PROCESSING DISTORTION MODELS: A COMPARATIVE STUDY

SÉBASTIEN DESTERCKE, IGNACIO MONTES, AND ENRIQUE MIRANDA

Abstract. When dealing with sets of probabilities, distortion or neighbourhood models are convenient practical tools, as they rely on very little parameters. In this paper, we study their behaviour when such models are combined and processed through some reasoning tools. More specifically, we study their behaviour when merging different distortion models quantifying uncertainty on the same quantity, and when manipulating distortion models defined over multiple variables.

Keywords: neighbourhood models, independence, information fusion, imprecise probabilities

1. Introduction

Among the several imprecise probability models that are representable by means of credal sets, distortion models, defined as a ball around an initial probability, are quite practical, as their specification requires only a distance and a bound on it. This makes them instrumental models for various tasks, such as robustness analysis.

However, the mathematical properties, as well as the interpretation of such neighbourhood models, heavily depend on the chosen distance. We recently conducted a thorough analysis [30, 31] of the probability sets induced by different distances, when those probability sets are convex polytopes[1] defined on finite spaces.

However, we did not explore what happens to those models when they are processed through some reasoning tool. This is what we aim to do in this paper, where we look in particular at two aspects often present in different applications of imprecise probabilistic models:

- Merging or fusing together multiple models bearing on the same quantity or domain [36]: such a process is typically required when probabilistic assessments are provided by multiple sources such as expert opinions in system modelling [38], or outputs from classifiers in machine learning [15].
- Combining multiple models defined on multiple quantities with different domains: such a process is typically required in risk analysis [3] or in general in applications involving multivariate domains.

Therefore, this paper can also be seen as a companion paper to the aforementioned studies [30, 31], filling out the comparative analysis of the models carried out in [31, Sec. 5].

To make this paper self contained, we first provide necessary notions and notations in Section 2, and then investigate in the following sections the behaviour

[1] Excluding therefore distances or divergences such as the Euclidean distance or Kullback-Leibler divergence, that do not induce polytopes.

of different models under merging and combination, reminding the basics of each distortion model in the corresponding section. More precisely:

- Section 3 deals with the *Pari mutuel model* (PMM), that originates from betting schemes on horse-racing. It has been studied from the point of view of imprecise probabilities in [29, 40, 53], and in [30] as a distortion model.
- Section 4 deals with the *Linear-Vacuous* (LV) *model*, that consists in taking a mixture between a precise probability measure and the set of all possible distributions. This model has been used for instance in robust statistics [21]. From the point of view of imprecise probabilities it was studied in [53], and as a distortion model in [30, Sec. 5]. Moreover, it corresponds to the well-known and basic discounting approach in evidence theory.
- Section 5 deals with the *constant-odds ratio* (COR) *model*, that has the advantage of being stable under Bayesian updating. The constant odds ratio was given a behavioural interpretation in [53, Sec. 2.9.4]. We refer to [4, 5, 41, 50] for some applications of this model, and to [30, Sec. 6] for a detailed study of some of its properties as a distortion model.
- Section 6 deals with the *total variation* (TV) *model*, which is the distortion model induced by the *total variation* distance, i.e., the maximum absolute difference between two probability measures. We refer to [53, Sec. 3.2.4], [40, Sec. 3.2] and [20], [31, Sec. 2] for some studies of this distortion model. The total variation distance has also been used in other setting, such as to find consensus between various probabilities [43] (a task close in spirit to the one of information merging).
- Section 7 deals with the *Kolmogorov* (K) *model*, a distortion model induced by the Kolmogorov distance between cumulative distributions. It is connected to imprecise cumulative distribution functions, also called *p-boxes*. This distortion model was analysed in detail in [31, Sec. 3].
- Section 8 deals with the L_1 model, a distortion model induced by the L_1 distance. While this distance has been used in robust statistics [42], it was studied as an imprecise model for the first time in [31, Sec. 4].

Finally, in Section 9 we provide our final comments and remarks.

2. Preliminary concepts

This section introduces the chosen notations, as well as the necessary general elements on distortion models and their processing. Readers interested in further details about those models can refer to [30, 31].

2.1. **Notation and basic notions about probability.** We consider in this paper finite possibility spaces, that will be denoted by $\mathcal{X}$, $\mathcal{Y}$ or their product space $\mathcal{X}\times\mathcal{Y}$. We denote by $\mathcal{P}(\mathcal{X})$ the power set of a space $\mathcal{X}$, by $\mathbb{P}(\mathcal{X})$ the set of probability measures on $\mathcal{X}$, and by $\mathbb{P}^*(\mathcal{X})$ the set of probability measures P satisfying $P(A)\in(0,1)$ for any $A\neq\emptyset,\mathcal{X}$.

Whenever $\mathcal{X}=\{x_1,\dots,x_n\}$ is equipped with a total order, we will assume that $x_1<\dots<x_n$. In that case, we denote by F_P the cumulative distribution function (cdf, for short) associated with the probability measure P, given by

$$F_P(x_i)=P(\{x_1,\dots,x_i\}) \text{ for every } i=1,\dots,n.$$

When we deal with two ordered spaces $\mathcal{X}$ and $\mathcal{Y}=\{y_1,\dots,y_m\}$ $(y_1<\dots<y_m)$, and we consider the product space $\mathcal{X}\times\mathcal{Y}$, every $P\in\mathbb{P}(\mathcal{X}\times\mathcal{Y})$ has an associated

bivariate cdf F_P given by

$$F_P(x_i, y_j) = P\big(\{x_1, \dots, x_i\} \times \{y_1, \dots, y_j\}\big) \quad \forall i = 1, \dots, n, \ \forall j = 1, \dots, m.$$

Given a probability measure $P \in \mathbb{P}(\mathcal{X} \times \mathcal{Y})$ or its associated bivariate cdf F_P, we denote by $P^{\mathcal{X}}, F^{\mathcal{X}}$ and $P^{\mathcal{Y}}, F^{\mathcal{Y}}$ its $\mathcal{X}$ and $\mathcal{Y}$ marginals, respectively, given by:

$$P^{\mathcal{X}}(A) = P(A \times \mathcal{Y}) \quad \forall A \subseteq \mathcal{X}, \quad F^{\mathcal{X}}(x_i) = F_P(x_i, y_m) \quad \forall i = 1, \dots, n.$$

$$P^{\mathcal{Y}}(B) = P(\mathcal{X} \times B) \quad \forall B \subseteq \mathcal{Y}, \quad F^{\mathcal{Y}}(y_j) = F_P(x_n, y_j) \quad \forall j = 1, \dots, m.$$

Also, every bivariate cdf F_P satisfies the Fréchet-Hoeffding inequalities:

$$\max\left\{F^{\mathcal{X}}(x_i) + F^{\mathcal{Y}}(y_j) - 1, 0\right\} \le F_P(x_i, y_j) \le \min\left\{F^{\mathcal{X}}(x_i), F^{\mathcal{Y}}(y_j)\right\}$$

for every $i = 1, \dots, n$ and $j = 1, \dots, m$.

2.2. **Imprecise probabilities.** Let us introduce the main notions from the theory of imprecise probabilities that we shall use in this paper. We refer to [2, 47, 53] for a deeper discussion of this theory.

A *lower probability* on a possibility space $\mathcal{X}$ is a function $\underline{P} : \mathcal{P}(\mathcal{X}) \to [0, 1]$ that is monotone ($A \subseteq B$ implies $\underline{P}(A) \le \underline{P}(B)$) and normalised ($\underline{P}(\emptyset) = 0, \underline{P}(\mathcal{X}) = 1$). To any lower probability, we can associate a *credal set*, which is a closed and convex set of probability measures defined as:

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

$\underline{P}$ is called *coherent* if and only if $\mathcal{M}(\underline{P})$ is non-empty and $\underline{P}(A) = \min_{P \in \mathcal{M}(\underline{P})} P(A)$ for every $A \subseteq \mathcal{X}$. We will assume that all the lower probabilities we consider in this paper satisfy this consistency requirement. Since the credal set $\mathcal{M}(\underline{P})$ is closed and convex, it is determined by its *extreme points*, which are those probability measures $P \in \mathbb{P}(\mathcal{X})$ such that

$$(\forall P_1, P_2 \in \mathcal{M}(\underline{P}), \alpha \in (0, 1))(P = \alpha P_1 + (1 - \alpha)P_2 \Rightarrow P_1 = P_2 = P).$$

We can associate with a coherent lower probability $\underline{P}$ its conjugate *coherent upper probability*, given by $\overline{P}(A) = 1 - \underline{P}(A^c)$ for every $A \subseteq \mathcal{X}$. In fact, every probability measure $P \in \mathcal{M}(\underline{P})$ also satisfies $P(A) \le \overline{P}(A)$ for every $A \subseteq \mathcal{X}$.

A more general notion than lower probability is that of *lower prevision*. A *gamble* on $\mathcal{X}$ is a real-valued function $f : \mathcal{X} \to \mathbb{R}$, and the set of all the gambles on $\mathcal{X}$ is denoted by $\mathcal{L}(\mathcal{X})$. A *lower prevision* is a map $\underline{P} : \mathcal{L}(\mathcal{X}) \to \mathbb{R}$. Its associated credal set is

$$\mathcal{M}(\underline{P}) = \{P \in \mathbb{P}(\mathcal{X}) \mid P(f) \ge \underline{P}(f) \ \forall f \in \mathcal{L}(\mathcal{X})\},$$

where, in order to ease the notation, we are using the same symbol to denote a probability measure P and its associated expectation operator. A lower prevision $\underline{P}$ is called *coherent* if and only if $\mathcal{M}(\underline{P})$ is non-empty and $\underline{P}(f) = \min_{P \in \mathcal{M}(\underline{P})} P(f)$ for every gamble $f \in \mathcal{L}(\mathcal{X})$. Given a coherent lower prevision $\underline{P}$, its associated conjugate *coherent upper prevision* is given by $\overline{P}(f) = -\underline{P}(-f)$ for any $f \in \mathcal{L}(\mathcal{X})$.

When a coherent lower prevision $\underline{P}$ is restricted to indicators[2] of events, i.e., we restrict to the set of gambles $\{I_A \mid A \subseteq \mathcal{X}\}$, the coherent lower prevision becomes a coherent lower probability, where we use the notation $\underline{P}(A)$ for $\underline{P}(I_A)$. However, two different coherent lower previsions may have the same restriction to indicators of events, and as a consequence may induce the same coherent lower probability.

[2]Recall that the indicator I_A of an event A is the gamble that takes the value 1 on the elements of A and 0 elsewhere.

A subfamily of particular interest within coherent lower probabilities is given by those that satisfy 2-monotonicity:

Definition 1. *A lower probability* $\underline{P} : \mathcal{P}(\mathcal{X}) \to [0,1]$ *is* 2-monotone *when it satisfies*

$$\underline{P}(A \cup B) + \underline{P}(A \cap B) \geq \underline{P}(A) + \underline{P}(B) \quad \forall A, B \subseteq \mathcal{X}.$$

We refer to [14, 51] for a detailed account of these models. In particular, they include those associated with probability boxes. Assuming that $\mathcal{X}$ is equipped with a total order, a (univariate) *probability box* [19] is a pair of cdfs $\underline{F}, \overline{F} : \mathcal{X} \to [0,1]$, called lower and upper cdfs, satisfying $\underline{F} \leq \overline{F}$. A p-box $(\underline{F}, \overline{F})$ defines a credal set $\mathcal{M}(\underline{F}, \overline{F})$ by:

$$\mathcal{M}(\underline{F}, \overline{F}) = \big\{P \in \mathbb{P}(\mathcal{X}) \mid \underline{F} \leq F_P \leq \overline{F}\big\}.$$

The lower and upper envelopes of $\mathcal{M}(\underline{F}, \overline{F})$, $\underline{P}_{(\underline{F},\overline{F})}$ and $\overline{P}_{(\underline{F},\overline{F})}$, are conjugate coherent lower and upper probabilities satisfying

$$\underline{P}_{(\underline{F},\overline{F})}\big(\{x_1, \ldots, x_i\}\big) = \underline{F}(x_i) \text{ and } \overline{P}_{(\underline{F},\overline{F})}\big(\{x_1, \ldots, x_i\}\big) = \overline{F}(x_i) \quad \forall i = 1, \ldots, n.$$

These coherent lower and upper probabilities can be computed following the results in [48], where it was proven that $\underline{P}_{(\underline{F},\overline{F})}$ is 2-monotone (in fact, it satisfies the stronger property of complete monotonicity). We refer to [17, 27, 48, 49] for detailed studies about (univariate) p-boxes.

2.3. **Distortion models.** In this paper, our focus is on a family of imprecise probability models that are usually referred to as *distortion models* [6, 9, 21]. They can arise by considering a neighbourhood model around some probability measure using some distorting function d and some distortion factor $\delta > 0$ (as in [22, 41, 50]), or making a transformation of a given (lower) probability (as in [7, 10, 44]). We showed in [30, Prop. 2] that this second approach can be embedded in the first, and for this reason we consider here distortion models defined in terms of neighbourhoods. Given a distorting function $d : \mathbb{P}(\mathcal{X}) \times \mathbb{P}(\mathcal{X}) \to [0, \infty)$, a distortion parameter $\delta > 0$ and a fixed probability measure $P_0 \in \mathbb{P}(\mathcal{X})$, we can define the following set of probabilities:

$$B_d^{\delta}(P_0) = \{P \in \mathbb{P}(\mathcal{X}) \mid d(P, P_0) \leq \delta\}.$$

Whenever d is convex and continuous, $B_d^{\delta}(P_0)$ is a convex and closed set of probabilities [30, Prop. 1]. This means that if we consider its lower envelope:

$$\underline{P}_d(f) = \min\big\{P(f) \mid P \in B_d^{\delta}(P_0)\big\} \quad \forall f \in \mathcal{L}(\mathcal{X}),$$

the credal sets $\mathcal{M}(\underline{P}_d)$ and $B_d^{\delta}(P_0)$ coincide, and $\underline{P}_d$ is a coherent lower prevision.

For the sake of simplicity, in [30, 31] we assumed that $P_0 \in \mathbb{P}^*(\mathcal{X})$, i.e. P_0 is strictly positive for every non-empty event, and that also δ is small enough so that $B_d^{\delta}(P_0) \subseteq \mathbb{P}^*(\mathcal{X})$. In this paper, we also assume that this simplifying hypothesis holds.

