Choquet integral versus weighted sum in multicriteria decision contexts

Thibaut Lust

Sorbonne Universités, UPMC Université Paris 06
CNRS, LIP6, UMR 7606, F-75005, Paris, France
thibaut.lust@lip6.fr

Abstract. In this paper, we address the problem of comparing the performances of two popular aggregation operators, the weighted sum and the Choquet integral, for selecting the best alternative among a set of alternatives, all evaluated according to different criteria. While the weighted sum is simple to use and very popular, the Choquet integral is still hard to use in practice but leads theoretically to better results in terms of concordance with the preferences of a decision maker. However, given the efforts needed to set the parameters of the Choquet integral, it is important to measure, for a given decision problem, if it is really worth defining the Choquet integral or if a simple weighted sum could have been used to determine the best alternative. We will compute the probability that a recommendation to a decision maker could only been obtained with the Choquet integral and not with a weighted sum. When the number of criteria increases, the results show that this probability tends to one. However, a high value of probability can only be attained for particular data sets.

Keywords: Choquet integral, weighted sum, multicriteria decision making, multiobjective optimization

1 Introduction

When a decision maker is confronted to a set of alternatives presenting each one meaningful advantages and disadvantages, the temptation to use a weighted sum to select the best alternative is high. However, it is well-known that this easy to use and simple aggregation method presents the disadvantage that, in some cases, alternatives cannot be elicited even if they correspond to the preferences of the decision maker.

For example, if we consider a set of four alternatives $\{y^1, \ldots, y^4\}$ where each alternative y^i is evaluated with three criteria to maximize: $y^1 = (18, 10, 10)$, $y^2 = (10, 18, 10)$, $y^3 = (10, 10, 18)$ and $y^4 = (14, 12, 11)$, the alternative y^4 could never be selected with a weighted sum (it is impossible to have at the same time $14\lambda_1 + 12\lambda_2 + 11\lambda_3 \geq 18\lambda_1 + 10\lambda_2 + 10\lambda_3$, $14\lambda_1 + 12\lambda_2 + 11\lambda_3 \geq 10\lambda_1 + 18\lambda_2 + 10\lambda_3$ and $14\lambda_1 + 12\lambda_2 + 11\lambda_3 \geq 10\lambda_1 + 10\lambda_2 + 18\lambda_3$, with $\lambda_1, \lambda_2, \lambda_3 \geq 0$ and $\lambda \neq 0$). However, the alternative y^4 is the most balanced alternative among the different criteria and is a good candidate for a decision maker who prefers well-balanced alternatives.

Other models could be used to handle this problem [1, 2] like outranking methods (ranking of the alternatives based on pairwise comparisons) [3], additive value function models [4] or more evolved aggregation functions like the weighted minimum, weighted maximum, ordered weighted average operators (OWA) [5], weighted ordered weighted averaging operator (WOWA) [6], Choquet integral [7], etc.

In this paper, we will focus on the Choquet integral.

The Choquet integral is a powerful tool in multicriteria decision making and decision under uncertainty [8–10]. A Choquet integral can be seen as an integral on a non-additive measure (or capacity or fuzzy measure). It presents extremely wide expressive capabilities and can model many specific aggregation operators, including, but not limited to, the weighted sum, the minimum, the maximum, all the statistic quantiles, OWA, WOWA, etc. The Choquet integral can also be used in the additive value function model instead of the weighted sum [11]. In this case, the performance vector is replaced by marginal utility values.

However, this high expressiveness capability has a price: while the definition of a simple weighted sum operator with m criteria requires $m-1$ parameters, the definition of the Choquet integral with m criteria requires setting of $2^m - 2$ values, which can be a problem even for low values of m.

Given the effort needed to set the parameters of the Choquet integral comparing to the weighted sum, we will measure in this paper the interest of using the Choquet integral instead of the weighted sum. More precisely, for different alternatives evaluated with m criteria, we will evaluate the probability that an alternative that optimizes a defined Choquet integral could not be obtained with a simple weighted sum. This is particularly important in the general context where the alternatives to compare are not explicitly given but are obtained from a multiobjective optimization problem. In different works [12–16], the authors define a Choquet integral and then search for an optimal solution according to the defined Choquet integral. Two difficulties are thus introduced: the elicitation of the Choquet integral and the optimization of the Choquet integral for the particular multiobjective problem studied. Given this complexity, it is worth studying the real strength of the Choquet integral and to see, if a simpler method (the weighted sum) could have been used to obtain the same optimal solution.

To our knowledge, only one group of authors have performed experiments to assess the powerfulness of the Choquet integral. In [17], Meyer and Pirlot compare the ability of related models to represent rankings of alternatives. They compare different aggregators, including the weighted sum, the Choquet integral and additive value functions. To do so, they randomly generate alternatives and define a ranking of the alternatives. Then they check if the models can represent the ranking. They show that the Choquet integral model can represent significantly more orders than the weighted sum and that the difference becomes quite large when the number of criteria is high. If their work can appear similar to our work, there are two important differences:

- They consider rankings of alternatives while we only check the ability of an aggregator to reach one optimal alternative.
- Given a set alternatives we will not pick up randomly a best alternative, as they do, but we will generate randomly a Choquet integral (with a uniform law), and check if the alternative optimizing the Choquet integral could not have been obtained with a weighted sum too. Therefore, a probability will be defined, according to the set of possible Choquet integrals, and not according to the set of possible alternatives. That gives a more general way to measure the strength of the Choquet integral since results defined independently from the set of alternatives considered could be given.

The paper is organized as follows: we first introduce the main notions of this paper: the multicriteria context and the two aggregation operators used (weighted sum and Choquet integral). In Sections 3 and 4, we present the main contributions of the paper: we expose how we have compared the weighted sum and the Choquet integral operators: the comparison is based on the probability to reach an alternative optimal for a Choquet integral but not for a weighted sum. In the results section (Section 5), lower and upper bounds are computed, according to the number of criteria considered, and independently from the problem studied. We will see that this probability tends to one according to the number of criteria considered. We expose then some experimental results on randomly generated data sets. We will see that some conditions have to be respected to attain high probability values.

2 Definitions

We consider a general model with a set $\mathcal{Y}$ of n alternatives $\{y^1, \ldots, y^n\}$ evaluated with a set $\mathcal{M}$ of m criteria $\{1, \ldots, m\}$. The performance vector associated to an alternative y^i is denoted $(y^i_1, \ldots, y^i_m)$. We will consider w.l.o.g that the performance values are in the $[0, 1]$ interval and that the criteria have to be maximized.

The representation of an alternative in the criteria space is called a point and these two notions will be considered as equivalent in the rest of the paper.

We first recall the notion of Pareto dominance.

