

# Approximations of coherent lower probabilities by 2-monotone capacities

Ignacio Montes<sup>1</sup>, Enrique Miranda<sup>1</sup>, and Paolo Vicig<sup>2</sup>

<sup>1</sup> University of Oviedo (Spain), Dep. of Statistics and Operations Research

`imontes@uniovi.es` orcid: 0000-0001-6534-1613

`mirandaenrique@uniovi.es` orcid: 0000-0003-0470-5318

<sup>2</sup> University of Trieste (Italy), DEAMS

`paolo.vicig@deams.units.it` orcid: 0000-0001-8842-442X

**Abstract.** We investigate the problem of approximating a coherent lower probability on a finite space by a 2-monotone capacity that is at the same time as close as possible while not including additional information. We show that this can be tackled by means of a linear programming problem, and investigate the features of the set of undominated solutions. While our approach is based on a distance proposed by Baroni and Vicig, we also discuss a number of alternatives. Finally, we show that our work applies to the more general problem of approximating coherent lower previsions.

**Keywords:** Coherent lower probabilities, 2-monotonicity, coherent lower previsions, distortion models, total variation distance.

## 1 Introduction

Among the many models of imprecise probabilities [1], one of the most general is that of *coherent lower previsions* [2], that can be regarded as sets of expectations with respect to a convex family of finitely additive probability measures. In addition to its generality, it also has a clear behavioural interpretation in terms of acceptable betting rates, as well as the epistemic interpretation in terms of sets of probability measures. Nevertheless, coherent lower previsions (or their restrictions to events, called *coherent lower probabilities*) also have a number of drawbacks that hinder their use in practice: for instance, they have no easy representation in terms of their extreme points in general, and they lack some attractive mathematical properties possessed by more specific models.

One alternative that somewhat solves these issues is to work with 2-monotone capacities, which can be easily determined by means of a finite number of extreme points [3] and that still include as particular cases many of the imprecise probability models from the literature, such as probability intervals [4], belief functions [5] or possibility measures [6]. It is therefore interesting to determine if we can approximate a coherent lower probability by a 2-monotone one with a minimal loss of information. This is the problem we are tackling in this paper.

After giving some preliminary concepts in Section 2, in Section 3 we study the problem of finding undominated outer approximations that minimize the distance to the original model, in the sense proposed by Baroni and Vicig in [7]. In Section 4, we focus on outer approximations by means of some particular subfamilies of 2-monotone capacities and prove that this problem has a unique solution. A comparison with some alternative approaches is given in Section 5. Finally, in Section 6 we show that our results allow us to solve the problem of outer approximating coherent lower previsions. Some additional comments are provided in Section 7. Due to the space limitations, proofs of the results have been omitted.

## 2 Preliminary concepts

Let  $\mathcal{X}$  be a finite space with cardinality  $n$ , and consider a lower probability  $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ . Its associated *credal set* is given by:

$$\mathcal{M}(\underline{P}) = \{P \text{ probability} \mid P(A) \geq \underline{P}(A) \quad \forall A \subseteq \mathcal{X}\},$$

Under an epistemic interpretation of uncertainty, we may regard  $\underline{P}$  as a model for the imprecise knowledge of a probability measure  $P$ , and then  $\mathcal{M}(\underline{P})$  would be the set of candidates for this unknown probability measure. The notion of coherence means that the bounds  $\underline{P}$  gives for the probabilities of the different events are tight:

**Definition 1.** [2] *A lower probability  $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$  is called coherent when  $\mathcal{M}(\underline{P}) \neq \emptyset$  and  $\underline{P}(A) = \min\{P(A) : P \in \mathcal{M}(\underline{P})\}$  for every  $A \subseteq \mathcal{X}$ .*

The conjugate of a coherent lower probability, given by  $\overline{P}(A) = 1 - \underline{P}(A^c)$  for every  $A \subseteq \mathcal{X}$ , is called coherent *upper* probability.

Coherent lower probabilities include as particular cases most of the models of non-additive measures in the literature; they correspond moreover to *balanced* games within game theory [8]. One particular case of coherent lower probabilities are the 2-monotone capacities.

**Definition 2.** [9] *A coherent lower probability  $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$  is called 2-monotone if for every  $A, B \subseteq \mathcal{X}$  it satisfies:*

$$\underline{P}(A \cup B) + \underline{P}(A \cap B) \geq \underline{P}(A) + \underline{P}(B). \quad (1)$$

2-monotone capacities are sometimes called *convex* in the literature. They possess a number of interesting properties that are not always shared with coherent lower probabilities: the extreme points of their credal set can be easily determined using the permutations of the possibility space [3]; moreover, they have a unique extension as an expectation operator that preserves 2-monotonicity: their Choquet integral [10].

