Necessary and possible interaction between criteria in a 2-additive Choquet integral model *

Brice Mayag †

Denis Bouyssou ‡

14 October 2019

Abstract

This paper deals with the interpretation of the 2-additive Choquet integral model in the context of Multiple Criteria Decision Making. When the set of alternatives is discrete, using classical interaction indices proposed in the literature may lead to interpretations that are not robust. Indeed, the sign of these indices may depend upon the arbitrary choice of a numerical representation within the set of all possible numerical representations. We tackle this problem in two ways. First, in the context of binary alternatives, we characterize the preference relations for which the problem does not occur. Outside the framework of binary alternatives, we propose a simple linear programming model allowing one to test for robust conclusions concerning the sign of interaction indices. We illustrate our results on a real world example in the domain of health.

Keywords: Multiple criteria decision analysis, Choquet integral, 2-additive Capacity, Interaction

1 Introduction

The dominant model in Multiple Criteria Decision Making (MCDM) is the additive value function model. It has quite solid theoretical foundations [28]. Moreover, many techniques have been proposed in order to elicit its parameters [26, 50]. This model makes central use of an independence hypothesis stating that tradeoffs between criteria

*We wish to thank Michel Grabisch, Christophe Labreuche, Patrice Perny and Marc Pirlot as well as three referees for their comments on an earlier draft of this text. We remain fully responsible for all errors.

†Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, UMR 7243, 75016 Paris, France, brice.mayag@dauphine.fr

‡Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, UMR 7243, 75016 Paris, France, denis.bouyssou@dauphine.fr

can be elicited independently of common evaluations on other criteria. In some contexts, this independence hypothesis may seem restrictive. Hence, alternative preference models were developed. Among the various models allowing weakening the independence hypothesis, the Choquet integral model stands as a central reference point. Its use in MCDM was popularized through the work of Michel Grabisch [13, 14], following its wide diffusion in the field of decision making under uncertainty [42, 51, 52]. The theoretical status of the Choquet integral model is less firmly established in the field of MCDM than it is in the field of decision making under uncertainty. Indeed, the model assumes that criteria are somehow "commensurate", while there is no consensus in the literature on the precise meaning of this hypothesis and the way to test it in practice. Nevertheless, this model is now considered as a central tool in MCDM when one wants to escape the independence hypothesis [16, 17, 18].

The Choquet integral model is quite flexible. The elicitation of its parameters (i.e., the capacity assigning a weight to all subsets of criteria) is therefore difficult without additional hypotheses. In the literature, the special case of 2-additive capacities has received much attention [15]. This case is often considered as a useful compromise between a fully additive model implying independence and a fully general Choquet integral model (i.e., using a capacity that is not restricted to be 2-additive) raising difficult elicitation issues [22]. This model is often used in applications (evaluation of discomfort [20], performance measurement in supply chains [3, 7], complex system design [39]).

This paper wishes to be a contribution to the study and interpretation of the 2-additive Choquet integral model (i.e., the Choquet integral model using a 2-additive capacity) in MCDM. We will not be concerned here with the commensurability hypothesis and we shall suppose that criteria has been built so as to be commensurate. Given this hypothesis, the Choquet integral model is often interpreted as an extension of the weighted sum model in which "weights" can also be allocated to sets of criteria containing more than one element. The interpretation of such weights requires the definition of adequate interaction indices among criteria [19]. It is widely believed that these indices can be used to interpret the type of interaction at work with a given capacity. A null interaction means "independence", a positive interaction implies "complementarity", while a negative interaction implies "substitutability" [27, 31]. The main purpose of this paper is to question the soundness of this received interpretation.

It is clear that using a 1-additive capacity in the Choquet integral model (i.e., a simple weighted sum) implies independence. It is known that the reverse implication is also true when the model is applied to a "continuous structure" (see Theorem 4.1 in [38] and the use of condition DC). When the structure is discrete, one may clearly not expect a representation in the Choquet integral model to be unique. When several (2-additive) capacities can represent the same preference relation, it would then be heroic to suppose that the sign of the interaction indices remains unchanged in all possible representations. This paper starts by showing that this is indeed the case. We then proceed by defining cases in which preference information allows for an unambiguous interpretation of interaction indices. This will lead us to define "necessary" and "possible" interactions. These notions are reminiscent of this idea of necessary and possible preference relations used in the context of robust ordinal regression (see UTA$^{\text{GMS}}$ [24], GRIP [11] or their extension to possibly non-additive models [1]). We use a similar idea applied to the sign of interaction indices.

We study necessary interactions in two contexts. First, within the binary alternatives framework used in [35], we characterize preference relations allowing for an unambiguous interpretation of interaction indices. Our conclusion is that, for a large variety of preference relations on binary alternatives, the use of negative interaction is not necessary, i.e., negative interaction does not occur in all feasible numerical representation of the preference information. The framework of binary alternatives may however seem restrictive because it contains too few alternatives. Hence, we also present a simple linear programming model allowing us to test whether the interpretation of interaction indices is or not ambiguous. This model is not limited to the case of preference information on binary alternatives.

The paper is organized as follows. Section 2 presents our setting and recalls some basic facts about the Choquet integral model in MCDM. Section 3 analyzes a classic example showing the difficulty to interpret interaction indices in the discrete case. Section 4 defines necessary (and possible) interactions in order to circumvent this ambiguity. Section 5 characterizes preference relation giving rise to necessary interactions, in the context of binary alternatives. Section 6 proposes a linear programming model allowing one to test the existence of necessary interactions outside the framework of binary alternatives. Section 7 illustrates our results using a real-world example coming from the medical field. A final section concludes.

2 Notation and definitions

2.1 The framework

We consider a set of alternatives X evaluated on a finite set of n criteria $N = \{1, \dots, n\}$. The set of all alternatives X is assumed to be a Cartesian product $X = X_1 \times \cdots \times X_n$.

The various criteria are recoded numerically using, for each $i \in N$, a function u_i from X_i into $\mathbb{R}$. It is supposed the use of these functions allow to assume that the various recoded criteria are “commensurate” and, hence, the application of the Choquet integral model is meaningful [21]. We will sometimes write $u(x)$ as a shorthand for $(u_1(x_1), \dots, u_n(x_n))$.

2.2 The Choquet integral

The Choquet integral [18] is an aggregation function known in MCDM as a tool generalizing the arithmetic mean. It is based on the notion of *capacity* μ defined as a function from the powerset 2^N into $[0, 1]$ such that,$\forall A, B \in 2^N$:

$$\begin{aligned}
&\mu(\varnothing) = 0,\\
&\mu(N) = 1,\\
&[A \subseteq B \Rightarrow \mu(A) \leq \mu(B)] \quad (\textit{monotonicity}).
\end{aligned}$$

The *Möbius transform* $m^\mu : 2^N \to \mathbb{R}$ of a capacity μ is defined, for all $T \in 2^N$, by:

$$m^\mu(T) := \sum_{K \subseteq T} (-1)^{|T \setminus K|} \mu(K). \tag{1}$$

Conversely, we have, for all $T \in 2^N$,

$$\mu(T) = \sum_{K \subseteq T} m^\mu(K). \tag{2}$$

A 2-additive capacity [35] is a capacity μ such that its *Möbius transform* satisfies the following two conditions:

- for all subset T of N such that $|T| > 2$, $m^\mu(T) = 0$,
- there exists a subset B of N such that $|B| = 2$ and $m^\mu(B) \neq 0$.

We simplify our notation for a 2-additive capacity μ by using the following shorthand: $\mu_i := \mu(\{i\})$, $\mu_{ij} := \mu(\{i,j\})$ for all $i, j \in N$, $i \neq j$. Whenever we use i and j together, it is always understood that they are distinct.

For an alternative $x := (x_1, \ldots, x_n) \in X$, the expression of the Choquet integral w.r.t. a capacity μ is given by:

$$\begin{aligned} &C_\mu(u(x)) := C_\mu(u_1(x_1), \ldots, u_n(x_n)) := \\ &\quad \sum_{i=1}^{n} (u_{\tau(i)}(x_{\tau(i)}) - u_{\tau(i-1)}(x_{\tau(i-1)}))\mu(\{\tau(i), \ldots, \tau(n)\}), \end{aligned}$$

where τ is a permutation on N such that $u_{\tau(1)}(x_{\tau(1)}) \leq u_{\tau(2)}(x_{\tau(2)}) \leq \cdots \leq u_{\tau(n-1)}(x_{\tau(n-1)}) \leq u_{\tau(n)}(x_{\tau(n)})$, and $u_{\tau(0)}(x_{\tau(0)}) := 0$.

In the case of the Choquet integral w.r.t. a 2-additive capacity, called for short the 2-additive Choquet integral, the above equation is equivalent to the following two expressions: [18, 33]:

$$\begin{aligned} C_\mu(u_1(x_1), \ldots, u_n(x_n)) &:= \sum_{i \in N} m^\mu(\{i\}) u_i(x_i), \\ &+ \sum_{i,j \in N} m^\mu(\{i,j\}) \min(u_i(x_i), u_j(x_j)) && (3) \\ &:= \sum_{i}^{n} V_i^\mu u_i(x_i) - \frac{1}{2} \sum_{\{i,j\} \subseteq N} I_{ij}^\mu |u_i(x_i) - u_j(x_j)|, && (4) \end{aligned}$$

where

$$V_i^\mu = \sum_{K \subseteq N \setminus \{i\}} \frac{(n - |K| - 1)!|K|!}{n!} [\mu(K \cup i) - \mu(K)] = \mu_i + \frac{1}{2} \sum_{j \in N \setminus \{i\}} [\mu_{ij} - \mu_i - \mu_j],$$

is an interpretation, according to the Shapley value of μ [43], of the importance of criterion i, and $I^{\mu}_{ij} = \mu_{ij} - \mu_i - \mu_j$ is the interaction index between the two criteria i and j as defined in [15, 37]. We have $I^{\mu}_{ij} = m^{\mu}(\{i, j\})$, when μ is a 2-additive capacity.

Only interactions between two criteria can exist when using a 2-additive Choquet integral, i.e., interaction among more than three criteria are ignored by this model. Equation (4) is equivalent to an arithmetic mean when there is no interaction between criteria. Therefore this operator appears as a compromise between the arithmetic mean and the Choquet integral. We can notice that, given a 2-additive capacity, the interaction between criteria i and j, measured by I^{μ}_{ij}, has a very simple expression that is much simpler than the ones that deal with the general Choquet integral [17, 18, 33]. It is usually interpreted as follows:

- there is *complementarity* among the criteria i and j when $I^{\mu}_{ij} > 0$, i.e., these criteria have some value by themselves, but put together they become even more important for the decision maker (DM),
- there is *substitutability* or *redundancy* among criteria i and j when $I^{\mu}_{ij} < 0$, i.e., these criteria have some value by themselves, but put together they become less important for the DM,
- it seems natural to consider that criteria i and j do not interact, i.e., that they have independent roles in the decision problem when $I^{\mu}_{ij} = 0$.

The interpretation given to the interaction index I^{μ}_{ij} between two criteria is clearly dependent upon the capacity μ. But when *several* capacities can represent the same preference information, the situation becomes more complex, as shown in the next section.