Definition 1. *The Pareto dominance relation (P-dominance for short) is defined, for all $y^1, y^2 \in \mathbb{R}^m$, by:*

$$y^1 \succ_P y^2 \iff [\forall k \in \mathcal{M}, y^1_k \geq y^2_k \text{ and } y^1 \neq y^2]$$

We will only work with sets $\mathcal{Y}$ of Pareto non-dominated alternatives, that is $\forall y^1 \in \mathcal{Y},\ \nexists\, y^2 \in \mathcal{Y} \mid y^2 \succ_P y^1$.

2.1 Weighted sum

The most popular aggregation operator is the weighted sum (WS), where positive importance weights $\lambda_i (i = 1, \ldots, m)$ are allocated to the criteria.

Definition 2. *Given a vector* $y \in \mathbb{R}^m$ *and a weight set* $\lambda \in \mathbb{R}^m$ *(with* $\lambda_i > 0$ *and* $\sum_{i=1}^m \lambda_i = 1$*), the WS* $f_\lambda^{ws}(y)$ *of* y *is equal to:*

$$f_\lambda^{ws}(y) = \sum_{i=1}^{m} \lambda_i y_i$$

In a set of $\mathcal{Y}$ of Pareto non-dominated alternatives, the alternatives that optimize a WS are called WS-optimal alternatives or supported Pareto-optimal alternatives [18] (SP alternatives). Note that there could exist alternatives that do not optimize a WS, and they are generally called non-supported Pareto-optimal alternatives (N-SP alternatives).

2.2 Choquet integral

The Choquet integral has been introduced by Choquet [7] in 1953 and has been intensively studied, especially in the field of multicriteria decision analysis, by several authors (see [9, 10, 19] for a brief review). Lately, the Choquet integral has also been used in the AI field, for classification problems [20, 21], constraint programming [22] or state space search [23].

We first define the notion of capacity, on which the Choquet integral is based.

Definition 3. *A capacity is a set function* $v\colon 2^{\mathcal{M}} \to [0, 1]$ *such that:*

- $v(\emptyset) = 0$, $v(\mathcal{M}) = 1$ *(boundary conditions)*
- $\forall \mathcal{A}, \mathcal{B} \in 2^{\mathcal{M}}$ *such that* $\mathcal{A} \subseteq \mathcal{B}, v(\mathcal{A}) \leq v(\mathcal{B})$ *(monotonicity conditions)*

Therefore, for each subset of criteria $\mathcal{A} \subseteq \mathcal{M}$, $v(\mathcal{A})$ represents the importance of the subset $\mathcal{A}$.

Definition 4. *A capacity is said additive if for each subset* $\mathcal{A}, \mathcal{B} \subseteq \mathcal{M}, v(\mathcal{A} \cup \mathcal{B}) = v(\mathcal{A}) + v(\mathcal{B})$.

Definition 5. *The Choquet integral of a vector* $y \in \mathbb{R}^m$ *with respect to a capacity* v *is defined by:*

$$\begin{aligned} f_v^C(y) &= \sum_{i=1}^{m} \big(v(Y_i^{\uparrow}) - v(Y_{i+1}^{\uparrow})\big) y_i^{\uparrow} \\ &= \sum_{i=1}^{m} (y_i^{\uparrow} - y_{i-1}^{\uparrow}) v(Y_i^{\uparrow}) \end{aligned}$$

where $y^{\uparrow} = (y_1^{\uparrow}, \ldots, y_m^{\uparrow})$ *is a permutation of the components of* y *such that* $0 = y_0^{\uparrow} \leq y_1^{\uparrow} \leq \ldots \leq y_m^{\uparrow}$ *and* $Y_i^{\uparrow} = \{j \in \mathcal{M}, y_j \geq y_i^{\uparrow}\} = \{i^{\uparrow}, (i+1)^{\uparrow}, \ldots, m^{\uparrow}\}$ *for* $i \leq m$ *and* $Y_{(m+1)}^{\uparrow} = \emptyset$.

The Choquet integral is a versatile aggregation operator, as it can express preferences to a wider set of solutions than a weighted sum, through the use of a non-additive capacity. For example, the Choquet integral can attain N-SP alternatives, while it is impossible with the weighted sum [24].

Example 1. Let us consider the four alternatives exposed in the introduction and the following capacity: $v(\{1\}) = v(\{2\}) = v(\{3\}) = 0.2$ and $v(\{1,2\}) = v(\{1,3\}) = v(\{2,3\}) = 0.4$. We obtain $f_v^C(y^1) = 10 + (10-10) * v(\{1,2\}) + (18-10) * v(\{1\}) = 11.6$, $f_v^C(y^2) = 10 + (10-10) * v(\{2,3\}) + (18-10) * v(\{2\}) = 11.6$, $f_v^C(y^3) = 10 + (10-10) * v(\{1,3\}) + (18-10) * v(\{3\}) = 11.6$, $f_v^C(y^4) = 11 + (12-11) * v(\{1,2\}) + (14-12) * v(\{1\}) = 11.8$. For this capacity, y^4 is thus the best alternative.

In this example, we see that the alternative y^4 can optimize a Choquet integral, but cannot optimize a WS. An alternative presenting this property will be called an *exclusive* Choquet optimal (C-optimal) alternative.

In the following, we will define more precisely the notion of exclusive C-optimal alternative.

3 Exclusive C-optimal alternatives

We first define the notion of WS-optimal set.

Definition 6. *Given a set $\mathcal{Y}$ of alternatives, the WS-optimal set, called $\mathcal{Y}_{ws}$, is the set containing an optimal alternative, for each possible WS, that is $\forall \lambda \in \mathcal{L}, \exists\, y^j \in \mathcal{Y}_{ws} \,|\, f_\lambda^{ws}(y^j) \geq f_\lambda^{ws}(y^i)\ \forall y^i \in \mathcal{Y}$, where $\mathcal{L}$ represents the set of possible weights defined over m criteria.*

Similarly, we can define the notion of Choquet-optimal set (C-optimal set).

Definition 7. *Given a set $\mathcal{Y}$ of alternatives, the C-optimal set, called $\mathcal{Y}_C$, is the set containing an optimal alternative, for each possible Choquet integral, that is $\forall v \in \mathcal{V}, \exists\, y^j \in \mathcal{Y}_C \,|\, f_v^C(y^j) \geq f_v^C(y^i)\ \forall y^i \in \mathcal{Y}$, where $\mathcal{V}$ represents the set of possible capacities defined over m criteria.*

The C-optimal set contains thus all potential C-optimal alternatives. A characterization of the C-optimal set has been proposed in [25]. We briefly recall it here.

