4OR manuscript No.
(will be inserted by the editor)

Quadratic distances for capacity and bi-capacity approximation and identification

Ivan Kojadinovic

LINA CNRS FRE 2729, Site école polytechnique de l'université de Nantes, Rue Christian Pauc, 44306 Nantes, France.
e-mail: `ivan.kojadinovic@polytech.univ-nantes.fr`

Received: date / Revised version: date

Abstract The application of multi-attribute utility theory based on the Choquet integral requires the prior identification of a capacity if the utility scale is unipolar, or of a bi-capacity if the utility scale is bipolar. In order to implement a minimum distance principle for capacity or bi-capacity approximation or identification, quadratic distances between capacities and bi-capacities are studied. The proposed approach, consisting in solving a strictly convex quadratic program, has been implemented within the GNU R `kappalab` package for capacity and non-additive integral manipulation. Its application is illustrated on two examples.

Keywords: Multi-attribute utility theory, Choquet integral, capacity, bi-capacity, quadratic programming (MSC: 90B50, 91B16, 90C20).

---

1 Introduction

With outranking methods (Roy and Bouyssou, 1993), multi-attribute utility theory (MAUT) (Keeney and Raiffa, 1976) is probably the most frequently applied approach to multi-criteria decision aiding (MCDA) (Vincke, 1992). Given a set $\mathcal{A} := \{a, b, c, \dots\}$ of *alternatives* and a set $N := \{1, \dots, n\}$ of *criteria*, MAUT roughly consists in synthesizing, for each alternative, the n different points of view quantified by the criteria in order, typically, to help the decision maker choose a subset of alternatives that can be considered the best for him. More precisely, in such a context, each alternative $a \in \mathcal{A}$ is identified with its vector of *scores* $(a_1, \dots, a_n) \in \mathbb{R}^n$ where, for any $i \in N$, a_i represents the *utility* of a for the decision maker with respect to (w.r.t.) criterion i. The preferences of the decision maker over the alternatives, represented by a binary relation $\succeq_{\mathcal{A}}$ supposed to be transitive and complete in the considered context, are then to be modeled by means of a *global utility function* $U : \mathbb{R}^n \to \mathbb{R}$ such that

$$a \succeq_{\mathcal{A}} b \iff U(a_1, \dots, a_n) \geq U(b_1, \dots, b_n), \qquad \forall\, a, b \in \mathcal{A}. \tag{1}$$

The form of the global utility function U depends on the hypotheses on which the MCDA problem is based. When *mutual preferential independence* (see e.g. Vincke, 1992) among criteria can be assumed, it is frequent to consider that the global utility function is additive and takes the form of a weighted arithmetic mean. This assumption is however rarely verified in practice. In order to be able to take interaction among criteria into account, it has been proposed to substitute a monotone set function on $N := \{1, \dots, n\}$, called *capacity* (Choquet, 1953) or *fuzzy measure* (Sugeno, 1974), to the weight vector involved in the calculation of weighted arithmetic means. Such an approach can be regarded as taking into account not only the importance of each criterion but also the importance of each subset of criteria. A natural extension of the weighted arithmetic mean in such a context is the *Choquet integral* w.r.t. the defined capacity (Grabisch, 1992; Marichal, 2000; Labreuche and Grabisch, 2003).

The use of the Choquet integral in Eq. (1) requires, as we shall see, the ability to compare utility levels on different criteria. In other terms, it is necessary that the utilities be *commensurable*, i.e. $a_i = a_j$ if and only if, for the decision maker, the alternative a is satisfied to the same extent on criteria i and j (see e.g. Grabisch et al., 2003, for a more complete discussion on commensurability). Obtaining commensurable utilities is a difficult problem that will not be dealt with in this paper. Note however that, in the considered context, it can be performed using the extension of the MACBETH methodology (Bana e Costa et al., 2005) proposed by Labreuche and Grabisch (2003).

Grabisch and Labreuche (2005a,b) have recently shown that even such a general aggregation function as the Choquet integral w.r.t. a capacity is not suited for situations where the utilities to be aggregated lie on a *bipolar* scale. Compared to a classical (*unipolar*) scale, a bipolar scale is characterized by the additional presence of a neutral value such that values above this neutral reference point are considered to be “good” or “positive” by the decision maker, whereas values below it are considered to be “bad” or “negative” (see Grabisch and Labreuche, 2005d, for a complete discussion on bipolarity). In order to derive aggregation models taking the specificity of bipolar scales into account, Grabisch and Labreuche (2005a,b,d) have recently introduced the notion of *bi-capacity*, extending that of capacity, and have proposed a natural generalization of the Choquet integral in that context.

The use of a Choquet integral as a global utility function clearly requires the prior identification of the underlying capacity if the utility scale is unipolar, or of the underlying bi-capacity if the utility scale is bipolar. The learning data from which the capacity or the bi-capacity is to be determined consists of what Marchant (2003) calls the *initial preferences* of the decision maker: usually, a partial weak order over the set of alternatives, a partial weak order over the set of criteria, intuitive judgments about the importance of the criteria, etc.

In this paper, generalizing the minimum variance approach to capacity identification recently put forward in (Kojadinovic, 2006) and following Marichal (1998, Chap. 7), we propose to use a minimum distance principle for capacity (resp. bi-capacity) identification based on natural distances between capacities (resp. bi-capacities). For practical purposes, we focus on quadratic distances between capacities and bi-capacities which enables us to implement the minimum distance principle under the form of a strictly convex quadratic program. Furthermore, as we shall see, the capacity (resp. bi-capacity) identification problem is closely related to the capacity approximation problem (Marichal, 1998, Chap. 7) (resp. bi-capacity approximation problem), which we will be able to address as well using the proposed minimum distance principle. The derived methodology has been implemented within the `kappalab` package (Grabisch et al., 2005) for the GNU R statistical system (R Development Core Team, 2005), an application of which will be presented in § 4.

In the second section, after recalling the notions of game, capacity and Choquet integral in the context of aggregation, we review uniformity measures that can be used for capacity identification and we propose new quadratic distances that can be used for similar purposes. In the third section, the concepts presented in the second section are generalized: we recall the notions of bi-cooperative game, bi-capacity and Choquet integral w.r.t. a bi-capacity and we extend the quadratic distances between capacities previously defined to bi-capacities. The last section is devoted to the presentation of two applications of the proposed minimum distance approach, one to capacity approximation, the other to capacity identification.

In order to avoid cumbersome notation, we will often omit braces for singletons and pairs, e.g., by writing $\mu(i)$, $N \setminus ij$ instead of $\mu(\{i\})$, $N \setminus \{i,j\}$. Furthermore, cardinalities of subsets $S, T, \dots$, will be denoted by the corresponding lower case letters $s, t, \dots$

2 Quadratic distances for capacity identification and approximation

In the context of aggregation by the Choquet integral, *capacities* (Choquet, 1953), also called *fuzzy measures* (Sugeno, 1974), can be regarded as generalizations of weighting vectors involved in the calculation of weighted sums. Throughout this section, the utility scale is considered to be unipolar.

2.1 Cooperative games and capacities

Let $\mathcal{P}(N)$ denote the power set of N.

Definition 1. *A function $\mu : \mathcal{P}(N) \to \mathbb{R}$ is a cooperative game on N if it satisfies $\mu(\emptyset) = 0$.*

In the sequel, the set of cooperative games on N shall be denoted by $\mathcal{G}_N$.

Definition 2. *A function $\mu : \mathcal{P}(N) \to \mathbb{R}$ is a capacity on N if it is a monotone game, i.e. if $\mu(\emptyset) = 0$, and, if, for any $S, T \subseteq N$, $S \subseteq T$ implies $\mu(S) \leq \mu(T)$.*

A capacity is additionally said to be normalized if $\mu(N) = 1$. Note that, in the context of MAUT based on the Choquet integral, normalized capacities are exclusively considered. As we continue, the set of capacities on N shall be denoted by $\mathcal{C}_N$, and the set of normalized capacity by $\mathcal{C}_N^*$.

A game $\mu \in \mathcal{G}_N$ is further said to be

- *additive* if $\mu(S \cup T) = \mu(S) + \mu(T)$ for all disjoint subsets $S, T \subseteq N$,
- *supermodular* (resp. *submodular*) if $\mu(S \cup T) + \mu(S \cap T) \geq \mu(S) + \mu(T)$ (resp. $\leq$) for all disjoint subsets $S, T \subseteq N$,
- *cardinality-based* if, for any $T \subseteq N$, $\mu(T)$ depends only on the cardinality of T.

Note that there is only one normalized capacity on N that is both additive and cardinality-based. We shall call it *the uniform capacity* and denote it by μ^*. It is easy to verify that μ^* is given by

$$\mu^*(T) = t/n, \qquad \forall T \subseteq N.$$

The *dual* (or *conjugate*) of a game μ on N is a game $\bar{\mu}$ on N defined by $\bar{\mu}(A) := \mu(N) - \mu(N \setminus A)$, for all $A \subseteq N$.