3 A motivating example

We consider a classic exemple in the literature [18]. Four students of a faculty are evaluated on three subjects Mathematics (M), Statistics (S) and Language skills (L). All marks are taken from the same scale, from 0 to 20. The evaluations of these students are given by the table below:

	1: Mathematics (M)	2: Statistics (S)	3: Language (L)
a	16	13	7
b	16	11	9
c	6	13	7
d	6	11	9

To select the best students, the Dean of the faculty expresses his/her preferences where the notation $x\ P\ y$ means x is strictly preferred to y. For a student good in Mathematics, Language is privileged compared to Statistics, so that

$$b\ P\ a, \tag{5}$$

for a student bad in Mathematics, Statistics is privileged, compared to Language, so that

$$c\ P\ d. \tag{6}$$

Let us denote by w_M, w_S and w_L the weights associated to Mathematics, Statistics and Language. It is not possible to model the two preferences $b\ P\ a$ and $c\ P\ d$ by an arithmetic mean model. Indeed we have:

$$\begin{aligned} b\ P\ a \Rightarrow\ & u_M(16)w_M + u_S(13)w_S + u_L(7)w_L < \\ & u_M(16)w_M + u_S(11)w_S + u_L(9)w_L, \\ c\ P\ d \Rightarrow\ & u_M(6)w_M + u_S(11)w_S + u_L(9)w_L < \\ & u_M(6)w_M + u_S(13)w_S + u_L(7)w_L. \end{aligned} \tag{7}$$

leading to the following contradiction:

$$\begin{aligned} & u_S(13)w_S + u_L(7)w_L < u_S(11)w_S + u_L(9)w_L \text{ and} \\ & u_S(11)w_S + u_L(9)w_L < u_S(13)w_S + u_L(7). \end{aligned}$$

Let us assume that the scale of evaluation $[0, 20]$ corresponds to the utility function associated to each subject, i.e., $u_M(16) = 16$, $u_M(6) = 6$, $u_S(13) = 13$, $u_S(11) = 11$, $u_L(7) = 7$ and $u_L(9) = 9$. Using these utility functions the preferences $b\ P\ a$ and $c\ P\ d$, are now representable by a 2-additive Choquet integral w.r.t. any capacity given in Table 1 below. Among all the capacities compatible with these preferences, we chose nine of them (called Parameter, Par. for short in Table 1) in order to illustrate the fact that the sign of an interaction index is strongly dependent upon to the chosen capacity.

In this example, it seems clear that it is not easy to interpret the interaction between two criteria. For instance, depending on the capacity, the interaction between Mathematics and Statistics, I^{μ}_{MS}, could be positive (Par. 3), null (Par. 1) or negative (Par. 4). In other words, from the preferences given by the Dean, could we conclude that the subjects Mathematics and Statistics are complementary, redundant or independent? Answering this question is not obvious. This conclusion is still valid concerning the interaction I^{μ}_{ML} between Mathematics and Language (see Par. 2, Par. 8 and Par. 9), and interaction I^{μ}_{SL} between Statistics and Language (see Par. 5, Par. 6 and Par. 7). In fact, the only information provided by Equation (7) is that: "the three criteria (subjects), taken together, are not without interaction", i.e., the three interaction indices cannot be simultaneously null. Roughly speaking, for each 2-additive capacity μ allowing to have $b\ P\ a$ and $c\ P\ d$ via the use of a Choquet integral, there are $i, j \in \{M, S, L\}$ such that $I^{\mu}_{ij} \neq 0$.

We have used a classic example to argue that the usual interpretation of interaction indices is not always convincing. This troubling observation may arise because, in the example, the DM gave only very poor information consisting in two strict preferences. We may expect that adding more preferences could help to have a consistent interpretation of an interaction index. Indeed, when the set of alternative has a continuous structure and preference is compatible with this rich structure, the recent work of M. Timonin

	Par. 1	Par. 2	Par.3	Par. 4	Par. 5	Par. 6	Par. 7	Par.8	Par. 9
$C_\mu(a)$	8.5	13.75	9.1	13.765	13.75	13.75	11.47	12.535	10.45
$C_\mu(b)$	9.5	14.25	9.7	13.995	14.25	14.25	11.93	12.785	10.75
$C_\mu(c)$	7.75	9.75	7.75	11.325	11.25	9.75	9.45	9.515	7.85
$C_\mu(d)$	7.25	9.25	7.25	10.295	9.75	9.25	8.91	9.265	7.55
μ_M	0	0.75	0	0.685	0.75	0.75	0.36	0.485	0.15
μ_S	0.25	0.5	0.25	0.73	0.75	0.5	0.465	0.455	0.25
μ_L	0	0.25	0	0.315	0	0	0.205	0.32	0
μ_{MS}	0.25	0.75	0.35	0.785	0.75	0.75	0.565	0.68	0.5
μ_{ML}	0.75	1	0.65	1	0.1	0.75	0.805	0.795	0.55
μ_{SL}	0.25	0.75	0.25	0.945	0.75	0.75	0.66	0.785	0.35
V_M^μ	0.375	0.5	0.375	0.37	0.5	0.5	0.35	0.35	0.4
V_S^μ	0.25	0.25	0.3	0.365	0.375	0.375	0.33	0.33	0.35
V_L^μ	0.375	0.25	0.325	0.265	0.125	0.125	0.32	0.32	0.25
I_{MS}^μ	$\boxed{0}$	-0.5	$\boxed{0.1}$	$\boxed{-0.63}$	-0.75	-0.5	-0.26	-0.26	0.1
I_{ML}^μ	0.75	$\boxed{0}$	0.65	0	0.25	0	0.24	$\boxed{-0.01}$	$\boxed{0.4}$
I_{SL}^μ	0	0	0	-0.1	$\boxed{0}$	$\boxed{0.25}$	$\boxed{-0.01}$	0.01	0.1

Table 1: A set of nine 2-additive capacities compatible with the preferences $b\ P\ a$ and $c\ P\ d$.

[46, 47, 48] shows that, in such a situation, the representing capacity becomes *unique*, so that there is no interpretation problem of the interaction index (one may also see [38]). But in common elicitation tasks, the DM only provides information on a discrete set of alternatives. We will show below that, when this discrete set of alternatives takes the form of binary alternatives, the problem exemplified in this section remains.

4 Necessary and possible interaction

In all what follows, we will suppose that: the DM has compared [1] a number of alternatives in terms of strict preference (P) or indifference (I), that the criteria have been made commensurable via the use of adequate utility functions $u_1, u_2, \ldots, u_n$, and that this preference information can be represented in the 2-additive Choquet integral model, i.e., that there is a 2-additive capacity μ such that:

$$\begin{aligned} (x,y) \in P &\Rightarrow C_\mu(u(x)) > C_\mu(u(y)), \\ (x,y) \in I &\Rightarrow C_\mu(u(x)) = C_\mu(u(y)). \end{aligned} \tag{8}$$

[1] We restrict ourselves in this paper to ordinal preference information, i.e., judgments of preference and indifference provided by the DM. In particular, we do not take into account information concerning the comparison of "preference differences" or the "intensity of preference". Although enlarging the type of preference information that is taken into account would clearly alleviate some of the difficulties encountered below, it is not completely clear how we could obtain such information in a clear and reliable way. This is further commented in Section 8.

The set of all 2-additive capacities that can be used to represent the preference information at hand will be denoted $\mathcal{C}_{2\text{-}add}(P, I)$. When there is no ambiguity on the underlying preference information, we will simply write $\mathcal{C}_{2\text{-}add}$.

The following definition of necessary and possible interactions will be central in the rest of this text.

Definition 1
Let $i, j \in N$ be two distinct criteria, We say that:

1. there exists a *possible* positive (resp. null, negative) interaction between i and j if there exists a capacity $\mu \in \mathcal{C}_{2\text{-}add}$ such that $I_{ij}^{\mu} > 0$ (resp. $I_{ij}^{\mu} = 0$, $I_{ij}^{\mu} < 0$),
2. there exists a *necessary* positive (resp. null, negative) interaction between i and j if $I_{ij}^{\mu} > 0$ (resp. $I_{ij}^{\mu} = 0$, $I_{ij}^{\mu} < 0$) for all capacity $\mu \in \mathcal{C}_{2\text{-}add}$. ⌟

This notion of necessary and possible interaction, defined here for interactions indices I_{ij}^{μ}, is related to but different from the necessary and possible preference relations defined in [1, 11, 24]. We are here concerned with the sign of interaction indices.

An obvious consequence of the above definition is spelled out in the following Remark.

Remark 1
Let $i, j \in N$ be two distinct criteria. If there exists a necessary positive (resp. null, negative) interaction between i and j, then there exists a possible positive (resp. null, negative) interaction between i and j.

If there is no necessary positive (resp. null, negative) interaction between i and j, then there exists a possible negative or null (resp. positive or negative, positive or null) interaction between i and j. •

Given preference information provided by the DM, the interpretation of interaction indices is only meaningful when interactions are necessary. An interaction that is possible but not necessary is meaningless since its interpretation is dependent upon the arbitrary choice a capacity μ in the set $\mathcal{C}_{2\text{-}add}$. The conditions under which preference information may lead to necessary interactions are investigated in the next two sections.

We conclude this section with a simple observation. In the discrete case, when there is no indifference, there are "holes" between all values of $C_{\mu}(u(x))$. If we slightly modify the capacity μ, so as to keep all values $C_{\mu}(u(x))$ within these "holes", we find that null interactions are never necessary. We formalize this as a simple Proposition. Its proof is elementary: it simply exploits the fact that when the structure has holes, it is possible to slightly modify the representing model while remaining in the holes.

Proposition 1
Suppose that we have a preference structure (P, I) on a set $Y \subseteq X$ that can be represented using the 2-additive Choquet integral model. If the relation I is empty then there is no necessary null interactions.

Proof
We only give the proof in the special case in which the DM has provided a linear order on a subset $Y \subseteq X$. It is easy to modify the proof to cover the other cases.

Suppose that we have a 2-additive Choquet integral model representing the preference relation P that linearly orders the set $Y = \{x^1, x^2, \dots, x^p\}$. We suppose w.l.o.g. that $x^p \; P \; x^{p-1} \; P \cdots P \; x^1$.

Let us suppose that this information can be represented using a 2-additive Choquet integral model using a capacity μ for which $I^\mu_{ij} = 0$, so that $m^\mu_{ij} = 0$. Let us first show that this possible null interaction is not necessary.

Let us denote by α^μ_* the smallest difference between the value of the Choquet integral C_μ for two consecutive elements in Y.

Let us build a capacity τ that has the same Möbius transform as μ but has $m^\tau_{ij} = \epsilon > 0$. Notice that this modification of m^μ_{ij} will require a normalization of the values of m^τ in order to have $\sum_i m^\tau_i + \sum_{i,j} m^\tau_{ij} = 1$, but this normalization plays no role in the ranking of the elements of Y.

Let $k_{ij}(x^h) = \min(x^h_i, x^h_j)$. Let $k^*_{ij} = \max_h k_{ij}(x^h)$. Obviously, if we choose ϵ in such a way that $k^*_{ij} \times \epsilon < \alpha_*$, the ranking of all alternatives in Y will remain unchanged with the new capacity in which $m^\tau_{ij} > 0$. The value of the 2-additive Choquet integral $C_\mu(u(x^h))$ is now increased to $C_\tau(u(x^h))$ but this increase is sufficiently small so as to remain smaller than the difference $C_\mu(u(x^{h+1})) - C_\mu(u(x^h))$. Hence there is no null interaction between i and j.