2.4. **Processing imprecise probabilistic models.** One of the criteria that may be used in order to choose one uncertainty model over another is that it is closed under a number of operations that we may perform. These operations may arise for instance from the combination of different sources of information, the extension to a different domain, or the updating under the presence of new information. Next, we introduce the procedures that shall be analysed in this paper.

2.4.1. *Merging.* The first operation we shall consider in this paper is that of *merging*. By this, we refer to the procedure where we aggregate a number of belief models, defined on the same domain $\mathcal{X}$, into a unique one. These models may arise as the opinion of different experts or several data sources, for instance. The problem of aggregating imprecise beliefs has been analysed from the axiomatic point of view by Walley [52]. Other relevant works on this topic are [36, 37].

In this paper, we shall focus on the three most fundamental merging procedures: those of *conjunction*, *disjunction* and *convex mixture*. If we model our beliefs in terms of two credal sets $\mathcal{M}_1, \mathcal{M}_2$, they will produce the sets $\mathcal{M}_1 \cap \mathcal{M}_2$, $\mathcal{M}_1 \cup \mathcal{M}_2$ and $\epsilon\mathcal{M}_1 + (1-\epsilon)\mathcal{M}_2 = \{\epsilon P_1 + (1-\epsilon)P_2 | P_i \in \mathcal{M}_i\}$ with $\epsilon \in [0,1]$, respectively.

In terms of the lower probabilities associated with these sets, it should be noted that, while $\mathcal{M}_1 \cup \mathcal{M}_2$ is not convex in general, its lower envelope, that coincides with the lower envelope of its convex hull $ch(\mathcal{M}_1 \cup \mathcal{M}_2)$, is given by $\underline{P} := \min\{\underline{P}_1, \underline{P}_2\}$, where $\underline{P}_1, \underline{P}_2$ denote the lower envelopes of $\mathcal{M}_1, \mathcal{M}_2$, respectively. Since we focus in this paper on lower probabilities and previsions, we can restrict ourselves to[3] $ch(\mathcal{M}_1 \cup \mathcal{M}_2)$. While this disjunction will always determine a coherent lower probability by considering its associated lower envelope of events, the latter may not belong to the same family as the original models $\underline{P}_1, \underline{P}_2$. In that case, one possibility would be to consider an outer approximation; in this respect, our earlier work in [24, 33, 34] shall be useful.

In contrast, while $\mathcal{M}_1 \cap \mathcal{M}_2$ is convex, its lower envelope $\underline{P}$ will dominate in general $\max\{\underline{P}_1, \underline{P}_2\}$. It is not difficult to see that for any linear prevision P it holds that $P \in \mathcal{M}_1 \cap \mathcal{M}_2$ if and only if $P \geq \max\{\underline{P}_1, \underline{P}_2\}$. This means that $\underline{P}$ is the *natural extension* (the smallest dominating coherent lower probability) of $\max\{\underline{P}_1, \underline{P}_2\}$. A sufficient condition for the equality between them is precisely the convexity of $\mathcal{M}_1 \cup \mathcal{M}_2$, as shown in [54, Thm. 6]. The equality between the lower envelope of $\mathcal{M}_1 \cap \mathcal{M}_2$ and $\max\{\underline{P}_1, \underline{P}_2\}$ was investigated in the case of possibility measures in [25], and it will be analysed here for distortion models.

Finally, $\epsilon\mathcal{M}_1 + (1-\epsilon)\mathcal{M}_2$ is always convex, and its lower envelope is such that $\underline{P} := \epsilon\underline{P}_1 + (1-\epsilon)\underline{P}_2$.

2.4.2. *Marginal and joint models in multivariate settings.* Another relevant scenario is the restriction of the model to a smaller domain or its extension to a larger one. In this paper, we shall focus on the case where our possibility space is the product $\mathcal{X} \times \mathcal{Y}$ of two finite spaces. In that case, we may move from the joint model to the marginals, or viceversa.

Marginalisation

In this first case, given a joint model $\underline{P}^{\mathcal{X},\mathcal{Y}}$ defined on the space $\mathcal{X} \times \mathcal{Y}$, we can consider the marginal models $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$, defined on $\mathcal{X}$ and $\mathcal{Y}$, respectively. Their corresponding credal sets $\mathcal{M}(\underline{P}^{\mathcal{X}})$ and $\mathcal{M}(\underline{P}^{\mathcal{Y}})$ are formed by the $\mathcal{X}$- and $\mathcal{Y}$-projections of the probability measures in $\mathcal{M}(\underline{P}^{\mathcal{X},\mathcal{Y}})$, respectively.

[3] It should however be noted that convexity is not desirable in every situation: for instance Seidenfeld *et al.* [45] show in a decision-making context that when considering pairs of imprecise probabilities and utilities, one may have to let go of convexity.

Independent products

Conversely, we may start from two marginal models $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$ defined on $\mathcal{X}$ and $\mathcal{Y}$, respectively, and build a joint model on $\mathcal{X} \times \mathcal{Y}$ that is compatible with them. When the sources are assumed to be independent, this leads us to consider an *independent product*. Out of the several extensions of the notion of independence to the imprecise case [11], we consider in this paper the *strong product* of $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$, that we shall denote $\underline{P}^{\mathcal{X}} \boxtimes \underline{P}^{\mathcal{Y}}$. It is the lower envelope of

$$\mathcal{M}(\underline{P}^{\mathcal{X}}) \boxtimes \mathcal{M}(\underline{P}^{\mathcal{Y}}) = \{P^{\mathcal{X}} \times P^{\mathcal{Y}} \mid P^{\mathcal{X}} \in \mathcal{M}(\underline{P}^{\mathcal{X}}),\ P^{\mathcal{Y}} \in \mathcal{M}(\underline{P}^{\mathcal{Y}})\}, \tag{1}$$

where $P^{\mathcal{X}} \times P^{\mathcal{Y}}$ denotes the joint probability defined using the marginals $P^{\mathcal{X}}$ and $P^{\mathcal{Y}}$ through stochastic independence. The strong product $\underline{P}^{\mathcal{X}} \boxtimes \underline{P}^{\mathcal{Y}}$ and its conjugate $\overline{P}^{\mathcal{X}} \boxtimes \overline{P}^{\mathcal{Y}}$ satisfy the following factorisation properties on events:

$$\begin{aligned} \underline{P}^{\mathcal{X}} \boxtimes \underline{P}^{\mathcal{Y}}(A \times B) &= \underline{P}^{\mathcal{X}}(A) \cdot \underline{P}^{\mathcal{Y}}(B) \quad \text{for any } A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}. \\ \overline{P}^{\mathcal{X}} \boxtimes \overline{P}^{\mathcal{Y}}(A \times B) &= \overline{P}^{\mathcal{X}}(A) \cdot \overline{P}^{\mathcal{Y}}(B) \quad \text{for any } A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}. \end{aligned} \tag{2}$$

Natural extension of marginal models

Alternatively, given two marginal models $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$ on $\mathcal{X}$ and $\mathcal{Y}$, respectively, we may look for the most conservative joint model in $\mathcal{X} \times \mathcal{Y}$ with these given marginals, imposing no dependence assumption whatsoever. Using Walley's terminology [53], this corresponds to considering the *natural extension* $\underline{E}$ of the coherent lower probability $\underline{P}$ that is defined on $\{A \times \mathcal{Y} : A \subseteq \mathcal{X}\} \cup \{\mathcal{X} \times B : B \subseteq \mathcal{Y}\}$ by

$$\underline{P}(A \times \mathcal{Y}) = \underline{P}^{\mathcal{X}}(A) \quad \text{and} \quad \underline{P}(\mathcal{X} \times B) = \underline{P}^{\mathcal{Y}}(B).$$

It can be equivalently obtained as the lower envelope of the credal set given by those probabilities whose marginals are compatible with the information provided by $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$:

$$\mathcal{E}(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}}) = \Big\{P \in \mathbb{P}(\mathcal{X} \times \mathcal{Y}) \mid P^{\mathcal{X}} \in \mathcal{M}(\underline{P}^{\mathcal{X}}),\ P^{\mathcal{Y}} \in \mathcal{M}(\underline{P}^{\mathcal{Y}})\Big\}. \tag{3}$$

If we consider the upper envelope $\overline{E}$ of this set we obtain the conjugate upper prevision of $\underline{E}$, that corresponds to the (upper) natural extension of the coherent upper probability $\overline{P}$ given by

$$\overline{P}(A \times \mathcal{Y}) = \overline{P}^{\mathcal{X}}(A) \quad \text{and} \quad \overline{P}(\mathcal{X} \times B) = \overline{P}^{\mathcal{Y}}(B).$$

Equivalently,

$$\underline{E}(C) = \inf_{P \in \mathcal{E}(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}})} P(C), \quad \overline{E}(C) = \sup_{P \in \mathcal{E}(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}})} P(C) \quad \forall C \subseteq \mathcal{X} \times \mathcal{Y}. \tag{4}$$

The study of a joint distribution with given marginals has received a long standing attention in the literature; see for instance [1, 23, 26, 46]. Its existence is trivial in situations like the one considered in this paper: where the marginals are established in disjoint sets of variables. In that case, we can use the techniques of natural extension to determine the lower and upper envelopes of all such joints. Our next proposition gives the expression of $\underline{E}, \overline{E}$ on Cartesian products of events:

Proposition 1. *Let $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$ be two coherent lower probabilities on $\mathcal{X}$ and $\mathcal{Y}$, respectively, with conjugates $\overline{P}^{\mathcal{X}}$ and $\overline{P}^{\mathcal{Y}}$. Then:*

$$\underline{E}(A \times B) = \max\left\{\underline{P}^{\mathcal{X}}(A) + \underline{P}^{\mathcal{Y}}(B) - 1, 0\right\}, \tag{5}$$

$$\overline{E}(A \times B) = \min\left\{\overline{P}^{\mathcal{X}}(A), \overline{P}^{\mathcal{Y}}(B)\right\} \quad \forall A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}. \tag{6}$$

Proof. This is a consequence of the result stated by Walley in [53, Sec. 3.1.1] for intersections of events. □

In the particular case where we start with precise marginals $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ on $\mathcal{X}, \mathcal{Y}$, we obtain

$$\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(A \times B) = \max\left\{P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B) - 1, 0\right\}, \tag{7}$$

$$\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(A \times B) = \min\left\{P_0^{\mathcal{X}}(A), P_0^{\mathcal{Y}}(B)\right\} \quad \forall A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}. \tag{8}$$

From Proposition 1 we obtain a simple procedure for computing the natural extension in Equation (3) for events of the type $A \times B$, which are simply the Fréchet-Hoeffding bounds. One may think that the expressions in Equations (5) and (6) also hold when considering any event $C \subseteq \mathcal{X} \times \mathcal{Y}$, and decomposing it into its $\mathcal{X}$- and $\mathcal{Y}$-projections. However, our next example shows that this is not always the case.

Example 1. *Let $\mathcal{X} = \{x_1, x_2, x_3\}$ and $\mathcal{Y} = \{y_1, y_2, y_3\}$ and let the coherent conjugate lower and upper probabilities be given by:*

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}^{\mathcal{X}}(A)$	0.1	0.2	0.5	0.4	0.7	0.8
$\overline{P}^{\mathcal{X}}(A)$	0.2	0.3	0.6	0.5	0.8	0.9

B	$\{y_1\}$	$\{y_2\}$	$\{y_3\}$	$\{y_1, y_2\}$	$\{y_1, y_3\}$	$\{y_2, y_3\}$
$\underline{P}^{\mathcal{Y}}(B)$	0.1	0.2	0.4	0.4	0.7	0.7
$\overline{P}^{\mathcal{Y}}(B)$	0.3	0.3	0.6	0.6	0.8	0.9

Consider the event $C_1 = \{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$, whose projections are $C_1^{\mathcal{X}} = \mathcal{X}$, $C_1^{\mathcal{Y}} = \mathcal{Y}$, and the probability mass function P given by

y_3	0	0	0.4
y_2	0	0.1	0.2
y_1	0.2	0.1	0
$P\big(\{(x_i, y_j)\}\big)$	x_1	x_2	x_3

Its marginals dominate $\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}}$, respectively. As a consequence, $\underline{E}(C_1) \leq P(C_1) = 0.7$, while

$$\max\left\{\underline{P}^{\mathcal{X}}(C_1^{\mathcal{X}}) + \underline{P}^{\mathcal{Y}}(C_2^{\mathcal{Y}}) - 1, 0\right\} = 1. \blacklozenge$$

2.5. **Aim of the paper.** In [30, 31], we carried out a comparative analysis of six distortion models: the pari mutuel, linear vacuous, constant odds ratio, total variation, Kolmogorov and L_1 distance models. The aim was to give guidelines about which model performs better on some respect or which one is more appropriate in each particular scenario. Specifically, we analysed the following features: (i) the amount of imprecision present in the model once the probability measure P_0 and the distortion factor $\delta > 0$ are fixed; (ii) the properties of the associated lower probability as a non-additive measure; (iii) the complexity of their associated neighbourhood models, in terms of the number of extreme points; and (iv) the behaviour of the model under conditioning.

Our goal in this paper is to complement the analysis performed in [31, Sec. 5] by investigating the behaviour of the different families of distortion models under the procedures described in Section 2.4. Specifically, we shall tackle the following problems:

Merging: We first analyse if the distortion models are closed under merging. For this aim, we consider two distortion models $B_d^{\delta_1}\big(P_0^1\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_d^{\delta_2}\big(P_0^2\big) \subseteq \mathbb{P}^*(\mathcal{X})$ in some specific family. Our aim is to know whether their conjunction $B_d^{\delta_1}\big(P_0^1\big) \cap B_d^{\delta_2}\big(P_0^2\big)$, their disjunction $B_d^{\delta_1}\big(P_0^1\big) \cup B_d^{\delta_2}\big(P_0^2\big)$ or their mixture $\epsilon B_d^{\delta_1}\big(P_0^1\big) + (1-\epsilon) B_d^{\delta_2}\big(P_0^2\big)$ belong to the same family, in the sense that it is equal to $B_d^{\delta^*}\big(P_0^*\big)$ for some appropriate δ^* and P_0^*. As we shall see, this is almost always the case for the convex mixture[4], sometimes the case for the conjunction, and never for the disjunction. In this last case, we might then consider the convex hull of the disjunction $ch\big(B_d^{\delta_1}(P_0^1) \cup B_d^{\delta_2}(P_0^2)\big)$ and investigate whether it has a unique outer approximation in the same family [8, 33, 34]. By an *outer approximation* of a coherent lower probability $\underline{P}$ in some family $\mathcal{C}$ we mean some $\underline{Q} \in \mathcal{C}$ such that $\underline{Q} \leq \underline{P}$. We will say that the outer approximation is *undominated* when there is no $\underline{Q}' \in \mathcal{C}$ such that $\underline{Q} \lneq \underline{Q}' \leq \underline{P}$.