We now recall an equivalent representation of a game. Any game $\mu \in \mathcal{G}_N$ (and, more generally, any set function on N) can be uniquely expressed in terms of its *Möbius representation* (Rota, 1964) by

$$\mu(T) = \sum_{S \subseteq T} m_\mu(S), \qquad \forall\, T \subseteq N, \tag{2}$$

where the set function $m_\mu : \mathcal{P}(N) \to \mathbb{R}$ is called the *Möbius transform* or *Möbius representation* of μ and is given by

$$m_\mu(S) = \sum_{T \subseteq S} (-1)^{s-t} \mu(T), \qquad \forall\, S \subseteq N.$$

From the previous results, it follows that a game μ on N is completely defined by the knowledge of $2^n - 1$ coefficients, for instance $(\mu(S))_{\emptyset \neq S \subseteq N}$ or $(m_\mu(S))_{\emptyset \neq S \subseteq N}$. Such a complexity may be prohibitive in certain applications. The fundamental notion of *k-additivity* proposed by Grabisch (1997) enables to find a trade-off between the complexity of the game and its modeling ability.

Definition 3. *Let $k \in \{1, \dots, n\}$. A game $\mu \in \mathcal{G}_N$ is said to be* k-additive *if its Möbius representation satisfies $m_\mu(T) = 0$ for all $T \subseteq N$ such that $t > k$ and there exists at least one subset T of cardinality k such that $m_\mu(T) \neq 0$.*

As one can easily check, the notion of 1-additivity coincides with that of additivity. Let $k \in \{1, \dots, n\}$ and let μ be a k-additive game on N. From Eq. (2), we immediately have that

$$\mu(S) = \sum_{\substack{\emptyset \neq T \subseteq S \\ t \leq k}} m_\mu(T), \qquad \forall\, S \subseteq N,$$

a k-additive game being thus completely defined by the knowledge of $\sum_{l=1}^{k} \binom{n}{l}$ coefficients.

2.2 The lattice $(\mathcal{P}(N), \subseteq)$

Denote by $\mathcal{M}_{(\mathcal{P}(N),\subseteq)}$ the set of maximal chains of the lattice $(\mathcal{P}(N), \subseteq)$ and by Π_N the set of permutations on N. The sets $\mathcal{M}_{(\mathcal{P}(N),\subseteq)}$ and Π_N are clearly in one-to-one correspondence. Indeed, with each permutation $\sigma \in \Pi_N$ can be associated a unique maximal chain $m_\sigma \in \mathcal{M}_{(\mathcal{P}(N),\subseteq)}$ defined by

$$m_\sigma := (\emptyset \subsetneq \{\sigma(n)\} \subsetneq \{\sigma(n-1), \sigma(n)\} \subsetneq \cdots \subsetneq \{\sigma(1), \dots, \sigma(n)\} = N).$$

Given a game $\mu \in \mathcal{G}_N$, with each permutation $\sigma \in \Pi_N$ (i.e. with each maximal chain m_σ), we associate a distribution of real numbers ω_σ^μ on N defined by

$$\omega_\sigma^\mu(i) := \mu\big(\{\sigma(i),\dots,\sigma(n)\}\big) - \mu\big(\{\sigma(i+1),\dots,\sigma(n)\}\big), \qquad \forall\, i \in N.$$

It follows that a game μ on N can be regarded as a set of $n!$ such distributions. Note that if μ is a capacity, then, from the monotonicity conditions, the numbers $\omega_\sigma^\mu(i)$ are necessarily non negative for all $\sigma \in \Pi_N$ and for all $i \in N$. If μ is further normalized, then ω_σ^μ is a probability distribution on N for all $\sigma \in \Pi_N$.

Equivalently, with the maximal chain $m_\sigma \in \mathcal{M}_{(\mathcal{P}(N),\subseteq)}$ is associated the distribution $\omega_{m_\sigma}^\mu := \omega_\sigma^\mu$.

2.3 The Choquet integral w.r.t. a game

In the context of aggregation, the Choquet integral can be regarded as a natural generalization of the weighted sum (Grabisch, 1992; Labreuche and Grabisch, 2003; Marichal, 2000).

Definition 4. *The Choquet integral of a function $x : N \to \mathbb{R}^+$, represented by the vector $(x_1,\dots,x_n)$, w.r.t. the game μ on N is defined by*

$$C_\mu(x) := \sum_{i=1}^{n} \omega_\sigma^\mu(i) x_{\sigma(i)},$$

where σ is a permutation on N such that $x_{\sigma(1)} \le \cdots \le x_{\sigma(n)}$.

It is easy to see that the Choquet integral is a piecewise linear function that coincides with a weighted sum on each set

$$\mathcal{O}_\sigma := \{x \in (\mathbb{R}^+)^n \mid x_{\sigma(1)} \le \cdots \le x_{\sigma(n)}\} \qquad (\sigma \in \Pi_N). \tag{3}$$

Indeed, the sets $\mathcal{O}_\sigma$ cover $(\mathbb{R}^+)^n$, and hence, for any $x \in (\mathbb{R}^+)^n$, there exists $\sigma \in \Pi_N$ such that $x \in \mathcal{O}_\sigma$ and $C_\mu(x) = \sum_{i\in N} \omega_\sigma^\mu(i) x_{\sigma(i)}$.

Let $A \subseteq N$. We denote by $(1_A, 0_{N\setminus A})$ an alternative, called *binary*, whose vector of scores is such that all criteria in A have a score equal to 1, and all the others have a score equal to 0. A fundamental property of the Choquet integral is that $C_\mu(1_A, 0_{N\setminus A}) = \mu(A)$ for all $A \subseteq N$. When restricted to vectors of $[0,1]^n$, a very nice geometrical interpretation of the Choquet integral has been given by Grabisch (2004): it is the simplest linear interpolation between vertices of the hypercube $[0,1]^n$.

In terms of the Möbius representation of a game μ, Chateauneuf and Jaffray (1989) showed that, for any $x = (x_1, \dots, x_n) \in (\mathbb{R}^+)^n$, the Choquet integral of x w.r.t μ is given by

$$C_{m_\mu}(x) = \sum_{T \subseteq N} m_\mu(T) \bigwedge_{i \in T} x_i, \tag{4}$$

where the symbol $\wedge$ denotes the minimum operator.

2.4 Uniformity measures

A first class of measures that can be used for capacity identification are *uniformity* measures. They were initially obtained by generalizing probabilistic entropy measures to normalized capacities.

The first extension of a probabilistic entropy measure to normalized capacities is due to Marichal (1998, 2002). For any normalized capacity μ on N, the generalized Shannon entropy is defined by

$$H_M(\mu) := -\sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) [\mu(S \cup i) - \mu(S)] \ln[\mu(S \cup i) - \mu(S)],$$

with the convention that $0 \ln 0 := 0$ and where

$$\gamma_s(n) := \frac{(n-s-1)!s!}{n!} \qquad (s = 0, 1, \dots, n-1).$$

A fundamental property of H_M is that it can be rewritten in terms of the maximal chains of $(\mathcal{P}(N), \subseteq)$. Using the notations introduced in § 2.2, we have

$$H_M(\mu) = \frac{1}{n!} \sum_{m \in \mathcal{M}_{(\mathcal{P}(N), \subseteq)}} H_S(\omega_m^\mu) = \frac{1}{n!} \sum_{\sigma \in \Pi_N} H_S(\omega_\sigma^\mu), \qquad \forall \mu \in \mathcal{C}_N^*,$$

where, for any probability distribution p on N, $H_S(p)$ denotes the Shannon (1948) entropy of p defined by

$$H_S(p) := -\sum_{i \in N} p(i) \ln p(i).$$

The quantity $H_M(\mu)$ can therefore simply be seen as an average of the uniformity values of the probability distributions ω_m^μ ($m \in \mathcal{M}_N$) calculated by means of the Shannon entropy. To stress the fact that H_M is an average of Shannon entropies, we equivalently denote it by $\overline{H}_S$.

It has been shown that $H_M = \overline{H}_S$ satisfies many properties that one would intuitively require from an entropy measure (Marichal, 2002; Kojadinovic et al., 2005). Two of the most important ones in the framework of capacity identification are the *maximality* property, stating that $\overline{H}_S(\mu)$ is always non negative and reaches its maximum if and only if $\mu = \mu^*$, and the strict concavity of $\overline{H}_S$ as a function of a normalized capacity, which implies that maximizing it over a convex subset of the set of normalized capacities on N always leads to a unique global maximum.

In (Kojadinovic, 2006), for practical purposes, the Havrda and Charvat (1967) entropy was similarly extended to capacities. For any normalized capacity μ on N, the generalized Havrda and Charvat entropy of order $\beta \neq 1$ is defined by

$$\overline{H}^{\beta}_{HC}(\mu) := \frac{1}{1-\beta}\left[\sum_{i\in N}\sum_{S\subseteq N\setminus i}\gamma_s(n)[\mu(S\cup i)-\mu(S)]^{\beta}-1\right],\ \beta>0,\ \beta\neq 1.$$

The quantity $\overline{H}^{\beta}_{HC}$ satisfies the same properties as $\overline{H}_S$, and in particular it is a strictly concave function of a normalized capacity.

