

OUTER APPROXIMATING COHERENT LOWER PROBABILITIES WITH BELIEF FUNCTIONS

IGNACIO MONTES, ENRIQUE MIRANDA AND PAOLO VICIG

ABSTRACT. From an epistemic point of view, coherent lower probabilities allow us to model the imprecise information about a partially unknown probability. However, there are some issues that hinder their use in practice. Since belief functions are easier to deal with, we propose to approximate the coherent lower probability by a belief function that is at the same time as close as possible to the initial coherent lower probability while not including additional information. We show that this problem can be tackled by means of linear programming, and investigate the features of the set of optimal solutions. Moreover, we emphasize the differences with the outer approximations by 2-monotone lower probabilities. We also study the problem for two particular cases of belief functions that are computationally easier to handle: necessity measures and probability boxes.

Keywords: Coherent lower probabilities, belief functions, possibility measures, probability boxes, 2-monotonicity.

1. INTRODUCTION

In an experiment with imprecise or ambiguous information, it may be difficult to elicit a unique probability model that accurately formalizes our uncertainty; instead, it is arguably more realistic to take an epistemic approach and to work with the set of probability distributions that are compatible with the available information. Such a set is referred to as a *credal set* [24], and it can be summarized by means of the *lower* and *upper* probabilities that can be obtained by taking lower and upper envelopes. These two functions give the tightest lower and upper bounds of the probability of any event, representing the information at our disposal, and are an instance of imprecise probability models [1].

The lower and upper probabilities associated with a credal set satisfy the property of *coherence* [36], that allows to give them a behavioural interpretation similar to the one considered by Bruno de Finetti [10] in the context of subjective probability. However, as an imprecise probability model they also have a number of drawbacks that hinder somewhat their use in practice: (a) they do not have a unique extension as an expectation operator, meaning that we must instead consider the model of coherent lower and upper *previsions*; (b) the verification of the property of coherence for some given lower and upper probabilities may be difficult; and (c) the structure of the

credal set associated with a pair of coherent lower and upper probabilities in terms of extreme points is not straightforward.

The reasons above motivate the use of other imprecise probability models that, while slightly less expressive, are easier to use in practice. In [3], the authors proposed to outer approximate a coherent lower probability by means of a *2-monotone* one. Following this idea, in a previous contribution [27] we studied in detail how to obtain outer approximations in terms of 2-monotone lower probabilities as well as their features. These have a number of advantages over coherent lower and upper probabilities: for instance, the property of 2-monotonicity determines a unique extension as an expectation operator (the Choquet integral [5]), and their associated credal set has a neat structure in terms of its extreme points [32].

This is not to say, however, that 2-monotone lower probabilities are without drawbacks: for instance, their interpretation is not as clear as that of coherent lower and upper probabilities, even if there exists a link with comonotone additivity [8]; moreover, the computational cost of working with them may still be high.

For this reason, in this paper we go a bit further and consider a particular case of 2-monotone lower probabilities that has a clearer interpretation and is at the same time computationally simpler: belief functions [31]. They were studied in detail by Shafer in the context of evidence theory, and had also appeared earlier in connection with multi-valued mappings [11] and with the property of complete monotonicity [5]. As particular cases of lower probabilities, belief functions are uniquely determined by their Möbius inverse; however, unlike arbitrary lower probabilities, the Möbius inverse of a belief function is non-negative, and can thus be interpreted as a basic probability assignment. Because of this, they benefit from the machinery established in evidence and in random set theory [26, 29, 31].

In this work we shall look for the closest (in a sense that shall be specified later on) belief function that outer approximates a given coherent lower probability, meaning that it does not introduce any additional information. While we shall prove that there is always one such approximation, we will also see that our problem does not have a unique solution in general, and that in fact there may be infinitely many. Moreover, we shall show that the infinite set of solutions may not even be convex. For this reason, we will study a number of particular cases of belief functions that, although entailing a loss of expressive power, are also computationally simpler, such as belief functions defined on ternary spaces, possibility/necessity measures and probability boxes. In these two cases, we shall establish that the set of optimal outer approximations is finite and can be easily determined.

Some earlier works in the literature also studied this, or related approaches: in [6], de Campos defined an inclusion relation between fuzzy measures that in certain cases is equivalent to being an outer approximation; in [14, 17, 18], Dubois and Prade investigated the problem of outer approximating a belief

function by means of a necessity measure; and some algorithms for approximating a coherent lower probability by means of a belief function were proposed in [21, 30]. We shall give more detailed information about these works at various points in the paper.

The contribution is organized as follows: after giving some preliminary concepts in Section 2, Section 3 summarizes our previous work [27] on the problem of outer approximating a coherent lower probability with a 2-monotone one. Section 4 contains the main bulk of our work: we investigate the same problem in terms of belief functions, showing that we can solve it by means of linear programming (Section 4.1), studying the properties of the set of solutions (Section 4.2) and characterizing this set on the particular case of ternary spaces (Section 4.3). Also, in Section 5 we compare our approach with some previous works on the matter. Finally, in Sections 6.1 and 6.2 we consider the particular cases of necessity measures and p-boxes, respectively. Some additional comments are given in Section 7.

2. PRELIMINARIES

In this section we introduce the main concepts we shall use in this paper: coherent lower previsions and probabilities, belief functions, possibility measures and p-boxes.

2.1. Lower previsions. Consider a finite possibility space $\mathcal{X} = \{x_1, \dots, x_n\}$ with cardinality n . A *gamble* is a real-valued function $f : \mathcal{X} \rightarrow \mathbb{R}$, and the set of all the gambles in \mathcal{X} is denoted by $\mathcal{L}(\mathcal{X})$. A *lower prevision* defined on $\mathcal{K} \subseteq \mathcal{L}(\mathcal{X})$ is a function $\underline{P} : \mathcal{K} \rightarrow \mathbb{R}$. In this paper, we follow an *epistemic* interpretation, so we regard $\underline{P}(f)$ as a lower bound of the expectation of f with respect to some partially unknown probability P_0 , $E_{P_0}(f)$. With this in mind, the set of probability measures that are compatible with the information given by \underline{P} is:

$$\mathcal{M}(\underline{P}) = \{P \text{ probability measure} \mid P(f) \geq \underline{P}(f) \forall f \in \mathcal{K}\}. \quad (1)$$

$\mathcal{M}(\underline{P})$ is usually called a *credal set* following the terminology of Levi [24].

Some particular cases of lower previsions have been considered, defined in terms of rationality criteria. The minimal requirement on \underline{P} we shall consider in this paper is that the bounds it provides on the expectations of the gambles of \mathcal{K} are tight:

Definition 1. [36] *A lower prevision \underline{P} is called coherent when its associated credal set $\mathcal{M}(\underline{P})$ is non-empty and \underline{P} is the lower envelope of this set: $\underline{P}(f) = \min_{P \in \mathcal{M}(\underline{P})} P(f)$ for every $f \in \mathcal{K}$.*

The conjugate of a lower prevision \underline{P} , denoted by \overline{P} , is called *upper prevision* and it is given by $\overline{P}(f) = -\underline{P}(-f)$ for every f such that $-f \in \mathcal{K}$. $\overline{P}(f)$ can be interpreted as an upper bound for the expectation of f with respect to the partially unknown probability P_0 . When \underline{P} is coherent, \overline{P} can be computed by $\overline{P}(f) = \max_{P \in \mathcal{M}(\underline{P})} P(f)$ for every f such that $-f \in \mathcal{K}$.

2.2. Lower probabilities. One particular case of lower previsions are those whose domain contains only indicator functions: I_A for some $A \subseteq \mathcal{X}$. In that case, we will use the notation $\underline{P}(A)$ instead of $\underline{P}(I_A)$, and the lower prevision will be called *lower probability*.

Formally, a *lower probability* on $\mathcal{K} \subseteq \mathcal{P}(\mathcal{X})$ is a function $\underline{P} : \mathcal{K} \rightarrow [0, 1]$. Following the epistemic interpretation, $\underline{P}(A)$ can be interpreted as a lower bound of some unknown probability $P_0(A)$. The credal set associated with the lower probability can then be determined using Eq. (1):

$$\mathcal{M}(\underline{P}) = \{P \text{ probability measure} \mid P(A) \geq \underline{P}(A) \forall A \in \mathcal{K}\}.$$

The conjugate of a lower probability \underline{P} , denoted by \overline{P} , is called *upper probability* and it is given by $\overline{P}(A) = 1 - \underline{P}(A^c)$ for every A such that $A^c \in \mathcal{K}$. Also, when \underline{P} is coherent, \overline{P} can be computed by $\overline{P}(A) = \max_{P \in \mathcal{M}(\underline{P})} P(A)$ for every A such that $A^c \in \mathcal{K}$.

From now on, we will assume that the lower previsions/probabilities are coherent, and that they are defined on $\mathcal{L}(\mathcal{X})$ or $\mathcal{P}(\mathcal{X})$, respectively. Note that if \underline{P} is a coherent lower prevision (defined on $\mathcal{L}(\mathcal{X})$), we can define its restriction to events: $\underline{P}'(A) = \underline{P}(I_A)$ for every $A \in \mathcal{P}(\mathcal{X})$, and \underline{P}' is a coherent lower probability.

On the other hand, there may be different coherent lower previsions on $\mathcal{L}(\mathcal{X})$ with the same coherent lower probability as their restriction to events, and in this sense coherent lower previsions are a more informative model than coherent lower probabilities; see [36, Section 2.7.3] for more details.

Any lower probability, coherent or not, defined on $\mathcal{P}(\mathcal{X})$ can be equivalently represented in terms of a function $m : \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ called its *Möbius inverse*, defined by:

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \underline{P}(B) \quad \forall A \subseteq \mathcal{X}. \quad (2)$$

This function allows to retrieve the initial lower probability by using the formula:

$$\underline{P}(A) = \sum_{B \subseteq A} m(B) \quad \forall A \subseteq \mathcal{X}. \quad (3)$$

2.3. 2- and completely monotone lower probabilities. One useful property that a coherent lower probability may satisfy is that of 2-monotonicity.

Definition 2. [5] *A lower probability $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ is called 2-monotone if it satisfies $\underline{P}(A \cup B) + \underline{P}(A \cap B) \geq \underline{P}(A) + \underline{P}(B)$ for every $A, B \subseteq \mathcal{X}$.*

The use of 2-monotone lower probabilities has some benefits with respect to coherent lower probabilities:

- For coherent lower probabilities, $\mathcal{M}(\underline{P})$ is a closed and convex set of probabilities, hence it is characterized by its extreme points. It is known that the maximal number of extreme points of $\mathcal{M}(\underline{P})$ is $n!$ [37]. However, there is no procedure for obtaining those extreme points. Instead, such a procedure does exist in the particular case

of 2-monotone lower probabilities \underline{P} : as shown by Shapley [32], the extreme points of $\mathcal{M}(\underline{P})$ are in correspondence with the permutations of $\{1, \dots, n\}$.

- We mentioned in the previous subsection that a coherent lower probability on $\mathcal{P}(\mathcal{X})$ may have more than one coherent extension to $\mathcal{L}(\mathcal{X})$. However, a 2-monotone lower probability \underline{P} on $\mathcal{P}(\mathcal{X})$ has a unique extension to $\mathcal{L}(\mathcal{X})$ that satisfies 2-monotonicity [8, 35], its Choquet integral [5]:

$$\underline{P}(f) = (C) \int f d\underline{P} = \inf f + \int_{\inf f}^{\sup f} \underline{P}(\{f > t\}) dt. \quad (4)$$

For additional discussion about 2-monotonicity, we refer to [8, 12].

The notion of 2-monotonicity can be extended to higher degrees:

Definition 3. [5] *A lower probability $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ is k -monotone if for every $p \leq k$, and for every $A_1, \dots, A_p \subseteq \mathcal{X}$ it holds that:*

$$\underline{P}(\cup_{i=1}^p A_i) \geq \sum_{\emptyset \neq I \subseteq \{1, \dots, p\}} (-1)^{|I|+1} \underline{P}(\cap_{i \in I} A_i).$$

If \underline{P} is k -monotone for every k , it is called completely monotone.

In Shafer's Evidence Theory [31], completely monotone lower probabilities are usually called *belief functions*, denoted by Bel , and their conjugate upper probabilities are usually called *plausibility functions* and denoted by Pl .

As lower probabilities, belief functions can also be equivalently expressed in terms of their Möbius inverse using Eq. (2). Interestingly, belief functions can be characterized as those lower probabilities whose Möbius inverse m is non-negative: we have $m(A) \in [0, 1]$ for every $A \subseteq \mathcal{X}$, $m(\emptyset) = 0$ and $\sum_{A \subseteq \mathcal{X}} m(A) = 1$. In the context of belief functions, the function m is usually called *basic probability assignment*, and an event A with strictly positive mass, $m(A) > 0$, is called *focal event*. The mass $m(A)$ may be interpreted as the amount of evidence supporting the occurrence of the event A .

2.4. Possibility measures. A particular case of plausibility functions are possibility measures.

Definition 4. [16, 39] *A possibility measure $\Pi : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ is a normalized and supremum preserving function: $\Pi(\mathcal{X}) = 1$ and $\Pi(\cup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i)$ for every family of events $A_i \subseteq \mathcal{X}, i \in I$.*

Since in this paper we are dealing with finite spaces, the above supremum becomes a maximum, whence there must be some $x \in \mathcal{X}$ such that $\Pi(\{x\}) = 1$. The conjugate function of a possibility measure is called *necessity measure* and is denoted by N . A necessity measure and its conjugate possibility measure are in particular belief and plausibility functions: they correspond to the case where the focal events are nested, meaning that for every two focal events E_1, E_2 , either $E_1 \subseteq E_2$ or $E_2 \subseteq E_1$.

Since a possibility measure is a particular case of coherent upper probability, its associated credal set is given by:

$$\mathcal{M}(\Pi) = \{P \text{ probability measure} \mid P(A) \leq \Pi(A) \quad \forall A \subseteq \mathcal{X}\}. \quad (5)$$

A possibility measure Π can be represented by means of a function $\pi : \mathcal{X} \rightarrow [0, 1]$ called *possibility distribution*, and defined by $\pi(x) = \Pi(\{x\}) \quad \forall x \in \mathcal{X}$. That is, π is the restriction of Π to singletons. Since Π is supremum-preserving (or maxitive in our finite framework), it holds that $\Pi(A) = \sup_{x \in A} \pi(x)$ for every $A \subseteq \mathcal{X}$.

2.5. Probability boxes. Probability boxes (p-boxes, for short) [19, 33] are useful models when there is imprecise information about a cumulative distribution function (cdf, for short) and we consider lower and upper bounds:

Definition 5. Let $\mathcal{X} = \{x_1, \dots, x_n\}$ be a totally ordered space, so that $x_1 \leq \dots \leq x_n$. A p-box $(\underline{F}, \overline{F})$ is a pair of cdfs $\underline{F}, \overline{F}$ satisfying $\underline{F} \leq \overline{F}$.

A p-box $(\underline{F}, \overline{F})$ defines a credal set by:

$$\mathcal{M}(\underline{F}, \overline{F}) = \{P \text{ probability measure} \mid \underline{F}(x) \leq F_P(x) \leq \overline{F}(x) \quad \forall x \in \mathbb{R}\}, \quad (6)$$

where F_P denotes the cdf associated with the probability P . The lower and upper envelopes of $\mathcal{M}(\underline{F}, \overline{F})$, that we shall denote $\underline{P}_{(\underline{F}, \overline{F})}, \overline{P}_{(\underline{F}, \overline{F})}$, are coherent lower and upper probabilities, and they satisfy:

$$\underline{P}_{(\underline{F}, \overline{F})}(\{x_1, \dots, x_i\}) = \underline{F}(x_i), \quad \overline{P}_{(\underline{F}, \overline{F})}(\{x_1, \dots, x_i\}) = \overline{F}(x_i) \quad \forall x_i \in \mathcal{X}. \quad (7)$$

These lower and upper probabilities are not only coherent, but also belief and plausibility functions [33]. Their focal events are ordered intervals, meaning that for every focal event E , if $\min E \leq x \leq \max E$, then $x \in E$, and if E_1, E_2 are two focal events, either $\min E_1 \leq \min E_2$ and $\max E_1 \leq \max E_2$ or $\min E_2 \leq \min E_1$ and $\max E_2 \leq \max E_1$. Conversely, if Bel is a belief function whose focal events are ordered intervals, and we define the p-box $(\underline{F}, \overline{F})$ by:

$$\underline{F}(x_i) = Bel(\{x_1, \dots, x_i\}), \quad \overline{F}(x_i) = Pl(\{x_1, \dots, x_i\}) \quad \forall x_i \in \mathcal{X},$$

then $Bel = \underline{P}_{(\underline{F}, \overline{F})}$.

The definition of p-box requires \mathcal{X} to be endowed with a total order. However, this is sometimes a rather strong condition. To be able to deal with the nice properties of p-boxes in non-ordered spaces, the following notion was introduced in [13].

Definition 6. Let \mathcal{X} be an arbitrary space. A generalized p-box $(\underline{F}, \overline{F})$ is a pair of comonotone¹ mappings such that there exists $x \in \mathcal{X}$ with $\underline{F}(x) = \overline{F}(x) = 1$ and \underline{F} is dominated by \overline{F} : $\underline{F} \leq \overline{F}$.

¹Two functions f, g are comonotone if for every $x, y \in \mathcal{X}$ it holds that $f(x) < f(y)$ implies $g(x) \leq g(y)$.

From [13], a generalized p-box $(\underline{F}, \overline{F})$ induces an order $\leq_{(\underline{F}, \overline{F})}$ in \mathcal{X} and a permutation σ of $\{1, \dots, n\}$ such that:

$$\begin{aligned} \underline{F}(x_{\sigma(1)}) &\leq_{(\underline{F}, \overline{F})} \dots \leq_{(\underline{F}, \overline{F})} \underline{F}(x_{\sigma(n)}) = 1, \\ \overline{F}(x_{\sigma(1)}) &\leq_{(\underline{F}, \overline{F})} \dots \leq_{(\underline{F}, \overline{F})} \overline{F}(x_{\sigma(n)}) = 1. \end{aligned}$$

We shall use this fact in Section 6.2 later on.

2.6. Aim of the paper. Because of the advantageous mathematical properties of 2-monotone lower probabilities over those that are merely coherent, following the initial steps in [3], in a recent paper [27] we investigated how to approximate a coherent lower probability $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ by a 2-monotone lower probability \underline{Q} that at the same time (a) introduces no new information, in the sense that the credal set associated with \underline{Q} includes all those probability measures in $\mathcal{M}(\underline{P})$, that were considered a suitable model for the available knowledge; and (b) is as close as possible to the original model.