For all these reasons, it becomes interesting in practice to approximate a coherent lower probability  $\underline{P}$  by a 2-monotone capacity  $\underline{Q}$  that at the same time

(a) does not introduce new information; and (b) is as close as possible to the original model.

The first constraint is modelled by requiring that the credal set determined by  $\underline{Q}$  includes that of  $\underline{P}$ , or in other words, that  $\underline{Q}(E) \leq \underline{P}(E)$  for every  $E \subseteq \mathcal{X}$ . In that case, we shall say that  $\underline{Q}$  is an *outer approximation* of  $\underline{P}$ .

With respect to the second, one preliminary idea would be to use the partial order associated with the credal set inclusion and to require  $\underline{Q}$  to be *undominated*, in the sense that there is no other 2-monotone capacity  $\underline{Q}'$  such that  $\mathcal{M}(\underline{P}) \subseteq \mathcal{M}(\underline{Q}') \subsetneq \mathcal{M}(\underline{Q})$ . However, this requirement alone does not determine a unique solution, nor does it provide us with a tool to determine the 2-monotone outer approximations, either.

### 3 Approximations by linear programming

In order to overcome the above issues, in this paper we shall consider the outer approximations  $\underline{Q}$  of the coherent lower probability  $\underline{P}$  that minimize the distance proposed by Baroni and Vicig in [7], given by

$$d(\underline{P}, \underline{Q}) := \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - \underline{Q}(E)). \quad (2)$$

If we interpret  $\underline{P}(E) - \underline{Q}(E)$  as the additional imprecision introduced on  $E$  when replacing  $\underline{P}(E)$  with  $\underline{Q}(E)$ , then  $d(\underline{P}, \underline{Q})$  can be understood as the total imprecision added by the outer approximation  $\underline{Q}$ .

To solve the minimization problem, we determine  $\underline{Q}$  through its *Möbius inverse*  $m_{\underline{Q}}$  by means of the formula  $\underline{Q}(E) = \sum_{B \subseteq E} m_{\underline{Q}}(B)$  for every  $E \subseteq \mathcal{X}$  and consider thus the following linear programming problem:

$$\min d(\underline{P}, \underline{Q}) \quad (\text{LP-2monot})$$

subject to:

$$\sum_{E \subseteq \mathcal{X}} m_{\underline{Q}}(E) = 1, \quad m_{\underline{Q}}(\emptyset) = 0. \quad (\text{LP-2monot.1})$$

$$\sum_{\{x_i, x_j\} \subseteq A \subseteq E} m_{\underline{Q}}(A) \geq 0, \quad \forall E \subseteq \mathcal{X}, \forall x_i, x_j \in E, x_i \neq x_j. \quad (\text{LP-2monot.2})$$

$$m_{\underline{Q}}(\{x_i\}) \geq 0, \quad \forall x_i \in \mathcal{X}. \quad (\text{LP-2monot.3})$$

$$\sum_{A \subseteq E} m_{\underline{Q}}(A) \leq \underline{P}(E) \quad \forall E \neq \emptyset, \mathcal{X}. \quad (\text{LP-2monot.4})$$

In fact, (LP-2monot.2) characterizes 2-monotonicity of  $\underline{Q}$  via its Möbius inverse  $m_{\underline{Q}}$  [11], while (LP-2monot.1) and (LP-2monot.3) ensure that  $\underline{Q}$  is also a coherent lower probability. Finally, (LP-2monot.4) means that  $\underline{Q}$  outer approximates  $\underline{P}$ . It is not difficult to check that the number of constraints in this linear programming problem is  $2^n + n + 2^{n-2} \binom{n}{2}$ .

The feasible region of this linear programming problem is non-empty: it suffices to take into account that the *vacuous* lower probability, given by  $\underline{Q}_v(E) = 0$  for every  $E$  that is not equal to  $\mathcal{X}$  and  $\underline{Q}_v(\mathcal{X}) = 1$ , is a 2-monotone outer approximation of any coherent lower probability.

Moreover, the linear programming problem above has an optimal solution by means of Weierstrass' theorem [12]. To see this, note that (i)  $d(\underline{P}, \underline{Q}) = \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - \sum_{B \subseteq E} m_{\underline{Q}}(B))$  is continuous on the variables  $m_{\underline{Q}}(B)$ ; (ii) the feasible region is bounded, since by [13, Theorem 1] the values of  $m_{\underline{Q}}$  are bounded when  $\underline{Q}$  is 2-monotone; and (iii) it is closed, being a polyhedral set in  $\mathbb{R}^{2^n}$ .

