Discrete Applied Mathematics 161 (2013) 184–199

Contents lists available at SciVerse ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Compact versus noncompact LP formulations for minimizing convex Choquet integrals

J. Lesca, M. Minoux, P. Perny *

LIP6 – Pierre et Marie Curie University, 4 Place Jussieu, 75005 Paris, France

ARTICLE INFO

Article history:
Received 28 March 2011
Received in revised form 24 June 2012
Accepted 8 July 2012
Available online 9 August 2012

Keywords:
Choquet integral
Submodular capacity
Linear programming
Multicriteria optimization
Compromise search

ABSTRACT

We address here the problem of minimizing Choquet Integrals (also known as "Lovász Extensions") over solution sets which can be either polyhedra or (mixed) integer sets. Typical applications of such problems concern the search of compromise solutions in multicriteria optimization. We focus here on the case where the Choquet Integrals to be minimized are convex, implying that the set functions (*or "capacities"*) underlying the Choquet Integrals considered are *submodular*. We first describe an approach based on a large scale LP formulation, and show how it can be handled via the so-called column-generation technique. We next investigate alternatives based on *compact LP formulations*, i.e. featuring a polynomial number of variables and constraints. Various potentially useful special cases corresponding to well-identified subclasses of underlying set functions are considered: quadratic and cubic submodular functions, and a more general class including set functions which, up to a sign, correspond to capacities which are both $(k+1)$–additive and k-monotone for $k \geq 3$. Computational experiments carried out on series of test instances, including transportation problems and knapsack problems, clearly confirm the superiority of compact formulations. As far as we know, these results represent the first systematic way of practically solving Choquet minimization problems on solution sets of significantly large dimensions.

1. Introduction

The purpose of the present paper is to investigate various linear programming-based formulations for the problem of minimizing the so-called *Choquet integral* w.r.t. a given capacity or, equivalently, the *Lovász extension* of a given set function. We mainly focus on the *convex case*, which corresponds to *submodular* set functions (referred to as concave capacities in the field of Decision Theory). In such a case, and assuming that we want to optimize on a polyhedron in $\mathbb{R}^n$ (represented by a given linear equality/inequality system), the problem reduces to minimizing a convex function $\mathbb{R}^n \to \mathbb{R}$ given through an oracle providing the value of the function and an associated subgradient in any $x \in \mathbb{R}^n$; it is well known that such a problem can be reformulated as a large scale linear program (*LP*) with exponentially many constraints (we call this a "*noncompact*" *formulation*) which is solvable using a polynomial number of calls to the oracle via the *ellipsoid algorithm* (see [12]). In practice, these large scale ("noncompact") linear programs are solved by applying the well-known column-generation approach to their duals.

However, as will be confirmed by the computational experiments reported in Section 5, the column-generation approach is computationally demanding, and thus only practicable for problems of limited size. In view of this, one is naturally led to investigating the possibility of deriving alternative formulations, with the objective of coming up with LP models featuring *polynomial number of variables and constraints*, referred to here as *compact* LP models.

* Corresponding author. Fax: +33 144277004.
E-mail addresses: Julien.Lesca@lip6.fr (J. Lesca), Michel.Minoux@lip6.fr (M. Minoux), Patrice.Perny@lip6.fr (P. Perny).

0166-218X/$ – see front matter
doi:10.1016/j.dam.2012.07.007

The main contribution of the present paper is to propose such compact formulations for various cases of interest, both theoretically and practically, in particular the case of degree 2 and degree 3 submodular set functions (2-additive and 3-additive capacities respectively); more general classes of set functions will also be investigated.

It is worth mentioning that all the results in the present paper would readily apply to the symmetric problem of maximizing a Choquet integral with respect to a supermodular set function (convex capacity).

The paper is organized as follows: Section 2 recalls the main definitions and basic properties related to Lovász extensions and Choquet integrals. In Section 3, a first (large scale) linear programming (LP) formulation for the problem of minimizing a Choquet integral is proposed, and it is shown how the column-generation principle can be applied, based on polynomial solvability of submodular function minimization. As possible alternatives to this column-generation approach, several compact formulations are investigated in Section 4, in connection with various classes of underlying set functions. The cases of quadratic and cubic submodular functions are addressed in Sections 4.2 and 4.3. Section 4.4 is then devoted to a more general class of submodular set functions (with degree higher than 3) which can be recognized in polynomial time and for which Choquet minimization can still be expressed in compact formulation. The results in Section 4.4 are shown to apply, in particular, to the whole subclass of problems of maximizing a Choquet integral w.r.t. any capacity which is both $(k+1)$-additive and k-monotone (for $k \geq 3$). Section 5 reports on computational experiments carried out on series of test instances of the transportation problem and of the knapsack problem, both in their continuous and discrete versions. The results obtained clearly confirm the attractiveness of compact formulations, both in terms of computational efficiency and of easy handling of possible integrality requirements on decision variables.

2. Definitions and basic properties

In the field of combinatorial optimization, the so-called *Lovász extension* of a set function has been introduced [16] in connection with the study of some remarkable polyhedra associated with submodular functions (polymatroids) and their links to the so-called "greedy algorithm" [6]. Almost the same concept was (independently) investigated in the field of Decision Analysis under the name of Choquet integral (CI). One of the important applications of the Choquet integral is in *multicriteria decision making* (MCDM) since it provides a systematic way of aggregating multiple criteria enjoying various nice properties (see [8]). Among others, it includes several other aggregators as particular instances (Min, Max and any order statistics, weighted sums and ordered weighted averages like OWA [38] and WOWA [35]); it can handle both positive and negative interactions among criteria [9], it offers the possibility to control andness or orness in aggregating values [18] which can be used to control the type of compromise looked for [7]. We also mention that Choquet integrals appear in the literature on decision making under risk or uncertainty, in models known as Yaari's model [37], "Rank-Dependent Expected Utility" (RDEU, [25]) and "Choquet Expected Utility" (CEU, [29]). These models generalize the so-called Savage's *Expected Utility* model and provide enhanced descriptive possibilities [30,4].

The close connections between the Lovász extension of a pseudo-Boolean function and the Choquet integral w.r.t. the associated set function have been observed by many authors (cf. e.g. [17]). Here we first introduce definitions and basic properties related to Lovász extensions, and then provide the corresponding terminology related to the Choquet Integral, commonly in use in the field of Decision Theory.

Let $E = \{1, 2, \dots, n\}$ a given n-element set ("ground set") and suppose that we are given a set function or pseudo-Boolean function $v : \{0,1\}^E \to \mathbb{R}$. Considering any subset $S \subseteq E$, and denoting $\mathbb{1}_S \in \{0,1\}^E$ the characteristic vector of S, the simplified notation $v(S)$ will be frequently used in the sequel instead of $v(\mathbb{1}_S)$. The Lovász extension of the set function v is the function $\mathcal{L}_v : [0,1]^n \to \mathbb{R}$ defined as follows.

For any $x \in [0,1]^n$, we denote $x^{\uparrow} = (x_1^{\uparrow}, x_2^{\uparrow}, \dots, x_n^{\uparrow})$ the n-vector the components of which are those of x sorted according to nondecreasing order: $x_1^{\uparrow} \leq x_2^{\uparrow} \leq \cdots \leq x_n^{\uparrow}$. Also, for any $i \in E$, we denote $X_i^{\uparrow}$ the set $\{j \in E, x_j \geq x_i^{\uparrow}\}$.

The Lovász extension is then defined as

$$\mathcal{L}_v(x) = \sum_{i=1}^{n-1} x_i^{\uparrow}[v(X_i^{\uparrow}) - v(X_{i+1}^{\uparrow})] + x_n^{\uparrow} v(X_n^{\uparrow}). \tag{1}$$

It can be observed that $\mathcal{L}_v$ is the unique function: $[0,1]^n \to \mathbb{R}$ which takes on the same values as v on $\{0,1\}^n$ and which linearly interpolates v on each canonical simplex S_σ involved in the standard triangulation of $[0,1]^n$. We recall that the canonical simplex associated with any permutation σ of $\{1, 2, \dots, n\}$ is

$$S_\sigma = [0,1]^n \cap \left\{x / x_{\sigma(1)} \leq x_{\sigma(2)} \leq \cdots \leq x_{\sigma(n)}\right\}.$$

Since the main focus of the present paper is on efficient algorithms for *minimizing* Lovász extensions (or Choquet integrals), we will restrict to the special case when $\mathcal{L}_v(x)$ in (1) is *convex* in x. The fact that this special case corresponds to the subclass of *submodular* set functions is a well-known result due to Lovász [16] (see also [33]).

In view of formula (1), assuming that for any subset $S \subseteq V$, $v(S)$ can be computed in time polynomial in $n = |E|$, it is clear that for any $x \in \mathbb{R}_+^n$ the value $\mathcal{L}_v(x)$ can be computed in polynomial time. Now, assuming v submodular such that $v(\emptyset) = 0$, we recall below how a *subgradient* of $\mathcal{L}_v(x)$ can be derived. For that, we use the so-called *base-polyhedron* associated with v which is defined as

$$B(v) = \left\{y \in \mathbb{R}^E / y(S) \le v(S)\ \forall S \subset E;\ y(E) = v(E)\right\}$$

(where, $\forall y \in \mathbb{R}^E$, $y(S) = \sum_{i \in S} y_i$) and the following property:

Property 2.1 (*Lovász [16], Murota [20]*)**.**

$$\mathcal{L}_v(x) = \sup_{y \in B(v)} \left\{y^T x\right\}. \tag{2}$$

From expression (2) it follows that a subgradient of $\mathcal{L}_v(x)$ in any x such that $\mathcal{L}_v(x) < +\infty$ is obtained as

$$y^* = \underset{y \in B(v)}{\operatorname{argmax}} \left\{y^T x\right\}. \tag{3}$$

Now it is well known (see [6]) that, for submodular v, the maximization in (2) can be carried out in *polynomial time* using the so-called *Greedy Algorithm* recalled below.

