International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems

The Total Variation distance for comparing non-additive measures

David Nieto-Barba, Enrique Miranda, Ignacio Montes
Department of Statistics and O.R., University of Oviedo
C/Leopoldo Calvo Sotelo, 18, 33007, Oviedo, Spain
nietodavid@uniovi.es, mirandaenrique@uniovi.es, imontes@uniovi.es

Received (received date)
Revised (revised date)

The Total Variation is a common distance between probability distributions that measures the maximum difference in probability among all events. When comparing non-additive measures, that can be represented by closed and convex sets of probability measures, the Total Variation distance can be extended in multiple ways. This paper explores three specific approaches in detail. The first approach considers the minimum Total Variation distance between the probability measures dominating the non-additive measures. The second approach replaces the minimum with the maximum of the Total Variation distances. The third approach modifies the first one by using the supremum distance instead of the Total Variation. We analyse the main properties of each approach and, in particular, apply them to distort non-additive measures. Finally, we demonstrate that the distortion of non-additive measures with the proposed extensions are closely connected to the strong and weak cores in coalition game theory.

Keywords: Total variation distance, supremum distance, lower probabilities, non-additive measures, 2-monotone lower probabilities, coalitional games.

1. Introduction

In many areas of probability theory, and in particular when analysing the robustness of a model, there is a need to measure the distance between probability measures. One popular approach is to do this by means of the Total Variation distance (TV-distance, for short[?]), which has two key features, its simple formulation and its intuitive interpretation: it measures the maximum absolute difference between the probabilities of each event.

The TV-distance also plays an important role when dealing with non-additive measures, also called capacities or, following the terminology of the broader theory of imprecise probabilities[?,?] , lower probabilities. To mention just a few of its applications, it has been used to distort a probability[?] , giving rise to a neighbourhood model called the TV-model[?] ; to measure the distance between imprecise Markov Chains[?] ; to extend the Radon-Nikodym Theorem to non-additive measures[?] ; or to determine a centroid of the set of probability measures compatible with a non-additive measure[?] .

Our aim in this contribution is to explore how the TV-distance can be generalised in order to be able to compare non-additive measures, going a step beyond the study presented in[?] . Given two non-additive measures or lower probabilities, we consider the set of probability measures they determine (i.e., their credal sets[?]) and follow two different approaches. On the one hand, we seek the minimum TV-distance between all the probabilities in the credal sets, giving rise to what we will call *Minimum TV*. On the other hand, we look for the maximum of the TV-distances instead of the minimum, leading to what we will call *Maximum Discrepancy*. This second approach aligns with the proposal in[?] . Additionally, we also consider the supremum distance between probability measures, that as we shall see is related to a sort of penalised or weighted TV-distance. This leads to our third approach, that we shall call *Minimum Supremum*, and that consists in finding the minimum of the supremum distances between the probability measures in the credal sets.

For each of these three proposals, we investigate: (i) whether they define a distance between lower probabilities; (ii) if alternative formulations can be derived in terms of the extreme points of the credal sets, in terms of the probability measures at the boundaries of the credal sets, or in terms of the direct comparison of lower and upper probabilities; and (iii) how they can be used to distort a lower probability by creating a neighbourhood around it.

There exists a formal connection between lower probabilities (or non-additive measures) and (normalised and monotone) coalitional games[?,?] : the domain of the lower probability can be seen as the set of coalitions between players and the lower probability of an event can be regarded as the minimum proportion of an available resource required by the players in the corresponding coalition. In this paper, we exploit this connection to show that the so-called strong and weak cores[?] , used to avoid games with an empty core, are closely connected to the distortion of lower probabilities. This complements our recent studies in[?,?] , where we analysed the invariance of probabilistic solutions after distorting games.

The paper is organised as follows. After providing an overview of lower probabilities in Section 2, in Section 3 we recall the definition of the TV-distance and introduce the three aforementioned extensions for comparing lower probabilities. Later, in Sections 4÷6 we explore in detail the three proposals, and in Section 7 we show the application of our results in the context of coalitional game theory, by means of a penalised version of the supremum distance. We conclude the paper in Section 8 with some final comments.

2. Preliminaries

In this paper we consider a finite possibility space $\mathcal{X} = \{x_1, \ldots, x_n\}$. We use $\mathbb{P}(\mathcal{X})$ to denote the probability measures defined on $\mathcal{P}(\mathcal{X})$, the power set of $\mathcal{X}$, and $\mathbb{P}^*(\mathcal{X})$ for those probability measures P satisfying $P(A) > 0$ for any $A \neq \emptyset$.

A *lower probability* is a function $\underline{P} : \mathcal{P}(\mathcal{X}) \to [0, 1]$ that is monotone increasing and satisfies $\underline{P}(\emptyset) = 0$ and $\underline{P}(\mathcal{X}) = 1$. It is also called non-additive measure or capacity[?] in the literature; here we follow the standard terminology in the imprecise probability theory[?,?] . By conjugacy, a lower probability determines an *upper probability* $\overline{P}$ using the formula $\overline{P}(A) = 1 - \underline{P}(A^c)$ for any $A \subseteq \mathcal{X}$. A lower probability determines a (possibly empty) closed and convex set of probability measures, usually called its associated *credal set*, by:

$$\mathcal{M}(\underline{P}) = \{Q \in \mathbb{P}(\mathcal{X}) \mid Q(A) \geq \underline{P}(A)\ \forall A \subseteq \mathcal{X}\}.$$

The credal set is determined by its extreme points, that are those probability measures $P \in \mathcal{M}(\underline{P})$ that cannot be expressed as a non-trivial convex combination of elements in $\mathcal{M}(\underline{P})$. We shall denote the set of extreme points by $\text{ext}\big(\mathcal{M}(\underline{P})\big)$.

A lower (respectively, upper) probability may be interpreted as a lower (respectively, upper) bound for a probability measure that is only partially known, and the credal set includes all those probability measures compatible with $\underline{P}$ (and $\overline{P}$). This leads to two common consistency requirements on lower probabilities. On the one hand, $\underline{P}$ *avoids sure loss* if $\mathcal{M}(\underline{P}) \neq \emptyset$, meaning that there is at least one probability measure compatible with $\underline{P}$. On the other hand, $\underline{P}$ is *coherent* when $\underline{P}(A) = \min\{P(A) \mid P \in \mathcal{M}(\underline{P})\}$ for any $A \subseteq \mathcal{X}$, meaning that the bounds given by $\underline{P}$ are tight.

A lower probability $\underline{P}$ avoiding sure loss satisfies $\underline{P}(A) \leq \overline{P}(A)$ for any $A \subseteq \mathcal{X}$, and it can be used to determine a coherent lower probability $\underline{Q}$ using the *natural extension* $\underline{Q}(A) = \min\{P(A) \mid P \in \mathcal{M}(\underline{P})\}$ for any $A \subseteq \mathcal{X}$. It holds that $\mathcal{M}(\underline{Q}) = \mathcal{M}(\underline{P})$, and $\underline{Q}$ is the smallest coherent lower probability that dominates $\underline{P}$. Obviously, if $\underline{P}$ is coherent itself then it coincides with its natural extension.

The natural extension can also be used to extend a lower probability from events to real-valued functions $g : \mathcal{X} \to \mathbb{R}$, usually called *gambles*, by taking $\underline{P}(g) = \min\{P(g) \mid P \in \mathcal{M}(\underline{P})\}$, where $P(g)$ denotes the expectation of g with respect to P. Using the terminology of the imprecise probability theory, $\underline{P}$ evaluated on gambles is a *coherent lower prevision*[?,?] . The set of all gambles defined on $\mathcal{X}$ shall be denoted by $\mathcal{L}(\mathcal{X})$.

We should also mention that any closed and convex set of probability measures $\mathcal{M}$ determines a coherent lower probability $\underline{P}$ by $\underline{P}(A) = \min\{P(A) \mid P \in \mathcal{M}\}$, but in general $\mathcal{M} \subseteq \mathcal{M}(\underline{P})$. In particular, different credal sets may determine the same coherent lower probability.

A useful property that a coherent lower probability may satisfy is that of 2-monotonicity, meaning that for any $A, B \subseteq \mathcal{X}$ it satisfies the inequality

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

2-monotone lower probabilities possess a number of interesting properties that coherent lower probabilities need not satisfy; some of these are summarised in[?] . For example, their natural extension to gambles can be easily computed using the Choquet integral[?] . For any gamble g, there is a permutation σ of $\{1, \ldots, n\}$ such that $g(x_{\sigma(1)}) \geq \ldots \geq g(x_{\sigma(n)})$; denoting $A_i = \{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\}$ for $i = 1, \ldots, n$, $\alpha_i = g(x_{\sigma(i)}) - g(x_{\sigma(i+1)})$ for $i = 1, \ldots, n-1$ and $\alpha_n = g(x_{\sigma(n)})$, the Choquet integral is given by:

$$\text{(C)} \int g d\underline{P} = \min_{x\in\mathcal{X}} g(x) + \int_{\min g}^{\max g} \underline{P}(\{x \mid g(x) \geq y\})dy = \sum_{i=1}^{n} \alpha_i \underline{P}(A_i). \tag{1}$$

In what follows, we shall use $\underline{\mathbb{P}}(\mathcal{X})$ to denote the set of lower probabilities defined on $\mathcal{P}(\mathcal{X})$ and $\underline{\mathbb{P}}^*(\mathcal{X})$ for those lower probabilities $\underline{P}$ satisfying $\underline{P}(A) > 0$ for any $A \neq \emptyset$.

3. Extensions of the Total Variation distance

In this section we propose three different extensions of the Total Variation and supremum distances as methods for comparing lower probabilities. For this aim, we briefly review these distances (Section 3.1) and later introduce our generalisations (Section 3.2). Recall that a *distance* d on a space M is a function $d : M \times M \to \mathbb{R}$ satisfying the following properties:

- Non-negativity: $d(x, y) \geq 0$ for any $x, y \in M$.
- $d(x, y) = 0$ if and only if $x = y$.
- Symmetry: $d(x, y) = d(y, x)$ for any $x, y \in M$.
- Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$ for any $x, y, z \in M$.

3.1. *The Total Variation distance*

The Total Variation distance (TV-distance, for short) between probability measures is defined as:

$$d_{\text{TV}}(P, Q) = \max_{A\subseteq\mathcal{X}} |P(A) - Q(A)| \quad \forall P, Q \in \mathbb{P}(\mathcal{X}). \tag{TV}$$

It is a convex (in particular, quasi-convex) and continuous distance [**?**, Prop.2.1]. The continuity is referred to the product topology on $\mathbb{P}(\mathcal{X}) \times \mathbb{P}(\mathcal{X})$ based on the topology induced by the Euclidean distance on $\mathbb{P}(\mathcal{X})$, while the convexity is referred to the convex linear combination inherited from the ambient Euclidean space. An alternative expression of the TV-distance is the following:

$$d_{\text{TV}}(P, Q) = \frac{1}{2} \sum_{x\in\mathcal{X}} |P(\{x\}) - Q(\{x\})|. \tag{2}$$

Given a probability measure P_0 and a distortion parameter $\delta > 0$, the TV-distance determines a neighbourhood or distortion model[**?**,**?**] as:

$$B^{\delta}_{d_{\text{TV}}}(P_0) = \{Q \in \mathbb{P}(\mathcal{X}) \mid d_{\text{TV}}(Q, P_0) \leq \delta\}.$$

The continuity and quasi-convexity of the TV-distance implies that $B^{\delta}_{d_{\text{TV}}}(P_0)$ is a closed and convex set of probability measures [**?**, Prop.1], i.e. a credal set. Moreover, the lower probability it determines is given by:

$$\underline{P}_{\text{TV}}(A) = \max\{0, P_0(A) - \delta\} \qquad \forall A \subset \mathcal{X}$$

and $\underline{P}_{\mathrm{TV}}(\mathcal{X}) = 1$, and the neighbourhood can be expressed as $B^{\delta}_{d_{\mathrm{TV}}}(P_0) = \mathcal{M}(\underline{P}_{\mathrm{TV}})$. Indeed, $\underline{P}_{\mathrm{TV}}$ is not only coherent but also 2-monotone. This model, known as the TV-model, was thoroughly investigated in [**?**, Sect.2].