Given this, our first result tells us that any solution of the linear programming problem is undominated:

**Proposition 1.** *Let  $\underline{P}$  be a coherent lower probability, and let  $\underline{Q}$  be an optimal solution of the linear programming problem (LP-2monot). Then,  $\underline{Q}$  is an undominated outer approximation of  $\underline{P}$ .*

Not surprisingly, (LP-2monot) may not have a unique solution:

*Example 1.* Consider  $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$  and let  $\underline{P}$  be the coherent lower probability that is the lower envelope of the probability mass functions  $P_1 = (0.5, 0.5, 0, 0)$ ,  $P_2 = (0, 0, 0.5, 0.5)$ . It is given by:

$$\underline{P}(A) = \begin{cases} 0 & \text{if } |A| = 1 \text{ or } A = \{x_1, x_2\}, \{x_3, x_4\} \\ 1 & \text{if } A = \mathcal{X} \\ 0.5 & \text{otherwise.} \end{cases}$$

To see that  $\underline{P}$  is not 2-monotone, note that, given  $A = \{x_1, x_3\}$  and  $B = \{x_2, x_3\}$ ,

$$\underline{P}(A \cup B) + \underline{P}(A \cap B) = 0.5 < 1 = \underline{P}(A) + \underline{P}(B).$$

To see that (LP-2monot) may have more than one solution, note that, if  $\underline{Q}$  is a 2-monotone outer approximation of  $\underline{P}$ , it must satisfy  $\underline{Q}(\{x_1, x_3\}) + \underline{Q}(\{x_2, x_3\}) \leq \underline{Q}(\{x_1, x_2, x_3\}) + \underline{Q}(\{x_3\}) \leq 0.5$ , whence  $\underline{P}(\{x_1, x_3\}) + \underline{P}(\{x_2, x_3\}) - \underline{Q}(\{x_1, x_3\}) - \underline{Q}(\{x_2, x_3\}) \geq 0.5$ ; similarly, we obtain that  $\underline{P}(\{x_1, x_4\}) + \underline{P}(\{x_2, x_4\}) - \underline{Q}(\{x_1, x_4\}) - \underline{Q}(\{x_2, x_4\}) \geq 0.5$ , and therefore  $d(\underline{P}, \underline{Q}) \geq 1$  for any 2-monotone outer approximation of  $\underline{P}$ . This distance is attained by the 2-monotone capacities  $\underline{Q}_1, \underline{Q}_2$  given by

$$\underline{Q}_1(A) = \begin{cases} 0 & \text{if } |A| = 1 \text{ or } A = \{x_1, x_2\}, \{x_3, x_4\} \\ 0.5 & \text{if } |A| = 3 \\ 1 & \text{if } A = \mathcal{X} \\ 0.25 & \text{otherwise.} \end{cases}$$

and

$$\underline{Q}_2(A) = \begin{cases} 0 & \text{if } |A| = 1 \text{ or } A = \{x_1, x_2\}, \{x_3, x_4\} \\ 0.5 & \text{if } |A| = 3 \\ 1 & \text{if } A = \mathcal{X} \\ 0.2 & \text{if } A = \{x_1, x_4\}, \{x_2, x_3\} \\ 0.3 & \text{otherwise.} \end{cases}$$

Their 2-monotonicity can easily be verified by means of Eq. (1). ♦

Obviously, if our initial model  $\underline{P}$  is not 2-monotone, any undominated 2-monotone capacity that outer approximates  $\underline{P}$  will not agree with  $\underline{Q}$  on some event  $A$ . Interestingly, it can be checked that both models always agree on singletons:

**Proposition 2.** *Let  $\underline{P}$  be a coherent lower probability. If  $\underline{Q}$  is an undominated 2-monotone capacity that outer approximates  $\underline{P}$ , then  $\underline{Q}(\{x\}) = \underline{P}(\{x\})$  for every  $x \in \mathcal{X}$ .*

As a consequence, both of them induce the same order on  $\mathcal{X}$ . It can be checked that this property does not extend to some particular subfamilies of 2-monotone capacities, such as belief functions.

## 4 Particular cases

In this section, we investigate the outer approximations of a coherent lower probability in some subfamilies of 2-monotone capacities associated with distortion models. With the term *distortion model* we refer to a model where an initial probability measure  $P_0$  is modified in some sense.