Greedy algorithm for (2).

(a) Sort the components of x according to non-increasing order:
$x_{k_1} \ge x_{k_2} \ge \cdots \ge x_{k_n}$.
(b) Compute

$$\begin{aligned} y^*_{k_1} &= v(\{k_1\}) - v(\emptyset) \\ y^*_{k_2} &= v(\{k_1, k_2\}) - v(\{k_1\}) \\ &\vdots \\ y^*_{k_n} &= v(\{k_1, k_2, \ldots, k_n\}) - v(\{k_1, k_2, \ldots, k_{n-1}\}). \end{aligned}$$

We note that, since $x_i^{\uparrow} = x_{k_{n-i+1}}$ and $X_i^{\uparrow} = \{k_1, k_2, \ldots, k_{n-i+1}\}$, the expression of $\mathcal{L}_v(x)$ in (1) can be identified with $\sum_{i=1}^n x_{k_i} y^*_{k_i} = (y^*)^T x$, i.e. the maximum value in (2).

As a consequence of the work by Grötschel et al. [12] it is well-known that minimizing a convex function $f : \mathbb{R}^n \to \mathbb{R}$ on a given polyhedron in $\mathbb{R}^n$ can be done in *polynomial time* using the so-called "Ellipsoid Algorithm" [14] (assuming that computing $f(x)$ and a corresponding subgradient $y \in \partial f(x)$ in any $x \in \mathbb{R}^n$ can be carried out in polynomial time).

Thus, for submodular $v : \{0, 1\}^E \to \mathbb{R}$ we know that, when X is a given nonempty polyhedron in $\mathbb{R}^n$ (specified by a given linear equality/inequality system) the problem:

$$(P) \begin{cases} \text{Minimize } \mathcal{L}_v(x) \\ \text{s.t. :} \\ x \in X \end{cases}$$

can be solved in polynomial time.

We note that in the literature on Decision Theory, (see e.g. [29,37,25,8,10]) the formula defining $C_v(x)$, the *Choquet integral* of $x \in \mathbb{R}^n$ with respect to a given set function v (referred to as a "capacity" in this context) is exactly the same as (1), the only difference being that $C_v(x)$ is possibly defined for any $x \in \mathbb{R}^n_+$ instead of being restricted to $x \in [0, 1]^n$. This difference is not really significant. The following straightforward property of the Choquet integral indeed shows that, when bounded domains for x are considered, it is always possible to scale the feasible set to $[0, 1]^n$, leading to an equivalent problem with a Lovász extension.

Property 2.2. *For any $x \in \mathbb{R}^n_+$, let $C_v(x) = \sum_{i=1}^{n-1} x_i^{\uparrow}[v(X_i^{\uparrow}) - v(X_{i+1}^{\uparrow})] + x_n^{\uparrow} v(X_n^{\uparrow})$. Then, for any pair of reals $\alpha > 0$ and β, $C_v(\alpha x + \beta \mathbb{1}_E) = \alpha C_v(x) + \beta v(E)$.*

Proof. Follows directly from the definition. □

In view of this, if $C_v(x)$ has to be minimized over $x \in X \subset [\ell, u]^n$ then carrying out the variable redefinition:

$$y = \frac{1}{u - \ell}(x - \ell \mathbb{1}_E)$$

the problem is reduced to minimizing $C_v(y)$ where $y \in [0, 1]^n$, which is exactly the Lovász extension of v. So, in all the developments to follow, it will be assumed w.l.o.g. that the solution sets X over which $C_v(x)$ is to be minimized are bounded polyhedra or MIP sets *included in the nonnegative orthant* $\mathbb{R}^E_+$. Also we note that, in Decision Theory, applications involving the Choquet integral as an aggregation function often introduce several additional restrictions on the set-function v, in particular: (i) monotonicity ($v(A) \le v(B)$ for all $A \subseteq E, B \subseteq E,\ A \subseteq B$), (ii) the normalization condition $v(E) = 1$ and sometimes (iii) the submodularity of v. Set functions satisfying (i) and (ii) are called *capacities* in this context. Also in Decision Theory, a capacity is called *convex* (resp. *concave*) if and only if the associated set function is *supermodular* (resp. submodular).

The rationality of conditions (i)–(iii) can easily be explained in the context of multicriteria optimization. Assuming that $x = (x_1, \dots, x_n)$ represents a cost vector where x_i is the cost of solution x with respect to criterion i for any $i \in E$, a cost vector x is preferred to a cost vector y if and only if $C_v(x) \leq C_v(y)$. In this context, it can easily be shown from Eq. (1) that condition (i) is equivalent to Pareto-monotonicity which reads as follows:

$$\forall x, y \in \mathbb{R}^n_+, \quad [\forall i \in E, x_i \leq y_i \Rightarrow C_v(x) \leq C_v(y)].$$

Moreover, assuming that (i) holds, (ii) is only a normalization condition ensuring that $C_v(\mathbb{1}_E) = 1$ and $\forall x \in \mathbb{R}^n_+$, $\mathrm{Min}_{i\in E} x_i \leq C_v(x) \leq \mathrm{Max}_{i\in E} x_i$. Now, assuming that (i) and (ii) hold, imposing (iii), i.e. submodularity of the capacity in the Choquet integral is equivalent to imposing the following property (see [4,7]):

$$\forall x^1, \dots, x^p \in \mathbb{R}^n_+, \qquad \forall k \in \{1, \dots, p\}, \qquad \forall \lambda \in \mathbb{R}^p_+ \quad \text{s.t.} \sum_{i=1}^{p} \lambda_i = 1,$$

$$C_v(x^1) = \cdots = C_v(x^p) \Rightarrow C_v\left(\sum_{i=1}^{p} \lambda_i x^i\right) \leq C_v(x^k). \tag{4}$$

Property (4) named *preference for well-balanced solutions* in [7] means that any compromise cost vector obtained by a convex combination of p equivalent cost vectors (i.e. vectors having the same C_v value) improves these vectors. Due to this property, whenever we are indifferent between (20, 0) and (0, 20), we should prefer (10, 10) = 0.5(20, 0) + 0.5(0, 20) to the two initial vectors. This is a way of enforcing equity of solutions minimizing $C_v(x)$. Let us also mention that, in the field of cooperative games, the *base polyhedron* $B(v)$ has also emerged as a central concept, referred to as the *core* of the game associated with the capacity $\bar{v}$ ($\bar{v}$, the so-called *dual capacity*, is defined by $\bar{v}(S) = 1 - v(E \setminus S)$, $\forall S \subseteq E$). For any submodular (concave) capacity v, the dual capacity $\bar{v}$ is supermodular (convex) and the core of $\bar{v}$ is known to be nonempty (see [32]).

In the developments to follow restrictions such as (i) and (ii) will not be assumed and general submodular set functions will be considered.

3. Large scale (noncompact) LP formulation and the column-generation approach

Keeping in mind that v is assumed to be submodular and $v(E) = 1$, the first formulation proposed here makes use of the LP duality theorem applied to the problem defining $C_v(x)$ for given x:

$$\text{(I)} \begin{cases} C_v(x) = \max x^T \cdot y \\ \text{s.t. :} \\ \sum_{i\in S} y_i \leq v(S) \quad \forall S \subseteq E \\ \sum_{i\in E} y_i = 1 \\ y \quad \text{unconstrained in sign.} \end{cases}$$

The LP dual to (I) reads

$$(D) \begin{cases} \mathrm{Min} \sum_{S\subseteq E} v(S)\lambda_S \\ \text{s.t. :} \\ \forall i = 1, \dots, n : \quad \sum_{S/i\in S} \lambda_S = x_i \\ \lambda_S \geq 0, \quad \forall S \subset E, \\ \lambda_E \lesseqgtr 0 \end{cases}$$

where there is one variable λ_S associated with each subset $S \subseteq E$.