Definition 7. Let \mathcal{C} denote a class of coherent lower probabilities. We say that $\underline{Q} \in \mathcal{C}$ is an outer approximation of \underline{P} in \mathcal{C} if $\underline{Q} \leq \underline{P}$. Moreover, $\underline{Q} \in \mathcal{C}$ is an undominated outer approximation of \underline{P} in \mathcal{C} if there is no $\underline{Q}' \in \mathcal{C}$ such that $\underline{Q} \preceq \underline{Q}' \leq \underline{P}$.

In terms of credal sets, \underline{Q} is an outer approximation of \underline{P} when $\mathcal{M}(\underline{P}) \subseteq \mathcal{M}(\underline{Q})$, and it is an undominated outer approximation in \mathcal{C} when there is no \underline{Q}' such that $\mathcal{M}(\underline{P}) \subseteq \mathcal{M}(\underline{Q}') \subsetneq \mathcal{M}(\underline{P})$.

In the remainder of this paper, we denote by $\mathcal{C}_2, \mathcal{C}_\infty, \mathcal{C}_\Pi, \mathcal{C}_{(\underline{F}, \overline{F})}$ and $\mathcal{C}_{(\underline{F}, \overline{F})}^*$ the families of 2-monotone lower probabilities, belief functions, possibility measures, p-boxes and generalized p-boxes, respectively, defined for a fixed possibility space \mathcal{X} .

The following lemma shall be useful later on.

Lemma 1 ([27]). Let \underline{P} be a coherent lower probability and denote by \underline{P}' an outer approximation in the class \mathcal{C} . If \underline{P}' is a solution to the problem:

$$\min_{\underline{P}' \in \mathcal{C}, \underline{P}' \leq \underline{P}} \sum_{E \subseteq \mathcal{X}} g(\underline{P}(E) - \underline{P}'(E)),$$

for some strictly increasing function $g : \mathbb{R} \rightarrow \mathbb{R}$, then \underline{P}' is an undominated outer approximation of \underline{P} in \mathcal{C} .

In the rest of the paper we aim to:

- investigate how to obtain undominated outer approximations of a coherent lower probability in \mathcal{C}_∞ , as well as the properties of the set of undominated outer approximations;
- compare the results with those in [27] where we studied how to obtain undominated outer approximations in \mathcal{C}_2 ;

- study how to outer approximate a coherent lower probability with some particular families of belief functions: possibility measures and p-boxes.

It may also be of interest to determine *inner* approximations in \mathcal{C}_∞ , which are those belief functions Bel satisfying $\underline{P} \leq Bel$. Although the linear programming approach we shall discuss below could easily be applied in that context, the use of inner approximations is in conflict with the goal of *not introducing new information*. Indeed, if \underline{Q} is an inner approximation of \underline{P} then we have the inclusion between the credal sets $\mathcal{M}(\underline{P}) \supseteq \mathcal{M}(\underline{Q})$, and the strict inclusion means that we are removing some elements from $\mathcal{M}(\underline{P})$, that according to our model are compatible with the available information.

In addition, the existence of an outer approximation is always guaranteed to hold both for belief functions and their subfamilies we shall consider later on in this paper, while the same cannot be said for inner approximations: it may not be possible for instance, to find a possibility measure that inner approximates a coherent upper probability. For all these reasons, in this paper we focus on outer approximations. For a deeper discussion about inner approximations we refer to [27, Section 7].

3. OUTER APPROXIMATIONS OF COHERENT LOWER PROBABILITIES BY MEANS OF 2-MONOTONE LOWER PROBABILITIES

In this section, we recall our results from [27] about the problem of outer approximating a coherent lower probability by means of a 2-monotone one.

In [27], we showed that the set of undominated outer approximations of a coherent lower probability in \mathcal{C}_2 may be too large, and as a consequence we need some criterion that allows us to choose some outer approximations over others. One possibility is to focus on those outer approximations that minimize the distance with the initial model \underline{P} . If we pursue this avenue, then we need to choose an appropriate distance between the two models. In [27] we considered the distance proposed by Baroni and Vicig in [2], given by²:

$$d(\underline{P}, \underline{Q}) := \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - \underline{Q}(E)). \quad (8)$$

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 $d(\underline{P}, \underline{Q})$ has the meaning of total imprecision added by the outer approximation \underline{Q} .

In our view, this distance makes more sense from the point of view of the interpretation than other alternatives, such as the quadratic distance we shall consider in Section 5.2, because it takes into account the differences on all subsets of the possibility space and treats all these differences in the same

²The distance by Baroni and Vicig could be normalized just dividing by the maximum value of $d(\underline{P}, \underline{Q})$. In this case, if $\underline{P} \geq \underline{Q}$, the maximum distance is attained taking $\underline{Q}(A) = 0$ for any $A \neq \mathcal{X}$ and $\underline{Q}(\mathcal{X}) = 1$, usually called *vacuous* lower probability, and \underline{P} a precise probability measure on \mathcal{X} . Thus, we obtain a maximum distance of $2^{n-1} - 1$.

manner (instead of giving less weight to small differences, as is implicitly done with the quadratic distance). In addition, it shall allow us to use the machinery of linear programming, for which the set of solutions is a bounded polyhedral set with a finite number of extreme points. See [27, Section 5.1] for further discussion.

We obtain thus the following minimization problem, where the 2-monotone lower probability \underline{Q} is expressed in terms of its Möbius inverse $m_{\underline{Q}}$ by means of Eq. (3):

$$\min d(\underline{P}, \underline{Q}) = \sum_{E \subseteq \mathcal{X}} \left(\underline{P}(E) - \sum_{B \subseteq E} m_{\underline{Q}}(B) \right) \quad (\text{LP-2monot})$$

subject to:

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

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

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

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

As discussed in [27], according to [4], (LP-2monot.1)÷(LP-2monot.3) assure that $m_{\underline{Q}}$ defines a 2-monotone lower probability by means of Eq. (3). Our next proposition summarizes some of the main results from [27]:

Proposition 2. *Let \underline{P} be a coherent lower probability. The following statements hold:*

- (1) *The optimal solutions of (LP-2monot) subject to the restrictions (LP-2monot.1)÷(LP-2monot.4) are undominated outer approximations of \underline{P} in \mathcal{C}_2 .*
- (2) *The optimal solutions of (LP-2monot) subject to the restrictions (LP-2monot.1)÷(LP-2monot.4) and also to:*

$$\sum_{B \subseteq A} m_{\underline{Q}}(B) = \underline{P}(A), \quad (\text{LP-2monot.A})$$

for a fixed event A , are undominated outer approximations of \underline{P} in \mathcal{C}_2 satisfying $\underline{Q}(A) = \underline{P}(A)$.

- (3) *If \underline{Q} is an undominated outer approximation in \mathcal{C}_2 , and \overline{Q} is its conjugate upper probability, then $\underline{Q}(\{x\}) = \underline{P}(\{x\})$ and $\overline{Q}(\{x\}) = \overline{P}(\{x\})$ for every $x \in \mathcal{X}$.*
- (4) *If $\{\underline{Q}_i\}_{i \in I}$ denotes the set of undominated outer approximations in \mathcal{C}_2 , it holds that $\underline{P}(A) = \max_{i \in I} \underline{Q}_i(A)$ for every $A \subseteq \mathcal{X}$.*

The first and second items mean that it is possible to find undominated outer approximations in \mathcal{C}_2 just by solving linear programming problems.

The third item shows that any undominated outer approximation in \mathcal{C}_2 and its conjugate upper probability preserve the order determined by \underline{P} and \overline{P} on \mathcal{X} . Finally, the fourth item says that the undominated outer approximations in \mathcal{C}_2 allow us to retrieve the initial lower probability. Also, in [27, Example 1] it is shown that the set of undominated outer approximations of \underline{P} in \mathcal{C}_2 may be infinite. When that is the case, the problem of characterizing all the undominated outer approximations arises.

Besides the linear programming approach, in [27] we also studied some other approaches to obtain undominated outer approximation in \mathcal{C}_2 : we considered for instance some extensions of the total variation distance to the imprecise case, and showed that none of them guarantees obtaining an undominated outer approximation; in addition, we also considered the quadratic distance, that, although determining a unique undominated outer approximation, is not so well justified as the distance by Baroni and Vicig we have considered, in our view.

4. OUTER APPROXIMATIONS OF COHERENT LOWER PROBABILITIES BY MEANS OF BELIEF FUNCTIONS

In this section, we consider the problem of outer approximating a coherent lower probability by means of a belief function. Although belief functions are particular cases of 2-monotone lower probabilities, and as a consequence their use as outer approximating models will entail a loss of information with respect to the ones we can obtain in \mathcal{C}_2 , they also have advantages over them: the most important of these is that they have a clearer interpretation from the point of view of evidential theory, because in the case of belief functions, the Möbius inverse $m(A)$ can be interpreted as a measure of the evidence supporting event A as the outcome of the experiment. This is no longer possible for 2-monotone lower probabilities, because in their case the Möbius inverse need not be non-negative.

In [27], we considered already one instance of belief functions: the so-called *ϵ -contamination models* [36], or linear-vacuous mixtures, given by

$$\underline{P}_\epsilon(A) = \begin{cases} (1 - \epsilon)P_0(A) & \text{if } A \neq \mathcal{X}, \\ 1 & \text{if } A = \mathcal{X}, \end{cases}$$

for a given probability measure P_0 and $\epsilon \in (0, 1)$. The lower probability \underline{P}_ϵ is completely monotone, because it is a convex combination of two completely monotone models: the probability measure P_0 and the vacuous lower probability. Its focal events are the singletons and the total space, with respective masses:

$$m(\{x_i\}) = (1 - \epsilon)P_0(\{x_i\}) \quad \forall i = 1, \dots, n \quad \text{and} \quad m(\mathcal{X}) = \epsilon.$$

The credal set of the ϵ -contamination model is given by those probability measures that are mixtures of P_0 with some probability Q , with respective weights $1 - \epsilon$ and ϵ . Thus, the ϵ -contamination corresponds to a particular

family of distortion models, and were also considered in outer approximation problems by Quaeghebeur in [30]. In this section we analyze the outer approximations in the whole class \mathcal{C}_∞ .

As shown in [27, Example 1], some of the solutions of the linear programming problem (LP-2monot) subject to (LP-2monot.1)÷(LP-2monot.4) may be in particular completely monotone lower probabilities, and therefore they are also undominated outer approximations in the smaller class \mathcal{C}_∞ . This is of course not always the case: if \underline{P} is 2-monotone but not completely monotone, then it is the only undominated outer approximation of itself in \mathcal{C}_2 . Another example can be seen in Example 1 later on.

4.1. Outer approximations in \mathcal{C}_∞ by means of linear programming.

Our proposal is to consider those outer approximations that minimize the distance in Eq. (8) between the initial lower probability \underline{P} and the belief function: $d(\underline{P}, Bel) = \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - Bel(E))$. This expression can be equivalently written using the basic probability assignment m :

$$d(\underline{P}, Bel) = \sum_{E \subseteq \mathcal{X}} \left(\underline{P}(E) - \sum_{B \subseteq E} m(B) \right).$$

Thus, our goal is to determine the basic probability assignment minimizing the previous expression; this will induce a belief function using Eq. (3).

Proposition 3. *Let \underline{P} be a coherent lower probability. Consider the following minimization problem:*

$$\min d(\underline{P}, Bel) = \sum_{E \subseteq \mathcal{X}} \left(\underline{P}(E) - \sum_{B \subseteq E} m(B) \right) \quad (\text{LP-bel})$$

subject to:

$$\sum_{B \subseteq \mathcal{X}} m(B) = 1, \quad m(B) \geq 0 \quad \forall B \subseteq \mathcal{X}. \quad (\text{LP-bel.1})$$

$$\sum_{B \subseteq E} m(B) \leq \underline{P}(E) \quad \forall E \subseteq \mathcal{X}. \quad (\text{LP-bel.2})$$

Then, the following statements hold:

- (1) *The feasible region of the linear programming problem is non-empty.*
- (2) *Any optimal solution to the linear programming problem is an undominated outer approximation of \underline{P} in \mathcal{C}_∞ .*
- (3) *Consider the problem (LP-bel) subject to (LP-bel.1)÷(LP-bel.2) and also to:*

$$\sum_{B \subseteq A} m(B) = \underline{P}(A) \quad (\text{LP-bel.A})$$

for a fixed event $A \subsetneq \mathcal{X}$. Then the feasible region is not empty and the optimal solutions are also undominated outer approximations of \underline{P} in \mathcal{C}_∞ satisfying $Bel(A) = \underline{P}(A)$.

Proof. We shall establish the last statement, and deduce the other two as particular cases.

- (3) Consider the problem (LP-bel) subject to (LP-bel.1)÷(LP-bel.2) and (LP-bel.A) for a fixed event A . The feasible region of this linear programming problem is non-empty (consider for instance the simple support function, given by the basic probability assignment $m(A) = \underline{P}(A)$, $m(\mathcal{X}) = 1 - \underline{P}(A)$ and 0 elsewhere), it forms a bounded polyhedral set, and since the objective function is continuous in the variables $m(B)$, we can apply Weierstrass' theorem to deduce that there exists an optimal solution. Let Bel be one such solution, and let m denote its basic probability assignment. Since $\underline{P}(\emptyset) = 0$, conditions (LP-bel.1) and (LP-bel.2) imply that $m(\emptyset) = 0$. This, together with condition (LP-bel.1) ensures that m is a basic probability assignment, so it defines a belief function Bel using Eq. (3). Also, (LP-bel.2) implies that Bel is an outer approximation of \underline{P} . Condition (LP-bel.A) guarantees that $Bel(A) = \underline{P}(A)$. The fact that Bel is undominated follows from Lemma 1, just considering the identity function $g(x) = x$ for every $x \in \mathbb{R}$.

(1),(2) These correspond to the particular case of $A = \emptyset$ in item (3). \square

This means that, as we did in the case of 2-monotone outer approximations, we can again obtain undominated outer approximations in \mathcal{C}_∞ by solving a linear programming problem. Also, if we want to obtain an undominated outer approximation satisfying $Bel(A) = \underline{P}(A)$ for a fixed A , we just need to impose the additional constraint (LP-bel.A). This could be useful in some situations where all the optimal solutions Bel of (LP-bel) satisfy $Bel(A) < \underline{P}(A)$ for a fixed event A , as next example shows.

Example 1. *Let us consider a four-element space \mathcal{X} and the coherent lower probability \underline{P} given in Table 1. To see that it is coherent, note that it is the lower envelope of the probability measures with mass functions $(0.1, 0, 0.4, 0.5)$, $(0.4, 0.1, 0.2, 0.3)$ and $(0.3, 0.3, 0, 0.4)$. Solving the linear programming problem (LP-bel) with constraints (LP-bel.1)÷(LP-bel.2), we obtain the optimal solutions Bel_0 and Bel_1 as well as their convex combinations $Bel_\alpha = \alpha Bel_0 + (1 - \alpha) Bel_1$ for $\alpha \in [0, 1]$.*

However, any of these convex combinations satisfies

$$Bel_\alpha(\{x_3, x_4\}) = 0.3 < 0.4 = \underline{P}(\{x_3, x_4\}).$$

If we want to obtain an outer approximation in \mathcal{C}_∞ whose value for the event $\{x_3, x_4\}$ equals $\underline{P}(\{x_3, x_4\})$, we can solve the same linear programming problem with the additional constraint (LP-bel.A) with $A = \{x_3, x_4\}$. In that case, we obtain infinite solutions, such as Bel_2 . Of course, since Bel_2 is not an optimal solution to the former linear programming problem (without the constraint (LP-bel.A)), we observe that:

$$d(\underline{P}, Bel_0) = d(\underline{P}, Bel_1) = 0.3 < 0.5 = d(\underline{P}, Bel_2).$$

| A | $\underline{P}(A)$ | \underline{Q} | Bel_0 | Bel_1 | Bel_2 |
|---------------------|--------------------|-----------------|---------|---------|---------|
| $\{x_1\}$ | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| $\{x_2\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_3\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_4\}$ | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 |
| $\{x_1, x_2\}$ | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| $\{x_1, x_3\}$ | 0.3 | 0.3 | 0.2 | 0.3 | 0.1 |
| $\{x_1, x_4\}$ | 0.6 | 0.5 | 0.6 | 0.5 | 0.6 |
| $\{x_2, x_3\}$ | 0.3 | 0.2 | 0.3 | 0.2 | 0.2 |
| $\{x_2, x_4\}$ | 0.4 | 0.4 | 0.3 | 0.4 | 0.3 |
| $\{x_3, x_4\}$ | 0.4 | 0.4 | 0.3 | 0.3 | 0.4 |
| $\{x_1, x_2, x_3\}$ | 0.5 | 0.5 | 0.5 | 0.5 | 0.4 |
| $\{x_1, x_2, x_4\}$ | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| $\{x_1, x_3, x_4\}$ | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 |
| $\{x_2, x_3, x_4\}$ | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| \mathcal{X} | 1 | 1 | 1 | 1 | 1 |

TABLE 1. Coherent lower probability of Example 1 and some of its outer approximations in \mathcal{C}_2 and \mathcal{C}_∞ .

Table 1 also shows the values of the belief functions Bel_0 , Bel_1 and Bel_2 , as well as an undominated 2-monotone lower probability \underline{Q} that outer approximates \underline{P} . It holds that Bel_1 is dominated by \underline{Q} , showing that the undominated outer approximations in \mathcal{C}_∞ may be dominated in the broader class \mathcal{C}_2 . \blacklozenge

In particular, this example shows that there may be infinitely many undominated outer approximations in \mathcal{C}_∞ .

We can easily see that we can retrieve the original lower probability from these undominated outer approximations in \mathcal{C}_∞ .

Proposition 4. *Let \underline{P} be a coherent lower probability, and let $\{Bel_i\}_{i \in I}$ be the set of undominated outer approximations in \mathcal{C}_∞ that are optimal solutions of (LP-bel) subject to either (LP-bel.1) \div (LP-bel.2) or (LP-bel.1) \div (LP-bel.2) and (LP-bel.A) for any fixed event $A \subseteq \mathcal{X}$. Then*

$$\underline{P}(E) = \max_{i \in I} Bel_i(E) \quad \forall E \subseteq \mathcal{X}.$$

Equivalently, if $\mathcal{M}(\underline{P})$ and $\mathcal{M}(Bel_i)$ are the credal sets of \underline{P} and Bel_i , respectively, it holds that $\mathcal{M}(\underline{P}) = \bigcap_{i \in I} \mathcal{M}(Bel_i)$.