### 4.1 Pari-mutuel models

We begin by considering the *Pari Mutuel Model* [2,14,15] (PMM, for short). This is a betting scheme originated in horse racing. It is determined by two elements: a probability measure  $P_0$  and a distortion factor  $\delta > 0$ . For every event  $A$  of  $\mathcal{P}(\mathcal{X})$ ,  $P_0(A)$  is interpreted as a fair prize for a bet on  $A$ , and  $\delta > 0$  denotes the loading of the house. They determine a coherent lower probability by:

$$\underline{P}(A) = \max\{0, (1 + \delta)P_0(A) - \delta\} \quad \forall A \subseteq \mathcal{X}. \quad (3)$$

The lower probability associated with a PMM is 2-monotone, as shown for instance in [15, Section 2]. Moreover, in [14] it is proven that PMMs correspond to particular instances of *probability intervals* [4].

Our next result gives the unique undominated outer approximation of a coherent lower probability in terms of pari mutuel models.

**Proposition 3.** *Let  $\underline{P}$  be a coherent lower probability with conjugate upper probability  $\bar{P}$ . Define the constant value  $\delta > 0$  and the probability  $P_0$  by:*

$$\delta = \sum_{i=1}^n \bar{P}(\{x_i\}) - 1, \quad P_0(\{x_i\}) = \frac{\bar{P}(\{x_i\})}{1 + \delta} \quad \forall i = 1, \dots, n.$$

Denote by  $\underline{Q}$  the coherent lower probability associated with the PMM  $(P_0, \delta)$  by means of Eq. (3). Then,  $\underline{Q}$  is the unique undominated pari mutuel model that outer approximates  $\underline{P}$ .

## 4.2 $\varepsilon$ -contamination models

Another distortion model is the  $\varepsilon$ -contamination model, also called *linear-vacuous mixture* in [2]. Given a probability measure  $P_0$  and  $\varepsilon \in (0, 1)$ , they determine the coherent lower probability

$$\underline{P}(A) = \begin{cases} (1 - \varepsilon)P_0(A) & \text{if } A \neq \mathcal{X}. \\ 1 & \text{if } A = \mathcal{X}. \end{cases} \quad (4)$$

Equivalently,  $\underline{P} = (1 - \varepsilon)P_0 + \varepsilon\underline{Q}_v$ . The lower probability induced by such a model is 2-monotone. This follows from the fact that it satisfies an even stronger property: complete monotonicity, as can be deduced for instance from [10, Theorems 5 and 11].

As with the PMM, we prove that there is only one undominated outer approximation for a coherent lower probability in terms of  $\varepsilon$ -contamination models.

**Proposition 4.** *Let  $\underline{P}$  be a coherent lower probability satisfying the condition  $\sum_{j=1}^n \underline{P}(\{x_j\}) > 0$ . Define  $\varepsilon \in (0, 1)$  and the probability  $P_0$  by:*

$$\varepsilon = 1 - \sum_{j=1}^n \underline{P}(\{x_j\}), \quad P_0(\{x_i\}) = \frac{\underline{P}(\{x_i\})}{\sum_{j=1}^n \underline{P}(\{x_j\})} \quad \forall i = 1, \dots, n.$$

Denote by  $\underline{P}_\varepsilon$  the  $\varepsilon$ -contamination model they determine by means of Eq. (4). Then,  $\underline{P}_\varepsilon$  is the unique undominated  $\varepsilon$ -contamination model that outer approximates  $\underline{P}$ .

Note that the assumption  $\sum_{j=1}^n \underline{P}(\{x_j\}) > 0$  in this proposition is necessary for the existence of some outer approximation: if  $\underline{P}(\{x_j\}) = 0$  for every  $x_j \in \mathcal{X}$ , any  $\varepsilon$ -contamination model that outer approximates  $\underline{P}_\varepsilon$  should also satisfy  $\underline{P}_\varepsilon(\{x_j\}) = 0$  for every  $x_j \in \mathcal{X}$ , whence

$$\underline{P}_\varepsilon(\{x_j\}) = (1 - \varepsilon)P_0(\{x_j\}) = 0 \quad \forall x_j \in \mathcal{X},$$

where  $P_0$  is the precise probability in the  $\varepsilon$ -contamination model. However, since  $\varepsilon \in (0, 1)$ , it follows that  $P_0(\{x_j\}) = 0$  for every  $x_j \in \mathcal{X}$  and  $P_0(\mathcal{X}) = \sum_{j=1}^n P_0(\{x_j\}) = 0$ , a contradiction.

