the scenario occurrence probabilities are precisely known. To illustrate them, we use the sample SHORTEST PATH problem with 4 scenarios shown in Fig. 1. In order to choose a solution we can use the well-known expected value criterion (see, e.g., [12], [13]):

$$E[F(X, p)] = \sum_{j \in [K]} p_j f(X, S_j).$$

(1)

In the example in Fig. 1, the smallest expected cost has the path $\{e_2, e_5\}$. This solution, however, may be questionable. Observe that the path $\{e_2, e_5\}$ is better than $\{e_1, e_4\}$ only under scenario $S_1$, whose probability is rather small (equal to $0.1$). Hence, after rejecting scenario $S_1$, the path $\{e_2, e_5\}$ should not be chosen. We can say, equivalently, that $\{e_2, e_5\}$ will be chosen by a pessimistic decision maker or in situations when it is very important to avoid bad scenarios such as $S_1$.

In many situations, decision makers would take into account the occurrence of additional information provided with $\mathcal{U}$. We now give a brief exposition of the tail mean quantity based criteria adopted to the problems under consideration. In Section III we show exact and approximation methods for solving the problems with the expected cost criterion. In Sections IV, V and VI we examine the problems with the tail mean quantity based criteria and show polynomially solvable, NP-hard and approximable cases of the problems.
distribution \( \Pr[F(X, p) \leq t] \) is right-continuous function of \( t \). Notice that \( \text{VaR}_\alpha[F(X, p)] = \max_{j \in [K]} f(X, S_j) \) and \( \text{VaR}_\alpha[F(X, p)] = \min_{j \in [K]} f(X, S_j) \). The value at risk criterion can be equivalently stated as follows. Let \( \mathcal{U}(\alpha) \subseteq \mathcal{U} \) be a nonempty subset of scenarios such that \( \sum_{S_j \in \mathcal{U}(\alpha)} p_j \geq \alpha \).

Then

\[
\text{VaR}_\alpha[F(X, p)] = \min_{\mathcal{U}(\alpha) \subseteq \mathcal{U}} \max_{S_j \in \mathcal{U}(\alpha)} f(X, S_j). \tag{3}
\]

Note that \( \mathcal{U}(1) = \mathcal{U} \) and \( \mathcal{U}(0^+) \) is any nonempty subset of \( \mathcal{U} \), which again shows that the value at risk is the maximum for \( \alpha = 1 \) and the minimum for \( \alpha = 0^+ \).

Consider the sample problem in Fig. 1. Let us choose \( \alpha = 0.8 \). Then for path \( X = \{e_2, e_3\} \) we get \( \text{VaR}_{0.8}[F(X, p)] = 4 \) (see Fig. 2a). We can provide the following interpretation of the obtained value. The maximum cost of path \( X \), after rejecting the worst scenarios whose probability of occurrence is less than \( 1 - \alpha = 0.2 \) is equal to 4. Equivalently, the decision maker applies the maximum criterion to the best subset of scenarios, whose probability of occurrence is at least \( \alpha = 0.8 \) (see (3)). The corresponding values at risk for paths \( \{e_1, e_4\} \) and \( \{e_1, e_3, e_5\} \) equals 3 and 2, respectively. In consequence the criterion suggests to choose the path \( \{e_1, e_3, e_5\} \), as this path is the most robust after ignoring some worst scenarios (scenario \( S_1 \) in this example). We thus can see that the parameter \( \alpha \) allows us to control the attitude towards a risk. When \( \alpha = 1 \), then decision maker is extremely risk averse and always minimizes the maximum solution cost, regardless of the probability distribution in \( \mathcal{U} \). On the other hand \( \alpha < 1 \) allows decision makers to reject some worst scenarios whose probability of occurrence is appropriately small. In the extreme case, when \( \alpha \to 0 \), then decision makers are extremely optimistic and assume that the best scenarios will occur, regardless of the probability distribution in \( \mathcal{U} \).