It is well-known in information theory that the Havrda and Charvat entropy of order 2 is closely linked to the notion of variance. In (Kojadinovic, 2006), the *variance* of a normalized capacity μ on N was defined as

$$\overline{V}(\mu) := \frac{1}{n}\sum_{i\in N}\sum_{S\subseteq N\setminus i}\gamma_s(n)\left(\mu(S\cup i)-\mu(S)-\frac{1}{n}\right)^2. \tag{5}$$

In the considered context, for any $\mu \in \mathcal{C}^*_N$, the Havrda and Charvat entropy of order 2 and the variance are linked by the following linear equation:

$$\overline{H}^{2}_{HC}(\mu) = \frac{n-1}{n} - n\overline{V}(\mu).$$

The quantity $\overline{V}$ is clearly a strictly convex function of a normalized capacity. It has been used in (Kojadinovic, 2006) as the objective function of a quadratic program whose aim is to obtain the minimum variance capacity, or equivalently, the maximum $\overline{H}^{2}_{HC}$ entropy capacity compatible with the initial preferences of a decision maker.

2.5 Quadratic distances between games

As a second class of measures that could be used for capacity identification or approximation, we propose to focus on *distances* between capacities. In order to define quadratic distances between games, we consider the following three functions from $\mathcal{G}_N \times \mathcal{G}_N$ to $\mathbb{R}$ defined, for any $\mu, \nu \in \mathcal{G}_N$, by

$$\langle \mu, \nu \rangle_1 := \frac{1}{2^n - 1} \sum_{\substack{T \subseteq N \\ T \neq \emptyset}} \mu(T)\nu(T),$$

$$\langle \mu, \nu \rangle_2 := \frac{1}{n!} \sum_{\sigma \in \Pi_N} \frac{1}{n} \sum_{i \in N} \omega_\sigma^\mu(i)\omega_\sigma^\nu(i),$$

$$\langle \mu, \nu \rangle_3 := \int_{[0,1]^n} C_\mu(x) C_\nu(x) dx.$$

The two following propositions give more operational expressions of these functions.

Proposition 1. *For any $\mu, \nu \in \mathcal{G}_N$, we have*

$$\langle \mu, \nu \rangle_2 = \frac{1}{n} \sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) \left(\mu(S \cup i) - \mu(S)\right) \left(\nu(S \cup i) - \nu(S)\right).$$

Proof. The result is a direct consequence of (Kojadinovic et al., 2005, Prop. 1). □

Proposition 2. *For any $\mu, \nu \in \mathcal{G}_N$, we have*

$$\langle \mu, \nu \rangle_3 = \sum_{T \subseteq N} \sum_{S \subseteq N} \frac{1}{|T \cup S| + 2} \left(\frac{1}{t+1} + \frac{1}{s+1} \right) m_\mu(T) m_\nu(S).$$

The previous proposition is a direct consequence of Eq. (4) and of the following lemma proved by Marichal (1998, Lemma 7.2.1).

Lemma 1. *For all $T, S \subseteq N$, we have*

$$\int_{[0,1]^n} \left(\bigwedge_{i \in T} x_i \right) \left(\bigwedge_{j \in S} x_j \right) dx = \frac{1}{|T \cup S| + 2} \left(\frac{1}{t+1} + \frac{1}{s+1} \right).$$

Proposition 3. *The functions $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$, and $\langle \cdot, \cdot \rangle_3$ are inner products in $\mathcal{G}_N$.*

Proof. It is straightforward that verify that $\langle \cdot, \cdot \rangle_1$ is bi-linear, symmetric, and positive-definite. From Proposition 1, the same clearly holds for $\langle \cdot, \cdot \rangle_2$. The function $\langle \cdot, \cdot \rangle_3$ is clearly symmetric. Its bi-linearity follows from the linearity of the Choquet integral w.r.t. a game. Finally, $\langle \mu, \mu \rangle_3$ is clearly non negative and $\langle \mu, \mu \rangle_3 = 0$ clearly implies that $C_\mu(x) = 0$ for all $x \in [0,1]^n$, which immediately implies that all the distributions ω_σ^μ, $\sigma \in \Pi_N$, are zero, and therefore that $\mu(S) = 0$ for all $S \subseteq N$. □

The quadratic distances between games obtained from the above inner products are respectively given, for any $\mu, \nu \in \mathcal{G}_N$, by

$$d_1^2(\mu,\nu) := \frac{1}{2^n - 1} \sum_{\substack{T \subseteq N \\ T \neq \emptyset}} [\mu(T) - \nu(T)]^2,$$

$$d_2^2(\mu,\nu) := \frac{1}{n!} \sum_{\sigma \in \Pi_N} \frac{1}{n} \sum_{i \in N} [\omega_\sigma^\mu(i) - \omega_\sigma^\nu(i)]^2,$$

$$d_3^2(\mu,\nu) := \int_{[0,1]^n} [C_\mu(x) - C_\nu(x)]^2 dx.$$

In the context of Choquet integral based MAUT, $d_1(\mu,\nu)$ can be interpreted as the average quadratic difference between the global scores of the binary alternatives $(1_T, 0_{N \setminus T})$, $T \subseteq N$, $T \neq \emptyset$, measured by μ and those measured by ν. With respect to the Choquet integrals C_μ and C_ν, it can be regarded as measuring the average quadratic difference between their values at the vertices of $[0,1]^n$, from which, as recalled in § 2.3, all their values at vectors of $[0,1]^n$ are computed by linear interpolation (Grabisch, 2004).

The second distance measures the average quadratic difference between the coefficients of the Choquet integrals C_μ and C_ν. Indeed, recall that, for any $x \in (\mathbb{R}^+)^n$, there exists $\sigma \in \Pi_N$ such that $x \in \mathcal{O}_\sigma$ (see Eq. (3)) and hence

$$C_\mu(x) = \sum_{i \in N} \omega_\sigma^\mu(i) x_{\sigma(i)}, \qquad \text{and} \qquad C_\nu(x) = \sum_{i \in N} \omega_\sigma^\nu(i) x_{\sigma(i)}.$$

The idea behind distance d_2 is simply to compute the average quadratic difference $[\omega_\sigma^\mu(i) - \omega_\sigma^\nu(i)]^2$, $i \in N$, $\sigma \in \Pi_N$, between Choquet integral weights in order to appraise the difference between μ and ν, or equivalently between C_μ and C_ν.

The last distance, thoroughly studied by Marichal (1998, Chap. 7) in the context of the extension of pseudo-Boolean functions, can be interpreted as the expected quadratic difference between global scores computed by C_μ and C_ν assuming that the alternatives are uniformly distributed over $[0,1]^n$.

The next results, which directly follow from Propositions 1 and 2, give more operational expressions for distances d_2 and d_3.

Corollary 1. *For any two games μ and ν on N, we have*

$$d_2^2(\mu,\nu) = \frac{1}{n} \sum_{i \in N} \sum_{T \subseteq N \setminus i} \gamma_t(n) \left[\mu(T \cup i) - \mu(T) - \nu(T \cup i) + \nu(T)\right]^2. \qquad (6)$$

Corollary 2. *For any two games μ and ν on N, we have*

$$d_3^2(\mu,\nu) = \sum_{T\subseteq N}\sum_{S\subseteq N} \frac{1}{|T\cup S|+2}\left(\frac{1}{t+1}+\frac{1}{s+1}\right)[m_\mu(T)m_\mu(S) \\ -2m_\mu(T)m_\nu(S)+m_\nu(T)m_\nu(S)]\,. \quad (7)$$

2.6 Capacity approximation and identification based on quadratic distances

We restrict ourselves to the situations where both the capacity approximation (Marichal, 1998, Chap. 7) and the capacity identification problems can be stated as the following strictly convex quadratic program:

$$\begin{array}{l} \min_{\nu\in\mathcal{G}_N} d^2(\mu,\nu) \\ \text{subject to } \begin{cases} \text{monotonicity constraints on } \nu, \\ \text{possible additional constraints,} \end{cases} \end{array} \quad (8)$$

where d is a quadratic distance on $\mathcal{G}_N$ and μ is a capacity on N. Note that, since the aim is to obtain a capacity, at least monotonicity constraints on ν should be imposed.

The difference between the approximation problem and the identification problem would come from the type of the additional constraints and the form of μ. The identification problem would be typically characterized by an additive μ. In the absence of clear requirements on the aggregation function, a very natural choice for μ would be μ^* since the Choquet integral w.r.t. μ^* is nothing else than the simple arithmetic mean, which can be regarded as the simplest possible aggregation operator. Interestingly enough, in that case, the identification of the closest capacity to μ^* w.r.t. d_2 coincides with the minimum variance approach proposed in (Kojadinovic, 2006).