## 5 Comparison with other approaches

In this section, we briefly explore other alternatives to the linear programming approach we have considered so far, in order to justify better our choice.

### 5.1 Quadratic problems

As Example 1 shows, the linear programming problem (LP-2monot) may not have a unique solution. One way to overcome this issue is to consider, instead of the distance given by Eq. (2), the quadratic distance given by:

$$\tilde{d}(\underline{P}, \underline{Q}) := \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - \underline{Q}(E))^2.$$

It is not difficult to prove that, for any coherent lower probability  $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ , there is a unique 2-monotone capacity  $\underline{Q} \leq \underline{P}$  that minimizes  $\tilde{d}(\underline{P}, \underline{Q})$ . From this it follows that  $\underline{Q}$  is therefore undominated in the family of outer approximations of  $\underline{P}$  by 2-monotone capacities. Note this outer approximation need not be one of the solutions of the linear programming problem (LP-2monot).

In spite of this positive result, while in our view the distance of Baroni and Vicig may be interpreted as the additional imprecision introduced by the outer approximation, a similar interpretation of the quadratic distance is not immediate; further, summing squares of differences in  $[0, 1]$  the solution of the quadratic problem may seem closer to the original model than it actually is.

## 5.2 The total variation distance

Another possibility would be to consider an extension of the *total variation distance* [16, Chapter 4.1] to the imprecise case. Recall that given two probability measures  $P_1$  and  $P_2$ , their *total variation* is defined as

$$\|P_1 - P_2\| = \max_{E \subseteq \mathcal{X}} |P_1(E) - P_2(E)|.$$

This definition can be equivalently expressed as:

$$\|P_1 - P_2\| = \frac{1}{2} \sum_{x \in \mathcal{X}} |P_1(\{x\}) - P_2(\{x\})|.$$

In an imprecise framework, given two coherent lower probabilities  $\underline{P}_1, \underline{P}_2$ , we can extend the definition above in a number of (not necessarily equivalent) ways:

$$\begin{aligned} d_1(\underline{P}_1, \underline{P}_2) &= \max_{E \subseteq \mathcal{X}} |\underline{P}_1(E) - \underline{P}_2(E)|, \\ d_2(\underline{P}_1, \underline{P}_2) &= \frac{1}{2} \sum_{x \in \mathcal{X}} |\underline{P}_1(\{x\}) - \underline{P}_2(\{x\})|, \\ d_3(\underline{P}_1, \underline{P}_2) &= \sup_{P_1 \in \mathcal{M}(\underline{P}_1), P_2 \in \mathcal{M}(\underline{P}_2)} \|P_1 - P_2\|, \end{aligned}$$

and we refer to [1, Section 11.4] for some comments on  $d_1$  in the context of imprecise Markov chains.

However, all these extensions may lead to outer approximations that are dominated, and therefore cannot be considered adequate for our problem, as the next examples show.

*Example 2.* Consider a four element space and the lower probability  $\underline{P}$  defined in the following table:

$A$	$\underline{P}(A)$	$\underline{Q}'_1(A)$	$\underline{Q}'_2(A)$
$\{x_1\}$	0.1	0.1	0.1
$\{x_2\}$	0	0	0
$\{x_3\}$	0	0	0
$\{x_4\}$	0.1	0.1	0.1
$\{x_1, x_2\}$	0.4	0.3	0.3
$\{x_1, x_3\}$	0.4	0.3	0.3
$\{x_1, x_4\}$	0.4	0.4	0.3
$\{x_2, x_3\}$	0.2	0.2	0.2
$\{x_2, x_4\}$	0.4	0.3	0.3
$\{x_3, x_4\}$	0.4	0.3	0.3
$\{x_1, x_2, x_3\}$	0.5	0.5	0.5
$\{x_1, x_2, x_4\}$	0.6	0.6	0.6
$\{x_1, x_3, x_4\}$	0.6	0.6	0.6
$\{x_2, x_3, x_4\}$	0.5	0.5	0.5
$\mathcal{X}$	1	1	1

Note that  $\underline{P}$  is a coherent lower probability because it is the lower envelope of the probability measures with mass functions

$$(0.4, 0, 0.2, 0.4), (0.3, 0.1, 0.1, 0.5), (0.3, 0.3, 0.3, 0.1), (0.1, 0.3, 0.3, 0.3) \\ (0.4, 0.2, 0, 0.4), (0.2, 0.2, 0.4, 0.2), (0.2, 0.4, 0.2, 0.2), (0.5, 0.1, 0.1, 0.3).$$

