



# Modeling criteria and project interactions in portfolio decision analysis with the Choquet integral<sup>☆</sup>

Matteo Brunelli<sup>a</sup>, Salvatore Corrente<sup>b,\*</sup>

<sup>a</sup> Department of Industrial Engineering, University of Trento, Italy

<sup>b</sup> Department of Economics and Business, University of Catania, Corso Italia, 55, 95129 Catania, Italy

## ARTICLE INFO

### Keywords:

Portfolio decision analysis  
Interactions between criteria  
Interactions between projects  
Choquet integral  
R&D portfolio evaluation

## ABSTRACT

We present a general framework to deal with multicriteria portfolio decision analysis problems in which between-projects independence or within-project independence do not necessarily hold. The Choquet integral preference model is a non-additive integral widely used in multicriteria decision analysis to take into account the possible interactions between criteria. In this case, we apply the Choquet integral, on the one hand, to deal with the interactions between criteria used to evaluate the projects and, on the other hand, to take into account the interactions between projects in the portfolio global evaluation. To reduce the number of parameters necessary to define the considered preference model and keep the problem tractable, we use the 2-additive Choquet integral that assigns a value to each entity and to each pair of entities only. An example shows how to apply our proposal to a multicriteria portfolio decision analysis problem.

## 1. Introduction

Portfolio selection has been a long standing issue in Operations Research and the well-known knapsack problem is probably the oldest and foremost example. In the knapsack problem, given a reference set, the goal is to find the subset of the reference set that satisfies one or more resource constraints and maximizes a profit function. As often happens, reality presents further elements of complexity such as the need to consider multiple periods [1,2], uncertainty [3], and spatial implications [4], to cite just three extensions of the knapsack problem.

In parallel, in the field of Decision Analysis, the seminal paper by Golabi et al. [5] built the foundations of, and triggered interest in, the problem of selecting the best subset (portfolio) of items, in contrast to the common use of value functions to find a single best alternative. Since then, considerations have been made on the possibility to account for *interdependencies* and *multiple criteria*. The proposal of Stummer and Heidenberger [6] is a good example of how interactions were considered. Their methodology allows for multiple-objectives and the existence of an interaction term in the objective function that is active only if a lower bound threshold on the number of selected projects with some given characteristics is exceeded. The optimization problem proposed by Stummer and Heidenberger [6], which had been solved by the authors by enumeration for up to 30 variables, was then object of some algorithmic studies which tried to solve it with particle swarm optimization [7] and scatter search [8]. The idea of adding a term

in the objective function was followed, among others, by de Almeida and Duarte [9] and Gutjahr and Pichler [10]. In their methods, the terms associated to synergies were not discussed in depth and were bounded from above by the values of the individual projects leading to the possible synergies. Other attempts to model interdependencies between projects or between criteria were made by Gomes et al. [11] and Barbati et al. [12], respectively, in both cases by means of the Choquet integral. Similarly, also Liesiö et al. [13] acknowledged the need to evaluate interactions in Robust Portfolio Management method.

It is safe to say that the formal definition and inception of *Portfolio Decision Analysis* (PDA) as a stand-alone field of study represented a turning point. According to Salo et al. [14], the goal of PDA is to study formulations of portfolio problems, especially with respect to the objective function, such that they are truly representative of the preferences of the Decision Maker (DM). The merit of the contribution by Salo et al. [14] is very much related to their emphasis on the proper use of decision analysis techniques.

More recently, Morton et al. [15] proposed to extend PDA to the case when elements of the reference set can be evaluated on multiple criteria and called it Multiple Criteria Portfolio Decision Analysis (MCPDA). Moreover, they also provided some conditions under which additive value functions can be representative of the preferences of the DM.

<sup>☆</sup> Area: Decision Analysis and Preference Modeling. This manuscript was processed by Associate Editor Sarah Ben Amor.

\* Corresponding author.

E-mail addresses: [matteo.brunelli@unitn.it](mailto:matteo.brunelli@unitn.it) (M. Brunelli), [salvatore.corrente@unict.it](mailto:salvatore.corrente@unict.it) (S. Corrente).

The idea of combining a multicriteria approach with PDA has generated interest and extensions in the fuzzy context [16]. Furthermore, it has pushed others to adapt well-known multiple criteria decision analysis (MCDA) methods, for instance weighing methods as the FI-Tradeoff [17] and the ANP [18], to the concept of portfolio selection. A review on extensions of MCDA methods to the case of portfolio selection was offered by Kandakoglu et al. [19]. Further research has also explored extensions of MCPDA for decisions under uncertainty [20,21] and the applicability of Dominance-based Rough Set Approach [22].

In spite of the independence conditions considered by Morton et al. [15] and the wide use of the additive value function, it has been observed that interactions can appear for subsets of projects as well as criteria. In the former case, especially relevant for portfolios, when multiple projects can be selected, Liesiö et al. [23] considered interactions between projects as “extra benefits and/or cost savings when particular project combinations are started”. Project and product cannibalization is a typical example of negative interaction: consider, for example, the launch of two competing products on the market. Similarly, interactions between criteria can happen when it is too simplistic to express the value of a project as a weighted sum of the contribution of all the relevant criteria, i.e. when the condition of preference independence does not hold [24].

As recalled by Liesiö et al. [25] in a recent survey on PDA, some attempts have been made to relax the conditions behind additive value functions in the search for more flexible models. Two approaches have been used: in the first case, interactions are modeled by means of constraints and/or dummy projects [23] whereas, in the second case, non-additive value functions, e.g. multilinear and multiplicative forms, were employed.

A more axiomatic approach was followed by Liesiö [26], when he considered specific sets of conditions leading to the use of the additive-linear, additive-multilinear, and multilinear forms, respectively. Following the same stream of research, Liesiö and Vilkkumaa [27] showed that two simple conditions imply another specific functional form. Such a form is invariant under permutations of the projects and has only  $(n - 1)$  characterizing parameters, each of which is a positive scaling constant associated to specific cardinalities of subsets of projects, and not to the projects themselves. Hence, this approach assumes equal importance of projects and also that the interaction does not depend on what subset of projects it refers to, but only on its cardinality.

The literature shows the need to account for possible interdependencies at both criteria and project levels. For example, Durbach et al. [28] considered problems with possible non-negative interactions between projects and tested some heuristics, both in terms of their cognitive burden and their accuracy to reach a portfolio that is close to the optimum. They concluded that also their heuristics “must use at least some interaction information”. In fact, the most pragmatic approaches, e.g. Stummer and Heidenberger [6], de Almeida and Duarte [9], only consider interactions between projects and not criteria. On the other hand, other approaches, in the search of axiomatic validation often lead to methods that require a large number of parameters and/or difficult interactions with the DM.

The discrete Choquet integral can be seen as a generalization of both the weighted average and the ordered weighted averaging operators [29]. In its more general formulation it can account for interactions for all possible subsets of a reference set with  $n$  elements, in which case it requires the definition of  $2^n - 2$  parameters, whose elicitation is a deterrent for its use in practical applications. For this reason, compromise solutions appeared in the literature and the most relevant goes under the name of 2-additive Choquet integral: interactions are limited to pairs of elements and yet such an aggregation function is often a very good approximation and satisfies a number of desirable properties [30]. Furthermore, Beliakov [31] studied the scalability of Choquet-based knapsack problems when they are solved using commercial software and showed the much greater tractability of 2-additive capacities with respect to general ones.

We acknowledge the need for computationally tractable and flexible tools for the evaluation of portfolios of projects which can consider interactions at both criteria and projects levels. To fill this gap, in this paper, we will not focus on *either* interactions between criteria *or* interactions between projects. Conversely, the contribution of this paper is a holistic model, based on the 2-additive Choquet integral, to evaluate projects and portfolios in the presence of interactions *both* within the set of criteria *and* the set of projects. The model represents a generalization of well-known methods in MCPDA based on additive value functions. In particular, we present two different approaches: in the first, we provide a unique optimal portfolio based on two representative measures compatible with the preference information provided by the DM; in the second, following the principles of Robust Portfolio Modeling [13] and Stochastic Multicriteria Acceptability Analysis [32], we compute different optimal portfolios taking into account several samples of models compatible with the preferences of the DM presenting the frequencies with which they are obtained.

The rest of the manuscript is organized as follows. Section 2 introduces the necessary notation and some examples motivating a departure from the commonly used additive value function in MCPDA. Section 3 recalls the notions of capacity and Choquet integral. Section 4 presents our proposal to apply the Choquet integral in MCPDA to relax some well-known independence conditions. In particular, Sections 4.4–4.6 will describe the first approach mentioned before aiming to compute a single optimal portfolio. Section 5 shows how to apply this approach to find the optimal portfolio under a number of budget constraints in a well-known example from the literature. Section 6 presents the robust approach showing its application to the same example. Finally, Section 7 summarizes the manuscript and draws some conclusions.

## 2. Notation, definitions and motivating examples

Given a finite non-empty reference set of  $q$  projects  $\{p_1, \dots, p_q\}$ , we call *portfolio* any of its subsets. Within this very general framework, MCPDA aims to employ the body of knowledge of decision analysis to identify portfolios such that the “value function” of the DM is maximized under several constraints [15]. In the following,  $Q$  is the index set of the projects, that is,  $Q = \{1, \dots, q\}$ . Let us also assume that projects are evaluated on  $n$  attributes  $G = \{g_1, \dots, g_n\}$  and let  $N$  be the index set of criteria, that is,  $N = \{1, \dots, n\}$ . Denoting by  $Y_j$  the set of possible values achievable by a project on attribute  $g_j$  we have:

- $Y^i = \prod_{j \in N} Y_j$  is the space of variation of the  $i$ th project. Note that, for sake of simplicity, we shall assume that the sets of attribute levels are the same for each project. Therefore, we denote by  $Y^P$  the space of variation of each project  $i \in Q$ . A vector  $\mathbf{y}^i = (y_1^i, \dots, y_n^i) \in Y^P$  can be interpreted as a description of the  $i$ th project with respect to its attribute levels as  $y_j^i$  is its evaluation on attribute  $g_j$ ;
- Let us consider the multiset<sup>1</sup> of projects  $\mathbf{Y} = \{\mathbf{y}^1, \dots, \mathbf{y}^q\} \in \overbrace{Y^P \times \dots \times Y^P}^{q \text{ times}}$ . We call  $2^{\mathbf{Y}}$  the power set of  $\mathbf{Y}$ , that is, the set composed of all subsets of  $\mathbf{Y}$ . This allows a (multicriterial) definition of a *portfolio* as an element of  $2^{\mathbf{Y}}$ . Before proceeding, let us observe that each portfolio can equivalently be represented by a binary vector in  $\{0, 1\}^q$  using the bijection  $\psi : 2^{\mathbf{Y}} \rightarrow \{0, 1\}^q$  that assigns each  $\mathbf{Y}^k \in 2^{\mathbf{Y}}$  to a binary vector  $\mathbf{z}^{\mathbf{Y}^k} \in \{0, 1\}^q$  such that

$$z_i^{\mathbf{Y}^k} = \begin{cases} 1 & \text{if } \mathbf{y}^i \in \mathbf{Y}^k, \\ 0 & \text{if } \mathbf{y}^i \notin \mathbf{Y}^k. \end{cases}$$

<sup>1</sup> We use multisets instead of sets as it cannot be excluded that two different projects could have the same criteria evaluations, then ending up as being indistinguishable elements of the same set.

**Note 2.1.** In the following, for the sake of simplicity, we shall use interchangeably the terms criterion  $g_j \in G$  and criterion  $j \in N$  as well as project  $y^i \in Y$  and project  $i \in Q$ .

Hereafter, we will assume the existence of a preference relation  $\succeq_Q$  over  $2^Y$ , and a preference relation  $\succeq_P$  over  $Y^P$ . That is, a DM has complete, reflexive, and transitive preferences over  $2^Y$  and  $Y^P$ .  $>_Q$  and  $>_P$  represent the asymmetric parts of  $\succeq_Q$  and  $\succeq_P$ , respectively, while  $\sim_Q$  and  $\sim_P$  represent their symmetric parts. Let us underline that the bijection  $\psi$  allows to define a preference relation  $\succeq_{\{0,1\}^q}$  over  $\{0,1\}^q$  such that, for all  $Y^{k_1}, Y^{k_2} \in 2^Y$  and, consequently,  $z^{Y^{k_1}}, z^{Y^{k_2}} \in \{0,1\}^q$ ,  $Y^{k_1} \succeq_Q Y^{k_2} \Leftrightarrow z^{Y^{k_1}} \succeq_{\{0,1\}^q} z^{Y^{k_2}}$ .