Let σ be a permutation on $\mathcal{M}$. Let $\mathcal{O}_\sigma$ be the subset of alternatives $y \in \mathbb{R}^m$ such that $y \in \mathcal{O}_\sigma \iff y_{\sigma_1} \geq y_{\sigma_2} \geq \ldots \geq y_{\sigma_m}$.

Let $a_{\mathcal{O}_\sigma}$ be the following application:

$$a_{\mathcal{O}_\sigma} : \mathbb{R}^m \to \mathbb{R}^m, (a_{\mathcal{O}_\sigma}(y))_{\sigma_i} = (\min(y_{\sigma_1}, \ldots, y_{\sigma_i})), \forall i \in \mathcal{M}$$

For example, if $m = 3$, for the permutation (2,3,1), we have:

$$a_{\mathcal{O}_\sigma}(y) = \big(\min(y_2, y_3, y_1), \min(y_2), \min(y_2, y_3)\big)$$

We denote by $\mathcal{A}_{\mathcal{O}_\sigma}(\mathcal{Y})$ the set containing the alternatives obtained by applying the application $a_{\mathcal{O}_\sigma}(y)$ to all the alternatives $y \in \mathcal{Y}$. As $(a_{\mathcal{O}_\sigma}(y))_{\sigma_1} \geq (a_{\mathcal{O}_\sigma}(y))_{\sigma_2} \geq \ldots \geq (a_{\mathcal{O}_\sigma}(y))_{\sigma_m}$, we have $\mathcal{A}_{\mathcal{O}_\sigma}(\mathcal{Y}) \subseteq \mathcal{O}_\sigma$.

In the following, we will denote $\mathcal{O}_\sigma$ as simply $\mathcal{O}$ for the sake of simplicity, and we will consider, w.l.o.g., that the permutation σ is equal to $(1, 2, \dots, m)$, that is $y \in \mathcal{O} \Leftrightarrow y_1 \geq y_2 \geq \cdots \geq y_m$.

In [25], Lust and Rolland show the following results in order to characterize the C-optimal set $\mathcal{Y}_C$ of a set $\mathcal{Y}$:

Theorem 1.

$$\mathcal{Y}_C \cap \mathcal{O}_\sigma = \mathcal{Y} \cap \mathcal{Y}_{ws}(\mathcal{A}_{\mathcal{O}_\sigma}(\mathcal{Y}))$$

where $\mathcal{Y}_{ws}(\mathcal{A}_{\mathcal{O}_\sigma}(\mathcal{Y}))$ designs the set of WS-optimal alternatives of the set $\mathcal{A}_{\mathcal{O}_\sigma}(\mathcal{Y})$.

This theorem characterizes the alternatives that can be C-optimal in a set $\mathcal{Y}$ of points as being, in each subspace of the criteria space $\mathcal{Y}$ where $y_{\sigma_1} \geq y_{\sigma_2} \geq \ldots \geq y_{\sigma_m}$, the WS-optimal points in $\mathcal{A}_{\mathcal{O}_\sigma}(\mathcal{Y})$.

Proof: see [25].

Property 1. $\mathcal{Y}_{ws} \subseteq \mathcal{Y}_C$

Proof. If the capacity v is additive, the Choquet integral of a vector $y \in \mathbb{R}^m$ is equal to $\sum_{i=1}^m \left(v(Y_i^\uparrow) - v(Y_{i+1}^\uparrow)\right)y_i^\uparrow = \sum_{i=1}^m v(\{i\})y^i$. Therefore, the WS is a particular Choquet integral for which the capacity is additive. All WS-optimal alternatives are thus also C-optimal alternatives.

Example 2. For the four alternatives of the introduction, we have $\mathcal{Y}_{ws} = \{y^1, y^2, y^3\}$ and $\mathcal{Y}_C = \{y^1, y^2, y^3, y^4\}$.

Definition 8. *Given a set $\mathcal{Y}$ of alternatives, the exclusive C-optimal set $\mathcal{Y}_{eC}$ is equal to $\{\mathcal{Y}_C \setminus \mathcal{Y}_{ws}\}$.*

The set $\mathcal{Y}_{eC}$ is thus composed of the alternatives that optimize a Choquet integral, without optimizing a WS.

Definition 9. *Given a set $\mathcal{Y}$ of alternatives, let us consider the exclusive C-optimal set $\mathcal{Y}_{eC}$. For all alternatives $y^i \in \mathcal{Y}_{eC}$, let $\mathcal{V}^i$ the set of capacities for which the alternative $y^i \in \mathcal{Y}_{eC}$ is C-optimal in $\mathcal{Y}$. Let $\mathcal{V}_e = \bigcup \mathcal{V}^i$ the union of these sets. The set $\mathcal{V}_e$ is called the exclusive capacity set. All exclusive C-optimal alternatives are optimal for capacities $v \in \mathcal{V}_e$ (it does not exist an exclusive C-optimal alternative for a capacity $v \notin \mathcal{V}_e$).*

We can now define the probability p to get an exclusive C-optimal alternative when a Choquet integral is randomly generated (that is a capacity v is randomly generated, with a uniform law).

Definition 10. *Let $v \in \mathcal{V}$ a capacity randomly generated with a uniform law. Let $F(v) = 1$ if $v \in \mathcal{V}^e$ and 0 otherwise and y the best alternative for the Choquet integral f_v^C. The probability p_e that y is exclusive C-optimal is equal to:*

$$p_e = \int_{\mathcal{V}} F(v)dv$$

Example 3. Let $m = 2$ and $\mathcal{Y} = \{(1,0),(0,1),(0.2,0.7)\}$. The point $(0.2, 0.7)$ cannot be WS-optimal as $(0.2 + 0.7 < 1)$. But it can be C-optimal if the capacity v respects the following conditions[1]: $0.2 + 0.5v_2 \geq v_2$ and $0.2 + 0.5v_2 \geq v_1$, that is v such that $v_1 \leq 0.2 + 0.5v_2$ and $v_2 \leq 0.4$. Therefore we have $\mathcal{Y}_{ws} = \{(1,0),(0,1)\}, \mathcal{Y}_C = \mathcal{Y}$ and $\mathcal{Y}_{eC} = \{(0.2, 0.7)\}$. If we generate randomly a capacity v, the probability p_e to get the point $(0.2, 0.7)$ optimal for the defined Choquet integral is thus equal to $\int_0^{0.4} \int_0^{0.2+0.5v_2} 1dv1\, dv2 = 0.12$. We see also through this example that the probability takes a different value compared to the probability to get an exclusive C-optimal alternative if the alternative is randomly selected (which is equal to 1/3) , as done in the work of Meyer and Pirlot [17].