To see that it is not 2-monotone, note that, taking  $A = \{x_1, x_2\}$  and  $B = \{x_1, x_3\}$  it holds that:

$$\underline{P}(\{x_1, x_2, x_3\}) + \underline{P}(\{x_1\}) = 0.6 < 0.8 = \underline{P}(\{x_1, x_2\}) + \underline{P}(\{x_1, x_3\}).$$

Therefore, any outer approximation  $\underline{Q}$  in the class of 2-monotone lower probabilities must satisfy:

$$\underline{Q}(\{x_1, x_2\}) + \underline{Q}(\{x_1, x_3\}) \leq \underline{P}(\{x_1, x_2\}) + \underline{P}(\{x_1, x_3\}) - 0.2.$$

Hence,  $d_1(\underline{P}, \underline{Q}) \geq 0.1$ . Also, the previous inequality is indeed an equality, which is attained, for example, by the 2-monotone capacities  $\underline{Q}'_1, \underline{Q}'_2$  in the table above. Thus, both  $\underline{Q}'_1, \underline{Q}'_2$  are optimal outer approximations with respect to the distance  $d_1$ , even if  $\underline{Q}'_2$  is dominated by  $\underline{Q}'_1$ .  $\blacklozenge$

*Example 3.* Consider again the coherent lower probability from Example 1. Any 2-monotone outer approximation  $\underline{Q}$  of  $\underline{P}$ , undominated or not, shall satisfy  $\underline{Q}(\{x_j\}) = 0$  for every  $j$ , and as a consequence  $d_2(\underline{P}, \underline{Q}) = 0$ . As for  $d_3$ , since  $\|P_1 - P_2\| = 1$  for the probability measures  $P_1 = (0.5, 0.5, 0, 0)$  and  $P_2 = (0, 0, 0.5, 0.5)$  from  $\mathcal{M}(\underline{P})$ , and by definition this is the maximum value of the total variation, we deduce that  $d_3(P_1, P_2) = 1$ . Because  $\mathcal{M}(\underline{P}) \subset \mathcal{M}(\underline{Q})$ , also  $d_3(\underline{P}, \underline{Q}) = 1$  for any 2-monotone outer approximation  $\underline{Q}$  of  $\underline{P}$ , even for the ‘most dominated’ vacuous lower probability  $\underline{Q}_v$ . Thus,  $d_2, d_3$  do not rule out the undominated solutions, either.  $\blacklozenge$



### 5.3 The Weber set

We have already mentioned that one of the advantages of 2-monotone capacities is the existence of a simple procedure to obtain the extreme points of the associated credal set. Let  $\underline{P}$  be a 2-monotone capacity, and for any permutation  $\sigma$  of  $\{1, \dots, n\}$ , define the precise probability  $P_\sigma$  by means of the constraints

$$P_\sigma(\{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}) = \underline{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}) \quad (5)$$

for  $k = 2, \dots, n$ . It was first proven by Shapley [3] that, if  $S_n$  denotes the set of permutations of  $\{1, \dots, n\}$ , it holds that  $\text{ext}(\mathcal{M}(\underline{P})) = \{P_\sigma \mid \sigma \in S_n\}$ . In general, even when  $\underline{P}$  is not 2-monotone but only coherent, we can define the set of probabilities:

$$W(\underline{P}) = \{P_\sigma \mid \sigma \in S_n\}, \quad (6)$$

where  $P_\sigma$  is defined as in Eq. (5). This set is called the *Weber set* of  $\underline{P}$ , and it holds that [17]  $\underline{P}$  is 2-monotone if and only if  $\text{ext}(\mathcal{M}(\underline{P})) = W(\underline{P})$ . Otherwise,  $\mathcal{M}(\underline{P})$  is a proper subset of  $\text{conv}(W(\underline{P}))$ . This implies that the lower envelope of  $\text{conv}(W(\underline{P}))$  is a coherent lower probability that outer approximates  $\underline{P}$ .