The TV-distance considers the maximum absolute difference between the probabilities of all the events. If we focus on the singletons, we obtain the distance associated with the supremum norm on the associated mass functions, i.e.:

$$d_\infty(P,Q) = \max_{x\in\mathcal{X}} |P(\{x\}) - Q(\{x\})| \qquad \forall P,Q \in \mathbb{P}(\mathcal{X}). \tag{3}$$

Interestingly, this distance, that is also continuous and convex, is related to a modified version of the TV-distance, since it can also be expressed as:

$$d_\infty(P,Q) = \max_{A\subseteq\mathcal{X}} \frac{|P(A) - Q(A)|}{|A|} \qquad \forall P,Q \in \mathbb{P}(\mathcal{X}). \tag{4}$$

To see the equality between (3) and (4), note that for any non-empty event A, it holds that:

$$\begin{aligned}\frac{|P(A) - Q(A)|}{|A|} &= \frac{|\sum_{x\in A} P(\{x\}) - \sum_{x\in A} Q(\{x\})|}{|A|} \\ &\le \sum_{x\in A} \frac{|P(\{x\}) - Q(\{x\})|}{|A|} \le \max_{x\in A} |P(\{x\}) - Q(\{x\})|,\end{aligned}$$

from which it follows that the expression in Eq. (4) is dominated by that in Eq. (3). The converse inequality follows considering that Eq. (3) corresponds to the case where the maximum on Eq. (4) is taken on singletons only.

We may thus interpret the supremum distance as a sort of penalised TV-distance, where the differences are weighted by the cardinality of the event.

3.2. *The TV distance for lower probabilities*

As we discussed in the introduction, our aim in this paper is to extend the TV-distance to compare lower probabilities. We shall explore the following possibilities:

Minimum Total Variation: Our first option computes the minimum of the TV-distances between any pair of probability measures in the credal sets:

$$d^{\min}_{\mathrm{TV}}(\underline{P},\underline{Q}) := \min_{\substack{P\in\mathcal{M}(\underline{P})\\ Q\in\mathcal{M}(\underline{Q})}} d_{\mathrm{TV}}(P,Q) \qquad \forall \underline{P},\underline{Q} \in \underline{\mathbb{P}}(\mathcal{X}). \tag{5}$$

Maximum Discrepancy: If we replace the minimum by a maximum in the previous proposal, we get:

$$d^{\max}_{\mathrm{TV}}(\underline{P},\underline{Q}) := \max_{\substack{P\in\mathcal{M}(\underline{P})\\ Q\in\mathcal{M}(\underline{Q})}} d_{\mathrm{TV}}(P,Q) \qquad \forall \underline{P},\underline{Q} \in \underline{\mathbb{P}}(\mathcal{X}). \tag{6}$$

This formula gives the maximum difference between any pair of probability measures included in the credal sets.

Minimum Supremum: Our third proposal takes into consideration the minimum of the supremum distances between each pair of probability measures in the credal sets. In other words:

$$d_{\infty}^{\min}(\underline{P},\underline{Q}) := \min_{\substack{P\in\mathcal{M}(\underline{P})\\ Q\in\mathcal{M}(\underline{Q})}} d_{\infty}(P,Q) \qquad \forall \underline{P},\underline{Q}\in\underline{\mathbb{P}}(\mathcal{X}). \tag{7}$$

Note that the expressions in Eqs. (5)÷(7) are well defined because the maxima and minima for $P \in \mathcal{M}(\underline{P})$ and $Q \in \mathcal{M}(\underline{Q})$ are indeed attained, since the product of the credal sets is compact and d_{TV} and d_{∞} are continuous. Clearly $d_{\mathrm{TV}}^{\min}$ and $d_{\mathrm{TV}}^{\max}$ are extensions of the TV-distance, in the sense that $d_{\mathrm{TV}}^{\min}(P,Q) = d_{\mathrm{TV}}^{\max}(P,Q) = d_{\mathrm{TV}}(P,Q)$ for $P,Q \in \mathbb{P}(\mathcal{X})$. Similarly, d_{∞} extends the supremum distance because $d_{\infty}^{\min}(P,Q) = d_{\infty}(P,Q)$.

As observed in [**?**, Sect.3.3], any distance between probability measures may be used to compare two credal sets by considering the minimum distance between pairs of elements $P \in \mathcal{M}(\underline{P})$ and $Q \in \mathcal{M}(\underline{Q})$. This is the path followed by the first and third options above. In fact, the first one is considered in[?] and it also appears in the particular case when one of the lower probabilities is in correspondence with a single probability measure in[?] . The latter situation was thoroughly analysed in our recent contributions[?,?] . The second option considers the maximum over the credal sets of the TV-distances, and appeared in[?] under the name of maximal (TV-)distance.

Also, these extensions are closely related to the Haussdorf distance induced by the TV-distance, given by[?] :

$$\max\left\{ \max_{P\in\mathcal{M}(\underline{P})} \min_{Q\in\mathcal{M}(\underline{Q})} d_{\mathrm{TV}}(P,Q), \max_{Q\in\mathcal{M}(\underline{Q})} \min_{P\in\mathcal{M}(\underline{P})} d_{\mathrm{TV}}(Q,P) \right\}.$$

Specifically, by replacing the maxima over the credal sets in the previous expression with minima and using the symmetry of the TV-distance, we obtain the first option. Moreover, if the TV-distance is replaced by the supremum distance, we get the third option. If instead we replace the minima with maxima in the expression of the Hausdorff distance we recover the Maximum Discrepancy.

Note that the above definitions depend entirely on the credal sets associated with the lower probabilities $\underline{P},\underline{Q}$; as a consequence, if $\underline{P}',\underline{Q}' \in \underline{\mathbb{P}}(\mathcal{X})$ satisfy $\mathcal{M}(\underline{P}) = \mathcal{M}(\underline{P}')$ and $\mathcal{M}(\underline{Q}) = \mathcal{M}(\underline{Q}')$, we obtain $d_{\mathrm{TV}}^{\min}(\underline{P},\underline{Q}) = d_{\mathrm{TV}}^{\min}(\underline{P}',\underline{Q}')$, $d_{\mathrm{TV}}^{\max}(\underline{P},\underline{Q}) = d_{\mathrm{TV}}^{\max}(\underline{P}',\underline{Q}')$ and $d_{\infty}^{\min}(\underline{P},\underline{Q}) = d_{\infty}^{\min}(\underline{P}',\underline{Q}')$. In particular, the discrepancies between two lower probabilities coincide with those between their natural extensions.

In the reminder of this manuscript we shall explore the properties of the functionals defined in Eqs.(5)÷(7) and analyse their suitability as a distortion model for lower probabilities.

4. Minimum Total Variation

Our analysis begins with the Minimum Total Variation (Minimum TV, for short).

We start by observing that whenever $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) \neq \emptyset$, we get $d_{\mathrm{TV}}^{\min}(\underline{P},\underline{Q}) = 0$; as a consequence, $d_{\mathrm{TV}}^{\min}(\underline{P},\underline{Q}) = 0$ does not imply $\underline{P} = \underline{Q}$, so the Minimum TV is not a distance, in spite of satisfying $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = 0$ whenever $\underline{P} = \underline{Q}$ and the symmetry axiom. It is also not difficult to show that the triangle inequality does not hold either: it suffices to consider three coherent lower probabilities $\underline{P}, \underline{Q}, \underline{R} \in \underline{\mathbb{P}}(\mathcal{X})$ such that $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{R})$ and $\mathcal{M}(\underline{Q}) \cap \mathcal{M}(\underline{R})$ are non-empty but $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) = \emptyset$. In this way, $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{R}) = d_{\mathrm{TV}}^{\min}(\underline{R}, \underline{Q}) = 0$ but $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) > 0$, meaning that the triangle inequality is violated.

Next, we delve into the properties of the Minimum TV, and establish some alternative expressions.

4.1. *Alternative expressions*

Since the value $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q})$ is attained on some $P \in \mathcal{M}(\underline{P}), Q \in \mathcal{M}(\underline{Q})$ and d_{TV} is a distance, it follows that $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = 0$ if and only if $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) \neq \emptyset$. We focus then our attention on lower probabilities whose associated credal sets are disjoint. Let us establish that in that case the Minimum TV is attained at the boundaries of the credal sets, that we denote in this paper using the symbol ∂.

Proposition 1. *Let $\underline{P}$ and $\underline{Q}$ be two lower probabilities avoiding sure loss such that $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) = \emptyset$. If $P \in \mathcal{M}(\underline{P})$ and $Q \in \mathcal{M}(\underline{Q})$ satisfy $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = d_{\mathrm{TV}}(P, Q)$, then $P \in \partial\mathcal{M}(\underline{P})$ and $Q \in \partial\mathcal{M}(\underline{Q})$. As a consequence,*

$$d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = \min_{\substack{P \in \partial\mathcal{M}(\underline{P}) \\ Q \in \partial\mathcal{M}(\underline{Q})}} d_{\mathrm{TV}}(P, Q).$$

Proof. Ex-absurdo, assume that $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = d_{\mathrm{TV}}(P, Q)$ where at least one of $P \in \mathcal{M}(\underline{P}) \setminus \partial\mathcal{M}(\underline{P})$ or $Q \in \mathcal{M}(\underline{Q}) \setminus \partial\mathcal{M}(\underline{Q})$ holds. Assume that $P \in \mathcal{M}(\underline{P}) \setminus \partial\mathcal{M}(\underline{P})$; the other case follows analogously. Then P belongs to the interior of $\mathcal{M}(\underline{P})$, and as a consequence it must be $\overline{P}(A) > P(A) > \underline{P}(A)$ for any $A \neq \emptyset, \mathcal{X}$, given that there must be some neighbourhood of P included within $\mathcal{M}(\underline{P})$.

Let us consider the following partition of $\mathcal{X}$:

$$\mathcal{X}_1 = \{x \in \mathcal{X} \mid P(\{x\}) > Q(\{x\})\}, \qquad \mathcal{X}_2 = \{x \in \mathcal{X} \mid P(\{x\}) < Q(\{x\})\},$$
$$\mathcal{X}_3 = \{x \in \mathcal{X} \mid P(\{x\}) = Q(\{x\})\}.$$

Since $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) = \emptyset$, it follows that $P \neq Q$, implying that both $\mathcal{X}_1$ and $\mathcal{X}_2$ are non-empty. Thus, applying Eq. (2) we obtain:

$$\begin{aligned} d_{\mathrm{TV}}^{\min}(P, Q) = \frac{1}{2} \sum_{x \in \mathcal{X}} |P(\{x\}) - Q(\{x\})| &= \frac{1}{2} \sum_{x \in \mathcal{X}_1} |P(\{x\}) - Q(\{x\})| \\ &+ \frac{1}{2} \sum_{x \in \mathcal{X}_2} |P(\{x\}) - Q(\{x\})| + \frac{1}{2} \sum_{x \in \mathcal{X}_3} |P(\{x\}) - Q(\{x\})| \\ &= \frac{1}{2} \sum_{x \in \mathcal{X}_1} \big(P(\{x\}) - Q(\{x\})\big) + \frac{1}{2} \sum_{x \in \mathcal{X}_2} \big(Q(\{x\}) - P(\{x\})\big). \end{aligned}$$

Take $\varepsilon > 0$ satisfying:

$$\varepsilon < \frac{1}{2} \min_{A \neq \emptyset, \mathcal{X}} \big(P(A) - \underline{P}(A)\big). \tag{8}$$

Fix $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$ and define:

$$P^*(\{x\}) = \begin{cases} P(\{x_1\}) - \varepsilon, & \text{if } x = x_1, \\ P(\{x_2\}) + \varepsilon, & \text{if } x = x_2, \\ P(\{x\}), & \text{if } x \neq x_1, x_2. \end{cases}$$

By construction $P^*(\{x_1\}) \geq \underline{P}(\{x_1\}) \geq 0$ and since $\sum_{x \in \mathcal{X}} P^*(\{x\}) = \sum_{x \in \mathcal{X}} P(\{x\}) = 1$ we deduce that P^* is a probability measure; moreover, from Eq. (8), it satisfies $P^*(A) \geq P(A) - \varepsilon \geq \underline{P}(A)$ for any $A \neq \emptyset, \mathcal{X}$. Thus, $P^* \in \mathcal{M}\big(\underline{P}\big)$. Moreover, using Eq. (2) we obtain:

$$\begin{aligned} d_{\mathrm{TV}}(P^*, Q) = \frac{1}{2} \sum_{x \in \mathcal{X}} |P^*(\{x\}) - Q(\{x\})| &= \frac{1}{2}\big((P(\{x_1\}) - \varepsilon) - Q(\{x_1\})\big) \\ &+ \frac{1}{2}\big(Q(\{x_2\}) - (P(\{x_2\}) + \varepsilon)\big) + \frac{1}{2} \sum_{x \neq x_1, x_2} |P(\{x\}) - Q(\{x\})| \\ &= \frac{1}{2} \sum_{x \in \mathcal{X}} |P(\{x\}) - Q(\{x\})| - \varepsilon < d_{\mathrm{TV}}(P, Q) = d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}), \end{aligned}$$

a contradiction with $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = d_{\mathrm{TV}}(P, Q)$. □

We should mention that the same result does not necessarily hold when $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) \neq \emptyset$, because in that case the value $d_{\mathrm{TV}}^{\min}(\underline{P}, \underline{Q}) = 0$ will be attained by considering $d_{\mathrm{TV}}(P, P)$ for any $P \in \mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q})$, be it in the boundary of these credal sets or not.