Proposition 4. *For any normalized capacity ν on N, we have*

$$d_2^2(\mu^*,\nu) = \overline{V}(\nu).$$

Proof. The result immediately follows from Eqs. (5) and (6). □

Additional constraints would typically include the normalization constraint $\nu(N) = 1$, as well as constraints resulting from preferential information, for instance a partial weak order over the alternatives in terms of C_ν, a partial weak order over the criteria in terms of their Shapley (1953) value, the behavior of some criteria as *veto* or *favor* (Marichal, 2004), etc.

The precise form of these constraints, all linear w.r.t. ν, is discussed in more detail in (Kojadinovic, 2006).

For the approximation problem, μ could be any capacity on N, typically one obtained by an identification procedure, and the additional constraints would usually not represent preferential information but rather simplicity requirements such as k-additivity, supermodularity or submodularity.

Of course, the k-additivity requirement could be imposed in the identification problem as well. In that case, it is natural to rewrite the problem given in (8) in terms of the Möbius transform of ν, which reduces the number of variables to $\sum_{l=1}^{k} \binom{n}{l}$.

2.7 Basic linear constraints on games

The most basic linear constraints on a game are monotonicity constraints.

Proposition 5. *A set function $\nu : \mathcal{P}(N) \to \mathbb{R}$ is monotone if and only if, for any $i \in N$, and any $S \subseteq N \setminus i$,*

$$\nu(S \cup i) \geq \nu(S).$$

Proof. Immediate. □

The following result is due to Chateauneuf and Jaffray (1989).

Proposition 6. *A set function $m : \mathcal{P}(N) \to \mathbb{R}$ corresponds to the Möbius representation of a capacity if and only if*

(i) $m(\emptyset) = 0$,
(ii) $\sum_{T \subseteq S} m(T \cup i) \geq 0, \qquad \forall\, i \in N,\ \forall\, S \subseteq N \setminus i.$

If, additionally, $\sum_{T \subseteq N} m(T) = 1$, then m is the Möbius representation of a normalized capacity.

The next proposition (Kojadinovic, 2002) gives the form of supermodularity and submodularity constraints. Such constraints may be imposed if the resulting game is to model exclusively complementarity (resp. substitutivity) among criteria.

Proposition 7. *A set function $\nu : \mathcal{P}(N) \to \mathbb{R}$ is supermodular (resp. submodular) if and only if, for any $ij \subseteq N$, and any $S \subseteq N \setminus ij$,*

$$\nu(ij \cup S) + \nu(S) \geq \nu(i \cup S) + \nu(j \cup S) \ \ \textit{(resp. } \leq \textit{).}$$

This last proposition (Chateauneuf and Jaffray, 1989) gives the form of supermodularity and submodularity constraints in terms of the Möbius transform.

Proposition 8. *A set function* $m : \mathcal{P}(N) \to \mathbb{R}$ *corresponds to the Möbius representation of a supermodular (resp. submodular) game if and only if*

(i) $m(\emptyset) = 0$,
(ii) $\sum_{K \subseteq S} m(ij \cup K) \geq 0$ *(resp.* ≤ 0*), for all* $ij \subseteq N$*, and all* $S \subseteq N \setminus ij$.

The form of higher monotonicity constraints can be obtained similarly (see Chateauneuf and Jaffray, 1989; Fujimoto and Murofushi, 2005).

2.8 Matrix formulation of the quadratic distances

Let us define a linear order on the set $\mathcal{P}(N) \setminus \{\emptyset\}$ so that it can be identified with $\{1, 2, \dots, 2^n - 1\}$. A natural choice is the following total ordering of the elements of $\mathcal{P}(N) \setminus \{\emptyset\}$:

$$\{1\}, \{2\}, \{3\}, \dots, \{1,2\}, \{1,3\}, \{1,4\}, \dots, \{1,2,3\}, \{1,2,4\}, \dots, N.$$

In other terms, first are given the n singletons, then the $\binom{n}{2}$ pairs, then the $\binom{n}{3}$ 3-element subsets, ..., and finally N itself.

Let μ and ν be two games on N and let us define the following vectors of $2^n - 1$ components respectively representing the Möbius transform of ν and of μ:

$$x := (m_\nu(S))_{\emptyset \neq S \subseteq N}, \qquad \text{and} \qquad y := (m_\mu(S))_{\emptyset \neq S \subseteq N}.$$

Now, let d be a quadratic distance on $\mathcal{G}_N$. The Möbius representation being a linear invertible transform of a game, there exists a symmetric positive-definite matrix H of order $2^n - 1$ such that

$$d^2(\mu, \nu) = (y - x)^t H (y - x) = y^t H y - 2 y^t H x + x^t H x,$$

where ^{t} denotes the matrix or vector transpose.

Assume now that ν and μ are k and k'-additive respectively so that the vectors x and y can be truncated to $\sum_{l=1}^{k} \binom{n}{l}$ and $\sum_{l=1}^{k'} \binom{n}{l}$ components respectively, i.e.

$$x = (m_\nu(S))_{\emptyset \neq S \subseteq N, s \leq k}, \qquad \text{and} \qquad y = (m_\mu(S))_{\emptyset \neq S \subseteq N, s \leq k'}. \tag{9}$$

It is then easy to verify that, in matrix form,

$$d_1^2(\mu,\nu) = \frac{1}{2^n-1}\left(x^t B_k^t B_k x - 2y^t B_{k'}^t B_k x + y^t B_{k'}^t B_{k'} y\right), \qquad (10)$$

where B_k (resp. $B_{k'}$) is an $(2^n-1) \times \sum_{l=1}^{k}\binom{n}{l}$ $\left(\text{resp. } (2^n-1) \times \sum_{l=1}^{k'}\binom{n}{l}\right)$ matrix such that

$$B_k x = (\nu(S))_{\emptyset \neq S \subseteq N} \qquad \left(\text{resp. } B_{k'} y = (\mu(S))_{\emptyset \neq S \subseteq N}\right).$$

Similarly, from Eq. (6), it can be verified (see Kojadinovic, 2006) that

$$d_2^2(\mu,\nu) = x^t M_k^t D_{Sh} M_k x - 2y^t M_{k'}^t D_{Sh} M_k x + y^t M_{k'}^t D_{Sh} M_{k'} y, \qquad (11)$$

where D_{Sh} is a diagonal matrix of order $n2^{n-1}$ whose diagonal elements are $(\gamma_s(n)/n)_{i\in N, S\subseteq N\setminus i}$ and M_k (resp. $M_{k'}$) is a matrix of order $n2^{n-1} \times \sum_{l=1}^{k}\binom{n}{l}$ $\left(\text{resp. } n2^{n-1} \times \sum_{l=1}^{k'}\binom{n}{l}\right)$ such that

$$M_k x = \left(\sum_{\substack{T\subseteq S\\ t\leq k-1}} m_\nu(T\cup i)\right)_{\substack{i\in N,\\ S\subseteq N\setminus i}} \left(\text{resp. } M_{k'} y = \left(\sum_{\substack{T\subseteq S\\ t\leq k'-1}} m_\mu(T\cup i)\right)_{\substack{i\in N,\\ S\subseteq N\setminus i}}\right).$$

Finally, from Eq. (7) and following Marichal (1998, Chap. 7), it is easy to check that

$$d_3^2(\mu,\nu) = x^t D_{k,k} x - 2y^t D_{k',k} x + y^t D_{k',k'} y, \qquad (12)$$

where, for any $r,m \in \{1,\dots,n\}$, $D_{r,m}$ is a $\sum_{l=1}^{r}\binom{n}{l} \times \sum_{l=1}^{m}\binom{n}{l}$ matrix such that

$$D_{r,m} = \left(\frac{1}{|T\cup S|+2}\left(\frac{1}{t+1}+\frac{1}{s+1}\right)\right)_{\substack{\emptyset\neq S\subseteq N, s\leq r\\ \emptyset\neq T\subseteq N, t\leq m}}.$$

2.9 Matrix formulation of the quadratic program

Most solvers for solving quadratic programming problems on $\mathbb{R}^p$ use the following matrix formulation:

$$\begin{array}{l} \min_{x\in\mathbb{R}^p} \frac{1}{2}x^t Q x - q^t x \\ \text{subject to } \begin{cases} A_e x = a_e \\ A_i x \geq a_i \end{cases} \end{array} \qquad (13)$$

where $q \in \mathbb{R}^p$, Q is symmetric matrix of order p, the matrix equality $A_e x = a_e$ represents the set of equality constraints and the matrix inequality $A_i x \geq a_i$ represents the set of inequality constraints.

In the considered context of capacity approximation and identification, the vector x to be determined represents the Möbius transform of the unknown k-additive game ν and is defined as in Eq. (9). The form of the matrix Q and of the vector q for each of the three considered quadratic distances immediately follows from Eqs. (10), (11) and (12). The matrices A_e and A_i can be easily obtained from the propositions given in § 2.7 and as discussed in (Kojadinovic, 2006). Examples of the application of the proposed minimum distance approach to capacity approximation and identification are presented in § 4.