Note that, if the modified capacity shows a possible null interaction between a different pair of criteria, the above process can be repeated. This will lead to exhibit a capacity in which there are only positive interactions. Hence, null interactions are never necessary when $I = \varnothing$. □

Remark 2
The condition that I is empty is likely to be met in most applications: indifference is indeed much less likely between alternative than strict preference, unless alternatives have been specially designed to be indifferent. Going through the above proof shows that the result can be generalized to some cases in which I is not empty. We leave the details to the interested reader. •

5 Necessary interaction with binary alternatives

5.1 Framework

We suppose that the DM has been able to identify on each criterion $i \in N$ two reference levels $\mathbf{1}_i$ and $\mathbf{0}_i$.

1. The level $\mathbf{1}_i \in X_i$ is considered as good and completely satisfying if it can be attained on criterion i, even though more attractive elements can exist. This reference level is reminiscent of the *satisficing level* in the theory of bounded rationality of [44].

2. The level $\mathbf{0}_i$ in X_i is considered to be a neutral level. The level is an element which is thought by the DM to be neither good nor bad, neither attractive nor repulsive relatively to his/her concerns with respect to the criterion i. The existence of such

a neutral level has roots in Psychology [45], and is used in bipolar models like Cumulative Prospect Theory [49].

With the definition of these two reference levels, we suppose that the commensurateness problem between criteria has been solved. We set $u_i(\mathbf{1}_i) = 1$ and $u_i(\mathbf{0}_i) = 0$. Therefore the previous reference levels can be used in order to define the same scale on each criterion [21, 29]. In defining $\mathbf{1}_i$ and $\mathbf{0}_i$, we have followed the interpretation favored in [16, 35]. It is not the only possible one and in all what follows, we only use the fact that the level $\mathbf{1}_i$ is above the level $\mathbf{0}_i$.

For a subset $A \subsetneq N$ and alternatives $x, y \in X$, we denote by $z = (x_A, y_{N-A})$ the element of X such that $z_i = x_i$ if $i \in A$ and $z_i = y_i$ otherwise.

We call *binary alternatives*, the elements of the set

$$\mathcal{B} = \{\mathbf{0}_N, (\mathbf{1}_i, \mathbf{0}_{N-i}), (\mathbf{1}_{ij}, \mathbf{0}_{N-ij}), i, j \in N, i \neq j\} \subseteq X,$$

where

- $\mathbf{0}_N = (\mathbf{1}_\varnothing, \mathbf{0}_N) =: a_0$ is an alternative considered neutral on all criteria,
- $(\mathbf{1}_i, \mathbf{0}_{N-i}) =: a_i$ is an alternative considered satisfactory on criterion i and neutral on the other criteria,
- $(\mathbf{1}_{ij}, \mathbf{0}_{N-ij}) =: a_{ij}$ is an alternative considered satisfactory on criteria i and j and neutral on the other criteria.

The map ϕ will indicate the bijection between $\mathcal{B}$ and $\mathcal{P}^2(N) = \{S \subseteq N : |S| \leq 2\}$ defined by, for all $S \in \mathcal{P}^2(N)$, $\phi((\mathbf{1}_S, \mathbf{0}_{N-S})) := S$. The number of binary alternatives is ${}^{n(n+1)}/_2 + 1$.

For any 2-additive capacity μ, we have:

$$\begin{aligned} C_\mu(u(a_0)) &= 0, \\ C_\mu(u(a_i)) &= \mu_i, \\ C_\mu(u(a_{ij})) &= \mu_{ij}. \end{aligned} \tag{9}$$

In order to compute all the parameters of the 2-additive Choquet integral, Mayag et al. [35] suggest to ask the DM for preference information $\{P, I\}$ on the set of binary alternatives, called *ordinal information on* $\mathcal{B}$, and given by:

$$\begin{aligned} P &= \{(x, y) \in \mathcal{B} \times \mathcal{B} : \text{DM strictly prefers } x \text{ to } y\}, \\ I &= \{(x, y) \in \mathcal{B} \times \mathcal{B} : \text{DM is indifferent between } x \text{ and } y\}. \end{aligned} \tag{10}$$

We add to this ordinal information a relation M modeling the relation of monotonicity between binary alternatives, and allowing us to ensure the satisfaction of the monotonicity conditions $\mu(\{i\}) \geq 0$ and $\mu(\{i, j\}) \geq \mu(\{i\})$ for a capacity μ. For $(x, y) \in \{(a_i, a_0), i \in N\} \cup \{(a_{ij}, a_i), i, j \in N, i \neq j\}$, we define

$$x \; M \; y \text{ if } \mathit{Not}[x \; (P \cup I) \; y]. \tag{11}$$

Finally, remember from Section 4 that a preference information $\{P, I\}$ is compatible with the 2-additive Choquet integral model if there is a 2-additive capacity μ such that (8) holds.

The characterization of the representation of $\{P, I\}$ by a 2-additive Choquet integral is given in [35]. The result is based on the property MOPI defined below and on the existence of cycles [2] in the relation $(P \cup I \cup M)$.

Definition 2 ([35, Def. 3.1, p. 305])
Let $i, j, k \in N$. We call *Monotonicity of Preferential Information* in $\{i, j, k\}$ w.r.t. i the following property (denoted by **($\{i,j,k\}$,i)-MOPI**):

$$\begin{array}{c} \left\{\begin{array}{l} a_{ij} \sim a_j \\ a_{ik} \sim a_i \end{array}\right. \Rightarrow Not[a_k \; TC_P \; a_0], \\ \text{and} \\ \left\{\begin{array}{l} a_{ij} \sim a_i \\ a_{ik} \sim a_k \end{array}\right. \Rightarrow Not[a_j \; TC_P \; a_0], \\ \text{and} \\ \left\{\begin{array}{l} a_{ij} \sim a_j \\ a_{ik} \sim a_k \end{array}\right. \Rightarrow Not[a_i \; TC_P \; a_0], \end{array} \tag{12}$$

where

1. $x \sim y$ if the elements x and y belong to a cycle of $(I \cup M)$,
2. $x \; TC_P \; y$ if there exists a strict path of $(P \cup I \cup M)$ from x to y.

We say that, the set $\{i, j, k\}$ satisfies the property of *Monotonicity of Preferential Information* (MOPI) if $\forall \ell \in \{i, j, k\}$, Condition $(\{i, j, k\}, \ell)$-MOPI is satisfied. ⌟

The MOPI condition can be interpreted as follows [35]. Suppose that a_{ij} and a_j are indifferent. This would suggest that i is not important for the DM, but this is *relatively to* j, or put differently, i is much less important than j. Suppose in addition that a_{ik} is indifferent to a_i. Again, this suggests that k is much less important than i. Since i is much less important than j, the conclusion is that k is quite unimportant, hence a_k is indifferent to a_0. This explains the first case in the MOPI condition. The second case (indifference between a_{ik} and a_k, and between a_{ij} and a_i) works exactly the same way. The third case says that a_{ij} and a_j are indifferent (i is much less important than j) as well as a_{ik} and a_k (i is much less important than k). Since i is much less important than both j and k, the conclusion is that i is quite unimportant, so that a_i is indifferent with a_0.

Theorem 1 ([35, Th. 1, p. 305])
An ordinal information $\{P, I\}$ is representable by a 2-additive Choquet integral on $\mathcal{B}$ if and only if the following two conditions are satisfied:

[2] Let T be a binary relation on $\mathcal{B}$ and $x, y \in \mathcal{B}$. We say that $\{x_1, x_2, \ldots, x_p\} \in \mathcal{B}$ is a *path* of T from x to y if $x = x_1 \; T \; x_2 \; T \cdots T \; x_{p-1} \; T \; x_p = y$. A path from x to x is a *cycle*. When $T = (P \cup I \cup M)$, we say that a path of T is *strict* if, for some $i \in \{1, \ldots, p-1\}$, we have $x_i \; P \; x_{i+1}$.

1. $(P \cup I \cup M)$ contains no strict cycle,

2. Any subset K of N such that $|K| = 3$ satisfies the MOPI property.

5.2 Results

Thereafter, we assume that $\{P, I\}$ is an ordinal information on $\mathcal{B}$ that can be represented by a 2-additive Choquet integral. Remember that we denote by $\mathcal{C}_{2\text{-}add}$, the set of all 2-additive capacities compatible with the preference information $\{P, I\}$.

Our first result, Theorem 2, says that if $I = \varnothing$, it is always possible to find a capacity μ belonging to $\mathcal{C}_{2\text{-}add}$ and such that all the interaction indices between two criteria are strictly positive. This fact implies that negative and null interactions can arise but are never necessary in this case. Theorem 2 improves over [35, Cor. 1, p. 306], that only dealt with nonnegative interactions. In following result, as well as others presented later in this section, positive and negative interactions are not treated in a symmetric way. This is indeed puzzling and is commented upon in Section 5.4.

Theorem 2
Let $\{P, I\}$ be an ordinal information on $\mathcal{B}$ such that $I = \varnothing$. Suppose that this information can be represented in the 2-additive Choquet integral model. In $\mathcal{C}_{2\text{-}add}$, there is a capacity μ such that, for all $i, j \in N$, $I^{\mu}_{ij} > 0$ and, hence, all pairs of criteria possibly interact positively.

PROOF
The proof consists in building a partition $\{\mathcal{B}_0, \mathcal{B}_1, \dots, \mathcal{B}_m\}$ of $\mathcal{B}$. Using this partition, a capacity μ belonging to $\mathcal{C}_{2\text{-}add}$ is built. It is such that $I^{\mu}_{ij} > 0$ for all $i, j \in N$.

The construction of the partition $\{\mathcal{B}_0, \mathcal{B}_1, \dots, \mathcal{B}_m\}$ is detailed in detailed[3] in [35, Sect. 5.2, p. 315].

Therefore, given $i, j \in N$, there exist $p, q, s \in \{1, \dots, m\}$ such that $a_{ij} \in \mathcal{B}_p$, $a_i \in \mathcal{B}_q$, $a_j \in \mathcal{B}_s$ with $p > q > 0$ and $p > s > 0$ (as illustrated in Figure 1).

Let us define the mapping $f : \mathcal{B} \to \mathbb{R}$ and $\mu : 2^N \to [0, 1]$ as follows: For $\ell \in \{0, \dots, m\}$, and for all $x \in \mathcal{B}_\ell$,

$$f(\phi(x)) = \begin{cases} 0 & \text{if } \ell = 0, \\ (2n)^{\ell} & \text{otherwise.} \end{cases} \tag{13}$$

[3]Notice that we use here *exactly* the same notation as in [35], which should facilitate the task of the reader willing to understand how the partition is built. We sketch the construction below. We know that $(P \cup I \cup M)$ contains no strict cycle.