Using the standard convention that (D) has optimal solution value $+\infty$ whenever its solution set is empty, then $C_v(x)$ is equal to the optimal dual solution value, therefore problem (P) can equivalently be stated as the (large scale) LP problem:

$$\text{(II)} \begin{cases} \text{Minimize} \sum_{S\subseteq E} v(S)\lambda_S \\ \text{s.t. :} \\ \sum_{S/i\in S} \lambda_S - x_i = 0 \quad (i = 1, \dots, n) \\ \lambda_S \geq 0 \quad \forall S \subset E \\ \lambda_E \lesseqgtr 0 \\ x \in X \end{cases} \tag{5}$$

(minimization in (II) is with respect to variables λ and x). It turns out that the above can be solved exactly (and efficiently in practice) by applying the so-called column-generation procedure because of the following property:

Property 3.1. *Let* $(\mathrm{II})'$ *any restricted version of problem* (II)*, where only a restricted number of variables* λ_S *are allowed to be nonzero. Let* $\pi = (\pi_1, \dots, \pi_n)$ *denote the optimal simplex multipliers (optimal dual variables) corresponding to the constraints* (5) *in* $(\mathrm{II})'$*. Then the problem of determining a subset* $S^* \subseteq E$ *achieving minimum reduced cost:*

$$\gamma\left(S^*\right) = v\left(S^*\right) - \sum_{i \in S^*} \pi_i = \underset{S \subseteq E}{\mathrm{Min}}\{\gamma(S)\} = \underset{S \subseteq E}{\mathrm{Min}}\left\{v(S) - \sum_{i \in S} \pi_i\right\}$$

can be solved in polynomial time, assuming that v *is computable via a polynomial-time oracle.*

Proof. v being submodular, the set function γ representing the reduced costs of the λ_S variables is also submodular as the sum of a submodular function and of a modular function: $S \to \sum_{i \in S} \pi_i$. The result then follows from the existence of efficient, strongly polynomial algorithms for minimizing submodular functions; see e.g. [23]. □

Computational results obtained with the column-generation approach will be discussed in Section 5, and compared with those obtained using the compact formulations presented in the next section.

4. Compact LP formulations

In the present section, we investigate the possibility of deriving alternative LP models for the problem of minimizing a convex Choquet integral over a given solution set (polyhedron or mixed integer set). More precisely the main focus is on exhibiting conditions under which it is possible to come up with *compact formulations* i.e. LP models requiring numbers of extra variables and constraints *polynomially bounded* in n (the cardinality of the ground set E) and K (the size of an explicit description of the set function v).

The various compact formulations proposed below rely on the availability of an explicit multilinear polynomial expression of the set function v of manageable size, i.e. featuring a number of terms polynomial in $n = |E|$. This amounts to assuming that we are given an expression of v of the form:

$$\forall u \in \{0, 1\}^n : v(u) = \sum_{p=1}^{P} \mu_p \left(\prod_{i \in S_p} u_i\right) \tag{6}$$

where, $\forall p = 1, \dots, P$, $\mu_p \in \mathbb{R}$ and $S_p \subseteq E$ are given. We recall here the well-known fact (see [13]) that any set function or pseudo-Boolean function $v : \{0, 1\}^n \to \mathbb{R}$ has a unique multilinear polynomial representation of the form (6).

4.1. Choquet integral and Möbius transforms

Consider a set function v and its associated pseudo-Boolean representation given by (6). Let $m : \mathcal{P}(E) \to \mathbb{R}$ be the set function defined as

$$\begin{cases} m(S_p) = \mu_p & \forall p = 1, \dots, P \\ m(S) = 0 & \forall S \in \mathcal{P}(E) \setminus \{S_1, \dots, S_P\}. \end{cases} \tag{7}$$

Then (6) shows that an equivalent expression of v in terms of m is

$$\forall A \subseteq E : v(A) = \sum_{B \subseteq A} m(B). \tag{8}$$

The set function $m : \mathcal{P}(E) \to \mathbb{R}$ defined by (7) is called the *Möbius transform* of v (see [28,1,10]). Indeed, to any set function v, it is possible to let correspond its Möbius transform m defined as

$$\forall A \subseteq E : m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} v(B)$$

and conversely, given m, v can be "reconstructed" via formula (8). The values $m(A)$ associated with the subsets of E are called *Möbius masses*, and in view of (7), it is seen that the coefficients μ_p involved in the multilinear polynomial expression (6) can be interpreted as the Möbius masses of the subsets $S_p (p = 1, \dots, P)$.

So, in this section, we will restrict to considering set functions featuring a polynomial number P of nonzero Möbius masses. Also we will investigate special cases of interest corresponding to set functions for which this associated multilinear polynomial expression (6) has small degree: $k = 2$ (see Section 4.2) or $k = 3$ (see Section 4.3).

Remark 1. In the literature on Decision Analysis, a set function having nonzero Möbius masses only for subsets of cardinality $\leq k$ is called k-additive (see e.g. [11]).

The following result, which can be traced back to Rosenmüller [27], Smets [34], Dubois and Prade [5], Chateauneuf and Jaffray [3], is key to the developments to follow. It applies to arbitrary set functions, not only to submodular ones, and provides a simple expression of $C_v(x)$, the value of the Choquet integral (or Lovász extension) in any $x \in \mathbb{R}_+^E$ in terms of the Möbius transforms of v:

Property 4.1. *Let $v : \mathcal{P}(E) \to \mathbb{R}$ and $m : \mathcal{P}(E) \to \mathbb{R}$ its Möbius transform. Then, $\forall x \in \mathbb{R}_+^E$:*

$$C_v(x) = \sum_{A \subseteq E} m(A) \ \underset{i \in A}{\mathrm{Min}}\{x_i\}. \tag{9}$$

In view of Property 4.1 the value of the Choquet integral w.r.t. a set function v specified by its multilinear polynomial expression (6) is given for any $x \in \mathbb{R}_+$ by

$$C_v(x) = \sum_{i=1}^{n} \mu_i x_i + \sum_{p=n+1}^{P} \mu_p \ \underset{i \in S_p}{\mathrm{Min}}\{x_i\} \tag{10}$$

where it is assumed w.l.o.g. that $S_i = \{i\}$ for $i = 1, \dots, n$. The terms indexed $p = n + 1, \dots, P$ thus correspond to the nonlinear terms in expression (6).

4.2. *First compact formulation* (CF1)*: the case of quadratic submodular v and of "belief functions"*

In this section, we exhibit a first compact formulation for the problem:

$$\begin{cases} \text{Minimize } C_v(x) \\ x \in X \subset \mathbb{R}_+^E \end{cases}$$

for the whole class of submodular set functions in the subclass $\mathcal{B}^-$ defined as follows:

Definition 1. A set function v is said to belong to the class $\mathcal{B}^-$ (resp. $\mathcal{B}^+$) if and only if its multilinear polynomial expression (6) contains a number of terms polynomial in n, only featuring:

(i) linear terms of the form $\alpha_i u_i$ without sign restriction on the α_i coefficients;
(ii) nonlinear terms of the form $\mu_S \prod_{i \in S} u_i$ with $\mu_S < 0$ (resp. $\mu_S > 0$).

We denote $\mathcal{B}_K^-$ (resp. $\mathcal{B}_K^+$) the subclass of $\mathcal{B}^-$ (resp. $\mathcal{B}^+$) of set functions having multilinear polynomial expressions featuring nonlinear terms of degree at most K.

We note that the class $\mathcal{B}^-$ includes as a special case *quadratic submodular* set functions for which the n linear terms (corresponding to $S_i = \{i\}$, for $i = 1, \dots, n$) are not sign-restricted, and the quadratic terms (indexed $p = n + 1, \dots, P$) are strictly negative.

We also note that set functions in the class $\mathcal{B}^-$ are sometimes referred to as Rhys functions or "positive–negative" pseudo-Boolean functions in the OR literature, and arise in connection with some interesting applications (see e.g. [26]). On the other hand, note that the class $\mathcal{B}^+$ includes all set functions for which all coefficients μ_P in (6) are nonnegative; such set functions, whose Möbius masses are non-negative, are called *belief functions* [31] in Artificial Intelligence and Decision Theory. So the compact formulation (CF1) discussed below directly applies in particular to the problem of maximizing Choquet integrals with respect to belief functions.

In the sequel please keep in mind that the linear terms in expression (6) are indexed $i = 1, \dots, n$ and the nonlinear terms are indexed $p = n + 1, \dots, P$.

In view of Property 4.1, and if we assume that $v \in \mathcal{B}^-$, for *any fixed* $x \in \mathbb{R}_+^n$, the value $C_v(x)$ is the minimum value of the linear program in y variables:

$$\begin{cases} \text{Minimize } \displaystyle\sum_{i=1}^{n} \mu_i x_i + \sum_{p=n+1}^{P} \mu_p y_p \\ \text{s.t. :} \\ 0 \le y_p \le x_i \qquad \forall p = n+1, \dots, P, \qquad \forall i \in S_p. \end{cases} \tag{11}$$

Observe that in the above LP, there are $P - n$ continuous variables y, one for each nonlinear term of the multilinear polynomial expression (6).

Now, the correctness of the above follows from the fact that $\mu_p < 0$ for $p = n + 1, \dots, P$, and therefore the minimum in the second term of expression (11) is obtained when each y_p takes its maximum possible value, which is exactly $\mathrm{Min}_{j \in S_p}\{x_j\}$.

In view of this, the problem of minimizing $C_v(x)$ over a polyhedron $X \subset \mathbb{R}^n_+$ described by a set of m linear constraints of the form $Ax \le b$, $x \ge 0$ can be stated as

$$(\mathbf{CF1}) \begin{cases} \underset{(x,y)}{\mathrm{Min}} \displaystyle\sum_{i=1}^{n} \mu_i x_i + \sum_{p=n+1}^{P} \mu_p y_p \\ \text{s.t. :} \\ y_p \le x_i \quad \forall p = n+1, \ldots, P, \ \forall i \in S_p \\ Ax \le b \\ x \ge 0, \qquad y \ge 0 \end{cases}$$

which is an ordinary linear program featuring P variables and $\mathcal{O}(m + n(P - n))$ constraints (in the case of quadratic submodular v, the number of constraints, not including nonnegativity conditions, is exactly $m + 2(P - n)$). This problem can therefore be solved in *polynomial time*, using an interior-point algorithm (see e.g. [39,36]). In practice, if polynomiality is not explicitly required, one can also consider the use of a state-of-the art implementation of the simplex algorithm.