The path \( \{e_1, e_3, e_5\} \) may be still questionable. After rejecting scenario \( S_1 \), some decision makers may feel that the path \( \{e_1, e_4\} \) is better. The value at risk criterion takes into account only the worst scenario, ignoring the information connected with the remaining scenarios. One can thus consider the following lower \( \alpha \)-mean criterion (the tail \( \alpha \)-mean [17]):

\[
\text{CVaR}_\alpha[F(X, p)] = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_\beta[F(X, p)] \, d\beta. \tag{4}
\]

Observe that if \( \alpha = 1 \), then the lower \( \alpha \)-mean is equal to the expected value and when \( \alpha \to 0 \), then it is equal to \( \min_{j \in [K]} f(X, S_j) \). The computation of \( \text{CVaR}_\alpha[F(X, p)] \) for \( \alpha = 0.8 \) and \( X = \{e_2, e_5\} \) is shown in Fig. 2b. The integral is represented as the grey area in Fig. 2b. We obtain the lower \( \alpha \)-mean by dividing the size of this area by 0.8. The lower \( \alpha \)-mean criterion will be used by optimistic decision makers, who reject bad scenarios occurring with a reasonable probability. One can also take the following symmetric upper \( \alpha \)-mean criterion into account (the conditional value at risk [16]), in which the decision maker rejects good scenarios occurring with a reasonable probability:

\[
\text{CVaR}_\alpha[F(X, p)] = \frac{1}{1 - \alpha} \int_0^1 \text{VaR}_\beta[F(X, p)] \, d\beta, \tag{5}
\]

where \( \alpha \in [0, 1) \). Notice that when \( \alpha = 0 \), then the upper \( \alpha \)-mean is the expected value and if \( \alpha \to 1 \), then it becomes the maximum, i.e. \( \max_{j \in [K]} f(X, S_j) \). Consequently, it will be used by pessimistic decision makers, who want to avoid bad scenarios.

In this paper we assume that the scenario probabilities are not precisely known. We only know that \( p_j \in [l_j, u_j] \), \( l_j \leq u_j \), for each \( j \in [K] \), where \( l_j, u_j \in [0, 1] \) are the lower and upper bounds on the probability \( p_j \). Let \( \mathcal{P} \) be the family of probability distributions defined as follows:

\[
\mathcal{P} = \{(p_1, \ldots, p_K) : p_1 + \cdots + p_K = 1, \ p_j \in [l_j, u_j], j \in [K]\}.
\]

The above model for coping with imprecise probabilities has been proposed in [18], [19]. Set \( U = \sum_{j \in [K]} u_j \) and \( L = \sum_{j \in [K]} l_j \). It must hold \( U \geq 1 \), \( L \leq 1 \) and \( L \leq U \), since otherwise \( \mathcal{P} = \emptyset \). In the next sections we will investigate the problem with imprecise probabilities and the criteria described in this section.

Since we will investigate the approximability of the considered problems, we recall the definition of \( \rho \)-approximation algorithm (see [20, Definition 1.1]): a \( \rho \)-approximation algorithm for an optimization problem is a polynomial time algorithm that for all instances of the problem returns a solution whose value is within a factor of \( \rho \) of the value of an optimal solution. For minimization problems \( \rho > 1 \).

### III. The Expected Cost Criterion

In this section we will be concerned with the problem \( \mathcal{P} \) with uncertain element costs under the discrete scenario uncertainty representation and imprecise probabilities and the expected cost criterion (1), i.e. the problem stated as follows:

\[
\text{MIN-EXP } \mathcal{P}: \min_{X \in \Phi} \max_{p \in \mathcal{P}} \mathbb{E}[F(X, p)]. \tag{6}
\]

In the same setting, linear programming problems have been studied in [18]. We will show a MIP-based method for solving the above problem and a general approximation algorithm under the assumption that \( \mathcal{P} \) is polynomially solvable.
A vector of probabilities \( p^X = (p_1^X, \ldots, p_K^X) \in \mathbb{P} \) is said to be the worst-case probabilities for a given solution \( X \) if it is an optimal solution of the following problem:

\[
E[F(X, p^X)] = \sum_{j \in [K]} p_j^X f(X, S_j) = \max_{p \in \mathbb{P}} E[F(X, p)]. \tag{7}
\]