It is useful to define some restrictions (subsets) with respect to the projects and the criteria. These can be used to formulate two independence conditions, which, in turn, will be employed to narrow down the search for value functions that can be representative of  $\succeq_Q$  and  $\succeq_P$ :

- Given a subset of attributes  $J \subseteq N$ , and two projects  $y^{i_1}, y^{i_2} \in Y^P$ ,  $(y^i_{J}, y^i_{N \setminus J}) \in Y^P$  is a project having the evaluations of  $y^{i_1}$  for criteria in  $J$  and the evaluations of  $y^{i_2}$  for criteria in  $N \setminus J$ ;

**Example 2.1.** Considering  $N = \{1, 2, 3, 4\}$ , the two projects  $y^{i_1} = (0.7, 0.2, 0.1, 0.4)$ ,  $y^{i_2} = (0.6, 0.5, 0.2, 0.3)$ , and the criteria subset  $J = \{1, 2\}$  one has

$$(y^i_{J}, y^i_{N \setminus J}) = (0.7, 0.2, 0.2, 0.3).$$

- Given a subset of projects  $I \subseteq Q$ , and two portfolios  $Y^{k_1}, Y^{k_2} \in 2^Y$ ,  $(Y^{k_1}_I, Y^{k_2}_{Q \setminus I})$  is a portfolio having the same projects of  $Y^{k_1}$  regarding  $I$  and the same projects of  $Y^{k_2}$  regarding  $Q \setminus I$ . The same definition can be equivalently stated using vectors of binary variables: if  $z^{Y^{k_1}}, z^{Y^{k_2}} \in \{0,1\}^q$ , then  $(z^{Y^{k_1}}_I, z^{Y^{k_2}}_{Q \setminus I})$  is a binary vector having the same components of  $z^{Y^{k_1}}$  with respect to  $I$  and the same components of  $z^{Y^{k_2}}$  with respect to  $Q \setminus I$ .

**Example 2.2.** Considering  $Q = \{1, 2, 3, 4, 5\}$  and the two portfolios

$$Y^{k_1} = \{y^1, y^2, y^4\} \quad \text{and} \quad Y^{k_2} = \{y^1, y^3, y^5\}$$

if  $I = \{1, 2, 3\} \subseteq Q$ , then,

$$(Y^{k_1}_I, Y^{k_2}_{Q \setminus I}) = \{y^1, y^2, y^5\}.$$

This restriction can be equivalently expressed thinking in terms of binary vectors. Considering the two binary vectors associated to  $Y^{k_1}$  and  $Y^{k_2}$ , we have

$$z^{Y^{k_1}} = (1, 1, 0, 1, 0) \quad \text{and} \quad z^{Y^{k_2}} = (1, 0, 1, 0, 1),$$

and if  $I = \{1, 2, 3\} \subseteq Q$ , then

$$(z^{Y^{k_1}}_I, z^{Y^{k_2}}_{Q \setminus I}) = (1, 1, 0, 0, 1)$$

### 2.1. Independence conditions

Drawing from the previous results by Morton et al. [15], we are ready to recall two independence conditions that can greatly simplify the search for a representative value function. If the first condition holds, then it is enough to evaluate projects with respect to the attributes on which the projects differ, as the attributes for which they have the same score, become irrelevant.

**Definition 2.1 (Within-Project Independence [15]).** Given a set of criteria  $J \subseteq N$ ,  $\succeq_P$  is *within-project independent w.r.t. J* if and only if  $\forall y^{i_1}, y^{i_2} \in Y^P$ ,

$$(y^i_{J}, y^i_{N \setminus J}) \succeq_P (y^j_{J}, y^j_{N \setminus J}) \Rightarrow (y^i_{J}, y^i_{N \setminus J}) \succeq_P (y^j_{J}, y^j_{N \setminus J}),$$

$$\text{for all } y^{i_3}, y^{i_4} \in Y^P.$$

If the condition holds for all  $J \subseteq N$ , then  $\succeq_P$  is *within-project independent over N*.

In the literature on MCDA, within-project independence is known as *mutual preference independence* [33,34]. It is also well-known that, if it holds, then there exist marginal value functions  $u_j : Y_j \rightarrow \mathbb{R}$  such that the value of each project  $y^i \in Y^P$  can be represented by an *additive value function*,

$$u(y^i) = \sum_{j=1}^n u_j(y^i_j). \tag{1}$$

That is, given any two projects  $y^{i_1} = (y^{i_1}_1, \dots, y^{i_1}_n)$  and  $y^{i_2} = (y^{i_2}_1, \dots, y^{i_2}_n)$ , the greater the value  $u(y^i)$  the better the project represented by  $y^i$ , i.e.,

$$u(y^{i_1}) \geq u(y^{i_2}) \Leftrightarrow y^{i_1} \succeq_P y^{i_2}.$$

Several methods have been proposed in the literature to obtain the functions  $u_j$ , starting from those presented by Fishburn [35] arriving to those applying the ordinal regression paradigm [36,37]. However, a description of these methods goes beyond the aim of the paper since our proposal is independent on the way the  $u_j$  functions are obtained.

The next example shows a simple case where within-project independence does not hold and the additive formulation (1) cannot represent the preferences of the DM.

**Example 2.3.** Consider two projects represented by the following two vectors of attribute levels,  $y^{i_1} = (0.7, 0.2, 0.3, 0.4)$  and  $y^{i_2} = (0.6, 0.5, 0.3, 0.4)$ , and the preference  $y^{i_1} \succeq_P y^{i_2}$ . If  $\succeq_P$  is between-project independent with respect to  $J = \{1, 2\}$ , then

$$(0.7, 0.2, y_3, y_4) \succeq_P (0.6, 0.5, y_3, y_4)$$

for any choice of  $y_3$  and  $y_4$ . Conversely, if the preference between  $(0.7, 0.2, y_3, y_4)$  and  $(0.6, 0.5, y_3, y_4)$  depends on  $y_3$  and  $y_4$ , then  $\succeq_P$  cannot be within-project independent with respect to  $J = \{1, 2\}$ . For example, if

$$(0.7, 0.2, 0.1, 0.4) \succeq_P (0.6, 0.5, 0.1, 0.4) \\ \text{and } (0.7, 0.2, 0.2, 0.3) <_P (0.6, 0.5, 0.2, 0.3)$$

then  $\succeq_P$  is not within-project independent with respect to  $J = \{1, 2\}$ . Trying to represent this preference by Eq. (1), we have

- $(0.7, 0.2, 0.1, 0.4) \succeq_P (0.6, 0.5, 0.1, 0.4) \Leftrightarrow u_1(0.7) + u_2(0.2) \geq u_1(0.6) + u_2(0.5)$ ,
- $(0.7, 0.2, 0.2, 0.3) <_P (0.6, 0.5, 0.2, 0.3) \Leftrightarrow u_1(0.7) + u_2(0.2) < u_1(0.6) + u_2(0.5)$ .

Constraints 1. and 2. are contradictory and, therefore, the additive value function in (1) cannot represent the DM's preferences.

The following condition, called between-project independence, can now be defined. Note that it can be interpreted as an extension of the original between-project independence proposed by Morton et al. [15].

**Definition 2.2 (Between-Project Independence).** Given  $Y, Q$  and  $I \subseteq Q$ ,  $\succeq_Q$  is *between-project independent with respect to I* if and only if,  $\forall Y^{k_1}, Y^{k_2} \in 2^Y$ ,

$$(z^{Y^{k_1}}_I, z^{Y^{k_2}}_{Q \setminus I}) \succeq_{\{0,1\}^q} (z^{Y^{k_2}}_I, z^{Y^{k_3}}_{Q \setminus I}) \Rightarrow (z^{Y^{k_1}}_I, z^{Y^{k_4}}_{Q \setminus I}) \succeq_{\{0,1\}^q} (z^{Y^{k_2}}_I, z^{Y^{k_4}}_{Q \setminus I}), \\ \text{for all } Y^{k_3}, Y^{k_4} \in 2^Y.$$

If the definition holds for all  $I \subseteq Q$ , then  $\succeq_Q$  is *between-project independent over Q*.

In simple words, Definition 2.2 implies that the global preference of a portfolio over another is not dependent on the common projects in the set  $Q \setminus I$ .

If  $\succeq_Q$  is between-project independent over  $Q$ , then, the preferences of a DM, can be represented by a function  $\bar{u} : 2^Y \rightarrow [0, 1]$  such that for all  $\mathbf{Y}^k \in 2^Y$

$$\bar{u}(\mathbf{Y}^k) = \sum_{y^i \in \mathbf{Y}^k} w_i \cdot u(y^i) \quad (2)$$

where  $u : Y^P \rightarrow \mathbb{R}$  is the projects value and  $w_i$ 's are suitable positive scaling constants.

In the following example, we shall illustrate a case in which between-project independence does not hold. Meanwhile, we reckon that, whereas criteria independence can sometimes be achieved by properly restructuring a decision problem, e.g. by splitting criteria into subcriteria, project independence cannot be obtained by manipulating the problem because the interdependencies among subsets of projects depend on their intrinsic nature and projects cannot be decomposed into subprojects.

**Example 2.4.** Let us consider 5 different projects  $\mathbf{Y} = \{y^1, \dots, y^5\}$  evaluated on  $n$  criteria and the following portfolios

$$\mathbf{Y}^{k_1} = \{y^1, y^2, y^4\}, \quad \mathbf{Y}^{k_2} = \{y^1, y^3, y^4\},$$

$$\mathbf{Y}^{k_3} = \{y^1, y^2, y^5\}, \quad \mathbf{Y}^{k_4} = \{y^1, y^3, y^5\}.$$

Let us suppose that, for the DM,  $\mathbf{Y}^{k_1}$  is at least as good as  $\mathbf{Y}^{k_2}$  ( $\mathbf{Y}^{k_1} \succeq_Q \mathbf{Y}^{k_2}$ ) and  $\mathbf{Y}^{k_4}$  is preferred to  $\mathbf{Y}^{k_3}$  ( $\mathbf{Y}^{k_4} \succ_Q \mathbf{Y}^{k_3}$ ). This means that

$$\bar{u}(\mathbf{Y}^{k_1}) \geq \bar{u}(\mathbf{Y}^{k_2}) \quad \text{and} \quad \bar{u}(\mathbf{Y}^{k_3}) < \bar{u}(\mathbf{Y}^{k_4}).$$

If we rewrite them using (2), we obtain

$$w_1 \cdot u(y^1) + w_2 \cdot u(y^2) + w_4 \cdot u(y^4) \geq w_1 \cdot u(y^1) + w_3 \cdot u(y^3) + w_4 \cdot u(y^4)$$

$$w_1 \cdot u(y^1) + w_2 \cdot u(y^2) + w_5 \cdot u(y^5) < w_1 \cdot u(y^1) + w_3 \cdot u(y^3) + w_5 \cdot u(y^5),$$

which can be further simplified into

$$w_2 \cdot u(y^2) \geq w_3 \cdot u(y^3) \quad \text{and} \quad w_2 \cdot u(y^2) < w_3 \cdot u(y^3),$$

respectively. This shows a contradiction and therefore there does not exist an additive value function, i.e. (2), representative of the preferences of the DM. The contradiction is obtained because  $\succeq_Q$  is not between-project independent with respect to  $I = \{1, 2, 3\}$  (see Definition 2.2).

If conditions 2.1 and 2.2 hold, then, there exist  $u_j : Y_j \rightarrow \mathbb{R}$  for all  $j \in N$  such that for all  $\mathbf{Y}^k \in 2^Y$

$$\bar{u}(\mathbf{Y}^k) = \sum_{y^i \in \mathbf{Y}^k} w_i \cdot u(y^i) = \sum_{y^i \in \mathbf{Y}^k} w_i \sum_{j=1}^n u_j(y_j^i). \quad (3)$$

Fig. 1 offers a snapshot of the portfolio selection problem when intra- and between-project independence conditions hold and (3) can be used. In addition, in a large number of applications, e.g. [38,39], it is assumed that projects are anonymous and a property of symmetry can be assumed. Consequently,  $w_1 = \dots = w_n$ . As we shall reckon, this greatly simplifies the search for a suitable value function  $\bar{u}$  and it has been considered an initial assumption by many modelization proposals as, e.g., the one by Liesiö [26].

### 3. Non-additive measures and the Choquet integral

Given the restricted applicability of the model based on additive value functions, it is natural to explore more general representations which can accommodate violations of intra- and between-project independence conditions.

