



Necessary and possibly optimal items in selecting problems

Sébastien Destercke, Romain Guillaume

► To cite this version:

Sébastien Destercke, Romain Guillaume. Necessary and possibly optimal items in selecting problems. 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2022), Jul 2022, Milan, Italy. pp.494-503, 10.1007/978-3-031-08971-8_41 . hal-03895740

HAL Id: hal-03895740

<https://hal.science/hal-03895740>

Submitted on 13 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Necessary and possibly optimal items in selecting problems.

Sébastien Destercke¹ and Romain Guillaume²

¹ UMR CNRS 7253 Heudiasyc, Sorbonne Université, Université de Technologie de Compiègne
CS 60319 - 60203 Compiègne cedex, France
`{surname.name}@hds.utc.fr`
² University of Toulouse 2, IRIT, France

Abstract. In this paper, we consider the classical and basic optimisation problem consisting in selecting a set of p items among n with minimal costs, but where the costs are ill-known or uncertain. As the optimal solution in such a case is no longer uniquely nor well-defined, we consider the issue of characterizing the set of optimal solutions, as well as those items that are necessarily or possibly in the final selection.

Keywords: Combinatorial optimisation · Interval uncertainty · Credal sets · Possibly optimal solutions.

1 Introduction

As one of the most basic problem in combinatorial optimisation, item selection constitute a special case of many other well-known combinatorial optimisation problems. For instance, it corresponds to a knapsack problem with items of equal weights, or to specific one-job scheduling or assignment problems. As such, its study as a basic building block remains of interest.

This problem goes as follows: given a set $\Omega = \{x_1, \dots, x_n\}$ of n objects, each associated with a cost c_i , find amongst those n objects the p objects that minimize the cost. In other words, if we denote by $F(X) = \sum_{x_i \in X} c_i$ the accumulated cost for the subset $X \subseteq \Omega$ of objects, then the selecting items problem consists in finding

$$X^* = \arg \min_{X \subseteq \Omega, |X|=p} F(X). \quad (1)$$

One can see that the optimal solution is obtained by taking the p objects with the smallest costs, which can be achieved in $O(n)$ time. We will denote by $\Phi = \{X \subseteq \Omega : |X| = p\}$ the set of feasible solutions of the problem.

However, previous studies have shown that this simplicity of the problem may not last when adding some uncertainty to the problem. For instance, it has been shown that computing min-max regret solutions when costs can take values in a discrete space makes the problem NP-hard [6].

It can also be argued that using min-max, regret min-max or other criterion that makes the optimal solution well-defined but unique again amounts to making a sub-

jective choice. An alternative, notably explored in decision making (both in multi-criteria [5] and uncertain situations [8]) and computational social choice [9], is to consider and characterize the set of possibly optimal solutions rather than searching for a unique one. While some work exists in this line for combinatorial optimisation problems [4, 1], it is much less investigated.

In this paper, we discuss what happens when costs become uncertain, and when we want to characterise sets of possibly optimal solutions given this uncertainty. Such a characterization may be helpful to know on which items one should focus its investigation, or to provide a decision maker with robust and guaranteed information about the solution.

We first look at the problem when costs are known to belong to intervals, and then proceed to a more general framework where the uncertainty about costs is described by convex sets of probabilities, a model that includes as specific case most uncertainty theories, such as belief functions, probabilities or possibilities.

Section 2 discusses the problem we consider in the interval case, and notably what we mean by set of possibly optimal solutions, as well as by necessary and possibly optimal items. Section 3 then proposes a compact representation of all the possibly optimal solution by using the previously established notion, showing that this set can be obtained in polynomial time. Section 4 discusses the case where the uncertainty is described by a set of probability.

2 The interval case

This section introduces the simple idea of sets of possibly optimal solution (which in practice reduces to define a Pareto front), as well as the idea of necessarily and possibly optimal item, in the case where uncertainty is described by an interval.

2.1 Setting

We now consider that costs are defined by intervals, that is for object x_i we only know that its cost belong to $\mathcal{C}_i := [\underline{c}_i, \bar{c}_i]$. In practice, this means that $F(X) \in \mathcal{F}(X) := [\underline{F}(X), \bar{F}(X)]$, and that the optimal solution is not unique any more.