When $I = \varnothing$, the construction is easy to explain. Because, the preference information can be represented using a 2-additive capacity, we know that $(P \cup M)$ has no cycle. Then $\mathcal{B}_0$ consists in all alternatives in $\mathcal{B}$ having no successor. These alternatives are then removed and the same process is applied repeatedly.

and we define μ letting:

$$\begin{array}{l} \mu_{\varnothing} = 0, \\ \mu_i = {}^{f_i}/_{\alpha}, \quad \forall i \in N, \\ \mu_{ij} = {}^{f_{ij}}/_{\alpha}, \quad \forall i,j \in N, \\ \mu(K) = \sum_{\{i,j\}\subseteq K} \mu_{ij} - (|K|-2)\sum_{i\in K}\mu_i, \quad \forall K \subseteq N, |K|>2, \end{array} \tag{14}$$

where

$$f_i := f(\phi(a_i)), f_{ij} := f(\phi(a_{ij})), \text{ and } \alpha = \sum_{\{i,j\}\subseteq N} f_{ij} - (n-2)\sum_{i\in N} f_i.$$

The capacity μ, defined like this, is 2-additive (see [35, Prop. 7, P. 317]). We have $f_{ij} = (2n)^p$, $f_i = (2n)^q$, $f_j = (2n)^s$ and then $\mu_{ij} > \mu_i + \mu_j$, i.e., $I^{\mu}_{ij} > 0$.

Hence, we proved that, if $I = \varnothing$ then there exists a capacity μ such that $\forall i,j \in N$, $I^{\mu}_{ij} > 0$, i.e., i and j possibly interact positively. In other words, there is no necessary null and negative interaction between criteria i and j. □

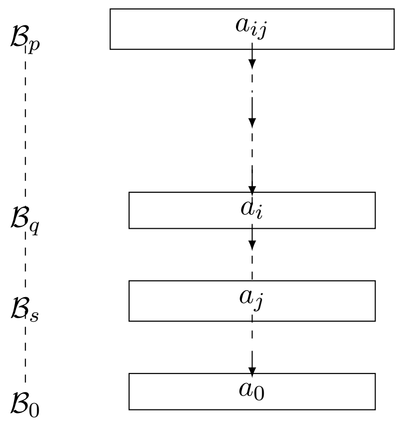

Figure 1: An illustration of the elements $\mathcal{B}_p$, $\mathcal{B}_q$, $\mathcal{B}_s$, and $\mathcal{B}_0$ such that $p > q > 0$ and $p > s > 0$

The above theorem is illustrated below.

Example 1

Let $N = \{1,2,3\}$, so that $\mathcal{B} = \{a_0, a_1, a_2, a_3, a_{12}, a_{13}, a_{23}\}$. Let $P = \{(a_{23}, a_2), (a_{23}, a_{12})\}$ and $I = \varnothing$.

This preference information is representable by a 2-additive Choquet since it contains no strict cycle (see [35, Cor. 1]).

The relation $(P \cup M)$ is as follows:

$$\begin{array}{l} a_{23}\ P\ a_{12}, a_{23}\ P\ a_2, a_{23}\ M\ a_3, \\ a_{12}\ M\ a_1, a_{12}\ M\ a_2, \\ a_{13}\ M\ a_1, a_{13}\ M\ a_3, \\ a_1\ M\ a_0, a_2\ M\ a_0, a_3\ M\ a_0. \end{array}$$

Using the technique described above, the only alternative without successor is a_0, so that $\mathcal{B}_0 = \{a_0\}$. Continuing the process, we find, $\mathcal{B}_1 = \{a_1, a_2, a_3\}$, $\mathcal{B}_2 = \{a_{12}, a_{13}\}$ and $\mathcal{B}_3 = \{a_{23}\}$. This is illustrated in Figure 2,

The capacity μ computed from the mapping f, as defined in the above proof, is: $\mu_1 = \mu_2 = \mu_3 = 6/270$, $\mu_{12} = \mu_{13} = 36/270$, $\mu_{13} = 216/270$. Hence we have $I^{\mu}_{12} = 36/270 - 6/270 - 6/270 = 24/270$, $I^{\mu}_{13} = 36/270 - 6/270 - 6/270 = 24/270$ and $I^{\mu}_{23} = 216/270 - 6/270 - 6/270 = 204/270$. $\diamond$

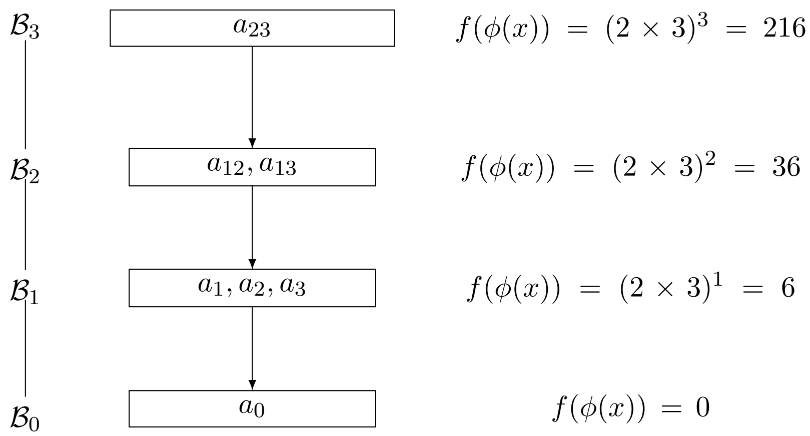

Figure 2: The partition of $\mathcal{B}$ elaborated from $P = \{(a_{23}, a_2), (a_{23}, a_{12})\}$

Theorem 2 states that negative and null interactions are never necessary if the preference information on $\mathcal{B}$ does not contain indifference. Such a situation in which the DM provides only strict preferences, most often happens in MCDA.

To investigate the situation when the DM expresses some indifference between binary alternatives, we need to define the following property, called 2-MOPI (2-Monotonicity Of Preferential Information), as first introduced in [35, Def. 4.2, p. 308].

Definition 3
A pair $\{i, j\} \subseteq N$ satisfies the 2-*MOPI*-$\{i, j\}$ *property* if

$$[a_{ij} \sim a_i \Rightarrow Not[a_j \; TC_P \; a_0]] \text{ and } [a_{ij} \sim a_j \Rightarrow Not[a_i \; TC_P \; a_0]]. \tag{15}$$

When the above equation hold, the alternatives a_j and a_i are said to be Neutral Binary Alternatives (NBA).

An ordinal information $\{P, I\}$ on $\mathcal{B}$ satisfies the 2-*MOPI property* if

$$\forall i, j \in N, i \neq j, \quad \text{2-MOPI-}\{i, j\} \text{ is satisfied.} \tag{16}$$

$\lrcorner$

The 2-MOPI property means that the contribution of the criterion j to the pair of criteria $\{i, j\}$ could be insignificant, whenever the satisfaction of the DM on the pair $\{i, j\}$ is equivalent to his satisfaction on the single criterion i. This is a strong condition, since it suffices that one such criterion i exists to infer the "nullity" of the criterion j. If this property holds for all pairs of criteria, then it is always possible to compute a capacity such that all the interaction indices are nonnegative, i.e., such that negative interactions are not necessary. This situation is characterized in Theorem 3 below, which is a reformulation of [35, Th. 3, p. 308], obtained in the case of 2-additive belief functions. We sketch its proof in order to be self-contained.

Theorem 3
Let $\{P, I\}$ be a preference information on $\mathcal{B}$, given by the DM and representable by a 2-additive Choquet integral. There exists a capacity $\mu \in \mathcal{C}_{2\text{-}add}$ such that $\forall i, j \in N, I^{\mu}_{ij} \geq 0$ iff the preference information satisfies the 2-MOPI property.

Proof
Necessity. Suppose that there exists a capacity $\mu \in \mathcal{C}_{2\text{-}add}$ such that $\forall i, j \in N, I^{\mu}_{ij} \geq 0$. If there exists $i_0, j_0 \in N$ such that $a_{i_0,j_0} \sim a_{i_0}$ and $a_{j_0}\ TC_P\ a_0$, we have $\mu_{i_0,j_0} = \mu_{i_0}$ and $\mu_{j_0} > 0$, i.e., $I^{\mu}_{i_0,j_0} < 0$, a contradiction.

Sufficiency. Assume that $\forall i, j \in N$, 2-MOPI-$\{i, j\}$ holds, i.e., $\forall i, j \in N$, $a_{ij} \sim a_i \Rightarrow$ $Not[a_j\ TC_P\ a_0]$. We build a partition $\{\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of $\mathcal{B}$ and build a capacity μ, belonging to $\mathcal{C}_{2\text{-}add}$, that is such that $I^{\mu}_{ij} = 0$ if $a_{ij} \sim a_i$ and $I^{\mu}_{ij} > 0$ if $Not[a_{ij} \sim a_i]$. This construction is detailed[4] in [35, Sect. 5.2, p. 315].

We consider the mapping $f : \mathcal{B} \to \mathbb{R}$ and the capacity $\mu : 2^N \to [0, 1]$ defined by Equations (13) and (14).

Let $i, j \in N$, with $i \neq j$. It is not difficult to check that if $a_{ij} \sim a_i$, then we have $Not[a_j\ TC_P\ a_0]$. In this case there exists $q \in \{0, \ldots, m\}$ such that $a_{ij}, a_i \in \mathcal{B}_q$ and $a_j \in \mathcal{B}_0$ Therefore we obtain $I^{\mu}_{i,j} = 0$. If $Not[a_{ij} \sim a_i]$, then there are $p, q, s \in \{0, \ldots, m\}$ such that $a_{ij} \in \mathcal{B}_p$, $a_i \in \mathcal{B}_q$, $a_j \in \mathcal{B}_s$ with $p > q$ and $p > s$. Hence, we have $f_{ij} = (2n)^p$, $f_i = (2n)^q$, $f_j = (2n)^s$ and then $\mu_{ij} \geq \mu_i + \mu_j$, i.e., $I^{\mu}_{ij} > 0$. □

We illustrate the above theorem below.

Example 2
Let $N = \{1, 2, 3\}$, so that $\mathcal{B} = \{a_0, a_1, a_2, a_3, a_{12}, a_{13}, a_{23}\}$. Let $P = \{(a_{23}, a_2), (a_{23}, a_{12})\}$ and $I = \{(a_{12}, a_1)\}$.

This preference information is representable by a 2-additive Choquet since it contains no strict cycle and it satisfies the MOPI property (see Theorem 1).

It is not difficult to check that every pair $\{i, j\}$ satisfies the 2-MOPI-$\{i, j\}$ property.

[4] Notice that we use here *exactly* the same notation as in [35], which should facilitate the task of the reader willing to understand how the construction works. We give a sketch below. We know that $(P \cup I \cup M)$ contains no strict cycle. When $I = \varnothing$, we have detailed above the construction. When $I \neq \varnothing$, the construction is slightly more involved. For each Neutral Binary Alternative a_i, we add to the relation $(P \cup I \cup M)$ an arc from a_0 to a_i (this is the relation $\mathcal{Z}$ in [35]). We first reduce the cycles of the relation $(I \cup M \cup \mathcal{Z})$. We then apply the process described for the case $I = \varnothing$ without change.