Note that since we are assuming that $B_d^{\delta_1}\big(P_0^1\big)$ and $B_d^{\delta_2}\big(P_0^2\big)$ are included in $\mathbb{P}^*(\mathcal{X})$, their convex mixture and their intersection will be also included in $\mathbb{P}^*(\mathcal{X})$. However, an undominated outer approximation of the disjunction need not be included in $\mathbb{P}^*(\mathcal{X})$.

Marginalisation: Given a distortion model $B_d^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big) \subseteq \mathbb{P}^*\big(\mathcal{X} \times \mathcal{Y}\big)$ with associated lower prevision $\underline{P}_d$, we want to know whether the marginal models $\underline{P}_d^{\mathcal{X}}$ and $\underline{P}_d^{\mathcal{Y}}$ correspond to distortion models of the same family on $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(\mathcal{Y})$, respectively. In other words, we want to know if $\mathcal{M}\big(\underline{P}_d^{\mathcal{X}}\big) = B_d^{\delta}\big(P_0^{\mathcal{X}}\big)$ and $\mathcal{M}\big(\underline{P}_d^{\mathcal{Y}}\big) = B_d^{\delta}\big(P_0^{\mathcal{Y}}\big)$.

Independent products: Consider two distortion models $B_d^{\delta}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_d^{\delta}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$ with the same distortion parameter. We want to build a joint model using independence, and we want to know whether the joint model belongs to the same family. Two solutions are then possible in the setting of this study:

- Combine $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ into a joint probability $P_0^{\mathcal{X},\mathcal{Y}} \in \mathbb{P}^*(\mathcal{X} \times \mathcal{Y})$ and apply the distortion to it. In this way, we obtain the distortion model $B_d^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big)$. We shall denote by $\underline{P}^{\mathcal{X}\times\mathcal{Y}}$ and $\overline{P}^{\mathcal{X}\times\mathcal{Y}}$ the resulting lower and upper probabilities obtained as the lower and upper envelope of the credal set $B_d^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big)$, respectively.
- Consider the distortion models $B_d^{\delta}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_d^{\delta}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$ and combine them using the strong product. By Equation (1), this produces the set $\mathcal{M}\big(\underline{P}^{\mathcal{X}}\big) \boxtimes \mathcal{M}\big(\underline{P}^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{X} \times \mathcal{Y})$, whose lower

[4]The constant odds ratio model is an exception.

and upper envelopes are the lower and upper probabilities $\underline{P}^{\mathcal{X}} \boxtimes \underline{P}^{\mathcal{Y}}$ and $\overline{P}^{\mathcal{X}} \boxtimes \overline{P}^{\mathcal{Y}}$.

We wonder if the lower envelopes of $B_d^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big)$ and $\mathcal{M}\big(\underline{P}^{\mathcal{X}}\big) \boxtimes \mathcal{M}\big(\underline{P}^{\mathcal{Y}}\big)$ coincide, or in case they do not, whether there is a dominance relationship between them, meaning that one of the procedures is more precise than the other. In other words, we shall analyse whether there is some dominance relation between $\underline{P}^{\mathcal{X}\times\mathcal{Y}}$ and $\underline{P}^{\mathcal{X}} \boxtimes \underline{P}^{\mathcal{Y}}$.

We will also check whether the different models are closed under the operation of independent product, that is, whether $\underline{P}^{\mathcal{X}} \boxtimes \underline{P}^{\mathcal{Y}}$ is also a distortion model of the same family.

Natural extension: Consider two marginal distortion models $B_d^{\delta_{\mathcal{X}}}\left(P_0^{\mathcal{X}}\right) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_d^{\delta_{\mathcal{Y}}}\left(P_0^{\mathcal{Y}}\right) \subseteq \mathbb{P}^*(\mathcal{Y})$ with associated lower probabilities $\underline{P}^{\mathcal{X}}$ and $\underline{P}^{\mathcal{Y}}$. We want to compute their least committal extension to $\mathcal{P}(\mathcal{X}\times\mathcal{Y})$, i.e. the natural extension given in Equation (3). In this case, the credal set of the natural extension is

$$\mathcal{E}\big(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}}\big) = \big\{P \in \mathbb{P}(\mathcal{X}\times\mathcal{Y}) \mid P^{\mathcal{X}} \in B_d^{\delta_{\mathcal{X}}}\left(P_0^{\mathcal{X}}\right),\ P^{\mathcal{Y}} \in B_d^{\delta_{\mathcal{Y}}}\left(P_0^{\mathcal{Y}}\right)\big\}.$$

We already know the expression of $\underline{E}$ and its conjugate $\overline{E}$ on events like $A\times B$ (see Equations (5) and (6)). We wonder whether we can give a simple expression of $\mathcal{E}(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}})$, $\underline{E}$ and $\overline{E}$ and also whether $\mathcal{E}(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}})$ is also the credal set of a distortion model of the same family.

Note that, although $B_d^{\delta_{\mathcal{X}}}\left(P_0^{\mathcal{X}}\right)$ and $B_d^{\delta_{\mathcal{Y}}}\left(P_0^{\mathcal{Y}}\right)$ are included in $\mathbb{P}^*(\mathcal{X})$ and $\mathbb{P}^*(\mathcal{Y})$, respectively, we cannot guarantee the resulting model $\mathcal{E}\big(\underline{P}^{\mathcal{X}}, \underline{P}^{\mathcal{Y}}\big)$ to be included in $\mathbb{P}^*(\mathcal{X}\times\mathcal{Y})$.

From now on, we devote one section to each of the six distortion models mentioned in the introduction and we analyse their behaviour under the previous operations.

3. Pari mutuel model

The first distortion model we analyse in this paper is the pari mutuel model (PMM, for short):

Definition 2. *Given a probability measure P_0 and a distortion factor $\delta > 0$, the associated* pari mutuel model *is determined by the following lower and upper probabilities:*

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

Since we are assuming that our initial model satisfies $P_0 \in \mathbb{P}^*(\mathcal{X})$ and that its lower probability $\underline{P}_{PMM}$ takes strictly positive values on non-empty events, the previous expressions simplify to:

$$\underline{P}_{PMM}(A) = (1+\delta)P_0(A) - \delta, \quad \overline{P}_{PMM}(A) = (1+\delta)P_0(A) \quad \forall A \neq \emptyset, \mathcal{X},$$

and taking the trivial values 0 and 1 for $\emptyset$ and $\mathcal{X}$, respectively.

It was shown in [30, Thm. 5] that the credal set $\mathcal{M}(\underline{P}_{PMM})$ coincides with $B^{\delta}_{d_{PMM}}(P_0)$, where $d_{PMM} : \mathbb{P}^*(\mathcal{X}) \times \mathbb{P}^*(\mathcal{X}) \to [0,\infty)$ is the distorting function given by

$$d_{PMM}(P,Q) = \max_{A\subset\mathcal{X}} \frac{Q(A) - P(A)}{1 - Q(A)}.$$

Next, we complement the work in [30] by investigating the behaviour of the family of pari mutuel models under a number of operations.

3.1. **Merging.** Let us first study the behaviour of the PMM under merging operators.

Conjunction

We start by analysing the conjunction of PMMs. Given two neighbourhood models $B^{\delta_1}_{d_{PMM}}(P^1_0) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta_2}_{d_{PMM}}(P^2_0) \subseteq \mathbb{P}^*(\mathcal{X})$, it was established in [29, Prop. 12] that their intersection is non-empty iff

$$\sum_{x\in\mathcal{X}} \min\left\{(1+\delta_1)P^1_0(\{x\}), (1+\delta_2)P^2_0(\{x\}), 1\right\} \geq 1. \tag{9}$$

This intersection is the PMM $B^{\delta^\cap}_{d_{PMM}}(P^\cap_0)$, for

$$\delta^\cap = \left(\sum_{x\in\mathcal{X}} \min\left\{(1+\delta_1)P^1_0(\{x\}), (1+\delta_2)P^2_0(\{x\})\right\}\right) - 1, \quad \text{and}$$
$$P^\cap_0(\{x\}) = \frac{\min\left\{(1+\delta_1)P^1_0(\{x\}), (1+\delta_2)P^2_0(\{x\})\right\}}{1+\delta^\cap} \quad \forall x \in \mathcal{X}.$$

Disjunction

Regarding the disjunction, the convex hull of $B^{\delta_1}_{d_{PMM}}(P^1_0) \cup B^{\delta_2}_{d_{PMM}}(P^2_0)$ will not be in general a PMM, as we show in the following example.

Example 2. *Consider $P^1_0 = (0.5, 0.3, 0.2), P^2_0 = (0.3, 0.5, 0.2)$ and $\delta_1 = \delta_2 = 0.1$. The associated PMMs $\underline{P}_{PMM_1}, \underline{P}_{PMM_2}$ and their disjunction $\underline{P}^\cup$ are given in the following table:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1,x_2\}$	$\{x_1,x_3\}$	$\{x_2,x_3\}$
$\underline{P}_{PMM_1}$	0.45	0.23	0.12	0.78	0.67	0.45
$\underline{P}_{PMM_2}$	0.23	0.45	0.12	0.78	0.45	0.67
$\underline{P}^\cup$	0.23	0.23	0.12	0.78	0.45	0.45

If it was $\mathcal{M}(\underline{P}^\cup) = B^{\delta}_{d_{PMM}}(P_0)$ for some P_0, δ, then we would obtain

$$\sum_{x\in\mathcal{X}} \underline{P}^\cup(\{x\}) = 0.58 = 1 - 2\delta \Rightarrow \delta = 0.21;$$

on the other hand, the equality

$$0.45 = \underline{P}^\cup(\{x_2,x_3\}) = (1+\delta)P_0(\{x_2,x_3\}) - \delta = \underline{P}^\cup(\{x_2\}) + \underline{P}^\cup(\{x_3\}) + \delta,$$

means that it should be $\delta = 0.1$. Thus, $\underline{P}^\cup$ is not a PMM. ♦

Interestingly, this disjunction has a unique undominated outer-approximation that is a PMM [33, Prop. 7]. It is given by the model $B^{\delta_\cup}_{d_{PMM}}(P_0^\cup)$ such that:

$$\delta^\cup = \left(\sum_{x \in \mathcal{X}} \max\left\{(1+\delta_1)P_0^1(\{x\}), (1+\delta_2)P_0^2(\{x\})\right\} \right) - 1, \quad \text{and}$$

$$P_0^\cup(\{x\}) = \frac{\max\left\{(1+\delta_1)P_0^1(\{x\}), (1+\delta_2)P_0^2(\{x\})\right\}}{1+\delta^\cup} \quad \forall x \in \mathcal{X}.$$

This is the greatest (in terms of $\underline{P}$), or more informative, PMM whose associated neighbourhood includes the disjunction $B^{\delta_1}_{d_{PMM}}(P_0^1) \cup B^{\delta_2}_{d_{PMM}}(P_0^2)$.

Convex mixture

The mixture operation was studied in [33, Sec. 5.1], where it was shown that the convex mixture of two PMM is again a PMM, given by $B^{\delta_\epsilon}_{d_{PMM}}(P_0^\epsilon)$ where

$$1+\delta_\epsilon = \epsilon(1+\delta_1) + (1-\epsilon)(1+\delta_2) \quad \text{and}$$

$$P_0^\epsilon(\{x\}) = \frac{\epsilon(1+\delta_1)P_0^1(\{x\}) + (1-\epsilon)(1+\delta_2)P_0^2(\{x\})}{1+\delta_\epsilon} \quad \forall x \in \mathcal{X}.$$

3.2. **Multivariate setting.** Let us now look at the behaviour of the PMM in a multivariate setting.

Marginalisation

In [29, Sec. 6.2], it was shown that the marginal lower probability $\underline{P}^{\mathcal{X}}$ obtained from a joint PMM $B^{\delta}_{d_{PMM}}(P_0^{\mathcal{X},\mathcal{Y}}) \subseteq \mathbb{P}^*(\mathcal{X}\times\mathcal{Y})$ is again a PMM $B^{\delta}_{d_{PMM}}(P_0^{\mathcal{X}}) \subseteq \mathbb{P}^*(\mathcal{X})$ with $P_0^{\mathcal{X}}$ the marginal probability of $P_0^{\mathcal{X},\mathcal{Y}}$ on $\mathcal{X}$ and the same distortion factor.

Independent products

When going from marginal models $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ to a joint one under the assumption of independence, it can be seen that there is no dominance relationship between $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}$ (combine through stochastic independence then distort) and $\underline{P}^{\mathcal{X}}_{PMM} \boxtimes \underline{P}^{\mathcal{Y}}_{PMM}$ (distort then combine through strong independence). To see this, note that on the one hand for the Cartesian product of events A, B, it holds that:

$$\begin{aligned}\overline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}(A\times B) &= (1+\delta)P_0^{\mathcal{X}}(A)P_0^{\mathcal{Y}}(B) \\ &\leq (1+\delta)P_0^{\mathcal{X}}(A)(1+\delta)P_0^{\mathcal{Y}}(B) = \overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(A\times B), \quad (10)\end{aligned}$$

where the inequality is strict whenever we consider non-trivial events A, B, i.e. $P_0^{\mathcal{X}}(A), P_0^{\mathcal{Y}}(B) \in (0,1)$. On the other hand, for events E that are not products, the relationship between $\overline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}(E)$ and $\overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(E)$ can be the reverse one, as we show in our next example:

Example 3. *Consider the spaces $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2\}$ and the probability measures $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ given by:*

$$P_0^{\mathcal{X}}(\{x_1\}) = 0.3, \quad P_0^{\mathcal{X}}(\{x_2\}) = 0.7, \quad P_0^{\mathcal{Y}}(\{y_1\}) = P_0^{\mathcal{Y}}(\{y_2\}) = 0.5,$$

and let $\delta = 0.1$. Given the event $E_1 = \{(x_2, y_2)\}^c$, it holds that:

$$\begin{aligned}\overline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}(E_1) &= 1 - \underline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}\big(\{(x_2, y_2)\}\big) = 0.715 \\ &> \overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(E_1) = 1 - \underline{P}^{\mathcal{X}}_{PMM} \boxtimes \underline{P}^{\mathcal{Y}}_{PMM}\big(\{(x_2, y_2)\}\big) \\ &= 1 - \underline{P}^{\mathcal{X}}_{PMM}(\{x_2\})\underline{P}^{\mathcal{Y}}_{PMM}(\{y_2\}) = 0.6985.\end{aligned}$$

On the other hand, if we consider the events $A = \{x_1\}$ and $B = \{y_1\}$, we obtain $\overline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}(A \times B) = 0.165$ and $\overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(A \times B) = 0.1815$, showing that the inequality in Equation (10) may be strict.