Our next example shows that Proposition 1 cannot be refined, in the sense that the Minimum TV is not necessarily reached at the extreme points of the credal sets.

Example 1. Let $\mathcal{X} = \{x_1, x_2, x_3\}$ and consider the coherent lower probabilities $\underline{P}$ and $\underline{Q}$ given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}(A)$	0.2	0.1	0.1	0.3	0.3	0.8
$\underline{Q}(A)$	0.3	0.3	0.4	0.6	0.7	0.7

Their credal sets are given by $\mathcal{M}(\underline{P}) = \{P_\alpha \mid \alpha \in [0,1]\}$, where $P_\alpha = (0.2, 0.6\alpha + 0.1, 0.7 - 0.6\alpha)$, and $\mathcal{M}(\underline{Q}) = \{Q\}$, where $Q = (0.3, 0.3, 0.4)$. We deduce that:

$$\operatorname{ext}\big(\mathcal{M}(\underline{P})\big) = \{P_0 = (0.2, 0.1, 0.7), P_1 = (0.2, 0.7, 0.1)\}, \quad \operatorname{ext}\big(\mathcal{M}(\underline{Q})\big) = \{Q\}.$$

Taking $P_{0.5} = (0.2, 0.4, 0.4) \in \mathcal{M}(\underline{P})$ we get that $d_{\mathrm{TV}}(P_{0.5}, Q) = 0.1$ while $d_{\mathrm{TV}}(P_0, Q) = 0.3$, $d_{\mathrm{TV}}(P_1, Q) = 0.4$. Thus, the minimum distance is not attained on the extreme points of the credal sets. ♦

Now, we look for an alternative expression of $d_{\mathrm{TV}}^{\min}(\underline{P},\underline{Q})$ in terms of the values that $\underline{P}$ and $\underline{Q}$ take on events. For this purpose, let us define:

$$d'_{\mathrm{TV}}(Q,\underline{P}) := \max_{A\subseteq\mathcal{X}} \big(\underline{P}(A) - Q(A)\big), \tag{9}$$

which is also an extension of the TV-distance to the case where its first argument is a precise probability and the second argument is a lower probability. Moreover, whenever $\underline{P}$ is coherent, it can be written as:

$$d'_{\mathrm{TV}}(Q,\underline{P}) = \max_{A\subseteq\mathcal{X}}\left(\min_{P\in\mathcal{M}(\underline{P})} P(A) - Q(A)\right) = \max_{A\subseteq\mathcal{X}}\min_{P\in\mathcal{M}(\underline{P})}\big(P(A) - Q(A)\big),$$

for each $Q \in \mathbb{P}(\mathcal{X})$. From this expression we see that d'_{TV} arises by exchanging the minimum and maximum in the Minimum TV.

We next prove, in virtue of a minimax theorem relating d'_{TV} and $d_{\mathrm{TV}}^{\min}$, that these two coincide under 2-monotonicity. Moreover, the maximum over events coincides with the maximum over the gambles taking values in $[0,1]$ of the respective analogous expressions for the natural extensions to gambles.

Proposition 2. *Let $\underline{P}$ and $\underline{Q}$ be two 2-monotone lower probabilities, and let $\mathcal{H} = \{g : \mathcal{X} \to [0,1]\}$ be the set of gambles taking values in $[0,1]$. It holds that:*

$$d_{\mathrm{TV}}^{\min}(\underline{P},\underline{Q}) = \max_{A\subseteq\mathcal{X}}\big(\underline{P}(A) - \overline{Q}(A)\big) = \max_{g\in\mathcal{H}}\big(\underline{P}(g) - \overline{Q}(g)\big), \tag{10}$$

where $\underline{P}(g), \underline{Q}(g)$ are the natural extensions of $\underline{P}, \underline{Q}$ to gambles.

Proof. First of all, it is a consequence of [**?**, Thm.1] and the 2-monotonicity of $\underline{P}$ that $d'_{\mathrm{TV}}(Q,\underline{P}) = d_{\mathrm{TV}}^{\min}(Q,\underline{P})$ for each $Q \in \mathbb{P}(\mathcal{X})$, which yields:

$$\begin{aligned} d_{\mathrm{TV}}^{\min}(\underline{P},\underline{Q}) &= \min_{\substack{P\in\mathcal{M}(\underline{P})\\ Q\in\mathcal{M}(\underline{Q})}} d_{\mathrm{TV}}(Q,P) = \min_{Q\in\mathcal{M}(\underline{Q})} d_{\mathrm{TV}}^{\min}(Q,\underline{P}) = \min_{Q\in\mathcal{M}(\underline{Q})} d'_{\mathrm{TV}}(Q,\underline{P}) \\ &= \min_{Q\in\mathcal{M}(\underline{Q})}\max_{A\subseteq\mathcal{X}}\big(\underline{P}(A) - Q(A)\big). \end{aligned}$$

We aim to prove that

$$\min_{Q\in\mathcal{M}(\underline{Q})}\max_{A\subseteq\mathcal{X}}\big(\underline{P}(A) - Q(A)\big) = \max_{A\subseteq\mathcal{X}}\min_{Q\in\mathcal{M}(\underline{Q})}\big(\underline{P}(A) - Q(A)\big), \tag{11}$$

which will give the first equality in Eq. (10). For this, note that $\mathcal{H}$ includes the set of indicator functions of the events $A \subseteq \mathcal{X}$. We define the map:

$$\begin{aligned} f_{\underline{P}} : \mathcal{M}(\underline{Q}) \times \mathcal{H} &\to \mathbb{R} \\ (Q,g) &\hookrightarrow f_{\underline{P}}(Q,g) = \underline{P}(g) - Q(g), \end{aligned}$$

where $\underline{P}(g)$ denotes the natural extension of the lower probability $\underline{P}$ to gambles and $Q(g)$ is the expectation of g with respect to the probability measure Q. We shall prove that the equality:

$$\max_{g\in\mathcal{H}}\min_{Q\in\mathcal{M}(\underline{Q})} f_{\underline{P}}(Q,g) = \min_{Q\in\mathcal{M}(\underline{Q})}\max_{g\in\mathcal{H}} f_{\underline{P}}(Q,g) \tag{12}$$

is satisfied and that it is equivalent to Eq. (11).

Steps 1÷3 below prove that the hypotheses of the minimax theorem in [**?**, App. E6] are fulfilled, implying that Eq. (12) holds. The remaining two steps show that the minimax and maximin on events, respectively, may be computed by taking the maximum over gambles in $\mathcal{H}$ rather than on events. From the minimax theorem over gambles, this implies that Eq. (11) is satisfied. Thus, the verification of these two equations trivially yields the desired result.

Step 1 $\mathcal{M}(\underline{Q})$ and $\mathcal{H}$ are convex and compact sets in $\mathbb{R}^n$, where $n = |\mathcal{X}|$.

Step 2 Given $g \in \mathcal{H}$ and $\mu \in \mathbb{R}$, we aim to prove that $C_{g,\mu} = \{Q \in \mathcal{M}(\underline{Q}) \mid f_{\underline{P}}(Q,g) \leq \mu\} = \{Q \in \mathcal{M}(\underline{Q}) \mid Q(g) \geq \underline{P}(g) - \mu\}$ is convex and closed. Let $\alpha \in [0,1]$ and $Q_1, Q_2 \in C_{g,\mu}$. Then,

$$\alpha Q_1(g) + (1-\alpha)Q_2(g) \geq \alpha(\underline{P}(g) - \mu) + (1-\alpha)(\underline{P}(g) - \mu) = \underline{P}(g) + \mu,$$

so $\alpha Q_1 + (1-\alpha)Q_2 \in C_{g,\mu}$, hence $C_{g,\mu}$ is convex.
Consider a sequence $(Q_k)_k \subset C_{g,\mu}$ for which $Q_k \to Q$ for some $Q \in \mathbb{P}(\mathcal{X})$ and let us prove that $Q \in C_{g,\mu}$. Since $(Q_k)_k$ is included in $\mathcal{M}(\underline{Q})$ and this is closed set, we have $Q \in \mathcal{M}(\underline{P})$. Also, $(Q_k)_k \to Q$ implies that $Q_k(g) \to Q(g)$ for every $g \in \mathcal{H}$. Thus, taking limits we get:

$$Q_k(g) \geq \underline{P}(g) - \mu \quad \forall k \in \mathbb{N} \Rightarrow Q(g) \geq \underline{P}(g) - \mu.$$

Hence, $Q \in C_{g,\mu}$ and $C_{g,\mu}$ is closed.

Step 3 Given $Q \in \mathcal{M}(\underline{Q})$ and $\mu \in \mathbb{R}$, let us prove that $S_{P,\mu} = \{g \in \mathcal{H} \mid f_{\underline{P}}(Q,g) \geq \mu\} = \{g \in \mathcal{H} \mid \underline{P}(g) \geq Q(g) + \mu\}$ is convex and closed. Let $\alpha \in [0,1]$ and $g_1, g_2 \in S_{P,\mu}$. Then, since $\underline{P}$ is concave [**?**, Sec.2.6.1(g)] and Q is linear,

$$\begin{aligned}\underline{P}(\alpha g_1 + (1-\alpha)g_2) &\geq \alpha\underline{P}(g_1) + (1-\alpha)\underline{P}(g_2) \\ &\geq \alpha(Q(g_1) + \mu) + (1-\alpha)(Q(g_2) + \mu) = Q(g) + \mu,\end{aligned}$$

so $\alpha g_1 + (1-\alpha)g_2 \in S_{P,\mu}$ and $S_{P,\mu}$ is convex.
Consider a sequence $(g_k)_k \subset S_{P,\mu}$ such that $g_k \to g$ for certain $g \in \mathcal{L}(\mathcal{X})$. Since $(g_k)_k \subset \mathcal{H}$ and $\mathcal{H}$ is closed, then $g \in \mathcal{H}$. Also, from [**?**, Sec.2.6.1(l)], Q being continuous and taking limits, we deduce:

$$\underline{P}(g_k) \geq Q(g_k) + \mu \quad \forall k \in \mathbb{N} \Rightarrow \underline{P}(g) \geq Q(g) + \mu,$$

so $g \in S_{P,\mu}$ and this is a closed set.

