On the elicitation of an optimal outer approximation of a coherent lower probability

Enrique Miranda[1][0000−0001−7763−3779]⋆, Ignacio Montes[1][0000−0001−6534−1613], and Paolo Vicig[2][0000−0001−8842−442X]

[1] University of Oviedo, Spain
mirandaenrique@uniovi.es, imontes@uniovi.es
[2] University of Trieste, Italy
paolo.vicig@deams.units.it

Abstract. The process of outer approximating a coherent lower probability by a more tractable model with additional properties, such as 2- or completely monotone capacities, may not have a unique solution. In this paper, we investigate whether a number of approaches may help in eliciting a unique outer approximation: minimising a number of distances with respect to the initial model, or maximising the specificity of the outer approximation. We apply these to 2- and completely monotone approximating lower probabilities, and also to possibility measures.

Keywords: Coherent lower probabilities · 2-monotonicity · Belief functions · Possibility measures · Specificity.

1 Introduction

The theory of imprecise probabilities [13] encompasses the different models that may be used as an alternative to probability theory in situations of imprecise or ambiguous information. Among them, we can find credal sets [7], coherent lower probabilities [13], belief functions [11] or possibility measures [15].

Within imprecise probabilities, one of the most general models are coherent lower and upper probabilities. However, this generality is at times harmed by the difficulties that arise when using them in practice. For example, there is no simple procedure for computing the extreme points of its associated credal set, and there is no unique coherent extension to gambles. These problems are solved when the coherent lower probability satisfies the additional property of 2-monotonicity [4, 12], or that of complete monotonicity.

For this reason, in previous papers [9, 10] we investigated the problem of transforming a coherent lower probability into a 2-monotone one that does not add information to the model while being as close as possible to it. This led us to the notion of undominated outer approximations, formerly introduced in [2]. In [9] we analysed the properties of the 2-monotone outer approximations, while in [10] we studied the completely monotone ones, considering in particular the

---

⋆ Corresponding author. C-Federico García Lorca, 18. 33007 Oviedo, Spain

outer approximations in terms of necessity measures. In both cases, we found out that there may be an infinity of undominated outer approximations, and that their computation may be quite involved. Nevertheless, in the case of necessity measures, we proved that there are a finite number of undominated ones and we introduced a procedure for determining them.

Since in any case there is not a unique undominated outer approximation in terms of 2- or completely monotone lower probabilities or even in terms of necessity measures, in this paper we explore a number of possibilities that may help single out a unique undominated outer approximation. After introducing some preliminary notions in Sec. 2, formalising the idea of outer approximation and summarising the main properties from [9, 10] in Sec. 3, in Sec. 4 and 5 we introduce and compare a number of different procedures to elicit an undominated outer approximation. We conclude the paper in Sec. 6 summarising the main contributions of the paper. Due to space limitations, proofs have been omitted.

2   Imprecise probability models

Consider an experiment taking values in a finite possibility space $\mathcal{X} = \{x_1, \ldots, x_n\}$. A *lower probability* on $\mathcal{P}(\mathcal{X})$ is a monotone function $\underline{P} : \mathcal{P}(\mathcal{X}) \to [0,1]$ satisfying $\underline{P}(\emptyset) = 0, \underline{P}(\mathcal{X}) = 1$. For every $A \subseteq \mathcal{X}$, $\underline{P}(A)$ is interpreted as a lower bound for the true (but unknown) probability of A. Any lower probability determines the *credal set* of probability measures that are compatible with it, given by $\mathcal{M}(\underline{P}) = \{P \mid P(A) \geq \underline{P}(A) \ \ \forall A \subseteq \mathcal{X}\}$. We say that $\underline{P}$ *avoids sure loss* when $\mathcal{M}(\underline{P})$ is non-empty, and that it is *coherent* when $\underline{P}(A) = \min\{P(A) \mid P \in \mathcal{M}(\underline{P})\}$ for every $A \subseteq \mathcal{X}$.

Associated with $\underline{P}$, we can consider its conjugate *upper probability*, given by $\overline{P}(A) = 1 - \underline{P}(A^c)$ for every $A \subseteq \mathcal{X}$. The value $\overline{P}(A)$ may be interpreted as an upper bound for the unknown probability of A, and it follows that $P \geq \underline{P}$ if and only if $P \leq \overline{P}$. This means that the probabilistic information given by a lower probability and its conjugate upper probability are equivalent, and so it suffices to work with one of them.

A coherent lower probability $\underline{P}$ is *k-monotone* if for every $1 \leq p \leq k$ and $A_1, \ldots, A_p \subseteq \mathcal{X}$ it satisfies $\underline{P}\big(\cup_{i=1}^p A_i\big) \geq \sum_{\emptyset \neq I \subseteq \{1, \ldots, p\}} (-1)^{|I|+1} \underline{P}\big(\cap_{i \in I} A_i\big)$. Of particular interest are the cases of *2-monotonicity* and *complete monotonicity*; the latter refers to those lower probabilities that are k-monotone for every k.

Any lower probability $\underline{P}$ can be represented in terms of a function called *Möbius inverse*, denoted by $m_{\underline{P}} : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$, and defined by:

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

Conversely, $m_{\underline{P}}$ allows to retrieve $\underline{P}$ using the expression $\underline{P}(A) = \sum_{B \subseteq A} m_{\underline{P}}(B)$. Moreover, $m_{\underline{P}}$ is the Möbius inverse associated with a 2-monotone lower proba-

bility $\underline{P}$ if and only if [3] $m_{\underline{P}}$ satisfies:

$$\sum_{A \subseteq \mathcal{X}} m_{\underline{P}}(A) = 1, \quad m_{\underline{P}}(\emptyset) = 0, \quad m_{\underline{P}}(\{x_i\}) \geq 0 \ \ \forall x_i \in \mathcal{X}, \qquad \text{(2monot.1)}$$

$$\sum_{\{x_i,x_j\} \subseteq B \subseteq A} m_{\underline{P}}(B) \geq 0, \quad \forall A \subseteq \mathcal{X}, \forall x_i, x_j \in A, \ x_i \neq x_j, \qquad \text{(2monot.2)}$$

while it is associated with a completely monotone lower probability if and only if [11] $m_{\underline{P}}$ satisfies:

$$\sum_{A \subseteq \mathcal{X}} m_{\underline{P}}(A) = 1, \quad m_{\underline{P}}(\emptyset) = 0, \quad m_{\underline{P}}(A) \geq 0 \ \ \forall A \subseteq \mathcal{X}. \qquad \text{(Cmonot.1)}$$

Completely monotone lower probabilities are also connected to Dempster-Shafer Theory of Evidence [11], where they are called *belief functions*. In that case, the events with strictly positive mass are called *focal events*.