In fact, in the case of cardinality four, the lower envelope of  $\text{conv}(W(\underline{P}))$  is indeed 2-monotone:

**Proposition 5.** *Let  $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$  be a coherent lower probability, where  $|\mathcal{X}| \leq 4$ , and denote by  $\underline{Q}$  the coherent lower probability defined by  $\underline{Q}(E) = \min\{P(E) \mid P \in \text{conv}(W(\underline{P}))\}$  for every  $E \subseteq \mathcal{X}$ , where  $W(\underline{P})$  is given by Eq. (6). Then,  $\underline{Q}$  is a 2-monotone outer approximation of  $\underline{P}$ .*

It can be checked that the lower envelope of the Weber set is not necessarily 2-monotone for cardinalities greater than four. Moreover, even in the case of cardinality four the lower envelope of the Weber set is not in general an undominated outer approximation:

*Example 4.* Consider a four-element space  $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ , and the lower probability  $\underline{P}$  given in the following table:

A	$\underline{P}(A)$	$\underline{Q}(A)$	$\underline{Q}'(A)$
$\{x_1\}$	0.1	0.1	0.1
$\{x_2\}$	0	0	0
$\{x_3\}$	0	0	0
$\{x_4\}$	0.3	0.3	0.3
$\{x_1, x_2\}$	0.1	0.1	0.1
$\{x_1, x_3\}$	0.3	0.2	0.3
$\{x_1, x_4\}$	0.6	0.5	0.5
$\{x_2, x_3\}$	0.3	0.2	0.2
$\{x_2, x_4\}$	0.4	0.3	0.4
$\{x_3, x_4\}$	0.4	0.3	0.4
$\{x_1, x_2, x_3\}$	0.5	0.5	0.5
$\{x_1, x_2, x_4\}$	0.6	0.6	0.6
$\{x_1, x_3, x_4\}$	0.7	0.7	0.7
$\{x_2, x_3, x_4\}$	0.6	0.6	0.6
$\mathcal{X}$	1	1	1

If we compute  $\underline{Q} = \min\{P \mid P \in \text{conv}(W(\underline{P}))\}$ , we obtain the values depicted in the table above. However, this 2-monotone capacity is dominated by the 2-monotone outer approximation  $\underline{Q}'$  given in the same table.  $\blacklozenge$

## 6 Approximations of coherent lower previsions

The problem considered in this paper could be generalized from coherent lower probabilities to the richer framework of coherent *lower previsions* [2]: if we denote by  $\mathcal{L}(\mathcal{X})$  the set of bounded real-valued functions on  $\mathcal{X}$ , a coherent lower prevision is a function  $\underline{P} : \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R}$  that satisfies

- $\underline{P}(f) \geq \inf f$
- $\underline{P}(\lambda f) = \lambda \underline{P}(f)$
- $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$

for every  $f, g \in \mathcal{L}(\mathcal{X})$  and every  $\lambda > 0$ . Equivalently,  $\underline{P}$  is coherent when it is the lower envelope of a set of expectation operators with respect to a family of probability measures on  $\mathcal{X}$ .

The notion of 2-monotonicity has also been extended to lower previsions:  $\underline{P}$  is a 2-monotone lower prevision if and only if

$$\underline{P}(f \wedge g) + \underline{P}(f \vee g) \geq \underline{P}(f) + \underline{P}(g) \quad \forall f, g \in \mathcal{L}(\mathcal{X}),$$

where  $\wedge$  and  $\vee$  denote the pointwise minimum and maximum. In general, a coherent lower probability  $\underline{P}$  on  $\mathcal{P}(\mathcal{X})$  may have more than one extension as a coherent lower prevision on  $\mathcal{L}(\mathcal{X})$ ; however, if  $\underline{P}$  is 2-monotone, then it has a unique extension to  $\mathcal{L}(\mathcal{X})$  as a 2-monotone lower prevision: its Choquet integral [10], that is also its least-committal or *natural extension* [2].

Similarly to what we have done in the rest of the paper, we could study the problem of outer approximating a coherent lower prevision by a 2-monotone one. Interestingly, this problem turns out to be equivalent to the one we are considering in this paper, as our next result shows:

**Theorem 1.** *Let  $\underline{P} : \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R}$  be a coherent lower prevision, and let  $\underline{P}'$  be its restriction to events. Then, there is a one-to-one correspondence between the sets*

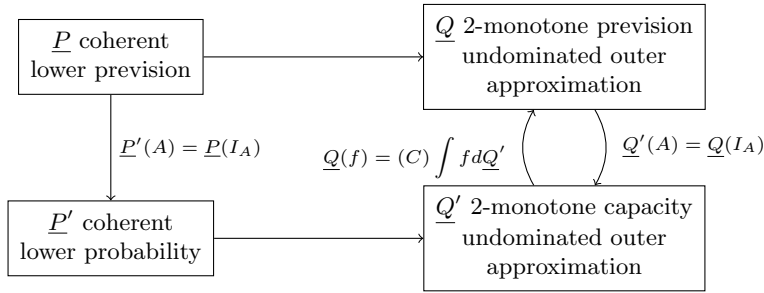
$$\{\underline{Q} : \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R} \text{ 2-monotone undominated outer approximation of } \underline{P}\}$$

and

$$\{\underline{Q}' : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1] \text{ 2-monotone undominated outer approx. of } \underline{P}'\}.$$

The key in this result is that if we want to outer approximate a coherent lower prevision, we can simply consider its restriction to events, outer approximate it and then apply the procedure of natural extension in [2]. Figure 1 illustrates the procedure.