A regular way to solve this issue is by considering a min-max or a regret min-max criterion, where one tries to find the best solution in the worst case scenario. Let us denote by $\Gamma = \times_{i=1}^n [\underline{c}_i, \bar{c}_i]$ the hyper-cube representing the set of possible scenarios, and by $\gamma \in \Gamma$ one possible scenario. We will denote by X_γ^* the optimal solution for a scenario γ . The min-max solution is

$$X_{mM}^* = \arg \min_{X \subseteq \Omega, |X|=p} \max_{\gamma \in \Gamma} F(X)$$

In the interval-valued case, such a solution is actually very easy to find, as we just have to solve Equation (1) for $c_i = \bar{c}_i$.

However, as said in the introduction, we will be interested in this paper in characterising the whole set \mathcal{O} of possibly optimal solutions. This set is defined as follows

$$\mathcal{O} = \{X \in \Phi : \exists \gamma \in \Gamma \text{ s.t. } X \in \arg \min_{X \in \Phi} F(X)\} \quad (2)$$

and includes all the solutions that, for what we know, could be optimal. In general, such a set is not obtainable through simple enumeration, nor is very practical to handle: any end-user confronted to a long list of solutions would simply not find it useful. It should however be noted that checking whether a given set $X \in \Phi$ is within \mathcal{O} is very easy: simply set $c_i = \underline{c}_i$ for all $x_i \in X$ and $c_i = \bar{c}_i$ for all $x_i \notin X$, and check that X is the solution to Equation (1). However, the number of subsets to check would still remain exponential.

For this reason, it is useful to easy to compute representations of \mathcal{O} , possibly approximate. This is what we do next, starting with the notion of necessary and optimal items.

2.2 Possibly and necessarily optimal items

An item x_i is said possibly optimal if it is in at least one solution of \mathcal{O} , and necessarily optimal if it is in all the solutions of \mathcal{O} . For the selecting item problem, we have the following propositions:

Proposition 1. x_i is necessarily optimal if and only if $\bar{c}_i > \underline{c}_j$ for at least $n - p$ items x_j , i.e.

$$|\{x_j : \bar{c}_i < \underline{c}_j, i \neq j\}| \geq n - p$$

Proof. If part: assume $|\{x_j : \bar{c}_i < \underline{c}_j, i \neq j\}| \geq n - p$, meaning that $|\{x_j : \bar{c}_i \geq \underline{c}_j, i \neq j\}| < p$ and take a solution X of p items where x_i is not present. Among the items of this solution is at least one item x_j such that $\bar{c}_i \leq \underline{c}_j$, hence replacing this item with x_i can only improve the solution.

Only if part: let us show, by contradiction, that when $|\{x_j : \bar{c}_i \geq \underline{c}_j, i \neq j\}| \geq p$, x_i cannot be necessarily optimal. If x_i was necessarily optimal, then every solution $X \in \mathcal{O}$ would contain it and include p items. This means that at least one item within $\{x_j : \bar{c}_i \geq \underline{c}_j, i \neq j\}$ is not included in X , say x_j . Clearly, since $\bar{c}_i \geq \underline{c}_j$, by picking $c_i = \bar{c}_i$ (since x_i is necessarily optimal, any replacement would lead to include x_i in an optimal solution) and $c_j = \underline{c}_j$, we could improve the solution while excluding x_i , thus x_i cannot be necessarily optimal.

Note that Proposition 1 tells us that checking whether x_i is necessarily optimal is computationally straightforward, as one can just compare \underline{c}_i to all \bar{c}_j . We will denote by $v \subseteq \Omega$ the set of necessarily optimal items.

Proposition 2. x_i is possibly optimal if and only if $\underline{c}_i \leq \bar{c}_j$ for at least $n - p$ items x_j , i.e.

$$|\{x_j : \underline{c}_i \leq \bar{c}_j, i \neq j\}| \geq n - p$$

Proof. If part: assume $|\{x_j : \underline{c}_i \leq \bar{c}_j, i \neq j\}| \geq n - p$, meaning that $|\{x_j : \underline{c}_i > \bar{c}_j, i \neq j\}| < p$. Now, consider the scenario γ where $c_i = \underline{c}_i$ and $c_j = \bar{c}_j$ for all other items x_j . Then $x_i \in X_\gamma^*$, as strictly less than p items will have a higher score than x_i in scenario γ .