Another usual imprecise model is that of necessity and possibility measures. A possibility measure [6, 15], denoted by Π, is a supremum-preserving function:

$$\Pi(\cup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i), \quad \forall A_i \subseteq \mathcal{X}, \ i \in I.$$

In our finite framework, the above condition is equivalent to $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$ for every $A, B \subseteq \mathcal{X}$. Every possibility measure is a coherent upper probability. Its conjugate lower probability, denoted by N and called *necessity measure*, is a completely monotone lower probability and its focal events are nested.

3   Outer approximations of coherent lower probabilities

Even if coherent lower probabilities are more general than 2-monotone ones, the latter have some practical advantages. For example, they can be easily extended to gambles [4] and the structure of their credal set can be easily determined [12]. Motivated by this, in [9] we proposed to replace a given coherent lower probability by a 2-monotone one that does not add information to the model while being as close as possible to the initial model. The first condition gives rise to the notion of *outer approximation*, and the second leads to the notion of *undominated* approximations. These concepts were formalised by Bronevich and Augustin [2]:

Definition 1. *Given a coherent lower probability $\underline{P}$ and a family $\mathcal{C}$ of coherent lower probabilities, $\underline{Q}$ is an* outer approximation *(OA, for short) of $\underline{P}$ if $\underline{Q} \leq \underline{P}$. Moreover, $\underline{Q}$ is* undominated *in $\mathcal{C}$ if there is no $\underline{Q}' \in \mathcal{C}$ such that $\underline{Q} \lneq \underline{Q}' \leq \underline{P}$.*

Similarly, given a coherent upper probability $\overline{P}$ and a family $\mathcal{C}$ of coherent upper probabilities, $\overline{Q} \in \mathcal{C}$ is an *outer approximation* of $\overline{P}$ if $\overline{Q}(A) \geq \overline{P}(A)$ for every $A \subseteq \mathcal{X}$, and it is called *non-dominating* in $\mathcal{C}$ if there is no $\overline{Q}' \in \mathcal{C}$ such that

$\overline{Q} \gneqq \overline{Q}' \geq \overline{P}$. It follows that $\overline{Q}$ is an outer approximation of $\overline{P}$ if and only if its conjugate $\underline{Q}$ is an outer approximation of the lower probability $\underline{P}$ conjugate of $\overline{P}$.

Let us consider now the families $\mathcal{C}_2, \mathcal{C}_\infty$ and $\mathcal{C}_\Pi$ of 2- and completely monotone lower probabilities and possibility measures. In [9, 10] we investigated several properties of the undominated (non-dominating for $\mathcal{C}_\Pi$) outer approximations in these families. We showed that determining the set of undominated OA in $\mathcal{C}_2$ and $\mathcal{C}_\infty$ is not immediate, and that these sets are infinite in general. The problem is somewhat simpler for the outer approximations in $\mathcal{C}_\Pi$, even if in this case there is not a unique non-dominating OA either. In this paper, we discuss different procedures to elicit a unique OA.

Before we proceed, let us remark that we may assume without loss of generality that all singletons have strictly positive upper probability.

Proposition 1. *Let $\overline{P}, \overline{Q} : \mathcal{P}(\mathcal{X}) \to [0, 1]$ be two coherent upper probabilities such that $\overline{P} \leq \overline{Q}$. Assume that $\overline{P}(\{x\}) = 0 < \overline{Q}(\{x\})$ for a given $x \in \mathcal{X}$, and let us define $\overline{Q}' : \mathcal{P}(\mathcal{X}) \to [0, 1]$ by $\overline{Q}'(A) = \overline{Q}(A \setminus \{x\})$ for every $A \subseteq \mathcal{X}$. Then:*

1. *$\overline{P} \leq \overline{Q}' \lneqq \overline{Q}$.*
2. *If $\overline{Q}$ is k-alternating, so is $\overline{Q}'$.*
3. *If $\overline{Q}$ is a possibility measure, so is $\overline{Q}'$.*

The proposition above allows us to deduce the following:

Corollary 1. *Let $\overline{P} : \mathcal{P}(\mathcal{X}) \to [0, 1]$ be a coherent upper probability and let $\overline{Q}$ be a non-dominating outer approximation of $\overline{P}$ in $\mathcal{C}_2$, $\mathcal{C}_\infty$ or $\mathcal{C}_\Pi$. If $\overline{P}(\{x\}) = 0$, then also $\overline{Q}(\{x\}) = 0$.*

As a consequence, we may assume without loss of generality that $\overline{P}(\{x\}) > 0$ for every $x \in \mathcal{X}$. This is relevant for the proofs of the results in Section 5.

4   Elicitation of an outer approximation in $\mathcal{C}_2$ and $\mathcal{C}_\infty$

From [9, 10], the number of undominated OAs in $\mathcal{C}_2$ and $\mathcal{C}_\infty$ is not finite in general. In [9, 10] we focused on those undominated OAs in $\mathcal{C}_2$ and $\mathcal{C}_\infty$ that minimise the BV-distance proposed in [1] with respect to the original coherent lower probability $\underline{P}$, given by $d_{BV}(\underline{P}, \underline{Q}) = \sum_{A \subseteq \mathcal{X}} |\underline{P}(A) - \underline{Q}(A)|$. It measures the amount of imprecision added to the model when replacing $\underline{P}$ by its OA $\underline{Q}$. Hence, its seems reasonable to minimise the imprecision added to the model.

Let $\mathcal{C}_2^{BV}(\underline{P})$ and $\mathcal{C}_\infty^{BV}(\underline{P})$ denote the set of undominated OAs in $\mathcal{C}_2$ and $\mathcal{C}_\infty$, respectively, that minimise the BV-distance with respect to $\underline{P}$. One advantage of focusing our elicitation to $\mathcal{C}_2^{BV}(\underline{P})$ and $\mathcal{C}_\infty^{BV}(\underline{P})$ is that these sets can be easily determined. Indeed, both $\mathcal{C}_2^{BV}(\underline{P})$ and $\mathcal{C}_\infty^{BV}(\underline{P})$ can be computed as the set of optimal solutions of a linear programming problem ([9, Prop. 1], [10, Prop. 3]). Hence, both sets are convex, and have an infinite number of elements in general. In the rest of the section we discuss different approaches to elicit an undominated OA within $\mathcal{C}_2^{BV}(\underline{P})$ and $\mathcal{C}_\infty^{BV}(\underline{P})$.