4 Maximal proportion of exclusive C-optimal alternatives

The value of the probability p_e is problem-dependent, but bounds, according to the number of criteria (and independent from the problem studied), can however be generated. The minimal value of p_e is zero, whatever the number the criteria, since it is always possible to generate sets composed of only WS-optimal alternatives. The maximal value of p_e is more difficult to compute, but also more interesting: given a problem with m criteria, what is the maximal value of p_e, that is the maximal chance to reach an exclusive C-optimal alternative with a Choquet integral?

In the following, we will construct artificial sets of alternatives, in order to estimate the maximal value of p_e. The WS-optimal set of the artificial sets will always have the same form and composed of only m points, in order to favor the Choquet integral. We will use m I^j points ($j \in \mathcal{M}$), such that $I^j_i = 1$ if $i = j$ and 0 otherwise ($i \in \mathcal{M}$). Such a WS-optimal set will be called $\mathcal{Y}^I_{ws}$. For example, for $m = 3$, $\mathcal{Y}^I_{ws} = \{(1,0,0),(0,1,0),(0,0,1)\}$.

We first introduce some new definitions.

Definition 11. *Let $\mathcal{Y}$ a set of alternatives. Let us consider two additional alternatives y^1 and y^2 and $\mathcal{V}_1$ and $\mathcal{V}_2$ the set of capacities for which y^1 and y^2 are C-optimal in $\mathcal{Y}$. We say that y^1 covers y^2 if $\mathcal{V}_2 \subseteq \mathcal{V}_1$.*

Definition 12. *Let $\mathcal{Y}$ a set of alternatives with $\mathcal{Y}_{ws} = \mathcal{Y}^I_{ws}$, $\mathcal{Y}_{eC}$ the exclusive C-optimal set and $\mathcal{V}_e$ the exclusive capacity set. The set $\mathcal{V}_e$ is called a maximal exclusive capacity set (denoted $\mathcal{V}^*_e$) if any other possible exclusive C-optimal alternatives (not necessarily in $\mathcal{Y}$) are covered by an alternative of $\mathcal{Y}_{eC}$. The set $\mathcal{Y}_{eC}$ associated with $\mathcal{V}^*_e$ is called a maximal cover set and is denoted $\mathcal{Y}^*_{eC}$.*

For a problem with m criteria, if we can generate $\mathcal{V}^*_e$, we can then compute the maximal value of p_e as follows:

[1] In the following, a capacity associated to a set will be simply written $v_{e_1,\ldots,e_n}$ where $e_1, \ldots, e_n$ denotes the elements belonging to the set.

Definition 13. *Let $v \in \mathcal{V}$ a capacity randomly generated with a uniform law. Let $G(v) = 1$ if $v \in \mathcal{V}_e^*$ and 0 otherwise and y the best alternative for the Choquet integral f_v^C. The maximal probability p_e^* that y is exclusive C-optimal is equal to:*

$$p_e^* = \int_{\mathcal{V}} G(v)dv$$

For a problem with m criteria, we have thus to generate $\mathcal{V}_e^*$ to obtain p_e^*.

We will start our study with alternatives evaluated with only two criteria and determine, in this case, the maximal value that p_e can take. We will then generalize for any number of criteria.

4.1 Two-criteria problems

We will try to generate $\mathcal{Y}_{eC}^*$ and $\mathcal{V}_e^*$, in order to express p_e^*. We have thus to generate a set $\mathcal{Y}$ of alternatives such that any other possible exclusive C-optimal alternatives will be covered by an alternative of $\mathcal{Y}_{eC}^*$. We first detail the results obtained with the simple case where $m = 2$.

Property 2. Let $m = 2$ and $\mathcal{Y}$ a set of alternatives with $\mathcal{Y}_{ws} = \mathcal{Y}_{ws}^I = \{(1, 0), (0, 1)\}$. The maximal exclusive capacity set $\mathcal{V}_e$ is equal to $\{(v_1, v_2) \,|\, v_1 < \frac{1}{2}$ and $v_2 < \frac{1}{2}\}$ and the maximal cover set $\mathcal{Y}_{eC}^*$ is equal to $\simeq (\frac{1}{2}, \frac{1}{2})$, where the notation $\simeq (\frac{1}{2}, \frac{1}{2})$ means that we have a point close to $(\frac{1}{2}, \frac{1}{2})$ (but the sum of its component is less than 1).

Proof. With two criteria, we have $\mathcal{V} = \{v_1, v_2\}$ with $v_1 \in [0, 1]$ and $v_2 \in [0, 1]$. We need at least three points to have one exclusive C-optimal alternative. Let us consider the two points of $\mathcal{Y}_{ws}^I$ ((1, 0) and (0, 1)) and the point (a, b) such that $a \geq b$ and $a + b < 1$. We have $f_v^C(1, 0) = v_1$, $f_v^C(0, 1) = v_2$ and $f_v^C(a, b) = b + (a - b)v_1$. To have (a, b) C-optimal we need to have $b + (a - b)v_1 \geq v_1$ and $b + (a - b)v_1 \geq v_2$, that is $v_1 \leq \frac{b}{1-a+b}$ and $v_2 \leq b + (a - b)v_1$.

To maximize $f_v^C(a, b)$, $a + b \to 1$, and thus $b \to 1 - a$. We obtain:

$$\begin{cases} v_1 < \frac{1-a}{1-a+1-a} = \frac{1}{2} \\ v_2 < (1 - a) + (2a - 1)v_1 \end{cases}$$

Let

$$\begin{cases} G(v) = 1 \text{ if } v_1 < \frac{1}{2} \text{ and } v_2 < (1 - a) + (2a - 1)v_1 \\ \quad\quad\;\; = 0 \text{ otherwise.} \end{cases}$$

We get:

$$\int_{\mathcal{V}} G(v)dv = \int_0^{\frac{1}{2}} \int_0^{(1-a)+(2a-1)v_1} 1 \; dv_2 \, dv_1 = \frac{3}{8} - \frac{1}{4}a$$

As $a \in [\frac{1}{2}, 1]$, the maximal value is obtained when $a = \frac{1}{2}$ and $b \to \frac{1}{2}$.

We see that it is enough to consider only one additional point $(a, b) \to (\frac{1}{2}, \frac{1}{2})$. Indeed, if we consider another points, we will still have $v_1 < \frac{1}{2}$ and $v_2 < (1 - a) + (2a - 1)v_1$ that is $v_2 < \frac{1}{2}$ when $v_1 \to \frac{1}{2}$. Therefore the maximal value of $\int_{\mathcal{V}} G(v)dv$ will still be bounded by $\frac{1}{4}$ $(\frac{1}{2} * \frac{1}{2})$; the value of p_e^* is thus equal to $\frac{1}{4}$.

Consequently, for two-criteria sets, $\mathcal{V}_e^* = \{(v_1, v_2) \,|\, v_1 < \frac{1}{2}$ and $v_2 < \frac{1}{2}\}$, $\mathcal{Y}_{eC}^* = \{\simeq (\frac{1}{2}, \frac{1}{2})\}$ and $p_e^* = \frac{1}{4}$.