Proof. On the one hand, since any Bel_i is an outer approximation of \underline{P} , it holds that $\underline{P} \geq Bel_i$ for every $i \in I$, and therefore $\underline{P} \geq \max_{i \in I} Bel_i$. On the other hand, take $A \subseteq \mathcal{X}$. Using Proposition 3(3) with this event A , there is an undominated belief function Bel_i such that $Bel_i(A) = \underline{P}(A)$. Therefore, for every A there is an undominated outer approximation Bel_i such that $Bel_i(A) = \underline{P}(A)$, and we conclude that $\underline{P} = \max_{i \in I} Bel_i$. \square

4.2. Properties of the undominated outer approximations in \mathcal{C}_∞ . Next, we investigate the properties of the set of undominated outer approximations in \mathcal{C}_∞ , that we shall denote from now on by $\tilde{\mathcal{C}}_\infty$.

From Proposition 3 we know that solving the linear programming problem (LP-bel) we obtain undominated belief functions. One may wonder whether all undominated outer approximations in \mathcal{C}_∞ are solutions of (LP-bel) subject to either (LP-bel.1)÷(LP-bel.2) or (LP-bel.1)÷(LP-bel.2) and (LP-bel.A) for a fixed event $A \subsetneq \mathcal{X}$. However, as next example shows, there are undominated belief functions that are not optimal solutions of any of these linear programming problems.

Example 2. Consider a four element space, the lower probability \underline{P} and the belief functions Bel_1, Bel_2 defined as follows:

| A | $\underline{P}(A)$ | m_1 | Bel_1 | m_2 | Bel_2 |
|---------------------|--------------------|-------|---------|-------|---------|
| $\{x_1\}$ | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| $\{x_2\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_3\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_4\}$ | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| $\{x_1, x_2\}$ | 0.4 | 0.3 | 0.4 | 0.3 | 0.4 |
| $\{x_1, x_3\}$ | 0.4 | 0.1 | 0.2 | 0.1 | 0.2 |
| $\{x_1, x_4\}$ | 0.4 | 0.1 | 0.3 | 0 | 0.2 |
| $\{x_2, x_3\}$ | 0.2 | 0 | 0 | 0 | 0 |
| $\{x_2, x_4\}$ | 0.4 | 0 | 0.1 | 0.1 | 0.2 |
| $\{x_3, x_4\}$ | 0.4 | 0.2 | 0.3 | 0.3 | 0.4 |
| $\{x_1, x_2, x_3\}$ | 0.5 | 0 | 0.5 | 0 | 0.5 |
| $\{x_1, x_2, x_4\}$ | 0.6 | 0 | 0.6 | 0 | 0.6 |
| $\{x_1, x_3, x_4\}$ | 0.6 | 0 | 0.6 | 0 | 0.6 |
| $\{x_2, x_3, x_4\}$ | 0.5 | 0.1 | 0.4 | 0 | 0.5 |
| \mathcal{X} | 1 | 0 | 1 | 0 | 1 |

On the one hand, \underline{P} is a coherent lower probability because it is the lower envelope of the probability measures with mass functions:

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

If we now compute the distance in Eq. (8) between \underline{P} and these belief functions, we obtain that $d(\underline{P}, Bel_1) = 1$, while $d(\underline{P}, Bel_2) = 0.8$. We deduce that, while $Bel_2 \not\leq Bel_1$, $d(Bel_2, \underline{P}) < d(Bel_1, \underline{P})$; moreover, Bel_2 agrees with \underline{P} in all events E such that $Bel_1(E) = \underline{P}(E)$. Thus, Bel_1 is not an optimal solution to any of the linear programming problems.

Let us show that Bel_1 is undominated. Any Bel such that $Bel_1 \leq Bel \leq \underline{P}$ should satisfy:

| A | m | Bel |
|---------------------|-----|-----------------------|
| $\{x_1\}$ | 0.1 | 0.1 |
| $\{x_2\}$ | 0 | 0 |
| $\{x_3\}$ | 0 | 0 |
| $\{x_4\}$ | 0.1 | 0.1 |
| $\{x_1, x_2\}$ | 0.3 | 0.4 |
| $\{x_1, x_3\}$ | a | $0.1 + a$ |
| $\{x_1, x_4\}$ | b | $0.2 + b$ |
| $\{x_2, x_3\}$ | c | c |
| $\{x_2, x_4\}$ | d | $0.1 + d$ |
| $\{x_3, x_4\}$ | e | $0.1 + e$ |
| $\{x_1, x_2, x_3\}$ | f | 0.5 |
| $\{x_1, x_2, x_4\}$ | g | 0.6 |
| $\{x_1, x_3, x_4\}$ | h | 0.6 |
| $\{x_2, x_3, x_4\}$ | i | $0.1 + c + d + e + i$ |
| \mathcal{X} | j | 1 |

Since $Bel_1(E) \leq Bel(E)$, taking $E = \{x_1, x_3\}$, $\{x_1, x_4\}$, $\{x_2, x_3\}$, $\{x_2, x_4\}$ or $\{x_3, x_4\}$, we obtain respectively that:

$$0.1 + a \geq 0.2 \Rightarrow a \geq 0.1. \quad (9)$$

$$0.2 + b \geq 0.3 \Rightarrow b \geq 0.1. \quad (10)$$

$$c \geq 0.$$

$$d \geq 0.$$

$$e \geq 0.2. \quad (11)$$

Adding these inequalities,

$$a + b + c + d + e \geq 0.4. \quad (12)$$

If we now apply the constraint $Bel_1(E) \leq Bel(E)$ for $E = \{x_1, x_2, x_3\}$, $\{x_1, x_2, x_4\}$ or $\{x_1, x_3, x_4\}$, we deduce that:

$$a + c + f = 0.1. \quad (13)$$

$$b + d + g = 0.1. \quad (14)$$

$$a + b + e + h = 0.4. \quad (15)$$

From Eqs. (9) and (13), we obtain that $a = 0.1$, $c = f = 0$; from Eqs. (10) and (14), we deduce that $b = 0.1$, $d = g = 0$. This, together with Eqs. (11), (12) and (15) implies that $e = 0.2$, $h = 0$. Also, since $Bel_1(\{x_2, x_3, x_4\}) \leq Bel(\{x_2, x_3, x_4\})$:

$$0.3 + i \geq 0.4 \Rightarrow i \geq 0.1.$$

Finally, since the sum of all the masses is 1, $i + j = 0.1$, so $i = 0.1$ and $j = 0$. We get that $m = m_1$, and so $Bel = Bel_1$, and as a consequence we conclude that Bel_1 is undominated. \blacklozenge

We conclude from this example that there are undominated outer approximations that are not optimal solutions of any of the previous linear programming problems. This is not surprising, since this feature was already observed for the outer approximations in \mathcal{C}_2 [27, Example 3]. In Section 5 we shall discuss some alternatives to the linear programming approach.

Next we show that, perhaps surprisingly, the set $\widetilde{\mathcal{C}}_\infty$ of undominated outer approximations in \mathcal{C}_∞ is not convex:

Example 3. Let $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ be a four-element space, and consider the lower probability \underline{P} given in Table 2.

| | \underline{P} | Bel | m | Bel_1 | m_1 | Bel_2 | m_2 | \widetilde{Bel} | \widetilde{m} |
|---------------------|-----------------|-------|-----|---------|-------|---------|-------|-------------------|-----------------|
| $\{x_1\}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\{x_2\}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\{x_3\}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\{x_4\}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\{x_1, x_2\}$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/12 | 1/12 | 1/8 | 1/8 |
| $\{x_1, x_3\}$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |
| $\{x_1, x_4\}$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |
| $\{x_2, x_3\}$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |
| $\{x_2, x_4\}$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |
| $\{x_3, x_4\}$ | 1/6 | 1/6 | 1/6 | 1/12 | 1/12 | 1/6 | 1/6 | 1/8 | 1/8 |
| $\{x_1, x_2, x_3\}$ | 7/12 | 1/2 | 0 | 7/12 | 1/12 | 5/12 | 0 | 1/2 | 1/24 |
| $\{x_1, x_2, x_4\}$ | 1/2 | 1/2 | 0 | 1/2 | 0 | 5/12 | 0 | 11/24 | 0 |
| $\{x_1, x_3, x_4\}$ | 1/2 | 1/2 | 0 | 5/12 | 0 | 1/2 | 0 | 11/24 | 0 |
| $\{x_2, x_3, x_4\}$ | 7/12 | 1/2 | 0 | 5/12 | 0 | 7/12 | 1/12 | 1/2 | 1/24 |
| \mathcal{X} | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |

TABLE 2. Lower probability \underline{P} from Example 3 and some of its outer approximations in \mathcal{C}_∞ .

This lower probability is coherent because it is the lower envelope of the probability measures associated with the mass functions:

$$(0, 1/3, 1/3, 1/3), \quad (1/3, 0, 1/3, 1/3), \quad (1/3, 1/3, 0, 1/3), \quad (1/3, 1/3, 1/3, 0).$$

$$(1/12, 1/12, 5/12, 5/12), \quad (5/12, 5/12, 1/12, 1/12), \quad (1/12, 1/2, 1/12, 1/3).$$

$$(1/12, 1/2, 1/3, 1/12), \quad (5/12, 1/12, 1/12, 5/12), \quad (1/3, 1/12, 1/2, 1/12).$$

In that table we also show some belief functions as well as their respective Möbius inverses. It is immediate to see that all of them are outer approximations of the coherent lower probability \underline{P} . In addition, they satisfy the following properties:

- $Bel \in \widetilde{\mathcal{C}}_\infty$: if there exists a belief function Bel' such that $Bel \leq Bel' \leq \underline{P}$, then Bel and Bel' should coincide in all the events of cardinality one and two, hence also m and m' coincide for those

events. Since $\sum_{|A|=2} m'(A) = \sum_{|A|=2} m(A) = 1$, it holds that $m = m'$, so Bel and Bel' coincide.

- $Bel_1 \in \widetilde{\mathcal{C}}_\infty$: if there exists a belief function Bel' such that $Bel_1 \leq Bel' \leq \underline{P}$, then Bel' and its Möbius inverse should satisfy:

$$\begin{aligned} Bel'(\{x_1, x_2\}) &= \frac{1}{6} \Rightarrow m'(\{x_1, x_2\}) = \frac{1}{6}. \\ Bel'(\{x_1, x_3\}) &= \frac{1}{6} \Rightarrow m'(\{x_1, x_3\}) = \frac{1}{6}. \\ Bel'(\{x_1, x_4\}) &= \frac{1}{6} \Rightarrow m'(\{x_1, x_4\}) = \frac{1}{6}. \\ Bel'(\{x_2, x_3\}) &= \frac{1}{6} \Rightarrow m'(\{x_2, x_3\}) = \frac{1}{6}. \\ Bel'(\{x_2, x_4\}) &= \frac{1}{6} \Rightarrow m'(\{x_2, x_4\}) = \frac{1}{6}. \\ Bel'(\{x_3, x_4\}) &\geq \frac{1}{12} \Rightarrow m'(\{x_3, x_4\}) \geq \frac{1}{12}. \\ Bel'(\{x_1, x_2, x_3\}) &= \frac{7}{12} \Rightarrow m'(\{x_1, x_2, x_3\}) = \frac{1}{12}. \end{aligned}$$

Since the sum of the masses of these events is greater than or equal to 1, it must be $m' = m_1$ and $Bel' = Bel_1$. Thus, Bel_1 is undominated.

- With an analogous reasoning, we see that $Bel_2 \in \widetilde{\mathcal{C}}_\infty$.
- The belief function \widetilde{Bel} can be expressed as a convex combination of Bel_1 and Bel_2 : $\widetilde{Bel} = \frac{1}{2}Bel_1 + \frac{1}{2}Bel_2$. However, \widetilde{Bel} is dominated by the belief function Bel .

Thus, the set $\widetilde{\mathcal{C}}_\infty$ of undominated outer approximations in the class of belief functions is not convex. \blacklozenge

Note that all the undominated belief functions in this example can be obtained as solutions of a linear programming problem of the type considered in Proposition 3: Bel is the unique optimal solution of (LP-bel) subject to (LP-bel.1)÷(LP-bel.2); Bel_1 is an optimal solution of the linear programming problem (LP-bel) subject to (LP-bel.1)÷(LP-bel.2) and (LP-bel.A) with $A = \{x_1, x_2, x_3\}$; and Bel_2 is an optimal solution of the linear programming problem (LP-bel) subject to (LP-bel.1)÷(LP-bel.2) and (LP-bel.A) with $A = \{x_2, x_3, x_4\}$. Hence, it alternatively follows from Proposition 3 that they are all undominated outer approximations in \mathcal{C}_∞ .

We observe then that the structure of the set of undominated outer approximations in \mathcal{C}_∞ is not simple. On the one hand, Example 1 shows that this set may be infinite, as it includes the (often) infinite set of optimal solutions of the linear programming problems considered in Proposition 3. However, Example 2 shows that our linear programming approach does not allow to retrieve all the undominated outer approximations, and Example 3 implies that the set of undominated solutions is not convex in general. As a consequence, we cannot summarize this infinite set by means of a family of

extreme points. All these examples are established on a space of cardinality four. As we shall see in Section 4.3, in the case of ternary spaces there is a simple procedure that determines all the undominated outer approximations.

One main difference with respect to optimal outer approximations in \mathcal{C}_2 is that in the case of undominated outer approximations $Bel \in \mathcal{C}_\infty$, Bel and \underline{P} , as well as their conjugate Pl and \overline{P} , may not coincide on some singleton. This means that the outer approximations in \mathcal{C}_∞ may not preserve the orders that \underline{P} and \overline{P} determine on \mathcal{X} .

Example 4. Consider the belief function Bel with Möbius inverse m and the dominating coherent lower probability \underline{P} given by:

| A | \underline{P} | m | Bel |
|---------------------|-----------------|-----|-------|
| $\{x_1\}$ | 0.1 | 0 | 0 |
| $\{x_2\}$ | 0 | 0 | 0 |
| $\{x_3\}$ | 0 | 0 | 0 |
| $\{x_4\}$ | 0.2 | 0.2 | 0.2 |
| $\{x_1, x_2\}$ | 0.1 | 0.1 | 0.1 |
| $\{x_1, x_3\}$ | 0.1 | 0.1 | 0.1 |
| $\{x_1, x_4\}$ | 0.3 | 0.1 | 0.3 |
| $\{x_2, x_3\}$ | 0.2 | 0.2 | 0.2 |
| $\{x_2, x_4\}$ | 0.3 | 0.1 | 0.3 |
| $\{x_3, x_4\}$ | 0.3 | 0.1 | 0.3 |
| $\{x_1, x_2, x_3\}$ | 0.5 | 0.1 | 0.5 |
| $\{x_1, x_2, x_4\}$ | 0.5 | 0 | 0.5 |
| $\{x_1, x_3, x_4\}$ | 0.5 | 0 | 0.5 |
| $\{x_2, x_3, x_4\}$ | 0.6 | 0 | 0.6 |
| \mathcal{X} | 1 | 0 | 1 |

\underline{P} is coherent because it is the lower envelope of the probability measures associated with the mass functions

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

Let us now prove that Bel is undominated. Since $Bel(E) = \underline{P}(E)$ for any event E except for $E = \{x_1\}$, it may only be dominated if we increase the

mass of $\{x_1\}$ and leave the belief of all the other events equal. This gives:

| A | m | Bel |
|---------------------|-----------|-------|
| $\{x_1\}$ | a | a |
| $\{x_2\}$ | 0 | 0 |
| $\{x_3\}$ | 0 | 0 |
| $\{x_4\}$ | 0.2 | 0.2 |
| $\{x_1, x_2\}$ | $0.1 - a$ | 0.1 |
| $\{x_1, x_3\}$ | $0.1 - a$ | 0.1 |
| $\{x_1, x_4\}$ | $0.1 - a$ | 0.3 |
| $\{x_2, x_3\}$ | 0.2 | 0.2 |
| $\{x_2, x_4\}$ | 0.1 | 0.3 |
| $\{x_3, x_4\}$ | 0.1 | 0.3 |
| $\{x_1, x_2, x_3\}$ | $0.1 + a$ | 0.5 |
| $\{x_1, x_2, x_4\}$ | a | 0.5 |
| $\{x_1, x_3, x_4\}$ | a | 0.5 |
| $\{x_2, x_3, x_4\}$ | 0 | 0.6 |
| \mathcal{X} | 0 | 1 |

But in this case $\sum_{E \subseteq \mathcal{X}} m(E) = 1 + a$, so Bel is a belief function if and only if $a = 0$. Therefore, Bel is undominated but does not coincide with \underline{P} on all the singletons.

To see that \overline{P} and Pl do not coincide in general, consider the coherent lower probability \underline{P} in Example 2 and its undominated outer approximation Bel_1 . It satisfies that

$$Bel_1(\{x_2, x_3, x_4\}) = 0.4 < 0.5 = \underline{P}(\{x_2, x_3, x_4\}),$$

and therefore by conjugacy $Pl(\{x_1\}) = 0.6 > 0.5 = \overline{P}(\{x_1\})$. \blacklozenge

We conclude that, as we did in [27] for \mathcal{C}_2 , we can find undominated outer approximations of a coherent lower probability in \mathcal{C}_∞ by solving linear programming problems. In comparison with the results in [27], the only main property that is lost when we outer approximate in \mathcal{C}_∞ is that the order between the singletons given by \underline{P} and \overline{P} may not be preserved.

However, the use of belief functions has two advantages over 2-monotone lower probabilities: on the one hand, their interpretation is clearer, and they allow us to use a number of tools from evidence theory; on the other, the number of constraints in the linear programming problem is smaller: in the case of (LP-bel.1)÷(LP-bel.2) we have 2^{n+1} constraints: $2^n + 1$ in (LP-bel.1) and $2^n - 1$ in (LP-bel.2), while solving (LP-2monot) subject to (LP-2monot.1)÷(LP-2monot.4) involves $2^n + n + 2^{n-2} \binom{n}{2}$ constraints (see [27] for more details). Hence, the number of constraints decreases significantly.