Only if part: again, assume by contradiction that $|\{x_j : \underline{c}_i > \bar{c}_j, i \neq j\}| \geq p$, then clearly any solution including x_i does not include at least one item within $\{x_j : \underline{c}_i > \bar{c}_j, i \neq j\}$, and replacing x_i with such an item would improve the solution, whatever the scenario γ . Therefore, x_i is not in any possible optimal solution, and is not a possibly optimal item.

As for the necessary optimal items, the procedure to detect possibly optimal items is pretty straightforward computationally speaking. We will denote by $\pi \subseteq \Omega$ the set of possibly optimal items. It should be noted that by construction, we have $v \subseteq \pi$, with the two being equal if and only if the set \mathcal{O} is reduced to a singleton. It is also easy to see that $\mathcal{O} \subseteq \{X \in \Phi : v \subseteq X \subseteq \pi\}$, as all necessary items must be in all possibly optimal solutions, and any non-possibly optimal items never appears in such solutions. A natural next question is then to wonder whether we have $\mathcal{O} = \{X \in \Phi : v \subseteq X \subseteq \pi\}$? The next simple example indicates that this is not the case in general.

Example 1. Consider the following example, with 5 objects x_1, \dots, x_5 and their associated interval costs:

$$c_1 \in [1, 10], \quad c_2 \in [2, 5], \quad c_3 \in [3, 6] \\ c_4 \in [4, 8], \quad c_5 \in [7, 9].$$

We now consider the problem of selecting 3 items among those, so that $n = 5$ and $p = 3$. One can easily check that none of those item is necessarily optimal, and that all are possibly optimal. However, the subset $\{x_1, x_4, x_5\}$ is not within \mathcal{O} . To see this, simply consider the scenario $c_1 = 1, c_2 = 5, c_3 = 6, c_4 = 4, c_5 = 7$, whose solution is the subset $\{x_1, x_2, x_4\}$ and not $\{x_1, x_4, x_5\}$.

3 Compact representation of possible optimal solution

In this section we present a polynomial representation of all possible optimal solutions \mathcal{O} , together with an algorithm to build it. Before focusing on the possible optimal solutions we present how to represent all feasible solutions by a directed layered graph $G = (V, E)$ of which one example is given by Figure 1.

The vertices of the graph are such that $V = \cup_{i=0, \dots, n} V_i$ with $V_0 = \{0_0\}$, $V_i = \{a_i, (a+1)_i, \dots, (b-1)_i, b_i\} \forall i = 1, \dots, n-1$ with a_i and b_i respectively the smaller (the bigger) number of selected items at layer i indexed by the layer i i.e. $a_i = \max(0, p - (n - i))$, $b_i = \min(i, p)$ and $V_n = \{p_n\}$. a_i correspond to the minimal number of selected item after considering whether to select the i th item, and b_i the maximal number of selected items. Each value i defines a vertical layer corresponding to object x_i , while each horizontal layer (corresponding to the values going from 0 to p) correspond to the number of objects added in the selection.

The edges $E = \cup_{i=1, \dots, n} E_i$ with $(u, v) \in E_i$ if $u \in V_{i-1}, v \in V_i$ and $u = v$ or $u = v + 1$ which means respectively *The object i is not selected* or *The object i is selected*. Each path from 0_0 to p_n is a feasible solution of the problem. We can see that the number of nodes as well as the number of edge is in $O(np)$. Figure 1 provides an illustration for the case of $n = 5$ objects among which $p = 3$ must be retained.

However, what we are aiming for is not to list all feasible solutions, but to get the set \mathcal{O} of possible optimal solutions, or to trim the graph from its unnecessary edges and vertices. To do this, we shall denote by $v^p(I)$ the set of necessary optimal objects for a given value p with given subset $I \subseteq \Omega$ of objects and $\pi^p(I)$ the set of possibly optimal objects for a given value p with given set of object I .