It should be observed that the same model can also be used to minimize $C_v(x)$ over a *mixed integer set* of the form:

$$X' = X \cap \{x/x_i \in \mathbb{Z}, \ \forall i \in I\}$$

or

$$X'' = X \cap \{x/x_i \in \{0, 1\}, \ \forall i \in I\}$$

where I denotes the index subset of those variables which are subject to integrality restrictions. Of course, in such cases, polynomial solvability is lost in general. Part of the computational experiments discussed in Section 5 concern instances featuring such integrality requirements.

4.3. *Second compact formulation* (CF2)*: the case of general cubic submodular v*

We now address a more complex situation when the set function v under consideration is *cubic submodular*. In this case, all the coefficients of the nonlinear terms in the multilinear expression (6) are not necessarily negative, and the technique shown in Section 4.2 does not apply anymore. As an example of this, consider the cubic pseudo-Boolean function:

$$v(u) = 6u_1 + 3u_2 + u_3 + 4u_4 - 5u_1u_2 - 3u_1u_3 - 2u_2u_3 - 3u_1u_4 - 4u_2u_4 - u_2u_3u_4 + 2u_1u_2u_3 + 3u_1u_2u_4.$$

Indeed this function, which includes positive as well as negative nonlinear terms, is submodular because it satisfies the following characterization of cubic submodular functions:

Property 4.2 (*Billionnet and Minoux [2]*)**.** *A cubic pseudo-Boolean function $v : \{0, 1\}^n \to \mathbb{R}$ given by its multilinear expression (6) is submodular if and only if, for any p such that $|S_p| = 2$:*

$$\tilde{\mu}_p = \mu_p + \sum_{p' \in L_p^+} \mu_{p'} \le 0 \tag{12}$$

with $L_p^+ = \left\{p'/S_p \subset S_{p'}, \ \mu_{p'} > 0\right\}$.

Condition (12) can easily be checked for the example:

For $S_p = \{1, 2\}$, $\mu_p = -5$, there are two $S_{p'}$ with positive $\mu_{p'}$ strictly containing $\{1, 2\}$, namely: $\{1, 2, 3\}$, with weight 2 and $\{1, 2, 4\}$ with weight 3; so (12) holds in this case.

For $S_p = \{1, 3\}$, $\mu_p = -3$, $\{1, 2, 3\}$ is the only $S_{p'}$ with positive weight 2 strictly containing $\{1, 3\}$, and (12) again holds.

The same conclusion would be obtained for $S_p = \{1, 4\}\{2, 3\}\{2, 4\}$ and $\{3, 4\}$, implying that v is indeed submodular.

Remark 2. There are close connections between Property 4.2 and the concept of k-monotonicity investigated in [3]: v is k-monotone if and only if its Möbius transform m satisfies

$$\sum_{L: A \subseteq L \subseteq B} m(L) \ge 0, \quad \forall A, B \subseteq E, \ A \subseteq B, \ 2 \le |A| \le k. \tag{13}$$

Property 4.2 is exactly the one above for $k = 2$ and the reverse inequality. Therefore, the results of this section are directly applicable to maximizing Choquet integrals with respect to capacities which are both 3-additive and 2-monotone.

Now, we are going to show how Property 4.2 can be used to derive a compact LP formulation (CF2).

The first step consists in transforming the expression of $C_v(x)$ given by (10) according to the following result.

Proposition 1. *Let v cubic submodular given by (6), and denote: $Q_2 \subset \{n+1,\dots,P\}$ the subset of indices p such that $|S_p| = 2$; Q_3^+ (resp. Q_3^-) the subset of indices in $\{n+1,\dots,P\}$ such that $|S_p| = 3$ and $\mu_p > 0$ (resp. $|S_p| = 3$ and $\mu_p < 0$). Then an equivalent expression for (10) is*

$$C_v(x) = \sum_{i=1}^{n} \mu_i x_i + \sum_{p\in Q_2} \tilde{\mu}_p \operatorname*{Min}_{i\in S_p}\{x_i\} + \sum_{p\in Q_3^-} \mu_p \operatorname*{Min}_{i\in S_p}\{x_i\} - \sum_{p\in Q_3^+} \mu_p \left[\operatorname*{Min}_{i\in S_p}\{x_i\} + \operatorname*{Min_2}_{i\in S_p}\{x_i\}\right] \tag{14}$$

where $\tilde{\mu}_p$ has been defined in (12) and Min_2 denotes the operator which selects the second minimum value among a finite set of real values.

Proof. Given any three real values z_1, z_2, z_3 it is easily checked that:

$$\mathrm{Min}\{z_1,z_2\} + \mathrm{Min}\{z_1,z_3\} + \mathrm{Min}\{z_2,z_3\} = 2\mathrm{Min}\{z_1,z_2,z_3\} + \mathrm{Min}_2\{z_1,z_2,z_3\}.$$

It follows that, for any $p \in Q_3^+$, assuming $S_p = \{j,k,l\}$, the term $\mu_p\,\mathrm{Min}\{x_j,x_k,x_l\}$ can be replaced with the expression:

$$\mu_p\left[\mathrm{Min}\{x_j,x_k\} + \mathrm{Min}\{x_j,x_l\} + \mathrm{Min}\{x_k,x_l\} - \mathrm{Min}\{x_j,x_k,x_l\} - \mathrm{Min}_2\{x_j,x_k,x_l\}\right].$$

Suppose we carry out this substitution for every $p \in Q_3^+$. We note that in the new resulting expression for $C_v(x)$, there are three quadratic terms originating from a cubic term of the form $\mu_p\,\mathrm{Min}\{x_j,x_k,x_l\}$ such that $\mu_p > 0$, namely: $\mu_p\,\mathrm{Min}\{x_j,x_k\}$; $\mu_p\,\mathrm{Min}\{x_j,x_l\}$; $\mu_p\,\mathrm{Min}\{x_k,x_l\}$.

Because v is submodular, Property 4.2 guarantees that each such "new" quadratic term involves a pair of variables corresponding to some existing quadratic term $\mu_q\,\mathrm{Min}_{i\in S_q}\{x_i\}$ in the original expression (10).

Moreover, in the transformed expression, the new coefficient of the term $\mathrm{Min}_{i\in S_q}\{x_i\}$ is equal to $\tilde{\mu}_q = \mu_q + \sum_{p\in L_q^+}\mu_q$, and $\tilde{\mu}_q \le 0$ in view of (12) (because of the submodularity property). So the quadratic part of the transformed expression reduces to: $\sum_{p\in Q_2}\tilde{\mu}_p\,\mathrm{Min}_{i\in S_p}\{x_i\}$ with all $\tilde{\mu}_p \le 0$. In addition to this, each cubic term $\mu_p\,\mathrm{Min}_{i\in S_p}\{x_i\}$ with $\mu_p > 0$ gives rise to the term $-\mu_p\left[\mathrm{Min}_{i\in S_p}\{x_i\} + \operatorname*{Min_2}_{i\in S_p}\{x_i\}\right]$ therefore resulting in expression (14). □

We observe that in expression (14), if we ignore the last summation:

$$-\sum_{p\in Q_3^+} \mu_p \left[\operatorname*{Min}_{i\in S_p}\{x_i\} + \operatorname*{Min_2}_{i\in S_p}\{x_i\}\right]$$

we find the right structure for applying the compact formulation (CF1), since all coefficients $\tilde{\mu}_p$ for $p \in Q_2$ and μ_p for $p \in Q_3^-$ are *nonpositive*.

The second step now consists in building a LP formulation for handling each of the terms of the form:

$$-\mu_p\left[\operatorname*{Min}_{i\in S_p}\{x_i\} + \operatorname*{Min_2}_{i\in S_p}\{x_i\}\right]$$

for $p \in Q_3^+$.

Proposition 2 shows how this can be done.