Next we show that it is possible to characterize those belief functions that outer approximate the coherent lower probability \underline{P} and agree with it on the singletons:

Proposition 5. *Let $\underline{P} : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ be a coherent lower probability, and let $\alpha = \sum_{x \in \mathcal{X}} \underline{P}(\{x\})$. Let P_1 be the probability measure given by*

$$P_1(\{x\}) = \frac{\underline{P}(\{x\})}{\alpha} \quad \forall x \in \mathcal{X},$$

and let \underline{P}_2 be the lower probability given by

$$\underline{P}_2(A) = \frac{\underline{P}(A) - \sum_{x \in A} \underline{P}(\{x\})}{1 - \alpha} = \frac{\underline{P}(A) - \alpha P_1(A)}{1 - \alpha}. \quad (16)$$

The following statements hold:

- (1) $\underline{P} = \alpha P_1 + (1 - \alpha) \underline{P}_2$.
- (2) \underline{P}_2 is a coherent lower probability.
- (3) $\text{Bel} \leq \underline{P}$ is a belief function such that $\text{Bel}(\{x\}) = \underline{P}(\{x\}) \quad \forall x \in \mathcal{X}$ if and only if there exists a belief function $\text{Bel}_2 \leq \underline{P}_2$ such that $\text{Bel} = \alpha P_1 + (1 - \alpha) \text{Bel}_2$. Moreover, Bel is an undominated outer approximation of \underline{P} if and only if Bel_2 is an undominated outer approximation of \underline{P}_2 .
- (4) In the above correspondence, we obtain $d(\underline{P}, \text{Bel}) = \frac{1}{1 - \alpha} d(\underline{P}_2, \text{Bel}_2)$.

Proof. Before we proceed, note that we may assume that $\alpha < 1$; when $\alpha = 1$ then $\underline{P} = P_1$ and the only belief function that agrees with \underline{P} on singletons is \underline{P} itself.

- (1) Consider $A \subseteq \mathcal{X}$. Then

$$\alpha P_1(A) + (1 - \alpha) \underline{P}_2(A) = \sum_{x \in A} \underline{P}(\{x\}) + \left(\underline{P}(A) - \sum_{x \in A} \underline{P}(\{x\}) \right) = \underline{P}(A).$$

- (2) To see that \underline{P}_2 is coherent, we are going to prove first that

$$\mathcal{M}(\underline{P}_2) = \left\{ \frac{P - \alpha P_1}{1 - \alpha} : P \in \mathcal{M}(\underline{P}) \right\}. \quad (17)$$

We begin with the direct inclusion. Consider $P_2 \in \mathcal{M}(\underline{P}_2)$, and let $P := \alpha P_1 + (1 - \alpha) P_2$. Then $P \geq \alpha P_1 + (1 - \alpha) \underline{P}_2 = \underline{P}$, and it is a probability measure because it is a convex combination of probability measures. Thus, $P \in \mathcal{M}(\underline{P})$, and by construction $P_2 = \frac{P - \alpha P_1}{1 - \alpha}$.

To see the converse inclusion, take $P \in \mathcal{M}(\underline{P})$. Then

$$P_2 := \frac{P - \alpha P_1}{1 - \alpha} \geq \frac{\underline{P} - \alpha P_1}{1 - \alpha} = \underline{P}_2,$$

where last equality follows by Eq. (16). To see that it is a probability measure, note that immediately $P_2(\emptyset) = 0, P_2(\mathcal{X}) = 1$, that it is additive because it is a linear combination of additive functions, and that $P_2(A) \geq \underline{P}_2(A) \geq 0$ for every $A \subseteq \mathcal{X}$.

Let us prove now that \underline{P}_2 is the lower envelope of $\mathcal{M}(\underline{P}_2)$. Take $A \subseteq \mathcal{X}$. Then since \underline{P} is coherent there is some $P \in \mathcal{M}(\underline{P})$ such that

$P(A) = \underline{P}(A)$. Let $P_2 := \frac{P - \alpha P_1}{1 - \alpha}$. By Eq. (17), $P_2 \geq \underline{P}_2$. In addition,

$$P_2(A) = \frac{P(A) - \alpha P_1(A)}{1 - \alpha} = \frac{\underline{P}(A) - \alpha P_1(A)}{1 - \alpha} = \underline{P}_2(A).$$

- (3) Let Bel be a belief function with Möbius inverse m_B that outer approximates \underline{P} and satisfies $Bel(\{x\}) = \underline{P}(\{x\})$ for every $x \in \mathcal{X}$. If we define the belief function Bel_2 with Möbius inverse

$$m_{B_2}(A) = \frac{m_B(A)}{1 - \alpha} \quad \forall A \subseteq \mathcal{X} \text{ with } |A| \geq 2$$

and $m_{B_2}(A) = 0$ otherwise, then

$$Bel(B) = \sum_{x \in B} m_B(\{x\}) + \sum_{A \subseteq B, |A| \geq 2} m_B(A) = \alpha P_1(B) + (1 - \alpha) Bel_2(B)$$

for every $B \subseteq \mathcal{X}$. Thus, $Bel_2 = \frac{Bel - \alpha P_1}{1 - \alpha}$ satisfies $Bel_2 \leq \frac{P - \alpha P_1}{1 - \alpha} = \underline{P}_2$.

Conversely, if Bel_2 is a belief function that outer approximates \underline{P}_2 , then $Bel := \alpha P_1 + (1 - \alpha) Bel_2$ is a belief function because it is a convex combination of belief functions and $Bel \leq \alpha P_1 + (1 - \alpha) \underline{P}_2 = \underline{P}$. Moreover, $\forall x \in \mathcal{X}$ it is $\underline{P}_2(\{x\}) = 0$, whence $Bel_2(\{x\}) = 0$ (because we assumed $Bel_2 \leq \underline{P}_2$) and as a consequence $Bel(\{x\}) = \alpha P_1(\{x\}) = \underline{P}(\{x\})$.

A similar proof shows that Bel is undominated if and only if Bel_2 is undominated.

- (4) Finally, note that

$$\begin{aligned} d(\underline{P}_2, Bel_2) &= \sum_{E \subseteq \mathcal{X}} (\underline{P}_2(E) - Bel_2(E)) \\ &= \sum_{E \subseteq \mathcal{X}} \left(\frac{\underline{P}(E) - \alpha P_1(E)}{1 - \alpha} - \frac{Bel(E) - \alpha P_1(E)}{1 - \alpha} \right) \\ &= \frac{1}{1 - \alpha} \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - Bel(E)) = \frac{d(\underline{P}, Bel)}{1 - \alpha}. \end{aligned}$$

This completes the proof. \square

This procedure shows that if we want to consider the undominated outer approximations that agree with the initial model \underline{P} on the singletons, it suffices to define the coherent lower probability \underline{P}_2 as in Eq. (16), determine undominated outer approximations of \underline{P}_2 by linear programming, as in Proposition 3, and then transform these into outer approximations of \underline{P} , using Proposition 5(3).

4.3. Outer approximations in ternary spaces. We now investigate the particular case of a three-element space $\mathcal{X} = \{x_1, x_2, x_3\}$. Any coherent lower probability \underline{P} in a three-element space is 2-monotone. Also, if we

denote by m its associated Möbius inverse by Eq. (2), it follows from the non-negativity of \underline{P} that:

$$\underline{P}(\{x_i\}) = m(\{x_i\}) \geq 0,$$

and from the property of superadditivity [36], $\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$ for any $A \cap B = \emptyset$, it holds that:

$$\begin{aligned} m(\{x_i, x_j\}) + m(\{x_i\}) + m(\{x_j\}) &= \underline{P}(\{x_i, x_j\}) \\ &\geq \underline{P}(\{x_i\}) + \underline{P}(\{x_j\}) = m(\{x_i\}) + m(\{x_j\}), \end{aligned}$$

which implies $m(\{x_i, x_j\}) \geq 0$. We conclude that a coherent lower probability \underline{P} in a three-element space is not a belief function if and only if $m(\mathcal{X}) < 0$.

We start with two preliminary results that provide necessary conditions for Bel to be an undominated outer approximation.

Lemma 6. *Let \underline{P} be a coherent lower probability on $\mathcal{P}(\mathcal{X})$ for $\mathcal{X} = \{x_1, x_2, x_3\}$ with associated Möbius inverse m . Assume that $m(\mathcal{X}) < 0$ and that $Bel \in \mathcal{C}_\infty$ is an undominated outer approximation of \underline{P} with Möbius inverse m_B . Then $m_B(\{x_i\}) = m(\{x_i\})$ for every $i = 1, 2, 3$.*

Proof. Assume that $m_B(\{x_i\}) < m(\{x_i\})$ for some $i = 1, 2, 3$, and define Bel^* by:

$$\begin{aligned} Bel^*(\{x_i\}) &= \underline{P}(\{x_i\}), \quad i = 1, 2, 3, \\ Bel^*(\{x_i, x_j\}) &= Bel(\{x_i, x_j\}) + \delta_{ij}, \quad 1 \leq i < j \leq 3, \end{aligned}$$

where $\delta_{ij} = \max\{0, m(\{x_i\}) + m(\{x_j\}) - Bel(\{x_i, x_j\})\}$. Then the Möbius inverse m^* of Bel^* satisfies:

- $m^*(\{x_i\}) = m(\{x_i\}) \geq 0$, for $i = 1, 2, 3$.
- If $\delta_{ij} = 0$, it holds that:

$$\begin{aligned} m^*(\{x_i, x_j\}) &= Bel^*(\{x_i, x_j\}) - Bel^*(\{x_i\}) - Bel^*(\{x_j\}) \\ &= Bel(\{x_i, x_j\}) - m(\{x_i\}) - m(\{x_j\}) \geq 0. \end{aligned}$$

If $\delta_{ij} > 0$, it holds that:

$$\begin{aligned} m^*(\{x_i, x_j\}) &= Bel^*(\{x_i, x_j\}) - Bel^*(\{x_i\}) - Bel^*(\{x_j\}) \\ &= Bel(\{x_i, x_j\}) + m(\{x_i\}) + m(\{x_j\}) \\ &\quad - Bel(\{x_i, x_j\}) - m(\{x_i\}) - m(\{x_j\}) = 0. \end{aligned}$$

In both cases, $m^*(\{x_i, x_j\}) \geq 0$.

- Let us see that $m^*(\mathcal{X}) \geq 0$. We distinguish the following cases:

– Case 1: $\delta_{12}, \delta_{13}, \delta_{23} > 0$.

$$\begin{aligned}
 m^*(\mathcal{X}) &= 1 + m(\{x_1\}) + m(\{x_2\}) + m(\{x_3\}) \\
 &\quad - (Bel(\{x_1, x_2\}) + Bel(\{x_1, x_3\}) + Bel(\{x_2, x_3\})) \\
 &\quad - (m(\{x_1\}) + m(\{x_2\}) - Bel(\{x_1, x_2\}) + m(\{x_1\}) + m(\{x_3\}) \\
 &\quad - Bel(\{x_1, x_3\}) + m(\{x_2\}) + m(\{x_3\}) - Bel(\{x_2, x_3\})) \\
 &= 1 - (m(\{x_1\}) + m(\{x_2\}) + m(\{x_3\})) \geq 0,
 \end{aligned}$$

from the superadditivity of \underline{P} .

– Case 2: $\delta_{12}, \delta_{13} > 0$ and $\delta_{23} = 0$.

$$\begin{aligned}
 m^*(\mathcal{X}) &= 1 + m(\{x_1\}) + m(\{x_2\}) + m(\{x_3\}) \\
 &\quad - (Bel(\{x_1, x_2\}) + (Bel(\{x_1, x_3\}) + Bel(\{x_2, x_3\}))) \\
 &\quad - (m(\{x_1\}) + m(\{x_2\}) - Bel(\{x_1, x_2\}) + m(\{x_1\}) \\
 &\quad + m(\{x_3\}) - Bel(\{x_1, x_3\})) \\
 &= 1 - Bel(\{x_2, x_3\}) - m(\{x_1\}) \\
 &\geq 1 - \underline{P}(\{x_2, x_3\}) - \underline{P}(\{x_1\}) \geq 0,
 \end{aligned}$$

from the superadditivity of \underline{P} .

– Case 3: $\delta_{12} > 0$ and $\delta_{13} = \delta_{23} = 0$.

$$\begin{aligned}
 m^*(\mathcal{X}) &= 1 + m(\{x_1\}) + m(\{x_2\}) + m(\{x_3\}) \\
 &\quad - (Bel(\{x_1, x_2\}) + Bel(\{x_1, x_3\}) + Bel(\{x_2, x_3\})) \\
 &\quad - (m(\{x_1\}) + m(\{x_2\}) - Bel(\{x_1, x_2\})) \\
 &= 1 + m(\{x_3\}) - Bel(\{x_1, x_3\}) - Bel(\{x_2, x_3\}) \\
 &\geq Bel(\mathcal{X}) + Bel(\{x_3\}) - Bel(\{x_1, x_3\}) - Bel(\{x_2, x_3\}) \geq 0,
 \end{aligned}$$

because Bel is a belief function and, in particular, 2-monotone.

– Case 4: $\delta_{12} = \delta_{13} = \delta_{23} = 0$.

$$\begin{aligned}
 m^*(\mathcal{X}) &= 1 + m(\{x_1\}) + m(\{x_2\}) + m(\{x_3\}) \\
 &\quad - (Bel(\{x_1, x_2\}) + Bel(\{x_1, x_3\}) + Bel(\{x_2, x_3\})) \\
 &\geq 1 + m_B(\{x_1\}) + m_B(\{x_2\}) + m_B(\{x_3\}) \\
 &\quad - (Bel(\{x_1, x_2\}) + Bel(\{x_1, x_3\}) + Bel(\{x_2, x_3\})) \\
 &= m_B(\mathcal{X}) \geq 0.
 \end{aligned}$$

The remaining cases for $\delta_{12}, \delta_{13}, \delta_{23}$ follow by analogy. We conclude that Bel^* is a belief function. Also, it satisfies the following properties:

- Bel^* is an outer approximation of \underline{P} . Trivially, $Bel^*(\{x_i\}) = \underline{P}(\{x_i\})$ for $i = 1, 2, 3$. Also, if $\delta_{ij} = 0$:

$$Bel^*(\{x_i, x_j\}) = Bel(\{x_i, x_j\}) \leq \underline{P}(\{x_i, x_j\}),$$

because Bel is an outer approximation of \underline{P} . If $\delta_{ij} > 0$:

$$\begin{aligned} Bel^*({x_i, x_j}) &= Bel({x_i, x_j}) + m({x_i}) + m({x_j}) - Bel({x_i, x_j}) \\ &\leq m({x_i}) + m({x_j}) + m({x_i, x_j}) = \underline{P}({x_i, x_j}). \end{aligned}$$

- Finally, Bel^* dominates Bel . On the one hand,

$$Bel^*({x_i}) = \underline{P}({x_i}) \geq Bel({x_i}) \quad \forall i,$$

with strict inequality in some ${x_i}$. Also:

$$Bel^*({x_i, x_j}) = Bel({x_i, x_j}) + \delta_{ij} \geq Bel({x_i, x_j}) \quad \forall i \neq j.$$

We conclude that Bel cannot be an undominated outer approximation. \square

The next auxiliary result shows that the undominated outer approximations in \mathcal{C}_∞ must assign zero mass to the sure event \mathcal{X} .

Lemma 7. *Let $\mathcal{X} = \{x_1, x_2, x_3\}$ and let \underline{P} be a coherent lower probability on $\mathcal{P}(\mathcal{X})$ with associated Möbius inverse m . If $m(\mathcal{X}) < 0$, and $Bel \in \mathcal{C}_\infty$ is an undominated outer approximation of \underline{P} with Möbius inverse m_B , then $m_B(\mathcal{X}) = 0$.*

Proof. Assume that $m_B(\mathcal{X}) > 0$. Since we can assume from the previous lemma that \underline{P} and Bel coincide in singletons,

$$m_B(\{x_1\}) + m_B(\{x_2\}) + m_B(\{x_3\}) = m(\{x_1\}) + m(\{x_2\}) + m(\{x_3\}),$$

whence:

$$\begin{aligned} m_B(\{x_1, x_2\}) + m_B(\{x_1, x_3\}) + m_B(\{x_2, x_3\}) + m_B(\mathcal{X}) \\ = m(\{x_1, x_2\}) + m(\{x_1, x_3\}) + m(\{x_2, x_3\}) + m(\mathcal{X}). \end{aligned}$$

As a consequence,

$$\begin{aligned} m(\{x_1, x_2\}) + m(\{x_1, x_3\}) + m(\{x_2, x_3\}) - m_B(\{x_1, x_2\}) \\ - m_B(\{x_1, x_3\}) - m_B(\{x_2, x_3\}) = m_B(\mathcal{X}) - m(\mathcal{X}), \end{aligned}$$

where both terms in the right hand side of the equality are strictly positive.

Let us define m^* by:

$$m^*({x_i}) = m({x_i}), \quad i = 1, 2, 3.$$

$$m^*({x_i, x_j}) = m_B({x_i, x_j}) + \frac{m_B(\mathcal{X})}{m_B(\mathcal{X}) - m(\mathcal{X})} (m({x_i, x_j}) - m_B({x_i, x_j})).$$

$$m^*(\mathcal{X}) = 0.$$

This function m^* is non-negative and also:

$$\begin{aligned} \sum_{A \subseteq \mathcal{X}} m^*(A) &= m_B(\{x_1\}) + m_B(\{x_2\}) + m_B(\{x_3\}) \\ &\quad + m_B(\{x_1, x_2\}) + m_B(\{x_1, x_3\}) + m_B(\{x_2, x_3\}) \\ &\quad + \frac{m_B(\mathcal{X})}{m_B(\mathcal{X}) - m(\mathcal{X})} (m_B(\mathcal{X}) - m(\mathcal{X})) = 1. \end{aligned}$$

We conclude that m^* is a basic probability assignment, so it defines a belief function Bel^* . This belief function Bel^* satisfies $Bel \leq Bel^* \leq \underline{P}$:

- $Bel(\{x_i\}) = Bel^*(\{x_i\}) = \underline{P}(\{x_i\})$ for every $i = 1, 2, 3$.
- $Bel(\{x_i, x_j\}) \leq Bel^*(\{x_i, x_j\}) \leq \underline{P}(\{x_i, x_j\})$ holds if and only if $m_B(\{x_i, x_j\}) \leq m^*(\{x_i, x_j\}) \leq m(\{x_i, x_j\})$. These inequalities hold because:

$$\begin{aligned} m_B(\{x_i, x_j\}) &\leq m_B(\{x_i, x_j\}) + \frac{m_B(\mathcal{X})}{m_B(\mathcal{X}) - m(\mathcal{X})} (m(\{x_i, x_j\}) - m_B(\{x_i, x_j\})) \\ &\leq m_B(\{x_i, x_j\}) + (m(\{x_i, x_j\}) - m_B(\{x_i, x_j\})) = m(\{x_i, x_j\}). \end{aligned}$$

Also, $Bel(\{x_i, x_j\}) < Bel^*(\{x_i, x_j\})$ for all $1 \leq i < j \leq 3$ such that $m(\{x_i, x_j\}) > m_B(\{x_i, x_j\})$ (and there exist at least some i, j for which this inequality holds, because $Bel \not\leq \underline{P}$), so Bel cannot be undominated. \square

Using these preliminary lemmas, we can characterize the set $\tilde{\mathcal{C}}_\infty$ of undominated outer approximations in the class of belief functions of \underline{P} .

