24th European Conference on Artificial Intelligence - ECAI 2020
Santiago de Compostela, Spain

New Computational Models for the Choquet Integral

Hugo Martin and Patrice Perny[1]

Abstract. Multiobjective optimization is a central problem in a wide range of contexts, such as multi-agent optimization and multicriteria decision support or decision under risk and uncertainty. The presence of several objectives leads to multiple non-dominated solutions and requires the use of a sophisticated decision model allowing various attitudes towards preference aggregation. The Choquet Integral is one of the most expressive parameterized models introduced in decision theory to scalarize performance vectors and support decision making. However, its use in optimization contexts raises computational issues. This paper proposes new computational models based on mathematical programming to optimize the Choquet integral on implicit sets. A new linearization of the Choquet integral exploiting the vertices of the core of the convex capacity is proposed, combined with a constraint generation algorithm. Then the computational model is extended to the bipolar Choquet Integral to allow asymmetric aggregation with respect to a specific reference point.

1 INTRODUCTION

Preference aggregation is one of the key subjects in decision theory and one of the critical problems to deal with in the design of intelligent systems for decision support. In particular, the aggregation problem appears in various decision situations involving multiple agents, for example to define the social welfare from individual values in a collective decision-making problem, or to derive a recommendation to a user from others' preferences in collaborative recommender systems. Preference aggregation is also a central problem in multicriteria analysis, to obtain decision models that are more decisive than simple Pareto dominance. Aggregation operators are also used to remove uncertainty when the outcomes attached to possible decisions are imperfectly known and depend on various possible scenarios.

One standard preference aggregation method widely used in multiobjective optimization is based on the use of a *scalarizing function* transforming a vector of criterion values (reflecting the quality of the alternatives with respect to various point of views) into an overall utility or disutility [22, 19]. This allows the reformulation of any multicriteria choice problem as a single-objective optimization problem. In particular, when the objective are expressed as linear functions of decision variables, using a weighted sum in the aggregation step leads to a linear optimization problem that is generally much easier to solve than a non-linear one. Moreover, the use of a weighted combination provide some control on the profile of the solution by playing with the weighting coefficients.

However linear scalarizing functions suffer from some descriptive weaknesses, well-known in Decision Theory, that considerably reduce their interest for decision support. First, when optimizing on non-convex sets (e.g., in discrete optimization), none of the Pareto-optimal solutions that are located in the interior of the convex hull of the feasible area can be obtained by optimizing a weighted sum of the objectives. Hence the simple fact of using a linear aggregator may eliminate arbitrarily a significant part of the Pareto set from the set of possible winners even before the weights were chosen. Moreover, the use of a linear weighted means amounts to assuming that the notion of weight attached to the objectives under consideration (representing agents, criteria or scenarios) is additive (i.e., the weight attached to a subset of objectives is the sum of the weights of these objectives). This assumption is not always relevant. In cooperative game theory, it is well known that the strength of coalitions of agents is not always representable by an additive measure. Similarly, in multicriteria problems, non-additive importance measures are widely used to model positive or negative interactions among the criteria (e.g., Choquet capacities [8]). Finally, in decision under uncertainty, beliefs about possible events impacting the outcomes of the possible acts are not always additive.

The above considerations have contributed to the development of decision models based on non-linear aggregation functions allowing positive and negative synergies among components, and are at the origin of the introduction of Choquet integral in Decision Theory [20, 7]. The precise definition of the discrete Choquet integral will be recalled later in the paper but roughly speaking, this is a kind of sophisticated weighted averaging operation involving a non-necessarily additive set function named *capacity* to define the weight of any group of components to be aggregated [8]. The Choquet integral explicitly enables interactions among components (agents, criteria, scenarios) and allows to model positive and negative synergies in the aggregation process [12]. It includes various simpler models as special cases such as weighted averages, ordered weighted averages (OWA) [34], weigthed OWA [28] and Yaari's model [33].

This ability to encompass a wide range of attitudes in the aggregation of criteria has motivated the use of the Choquet integral for decision aid in various application contexts involving multiple criteria [11]. The Choquet integral has also been used in various domains of artificial intelligence such as machine learning [23, 24], recommender systems [1], multiagent decision making [4], information fusion [29], multiobjective state-space search [5], preference elicitation [2] and multiattribute evaluation [13]. However the high expressivity of the model comes at a computational cost for optimization tasks. The Choquet integral is a piecewise linear scalarizing functions including a number of pieces that is exponential in the number arguments. Substituting a linear aggregation by a Choquet aggregation in multiobjective combinatorial optimization can easily transform an easy problem to a NP-hard problem (see e.g., [6] for examples on shortest paths and spanning tree problems).

Optimizing a Choquet integral is a challenging problem that was the topic of several contributions in the last decade. A first lineariza-

[1] Sorbonne Université, CNRS, LIP6, F-75005 Paris, France, email: firstname.name@lip6.fr.

tion of the Choquet integral was proposed for convex capacities in [15] with an application to fair optimization. Another approach exploiting a decomposition of the Choquet integral as the maximum of integrals w.r.t. belief functions was proposed in [25]. Another decomposition method for optimizing the Choquet integral over a convex set with an application to resource allocation problems is proposed in [26]. The problem of optimizing a Choquet integral with an imprecisely defined capacity is studied in [27]. Finally some compact models (i.e., featuring a polynomial number of variables and constraints) for specific class of convex capacities have been proposed, e.g., k-additive belief functions [15] or $k+1$-additive and k-monotone capacities with applications to transportation and knapsack problems [14]. Now, we need to consider wider classes of capacities in Choquet optimization.

Furthermore, bipolar extensions of various scalarizing functions have been proposed in the literature to model asymmetric aggregation with respect to a specific reference point e.g., for decision making under risk [30], or for multicriteria decision making [9, 18]. However, none of these contributions addresses the computational aspects. More recently, the problem of optimizing the bipolar Choquet integral has been considered in [16, 17] but only for very specific classes of capacities. We will address this problem for more general capacities in the second part of the paper.

Although mutlicriteria optimization rarely considers problems involving more than ten criteria, multiagent-optimization may involve significantly larger sets of agents. Hence, we need scalable solutions methods for the optimization of the Choquet integral. The aim of this paper is to go one step further in this direction and to propose a new approach for general convex capacities that is more scalable to large sets of objectives. We will then extend the approach to the bipolar Choquet Integral.