4.2 m-criteria problems

For three criteria, we have followed the same reasoning as for two-criteria problems, by adding points (a, b, c) such that $a+b+c < 1$, in an initial set composed of the three points of $\mathcal{Y}^I_{ws}$ ((1,0,0), (0,1,0) and (0,0,1)). We have obtained $\mathcal{V}^*_e = \{(v_1, v_2, v_3, v_{12}, v_{13}, v_{23}) \,|\, (v_1 < \frac{1}{3}, v_2 < \frac{1}{3}, v_3 < \frac{1}{3})$ or $(v_1 < \frac{1}{2}v_{12}, v_2 < \frac{1}{2}v_{12}, v_3 < \frac{1}{2}v_{12})$ or $(v_1 < \frac{1}{2}v_{13}, v_2 < \frac{1}{2}v_{13}, v_3 < \frac{1}{2}v_{13})$ or $(v_1 < \frac{1}{2}v_{23}, v_2 < \frac{1}{2}v_{23}, v_3 < \frac{1}{2}v_{23})\}$, $\mathcal{Y}^*_{eC} = \{\simeq (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), \simeq (\frac{1}{2}, \frac{1}{2}, 0), \simeq (\frac{1}{2}, 0, \frac{1}{2}), \simeq (0, \frac{1}{2}, \frac{1}{2})\}$.

By solving analytically different sextuples integrals defined on $\mathcal{V}^*_e$, we have computed p^*_e and obtained a value equal to $\frac{20323}{6^6} = 0.4356$.

Generally, for m criteria, we have the following property:

Property 3. Let $\mathcal{Y}$ a set of alternatives with $\mathcal{Y}_{ws} = \mathcal{Y}^I_{ws}$. The maximal cover set $\mathcal{Y}^*_{eC}$ for a number m of criteria is composed of the following points: $\simeq (\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m})$, all permutations of $\simeq (0, \frac{1}{m-1}, \frac{1}{m-1}, \ldots, \frac{1}{m-1})$, all permutations of $\simeq (0, 0, \frac{1}{m-2}, \frac{1}{m-2}, \ldots, \frac{1}{m-2})$, ..., and all permutations of $\simeq (0, \ldots, 0, \frac{1}{2}, \frac{1}{2})$.

Proof. We have to show that any additional exclusive C-optimal point in $[0,1]^m$ will be covered by a point of $\mathcal{Y}^*_{eC}$. In Property 3, we have that $\mathcal{Y}$ is composed of $\mathcal{Y}^I_{ws}$ and $\mathcal{Y}^*_{eC}$. Let us consider an additional alternative z in $\mathcal{O}$, that is $z_1 \geq z_2 \geq \cdots \geq z_m$. As $\mathcal{Y}_{ws}$ is composed of the I^j points, to be an exclusive C-optimal point, z has to fulfill the following constraint: $\sum_{i=1}^m z_i < 1$ (otherwise that point would be a WS-optimal point). Moreover, to be C-optimal, z has to be WS-optimal in $\mathcal{A}_{\mathcal{O}}(\mathcal{Y})$, where $\mathcal{A}_{\mathcal{O}}(\mathcal{Y})$ is obtained by applying the application $(a_{\mathcal{O}}(y))_i = (\min(y_1, \ldots, y_i)), \forall i \in \mathcal{M}$, to all $y \in \mathcal{Y}$ (see Theorem 1). $\mathcal{A}_{\mathcal{O}}(\mathcal{Y})$ is thus equal to $\{\simeq (\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}), \simeq (\frac{1}{m-1}, \ldots, \frac{1}{m-1}, 0), \simeq (\frac{1}{m-2}, \ldots, \frac{1}{m-2}, 0, 0), \ldots, \simeq (\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0), (1, 0, \ldots, 0)\}$ (the point $(0, \ldots, 0)$ belongs also to this set but it can be removed since this point is P-dominated in $\mathcal{A}_{\mathcal{O}}(\mathcal{Y})$). We can remark that the vertices of the polyhedron defined by the constraint inequalities $(z_1 \geq z_2 \geq \cdots \geq z_m)$ and $(\sum_{i=1}^m z_i \leq 1)$ are the points $\{(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}), (\frac{1}{m-1}, \ldots, \frac{1}{m-1}, 0), (\frac{1}{m-2}, \ldots, \frac{1}{m-2}, 0, 0), \ldots, (\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0), (1, 0, \ldots, 0)\}$. Therefore, if the points of $\mathcal{A}_{\mathcal{O}}(\mathcal{Y})$ are sufficiently close to the vertices of the polyhedron, it will not possible to find a point z in $\mathcal{A}_{\mathcal{O}}(\mathcal{Y})$ which is WS-optimal [26]. Any additional points that do not optimize a WS will not be C-optimal and thus covered by a point of $\mathcal{Y}^*_{eC}$. Therefore $\mathcal{Y}^*_{eC}$ is a maximal cover set.

5 Results

5.1 Computation of $p^*_e(m)$

In this section, we estimate the probability p^*_e that some alternative might be exclusively Choquet-optimal, as the number of criteria varies. That is, we estimate $p^*_e(m)$ for varying m. We have already analytically determined this probability for the cases of 2 and 3 criteria ($p^*_e(2) = 0.25$ and $p^*_e(3) = 0.4356$); but for larger number of criteria we turn to numerical estimation, because analytically solving the integrals is exponential in the number of criteria. To do so, roughly speaking, we randomly generate candidate Choquet integrals with a uniform distribution, and count the number of times that some alternative of the maximal cover set $\mathcal{Y}^*_{eC}$ (see Property 3) is superior to some alternative of the initial set of alternatives ($\mathcal{Y}^I_{ws}$). Dividing that by the number of samples gives our numerical estimate.

However, because of the monotonicity constraints, it is not trivial to generate randomly (with a uniform distribution) capacities. One way to deal with this problem is to generate $(2^m - 2)$ random values between 0 and 1, and to check if the values respect the monotonicity constraints of a capacity. If it works for three or four criteria, it is quickly unusable since the monotonicity constraints become harder to satisfy when the number of criteria increases.

Recently, Combarro *et al.* [27] have proposed a way to generate randomly capacities. They established that the random generation of capacities involves the generation of random linear extensions of capacities (that is linear extensions of the monotonocity constraints). Once a linear extension is generated, it is enough to compute a point that respects the linear extension, which can be easily done (see [28]). A method to generate linear extensions in a random way appears in [29]. However, this procedure implies the generation of graphs (lattice of ideals) whose the number of vertices increases according to the sequence of numbers defined by Dedekind [30]. This procedure can only be applied to generate capacities until $m = 5$ [27].