Step 4 Let us next prove that the equality

$$\max_{g \in \mathcal{H}} \min_{Q \in \mathcal{M}(\underline{Q})} f_{\underline{P}}(Q,g) = \max_{A \subseteq \mathcal{X}} \min_{Q \in \mathcal{M}(\underline{Q})} \big(\underline{P}(A) - Q(A)\big).$$

holds. Since the indicator functions belong to $\mathcal{H}$, we clearly have that:

$$\max_{g \in \mathcal{H}} \min_{Q \in \mathcal{M}(\underline{Q})} f_{\underline{P}}(Q,g) \geq \max_{A \subseteq \mathcal{X}} \min_{Q \in \mathcal{M}(\underline{Q})} \big(\underline{P}(A) - Q(A)\big).$$

For the converse inequality, note that

$$\max_{A \subseteq \mathcal{X}} \min_{Q \in \mathcal{M}(\underline{P})} (\underline{P}(A) - Q(A)) \leq \delta \Leftrightarrow \forall A \subseteq \mathcal{X} \quad \underline{P}(A) - \overline{Q}(A) \leq \delta;$$

if we decompose $g \in \mathcal{H}$ in the manner described in Eq. (1) and use the expression of the Choquet integral, we conclude that

$$\underline{P}(g) - \overline{Q}(g) = \sum_{i=1}^{n} \alpha_i \big(\underline{P}(A_i) - \overline{Q}(A_i)\big) \leq \delta \sum_{i=1}^{n} \alpha_i \leq \delta.$$

Hence, $\max_{g\in\mathcal{H}} \min_{Q\in\mathcal{M}(\underline{Q})} f_{\underline{P}}(Q,g) \leq \delta$ for every $\delta \geq 0$.

Step 5 Let us prove that

$$\min_{Q\in\mathcal{M}(\underline{Q})} \max_{g\in\mathcal{H}} f_{\underline{P}}(Q,g) = \min_{Q\in\mathcal{M}(\underline{Q})} \max_{A\subseteq\mathcal{X}} \big(\underline{P}(A) - Q(A)\big).$$

One of the inequalities is trivial, while expressing any given $g = \sum_{i=1}^{n} \alpha_i I_{A_i} \in \mathcal{H}$ as in the previous step, an analogous procedure allows us to deduce the converse inequality, since:

$$\min_{Q\in\mathcal{M}(\underline{Q})} \max_{A\subseteq\mathcal{X}} \big(\underline{P}(A) - Q(A)\big) \leq \delta$$

is equivalent to the existence of $Q \in \mathcal{M}(\underline{Q})$ such that $\underline{P}(A) - Q(A) \leq \delta$ for every $A \subseteq \mathcal{X}$ and, for this probability measure, we deduce, again using Eq. (1), that:

$$f_{\underline{P}}(Q,g) = \underline{P}(g) - Q(g) = \sum_{i=1}^{n} \alpha_i \big(\underline{P}(A_i) - Q(A_i)\big) \leq \delta \quad \forall g \in \mathcal{H}.$$

Thus, $\min_{P\in\mathcal{M}(\underline{P})} \max_{g\in\mathcal{H}} f_{\underline{P}}(Q,g) \leq \delta$, for every $\delta \geq 0$. □

Despite the previous result, the equality between $d_{\mathrm{TV}}^{\min}$ and d'_{TV} does not hold when one of the lower probabilities is not 2-monotone. For an explicit counterexample we refer to [**?**, Ex.3], where one of the arguments is a probability measure and the other one is a coherent lower probability that is not 2-monotone.

4.2. *Distorting lower probabilities with the Minimum TV*

In our previous contributions[?,?] we investigated how $d_{\mathrm{TV}}^{\min}$ can be used to distort a lower probability by creating a neighbourhood around a lower probability $\underline{P}$. We next summarise our results from[?] . In order to distort $\underline{P} \in \underline{\mathbb{P}}(\mathcal{X})$, we may apply the Minimum TV to obtain the function:

$$d_{\mathrm{TV}}^{\min}(Q,\underline{P}) := \min_{P\in\mathcal{M}(\underline{P})} d_{\mathrm{TV}}(Q,P) \qquad \forall Q \in \mathbb{P}(\mathcal{X}).$$

It determines a neighbourhood around the lower probability $\underline{P}$ using a distortion parameter $\delta \geq 0$ as:

$$B^{\delta}_{d_{\mathrm{TV}}^{\min}}(\underline{P}) = \big\{Q \in \mathbb{P}(\mathcal{X}) \mid d_{\mathrm{TV}}^{\min}(Q,\underline{P}) \leq \delta\big\},$$

which is closed and convex, hence a credal set, since d_{TV} is quasi-convex and continuous [**?**, Prop.2.1]. We observe that by construction $B^{\delta}_{d_{\mathrm{TV}}^{\min}}(\underline{P})$ is equal to

$$\{Q \in \mathbb{P}(\mathcal{X}) \mid \exists P \in \mathcal{M}(\underline{P}) : d_{\mathrm{TV}}(Q,P) \leq \delta\} = \bigcup_{P\in\mathcal{M}(\underline{P})} B^{\delta}_{d_{\mathrm{TV}}}(P),$$

from which it follows that $B^{\delta}_{d^{\min}_{\mathrm{TV}}}(\underline{P}) \supseteq \mathcal{M}(\underline{P})$ for any $\delta \geq 0$.

Similarly, if we consider d'_{TV}, we may define a neighbourhood around $\underline{P}$ with distortion parameter δ as:

$$B^{\delta}_{d'_{\mathrm{TV}}}(\underline{P}) = \{Q \in \mathbb{P}(\mathcal{X}) \mid d'_{\mathrm{TV}}(Q, \underline{P}) \leq \delta\} = \Big\{Q \in \mathbb{P}(\mathcal{X}) \mid \max_{A \subseteq \mathcal{X}} \big(\underline{P}(A) - Q(A)\big) \leq \delta\Big\}.$$

From [?, Cor.1], the lower probabilities associated with $B^{\delta}_{d^{\min}_{\mathrm{TV}}}(\underline{P})$ and $B^{\delta}_{d'_{\mathrm{TV}}}(\underline{P})$ coincide, for $\underline{P}$ coherent, and are given by:

$$\underline{Q}_{d'_{\mathrm{TV}}}(A) = \max\big\{\underline{P}(A) - \delta, 0\big\} \quad \forall A \neq \mathcal{X}, \quad \underline{Q}_{d'_{\mathrm{TV}}}(\mathcal{X}) = 1.$$

In general, this expression completely characterises the neighbourhood of d'_{TV} [?, Prop.8], in the sense that $B^{\delta}_{d'_{\mathrm{TV}}}(\underline{P}) = \mathcal{M}\big(\underline{Q}_{d'_{\mathrm{TV}}}\big)$. However, even if both $B^{\delta}_{d^{\min}_{\mathrm{TV}}}(\underline{P})$ and $B^{\delta}_{d'_{\mathrm{TV}}}(\underline{P})$ induce the same lower probability $\underline{Q}_{d'_{\mathrm{TV}}}$, these two credal sets may not coincide (see [?, Ex.3]). On the other hand, when $\underline{P}$ is 2-monotone we do have $B^{\delta}_{d^{\min}_{\mathrm{TV}}}(\underline{P}) = B^{\delta}_{d'_{\mathrm{TV}}}(\underline{P})$ [?, Thm.1]. Moreover, in that case [?, Prop.5] $\underline{Q}_{d'_{\mathrm{TV}}}$ is 2-monotone as well. The equality between the credal sets goes in line with the statement in Proposition 2, where we proved that under 2-monotonicity $d^{\min}_{\mathrm{TV}}$ and d'_{TV} are connected.

5. Maximum Discrepancy

We now shift our attention to the Maximum Discrepancy. We must first of all remark that $d^{\max}_{\mathrm{TV}}$ is not a distance between coherent lower probabilities: while it clearly satisfies symmetry and (in contradistinction with the Minimum TV) the triangle inequality, $d^{\max}_{\mathrm{TV}}(\underline{P}, \underline{P})$ equals 0 if and only if $\mathcal{M}(\underline{P})$ is a singleton, being $d^{\max}_{\mathrm{TV}}(\underline{P}, \underline{P}) > 0$ otherwise.

In what follows, we establish more operative expressions for the Maximum Discrepancy and use it as a tool to distort lower probabilities.

5.1. *Alternative expressions*

The following result, shown in[?] , provides an alternative expression for the Maximum Discrepancy in terms of the direct comparison of the lower and upper probabilities, and also in terms of that of the lower and upper previsions of gambles taking values in $[0, 1]$.

Proposition 3. *[?, Sec.3.1] Let $\underline{P}$ and $\underline{Q}$ be two coherent lower probabilities. Taking $\mathcal{H} = \{g : \mathcal{X} \to [0, 1]\}$, it holds that:*

$$d^{\max}_{\mathrm{TV}}(\underline{P}, \underline{Q}) = \max_{A \subseteq \mathcal{X}} |\underline{P}(A) - \overline{Q}(A)| = \max_{g \in \mathcal{H}} |\underline{P}(g) - \overline{Q}(g)| \,. \tag{13}$$

As we mentioned in Section 3, the Maximum Discrepancy between two lower probabilities depends on their credal sets, whence it follows that it does not vary when we consider two pairs of lower probabilities associated with the same credal sets.

For this reason, Eq. (13) does not hold when either $\underline{P}$ or $\underline{Q}$ avoids sure loss but is not coherent. We illustrate this in our next example.

Example 2. Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and let $\underline{P}$ be the lower probability given by

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}(A)$	0.2	0.3	0.4	0.6	0.7	0.7

Then $\mathcal{M}(\underline{P}) = \{P\}$, where P is given by the mass function $(0.3, 0.3, 0.4)$. As a consequence, $\underline{P}$ avoids sure loss but is not coherent. It follows that

$$d_{\mathrm{TV}}^{\max}(\underline{P}, P) = 0 < \max_{A \subseteq \mathcal{X}} |\underline{P}(A) - P(A)| = |\underline{P}(\{x_1\}) - P(\{x_1\})| = 0.1,$$

and therefore Eq. (13) does not hold. ♦

However, if $\underline{P}'$ and $\underline{Q}'$ are the natural extensions on events of $\underline{P}$ and $\underline{Q}$, we deduce from Proposition 3 that

$$d_{\mathrm{TV}}^{\max}(\underline{P}, \underline{Q}) = d_{\mathrm{TV}}^{\max}(\underline{P}', \underline{Q}') = \max_{A \subseteq \mathcal{X}} \left| \underline{P}'(A) - \overline{Q}'(A) \right|.$$

Since any coherent lower probability can be obtained as the lower envelope of the extreme points of its credal set, this allows us to conclude the following:

Corollary 1. *If $\underline{P}$ and $\underline{Q}$ are two lower probabilities avoiding sure loss, then:*

$$d_{\mathrm{TV}}^{\max}(\underline{P}, \underline{Q}) = \max_{\substack{P \in \mathrm{ext}(\mathcal{M}(\underline{P})) \\ Q \in \mathrm{ext}(\mathcal{M}(\underline{Q}))}} d_{\mathrm{TV}}(P, Q). \tag{14}$$

The expression in Eq. (14) notably simplifies the computation of the Maximum Discrepancy, as it implies that we only need to compute the TV-distance between pairs of extreme points. It is known that the maximum number of extreme points of a coherent lower probability is $n!$[?, ?] , and they can be obtained by computing the feasible region of a linear programming problem. Moreover, for some particular cases of coherent lower probabilities (e.g. 2-monotone[?] , completely monotone[?] , p-boxes[?] , possibility measures[?] , ...) the maximum number of extreme points is smaller and they can moreover be easily computed.

5.2. *Distorting lower probabilities with the Maximum Discrepancy*

Following the same reasoning as in Section 4.2 with the Minimum TV, we use the Maximum Discrepancy for creating a neighbourhood around a coherent lower probability. For this aim, we compare probability measures Q and lower probabilities $\underline{P}$ as:

$$d_{\mathrm{TV}}^{\max}(Q, \underline{P}) := \max_{P \in \mathcal{M}(\underline{P})} d_{\mathrm{TV}}(Q, P).$$

In this way, the neighbourhood around the lower probability $\underline{P}$ with distortion parameter δ is given by:

$$B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) = \{Q \in \mathbb{P}(\mathcal{X}) \mid d^{\max}_{\mathrm{TV}}(Q,\underline{P}) \leq \delta\}.$$

Since $d^{\max}_{\mathrm{TV}}(Q,\underline{P}) > 0$ unless $\underline{P}$ is a precise probability and coincides with Q, we will focus in this section on the case of $\delta > 0$.