2 BACKGROUND ON THE CHOQUET INTEGRAL, CAPACITIES AND THE CORE

Let $N=\{1,\dots,n\}$ denotes the set of points of view under consideration to assess the value of a solution in the decision problem. Depending on the context, the elements of N may represent a set of agents (collective decision-making), a set of criteria (multicriteria decision making), or a set of scenarios (decision under uncertainty). In the context of multiobjective optimization, the set N may also be seen as the set of objective functions to be optimized. In all cases, we assume that any feasible solution is characterized by a performance vector $x=(x_1,\dots,x_n)$ where x_i represent the value of x w.r.t. the ith point of view.

We recall now some formal definitions related to Choquet capacities and Choquet integrals.

Definition 1. *A* capacity *on N is a set function $v:2^N\to[0,1]$ such that $v(\emptyset)=0$ and for all $A,B\subseteq N, A\subseteq B\Rightarrow v(A)\le v(B)$. It is a normalized capacity if $v(N)=1$.*

Throughout the paper we will always assume that the capacities under consideration are normalized. A capacity v is said to be concave if $v(A\cup B)+v(A\cap B)\le v(A)+v(B)\,\forall A,B\subseteq N$ and convex if $v(A\cup B)+v(A\cap B)\ge v(A)+v(B)\,\forall A,B\subseteq N$. A capacity is additive if $v(A\cup B)+v(A\cap B)=v(A)+v(B)\,\forall A,B\subseteq N$. Therefore, an additive capacity is completely characterized by a vector λ such as $\lambda_i=v(\{i\}),\ i=1,\dots,n$ since $v(A)=\sum_{i\in A}\lambda_i$. Then the Choquet integral can be defined from any capacity as follows:

Definition 2. *For any vector $x=(x_1,\dots,x_n)\in\mathbb{R}^n$, the Choquet integral w.r.t. capacity v is a scalarizing function $C_v:\mathbb{R}^n\to\mathbb{R}$ defined by :*

$$C_v(x)=\sum_{i=1}^{n}\big[v(X_{(i)})-v(X_{(i+1)})\big]x_{(i)} \qquad (1)$$

$$=\sum_{i=1}^{n}\big[x_{(i)}-x_{(i-1)}\big]v(X_{(i)}) \qquad (2)$$

where (.) is any permutation such that $x_{(1)}\le\dots\le x_{(n)}$, and $X_{(i)}=\{x_{(i)},\dots,x_{(n)}\}$ is the set of objectives where the performance is at least as good as $x_{(i)}$, for $i=1,\dots,n$. Furthermore we assume that $x_{(0)}=0$ and $X_{(n+1)}=\emptyset$.

When the capacity is additive, the Choquet integral boils down to the weighted sum. The Choquet integral allows to model a wide range of behaviors as illustrated in the following example.

Example 1. *Let us consider a bi-objective optimization problem ($N=\{1,2\}$) with 5 different feasible solutions the performance of which are given in the following table.*

	a	*b*	*c*	*d*	*e*
x_1	9	7	4	2	1
x_2	1	2	3	7	9

Note that all solutions are Pareto-optimal (no solution is beaten on both objectives by another solution). Let us assume that objective 1 is more important than objective 2 for the decision maker. Let us consider a first capacity v_1 defined by $v_1(1)=0.7$ and $v_1(2)=0.3$. The optimal solution according to C_{v_1} is a, with an overall value given by $C_{v_1}(a)=[1-0.7]\times1+0.7\times9=6.6$. Since v_1 is additive C_{v_1} returns exaclty the weighted sum of the performance vector. Now it can easily be checked that the optimal solution would not change if we set $v_1(1)=t$ and $v_1(2)=1-t$ for any $t>1/2$. Indeed, as long as we use an additive capacity, the Choquet integral is actually a linear weighted mean and the only possible optima are the extreme points of the convex hull of the feasible set, namely a and e in the present case (and only a under the constraint that $t>1/2$).

To overcome the limitation of linear aggregation criteria and to obtain more balanced compromise solutions, the Choquet integral must be used with a convex capacity. For example, let us consider capacity v_2 defined by $v_2(1)=0.3$ and $v_2(2)=0.2$ which is not additive but convex. The optimal solution according to the Choquet integral is then b, with $C_{v_2}(b)=[1-0.3]\times2+0.3\times7=3.5$. We observe that this solution proposes a better compromise between both objectives. Now if we move a little further from additivity by considering v_3 defined by $v_3(1)=0.2$ and $v_3(2)=0.1$ which is also convex, the optimal solution according to the Choquet Integral is then c, with $C_{v_3}(c)=[1-0.3]\times2+0.3\times9=3.2$.

We see on the above example that the Choquet integral provides a much better control on the profile of the solution than a linear criterion due to the use of convex capacities. A similar statement could be made with concave capacities in minimization problems. We also observe that, in the previous example, convex capacities seem to favor solutions having a balanced profile. This is strongly related to the notion of *preference for interior points* introduced below:

Definition 3. *A scalarizing function $f:\mathbb{R}^n\to\mathbb{R}$ satisfies preference for interior points if and only if, for all vectors $x^1,\dots,x^m$ such that $f(x^1)=f(x^2)=\dots=f(x^m)$, for all $\lambda\in\mathbb{R}_+^m$ such that $\sum_{i=1}^m\lambda_i=1$ we have:*

$$f(\sum_{i=1}^{m}\lambda_i x^i)\ge f(x^k),\ \forall k\in\{1,\dots,m\}$$

Example 2. *Let us consider two solutions* $a = (10, 0)$ *and* $b = (0, 10)$*. If we use a scalarizing function satisfying the preference for interior points, then, the point* $c = \frac{a+b}{2} = (5, 5)$ *is preferred to* a *and* b *since* $f(c) \geq f(a)$ *and* $f(c) \geq f(b)$*.*

Preference for interior points favors solutions having a balanced profile which is natural in a number of different contexts such as multi-agent decision making (fairness), uncertainty (robustness), and multicriteria analysis (compromise solutions). Interestingly, the condition for a Choquet integral to satisfy the preference for interior points is known due to a result established in the context of portofolio optimization [3] (where preference for interior points is named preference for diversification). The result can be reformulated in our context as follows:

Property 1. *The Choquet Integral* C_v *satisfies preference for interior points if and only if* v *is convex.*

In several optimization contexts, it is useful to consider the set of additive measures dominating a given capacity. This set is named the core and is defined as follows:

Definition 4. *The* core *of a capacity* v *is the set of all additive capacities dominating* v*, more formally we have:*

$$core(v) = \{\lambda : 2^N \to [0, 1] \text{ additive} \mid \lambda(S) \geq v(S)\ \forall S \subseteq N\}$$

It is well known that a convex capacity has a non empty core [21]. Moreover we have the following useful property [20]:

Property 2. *If* v *is convex we have* $C_v(x) = \min_{\lambda \in core(v)} \lambda \cdot x$

This property has led to a first mathematical program proposed in [15] for Choquet optimization. Let us recall it briefly hereafter. Using Property 2 we know that for any fixed $x \in \mathbb{R}^n$, $C_v(x)$ is the optimal value of the following mathematical program:

$$(\mathcal{P}_1) \quad \begin{array}{l} \min \sum_{i=1}^{n} \lambda_i x_i \\ \begin{cases} \sum_{i \in A} \lambda_i \geq v(A) & \forall A \subseteq N \\ x \in X \end{cases} \\ \lambda_i \geq 0,\ \forall i = 1, .., n \end{array}$$

Hence, in order to maximize a Choquet Integral on a set of vectors X, one can consider the dual of $\mathcal{P}_1$ and define x as a variable of the problem subject to $x \in X$. This leads to the following mathematical program [15]:

$$(\mathcal{P}_2) \quad \begin{array}{l} \max \sum_{A \subseteq N} v(A) \times d_A \\ \begin{cases} \sum_{A \subseteq N : i \in A} d_A \leq x_i & \forall i = 1, .., n \\ x \in X \end{cases} \\ d_A \geq 0,\ \forall A \subseteq N \end{array}$$

Problem $\mathcal{P}_2$ has $2^n + n$ continuous variables and n constraints. It can be specialized to solve any combinatorial optimization program, by replacing the constraint $x \in X$ by the ones specifying the multi-objective problem. For example, to solve a multiagent knapsack problem with m items, we consider the binary variables y_j, for all $j = 1, \ldots, m$; y_j takes value 1 if item j is chosen and 0 otherwise. Then variables x_i are linked to y_j by linear constraints of type $x_i = \sum_{j=1}^{m} y_j u_{ij}$, with u the $n \times m$ matrix giving the utilities of the items for all agents. We also add the budget constraint $\sum_{j=1}^{m} w_j y_j \leq K$, w_j denoting the weight of item j and K being the maximal admissible weight for the knapsack. We will come back to this problem in Section 4 dedicated to numerical tests.

3 A NEW LP FORMULATION OF CHOQUET OPTIMIZATION

In this section we are going to introduce an alternative approach for Choquet optimization, also based on Property 2. Let us first recall the polyhedral properties of the core of convex capacities (for more details see [8]). On the lattice $(2^N, \subseteq)$, a *maximal chain* is defined as a sequence of subsets $\emptyset = A_0 \subset A_1 \subset .. \subset A_n = N$, (therefore we have $|A_{i+1} \backslash A_i| = 1$ for all $i = 1, \ldots, n-1$). Let Π be the set of permutation on N, these permutations correspond bijectively to maximal chains in the lattice $(2^N, \subseteq)$. Indeed, we associate to a permutation π a maximal chain defined by $A_0^\pi = \emptyset$ and $A_i^\pi = \{\pi(1), \ldots, \pi(i)\}$ for $i = 1, \ldots, n$. We can now associate to every $\pi \in \Pi$ and any given capacity v the marginal vector $\lambda^{\pi,v} \in \mathbb{R}^n$, defined by

$$\lambda_{\pi(i)}^{\pi,v} = v(A_i^\pi) - v(A_{i-1}^\pi),\ i = 1, \ldots, n \qquad (3)$$

Let us illustrate these properties on a short example.

Example 3. *Let us consider the lattice* $(2^{\{1,2,3\}}, \subseteq)$ *and the convex capacity* v *defined as follows:*

	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
v	*0.2*	*0.3*	*0.4*	*0.6*	*0.7*	*0.8*	*1*

Let π *be the permutation defined by* $\pi(1) = 2$*,* $\pi(2) = 1$ *and* $\pi(3) = 3$*. The maximal chain associated to* π *corresponds to the sequence* $\emptyset, \{2\}, \{1,2\}, \{1,2,3\}$*. Therefore, the marginal vector associated to* π *and* v *is defined by:*
$\lambda_{\pi(1)}^{\pi,v} = \lambda_2^{\pi,v} = v(\{2\}) - v(\emptyset) = 0.3$
$\lambda_{\pi(2)}^{\pi,v} = \lambda_1^{\pi,v} = v(\{1,2\}) - v(\{2\}) = 0.3$
$\lambda_{\pi(3)}^{\pi,v} = \lambda_3^{\pi,v} = v(\{1,2,3\}) - v(\{1,2\}) = 0.4.$

We define the convex hull of the marginal vectors, often known as the Weber set, such as:

$$\text{Web}(v) = \text{conv}(\lambda^{\pi,v}, \pi \in \Pi) \qquad (4)$$

The following theorem [31, 8] asserts that, for any capacity v, the convex hull of marginal vectors always contains the core:

Theorem 1. *For any capacity* v*, core*$(v) \subseteq$ *Web*(v)

The next theorem (see [8] Chapter 3) asserts that the converse holds for convex capacities and details the structure of the extreme points of such capacities.

Theorem 2. *The following statements are equivalent:*

1. v *is a convex capacity*
2. *for all* $\pi \in \Pi$*,* $\lambda^{\pi,v} \in$ *core*(v)
3. *core*(v) *=* *Web*(v)
4. *ext*$($*core*$(v)) = \{\lambda^{\pi,v}, \forall \pi \in \Pi\}$

with ext the set of extreme points for a given polyhedron.

Using Theorem 2 and Property 2, we can propose a new formulation of the Choquet Integral.

Proposition 1. *Let* v *be a convex capacity and* $x \in \mathbb{R}^n$ *a solution vector, the Choquet Integral of* x *is defined as follows:*

$$C_v(x) = \min_{\pi \in \Pi} \lambda^{\pi,v} \cdot x$$

Proof. According to Theorem 2, the core of v is a non empty convex polyhedron, whose extreme points are the vectors $\lambda^{\pi,v}$ associated to permutations $\pi \in \Pi$ and the capacity v. Thus, it always exists an extreme point such that $\lambda^{\pi,v} \cdot x = \min_{\lambda \in core(v)} \lambda \cdot x = C_v(x)$. □

This proposition shows that, when v is convex, function $C_v(x)$ is the minimum of a *finite* set of linear functions of the form $\lambda^{\pi,v} \cdot x$, $\pi \in \Pi$. This suggest a new linearization of the Choquet integral for convex capacities:

$$(\mathcal{P}_3) \quad \begin{array}{l} \max\ y \\ \left\{ \begin{array}{ll} y \leq \lambda^{\pi,v} \cdot x & \forall \pi \in \Pi \\ x \in X \end{array} \right. \end{array}$$

This linear problem has $n+1$ continuous variables and $n!$ constraints, beside the constraints necessary to define the feasible space ($x \in X$). It can be specialized to solve any combinatorial optimization program, by replacing $x \in X$ by the relevant constraints. The number of variables remains polynomial contrary to problem $\mathcal{P}_2$ but there is still a problem of scalability since the number of constraints grows exponentially with n. This formulation can be used in combinatorial multicriteria optimization where the number of criteria is usually limited but would not be sufficient for multiagent optimization where the number of agents can be more important. In order to overcome the problem we introduced below a constraint generation method for this problem.