So, for greater number of criteria ($m > 5$), we exploit a heuristic method. We have used the Markov chain Monte Carlo (MCMC) method introduced in [31]. The method works with iterative modifications of a starting admissible linear extension. It has been shown that this procedure evolves in limit to a uniform linear extension [32], no matter the initial linear extension.

The results obtained are shown in Figure 1. We vary the number of criteria m from 2 to 8. For $m \geq 4$ we have used random generations, with 1000000 trials.

We see that if $p^*_e(m)$ was quite low for $m = 2$ and $m = 3$, the value of $p^*_e(m)$ increases rapidly with m. We have $p^*_e(4) = 0.659$, $p^*_e(5) = 0.868$, $p^*_e(6) = 0.977$, $p^*_e(7) = 0.997$ and for $p = 8$, we are really close to 1 ($p^*_e(8) = 0.999$).

These results testify of the strength of the Choquet integral and its ability to attain alternatives that are not possible to reach with the WS, especially when the number of criteria increases.

However these results have been obtained for the "best possible data set" for the Choquet integral, composed of well-located exclusive C-optimal points and presenting few WS-optimal points (the m points I^j).

We study in the next section how the probability p_e to get exclusive Choquet optimal solutions evolves when the number of WS-optimal points increases.

5.2 Influence of the number of WS-optimal points

We will now increase the number of WS-optimal points (or supported points) and see how p_e evolves. We have generated sets of alternatives composed of n_{SP} supported points and n_{NSP} non-supported points. We first generate $\mathcal{Y}^I_{ws}$

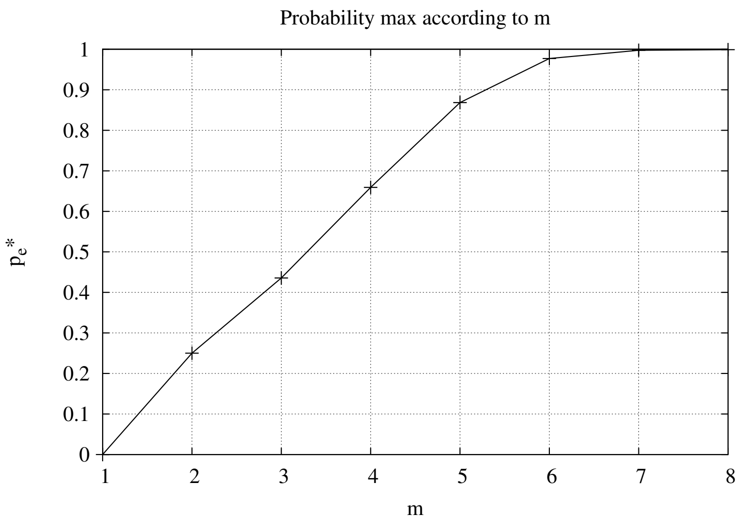

Fig. 1: Evolution of p_e^* according to the number of criteria.

(all the m points of $\mathcal{Y}^I_{ws}$ are supported). We add $(n_{SP} - m)$ points by generating randomly different points y in $[0,1]^m$. We add the following constraint: $\sum_{i=1}^{m} y_i^2 < 1$ (otherwise all the points will be closed to the point (1,. . . ,1)). We check with linear programming that the point optimizes a WS. If not, we try another point, until $n_{SP} - m$ points have been generated. Then, we generate n_{NSP} non-supported points. A point will be added in the set if it does not optimize a WS (also checked with linear programming) and if it is a non-dominated point (and if the point dominates another non-supported points, these points are removed from the set). The procedure is repeated until n_{NSP} non-supported points are produced. An example of a set obtained, for two criteria, and for $n_{SP} = 22$ and $n_{NSP} = 20$ is represented in Figure 2.

The results for $m = 2$ to $m = 6$ are given in Figures 3, 4, 5, 6 and 7. We vary the number of n_{SP} points and for each value of n_{SP}, we also vary the number of n_{NSP} points. For each combination (n_{SP},n_{NSP}), 100 different sets are randomly produced and we average the probability p_e over these sets.

For $m = 2$ and $m = 3$, adding SP alternatives decreases p_e, but if enough N-SP alternatives are added (about 400), the decreasing remains reasonable (from 0.25 to 0.2 for $m = 2$ and from 0.42 to 0.3 for $m = 3$). Quite surprisingly, the decreasing of p_e is higher for $m = 4$, $m = 5$ and $m = 6$. Adding only few more SP alternatives reduces p_e of about 50%. The results are quite impressive for $m = 6$: for 6 SP alternatives and 400 N-SP alternatives, we have p_e equal to 0.91. If we add only one SP alternatives, p_e drops to 0.43.

This phenomenon can be explained by the fact that the Choquet integral is attracted by particular points, that is the points that composed the set $\mathcal{Y}^*_{eC}$.

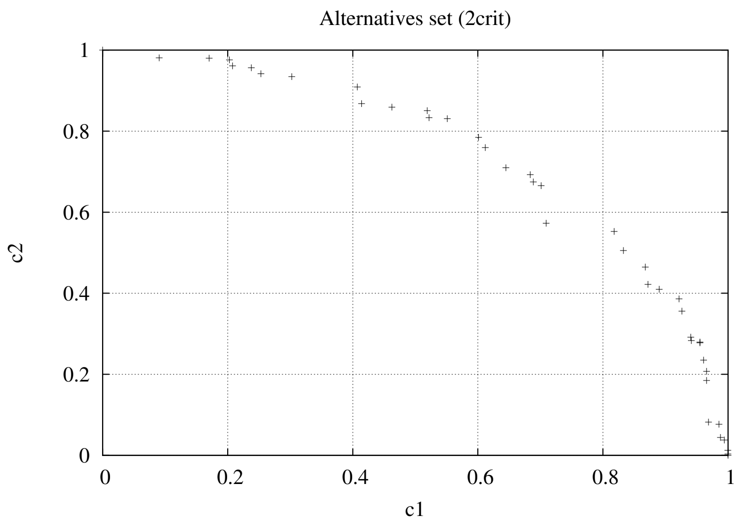

Fig. 2: Example of alternatives set obtained for $n_{SP} = 22$ (supported points) and $n_{NSP} = 20$ (non-supported points) (2 criteria).