Let us now discuss the complexity of the problem (6). When \( \mathbb{P} \) contains one vector of probabilities, say \( p \), then problem (6) is polynomially solvable if \( \mathbb{P} \) is polynomially solvable. It is enough to solve \( \mathbb{P} \) for the aggregated element costs \( \hat{c}_i = \sum_{j \in [K]} p_j c_{ij} \), \( i \in [n] \). However, when \( [j, u_j] = [0, 1] \) for each \( j \in [K] \), then the problem (7) becomes the following one:

\[
\max_{p \in \mathbb{P}} E[F(X, p)] = \max_{j \in [K]} f(X, S_j),
\]

which follows from the fact that we assign the probability equal to 1 to the worst scenario for \( X \) and the probability 0 to the remaining scenarios. Therefore, MIN-EXP \( \mathbb{P} \) is at least as hard as the traditional robust MIN-MAX \( \mathbb{P} \) problem (see, e.g., [2]). The complexity of MIN-MAX \( \mathbb{P} \) is well established in the existing literature. This problem is typically NP-hard when \( K = 2 \) and becomes strongly NP-hard and also hard to approximate when \( K \) is a part of input [21]–[24]. It is also well known (see, e.g. [3]) that MIN-MAX \( \mathbb{P} \) is approximable within \( K \) when \( \mathbb{P} \) is polynomially solvable. The idea is just to solve \( \mathbb{P} \) for the aggregated costs \( \hat{c}_i = \max_{j \in [K]} c_{ij}, i \in [n] \).

In the following we will generalize this idea to MIN-EXP \( \mathbb{P} \).

A. A MIP Formulation

We now build a MIP formulation for MIN-EXP \( \mathbb{P} \). Let \( ch(\Phi) \) be the set of characteristic vectors of \( \Phi \). We associate binary variable \( x_i \in \{0, 1\} \) with each element \( e_i \in E \). Given \( x = (x_1, \ldots, x_n) \in ch(\Phi) \), which indicates a solution \( X \in \Phi \), the value of \( E[F(X, p^X)] \) (see (7)) can be computed as follows:

\[
\max \sum_{j \in [K]} p_j \sum_{i \in [n]} x_i c_{ij} \quad p_1 + \cdots + p_K = 1 \quad l_j \leq p_j \leq u_j \quad j \in [K]
\]

(8)

It is easily seen that the problem (8) can be solved efficiently by a greedy method. An equivalent dual linear programming formulation of (8) is as follows (see [17]):

\[
\min \gamma + \sum_{j \in [K]} (u_j a_j - l_j \beta_j) \\
\quad + \alpha_j + \beta_j \geq \sum_{i \in [n]} x_i c_{ij} \quad j \in [K]
\]

(9)

So, the MIP model for MIN-EXP \( \mathbb{P} \) is:

\[
\min \gamma + \sum_{j \in [K]} (u_j a_j - l_j \beta_j) \\
\quad + \alpha_j + \beta_j \geq \sum_{i \in [n]} x_i c_{ij} \quad j \in [K]
\]

(9)

Note that before applying (9) one should replace \( (x_1, \ldots, x_n) \in ch(\Phi) \) with a system of linear constraints describing \( ch(\Phi) \) for a particular problem \( \mathbb{P} \).

B. Approximation Algorithm

In this section we propose an approximation algorithm for MIN-EXP \( \mathbb{P} \). Let \( C_i(p) \) be a discrete random variable, where \( p \in \mathbb{P} \), with the probability defined as follows: \( \Pr[C_i(p) = c_{ij}] = p_j \) for \( j \in [K] \). Hence \( C_i(p) \) is the random cost of element \( e_i \) according to the vector \( p \). Let

\[
\sum_{j \in [K]} p_j c_{ij} = E[C_i(p)] = \max_{p \in \mathbb{P}} E[C_i(p)],
\]

(10)

where vector \( p^i = (p_1^i, \ldots, p_K^i) \in \mathbb{P} \) represents the worst-case probabilities for element \( e_i \). The values of \( E[C_i(p^i)] \) for all \( e_i \in E \) can be computed in polynomial time by a greedy method (see the linear programming model (8)). Given \( \hat{X} \in \Phi \), define \( \hat{F}(X) = \sum_{e_i \in X} E[C_i(p^i)] \). Thus

\[
E[F(X, p^X)] = \sum_{j \in [K]} p_j^X f(X, S_j) = \sum_{j \in [K]} p_j^X \sum_{e_i \in X} c_{ij}
\]

(11)

Hence \( \hat{F}(X) \) is an upper bound on \( E[F(X, p^X)] \). Let \( \hat{X} \) be an optimal solution for the element costs \( E[C_i(p^i)] \), \( e_i \in E \). Note that \( \hat{X} \) can be computed in polynomial time if \( \mathbb{P} \) is polynomially solvable.