Proposition 8. *Let \underline{P} be a coherent lower probability in a ternary space with Möbius inverse m , and assume that it is not a belief function. Then, a belief function Bel is an undominated outer approximation in \mathcal{C}_∞ of \underline{P} if and only if its Möbius inverse m_B satisfies the following properties:*

- (1) $m_B(\{x_i\}) = m(\{x_i\})$, $i = 1, 2, 3$.
- (2) $0 \leq m_B(\{x_i, x_j\}) \leq m(\{x_i, x_j\})$, $1 \leq i < j \leq 3$.
- (3) $\sum_{|A|=2} m_B(A) = m(\mathcal{X}) + \sum_{|A|=2} m(A)$.
- (4) $m_B(\mathcal{X}) = 0$.

Proof. On the one hand, from Lemma 6 we know that the undominated outer approximations in \mathcal{C}_∞ must coincide with \underline{P} in singletons. Also, from Lemma 7, they should assign zero mass to the sure event, so every undominated outer approximation in \mathcal{C}_∞ should satisfy properties (1)–(4).

On the other hand, assume that Bel satisfies the given properties, and let us see that it is an undominated outer approximation of \underline{P} in \mathcal{C}_∞ . First of all, Bel satisfies:

$$\begin{aligned} Bel(\{x_i\}) &= \underline{P}(\{x_i\}), \quad i = 1, 2, 3. \\ Bel(\{x_i, x_j\}) &\leq \underline{P}(\{x_i, x_j\}), \quad 1 \leq i < j \leq 3. \\ Bel(\mathcal{X}) &= \underline{P}(\mathcal{X}) = 1. \end{aligned}$$

This means that Bel is an outer approximation of \underline{P} , and since all the masses are non-negative and they sum up to 1, it is a belief function. Let us now assume that there exists another belief function Bel' , with associated Möbius inverse m' , such that $Bel \leq Bel' \leq \underline{P}$. Then, they should satisfy:

$$\begin{aligned} Bel(\{x_i\}) = Bel'(\{x_i\}) = \underline{P}(\{x_i\}) &\Rightarrow m_B(\{x_i\}) = m'(\{x_i\}), \forall i = 1, 2, 3. \\ Bel(\{x_i, x_j\}) \leq Bel'(\{x_i, x_j\}) &\Rightarrow m_B(\{x_i, x_j\}) \leq m'(\{x_i, x_j\}) \end{aligned}$$

for every $1 \leq i < j \leq 3$. Then, it holds that:

$$\begin{aligned} 1 &\geq \sum_{i=1,2,3} m'(\{x_i\}) + \sum_{1 \leq i < j \leq 3} m'(\{x_i, x_j\}) \\ &\geq \sum_{i=1,2,3} m_B(\{x_i\}) + \sum_{1 \leq i < j \leq 3} m_B(\{x_i, x_j\}) = 1, \end{aligned}$$

which implies that $m_B = m'$, so $Bel = Bel'$. We conclude that Bel is undominated in \mathcal{C}_∞ . \square

From the previous result, we deduce that all undominated outer approximations are at the same distance from the original coherent lower probability:

$$d(\underline{P}, Bel) = \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - Bel(E)) = \sum_{|E|=2} (\underline{P}(E) - Bel(E)) = -m(\mathcal{X}).$$

From this fact, we deduce the following properties:

Corollary 9. *Let \underline{P} be a coherent lower probability on a three-element space. The undominated outer approximations of \underline{P} in \mathcal{C}_∞ coincide with the optimal solutions of (LP-bel) subject to (LP-bel.1) \div (LP-bel.2). Hence, the set $\tilde{\mathcal{C}}_\infty$ is convex.*

We conclude that for ternary spaces, all the undominated outer approximations in \mathcal{C}_∞ are in particular solutions of the linear programming problem (LP-bel)³. Further, the set $\tilde{\mathcal{C}}_\infty$ is convex and all the undominated outer approximations Bel in \mathcal{C}_∞ coincide with \underline{P} in singletons. These facts are in contrast with the general results established in Section 4.2, where we have proven that not all the undominated outer approximations in \mathcal{C}_∞ are obtained solving a linear programming problem (see Example 2), $\tilde{\mathcal{C}}_\infty$ is not convex in general (Example 3) and the undominated outer approximations Bel in \mathcal{C}_∞ may not coincide with \underline{P} in singletons (see Example 4). Since all these examples are established in spaces of cardinality four, we believe it would be hard to find other cases where the structure of the set of undominated outer approximations is easy to work with.

5. COMPARISON WITH OTHER APPROACHES

Next we consider other approaches to compute undominated outer approximations. First of all, we discuss two algorithms given in [21, 30] which can be used to compute outer approximations, and also we consider quadratic linear problems instead of linear ones.

³Note that we do not even need to run the linear programming problem (LP-bel): its optimality set is characterized by conditions (1) \div (4) in Proposition 8. Because of this, it is easy to see that it is a singleton (i.e., that there is a unique solution to (LP-bel)) if and only if the mass function of \underline{P} is zero in two of the three events of $\mathcal{P}(\mathcal{X})$ with cardinality two.

5.1. Iterative and minimal rescaling methods. The problem of outer approximating a coherent lower probability by means of a belief function was also investigated in [21, 30]. In both references, the authors considered the linear programming problem (LP-bel) subject to (LP-bel.1)÷(LP-bel.2), and they discussed how to solve it efficiently by means of a heuristic procedure. In [21], Hall and Lawry proposed an algorithm called *Iterative Rescaling Method* (IRM) to find outer approximations in \mathcal{C}_∞ . This algorithm works as follows:

Step 0: \underline{P} is the coherent lower probability with Möbius inverse m . Define an order among the non-empty subsets of \mathcal{X} such that $i < j$ whenever $|A_i| < |A_j|$. Set $i = 1$.

Step 1: For A_i , evaluate $m(A_i) = \underline{P}(A_i) - \sum_{B \subsetneq A_i} m(B)$.

Step 2: If $m(A_i) \geq 0$, then let $i = i + 1$ and go to **Step 3**.

Else:

Step 2.1: Determine the value k :

$$k = \max \left\{ j \mid j < |A_i|, \sum_{B \subsetneq A_i: |B| \geq j} m(B) > \underline{P}(A_i) \right\}, \quad (18)$$

Step 2.2: Set

$$\alpha = \sum_{B \subsetneq A_i: |B| \geq k} m(B). \quad (19)$$

Step 2.3: For each $B \subsetneq A_i$ with $|B| \geq k$, rescale $m(B)$ using the formula:

$$m(B) = \frac{\alpha + m(A_i)}{\alpha} m(B).$$

Step 2.4: Set $m(A_i) = 0$, let $i = i + 1$ and go to **Step 3**.

Step 3: If $A_i = \mathcal{X}$, end the algorithm.

Else: go to **Step 1**.

The basic idea behind this algorithm is to compute the Möbius inverses until we find a negative value for the event A_i . At that point, positive masses of some subsets of A_i are transferred to A_i . In this way, the masses of some subsets are slightly decreased, while the mass of A_i becomes 0. It was proven in [21, Thm. 1] that this algorithm produces a belief function that outer approximates \underline{P} .

Afterwards, Quaeghebeur [30] pointed out two flaws of the IRM: on the one hand, the solution of the algorithm depends on the order chosen among the non-empty subsets, in the sense that different orders may produce different solutions (see Table 1 in [30]). On the other hand, even if all the optimal solutions of the linear programming problem (LP-bel) satisfy $Bel(A) = \underline{P}(A)$ for an event A , the solution given by the IRM may not attain the value $\underline{P}(A)$ (see Table 2 in [30]).

Inspired by these comments, Quaeghebeur [30] compared the IRM with a number of alternative approaches:

- Solving the linear programming problem by means of either a dual simplex solver or a criss-cross solver.
- The restriction of the set of outer approximations to linear-vacuous mixtures. This coincides with the ϵ -contamination outer approximation we gave in [27].
- Iterative minimal rescaling method: a slightly modified version of the IRM, that performs better than the IRM.

The modified version of the IRM works as follows:

Step 0: \underline{P} is the coherent lower probability with Möbius inverse m . Set $i = 1$.

Step 1: For each A satisfying $|A| = i$, evaluate $m(A) = \underline{P}(A) - \sum_{B \subsetneq A} m(B)$, and consider the set $\mathcal{A} = \{A \subseteq \mathcal{X} \mid |A| = i, m(A) < 0\}$.

Step 2: If $\mathcal{A} = \emptyset$, let $i = i + 1$ and go to **Step 1**.

Else:

Step 2.1: For each $A \in \mathcal{A}$, use Eqs. (18) and (19) to obtain the values k_A and α_A .

Step 2.2: Set $k = \min_{A \in \mathcal{A}} k_A$ and $\mathcal{B} = \{A \in \mathcal{A} \mid k_A = k\}$.

Step 2.3: For each $A \in \mathcal{B}$, set $\beta_A = \sum_{B \subsetneq A: |B|=k} m(B)$.

Step 2.4: For each $B \in \cup_{A \in \mathcal{B}} \mathcal{P}(A)$ such that $|B| = k$, rescale $m(B)$ using the formula:

$$\mu(B) = \max_{A \in \mathcal{B}: B \subsetneq A} \frac{\alpha_A + m(A)}{\beta_A} m(B).$$

Step 2.5: Let $i = i + 1$ and go to **Step 3**.

Step 3: If $i > n$, end the algorithm.

Else: go to **Step 1**.

In contrast to the IRM, the IMRM rescales the negative masses of all the events of the same cardinality at the same time. In doing so, the algorithm limits the mass loss for events with low cardinality. The reason is that an event of low cardinality loses mass only if it cannot be compensated with events of higher cardinality.

In any case, Quaeghebeur [30] showed that neither IRM nor its modified version performs so well as the linear programming approach, because they may produce suboptimal solutions. Our next example also shows that both the IRM and its modified version may produce dominated outer approximations in \mathcal{C}_∞ .

Example 5. *Let us consider the example given in [30, Table 1]. There, \mathcal{X} is a four element possibility space, and the following coherent lower probability*

is defined:

| A | \underline{P} | $\underline{P}_{\text{IRM}}$ | m_{IRM} | $\underline{P}_{\text{IMRM}}$ | m_{IMRM} |
|---------------------|-----------------|------------------------------|------------------|-------------------------------|-------------------|
| $\{x_1\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_2\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_3\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_4\}$ | 0 | 0 | 0 | 0 | 0 |
| $\{x_1, x_2\}$ | 0.117 | 0.046 | 0.046 | 0.066 | 0.066 |
| $\{x_1, x_3\}$ | 0.358 | 0.185 | 0.185 | 0.211 | 0.211 |
| $\{x_1, x_4\}$ | 0.381 | 0.242 | 0.242 | 0.244 | 0.244 |
| $\{x_2, x_3\}$ | 0.162 | 0.074 | 0.074 | 0.082 | 0.082 |
| $\{x_2, x_4\}$ | 0.391 | 0.219 | 0.219 | 0.227 | 0.227 |
| $\{x_3, x_4\}$ | 0.132 | 0.099 | 0.099 | 0.081 | 0.081 |
| $\{x_1, x_2, x_3\}$ | 0.358 | 0.305 | 0 | 0.358 | 0 |
| $\{x_1, x_2, x_4\}$ | 0.579 | 0.507 | 0 | 0.579 | 0.042 |
| $\{x_1, x_3, x_4\}$ | 0.550 | 0.526 | 0 | 0.550 | 0.014 |
| $\{x_2, x_3, x_4\}$ | 0.391 | 0.391 | 0 | 0.391 | 0 |
| \mathcal{X} | 1 | 1 | 0.136 | 1 | 0.033 |

Let us see that none of $\underline{P}_{\text{IRM}}$ and $\underline{P}_{\text{IMRM}}$ is undominated. First of all, with respect to $\underline{P}_{\text{IRM}}$, note that $\underline{P}_{\text{IRM}}(\{x_1, x_2, x_3\}) < \underline{P}(\{x_1, x_2, x_3\})$. Also, since $m_{\text{IRM}}(\mathcal{X}) > 0$, we can transfer some positive mass from \mathcal{X} to $\{x_1, x_2, x_3\}$, getting a greater belief function. Formally, take

$$\varepsilon = \underline{P}(\{x_1, x_2, x_3\}) - \underline{P}_{\text{IRM}}\{x_1, x_2, x_3\} = 0.053,$$

and define the basic probability assignment m_1 by:

$$m_1(A) = \begin{cases} \varepsilon & \text{if } A = \{x_1, x_2, x_3\}. \\ m(\mathcal{X}) - \varepsilon & \text{if } A = \mathcal{X}. \\ m(A) & \text{otherwise.} \end{cases}$$

It can be easily seen that its associated belief function, Bel_1 , satisfies the equality $Bel_1(A) = \underline{P}_{\text{IRM}}(A)$ for every $A \neq \{x_1, x_2, x_3\}$ and

$$Bel_1(\{x_1, x_2, x_3\}) = \varepsilon + \underline{P}_{\text{IRM}}(\{x_1, x_2, x_3\}) > \underline{P}_{\text{IRM}}(\{x_1, x_2, x_3\}).$$

Also, $Bel_1 \not\leq \underline{P}$, so we conclude that $\underline{P}_{\text{IRM}}$ is not undominated.

Similarly, $\underline{P}_{\text{IMRM}}(A) < \underline{P}(A)$ for every event A of cardinality 2, and also $m_{\text{IMRM}}(\{x_1, x_2, x_4\}), m_{\text{IMRM}}(\{x_1, x_3, x_4\}) > 0$. This means that we can increase the mass of the events $\{x_1, x_4\}$ and \mathcal{X} by ε small enough and decrease the mass of the events $\{x_1, x_2, x_4\}, \{x_1, x_3, x_4\}$ by the same ε . Formally, take:

$$\varepsilon = \min \left\{ m_{\text{IMRM}}(\{x_1, x_2, x_4\}), m_{\text{IMRM}}(\{x_1, x_3, x_4\}) \right\},$$

$$\underline{P}(\{x_1, x_4\}) - \underline{P}_{\text{IMRM}}(\{x_1, x_4\}) = \min\{0.042, 0.014, 0.137\} = 0.014.$$

Now, we can define the basic probability assignment m_2 by:

$$m_2(A) = \begin{cases} m_{\text{IMRM}}(\{x_1, x_4\}) + \varepsilon & \text{if } A = \{x_1, x_4\}. \\ m_{\text{IMRM}}(\mathcal{X}) + \varepsilon & \text{if } A = \mathcal{X}. \\ m_{\text{IMRM}}(\{x_1, x_2, x_4\}) - \varepsilon & \text{if } A = \{x_1, x_2, x_4\}. \\ m_{\text{IMRM}}(\{x_1, x_3, x_4\}) - \varepsilon & \text{if } A = \{x_1, x_3, x_4\}. \\ m_{\text{IMRM}}(A) & \text{otherwise.} \end{cases}$$

Its associated belief function, Bel_2 , satisfies $Bel_2(A) = \underline{P}_{\text{IMRM}}(A)$ for every $A \neq \{x_1, x_4\}$ and

$$Bel_2(\{x_1, x_4\}) = \underline{P}_{\text{IMRM}}(\{x_1, x_4\}) + \varepsilon > \underline{P}_{\text{IMRM}}(\{x_1, x_4\}).$$

Also, $Bel_2 \not\leq \underline{P}$, so we conclude that $\underline{P}_{\text{IMRM}}$ is dominated by Bel_2 in \mathcal{C}_∞ , hence it is not undominated. \blacklozenge

5.2. Outer approximations using quadratic programming problems.

In Section 4 we have studied how to obtain undominated outer approximations by means of linear programming. One of the drawbacks of this approach is that the optimization problem has no unique solution in general. This issue can be overcome by means of quadratic programming, so that instead of minimizing the sum of the distances over all events, we minimize the sum of the square distances⁴:

$$\min \sum_{E \subseteq \mathcal{X}} (\underline{P}(E) - Bel(E))^2. \quad (\text{QP-bel})$$

By expanding this expression, we obtain that this is equivalent to:

$$\min \sum_{E \subseteq \mathcal{X}} (Bel(E)^2 - 2\underline{P}(E)Bel(E)).$$

Of course, we need to add the constraints (LP-bel.1) and (LP-bel.2). The quadratic programming problem in (QP-bel) can be expressed in the usual matrix form:

$$\min \frac{1}{2} \vec{Bel}^t H \vec{Bel} + c^t \vec{Bel},$$

where \vec{Bel} denotes a vector with the values of Bel , $H = 2I_{2^n-1}$ is twice the identity matrix of size $2^n - 1$ and $c = -2\vec{P}$, where again \vec{P} is a vector with the values of \underline{P} . Since H is positive definite, the quadratic programming problem (QP-bel) subject to constraints (LP-bel.1) and (LP-bel.2) has an optimal solution and it is unique.

Next, we prove that the unique optimal solution of (QP-bel) is an undominated outer approximation of \underline{P} .

⁴The quadratic distance could be normalized just dividing by the maximum quadratic distance. If $\underline{P} \geq Q$, the maximum distance turns out to be $\sum_{k=1}^{n-1} \binom{n-1}{k-1} \frac{k}{n}$, that is attained taking Q the vacuous lower probability and \underline{P} a precise probability uniformly distributed on \mathcal{X} .