The principle of constraint generation is as follows: instead of considering from the begining the entire set of constraint $y \leq \lambda^{\pi,v} \cdot x$ for all $\pi \in \Pi$, we only insert one of these constraints, chosen randomly and we launch the optimization. Once an optimal solution to the restricted problem is found, we have to check whether the omitted constraints are satisfied by the current solutions. If it happens to be the case, we have found the optimal solution. Otherwise we must identify one of the violated constraints and insert it in the current optimization problem. In order to implement this principle we first establish the following proposition:

Proposition 2. *For any fixed vector $x \in \mathbb{R}^n$, and any capacity v, the marginal vector $\lambda^{[.],v}$, associated to a permutation $[.]$ sorting the elements of x in decreasing order ($x_{[1]} \geq \ldots \geq x_{[n]}$) is such that $C_v(x) = \lambda^{[.],v} \cdot x$.*

Proof. We remind that $X_{(i)}$ is the set of objectives whose performance level is greater than or equal to $x_{(i)}$ and that A_i^π is the i-th set in the maximal chains corresponding to a given permutation $\pi \in \Pi$. We have:

$$C_v(x) = \sum_{i=1}^{n} (v(X_{(i)}) - v(X_{(i+1)}))\, x_{(i)} \tag{5}$$

$$= \sum_{i=1}^{n} (v(X_{(n+1-i)}) - v(X_{(n-i+)}))\, x_{(n+1-i)} \tag{6}$$

Now, the chain associated to permutation $[.]$ is such that $A_i^\pi = X_{n+1-i}$. Hence we have $\lambda_{[i]}^{[.],v} = v(X_{(n+1-i)}) - v(X_{(n-i)})$ from Equation (3). Moreover $x_{[i]} = x_{(n+1-i)}$. Hence, from Equation (6) we have $C_v(x) = \sum_{i=1}^{n} \lambda_{[i]}^{[.],v} x_{[i]} = \lambda^{[.],v} \cdot x$.

□

This result is used as follows. At any step of the constraint generation process, when an optimal solution x of the current linear programm is found, with objective value y, we consider a permutation $[.]$ reordering the components of x by decreasing order and the corresponding marginal vector $\lambda^{[.],v}$. Then we check whether $y \leq \lambda^{[.],v} \cdot x$ (this is the more demanding constraint due to Proposition 2). If this constraint is violated, we add it to the current model for the next step and we relaunch the optimization process. If the constraint is satisfied, the optimization process is over and the following proposition ensures that solution x is feasible, and therefore optimal.

Proposition 3. *The constraint generation algorithm described above yields an optimal solution of $\mathcal{P}_3$.*

Proof. Let x be the optimal solution of the linear programm solved at the final iteration, with objective value y. Since the algorithm terminates at this step we have $y \leq \lambda^{[.],v} \cdot x = C_v(x) \leq \lambda^{\pi,v} \cdot x$ for all $\pi \in \Pi$ due to Proposition 1. Hence all constraints of $\mathcal{P}_3$ are satisfied. This establishes the feasibility of x for $\mathcal{P}_3$ and therefore its optimality. □

Note that the determination of a violated constraint (if any) is performed in $O(n \log n)$ at every step. It is indeed sufficient to sort current vector x to determine $[.]$.

4 COMPUTATIONAL TESTS FOR THE CHOQUET INTEGRAL

We implemented the presented models using the Gurobi 8.1.1 solver on a computer with 12GB of RAM and an Intel(R) Core(TM) i7 CPU 950 @ 3.07GHz processor. These models were tested on randomly generated instances of the Choquet-knapsack problem that is defined by the following mathematical program:

$$\max C_v(x_1, \ldots, x_n)$$

$$s.t. \left\{ \begin{array}{l} x_i = \sum_{j=1}^{m} u_{ij} y_j,\ i = 1, \ldots, n \\ \sum_{j=1}^{n} w_j y_j \leq K \end{array} \right.$$

$$x_i \in \mathbb{R}, y_i \in \{0, 1\}$$

where $u_{ij} \in [\![1, 10]\!]$ denotes the utility of item j w.r.t. criterion i (or agent i) and $w_j \in [\![1, 100]\!]$ denotes the weight of item j, and $K = \sum_{i=1}^{m} w_i/2$. We generated instances of different sizes, with m the number of items varying from 100 to 1000 and n the number of criteria (or agents) varying from 5 to for 500. The convex capacities v used in the tests are randomly drawn. For every size, we solved 20 instances of the Choquet optimal knapsack problem of the same size and the average computation times are given in the tables given below (a time limit set to 7200 seconds for each instance was used and the average time is computed only when all the 20 instances are solved within the time constraint).

Table 1. Times (s) obtained by model $\mathcal{P}_2$ for the Choquet-knapsack

m	n=5	n=7	n=10	n=15	n=17
100	0.03	0.04	0.13	4.11	14.72
250	0.1	0.09	0.23	4.55	22.4
500	0.1	0.2	0.64	26.33	106.57
750	0.58	0.92	0.91	25.95	362.55
1000	1.32	2.26	14.14	263.28	395.3

We observe that $\mathcal{P}_2$ is able to solve instances with a large number of items. Nonetheless, the exponential number of variables limits the model to a restrained number of objectives, which can be problematic in several contexts, such as multi-agent optimization for example.

Table 2. Times (s) obtained by model $\mathcal{P}_3$, with the constraint generation algorithm, for the Choquet-knapsack

m	n=25	n=50	n=100	n=250	n=500
100	0.14	0.24	0.17	3.42	26.76
250	0.44	1.01	1.58	4.11	47.96
500	2.34	2.56	9.73	11.3	12.90
750	3.85	5.3	22.82	36.83	41.65
1000	6.19	9.74	50.45	111.24	57.11

We did not report the results beyond $n = 17$ because some instances were not solved within the time constraint.