By construction, the above neighbourhood can be expressed as:

$$\{Q \in \mathbb{P}(\mathcal{X}) \mid \forall P \in \mathcal{M}(\underline{P}) : d_{\mathrm{TV}}(Q,P) \leq \delta\} = \bigcap_{P \in \mathcal{M}(\underline{P})} B^{\delta}_{d_{\mathrm{TV}}}(P). \tag{15}$$

Next, we investigate some properties of the neighbourhood $B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$. First of all, note that whenever $\underline{P}$ is coherent, from Proposition 3 we get:

$$d^{\max}_{\mathrm{TV}}(Q,\underline{P}) = \max_{A \subseteq \mathcal{X}} \left|Q(A) - \overline{P}(A)\right|,$$

leading to an alternative expression for $B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$:

$$B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) = \left\{Q \in \mathbb{P}(\mathcal{X}) \mid \max_{A \subseteq \mathcal{X}} |Q(A) - \overline{P}(A)| \leq \delta\right\}.$$

Our next example uses the above expression to show two facts that may appear counter-intuitive at first: that $B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$ may be empty, and that, when it is not, it may not be a superset of $\mathcal{M}(\underline{P})$, not even for a 2-monotone $\underline{P}$.

Example 3. Let $\mathcal{X} = \{x_1, x_2, x_3\}$, and consider the following lower probability $\underline{P}$ and its conjugate upper probability $\overline{P}$:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_2,x_3\}$
$\underline{P}(A)$	0.1	0.1	0.3	0.3	0.6	0.7
$\overline{P}(A)$	0.3	0.4	0.7	0.7	0.9	0.9

$\underline{P}$ is coherent and, since $\mathcal{X}$ is a 3-element possibility space, it is 2-monotone as well. The extreme points of $\mathcal{M}(\underline{P})$ correspond to the following mass functions:

$$(0.1, 0.2, 0.7), \quad (0.1, 0.4, 0.5), \quad (0.2, 0.1, 0.7), \quad (0.3, 0.1, 0.6), \quad (0.3, 0.4, 0.3).$$

On the one hand, taking $\delta = 0.1$, any $P \in B^{0.1}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$ should satisfy:

$$P(\{x_1\}) \geq \overline{P}(\{x_1\}) - \delta = 0.2, \qquad P(\{x_2\}) \geq \overline{P}(\{x_2\}) - \delta = 0.3,$$
$$P(\{x_3\}) \geq \overline{P}(\{x_3\}) - \delta = 0.6,$$

but there is not probability P satisfying these three inequalities, so $B^{0.1}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) = \emptyset$.

On the other hand, taking $\delta = 0.3$, any $P \in B^{0.3}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$ should satisfy:

$$\begin{aligned}
0.6 = \overline{P}(\{x_1\}) + \delta \geq P(\{x_1\}) &\geq \overline{P}(\{x_1\}) - \delta = 0, \\
0.7 = \overline{P}(\{x_2\}) + \delta \geq P(\{x_2\}) &\geq \overline{P}(\{x_2\}) - \delta = 0.1, \\
1 = \overline{P}(\{x_3\}) + \delta \geq P(\{x_3\}) &\geq \overline{P}(\{x_3\}) - \delta = 0.4, \\
1 = \overline{P}(\{x_1,x_2\}) + \delta \geq P(\{x_1\}) + P(\{x_2\}) &\geq \overline{P}(\{x_1,x_2\}) - \delta = 0.4, \\
1 = \min\left\{\overline{P}(\{x_1,x_3\}) + \delta, 1\right\} \geq P(\{x_1\}) + P(\{x_3\}) &\geq \overline{P}(\{x_1,x_3\}) - \delta = 0.6, \\
1 = \min\left\{\overline{P}(\{x_2,x_3\}) + \delta, 1\right\} \geq P(\{x_2\}) + P(\{x_3\}) &\geq \overline{P}(\{x_2,x_3\}) - \delta = 0.6.
\end{aligned}$$

Hence, $B^{0.3}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) = \mathcal{M}(\underline{Q})$, where $\underline{Q}$ is the coherent lower probability given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_2,x_3\}$
$\underline{Q}(A)$	0	0.1	0.4	0.4	0.6	0.6

However, $\mathcal{M}(\underline{P}) \not\subseteq B^{0.3}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$, since for instance $(0.1, 0.2, 0.7)$ dominates $\underline{P}$ but not $\underline{Q}$. ♦

This example shows some unwanted behaviour of the neighbourhoods determined by the Maximum Discrepancy, in that they may be empty or may not be a neighbourhood of the original credal set if the distortion parameter is small enough. Next, we identify which is the minimum distortion parameter that should be considered to avoid these problems.

Lemma 1. *Let $\underline{P}$ be a coherent lower probability and consider $\delta^* = d^{\max}_{\mathrm{TV}}(\underline{P}, \underline{P})$. Then $B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) \supseteq \mathcal{M}(\underline{P})$ if and only if $\delta \geq \delta^*$.*

Proof. On the one hand, if $\delta \geq \delta^*$ it holds that $d_{\mathrm{TV}}(P,Q) \leq \delta^* \leq \delta$ for any $P, Q \in \mathcal{M}(\underline{P})$, meaning that $B^{\delta}_{d_{\mathrm{TV}}}(P) \supseteq \mathcal{M}(\underline{P})$. Also, from Eq. (15), it holds that:

$$B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) = \bigcap_{P \in \mathcal{M}(\underline{P})} B^{\delta}_{d_{\mathrm{TV}}}(P) \supseteq \mathcal{M}(\underline{P}).$$

Conversely, assume that $B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) \supseteq \mathcal{M}(\underline{P})$. Ex-absurdo, if $\delta < \delta^*$, from Corollary 1 there are $P, Q \in \mathrm{ext}(\mathcal{M}(\underline{P}))$ such that $d_{\mathrm{TV}}(P,Q) = \delta^* > \delta$, implying from Eq. (15) that either $P \notin B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$ or $Q \notin B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$, a contradiction with the inclusion $B^{\delta}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) \supseteq \mathcal{M}(\underline{P})$. □

Perhaps surprisingly, if we take δ^* as distortion parameter, the ball $B^{\delta^*}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$ may not coincide with $\mathcal{M}(\underline{P})$, even if it will include it because of the previous lemma.

Example 4. Let us continue with Example 3. Using Proposition 3 we get:

$$\delta^* = d^{\max}_{\mathrm{TV}}(\underline{P}, \underline{P}) = \max_{A \subseteq \mathcal{X}} |\underline{P}(A) - \overline{P}(A)| = 0.4.$$

Taking this distortion parameter, any $P \in B^{\delta^*}_{d^{\max}_{\mathrm{TV}}}(\underline{P})$ should satisfy:

$$\begin{aligned}
0.7 = \overline{P}(\{x_1\}) + \delta \geq P(\{x_1\}) \geq \max\left\{\overline{P}(\{x_1\}) - \delta, 0\right\} = 0,\\
0.8 = \overline{P}(\{x_2\}) + \delta \geq P(\{x_2\}) \geq \max\left\{\overline{P}(\{x_2\}) - \delta, 0\right\} = 0,\\
1 = \min\{\overline{P}(\{x_3\}) + \delta, 1\} \geq P(\{x_3\}) \geq \overline{P}(\{x_3\}) - \delta = 0.3,\\
1 = \min\left\{\overline{P}(\{x_1,x_2\}) + \delta, 1\right\} \geq P(\{x_1\}) + P(\{x_2\}) \geq \overline{P}(\{x_1,x_2\}) - \delta = 0.3,\\
1 = \min\left\{\overline{P}(\{x_1,x_3\}) + \delta, 1\right\} \geq P(\{x_1\}) + P(\{x_3\}) \geq \overline{P}(\{x_1,x_3\}) - \delta = 0.5,\\
1 = \min\left\{\overline{P}(\{x_2,x_3\}) + \delta, 1\right\} \geq P(\{x_2\}) + P(\{x_3\}) \geq \overline{P}(\{x_2,x_3\}) - \delta = 0.5.
\end{aligned}$$

This means that $B^{\delta^*}_{d^{\max}_{\mathrm{TV}}}(\underline{P}) = \mathcal{M}(\underline{Q})$, where:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_2,x_3\}$
$\underline{Q}(A)$	0	0	0.3	0.3	0.5	0.5

Since $(0, 0.5, 0.5)$ belongs to $\mathcal{M}(\underline{Q})$ but not to $\mathcal{M}(\underline{P})$, we deduce that $B^{\delta^*}_{d^{\max}_{TV}}(\underline{P})$ and $\mathcal{M}(\underline{P})$ do not coincide. ♦

The next result gives an intuitive expression of $B^{\delta^*}_{d^{\max}_{TV}}(\underline{P})$ in terms of δ^* and $\overline{P}$.

Proposition 4. *Let $\underline{P}$ be a coherent lower probability and consider $\delta^* = d^{\max}_{TV}(\underline{P}, \underline{P})$. It holds that $B^{\delta^*}_{d^{\max}_{TV}}(\underline{P}) = \mathcal{M}(\underline{P}^*)$, where $\underline{P}^*(A) = \max\left\{\overline{P}(A) - \delta^*, 0\right\}$ for any $A \neq \mathcal{X}$ and $\underline{P}^*(\mathcal{X}) = 1$.*

Proof. On the one hand, it is clear that $B^{\delta^*}_{d^{\max}_{TV}}(\underline{P}) \subseteq \mathcal{M}(\underline{P}^*)$. Conversely, take $P \in \mathcal{M}(\underline{P}^*)$, and let us see that $d^{\max}_{TV}(P, \underline{P}) \leq \delta^*$. We consider two cases:

Case 1: If $P(A) < \overline{P}(A)$, we get:

$$|P(A) - \overline{P}(A)| = \overline{P}(A) - P(A) = \left(\overline{P}(A) - \delta^*\right) + \delta^* - P(A) \leq \delta^*.$$

Case 2: If $P(A) \geq \overline{P}(A)$, we get:

$$\begin{aligned}|P(A) - \overline{P}(A)| &= P(A) - \overline{P}(A) = \underline{P}(A^c) - P(A^c) \\ &\leq \underline{P}(A^c) - \max\left\{0, \overline{P}(A^c) - \delta^*\right\} = \underline{P}(A^c) + \min\left\{0, \delta^* - \overline{P}(A^c)\right\} \\ &= \min\{\underline{P}(A^c), \delta^* + \underline{P}(A^c) - \overline{P}(A^c)\} \leq \delta^*\end{aligned}$$

given that $\underline{P} \leq \overline{P}$.

We therefore conclude that $d^{\max}_{TV}(P, \underline{P}) = \max_{A \subseteq \mathcal{X}} |P(A) - \overline{P}(A)| \leq \delta^*$. □

Notwithstanding, the lower probability $\underline{P}^*$ considered in the previous result may not be coherent.