**Definition 3.1 (Normalized Capacity [40]).** A normalized capacity over a set  $S$  is a function  $\mu : 2^S \rightarrow [0, 1]$  such that

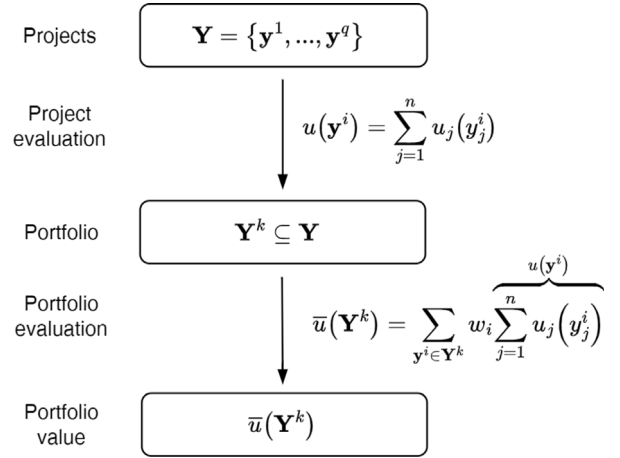


Fig. 1. Portfolio evaluation under intra- and between-project independence conditions.

(1a)  $\mu(\emptyset) = 0$  and  $\mu(S) = 1$ ,

(2a)  $A \subseteq B \subseteq S \Rightarrow \mu(A) \leq \mu(B)$ .

**Definition 3.2 (Discrete Choquet Integral [41]).** Given a list  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  and a normalized capacity  $\mu$  defined over  $S = \{1, \dots, n\}$ , the discrete Choquet integral of  $\mathbf{x}$  with respect to  $\mu$  is defined as

$$Ch_\mu(\mathbf{x}) = \sum_{i=1}^n (x_{\sigma(i)} - x_{\sigma(i-1)}) \mu(A_{\sigma(i)}) \quad (4)$$

where  $\sigma$  is permutation on  $S = \{1, \dots, n\}$  such that  $x_{\sigma(1)} \leq \dots \leq x_{\sigma(n)}$ , with the convention  $x_{\sigma(0)} := 0$  and  $A_{\sigma(i)} := \{\sigma(i), \dots, \sigma(n)\}$ .

Interestingly, weighted averages and ordered weighted averaging (OWA) functions are special cases of the Choquet integral, when the normalized capacity is additive and symmetric, respectively [41]. Hence, the Choquet integral can combine the expressive power of both.

To make things easier, in general, a Möbius transformation  $m$  of the capacity  $\mu$  and  $k$ -additive capacities are used:

- The Möbius transformation  $m$  of the capacity  $\mu$  [42] is a set function  $m : 2^S \rightarrow \mathbb{R}$  such that, for all  $T \subseteq S$

$$\mu(T) = \sum_{R \subseteq T} m(R) \quad \text{and} \quad m(T) = \sum_{R \subseteq T} (-1)^{|T \setminus R|} \mu(R),$$

that is, each normalized capacity is associated to its Möbius representation, and vice-versa;

- A capacity  $\mu$  is  $k$ -additive [43] iff its Möbius transformation  $m$  is such that  $m(T) = 0$  for all  $T \subseteq S : |T| > k$ . In words, a  $k$ -additive capacity is a capacity which can model interactions for subsets of cardinality at most  $k$  and, conversely, shows an additive behavior for subsets with greater cardinalities.

2-additive capacities are a flexible approach for representing non-additive preferences between criteria that have been applied in a number of real-world applications, including citation-based journal ranking [44], preference modeling in hotel selection [45], evaluation of projects for the requalification of abandoned buildings [46], warehouse location selection [47], location selection for underground natural gas storage [48], analysis of SME's propensity for open innovation [49], and selection of logistics operating models [50]. For this reason, we model interactions between both criteria and projects using the Choquet integral equipped with 2-additive capacities. In this case, thanks to the Möbius transformation, constraints (1a) and (2a) can be rewritten as

(1b)  $m(\emptyset) = 0$  and  $\sum_{i \in S} m(\{i\}) + \sum_{\{i,j\} \subseteq S} m(\{i,j\}) = 1$ ,

$$(2b) \quad \forall i \in S \text{ and } \forall T \subseteq S \setminus \{i\}, \quad m(\{i\}) + \sum_{j \in T} m(\{i, j\}) \geq 0,$$

while the Choquet integral in Eq. (4) can now be equivalently rewritten as:

$$Ch_m(x) = \sum_{i \in S} m(\{i\}) \cdot x_i + \sum_{\{i, j\} \subseteq S} m(\{i, j\}) \cdot \min\{x_i, x_j\}. \quad (5)$$

In this context, the importance of  $i \in S$  is not related only to itself but also to its contribution to all possible  $T \subseteq S \setminus \{i\}$ . Analogously, the importance attached to the pair of elements  $\{i, j\}$  depends on their contribution to all possible  $T \subseteq S \setminus \{i, j\}$ . For this reason, the *Shapley index*  $\varphi(\{i\})$  [51] and the *Murofushii index*  $\varphi(\{i, j\})$  [52] are used to evaluate the importance of  $i \in S$  and the importance of  $\{i, j\} \subseteq S$ , respectively. In the case of a 2-additive measure, the two indices can be written in the simple forms,

$$\varphi(\{i\}) = m(\{i\}) + \sum_{j \in S \setminus \{i\}} \frac{m(\{i, j\})}{2}, \quad \varphi(\{i, j\}) = m(\{i, j\}). \quad (6)$$

#### 4. 2-Additive Choquet integral for project and portfolio evaluation

In this section, we shall explain how the 2-additive Choquet integral can be used to compute the value of projects  $y^i \in Y$  and the value of portfolios  $Y^k$  when  $\succsim_P$  is not within-project independent over  $N$  (see Definition 2.1) and  $\succsim_Q$  is not between-project independent over  $Q$  (see Definition 2.2).

##### 4.1. Projects' values computed by means of the 2-additive Choquet integral

Considering the set of criteria  $G = \{g_1, \dots, g_n\}$  on which projects  $y^i$  are evaluated, we consider a 2-additive capacity  $\mu$  defined over  $G$  and its Möbius transformation  $m$ . The value of  $y^i \in Y$  is therefore obtained by

$$\begin{aligned} Ch_m(y^i) &= Ch_m(y^i_1, \dots, y^i_n) \\ &= \sum_{j \in N} m(\{g_j\}) \cdot u_j(y^i_j) \\ &\quad + \sum_{\{j_1, j_2\} \subseteq N} m(\{g_{j_1}, g_{j_2}\}) \cdot \min\{u_{j_1}(y^i_{j_1}), u_{j_2}(y^i_{j_2})\} \end{aligned} \quad (7)$$

and the Möbius parameters have to satisfy the constraints (1b) and (2b) that, in this case, are rewritten as follows:

$$(1c) \quad m(\emptyset) = 0 \text{ and } \sum_{g_j \in G} m(\{g_j\}) + \sum_{\{g_{j_1}, g_{j_2}\} \subseteq G} m(\{g_{j_1}, g_{j_2}\}) = 1,$$

$$(2c) \quad \forall g_{j_1} \in G \text{ and } \forall T \subseteq G \setminus \{g_{j_1}\}, \quad m(\{g_{j_1}\}) + \sum_{g_{j_2} \in T} m(\{g_{j_1}, g_{j_2}\}) \geq 0.$$

In this context, for all  $g_j \in G$ ,  $m(\{g_j\})$  has to be interpreted as the importance of criterion  $g_j$  (when it is considered alone), while for all  $\{g_{j_1}, g_{j_2}\} \subseteq G$ ,  $m(\{g_{j_1}, g_{j_2}\})$  is a value representing the magnitude and the polarity of the interaction between criteria  $g_{j_1}$  and  $g_{j_2}$ : positive (if  $m(\{g_{j_1}, g_{j_2}\}) > 0$ ), negative (if  $m(\{g_{j_1}, g_{j_2}\}) < 0$ ) or null (if  $m(\{g_{j_1}, g_{j_2}\}) = 0$ ). Let us observe that if  $m(\{g_{j_1}, g_{j_2}\}) = 0$  for all  $\{g_{j_1}, g_{j_2}\} \subseteq G$ , then, the 2-additive Choquet integral in Eq. (7) boils down to the additive value function shown in Eq. (1) and, therefore, can be considered its generalization.

**Example 4.1.** Let us reconsider Example 2.3 and try to replicate the preferences of the DM using the 2-additive Choquet integral recalled above. In this case, considering the following DM's preferences

$$y^1 = (0.7, 0.2, 0.1, 0.4) \succsim_P (0.6, 0.5, 0.1, 0.4) = y^2$$

$$y^3 = (0.7, 0.2, 0.2, 0.3) \prec_P (0.6, 0.5, 0.2, 0.3) = y^4$$

and considering a Möbius measure over the set of criteria  $G = \{g_1, g_2, g_3, g_4\}$ , these preferences are translated into the two following constraints:

$$\begin{aligned} Ch_m(y^1) \geq Ch_m(y^2) &\Leftrightarrow 0.1 \cdot m(\{g_1\}) - 0.3 \cdot m(\{g_2\}) \\ &\quad - 0.3 \cdot m(\{g_1, g_2\}) - 0.2 \cdot m(\{g_2, g_4\}) \geq 0 \end{aligned}$$

$$\begin{aligned} Ch_m(y^3) < Ch_m(y^4) &\Leftrightarrow 0.1 \cdot m(\{g_1\}) - 0.3 \cdot m(\{g_2\}) \\ &\quad - 0.3 \cdot m(\{g_1, g_2\}) - 0.1 \cdot m(\{g_2, g_4\}) < 0. \end{aligned}$$

They are satisfied, for example, by the Möbius measure

$$m(\{g_1\}) = 0.5893, \quad m(\{g_2\}) = 0.3750, \quad m(\{g_3\}) = 0.0178$$

$$m(\{g_4\}) = 0.3750, \quad m(\{g_2, g_4\}) = -0.3571$$

considering null the values corresponding to all other interactions. One can therefore observe that, differently from an additive value function, the 2-additive Choquet integral can represent the preference information provided by the DM.

##### 4.2. Portfolios' values and the 2-additive Choquet integral

Let us consider the set of projects  $Y = \{y^1, \dots, y^q\}$  and a portfolio  $Y^k \in 2^Y$ . To compute the value of  $Y^k$  using the 2-additive Choquet integral one has to define a capacity  $\bar{\mu}$  over  $Y$  and, consequently, its Möbius transformation  $\bar{m} : 2^Y \rightarrow \mathbb{R}$ . Known the values of projects  $y^1, \dots, y^q$  computed as described in Eq. (7), that is,  $Ch_m(y^1), \dots, Ch_m(y^q)$ , these values are aggregated by the 2-additive Choquet integral

$$\begin{aligned} Ch_{\bar{m}}(Y^k) &= \sum_{y^i \in Y^k} \bar{m}(\{y^i\}) \cdot Ch_m(y^i) + \\ &\quad + \sum_{\{y^i, y^{i_2}\} \subseteq Y^k} \bar{m}(\{y^i, y^{i_2}\}) \cdot \min\{Ch_m(y^i), Ch_m(y^{i_2})\} \end{aligned} \quad (8)$$

in which  $\bar{m}$  has to satisfy constraints (1b) and (2b) that, in this context, are reformulated as:

$$(1d) \quad \bar{m}(\emptyset) = 0 \text{ and } \sum_{y^i \in Y} \bar{m}(\{y^i\}) + \sum_{\{y^i, y^{i_2}\} \subseteq Y} \bar{m}(\{y^i, y^{i_2}\}) = 1,$$

$$(2d) \quad \forall y^{i_1} \in Y \text{ and } \forall T \subseteq Y \setminus \{y^{i_1}\}, \quad \bar{m}(\{y^{i_1}\}) + \sum_{y^{i_2} \in T} \bar{m}(\{y^{i_1}, y^{i_2}\}) \geq 0.$$

The Möbius parameters in Eq. (8) can be interpreted as follows:

- for all  $y^i \in Y$ ,  $\bar{m}(\{y^i\})$  is the importance of project  $y^i$  in the family of projects  $Y$  (when it is considered alone, therefore, without taking into account interactions between projects);
- for all  $y^{i_1}, y^{i_2} \in Y$ ,  $\bar{m}(\{y^{i_1}, y^{i_2}\})$  is a value representing the interaction between projects  $y^{i_1}$  and  $y^{i_2}$  that can be positive (if  $\bar{m}(\{y^{i_1}, y^{i_2}\}) > 0$ ), negative (if  $\bar{m}(\{y^{i_1}, y^{i_2}\}) < 0$ ) or null (if  $\bar{m}(\{y^{i_1}, y^{i_2}\}) = 0$ ). Let us observe that if  $\bar{m}(\{y^{i_1}, y^{i_2}\}) = 0$  for all  $\{y^{i_1}, y^{i_2}\} \subseteq Y$ , then, the 2-additive Choquet integral boils down to the additive portfolio value shown in Eq. (2) in which  $w_i = m(\{g_i\})$ ,  $\forall y^i \in Y$ .

In this way, we have a more general method to evaluate portfolios, which considers pairwise interactions within the sets of criteria and projects. Fig. 2 presents a sketch of this approach.