Proposition 10. *Let \underline{P} be a coherent lower probability, and denote by Bel the optimal solution to the quadratic programming problem (QP-bel) subject to (LP-bel.1) and (LP-bel.2). Then, Bel is an undominated belief function that outer approximates \underline{P} .*

Proof. That Bel is a belief function trivially follows from constraint (LP-bel.1). Also, from and (LP-bel.2) Bel is an outer approximation of \underline{P} in \mathcal{C}_∞ . It only remains to see that it is undominated in \mathcal{C}_∞ , but it suffices to use Lemma 1 with the function $g(x) = x^2$, that is increasing for $x \in [0, +\infty)$. \square

Example 6. *Let us continue Example 1. If we solve the corresponding quadratic programming problem, we obtain the following undominated belief function:*

| A | $Bel(A)$ |
|---------------------|----------|
| $\{x_1\}$ | 0.1 |
| $\{x_2\}$ | 0 |
| $\{x_3\}$ | 0 |
| $\{x_4\}$ | 0.3 |
| $\{x_1, x_2\}$ | 0.1 |
| $\{x_1, x_3\}$ | 0.25 |
| $\{x_1, x_4\}$ | 0.55 |
| $\{x_2, x_3\}$ | 0.25 |
| $\{x_2, x_4\}$ | 0.35 |
| $\{x_3, x_4\}$ | 0.3 |
| $\{x_1, x_2, x_3\}$ | 0.5 |
| $\{x_1, x_2, x_4\}$ | 0.6 |
| $\{x_1, x_3, x_4\}$ | 0.7 |
| $\{x_2, x_3, x_4\}$ | 0.6 |
| \mathcal{X} | 1 |

None of the undominated belief functions given in Example 1 dominates this belief function: in fact, if we consider the optimal solutions Bel_0, Bel_1 of the linear programming problem, it holds that $Bel = Bel_{0.5} = (Bel_0 + Bel_1)/2$, whence they are all at the same distance from \underline{P} . This means that the solution of the quadratic problem is also an optimal solution of the linear programming problem (and is therefore undominated). \blacklozenge

We see then that we can obtain undominated outer approximations by means of both the linear or quadratic programming approaches. However, these two approaches are not without drawbacks: on the one hand, the linear programming problem does not produce a unique solution; on the other, the quadratic distance lacks a compelling interpretation and it also entails a heavier computational cost. For these reasons, in the remainder of the paper we investigate the problem for the two particular cases of belief functions introduced in Sections 2.4 and 2.5.

6. PARTICULAR CASES: POSSIBILITY MEASURES AND P-BOXES

In this section we consider outer approximations in two particular families of belief functions: necessity measures (or their conjugate possibility measures) and p-boxes. These two families are computationally simpler; moreover, we shall prove that in these cases, either the undominated outer approximation is unique, or we can determine the finite set of undominated outer approximations.

6.1. Outer approximations using possibility measures. In this subsection we consider the problem of outer approximating a coherent lower probability using a necessity measure. Since possibility measures appear more frequently in the literature than their conjugate necessity measures, due for instance to their connection with fuzzy sets [16, 39], throughout this section we shall work with them and use the following notation: we say that a possibility measure Π outer approximates a coherent upper probability \overline{P} with conjugate lower probability \underline{P} when $\overline{P} \leq \Pi$. Of course, this is equivalent to $N \leq \underline{P}$, where N is the conjugate necessity measure of Π , and also to $\mathcal{M}(\underline{P}) \subseteq \mathcal{M}(\Pi)$, where these two sets are given by Eqs. (2.2) and (5), respectively. Because of the conjugacy relation, we will call such an outer approximation Π *non-dominating* when its conjugate necessity measure is an undominated outer approximation of N : that is, when there is no other possibility measure Π' such that $\overline{P} \leq \Pi' \leq \Pi$.

Our next result provides non-dominating outer approximations of a coherent upper probability \overline{P} in \mathcal{C}_Π ; we will later establish (in Corollary 13) that they are the only ones.

Proposition 11. *Let \overline{P} be a coherent upper probability on $\mathcal{P}(\mathcal{X})$. Take $\sigma \in S_n$ and define a possibility measure $\Pi : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ by $\Pi(\emptyset) = 0$,*

$$\Pi(\{x_{\sigma(1)}\}) = \overline{P}(\{x_{\sigma(1)}\}) \text{ and} \quad (20)$$

$$\Pi(\{x_{\sigma(i)}\}) = \max_{A \in \mathcal{A}_{\sigma(i)}} \overline{P}(A \cup \{x_{\sigma(i)}\}), \text{ where for every } i > 1: \quad (21)$$

$$\mathcal{A}_{\sigma(i)} = \left\{ A \subseteq \{x_{\sigma(1)}, \dots, x_{\sigma(i-1)}\} \mid \overline{P}(A \cup \{x_{\sigma(i)}\}) > \max_{x \in A} \Pi(\{x\}) \right\}, \quad (22)$$

and let $\Pi(A) = \max_{x \in A} \Pi(\{x\})$ for every other $A \subseteq \mathcal{X}$. Then, Π is a non-dominating outer approximation of \overline{P} in \mathcal{C}_Π .

Proof. Let us prove that Π is a possibility measure, that it outer approximates \overline{P} and that it is non-dominating.

- (i) To see that Π is a possibility measure, since it is maxitive by construction we only need to show that $\max_{x \in \mathcal{X}} \Pi(\{x\}) = 1$. Now, if $\max_{i < n} \Pi(\{x_{\sigma(i)}\}) < 1$, it follows from Eq. (22) that

$$A = \{x_{\sigma(1)}, \dots, x_{\sigma(n-1)}\} \in \mathcal{A}_{\sigma(n)},$$

and then by Eq. (21), $\Pi(\{x_n\}) = \overline{P}(\mathcal{X}) = 1$.

- (ii) Let us now see that Π is an outer approximation of \underline{P} , that is, $\overline{P}(A) \leq \Pi(A)$ for every $A \subseteq \mathcal{X}$. First of all, note that $\overline{P}(\{x_{\sigma(1)}\}) = \Pi(\{x_{\sigma(1)}\})$ by Eq. (20). Take $A \neq \{x_{\sigma(1)}\}$, and denote by i the element:

$$i = \max\{j = 1, \dots, n \mid x_{\sigma(j)} \in A\}.$$

Then, $\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\} \supseteq A$. There are two possibilities:

- (a) If $A \setminus \{x_{\sigma(i)}\} \notin \mathcal{A}_{\sigma(i)}$, then by definition of $\mathcal{A}_{\sigma(i)}$,

$$\overline{P}(A) \leq \max_{x \in A \setminus \{x_{\sigma(i)}\}} \Pi(\{x\}) = \Pi(A \setminus \{x_{\sigma(i)}\}) \leq \Pi(A),$$

and as a consequence $\overline{P}(A) \leq \Pi(A)$.

- (b) If $A \setminus \{x_{\sigma(i)}\} \in \mathcal{A}_{\sigma(i)}$, then by Eq. (21)

$$\Pi(\{x_{\sigma(i)}\}) = \max_{B \in \mathcal{A}_{\sigma(i)}} \overline{P}(B \cup \{x_{\sigma(i)}\}) \geq \overline{P}(A),$$

whence

$$\Pi(A) = \max_{x \in A} \Pi(\{x\}) \geq \Pi(\{x_{\sigma(i)}\}) \geq \overline{P}(A).$$

Thus, Π outer approximates \overline{P} .

- (iii) To see that it is a non-dominating outer approximation in \mathcal{C}_{Π} , let Π' denote a possibility measure satisfying $\overline{P} \leq \Pi' \leq \Pi$. By definition, $\Pi(\{x_{\sigma(1)}\}) = \overline{P}(\{x_{\sigma(1)}\})$, and therefore $\Pi(\{x_{\sigma(1)}\}) = \Pi'(\{x_{\sigma(1)}\})$. Assume that

$$\Pi(\{x_{\sigma(1)}\}) = \Pi'(\{x_{\sigma(1)}\}), \dots, \Pi(\{x_{\sigma(i-1)}\}) = \Pi'(\{x_{\sigma(i-1)}\})$$

and let us prove that $\Pi(\{x_{\sigma(i)}\}) = \Pi'(\{x_{\sigma(i)}\})$. Taking Eqs. (21)–(22) into account, we have the following possibilities:

- (a) If there is no $A \neq \emptyset$ in $\mathcal{A}_{\sigma(i)}$, we deduce from Eq. (21) that:

$$\Pi(\{x_{\sigma(i)}\}) = \overline{P}(\{x_{\sigma(i)}\}),$$

whence $\Pi(\{x_{\sigma(i)}\}) = \Pi'(\{x_{\sigma(i)}\})$.

- (b) If there exists $A \neq \emptyset$ such that $A \in \mathcal{A}_{\sigma(i)}$, let B denote an event in $\mathcal{A}_{\sigma(i)}$ such that $\Pi(\{x_{\sigma(i)}\}) = \overline{P}(B \cup \{x_{\sigma(i)}\})$. Since $B \in \mathcal{A}_{\sigma(i)}$, it holds that

$$\overline{P}(B \cup \{x_{\sigma(i)}\}) > \max_{x \in B} \Pi(\{x\}) = \Pi(B). \quad (23)$$

Therefore:

$$\Pi(B \cup \{x_{\sigma(i)}\}) = \max\{\Pi(\{x_{\sigma(i)}\}), \Pi(B)\} = \Pi(\{x_{\sigma(i)}\}) = \overline{P}(B \cup \{x_{\sigma(i)}\}),$$

whence $\Pi(B \cup \{x_{\sigma(i)}\}) = \Pi'(B \cup \{x_{\sigma(i)}\}) = \overline{P}(B \cup \{x_{\sigma(i)}\})$. On the other hand,

$$\begin{aligned} \overline{P}(B \cup \{x_{\sigma(i)}\}) &\leq \Pi'(B \cup \{x_{\sigma(i)}\}) = \max\{\Pi'(\{x_{\sigma(i)}\}), \Pi'(B)\} \\ &\leq \max\{\Pi'(\{x_{\sigma(i)}\}), \Pi(B)\} = \Pi'(\{x_{\sigma(i)}\}), \end{aligned}$$

taking into account that $\Pi(B) < \bar{P}(B \cup \{x_{\sigma(i)}\})$ by Eq. (23).

Thus, $\Pi'(\{x_{\sigma(i)}\}) \geq \bar{P}(B \cup \{x_{\sigma(i)}\}) = \Pi(\{x_{\sigma(i)}\})$ and therefore $\Pi'(\{x_{\sigma(i)}\}) = \Pi(\{x_{\sigma(i)}\})$.

We conclude that $\Pi(\{x\}) = \Pi'(\{x\})$ for every $x \in \mathcal{X}$, and as a consequence

$$\Pi(A) = \max_{x \in A} \Pi(\{x\}) = \max_{x \in A} \Pi'(\{x\}) = \Pi'(A) \quad \forall A \subseteq \mathcal{X}. \quad \square$$

Similarly to the comment we made in Example 1 about the outer approximations in \mathcal{C}_∞ and \mathcal{C}_2 , a non-dominating outer approximation of \bar{P} in the class \mathcal{C}_Π may not be so if we consider the broader class of the plausibility functions: to see this, it suffices to consider \bar{P} that is a plausibility function and not a possibility measure.

Let us show next that the possibility measures determined by Proposition 11 allow us to recover the initial coherent lower probability.

Proposition 12. *Let \underline{P} be a coherent lower probability with associated credal set $\mathcal{M}(\underline{P})$, and let \bar{P} be its conjugate coherent upper probability. Denote by Π_1, \dots, Π_k the possibility measures determined by Eqs. (20)–(22) for $\sigma \in S_n$. Then $\mathcal{M}(\underline{P}) = \mathcal{M}(\Pi_1) \cap \dots \cap \mathcal{M}(\Pi_k)$, or equivalently,*

$$\bar{P}(A) = \min\{\Pi_1(A), \dots, \Pi_k(A)\} \quad \forall A \subseteq \mathcal{X}.$$

Proof. Since Π_i is an outer approximation of \bar{P} for $i \in \{1, \dots, k\}$, it follows that $\bar{P} \leq \min\{\Pi_1, \dots, \Pi_k\}$. In order to establish the converse, we shall prove that for every $A \subseteq \mathcal{X}$ there exists some $i \in \{1, \dots, k\}$ such that $\Pi_i(A) = \bar{P}(A)$. First of all, if A is a singleton, $A = \{x\}$, we just need to consider a permutation such that x is the first element, since then Eq. (20) implies that $\bar{P}(A) = \bar{P}(\{x\}) = \Pi(\{x\})$. Assume that $|A| \geq 2$, and let σ be a permutation such that $A = \{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}$. Let Π be the possibility measure it determines by means of Eqs. (20)–(22).

By monotonicity, for every $j \leq i$ it holds that

$$\Pi(\{x_{\sigma(j)}\}) = \max_{B \in \mathcal{A}_{\sigma(j)}} \bar{P}(B \cup \{x_{\sigma(j)}\}) \leq \bar{P}(A),$$

whence $\Pi(A) = \max_{j \leq i} \Pi(\{x_{\sigma(j)}\}) \leq \bar{P}(A)$. Since the converse inequality holds because Π is an outer approximation of \bar{P} , we conclude that $\Pi(A) = \bar{P}(A)$. \square

This result allows us to deduce that the procedure in Proposition 11 determines all the non-dominating outer approximations of \bar{P} in \mathcal{C}_Π :

Corollary 13. *Let \bar{P} be a coherent upper probability. A possibility measure Π is a non-dominating outer approximation of \bar{P} in \mathcal{C}_Π if and only if there exists $\sigma \in S_n$ inducing Π by means of Eqs. (20)–(22).*

Proof. The ‘if’ part has been established in Proposition 11; let us then prove the ‘only if’ part.

Let π be the possibility distribution of Π , and let $\sigma \in S_n$ be a permutation satisfying $\pi(x_{\sigma(1)}) \leq \pi(x_{\sigma(2)}) \leq \dots \leq \pi(x_{\sigma(n)}) = 1$. Denote by Π_σ the possibility measure σ determines by means of Eqs. (20)–(22). We shall prove that $\Pi_\sigma(\{x_{\sigma(i)}\}) \leq \Pi(\{x_{\sigma(i)}\})$ for every $i \in \{1, \dots, n\}$.

From the proof of Proposition 12, for every $i \in \{1, \dots, n\}$ it holds that

$$\begin{aligned} \Pi_\sigma(\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}) &= \bar{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}) \\ &\leq \Pi(\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}) = \Pi(\{x_{\sigma(i)}\}), \end{aligned}$$

from which it follows in particular that $\Pi_\sigma(\{x_{\sigma(i)}\}) \leq \Pi(\{x_{\sigma(i)}\})$ for every i . This implies that $\Pi_\sigma \leq \Pi$, and since Π is assumed to be non-dominating we conclude that they are equal. \square

As a consequence, there are at most $n!$ non-dominating outer approximations of \bar{P} in \mathcal{C}_Π . Our next example shows that this bound is tight:

Example 7. Let \mathcal{X} be a possibility space with cardinality n , and consider the credal set $\mathcal{M} = \{P\}$ where P is the uniform distribution in \mathcal{X} . Then, $P(A) = \frac{|A|}{n}$ for every $A \subseteq \mathcal{X}$. Given a permutation σ of $\{1, \dots, n\}$, using Eqs. (20)–(22) we obtain the possibility measure Π_σ given by $\Pi_\sigma(\{x_{\sigma(i)}\}) = \frac{i}{n}$ for every $i \in \{1, \dots, n\}$. Therefore, every permutation produces a different possibility measure. \blacklozenge

6.1.1. *Comparison with the Optimal Mass Allocation Procedure.* A procedure similar to that of Proposition 11 was considered earlier by Dubois and Prade in [14] and [17, Section 3.3] in the particular case where \underline{P} is a belief function, and it was given the name of *Optimal Mass Allocation Procedure*. Dubois and Prade later adapted it to outer approximate a coherent upper (or lower) probability [18]. The generalized procedure works as follows: for every permutation $\sigma \in S_n$, define $E_j^\sigma = \{x_{\sigma(1)}, \dots, x_{\sigma(j)}\}$ for $j = 1, \dots, n$. Then, let the possibility distribution π_σ^{DP} be given by:

$$\pi_\sigma^{DP}(x_{\sigma(i)}) = 1 - \underline{P}(E_{i-1}^\sigma) = \bar{P}(\{x_{\sigma(i)}, \dots, x_{\sigma(n)}\}) \quad \forall i \in \{1, \dots, n\}, \quad (24)$$

where $E_0^j := \emptyset$, and then $\Pi_\sigma^{DP}(A) = \max_{x \in A} \pi_\sigma^{DP}(x)$ for any $A \subseteq \mathcal{X}$.

In [17], it was proven that among the set $\{\Pi_\sigma^{DP} : \sigma \in S_n\}$ we can find all the undominated outer approximations of \underline{P} .

Let us prove that the possibility measures defined with the Optimal Mass Allocation Procedure always dominate some of the possibility measures defined using our procedure from Proposition 11. For this aim, we use the following notation: if $\sigma \in S_n$, $\bar{\sigma}$ is the permutation defined by $\bar{\sigma}(i) = \sigma(n-i+1)$ for every $i \in \{1, \dots, n\}$.

Proposition 14. Let \underline{P} be a coherent lower probability and $\sigma \in S_n$, and let \bar{P} denote its conjugate coherent upper probability. Let Π_σ be the possibility measure built using Eqs. (20)–(22), and let Π_σ^{DP} be the possibility measure defined using the Optimal Mass Allocation Procedure with the permutation $\bar{\sigma}$. Then, $\Pi_\sigma \leq \Pi_\sigma^{DP}$.

Proof. By Eq. (24), it holds that:

$$\begin{aligned}\Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(n)}\}) &= \bar{P}(\{x_{\bar{\sigma}(n)}\}) = \bar{P}(\{x_{\sigma(1)}\}), \\ \Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(n-1)}\}) &= \bar{P}(\{x_{\bar{\sigma}(n-1)}, x_{\bar{\sigma}(n)}\}) = \bar{P}(\{x_{\sigma(1)}, x_{\sigma(2)}\}), \\ &\dots \\ \Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(2)}\}) &= \bar{P}(\{x_{\bar{\sigma}(2)}, \dots, x_{\bar{\sigma}(n)}\}) = \bar{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(n-1)}\}), \\ \Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(1)}\}) &= \bar{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(n)}\}) = 1,\end{aligned}$$

whence

$$\Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(n)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(n-1)}\}) \leq \dots \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(2)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(1)}\}) = 1. \quad (25)$$

Recall that by definition of $\bar{\sigma}$, $\Pi_{\bar{\sigma}}^{DP}(\{x_{\bar{\sigma}(i)}\}) = \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(n-i+1)}\})$.

Let us now prove by induction that $\Pi_{\sigma}(\{x_{\sigma(i)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(i)}\})$ for every $i \in \{1, \dots, n\}$.

- For $i = 1$, it holds that:

$$\Pi_{\sigma}(\{x_{\sigma(1)}\}) = \bar{P}(\{x_{\sigma(1)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(1)}\}),$$

where the equality follows by Eq. (20) and the inequality follows because $\Pi_{\bar{\sigma}}^{DP}$ is an outer approximation of \underline{P} .