The relation $(P \cup I \cup M \cup \mathcal{Z})$ is as follows:

$$\begin{aligned}
&a_{23}\ P\ a_{12}, a_{23}\ P\ a_2, a_{23}\ M\ a_3,\\
&a_{12}\ I\ a_1, a_{12}\ M\ a_2,\\
&a_{13}\ M\ a_1, a_{13}\ M\ a_3,\\
&a_1\ I\ a_{12}, a_1\ M\ a_0,\\
&a_2\ M\ a_0,\\
&a_3\ M\ a_0,\\
&a_0\ \mathcal{Z}\ a_2.
\end{aligned}$$

By using the process described above, we first reduce the circuits in the relation $(I \cup M \cup \mathcal{Z})$, so that a_{12} and a_1 on the one hand and a_2 and a_0 are now merged, so that we now deal with a relation without cycle.

Using the technique described above, we obtain: $\mathcal{B}_0 = \{a_0, a_2\}$. Continuing the process, we find, $\mathcal{B}_1 = \{a_1, a_3, a_{12}\}$ and $\mathcal{B}_2 = \{a_{13}, a_{23}\}$.

The capacity μ computed from the mapping f, as defined in the proof of Theorem 2 (see Equations (13) and (14)), is such that: $\mu_2 = 0$, $\mu_1 = \mu_3 = \mu_{12} = 6/66$, $\mu_{23} = \mu_{13} = 36/66$. Hence we have $I^{\mu}_{12} = 6/66 - 6/66 - 0 = 0$, $I^{\mu}_{13} = 36/66 - 6/66 - 6/66 = 24/66$ and $I^{\mu}_{23} = 36/66 - 0 - 6/66 = 30/66$. $\diamond$

Theorem 3 implies that, in order to be sure that there is no necessary negative interaction, the 2-MOPI condition needs to be satisfied for all pairs of criteria. Hence, compared to the result of Theorem 2, the test related to the presence of necessary interactions is more complex when indifference is allowed. In this context, how to detect a necessary negative interaction between two given criteria, without trying to know what is happening among the other pairs of criteria? Our main result, Theorem 4 below, answers this question: a negative interaction between two criteria is necessary if and only if these two criteria i and j do not satisfy the 2-MOPI-$\{i, j\}$ property. Therefore, this condition is seen as a “local” condition.

Theorem 4
Let $\{P, I\}$ be an ordinal information on $\mathcal{B}$, given by the DM that can be represented by a 2-additive Choquet integral. Let $i, j \in N$ be a pair of distinct criteria.

There exists a capacity $\mu \in \mathcal{C}_{2\text{-}add}$ such that $I^{\mu}_{ij} \geq 0$ iff the condition 2-MOPI-$\{i, j\}$ holds.

PROOF
The proof is entirely similar to that of Theorem 3, except the fact that, here the two criteria i, j are fixed. $\square$

Remark 3
In Theorem 3, the capacity μ computed is identical for all the pairs of criteria satisfying the 2-MOPI property. All the interaction indices w.r.t. μ and associated to all these pairs are positive or null, i.e., $I^{\mu}_{ij} \geq 0$ for all $i, j \in N$.

In Theorem 4, the capacity μ is related on a fixed pair i_0, j_0 satisfying the 2-MOPI property. Only the interaction index w.r.t. μ and associated to i_0, j_0 must be positive or null. Here the sign of the interaction index associated to the other pairs of criteria is unknown. That is why the 2-MOPI-$\{i,j\}$ property plays the role of a local condition in Theorem 4. $\bullet$

The next remark gives the contrapositive of the Theorem 4. It gives necessary and sufficient conditions for the existence of necessary negative interaction between two criteria.

Remark 4
Let $\{P, I\}$ be an ordinal information on $\mathcal{B}$, given by the DM that can be represented using a 2-additive Choquet integral. Let $i, j \in N$ be a fixed pair of distinct criteria.

There is a necessary negative interaction between i and j iff $\{i,j\}$ does not satisfy the 2-MOPI-$\{i,j\}$ property, i.e.,

$$\big(a_{ij} \sim a_i \text{ and } a_j \; TC_P \; a_0\big) \text{ or } \big(a_{ij} \sim a_j \text{ and } a_i \; TC_P \; a_0\big).$$

This is illustrated below.

Example 3
Let $N = \{1,2,3\}$, so that $\mathcal{B} = \{a_0, a_1, a_2, a_3, a_{12}, a_{13}, a_{23}\}$. Let $P = \{(a_{23}, a_2), (a_{23}, a_{12}), (a_2, a_0)\}$ and $I = \{(a_{12}, a_1)\}$.

This preference information is representable by a 2-additive Choquet since it contains no strict cycle and it satisfies the MOPI property (see Theorem 1).

It is not difficult to check that:

- The pairs $\{1,3\}$ and $\{2,3\}$ satisfy the 2-MOPI-$\{i,j\}$ property,
- The pair $\{1,2\}$ does not satisfy 2-MOPI-$\{i,j\}$ property since $a_{12} \; I \; a_1$ and $a_2 \; P \; a_0$.

It follows from Remark 4 that only the interaction between 1 and 2 is necessarily negative.

The relation $(P \cup I \cup M \cup \mathcal{Z})$ is as follows:

$$\begin{aligned}
&a_{23} \; P \; a_{12}, a_{23} \; P \; a_2, a_{23} \; M \; a_3,\\
&a_{12} \; I \; a_1, a_{12} \; M \; a_2,\\
&a_{13} \; M \; a_1, a_{13} \; M \; a_3,\\
&a_1 \; I \; a_{12}, a_1 \; M \; a_0,\\
&a_2 \; P \; a_0,\\
&a_3 \; M \; a_0.
\end{aligned}$$

Using the technique described above, we obtain: $\mathcal{B}_0 = \{a_0\}$, $\mathcal{B}_1 = \{a_2, a_3\}$, $\mathcal{B}_2 = \{a_1, a_{12}\}$ and $\mathcal{B}_3 = \{a_{13}, a_{23}\}$.

Using Equations (13) and (14), we obtain $\mu_2 = \mu_3 = {}^{6}/_{420}$, $\mu_1 = \mu_{12} = {}^{36}/_{420}$, $\mu_{13} = \mu_{23} = {}^{216}/_{429}$. This implies $I^{\mu}_{12} = {}^{36}/_{420} - {}^{36}/_{420} - {}^{6}/_{420} = {}^{-6}/_{420}$, $I^{\mu}_{13} = {}^{216}/_{420} - {}^{36}/_{420} - {}^{6}/_{420} = {}^{174}/_{420}$ and $I^{\mu}_{23} = {}^{216}/_{420} - {}^{6}/_{420} - {}^{6}/_{420} = {}^{204}/_{420}$. $\diamond$

5.3 A procedure identifying necessary interactions on binary alternatives

In many MCDA applications, capturing an interaction phenomenon seems important. We have seen that, when a capacity is inferred from preference information, it is advisable not to interpret interactions that are not necessary.

For preference information obtained on binary alternatives, it is possible to summarize our results in the form of an algorithm. This algorithm deals with the detection of necessary negative interactions between criteria i and j. It was elaborated following the conditions used in the previous three theorems. Its justification is provided by the contrapositive of Theorem 4 (see Remark 4).

The input of the algorithm is a preference information $\{P, I\}$ on the set $\mathcal{B}$. We first test if I is empty. If YES, then the algorithm outputs that "i and j possibly interact positively" and stops. If NO, then we test if $a_{ij} \sim a_i$. If NO , then the algorithm outputs that "i and j do not necessarily interact negatively" and stops. If YES, then we test if $a_j \; TC_P \; a_0$. If NO , then the algorithm outputs that "i and j do not necessarily interact negatively" and stops. If YES , then the algorithm outputs that "i and j necessarily interact negatively" and stops.

5.4 Remarks

Because the results above may appear rather negative, we would like to emphasize here several features of the above analysis that call for caution in their interpretation.

First of all, the framework of binary alternatives *is* quite restrictive. Outside this framework, it is simple to devise examples in which preference information leads to a necessary negative interaction. The following example was suggested to us by Patrice Perny and is based on the analysis in [6, Th. 1].

Example 4
Suppose that 2 criteria expressed on a commensurate scale form 0 to 1. Suppose now the decision maker has given the following information:

$$(1, 0) \; P \; (0.5, 0.5),$$
$$(0, 1) \; P \; (0.5, 0.5).$$

This very poor preference information is nevertheless sufficient to imply the existence of a necessary negative interaction between the two criteria. Indeed, these two relations imply:

$$\mu_1 > 1/2\mu_{12},$$
$$\mu_2 > 1/2\mu_{12},$$

which implies $\mu_1 + \mu_2 > \mu_{12}$, so that $I^{\mu}_{12} < 0$. $\diamond$

Second, our results on binary alternatives exhibit a strange asymmetry. We have been able to prove that necessary negative interactions are rare. It would be tempting to conclude that a similar conclusion holds for necessary positive interactions. However, we do not have such results at hand. Indeed, the proof technique that was used above is not easily adapted to cover the case of positive interactions.

Hence, when we say that necessary negative interactions are rare, this is only valid when we work with binary alternatives and we do not suggest that the same conclusion holds for necessary positive interaction. Hence there is a real need to have tools to test for the existence of positive and negative interactions outside the framework of binary alternatives. We do so in the next section.

6 A LP model testing for necessary interaction

In this section, we drop the hypothesis that we only ask preference information on binary alternatives. We show how to test for the existence of necessary positive and negative interactions on the basis of preference information given on a subset of X that is not necessarily $\mathcal{B}$. Our approach is based on a linear program that was initially proposed in [34] for the elicitation of a capacity. We are only interested here in testing for the existence of a capacity with given properties and not in choosing the more "adequate" one (e.g., the one maximizing entropy, as done in [32, Sect. 5]).

Assume that the DM provides a strict preference P and an indifference I relations on a subset of X. Let i, j be two distinct criteria in N. Our approach consists in testing first, in two steps, the compatibility of this preference information with a 2-additive Choquet integral, and then, in the last step, the existence of necessary positive or negative interaction between i and j. These three steps are as follows [5].