4.1   Approach based on a quadratic distance

One possibility for obtaining a unique solution to our problem could be to use the quadratic distance, i.e., to consider the OA in $\mathcal{C}_2^{BV}$ or $\mathcal{C}_\infty^{BV}$ minimising

$$d_q(\underline{P}, \underline{Q}) = \sum_{A \subseteq \mathcal{X}} \left(\underline{P}(A) - \underline{Q}(A) \right)^2. \tag{1}$$

Given $\delta_2^{BV} = \min_{\underline{Q} \in \mathcal{C}_2} d_{BV}(\underline{P}, \underline{Q})$ and $\delta_\infty^{BV} = \min_{\underline{Q} \in \mathcal{C}_\infty} d_{BV}(\underline{P}, \underline{Q})$, in $\mathcal{C}_2^{BV}$ we may set up the quadratic problem of minimising Eq. (1) subject to the constraints (2monot.1)$-$(2monot.2), and also to

$$\sum_{B \subseteq A} m_{\underline{Q}}(B) \leq \underline{P}(A) \quad \forall A \neq \emptyset, \mathcal{X}. \tag{OA}$$

$$\sum_{A \subseteq \mathcal{X}} \left(\underline{P}(A) - \sum_{B \subseteq A} m_{\underline{Q}}(B) \right) = \delta_2^{BV}. \tag{2monot-δ}$$

Analogously, in $\mathcal{C}_\infty^{BV}$ we can minimise Eq. (1) subject to (Cmonot.1), (OA) and:

$$\sum_{A \subseteq \mathcal{X}} \left(\underline{P}(A) - \sum_{B \subseteq A} m_{\underline{Q}}(B) \right) = \delta_\infty^{BV}. \tag{Cmonot-δ}$$

Proposition 2. *Let $\underline{P}$ be a coherent lower probability. Then:*

1. *The problem of minimising Eq. (1) subject to (2monot.1)$\div$(2monot.2), (OA) and (2monot-δ) has a unique solution, which is an undominated OA of $\underline{P}$ in $\mathcal{C}_2^{BV}(\underline{P})$.*
2. *Similarly, the problem of minimising Eq. (1) subject to (Cmonot.1), (OA) and (Cmonot-δ) has a unique solution, which is an undominated OA of $\underline{P}$ in $\mathcal{C}_\infty^{BV}(\underline{P})$.*

The following example illustrates this result.

Example 1. Consider the coherent $\underline{P}$ given on $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ by [9, Ex.1]:

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

For this coherent lower probability, $\delta_2^{BV} = \delta_\infty^{BV} = 1$, the sets $\mathcal{C}_2^{BV}(\underline{P})$ and $\mathcal{C}_\infty^{BV}(\underline{P})$ coincide and they are given by $\left\{ \underline{Q}_\alpha \mid \alpha \in [0, 0.5] \right\}$, where:

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

Therefore, if among these $\underline{Q}_\alpha$ we minimise the quadratic distance with respect to $\underline{P}$, the optimal solution is $\underline{Q}_{0.25}$, in both $\mathcal{C}_2$ and $\mathcal{C}_\infty$. ♦

Note that the solution we obtain in Prop. 2 is *not* the OA that minimises the quadratic distance in $\mathcal{C}_2$ (or $\mathcal{C}_\infty$), but the one that minimises it in $\mathcal{C}_2^{BV}$ (or $\mathcal{C}_\infty^{BV}$).

While the quadratic distance is in our view the most promising approach in order to elicit a unique undominated OA of $\underline{P}$ in $\mathcal{C}_2$ and $\mathcal{C}_\infty$, it is not the only possibility. In the rest of the section we explore other approaches.

4.2   Approach based on the total variation distance

Instead of considering the quadratic distance, we may consider some extensions of the total variation distance [8, Ch.4.1] defined between lower probabilities:

$$d_1(\underline{P}_1, \underline{P}_2) = \max_{A \subseteq \mathcal{X}} |\underline{P}_1(A) - \underline{P}_2(A)|, \tag{2}$$

$$d_2(\underline{P}_1, \underline{P}_2) = \frac{1}{2} \sum_{x \in \mathcal{X}} |\underline{P}_1(\{x\}) - \underline{P}_2(\{x\})|, \tag{3}$$

$$d_3(\underline{P}_1, \underline{P}_2) = \sup_{P_1 \geq \underline{P}_1, P_2 \geq \underline{P}_2} \left(\max_{A \subseteq \mathcal{X}} |P_1(A) - P_2(A)| \right). \tag{4}$$

Thus, instead of minimising Eq. (1) we may consider the OA in $\mathcal{C}_2^{BV}(\underline{P})$ or $\mathcal{C}_\infty^{BV}(\underline{P})$ that minimises one of $d_i(\underline{P}, Q)$. However, none of d_1, d_2, d_3 determines a unique OA in $\mathcal{C}_2^{BV}(\underline{P})$ or $\mathcal{C}_\infty^{BV}(\underline{P})$, as we next show.

Example 2. Consider the coherent $\underline{P}$ in a four-element space given by:

A	$\underline{P}(A)$	$\underline{Q}_0$	$\underline{Q}_1$	A	$\underline{P}(A)$	$\underline{Q}_0$	$\underline{Q}_1$
$\{x_1\}$	0.1	0.1	0.1	$\{x_2, x_3\}$	0.3	0.3	0.2
$\{x_2\}$	0	0	0	$\{x_2, x_4\}$	0.4	0.3	0.4
$\{x_3\}$	0	0	0	$\{x_3, x_4\}$	0.4	0.3	0.3
$\{x_4\}$	0.3	0.3	0.3	$\{x_1, x_2, x_3\}$	0.5	0.5	0.5
$\{x_1, x_2\}$	0.1	0.1	0.1	$\{x_1, x_2, x_4\}$	0.6	0.6	0.6
$\{x_1, x_3\}$	0.3	0.2	0.3	$\{x_1, x_3, x_4\}$	0.7	0.7	0.7
$\{x_1, x_4\}$	0.6	0.6	0.5	$\{x_2, x_3, x_4\}$	0.6	0.6	0.6
				$\mathcal{X}$	1	1	1