When a node u_i is reached in the graph, we know that we have considered the $i - 1$ first objects, and selected u of them. We therefore know that we still need to select $p - u$

objects in the remaining set $I = \{i, \dots, n\}$. From this situation, we will see which edges must be added (if we start from the empty graph), or which edge should be removed from the complete graph, depending on the situation of object i in I when $p - u$ objects must be selected. More precisely:

- if $x_i \in v^{p-u}(I)$, only the edge *object is selected* is added to the graph, as it is necessarily selected;
- if $x_i \in \pi^{p-u}(I) \setminus v^{p-u}(I)$, both edges are added to the graph, as it is a possibly but not necessarily optimal selection;
- and if $x_i \notin \pi^{p-u}(I)$, only the edge *object is not selected* is added to the graph.

The procedure to compute the graph of all possible optimal solutions is described in algorithm.1.

Example 2. The graph of all feasible solutions of Example 1 is represented on Figure 1. But each path of the current graph is not possibly optimal for instance the path

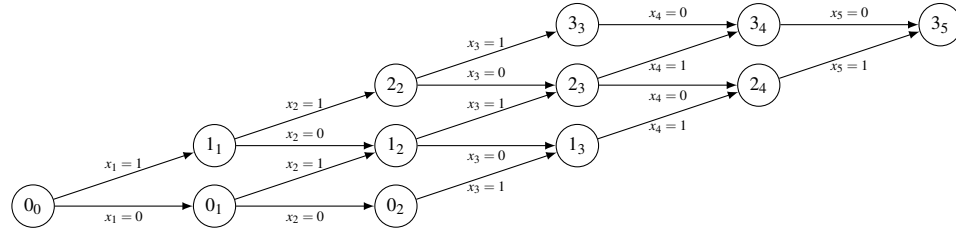


Fig. 1. Set of feasible solutions

$(0_0, 1_1, 1_2, 1_3, 2_4, 3_5)$ which represent the solution $x_1 = 1, x_4 = 1, x_5 = 1$. Let us illustrate the algorithm on this path. At node 0_0 the object 1 is possibly optimal, hence the arc $(0_0, 1_1)$ is in the graph of possible optimal solution (we can add object x_1 to the solution). Since it is not necessarily optimal, we also add $(0_0, 0_1)$. At node 1_1 we check if object 2 is possibly optimal for the updated constraint $\sum_{i=2}^5 x_i \geq 2$ (we still have to select 2 objects). we can see that it is possibly optimal but not necessary. Hence the path $(1_1, 1_2)$ is possible and added. At node 1_2 the constraint becomes $\sum_{i=3}^5 x_i \geq 2$ so object 3 is necessarily optimal since $\bar{c}_3 < \underline{c}_5$, and we still have 2 objects to select. So the edge $(1_2, 1_3)$ is not in the graph of possible optimal solution. The set of all possible optimal solutions is represented on Figure 2

Proposition 3. All paths of graph returned by Algorithm 1 represent all possible optimal solution in space $O(np)$.

Proof. We give a proof by induction on $i = 1, \dots, n$.

Base case with $i = 1$: by definition if x_i is possibly optimal but not necessary then $\exists X^* \in \mathcal{O}$ with $x_1 = 1$ and $\exists X^* \in \mathcal{O}$ with $x_1 = 0$, if it is necessary optimal then $\exists X^* \in \mathcal{O}$ with $x_1 = 1$ and $\nexists X^* \in \mathcal{O}$ with $x_1 = 0$ else $\exists X^* \in \mathcal{O}$ with $x_1 = 0$ and $\nexists X^* \in \mathcal{O}$ with

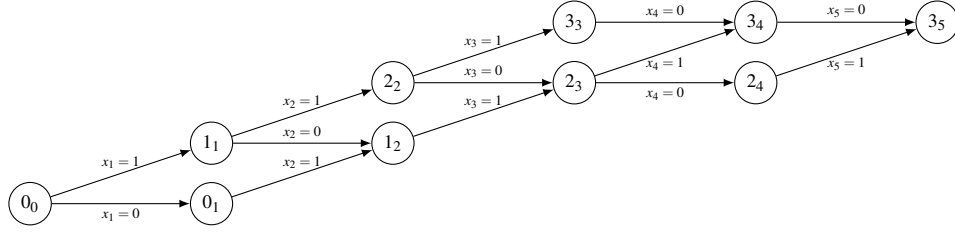


Fig. 2. Set of feasible solutions

$x_1 = 1$. So all paths from node 0_0 to node of the layer 1 are all the partial possibly optimal solutions of x_1 .