Therefore, there is not a dominance relationship between $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}$ and $\underline{P}^{\mathcal{X}}_{PMM} \boxtimes \underline{P}^{\mathcal{Y}}_{PMM}$ (or equivalently between $\overline{P}^{\mathcal{X}\times\mathcal{Y}}_{PMM}$ and $\overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}$). ♦

We may also wonder whether the family of PMM is closed under the strong product, or in other words if $\underline{P}^{\mathcal{X}}_{PMM} \boxtimes \underline{P}^{\mathcal{Y}}_{PMM}$ is still a PMM. The next example shows that this is not the case.

Example 4. *Consider the setting of Example 3 and the events $\{(x_1, y_1)\}, \{(x_1, y_2)\}$ and $\{x_1\} \times \mathcal{Y}$. We obtain*

$$\overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(\{(x_1, y_1)\}) = \overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(\{(x_1, y_2)\}) = 0.33 \cdot 0.55 = 0.1815$$

while

$$\overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}(\{x_1\} \times \mathcal{Y}) = \overline{P}^{\mathcal{X}}_{PMM}(\{x_1\}) = 0.33.$$

Since $0.33 \neq 2 \cdot 0.1815$ and any PMM satisfies $\overline{P}(A \cup B) = \overline{P}(A) + \overline{P}(B)$ if $A \cup B \subset \mathcal{X}$ and $A \cap B = \emptyset$, we conclude that $\overline{P}^{\mathcal{X}}_{PMM} \boxtimes \overline{P}^{\mathcal{Y}}_{PMM}$ is not a PMM. ♦

Natural extension of marginal models

Consider the lower and upper probabilities that are the lower and upper envelopes of $B^{\delta_{\mathcal{X}}}_{d_{PMM}}(P^{\mathcal{X}}_0) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta_{\mathcal{Y}}}_{d_{PMM}}(P^{\mathcal{Y}}_0) \subseteq \mathbb{P}^*(\mathcal{Y})$. Using Equations (5) and (6), we can give the form of $\underline{E}_{PMM}$ and $\overline{E}_{PMM}$ for the events $A \times B \neq \emptyset$, for $A \subset \mathcal{X}$, $B \subset \mathcal{Y}$:

$$\begin{aligned}\overline{E}_{PMM}(A \times B) &= \min\Big\{\overline{P}^{\mathcal{X}}_{PMM}(A), \overline{P}^{\mathcal{Y}}_{PMM}(B)\Big\} \\ &= \min\big\{(1+\delta_{\mathcal{X}})P^{\mathcal{X}}_0(A), (1+\delta_{\mathcal{Y}})P^{\mathcal{Y}}_0(B), 1\big\}. \\ \underline{E}_{PMM}(A \times B) &= \max\Big\{\underline{P}^{\mathcal{X}}_{PMM}(A) + \underline{P}^{\mathcal{Y}}_{PMM}(B) - 1, 0\Big\} \\ &= \max\big\{(1+\delta_{\mathcal{X}})P^{\mathcal{X}}_0(A) - \delta_{\mathcal{X}} + (1+\delta_{\mathcal{Y}})P^{\mathcal{Y}}_0(B) - \delta_{\mathcal{Y}} - 1, 0\big\}.\end{aligned}$$

Moreover, when the distortion parameters coincide, $\delta_{\mathcal{X}} = \delta_{\mathcal{Y}} = \delta$, the previous equations become:

$$\begin{aligned}\overline{E}_{PMM}(A \times B) &= \min\big\{(1+\delta)P^{\mathcal{X}}_0(A), (1+\delta)P^{\mathcal{Y}}_0(B), 1\big\} \\ &= \min\big\{1, (1+\delta)\min\big\{P^{\mathcal{X}}_0(A), P^{\mathcal{Y}}_0(B)\big\}\big\}. \qquad (11) \\ \underline{E}_{PMM}(A \times B) &= \max\big\{(1+\delta)\big(P^{\mathcal{X}}_0(A) + P^{\mathcal{Y}}_0(B)\big) - 2\delta - 1, 0\big\} \\ &= \max\big\{(1+\delta)\big(P^{\mathcal{X}}_0(A) + P^{\mathcal{Y}}_0(B) - 1\big) - \delta, 0\big\} \\ &= (1+\delta)\max\left\{P^{\mathcal{X}}_0(A) + P^{\mathcal{Y}}_0(B) - 1, \frac{\delta}{1+\delta}\right\} - \delta. \qquad (12)\end{aligned}$$

We note that the above expressions for $\overline{E}_{PMM}(A \times B)$ and $\underline{E}_{PMM}(A \times B)$ recall those of the upper and lower probabilities of a PMM.

Even if the expressions in Equations (11) and (12) are only valid for the events of the type $A \times B$, one may think that the natural extension is somehow related to a PMM. Our next result shows that indeed such a connection can be established.

Theorem 2. *Let $B^{\delta}_{d_{PMM}}(P_0^{\mathcal{X}}) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta}_{d_{PMM}}(P_0^{\mathcal{Y}}) \subseteq \mathbb{P}^*(\mathcal{Y})$ be two PMMs with associated lower and upper probabilities $\underline{P}^{\mathcal{X}}_{PMM}, \overline{P}^{\mathcal{X}}_{PMM}$ and $\underline{P}^{\mathcal{Y}}_{PMM}, \overline{P}^{\mathcal{Y}}_{PMM}$. Then, the credal set of the natural extension defined in Equation* (3) *can be expressed as:*

$$\mathcal{E}\big(\underline{P}^{\mathcal{X}}_{PMM}, \underline{P}^{\mathcal{Y}}_{PMM}\big) = \Big\{P \in \mathbb{P}(\mathcal{X} \times \mathcal{Y}) \mid P \leq (1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}\Big\}; \tag{13}$$

equivalently,

$$\overline{E}_{PMM}(C) = \min\big\{(1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(C), 1\big\} \quad \forall C \subseteq \mathcal{X} \times \mathcal{Y}, \tag{14}$$

where $\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}$ corresponds to the upper envelope of the credal set in Equation (3) *applied to the particular case of precise marginals $P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}$.*

Proof. Consider first of all P satisfying $P \leq (1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}$, and let us prove that $P \in \mathcal{E}\big(\underline{P}^{\mathcal{X}}_{PMM}, \underline{P}^{\mathcal{Y}}_{PMM}\big)$. Since

$$\begin{aligned} P(A \times \mathcal{Y}) &\leq (1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(A \times \mathcal{Y}) \\ &= (1+\delta)\min\big\{P_0^{\mathcal{X}}(A), P_0^{\mathcal{Y}}(\mathcal{Y})\big\} = (1+\delta)P_0^{\mathcal{X}}(A) \quad \forall A \subset \mathcal{X}, \\ P(\mathcal{X} \times B) &\leq (1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(\mathcal{X} \times B) \\ &= (1+\delta)\min\big\{P_0^{\mathcal{X}}(\mathcal{X}), P_0^{\mathcal{Y}}(B)\big\} = (1+\delta)P_0^{\mathcal{Y}}(B) \quad \forall B \subset \mathcal{Y}, \end{aligned}$$

we deduce that the $\mathcal{X}$ and $\mathcal{Y}$ marginals of P are dominated by the upper probability determined by $\overline{P}^{\mathcal{X}}_{PMM}$ and $\overline{P}^{\mathcal{Y}}_{PMM}$, and thus that $P \in \mathcal{E}\big(\underline{P}^{\mathcal{X}}_{PMM}, \underline{P}^{\mathcal{Y}}_{PMM}\big)$. Therefore,

$$\mathcal{E}\big(\underline{P}^{\mathcal{X}}_{PMM}, \underline{P}^{\mathcal{Y}}_{PMM}\big) \supseteq \Big\{P \in \mathbb{P}(\mathcal{X} \times \mathcal{Y}) \mid P \leq (1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}\Big\},$$

or equivalently

$$\overline{E}_{PMM}(C) \geq \min\big\{(1+\delta)\overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(C), 1\big\} \quad \forall C \subseteq \mathcal{X} \times \mathcal{Y}. \tag{15}$$

Fix now $C \subseteq \mathcal{X} \times \mathcal{Y}$, take $P \in \mathcal{E}\big(\underline{P}^{\mathcal{X}}_{PMM}, \underline{P}^{\mathcal{Y}}_{PMM}\big)$ such that $P(C) = \overline{E}_{PMM}(C)$, and denote by $P^{\mathcal{X}}$ and $P^{\mathcal{Y}}$ its marginals.

Let us define the probability measures $Q^{\mathcal{X}}, Q^{\mathcal{Y}}$ by means of the equalities

$$Q^{\mathcal{X}}(\{x\}) = \frac{(1+\delta)P_0^{\mathcal{X}}(\{x\}) - P^{\mathcal{X}}(\{x\})}{\delta}, \quad Q^{\mathcal{Y}}(\{y\}) = \frac{(1+\delta)P_0^{\mathcal{Y}}(\{y\}) - P^{\mathcal{Y}}(\{y\})}{\delta}.$$

Note that by construction

$$\sum_{x \in \mathcal{X}} Q^{\mathcal{X}}(\{x\}) = \frac{1+\delta-1}{\delta} = 1,$$

and also $Q^{\mathcal{X}}(\{x\}) \geq 0$ because $P^{\mathcal{X}} \leq (1+\delta)P_0^{\mathcal{X}}$. Therefore, $Q^{\mathcal{X}}$ is a probability measure on $\mathcal{X}$. Similar considerations imply that $Q^{\mathcal{Y}}$ is a probability measure on $\mathcal{Y}$. Let $Q := Q^{\mathcal{X}} \times Q^{\mathcal{Y}}$ denote their independent product, and consider the probability measure P' given by

$$P' = \frac{1}{1+\delta}P + \frac{\delta}{1+\delta}Q.$$

P' is a probability measure because it is a convex combination of probability measures. Moreover, for any $x \in \mathcal{X}$:

$$\begin{aligned}\sum_{y\in\mathcal{Y}} P'\big(\{(x,y)\}\big) &= \sum_{y\in\mathcal{Y}} \left(\frac{1}{1+\delta}P\big(\{(x,y)\}\big) + \frac{\delta}{1+\delta}Q\big(\{(x,y)\}\big)\right) \\ &= \frac{P^{\mathcal{X}}\big(\{x\}\big)}{1+\delta} + \frac{\delta}{1+\delta}Q^{\mathcal{X}}\big(\{x\}\big) \\ &= \frac{P^{\mathcal{X}}\big(\{x\}\big)}{1+\delta} + \frac{(1+\delta)P_0^{\mathcal{X}}\big(\{x\}\big) - P^{\mathcal{X}}\big(\{x\}\big)}{1+\delta} = P_0^{\mathcal{X}}\big(\{x\}\big).\end{aligned}$$

Similarly, the $\mathcal{Y}$-marginal of P' coincides with $P_0^{\mathcal{Y}}$. As a consequence, $P' \in \mathcal{E}\big(P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}\big)$, whence $P'(C) \leq \overline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}(C)$. Moreover

$$\begin{aligned}P'(C) = \sum_{(x,y)\in C} P'\big(\{(x,y)\}\big) &\geq \frac{1}{1+\delta}\sum_{(x,y)\in C} P\big(\{(x,y)\}\big) \\ &= \frac{1}{1+\delta}P(C) = \frac{1}{1+\delta}\overline{E}_{PMM}(C),\end{aligned}$$

so we deduce that

$$\overline{E}_{PMM}(C) \leq \min\Big\{(1+\delta)\overline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}(C), 1\Big\}. \tag{16}$$

By putting together Equations (15) and (16) we conclude that Equations (13) and (14) hold. □

This result shows that the procedures of natural extension and the distortion produced by the PMM commute, in the sense that the natural extension of the coherent upper probability determined by the two marginal PMMs can also be obtained as a PMM starting from the joint probability the two marginals determine on the product events[5]. This is illustrated in Figure 1.

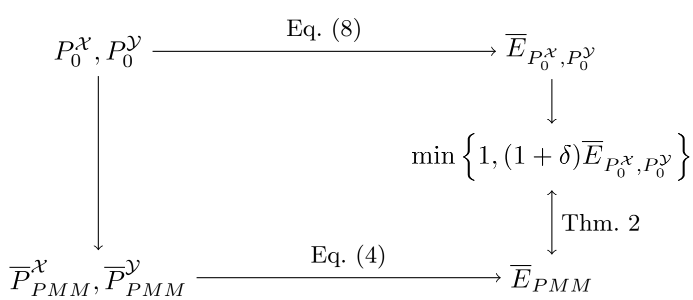

FIGURE 1. Graphical representation of the natural extension of two PMMs.

[5]This second approach is reminiscent of that considered by Moral in [35] for the distortion of credal sets (there with the name *discounting*): he distorts each of the elements in the initial credal sets, and then takes the closure of the union of those credal sets obtained. In particular, he investigated the cases of the total variation distance and the linear vacuous mixtures we shall consider later on in this paper.

4. Linear vacuous mixtures

Our next model is the so-called ϵ-contamination model, or linear vacuous mixture (LV, for short):

Definition 3. *Given a probability measure P_0 and a distortion factor $\delta \in (0,1)$, its associated* linear vacuous mixture *is given by the following conjugate lower and upper probabilities:*

$$\underline{P}_{LV}(A) = (1-\delta)P_0(A), \quad \overline{P}_{LV}(A) = (1-\delta)P_0(A) + \delta \quad \forall A \neq \emptyset, \mathcal{X},$$

$\underline{P}_{LV}(\emptyset) = \overline{P}_{LV}(\emptyset) = 0$ *and* $\underline{P}_{LV}(\mathcal{X}) = \overline{P}_{LV}(\mathcal{X}) = 1$.