Proposition 2. *For $\mu_p > 0$ any term of the form:*

$$-\mu_p\left[\mathrm{Min}\{x_j,x_k,x_l\} + \mathrm{Min}_2\{x_j,x_k,x_l\}\right]$$

can be represented as the optimal value of the linear program (in variables $\lambda_0, \lambda_1, \lambda_2, \lambda_3$):

$$(\mathrm{III})\begin{cases}\mathrm{Min}\, 2\mu_p\lambda_0 + \mu_p\lambda_1 + \mu_p\lambda_2 + \mu_p\lambda_3\\ s.t.:\\ \lambda_0 + \lambda_1 \ge -x_j\\ \lambda_0 + \lambda_2 \ge -x_k\\ \lambda_0 + \lambda_3 \ge -x_l\\ \lambda_1 \ge 0, \qquad \lambda_2 \ge 0, \qquad \lambda_3 \ge 0\\ \lambda_0 \lesseqgtr 0 \quad (\textit{i.e. without sign restriction}).\end{cases}$$

Proof. For any fixed $x_j, x_k, x_l \ge 0$ let $\alpha = \mathrm{Min}\{x_j,x_k,x_l\} + \mathrm{Min}_2\{x_j,x_k,x_l\}$. Thus α represents the sum of the two smallest values among x_j, x_k, x_l and therefore is equal to the optimal solution value of the continuous knapsack problem in bounded variables u_j, u_k, u_l:

$$\begin{cases}\mathrm{Min}\, x_j u_j + x_k u_k + x_l u_l\\ \text{s.t.}:\\ u_j + u_k + u_l = 2\\ 0 \le u_j \le 1, \qquad 0 \le u_k \le 1, \qquad 0 \le u_l \le 1.\end{cases}$$

Equivalently $-\alpha$ is equal to the optimal solution value to:

$$(\text{IV})\begin{cases}\text{Max} - x_j u_j - x_k u_k - x_l u_l\\ \text{s.t. :}\\ u_j + x_k + u_l = 2\\ u_j \leq 1; \qquad u_k \leq 1; \qquad u_l \leq 1\\ u_j \geq 0 \qquad u_k \geq 0 \qquad u_l \geq 0.\end{cases}$$

Denoting $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ the dual variables associated with the constraints of (IV), and using the LP duality theorem, α is also equal to the optimal solution value of the dual:

$$(\text{V})\begin{cases}\text{Min} 2\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3\\ \text{s.t. :}\\ \lambda_0 + \lambda_1 \geq -x_j\\ \lambda_0 + \lambda_2 \geq -x_k\\ \lambda_0 + \lambda_3 \geq -x_l\\ \lambda_0 \lesseqgtr 0, \qquad \lambda_1 \geq 0, \qquad \lambda_2 \geq 0, \qquad \lambda_3 \geq 0\end{cases}$$

which proves the claim. $\square$

In view of Proposition 2, for any fixed $x \geq 0$, it is possible to express the last summation in (14) as the minimum value of a linear program deduced from (V) by associating with each $p \in Q_3^+$ four continuous variables $\lambda_0^p, \lambda_1^p, \lambda_2^p, \lambda_3^p$, which reads:

$$\begin{cases}\sum_{p\in Q_3^+} \mu_p \left[2\lambda_0^p + \lambda_1^p + \lambda_2^p + \lambda_3^p\right]\\ \text{s.t. :}\\ \left.\begin{array}{l}\lambda_0^p + \lambda_1^p \geq -x_{\alpha[p]}\\ \lambda_0^p + \lambda_2^p \geq -x_{\beta[p]}\\ \lambda_0^p + \lambda_3^p \geq -x_{\gamma[p]}\end{array}\right\} \quad \forall p \in Q_3^+\\ \lambda_0^p \lesseqgtr 0, \qquad \lambda_1^p \geq 0, \qquad \lambda_2^p \geq 0, \qquad \lambda_3^p \geq 0\end{cases}$$

where $\alpha[p], \beta[p], \gamma[p]$ denote the three indices in the subset S_p.

We can now summarize the resulting compact LP formulation (CF2) proposed for minimizing the Choquet integral $C_v(x)$ over a given set $X \subset \mathbb{R}_+^n$ as follows:

$$(\textbf{CF2})\begin{cases}\text{Minimize } \sum_{i=1}^{n} \mu_i x_i + \sum_{p\in Q_2} \tilde{\mu}_p y_p + \sum_{p\in Q_3^-} \mu_p y_p + \sum_{p\in Q_3^+} \mu_p \left[2\lambda_0^p + \lambda_1^p + \lambda_2^p + \lambda_3^p\right]\\ \text{s.t. :}\\ y_p \leq x_i \quad \forall p \in Q_2 \cup Q_3^-, \ \forall i \in S_p\\ \left.\begin{array}{l}\lambda_0^p + \lambda_1^p \geq -x_{\alpha[p]}\\ \lambda_0^p + \lambda_2^p \geq -x_{\beta[p]}\\ \lambda_0^p + \lambda_3^p \geq -x_{\gamma[p]}\end{array}\right\} \quad \forall p \in Q_3^+\\ \lambda_0^p \lesseqgtr 0, \qquad \lambda_1^p \geq 0, \qquad \lambda_2^p \geq 0, \qquad \lambda_3^p \geq 0\\ x \in X \subset \mathbb{R}_+^n.\end{cases}$$

If we assume that X is a polyhedron described by a set of m linear equality/inequality constraints we have the following:

Proposition 3. *The compact formulation* (CF2) *involves* $P + 4|Q_3^+|$ *variables and* $m + 2|Q_2| + 3|Q_3|$ *equality/inequality constraints. It is thus solvable in polynomial time using interior point algorithms.*

Additional integrality conditions on the decision variables are readily handled within (CF2). Computational experiments illustrating this capability are provided in Section 5.

4.4. Third compact formulation (CF3) for a class of submodular v extending the cubic case

We now turn to show that the approach described in the previous section for the cubic submodular case can be extended to a wider class of submodular set functions v including set functions of degree higher than 3. This extension originates from the following basic remark. If we refer to the proof of Proposition 1, we observe that the transformation carried out on each positive cubic term of the form

$$\mu_p \ \underset{i\in S_p}{\text{Min}}\{x_i\} \quad \text{with } \mu_p > 0$$

consists in rewriting it as a sum of positive quadratic terms (eventually "absorbed" by the already existing negative quadratic terms) with a new term popping up, of the form:

$$-\mu_p\left[\underset{i\in S_p}{\mathrm{Min}}\{x_i\}+\underset{i\in S_p}{\mathrm{Min}_2}\{x_i\}\right]\quad\text{with } \mu_p>0.$$

Proposition 2 then shows how this last term can be "linearized", *i.e.*, represented as the optimal value of a small associated linear program with the x variables (the variables with respect to which the Choquet integral has to be minimized) involved linearly in the constraints.

In the generalization proposed in the present section we will address the case of submodular set functions v belonging to the class $\mathcal{C}_K$ defined as follows:

Definition 2. A set function v is said to belong to the class $\mathcal{C}_K$ if and only if the associated Choquet integral $C_v(x)$ can be expressed as a sum of

(i) linear terms of the form $\alpha_i x_i$ without sign restriction on the α_i coefficients,
(ii) nonlinear terms of the form $-\beta_S\,\mathrm{Min}_{i\in S}\{x_i\}$ with $S\subset E$, $|S|\le K$ and $\beta_S>0$,
(iii) nonlinear terms of the form $-q_p^S L_p^S(x)$ where $S\subset E$, $|S|\le K, p\le|S|, q_p^S>0$ and $L_p^S(x)=\sum_{i=1}^p x_i^{\uparrow}(S)$ where $x_i^{\uparrow}(S)\triangleq\underset{j\in S}{\mathrm{Min}_i}\{x_j\}$ is the ith minimum value among the set $\{x_j : j\in S\}$.

Remark 3. Submodularity of the set functions in $\mathcal{C}_K$, though not immediately obvious, will appear as a consequence of the analysis to follow; see Proposition 7. Additional comments on the relevance of the $\mathcal{C}_K$ class will be given later (Remark 4) where an interesting and well-characterized subclass is identified.

Observe that $\mathcal{B}_K^-$ introduced in Definition 1 is the subclass of $\mathcal{C}_K$ satisfying only (i) and (ii). Also observe that for $K=3$, the class $\mathcal{C}_3$ as defined above includes all cubic submodular set functions, since expression (14) is indeed a special case of (i)–(iii) involving subsets S of cardinality no greater than 3, and only terms of the form L_2^S. Also we note that $L_p^S(x)$ can be interpreted as the pth coordinate of the so-called *Lorenz vector* [15] associated with components $x_j, j\in S$ (see [19,24]).

Of course, a basic issue in connection with the practical usefulness of class $\mathcal{C}_K$ is the *recognition problem*. We are going to show that both testing membership in $\mathcal{C}_K$ of a set function v, given as a multilinear polynomial expression of the form (6), and constructing an equivalent expression of its Choquet integral $C_v(x)$ complying with Definition 2 reduces to checking feasibility of a linear inequality system of size polynomial in n, for fixed maximum degree K. To achieve this, we need a few preliminary results. For any $C\subset E=\{1,2,\dots,n\}$ and for any $k\le|C|$ let:

$$\sigma_k^C(x)=\sum_{S\subseteq C,|S|=k}\underset{j\in S}{\mathrm{Min}}\{x_j\}.$$

We note that $\sigma_k^C(x)$ is a sum of terms only involving the Min operator. The following result shows that any term of the form $L_p^C(x)$ arising from Definition 2-(iii) can also be expressed as a sum of terms only involving the Min operator applied to *subsets* S of C.

Proposition 4. *For $p\le|C|-1$, $L_p^C(x)\triangleq\sum_{i=1}^p\underset{j\in C}{\mathrm{Min}_i}\{x_j\}$ can be rewritten as*

$$L_p^C(x)=\sum_{k=1}^p(-1)^{k-1}\binom{|C|-p-2+k}{|C|-p-1}\sigma_{|C|-p+k}^C(x).\tag{15}$$

For $p=|C|$, $L_{|C|}^C(x)=\sigma_1^C(x)$.