Theorem 1: If \( U > L \) and \( L < 1 \), then for any solution \( X \in \Phi \) the following inequality holds:

\[
E[F(\hat{X}, p^X)] \leq \frac{U - L}{1 - L} E[F(X, p^X)].
\]

Proof: Taking \( p_j = q_j + l_j \) and \( \sum_{i \in [n]} x_i c_{ij} = f(X, S_j) \) in (8) yields the following equivalent model for evaluating \( E[F(X, p^X)] \):

\[
\max \sum_{j \in [K]} (q_j + l_j)f(X, S_j) \quad q_1 + \cdots + q_K = 1 - L \quad 0 \leq q_j \leq u_j - l_j \quad j \in [K]
\]

(12)

By (11) and the definition of \( \hat{X} \), we get:

\[
E[F(\hat{X}, p^X)] \leq \hat{F}(\hat{X}) \leq \hat{F}(\hat{X}) = \sum_{j \in [K]} p_j^X c_{ij}
\]

(13)

Set \( q_j = (u_j - l_j)(1 - L) \), \( j \in [K] \). Observe that \( q_j, j \in [K] \) are feasible in (12). Indeed \( q_j \leq u_j - l_j \), because \( \frac{1 - L}{U - L} \leq 1 \) and

\[
\sum_{j \in [K]} q_j = \frac{1 - L}{U - L} \sum_{j \in [K]} (u_j - l_j) = 1 - L.
\]

In consequence, (12) leads to

\[
E[F(X, p^X)] \geq \sum_{j \in [K]} \left( \frac{(u_j - l_j)(1 - L) + l_j}{U - L} \right) f(X, S_j).
\]
TABLE I
A HARD INSTANCE FOR THE APPROXIMATION ALGORITHM

<table>
<thead>
<tr>
<th>[l₁, u₁]</th>
<th>[l₁, L]</th>
<th>[0, 1 - L]</th>
<th>0.1 - L</th>
<th>...</th>
<th>0.1 - L</th>
</tr>
</thead>
<tbody>
<tr>
<td>e₁</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>e₂</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>eₖ</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>f₁</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>f₂</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>fₖ</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>1</td>
</tr>
</tbody>
</table>

Since \( \frac{1}{1 - L} \leq 1 \),
\[
E[F(X, p^X)] \geq \sum_{j \in [K]} \left( \frac{(u_j - l_j)(1 - L)}{U - L} + \frac{1 - L}{U - L} \right) f(X, S_j)
\]
\[
= \sum_{j \in [K]} \frac{1 - L}{U - L} u_j f(X, S_j) = \frac{1 - L}{U - L} \sum_{j \in [K]} u_j f(X, S_j),
\]
which, together with (13), complete the proof.

The ratio \( \frac{1}{U - L} \) \( \rightarrow \) \( +\infty \) when \( U > L \) and \( L \rightarrow 1 \). However, the value of \( U \) is closely related to the value of \( L \), which allows us to estimate this ratio from above. Observe that \( p_j \leq 1 - L + l_j \) for each \( j \in [K] \). Otherwise, \( p_j + L - l_j = l_1 + \cdots + l_{j-1} + p_j + l_{j+1} + \cdots + l_n > 1 \), a contradiction. In consequence, MIN-EXP \( \mathcal{P} \) remains unaffected when we replace \( u_j \) with \( u_j' = \min\{u_j, 1 - L + l_j\} \) for each \( j \in [K] \) in the definition of \( \mathcal{P} \).

In this case \( U \leq \sum_{j \in [K]} (1 - L + l_j) = K - KL + L \). Hence, for any \( L < 1 \) we have
\[
\frac{U - L}{1 - L} \leq \frac{K - KL}{1 - L} \leq K.
\]

We thus get the following approximation result:

**Theorem 2:** If \( \mathcal{P} \) is polynomially solvable, then MIN-EXP \( \mathcal{P} \) is approximable within \( \frac{U - L}{1 - L} \leq K \).