The credal set $\mathcal{M}(\underline{P}_{LV})$ coincides with $B^{\delta}_{d_{LV}}(P_0)$, where $d_{LV} : \mathbb{P}^*(\mathcal{X}) \times \mathbb{P}^*(\mathcal{X}) \to [0,\infty)$ is the distorting function given by [30, Thm. 9]:

$$d_{LV}(P,Q) = \max_{A \neq \emptyset} \frac{Q(A) - P(A)}{Q(A)}.$$

Let us analyse the behaviour of the LV model under the different operations introduced in Section 2.4.

4.1. Merging.

Let us first look at the behaviour of LV models under merging.

Conjunction

Similarly to the PMM, the intersection of two LV models is again a LV model, when this intersection is non-empty.

Proposition 3. *Given the distortion models $B^{\delta_1}_{d_{LV}}(P^1_0) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta_2}_{d_{LV}}(P^2_0) \subseteq \mathbb{P}^*(\mathcal{X})$, the set $B^{\delta_1}_{d_{LV}}(P^1_0) \cap B^{\delta_2}_{d_{LV}}(P^2_0)$ is non-empty if and only if*

$$\sum_{x \in \mathcal{X}} \max\left\{(1-\delta_1)P^1_0(\{x\}), (1-\delta_2)P^2_0(\{x\})\right\} \leq 1. \tag{17}$$

In that case, it is induced by the LV model generated by the following probability measure $P^{\cap}_0$ and the distortion parameter $\delta^{\cap}$:

$$\delta^{\cap} = 1 - \sum_{x \in \mathcal{X}} \max\left\{(1-\delta_1)P^1_0(\{x\}), (1-\delta_2)P^2_0(\{x\})\right\}, \quad \textit{and}$$

$$P^{\cap}_0(\{x\}) = \frac{\max\left\{(1-\delta_1)P^1_0(\{x\}), (1-\delta_2)P^2_0(\{x\})\right\}}{1-\delta^{\cap}} \quad \forall x \in \mathcal{X}.$$

Proof. It suffices to notice [30, Sec. 5.1] that a linear vacuous model is equivalent to specific probability intervals that are only lower bounded, i.e. given P_0, δ, the credal set $B^{\delta}_{LV}(P_0)$ is the set of those probability measures P satisfying the constraints

$$(1-\delta)P_0(\{x\}) \leq P(\{x\}) \quad \forall x \in \mathcal{X}.$$

It is then known [13, Sec. 3.2] that the intersection of two such models $B^{\delta_1}_{d_{LV}}(P^1_0) \cap B^{\delta_2}_{d_{LV}}(P^2_0)$ corresponds to the probability interval whose lower bounds are the maximum of their respective lower bounds. From this, it follows that this conjunction is a linear vacuous model when this intersection is non-empty. Equation (17) follows then from [13, Eq. (2)]. □

Disjunction

Regarding the disjunction, the convex hull of $B^{\delta_1}_{d_{LV}}(P^1_0) \cup B^{\delta_2}_{d_{LV}}(P^2_0)$ will in general not be a LV model, not even when $\delta_1 = \delta_2$ as we show in next example.

Example 5. *As in Example 2, take the probability measures* $P^1_0 = (0.5, 0.3, 0.2)$ *and* $P^2_0 = (0.3, 0.5, 0.2)$ *and the distortion factor* $\delta_1 = \delta_2 = 0.1$. *The associated LV models* $\underline{P}_{LV_1}, \underline{P}_{LV_2}$ *and their disjunction* $\underline{P}^{\cup}_{LV}$ *are given in the following table:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{LV_1}$	0.45	0.27	0.18	0.72	0.63	0.45
$\underline{P}_{LV_2}$	0.27	0.45	0.18	0.72	0.45	0.63
$\underline{P}^{\cup}_{LV}$	0.27	0.27	0.18	0.72	0.45	0.45

If there was some probability measure P_0 *and* $\delta > 0$ *such that* $\mathcal{M}(\underline{P}_{LV}) = B^{\delta}_{d_{LV}}(P_0)$, *then we would obtain*

$$\underline{P}^{\cup}_{LV}(\{x_1, x_2\}) = (1-\delta)P_0(\{x_1, x_2\}) = \underline{P}^{\cup}_{LV}(\{x_1\}) + \underline{P}^{\cup}_{LV}(\{x_2\}),$$

which does not hold. As a consequence, the disjunction of $B^{\delta_1}_{d_{LV}}(P^1_0) \cup B^{\delta_2}_{d_{LV}}(P^2_0)$ *does not produce a LV model.* ♦

This disjunction has a unique undominated outer approximation as a LV model, since the greatest LV outer approximation (in terms of $\underline{P}$) of any given credal set is unique [33, Prop. 8]. This undominated outer approximation is the model $B^{\delta_{\cup}}_{d_{LV}}(P^{\cup}_0)$ such that

$$\delta^{\cup} = 1 - \left(\sum_{x \in \mathcal{X}} \min\left\{(1-\delta_1)P^1_0(\{x\}), (1-\delta_2)P^2_0(\{x\})\right\} \right), \quad \text{and}$$

$$P^{\cup}_0(\{x\}) = \frac{\min\left\{(1-\delta_1)P^1_0(\{x\}), (1-\delta_2)P^2_0(\{x\})\right\}}{1-\delta^{\cup}} \quad \forall x \in \mathcal{X}.$$

Convex mixture

The mixture of two LV models, that is, the computation of $B^{\delta_\epsilon}_{d_{LV}}(P^{\epsilon}_0)$ for a given $\epsilon \in (0, 1)$ can be established by a reasoning similar to the one done for the PMM in [33, Sec. 5.1]. In particular, using in a straightforward way results established for probability intervals [36], $B^{\delta_\epsilon}_{d_{LV}}(P^{\epsilon}_0)$ is given by the probability measures P satisfying

$$\epsilon(1-\delta_1)P^1_0(\{x\}) + (1-\epsilon)(1-\delta_2)P^2_0(\{x\}) \le P(\{x\}) \;\; \forall x \in \mathcal{X}.$$

From this, we deduce that

$$\begin{aligned} 1 - \delta_\epsilon &= \sum_{x \in \mathcal{X}} \Big(\epsilon(1-\delta_1)P^1_0(\{x\}) + (1-\epsilon)(1-\delta_2)P^2_0(\{x\})\Big) \\ &= \epsilon(1-\delta_1)\sum_{x \in \mathcal{X}} P^1_0(\{x\}) + (1-\epsilon)(1-\delta_2)\sum_{x \in \mathcal{X}} P^2_0(\{x\}) \\ &= \epsilon(1-\delta_1) + (1-\epsilon)(1-\delta_2), \end{aligned}$$

and

$$P_0^{\epsilon}(\{x\}) = \frac{\epsilon(1-\delta_1)P_0^1(\{x\}) + (1-\epsilon)(1-\delta_2)P_0^2(\{x\})}{1-\delta_\epsilon} \quad \forall x \in \mathcal{X}.$$

4.2. **Multivariate setting.** Let us now look at the behaviour of the LV in a multivariate setting.

Marginalisation

We first show that the marginal model of a joint LV is again a LV model, with the same distortion factor δ applied to the marginal probability.

Proposition 4. *Consider the distortion model $B_{d_{LV}}^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{X}\times\mathcal{Y})$ and its induced lower probability $\underline{P}_{LV}$. Then, the marginal model $\underline{P}_{LV}^{\mathcal{X}}$ induces the credal set $B_{d_{LV}}^{\delta}\big(P_0^{\mathcal{X}}\big)$ with $P_0^{\mathcal{X}}$ the marginal probability of $P_0^{\mathcal{X},\mathcal{Y}}$ on $\mathcal{X}$.*

Proof. Using again that, from [30, Sec. 5.1] $B_{d_{LV}}^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big)$ is defined by lower bounds $(1-\delta)P_0^{\mathcal{X},\mathcal{Y}}\big(\{(x,y)\}\big)$ on singletons, it is sufficient to notice that the marginal model on $\mathcal{X}$ is described by the constraints

$$P(\{x\}) = \sum_{y\in\mathcal{Y}} P\big(\{(x,y)\}\big) \geq \sum_{y\in\mathcal{Y}}(1-\delta)P_0^{\mathcal{X},\mathcal{Y}}\big(\{(x,y)\}\big) = (1-\delta)P_0^{\mathcal{X}}(\{x\}).$$

As a consequence, it is a LV model. □

Independent products

Regarding the problem of going from marginal models $P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}$ to multivariate ones, we can first notice that on Cartesian products of events, we have

$$\begin{aligned}\underline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}}(A\times B) &= (1-\delta)P_0^{\mathcal{X}}(A)P_0^{\mathcal{Y}}(B)\\ &\geq (1-\delta)P_0^{\mathcal{X}}(A)(1-\delta)P_0^{\mathcal{Y}}(B) = \underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}(A\times B) \qquad (18)\end{aligned}$$

where last equality follows from the factorization property in Equation (2). Note that the inequality is strict for any $\delta > 0$. We may then wonder if $\underline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}} \geq \underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}$ in general. The next example shows that this is not the case, and hence that we have no dominance relation between the joint models $\underline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}}$ and $\underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}$.

Example 6. *Let us continue with Example 3. Given $E_1 = \{(x_2,y_2)\}^c$, we obtain*

$$\begin{aligned}\underline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}}(E_1) &= 1 - \overline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}}\big(\{(x_2,y_2)\}\big) = 0.585\\ &< \underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}(E_1) = 1 - \overline{P}_{LV}^{\mathcal{X}} \boxtimes \overline{P}_{LV}^{\mathcal{Y}}\big(\{(x_2,y_2)\}\big)\\ &= 1 - \overline{P}_{LV}^{\mathcal{X}}(\{x_2\})\overline{P}_{LV}^{\mathcal{Y}}(\{y_2\}) = 0.5985,\end{aligned}$$

and therefore it cannot be $B_{d_{LV}}^{\delta}\big(P_0^{\mathcal{X},\mathcal{Y}}\big) \subseteq \mathcal{M}\big(\underline{P}_{LV}^{\mathcal{X}}\big) \boxtimes \mathcal{M}\big(\underline{P}_{LV}^{\mathcal{Y}}\big)$.

On the other hand, taking the events $A = \{x_1\}$ and $B = \{y_1\}$, we obtain $\underline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}}(A\times B) = 0.135$ and $\underline{P}_{LV}^{\mathcal{X}}(A) \boxtimes \underline{P}_{LV}^{\mathcal{Y}}(B) = 0.1215$, showing that the inequality in Equation (18) may be strict. We therefore conclude that there is no dominance relationship between $\underline{P}_{LV}^{\mathcal{X}\times\mathcal{Y}}$ and $\underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}$. ♦

Our next example shows that the family of LV models is not closed under strong products.

Example 7. *Consider the same probability measures and distortion factor as in Example 3, and the events* $\{(x_1,y_1)\},\{(x_1,y_2)\}$ *and* $\{x_1\}\times\mathcal{Y}$*. Then*

$$\underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}(\{(x_1,y_1)\}) = \underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}(\{(x_1,y_2)\}) = 0.27\cdot 0.45 = 0.1215$$

and

$$\underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}(\{x_1\}\times\mathcal{Y}) = \underline{P}_{PMM}^{\mathcal{X}}(\{x_1\}) = 0.27.$$

The fact that $0.27 \neq 2\cdot 0.1215$ *contradicts the fact that a LV model should satisfy* $\underline{P}(A\cup B) = \underline{P}(A)+\underline{P}(B)$ *if* $A\cap B=\emptyset$ *and* $\min(\underline{P}(A),\underline{P}(B))>0$*; thus,* $\underline{P}_{LV}^{\mathcal{X}} \boxtimes \underline{P}_{LV}^{\mathcal{Y}}$ *is not a LV model.* ♦

Natural extension of marginal models

We now consider the lower and upper probabilities that are the lower and upper envelopes of $B_{d_{LV}}^{\delta_{\mathcal{X}}}(P_0^{\mathcal{X}}) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_{d_{LV}}^{\delta_{\mathcal{Y}}}(P_0^{\mathcal{Y}}) \subseteq \mathbb{P}^*(\mathcal{Y})$. Using Equations (5) and (6), we can give the form of $\underline{E}_{LV}$ and $\overline{E}_{LV}$ for the events $A\times B$, for $A\subset\mathcal{X}$, $B\subset\mathcal{Y}$:

$$\begin{aligned}
\underline{E}_{LV}(A\times B) &= \max\left\{\underline{P}_{LV}^{\mathcal{X}}(A) + \underline{P}_{LV}^{\mathcal{Y}}(B) - 1, 0\right\}\\
&= \max\left\{(1-\delta_{\mathcal{X}})P_0^{\mathcal{X}}(A) + (1-\delta_{\mathcal{Y}})P_0^{\mathcal{Y}}(B) - 1, 0\right\},\\
\overline{E}_{LV}(A\times B) &= \min\left\{\overline{P}_{LV}^{\mathcal{X}}(A), \overline{P}_{LV}^{\mathcal{Y}}(B)\right\}\\
&= \min\left\{(1-\delta_{\mathcal{X}})P_0^{\mathcal{X}}(A) + \delta_{\mathcal{X}}, (1-\delta_{\mathcal{Y}})P_0^{\mathcal{Y}}(B) + \delta_{\mathcal{Y}}\right\}.
\end{aligned}$$

When the distortion parameters coincide, $\delta_{\mathcal{X}} = \delta_{\mathcal{Y}} = \delta$, the previous expressions simplify to:

$$\begin{aligned}
\underline{E}_{LV}(A\times B) &= \max\left\{(1-\delta)P_0^{\mathcal{X}}(A) + (1-\delta)P_0^{\mathcal{Y}}(B) - 1, 0\right\}\\
&= \max\left\{(1-\delta)\big(P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B)\big) - 1, 0\right\}\\
&= \max\left\{(1-\delta)\Big(P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B) - \frac{1}{1-\delta}\Big), 0\right\}\\
&= (1-\delta)\max\left\{P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B) - \frac{1}{1-\delta}, 0\right\}. \qquad (19)\\
\overline{E}_{LV}(A\times B) &= \min\left\{(1-\delta)P_0^{\mathcal{X}}(A) + \delta, (1-\delta)P_0^{\mathcal{Y}}(B) + \delta\right\}\\
&= (1-\delta)\min\left\{P_0^{\mathcal{X}}(A), P_0^{\mathcal{Y}}(B)\right\} + \delta. \qquad (20)
\end{aligned}$$

The expressions in Equations (19) and (20) are somewhat similar to the lower and upper probabilities of a LV model. However, unlike what happened in the case of the PMM, the equality

$$\underline{E}_{LV} = (1-\delta)\underline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}$$

does not hold:

Example 8. *Consider the probability measures from Example 3, and let* $\delta = 0.2$*. Then Equation* (19) *gives*

$$\underline{E}_{LV}\big(\{(x_2,y_2)\}\big) = \max\{0.8\cdot 0.7 + 0.8\cdot 0.5 - 1, 0\} = 0,$$

while from Equation (7) *we obtain*

$$\underline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}\big(\{(x_2,y_2)\}\big) = \max\{0.7+0.5-1,0\} = 0.2,$$

meaning that $(1-\delta)\underline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}\big(\{(x_2,y_2)\}\big) = 0.16$. ♦

5. Constant odds ratio

Our next distortion model is the constant odds ratio model (COR, for short):

Definition 4. *Given a probability measure* P_0 *and a distortion factor* $\delta \in (0,1)$*, the associated* constant odds ratio model *is the coherent lower prevision* $\underline{P}_{COR}$ *that, on any gamble* f*, is defined as the unique solution to the implicit equation:*

$$(1-\delta)P_0\big((f-\underline{P}_{COR}(f))^+\big) = P_0\big((f-\underline{P}_{COR}(f))^-\big), \tag{21}$$

where $g^+ = \max\{g,0\}$ *and* $g^- = \max\{-g,0\}$.