- Assume that $\Pi_{\sigma}(\{x_{\sigma(j)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(j)}\})$ for every $j = 2, \dots, i-1$, and let us prove that $\Pi_{\sigma}(\{x_{\sigma(i)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(i)}\})$. There are two possible scenarios:

- If there exists $j < i$ such that $\Pi_{\sigma}(\{x_{\sigma(i)}\}) \leq \Pi_{\sigma}(\{x_{\sigma(j)}\})$, then

$$\Pi_{\sigma}(\{x_{\sigma(i)}\}) \leq \Pi_{\sigma}(\{x_{\sigma(j)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(j)}\}) \leq \Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(i)}\}),$$

where the second inequality follows by the induction hypothesis and the third one follows from Eq. (25).

- If $\Pi_{\sigma}(\{x_{\sigma(i)}\}) > \max_{j < i} \Pi_{\sigma}(\{x_{\sigma(j)}\})$ then by definition (see Eq. (21))

$$\Pi_{\sigma}(\{x_{\sigma(i)}\}) = \max_{A \in \mathcal{A}_{\sigma(i)}} \bar{P}(A \cup \{x_{\sigma(i)}\}).$$

Let $A^* \in \mathcal{A}_{\sigma(i)}$ be an event such that $\Pi_{\sigma}(\{x_{\sigma(i)}\}) = \bar{P}(A^* \cup \{x_{\sigma(i)}\})$; note that there is one such event because otherwise it would be $\Pi_{\sigma}(\{x_{\sigma(i)}\}) = 0$. This implies that:

$$\begin{aligned}\Pi_{\sigma}(A^* \cup \{x_{\sigma(i)}\}) &= \max\{\Pi_{\sigma}(\{x_{\sigma(i)}\}), \Pi_{\sigma}(\{x_{\sigma(j)}\}) \mid x_{\sigma(j)} \in A^*\} \\ &= \Pi_{\sigma}(\{x_{\sigma(i)}\}) = \bar{P}(A^* \cup \{x_{\sigma(i)}\}),\end{aligned} \quad (26)$$

where the second equality follows by our assumption. Using Eqs. (25) and (26), we deduce that:

$$\Pi_{\bar{\sigma}}^{DP}(\{x_{\sigma(i)}\}) = \Pi_{\bar{\sigma}}^{DP}(A^* \cup \{x_{\sigma(i)}\}) \geq \bar{P}(A^* \cup \{x_{\sigma(i)}\}) = \Pi_{\sigma}(\{x_{\sigma(i)}\}),$$

where the inequality follows because $\Pi_{\bar{\sigma}}^{DP}$ is an outer approximation of \underline{P} .

This completes the proof. \square

Thus, if \bar{P} is a coherent upper probability, $\Pi_{\bar{\sigma}}^{DP}$ dominates Π_{σ} . In particular, for $\Pi_{\bar{\sigma}}^{DP}$ to be non-dominating it must coincide with Π_{σ} . Otherwise, $\Pi_{\bar{\sigma}}^{DP}$ would be more imprecise than Π_{σ} . This means that our procedure in Proposition 11 gives all the non-dominating possibility measures, while Dubois and Prade's procedure gives not only the non-dominating possibility measures, but maybe also some dominating ones.

The following example shows that the inequality established in Proposition 14 may be strict.

Example 8. Consider a three-element space and the belief and plausibility functions Bel, Pl determined by the basic probability assignment:

$$m(\{x_1\}) = 0.3, \quad m(\{x_3\}) = 0.3, \quad m(\{x_1, x_2\}) = 0.1, \quad m(\{x_2, x_3\}) = 0.3, \\ \text{and } m(E) = 0 \text{ otherwise. Then, } Bel, Pl \text{ are given by:}$$

| A | $Bel(A)$ | $Pl(A)$ |
|----------------|----------|---------|
| $\{x_1\}$ | 0.3 | 0.4 |
| $\{x_2\}$ | 0 | 0.4 |
| $\{x_3\}$ | 0.3 | 0.6 |
| $\{x_1, x_2\}$ | 0.4 | 0.7 |
| $\{x_1, x_3\}$ | 0.6 | 1 |
| $\{x_2, x_3\}$ | 0.6 | 0.7 |
| \mathcal{X} | 1 | 1 |

Considering the permutation $\sigma = (3, 1, 2)$, our procedure from Proposition 11 gives the possibility measure Π_{σ} with possibility distribution:

$$\pi_{\sigma}(x_1) = 1, \quad \pi_{\sigma}(x_2) = 0.7, \quad \pi_{\sigma}(x_3) = 0.6.$$

If we apply Dubois and Prade's procedure with the permutation $\bar{\sigma} = (2, 1, 3)$, we obtain the possibility distribution $\pi_{\bar{\sigma}}^{DP}$ given by:

$$\pi_{\bar{\sigma}}^{DP}(x_1) = \pi_{\bar{\sigma}}^{DP}(x_2) = 1, \quad \pi_{\bar{\sigma}}^{DP}(x_3) = 0.6.$$

We deduce that $\pi_{\bar{\sigma}}^{DP}$ dominates π_{σ} .

In Figure 1 we have graphically depicted the credal sets of Bel, Π_{σ} and $\Pi_{\bar{\sigma}}^{DP}$. In that figure it can be seen that their credal sets are nested: $\mathcal{M}(Bel) \subsetneq \mathcal{M}(\Pi_{\sigma}) \subsetneq \mathcal{M}(\Pi_{\bar{\sigma}}^{DP})$. \blacklozenge

6.1.2. *Preserving preferences with outer approximations.* When outer approximating a coherent lower probability with either a 2-monotone one or a belief function, the number of distinct numerical evaluations theoretically possible for the elements of $\mathcal{P}(\mathcal{X})$ remains unchanged and bounded above by 2^n . By contrast, replacing a coherent upper probability \bar{P} by a possibility measure decreases this number to $n+1$, with the additional constraint that at least one $x_i \in \mathcal{X}$ must satisfy $\Pi(\{x_i\}) = 1$. Because of this, one cannot expect an outer approximation in \mathcal{C}_{Π} to be really close to the original coherent upper probability. One might arguably regard possibility measures

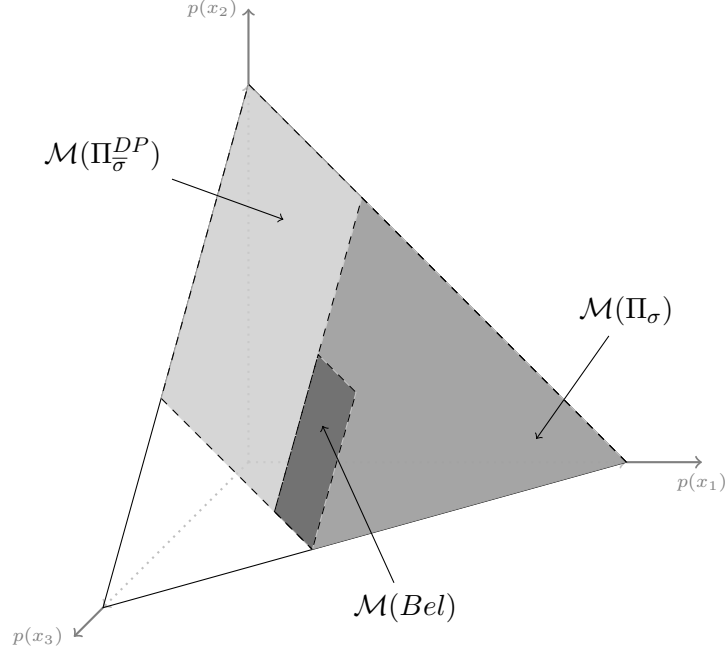


FIGURE 1. Credal sets of the belief function Bel (dark gray area) and the possibility measures Π_σ (the union of dark and medium gray areas) and Π_σ^{DP} (the union of the dark, medium and light gray areas) from Example 8.

as little more than a qualitative judgment, and with these considerations in mind, evaluate that the possibility measure should preserve the order of \bar{P} , rather than aiming at a difficult to reach minimal numerical distance. As we shall now see, there is a trade-off between these goals.

We say that Π weakly preserves the ordering given by \bar{P} when, $\forall A, B \in \mathcal{P}(\mathcal{X})$,

$$\bar{P}(A) = \bar{P}(B) \Rightarrow \Pi(A) = \Pi(B) \quad (\text{N1})$$

$$\bar{P}(A) > \bar{P}(B) \Rightarrow \Pi(A) \geq \Pi(B). \quad (\text{N2})$$

These two conditions formalize a form of the *Weak Preference Preservation* (WPP) principle.

We have already seen that 2-monotone outer approximations ensure a strong form of WPP on singletons (Proposition 2(3)), while the outer approximations in \mathcal{C}_∞ do not (see Example 4 and also Proposition 5). As for possibility measures outer approximating \bar{P} by either Proposition 11 or the Optimal Mass Allocation Procedure, it follows easily from Example 8 that they do not satisfy WPP on singletons in general. Because of this, we say

that a permutation σ^* is *order basic* when

$$\overline{P}(\{x_{\sigma^*(1)}\}) \leq \overline{P}(\{x_{\sigma^*(2)}\}) \leq \dots \leq \overline{P}(\{x_{\sigma^*(n)}\});$$

the possibility measure Π_{σ^*} it induces by means of Proposition 11 is then termed order basic, too. It is easy to see that, since Eqs. (20)–(22) determine a non-decreasing sequence $\Pi(\{x_{\sigma^*(i)}\})_{i=1,\dots,n}$, an order basic possibility measure induces the same order on the atoms of \mathcal{X} , and therefore when \overline{P} has no ties on the singletons:

$$\begin{aligned} \overline{P}(\{x_{\sigma^*(1)}\}) &< \overline{P}(\{x_{\sigma^*(2)}\}) < \dots < \overline{P}(\{x_{\sigma^*(n)}\}) \\ \Rightarrow \Pi(\{x_{\sigma^*(1)}\}) &\leq \Pi(\{x_{\sigma^*(2)}\}) \leq \dots \leq \Pi(\{x_{\sigma^*(n)}\}). \end{aligned}$$

This means that Π satisfies WPP on singletons; as we shall see in Example 10 later on, it may not satisfy WPP on events.

This result may be a motivation for selecting order basic permutations when outer approximating \overline{P} with a possibility measure. Yet, doing so does not guarantee WPP on all events of the power set of \mathcal{X} , nor on the atomic events of \mathcal{X} if \overline{P} has ties there. More generally, it is possible that, whatever permutation is chosen, order basic or not, WPP is violated for at least a couple of events in $\mathcal{P}(\mathcal{X})$. Our next example illustrates this.

Example 9. Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and the upper probability \overline{P} given in the following table. It is coherent, being the upper envelope of the probability measures P_1, P_2 . We also depict in Table 2 the possibility measures determined by Eqs. (20)–(22) for the different permutations.

| | $\{x_1\}$ | $\{x_2\}$ | $\{x_3\}$ | $\{x_1, x_2\}$ | $\{x_1, x_3\}$ | $\{x_2, x_3\}$ |
|-----------------|-----------|-----------|-----------|----------------|----------------|----------------|
| P_1 | 0.5 | 0.25 | 0.25 | 0.75 | 0.75 | 0.5 |
| P_2 | 0.4 | 0.3 | 0.3 | 0.7 | 0.7 | 0.6 |
| \overline{P} | 0.5 | 0.3 | 0.3 | 0.75 | 0.75 | 0.6 |
| $\Pi_{(1,2,3)}$ | 0.5 | 0.75 | 1 | 0.75 | 1 | 1 |
| $\Pi_{(1,3,2)}$ | 0.5 | 1 | 0.75 | 1 | 0.75 | 1 |
| $\Pi_{(2,1,3)}$ | 0.75 | 0.3 | 1 | 0.75 | 1 | 1 |
| $\Pi_{(2,3,1)}$ | 1 | 0.3 | 0.6 | 1 | 1 | 0.6 |
| $\Pi_{(3,1,2)}$ | 0.75 | 1 | 0.3 | 1 | 0.75 | 1 |
| $\Pi_{(3,2,1)}$ | 1 | 0.6 | 0.3 | 1 | 1 | 0.6 |

We see for instance that the possibility measure originated by the permutation $\sigma = (1, 2, 3)$ is at conflict with WPP at (x_1, x_2) : we have that $\overline{P}(\{x_1\}) > \overline{P}(\{x_2\})$ while $\Pi_{(1,2,3)}(\{x_1\}) < \Pi_{(1,2,3)}(\{x_2\})$. Similar violations can be observed for the other outer approximating possibility measures in the table above.

In this case, the order basic possibility measures are the ones associated with the permutations $(2, 3, 1)$ and $(3, 2, 1)$; they violate WPP with $A = \{x_2\}$ and $B = \{x_3\}$: we have that $\overline{P}(A) = \overline{P}(B)$ while $\Pi(A) \neq \Pi(B)$ in both cases. ♦

A procedure to outer approximate an upper probability with a possibility measure that guarantees WPP was introduced by Baroni and Vicig in [2, Section 6.3]; let us call it BV-procedure. It determines a unique possibility measure that outer approximates \bar{P} , satisfies WPP and is non-dominating in the subclass of \mathcal{C}_Π satisfying WPP [2, Proposition 10]⁵. However, while offering a solution to the WPP problem, the BV-procedure may return a possibility measure that is dominating in the class of outer approximations in \mathcal{C}_Π , as we can see applying Corollary 13 to Example 9, where all the non-dominating outer approximations of \bar{P} violate WPP.

We conclude thus that a trade-off arises here as for the kind of information we believe the outer approximating possibility should retain. If (weakly) preserving the ordering given by \bar{P} is the most important feature, the BV-procedure should be used. If being numerically as close as possible to \bar{P} is what mostly matters, the procedure of Proposition 11 should be applied. It is possible to achieve both goals in some, but not all instances. In this respect, it is interesting to see that, unlike Example 9, order basic possibility measures may not minimize the distance with respect to the original coherent upper probability:

Example 10. Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and let \bar{P} be the coherent upper probability that is the upper envelope of the probability measures with the mass functions $(0.3, 0.2, 0.5)$, $(0.31, 0.45, 0.24)$, $(0.4, 0.3, 0.3)$. Its values are depicted in the table below.

| A | $\bar{P}(A)$ | $\Pi_\sigma(A)$ | $\Pi_{\sigma^*}(A)$ |
|----------------|--------------|-----------------|---------------------|
| $\{x_1\}$ | 0.4 | 0.4 | 1 |
| $\{x_2\}$ | 0.45 | 0.76 | 0.45 |
| $\{x_3\}$ | 0.5 | 1 | 0.7 |
| $\{x_1, x_2\}$ | 0.76 | 0.76 | 1 |
| $\{x_1, x_3\}$ | 0.8 | 1 | 1 |
| $\{x_2, x_3\}$ | 0.7 | 1 | 0.7 |

Since $\bar{P}(\{x_1\}) < \bar{P}(\{x_2\}) < \bar{P}(\{x_3\})$, the only order basic permutation is $\sigma = (1, 2, 3)$. Its associated possibility measure is given in the table above. However, if we consider $\sigma^* = (2, 3, 1)$ we obtain the possibility measure Π_{σ^*} in the same table, and it holds that

$$d(\Pi_{\sigma^*}, \bar{P}) = 1.24 < 1.31 = d(\Pi, \bar{P}).$$

Thus, the minimal distance is not always attained by an order basic possibility measure. \blacklozenge

6.2. Outer approximations using p-boxes. As we mentioned in Section 2.5, a p-box is a particular case of belief function where the focal events are ordered intervals. In our next result we assume that the possibility space

⁵The procedure was more generally devised to outer approximate a coherent upper probability defined on a subset $\mathcal{K} \subseteq \mathcal{P}(\mathcal{X})$ by means of a possibility measure on the same domain.

\mathcal{X} is totally ordered and derive an optimal outer approximation by means of p-boxes.

Proposition 15. *Let \mathcal{X} be a totally ordered space, with $x_1 \leq \dots \leq x_n$ and let \underline{P} be a coherent lower probability on $\mathcal{P}(\mathcal{X})$. Define the p-box $(\underline{F}, \overline{F})$ by:*

$$\underline{F}(x_i) = \underline{P}(\{x_1, \dots, x_i\}), \quad \overline{F}(x_i) = \overline{P}(\{x_1, \dots, x_i\}) \quad \forall x_i \in \mathcal{X}. \quad (27)$$

Then, the belief function $\underline{P}_{(\underline{F}, \overline{F})}$ this p-box determines via Eq. (7) is the unique undominated outer approximation of \underline{P} in $\mathcal{C}_{(\underline{F}, \overline{F})}$.

Proof. The inequality $\underline{P}_{(\underline{F}, \overline{F})} \leq \underline{P}$ follows from [33, Theorem 2]. Let us prove next that $\underline{P}_{(\underline{F}, \overline{F})}$ is an undominated outer approximation. Indeed, if $(\underline{F}', \overline{F}')$ is a p-box satisfying $\underline{P}_{(\underline{F}, \overline{F})} \leq \underline{P}_{(\underline{F}', \overline{F}')} \leq \underline{P}$, for every $i \in \{1, \dots, n\}$ it holds that:

$$\underline{P}_{(\underline{F}', \overline{F}')}(\{x_1, \dots, x_i\}) = \underline{P}_{(\underline{F}, \overline{F})}(\{x_1, \dots, x_i\}) = \underline{P}(\{x_1, \dots, x_i\}),$$

whence $\underline{P}_{(\underline{F}, \overline{F})}$ and $\underline{P}_{(\underline{F}', \overline{F}')}$ coincide on the events $\{x_1, \dots, x_i\}$ for every $i \in \{1, \dots, n\}$. Applying Eq. (6), $\mathcal{M}(\underline{F}, \overline{F}) = \mathcal{M}(\underline{F}', \overline{F}')$ and therefore $\underline{P}_{(\underline{F}, \overline{F})} = \underline{P}_{(\underline{F}', \overline{F}')}$.

To see that $\underline{P}_{(\underline{F}, \overline{F})}$ is the unique undominated outer approximation of \underline{P} in $\mathcal{C}_{(\underline{F}, \overline{F})}$, note that any p-box $(\underline{F}', \overline{F}')$ outer approximating \underline{P} must satisfy $\underline{F}' \leq \underline{F} \leq \overline{F} \leq \overline{F}'$. From Eq. (6) this means that $\mathcal{M}(\underline{F}, \overline{F}) \subseteq \mathcal{M}(\underline{F}', \overline{F}')$. \square

Since $\mathcal{C}_{(\underline{F}, \overline{F})} \subset \mathcal{C}_\infty$ (for $n \geq 3$), the outer approximation given in the previous result is also an outer approximation that belongs to the broader class \mathcal{C}_∞ . However, it may not be an undominated outer approximation in \mathcal{C}_∞ : it suffices to outer approximate a belief function that is not induced by a p-box. This is also illustrated by our next example, that shows that the unique outer approximation may not coincide with the initial model on singletons.