In [10, Ex.1] we showed that $\mathcal{C}_\infty^{BV}(\underline{P})$ is given by $\left\{\underline{Q}_0, \underline{Q}_1, \underline{Q}_\alpha \mid \alpha \in (0,1)\right\}$, where $\underline{Q}_\alpha = \alpha\underline{Q}_0 + (1-\alpha)\underline{Q}_1$. In all the cases it holds that:

$d_1(\underline{P}, \underline{Q}_0) = d_1(\underline{P}, \underline{Q}_1) = d_1(\underline{P}, \underline{Q}_\alpha) = 0.1 \quad \forall \alpha \in (0,1).$
$d_2(\underline{P}, \underline{Q}_0) = d_2(\underline{P}, \underline{Q}_1) = d_2(\underline{P}, \underline{Q}_\alpha) = 0 \quad \forall \alpha \in (0,1).$
$d_3(\underline{P}, \underline{Q}_0) = d_3(\underline{P}, \underline{Q}_1) = d_3(\underline{P}, \underline{Q}_\alpha) = 0.6 = \max_{A \subseteq \mathcal{X}} |\overline{P}(A) - \underline{Q}_\alpha(A)| \quad \forall \alpha \in (0,1).$

This means that none of d_1, d_2 and d_3 allows to elicit a unique OA from $\mathcal{C}_\infty^{BV}$. ♦

Consider now the undominated OAs in $\mathcal{C}_2$. First of all, note we may disregard d_2, because from [9, Prop.2] every undominated OA Q in $\mathcal{C}_2$ satisfies $Q(\{x\}) = \underline{P}(\{x\})$ for every $x \in \mathcal{X}$, and therefore $d_2(\underline{P}, Q) = 0$. The following example shows that d_1 and d_3 do not allow to elicit a unique undominated OA from $\mathcal{C}_2^{BV}(\underline{P})$, either.

Example 3. Consider now the coherent lower probability $\underline{P}$ given by:

A	$\underline{P}(A)$	$\underline{Q}_0$	$\underline{Q}_1$	A	$\underline{P}(A)$	$\underline{Q}_0$	$\underline{Q}_1$
$\{x_1\}$	0	0	0	$\{x_2,x_3\}$	0.3	0.3	0.3
$\{x_2\}$	0	0	0	$\{x_2,x_4\}$	0.3	0.2	0.3
$\{x_3\}$	0.1	0.1	0.1	$\{x_3,x_4\}$	0.4	0.4	0.3
$\{x_4\}$	0	0	0	$\{x_1,x_2,x_3\}$	0.5	0.5	0.5
$\{x_1,x_2\}$	0.3	0.2	0.2	$\{x_1,x_2,x_4\}$	0.6	0.6	0.6
$\{x_1,x_3\}$	0.3	0.3	0.3	$\{x_1,x_3,x_4\}$	0.6	0.6	0.6
$\{x_1,x_4\}$	0.2	0.2	0.2	$\{x_2,x_3,x_4\}$	0.6	0.6	0.6
				$\mathcal{X}$	1	1	1

It holds that $\mathcal{C}_2^{BV}(\underline{P}) = \mathcal{C}_\infty^{BV}(\underline{P}) = \left\{\underline{Q}_0, \underline{Q}_1, \underline{Q}_\alpha \mid \alpha \in (0,1)\right\}$, where $\underline{Q}_\alpha = \alpha\underline{Q}_0 + (1-\alpha)\underline{Q}_1$. However:

$$d_1\left(\underline{P}, \underline{Q}_0\right) = d_1\left(\underline{P}, \underline{Q}_1\right) = d_1\left(\underline{P}, \underline{Q}_\alpha\right) = 0.1 \quad \forall\alpha \in (0,1), \quad \text{and}$$
$$d_3\left(\underline{P}, \underline{Q}_0\right) = d_3\left(\underline{P}, \underline{Q}_1\right) = d_3\left(\underline{P}, \underline{Q}_\alpha\right) = 0.5 = \max_{A \subseteq \mathcal{X}}\left|\overline{P}(A) - \underline{Q}_\alpha(A)\right|, \ \forall\alpha \in (0,1).$$

Thus, neither d_1 nor d_3 determines a unique undominated OA in $\mathcal{C}_2^{BV}(\underline{P})$. ♦

4.3   Approach based on measuring specificity

When we consider the OAs of $\underline{P}$ in $\mathcal{C}_\infty$, we may compare them by measuring their *specificity*. We consider here the specificity measure defined by Yager [14], that splits the mass of the focal events among its elements.

Definition 2. *Let $\underline{Q}$ be a completely monotone lower probability on $\mathcal{P}(\mathcal{X})$ with Möbius inverse $m_{\underline{Q}}$. Its* specificity *is given by*

$$S(\underline{Q}) = \sum_{\emptyset \neq A \subseteq \mathcal{X}} \frac{m_{\underline{Q}}(A)}{|A|} = \sum_{i=1}^{n} \frac{1}{i} \sum_{A:|A|=i} m_{\underline{Q}}(A).$$

Hence, we can choose an undominated OA in $\mathcal{C}_\infty^{BV}(\underline{P})$ with the greatest specificity. The next example shows that this criterion does not give rise to a unique undominated OA.