While Equation (21) does not have a explicit expression, the restriction to (indicators of) events of the constant odds ratio can be more easily computed as:

$$\underline{P}_{COR}(A) = \frac{(1-\delta)P_0(A)}{1-\delta P_0(A)} \quad \forall A \subseteq \mathcal{X}. \tag{22}$$

When $P_0 \in \mathbb{P}^*(\mathcal{X})$ and δ is small enough, the credal set $\mathcal{M}(\underline{P}_{COR})$ coincides with $B^{\delta}_{d_{COR}}(P_0)$, where $d_{COR} : \mathbb{P}^*(\mathcal{X}) \times \mathbb{P}^*(\mathcal{X}) \to [0,\infty)$ is the distorting function given by [30, Thm. 14]:

$$d_{COR}(P,Q) = \max_{A,B\neq\emptyset}\left\{1-\frac{P(A)\cdot Q(B)}{P(B)\cdot Q(A)}\right\}.$$

Also, the credal set $\mathcal{M}(\underline{P}_{COR})$ can be expressed as [53, Sec. 3.3.5]:

$$\mathcal{M}(\underline{P}_{COR}) = \{P \in \mathbb{P}(\mathcal{X}) \mid P(A)P_0(B) \geq (1-\delta)P_0(A)P(B) \ \ \forall A,B \subseteq \mathcal{X}\}. \tag{23}$$

Also, the COR model is more informative than the PMM and the LV models, in the sense that once we fix P_0 and δ, it holds that $B^{\delta}_{d_{COR}}(P_0) \subseteq B^{\delta}_{d_{PMM}}(P_0) \cap B^{\delta}_{d_{LV}}(P_0)$ (see [53, Sec. 4.6.5] for more comments in this direction).

5.1. Merging.

Let us first look at the behaviour of the family of COR models under merging.

Conjunction

Unlike the PMM and LV models, the intersection of two constant odds ratio models is not a COR model in general, as next example shows.

Example 9. *Consider* $\mathcal{M}_1 = B^{\delta_1}_{d_{COR}}(P_0^1) \subseteq \mathbb{P}^*(\mathcal{X})$ *with* $P_0^1 = (0.5,0.3,0.2)$ *and* $\delta_1 = 0.2$*, and* $\mathcal{M}_2 = B^{\delta_2}_{d_{COR}}(P_0^2) \subseteq \mathbb{P}^*(\mathcal{X})$ *such that* $P_0^2 = (0.35,0.3,0.35)$ *with* $\delta_2 = 0.5$*. From Equation* (23)*, the ratio* $P(\{x_1\})/P(\{x_3\})$ *is constrained by the inequalities*

$$3.125 \geq \frac{P(\{x_1\})}{P(\{x_3\})} \geq 2, \qquad 2 \geq \frac{P(\{x_1\})}{P(\{x_3\})} \geq 0.5,$$

respectively for $\mathcal{M}_1$ *and* $\mathcal{M}_2$*. From this, we can deduce that any* $P \in \mathcal{M}_1 \cap \mathcal{M}_2$ *must satisfy the constraint* $\frac{P(\{x_1\})}{P(\{x_3\})} = 2$*. As a consequence, the credal set* $\mathcal{M}_1 \cap \mathcal{M}_2$ *has at most two extreme points:*

$$\left(\frac{2(1-\underline{P}(\{x_2\}))}{3}, \underline{P}(\{x_2\}), \frac{1-\underline{P}(\{x_2\})}{3}\right)$$

and

$$\left(\frac{2(1-\overline{P}(\{x_2\}))}{3}, \overline{P}(\{x_2\}), \frac{1-\overline{P}(\{x_2\})}{3}\right),$$

where $\underline{P}, \overline{P}$ *denote the lower and upper probabilities associated with* $\mathcal{M}_1 \cap \mathcal{M}_2$. *Since it was proven in* [30, Prop. 13] *that the number of extreme points of the credal set of a COR model is equal to* $2^n - 2$, *where* n *is the cardinality of* $\mathcal{X}$, *this implies that the conjunction* $\mathcal{M}_1 \cap \mathcal{M}_2$ *does not determine a constant odds-ratio model, as it contains less than* $2^n - 2 = 6$ *extreme points.* ♦

For the PMM and LV models, we had easy ways to check whether a conjunction was empty (Equations (9) and (17), respectively). This resulted from the fact that their conjunctions are specific probability intervals, which are models for which checking non-emptiness is easy. This is not the case for the COR model, that is not closed under conjunction. A possibility is to use the constraints induced by the models $\mathcal{M}_1 = B^{\delta_1}_{d_{COR}}(P^1_0)$ and $\mathcal{M}_2 = B^{\delta_2}_{d_{COR}}(P^2_0)$, to check that taken together they still have a solution (i.e., that there is at least a probability P within $\mathcal{M}_1 \cap \mathcal{M}_2$). However, as those are defined implicitly by Equation (21), they would have to be made explicit, for instance by enumerating the extreme points induced by $\mathcal{M}_1$, $\mathcal{M}_2$ (see for example [30, Prop. 12]) and extracting the corresponding constraints.

Disjunction

Similarly, the disjunction of two COR models will not produce a COR model in general, not even when $\delta_1 = \delta_2$:

Example 10. *Consider* $P^1_0 = (0.4, 0.3, 0.3), P^2_0 = (0.3, 0.4, 0.3)$ *and* $\delta_1 = \delta_2 = 0.1$. *Using Equation* (22)*, the associated COR models* $\underline{P}_{COR_1}, \underline{P}_{COR_2}$ *and their disjunction* $\underline{P}^{\cup} = \min\{\underline{P}_{COR_1}, \underline{P}_{COR_2}\}$ *are given in the following table:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{COR_1}$	$3/8$	$27/97$	$27/97$	$21/31$	$21/31$	$27/47$
$\underline{P}_{COR_2}$	$27/97$	$3/8$	$27/97$	$21/31$	$27/47$	$21/31$
$\underline{P}^{\cup}$	$27/97$	$27/97$	$27/97$	$21/31$	$27/47$	$27/47$

If $\underline{P}^{\cup}$ *was a COR model, i.e. if* $\mathcal{M}(\underline{P}^{\cup}) = B^{\delta}_{d_{COR}}(P_0)$ *for some* P_0 *and* δ*, since* $\underline{P}^{\cup}(\{x_1\}) = \underline{P}^{\cup}(\{x_2\}) = \underline{P}^{\cup}(\{x_3\})$*, it should be*

$$P_0(\{x_1\}) = P_0(\{x_2\}) = P_0(\{x_3\}) = \frac{1}{3}.$$

But in that case, regardless of the value of δ, $\underline{P}^{\cup}$ *must take the same value for all the events of cardinality two, a contradiction.*

This example also allows us to show that $\underline{P}^{\cup}$ *does not have a unique undominated outer approximation in terms of COR models. Consider* $P_A = (31/80, 31/80, 18/80)$, $P_B = (35/124, 35/124, 27/62)$, $\delta_A = 121/310$, $\delta_B = \frac{1}{2}$ *and the COR models* $B^{\delta_A}_{d_{COR}}(P_A)$ *and* $B^{\delta_B}_{d_{COR}}(P_B)$ *they induce. These produce the following coherent lower probabilities:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{COR_A}$	$27/97$	$27/97$	$1701/11311$	$21/31$	$9261/18871$	$9261/18871$
$\underline{P}_{COR_B}$	$35/213$	$35/213$	$27/97$	$35/89$	$89/159$	$89/159$

Both $\underline{P}_{COR_A}$ and $\underline{P}_{COR_B}$ are outer approximations of $\underline{P}^{\cup}$. Moreover, if there was a unique undominated outer approximation $\underline{Q}_{COR}$ of $\underline{P}^{\cup}$ in terms of COR models, then it should be $\underline{P}_{COR_A}, \underline{P}_{COR_B} \leq \underline{Q}_{COR}$, that implies

$$\underline{Q}_{COR}(\{x_i\}) = \underline{P}^{\cup}(\{x_i\}) = \frac{27}{97} \quad \text{for } i = 1, 2, 3,$$

meaning that $\underline{Q}_{COR}$ is defined through $P_0 = (1/3, 1/3, 1/3)$ and $\delta = 8/35$. However, on the event $\{x_1, x_3\}$ this distortion model satisfies

$$\underline{Q}_{COR}\big(\{x_1, x_3\}\big) = \frac{(1-\delta)P_0(\{x_1, x_3\})}{1 - \delta P_0(\{x_1, x_3\})} = \frac{54}{89} > \frac{27}{47} = \underline{P}^{\cup}\big(\{x_1, x_3\}\big),$$

so $\underline{Q}_{COR}$ is not an outer approximation of $\underline{P}^{\cup}$. ♦

We therefore conclude that the COR model is neither preserved by conjunction nor by disjunction, and also that its disjunction has not a unique undominated outer approximation.

Convex mixture

As for the previous models, given the fact that two COR models $B^{\delta_1}_{d_{COR}}(P^1_0) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta_2}_{d_{COR}}(P^2_0) \subseteq \mathbb{P}^*(\mathcal{X})$ are described by the same set of constraints over $P(A)/P(B)$, their convex mixture is a credal set described by the constraints

$$\frac{P(A)}{P(B)} \geq \epsilon(1-\delta_1)\frac{P_1(A)}{P_1(B)} + (1-\epsilon)(1-\delta_2)\frac{P_2(A)}{P_2(B)}.$$

However, next example shows that such constraints will not lead, in general, to a COR model.

Example 11. *Consider $P^1_0 = (1/4, 1/4, 1/2), P^2_0 = (1/2, 1/4, 1/4)$ and $\delta_1 = \delta_2 = 0.5$. Using Equation (22), the associated COR models $\underline{P}_{COR_1}, \underline{P}_{COR_2}$ and their average $\underline{P}^{0.5}$ obtained for $\epsilon = 0.5$ are given in the following table:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{COR_1}$	$1/7$	$1/7$	$1/3$	$1/3$	$3/5$	$3/5$
$\underline{P}_{COR_2}$	$1/3$	$1/7$	$1/7$	$3/5$	$3/5$	$1/3$
$\underline{P}^{0.5}$	$5/21$	$1/7$	$5/21$	$7/15$	$3/5$	$7/15$

Should $\underline{P}^{\varepsilon}$ be the lower probability of a COR model $B^{\delta_\varepsilon}_{d_{COR}}(P^{\varepsilon}_0)$, we should have $P^{\varepsilon}_0(\{x_1\}) = P^{\varepsilon}_0(\{x_3\}) = p$, hence $P^{\varepsilon}_0(\{x_2\}) = 1 - 2p$. Using this observation and Equation (22) on events $\{x_1\}$ and $\{x_1, x_3\}$, we should have $\delta_\varepsilon = 13/28$ and $p = 7/19$, and applying again Equation (22) with these values on $\{x_2\}$ would give $\underline{P}(\{x_2\}) = 75/467$ for COR model, which is close but not equal to the $1/7$ reported in the table. ♦

The COR model is therefore also not preserved under convex mixture, making it a not very convenient model when having to merge distortion models.

5.2. **Multivariate setting.**

Marginalisation

As for the PMM and LV models, we can show that the marginal distribution of a joint constant odds ratio model is also a constant odds ratio model.