Proof.
First we observe that for $i\le|C|$ and $k\le|C|-i+1$, the number of subsets $B\subseteq C$ of size k such that: $\arg\min_{j\in B}\{x_j\}=\underset{j\in C}{\arg\min_i}\{x_j\}$ is $\binom{|C|-i}{k-1}$. So for $k\le|C|, k\ge1$ we can write:

$$\begin{aligned}\sigma_k^C(x)&=\sum_{S\subseteq C,|S|=k}\underset{j\in S}{\mathrm{Min}}\{x_j\}\\&=\binom{|C|-1}{k-1}x_1^{\uparrow}(C)+\binom{|C|-2}{k-1}x_2^{\uparrow}(C)+\cdots+\binom{k-1}{k-1}x_{|C|-k+1}^{\uparrow}(C).\end{aligned}$$

By choosing $k=|C|-p+1$, we can therefore express $x_p^{\uparrow}(C)$ as

$$x_p^{\uparrow}(C)=\sigma_{|C|-p+1}^C(x)-\binom{|C|-1}{|C|-p}x_1^{\uparrow}(C)-\cdots-\binom{|C|-p+1}{|C|-p}x_{p-1}^{\uparrow}(C).$$

Similarly, $x^{\uparrow}_{p-1}(C)$ can be expressed as

$$x^{\uparrow}_{p-1}(C) = \sigma^C_{|C|-p+2}(x) - \binom{|C|-1}{|C|-p+1} x^{\uparrow}_1(C) - \cdots - \binom{|C|-p+2}{|C|-p+1} x^{\uparrow}_{p-2}(C)$$

and so on until:

$$x^{\uparrow}_2(C) = \sigma^C_{|C|-1}(x) - \binom{|C|-1}{|C|-2} x^{\uparrow}_1(C).$$

Eq. (15) then follows by a sequence of substitutions using the above relations. □

For any $C \subseteq \{1, 2, \dots, n\}$ and $p \leq |C|$, let us denote φ^C_p the set function $\mathscr{P}(E) \to \mathbb{R}$ having nonzero Möbius masses only for subsets $S \subseteq C$ of cardinality $|S|$ in the range $[|C| - p + 1; |C|]$, with the Möbius masses of all subsets having the same cardinality $|S| = k$ being equal to $m_k = (-1)^{k-|C|+p-1} \binom{k-2}{|C|-p-1}$.

Corollary 1. *$L^C_p(x)$ is the Choquet integral of x with respect to the set function φ^C_p.*

Proof. By definition of the set function φ^C_p and using Eq. (9), the Choquet integral with respect to φ^C_p is

$$\begin{aligned} C_{\varphi^C_p}(x) &= \sum_{k=|C|-p+1}^{|C|} \sum_{S\subseteq C, |S|=k} m_k \operatorname*{Min}_{j\in S}\{x_j\} \\ &= \sum_{k=|C|-p+1}^{|C|} m_k \sum_{S\subseteq C, |S|=k} \operatorname*{Min}_{j\in S}\{x_j\} \\ &= \sum_{k=|C|-p+1}^{|C|} m_k \sigma^C_k(x). \end{aligned}$$

So

$$C_{\varphi^C_p}(x) = \sum_{k=|C|-p+1}^{|C|} (-1)^{k-|C|+p-1} \binom{k-2}{|C|-p-1} \sigma^C_k(x).$$

By setting $k = |C| - p + k'$, the above can be rewritten as

$$\begin{aligned} C_{\varphi^C_p}(x) &= \sum_{k'=1}^{p} (-1)^{k'-1} \binom{|C|-p-2+k'}{|C|-p-1} \sigma^C_{|C|-p+k'}(x) \\ &= L^C_p(x). \quad \square \end{aligned}$$

Now, let us consider a set function v of given maximum degree $K \geq 4$ given in the form (6) with $\mu_p \leq 0$ for all p such that $|S_p| = 2$ and $\mu_p > 0$ for some p such that $3 \leq |S_p| \leq K$. Thus its Choquet integral given by (10):

$$C_v(x) = \sum_{i=1}^{n} \mu_i x_i + \sum_{p=n+1}^{P} \mu_p \operatorname*{Min}_{i\in S_p}\{x_i\}$$

cannot be readily linearized using the techniques of Section 4.2 or 4.3. The main result of this section is to show that linearization is nevertheless possible for set functions belonging to the class $\mathcal{C}_K$.

Proposition 5. *A set function v belongs to the class $\mathcal{C}_K$ if and only if we can find nonnegative weights q^C_i such that:*

$$v + \sum_{\substack{C\subseteq E \\ 2\leq|C|\leq K}} \sum_{i=2}^{|C|-1} q^C_i \varphi^C_i \in \mathscr{B}^-_K \tag{16}$$

where we recall that $\mathscr{B}^-_K$ is the subclass of set functions complying with Definition 1.

Proof. In view of the straightforward property $C_{v+v'}(x) = C_v(x) + C_{v'}(x)$ for any two set functions v and v' and using Corollary 1, (16) is equivalent to

$$C_v(x) = C_w(x) - \sum_{\substack{C\subseteq E \\ 2\leq|C|\leq K}} \sum_{i=2}^{|C|-1} q^C_i L^C_i(x)$$

for some set function $w \in \mathscr{B}^-_K$. The result follows. □

For any $S \subseteq E$, $|S| \leq K$, let $m(S)$ denote, consistently with (7), the Möbius mass of S for set function v.

Now, for $S \subset E$, $2 \leq |S| \leq K$, the Möbius mass of S for set function $\sum_{\substack{C\subseteq E\\ 2\leq|C|\leq K}} \sum_{i=2}^{|C|-1} q_i^C \varphi_i^C$ is

$$\Phi^S(q) = \sum_{\substack{C\subseteq E, C\supseteq S\\ 3\leq|C|\leq K}} \sum_{i=|C|-|S|+1}^{|C|-1} (-1)^{|S|-|C|+i-1} \binom{|S|-2}{|C|-i-1} q_i^C. \tag{17}$$

Expression (17) is a linear function of the (unknown) coefficients q_i^c which we denote $\Phi^S(q)$. We therefore deduce the following:

Proposition 6. *A given set function v belongs to the class $\mathcal{C}_K$ if and only if the linear system:*

$$m(S) + \Phi^S(q) \leq 0 \tag{18}$$

for all $S \subseteq E$ such that $2 \leq |S| \leq K$, has a nonnegative solution in the q_i^C variables.

For fixed K such that $4 \leq K < n$, the linear system (18) has a polynomial number of variables and constraints. We can therefore deduce the following:

Corollary 2. *Testing whether a given set function v of maximum degree K belongs to the class $\mathcal{C}_K$ and constructing an equivalent expression for its Choquet integral $C_v(x)$ fitting Definition 2 can be done in time polynomial in n for fixed K.*

We note that, for fixed $K \geq 4$, polynomial-time recognition is a remarkable property of the class $\mathcal{C}_K$ in view of the well-known fact that recognizing submodularity for general set functions of degree greater or equal to 4 is NP-complete.

Proposition 7 below generalizes Proposition 2 and shows how any term of the form $-q_p^C L_p^C(x)$ (with $q_p^C > 0$) can be linearized. It is interesting to remark that such terms also appear in the linearization of OWA and WOWA operators (which are particular instances of the Choquet integral) proposed by Ogryczak and Sliwinski [21,22].

Proposition 7. *Any term of the form $-L_p^C(x)$ can be represented as the optimal value of the linear program*

$$\begin{cases} \text{Minimize } p\lambda_0 + \sum_{j\in C} \lambda_j \\ \text{s.t. :} \\ \lambda_0 + \lambda_j \geq -x_j \quad (j \in C) \\ \lambda_0 \gtrless 0, \qquad \lambda_j \geq 0 \quad (j \in C). \end{cases}$$

As a consequence, $-L_p^C(x)$ is convex in x and the set functions in the class $\mathcal{C}_K$ are submodular.

Proof. For any fixed $x \geq 0$, $L_p^C(x)$ is the optimal solution value of the following continuous knapsack problem in bounded variables u_j $(j \in C)$

$$\text{(VI)} \begin{cases} \text{Min} \sum_{j\in C} x_j u_j = -\text{Max}\left(-\sum_{j\in C} x_j u_j\right) \\ \text{s.t. :} \\ \sum_{j\in C} u_j = p \\ 0 \leq u_j \leq 1 \quad (j \in C). \end{cases}$$

Denoting λ_0 and $\lambda_j (j \in C)$ the dual variables, we can see that $-L_p^C(x)$ is equal to the optimal solution value of the dual:

$$\begin{cases} \text{Min } p\lambda_0 + \sum_{j\in C} \lambda_j \\ \text{s.t. :} \\ \lambda_0 + \lambda_j \geq -x_j \quad (j \in C) \\ \lambda_0 \gtrless 0, \qquad \lambda_j \geq 0 \quad (j \in C) \end{cases}$$

which proves the claim. Now, the above shows that $-L_p^C(x)$ can be interpreted as the *perturbation function* of a linear minimization problem, from which convexity follows. Since $-L_p^C(x)$ is the Choquet integral (Lovász extension) of the set function $-\varphi_p^C$, submodularity of the latter is deduced (see [16]). Now (16) shows that any $v \in \mathcal{C}_K$ is submodular, since it is the sum of a submodular function $w \in \mathcal{B}_K^-$ and a weighted sum (with nonnegative weights) of submodular set functions of the form $-\varphi_i^C$. □

Let $\bar{q} = (\bar{q}_i^C)$ be the solution to system (18) obtained after checking membership of v in $\mathcal{C}_K$. For any $2 \le |S| \le K$, denote $\theta(S) = m(S) + \Phi^S(\bar{q})$. Note that $\theta(S) \le 0$, $\forall S \subseteq E$. Minimizing the Choquet integral $C_v(x)$ for $v \in \mathcal{C}_K$ can be reformulated as minimizing over $x \in X$ the function:

$$\sum_{i=1}^{n} \mu_i x_i + \sum_{2 \le |S| \le K} \theta(S) \operatorname*{Min}_{j \in S}\{x_j\} - \sum_{\substack{C \subseteq E \\ 2 \le |C| \le K}} \sum_{i=2}^{|C|-1} \bar{q}_i^C L_i^C(x). \tag{19}$$

The new compact formulation (**CF3**) is then deduced as follows:

- Each term in (19) of the form $\theta(S)\ \mathrm{Min}_{j \in S}\{x_j\}$ is linearized using the same technique as for (CF1);
- Each term in (19) of the form $-\bar{q}_i^C L_i^C(x)$ is linearized using Proposition 7.