**Proof:** If \( U = L = 1 \), then \( \mathcal{P} \) contains at most one probability vector \( p \). In this case MIN-EXP \( \mathcal{P} \) boils down to computing a solution minimizing \( E[F(X, p)] \), which can be done in polynomial time. If \( U > L > 1 \), then the bound \( U - L \) holds by Theorem 1. This ratio is not greater than \( K \) by the remark shown after Theorem 1.

The bound in Theorem 2 is tight. To see this, consider an instance of the MIN-EXP SELECTION problem shown in Table I. In this problem \( E = \{e₁, \ldots, eₖ, f₁, \ldots, fₖ\} \), \( \Phi = \{X \subseteq E : |X| = K\} \), \( |E| = 2K \). The element costs under scenarios and the corresponding intervals for probabilities are shown in Table I. The value of \( L \) is such that \( L \in (0, 1) \). Observe that the aggregated cost (see (10)) of each element is the same and equals 1 - L, for \( X = \{e₁, \ldots, eₖ\} \) we have \( E[F(X, p^X)] = K \cdot L \cdot (1 - L)/L = KL(1 - L) \) and for \( X' = \{f₁, \ldots, fₖ\} \) we obtain \( E[F(X', p^{X'})] = 1 - L \). Clearly, \( E[F(X, p^X)] = K \cdot E[F(X', p^{X'})] \). Thus the bound of \( K \) is tight for any \( L \in (0, 1) \).

The worst case ratio of the algorithm may be less than \( K \) if \( U \) and \( L \) are sufficiently small. In particular, if \( L = 0 \), then MIN-EXP \( \mathcal{P} \) is approximable within \( U \in [1, K] \). Notice also that if \( [l_j, u_j] = [0, 1] \) for each \( j \in [K] \), then the algorithm boils down to the \( K \)-approximation algorithm, known for MIN-MAX \( \mathcal{P} \) (see, e.g. [3]).

Theorem 1 can be generalized to problems whose deterministic versions are NP-hard. In this case we often have a \( \rho \)-approximation algorithm for some \( \rho > 1 \). For example, when \( \mathcal{P} \) is the TRAVELING SALESMAN problem that obeys the triangle inequality. We can apply the \( \rho \)-approximation algorithm to compute \( X \) such that \( F(X) \leq \rho F(X) \) for any \( X \in \Phi \). Combining this with (13) and (14) yields the following result:

**Theorem 3:** If \( \mathcal{P} \) is approximable within \( \rho > 1 \), then MIN-EXP \( \mathcal{P} \) is approximable within \( \rho\frac{U - L}{1 - L} \leq \rho K \).

**IV. THE VALUE AT RISK CRITERION**

In this section we discuss the problem \( \mathcal{P} \) with uncertain element costs under the discrete scenario uncertainty representation and imprecise probabilities and the the value at risk criterion (2):

\[
\text{MIN-VAR } \mathcal{P} : \min_{X \in \Phi, p \in \mathcal{P}} \text{VaR}_α[F(X, p)].
\]

We first investigate the boundary cases of \( α \). Namely, if \( α = 1 \), then \( \text{VaR}_α[F(X, p)] \) becomes the maximum criterion, i.e.
\[
\text{VaR}_α[F(X, p)] = \max_{j \in [K]} f(X, S_j)
\]
for any probability \( p \in \mathcal{P} \). Hence, in this case the problem (15) is typically NP-hard ([21]–[24]) but it is also approximable within \( K \) if \( \mathcal{P} \) is polynomially solvable ([3]) approximation algorithms with better worst case ratios are known for particular problems, see, e.g., ([23], [25]). If \( α \to 0 \), then \( \text{VaR}_α[F(X, p)] \) is the minimum criterion, i.e.
\[
\text{VaR}_α[F(X, p)] = \min_{j \in [K]} f(X, S_j)
\]
for any probability \( p \in \mathcal{P} \). In this case the problem (15) is polynomially solvable if \( \mathcal{P} \) is polynomially solvable. We now study the remaining case, when \( α \) is a fixed constant in the interval \((0, 1)\). Unfortunately in this case the problem (15) turns out to be computationally hard.

**Theorem 4:** For any fixed \( α \in (0, 1) \), MIN-VAR SHORTEST PATH is strongly NP-hard and not at all approximable unless \( \mathcal{P}=\mathcal{NP} \). This remains true even if \( \mathcal{P} \) contains only uniform probability distribution.