Example 4. Consider again Ex. 1, where $\mathcal{C}_\infty^{BV} = \left\{\underline{Q}_\alpha \mid \alpha \in [0, 0.5]\right\}$. The Möbius inverse of $\underline{Q}_\alpha$ is given by

$$m_{\underline{Q}_\alpha}(\{x_1, x_4\}) = m_{\underline{Q}_\alpha}(\{x_2, x_3\}) = \alpha, \ m_{\underline{Q}_\alpha}(\{x_1, x_3\}) = m_{\underline{Q}_\alpha}(\{x_2, x_4\}) = 0.5 - \alpha$$

and zero elsewhere. Hence, the specificity of $\underline{Q}_\alpha$ is

$$S(\underline{Q}_\alpha) = \frac{1}{2}\left(\alpha + \alpha + 0.5 - \alpha + 0.5 - \alpha\right) = 0.5,$$

regardless of the value of $\alpha \in [0, 0.5]$. We conclude that all the undominated OAs in $\mathcal{C}_\infty^{BV}(\underline{P})$ have the same specificity. ♦

5   Elicitation of an outer approximation in $\mathcal{C}_\Pi$

In [10, Sec.6] we showed that the set of non-dominating OAs in $\mathcal{C}_\Pi$ is finite and that we have a simple procedure for determining them. Given the conjugate upper probability $\overline{P}$ of $\underline{P}$, each permutation σ in the set of all permutations S_n of $\{1, 2, \ldots, n\}$ defines a possibility measure by:

$$\Pi_\sigma(\{x_{\sigma(1)}\}) = \overline{P}(\{x_{\sigma(1)}\}), \text{ and} \tag{5}$$

$$\Pi_\sigma(\{x_{\sigma(i)}\}) = \max_{A \in \mathcal{A}_{\sigma(i)}} \overline{P}(A \cup \{x_{\sigma(i)}\}), \text{ where for every } i > 1: \tag{6}$$

$$\mathcal{A}_{\sigma(i)} = \left\{ A \subseteq \{x_{\sigma(1)}, \ldots, x_{\sigma(i-1)}\} \mid \overline{P}(A \cup \{x_{\sigma(i)}\}) > \max_{x \in A} \Pi_\sigma(\{x\}) \right\}, \tag{7}$$

and $\Pi_\sigma(A) = \max_{x \in A} \Pi_\sigma(\{x\})$ for every $A \subseteq \mathcal{X}$. Then, the set of non-dominating OAs of $\overline{P}$ is $\{\Pi_\sigma \mid \sigma \in S_n\}$ (see [10, Prop.11, Cor.13]).

Next we propose a number of approaches to elicit a unique OA of $\overline{P}$ among the Π_σ determined by Eqs. (5)÷(7). Note that the procedure above may determine the same possibility measure using different permutations. The next result is concerned with such cases, and will be helpful for reducing the candidate possibility measures.

Proposition 3. *Let $\{\Pi_\sigma \mid \sigma \in S_n\}$ be the set of non-dominating OAs of $\overline{P}$ in $\mathcal{C}_\Pi$. Consider $\sigma \in S_n$ and its associated Π_σ. Assume that $\exists i \in \{2, \ldots, n\}$ such that $\Pi_\sigma(\{x_{\sigma(i)}\}) \neq \overline{P}(\{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\})$. Then, there exists $\sigma' \in S_n$ such that*

$$\Pi_\sigma(A) = \Pi_{\sigma'}(A) \quad \forall A \subseteq \mathcal{X} \quad and$$

$$\Pi_{\sigma'}(\{x_{\sigma'(j)}\}) = \overline{P}(\{x_{\sigma'(1)}, \ldots, x_{\sigma'(j)}\}) \quad \forall j = 1, \ldots, n. \tag{8}$$

5.1   Approach based on the BV-distance

Our first approach consists in looking for a possibility measure, among $\{\Pi_\sigma \mid \sigma \in S_n\}$, that minimises the BV-distance with respect to the original model. If we denote by N_σ the conjugate necessity measure of Π_σ, the BV-distance can be expressed by:

$$d_{BV}(\underline{P}, N_\sigma) = \sum_{A \subseteq \mathcal{X}} (\Pi_\sigma(A) - \overline{P}(A)) = \sum_{A \subseteq \mathcal{X}} \Pi_\sigma(A) - \sum_{A \subseteq \mathcal{X}} \overline{P}(A).$$

To ease the notation, for each $\sigma \in S_n$ we denote by $\vec{\beta}_\sigma$ the ordered vector determined by the values $\Pi_\sigma(\{x_{\sigma(i)}\})$, $i = 1, \ldots, n$, so that $\beta_{\sigma,1} \leq \ldots \leq \beta_{\sigma,n}$. Using this notation:

$$\sum_{A \subseteq \mathcal{X}} \Pi_\sigma(A) = \beta_{\sigma,1} + 2\beta_{\sigma,2} + \ldots + 2^{n-1}\beta_{\sigma,n} = \sum_{i=1}^{n} 2^{i-1} \beta_{\sigma,i}. \tag{9}$$

This means that, in order to minimise $d_{BV}(\underline{P}, N_\sigma)$, we must minimise Eq. (9). Our next result shows that if a dominance relation exists between $\vec{\beta}_\sigma$ and $\vec{\beta}_{\sigma'}$, this induces an order between the values in Eq. (9).

Lemma 1. *Let $\vec{\beta}_\sigma$ and $\vec{\beta}_{\sigma'}$ be two vectors associated with two possibility measures Π_σ and $\Pi_{\sigma'}$. Then $\vec{\beta}_\sigma \leq \vec{\beta}_{\sigma'}$ implies that $d_{BV}(\underline{P}, N_\sigma) \leq d_{BV}(\underline{P}, N_{\sigma'})$, and $\vec{\beta}_\sigma \lneq \vec{\beta}_{\sigma'}$ implies that $d_{BV}(\underline{P}, N_\sigma) < d_{BV}(\underline{P}, N_{\sigma'})$.*

This result may contribute to rule out some possibilities in S_n, as illustrated in the next example.

Example 5. Consider the following coherent conjugate lower and upper probabilities, as well as their associated possibility measures Π_σ and vectors β_σ:

A	$\underline{P}(A)$	$\overline{P}(A)$		σ	$\Pi_\sigma(\{x_1\})$	$\Pi_\sigma(\{x_2\})$	$\Pi_\sigma(\{x_3\})$	$\vec{\beta}_\sigma$
$\{x_1\}$	0.25	0.4		$\sigma_1 = (1,2,3)$	0.4	0.8	1	$(0.4, 0.8, 1)$
$\{x_2\}$	0.2	0.5		$\sigma_2 = (1,3,2)$	0.4	1	0.8	$(0.4, 0.8, 1)$
$\{x_3\}$	0.2	0.5		$\sigma_3 = (2,1,3)$	0.8	0.5	1	$(0.5, 0.8, 1)$
$\{x_1, x_2\}$	0.5	0.8		$\sigma_4 = (2,3,1)$	1	0.5	0.75	$(0.5, 0.75, 1)$
$\{x_1, x_3\}$	0.5	0.8		$\sigma_5 = (3,1,2)$	0.8	1	0.5	$(0.5, 0.8, 1)$
$\{x_2, x_3\}$	0.6	0.75		$\sigma_6 = (3,2,1)$	1	0.75	0.5	$(0.5, 0.75, 1)$
$\mathcal{X}$	1	1						