Step 1. The following linear program (PL_1) models each preference of $\{P, I\}$ by introducing two nonnegative slack variables Γ^+_{xy} and Γ^-_{xy} in the corresponding constraint (Equation (17a) or (17b)). To ensure the normalization and monotonicity of the capacity μ, Equations (17e) to (17g), expressed in terms of the Möbius transform, need to be satisfied. The objective function Z_1 minimizes all the nonnegative variables introduced in (17a) and (17b).

$$\text{Minimize } Z_1 = \sum_{(x,y)\in P\cup I} (\Gamma^+_{xy} + \Gamma^-_{xy}), \tag{PL$_1$}$$

subject to

$$C_\mu(u(x)) - C_\mu(u(y)) + \Gamma^+_{xy} - \Gamma^-_{xy} \geq \varepsilon \quad \forall x, y \in X \text{ such that } x\ P\ y, \tag{17a}$$

$$C_\mu(u(x)) - C_\mu(u(y)) + \Gamma^+_{xy} - \Gamma^-_{xy} = 0 \quad \forall x, y \in X \text{ such that } x\ I\ y, \tag{17b}$$

$$\Gamma^+_{xy} \geq 0, \quad \Gamma^-_{xy} \geq 0 \quad \forall x, y \in X \text{ such that } x\ (P \cup I)\ y, \tag{17c}$$

[5] It is customary to convert strict inequalities into non-strict ones, using a constant ϵ, chosen to be "small". This is perfectly legitimate. Doing so would clearly allow to simplify the presentation below, converting Steps 1 and 2 into a single step. However, we think that using such a constant ϵ should be avoided, whenever possible. An infeasible LP using such a constant has an ambiguous interpretation: either the constraints are indeed incompatible or ϵ has not been chosen small enough. With our use of Steps 1 and 2, we avoid this potential ambiguity. Notice that Steps 1 and 2 below can be treated simultaneously, using the lexicographic optimization options offered by most solvers.

$$\varepsilon \geq 0, \tag{17d}$$

$$\sum_{\{i,j\} \subseteq N} m^{\mu}(\{i,j\}) + \sum_{i \in N} m^{\mu}(\{i\}) = 1, \tag{17e}$$

$$m^{\mu}(\{i\}) \geq 0 \quad \text{for all } i \in N, \tag{17f}$$

$$m^{\mu}(\{i\}) + \sum_{j \in A \setminus \{i\}} m^{\mu}(\{i,j\}) \geq 0 \quad \forall A \setminus \{i\}, \forall i \in N. \tag{17g}$$

Observe that the linear program (PL_1) is always feasible due to the introduction of the nonnegative variables Γ^{+}_{xy} and Γ^{-}_{xy}. There are two cases:

1. If the optimal solution of (PL_1) is $Z_1^* = 0$, then we can conclude that, depending on the sign of the variable ε, the preference information $\{P, I\}$ may be representable by a 2-additive Choquet integral. The next step of the procedure, Step 2 hereafter, will confirm or not this possibility.

2. If the optimal solution of (PL_1) is $Z_1^* > 0$, then there is no 2-additive Choquet integral model compatible with $\{P, I\}$.

Step 2. The linear program (PL_2) ensures the existence of a 2-additive Choquet integral model compatible with $\{P, I\}$, when the optimal solution of (PL_1) is $Z_1^* = 0$. Compared to the previous linear program, in this formulation, we only removed the nonnegative variables Γ^{+}_{xy} and Γ^{-}_{xy} (or put them equal to zero) and change the objective function by maximizing the value of the variable ε, in order to satisfy the strict preference relation.

$$\text{Maximize } Z_2 = \varepsilon, \tag{PL_2}$$

subject to

$$C_{\mu}(u(x)) - C_{\mu}(u(y)) \geq \varepsilon \quad \forall x, y \in X \text{ such that } x\ P\ y, \tag{18a}$$

$$C_{\mu}(u(x)) - C_{\mu}(u(y)) = 0 \quad \forall x, y \in X \text{ such that } x\ I\ y, \tag{18b}$$

$$\varepsilon \geq 0, \tag{18c}$$

$$\sum_{\{i,j\} \subseteq N} m^{\mu}(\{i,j\}) + \sum_{i \in N} m^{\mu}(\{i\}) = 1, \tag{18d}$$

$$m^{\mu}(\{i\}) \geq 0 \quad \text{for all } i \in N, \tag{18e}$$

$$m^{\mu}(\{i\}) + \sum_{j \in A \setminus \{i\}} m^{\mu}(\{i,j\}) \geq 0 \quad \forall A \setminus \{i\}, \forall i \in N. \tag{18f}$$

Notice that (PL_2) is solved only if $Z_1^* = 0$. Hence, the linear program (PL_2) is always feasible and it does not have an unbounded solution (it is not restrictive to suppose that $C_{\mu}(u(x)) \in [0, 1], \forall x \in X$). Hence, there are two cases.

1. If the optimal solution of (PL_2) is $Z_2^* = 0$, then there is no 2-additive Choquet integral model compatible with $\{P, I\}$.

2. If the optimal solution of (PL_2) is $Z_2^* > 0$, then $\{P, I\}$ is representable by a 2-additive Choquet integral.

Step 3. At this step, we suppose the preference information $\{P, I\}$ representable by a 2-additive Choquet integral, i.e., $Z_2^* > 0$. In order to know if the interaction between i and j is necessarily negative (resp. positive) w.r.t. the provided preference information, we solve the following linear program denoted by $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$):

$$\text{Maximize } Z_3 = \varepsilon, \tag{$PL^{ij}_{\text{N-N}}$}$$
subject to
$$m^\mu(\{i,j\}) \geq 0 \quad (\text{resp. } m^\mu(\{i,j\}) \leq 0), \tag{19a}$$
$$C_\mu(u(x)) - C_\mu(u(y)) \geq \varepsilon \quad \forall x, y \in X \text{ such that } x \; P \; y, \tag{19b}$$
$$C_\mu(u(x)) - C_\mu(u(y)) = 0 \quad \forall x, y \in X \text{ such that } x \; I \; y, \tag{19c}$$
$$\varepsilon \geq 0, \tag{19d}$$
$$\sum_{\{i,j\} \subseteq N} m^\mu(\{i,j\}) + \sum_{i \in N} m^\mu(\{i\}) = 1, \tag{19e}$$
$$m^\mu(\{i\}) \geq 0 \quad \text{for all } i \in N, \tag{19f}$$
$$m^\mu(\{i\}) + \sum_{j \in A \setminus \{i\}} m^\mu(\{i,j\}) \geq 0 \quad \forall A \setminus \{i\}, \forall i \in N. \tag{19g}$$

This linear program tries to find a possible positive (resp. a negative) interaction index between i and j, such that the preference $\{P, I\}$ is representable by a 2-additive Choquet integral. To do this, as in (PL_2), the positive value ε, allowing to satisfy the strict preference relation P, is maximized. We keep also the other constraints of the linear program (PL_2) and add only the sign of the interaction index $I^\mu_{ij} = m^\mu(\{i,j\}) = \mu_{ij} - \mu_i - \mu_j$ (positive or negative), between i and j, tested by the Equation (19a).

After a resolution of the linear program $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$), we have one of the following three conclusions:

1. If $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$) is not feasible, then there is a necessary negative (resp. positive) interaction between i and j. Indeed, as the program (PL_2) is feasible with an optimal solution $Z_2^* > 0$, the contradiction about the representation of $\{P, I\}$ only comes from the introduction of the constraint $I^\mu_{ij} = m^\mu(\{i,j\}) = \mu_{ij} - \mu_i - \mu_j \geq 0$ (resp. $I^\mu_{ij} = m^\mu(\{i,j\}) = \mu_{ij} - \mu_i - \mu_j \leq 0$) in $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$).

2. If $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$) is feasible and the optimal solution $Z_3^* = 0$, then the constraint (19b) is satisfied with $\varepsilon = 0$, i.e., it is not possible to model strict preference by adding the constraint $I^\mu_{ij} = m^\mu(\{i,j\}) = \mu_{ij} - \mu_i - \mu_j \geq 0$ (resp. $I^\mu_{ij} = m^\mu(\{i,j\}) = \mu_{ij} - \mu_i - \mu_j \leq 0$) in $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$). Therefore, we can conclude that there is a necessary negative (resp. positive) interaction between i and j.

3. If $PL^{ij}_{\text{N-N}}$ (resp. $PL^{ij}_{\text{N-P}}$) is feasible and the optimal solution $Z_3^* > 0$, then there is no necessary negative (resp. positive) interaction between i and j.

This procedure is illustrated below.

Example 5

We consider the preferences given by the DM in the classic example of students evaluation where $N = \{M, S, L\}$. We proved in Section 3 that these preferences are representable by a 2-additive Choquet integral. The linear program $PL^{MS}_{\text{N-N}}$ (resp. $PL^{MS}_{\text{N-P}}$) corresponding to the test of the existence of a necessary negative (resp. positive) interaction between the Mathematics (M) and Statistics (S) is the following:

$$
\begin{aligned}
&\text{Maximize} \quad Z_3 = \varepsilon,\\
&\text{subject to}\\
&m^\mu(\{M,S\}) \geq 0 \quad (\text{resp. } m^\mu(\{M,S\}) \leq 0),\\
&\varepsilon \geq 0,\\
&C_\mu(u(b)) - C_\mu(u(a)) \geq \varepsilon,\\
&C_\mu(u(c)) - C_\mu(u(d)) \geq \varepsilon,\\
&m^\mu(\{M,S\}) + m^\mu(\{M,L\}) + m^\mu(\{S,L\}) + m^\mu(\{M\}) + m^\mu(\{S\}) + m^\mu(\{L\}) = 1,\\
&m^\mu(\{M\}) \geq 0 \quad m^\mu(\{S\}) \geq 0 \quad m^\mu(\{L\}) \geq 0,\\
&m^\mu(\{M\}) + m^\mu(\{M,S\}) \geq 0,\\
&m^\mu(\{M\}) + m^\mu(\{M,L\}) \geq 0,\\
&m^\mu(\{M\}) + m^\mu(\{M,S\}) + m^\mu(\{M,L\}) \geq 0,\\
&m^\mu(\{S\}) + m^\mu(\{M,S\}) \geq 0,\\
&m^\mu(\{S\}) + m^\mu(\{S,L\}) \geq 0,\\
&m^\mu(\{S\}) + m^\mu(\{M,S\}) + m^\mu(\{S,L\}) \geq 0,\\
&m^\mu(\{L\}) + m^\mu(\{M,L\}) \geq 0,\\
&m^\mu(\{L\}) + m^\mu(\{S,L\}) \geq 0,\\
&m^\mu(\{L\}) + m^\mu(\{S,L\}) + m^\mu(\{M,L\}) \geq 0.
\end{aligned}
$$

Replacing $C_\mu(u(a))$, $C_\mu(u(b))$, $C_\mu(u(c))$ and $C_\mu(u(d))$ by their expression (in terms of interaction indices I^μ_{ij} and Shapley values V_i, $i, j \in N = \{M, S, L\}$) in the above linear program, we obtain this equivalent formulation:

$$
\begin{aligned}
&\text{Maximize} \quad Z_3 = \varepsilon,\\
&\text{subject to}\\
&16V^\mu_M + 11V^\mu_S + 9V^\mu_L - 2.5I^\mu_{MS} - 3.5I^\mu_{ML} - I^\mu_{SL} -\\
&\quad [16V^\mu_M + 13V^\mu_S + 7V^\mu_L - 1.5I^\mu_{MS} - 4.5I^\mu_{ML} - 3I^\mu_{SL}] \geq \varepsilon,\\
&6V^\mu_M + 13V^\mu_S + 7V^\mu_L - 3.5I^\mu_{MS} - 0.5I^\mu_{ML} - 3I^\mu_{SL} -\\
&\quad [6V^\mu_M + 11V^\mu_S + 9V^\mu_L - 2.5I^\mu_{MS} - 1.5I^\mu_{ML} - I^\mu_{SL}] \geq \varepsilon,\\
&m^\mu(\{M,S\}) \geq 0 \qquad (\text{resp. } m^\mu(\{M,S\}) \leq 0),\\
&\varepsilon \geq 0,
\end{aligned}
$$

$$
\begin{aligned}
&m^\mu(\{M,S\})+m^\mu(\{M,L\})+m^\mu(\{S,L\})+m^\mu(\{M\})+m^\mu(\{S\})+m^\mu(\{L\})=1,\\
&m^\mu(\{M\})\geq 0 \quad m^\mu(\{S\})\geq 0 \quad m^\mu(\{L\})\geq 0,\\
&m^\mu(\{M\})+m^\mu(\{M,S\})\geq 0,\\
&m^\mu(\{M\})+m^\mu(\{M,L\})\geq 0,\\
&m^\mu(\{M\})+m^\mu(\{M,S\})+m^\mu(\{M,L\})\geq 0,\\
&m^\mu(\{S\})+m^\mu(\{M,S\})\geq 0,\\
&m^\mu(\{S\})+m^\mu(\{S,L\})\geq 0,\\
&m^\mu(\{S\})+m^\mu(\{M,S\})+m^\mu(\{S,L\})\geq 0,\\
&m^\mu(\{L\})+m^\mu(\{M,L\})\geq 0,\\
&m^\mu(\{L\})+m^\mu(\{S,L\})\geq 0,\\
&m^\mu(\{L\})+m^\mu(\{S,L\})+m^\mu(\{M,L\})\geq 0,\\
&I^\mu_{MS}=m^\mu(\{M,S\}),\\
&I^\mu_{ML}=m^\mu(\{M,L\}),\\
&I^\mu_{SL}=m^\mu(\{S,L\}),\\
&V_M=m^\mu(\{M\})+0.5I^\mu_{MS}+0.5I^\mu_{ML},\\
&V_S=m^\mu(\{S\})+0.5I^\mu_{MS}+0.5I^\mu_{SL},\\
&V_L=m^\mu(\{L\})+0.5I^\mu_{ML}+0.5I^\mu_{SL}.
\end{aligned}
$$