**Example 4.2.** We reprise Example 2.4 and try to represent the preferences  $Y^{k_1} \succsim_Q Y^{k_2}$  and  $Y^{k_4} \succ_Q Y^{k_3}$  using the 2-additive Choquet integral described above. Without loss of generality and for the sake of simplicity, let us assume that the five projects  $y^1, \dots, y^5$  have the same value as computed by Eq. (7). The two preferences provided by the DM are therefore translated by the following constraints

$$Ch_{\bar{m}}(Y^{k_1}) \geq Ch_{\bar{m}}(Y^{k_2})$$

$$Ch_{\bar{m}}(Y^{k_3}) < Ch_{\bar{m}}(Y^{k_4})$$

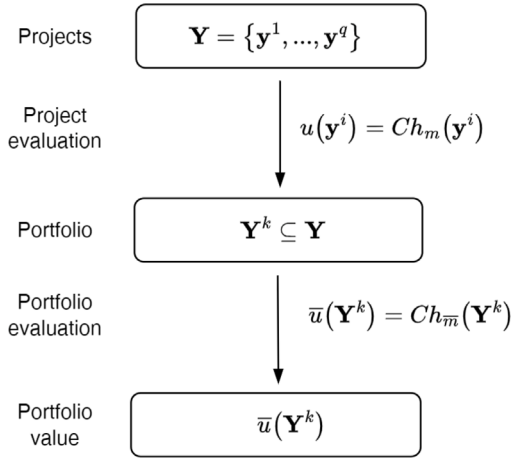


Fig. 2. Portfolio evaluation with the Choquet integrals equipped with 2-additive normalized capacities.

which can be expanded and rearranged into

$$\begin{aligned} & \bar{m}(\{y^2\}) - \bar{m}(\{y^3\}) + \bar{m}(\{y^1, y^2\}) - \bar{m}(\{y^1, y^3\}) \\ & + \bar{m}(\{y^2, y^4\}) - \bar{m}(\{y^3, y^4\}) \geq 0 \\ & \bar{m}(\{y^2\}) - \bar{m}(\{y^3\}) + \bar{m}(\{y^1, y^2\}) - \bar{m}(\{y^1, y^3\}) \\ & + \bar{m}(\{y^2, y^5\}) - \bar{m}(\{y^3, y^5\}) < 0, \end{aligned}$$

respectively. Such inequalities are satisfied, for example, by the Möbius measure

$$\bar{m}(\{y^1\}) = \dots = \bar{m}(\{y^5\}) = \bar{m}(\{y^2, y^4\}) = \bar{m}(\{y^3, y^5\}) = 0.1429$$

considering null the values corresponding to all other pairs of projects. One can therefore observe that, differently from the additive value function in Eq. (2), the 2-additive Choquet integral in Eq. (8) can represent the preferences given by the DM on the four considered portfolios.

#### 4.3. Estimation of parameters

As shown in the previous sections by two illustrative examples, the 2-additive Choquet integral could represent preferences provided by the DM accounting for, on the one hand, interactions between criteria and, on the other hand, interactions between projects. The greater flexibility of the method is nevertheless counterbalanced by the necessity to define a greater number of parameters than those of simpler models such as the weighted sum. Indeed, the use of the 2-additive Choquet integral to compute the comprehensive value of a project  $y^i$  (see Eq. (7)) implies the knowledge of  $n + \binom{n}{2}$  parameters, where  $n$  is the number of criteria on which the projects are evaluated.<sup>2</sup> Analogously, the use of the 2-additive Choquet integral to compute the comprehensive value of a portfolio  $Y^k$  (see Eq. (8)) implies the knowledge of  $q + \binom{q}{2}$  parameters, where  $q$  is the number of projects that can be included in the portfolio.<sup>3</sup> This means that in a small problem as the one we considered in Examples 2.3 and 2.4 where five projects were evaluated on four criteria the number of necessary parameters is twenty-five. Hence, a direct elicitation of the parameters appears prohibitive even for small-scale problems, let alone in settings where possibly hundreds of projects are considered, as shown, for instance, by Mild et al. [38]. For this reason, and for the difficulty of

<sup>2</sup> One value  $m(\{g_j\})$  for each  $g_j \in G = \{g_1, \dots, g_n\}$  and one value  $m(\{g_i, g_j\})$  for each  $\{g_i, g_j\} \subseteq G$ .

<sup>3</sup> One value  $\bar{m}(\{y^i\})$  for each  $y^i \in Y = \{y^1, \dots, y^q\}$  and one value  $\bar{m}(\{y^i, y^j\})$  for each  $\{y^i, y^j\} \subseteq Y$ .

the interpretation of the values of parameters, we do not use a direct preference elicitation, but an indirect one based on the aggregation–disaggregation paradigm [37]. We ask the DM some comprehensive preference statements regarding projects and portfolios from which compatible parameters of the model can be inferred. In this way, the DM is more comfortable in providing such a type of preference in an easy language instead of providing exact numbers to the parameters of the model and, moreover, this task may require a smaller cognitive effort.

To infer the parameters of the 2-additive Choquet integral to compute the projects’ value and the Portfolios’ value, we propose a two-step procedure. In the first step, we infer the Möbius measure  $m$  over the set of criteria  $G$  useful to compute the value of each project (see Section 4.4); on the basis of the projects value computed in this way, in the next step, we compute the Möbius measure  $\bar{m}$  over  $Y$  (see Section 4.5). The reason for not inferring both measures simultaneously will be later clarified.

#### 4.4. Inferring a Möbius measure over the set of criteria $G$

As already suggested, the aggregation–disaggregation approach is used to infer the Möbius measure over  $G = \{g_1, \dots, g_n\}$ . The DM can provide the following preference information in terms of comparison between projects in  $Y$ :

- $y^{i1}$  is at least as good as  $y^{i2}$ , denoted by  $y^{i1} \succeq_P y^{i2}$  (translated to the constraint  $Ch_m(y^{i1}) \geq Ch_m(y^{i2})$ ),
- $y^{i1}$  is preferred to  $y^{i2}$ , denoted by  $y^{i1} \succ_P y^{i2}$  ( $Ch_m(y^{i1}) > Ch_m(y^{i2})$ ),
- $y^{i1}$  is indifferent to  $y^{i2}$ , denoted by  $y^{i1} \sim_P y^{i2}$  ( $Ch_m(y^{i1}) = Ch_m(y^{i2})$ ),
- $y^{i1}$  is preferred to  $y^{i2}$  more than  $y^{i3}$  is preferred to  $y^{i4}$  ( $Ch_m(y^{i1}) - Ch_m(y^{i2}) > Ch_m(y^{i3}) - Ch_m(y^{i4})$  and  $Ch_m(y^{i3}) - Ch_m(y^{i4}) > 0$ ),
- the intensity of preference of  $y^{i1}$  over  $y^{i2}$  is the same as the one of  $y^{i3}$  over  $y^{i4}$  ( $Ch_m(y^{i1}) - Ch_m(y^{i2}) = Ch_m(y^{i3}) - Ch_m(y^{i4})$  and  $Ch_m(y^{i3}) - Ch_m(y^{i4}) > 0$ ),

where  $y^{i1}, y^{i2}, y^{i3}, y^{i4} \in Y$ . Additionally, the DM can also provide further statements on the criteria in  $G$ :

- $g_{j1}$  is at least as important as  $g_{j2}$  (translated to the constraint  $\varphi(\{g_{j1}\}) \geq \varphi(\{g_{j2}\})$ ),
- $g_{j1}$  is more important than  $g_{j2}$  ( $\varphi(\{g_{j1}\}) > \varphi(\{g_{j2}\})$ ),
- $g_{j1}$  and  $g_{j2}$  are equally important ( $\varphi(\{g_{j1}\}) = \varphi(\{g_{j2}\})$ ),
- the difference of importance between  $g_{j1}$  and  $g_{j2}$  is greater than the difference of importance between  $g_{j3}$  and  $g_{j4}$  ( $\varphi(\{g_{j1}\}) - \varphi(\{g_{j2}\}) > \varphi(\{g_{j3}\}) - \varphi(\{g_{j4}\})$  and  $\varphi(\{g_{j3}\}) - \varphi(\{g_{j4}\}) > 0$ ),
- the difference of importance between  $g_{j1}$  and  $g_{j2}$  is the same as the difference of importance between  $g_{j3}$  and  $g_{j4}$  ( $\varphi(\{g_{j1}\}) - \varphi(\{g_{j2}\}) = \varphi(\{g_{j3}\}) - \varphi(\{g_{j4}\})$  and  $\varphi(\{g_{j3}\}) - \varphi(\{g_{j4}\}) > 0$ ),
- $g_{j1}$  and  $g_{j2}$  are positively [negatively] interacting ( $\varphi(\{g_{j1}, g_{j2}\}) > 0$  [ $< 0$ ]),
- $g_{j1}$  and  $g_{j2}$  are not interacting ( $\varphi(\{g_{j1}, g_{j2}\}) = 0$ ),

where  $g_{j1}, g_{j2}, g_{j3}, g_{j4} \in G$ . Let us observe that, to get more precise information about the type of interactions between criteria and about their magnitude, one could also use the methodological background presented by Siskos and Burgherr [53]. Indeed, the type of information asked in that case can easily be translated into the constraints of our model.

At this point we call *compatible model* a Möbius measure  $m$  over  $2^G$  satisfying the following set of constraints

$$E^{DM}, \left. \begin{aligned} \sum_{g_j \in G} m(\{g_j\}) + \sum_{\{g_{j_1}, g_{j_2}\} \subseteq G} m(\{g_{j_1}, g_{j_2}\}) &= 1, \\ m(\{g_{j_1}\}) + \sum_{g_{j_2} \in T} m(\{g_{j_1}, g_{j_2}\}) &\geq 0, \quad \forall g_{j_1} \in G \quad \text{and} \quad \forall T \subseteq G \setminus \{g_{j_1}\} \end{aligned} \right\} E_m \quad (9)$$

where  $E^{DM}$  is the set of constraints translating the preference information described above. After transforming strict inequalities into weak ones using an auxiliary variable  $\epsilon$  (for example,  $Ch_m(y^{i_1}) > Ch_m(y^{i_2})$  is transformed into  $Ch_m(y^{i_1}) \geq Ch_m(y^{i_2}) + \epsilon$ ), to check for the existence of at least one compatible model one has to solve the following LP problem

$$\begin{aligned} \epsilon_m = \max \epsilon \quad \text{subject to,} \\ E_m. \end{aligned} \quad (10)$$

If  $E_m$  is feasible and  $\epsilon_m > 0$ , then, at least one compatible model exists. In the opposite case ( $E_m$  is infeasible or  $\epsilon_m \leq 0$ ), there is not any compatible model and the reason can be checked, for example, using one of the methods presented in [54].<sup>4</sup> If  $E_m$  is feasible and  $\epsilon_m > 0$ , in general, more than one compatible model exists, and, consequently, choosing only one of them could be considered arbitrary. Denoting by  $S_m$  the space of compatible models defined by the set of constraints  $E_m \cup \{0 \leq \epsilon \leq \epsilon_m\}$ , to make the inference procedure more robust, we propose to compute the barycenter of a large sample of compatible models in  $S_m$ . At first, since the constraints in  $E_m \cup \{0 \leq \epsilon \leq \epsilon_m\}$  define a convex and bounded polytope, the Hit-And-Run (HAR) method [55,56] can be applied to sample a large number of compatible models. We then estimate its barycenter [57]<sup>5</sup> and use it to compute the value of each project  $y^i \in Y$ , that is,  $Ch_m(y^i)$ .

#### 4.5. Inferring a Möbius measure over the set of projects $Y$

Once the projects' values have been obtained, the computation of the value of each portfolio  $Y^k \in 2^Y$ , as shown in Eq. (8), implies the knowledge of a Möbius measure over  $Y$ . To get such a measure, we apply again the aggregation–disaggregation approach asking the DM to provide some preference information on projects in terms of their importance or interaction:

- $y^{i_1}$  is at least as important as  $y^{i_2}$  (translated to the constraint  $\varphi(\{y^{i_1}\}) \geq \varphi(\{y^{i_2}\})$ ),
- $y^{i_1}$  is more important than  $y^{i_2}$  ( $\varphi(\{y^{i_1}\}) > \varphi(\{y^{i_2}\})$ ),
- $y^{i_1}$  and  $y^{i_2}$  have the same importance ( $\varphi(\{y^{i_1}\}) = \varphi(\{y^{i_2}\})$ ),
- the difference of importance between  $y^{i_1}$  and  $y^{i_2}$  is greater than the difference of importance between  $y^{i_3}$  and  $y^{i_4}$  ( $\varphi(\{y^{i_1}\}) - \varphi(\{y^{i_2}\}) > \varphi(\{y^{i_3}\}) - \varphi(\{y^{i_4}\})$  and  $\varphi(\{y^{i_3}\}) - \varphi(\{y^{i_4}\}) > 0$ ),
- the difference of importance between  $y^{i_1}$  and  $y^{i_2}$  is the same of the difference of importance between  $y^{i_3}$  and  $y^{i_4}$  ( $\varphi(\{y^{i_1}\}) - \varphi(\{y^{i_2}\}) = \varphi(\{y^{i_3}\}) - \varphi(\{y^{i_4}\})$  and  $\varphi(\{y^{i_3}\}) - \varphi(\{y^{i_4}\}) > 0$ ),

<sup>4</sup> If  $\epsilon$  does not appear in any of the constraints in  $E^{DM}$  (because the decision-maker did not provide any strict preference regarding projects or criteria as well as any information about the possible interaction between criteria), the space defined by the set of constraints  $E_m$  is unbounded since  $\epsilon$  is not constrained (neither from below nor from above). To avoid this, it is sufficient to add the constraint  $-M \leq \epsilon \leq M$ , with  $M$  being a positive number, to  $E_m$ .