Inductive step: We denote P_{i-1} the set of paths from node 0_0 to each the nodes of layer $i-1$. Now we show that if all paths in P_{i-1} are all the partial possibly optimal solution then all paths in P_i are all the partial possibly optimal solution. We assume that is true for $i-1$, more formally $\forall path \in P_{i-1}, \exists X^* \in \mathcal{O}$ such that $X^* = (x_1^{path}, \dots, x_{i-1}^{path}, x_i^*, \dots, x_n^*)$ and $\nexists X \in \mathcal{O}$ such that $(x_1, \dots, x_{i-1}) \in \cup_{path \in P_{i-1}} (x_1^{path}, \dots, x_{i-1}^{path})$. Hence $\exists \gamma^* \in \Gamma$ such that X^* is optimal. Since Γ is hyper-cube the set of scenarios form all nodes of the layer i is $\times_{j=1}^{i-1} \gamma_j^* \times_{j=i}^n [\underline{c}_j, \bar{c}_j]$. Using the Bellman's principle of optimality $X_i^* = (x_i^*, \dots, x_n^*) = \arg \min_{X_i \subseteq I, |X_i|=p-|X^{path}|} F(X_i)$. Since for each nodes of layer i , all paths to it have the same value of $|X^{path}| = K$ and that the scenario set on the remaining sub-set is a hyper-cube $\times_{j=i}^n [\underline{c}_j, \bar{c}_j]$ so what ever the path the possibly optimality of selected item depend only on the restricted problem to I and $|X_i| = p - K$. We can apply at each node of the layer i the propositions 2 and 1. So if x_i is possibly optimal and not necessary optimal for a given node then they exist for each path to it a $X^* \in \mathcal{O}$ such that $X^* = (x_1^{path}, \dots, x_{i-1}^{path}, 1, x_{i+1}^*, \dots, x_n^*)$ and if x_i is necessary optimal for a given node then they not exist for each path to it $X^* \in \mathcal{O}$ such that $X^* = (x_1^{path}, \dots, x_{i-1}^{path}, 0, x_{i+1}^*, \dots, x_n^*)$. So all paths from node 0 to node of the layer 1 are all the partial possibly optimal solution of (x_1, \dots, x_i) . Since both the base case and the inductive step have been proved as true, by mathematical induction All the paths of graph returned by algorithm.1 represent all possible optimal solution. The graph returned by the algorithm has $n+1$ layers with maximum p nodes and for each node maximal two edges.

Proposition 3 confirms that our algorithm return the expected set \mathcal{O} , and that this representation remains compact while listing all the elements of \mathcal{O} , and only those elements. Of course the number of possible paths remain exponential, but this becomes a minor issue, as we can represent them compactly. Note that the necessity or possibility of an object being in a solution can also be directly read from the graph: an object x_i is necessary if all edges from layer $i-1$ to i go upward, and impossible if all edges from layer $i-1$ to i are horizontal. In Figure 2, for instance, one can check that it is not the case for any of the object/layers.

The next proposition is a simple expression of the complexity of our algorithm, which remains polynomial.

Proposition 4. *The set of all possible optimal solutions can be computed in $O(n^2p)$ time.*

Proof. We need to answer at each node of the graph ($O(np)$ nodes) if the x_i is possible optimal, necessary optimal solution that takes $O(n)$ times.

Algorithm 1: Set of all possible optimal solution

Input: cost intervals $[\underline{c}_i, \bar{c}_i], \forall i \in \{1, \dots, n\}$, minimal number p