Example 5. Consider a possibility space $\mathcal{X}$ with cardinality five and the coherent lower probability $\underline{P}$ given by:

$\lvert A\rvert$	1	2	3	4	5
$\underline{P}(A)$	0.19	0.38	0.58	0.78	1

$\underline{P}$ is coherent because it is the lower envelope of the mass function $(0.19, 0.19, 0.2, 0.2, 0.22)$ and its permutations. We get $\delta^* = d^{\max}_{TV}(\underline{P}, \underline{P}) = 0.04$, and when computing $\underline{P}^*$ we obtain the following values:

$\lvert A\rvert$	1	2	3	4	5
$\underline{P}^*(A)$	0.18	0.38	0.58	0.77	1

However, $\underline{P}^*$ is not coherent. For example, taking $A = \{x_1, x_2, x_3, x_4\}$, it is not possible to find $P \in \mathcal{M}(\underline{P}^*)$ such that $P(A) = \underline{P}^*(A) = 0.77$. If it existed, then $P(\{x_5\}) = 0.23$, and since $P(\{x_i, x_5\}) \in [0.38, 0.42]$ for $i = 1, \ldots, 4$, we should have $P(\{x_i\}) \in [0.18, 0.19]$ for $i = 1, \ldots, 4$. However, this implies that:

$$1 = \sum_{i=1}^{5} P(\{x_i\}) \leq 4 \cdot 0.19 + 0.23 = 0.76 + 0.23 = 0.99,$$

a contraction. ♦

Finally, if the distortion parameter is greater than δ^*, we may also find a simple expression for the distortion model.

Proposition 5. *Let $\underline{P}$ be a coherent lower probability, $\delta^* = d_{\mathrm{TV}}^{\max}(\underline{P}, \underline{P})$ and $\delta > 0$. It holds that $B_{d_{\mathrm{TV}}^{\max}}^{\delta^*+\delta}(\underline{P}) = \mathcal{M}(\underline{Q})$, where $\underline{Q}(A) = \max\{\underline{P}^*(A) - \delta, 0\}$ for any $A \subset \mathcal{X}$, $\underline{Q}(\mathcal{X}) = 1$ and $\underline{P}^*$ is given by Proposition 4.*

Proof. On the one hand, assume that $P \in B_{d_{\mathrm{TV}}^{\max}}^{\delta^*+\delta}(\underline{P})$. Thus, for any $A \subset \mathcal{X}$ it holds that $|P(A) - \overline{P}(A)| \leq \delta^* + \delta$, meaning that

$$-\delta^* - \delta \leq P(A) - \overline{P}(A) \leq \delta^* + \delta.$$

Hence, $P(A) \geq \overline{P}(A) - \delta^* - \delta$, or equivalently, $P(A) \geq \max\{\underline{P}^*(A) - \delta, 0\} = \underline{Q}(A)$.

Conversely, take $P \in \mathcal{M}(\underline{Q})$, and let us see that $d_{\mathrm{TV}}^{\max}(P, \underline{P}) \leq \delta + \delta^*$. We consider two cases:

Case 1: If $P(A) < \overline{P}(A)$ we get:

$$|P(A) - \overline{P}(A)| = \overline{P}(A) - P(A) \leq \overline{P}(A) - \big(\overline{P}(A) - \delta^* - \delta\big) = \delta^* + \delta.$$

Case 2: If $P(A) \geq \overline{P}(A)$, assume ex-absurdo that $|P(A) - \overline{P}(A)| > \delta^* + \delta$. We get:

$$|P(A) - \overline{P}(A)| = P(A) - \overline{P}(A) > \delta^* + \delta,$$

which is equivalent to $\underline{P}(A^c) - P(A^c) > \delta^* + \delta$, and then:

$$P(A^c) < \underline{P}(A^c) - \delta^* - \delta \leq \overline{P}(A^c) - \delta^* - \delta,$$

a contradiction.

We conclude that $d_{\mathrm{TV}}^{\max}(P, \underline{P}) = \max_{A \subseteq \mathcal{X}} |P(A) - \overline{P}(A)| \leq \delta^* + \delta$. □

Again, the lower probability $\underline{Q}$ may not be coherent.

Example 6. Let us continue with Example 5. For $\delta \in (0, 0.01)$, we get:

$\lvert A\rvert$	1	2	3	4	5
$\underline{Q}(A)$	$0.18 - \delta$	$0.38 - \delta$	$0.58 - \delta$	$0.77 - \delta$	1

Following the same steps as in Example 5, $\underline{Q}$ is not coherent because there is no $P \in \mathcal{M}(\underline{Q})$ such that $P(\{x_1, x_2, x_3, x_4\}) = \underline{Q}(\{x_1, x_2, x_3, x_4\})$. ♦

6. Minimum Supremum

The third distortion procedure we analyse in this paper is based on the minimum of the supremum distances between the elements of the credal set.

We begin our analysis by remarking that some properties of $d_{\infty}^{\min}$ are analogous to those of $d_{\mathrm{TV}}^{\min}$. This is not surprising, since both of them consist in taking the minimum over the credal sets of a convex and continuous distance between probability measures. We observe then that the Minimum Supremum between two lower probabilities is zero if and only if their credal sets have non-empty intersection, and therefore $d_\infty^{\min}$ is not a distance between lower probabilities. The same reasoning used in Section 4 to show that the Minimum TV violates the triangle inequality is valid here, while the rest of the axioms of a distance are satisfied.

Next, as we did for the Minimum TV and Maximum Discrepancy, we look for more operative expressions for $d_\infty^{\min}$.

6.1. *Alternative expressions*

We start by showing in the following example that the Minimum Supremum is not attained at the extreme points of the credal set.

Example 7. Considering again the setting of Example 1, we obtain that $d_\infty^{\min}(P_0, Q) = 0.3$, $d_\infty^{\min}(P_1, Q) = 0.4$ and $d_\infty(P_{0.5}, Q) = 0.1$, meaning that the Minimum Supremum is not reached at the extreme points of the credal sets. ♦

This same example shows that $d_\infty^{\min}(\underline{P}, \underline{Q})$ cannot be expressed as

$$\max_{x\in\mathcal{X}} \min\left\{|\underline{P}(\{x\}) - \overline{Q}(\{x\})|, |\overline{P}(\{x\}) - \underline{Q}(\{x\})|\right\},$$

which for this example gives the value 0.3, different from $d_\infty^{\min}(\underline{P}, \underline{Q}) = 0.1$.

Despite the previous counterexample, we next prove that, under some conditions, the Minimum Supremum is attained at the boundary of the credal sets:

Proposition 6. *Let $\underline{P}$ and $\underline{Q}$ be two lower probabilities avoiding sure loss such that $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) = \emptyset$. If $P \in \mathcal{M}(\underline{P})$ and $Q \in \mathcal{M}(\underline{Q})$ satisfy $d_\infty^{\min}(\underline{P}, \underline{Q}) = d_\infty(P, Q)$, then $P \in \partial\mathcal{M}(\underline{P})$ and $Q \in \partial\mathcal{M}(\underline{Q})$. As a consequence,*

$$d_\infty^{\min}(\underline{P}, \underline{Q}) = \min_{\substack{P\in\partial\mathcal{M}(\underline{P})\\ Q\in\partial\mathcal{M}(\underline{Q})}} d_\infty(P, Q).$$

Proof. The proof is quite similar to that of Proposition 1. Ex-absurdo, assume that $d_\infty^{\min}(\underline{P}, \underline{Q}) = d_\infty(P, Q)$ where at least one of $P \in \mathcal{M}(\underline{P}) \setminus \partial\mathcal{M}(\underline{P})$ or $Q \in \mathcal{M}(\underline{Q}) \setminus \partial\mathcal{M}(\underline{Q})$ holds. Assume that $P \in \mathcal{M}(\underline{P}) \setminus \partial\mathcal{M}(\underline{P})$; the other case follows analogously. Then P belongs to the interior of $\mathcal{M}(\underline{P})$, and as a consequence it must be $\overline{P}(A) > P(A) > \underline{P}(A)$ for any $A \neq \emptyset, \mathcal{X}$, given that there must be some neighbourhood of P included within $\mathcal{M}(\underline{P})$.

Let us consider the following partition of $\mathcal{X}$:

$$\mathcal{X}_1 = \{x \in \mathcal{X} \mid P(\{x\}) > Q(\{x\})\}, \qquad \mathcal{X}_2 = \{x \in \mathcal{X} \mid P(\{x\}) < Q(\{x\})\},$$
$$\mathcal{X}_3 = \{x \in \mathcal{X} \mid P(\{x\}) = Q(\{x\})\}.$$

Since $\mathcal{M}(\underline{P}) \cap \mathcal{M}(\underline{Q}) = \emptyset$, it follows that $P \neq Q$, implying that both $\mathcal{X}_1$ and $\mathcal{X}_2$ are non-empty. Thus, applying Eq. (3) we obtain:

$$\begin{aligned} d_\infty^{\min}(P,Q) &= \max_{x\in\mathcal{X}} |P(\{x\}) - Q(\{x\})| \\ &= \max_{x\in\mathcal{X}_1\cup\mathcal{X}_2} |P(\{x\}) - Q(\{x\})| = |P(\{x^*\}) - Q(\{x^*\})| \end{aligned}$$

for some $x^* \in \mathcal{X}_1 \cup \mathcal{X}_2$. Assume that $x^* \in \mathcal{X}_1$; the other case follows similarly. Take $\varepsilon > 0$ satisfying:

$$\varepsilon < \min_{A\neq\emptyset,\mathcal{X}} \big(P(A) - \underline{P}(A)\big), \tag{16}$$

fix $x_2 \in \mathcal{X}_2$ and define:

$$P^*(\{x\}) = \begin{cases} P(\{x^*\}) + \varepsilon, & \text{if } x = x^*, \\ P(\{x_2\}) - \varepsilon, & \text{if } x = x_2, \\ P(\{x\}), & \text{if } x \neq x^*, x_2. \end{cases}$$

By construction $P^*(\{x_2\}) \geq \underline{P}(\{x_2\}) \geq 0$ and $P^*(\{x^*\}) \leq \overline{P}(\{x^*\}) \leq 1$. From this we deduce that P^* is a probability measure that, from Eq. (16), satisfies $P^*(A) \geq P(A) - \varepsilon > \underline{P}(A)$ for any $A \neq \emptyset, \mathcal{X}$, meaning that $P^* \in \mathcal{M}\big(\underline{P}\big)$. Moreover, using Eq. (3) we obtain:

$$\begin{aligned} d_\infty(P^*,Q) &= \max_{x\in\mathcal{X}} |P^*(\{x\}) - Q(\{x\})| \geq P^*(\{x^*\}) - Q(\{x^*\}) \\ &= P(\{x^*\}) - Q(\{x^*\}) + \varepsilon = d_\infty(P,Q) + \varepsilon > d_\infty(P,Q) = d_\infty^{\min}(\underline{P},\underline{Q}). \end{aligned}$$

a contradiction. □

6.2. *Distorting lower probabilities with the Minimum Supremum*

Next we analyse the distortion of a lower probability using the Minimum Supremum. If we consider $d_\infty^{\min}$, a distortion factor $\delta \geq 0$ and a lower probability $\underline{P}$, we can determine the following neighbourhood:

$$B^\delta_{d_\infty^{\min}}(\underline{P}) = \big\{Q \in \mathbb{P}(\mathcal{X}) \mid d_\infty^{\min}(Q,\underline{P}) \leq \delta\big\}.$$

Let us denote by $\underline{Q}_{d_\infty^{\min}}$ the coherent lower probability it induces by taking lower envelopes.

Given a probability measure P, we consider also the neighbourhood $B^\delta_{d_\infty^{\min}}(P)$ and denote by $\underline{Q}^P_{d_\infty^{\min}}$ the coherent lower probability it induces. By definition:

$$B^\delta_{d_\infty^{\min}}(\underline{P}) = \bigcup_{P\in\mathcal{M}(\underline{P})} B^\delta_{d_\infty^{\min}}(P) \ \Rightarrow\ \underline{Q}_{d_\infty^{\min}} = \min_{P\in\mathcal{M}(\underline{P})} \underline{Q}^P_{d_\infty^{\min}}.$$

Taking into account the comments above, we look for an explicit formula for $\underline{Q}_{d_\infty^{\min}}$.