The results obtained by solving $PL^{MS}_{\text{N-N}}$ (resp. $PL^{MS}_{\text{N-P}}$) are given in Table 2 (resp. Table 3). We can conclude that the interaction between Mathematics and Statistics is neither not necessarily negative nor necessarily positive, because the optimal solution of the program $PL^{MS}_{\text{N-N}}$ and $PL^{MS}_{\text{N-P}}$ is respectively $Z^*_3=0.667$ and $Z^*_3=1$. $\diamond$

	$\mathbf{Z_3}=\varepsilon$	M	S	L	$\{\mathbf{M},\mathbf{S}\}$	$\{M,L\}$	$\{S,L\}$
Optimal solution Z^*_3	**0.667**	–	–	–	–	–	–
Möbius transform m^μ		0	0.33	0.33	0	0.67	−0.33
Importance index V^μ_i		0.33	0.17	0.5	–	–	–
Interaction index I^μ_{ij}		–	–	–	**0**	0.67	−0.33

Table 2: Results of $PL^{MS}_{\text{N-N}}$ testing necessary negative interaction between Mathematics and Statistics

7 Interactions in a ranking of hospitals for weight loss surgery

In this section, we illustrate our results using a real-world application. At the time the problem was tackled, it was thought that the MCDA model elaborated would allow for a solid interpretation of interaction phenomena. The problem is about a ranking of French hospitals for weight loss surgery elaborated in [36] using a 2-additive Choquet

	$\mathbf{Z_3} = \varepsilon$	M	S	L	$\{\mathbf{M}, \mathbf{S}\}$	$\{M, L\}$	$\{S, L\}$
Optimal solution Z_3^*	**1**	–	–	–	–	–	–
Möbius transform m^μ		0.5	0.5	0	-0.5	0.5	0
Importance index V_i^μ		0.5	0.25	0.25	–	–	–
Interaction index I_{ij}^μ		–	–	–	$\mathbf{-0.5}$	0.5	0

Table 3: Results of $PL_{\text{N-P}}^{MS}$ testing necessary positive interaction between Mathematics and Statistics

integral model. Weight loss surgery (bariatric surgery) includes a variety of procedures performed on people who are obese [6].

To identify the "best" hospitals in weight loss surgery, four criteria [7] are considered:

1. *Activity*: number of procedures performed during one year. In general, if a hospital has a good score on activity then its teams are more trained and often have good results. Therefore this criterion has to be maximized. The relevance of this criterion is not totally shared by all the experts in medicine.

2. *Notoriety*: Its corresponds to the reputation and attractiveness of the hospital. It is a percentage of patients treated in the hospital but living in another French administrative department. The more the percentage increases, the more the hospital is attractive.

3. *Average Length Of Stay (ALOS)*: a mean calculated by dividing the sum of inpatient days by the number of patients admissions with the same diagnosis-related group classification. The more the hospital is organized in terms of ressources, the more the ALOS score decreases.

4. *Technicality*: this particular indicator measures the ratio of procedures performed with an efficient technology compared to the same procedures performed with obsolete technology. The higher the percentage is, the more the team is trained in advanced technologies or complex surgeries.

We denote this set of criteria by $N = \{1, 2, 3, 4\}$. We have the following sets of values on each criterion $X_1 = [0, 500]$, $X_2 = [0, 100]$, $X_3 = [5, 0]$ and $X_4 = [0, 100]$.

The reference levels $\mathbf{1}_i$ and $\mathbf{0}_i$ associated to each criterion i, and identified by the DM, are given in Table 4 below. Of course, these reference elements could be different from the bounds defined in X_i, even if it is not the here. We have $\mathcal{B} = \{a_0, a_1, a_2, a_3, a_4, a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34}\}$.

[6] Weight loss is achieved by reducing the size of the stomach with a gastric band or through removal of a portion of the stomach (sleeve gastrectomy or biliopancreatic diversion with duodenal switch) or by resecting and re-routing the small intestines to a small stomach pouch (gastric bypass surgery).

[7] These four criteria are also used by the French magazine "Le Point" in their ranking of hospitals using the arithmetic mean.

	Activity	Notoriety	ALOS	Technicality
Satisfactory level $\mathbf{1}_i$	500	100	0	100
Neutral level $\mathbf{0}_i$	0	0	5	0

Table 4: The reference levels associated to each criterion in weight loss surgery.

The following interpretation of the binary alternatives, understandable by the experts, was proposed:

- $\mathbf{0}_N = (\mathbf{1}_\varnothing, \mathbf{0}_N) := a_0$ is a hospital considered neutral on all the four criteria.
- $(\mathbf{1}_i, \mathbf{0}_{N-i}) := a_i$ is a hospital considered satisfactory on criterion i and neutral on the other criteria.
- $(\mathbf{1}_{ij}, \mathbf{0}_{N-ij}) := a_{ij}$ is a hospital considered satisfactory on criteria i and j and neutral on the other criteria.

The preferences, on $\mathcal{B}$, provided by the DM are:

- A satisfactory hospital on Activity and ALOS (neutral on the other criteria) is better than a satisfactory hospital on Notoriety and Technicality (neutral on the other criteria), i.e., $a_{13}\ P\ a_{24}$.
- A satisfactory hospital on Activity (neutral on the other criteria) is better that a satisfactory hospital only on Notoriety and ALOS (neutral on the other criteria), i.e., $a_1\ P\ a_{23}$.
- A hospital only better in Activity (neutral on the other criteria) is judged indifferent to a hospital better on Activity and ALOS (neutral on the other criteria), i.e., $a_1\ I\ a_{13}$.
- If a hospital is fully satisfying on the criterion Technicality (neutral on the other criteria), then it will be preferred to a hospital satisfactory on Notoriety (neutral on the other criteria), i.e., $a_4\ P\ a_2$.
- A satisfactory hospital on Activity and Technicality (neutral on the other criteria) is better than a satisfactory hospital on ALOS and Technicality (neutral on the other criteria), i.e., $a_{14}\ P\ a_{34}$.

In [36], it is shown that this preference information can be represented by a 2-additive Choquet integral w.r.t. the capacity μ given in Table 5.

Based on the capacity computed in Table 5, Activity and Notoriety were judged complementary (positive interaction) while Activity and ALOS were judged redundant (negative interaction). The results presented above show that this interpretation is not fully warranted, since it does not correspond to necessary interactions. Indeed, it can be noticed that each pair of criteria satisfies the 2-MOPI property. Therefore, the only

	1	2	3	4	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
μ	0.2	0	0.1	0.1	0.9	0.2	0.3	0.1	0.1	0.2
V_i^μ	0.5	0.35	0.05	0.1	–	–	–	–	–	–
I_{ij}^μ	–	–	–	–	0.7	−0.1	0	0	0	0

Table 5: The parameters computed for our 2-additive Choquet integral model in weight loss surgery.

valid conclusion is: "the negative interaction related to these preference information is not necessary". In other words, the criteria Activity and Notoriety possibly interact positively while criteria Activity and ALOS possibly interact negatively.

8 Discussion

This paper has discussed the use of interaction indices in order to interpret interaction phenomena at work within a 2-additive Choquet integral model. We have concentrated on the case in which the 2-additive capacity is assessed on the basis of preference information provided by the DM. Our emphasis is on the common case in which the DM expresses preference information on a finite number of alternatives, as opposed to the continuous setting used in [46, 47, 48]. Unsurprisingly, the capacity that is elicited in such a setting is not unique. Moreover, the interpretation of the interaction effects between criteria requires some caution. Indeed, we have exhibited simple examples in which the sign of the interaction index depends upon the arbitrary choice of a capacity within the polyhedron of all capacities compatible with the preference information. This has led us to define the notion of necessary and possible interactions, given a preference information. Only necessary interactions are robust since their sign and, hence, interpretation, does not vary within the set of all representing capacities.

In the context of binary alternatives, we have characterized the situations in which negative interactions are necessary. Quite surprisingly, when the preference information does not contain indifference, there are no such situations. Negative interaction is possible but never necessary. This extends previous results in [35]. It is important to realize that these results do not carry over to the case of positive interaction and are only valid with the framework of binary alternatives.

Outside the framework of binary alternatives, we extend the linear programming formulation proposed in [34] to test for the existence of necessary interactions of various kinds. The central message of the paper is that interpreting interaction indices requires some care.

The subject of the paper offers several avenues for future research.

First, our results show a curious asymmetry. We have been able to show that necessary negative interactions are rare in the context of binary alternatives. It would be tempting to conjecture similar results for necessary positive interactions. But we do not have such results at hand for the time being. The suggestive duality between positive and negative interactions is not easy to exploit in view of the proof technique used in the results presented in Section 5. This clearly deserves further study.

Second, the framework of an *ordinal* preference information $\{P, I\}$ provided the decision-maker may seem unnecessarily poor. As pointed out to us by Christophe Labreuche, many decision-makers in the real world seem to be confortable with the idea of expressing more than only an ordinal preference information and ready to express preference differences or even intensity of preference when comparing alternatives with which they are familiar. If this kind of information is taken into account, this potentially drastically restricts the set of representing 2-additive capacity, which would increase the likelihood of obtaining necessary interactions. Using information of preference differences is indeed commonly used in the field of decision with multiple attributes [2, 11, 50]. We refer to [1] for the use of information concerning preference difference together with the use of a Choquet integral.