<sup>5</sup> Let us observe that a compatible model can be represented by a vector of Möbius parameters satisfying constraints in  $E_m$ . Therefore, the barycenter can be computed averaging, component by component, the sampled compatible vectors [57].

- $y^{i_1}$  and  $y^{i_2}$  are positively [negatively] interacting ( $\varphi(\{y^{i_1}, y^{i_2}\}) > 0$  [ $< 0$ ]),
- $y^{i_1}$  and  $y^{i_2}$  are not interacting ( $\varphi(\{y^{i_1}, y^{i_2}\}) = 0$ ),

where  $y^{i_1}, y^{i_2}, y^{i_3}, y^{i_4} \in Y$ .

Analogously, the DM can provide some information on portfolios they know well:

- $Y^{k_1}$  is at least as good as  $Y^{k_2}$ , denoted by  $Y^{k_1} \succeq_Q Y^{k_2}$  ( $Ch_{\bar{m}}(Y^{k_1}) \geq Ch_{\bar{m}}(Y^{k_2})$ ),
- $Y^{k_1}$  is preferred to  $Y^{k_2}$ , denoted by  $Y^{k_1} \succ_Q Y^{k_2}$  ( $Ch_{\bar{m}}(Y^{k_1}) > Ch_{\bar{m}}(Y^{k_2})$ ),
- $Y^{k_1}$  is indifferent to  $Y^{k_2}$ , denoted by  $Y^{k_1} \sim_Q Y^{k_2}$  ( $Ch_{\bar{m}}(Y^{k_1}) = Ch_{\bar{m}}(Y^{k_2})$ ),
- $Y^{k_1}$  is preferred to  $Y^{k_2}$  more than  $Y^{k_3}$  is preferred to  $Y^{k_4}$  ( $Ch_{\bar{m}}(Y^{k_1}) - Ch_{\bar{m}}(Y^{k_2}) > Ch_{\bar{m}}(Y^{k_3}) - Ch_{\bar{m}}(Y^{k_4})$  and  $Ch_{\bar{m}}(Y^{k_3}) - Ch_{\bar{m}}(Y^{k_4}) > 0$ ),
- the intensity of preference of  $Y^{k_1}$  over  $Y^{k_2}$  is the same as the one of  $Y^{k_3}$  over  $Y^{k_4}$  ( $Ch_{\bar{m}}(Y^{k_1}) - Ch_{\bar{m}}(Y^{k_2}) = Ch_{\bar{m}}(Y^{k_3}) - Ch_{\bar{m}}(Y^{k_4})$  and  $Ch_{\bar{m}}(Y^{k_3}) - Ch_{\bar{m}}(Y^{k_4}) > 0$ ),

where  $Y^{k_1}, Y^{k_2}, Y^{k_3}, Y^{k_4} \in 2^Y$ .

If we use  $E_Q^{DM}$  to denote the set of constraints translating the preference information expressed by a DM as the above mentioned linear equalities and inequalities, then a *compatible model* is a Möbius measure  $\bar{m}$  over  $Y$  satisfying the following set of constraints

$$E_Q^{DM}, \left. \begin{aligned} \sum_{y^i \in Y} \bar{m}(\{y^i\}) + \sum_{\{y^{i_1}, y^{i_2}\} \subseteq Y} \bar{m}(\{y^{i_1}, y^{i_2}\}) &= 1, \\ \bar{m}(\{y^{i_1}\}) + \sum_{y^{i_2} \in T} \bar{m}(\{y^{i_1}, y^{i_2}\}) &\geq 0, \quad \forall y^{i_1} \in Y \quad \text{and} \quad \forall T \subseteq Y \setminus \{y^{i_1}\}. \end{aligned} \right\} E_{\bar{m}} \quad (11)$$

Considering that, also in this case, the strict inequalities are converted into weak ones by the use of the auxiliary variable  $\epsilon$ , to check for the existence of a compatible model one has to solve the following LP problem:

$$\begin{aligned} \epsilon_{\bar{m}} = \max \epsilon, \quad \text{subject to,} \\ E_{\bar{m}}. \end{aligned} \quad (12)$$

If  $E_{\bar{m}}$  is feasible and  $\epsilon_{\bar{m}} > 0$ , then, at least one compatible model exists, while, in the opposite case ( $E_{\bar{m}}$  is infeasible or  $\epsilon_{\bar{m}} \leq 0$ ), there is not any compatible model and the reasons can be checked by using one of the two methods presented in [54].<sup>6</sup>

If  $E_{\bar{m}}$  is feasible and  $\epsilon_{\bar{m}} > 0$ , in general, more than one compatible model  $\bar{m}$  exists. Also in this case, denoting by  $S_{\bar{m}}$  the space of compatible models defined by the set of constraints  $E_{\bar{m}} \cup \{0 \leq \epsilon \leq \epsilon_{\bar{m}}\}$ , to choose one representative compatible model, we apply the same procedure used in the previous section to select one Möbius measure  $m$  over  $G$  among those in  $S_m$ . At first, we sample a large number of compatible models from  $S_{\bar{m}}$  by using, for example, HAR. After that, its barycenter is estimated by averaging the sampled compatible models, and then used to compute the value of each portfolio  $Y^k \in 2^Y$  as shown in Eq. (8).

Other strategies could be followed to estimate  $m$  and  $\bar{m}$ . For example, the two steps presented above can be merged into a unique step. However, this would result in a unique problem where  $\binom{n}{2} + \binom{q}{2}$  parameters had to be estimated at once. Compared to the divide and conquer logic inspiring the two steps procedure, this would be both cognitively and computationally more demanding. In fact, (i) more questions on pairs of portfolios may be necessary, and these are more cognitively demanding than questions between pairs of projects,

<sup>6</sup> The considerations made in footnote 4 hold also for the LP problem (12).

and (ii) sampling would not be done from a convex and bounded polytope (for which efficient algorithms, such as HAR, exist) but from a more general feasible set stemming from inequalities using Eq. (8).

Let us conclude this section observing that, if  $E_{\bar{m}}$  is infeasible, or  $\epsilon_{\bar{m}} \leq 0$  then, the absence of 2-additive measure  $\bar{m}$  over  $\mathbf{Y}$  compatible with the preferences provided by the DM could be explained by the necessity to take into account interactions between more than two projects.<sup>7</sup> In this case, one could overcome the problem by taking into account a  $k$ -additive capacity with  $k > 2$ . In contrast, we should note that a significantly larger number of parameters need to be specified, resulting in a need for a substantial amount of preference information to be provided by the DM to assign them a value.

#### 4.6. Finding an optimal portfolio

Having presented the two-step procedure for computing the value of a given portfolio with the 2-additive Choquet integral, we can now formulate an optimization problem to obtain the optimal portfolio; that is, the portfolio with the greatest value which satisfies a number of constraints. In other words, the entries of the binary vector  $\mathbf{z} = (z_1, \dots, z_q)$  can be considered as variables and therefore, based on Eq. (8), the objective function can be written as

$$Ch_{\bar{m}}(\mathbf{Y}, \mathbf{z}) = \sum_{i=1}^q z_i \cdot \bar{m}(\{y^i\}) \cdot Ch_m(y^i) + \sum_{\{i_1, i_2\} \subseteq Q} z_{i_1} \cdot z_{i_2} \cdot \bar{m}(\{y^{i_1}, y^{i_2}\}) \cdot \min\{Ch_m(y^{i_1}), Ch_m(y^{i_2})\} \quad (13)$$

where  $z_i$ 's are the only variables contained in the function since  $Ch_m(y^i)$ , for all  $i \in Q$ , were computed in the first step of the procedure, while,  $\bar{m}(\{y^i\})$  and  $\bar{m}(\{y^{i_1}, y^{i_2}\})$ , with  $i, i_1, i_2 \in Q$  were computed in the second step of the procedure.

Assuming that the DM provides some linear constraints of different nature (resource constraints, cost constraints, logic constraints, etc.) [9], the optimization problem becomes

$$\begin{aligned} \max_{\mathbf{z} \in \{0,1\}^q} Ch_{\bar{m}}(\mathbf{Y}, \mathbf{z}) \quad \text{subject to} \\ \left. \begin{aligned} \mathbf{C}_{Ineq} \mathbf{z} &\leq \mathbf{b}_{Ineq} \\ \mathbf{C}_{Equal} \mathbf{z} &= \mathbf{b}_{Equal} \end{aligned} \right\} \quad (14) \end{aligned}$$

where  $\mathbf{C}_{Ineq} \mathbf{z} \leq \mathbf{b}_{Ineq}$  and  $\mathbf{C}_{Equal} \mathbf{z} = \mathbf{b}_{Equal}$  are matrix-form formulations of inequality and equality constraints, respectively. We call  $\mathbf{z}^{Opt}$  the solution of the optimization problem above,  $\mathbf{Y}^{Opt} \in 2^{\mathbf{Y}}$  its related portfolio composed of projects  $y^i \in \mathbf{Y}$  such that  $z_i^{Opt} = 1$  and  $Ch_{\bar{m}}(\mathbf{Y}, \mathbf{z}^{Opt})$  the value of the optimal portfolio.

### 5. An example

Let us consider an extension of the portfolio selection problem presented by de Almeida and Duarte [9] where R&D projects were evaluated on four criteria ( $g_1$ : Expected Return,  $g_2$ : Probability of Success Associated with the project,  $g_3$ : Degree of Strategic Impact on the organization,  $g_4$ : Degree of impact on the operational processes) to build a portfolio of projects. Using the notation introduced in the previous sections,  $G = \{g_1, g_2, g_3, g_4\}$ ,  $N = \{1, 2, 3, 4\}$ ,  $\mathbf{Y} = \{y^1, \dots, y^{10}\}$ , and  $Q = \{1, \dots, 10\}$ .

The evaluations of the projects on the considered criteria are shown in Table 1.

Let us observe that projects  $y^3$ - $y^{10}$  are the same considered by de Almeida and Duarte [9], whereas the first two projects,  $y^1$  and  $y^2$ , were introduced by us.

Table 1

Performances of the ten R&D projects on the four considered criteria and their value.

Project	$g_1$	$g_2$	$g_3$	$g_4$	$Ch_m(y^i)$
$y^1$	0.25	0.2	0.2	0.6	<b>0.2765</b>
$y^2$	0.15	0.2	0.1	0.8	<b>0.2581</b>
$y^3$	0.7	0.15	0.7	0.4	0.5305
$y^4$	0.9	0.05	0.9	0.9	0.6733
$y^5$	0.4	0.5	0.35	0.3	0.4029
$y^6$	0.4	0.6	0.3	0.4	0.4494
$y^7$	0.3	0.7	0.1	0.7	0.5066
$y^8$	0.25	0.8	0.2	0.6	<b>0.5203</b>
$y^9$	0.15	0.8	0.1	0.8	<b>0.5348</b>
$y^{10}$	0.05	0.9	0.05	0.7	0.5109

Applying the aggregation–disaggregation approach, let us assume that the DM provides the following preference information:  $y^1$  is preferred to  $y^2$  and  $y^9$  is preferred to  $y^8$ . Formally, considering the binary relation  $\succ_P$  between projects, this preference can be written as:

$$y^1 \succ_P y^2, \quad \text{and} \quad y^9 \succ_P y^8.$$

It can be seen that, considering  $J = \{1, 3, 4\} \subseteq N$ ,  $\succ_P$  is not within-project independent with respect to  $J$  (see Definition 2.1) and, therefore, over  $N$ . Consequently, an additive value function as the one in Eq. (1) cannot represent this preference information.