Therefore, it suffices to focus on outer approximations of coherent lower probabilities instead of lower previsions.



**Fig. 1.** Correspondence between the 2-monotone outer approximations.

## 7 Conclusions

Our results allow us to conclude that we can find undominated outer approximations of a coherent lower probability that are at the same time as close as possible, in the sense of Baroni and Vicig, by means of a suitable linear programming problem. Although the approximation is not unique in general, it is so if we focus on some particular subfamilies of 2-monotone capacities, such as those associated with distortion models. Moreover, the problem can be immediately applied to the approximation of coherent lower previsions by 2-monotone ones.

While in our view the distance we have considered is the most meaningful for the problem at hand and the results in Section 5 support this somewhat, we should also make a more thorough comparison with other distances from the literature, and also with the related study made in [18] about outer approximations with possibility measures.

As other future lines of research, we would like to study in more detail the loss of information entailed by the outer approximations, as well as the elicitation among them when there is more than one solution. In addition, we would also like to investigate more deeply the features of the solutions obtained by means of the quadratic approach. Finally, it may be interesting to consider the problem of the inner approximations of a coherent lower probability, even if they entail removing, perhaps unjustifiedly, some imprecision from our model.

## Acknowledgements

The research in this paper has been supported by project TIN2014-59543-P. We would also like to thank Sébastien Destercke for some helpful suggestions.

## References

1. Augustin, T., Coolen, F., de Cooman, G., Troffaes, M., eds.: Introduction to Imprecise Probabilities. Wiley Series in Probability and Statistics. Wiley (2014)

2. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)
3. Shapley, L.S.: Cores of convex games. *International Journal of Game Theory* **1** (1971) 11–26
4. de Campos, L.M., Huete, J.F., Moral, S.: Probability intervals: a tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **2** (1994) 167–196
5. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ (1976)
6. Dubois, D., Prade, H.: Possibility theory: qualitative and quantitative aspects. In Smets, P., ed.: *Handbook on Defeasible Reasoning and Uncertainty Management Systems. Volume 1: Quantified Representation of Uncertainty and Imprecision.* Kluwer Academic Publishers, Dordrecht (1998) 169–226
7. Baroni, P., Vicig, P.: An uncertainty interchange format with imprecise probabilities. *International Journal of Approximate Reasoning* **40** (2005) 147–180
8. Grabisch, M.: Set functions, games and capacities in decision making. Springer (2016)
9. Choquet, G.: Theory of capacities. *Annales de l’Institut Fourier* **5** (1953–1954) 131–295
10. de Cooman, G., Troffaes, M.C.M., Miranda, E.:  $n$ -Monotone exact functionals. *Journal of Mathematical Analysis and Applications* **347** (2008) 143–156
11. Chateauneuf, A., Jaffray, J.Y.: Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences* **17**(3) (1989) 263–283
12. Rudin, W.: Principles of mathematical analysis. McGraw Hill (1976)
13. Grabisch, M., Miranda, P.: Exact bounds of the Möbius inverse of monotone set functions. *Discrete Applied Mathematics* **186** (2015) 7–12
14. Montes, I., Miranda, E., Destercke, S.: A study of the Pari-Mutuel Model from the point of view of imprecise probabilities. In: *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA’2017).* (2017)
15. Pelessoni, R., Vicig, P., Zaffalon, M.: Inference and risk measurement with the pari-mutuel model. *International Journal of Approximate Reasoning* **51** (2010) 1145–1158
16. Levin, D.A., Peres, Y., Wilmer, E.: Markov Chains and Mixing Times. American Mathematical Society (2009)
17. Ichiishi, T.: Supermodularity: Applications to convex games and to the greedy algorithm for LP. *Journal of Economic Theory* **25** (1981) 283–286
18. Dubois, D., Prade, H.: Consonant approximations of belief functions. *International Journal of Approximate Reasoning* **4**(5-6) (1990) 419–449