**Proof:** Consider the MIN-2SAT problem which is known to be strongly NP-hard ([26]). We are given \( n \) boolean variables \( x₁, \ldots, xₙ \) and a set of clauses \( C₁, \ldots, Cₘ \), where each clause is disjunction of at most two literals (variables or their negations). We ask if there is an assignment to the variables in which at least \( l \) clauses are not satisfied, where \( l \in [m] \). Note that it is equivalent to the question whether there is an assignment to the variables which satisfies at most \( l' \) clauses (it holds \( l = m - l' \)). Fix \( α \in (0, 1) \).

Given an instance of MIN-2SAT we construct an instance of MIN-VAR SHORTEST PATH in the following way. The graph \( G \) is shown in Fig. 3. The arcs \( e₁, \ldots, eₙ \) correspond to literals \( x₁, \ldots, xₙ \) and the arcs \( f₁, \ldots, fₙ \) correspond to literals \( \bar{x₁}, \ldots, \bar{xₙ} \). It is easily seen that there is a natural one-to-one correspondence between the \( s \to t \) paths is \( G \) and the assignments to \( x₁, \ldots, xₙ \). Indeed, if \( e_i \in X \), then \( x_i = 1 \) and
Therefore, \( l \) positive under at most \( m \) clauses. By the construction, there is a path \( \mathbb{P} \). We need to consider two cases.

The first one, when \( l/m \geq \alpha \). We will restrict to the instances of MIN-2SAT such that \( l/\alpha \) is integer. This involves no loss of generality since we can always add a constant (dependent on \( \alpha \)) number of variables \( x_0', \ldots, x_r' \) and clauses \( (x_{i-1} \lor x_i'), i \in [r] \), obtaining an equivalent problem with \( m+r \) clauses and \( l' = l + r \). For sufficiently large (but constant \( r \)), \( l' / \alpha \) will be integer. We now form additional \((l - \alpha m)/\alpha \) scenarios under which the costs of all arcs are 1. Hence the number of scenarios is \( m' = m + (l - \alpha m)/\alpha = l/\alpha \). For each scenario \( j \in [m'] \), we define an uniform probability distribution \( \mathbb{P} \) in the scenario set. Assume that the answer to MIN-2SAT is yes. So, there is a path \( \mathbb{P} \) whose cost is positive under at most \( m - l \) scenarios. Hence

\[
\Pr[F(X, \mathbb{P}) > 0] \geq \frac{1 - \alpha}{m - l + 1} < \\
\left< \frac{1 - \alpha}{m - l + 1} = \alpha,
\right.
\]

where the second inequality follows from the fact that \( \alpha \in (0, 1) \). Thus \( \text{VaR}_\alpha[F(X, \mathbb{P})] > 0 \).

Theorem 4 holds for the SHORTEST PATH problem. However, it is not difficult to modify it and to show the same hardness result for other basic network problems such as the MINIMUM S-T CUT or MINIMUM ASSIGNMENT (see, e.g. [11]).

V. THE UPPER \( \alpha \)-MEAN CRITERION

In this section we examine a generalization of the MIN-EXP \( \mathcal{P} \) problem, i.e. the problem \( \mathcal{P} \) with uncertain element costs under the discrete scenario uncertainty representation and imprecise probabilities and the upper \( \alpha \)-mean criterion (5):

\[
\text{MIN} - \text{CVaR} \mathcal{P} : \min_{X \in \Phi} \max_{\mathbb{P} \in \mathbb{P}} \text{CVaR}_\alpha[F(X, \mathbb{P})],
\]

where \( \alpha \) is a fixed value in \( [0, 1] \). Of course, \( \text{MIN} - \text{CVaR} \mathcal{P} \) is at least as hard as \( \text{MIN-EXP} \mathcal{P} \) as it reduces to this problem when \( \alpha = 0 \).