As pointed out to us by a referee, some authors think that the DM can be confortable in expressing directly information about the sign of interactions between pairs of criteria when eliciting the parameters of a Choquet integral. This was indeed suggested in [1] and in the application reported in [4] (for similar ideas in other contexts, one may also refer to [10, 40]). Without questioning the potential usefulness of such an approach in a constructive approach to decision-aiding [41], our results seem to show that interactions between criteria are a concept that is less easy to define and to grasp than is usually thought. Indeed as shown, for instance, by the analysis of the classic example presented in Section 3, imposing of constraint on the sign of the interaction index I^{μ}_{ij} does not always ensure that the underlying logic of the preference model that is built will be compatible with the constraint. Although the preference given by the Dean is justified by the fact that Mathematics and Statistics are somewhat redundant, this preference can be explained using a model in which the interaction index between Mathematics and Statistics is strictly positive (see Table 1). Hence, imposing in an LP model of the type presented in Section 6 a constraint such as $I^{\mu}_{ij} < 0$ (resp. > 0) is not a guarantee that the resulting preference model can be interpreted as if there is redundancy (resp. complementarity) between i and j.

Third, our analysis only deals with the case of 2-additive Choquet integrals. Although dominant, this model is not the only one that have been suggested to capture interaction phenomena. Among them the Sugeno integral model [5, 8, 9, 23] and the GAI model [12, 25, 30] deserve mention. This raises several questions. Is the interaction index mainly developed for the case of the Choquet integral [15, 19, 37] adapted to these new model? On what basis is it possible to answer this question? Given adequate Interaction indices for these models, can we obtain results similar to the one presented here.

Finally, the notion of interaction would deserve further study. In particular, it would be interesting to have a definition that would not depend on a particular aggregation technique or on a particular index.

We have started working on all these points.

References

[1] S. Angilella, S. Greco, and B. Matarazzo. Non-additive robust ordinal regression: A multiple criteria decision model based on the Choquet integral. *European Journal of Operational Research*, 41(1):277–288, 2010.

[2] C. A. Bana e Costa and J.-C. Vansnick. MACBETH – An interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research*, 1:489–500, 1994.

[3] L. Berrah and V. Clivillé. Towards an aggregation performance measurement system model in a supply chain context. *Computers in Industry*, 58(7):709–719, 2007.

[4] M. Bottero, V. Ferretti, J. R. Figueira, S. Greco, and B. Roy. On the Choquet multiple criteria preference aggregation model: Theoretical and practical insights from a real-world application. *European Journal of Operational Research*, 271(1):120–140, 2018.

[5] D. Bouyssou, T. Marchant, and M. Pirlot. A conjoint measurement approach to the discrete Sugeno integral. In S. Brams, W. V. Gehrlein, and F. S. Roberts, editors, *The Mathematics of Preference, Choice and Order. Essays in Honor of Peter C. Fishburn*, pages 85–109. Springer, Boston, 2009. ISBN 978-3-540-79127-0.

[6] A. Chateauneuf and J.-M. Tallon. Diversification, convex preferences and non-empty core in the Choquet expected utility model. *Economic Theory*, 19, 2002.

[7] V. Clivillé, L. Berrah, and G. Mauris. Quantitative expression and aggregation of performance measurements based on the MACBETH multi-criteria method. *International Journal of Production economics*, 105:171–189, 2007.

[8] M. Couceiro, D. Dubois, H. Fargier, M. Grabisch, H. Prade, and A. Rico. New directions in ordinal evaluation: Sugeno integrals and beyond. In Michalis Doumpos, José Rui Figueira, Salvatore Greco, and Constantin Zopounidis, editors, *New Perspectives in Multiple Criteria Decision Making: Innovative Applications and Case Studies*, pages 177–228. Springer International Publishing, 2019. ISBN 978-3-030-11482-4.

[9] D. Dubois, J. L. Marichal, H. Prade, M. Roubens, and R. Sabbadin. The use of the discrete Sugeno integral in decision-making: a survey. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 9(5):539–561, 2001.

[10] J. R. Figueira, S. Greco, and B. Roy. ELECTRE methods with interaction between criteria: An extension of the concordance index. *European Journal of Operational Research*, 199(2): 478–495, 2009.

[11] J. R. Figueira, S. Greco, and R. Słowiński. Building a set of additive value functions representing a reference preorder and intensities of preference: GRIP method. *European Journal of Operational Research*, 195(2):460–486, 2009.

[12] C. Gonzales, P. Perny, and J.-P. Dubus. Decision making with multiple objectives using GAI networks. *Artificial Intelligence*, 175(7–8):1153–1179, 2011.

[13] M. Grabisch. Fuzzy integral in multicriteria decision making. *Fuzzy Sets and Systems*, 69 (3):279–298, 1995.

[14] M. Grabisch. The application of fuzzy integrals to multicriteria decision making. *European Journal of Operational Research*, 89(1):445–456, 1996.

[15] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92(2):167–189, 1997.

[16] M. Grabisch. *Set Functions, Games and Capacities in Decision Making.* Theory and Decision Library C. Springer, 2016.

[17] M. Grabisch and C.Labreuche. A decade of application of the Choquet and Sugeno integrals in multi-criteria, decision aid. *4OR*, 6(1):1–44, 2008.

[18] M. Grabisch and C. Labreuche. Fuzzy measures and integrals in MCDA. In Salvatore Greco, Matthias Ehrgott, and José R. Figueira, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, volume 233 of *International Series in Operations Research & Management Science*, pages 553–603. Springer-Verlag New York, 2016.

[19] M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *International Journal of Game Theory*, 28(4):547–565, 1999.

[20] M. Grabisch, J. Duchêne, F. Lino, and P. Perny. Subjective evaluation of discomfort in sitting position. *Fuzzy Optimization and Decision Making*, 1(3):287–312, 2002.

[21] M. Grabisch, C. Labreuche, and J.-C. Vansnick. On the extension of pseudo-boolean functions for the aggregation of interacting criteria. *European Journal of Operational Research*, 148(1):28–47, 2003.

[22] M. Grabisch, I. Kojadinovic, and P. Meyer. A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, 186(2):766–785, 2008.

[23] S. Greco, B. Matarazzo, and R. Słowiński. Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European Journal of Operational Research*, 158(2):271–292, 2004.

[24] S. Greco, V. Mousseau, and R. Słowiński. Ordinal regression revisited: Multiple criteria ranking using a set of additive value functions. *European Journal of Operational Research*, 51(2):416–436, 2008.

[25] S. Greco, V. Mousseau, and R. Słowiński. Robust ordinal regression for value functions handling interacting criteria. *European Journal of Operational Research*, 239(3):711–730, 2014.

[26] R. L. Keeney and H. Raiffa. *Decisions with multiple objectives: Preferences and value tradeoffs*. Wiley, New York, 1976.

[27] Ivan Kojadinovic. Modeling interaction phenomena using fuzzy measures: on the notions of interaction and independence. *Fuzzy Sets and Systems*, 135(3):317–340, 2003.

[28] D. H. Krantz, R. D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement*, volume 1: Additive and polynomial representations. Academic Press, New York, 1971.

[29] C. Labreuche and M. Grabisch. The Choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems*, 137:11–26, 2003.

[30] C. Labreuche and M. Grabisch. Using multiple reference levels in Multi-criteria decision aid: The Generalized-Additive Independence model and the Choquet integral approaches. *European Journal of Operational Research*, 267(2):598–611, 2018.

[31] J.-L. Marichal. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, 8(6):800–807, 2000.

[32] J.-L. Marichal. Entropy of discrete Choquet capacities. *European Journal of Operational Research*, 137(3):612–624, 2002.

[33] J.-L. Marichal. Aggregation of interacting criteria by means of the discrete Choquet integral. In T. Calvo, G. Mayor, and R. Mesiar, editors, *Aggregation operators: new trends and applications*, volume 97 of *Studies in Fuzziness and Soft Computing*, pages 224–244. Physica Verlag, 2002.

[34] J.-L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *Eur. J. of Operational Research*, 124:641–650, 2000.

[35] B. Mayag, M. Grabisch, and C. Labreuche. A representation of preferences by the Choquet integral with respect to a 2-additive capacity. *Theory and Decision*, 71(3):297–324, 2011.

[36] Brice Mayag. A 2-additive Choquet integral model for French hospitals rankings in weight loss surgery. In J. P. Carvalho, M.-J. Lesot, U. Kaymak, S. M. Vieira, B. Bouchon-

Meunier, and R/ R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 16th International Conference, IPMU 2016, Eindhoven, The Netherlands, June 20-24, 2016, Proceedings, Part I*, volume 610 of *Communications in Computer and Information Science*, pages 101–116. Springer, 2016.

[37] T. Murofushi and S. Soneda. Techniques for reading fuzzy measures (III): interaction index. In *9th Fuzzy System Symposium*, pages 693–696, Sapporo, Japan, May 1993. In Japanese.

[38] T. Murofushi and M. Sugeno. Choquet integral models and independence concepts in multiattribute utility theory. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, 08, 2000.

[39] J.P. Pignon and C. Labreuche. A methodological approach for operational and technical experimentation based evaluation of systems of systems architectures. In *Int. Conference on Software & Systems Engineering and their Applications (ICSSEA)*, Paris, France, December 4-6 2007.

[40] B. Roy. À propos de la signification des dépendances entre critères : quelle place et quels modes de prise en compte pour l'aide à la décision ? *RAIRO - Operations Research - Recherche Opérationnelle*, 43(3):255–275, 2009.

[41] B. Roy. Two conceptions of decision aiding. *International Journal of Multicriteria Decision Making*, 1(1):74–79, 2010.

[42] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[43] L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games, Vol. II*, number 28 in Annals of Mathematics Studies, pages 307–317. Princeton University Press, 1953.

[44] H. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63 (2):129–138, 1956.

[45] P. Slovic, M. Finucane, E. Peters, and D. G. MacGregor. The affect heuristic. In T. Gilovitch, D. Griffin, and D. Kahneman, editors, *Heuristics and biases: the psychology of intuitive judgment*, pages 397–420. Cambridge University Press, 2002.

[46] M. Timonin. Axiomatization of the Choquet integral for 2-dimensional heterogeneous product sets. *ArXiv e-prints*, July 2015. `http://adsabs.harvard.edu/abs/2015arXiv150704167T`.

[47] M. Timonin. Conjoint axiomatization of the Choquet integral for heterogeneous product sets. *ArXiv e-prints*, March 2016. `http://adsabs.harvard.edu/abs/2016arXiv160308142T`.

[48] M. Timonin. Conjoint axiomatization of the Choquet integral for heterogeneous product sets. In J. P. Carvalho, M.-J. Lesot, U. Kaymak, S. M. Vieira, B. Bouchon-Meunier, and R. R. Yager, editors, *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 16th International Conference, IPMU 2016, Eindhoven, The Netherlands, June 20-24, 2016, Proceedings, Part I*, volume 610 of *Communications in Computer and Information Science*, pages 46–57. Springer, 2016.

[49] A. Tversky and D. Kahneman. Advances in prospect theory: cumulative representation of uncertainty. *J. of Risk and Uncertainty*, 5:297–323, 1992.

[50] D. von Winterfeldt and W. Edwards. *Decision analysis and behavioral research.* Cambridge University Press, Cambridge, 1986.

[51] P. P. Wakker. *Additive representations of preferences: A new foundation of decision analysis.* Kluwer, Dordrecht, 1989.

[52] P. P. Wakker. *Prospect Theory for Risk and Ambiguity.* Cambridge University Press, Cambridge, 2010.