3 Quadratic distances for bi-capacity identification and approximation

As mentioned in the introduction, the concept of capacity can be further generalized. In the context of aggregation, *bi-capacities* arise as a natural extension of capacities when the underlying evaluation scale is bipolar (Grabisch and Labreuche, 2005a). In the sequel, we extend the three distances studied in the previous section to bi-cooperative games and bi-capacities, which will allow us to propose quadratic objective functions for bi-capacity approximation and identification.

3.1 Bi-cooperative games and bi-capacities

Let $\mathcal{Q}(N) := \{(A, B) \in \mathcal{P}(N) \times \mathcal{P}(N) \mid A \cap B = \emptyset\}$.

Definition 5. *A function $v : \mathcal{Q}(N) \to \mathbb{R}$ is a bi-cooperative game on N if it satisfies $v(\emptyset, \emptyset) = 0$.*

The set of bi-cooperative games on N shall be denoted by $\mathcal{BG}_N$ as we continue.

Definition 6. *A function $v : \mathcal{Q}(N) \to \mathbb{R}$ is a bi-capacity on N if it is a bi-cooperative game such that, additionally, for any $A, B \subseteq N$, $A \subseteq B$ implies $v(A, \cdot) \leq v(B, \cdot)$ and $v(\cdot, A) \geq v(\cdot, B)$.*

A bi-capacity v is said to be normalized if additionally $v(N, \emptyset) = 1$ and $v(\emptyset, N) = -1$. In the sequel, the set of bi-capacities on N will be denoted by $\mathcal{BC}_N$, and the set of normalized bi-capacities by $\mathcal{BC}_N^*$.

A bi-cooperative game v on N is further said to be of the *Cumulative Prospect Theory (CPT) type* (Grabisch and Labreuche, 2005a,b; Tversky and Kahneman, 1992) if there exist two games μ_1, μ_2 on N such that

$$v(A, B) = \mu_1(A) - \mu_2(B), \qquad \forall\, (A, B) \in \mathcal{Q}(N).$$

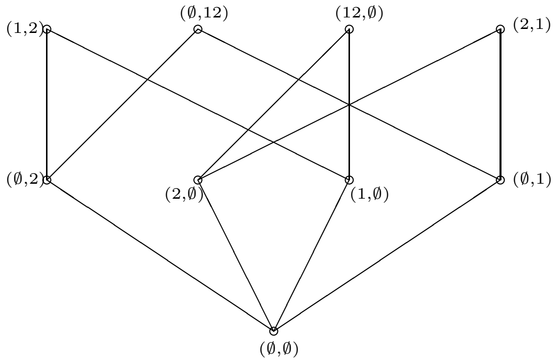

Fig. 1. $(\mathcal{Q}(N), \subseteq)$ with $n = 2$.

When $\mu_2 = \mu_1$ (resp. $\mu_2 = \bar{\mu}_1$), the bi-cooperative game is additionally said to be *symmetric* (resp. *asymmetric*).

We shall further say that a bi-cooperative game v on N is *difference cardinality-based* if, for any $(A, B) \in \mathcal{Q}(N)$, $v(A, B)$ only depends on the difference $a - b$ of cardinalities. It is easy to verify that there is only one normalized bi-capacity on N that is both additive and difference cardinality-based. We shall call it the *uniform* bi-capacity and denote it by v^*. It is defined by

$$v^*(A, B) := \frac{a - b}{n} = \mu^*(A) - \mu^*(B), \qquad \forall\, (A, B) \in \mathcal{Q}(N).$$

The uniform bi-capacity v^* plays the role of the uniform capacity μ^* since, as shall become clear in § 3.3, the Choquet integral w.r.t. v^* is merely the simple arithmetic mean.

3.2 The inf-semilattice $(\mathcal{Q}(N), \subseteq)$

Grabisch and Labreuche (2005a,d,c) and Bilbao et al. (2004) have introduced several orders on $\mathcal{Q}(N)$. As discussed in (Grabisch and Labreuche, 2005c,d), in the considered context, the most appropriate one may simply be the product order:

$$(A, A') \subseteq (B, B') \iff A \subseteq B \text{ and } A' \subseteq B'.$$

The ordered set $(\mathcal{Q}(N), \subseteq)$ is then an inf-semilattice with bottom element $(\emptyset, \emptyset)$ as shown in Figure 1 (see Grabisch and Labreuche, 2005d, for more details).

Let $\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}$ denote the set of maximal chains of $(\mathcal{Q}(N),\subseteq)$. It is easy to verify that the sets $\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}$ and $\mathcal{P}(N)\times\Pi_N$ are in one-to-one correspondence, which implies that $|\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}| = 2^n n!$. Proceeding as in § 2.2, with each subset $N^+\subseteq N$ and each permutation $\sigma\in\Pi_N$, we associate a unique maximal chain $m_{N^+,\sigma}\in\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}$ defined by

$$\begin{aligned} m_{N^+,\sigma} := \big((\emptyset,\emptyset) &\subsetneq \big(\{\sigma(n)\}\cap N^+,\{\sigma(n)\}\cap N^-\big) \\ &\subsetneq \big(\{\sigma(n-1),\sigma(n)\}\cap N^+,\{\sigma(n-1),\sigma(n)\}\cap N^-\big)\subsetneq\dots \\ \dots\subsetneq \big(\{\sigma(1),\dots,\sigma(n)\}&\cap N^+,\{\sigma(1),\dots,\sigma(n)\}\cap N^-\big) = (N^+,N^-)\big), \end{aligned}$$

where $N^- := N\setminus N^+$.

Now, given a bi-cooperative game v on N, with each subset $N^+\subseteq N$ and each permutation $\sigma\in\Pi_N$ (i.e. with each maximal chain $m_{N^+,\sigma}$), we associate a distribution of real numbers $\omega^v_{N^+,\sigma}$ on N defined by

$$\omega^v_{N^+,\sigma}(i) := \nu^v_{N^+}\big(\{\sigma(i),\dots,\sigma(n)\}\big)-\nu^v_{N^+}\big(\{\sigma(i+1),\dots,\sigma(n)\}\big),\qquad \forall\, i\in N, \tag{14}$$

where $\nu^v_{N^+}$ is a game on N defined by

$$\nu^v_{N^+}(C) := v(C\cap N^+, C\cap N^-),\qquad \forall\, C\subseteq N. \tag{15}$$

Equivalently,

$$\omega^v_{N^+,\sigma}(i) = v(m_{n-i+1}) - v(m_{n-i}),\qquad \forall\, i\in N,$$

where $m := m_{N^+,\sigma}$ and, for any $i\in N$, m_i denotes the i-th element of the maximal chain m.

It is important to note that, in general, $\omega^v_{N^+,\sigma}$ is neither a probability distribution nor a distribution of non negative numbers. One exception is when v is an asymmetric normalized bi-capacity (Kojadinovic and Marichal, 2006, Lemma 9).

In (Fujimoto and Murofushi, 2005) and (Grabisch and Labreuche, 2005d), the *Möbius transform* of a bi-cooperative game w.r.t. the inf-semilattice $(\mathcal{Q}(N),\subseteq)$ was obtained. For a bi-capacity v on N, it is given by

$$m_v(S,T) = \sum_{\substack{S'\subseteq S\\ T'\subseteq T}} (-1)^{|S\setminus S'|+|T\setminus T'|} v(S',T'),\qquad \forall\,(S,T)\in\mathcal{Q}(N).$$

The bi-cooperative game can be recovered from the Möbius representation using the following equation:

$$v(S,T) = \sum_{\substack{S'\subseteq S\\ T'\subseteq T}} m_v(S',T'),\qquad \forall\,(S,T)\in\mathcal{Q}(N).$$

3.3 The Choquet integral w.r.t. a bi-cooperative game

In the context of Choquet integral based MAUT, when the utility scale is bipolar, the following natural extension of the Choquet integral has recently been proposed by Grabisch and Labreuche (2005b).

Definition 7. *The Choquet integral of a function $x : N \to \mathbb{R}$, represented by the vector $(x_1, \dots, x_n)$, w.r.t. the bi-cooperative game v on N is defined by*

$$C_v(x) := C_{\nu^v_{N^+}}(|x|),$$

where $\nu^v_{N^+}$ is the game on N defined by Eq. (15) and $N^+ := \{i \in N \mid x_i \geq 0\}$, $N^- := N \setminus N^+$.

As for the Choquet integral w.r.t. a game, the Choquet integral w.r.t. a bi-cooperative game is a piecewise linear function on each set

$$\mathcal{O}_{N^+,\sigma} := \{x \in \mathbb{R}^n \mid x_i \geq 0, \forall\, i \in N^+, x_i < 0, \forall\, i \in N^-, \\ \text{and } |x_{\sigma(1)}| \leq \cdots \leq |x_{\sigma(n)}|\}, \quad (N^+ \subseteq N,\, \sigma \in \Pi_N),$$

which cover $\mathbb{R}^n$. Indeed, for any $x \in \mathbb{R}^n$, there exist $N^+ \subseteq N$ and $\sigma \in \Pi_N$ such that $x \in \mathcal{O}_{N^+,\sigma}$ and,

$$C_v(x) = \sum_{i \in N} \omega^v_{N^+,\sigma}(i)|x_{\sigma(i)}|.$$

The Choquet integral w.r.t. a bi-cooperative game generalizes the symmetric and asymmetric Choquet integrals (Grabisch and Labreuche, 2002) which are obtained by considering symmetric and asymmetric bi-cooperative games respectively. It further generalizes the CPT model of Tversky and Kahneman (1992) which is obtained by considering CPT type bi-capacities.