Taking σ_1 and σ_3, it holds that $\vec{\beta}_{\sigma_1} \lneq \vec{\beta}_{\sigma_3}$, so from Lemma 1 $d_{BV}(\underline{P}, N_{\sigma_1}) < d_{BV}(\underline{P}, N_{\sigma_3})$. Hence, we can discard Π_{σ_3}. The same applies to $\vec{\beta}_{\sigma_1}$ and $\vec{\beta}_{\sigma_5}$, whence $d_{BV}(\underline{P}, N_{\sigma_1}) < d_{BV}(\underline{P}, N_{\sigma_5})$. ♦

In the general case, the set of all vectors $\vec{\beta}_\sigma$ is not totally ordered. Then, the problem of minimising the BV-distance is solved by casting it into a shortest path problem, as we shall now illustrate.

As we said before, the possibility measure(s) in $\{\Pi_\sigma : \sigma \in S_n\}$ that minimise the BV-distance to $\underline{P}$ are the ones minimising $\sum_{A \subseteq \mathcal{X}} \Pi_\sigma(A)$. In turn, this sum can be computed by means of Eq. (9), once we order the values $\Pi_\sigma(\{x_{\sigma(i)}\})$, for $i = 1, \ldots, n$. Our next result will be useful for this aim:

Proposition 4. *Let $\{\Pi_\sigma \mid \sigma \in S_n\}$ be the set of non-dominating OAs of $\overline{P}$ in $\mathcal{C}_\Pi$. Then $\sum_{A \subseteq \mathcal{X}} \Pi_\sigma(A) \leq \sum_{i=1}^{n} 2^{i-1} \overline{P}(\{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\})$, and the equality holds if and only if Π_σ satisfies Eq. (8).*

From Prop. 3, if Π_σ does not satisfy Eq. (8) then there exists another permutation σ' that does so and such that $\Pi_\sigma = \Pi_{\sigma'}$. This means that we can find Π_σ minimising the BV-distance by solving a shortest path problem. For this aim, we consider the Hasse diagram of $\mathcal{P}(\mathcal{X})$, and if $x_i \notin A$, we assign the weight $2^{|A|} \overline{P}(A \cup \{x_i\})$ to the edge $A \rightarrow A \cup \{x_i\}$, and the fictitious weight 0 to $\mathcal{X} \setminus \{x_i\} \rightarrow \mathcal{X}$. Since these weights are positive, we can find the optimal solution using Dijkstra's algorithm [5]. In this diagram, there are two types of paths:

(a) Paths whose associated Π_σ satisfies Eq. (8); then $\sum_{A \subseteq \mathcal{X}} \Pi_\sigma(A)$ coincides with the value of the path.
(b) Paths whose associated Π_σ does not satisfy Eq. (8); then $\sum_{A \subseteq \mathcal{X}} \Pi_\sigma(A)$ shall be strictly smaller than the value of the path, and shall moreover coincide with the value of the path determined by another permutation σ', as established in Prop. 3. Then the shortest path can never be found among these ones.

As a consequence, the shortest path determines a permutation σ whose associated Π_σ satisfies Eq. (8). Moreover, this Π_σ minimises the BV-distance with respect to $\overline{P}$ among all the non-dominating OAs in $\mathcal{C}_\infty$. And in this manner we shall obtain all such possibility measures.

Example 6. Consider the coherent conjugate lower and upper probability $\underline{P}$ and $\overline{P}$ from Ex. 5. The following figure pictures the Hasse diagram with weights of the edges we discussed before:
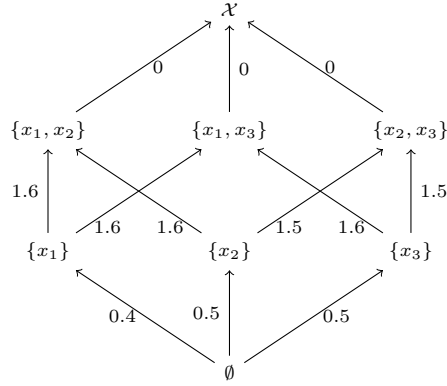
Solving the shortest path problem from $\emptyset$ to $\mathcal{X}$ using Dijkstra's algorithm, we obtain an optimal value of 2 that is attained with the following paths:

$$\emptyset \to \{x_1\} \to \{x_1, x_2\} \to \mathcal{X}, \qquad \emptyset \to \{x_1\} \to \{x_1, x_3\} \to \mathcal{X}.$$
$$\emptyset \to \{x_2\} \to \{x_2, x_3\} \to \mathcal{X}, \qquad \emptyset \to \{x_3\} \to \{x_2, x_3\} \to \mathcal{X}.$$

These four paths correspond to the permutations $\sigma_1 = (1, 2, 3)$, $\sigma_2 = (1, 3, 2)$, $\sigma_4 = (2, 3, 1)$ and $\sigma_6 = (3, 2, 1)$. Even if they induce four different possibility measures, all of them are at the same BV-distance to $\overline{P}$. Note also that the other two possibility measures are those that were discarded in Ex. 5. ♦

This example shows that with this approach we obtain the Π_σ at minimum BV-distance. It also shows that the solution is not unique, and that the vectors $\vec{\beta}_\sigma$ and $\vec{\beta}_{\sigma'}$ that are not pointwisely ordered may be associated with two different possibility measures Π_σ and $\Pi_{\sigma'}$ minimising the BV-distance (such as σ_1 and σ_6 in the example). Nevertheless, we can determine situations in which the BV-distance elicits one single Π_σ, using the following result:

Proposition 5. *Let $\underline{P}$ and $\overline{P}$ be coherent conjugate lower and upper probabilities. If there is a permutation $\sigma \in S_n$ satisfying*

$$\overline{P}\left(\{x_{\sigma(1)}, \ldots, x_{\sigma(j)}\}\right) = \min_{|A|=j} \overline{P}(A) \quad \forall j = 1, \ldots, n, \tag{10}$$

then Π_σ minimises the BV-distance.

As a consequence of this result, if there is only one permutation satisfying Eq. (10), this approach allows to elicit a unique undominated OA.