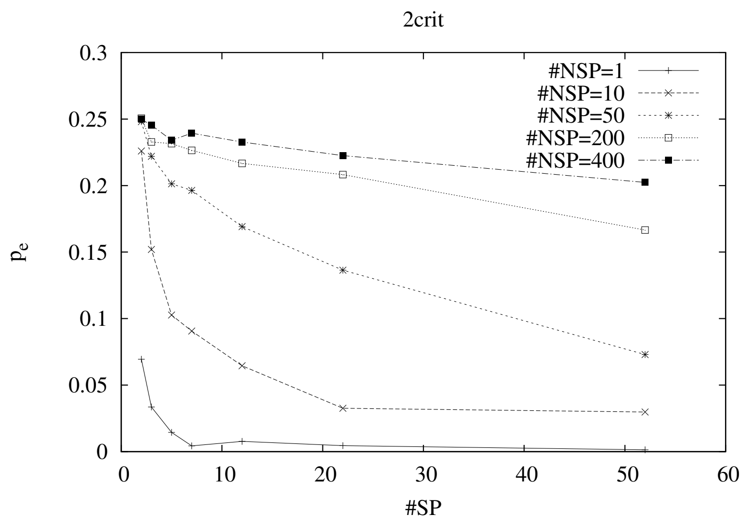

Fig. 3: 2 criteria: p_e according to n_{SP} (supported points) and n_{NSP} (non-supported points).

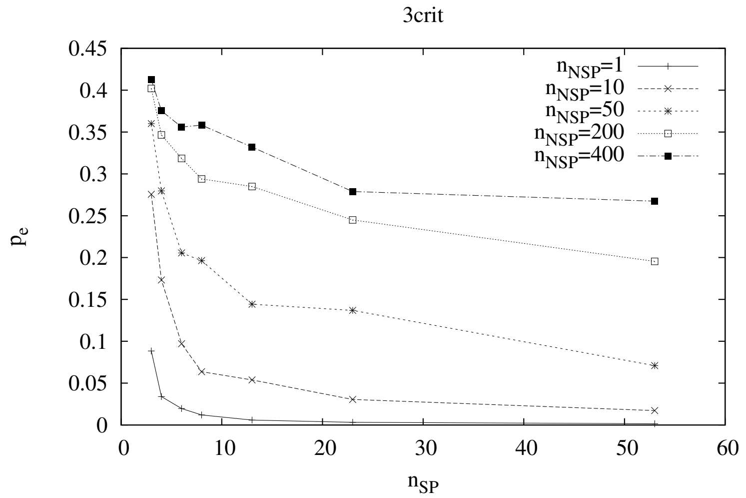

Fig. 4: 3 criteria: p_e according to n_{SP} (supported points) and n_{NSP} (non-supported points).

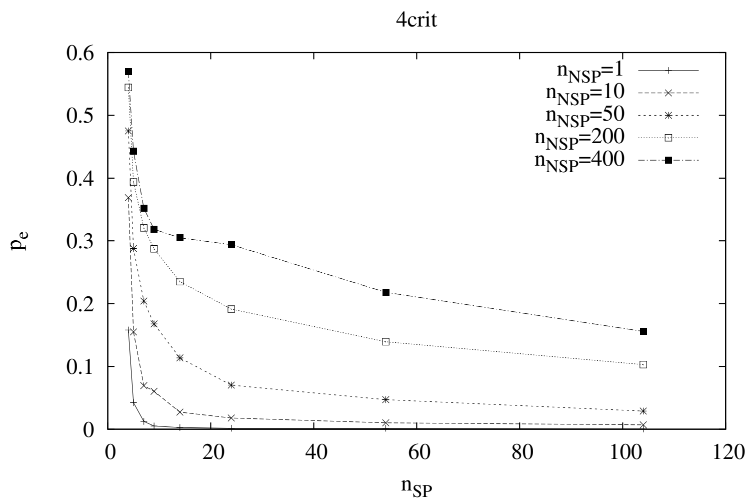

Fig. 5: 4 criteria: p_e according to n_{SP} (supported points) and n_{NSP} (non-supported points).

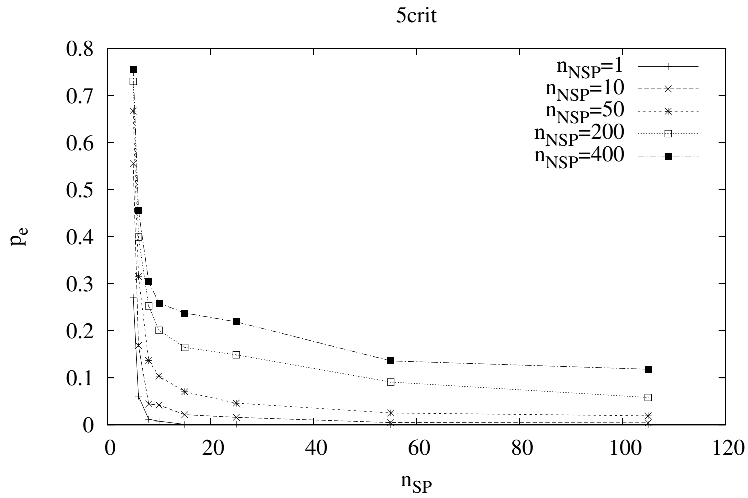

Fig. 6: 5 criteria: p_e according to n_{SP} (supported points) and n_{NSP} (non-supported points).

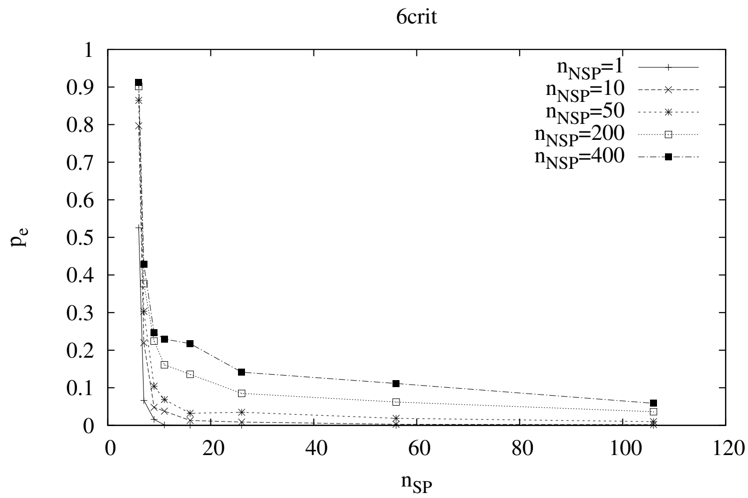

Fig. 7: 6 criteria: p_e according to n_{SP} (supported points) and n_{NSP} (non-supported points).

When the number of criteria m increases, it could be more difficult to have alternatives close to these points and the WS becomes a tougher opponent for the Choquet integral.

Similar results, not reported here, have been obtained for other constraint functions to produce sets of alternatives ($\sum_{i=1}^{m} y_i^{\alpha} < 1$ with $1 \leq \alpha \leq 4$) and for alternatives representing solutions of the multiobjective knapsack problem or the multiobjective traveling salesman problem.