Let $(A, B) \in \mathcal{Q}(N)$. In the sequel, $(1_A, 1_B, 0_{N\setminus(A\cup B)})$ shall denote an alternative, called *ternary*, whose vector of scores is such that all criteria in A have a score equal to 1, all criteria in B have a score equal to -1, and all the others have a score equal to 0. As an extension of the Choquet integral w.r.t. a game, the Choquet integral w.r.t. a bi-cooperative game v satisfies $C_v(1_A, 1_B, 0_{N\setminus(A\cup B)}) = v(A, B)$ for all $(A, B) \in \mathcal{Q}(N)$. When restricted to vectors of $[-1, 1]^n$, it can be interpreted as the simplest linear interpolation over $[-1, 1]^n$ (Grabisch, 2004).

We end this subsection by providing the expression of the Choquet integral w.r.t the Möbius transform on $(\mathcal{Q}(N), \subseteq)$ obtained in (Fujimoto and

Murofushi, 2005; Grabisch and Labreuche, 2005d). For any bi-cooperative game v on N, we have

$$C_{m_v}(x) = \sum_{(S,T)\in\mathcal{Q}(N)} m_v(S,T)\left[\bigwedge_{i\in S} x_i^+ \wedge \bigwedge_{i\in T} x_i^-\right], \qquad \forall\, x \in \mathbb{R}^n, \tag{16}$$

where $x_i^+ := x_i \vee 0$ and $x_i^- := (-x_i)^+$ for all $i \in N$.

3.4 Uniformity measures

The entropy measures presented in § 2.4 can be further generalized. In (Kojadinovic and Marichal, 2006), the extension of the Shannon entropy to bi-capacities has been defined by

$$\overline{\overline{\overline{H}}}_S(v) := \frac{1}{2^n}\sum_{N^+\subseteq N}\frac{1}{n!}\sum_{\sigma\in\Pi_N} H_S(p^v_{N^+,\sigma}), \qquad \forall\, v \in \mathcal{BC}_N, \tag{17}$$

where, for any $N^+ \subseteq N$, and any $\sigma \in \Pi_N$, $p^v_{N^+,\sigma}$ is the probability distribution on N defined by

$$p^v_{N^+,\sigma}(i) := \frac{|\omega^v_{N^+,\sigma}(i)|}{\sum_{j\in N}|\omega^v_{N^+,\sigma}(j)|}, \qquad \forall\, i \in N. \tag{18}$$

As in the case of capacities, the extended Shannon entropy $\overline{\overline{\overline{H}}}_S(v)$ is nothing else than an average of the uniformity values of the probability distributions $p^v_{N^+,\sigma}$ obtained along the maximal chains of $(\mathcal{Q}(N), \subseteq)$. It has been further shown that $\overline{\overline{\overline{H}}}_S$ has a natural interpretation in the context of aggregation by the Choquet integral w.r.t. a bi-capacity and that it satisfies very appealing properties that make it a natural measure of uniformity.

Clearly, by analogy with the definition given in Eq. (5), the variance of a bi-capacity could be immediately defined as an average of the variances of the probability distributions $p^v_{N^+,\sigma}$, $N^+ \subseteq N$, $\sigma \in \Pi_N$. The resulting quantity would not however be a quadratic form because of the normalization in Eq. (18).

To obtain objective functions that can be used within a quadratic program, we extend the three distances studied in § 2.5 to bi-cooperative games.

3.5 Quadratic distances between bi-cooperative games

We consider the following three functions from $\mathcal{BG}_N \times \mathcal{BG}_N$ to $\mathbb{R}$ defined, for any $v, w \in \mathcal{BG}_N$, by

$$\langle v,w\rangle_1 := \frac{1}{3^n-1} \sum_{\substack{(S,T)\in\mathcal{Q}(N)\\(S,T)\neq(\emptyset,\emptyset)}} v(S,T)w(S,T),$$

$$\langle v,w\rangle_2 := \frac{1}{2^n} \sum_{N^+\subseteq N} \frac{1}{n!} \sum_{\sigma\in\Pi_N} \frac{1}{n} \sum_{i\in N} \omega^v_{N^+,\sigma}(i)\omega^w_{N^+,\sigma}(i),$$

$$\langle v,w\rangle_3 := \frac{1}{2^n} \int_{[-1,1]^n} C_v(x)C_w(x)dx.$$

The two following propositions give more operational expressions of these functions.

Proposition 9. *For any $v, w \in \mathcal{BG}_N$, we have*

$$\langle v,w\rangle_2 = \frac{1}{n} \sum_{i\in N} \sum_{(S,T)\in\mathcal{Q}(N\setminus i)} \frac{\gamma_{s+t}(n)}{2^{s+t+1}} \left[(v(S\cup i,T) - v(S,T))\right.$$
$$\left.\times (w(S\cup i,T) - w(S,T)) + (v(S,T) - v(S,T\cup i))(w(S,T) - w(S,T\cup i))\right].$$

Proof. Let $v, w \in \mathcal{BG}_N$. Using the definitions adopted in § 3.2 and starting from the expression of $\langle v,w\rangle_2$, we can rewrite it as

$$\langle v,w\rangle_2 = \frac{1}{n!2^n n} \sum_{m\in\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}} \sum_{j\in N} (v(m_j) - v(m_{j-1}))\,(w(m_j) - w(m_{j-1}))\,.$$

Let $i \in N$ and $(S,T) \in \mathcal{Q}(N \setminus i)$. The number of maximal chains of $\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}$ containing (S,T) and $(S\cup i, T)$ is $(s+t)!(n-s-t-1)!2^{n-s-t-1}$. Indeed, there are $(s+t)!$ ways to reach (S,T) from $(\emptyset,\emptyset)$. Similarly, from $(S\cup i, T)$, there are $(n-s-t-1)!$ ways to reach an element of the form $(K, N\setminus K)$, where $K \subseteq N \setminus T$, $K \supseteq S \cup i$, and there are $2^{n-s-t-1}$ such elements in $\mathcal{Q}(N)$. The calculation is the same for maximal chains of $\mathcal{M}_{(\mathcal{Q}(N),\subseteq)}$ containing (S,T) and $(S, T\cup i)$.

It follows that, for a given $i \in N$ and a given $(S,T) \in \mathcal{Q}(N \setminus i)$, when summing the terms $\sum_{j\in N} (v(m_j) - v(m_{j-1}))\,(w(m_j) - w(m_{j-1}))$ over the set of maximal chains of $(\mathcal{Q}(N), \subseteq)$, the terms $(v(S\cup i,T) - v(S,T))(w(S\cup i,T) - w(S,T))$ and $(v(S,T) - v(S,T\cup i))(w(S,T) - w(S,T\cup i))$ will appear $(s+t)!(n-s-t-1)!2^{n-s-t-1}$ times. □

Proposition 10. *For any* $v, w \in \mathcal{BG}_N$*, we have*

$$\langle v,w\rangle_3 = \sum_{\substack{(S,T)\in\mathcal{Q}(N)\\(S',T')\in\mathcal{Q}(N)\\S\cup S'\subseteq N\setminus(T\cup T')}} m_v(S,T)m_w(S',T')\frac{1}{2^{|T\cup T'|+|S\cup S'|}} \times \frac{1}{|S\cup S'|+|T\cup T'|+2}\left(\frac{1}{s+t+1}+\frac{1}{s'+t'+1}\right).$$

Proof. Let $v, w \in \mathcal{BG}_N$. From Eq. (16), we can write

$$\langle v,w\rangle_3 = \frac{1}{2^n}\sum_{\substack{(S,T)\in\mathcal{Q}(N)\\(S',T')\in\mathcal{Q}(N)}} m_v(S,T)m_w(S',T') \times \int_{[-1,1]^n}\left[\bigwedge_{i\in S}x_i^+\wedge\bigwedge_{i\in T}x_i^-\right]\left[\bigwedge_{i\in S'}x_i^+\wedge\bigwedge_{i\in T'}x_i^-\right]dx.$$

Now, for any $N^+\subseteq N$, let

$$\mathcal{O}_{N^+} := \{x\in[-1,1]^n \mid \forall\, i\in N^+,\, x_i\geq 0,\ \text{and}\ \forall\, i\in N^-,\, x_i<0\}.$$