We shall therefore use the procedure described in Section 4.4 to check for a Möbius measure over  $N$  so that the 2-additive Choquet integral can represent the preferences given by the DM. Hence, in  $E_m$ , in addition to the two linear constraints translating the two pieces of preference given by the DM, following what was considered in [9], we add also the constraints translating the preference information over the set of criteria as follows:

- “Expected Return” and “Probability of Success Associated with the project” are equally important (translated to the constraint  $\varphi(\{g_1\}) = \varphi(\{g_2\})$ ),
- “Probability of Success Associated with the project” is more important than “Degree of Strategic Impact on the organization” ( $\varphi(\{g_2\}) > \varphi(\{g_3\})$ ),
- “Degree of Strategic Impact on the organization” is more important than “Degree of impact on the operational processes” ( $\varphi(\{g_3\}) > \varphi(\{g_4\})$ ).

Solving the LP feasibility problem (10), we find that  $E_m$  is feasible and  $\epsilon_m > 0$ . Therefore, there are infinitely many models compatible with the preferences given by the DM. We sampled 100,000 Möbius measures from the space defined by the constraints in  $E_m$  finding, consequently, an estimation of its barycenter. More precisely we have  $m(\{g_1\}) = 0.4055$ ,  $m(\{g_2\}) = 0.3362$ ,  $m(\{g_3\}) = 0.2706$ ,  $m(\{g_4\}) = 0.1446$ . Similarly, the Möbius values for pairs of criteria, which represents the polarity and the intensity of interactions, can be found in Fig. 3

Computing the Choquet integral of each project using Eq. (7) and considering the estimation of the barycenter of  $E_m$ , the values of the projects are obtained and shown on the last column of Table 1. As one can see, the two preference statements  $y^1 \succ_P y^2$  and  $y^9 \succ_P y^8$  are satisfied.

At this point, to compute the values of portfolios, we proceed as explained in Section 4.5. We assume that the DM is able to express some preferences on the following four portfolios

$$\begin{aligned} \mathbf{Y}^1 &= \{y^1, y^2, y^4\}, \quad \mathbf{Y}^2 = \{y^1, y^3, y^4\}, \\ \mathbf{Y}^3 &= \{y^1, y^2, y^5\}, \quad \mathbf{Y}^4 = \{y^1, y^3, y^5\} \end{aligned}$$

stating that  $\mathbf{Y}^1$  is preferred to  $\mathbf{Y}^2$  and  $\mathbf{Y}^4$  is preferred to  $\mathbf{Y}^3$ . Formally, considering the preference relation  $\succ_Q$  over  $2^{\mathbf{Y}}$ , this can be written

<sup>7</sup> The same applies in case  $E_m$  is infeasible or  $\epsilon_m \leq 0$  meaning that there is not any 2-additive measure over  $G$  able to represent the preferences of the DM.



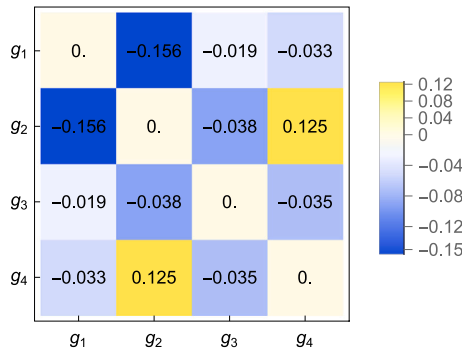


Fig. 3. Möbius values for pairs of criteria.

as  $Y^1 \succ_Q Y^2$  and  $Y^4 \succ_Q Y^3$ . Also in this case, independently of the values associated to the ten projects, an additive value function, as the one shown in Eq. (2), is not able to represent this preference. Indeed, considering  $I = \{1, 2, 3\}$ , it can be seen that  $\succ_Q$  is not between-project independent with respect to  $I$  and, consequently, over  $Q$  (see Definition 2.2 and Example 2.4).

The DM includes further preference information on projects:

- all projects have the same importance when considered alone: this piece of preference information is translated to the constraints  $\bar{m}(y^1) = \dots = \bar{m}(y^{10})$ ,
- $y^6, y^7, y^8, y^9, y^{10}$  do not interact each other: this piece of preference information is translated to the following constraints:

$$\bar{m}(y^{i_1}, y^{i_2}) = 0, \quad \text{for all } (i_1, i_2) \\ \text{such that } i_1 = 6, \dots, 9 \text{ and } i_2 = i_1 + 1, \dots, 10.$$

All these constraints will therefore compose the set  $E_Q^{DM}$  in the LP problem (12).

To check if the 2-additive Choquet integral in Eq. (8) is able to represent this preference information, we solved the LP feasibility problem (12) finding that  $E_m$  is feasible and  $\varepsilon_m > 0$ . This means that it is possible to define infinitely many Möbius measure over  $Y$  such that the 2-additive Choquet integral can represent the preferences given by the DM. Again, we sampled and averaged 100,000 compatible models from the space defined by the constraints in  $E_m$  to estimate its barycenter. The Möbius values of projects are all equal to 0.0522 while the Möbius values for pairs, representing interactions, are reported in Fig. 4.

Considering the constraints presented in [9] regarding Work Force, Equipment, Energy and Cost associated to the 10 projects as given in Table 2, we can formulate the optimization problem to be solved to obtain the optimal portfolio consisting, therefore, on the choice of the projects in  $Y$  maximizing the portfolio value and such that the four constraints mentioned above are satisfied.

The optimization problem can therefore be formulated as follows:

$$\begin{aligned} \max_{z \in \{0,1\}^{10}} Ch_{\bar{m}}(Y, z) \quad \text{subject to} \\ \left. \begin{aligned} 10z_1 + 15z_2 + \dots + 3z_{10} &\leq 60, \\ 39z_1 + 30z_2 + \dots + 12z_{10} &\leq 160, \\ 65z_1 + 70z_2 + \dots + 55z_{10} &\leq 380, \\ 190z_1 + 160z_2 + \dots + 40z_{10} &\leq 1000. \end{aligned} \right\} E_{Constraints} \end{aligned} \quad (15)$$

Solving it, we found  $z^{Opt} = (0, 0, 0, 1, 1, 0, 1, 1, 1, 1) \in \{0, 1\}^{10}$  and, consequently, the optimal portfolio is  $Y^{Opt} = \{y^4, y^5, y^7, y^8, y^9, y^{10}\} \subseteq Y$  with value  $Ch_{\bar{m}}(Y, z^{Opt}) = 0.2174$ . Checking if there exists another portfolio presenting the same maximal value, we solve again problem (15) adding

$$z_4 + z_5 + z_7 + z_8 + z_9 + z_{10} \leq 5$$

Table 2  
Constraints on the ten R&D projects under consideration.

Project	Work force	Equipment	Energy	Cost
$y^1$	10	39	65	190
$y^2$	15	30	70	166
$y^3$	18	38	63	205
$y^4$	35	45	80	250
$y^5$	8	20	53	107
$y^6$	8	18	58	112
$y^7$	5	20	58	97
$y^8$	5	12	60	83
$y^9$	3	16	54	85
$y^{10}$	3	12	55	40
Availability	60	160	380	1000

to  $E_{Constraints}$  to make the previously obtained optimal solution unfeasible. This yields the optimal value 0.2171, which is lower than the original optimum, meaning that the optimal portfolio  $Y^{Opt}$  is unique.

### 5.1. Analysis of the results

Given the relatively small size of the problem, its solutions can be enumerated. Fig. 5 shows that the optimal point saturates, or almost saturates, the constraints on workforce and energy (see Figs. 5(a) and 5(c)), while the other constraints have a significant slack (see Figs. 5(b) and 5(d)). Thus, this suggests that there is probably space for rearranging resources. In fact, no other point in Fig. 5(d) exists with (i) a higher value (ii) that satisfies all the constraints. If we consider a point with higher intercept, as the red dot in Fig. 5(c), we can see that it yields a greater value, but violates the constraints.

Many applications in portfolio decision analysis consider only one budget constraint as, very often, other resources can be acquired at a given cost and therefore multiple constraints can be aggregated into one expressed in monetary terms. In this case, we can consider the optimization problem with the budget constraint only and study the inclusion/exclusion of projects from the optimal portfolio as a function of the budget. Fig. 6 shows the results. Clearly, the optimal portfolio is empty when the budget is null, and the first project to be added is the tenth, as it is the least expensive. Note also that, due to this characteristic, the tenth project is often included and then excluded in the optimal portfolio and acts as a “filler project”.

## 6. An alternative approach

It is important to acknowledge that the methodology presented until now relied on the accuracy of the barycenters of the spaces  $S_m$  and  $S_{\bar{m}}$  in representing the DM’s preferences. This holds true when the expressed preference information is specific enough to narrow down the explored spaces  $S_m$  and  $S_{\bar{m}}$ . However, in all other cases, this assumption may not be reliable.

A more cautious and exploratory approach—which includes a variability analysis of the optimal portfolios that can be obtained by probing the two spaces—can be inspired by the principles of Robust Portfolio Modeling [13] and Stochastic Multicriteria Acceptability Analysis [32,58]. This approach involves a sampling phase in  $S_m$  to search for a compatible model  $m$  over  $G$ , followed by a sampling phase in  $S_{\bar{m}}$  to find compatible models  $\bar{m}$ . Next, the optimal portfolio is computed for each pair  $(m, \bar{m})$ . Finally, the resulting set of portfolios is analyzed and used to aid the selection of a final portfolio. This procedure is sketched and compared to the original one in Fig. 7.

Let us observe that the sampling of  $m$  from  $S_m$  and the sampling of  $\bar{m}$  from  $S_{\bar{m}}$  are not independent since the second is dependent on the first.

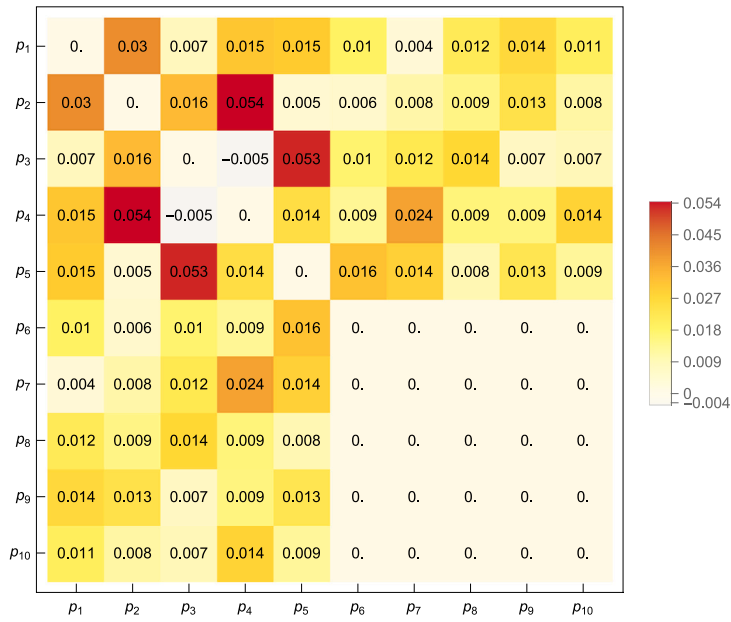


Fig. 4. Interaction values represented by Möbius transforms for pairs of projects.

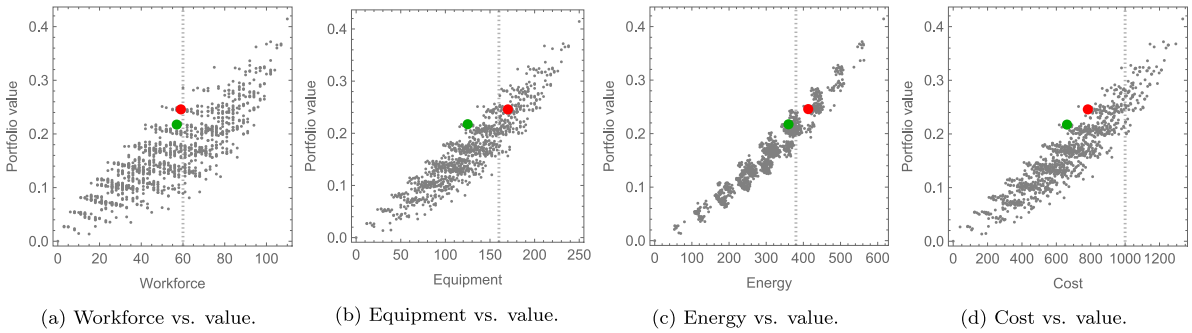


Fig. 5. Projections of all 1024 ( $2^{10}$ ) possible portfolios, feasible and unfeasible, onto 2-dimensional planes where one dimension is the portfolio value and the other is its demand in terms of one of the problem constraints. Availability thresholds of resources are denoted by a vertical dashed line, and the optimal portfolio with a green larger dot. For a correct visualization of colors the reader may have to refer to the online version.