Proposition 7. *Let P and Q be two probability measures and $\delta > 0$. It holds that:*

$$d_\infty(Q,P) \leq \delta \Leftrightarrow Q(A) \geq P(A) - \delta \min(|A|, |A^c|).$$

Proof. For the direct implication, it suffices to observe that, by Eq. (4),

$$d_\infty(Q,P) \leq \delta \Rightarrow |Q(A)-P(A)| \leq \delta|A| \text{ and } |Q(A^c)-P(A^c)| \leq \delta|A^c| \quad \forall A \neq \emptyset,$$

whence

$$Q(A) \geq \max\left\{P(A)-\delta|A|, P(A)-\delta|A^c|\right\} = P(A) - \delta\min(|A|,|A^c|) \quad \forall A \neq \emptyset.$$

Conversely, take any event $A \subseteq \mathcal{X}$. If $Q(A) \geq P(A)$, then by assumption $(Q(A) - P(A)) \geq -\delta|A|$ and also $(Q(A^c) - P(A^c)) \geq -\delta|A|$, and this second inequality is equivalent to $(Q(A) - P(A)) \leq \delta|A|$. Thus, $|Q(A) - P(A)| \leq \delta|A|$. The reasoning when $|Q(A) - P(A)| = P(A) - Q(A)$ is similar. □

As a consequence, the lower probability $\underline{Q}^P_{d_\infty^{\min}}$ corresponds to the natural extension of the functional given by:

$$\underline{Q}^P(A) = \max\{P(A) - \delta\min(|A|,|A^c|), 0\} \quad \forall A \neq \emptyset, \mathcal{X}, \tag{17}$$

$\underline{Q}^P(\emptyset) = 0$ and $\underline{Q}^P(\mathcal{X}) = 1$. This immediately allows to deduce the following result:

Corollary 2. *Let $\underline{P}$ be a coherent lower probability and $\delta > 0$. For any $P \in \mathcal{M}(\underline{P})$, it holds that $B^\delta_{d_\infty^{\min}}(P) = \mathcal{M}(\underline{Q}^P)$, where $\underline{Q}^P$ is given by Eq. (17). Consequently, $B^\delta_{d_\infty^{\min}}(\underline{P}) = \cup_{P\in\mathcal{M}(\underline{P})}\mathcal{M}(\underline{Q}^P)$, and if $\underline{Q}^P$ is coherent for every P then $\underline{Q}_{d_\infty^{\min}}$ coincides with the functional $\underline{Q}^{\underline{P}}$ given by*

$$\underline{Q}^{\underline{P}}(A) = \max\{\underline{P}(A) - \delta\min(|A|,|A^c|), 0\} \quad \forall A \neq \emptyset, \mathcal{X}, \tag{18}$$

$\underline{Q}^{\underline{P}}(\emptyset) = 0$ and $\underline{Q}^{\underline{P}}(\mathcal{X}) = 1$, that is then a coherent lower probability.

Proof. We shall only establish the equality $\underline{Q}_{d_\infty^{\min}} = \underline{Q}^{\underline{P}}$ when $\underline{Q}^P$ is coherent for every $P \geq \underline{P}$, the rest of the proof being immediate. For this, note that the coherence of $\underline{P}$ implies that for every $A \subseteq \mathcal{X}$ there is some $P \geq \underline{P}$ such that $P(A) = \underline{P}(A)$, and then the coherence of $\underline{Q}^P$ implies the existence of $Q \in \mathcal{M}(\underline{Q}^P) = B^\delta_{d_\infty^{\min}}(P)$ such that $Q(A) = \underline{Q}^P(A) = \underline{Q}^{\underline{P}}(A)$. Since $B^\delta_{d_\infty^{\min}}(P) \subseteq B^\delta_{d_\infty^{\min}}(\underline{P})$, this implies that $\underline{Q}_{d_\infty^{\min}} \leq \underline{Q}^{\underline{P}}$; and taking into account that the converse inequality is immediate, we deduce the equality. □

Moreover, under some relatively mild conditions we can guarantee the 2-monotonicity:

Proposition 8. *Let $\underline{P}$ be a 2-monotone lower probability in $\underline{\mathbb{P}}^*(\mathcal{X})$ and let $0 < \delta < \min_{A\neq\mathcal{X},\emptyset} \frac{\underline{P}(A)}{\min(|A|,|A^c|)}$. Then the functional given by Eq. (18) is a 2-monotone lower probability if and only if it is monotone.*

Proof. Assume that $\underline{Q}^{\underline{P}}$ is monotone, and let us prove that it is 2-monotone. Given the constraint in δ, we must show that for any two events A, B it holds that

$$\underline{Q}^{\underline{P}}(A\cup B) + \underline{Q}^{\underline{P}}(A\cap B) \geq \underline{Q}^{\underline{P}}(A) + \underline{Q}^{\underline{P}}(B).$$

Since this equation is trivial when either A or B is equal to $\mathcal{X}$, we can assume w.l.o.g. that $A, B \neq \mathcal{X}$. We shall also use that $\underline{Q}^{\underline{P}}(A \cup B) \geq \underline{P}(A \cup B) - \delta \min(|A \cup B|, |(A \cup B)^c|)$, with equality whenever $A \cup B \neq \mathcal{X}$. Using the assumption on δ we shall then prove that

$$\underline{P}(A \cup B) - \delta \min(|A \cup B|, |(A \cup B)^c|) + \underline{P}(A \cap B) - \delta \min(|A \cap B|, |(A \cap B)^c|) \\ \geq \underline{P}(A) - \delta \min(|A|, |A^c|) + \underline{P}(B) - \delta \min(|B|, |B^c|).$$

Since $\underline{P}$ is assumed to be 2-monotone, it satisfies

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

Let us prove that also

$$-\delta \min(|A \cup B|, |(A \cup B)^c|) - \delta \min(|A \cap B|, |(A \cap B)^c|) \\ \geq -\delta \min(|A|, |A^c|) - \delta \min(|B|, |B^c|),$$

or, equivalently, that

$$\min(|A \cup B|, |(A \cup B)^c|) + \min(|A \cap B|, |(A \cap B)^c|) \leq \min(|A|, |A^c|) + \min(|B|, |B^c|). \tag{19}$$

We consider a number of cases:

- If $\min(|A \cup B|, |(A \cup B)^c|) = |A \cup B|$, then $\min(|A \cap B|, |(A \cap B)^c|) = |A \cap B|$, $\min(|A|, |A^c|) = |A|$ and $\min(|B|, |B^c|) = |B|$, whence (19) holds with equality.
- Similarly, if $\min(|A \cap B|, |(A \cap B)^c|) = |(A \cap B)^c|$, then $\min(|A \cup B|, |(A \cup B)^c|)$, $\min(|A|, |A^c|) = |A^c|$ and $\min(|B|, |B^c|) = |B^c|$, and (19) holds with equality.
- Assume now that $|A \cap B| < |(A \cap B)^c|$ and $|A \cup B| > |(A \cup B)^c|$. Then if for instance $|A| \leq |A^c|, |B| \geq |B^c|$, we obtain

  $$|A^c \cap B^c| + |A \cap B| \leq |B^c| + |A|;$$

  if $|A| \geq |A^c|, |B| \leq |B^c|$, we obtain

  $$|A^c \cap B^c| + |A \cap B| \leq |A^c| + |B|;$$

  if $|A| \geq |A^c|, |B| \geq |B^c|$, we obtain

  $$|A^c \cap B^c| + |A \cap B| \leq |A^c \cap B^c| + |A^c \cup B^c| = |A^c| + |B^c|;$$

  and if $|A| \leq |A^c|, |B| \leq |B^c|$, we obtain

  $$|A^c \cap B^c| + |A \cap B| \leq |A \cup B| + |A \cap B| = |A| + |B|;$$

  in any of these cases we conclude that (19) holds. □

We see from the proof above that the assumption of $\underline{P} \in \underline{\mathbb{P}}^*(\mathcal{X})$ is critical to be able to choose a δ small enough so as to express the functional in Eq. (18) in terms of $\underline{P}$.

7. Penalised total variation and coalitional games

In this section, we shall make a connection between the distortion of imprecise probabilities and coalitional games. More specifically, we shall first of all introduce and analyse a slight modification of the Minimum Supremum distance, which we shall call *Penalised Total Variation*, and then show its relation with the weak delta-core from coalitional games.

7.1. *The penalised total variation*

We recall that, in the particular case where we want to compare a probability measure and a lower probability, the Minimum Supremum becomes:

$$d_\infty^{\min}(Q,\underline{P}) = \min_{P\in\mathcal{M}(\underline{P})}\max_{x\in\mathcal{X}}|P(\{x\})-Q(\{x\})| = \min_{P\in\mathcal{M}(\underline{P})}\max_{A\subseteq\mathcal{X}}\frac{|P(A)-Q(A)|}{|A|},$$

where the second equality follows from Eq. (4). This leads us to consider the function:

$$d_{\text{PTV}}^{\min}(Q,\underline{P}) := \max_{A\subseteq\mathcal{X}}\frac{\underline{P}(A)-Q(A)}{|A|},$$

that will be called the Penalised TV (PTV, for short), and that is similar to the function d'_{TV} given in Eq. (9), taking also into account the cardinality of the events. Note that $d_{\text{PTV}}^{\min}$ does not coincide with $d_\infty^{\min}$: to see it, consider again Examples 1 and 7. It holds that

$$d_{\text{PTV}}^{\min}(Q,\underline{P}) = \frac{\underline{P}(\{x_2,x_3\})-Q(\{x_2,x_3\})}{2} = 0.05 \neq 0.1 = d_\infty(Q,\underline{P}).$$

The neighbourhood model induced by $d_{\text{PTV}}^{\min}$ is given by:

$$B^\delta_{d_{\text{PTV}}^{\min}}(\underline{P}) = \{Q\in\mathbb{P}(\mathcal{X}) \mid d_{\text{PTV}}^{\min}(Q,\underline{P})\leq\delta\}.$$

The next result gives a simple expression for the coherent lower probability $\underline{Q}_{d_{\text{PTV}}^{\min}}$ it determines.

Proposition 9. *Let $\underline{P}$ be a coherent lower probability and $\delta\geq 0$. It holds that $\underline{Q}_{d_{\text{PTV}}^{\min}}$ is the natural extension of the functional $\underline{Q}$ given by:*

$$\underline{Q}(A) = \max\left\{\underline{P}(A)-\delta|A|,0\right\} \quad \forall A\subset\mathcal{X}, \tag{20}$$

and $\underline{Q}(\mathcal{X})=1$. Also, $B^\delta_{d_{\text{PTV}}^{\min}}(\underline{P}) = \mathcal{M}(\underline{Q})$.

Proof. The equality $B^\delta_{d_{\text{PTV}}^{\min}}(\underline{P}) = \mathcal{M}(\underline{Q})$ follows from the fact that $Q\in B^\delta_{d_{\text{PTV}}^{\min}}(\underline{P})$ if and only if $\max_{A\subseteq\mathcal{X}}\frac{\underline{P}(A)-Q(A)}{|A|}\leq\delta$, which is equivalent to $Q(A)\geq\underline{P}(A)-\delta|A|$ for any $A\subseteq\mathcal{X}$. □

Regarding the properties of this neighbourhood, by construction $\mathcal{M}(\underline{P}) \subseteq B^\delta_{d_{\text{PTV}}^{\min}}(\underline{P})$, meaning that $\underline{Q}_{d_{\text{PTV}}^{\min}}$ avoids sure loss when $\underline{P}$ does. However, $\underline{Q}$ need not be a coherent lower probability even if $\underline{P}$ is:

Example 8. Let $\mathcal{X}$ be a four element possibility space, $\underline{P}$ be the lower envelope of the probability mass functions $(0.5, 0.5, 0, 0)$ and $(0.25, 0.25, 0.25, 0.25)$, and consider $\delta = 0.2$. It holds that:

$$\underline{Q}(\{x_2, x_3, x_4\}) = \max\left\{\underline{P}(\{x_2, x_3, x_4\}) - 3\delta, 0\right\} = \max\{0.5 - 3 \cdot 0.2, 0\} = 0.$$
$$\underline{Q}(\{x_2, x_3\}) = \max\left\{\underline{P}(\{x_2, x_3\}) - 2\delta, 0\right\} = \max\{0.5 - 2 \cdot 0.2, 0\} = 0.1.$$

This means that $\underline{Q}$ is not monotone, hence it cannot be coherent (it is not even a lower probability). ♦

With a similar reasoning as the one in Proposition 8, we can show that the distortion by means of the Penalised Total Variation preserves 2-monotonicity:

Proposition 10. *Let* $\underline{P}$ *be a 2-monotone lower probability in* $\underline{\mathbb{P}}^*(\mathcal{X})$ *and let* $0 < \delta < \min_{A \neq \mathcal{X}, \emptyset} \frac{\underline{P}(A)}{|A|}$. *Then the functional given by Eq.* (20) *is a* 2*-monotone lower probability if and only if it is monotone.*

On the other hand, a sufficient condition for the monotonicity of the functionals $\underline{Q}^{\underline{P}}, \underline{Q}$ given in Eqs. (18),(20) is that for any event $A \neq \emptyset, \mathcal{X}$ and any $x \in A$

$$\delta \leq \underline{P}(A) - \underline{P}(A \setminus \{x\}).$$

In other words, provided $\underline{P}$ is strictly monotone and δ is sufficiently small, we have an expression for the lower probabilities associated with the Minimum Supremum and with $d_{\text{PTV}}^{\min}$, and the distortion procedures preserve 2-monotonicity. In particular, if

$$\delta \leq \min_{P \in \mathcal{M}(\underline{P})} \min_{A \neq \emptyset, \mathcal{X}, x \in A} \big(P(A) - P(A \setminus \{x\})\big),$$

we would deduce that $\underline{Q}^P$ is 2-monotone, and therefore coherent, for every $P \geq \underline{P}$, and then Corollary 2 implies that $\underline{Q}^{\underline{P}}$ coincides with $\underline{Q}_{d_\infty^{\min}}$.