We provide below an example of recognition of a submodular degree 4 set function admitting positive and negative Möbius masses.

Example 1. Consider the submodular degree 4 set function v, defined on a 5-element set ($n = 5$) having Möbius masses m with values:

$$\begin{aligned}
&m(\{5\}) = 3\\
&m(\{1\}) = m(\{4\}) = m(\{1,2,4,5\}) = 2\\
&m(\{2\}) = m(\{3\}) = m(\{1,2,3\}) = m(\{1,3,4\}) = m(\{2,3,4\}) = m(\{2,3,5\}) = 1\\
&m(\{1,5\}) = m(\{1,2,4\}) = m(\{1,3,5\}) = m(\{3,4,5\}) = 0\\
&m(\{1,2,3,5\}) = m(\{1,3,4,5\}) = m(\{2,3,4,5\}) = m(\{1,2,3,4,5\}) = 0\\
&m(\{1,2\}) = m(\{1,3\}) = m(\{1,4\}) = m(\{2,4\}) = m(\{2,5\}) = m(\{3,5\}) = m(\{4,5\}) = -1\\
&m(\{1,2,5\}) = m(\{1,4,5\}) = m(\{2,4,5\}) = m(\{1,2,3,4\}) = -1\\
&m(\{2,3\}) = m(\{3,4\}) = -2.
\end{aligned}$$

Now, testing membership of v in $\mathcal{C}_4$ is done using Proposition 6 (the resulting system has 20 variables and 25 constraints). It can be checked that the following nonnegative coefficients solve system (18): $q_2^{\{2,3,5\}} = 1$; $q_3^{\{1,2,3,4\}} = 1$; $q_2^{\{1,2,4,5\}} = 1$. In view of this, the Choquet integral $C_v(x)$ can be rewritten as

$$C_v(x) = 2x_1 + x_2 + x_3 + 2x_4 + 3x_5 - \mathrm{Min}\{x_4, x_5\} - \mathrm{Min}\{x_3, x_4\} - L_2^{\{2,3,5\}}(x) - L_3^{\{1,2,3,4\}}(x) - L_2^{\{1,2,4,5\}}(x).$$

This makes explicit the possibility of linearizing function $C_v(x)$ in a minimization problem since all non-linear terms appearing in the above expression have the right sign to be linearized using both the results of Section 4.2 and Proposition 7 above.

Remark 4. It is worth mentioning here an interesting subclass of problems to which the results of Section 4.4 readily apply: this corresponds to maximizing Choquet integrals with respect to set functions v *which are both* $k+1$*-additive and* k*-monotone* for an arbitrary k such that $3 \le k \le n-1$. Let us briefly explain the reasons for this. According to the definition of k-monotonicity (13) we know that $m(L) \ge 0$ for all L such that $|L| \le k$ (just use (13) with $A = B$ and $|A| \le k$). Therefore the only subsets possibly having negative Möbius masses are those of cardinality $k+1$. We also deduce from (13) that for any subset $A \subseteq E$ of cardinality k we have

$$m(A) + \sum_{i: m(A \cup \{i\}) < 0} m(A \cup \{i\}) \ge 0.$$

Thanks to this property, it is easily seen that $-v$ belongs to the class $\mathcal{C}_{k+1}$. A proof of this would be very similar to the proof of Proposition 1 using the following identity: for all $S \subseteq E$, $|S| = k+1$ and $z_j \in \mathbb{R}$ ($\forall j \in S$)

$$\sum_{i \in S} \operatorname*{Min}_{j \in S \setminus \{i\}}\{z_j\} = k \operatorname*{Min}_{j \in S}\{z_j\} + \operatorname*{Min_2}_{j \in S}\{z_j\}.$$

Using the above allows one to transform the expression of the Choquet integral $C_{-v}(x)$ into an expression involving non-linear terms of the form $\mathrm{Min}_{j \in S}\{x_j\}$ and $L_2^S(x)$ with only negative weights, which can be subsequently linearized.

5. Numerical tests

In this section, we present numerical tests showing the potential of the linearization of the Choquet integral for determining compromise solutions in multicriteria transportation problems and knapsack problems, both in their continuous and discrete versions. In such problems, n linear criteria (cost functions) must be minimized simultaneously and we look for a well-balanced Pareto-optimal solution. The worth of any feasible cost vector $x = (x_1, \ldots, x_n)$ is defined

Table 1
Transportation problems solved with column-generation applied to formulation (II).

n	t	# c	t_{cg}
20	0.30	41	0
40	16.55	111	0.15
60	240.05	187	1.28
80	2032.45	298	7.19

Table 2
Transportation problems solved with (CF2).

n	nc	nv	t	t'
20	782	635	0.01	0.03
40	2 694	2 140	0.10	0.24
60	5 815	4 566	0.44	1.39
80	10 199	7 959	0.71	4.46
100	15 736	12 192	1.74	22.17
120	22 807	17 644	3.60	77.05
140	30 904	23 890	7.35	679.14
160	39 963	30 850	14.13	2108.09

by a convex Choquet integral $C_v(x)$ (v is submodular and monotonic). The optimal compromise solution is obtained by minimizing $C_v(x)$ which ensures both preference for well-balanced solutions as defined in Eq. (4) and Pareto-optimality of the solution found.

The first set of tests is performed on the transportation problem. We consider a problem with n agents (clients) and $m = n$ providers. Each provider is able to produce 2 units of a given product and each agent needs exactly 2 units of this product. Each agent can be supplied from 3 different providers (the underlying graph is built by randomly generating 3 distinct matchings between agents and providers). The reason for generating instances that way is that this structure reduces the possibility to get feasible solutions with ideally balanced cost profiles $(x_1, \ldots, x_n)$, thus making the instances of Choquet minimization harder to solve. Let y_{ij} be the decision variables representing, for every pair (i, j), the quantity delivered to agent j from provider i. There is one criterion for each agent i, that assigns to any solution y the cost $x_i = \sum_{j=1}^{n} c_{ij}y_{ij}$ where c_{ij} are unit transportation costs randomly generated within the [0, 100] interval. Thus, to each feasible solution characterized by continuous variables y_{ij} is associated a cost vector $(x_1, \ldots, x_n)$. The value of each cost vector is then defined by $C_v(x_1, \ldots, x_n)$ for some degree 3 submodular capacity v (Möbius masses are zero for subsets of size greater than 3). For the tests, capacity v is defined from its Möbius masses. These masses are randomly drawn so as to preserve submodularity and monotonicity of v and constructed in such a way that the number of subsets of size 3 with nonzero Möbius masses is nearly equal to the number of subsets of size 2 with nonzero Möbius masses. These instances have been used to test the two linearizations presented in the paper, i.e. the one based on column-generation applied to the large scale formulation (II) introduced in Section 3; and the compact formulation (CF2) introduced in Section 4. The experimental results obtained using column-generation are shown in Table 1 where n is the number of criteria, t is the average solution time over 20 randomly generated instances, $\#c$ the average number of iterations performed to solve the instance (i.e. the total number of columns generated), and t_{cg} the average solution time to find each generated column. We performed these tests using ILOG CPLEX 12.1 on a computer with 8 Gb of memory and an Intel Core 2 Duo 3.33 GHz processor. All computation times shown are in seconds.

The results obtained using the compact formulation (CF2) are shown in Table 2 where t is the average solution time over 20 randomly generated instances, nc, nv are the number of columns and the number of variables respectively in the LP to be solved. In addition, we have solved the same instances after adding integrality constraints on variables y_{ij} ($y_{ij} \in \{0, 1\} \forall i, j$) to test the efficiency of the approach on the integer version of the problem. Thus we also give t', the average solution time for the integer version of the problem.

The comparison of values obtained for t in Tables 1 and 2 confirm the intuition. We observe indeed that for 3-additive capacities, using the compact formulation is much more efficient than column-generation on the large scale formulation. Moreover, (CF2) allows one to solve the integer version of the problem (using a MIP) in reasonable times as can be seen in the right column of Table 2. Note that, contrary to the column-generation approach, it is quite easy to handle integrality restrictions on the variables in the compact formulation.