A. A MIP Formulation

We propose an exact MIP-based method for solving (18). Define

\[
\text{CVaR}_\alpha[F(X, \mathbb{P}^X)] = \max_{\mathbb{P} \in \mathbb{P}} \text{CVaR}_\alpha[F(X, \mathbb{P})],
\]

Given a feasible solution \( X \in \Phi \), described by a characteristic vector \( (x_1, \ldots, x_n) \in \text{ch}(\Phi) \), the value of \( \text{CVaR}_\alpha[F(X, \mathbb{P}^X)] \) can be computed by solving the following linear programming problem (similar as a model for \( \text{CVaR}_\alpha[F(X, \mathbb{P}^X)] \) in [17]):

\[
\max \quad \frac{1 - \alpha}{m - l + 1} \sum_{j \in [K]} p'_j \sum_{i \in [n]} x_i c_{ij} \quad \text{subject to}
\]

\[
p'_1 + \cdots + p'_K = 1 - \alpha \quad p_1 + \cdots + p_K = 1 \quad 0 \leq p'_j \leq p_j \quad j \in [K] \quad l_j \leq p_j \leq u_j \quad j \in [K]
\]

The vector \( (p_1, \ldots, p_K) \) represents a feasible probability and \( (p'_1, \ldots, p'_K) \) represents the parts of the probabilities assigned...
to the upper \(\alpha\)-tail. Using the dual to (20), we arrive to the following MIP formulation for \(\text{MIN} - \text{CVaR} \mathcal{P}\):
\[
\begin{align*}
\min \ (1 - \alpha)\delta + \gamma + \sum_{j \in [K]} u_j \alpha_j - \sum_{j \in [K]} l_j \beta_j \\
\delta + \phi_j \geq \frac{1}{1 - \alpha} \sum_{i \in [n]} c_{ij} x_i & \quad j \in [K] \\
\gamma - \phi_j + \alpha_j - \beta_j \geq 0 & \quad j \in [K] \\
(x_1, \ldots, x_n) \in \mathcal{E}(\Phi) \\
\alpha_j, \beta_j, \phi_j \geq 0 & \quad j \in [K]
\end{align*}
\]

(21)

B. Approximation Algorithm

We now construct an approximation algorithm for \(\text{MIN} - \text{CVaR} \mathcal{P}\) under the assumption that \(l_j = 0\) for each \(j \in [K]\) in \(\mathcal{P}\). It is easy to check that in this case the model (20) is equivalent to the following one:
\[
\begin{align*}
\max \ \frac{1}{1 - \alpha} \sum_{j \in [K]} p'_j \sum_{i \in [n]} x_i c_{ij} \\
p'_1 + \cdots + p'_K = 1 - \alpha \\
0 \leq p'_j \leq u_j & \quad j \in [K]
\end{align*}
\]

(22)

Writing \(p'_j = q_j (1 - \alpha)\) for \(j \in [K]\) we get
\[
\begin{align*}
\max \ \sum_{j \in [K]} q_j \sum_{i \in [n]} x_i c_{ij} \\
q_1 + \cdots + q_K = 1 \\
0 \leq q_j \leq u_j/(1 - \alpha) & \quad j \in [K]
\end{align*}
\]

(23)

which is equivalent to
\[
\begin{align*}
\max \ \sum_{j \in [K]} q_j \sum_{i \in [n]} x_i c_{ij} \\
q_1 + \cdots + q_K = 1 \\
0 \leq q_j \leq u^*_j & \quad j \in [K]
\end{align*}
\]

(24)

where \(u^*_j = \min \{u_j/(1 - \alpha), 1\}\). Clearly, (24) is the same as (8) when \(l_j = 0\) for each \(j \in [K]\). Hence the following proposition is true:

**Proposition 1:** If \(l_j = 0\) for each \(j \in [K]\), then \(\text{MIN} - \text{CVaR} \mathcal{P}\) is equivalent to \(\text{MIN}-\text{EXP} \mathcal{P}\) with \(l_j = 0\) and the upper bounds equal to \(\min \{u_j/(1 - \alpha), 1\}\) for each \(j \in [K]\). Theorem 2 now leads to the following approximation result:

**Theorem 5:** If \(l_j = 0\) for each \(j \in [K]\), then \(\text{MIN} - \text{CVaR} \mathcal{P}\) is approximable within \(U^*\), where \(U^* = \sum_{j \in [K]} \min \{u_j/(1 - \alpha), 1\}\).