Proposition 5. *Consider the distortion model $B^{\delta}_{d_{COR}}(P_0^{\mathcal{X},\mathcal{Y}}) \subseteq \mathbb{P}^*(\mathcal{X}\times\mathcal{Y})$ and its induced lower prevision $\underline{P}_{COR}$. Then, the marginal model $\underline{P}^{\mathcal{X}}_{COR}$ induces the credal set $B^{\delta}_{d_{COR}}(P_0^{\mathcal{X}}) \subseteq \mathbb{P}^*(\mathcal{X})$ with $P_0^{\mathcal{X}}$ the marginal probability of $P_0^{\mathcal{X},\mathcal{Y}}$ on $\mathcal{X}$.*

Proof. By definition, $\underline{P}_{COR}(f)$ is determined as the unique solution of the equation:

$$(1-\delta)P_0\Big(\big(f-\underline{P}_{COR}(f)\big)^+\Big) = P_0\Big(\big(f-\underline{P}_{COR}(f)\big)^-\Big).$$

Consider a gamble f that only depends on the values in $\mathcal{X}$, in the sense that $f(x,y)=f(x,y')$ for every $y\neq y'\in\mathcal{Y}$ and every $x\in\mathcal{X}$, and let us define the gamble $f_{\mathcal{X}}:\mathcal{X}\to\mathbb{R}$ by $f_{\mathcal{X}}(x)=f(x,y)$. Then $\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})=\underline{P}_{COR}(f)$. Moreover,

$$(1-\delta)P_0\Big(\big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\big)^+\Big) = (1-\delta)P_0^{\mathcal{X}}\Big(\big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\big)^+\Big),$$

and also:

$$P_0\Big(\big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\big)^-\Big) = P_0^{\mathcal{X}}\Big(\big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\big)^-\Big),$$

because $\Big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\Big)^+$ only depends on the values of $x\in\mathcal{X}$. This means that $\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})$ is the unique solution of the equation:

$$(1-\delta)P_0^{\mathcal{X}}\Big(\big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\big)^+\Big) = P_0^{\mathcal{X}}\Big(\big(f_{\mathcal{X}}-\underline{P}^{\mathcal{X}}_{COR}(f_{\mathcal{X}})\big)^-\Big),$$

meaning that $\underline{P}^{\mathcal{X}}_{COR}$ is a constant odds ratio with respect to the parameter δ and probability $P_0^{\mathcal{X}}$ (the restriction of P_0 to the first component). □

Independent products

Consider now the marginal models $B^{\delta}_{d_{COR}}(P_0^{\mathcal{X}}) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta}_{d_{COR}}(P_0^{\mathcal{Y}}) \subseteq \mathbb{P}^*(\mathcal{Y})$. Regarding the problem of going from marginal models $P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}$ to joint ones, we can first notice that on Cartesian products of events, we have

$$\begin{aligned}\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{COR}(A\times B) &= \frac{(1-\delta)P_0^{\mathcal{X}}(A)P_0^{\mathcal{Y}}(B)}{1-\delta P_0^{\mathcal{X}}(A)P_0^{\mathcal{Y}}(B)}\\ &\geq \frac{(1-\delta)P_0^{\mathcal{X}}(A)(1-\delta)P_0^{\mathcal{Y}}(B)}{(1-\delta P_0^{\mathcal{X}}(A))(1-\delta P_0^{\mathcal{Y}}(B))} = \underline{P}^{\mathcal{X}}_{COR}\boxtimes\underline{P}^{\mathcal{Y}}_{COR}(A\times B) \qquad (24)\end{aligned}$$

where the last equality follows from the factorisation property in Equation (2). We can then wonder if $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{COR}(C) \geq \underline{P}^{\mathcal{X}}_{COR}\boxtimes\underline{P}^{\mathcal{Y}}_{COR}(C)$ for any event $C\subseteq\mathcal{X}\times\mathcal{Y}$. The next example shows that this is not the case, and that the inequality in Equation (24) may be strict; therefore it shows that there is no dominance relation between $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{COR}$ and $\underline{P}^{\mathcal{X}}_{COR}\boxtimes\underline{P}^{\mathcal{Y}}_{COR}$.

Example 12. *Consider our running Example 3. Given $E_2=\{(x_1,y_1),(x_2,y_2)\}$, we obtain*

$$\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{COR}(E_2) = 0.4737 < \underline{P}^{\mathcal{X}}_{COR}\boxtimes\underline{P}^{\mathcal{Y}}_{COR}(E_2) = 0.4883;$$

to see the last equality, note that

$$\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2) = \min_{P_{\mathcal{X}} \in \mathcal{M}(\underline{P}^{\mathcal{X}}_{COR}), P_{\mathcal{Y}} \in \mathcal{M}(\underline{P}^{\mathcal{Y}}_{COR})} \big(P_{\mathcal{X}}(\{x_1\})P_{\mathcal{Y}}(\{y_1\}) + P_{\mathcal{X}}(\{x_2\})P_{\mathcal{Y}}(\{y_2\})\big),$$

that $\mathcal{M}(\underline{P}^{\mathcal{X}}_{COR})$ *consists of the probabilities* $P_{\mathcal{X}}$ *for which* $P_X(\{x_1\}) \in [\frac{27}{97}, \frac{10}{31}]$ *and, similarly, that* $\mathcal{M}(\underline{P}^{\mathcal{Y}}_{COR})$ *consists of the probabilities* $P_{\mathcal{Y}}$ *for which* $P_Y(\{y_1\}) \in [\frac{9}{19}, \frac{10}{19}]$*; the minimum in the equation above is obtained for* $P_X(\{x_1\}) = \frac{27}{97}$ *and* $P_Y(\{y_1\}) = \frac{10}{19}$*, yielding*

$$\frac{27}{97} \cdot \frac{10}{19} + \frac{70}{97} \cdot \frac{9}{19} = \frac{900}{1843} = 0.4883.$$

Therefore, $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{COR}$ *does not dominate* $\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}$ *on all events.*

Finally, taking the events $A = \{x_1\}$ *and* $B = \{y_1\}$*, it holds that* $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{COR}(A \times B) \approx 0.137$*, while* $\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(A \times B) \approx 0.132$*, showing that the inequality in Equation* (24) *may be strict.* ♦

On the other hand, the family of COR models is not closed under strong products.

Example 13. *Consider again the running Example 3. Let* $E_2 = \{(x_1, y_1), (x_2, y_2)\}$ *and* $E_3 = \{x_1\} \times \mathcal{Y}$*. Then*

$$\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2) = 0.4883 \text{ and } \underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_3) = \underline{P}^{\mathcal{X}}_{COR}(\{x_1\}) = \frac{27}{97},$$

where the first equality follows from Example 12. On the other hand,

$$\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2 \cup E_3) = 1 - \overline{P}^{\mathcal{X}}_{COR} \boxtimes \overline{P}^{\mathcal{Y}}_{COR}\big(\{(x_2, y_1)\}\big) = 1 - \frac{70}{97} \cdot \frac{10}{19} = 0.6202$$
$$\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2 \cap E_3) = \underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}\big(\{(x_1, y_1)\}\big) = \frac{27}{97} \cdot \frac{9}{19} = 0.1319,$$

whence

$$\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2 \cup E_3) + \underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2 \cap E_3) = 0.7520 < \underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_2) + \underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}(E_3) = 0.7667$$

contradicting the fact that $\underline{P}^{\mathcal{X}}_{COR} \boxtimes \underline{P}^{\mathcal{Y}}_{COR}$ *should be 2-monotone on events if it was a COR model.* ♦

Natural extension of marginal models

We have already mentioned that there is not an explicit equation for the lower prevision of the COR model in gambles (see Equation (21)), and it can only be given for events (see Equation (22)). This hinders a bit the computation of the natural extension of this model. In addition, even if we consider only the values in events, this model is more difficult to handle than the PMM or the LV.

First of all, using Equations (5) and (6), we give the explicit form of the lower and upper natural extension on the Cartesian products $A \times B$, for $A \subseteq \mathcal{X}$, $B \subseteq \mathcal{Y}$:

$$\overline{E}_{COR}(A \times B) = \min\left\{\frac{(1-\delta_{\mathcal{X}})P_0^{\mathcal{X}}(A)}{1-\delta_{\mathcal{X}}P_0^{\mathcal{X}}(A)}, \frac{(1-\delta_{\mathcal{Y}})P_0^{\mathcal{Y}}(B)}{1-\delta_{\mathcal{Y}}P_0^{\mathcal{Y}}(B)}\right\}.$$

$$\underline{E}_{COR}(A \times B) = \max\left\{\frac{(1-\delta_{\mathcal{X}})P_0^{\mathcal{X}}(A)}{1-\delta_{\mathcal{X}}P_0^{\mathcal{X}}(A)} + \frac{(1-\delta_{\mathcal{Y}})P_0^{\mathcal{Y}}(B)}{1-\delta_{\mathcal{Y}}P_0^{\mathcal{Y}}(B)} - 1, 0\right\}.$$

When the distortion parameters coincide, $\delta_{\mathcal{X}} = \delta_{\mathcal{Y}} = \delta$, these two equations become:

$$\begin{aligned}\overline{E}_{COR}(A \times B) &= \min\left\{\frac{(1-\delta)P_0^{\mathcal{X}}(A)}{1-\delta P_0^{\mathcal{X}}(A)}, \frac{(1-\delta)P_0^{\mathcal{Y}}(B)}{1-\delta P_0^{\mathcal{Y}}(B)}\right\} \\ &= (1-\delta)\min\left\{\frac{P_0^{\mathcal{X}}(A)}{1-\delta P_0^{\mathcal{X}}(A)}, \frac{P_0^{\mathcal{Y}}(B)}{1-\delta P_0^{\mathcal{Y}}(B)}\right\}. \\ \underline{E}_{COR}(A \times B) &= \max\left\{\frac{(1-\delta)P_0^{\mathcal{X}}(A)}{1-\delta P_0^{\mathcal{X}}(A)} + \frac{(1-\delta)P_0^{\mathcal{Y}}(B)}{1-\delta P_0^{\mathcal{Y}}(B)} - 1, 0\right\}\end{aligned}$$

Although these expressions do not seem to resemble a COR model, we may wonder if, similarly to what happened with the PMM (see Theorem 2), the equality $\mathcal{E}\big(\underline{P}_{COR}^{\mathcal{X}}, \underline{P}_{COR}^{\mathcal{Y}}\big) = B_{d_{COR}}\big(\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}\big)$ holds. As we see next, this is not the case.

Example 14. *Consider our running Example 3. Given $E_4 = \{(x_2, y_2)\}$, we obtain*

$$\underline{E}_{COR}(E_4) = \max\{0.6774 + 0.4737 - 1, 0\} = 0.1511.$$

On the other hand, from Equation (7) we have that $\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(E_4) = 0.2$, whence

$$\frac{(1-\delta)\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(E_4)}{1-\delta\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}(E_4)} = 0.1837.$$

Thus, the two values do not coincide. ♦

6. Total variation model

The last three distortion models we shall analyse in this paper are defined directly from some distance between probability measures. The first of them is the total variation model (TV, for short): given two probability measures P, Q, their total variation distance is

$$d_{TV}(P, Q) = \max_{A \subseteq \mathcal{X}} |P(A) - Q(A)|.$$

By taking the lower and upper envelopes of the neighbourhood model it produces, we obtain the following:

Definition 5. *Given a probability measure P_0 and a distortion factor $\delta > 0$, the* total variation model *is given by the following lower and upper probabilities:*

$$\underline{P}_{TV}(A) = \max\{P_0(A) - \delta, 0\}, \quad \overline{P}_{TV}(A) = \min\{P_0(A) + \delta, 1\} \quad \forall A \subseteq \mathcal{X}. \tag{25}$$

Since we are assuming that $P_0 \in \mathbb{P}^*(\mathcal{X})$ and that δ is small enough so that $B_{d_{TV}}^{\delta}(P_0) \subseteq \mathbb{P}^*(\mathcal{X})$, Equation (25) simplifies to:

$$\underline{P}_{TV}(A) = P_0(A) - \delta, \quad \overline{P}_{TV}(A) = P_0(A) + \delta \quad \forall A \neq \emptyset, \mathcal{X}. \tag{26}$$

6.1. **Merging.** Let us now consider the problem of merging two TV models.

Conjunction and disjunction

Our next example shows that neither the conjunction nor the disjunction of two such models will produce in general a total variation model.

Example 15. *From Equation* (26)*, we know that a TV model satisfies, for any event A such that $\overline{P}_{TV}(A), \underline{P}_{TV}(A) \in (0,1)$, the following equality:*

$$\overline{P}_{TV}(A) - \underline{P}_{TV}(A) = (P_0(A) + \delta) - (P_0(A) - \delta) = 2\delta.$$

In particular, since we are assuming that $B^{\delta}_{d_{TV}}(P_0) \subseteq \mathbb{P}^*(P_0)$, this equality holds for any $A \neq \emptyset, \mathcal{X}$. Let us use this to derive that the family of TV models is not closed under conjunction or disjunction.

Let $B^{\delta_1}_{d_{TV}}(P_0^1)$ be induced by $P_0^1 = (0.41, 0.37, 0.22)$ and $\delta_1 = 0.12$, and $B^{\delta_2}_{d_{TV}}(P_0^2)$ be determined by $P_0^2 = (0.37, 0.41, 0.22)$ and $\delta_2 = 0.12$. The lower probabilities $\underline{P}_{TV_1}$ and $\underline{P}_{TV_2}$, their conjunction $\underline{P}^{\cap}$ and disjunction $\underline{P}^{\cup}$, are given by:

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{TV_1}$	0.29	0.25	0.1	0.66	0.51	0.47
$\underline{P}_{TV_2}$	0.25	0.29	0.1	0.66	0.47	0.51
$\underline{P}^{\cap}$	0.29	0.29	0.1	0.66	0.51	0.51
$\underline{P}^{\cup}$	0.25	0.25	0.1	0.66	0.47	0.47

For the third line, use that $\underline{P}^{\cap}$ is the natural extension of $\max\{\underline{P}_{TV_1}, \underline{P}_{TV_2}\}$, i.e., the smallest coherent lower probability that dominates $\max\{\underline{P}_{TV_1}, \underline{P}_{TV_2}\}$; but the latter is coherent since it is the lower envelope of

$$(0.29, 0.37, 0.34), (0.37, 0.29, 0.34), (0.41, 0.49, 0.1), (0.49, 0.41, 0.1),$$

and therefore it coincides with $\underline{P}^{\cap}$. We observe that

$$\overline{P}^{\cap}(\{x_1\}) - \underline{P}^{\cap}(\{x_1\}) = 0.2 \neq 0.24 = \overline{P}^{\cap}(\{x_3\}) - \underline{P}^{\cap}(\{x_3\}),$$

$$\overline{P}^{\cup}(\{x_1\}) - \underline{P}^{\cup}(\{x_1\}) = 0.28 \neq 0.24 = \overline{P}^{\cup}(\{x_3\}) - \underline{P}^{\cup}(\{x_3\}),$$

concluding that neither $\underline{P}^{\cap}$ nor $\underline{P}^{\cup}$ are TV models. ♦

As we know that a TV model is described by a 2-monotone lower probability [31, Prop. 4], a simple way to check whether the intersection of two TV models is non-empty is simply to take the constraints (26) induced by both models $B^{\delta_1}_{d_{TV}}(P_0^1)$ and $B^{\delta_2}_{d_{TV}}(P_0^2)$, and to check whether they have a solution. This can be achieved through standard linear programming, with the caveat that the number of constraints will increase exponentially with n.