Example 11. *Consider again the lower probability in Example 1, and assume that the total ordering in \mathcal{X} is given by $x_1 \leq x_2 \leq x_4 \leq x_3$. Then, the p-box defined from Eq. (27) is*

| | | | | |
|-----------------|-------|-------|-------|-------|
| | x_1 | x_2 | x_4 | x_3 |
| \underline{F} | 0.1 | 0.1 | 0.6 | 1 |
| \overline{F} | 0.4 | 0.6 | 1 | 1 |

In Figure 2 we have depicted this p-box emphasizing its focal events.

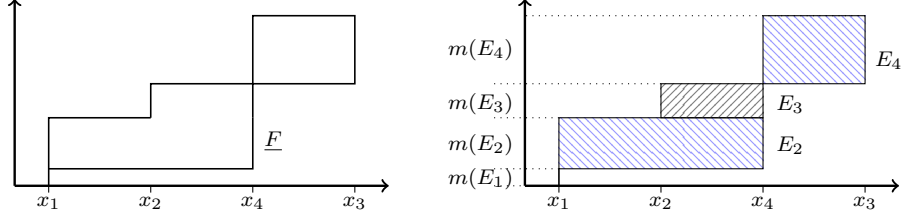


FIGURE 2. P-box (left) from Example 11 and its associated belief function (right), with focal events $E_1 = \{x_1\}$, $E_2 = \{x_1, x_2, x_4\}$, $E_3 = \{x_2, x_4\}$ and $E_4 = \{x_4, x_3\}$.

Using the results in [13], the belief function $\underline{P}_{(\underline{F}, \overline{F})}$ associated with $(\underline{F}, \overline{F})$ is given by:

| A | $\underline{P}_{(\underline{F}, \overline{F})}(A)$ |
|---------------------|--|
| $\{x_1\}$ | 0.1 |
| $\{x_2\}$ | 0 |
| $\{x_3\}$ | 0 |
| $\{x_4\}$ | 0 |
| $\{x_1, x_2\}$ | 0.1 |
| $\{x_1, x_3\}$ | 0.1 |
| $\{x_1, x_4\}$ | 0.1 |
| $\{x_2, x_3\}$ | 0 |
| $\{x_2, x_4\}$ | 0.2 |
| $\{x_3, x_4\}$ | 0.4 |
| $\{x_1, x_2, x_3\}$ | 0.1 |
| $\{x_1, x_2, x_4\}$ | 0.6 |
| $\{x_1, x_3, x_4\}$ | 0.5 |
| $\{x_2, x_3, x_4\}$ | 0.6 |
| \mathcal{X} | 1 |

We see then that $\underline{P}_{(\underline{F}, \overline{F})} \leq Bel_2$, where Bel_2 from Example 1 is an undominated outer approximation of \underline{P} in \mathcal{C}_∞ , and the inequality is strict for some events, such as $\{x_4\}$. \blacklozenge

While in Proposition 15 we have assumed that the possibility space \mathcal{X} is endowed with a total order, this may not always be the case. One alternative would be then to consider the outer approximations in terms of generalized p-boxes. As we mentioned in Section 2.5, a generalized p-box $(\underline{F}, \overline{F})$ induces a permutation σ and an order $\leq_{(\underline{F}, \overline{F})}$ such that $x_{\sigma(1)} \leq_{(\underline{F}, \overline{F})} \dots \leq_{(\underline{F}, \overline{F})} x_{\sigma(n)}$. Hence, looking for generalized p-boxes outer approximating \underline{P} is equivalent to looking for a p-box $(\underline{F}, \overline{F})$ and a permutation σ of \mathcal{X} so that:

$$x_{\sigma(1)} \leq_{\sigma} \dots \leq_{\sigma} x_{\sigma(n)}. \quad (28)$$

If we then apply Proposition 15, we obtain a p-box $(\underline{F}_\sigma, \overline{F}_\sigma)$, given by:

$$\begin{aligned}\underline{F}_\sigma(x_{\sigma(i)}) &= \underline{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}) \quad \text{and} \\ \overline{F}_\sigma(x_{\sigma(i)}) &= \overline{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}) \quad \forall i \in \{1, \dots, n\},\end{aligned}\tag{29}$$

whose associated belief function outer approximates the initial coherent lower probability \underline{P} . Moreover, by considering all the generalized p-boxes determined by these orders we can recover \underline{P} :

Proposition 16. *Let \underline{P} be a coherent lower probability. For every permutation $\sigma \in S_n$ define the total order \leq_σ as in Eq. (28), and let $(\underline{F}_\sigma, \overline{F}_\sigma)$ be the generalized p-box defined as in Eq. (29). Then:*

$$\underline{P}(A) = \max_{\sigma \in S_n} \underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)}(A) \quad \forall A \subseteq \mathcal{X}.$$

Equivalently, if $\mathcal{M}(\underline{P})$ and $\mathcal{M}(\underline{F}_\sigma, \overline{F}_\sigma)$ denote the credal sets of \underline{P} and $\underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)}$, respectively, it holds that:

$$\mathcal{M}(\underline{P}) = \bigcap_{\sigma \in S_n} \mathcal{M}(\underline{F}_\sigma, \overline{F}_\sigma).$$

Proof. First of all, from Proposition 15, $\underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)}$ is an outer approximation of \underline{P} for every $\sigma \in S_n$, whence $\underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)} \leq \underline{P}$ and also:

$$\max_{\sigma \in S_n} \underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)} \leq \underline{P}.$$

In order to prove the equality, we are going to establish that for every $A \subseteq \mathcal{X}$, there exists a permutation $\sigma \in S_n$ such that $\underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)}(A) = \underline{P}(A)$. For this aim, given $A \subseteq \mathcal{X}$, take a permutation σ such that $A = \{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}$. By construction, the p-box defined from the total ordering \leq_σ satisfies

$$\underline{F}(x_{\sigma(k)}) = \underline{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}),$$

so

$$\underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)}(A) = \underline{P}_{(\underline{F}_\sigma, \overline{F}_\sigma)}(\{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}) = \underline{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}) = \underline{P}(A).$$

This completes the proof. \square

Finally, we prove that all undominated outer approximations in $\mathcal{C}_{(\underline{F}, \overline{F})}^*$ are of the form $(\underline{F}_\sigma, \overline{F}_\sigma)$.

Theorem 17. *Let \underline{P} be a coherent lower probability and let $(\underline{F}, \overline{F})$ be an undominated outer approximation of \underline{P} in $\mathcal{C}_{(\underline{F}, \overline{F})}^*$. Then, there is a permutation σ such that $(\underline{F}, \overline{F}) = (\underline{F}_\sigma, \overline{F}_\sigma)$.*

Proof. The generalized p-box $(\underline{F}, \overline{F})$ has an associated order on \mathcal{X} . Assume that the order corresponds to the permutation $\sigma \in S_n$ such that:

$$x_{\sigma(1)} \leq_\sigma \dots \leq_\sigma x_{\sigma(n)}.$$

For this permutation σ , define the p-box $(\underline{F}_\sigma, \overline{F}_\sigma)$ as in Eq. (29). On the one hand, for every $i \in \{1, \dots, n\}$, it holds that:

$$\underline{F}_\sigma(x_{\sigma(i)}) = \underline{P}(\{x_{\sigma(1)}, \dots, x_{\sigma(i)}\}) \geq \underline{F}(x_{\sigma(i)}),$$

and similarly, $\overline{F}_\sigma \leq \overline{F}$. By Eq. (6), $\mathcal{M}(\underline{F}_\sigma, \overline{F}_\sigma) \subseteq \mathcal{M}(\underline{F}, \overline{F})$; since on the other hand $(\underline{F}, \overline{F})$ is an undominated outer approximation by assumption, we conclude that $(\underline{F}, \overline{F})$ and $(\underline{F}_\sigma, \overline{F}_\sigma)$ coincide. \square

7. CONCLUSIONS

In this paper, we have investigated the problem of outer approximating a coherent lower probability by means of a belief function with a minimal loss of information. We have considered those belief functions that minimize the distance proposed by Baroni and Vicig in [2], and showed that, while the problem always has a solution, it may not be unique. In addition, the structure of the set of undominated solutions is not straightforward: for instance, we have seen that this set is generally not convex (Example 3) and it may be formed by infinitely many belief functions (Example 1). For these reasons, we have also considered three particular cases for which we can determine the set of all undominated outer approximations and that are also computationally simpler: ternary possibility spaces, possibility measures and p-boxes.

By comparing our results with those in [27], it becomes clear that the use of belief functions, while having some computational advantages over 2-monotone lower probabilities, also entails some loss of information. To illustrate this, recall that, from Example 1, an undominated outer approximation in the class of belief functions may be dominated in the larger class of 2-monotone lower probabilities. In other words, the inclusion $\mathcal{C}_\infty \subset \mathcal{C}_2$ does not imply that \mathcal{C}_∞ is a subset of the set of undominated 2-monotone outer approximations.

The following table summarizes some of the results we have obtained in this paper and in our previous contribution [27]. We see for instance that the uniqueness of the solution is obviously incompatible with the ability to retrieve the initial model. Moreover, the use of more restrictive families (a) improves the mathematical structure of the set of solutions and decreases the computational cost; but (b) also entails a loss of information that can be avoided with the more general models.

| Outer approximation in the class \mathcal{C} | Unique undominated outer approximation? | The undominated outer approximation(s) retrieve \underline{P} ? | Characterization of the undominated outer approximation(s) | Coincide with $\underline{P}, \overline{P}$ on singletons? |
|--|---|---|--|--|
| \mathcal{C}_∞ | NO | YES (Prop. 4) | Open problem | NO (Ex. 4) |
| $\mathcal{C}_{(E, \overline{F})}$ | YES (Prop. 15) | NO | YES (Prop. 15) | NO (Ex. 11) |
| $\mathcal{C}_{(E, \overline{F})}^*$ | NO | YES (Prop. 16) | YES (Thm. 17) | NO (Ex. 11) |
| \mathcal{C}_Π | NO | YES (Prop.12) | YES (Cor. 13) | NO (Ex. 8) |
| \mathcal{C}_2 | NO | YES (Prop. 2) | Open problem | YES ([27]) |
| Prob. intervals | YES ([27]) | NO | YES ([27]) | YES ([27]) |
| Distortion models | YES ([27]) | NO | YES ([27]) | PARTIALLY ([27]) |

Finally, let us mention that in this paper we investigate how to outer approximate a coherent lower probability by means of a belief function. As was proved in [27, Theorem 1], if we want to outer approximate a coherent lower prevision, we just need to consider its restriction to events, which is a coherent lower probability, outer approximate it using some of the procedures presented in this paper, and lastly extend it to gambles using the Choquet integral as in Eq. (4). This is because there is a one-to-one correspondence between the undominated outer approximations and the ones we obtain if we consider the problem on the restrictions to events, as we have done in this paper.

Although the results are promising, there are a few points still pending, that we would like to tackle in the near future: first and foremost, we would like to study how to choose among the different optimal outer approximations, when there is more than one. This could be done, for instance, by comparing them by means of other distances to the original model, or by other tools such as specificity measures, measures of information, entropies, etc. [6, 7, 15, 38]. We could also investigate how the loss of information entailed by the use of outer approximations propagates when making inferences with the credal sets. In addition, we may also compare the undominated outer approximations obtained via linear programming by means of other distances, such as the quadratic or the L2-distance; and it might be worth considering distances with different weights on the differences on the subsets, depending on their cardinality.

There are also other transformations between imprecise probability models that may be interesting; as we mentioned, Dubois and Prade studied the transformation of belief functions into possibility measures in [14]; we could also analyze if our results provide some advantages in this respect. In particular, if we focus on belief functions we may consider several other distance measures [22]; the choice between them may be made in terms of their compatibility with some order between belief models, in the manner discussed in [23, Section 3.3]; see also Proposition 5.

We may also follow the converse path and look for *inner* approximations, that is, more informative models than a coherent lower probability that satisfy some desirable properties. This is in line with the procedure of natural extension of Walley [36] as a minimal correction procedure, but has the

drawback that, unlike with the problem we have considered in this paper, in some cases the inner approximations may not exist. Some early results were established in [27], where we studied how to inner approximate a coherent lower probability in \mathcal{C}_2 , and in [20], where inner approximations in the family of k -additive measures were considered. Other relevant results may be found in [2, 18, 25, 34].

ACKNOWLEDGEMENTS

The research in this paper has been supported by project TIN2014-59543-P. We would like to thank Sébastien Destercke, Serafin Moral, John Klein, Anne-Laure Jousselme and the anonymous reviewers for some helpful comments. This paper presents an extension, with proofs and additional results, of a previous work [28] presented at the BELIEF'2018 conference.

REFERENCES

- [1] T. Augustin, F. Coolen, G. de Cooman, and M. Troffaes, editors. *Introduction to Imprecise Probabilities*. Wiley Series in Probability and Statistics. Wiley, 2014.
- [2] P. Baroni and P. Vicig. An uncertainty interchange format with imprecise probabilities. *International Journal of Approximate Reasoning*, 40:147–180, 2005.
- [3] A. Bronevich and T. Augustin. Approximation of coherent lower probabilities by 2-monotone measures. In Thomas Augustin, Frank P. A. Coolen, Serafin Moral, and Matthias C. M. Troffaes, editors, *ISIPTA '09 – Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications*, pages 61–70, Durham, United Kingdom, 2009. SIPTA.
- [4] 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.
- [5] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953–1954.
- [6] L. M. de Campos. *Caracterización y estudio de medidas e integrales difusas a partir de probabilidades*. PhD thesis, University of Granada, 1988.
- [7] L.M. de Campos, M.T. Lamata, and S. Moral. Distances between fuzzy measures through associated probabilities: some applications. *Fuzzy Sets and Systems*, 35:57–68, 1990.
- [8] G. de Cooman, M. C. M. Troffaes, and E. Miranda. n -Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347:143–156, 2008.
- [9] B. de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.
- [10] B. de Finetti. *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons, Chichester, 1974–1975. English translation of [9], two volumes.
- [11] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
- [12] S. Destercke. Independence and 2-monotonicity: Nice to have, hard to keep. *International Journal of Approximate Reasoning*, 54(4):478–490, 2013.
- [13] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: I. Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.
- [14] D. Dubois and H. Prade. Fuzzy sets and statistical data. *European Journal of Operational Research*, 25(3):345–356, 1986.
- [15] D. Dubois and H. Prade. Properties of measures of information in evidence and possibility theories. *Fuzzy Sets and Systems*, 24:161–182, 1987.

- [16] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.
- [17] D. Dubois and H. Prade. Consonant approximations of belief functions. *International Journal of Approximate Reasoning*, 4(5-6):419–449, 1990.
- [18] D. Dubois and H. Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49:65–74, 1992.
- [19] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002–4015, Sandia National Laboratories, January 2003.
- [20] M. Grabisch. Upper approximation of non-additive measures by k -additive measures - The case of belief functions. In *Proceedings of 1st International Symposium on Imprecise Probabilities and Their Applications*, pages 158–164. ISIPTA, 1999.
- [21] J. W. Hall and J. Lawry. Generation, combination and extension of random set approximations to coherent lower and upper probabilities. *Reliability Engineering & System Safety*, 85:89–101, 2004.
- [22] A-L. Jousselme and P. Maupin. Distances in evidence theory: comprehensive survey and generalizations. *International Journal of Approximate Reasoning*, 53(2):118–145, 2012.
- [23] J. Klein, S. Destercke, and O. Colot. Interpreting evidential distances by connecting them to partial orders: application to belief function approximation. *International Journal of Approximate Reasoning*, 71:15–33, 2016.
- [24] I. Levi. *The enterprise of knowledge*. MIT Press, Cambridge, 1980.
- [25] E. Miranda and I. Montes. Shapley and Banzhaf values as probability transformations. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 26/6, 917–947, 2018.
- [26] I. Montes, S. Destercke. Extreme points of p-boxes and belief functions. *Annals of Mathematics and Artificial Intelligence*, 81/3, 405–428, 2017.
- [27] I. Montes, E. Miranda, and P. Vicig. 2-monotone outer approximations of coherent lower probabilities. *International Journal of Approximate Reasoning*, 101:181–205, 2018.
- [28] I. Montes, E. Miranda, and P. Vicig. Outer approximations of coherent lower probabilities using belief functions. In S. Destercke, T. Denoeux, F. Cuzzolin, and A. Martin, editors, *Belief Functions: Theory and Applications*, volume 11069 of *Lecture Notes in Artificial Intelligence. Subseries of Lecture Notes in Computer Science*, pages 190–198. Springer, 2018.
- [29] H. T. Nguyen. *An introduction to random sets*. Chapman and Hall, 2006.
- [30] E. Quaeghebeur. Completely monotone outer approximations of lower probabilities on finite possibility spaces. In Okazaki Y. Kawabe J. Murofushi T. Guan L. Li S., Wang X., editor, *Nonlinear Mathematics for Uncertainty and its Applications. Advances in Intelligent and Soft Computing*, volume 100. Springer, 2011.
- [31] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey, 1976.
- [32] L. S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1:11–26, 1971.
- [33] M. C. M. Troffaes and S. Destercke. Probability boxes on totally preordered spaces for multivariate modelling. *International Journal of Approximate Reasoning*, 52(6):767–791, 2011.
- [34] F. Voorbraak. A computationally efficient approximation of Dempster Shafer theory. *International Journal of Man-Machine Studies*, 30:525–536, 1989.
- [35] P. Walley. Coherent lower (and upper) probabilities. Statistics Research Report 22, University of Warwick, Coventry, 1981.
- [36] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

- [37] A. Wallner. Extreme points of coherent probabilities in finite spaces. *International Journal of Approximate Reasoning*, 44(3):339–357, 2007.
- [38] R.R. Yager. Entropy and specificity in a mathematical theory of evidence. *International Journal of General Systems*, 9:249–260, 1983.
- [39] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.

UNIVERSITY OF OVIEDO (SPAIN), DEPARTMENT OF STATISTICS AND OPERATIONS
RESEARCH

Email address: (imontes,mirandaenrique)@uniovi.es

UNIVERSITY OF TRIESTE (ITALY), DEAMS

Email address: paolo.vicig@deams.units.it