Using similar instances, we have investigated the impact on solution times of the value of K, i.e. the degree of the set function v considered. Thus, for various values of n ($n = 15, 20, 25$), we have generated submodular degree K set functions for increasing values of K. The set functions in $\mathcal{C}_K$ considered in this experiment are obtained by randomly generating the various coefficients (with the right sign) in Expression (19). Table 3 gives the average computation time t over 20 instances for minimizing the Choquet integral with respect to such set functions (using the compact formulation (CF3)). We also give into brackets the number of nonzero coefficients in the Expression (19) of Choquet integral. We can observe that the compact formulation CF3 is practically efficient at least for $K = 4$ and $K = 5$ up to $n = 25$. We also note that these computation times significantly increase with the value of K, but this seems essentially due to the rapid increase in size of the number

Table 3
Solution times using (CF3) as a function of K.

K	2	3	4	5	6	7
$n = 15$	0.01 (81)	0.02 (536)	0.23 (2584)	2.49 (8590)	12.74 (21 102)	242.55 (40 407)
$n = 20$	0.01 (142)	0.04 (1282)	1.18 (8549)	26.75 (39 557)	393.35 (136 457)	>1000 (369 017)
$n = 25$	0.03 (219)	0.08 (2519)	2.27 (21 494)	271.5 (127 754)	>1000 (570 504)	>1000 (2 012 604)

Table 4
Knapsack problems solved with (CF2).

n	nc	nv	t	t'
20	605	482	0.01	0.03
40	2 256	1 766	0.16	0.33
60	5 209	4 068	0.84	1.38
80	9 338	7 235	1.20	4.98
100	14 642	11 296	2.35	9.08
120	20 078	15 492	4.40	18.04
140	28 597	22 033	7.88	53.47
160	37 350	28 742	13.19	100.98

of nonzero coefficients in Expression (19) as K increases. We do not provide here comparison with the column-generation approach because the best known strongly polynomial algorithm [23] requires $O(n^5)$ function evaluations, which, in view of the sizes of the set functions considered in our experiments (see Table 3) would result in redhibitory computation times in the context of a column-generation procedure.

The second series of tests performed concern knapsack problems in both their continuous and 0–1 versions (such problems can be found e.g. in portfolio optimization or optimal project selection). The decision variables of the knapsack problem are z_j, $j = 1, \ldots, n$. The associated weights are integers randomly generated in [0, 10 000]. We consider n linear criteria with coefficient randomly generated in [0, 1000]. With each solution z we associate its image $(x_1, \ldots, x_n)$ in the space of criteria and we look for z minimizing $C_v(x)$. The set functions v are randomly generated using the same procedure as above. Tests are only provided for the second linearization (CF2) (see Table 4) with the same conventions for notation as in Table 2.

Again, it is observed that the compact formulation (CF2) enables one to compute Choquet-optimal knapsack solutions with many criteria very efficiently, even in the discrete (0–1) version of the problem.

6. Conclusions

The focus of the present paper has been on minimizing Choquet integrals (or Lovász extensions) over polyhedra or (mixed) integer sets in the convex case (i.e. when the underlying set functions are submodular). In theory, when the solution sets considered are polyhedra, polynomiality of the problem follows from a general result by Grötschel et al. (using the so-called "ellipsoid algorithm"). However, solving practical instances requires resorting to the simplex algorithm in conjunction with a column-generation procedure, a computationally demanding approach; as suggested by the computational experiments reported in Section 5, this approach tends to become impracticable beyond, say, about 80 decision variables. Moreover, possible integrality requirements are difficult to handle within such an approach. In view of this, we have investigated an alternative approach based on compact formulations which are obtained here for several classes of submodular set functions, thus encompassing many potential practical applications. The experimental results obtained (involving instances in dimensions up to $n = 160$) confirm the superiority of the compact formulations proposed, both in terms of computational efficiency and easy handling of possible integrality requirements on decision variables. To realize that the value $n = 160$ is indeed significant for Choquet minimization, please keep in mind that a piecewise affine description of a Choquet integral (or Lovász extension) in n-dimensional space involves $n!$ pieces, one for each canonical simplex associated with a permutation of $\{1, \ldots, n\}$. So, to the best of our knowledge, the present paper can be viewed as the first systematic way of solving Choquet minimization problems on solution sets of significantly large dimensions.

Acknowledgments

Professor Endre Boros is gratefully acknowledged for his participation in early discussions on the subject of this paper during his visit in Paris in March 2010. Thanks are also due to a referee for many insightful comments on a preliminary version of the paper.

References

[1] C. Berge, Principles of Combinatorics, Academic Press, New York, 1971.
[2] A. Billionnet, M. Minoux, Maximizing a supermodular pseudoBoolean function: a polynomial algorithm for supermodular cubic functions, Discrete Applied Mathematics 12 (1985) 1–11.
[3] A. Chateauneuf, J. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, Mathematical Social Sciences 17 (1989) 263–283.
[4] A. Chateauneuf, J. Tallon, Diversification, convex preferences and non-empty core in the Choquet expected utility model, Economic Theory 19 (2002) 509–523.
[5] D. Dubois, H. Prade, Evidence measures based on fuzzy information, Automatica 21 (1985) 547–862.
[6] J. Edmonds, Submodular functions, matroids and certain polyhedra, in: R. Guy, H. Hamani, N. Sauer, J. Schönheim (Eds.), Combinatorial Structures and their Applications, vol. 17, Gordon and Breach, 1989, pp. 69–87.
[7] L. Galand, P. Perny, O. Spanjaard, Choquet-based optimisation in multiobjective shortest path and spanning tree problems, European Journal of Operational Research 204 (2) (2010) 303–315.
[8] M. Grabisch, The application of fuzzy integrals in multicriteria decision making, European Journal of Operational Research 89 (3) (1996) 445–456.
[9] M. Grabisch, Fuzzy Measures and Integrals, Springer Verlag, 2000.
[10] M. Grabisch, J. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Cambridge University Press, 2009.
[11] M. Grabisch, T. Murofushi, M. Sugeno, *k*-order additive discrete fuzzy measures and their representation, Fuzzy Sets and Systems 92 (1997) 167–189.
[12] M. Grötschel, L. Lovász, A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, Combinatorica 1 (1981) 169–197.
[13] P.L. Hammer, S. Rudeanu, Boolean Methods in Operations Research, Springer Verlag, 1968.
[14] L.G. Khachian, A polynomial algorithm in linear programming, Soviet Mathematics Doklady 20 (1979) 191–194.
[15] M.O. Lorenz, Methods of measuring the concentration of wealth, Publications of the American Statistical Association 9 (1905) 209–219.
[16] L. Lovász, Submodular functions and convexity, in: A. Bachem, M. Grötschel, B. Korte (Eds.), Mathematical Programming, the State of the Art, 1983, pp. 235–257.
[17] J. Marichal, On Sugeno integral as an aggregation function, Fuzzy Sets and Systems 114 (3) (2000) 347–365.
[18] J. Marichal, Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral, European Journal of Operational Research 155 (3) (2004) 771–791.
[19] W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, London, 1979.
[20] K. Murota, Discrete convex analysis, SIAM MOnographs on Discrete Mathematics and Applications.
[21] W. Ogryczak, T. Sliwinski, On solving linear programs with the ordered weighted averaging objective, European Journal of Operational Research 148 (1) (2003) 80–91.
[22] W. Ogryczak, T. Sliwinski, On efficient WOWA optimization for decision support under risk, International Journal of Approximate Reasoning 50 (2009) 915–928.
[23] J. Orlin, A faster strongly polynomial time algorithm for submodular function minimization, Mathematical Programming 118 (2009) 237–251.
[24] P. Perny, O. Spanjaard, L.X. Storme, A decision–theoretic approach to robust optimization in multivalued graphs, Annals of Operations Research 147 (1) (2006) 317–341.
[25] J. Quiggin, Generalized Expected Utility Theory—The Rank-Dependent Model, Kluwer Academic, Dordrecht, 1993.
[26] J. Rhys, A selection problem of shared fixed costs and network flows, Management Science 17 (1970) 200–207.
[27] J. Rosenmüller, On core and value, Methods of Operations Research 9 (1971) 84–104.
[28] G. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, Probability Theory and Related Fields 2 (1964) 340–368.
[29] D. Schmeidler, Integral representation without additivity, Proceedings of the American Mathematical Society 97 (2) (1986) 255–261.
[30] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.
[31] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[32] L. Shapley, Cores of convex games, International Journal of Game Theory 1 (1971) 11–22.
[33] I. Singer, Extensions of functions of 0–1 variables and applications to combinatorial optimization, Numerical Functional Analysis and Optimization 7 (1) (1984) 23–62.
[34] P. Smets, The degree of belief in a fuzzy event, Information Sciences 25 (1) (1981) 1–19.
[35] V. Torra, The weighted OWA operator, International Journal of Intelligent Systems 12 (1997) 153–166.
[36] R.J. Vanderbei, Linear Programming: Foundations and Extensions, third ed., Springer, 2007.
[37] M. Yaari, The dual theory of choice under risk, Econometrica 55 (1987) 95–115.
[38] R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE Transactions on Systems, Man, and Cybernetics 18 (1) (1988) 183–190.
[39] Y. Ye, Interior Point Algorithms: Theory and Analysis, John Wiley & Sons, Inc., 1997.