Output: A graph $G = (V, E)$

$V_0 = \{0_0\}$

$a = \max(0, p - (n - i))$

$b = \min(i, p)$

$V_i = \{a_i, \dots, b_i\} \forall i = 1, \dots, n - 1$

$V_n = \{p_i\}$

$E_i = \emptyset, \forall i = 1, \dots, n$

for $i \in \{1, \dots, n\}$ **do**

$I = \{i, \dots, n\}$

for $j_{i-1} \in V_{i-1}$ **do**

$k = p - j$

if $x_i \in \pi^k(I) \setminus v^k(I)$ **then**

$E_i = E_i \cup \{(j_{i-1}, j + 1_i)\}$

$E_i = E_i \cup \{(j_{i-1}, j_i)\}$

if $x_i \in v^k(I)$ **then**

$E_i = E_i \cup \{(j_{i-1}, j + 1_i)\}$

if $x_i \notin \pi^k(I)$ **then**

$E_i = E_i \cup \{(j_{i-1}, j_i)\}$

$V = \cup_{i=0, \dots, n} V_i$

$E = \cup_{i=1, \dots, n} E_i$

4 Extension to the credal set case

In this section, we no longer assume that c_i uncertainty is described by an interval, but by a convex set \mathcal{P}_i of probabilities having for support a sub-interval of \mathbb{R} , often referred to as *credal sets*. Although more general models of uncertainty do exist, convex probability sets or credal sets are quite generic, as they include most uncertainty representations commonly used in practice, such as intervals, possibility distributions, belief functions, etc.

4.1 Precise probabilities

When the sets \mathcal{P}_i reduce to singletons P_i , i.e., are equivalent to specify probabilities for the costs c_i , a common way justified axiomatically to estimate the quality of a solution is by considering its expected cost. For a single item x_i , this expected cost is $\mathbb{E}_{P_i}(c_i) = \int_{\mathbb{R}} c_i p_i(c_i) dc_i$, where $p(c_i)$ is the probability density.

For the item selection issue and given a set $X \in \Phi$, this means that the solution would be evaluated as

$$\mathbb{E}[F(X)] = \mathbb{E} \left[\sum_{x_i \in X} c_i \right] = \sum_{x_i \in X} \mathbb{E}_{P_i}(c_i),$$

as the expectation is a linear operator, not depending on any possible dependencies between the distributions P_i . The problem is therefore not more complicated than in the case where costs are precisely known, and one can use all previous results on expectations, obtaining the optimal result as

$$X_P^* = \arg \min_{X \in \Phi} \mathbb{E}[F(X)] \quad (3)$$

4.2 Generic credal sets

We now consider the case where each price is defined by a convex set of probability distributions \mathcal{P}_i . While precise probabilities are retrieved when those are singletons, the interval case is recovered when \mathcal{P}_i contains all probabilities whose support is included in $[\underline{c}_i, \bar{c}_i]$. From a set \mathcal{P}_i , one can define lower and upper expectations of the cost c_i as

$$\underline{\mathbb{E}}_{\mathcal{P}_i}(c_i) = \inf_{P_i \in \mathcal{P}_i} \mathbb{E}_{P_i}(c_i), \quad \bar{\mathbb{E}}_{\mathcal{P}_i}(c_i) = \sup_{P_i \in \mathcal{P}_i} \mathbb{E}_{P_i}(c_i).$$

The question is then to know what becomes of $\mathbb{E}[F(X)]$, as in general lower and upper expectations are not additive any more³. However, we can easily show that, in the case where we have marginal credal sets for each costs, then they remain additive, on the proviso that the associated join model include every combinations of distributions within \mathcal{P}_i .

Proposition 5. *Given two credal sets \mathcal{P}_X and \mathcal{P}_Y defined for real-valued quantities X and Y , we have*

$$\underline{\mathbb{E}}(X + Y) = \underline{\mathbb{E}}(X) + \underline{\mathbb{E}}(Y)$$

for any joint credal set $\mathcal{P}_{X,Y}$ whose joints are built from all marginals within \mathcal{P}_X and \mathcal{P}_Y , i.e., $\forall P_X \in \mathcal{P}_X, P_Y \in \mathcal{P}_Y, \exists P_{X,Y} \in \mathcal{P}_{X,Y}$ with those marginals.