Example 7. Consider again $\underline{P}$ and $\overline{P}$ from Ex. 2. We can see that:

$$\overline{P}(\{x_2\}) = 0.3 = \min_{|A|=1} \overline{P}(A), \quad \overline{P}(\{x_2, x_3\}) = 0.4 = \min_{|A|=2} \overline{P}(A),$$

$$\overline{P}(\{x_1, x_2, x_3\}) = 0.7 = \min_{|A|=3} \overline{P}(A).$$

There is a (unique) chain of events satisfying Eq. (10), namely $\{x_2\} \subseteq \{x_2, x_3\} \subseteq \{x_1, x_2, x_3\} \subseteq \mathcal{X}$, that is associated with the permutation $\sigma = (2, 3, 1, 4)$. From Prop. 5, Π_σ is the unique undominated OA in $\mathcal{C}_\Pi$ minimising the BV-distance. ♦

5.2 Approach based on measuring specificity

Since any possibility measure is in particular the conjugate of a belief function, it is possible to compare them by means of specificity measures. In this section, we investigate which possibility measure(s) among $\{\Pi_\sigma \mid \sigma \in S_n\}$ are the most specific.

With each Π_σ in $\{\Pi_\sigma \mid \sigma \in S_n\}$, we consider its associated vector $\vec{\beta}_\sigma$. For possibility measures, the focal events A_i are nested: $A_i := \{x_{\sigma(n-i+1)}, \ldots, x_{\sigma(n)}\}$, with $m(A_i) = \beta_{\sigma,n-i+1} - \beta_{\sigma,n-i}$. Hence specificity simplifies to:

$$S(\Pi_\sigma) = 1 - \frac{\beta_{\sigma,n-1}}{2} - \frac{\beta_{\sigma,n-2}}{2 \cdot 3} - \ldots - \frac{\beta_{\sigma,1}}{n(n-1)}.$$

Thus, a most specific possibility measure will minimise

$$\frac{\beta_{\sigma,1}}{n(n-1)} + \frac{\beta_{\sigma,2}}{(n-1)(n-2)} + \ldots + \frac{\beta_{\sigma,n-1}}{2}. \tag{11}$$

Our first result is similar to Lemma 1, and allows to discard some of the possibility measures Π_σ.

Lemma 2. *Let $\vec{\beta}_\sigma$ and $\vec{\beta}_{\sigma'}$ be the vectors associated with the possibility measures Π_σ and $\Pi_{\sigma'}$. Then $\vec{\beta}_\sigma \leq \vec{\beta}_{\sigma'}$ implies that $S(\Pi_\sigma) \geq S(\Pi_{\sigma'})$ and $\beta_\sigma \lneq \beta_{\sigma'}$ implies that $S(\Pi_\sigma) > S(\Pi_{\sigma'})$.*

Example 8. Let us continue with Ex. 5 and 6. In Ex. 5 we showed the possibility measures $\{\Pi_\sigma \mid \sigma \in S_n\}$ and their associated vectors $\vec{\beta}_\sigma$. As we argued in Ex. 5, $\vec{\beta}_{\sigma_1} \lneq \vec{\beta}_{\sigma_3}$, where $\sigma_1 = (1, 2, 3)$ and $\sigma_3 = (2, 1, 3)$. Hence from Lemma 2, $S(\Pi_{\sigma_1}) > S(\Pi_{\sigma_3})$, meaning that we can discard Π_{σ_3}. A similar reasoning allows us to discard Π_{σ_5}. ♦

In order to find those possibility measures maximising the specificity, we have to minimise Eq. (11). Here we can make the same considerations as in Sec. 5.1.

Proposition 6. *Let $\{\Pi_\sigma \mid \sigma \in S_n\}$ be the set of non-dominating OAs of $\overline{P}$ in $\mathcal{C}_\Pi$. Then $S(\Pi_\sigma) \geq 1 - \sum_{i=1}^{n-1} \frac{\overline{P}(\{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\})}{(n-i)(n-i+1)}$, and the equality holds if and only if Π_σ satisfies Eq. (8).*

Moreover, from Prop. 3 we know that if Π_σ does not satisfy Eq. (8) then it is possible to find another permutation σ' that does so and such that $\Pi_\sigma = \Pi_{\sigma'}$.

This means that we can find the Π_σ maximising the specificity by solving a shortest path problem, similarly to what we did for the BV-distance. For this aim, we consider the Hasse diagram of $\mathcal{P}(\mathcal{X})$; if $x_i \notin A$, we assign the weight

$$\frac{\overline{P}(\{A \cup \{x_i\}\})}{(n - |A|)(n - |A| - 1)} \tag{12}$$

to the edge $A \to A \cup \{x_i\}$, and we give the fictitious weight 0 to $\mathcal{X} \setminus \{x_i\} \to \mathcal{X}$. In this diagram, there are two types of paths:

(a) Paths whose associated possibility measure Π_σ satisfies Eq. (8); then the value of Eq. (11) for Π_σ coincides with the value of the path.
(b) Paths whose associated possibility measure Π_σ does not satisfy Eq. (8); then the value of Eq. (11) for Π_σ is strictly smaller than the value of the path, and shall moreover coincide with the value of the path determined by another permutation σ', as established in Prop. 3. Then the shortest path can never be found among these ones.

As a consequence, if we find the shortest path we shall determine a permutation σ whose associated Π_σ satisfies Eq. (8), and therefore that maximises the specificity; and in this manner we shall obtain all such possibility measures.

Example 9. Consider again the running Ex. 5, 6 and 8. In the next figure we can see the Hasse diagram of $\mathcal{P}(\mathcal{X})$ with the weights from Eq. (12).
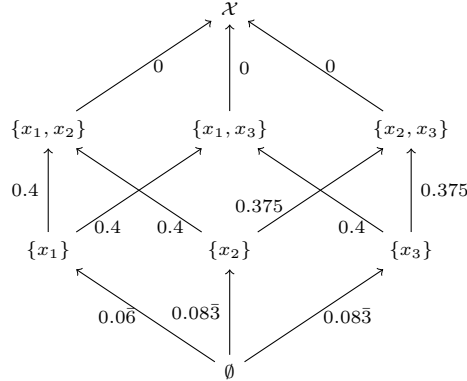
The optimal solutions of the shortest path problem are $\emptyset \to \{x_2\} \to \{x_2, x_3\} \to \mathcal{X}$ and $\emptyset \to \{x_3\} \to \{x_2, x_3\} \to \mathcal{X}$, which correspond to the permutations $\sigma_4 = (2, 3, 1)$ and $\sigma_6 = (3, 2, 1)$. ♦

These examples also show that the approach based on minimising the BV-distance and that based on maximising the specificity are not equivalent: in Ex. 5 we have seen that the possibility measures minimising the BV-distance are the ones associated with the permutations $(3, 1, 2)$ and $(3, 2, 1)$, while those maximising the specificity are the ones associated with $(2, 3, 1)$ and $(3, 2, 1)$.