The model $\mathcal{P}_3$ used with the constraint generation algorithm is much faster and can solve instances with large number of items and objectives within the time constraint, as we can see in the above table. Another measure that can be useful to assess the potential of the proposed approach is the average number of constraints inserted during the constraint generation process to reach the optimum. The results observed in our experiments are given in the following table:

Table 3. Number of constraints added by the constraint generation algorithm in model $\mathcal{P}_3$ for the Choquet-knapsack

m	n=25	n=50	n=100	n=250	n=500
100	9.3	9.25	15.35	91.15	165.1
250	19.3	17.6	21.15	82.7	236.8
500	37.15	34.45	35.1	51.35	75.05
750	42.85	29.35	24.85	32.35	19.9
1000	55.65	51.1	48.5	23.5	24.0

In our tests, the number of added constraints remains small whatever the size of the instance. For example, when $n = 500$, the total number of constraints is of order 3^{150} and our tests show that we only have to insert a few hundreds of constraints to find the optimal solution. Hence the implementation of a constraint generation algorithm on formulation $\mathcal{P}_3$ appears to be particularly efficient. It significantly improves the performance obtained from model $\mathcal{P}_2$.

It is worth noting that model $\mathcal{P}_2$ could be solved more efficiently by using column generation techniques as suggested in [14]. However, the authors of this suggestion remark that the column generation approach for $\mathcal{P}_2$ tends to become impracticable beyond $n = 80$. They also remark that possible integrality requirements are difficult to handle within such an approach. Hence, it is unlikely that this approach could provide better results than those obtained with our approach for multiobjective combinatorial optimization problems. More compact models are proposed in the same paper but they concern restricted classes of convex capacities. The advantage of our approach is that it applies to any convex capacity. This class is of special interest because, as recalled earlier, this is by using convex capacity that we can promote balanced solutions when optimizing a Choquet integral (preference of interior points).

5 THE CASE OF THE BIPOLAR CHOQUET INTEGRAL

In this Section, we extend our computational model to bipolar Choquet integrals motivated by the following statement. In several contexts, decision makers may refer to a specific point in the valuation scale to assess the performance under a specific objective. These scales are known as bipolar scales, with 0 defined as the reference point. This behavior cannot be represented by the standard Choquet Integral with preferences modeled by a capacity, as shown in the following example:

Example 4. *A decision maker ranks the alternatives of $X = \{a, b, c, d\}$ in the order $a \succ b \succ c \succ d$. Knowing that the performance of the alternative on two criteria are given in the table given below let us check if this preference order is representable by a discrete Choquet integral.*

	a	*b*	*c*	*d*
x_1	*3*	*5*	*-5*	*-7*
x_2	*9*	*5*	*-5*	*-1*

Note that such preferences have their internal consistency that can be explicited as follows: when performances are positive, the decision maker prefers the solution maximizing the sum of objectives. However, when performances are negative, he adopts a more cautious behavior towards inequalities and favors a solution having a more balanced profile. A suitable representation of the preference order by the Choquet integral should satisfy: $x \succ y \Leftrightarrow C_v(x) > C_v(y)$ for all $x, y \in X$.

We have $a \succ b$, therefore:

$$3 + 6v(\{2\}) > 5$$
$$\Leftrightarrow 6v(\{2\}) > 2$$
$$\Leftrightarrow v(\{2\}) > 2/6$$

We have $c \succ d$, therefore:

$$-7 + 6v(\{2\}) < -5$$
$$\Leftrightarrow 6v(\{2\}) < 2$$
$$\Leftrightarrow v(\{2\}) < 2/6$$

We obtain a contradiction showing that no capacity exists to represent the prescribed ranking. Therefore, function C_v cannot model these preferences.

To overcome this limitation, a bipolar extension of Choquet Integral has been proposed in [10].

Definition 5. *Let $x \in \mathbb{R}^n$ and u and v two capacities. The bipolar extension of the Choquet (biChoquet integral for short) is defined as follows:*

$$C_{u,v}(x) = C_u(x^+) - C_v(x^-) \tag{7}$$

where $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$.

We present below a set of necessary conditions for a biChoquet integral to satisfy this property. Let us first recall a well known property of Choquet integrals [8] :

Property 3. *For v a given capacity, let $\bar{v}$ be its dual capacity defined by $\bar{v}(A) = 1 - v(N \backslash A)$ for all $A \subseteq N$. We have the following equality: $C_v(x) = -C_{\bar{v}}(-x)$*

Proposition 4. *If the biChoquet Integral satisfies preference for interior points, then u is convex and v is concave.*

Proof. When $x \in \mathbb{R}^n_+$ then $C_{u,v}(x) = C_u(x) - C_v(0) = C_u(x)$. Thus, u must be convex to model preference for interior points on positive vectors. When $x \in \mathbb{R}^n_-$ then $C_{u,v}(x) = C_u(0) - C_v(-x) = -C_v(-x) = C_{\bar{v}}(x)$, with $\bar{v}$ is the dual capacity of v. Thus, $\bar{v}$ must be convex to model preference for interior points on negative vectors. It is well known that the dual of a convex capacity is concave. Thus, v must be concave to model preference for interior points. $\square$

As discussed earlier, preference for interior points is a desirable property in many decision contexts because it favors balanced solution vectors. In the sequel, we will therefore use a convex u and a concave v which are necessary to satisfy this property. As concave capacities appear in the previous result, we now make explicit a counterpart of Property 2 for concave capacities.

Proposition 5. *If v is concave, then $C_v(x) = \max_{\lambda \in core(\bar{v})} \lambda \cdot x$*

Proof. According to Property 3 we have: $C_v(x) = -C_{\bar{v}}(-x) = -\min_{\lambda \in core(\bar{v})} \lambda.(-x) = \max_{\lambda \in core(\bar{v})} \lambda \cdot x$. □

Hence we obtain the following formulation of a biChoquet integral :

Proposition 6. *Let $x \in \mathbb{R}^n$. If u is convex and v is concave then we have:*

$$C_{u,v}(x) = \min_{\lambda \in core(u)} \lambda \cdot x^+ - \max_{\lambda \in core(\bar{v})} \lambda \cdot x^-$$

Proof. The result directly derives from Propositions 2 and 5 and Definition 5. □

Proposition 6 allows to extend linear program $\mathcal{P}_2$ to optimize the bipolar Choquet Integral. From Proposition 5 the value of $C_v(x)$ can indeed be obtained, for any fixed outcome vector $x \in \mathbb{R}^n$, as the optimal value of the following linear program, when v is concave:

$$(\mathcal{P}_4) \quad \begin{array}{l} \max \sum_{i=1}^{n} \lambda_i x_i \\ \left\{ \begin{array}{ll} \sum_{i \in A} \lambda_i \leq v(A) & A \subseteq N \end{array} \right. \\ \lambda_i \geq 0 \; i = 1, .., n \end{array}$$