VI. THE LOWER \(\alpha\)-MEAN CRITERION

In this section we will be concerned with the problem \(\mathcal{P}\) with uncertain element costs under the discrete scenario uncertainty representation and imprecise probabilities and the lower \(\alpha\)-mean criterion (4):
\[
\begin{align*}
\text{MIN} - \text{CVaR} \mathcal{P} : \min_{\mathcal{X} \in \Phi} \max_{\mathcal{P} \in \mathcal{P}} \text{CVaR}_\alpha [F(X, \mathcal{P})],
\end{align*}
\]

(25)

where \(\alpha\) is a fixed value in \((0, 1]\). \(\text{MIN} - \text{CVaR} \mathcal{P}\) is at least as hard as \(\text{MIN}-\text{EXP} \mathcal{P}\) as it reduces to this problem when \(\alpha = 1\). Furthermore, the problem is polynomially solvable when \(\alpha \to 0\) as \(\text{CVaR}_\alpha [F(X, \mathcal{P})] = \min_{j \in [K]} f(X, S_j)\) for any probability distribution \(\mathcal{P} \in \mathcal{P}\) (the lower \(\alpha\)-mean criterion becomes the minimum criterion). For the case when \(\alpha \in (0, 1)\) \(\text{MIN} - \text{CVaR} \mathcal{P}\) turns out to be computationally intractable.

**Theorem 6:** For any fixed \(\alpha \in (0, 1)\), \(\text{MIN} - \text{CVaR} \mathcal{P}\) is NP-hard and not at all approximable unless \(\mathcal{P} = \text{NP}\). This remains true even when \(\mathcal{P}\) contains only uniform probability distribution.

**Proof:** The result follows directly from the proof of Theorem 4. In the construction, if the answer to \(\text{MIN}-\text{SAT}\) is yes, then \(\text{CVaR}_\alpha [F(X, \mathcal{P})] = 0\) for some path \(X \in \Phi\). Hence
\[
\int_0^\alpha \text{CVaR}_\beta [F(X, \mathcal{P})] d\beta = 0
\]

and \(\text{CVaR}_\alpha [F(X, \mathcal{P})] = 0\). On the other hand, if the answer is no, then \(\text{CVaR}_\alpha [F(X, \mathcal{P})] > 0\) and
\[
\int_0^\alpha \text{CVaR}_\beta [F(X, \mathcal{P})] d\beta > 0,
\]

which implies \(\text{CVaR}_\alpha [F(X, \mathcal{P})] > 0\).

VII. CONCLUSION

In this paper we have investigated a class of discrete optimization problems with uncertain costs, modeled by discrete scenarios, and imprecise scenario occurrence probabilities, modeled by intervals. We have applied the robust min-max concept to compute a solution. Namely, each solution is evaluated by using the expected value or a risk criterion for the worst probability distribution that may occur. We have shown various complexity and approximation results for the considered problems. The positive approximation results hold for all discrete optimization problems, whose feasible solutions can be represented as subsets of a given finite set. Also, the negative results remain true for many basic problems such the \text{SHORTEST PATH}, \text{MINIMUM S-T CUT} of \text{MINIMUM ASSIGNMENT}.

There is a number of open questions regarding the considered problems. For example, the approximation algorithm proposed for the upper \(\alpha\)-mean criterion is valid only for zero lower bounds of the probability intervals. It would be interesting to investigate this problem for general values of the lower bounds. It would be also interesting to provide an exact algorithm for solving the problem with the value at risk criterion for any fixed \(\alpha \in (0, 1)\). In this paper, we have shown only negative result for this problem. Another interesting research direction is to provide a probabilistic interpretation of the problem with imprecise probabilities (see, e.g. [27]–[30]). Observe that when \(p_j \leq u_j\) and \(u_j = 1\) for at least one \(j \in [K]\), then \(u_j\) can be seen as a possibility of the event that scenario \(S_j\) will occur. In other words, this possibility is an upper bound on the probability that scenario \(S_j\) will occur. Then we can fix \(l_j\) as the necessity of the event that \(S_j\) will occur (a lower bound on the probability). According to the possibility theory, \(l_j = 1 - \max_{i \in [K] \setminus \{j\}} u_i\) for each \(j \in [K]\). Consequently, given \(u_j, j \in [K]\), we obtain a particular instance of the imprecise probabilities. A more deep investigation of the properties of such a problem is an interesting research direction.
ACKNOWLEDGMENT
This work was partially supported by the National Center for Science (Narodowe Centrum Nauki), grant 2013/09/B/ST6/01525.

REFERENCES