Then, for any $(S,T)\in\mathcal{Q}(N)$, and any $(S',T')\in\mathcal{Q}(N\setminus i)$, we have

$$\begin{aligned}
&\int_{[-1,1]^n}\left[\bigwedge_{i\in S}x_i^+\wedge\bigwedge_{i\in T}x_i^-\right]\left[\bigwedge_{i\in S'}x_i^+\wedge\bigwedge_{i\in T'}x_i^-\right]dx\\
&=\int_{[-1,1]^n}\left[\bigwedge_{i\in S}(x_i\vee 0)\wedge\bigwedge_{i\in T}((-x_i)\vee 0)\right]\left[\bigwedge_{i\in S'}(x_i\vee 0)\wedge\bigwedge_{i\in T'}((-x_i)\vee 0)\right]dx\\
&=\sum_{\substack{N^+\subseteq N\setminus(T\cup T')\\N^+\supseteq(S\cup S')}}\int_{O_{N^+}}\left[\bigwedge_{i\in S}x_i\wedge\bigwedge_{i\in T}-x_i\right]\left[\bigwedge_{i\in S'}x_i\wedge\bigwedge_{i\in T'}-x_i\right]dx\\
&=\sum_{\substack{N^+\subseteq N\setminus(T\cup T')\\N^+\supseteq(S\cup S')}}\int_{[0,1]^n}\left[\bigwedge_{i\in S}x_i\wedge\bigwedge_{i\in T}x_i\right]\left[\bigwedge_{i\in S'}x_i\wedge\bigwedge_{i\in T'}x_i\right]dx,
\end{aligned}$$

after a change of variables. Thus, we obtain

$$\int_{[-1,1]^n}\left[\bigwedge_{i\in S}x_i^+\wedge\bigwedge_{i\in T}x_i^-\right]\left[\bigwedge_{i\in S'}x_i^+\wedge\bigwedge_{i\in T'}x_i^-\right]dx = \sum_{\substack{N^+\subseteq N\setminus(T\cup T')\\N^+\supseteq(S\cup S')}}\int_{[0,1]^n}\left[\bigwedge_{i\in S\cup T}x_i\wedge\bigwedge_{i\in S'\cup T'}x_i\right]dx,$$

which, using Lemma 1, is equal to

$$\sum_{\substack{N^+\subseteq N\setminus(T\cup T')\\ N^+\supseteq(S\cup S')}} \frac{1}{|S\cup T\cup S'\cup T'|+2}\left(\frac{1}{s+t+1}+\frac{1}{s'+t'+1}\right).$$

If $S\cup S'\subseteq N\setminus(T\cup T')$, then $\sum_{\substack{N^+\subseteq N\setminus(T\cup T')\\ N^+\supseteq(S\cup S')}} 1 = 2^{n-|T\cup T'|-|S\cup S'|}$, and

$$\int_{[-1,1]^n}\left[\bigwedge_{i\in S} x_i^+\wedge\bigwedge_{i\in T} x_i^-\right]\left[\bigwedge_{i\in S'} x_i^+\wedge\bigwedge_{i\in T'} x_i^-\right]dx$$
$$=\frac{2^{n-|T\cup T'|-|S\cup S'|}}{|S\cup S'|+|T\cup T'|+2}\left(\frac{1}{s+t+1}+\frac{1}{s'+t'+1}\right).$$

If $S\cup S'\not\subseteq N\setminus(T\cup T')$, then

$$\int_{[-1,1]^n}\left[\bigwedge_{i\in S} x_i^+\wedge\bigwedge_{i\in T} x_i^-\right]\left[\bigwedge_{i\in S'} x_i^+\wedge\bigwedge_{i\in T'} x_i^-\right]dx=0,$$

and we obtain the desired result. □

Proposition 11. *The functions $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$, and $\langle\cdot,\cdot\rangle_3$ are inner products in $\mathcal{BG}_N$.*

Proof. It is straightforward that verify that $\langle\cdot,\cdot\rangle_1$ is bi-linear, symmetric, and positive-definite. From Proposition 9, the same clearly holds for $\langle\cdot,\cdot\rangle_2$. The function $\langle\cdot,\cdot\rangle_3$ is clearly symmetric. Its bi-linearity follows from the linearity of the Choquet integral w.r.t. a bi-cooperative game. Finally, $\langle v,v\rangle_3$ is clearly non negative and $\langle v,v\rangle_3=0$ clearly implies that $C_v(x)=0$ for all $x\in[-1,1]^n$, which immediately implies that all the distributions $\omega^v_{N^+,\sigma}$, $N^+\subseteq N$, $\sigma\in\Pi_N$, are zero, and therefore that $v(S,T)=0$ for all $(S,T)\in\mathcal{Q}(N)$. □

The quadratic distances between bi-cooperative games obtained from the previously defined inner products are respectively given, for any $v,w\in\mathcal{BG}_N$, by

$$d_1^2(v,w):=\frac{1}{3^n-1}\sum_{\substack{(S,T)\in\mathcal{Q}(N)\\(S,T)\neq(\emptyset,\emptyset)}}[v(S,T)-w(S,T)]^2,$$

$$d_2^2(v,w):=\frac{1}{2^n}\sum_{N^+\subseteq N}\frac{1}{n!}\sum_{\sigma\in\Pi_N}\frac{1}{n}\sum_{i\in N}[\omega^v_{N^+,\sigma}(i)-\omega^w_{N^+,\sigma}(i)]^2,$$

$$d_3^2(v,w):=\frac{1}{2^n}\int_{[-1,1]^n}[C_v(x)-C_w(x)]^2dx.$$

These three distances being clearly generalizations of those between games presented in § 2.5, their interpretation is almost identical. In the framework of MAUT based on bi-capacities, $d_1(v,w)$ can be regarded as the average quadratic distance between the global evaluations of ternary alternatives; $d_2(v,w)$ corresponds to the average quadratic difference between Choquet integral coefficients; $d_3(v,w)$ measures the average quadratic difference between global evaluations assuming that the alternatives are uniformly distributed over $[-1,1]^n$.

These two last results directly follow from Propositions 9 and 10.

Corollary 3. *For any $v, w \in \mathcal{BG}_N$, we have*

$$d_2^2(v,w) = \frac{1}{n}\sum_{i\in N}\sum_{(S,T)\in\mathcal{Q}(N\setminus i)} \frac{\gamma_{s+t}(n)}{2^{s+t+1}} \Big[(v(S\cup i,T) - v(S,T) - w(S\cup i,T) + w(S,T))^2 + (v(S,T) - v(S,T\cup i) - w(S,T) + w(S,T\cup i))^2\Big].$$

Corollary 4. *For any $v, w \in \mathcal{BG}_N$, we have*

$$d_3^2(v,w) = \sum_{\substack{(S,T)\in\mathcal{Q}(N)\\(S',T')\in\mathcal{Q}(N)\\S\cup S'\subseteq N\setminus(T\cup T')}} \frac{1}{2^{|T\cup T'|+|S\cup S'|}}\,\frac{1}{|S\cup S'|+|T\cup T'|+2} \times\left(\frac{1}{s+t+1}+\frac{1}{s'+t'+1}\right) \times\left[m_v(S,T)m_v(S',T') - 2m_v(S,T)m_w(S',T') + m_w(S,T)m_w(S',T')\right].$$

3.6 Towards bi-capacity approximation and identification

Using the three quadratic distances defined in the previous subsection, the minimum distance approach proposed in § 2.6 for capacities can be straightforwardly extended to bi-capacities. Indeed, the notion of k-additive bi-capacity has been defined by Grabisch and Labreuche (2005d) and the monotonicity and boundary constraints in terms of the Möbius transform have been derived by Fujimoto and Murofushi (2005). For bi-capacity identification, typical objective functions would be $d_1^2(v^*,\cdot)$, $d_2^2(v^*,\cdot)$ or $d_3^2(v^*,\cdot)$. For bi-capacity approximation, besides k-additivity, additional constraints could involve higher monotonicity requirements whose form was recently studied by Fujimoto and Murofushi (2005).

4 Application to capacity identification and approximation

The minimum distance approach based on the three distances among games proposed in § 2.5 was implemented within the `kappalab` package (Grabisch et al., 2005) for the GNU R statistical system (R Development Core Team, 2005). The package is distributed as free software and can be downloaded from the Comprehensive R Archive Network (`http://cran.r-project.org`). The quadratic program is solved using the R `quadprog` package (Turlach and Weingessel, 2004) which implements the dual method of Goldfarb and Idnani (1983) for solving strictly convex quadratic programming problems.

4.1 A simple capacity approximation example

Consider the normalized capacity μ defined on $\{1,2,3,4\}$ given in Table 1 and assume that a 2-additive approximation of it is required.

Table 1. The capacity μ to be approximated and its Möbius representation m_μ.

	1	2	3	4	12	13	14	23	24	34
μ	0.3	0.2	0.3	0.1	0.4	0.5	0.3	0.4	0.4	0.4
m_μ	0.3	0.2	0.3	0.1	-0.1	-0.1	-0.1	-0.1	0.1	0.0

	123	124	134	234	1234
μ	0.8	0.9	1.0	0.7	1.0
m_μ	0.3	0.4	0.5	0.1	-0.9

For each of the three quadratic distances studied in § 2.5, the Möbius representation of the minimum distance 2-additive normalized capacity is given in Table 2.