Program $\mathcal{P}_4$ derives from Proposition 5 after observing that the constraints for all $B \subseteq N, \sum_{i \in B} \lambda_i \geq \bar{v}(B)$ are equivalent to $\forall A \subseteq N, \sum_{i \in A} \lambda_i \leq v(A)$ (by setting $A = N \setminus B$). Now, if we consider x as a variable vector and we pass to the dual, then we obtain the following mathematical program to linearize the bipolar Choquet Integral, provided that u is convex and v is concave:

$$(\mathcal{P}_5) \quad \begin{array}{l} \max \sum_{A \subseteq N} u(A) \times d_A^+ - \sum_{A \subseteq N} v(A) \times d_A^- \\ \left\{ \begin{array}{ll} \sum_{A \subseteq N : i \in A} d_A^+ \leq x_i^+ & i = 1, \dots, n \\ \sum_{A \subseteq N : i \in A} d_A^- \geq x_i^- & i = 1, \dots, n \\ x_i = x_i^+ - x_i^- & i = 1, \dots, n \\ 0 \leq x_i^+ \leq p_i \times M & i = 1, \dots, n \\ 0 \leq x_i^- \leq (1 - p_i) \times M & i = 1, \dots, n \\ x \in X & \end{array} \right. \\ x_i^-, x_i^+, d_A^+, d_A^- \geq 0 \; i = 1, .., n, \; \forall A \subseteq N \\ p_i \in \{0, 1\} \; i = 1, \dots, n \end{array}$$

Program $\mathcal{P}_5$ has $2^{n+1} + 3n$ constraints, $3n + 1$ continuous variables and n binary variables. Variables $p_i, \; i = 1, \dots, n$ are used to decide whether x_i is positive or not. The M is used as usual to model disjunctive constraints. Here also, $\mathcal{P}_5$ can be specialized to solve any combinatorial optimization program, by replacing the constraint $x \in X$ by the ones defining the set of feasible solutions.

Thus, we obtained a first linear formulation for biChoquet optimization under the assumption that u is convex and v is concave. This formulation generalizes the linearization proposed for Cumulative Prospect Theory in [17]. Of course, the formulation is not compact due to the number of variables involved and this formulation is only convenient for small values of n (as already observed for $\mathcal{P}_2$). For this reason, the remaining part of this section is devoted to the extension of the approach based on $\mathcal{P}_3$ to the bipolar case.

Let us first remark that using Property 3 we obtain the following reformulation of the biChoquet integral:

$$C_{u,v}(x) = C_u(x^+) + C_{\bar{v}}(-x^-) \tag{8}$$

where both capacities u and $\bar{v}$ are convex to satisfy preference for interior points. Hence we get the following proposition:

Proposition 7. *Let $x \in \mathbb{R}^n$ and let u and v be two capacities respectively convex and concave. We have:*

$$C_{u,v}(x) = \min_{\pi \in \Pi} \left(\lambda^{\pi,u} \cdot x^+ - \lambda^{\pi,\bar{v}} \cdot x^- \right) \tag{9}$$

Proof. Due to Proposition 1 we have:

$$\begin{array}{rcl} C_{u,v}(x) & = & \min_{\pi \in \Pi} \lambda^{\pi,u} \cdot x^+ + \min_{\pi \in \Pi} \lambda^{\pi,\bar{v}} \cdot (-x^-) \\ & \leq & \min_{\pi \in \Pi} \left(\lambda^{\pi,u} \cdot x^+ + \lambda^{\pi,\bar{v}} \cdot (-x^-) \right) \\ & = & \min_{\pi \in \Pi} \left(\lambda^{\pi,u} \cdot x^+ - \lambda^{\pi,\bar{v}} \cdot x^- \right) \end{array}$$

Now we establish the reverse inequality. Let π_* a permutation such that $C_u(x^+) = \lambda^{\pi_*,u} \cdot x^+$ and π'_* a permutation such that $C_{\bar{v}}(-x^-) = \lambda^{\pi'_*,\bar{v}} \cdot (-x^-)$. Now, let us consider any permutation τ_* of $(1, \dots, n)$ such that $\tau_*(i) = \pi_*(i)$ if $x_i > 0$ and $\tau_*(i) = \pi'_*(i)$ if $x_i < 0$, $\tau_*(i)$ being chosen arbitrarily for all i such that $x_i = 0$ to complete the permutation. By construction, we have:

$$\begin{array}{rcl} C_{u,v}(x) & = & C_u(x^+) + C_{\bar{v}}(-x^-) \text{ due to (8)} \\ & = & \lambda^{\tau_*,u} \cdot x^+ + \lambda^{\tau_*,\bar{v}} \cdot (-x^-) \\ & = & \lambda^{\tau_*,u} \cdot x^+ - \lambda^{\tau_*,\bar{v}} \cdot x^- \\ & \geq & \min_{\pi \in \Pi} \left(\lambda^{\pi,u} \cdot x^+ - \lambda^{\pi,\bar{v}} \cdot x^- \right) \end{array}$$

which completes the proof. □

We propose then the following mixed integer program for biChoquet optimization under the assumption that u and v are respectively convex and concave.

$$(\mathcal{P}_6) \quad \begin{array}{l} \max \; y \\ \left\{ \begin{array}{ll} y \leq \lambda^{\pi,u} \cdot x^+ - \lambda^{\pi,\bar{v}} \cdot x^- & \forall \pi \in \Pi \\ x_i = x_i^+ - x_i^- & \forall i \in N \\ 0 \leq x_i^+ \leq p_i \times M & \forall i \in N \\ 0 \leq x_i^- \leq (1 - p_i) \times M & \forall i \in N \\ x \in X & \end{array} \right. \\ p_i \in \{0, 1\}, \; \forall i \in N \end{array}$$

Problem $\mathcal{P}_6$ has $n! + 3n$ constraints, $3n + 1$ continuous variables and n binary variables. Here also, variables $p_i, \; i = 1, \dots, n$ are used to decide whether x_i is positive or not and M is used to model disjunctive constraints. Formulation $\mathcal{P}_6$ can also be specialized to solve any combinatorial optimization program, by replacing the constraint $x \in X$ by the relevant feasibility constraints. Similarly to $\mathcal{P}_2$, problem $\mathcal{P}_6$ is not compact due to the exponential number of constraints. Thus, we propose a constraint generation method for this problem. To this end, we first establish the following proposition:

Proposition 8. *For any fixed vector $x \in \mathbb{R}^n$, and any capacity v, the marginal vector $\lambda^{[.],v}$, associated to a permutation $[.]$sorting the elements of x in decreasing order ($x_{[1]} \geq \dots \geq x_{[n]}$) is such that $C_{u,v}(x) = \lambda^{[.],u} \cdot x^+ - \lambda^{[.],\bar{v}} \cdot x^-$.*

Proof. We have $C_{u,v}(x) = C_u(x^+) + C_{\bar{v}}(-x^-) = \lambda^{[.],u} x^+ + \lambda^{[.],\bar{v}}(-x^-)$ by Proposition 2. Hence $C_{u,v}(x) = \lambda^{[.],u} x^+ - \lambda^{[.],\bar{v}} x^-$. □

This proposition allows to extend the constraint generation approach introduced in Section 3 to the case of biChoquet integrals.