Therefore, the Choquet integral should be carefully used if the set of alternatives present supported alternatives, at least four criteria, and not too many non-supported alternatives.

6 Conclusion

We have proposed in this work a comparison of two successful and popular aggregation operators: the weighted sum and the Choquet integral. If it is clear than the Choquet integral is more powerful than the weighted sum, given the hard work that the elicitation of its parameters asks, it is still relevant to compare both operators and to measure the probability that the best alternative determined with a Choquet integral could not been obtained with a simple weighted sum. The results show that the maximal value that this probability can take is close to one, when the number of criteria is higher than four. However, to reach this high probability, particular data sets have been constructed, in favor of the Choquet integral. When the number of WS-optimal alternatives increases in a set, the probability decreases rapidly, and especially if the number of criteria is higher than four. This work opens many new perspectives: different operators could be compared using this framework, like WOWA and the Choquet integral, for example. Also, we could restrict the capacities of the Choquet integral to certain families (convex capacities, k-additive capacities, etc.) and to examine which families of capacities enable to mainly reach exclusive C-optimal alternatives. Finally, extending this study in a combinatorial space where the alternatives are not explicitly given would be worth studying.

References

1. Torra, V.: Aggregation operators and models. Fuzzy Sets and Systems **156**(3) (2005) 407–410
2. Beliakov, G., Pradera, A., Calvo, T.: Aggregation Functions: A Guide for Practitioners. Volume 221 of Studies in Fuzziness and Soft Computing. Springer (2007)
3. Bouyssou, D.: Outranking methods. In Floudas, C., Pardalos, P., eds.: Encyclopedia of optimization, Kluwer (2001)
4. Fishburn, P.: Utility Theory for Decision Making. Wiley, New York (1970)
5. Yager, R.: On ordered weighted averaging aggregation operators in multicriteria decision making. In: IEEE Trans. Systems, Man and Cybern. Volume 18. (1998) 183–190
6. Torra, V.: The weighted OWA operator. Int. J. Intell. Syst. (12) (1997) 153–166
7. Choquet, G.: Theory of capacities. Annales de l'Institut Fourier (5) (1953) 131–295
8. Chateauneuf, A.: Modeling attitudes towards uncertainty and risk through the use of choquet integral. Annals of Operations Research **52**(1) (1994) 1–20
9. Grabisch, M., Marichal, J.C., Mesiar, R., Pap, E.: Aggregation functions. Volume 127 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, UK (2009)
10. Grabisch, M., Labreuche, C.: A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. Annals OR **175**(1) (2010) 247–286
11. Grabisch, M., Labreuche, C.: Fuzzy measures and integrals in mcda. In: Multiple Criteria Decision Analysis: State of the Art Surveys. Volume 78 of International Series in Operations Research and Management Science. Springer New York (2005) 563–604
12. Galand, L., Perny, P., Spanjaard, O.: Choquet-based optimisation in multiobjective shortest path and spanning tree problems. European Journal of Operational Research **204**(2) (2010) 303–315
13. Galand, L., Perny, P., Spanjaard, O.: A branch and bound algorithm for Choquet optimization in multicriteria problems. In: Proceedings Lecture Notes in Economics and Mathematical Systems. Volume 634. (2011) 355–365
14. Fouchal, H., Gandibleux, X., Lehuédé, F.: Preferred solutions computed with a label setting algorithm based on Choquet integral for multi-objective shortest paths. In: 2011 IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making (MDCM), Paris, France, IEEE (2011) 143 – 150
15. Galand, L., Lesca, J., Perny, P.: Dominance rules for the Choquet integral in multiobjective dynamic programming. In: IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013. (2013)
16. Branke, J., Corrente, S., Grecco, S., Slowinski, R., Zielniewicz, P.: Using Choquet integral as preference model in interactive evolutionary multiobjective optimization. Technical report, Warwick Business School, Warwick, Uninted Kingdom (2014)
17. Meyer, P., Pirlot, M.: On the expressiveness of the additive value function and the Choquet integral models. In: DA2PL 2012 : from Multiple Criteria Decision Aid to Preference Learning. (2012) 48 – 56
18. Ehrgott, M.: Multicriteria Optimization. Second edition. Springer, Berlin (2005)
19. Grabisch, M.: The application of fuzzy integrals in multicriteria decision making. European Journal of Operational Research **89** (1996) 445–456

20. Tehrani, A., Cheng, W., Hüllermeier, E.: Choquistic regression: Generalizing logistic regression using the Choquet integral. In: In Proc. EUSFLAT2011, 7th Int. Conf. of the European Society for Fuzzy Logic and Technology. (2011) 868–875
21. Valls, A., Medina, C., Moreno, A., Puig, D.: Selection of gabor filters with Choquet integral for texture analysis in mammogram images. In Gibert, K., Botti, V.J., Bolao, R.R., eds.: CCIA. Volume 256 of Frontiers in Artificial Intelligence and Applications., IOS Press (2013) 67–76
22. Le Huédé, F., Grabisch, M., Labreuche, C., Savéant, P.: Integration and propagation of a multi-criteria decision making model in constraint programming. Journal of Heuristics **12**(4-5) (2006) 329–346
23. Galand, L., Perny, P.: Search for Choquet-optimal paths under uncertainty. In: in UAI07, 2007. (2007) 125–132
24. Lust, T., Rolland, A.: Choquet optimal set in biobjective combinatorial optimization. Computers & Operations Research **40**(10) (2013) 2260–2269
25. Lust, T., Rolland, A.: On the computation of Choquet optimal solutions in multicriteria decision contexts. In Ramanna, S., Lingras, P., Sombattheera, C., Krishna, A., eds.: MIWAI. Volume 8271 of Lecture Notes in Computer Science., Springer (2013) 131–142
26. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, New York, NY, USA (2004)
27. Combarro, E.F., DíAz, I., Miranda, P.: On random generation of fuzzy measures. Fuzzy Sets Syst. **228** (October 2013) 64–77
28. Rubinstein, R., Kroese, D.: Simulation and the Monte Carlo Method, 2nd Edition. Wiley (2008)
29. Loof, K.D., Meyer, H.D., Baets, B.D.: Exploiting the lattice of ideals representation of a poset. Fundam. Inf. **71**(2,3) (February 2006) 309–321
30. Dedekind, R.: Ueber zerlegungen von zahlen durch ihre grossten gemeinsamen teiler. Fest. Hoch. Braunschweig u. ges. Werke II (1897) 103–148
31. Karzanov, A., Khachiyan, L.: On the conductance of order Markov chains. Order **8**(1) (1991) 7–15
32. Bubley, R., Dyer, M.: Faster random generation of linear extensions. Discrete Mathematics **201** (1999) 81–88