To conclude this subsection, we establish a result analogous to Prop. 5.

Proposition 7. *Let $\underline{P}$ and $\overline{P}$ be coherent conjugate lower and upper probabilities. If there is a permutation σ satisfying Eq. (10), then Π_σ maximises the specificity.*

We arrive at the same conclusion of Prop. 5: if there is a unique permutation satisfying Eq. (10), then there is a unique possibility measure maximising the specificity; and in that case the chosen possibility measure maximises the specificity and at the same time minimises the BV-distance.

5.3   Approach based on the total variation distance

As we did in Sec. 4.2, we could elicit a possibility measure among $\{\Pi_\sigma \mid \sigma \in S_n\}$ by minimising one of the extensions of the TV-distance. When we focus on upper probabilities, the distances given in Eqs. (2)÷(4) can be expressed by:

$$d_1(\overline{P}_1, \overline{P}_2) = \max_{A \subseteq \mathcal{X}} |\overline{P}_1(A) - \overline{P}_2(A)|, \quad d_2(\overline{P}_1, \overline{P}_2) = \frac{1}{2} \sum_{x \in \mathcal{X}} |\overline{P}_1(\{x\}) - \overline{P}_2(\{x\})|,$$

$$d_3(\overline{P}_1, \overline{P}_2) = \sup_{P_1 \leq \overline{P}_1, P_2 \leq \overline{P}_2} \left(\max_{A \subseteq \mathcal{X}} |P_1(A) - P_2(A)| \right).$$

As in the case of $\mathcal{C}_2$ and $\mathcal{C}_\infty$, this approach is not fruitful:

Example 10. Consider our running Ex. 5. The values $d_i(\overline{P}, \Pi_j)$ are given by:

	Π_{σ_1}	Π_{σ_2}	Π_{σ_3}	Π_{σ_4}	Π_{σ_5}	Π_{σ_6}
$d_1(\overline{P}, \Pi_{\sigma_i})$	0.5	0.5	0.5	0.6	0.5	0.6
$d_2(\overline{P}, \Pi_{\sigma_i})$	0.4	0.4	0.45	0.425	0.45	0.425
$d_3(\overline{P}, \Pi_{\sigma_i})$	0.8	0.8	0.8	0.75	0.8	0.75

Thus, none of d_1, d_2 or d_3 allow to elicit a single possibility measure. ♦

6   Conclusions

In this paper, we have explored a number of approaches to elicit a unique undominated OA of a given coherent lower probability. When the OA belongs to the families $\mathcal{C}_2$ and $\mathcal{C}_\infty$, we first focus on the ones minimising the BV-distance. Among the approaches we have then considered, it seems that the best one is to consider the OA in $\mathcal{C}_2$ and $\mathcal{C}_\infty$ that minimises the quadratic distance: it singles out a unique undominated OA, while this is not the case when we minimise the TV-distance or maximise the specificity.

In the case of $\mathcal{C}_\Pi$, we know from [10] that there are at most $n!$ non-dominating OA of a coherent upper probability, and these are determined by Eqs. (5)÷(7). In order to elicit a unique possibility measure we have considered the approaches based on minimising the BV-distance, maximising the specificity and minimising the TV-distance. While none of them elicits a unique OA in general, we have given a sufficient condition for the uniqueness in Props. 5 and 7. Moreover, we have seen that we can find the optimal OA according to the BV-distance and the specificity approaches by solving a shortest path problem.

In a future work, we intend to make a thorough comparison between the main approaches and to report on additional results that we have not included in this paper due to space limitations, such as the comparison between the OA in terms of the preservation of the preferences encompassed by the initial model, and the analysis of other particular cases of imprecise probability models, such as probability boxes.

Acknowledgements

References

1. Baroni, P., Vicig, P.: An uncertainty interchange format with imprecise probabilities. International Journal of Approximate Reasoning **40**, 147–180 (2005)https://doi.org/10.1016/j.ijar.2005.03.001
2. Bronevich, A., Augustin, T.: Approximation of coherent lower probabilities by 2-monotone measures. In: Augustin, T., Coolen, F.P.A., Moral, S., Troffaes, M.C.M. (eds.) Proceedings of the Sixth ISIPTA. pp. 61–70 (2009)
3. Chateauneuf, A., Jaffray, J.Y.: Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. Mathematical Social Sciences **17**(3), 263–283 (1989) https://doi.org/10.1016/0165-4896(89)90056-5
4. Choquet, G.: Theory of capacities. Annales de l'Institut Fourier **5**, 131–295 (1954)
5. Dijkstra, E.: A note on two problems in connection with graphs. Numerische Mathematik pp. 269–271 (1959) https://doi.org/10.1007/BF01386390
6. Dubois, D., Prade, H.: Possibility Theory. Plenum Press, New York (1988)
7. Levi, I.: The enterprise of knowledge. MIT Press, Cambridge (1980)
8. Levin, D.A., Peres, Y., Wilmer, E.: Markov Chains and Mixing Times. American Mathematical Society (2009)
9. Montes, I., Miranda, E., Vicig, P.: 2-monotone outer approximations of coherent lower probabilities. International Journal of Approximate Reasoning **101**, 181–205 (2018) https://doi.org/10.1016/j.ijar.2018.07.004
10. Montes, I., Miranda, E., Vicig, P.: Outer approximating coherent lower probabilities with belief functions. International Journal of Approximate Reasoning **110**, 1–30 (2019) https://doi.org/10.1016/j.ijar.2019.03.008
11. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)
12. Shapley, L.S.: Cores of convex games. International Journal of Game Theory **1**, 11–26 (1971) https://doi.org/10.1007/BF01753431
13. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)
14. Yager, R.: Entropy and specificity in a mathematical theory of evidence. International Journal of General Systems **9**, 249–260 (1983)
15. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems **1**, 3–28 (1978) https://doi.org/10.1016/S0165-0114(99)80004-9