We start with a relaxed version of problem $\mathcal{P}_6$ including a single constraint limitating y using a permutation π arbitrary chosen. Then, at any step of the constraint generation process, when an optimal solution x of the current linear program is found, with objective value y, we determine a permutation $[.]$ reordering the components of x by decreasing order and the corresponding marginal vectors $\lambda^{[.],u}$ and $\lambda^{[.],\bar{v}}$. Then we check whether $y \leq \lambda^{[.],u} \cdot x^+ - \lambda^{[.],\bar{v}} \cdot x^-$. If this constraint is violated, we add it to the current model for the next step and we relaunch the optimization process. If the constraint is satisfied, x is feasible, then we can stop as proved in the following proposition.

Proposition 9. *The constraint generation algorithm described above yields an optimal solution of* $\mathcal{P}_6$.

Proof. Let x be the optimal solution of the linear program solved at the final iteration, with objective value y. Since the algorithm terminates at this step we have $y \leq \lambda^{[.],u} \cdot x^+ - \lambda^{[.],\bar{v}} \cdot x^- = C_{u,v}(x) \leq \lambda^{\pi,u} \cdot x^+ + \lambda^{\pi,\bar{v}} \cdot (-x^-)$ for all $\pi \in \Pi$ due to Proposition 7. Hence all constraints of $\mathcal{P}_6$ are satisfied. This establishes the feasibility of x for $\mathcal{P}_6$ and therefore its optimality. □

6 COMPUTATIONAL TESTS FOR THE BICHOQUET INTEGRAL

We implemented program $\mathcal{P}_5$ and $\mathcal{P}_6$ (the latter using the proposed constraint generation method) on the biChoquet optimal Knapsack problem. This is a variant of the problem introduced in Section 4 where the objective function is replaced by $\max C_{u,v}(x_1, \ldots, x_n)$. The experimental environment is the same as the one presented in Section 4. Here are the computation times obtained with $\mathcal{P}_5$:

Table 4. Times (s) obtained by MIP $\mathcal{P}_5$ for the BiChoquet-knapsack

m	$n=3$	$n=5$	$n=7$
100	0.03	0.21	0.67
500	0.05	1.31	45.60
750	0.08	0.87	125.72
1000	0.13	3.28	150.48

We observe that the model $\mathcal{P}_5$ is able to solve instances with a large number of items. Nonetheless, the number of objectives has to remain small, due to the exponential number of variables and the adds of binary variables to manage the sign of variables $x_i, i = 1, \ldots, n$. The optimization of the BiChoquet integral seems more challenging than the optimization of a simple Choquet integral. We now give the times obtained with $\mathcal{P}_6$ and constraint generation.

Table 5. Times (s) obtained by $\mathcal{P}_6$, with a constraint generation algorithm, for the BiChoquet-knapsack

m	n=3	n=5	n=7	n=10	n=15
100	0.02	0.06	0.13	0.17	1.19
250	0.04	0.12	0.47	1.88	19.09
500	0.06	0.37	1.87	29.01	785.46
750	0.06	0.51	2.70	114.08	-
1000	0.1	0.57	4.52	344	-

Model $\mathcal{P}_6$ with the constraint generation algorithm solves instances with larger number of objects and objectives compared to $\mathcal{P}_5$. Moreover, instances of the same size are solved up to 35 times faster. Let us have a look at the average number of constraints inserted during the optimization process.

Table 6. Number of constraints inserted by the constraint generation algorithm in $\mathcal{P}_6$ for the BiChoquet-knapsack

m	n=3	n=5	n=7	n=10	n=15
100	3.3	9.35	16.1	30	84.05
250	3.05	7.75	19.5	47.8	112.95
500	3.2	8.75	16.75	50.91	154.2
750	3.15	8.35	19.15	52.05	-
1000	3.15	8.35	16.15	50.45	-

The number of added constraints remains small independently of the size of the instance. For example, for $n = 10$ we only add about 4% of constraints to find the optimal solution.

7 CONCLUSION

We have proposed new computational models combining decision theory and linear programming to find compromise solutions in combinatorial multiobjective problems. We first used the Choquet Integral, well-known for its expressivity, with a convex capacity to promote solutions having a balanced performance profile (preference for interior points). The first model we proposed for the Choquet integral (Section 3) has an exponential number of constraints but we have shown that this weakness can be partly overcomed using a constraint generation algorithm taylored to the Choquet integral.We have implemented this model on various instances of the multi-objective knapsack problem. This leads to solve a mixed-integer linear program that appears to be efficient in terms of computation time (and number of added constraints) on instances of large size.

Then, we considered the bipolar extension of the Choquet Integral with respect to a convex capacity u and a concave capacity v. These conditions are proved to be necessary to satisfy the so-called *preference for interior points*. We proposed a first mixed integer program to optimize a bipolar Choquet integral in combinatorial multiobjective problems. As in the previous case, we used a constraint generation algorithm taylored to the optimization of a biChoquet integral. We also tested this model on various instances of the multiobjective knapsack problem. The computation times are still good but significantly higher than for a standard Choquet integral. It is worth noting that all computational models proposed in the paper are also directly applicable to mutliobjective optimization on continuous domains.

Let us finally mention some possible directions for further research. A dual approach and a useful complement to the material proposed here would be to design an efficient column generation algorithm for solving $\mathcal{P}_2$ for convex capacities. However the first steps made in this direction (see e.g., [14]) have shown some difficulties that have been discussed at the end of Section 4. Besides, a natural extension of our results would be to propose computational models for Choquet maximization with respect to a *non-convex* capacity. Although preference for interior points would not be longer guaranteed, such instances of the Choquet model are worth considering because they may correspond to sophisticated decision behaviors that could be observed in practice. Finally, another natural extension of our work concerns the bipolar case. It is worth extending the proposed computational model to general biChoquet integrals with respect to a bicapacity $w(X, Y)$ not necessarily decomposable under the form $u(X) - v(Y)$. This would certainly be the opportunity to exploit the notion of core of a bicapacity [32].