Proof. Given a joint credal set $\mathcal{P}_{X,Y}$, we want to find

$$\begin{aligned} \underline{\mathbb{E}}(X + Y) &= \inf_{P_{X,Y} \in \mathcal{P}_{X,Y}} \int_X \int_Y x + yp(x,y) dx dy \\ &= \inf_{P_{X,Y} \in \mathcal{P}_{X,Y}} \int_X \int_Y xp(x,y) dx dy + \int_X \int_Y yp(x,y) dx dy \\ &= \inf_{P_{X,Y} \in \mathcal{P}_{X,Y}} \int_X xp(x) dx + \int_Y yp(y) dy \\ &= \inf_{P_X \in \mathcal{P}_X} \int_X xp(x) dx + \inf_{P_Y \in \mathcal{P}_Y} \int_Y yp(y) dy \end{aligned}$$

³ They are usually super- and sub-additive, i.e., $\underline{\mathbb{E}}(X + Y) \geq \underline{\mathbb{E}}(X) + \underline{\mathbb{E}}(Y)$

with the last inequality being true because $\mathcal{P}_{X,Y}$ is convex and has \mathcal{P}_X and \mathcal{P}_Y for marginals, hence there is a joint within $\mathcal{P}_{X,Y}$ reaching simultaneously $\underline{\mathbb{E}}(X)$ and $\underline{\mathbb{E}}(Y)$

This includes most of the ways used to build joint credal sets, i.e., from independence statements [2, 3] or by using sets of copulas to model dependencies [7]. As the same holds true for upper expectations, we have that for a given solution $X \in \Phi$, we have

$$\underline{\mathbb{E}}(F(X)) = \sum_{x_i \in X} \underline{\mathbb{E}}_{\mathcal{P}_i}(c_i), \quad \overline{\mathbb{E}}(F(X)) = \sum_{x_i \in X} \overline{\mathbb{E}}_{\mathcal{P}_i}(c_i).$$

Given Equation (3), the set of possibly optimal solution readily becomes

$$\mathcal{O} = \{X \in \Phi : \exists P_i \in \mathcal{P}_i \text{ s.t. } X_p^* \in \arg \min_{X \in \Phi} \mathbb{E}[F(X)]\}$$

The transposition of all results obtained for the interval case is then direct, simply replacing the interval bounds by their lower/upper expectation counterparts.

5 Conclusion

In this paper, we investigated the problem of characterising possibly optimal solutions, and reducing them, for the simple problem of selecting optimal items when costs were interval-valued. We have also extended this analysis to much more generic uncertainty models, showing that in this case those more generic uncertainty models did not induce an increased complexity.

One obvious extension of the current work is to look for problems directly extending the selecting item problem, such as the knapsack or the assignment problem. We could also look at other simple problems issued from combinatorial optimisation where the objective function has an additive form, such as the minimum spanning tree problem. In this latter case, some existing results [10] in the vein of those of Section 2 could help us to obtain an efficient compact representation of possibly optimal solutions.

Bibliography

- [1] N. Benabbou and P. Perny. Interactive resolution of multiobjective combinatorial optimization problems by incremental elicitation of criteria weights. *EURO journal on decision processes*, 6(3):283–319, 2018.
- [2] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk, Decision and Policy*, 5(2):165–181, 2000.
- [3] S. Destercke. Independence and 2-monotonicity: Nice to have, hard to keep. *International Journal of Approximate Reasoning*, 54(4):478–490, 2013.
- [4] J. Fortin, P. Zieliński, D. Dubois, and H. Fargier. Interval analysis in scheduling. In *International Conference on Principles and Practice of Constraint Programming*, pages 226–240. Springer, 2005.
- [5] M. Kadziński, T. Tervonen, and J. R. Figueira. Robust multi-criteria sorting with the outranking preference model and characteristic profiles. *Omega*, 55:126–140, 2015.
- [6] A. Kasperski, A. Kurpisz, and P. Zieliński. Approximating the min–max (regret) selecting items problem. *Information Processing Letters*, 113(1-2):23–29, 2013.
- [7] I. Montes, E. Miranda, R. Pelessoni, and P. Vicig. Sklar’s theorem in an imprecise setting. *Fuzzy Sets and Systems*, 278:48–66, 2015.
- [8] M. C. Troffaes. Decision making under uncertainty using imprecise probabilities. *International journal of approximate reasoning*, 45(1):17–29, 2007.
- [9] L. Xia and V. Conitzer. Determining possible and necessary winners given partial orders. *Journal of Artificial Intelligence Research*, 41:25–67, 2011.
- [10] H. Yaman, O. E. Karaşan, and M. Ç. Pınar. The robust spanning tree problem with interval data. *Operations research letters*, 29(1):31–40, 2001.