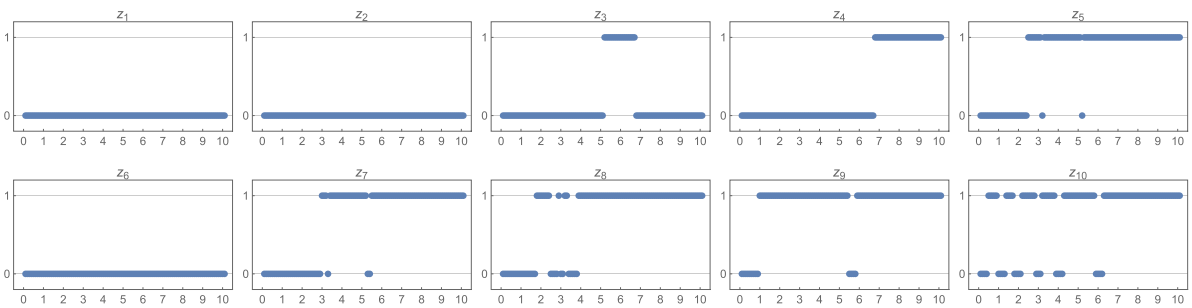


Fig. 6. Inclusion of projects into the optimal portfolios with respect to different budget levels. On the x-axis one find the budget  $\times 100$ .

Indeed, as evident in Eq. (8), the computation of the Portfolios' values is dependent on the Projects' values that are computed by Eq. (7). From a formal point of view, each  $m \in S_m$  defines a space  $S_{\bar{m}}$  from which a model  $\bar{m}$  could be sampled. To take into account this dependence, first, we sample 1,000 compatible models  $m$  from  $S_m$ . Then, for each  $m$  sampled in  $S_m$  we sample another 1,000  $\bar{m}$  from  $S_{\bar{m}}$  so that a total of 1,000,000 pairs  $(m, \bar{m})$  compatible with the DM's preferences are available. The portfolio optimization problem (14) is solved for each of them. The result is a list of potentially optimal portfolios with their frequencies.

If the set of potentially optimal portfolios with their relative frequencies is considered too dispersed, two strategies can be employed. The first strategy contemplates the elicitation of further preference information from the DM: such information, expressed in the form of equality and inequality constraints, reduces the size of the spaces  $S_m$  and  $S_{\bar{m}}$  from which the pairs  $(m, \bar{m})$  are sampled, and possibly also the number of the potentially optimal portfolios. The second strategy, loosely based on Occam's razor, selects only pairs  $(m, \bar{m})$  which are the most parsimonious in terms of interactions between criteria and between projects. That is, even though we know that, most likely,

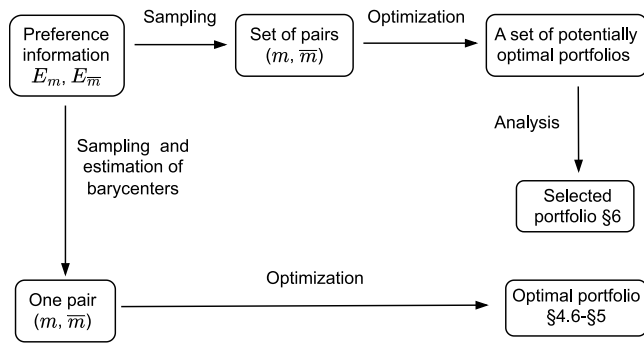


Fig. 7. A comparison between two approaches: optimization based on barycenters vs. Monte Carlo study and analysis of potentially optimal solutions.

an additive model is not compatible with the expressed preferences, we still want to find the most parsimonious (in terms of required interactions) non-additive model, namely, a non-additive model presenting the minimum necessary number of interactions. Having a lower number of parameters the model becomes more interpretable and, moreover, the space from which the vectors of parameters have to be sampled becomes smaller, diminishing therefore the variability of the parameters themselves. In the following, let us remind the procedure presented by Arcidiacono et al. [59] and apply it, on the one hand, to find a parsimonious 2-additive measure defined over the set of criteria  $G$  compatible with the preferences given by the DM and, on the other hand, to find a parsimonious 2-additive measure defined over the space of portfolios  $Y$  compatible with the preferences given by the DM:

- If we consider the set of constraints  $E_m$ , then to find a parsimonious 2-additive measure over  $G$  presenting the minimum number of necessary interactions between criteria, one has to solve the following MILP problem:

$$\left. \begin{aligned} \min \sum_{\{j_1, j_2\} \subseteq N} \gamma_{j_1 j_2}, \quad \text{subject to} \\ E_m \cup \{\varepsilon = \varepsilon_m\}, \\ -\gamma_{j_1 j_2} \leq m(\{g_{j_1}, g_{j_2}\}) \leq \gamma_{j_1 j_2}, \quad \forall \{j_1, j_2\} \subseteq N, \\ \gamma_{j_1 j_2} \in \{0, 1\}, \quad \forall \{j_1, j_2\} \subseteq N. \end{aligned} \right\} \quad (16)$$

Denoting by  $\{\gamma_{j_1 j_2}^* : \{j_1, j_2\} \subseteq N\}$  the solution of the previous MILP problem, the pairs of criteria  $\{g_{j_1}, g_{j_2}\}$  for which the corresponding  $\gamma_{j_1 j_2}^*$  is equal to 1 are those necessary to represent the DM's preferences. Vice versa, the pairs of criteria  $\{g_{j_1}, g_{j_2}\}$  for which the corresponding  $\gamma_{j_1 j_2}^*$  is null are not necessary to represent the DM's preferences and, therefore, can be removed;

- If we consider the set of constraints  $E_{\bar{m}}$ , then, to find a parsimonious 2-additive measure over  $Y$  presenting the minimum number of necessary interactions between projects, one has to solve the following MILP problem:

$$\left. \begin{aligned} \min \sum_{\{i_1, i_2\} \subseteq Q} \gamma_{i_1 i_2}, \quad \text{subject to} \\ E_{\bar{m}} \cup \{\varepsilon = \varepsilon_{\bar{m}}\}, \\ -\gamma_{i_1 i_2} \leq \bar{m}(\{y^{i_1}, y^{i_2}\}) \leq \gamma_{i_1 i_2}, \quad \forall \{i_1, i_2\} \subseteq Q, \\ \gamma_{i_1 i_2} \in \{0, 1\}, \quad \forall \{i_1, i_2\} \subseteq Q. \end{aligned} \right\} \quad (17)$$

Denoting by  $\{\gamma_{i_1 i_2}^* : \{i_1, i_2\} \subseteq Q\}$  the solution of the previous MILP problem, the pairs of projects  $\{y^{i_1}, y^{i_2}\}$  for which the corresponding  $\gamma_{i_1 i_2}^*$  is equal to 1 are those necessary to represent the DM's preferences. Vice versa, the pairs of projects  $\{y^{i_1}, y^{i_2}\}$  for which the corresponding  $\gamma_{i_1 i_2}^*$  is null are not necessary to represent the DM's preferences and, therefore, can be removed.

### 6.1. Example

In the example presented in Section 5, the total number of possible portfolios is  $2^{10}$ . Taking into account the technical constraints on the criteria at hand, by enumeration, we checked that 563 of them are feasible. Considering the DM's preferences over criteria and projects, the sampling procedure presented in the previous section shows that 98 out of the 563 feasible portfolios are potentially optimal with normalized frequencies varying between  $10^{-6}$  and 0.0691. This number is large, especially discounting the fact that a number of feasible portfolios are clearly non-optimal. Let us assume that the DM is able to provide the following additional preferences over projects:

- project  $y^5$  is preferred to project  $y^4$  ( $y^5 \succ_P y^4$ ),
- project  $y^3$  is preferred to project  $y^6$  ( $y^3 \succ_P y^6$ ).

Moreover, let us add the following technical requirements about the used 2-additive Choquet integrals:

- The 2-additive measure  $m$  defined over the set of criteria  $G$  has to be parsimonious (it has to present the minimum necessary number of interactions),
- The 2-additive measure  $\bar{m}$  defined over the set of projects  $Y$  has to be parsimonious (it has to present the minimum necessary number of interactions).

The procedure we used is presented in Algorithm 1 and its steps are articulated as follows:

---

#### Algorithm 1

---

- 1: **Require:** Preference information over criteria ( $E^{DM}$ ), Preference information over portfolios ( $E_Q^{DM}$ )
  - 2: Check for a compatible 2-additive measure over  $G$
  - 3: Check for a compatible parsimonious 2-additive measure over  $G$
  - 4: Sample 1,000 compatible parsimonious 2-additive measures  $m$  over  $G$
  - 5: **for** each  $m$  sampled in step 4: **do**
  - 6:   Check for a compatible 2-additive measure over  $Y$
  - 7:   Check for a compatible parsimonious 2-additive measure over  $Y$
  - 8: Sample 1,000 compatible parsimonious 2-additive measures  $\bar{m}$  over  $Y$
  - 9: **for** each  $\bar{m}$  sampled in step 8: **do**
  - 10:   Compute the optimal portfolio
  - 11: **end for**
  - 12: **end for**
  - 13: Give statistical information on the optimal portfolios obtained
- 

**Step 1:** We take into account the preference information on projects, criteria and portfolios specified in Section 5. This preference information is translated into inequality and equality constraints as explained in Sections 4.4 and 4.6, respectively. These constraints, together with technical constraints related to the 2-additive Choquet integral generate the sets of constraints  $E_m$  (see Eq. (9)) and  $E_{\bar{m}}$  (see Eq. (11)),

**Step 2:** Solve the LP problem (10) to check for the existence of at least one 2-additive measure over  $G$  such that the Choquet integral is able to represent the preferences of the DM over criteria and projects. In this case, we find that  $E_m$  is feasible and  $\varepsilon_m > 0$  meaning that there exists such a measure,

**Step 3:** Solve the MILP problem (16) to check for a 2-additive measure over  $G$  presenting the minimum number of interactions necessary to represent the preferences of the DM,

**Step 4:** Sample 1,000 parsimonious 2-additive measures over  $G$  from the space defined by the following set of constraints,

$$\left. \begin{aligned} E_m \cup \{0 \leq \varepsilon \leq \varepsilon_m\}, \\ m(\{g_{j_1}, g_{j_2}\}) = 0, \quad \forall \{j_1, j_2\} \subseteq N : \gamma_{j_1 j_2}^* = 0. \end{aligned} \right\}$$

Let us observe that the last constraint is used to impose that the interactions between the pairs of criteria not necessary to represent the preferences of the DM are null,

**Step 5:** As already observed in the previous section, each compatible parsimonious 2-additive measure over  $G$  sampled in the previous step constraints the space from which a parsimonious 2-additive measure over  $Y$  can be taken,

**Step 6:** Solve the LP problem (12) to check for the existence of at least one 2-additive measure over  $Y$  such that the Choquet integral is able to represent the preference of the DM over the portfolios. We find always that  $E_{\bar{m}}$  is feasible and  $\varepsilon_{\bar{m}} > 0$  meaning that there exists such measure,

**Step 7:** Solve the MILP problem (17) to check for a 2-additive measure over  $Y$  presenting the minimum number of interactions necessary to represent the preferences of the DM,

**Step 8:** Sample 1,000 parsimonious 2-additive measures over  $Y$  from the space defined by the following set of constraints:

$$\left. \begin{aligned} E_{\bar{m}} \cup \{0 \leq \varepsilon \leq \varepsilon_{\bar{m}}\}, \\ \bar{m}(\{y^{i_1}, y^{i_2}\}) = 0, \quad \forall \{i_1, i_2\} \subseteq Q : \gamma_{i_1 i_2}^* = 0. \end{aligned} \right\}$$

Let us observe that the last constraint imposes that the interactions between projects that are not necessary to represent the preferences of the DM are null,

**Steps 9–10:** For each pair of measures  $(m, \bar{m})$  sampled in steps 4: and 8:, respectively, solve the MILP problem (14) to find the optimal portfolio,

**Step 13:** Give information on how many different portfolios have been obtained through the considered procedure in the 1,000,000 simulations and the frequencies with which they have been obtained.

Applying the procedure described in Algorithm 1, we end up with the following three different portfolios

$$z_1^{opt} = (0, 0, 1, 0, 1, 1, 1, 1, 0)$$

$$z_2^{opt} = (0, 1, 0, 1, 0, 0, 1, 0, 1, 0)$$

$$z_3^{opt} = (0, 1, 0, 1, 1, 0, 0, 0, 0, 0)$$

that are obtained with normalized frequencies equal to 0.6134, 0.3860, and 0.0006, respectively. This means that the addition of the constraints translating the new pieces of preference information over projects and the technical constraints related to the parsimonious 2-additive Choquet integral permits to reduce the number of potentially optimal portfolios from 98 to 3.