REFERENCES

[1] Gleb Beliakov, Tomasa Calvo, and Simon James, 'Aggregation functions for recommender systems', in *Recommender Systems Handbook*, 777–808, Springer, (2015).

[2] Nawal Benabbou, Patrice Perny, and Paolo Viappiani, 'Incremental elicitation of choquet capacities for multicriteria choice, ranking and sorting problems', *Artificial Intelligence Journal*, **246**, 152–180, (2017).

[3] Alain Chateauneuf and Jean-Marc Tallon, 'Diversification, convex preferences and non-empty core in the choquet expected utility model', *Economic Theory*, (2002).

[4] Jean-Philippe Dubus, Christophe Gonzales, and Patrice Perny, 'Choquet optimization using gai networks for multiagent/multicriteria decision-making', in *International Conference on Algorithmic DecisionTheory*, pp. 377–389. Springer, (2009).

[5] Lucie Galand and Patrice Perny, 'Search for Choquet-optimal paths under uncertainty', in *Proceedings of the 23rd conference on Uncertainty in Artificial Intelligence*, pp. 125–132, (July 2007).

[6] Lucie Galand, Patrice Perny, and Olivier Spanjaard, 'Choquet-based optimisation in multiobjective shortest path and spanning tree problems', *European Journal of Operational Research*, **204**(2), 303–315, (2010).

[7] Michel Grabisch, 'The application of fuzzy integrals in multicriteria decision making', *European journal of operational research*, **89**(3), 445–456, (1996).

[8] Michel Grabisch, *Set functions, games and capacities in decision making*, Springer, 2016.

[9] Michel Grabisch and Christophe Labreuche, 'Bi-capacities—ii: the choquet integral', *Fuzzy sets and systems*, **151**(2), 237–259, (2005).

[10] Michel Grabisch and Christophe Labreuche, 'Bi-capacities – part II: the choquet integral', *CoRR*, **abs/0711.2112**, (2007).

[11] Michel Grabisch and Christophe Labreuche, 'A decade of application of the choquet and sugeno integrals in multi-criteria decision aid', *Annals OR*, **175**(1), 247–286, (2010).

[12] Michel Grabisch, Jean-Luc Marichal, Radko Mesiar, and Endre Pap, *Aggregation functions*, volume 127, Cambridge University Press, 2009.

[13] Christophe Labreuche and Sébastien Destercke, 'How to handle missing values in multi-criteria decision aiding?', in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pp. 1756–1763. International Joint Conferences on Artificial Intelligence Organization, (2019).

[14] Julien Lesca, Michel Minoux, and Patrice Perny, 'Compact versus noncompact LP formulations for minimizing convex choquet integrals', *Discrete Applied Mathematics*, **161**(1-2), 184–199, (2013).

[15] Julien Lesca and Patrice Perny, 'LP solvable models for multiagent fair allocation problems', in *ECAI 2010 - 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16-20, 2010, Proceedings*, pp. 393–398, (2010).

[16] Hugo Martin and Patrice Perny, 'Biowa for preference aggregation with bipolar scales: Application to fair optimization in combinatorial domains', in *Proceedings of IJCAI 2019*, pp. 1822–1828, (2019).

[17] Hugo Martin and Patrice Perny, 'Computational models for cumulative prospect theory: Application to the knapsack problem under risk', in *proceedings of SUM (Scalable Uncertainty Management)*, pp. 52–65. Springer International Publishing, (2019).

[18] Brice Mayag, Antoine Rolland, and Julien Ah-Pine, 'Elicitation of a 2-additive bi-capacity through cardinal information on trinary actions', in *Advances in Computational Intelligence - 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2012, Catania, Italy, July 9-13, 2012, Proceedings, Part IV*, pp. 238–247, (2012).

[19] Kaisa Miettinen and Marko M Mäkelä, 'On scalarizing functions in multiobjective optimization', *OR spectrum*, **24**(2), 193–213, (2002).

[20] D. Schmeidler, 'Integral representation without additivity', *Proceedings of the American Mathematical Society*, **97**(2), 255–261, (1986).

[21] L.S. Shapley, 'Cores of convex games', *International Journal of Game Theory*, **1**, 11–22, (1971).

[22] Ralph E Steuer and RE Steuer, *Multiple criteria optimization: theory, computation, and application*, volume 233, Wiley New York, 1986.

[23] Ali Fallah Tehrani, Weiwei Cheng, Krzysztof Dembczyński, and Eyke Hüllermeier, 'Learning monotone nonlinear models using the choquet integral', *Machine Learning*, **89**(1-2), 183–211, (2012).

[24] Ali Fallah Tehrani and Eyke Hüllermeier, 'Ordinal choquistic regression', in *Proceedings of the 8th conference of the European Society for Fuzzy Logic and Technology, EUSFLAT-13, Milano, Italy, September 11-13, 2013*, (2013).

[25] Mikhail Timonin, 'Choquet integral as maximum of integrals with respect to belief functions', in *Belief Functions: Theory and Applications - Proceedings of the 2nd International Conference on Belief Functions, Compiègne, France, 9-11 May 2012*, pp. 117–124, (2012).

[26] Mikhail Timonin, 'Maximization of the choquet integral over a convex set and its application to resource allocation problems', *Annals OR*, **196**(1), 543–579, (2012).

[27] Mikhail Timonin, 'Robust optimization of the choquet integral', *Fuzzy Sets and Systems*, **213**, 27–46, (2013).

[28] Vicenç Torra, 'The weighted OWA operator', *Int. J. Intell. Syst.*, **12**(2), 153–166, (1997).

[29] Vicenç Torra and Yasuo Narukawa, *Modeling decisions: information fusion and aggregation operators*, Springer Science & Business Media, 2007.

[30] Amos Tversky and Daniel Kahneman, 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and uncertainty*, **5**(4), 297–323, (1992).

[31] Robert J Weber, 'Probabilistic values for games', *The Shapley Value. Essays in Honor of Lloyd S. Shapley*, 101–119, (1988).

[32] Lijue Xie and Michel Grabisch, 'The core of bicapacities and bipolar games', *Fuzzy Sets and Systems*, **158**(9), 1000–1012, (2007).

[33] M.E. Yaari, 'The dual theory of choice under risk', *Econometrica*, **55**, 95–115, (1987).

[34] Ronald R. Yager, 'On ordered weighted averaging aggregation operators in multicriteria decision making', *IEEE Trans. Systems, Man, and Cybernetics*, **18**(1), 183–190, (1988).