Table 2. Möbius representations of the three 2-additive solutions for the simple capacity approximation problem.

	1	2	3	4	12	13	14	23	24	34
d_1	0.24	0.24	0.31	0.09	-0.01	0.04	0.09	-0.16	0.09	0.04
d_2	0.28	0.25	0.33	0.11	-0.02	0.03	0.08	-0.17	0.08	0.03
d_3	0.27	0.25	0.33	0.11	-0.02	0.03	0.08	-0.17	0.08	0.03

4.2 A capacity identification example

To illustrate the proposed minimum distance identification approach, we consider an extended version of the problem presented in (Kojadinovic, 2006) concerning the evaluation of students in an institute training econometricians. The students are evaluated w.r.t. five subjects: statistics (S), probability (P), economics (E), management (M) and English (En). The utilities of seven students a, b, c, d, e, f, g on a $[0, 20]$ scale are given in Table 3.

Table 3. Partial evaluations of the five students.

Student	S	P	E	M	En	Mean	d_1	d_2	$d3$
a	18	11	11	11	18	13.80	15.05	15.25	14.95
b	18	11	18	11	11	13.80	14.55	14.75	14.45
c	11	11	18	11	18	13.80	14.05	14.25	13.95
d	18	18	11	11	11	13.80	13.55	13.75	13.45
e	11	11	18	18	11	13.80	13.05	13.25	12.95
f	11	11	18	11	11	12.40	12.55	12.75	12.45
g	11	11	11	11	18	12.40	12.05	12.25	11.95

Assume that the institute is slightly more oriented towards statistics and probability and suppose that the DM considers that there are 3 groups of subjects: statistics and probability, economics and management, and English. Furthermore, he/she considers that within the two first groups, subjects are somewhat substitutive, i.e. they overlap to a certain extent. Finally, if a student is good in statistics or probability (resp. bad in statistics and probability), it is better that he/she is good in English (resp. economics or management) rather than in economics or management (resp. English). This reasoning leads to the following ranking:

$$a \succ_{\mathcal{O}} b \succ_{\mathcal{O}} c \succ_{\mathcal{O}} d \succ_{\mathcal{O}} e \succ_{\mathcal{O}} f \succ_{\mathcal{O}} g.$$

As far as the aggregation function is concerned, the decision maker shows no particular preference. For the sake of simplicity, he/she considers that the simple arithmetic mean would be a good starting point. Given one of the three quadratic distances studied in § 2.5, the minimum distance approach here consists in finding the Möbius representation of a k-additive capacity ν which is the closest to μ^* and that additionally satisfies

$$C_{m_\nu}(a) > C_{m_\nu}(b) > C_{m_\nu}(c) > C_{m_\nu}(d) > C_{m_\nu}(e) > C_{m_\nu}(f) > C_{m_\nu}(g).$$

To be practically usable, the above strict inequalities have to be transformed into non-strict inequalities involving a decision maker defined threshold which we here assume to be equal to 0.5, i.e. an alternative is considered better than another one by the decision maker if their difference in global score is greater than half a unit. From a more technical perspective, these inequalities were implemented using Eq. (4).

By considering students a and b, and f and g, it is easy to see that the criteria do not satisfy mutual preferential independence, which implies that there is no additive model that can numerically represent the above weak order. However, 2-additive models exist. For each of the three quadratic distances, the Möbius representation of the minimum distance 2-additive normalized capacity compatible with the above ranking of the students is given in Table 2. The global scores of the students computed w.r.t. these capacities are given in the last three columns of Table 3. As expected, the Choquet integral w.r.t. all three solutions preserves the ranking of the students provided by the decision maker.

Table 4. Möbius representations of the three 2-additive solutions for the capacity identification problem.

	S	P	E	M	En
d_1	0.35	0.17	0.22	0.14	0.15
d_2	0.34	0.18	0.25	0.15	0.18
d_3	0.33	0.19	0.21	0.16	0.14

	S,P	S,E	S,M	S,En	P,E	P,M	P,En	E,M	E,En	M,En
d_1	-0.15	-0.06	-0.05	0.08	0.03	0.12	-0.01	-0.07	0.06	0.01
d_2	-0.13	-0.06	-0.04	0.09	0.02	0.11	-0.03	-0.08	0.04	-0.01
d_3	-0.17	-0.05	-0.06	0.10	0.03	0.15	-0.02	-0.08	0.08	0.00

One question that remains to be answered is which distance to choose. As in most fields in which distances are used, it is difficult to give an objective answer to that question. In the context of capacity identification, one may claim that distances d_2 and d_3 are the most natural because of their (more obvious) link with the Choquet integral. In the context of capacity approximation however, d_1 might be considered as being the most appropriate because of its simple and natural form.

Acknowledgements The author is very grateful to Jean-Luc Marichal for fruitful discussions.

References

C.A. Bana e Costa, J.M. De Corte, and J.C. Vansnick. On the mathematical foundations of MACBETH. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 409–442. Springer, 2005.

J.M. Bilbao, J.R. Fernández, N. Jimńez, and J.J. López. Probabilistic values for bicooperative games. Technical report, University of Seville, Sevilla, Spain, 2004.

A. Chateauneuf and J-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17(3):263–283, 1989.

G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.

K. Fujimoto and T. Murofushi. Some characterizations of k-monotonicity through the bipolar Möbius transform in bi-capacities. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 9(5):484–495, 2005.

D. Goldfarb and A. Idnani. A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, 27:1–33, 1983.

M. Grabisch. The Choquet integral as a linear interpolator. In *10th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2004)*, pages 373–378, Perugia, Italy, July 2004.

M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1992.

M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92(2):167–189, 1997. ISSN 0165-0114.

M. Grabisch and C. Labreuche. The symmetric and asymmetric Choquet integrals on finite spaces for decision making. *Statistical Papers*, 43:37–52, 2002.

M. Grabisch and C. Labreuche. Bi-capacities I : definition, Möbius transform and interaction. *Fuzzy Sets and Systems*, 151:211–236, 2005a.

M. Grabisch and C. Labreuche. Bi-capacities II : the Choquet integral. *Fuzzy Sets and Systems*, 151:237–259, 2005b.

M. Grabisch and C. Labreuche. A general construction for unipolar and bipolar interpolative aggregation. In *Joint 4th EUSFLAT and 11th LFA Conference (EUSFLAT 2005)*, pages 916–921, Barcelona, Spain, September 2005c.

M. Grabisch and C. Labreuche. Bi-capacities: towards a generalization of cumulative prospect theory. working paper, 2005d.

M. Grabisch, Ch. Labreuche, and J.C. Vansnick. On the extension of pseudo-Boolean functions for the aggregation of interacting criteria. *European Journal of Operational Research*, 148:28–47, 2003.

M. Grabisch, I. Kojadinovic, and P. Meyer. `kappalab`*: Non additive measure and integral manipulation functions*, 2005. URL `http://www.polytech.univ-nantes.fr/kappalab`. R package version 0.2.

J. Havrda and F. Charvat. Quantification method in classification processes: concept of structural α-entropy. *Kybernetika*, 3:30–35, 1967.

R. L. Keeney and H. Raiffa. *Decision with multiple objectives*. Wiley, New-York, 1976.

I. Kojadinovic. *Modeling interaction phenomena using non additive measures : applications in data analysis*. PhD thesis, Université de La Réunion, France, 2002.

I. Kojadinovic. Minimum variance capacity identification. *European Journal of Operational Research*, page in press, 2006.

I. Kojadinovic and J-L. Marichal. Entropy of bi-capacities. *European Journal of Operational Research*, page in press, 2006.

I. Kojadinovic, J-L. Marichal, and M. Roubens. An axiomatic approach to the definition of the entropy of a discrete Choquet capacity. *Information Sciences*, 172:131–153, 2005.

C. Labreuche and M. Grabisch. The Choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems*, 137:11–16, 2003.

T. Marchant. Towards a theory of MCDM: Stepping away from social choice theory. *Mathematical Social Sciences*, 45:343–363, 2003.

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

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

J-L. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European Journal of Operational Research*, 155(3):771–791, 2004.

J-L. Marichal. *Aggregation operators for multicriteria decison aid*. PhD thesis, University of Liège, Liège, Belgium, 1998.

R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2005. URL `http://www.R-project.org`. ISBN 3-900051-00-3.

G-C. Rota. On the foundations of combinatorial theory. I. Theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2:340–368 (1964), 1964.

B. Roy and D. Bouyssou. *Aide multicritère à la décision : méthodes et cas*. Economica, 1993.

C. E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–623, 1948.

L. S. Shapley. A value for n-person games. In *Contributions to the theory of games, vol. 2*, Annals of Mathematics Studies, no. 28, pages 307–317. Princeton University Press, Princeton, N. J., 1953.

M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974.

B.A. Turlach and A. Weingessel. `quadprog`*: Functions to solve quadratic programming problems.*, 2004. R package version 1.4-7.

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

P. Vincke. *Multicriteria Decision-aid*. Wiley, 1992.