Let us observe that if we had applied the “less robust” procedure described in Sections 4.4–4.6, that is, (i) compute an estimate of the barycenter of a sample of parsimonious 2-additive measures over  $G$  compatible with the preferences of the DM and, consequently, the projects’ values, (ii) compute an estimate of the barycenter of a sample of parsimonious 2-additive measures over  $Y$  compatible with the preferences of the DM, and (iii) compute the optimal portfolio solving the MILP problem (14) considering the two estimated barycenters, then we would have got the optimal portfolio

$$z_{bb}^{opt} = (0, 0, 1, 0, 1, 1, 1, 1, 1, 0),$$

that is also the most frequently obtained by using the “robust” procedure described in detail before.

## 7. Conclusions

We introduced a model for portfolio decision analysis and selection that is sufficiently general to consider, simultaneously, interactions between criteria and projects. By doing so we overcame the limitations of using an additive value function, that is, the intra- and between-project independence conditions presented by Morton et al. [15]. We proposed the use of the Choquet integral, and in the attempt to strike a reasonable tradeoff between generality of the model and computational tractability, we adopted 2-additive capacities that assign a value to each item and to each pair of items only. Both capacities (on the set of criteria and on the set of projects) are obtained using the aggregation-disaggregation approach in a 2-steps procedure that takes into account the preference information provided by the Decision Maker (DM). Either way, the approach based on 2-additive capacities remains more general than the additive value function: in the absence of interactions between criteria and projects, our model collapses into (and therefore is compatible) with the one employing additive value functions.

We proposed two alternative ways to operationalize this model: (i) the two capacities necessary to compute, on the one hand, the value of each project and, on the other hand, the value of each portfolio, are the estimations of the barycenter of the space defined by this preference information; (ii) the space of all compatible models is probed, and an optimal portfolio is computed for each sampled model. This second approach allows for an analysis of potentially optimal portfolios and frequencies with which they are obtained.

To show the applicability of our proposal, we considered an example from the literature in which the optimal portfolio of R&D projects has to be selected.

Let us observe that even if the proposed approach considers 2-additive capacities, this is not a strict requirement. Indeed, if the 2-additive Choquet integral is not able to represent the preferences of the DM, the approach can easily be extended to a  $k$ -additive capacity with  $k > 2$ .

A possible improvement could be adding flexibility in the constraints. Future works could, for instance, use fuzzy optimization to consider the boundaries of the regions of feasible portfolios fuzzy instead of crisp.

### CRedit authorship contribution statement

**Matteo Brunelli:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Salvatore Corrente:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

No data was used for the research described in the article.

### Acknowledgments

Salvatore Corrente acknowledges the support of the Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR), Italy - PRIN 2017, project “Multiple Criteria Decision Analysis and Multiple Criteria Decision Theory”, grant 2017CY2NCA.

## References

- [1] Faaland BH. The multiperiod Knapsack problem. *Oper Res* 1981;29(3):612–6.
- [2] Sarnataro M, Barbati M, Greco S. A portfolio approach for the selection and the timing of urban planning projects. *Socio-Econ Plan Sci* 2021;(75):100908.
- [3] Ross KW, Tsang DH. The stochastic Knapsack problem. *IEEE Trans Commun* 1989;37(7):740–7.
- [4] Barbati M, Corrente S, Greco S. A general space-time model for combinatorial optimization problems (and not only). *Omega* 2020;96:102067.
- [5] Golabi K, Kirkwood C, Sicherman A. Selecting a portfolio of solar energy projects using multiattribute preference theory. *Manage Sci* 1982;27(2):174–89.
- [6] Stummer C, Heidenberger K. Interactive R&D portfolio analysis with project interdependencies and time profiles of multiple objectives. *IEEE Trans Eng Manage* 1997;50(2):175–83.
- [7] Doerner K, Gutjahr W, Hartl R, Strauss C, Stummer C. Pareto ant colony optimization with ILP preprocessing in multiobjective project portfolio selection. *European J Oper Res* 2006;171:830–41.
- [8] Carazo A, Gómez T, Molina J, Hernández-Díaz A, Guerrero F, Caballero R. Solving a comprehensive model for multiobjective project portfolio selection. *Comput Oper Res* 2010;37:630–9.
- [9] de Almeida A, Duarte M. A multi-criteria decision model for selecting project portfolio with consideration being given to a new concept for synergies. *Pesquisa Operacional* 2011;31(2):301–18.
- [10] Gutjahr W, Pichler A. Project portfolio selection under uncertainty with outsourcing opportunities. *Flex Serv Manuf J* 2013;25:255–81.
- [11] Gomes L, Machado M, da Costa F, Rangel L. Criteria interactions in multiple criteria decision aiding: A Choquet formulation for the TODIM method. *Procedia Comput Sci* 2013;17:324–31.
- [12] Barbati M, Greco S, Lami IM. A new ordinal regression procedure for multiple criteria decision aiding: the case of the space time model for a sustainable ecovillage. 2023, arXiv preprint arXiv:2305.00940.
- [13] Liesiö J, Mild P, Salo A. Preference programming for robust portfolio modeling and project selection. *European J Oper Res* 2007;181(3):1488–505.
- [14] Salo A, Keisler J, Morton A. Portfolio decision analysis: improved methods for resource allocation, vol. 162, Springer; 2011.
- [15] Morton A, Keisler J, Salo A. Multicriteria portfolio decision analysis for project selection. In: Greco S, Ehrgott M, Figueira J, editors. Multiple criteria decision analysis: state of the art surveys. Springer New York, NY; 2016, p. 1273–302.
- [16] Zhang X, Hipel KW, Tan Y. Project portfolio selection and scheduling under a fuzzy environment. *Memetic Comput* 2019;11:391–406.
- [17] Marques AC, Frej EA, de Almeida AT. Multicriteria decision support for project portfolio selection with the FITradeoff method. *Omega* 2022;111:102661.
- [18] Demircan Keskin F. A two-stage fuzzy approach for Industry 4.0 project portfolio selection within criteria and project interdependencies context. *J Multi-Criteria Decis Anal* 2020;27(1–2):65–83.
- [19] Kandakoglu M, Walther G, Ben Amor S. The use of multi-criteria decision-making methods in project portfolio selection: A literature review and future research directions. *Ann Oper Res* 2023. <http://dx.doi.org/10.1007/s10479-023-05564-3>.
- [20] Kandakoglu M, Walther G, Ben Amor S. A robust multicriteria clustering methodology for portfolio decision analysis. *Comput Ind Eng* 2022;174:108803.
- [21] Lopes Y, A.T. d. Assessment of synergies for selecting a project portfolio in the petroleum industry based on a multi-attribute utility function. *J Pet Sci Eng* 2015;126:131–40.
- [22] Barbati M, Greco S, Kadziński M, Słowiński R. Optimization of multiple satisfaction levels in portfolio decision analysis. *Omega* 2018;78:192–204.
- [23] Liesiö J, Mild P, Salo A. Robust portfolio modeling with incomplete cost information and project interdependencies. *European J Oper Res* 2008;190:679–95.
- [24] Keeney RL, von Winterfeldt D. Practical value models. In: Edwards W, Miles Jr. RF, von Winterfeldt D, editors. *Advances in decision analysis: from foundations to applications*. Cambridge University Press; 2007, p. 232–52.
- [25] Liesiö J, Salo A, Keisler J, Morton A. Portfolio decision analysis: Recent developments and future prospects. *European J Oper Res* 2021;293:811–25.
- [26] Liesiö J. Measurable multiattribute value functions for portfolio decision analysis. *Decis Anal* 2014;11(1):1–20.
- [27] Liesiö J, Vilkkumaa E. Nonadditive multiattribute utility functions for portfolio decision analysis. *Oper Res* 2021;69(6):1886–908.
- [28] Durbach IN, Algorta S, Kantu DK, Katsikopoulos KV, Şimşek Ö. Fast and frugal heuristics for portfolio decisions with positive project interactions. *Decis Support Syst* 2020;138:113399.
- [29] Grabisch M, Marichal J-L, Mesiar R, Pap E. Aggregation functions: Means. *Inform Sci* 2011;181(1):1–22.
- [30] Mayag B, Grabisch M, Labreuche C. A representation of preferences by the Choquet integral with respect to a 2-additive capacity. *Theory and Decision* 2011;71(3):297–324.
- [31] Beliakov G. Knapsack problems with dependencies through non-additive measures and Choquet integral. *European J Oper Res* 2022;301(1):277–86.
- [32] Lahdelma R, Hokkanen J, Salminen P. SMAA - Stochastic multiobjective acceptability analysis. *European J Oper Res* 1998;106(1):137–43.
- [33] Keeney R, Raiffa H. *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: J. Wiley; 1976.
- [34] Wakker P. *Additive representations of preferences: A new foundation of decision analysis*. Springer; 1989.
- [35] Fishburn P. Methods of estimating additive utilities. *Manage Sci* 1967;13(7):435–53.
- [36] Jacquet-Lagrez E, Siskos Y. Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *European J Oper Res* 1982;10(2):151–64.
- [37] Jacquet-Lagrez E, Siskos Y. Preference disaggregation: 20 years of MCDA experience. *European J Oper Res* 2001;130(2):233–45.
- [38] Mild P, Liesiö J, Salo A. Selecting infrastructure maintenance projects with Robust Portfolio Modeling. *Decis Support Syst* 2015;77:21–30.
- [39] Gasparini G, Brunelli M, Chiriach MD. Multi-period portfolio decision analysis: A case study in the infrastructure management sector. *Oper Res Perspect* 2022;9:100213.
- [40] Grabisch M. *Set functions, games and capacities in decision making*, vol. 46, Springer; 2016.
- [41] Grabisch M. The application of fuzzy integrals in multicriteria decision making. *European J Oper Res* 1996;89(3):445–56.
- [42] Rota G. On the foundations of combinatorial theory I. Theory of Möbius functions. *Z Wahrscheinlichkeitstheor Verwandte Geb* 1964;2:340–68.
- [43] Grabisch M. *k*-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems* 1997;92(2):167–89.
- [44] Beliakov G, James S. Citation-based journal ranks: The use of fuzzy measures. *Fuzzy Sets and Systems* 2011;167(1):101–19.
- [45] Bilbao-Terol A, Bilbao-Terol C. The Choquet integral supported by a hedonic approach for modelling preferences in hotel selection. *Omega* 2024;122:102971.
- [46] Bottero M, Ferretti V, Figueira J, Greco S, Roy B. On the Choquet multiple criteria preference aggregation model: Theoretical and practical insights from a real-world application. *European J Oper Res* 2018;271:120–40.
- [47] Demirel T, Demirel N, Kahraman C. Multi-criteria warehouse location selection using Choquet integral. *Expert Syst Appl* 2010;37(5):3943–52.
- [48] Demirel N, Demirel T, Devci M, Vardar G. Location selection for underground natural gas storage using Choquet integral. *J Nat Gas Sci Eng* 2017;45:368–79.
- [49] Silva A, Ferreira F, Carayannis E, Ferreira J. Measuring SMEs' propensity for open innovation using cognitive mapping and MCDA. *IEEE Trans Eng Manage* 2019;68(2):396–407.
- [50] Büyükoçkan G, Feyzioğlu O, Ersoy M. Evaluation of 4PL operating models: A decision making approach based on 2-additive Choquet integral. *Int J Prod Econ* 2009;121(1):112–20.
- [51] Shapley L. A value for *n*-person games. In: Kuhn HW, Tucker AW, editors. *Contributions to the theory of games II*. Princeton: Princeton University Press; 1953, p. 307–17.
- [52] Murofushi S. Techniques for reading fuzzy measures (III): Interaction index. In: 9th fuzzy systems symposium, Sapporo, Japan, 1993, p. 693–6.
- [53] Siskos E, Burgherr P. Multicriteria decision support for the evaluation of electricity supply resilience: Exploration of interacting criteria. *European J Oper Res* 2022;298:611–26.
- [54] Mousseau V, Figueira J, Dias L, Gomes da Silva C, Climaco J. Resolving inconsistencies among constraints on the parameters of an MCDA model. *European J Oper Res* 2003;147(1):72–93.
- [55] Smith R. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper Res* 1984;32:1296–308.
- [56] Tervonen T, Van Valkenhoef G, Bastürk N, Postmus D. Hit-and-run enables efficient weight generation for simulation-based multiple criteria decision analysis. *European J Oper Res* 2013;224:552–9.
- [57] Arcidiacono S, Corrente S, Greco S. Scoring from pairwise winning indices. *Comput Oper Res* 2023;157:106268.
- [58] Pelissari R, Oliveira M, Ben Amor S, Kandakoglu A, Helleno A. SMAA methods and their applications: A literature review and future research directions. *Ann Oper Res* 2020;293:433–93.
- [59] Arcidiacono S, Corrente S, Greco S. Robust stochastic sorting with interacting criteria hierarchically structured. *European J Oper Res* 2021;292(2):735–54.