7.2. *Distortion of coalitional games*

A (monotone and normalised) coalitional game[?] is formally equivalent to a lower probability $\underline{P} : \mathcal{P}(\mathcal{X}) \to [0, 1]$. In this context, $\mathcal{X}$ is interpreted as a set of players that may form coalitions, represented by subsets $A \subseteq \mathcal{X}$, in order to obtain a higher proportion of the resource for which they compete. Thus, $\underline{P}(A)$ represents the minimum proportion of the reward guaranteed by the coalition A, and $\mathcal{M}(\underline{P})$ contains all the additive distributions of the total reward compatible with the requirements imposed by each coalition. These distributions are called *solutions* of the game. All the notions related to lower probabilities (avoiding sure loss, coherence, 2-monotonicity, credal set, . . .) have a counterpart in coalitional game theory (balanced, exact or convex games and core, respectively) [**?**, Table 1]: even though the notation and interpretation may differ, the mathematical formulation is equivalent. We refer to[?] for a thorough analysis of coalitional games, and to[?,?] for some recent advances in the field.

A problem that may arise in coalitional game theory is the existence of an empty core. This means that there is no way to distribute the total reward while simultaneously satisfying the requirements imposed by all coalitions. In the literature[?], two common approaches are used to address this problem: the *strong δ-core* and the *weak δ-core*, given by:

$$\text{core}^{\text{S}}_{\delta}(\underline{P}) = \left\{Q \in \mathbb{P}(\mathcal{X}) \mid Q(A) \geq \underline{P}(A) - \delta \quad \forall A \subseteq \mathcal{X}\right\}.$$
$$\text{core}^{\text{W}}_{\delta}(\underline{P}) = \left\{Q \in \mathbb{P}(\mathcal{X}) \mid Q(A) \geq \underline{P}(A) - \delta|A| \quad \forall A \subseteq \mathcal{X}\right\}.$$

While the strong δ-core uniformly reduces the requirements of all the (non-trivial) coalitions, the weak δ-core reduces the requirements proportionally to the size of the coalition.

It is possible to create a connection between the neighbourhoods created in the previous sections and the strong and weak cores. On the one hand, it was proven in [**?**, Prop.11] that $B^{\delta}_{d'_{\text{TV}}}(\underline{P}) = \text{core}^{S}_{\delta}(\underline{P})$. On the other hand, the following result relates the weak core with one of the neighbourhoods previously introduced.

Corollary 3. *Let $\underline{P}$ be a lower probability and $\delta > 0$. It holds that:*

$$\textit{core}^{W}_{\delta}(\underline{P}) = B^{\delta}_{d^{\min}_{\text{PTV}}}(\underline{P}).$$

We therefore conclude that the weak and strong cores are related to the distortion of a lower probability using the functions d'_{TV} and $d^{\min}_{\text{PTV}}$, respectively.

8. Conclusions

In this work, we have explored two different extensions of the TV-distance for comparing non-additive measures or lower probabilities, the Minimum TV and the Maximum Discrepancy, and one extension of the supremum distance, the Minimum Supremum.

For each of these proposals we have investigated if they are distances between lower probabilities and alternative expressions in terms of the extreme points, elements in the frontier of the credal sets or in terms of the direct comparison of the lower and upper probabilities involved. Moreover, we have used the three approaches to distort a lower probability creating a neighbourhood around a lower probability. A summary of the results can be seen in Table 1, where a double check (✓✓) means that the property is satisfied, a single check (✓) means that the property is satisfied under some additional condition indicated at the bottom of the table, and a cross (✗) means that the property is not fulfilled.

In addition, we have shown an interesting connection between the distortion of lower probabilities by the extensions of the TV-distance and the strong and weak core in game theory. On the one hand, in[?] we showed that the neighbourhood created by the function d'_{TV} coincides with the strong core. On the other hand, we have seen that a function called PTV induces a neighbourhood that coincides with the weak core. Our results allow us also to deduce that the weak core does not

	Property	$d_{\mathrm{TV}}^{\min}$	$d_{\mathrm{TV}}^{\max}$	$d_{\infty}^{\min}$
Distance	$d_{\cdot}(\underline{P}, \underline{Q}) = 0 \Rightarrow \underline{P} = \underline{Q}$	✗	✓✓	✗
	$d_{\cdot}(\underline{P}, \overline{P}) = 0$	✓✓	✗	✓✓
	Symmetry	✓✓	✓✓	✓✓
	Triangle inequality	✗	✓✓	✗
Expression	Characterised by the extremes	✗	✓✓	✗
	Characterised by the boundary	✓✓	✓✓	✓✓
	Characterised by the lower/upper probs	✓[2]	✓[1]	✗
Distorting $\underline{P}$	$B^{\delta}_{d_{\cdot}}(\underline{P}) \neq \emptyset$	✓✓	✓[1,3]	✓✓
	$B^{\delta}_{d_{\cdot}}(\underline{P}) \supseteq \mathcal{M}(\underline{P})$	✓✓	✓[1,3]	✓✓
	$B^{\delta}_{d_{\cdot}}(\underline{P})$ determined by a lower probability	✓[2]	✓[1,3]	✓[1]

[1]: coherence. [2]: 2-monotonicity. [3]: $\delta \geq \delta^*$.

Table 1. Summary of the properties satisfied by $d_{\mathrm{TV}}^{\min}$, $d_{\mathrm{TV}}^{\max}$ and $d_{\infty}^{\min}$.

induce a monotone model in general, and to have sufficient conditions for it to be a 2-monotone one.

A number of interesting open problems arise from this contribution. First of all, similarly to the Maximum Discrepancy, we may analyse the Maximum Supremum, paying special attention to the neighbourhood it induces. Second, while we have focused on the TV and supremum distances, it would be worthwhile to explore other distances between probability measures, such as the Kolmogorov distance. For each of them, we believe that similar extensions could be applied to the comparison of lower probabilities. In this respect, a somewhat related approach in the context of optimal transport between belief functions and with respect to the Wasserstein distance has been considered in[?] ; it would be worth analysing whether our extensions of the total variation are of interest in that context. Thirdly, we should further explore the connection between the distortion of lower probabilities and coalitional games, and in particular on the preservation of other properties.

Acknowledgements

This contribution is part of grant PID2022-140585NB-I00 funded by MICIU/AEI/10.13039/501100011033 and “FEDER/UE”. We would also like to thank the reviewers for the careful reading of the paper and the valuable suggestions.

References

1. D. A. Levin, Y. Peres and E. Wilmer, *Markov Chains and Mixing Times* (American Mathematical Society, 2009).
2. T. Augustin, F. Coolen, G. de Cooman and M. Troffaes, eds., *Introduction to Imprecise Probabilities*, Wiley Series in Probability and Statistics (Wiley, 2014).
3. P. Walley, *Statistical Reasoning with Imprecise Probabilities* (Chapman and Hall, London, 1991).

4. T. Herron, T. Seidenfeld and L. Wasserman, "Divisive conditioning: further results on dilation", *Philosophy of Science* **64** (1997) 411–444.
5. I. Montes, E. Miranda and S. Destercke, "Unifying neighbourhood and distortion models: Part II- New models and synthesis", *International Journal of General Systems* **49** (2020) 636–674.
6. D. Škulj and R. Hable, "Coefficients of ergodicity for Markov chains with uncertain parameters", *Metrika* **76** (2013) 107–133.
7. Y. Rébillé, "A super Radom-Nikodym derivative for almost derivative subadditive set functions", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **21** (2013) 347–365.
8. E. Miranda and I. Montes, "Centroids of the core of exact capacities: a comparative study", *Annals of Operations Research* **321** (2023) 409–449.
9. I. Levi, *The enterprise of knowledge* (MIT Press, Cambridge, 1980).
10. M. Grabisch, *Set functions, games and capacities in decision making* (Springer, 2016).
11. E. Miranda and I. Montes, "Shapley and Banzhaf values as probability transformations", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **26** (2018) 917–947.
12. L. Shapley and M. Shubik, "Quasi-cores in a monetary economy with nonconvex preferences", *Econometrica* **34** (1966) 805–827.
13. D. Nieto-Barba, I. Montes and E. Miranda, "Distortions of imprecise probabilities", in *Proceedings of IPMU 2004* (Springer, 2024), volume 1175 of *Lecture Notes in Networks and Systems*, pp. 79–90.
14. D. Nieto-Barba, I. Montes and E. Miranda, "The imprecise total variation model and its connections with game theory", *Fuzzy Sets and Systems* **517** (2025) 109448.
15. M. Troffaes and G. de Cooman, *Lower previsions*, Wiley Series in Probability and Statistics (Wiley, 2014).
16. S. Destercke, "Independence and 2-monotonicity: Nice to have, hard to keep", *International Journal of Approximate Reasoning* **54(4)** (2013) 478–490.
17. G. Choquet, "Theory of capacities", *Annales de l'Institut Fourier* **5** (1953–1954) 131–295.
18. I. Montes, E. Miranda and S. Destercke, "Unifying neighbourhood and distortion models: Part I- New results on old models", *International Journal of General Systems* **49** (2020) 602–635.
19. J. Cano, S. Moral and J. Verdegay-López, "Partial inconsistency of probability envelopes", *Fuzzy Sets and Systems* **52** (1992) 201–216.
20. P. Walley, "The elicitation and aggregation of beliefs", Technical report, University of Warwick, Coventry, 1982, statistics Research Report 23.
21. J. Derks and J. Kuipers, "On the number of extreme points of the core of a transferable utility game", in *Chapters in Game Theory in Honour of Stef Tijs* (Kluwer, 2001), pp. 83–97.
22. A. Wallner, "Extreme points of coherent probabilities in finite spaces", *International Journal of Approximate Reasoning* **44** (2007) 339–357.
23. L. Shapley, "Cores of convex games", *International Journal of Game Theory* **1** (1971) 11–26.
24. A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping", *Annals of Mathematical Statistics* **38** (1967) 325–339.
25. I. Montes and S. Destercke, "Extreme points of p-boxes and belief functions", *Annals of Mathematics and Artificial Intelligence* **81** (2017) 405–428.
26. E. Miranda, I. Couso and P. Gil, "Extreme points of credal sets generated by 2-alternating capacities", *International Journal of Approximate Reasoning* **33** (2003)

95–115.
27. S. Lorenizini, D. Petturiti and B. Vantaggi, "Choquet-Wasserstein pesudo-distances via optimal transport under partially specified marginal probabilities", *Fuzzy Sets and Systems* **515** (2025) 109429.