The same example allows us to show that the disjunction does not have a unique undominated outer approximation:

Example 16. *Consider the model $\underline{P}^{\cup}$ from the previous example, and let us consider the TV models $B^{\delta_A}_{d_{TV}}(P_0^A)$ and $B^{\delta_B}_{d_{TV}}(P_0^B)$, where $P_A = (0.31, 0.31, 0.38)$, $P_B = (0.41, 0.41, 0.18)$, $\delta_A = 0.28$ and $\delta_B = 0.16$. The lower probabilities $\underline{P}_{TV_A}, \underline{P}_{TV_B}$ are given by:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{TV_A}$	0.03	0.03	0.1	0.34	0.41	0.41
$\underline{P}_{TV_B}$	0.25	0.25	0.02	0.66	0.43	0.43

Both $\underline{P}_{TV_A}, \underline{P}_{TV_B}$ are outer approximations of $\underline{P}^{\cup}_{TV}$ in the TV family. If there was a unique undominated outer approximation of $\underline{P}^{\cup}$, denoted by $B^{\delta}_{d_{TV}}(P_0)$ and with associated lower probability $\underline{Q}_{TV}$, then $\underline{P}_{TV_A}, \underline{P}_{TV_B} \leq \underline{Q}_{TV} \leq \underline{P}^{\cup}$, whence $\underline{Q}_{TV}(\{x_1\}) = \underline{P}^{\cup}(\{x_1\}) = 0.25$, $\underline{Q}_{TV}(\{x_2\}) = \underline{P}^{\cup}(\{x_2\}) = 0.25$, $\underline{Q}_{TV}(\{x_3\}) = \underline{P}^{\cup}(\{x_3\}) = 0.1$. Therefore,

$$\underline{Q}_{TV}(\{x_1\}) + \underline{Q}_{TV}(\{x_2\}) + \underline{Q}_{TV}(\{x_3\}) = 1 - 3\delta = 0.6,$$

whence $\delta = {}^{0.4}/_{3}$ and $P_0 = (0.25 + \delta, 0.25 + \delta, 0.1 + \delta)$. However, this means that:

$$\underline{Q}_{TV}\big(\{x_1, x_3\}\big) = P_0\big(\{x_1, x_3\}\big) - \delta = 0.35 + \delta > 0.47 = \underline{P}^{\cup}\big(\{x_1, x_3\}\big),$$

a contradiction. ♦

Convex mixture

It is rather direct to check that the convex mixture of two TV models $B^{\delta_1}_{d_{TV}}\big(P^1_0\big)$ and $B^{\delta_2}_{d_{TV}}\big(P^2_0\big)$ is again a TV model as we have

$$\epsilon \underline{P}^1_{TV}(A) + (1-\epsilon)\underline{P}^2_{TV}(A) = \epsilon P^1_0(A) + (1-\epsilon)P^2_0(A) - \epsilon\delta_1 - (1-\epsilon)\delta_2$$

which are lower probabilities induced by the TV model $B^{\delta_\epsilon}_{d_{TV}}\big(P^{\epsilon}_0\big)$ with

$$\delta_\epsilon = \epsilon\delta_1 + (1-\epsilon)\delta_2 \quad \text{and} \quad P^{\epsilon}_0(\{x\}) = \epsilon P^1_0(\{x\}) + (1-\epsilon)P^2_0(\{x\}) \quad \forall x \in \mathcal{X}.$$

6.2. **Multivariate setting.**

Marginalisation

Let us now look at the behaviour of the TV model in a multivariate setting. We can first show that the marginal model of a joint TV model is again a TV model, with the same distortion factor δ applied to the marginal probability.

Proposition 6. *Consider the distortion model $B^{\delta}_{d_{TV}}\big(P^{\mathcal{X},\mathcal{Y}}_0\big) \subseteq \mathbb{P}^*(\mathcal{X} \times \mathcal{Y})$ and its associated lower probability $\underline{P}_{TV}$. Then, the marginal model $\underline{P}^{\mathcal{X}}_{TV}$ induces the credal set $B^{\delta}_{d_{TV}}\big(P^{\mathcal{X}}_0\big) \subseteq \mathbb{P}^*(\mathcal{X})$ with $P^{\mathcal{X}}_0$ the marginal probability of $P^{\mathcal{X},\mathcal{Y}}_0$ on $\mathcal{X}$.*

Proof. From Equation (25), we obtain that for every non-empty $B \subset \mathcal{X}$

$$\underline{P}^{\mathcal{X},\mathcal{Y}}(B \times \mathcal{Y}) = P^{\mathcal{X},\mathcal{Y}}_0(B \times \mathcal{Y}) - \delta = P^{\mathcal{X}}_0(B) - \delta,$$

the last term being the lower probability induced by $B^{\delta}_{d_{TV}}(P^{\mathcal{X}}_0)$. □

Independent products

Consider now two marginal models $B^{\delta}_{d_{TV}}(P^{\mathcal{X}}_0) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta}_{d_{TV}}(P^{\mathcal{Y}}_0) \subseteq \mathbb{P}^*(\mathcal{Y})$. Regarding the problem of going from marginal models $P^{\mathcal{X}}_0, P^{\mathcal{Y}}_0$ to joint ones, we can first notice that on Cartesian products of events, we have

$$\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{TV}(A \times B) = P^{\mathcal{X}}_0(A)P^{\mathcal{Y}}_0(B) - \delta, \text{ while}$$
$$\underline{P}^{\mathcal{X}}_{TV} \boxtimes \underline{P}^{\mathcal{Y}}_{TV}(A \times B) = \big(P^{\mathcal{X}}_0(A) - \delta\big)\big(P^{\mathcal{Y}}_0(B) - \delta\big)$$

where the last equality follows from the factorization property in Equation (2). Clearly, the equality $\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{TV}(A \times B) = \underline{P}^{\mathcal{X}}_{TV} \boxtimes \underline{P}^{\mathcal{Y}}_{TV}(A \times B)$ does not generally hold. We show in our next example that the inequality can go both ways (either $\underline{P}_{TV}^{\mathcal{X}\times\mathcal{Y}}(A\times B) < \underline{P}_{TV}^{\mathcal{X}} \boxtimes \underline{P}_{TV}^{\mathcal{Y}}(A\times B)$ or the reverse).

Example 17. *Take the distortion models $B^{\delta}_{d_{TV}}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta}_{d_{TV}}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$ induced by the probability measures $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ and the distortion factor $\delta = 0.1$. On the one hand, assume that there are some events $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$ such that $P_0^{\mathcal{X}}(A) = P_0^{\mathcal{Y}}(B) = 0.5$. We obtain that:*

$$\underline{P}_{TV}^{\mathcal{X}\times\mathcal{Y}}(A\times B) = 0.15 < 0.16 = \underline{P}_{TV}^{\mathcal{X}} \boxtimes \underline{P}_{TV}^{\mathcal{Y}}(A\times B).$$

On the other hand, assume that there are $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$ satisfying $P_0^{\mathcal{X}}(A) = P_0^{\mathcal{Y}}(B) = 0.6$. We obtain that:

$$\underline{P}_{TV}^{\mathcal{X}\times\mathcal{Y}}(A\times B) = 0.26 > 0.25 = \underline{P}_{TV}^{\mathcal{X}} \boxtimes \underline{P}_{TV}^{\mathcal{Y}}(A\times B).$$

Therefore, there is no dominance relationship between the lower probabilities obtained with the two approaches. ♦

Let us now show through an example that the TV model is not closed under strong products.

Example 18. *Consider again the setting of our running Example 3. Given the events $\{(x_1, y_1)\}$ and $\{x_1\} \times \mathcal{Y}$,*

$$\begin{aligned}
\underline{P}_{TV}^{\mathcal{X}} \boxtimes \underline{P}_{TV}^{\mathcal{Y}}\big(\{(x_1,y_1)\}\big) &= 0.2 \cdot 0.4 = 0.08,\\
\overline{P}_{TV}^{\mathcal{X}} \boxtimes \overline{P}_{TV}^{\mathcal{Y}}\big(\{(x_1,y_1)\}\big) &= 0.4 \cdot 0.6 = 0.24,\\
\underline{P}_{TV}^{\mathcal{X}} \boxtimes \underline{P}_{TV}^{\mathcal{Y}}\big(\{x_1\} \times \mathcal{Y}\big) &= \underline{P}_{TV}^{\mathcal{X}}(\{x_1\}) = 0.2,\\
\overline{P}_{TV}^{\mathcal{X}} \boxtimes \overline{P}_{TV}^{\mathcal{Y}}\big(\{x_1\} \times \mathcal{Y}\big) &= \overline{P}_{TV}^{\mathcal{X}}(\{x_1\}) = 0.4.
\end{aligned}$$

Since the differences between the upper and lower probabilities of the two events do not coincide (0.16 *and* 0.2*, respectively) and they are all strictly positive, we deduce that $\underline{P}_{TV}^{\mathcal{X}} \boxtimes \underline{P}_{TV}^{\mathcal{Y}}$ is not a TV model.* ♦

Natural extension of marginal models

Consider now two TV models $B^{\delta_{\mathcal{X}}}_{d_{TV}}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta_{\mathcal{Y}}}_{d_{TV}}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$. Using Equations (5) and (6), we can give the form of the natural extension in the events $A \times B$ for $A \subset \mathcal{X}$ and $B \subset \mathcal{Y}$:

$$\begin{aligned}
\underline{E}_{TV}(A\times B) &= \max\big\{\underline{P}_{TV}^{\mathcal{X}}(A) + \underline{P}_{TV}^{\mathcal{Y}}(B) - 1, 0\big\}\\
&= \max\big\{P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B) - 1 - (\delta_{\mathcal{X}} + \delta_{\mathcal{Y}}), 0\big\}.\\
\overline{E}_{TV}(A\times B) &= \min\Big\{\overline{P}_{TV}^{\mathcal{X}}(A), \overline{P}_{TV}^{\mathcal{Y}}(B)\Big\} = \min\big\{P_0^{\mathcal{X}}(A) + \delta_{\mathcal{X}}, P_0^{\mathcal{Y}}(B) + \delta_{\mathcal{Y}}\big\}.
\end{aligned}$$

When the distortion parameters coincide, $\delta_{\mathcal{X}} = \delta_{\mathcal{Y}} = \delta$, these expressions simplify to:

$$\begin{aligned}
\underline{E}_{TV}(A\times B) &= \max\big\{P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B) - 1 - 2\delta, 0\big\}.\\
&= \max\big\{P_0^{\mathcal{X}}(A) + P_0^{\mathcal{Y}}(B) - 1, 2\delta\big\} - 2\delta, \qquad (27)\\
\overline{E}_{TV}(A\times B) &= \min\big\{P_0^{\mathcal{X}}(A) + \delta, P_0^{\mathcal{Y}}(B) + \delta\big\} = \min\big\{P_0^{\mathcal{X}}(A), P_0^{\mathcal{Y}}(B)\big\} + \delta. \qquad (28)
\end{aligned}$$

The lower and upper natural extension have a similar form as a TV model. However, the lower bound of the natural extension in Equation (27) has a distortion parameter of 2δ, while the upper bound of the natural extension in Equation (28) has a distortion parameter of δ. This fact suggests that, as happened with the COR model, the natural extension of the TV model can neither be expressed as $\mathcal{E}(\underline{P}_{TV}^{\mathcal{X}}, \underline{P}_{TV}^{\mathcal{Y}}) = B_{d_{TV}}^{\delta}(\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}})$ nor as $\mathcal{E}(\underline{P}_{TV}^{\mathcal{X}}, \underline{P}_{TV}^{\mathcal{Y}}) = B_{d_{TV}}^{2\delta}(\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}})$. This is illustrated in our next example.

Example 19. *Consider the spaces $\mathcal{X}$ and $\mathcal{Y}$ and the probabilities $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ from Example 3. Given the event $E_4 = \{(x_2, y_2)\}$, we deduce from Equations (27) and (7) that*

$$\begin{aligned}\underline{E}_{TV}\big(\{(x_2,y_2)\}\big) &= \max\big\{P_0^{\mathcal{X}}(\{x_2\}) + P_0^{\mathcal{Y}}(\{y_2\}) - 1 - 2\delta, 0\big\} \\ &= 0.2 - 2\delta = \underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}\big(\{(x_2,y_2)\}\big) - 2\delta\end{aligned}$$

for every $\delta \in (0, 0.1)$. On the other hand, if we consider the event $E_1 = \{(x_2, y_2)\}^c$, we deduce from Equations (28) and (8) that

$$\begin{aligned}\underline{E}_{TV}\big(\{(x_2,y_2)\}^c\big) &= 1 - \overline{E}_{TV}\big(\{(x_2,y_2)\}\big) = 0.5 - \delta \quad \textit{while} \\ \underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}\big(\{(x_2,y_2)\}^c\big) &= 1 - \overline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}\big(\{(x_2,y_2)\}\big) = 0.5\end{aligned}$$

meaning that we should distort $\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}$ by δ. We conclude from this that $\underline{E}_{TV}$ is not a TV model starting from $\underline{E}_{P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}}$.

To see that it is not a TV model starting from any probability measure on $\mathcal{X} \times \mathcal{Y}$, recall that, if that was the case, it should be $\overline{E}_{TV}(C) - \underline{E}_{TV}(C) = 2\delta$ whenever $0 < \underline{E}_{TV}(C) < \overline{E}_{TV}(C) < 1$. But in this case we have

$$\begin{aligned}\overline{E}_{TV}(\{(x_2,y_2)\}) - \underline{E}_{TV}(\{(x_2,y_2)\}) &= 0.5 + \delta - (0.2 - 2\delta) = 0.3 + 3\delta \\ \overline{E}_{TV}(\mathcal{X} \times \{y_2\}) - \underline{E}_{TV}(\mathcal{X} \times \{y_2\}) &= 0.5 + \delta - (0.5 - 2\delta) = 3\delta,\end{aligned}$$

a contradiction. ♦

7. Kolmogorov model

Our next distortion model is very much related to the total variation, because it can be regarded as the case when instead of comparing the probability measures we compare their associated distribution functions. It is referred to as Kolmogorov's distortion model (K model, for short).

Assuming that $\mathcal{X}$ is an ordered space, the Kolmogorov distance between two probability measures is defined by

$$d_K(P, Q) = \max_{x \in \mathcal{X}} |F_P(x) - F_Q(x)|.$$

Following our initial assumption that $P_0 \in \mathbb{P}^*(\mathcal{X})$ and that δ is small enough so that $B_{d_K}^{\delta}(P_0) \subseteq \mathbb{P}^*(\mathcal{X})$, the credal set $B_{d_K}^{\delta}(P_0)$ can be expressed as:

$$B_{d_K}^{\delta}(P_0) = \big\{P \in \mathbb{P}(\mathcal{X}) \mid \underline{F}_K(x) \leq F_P(x) \leq \overline{F}_K(x) \ \forall x \in \mathcal{X}\big\},$$

where $\underline{F}_K$ and $\overline{F}_K$ are given by:

$$\begin{aligned}\underline{F}_K(x_i) &= \max\{0, F_{P_0}(x_i) - \delta\} = F_{P_0}(x_i) - \delta \\ \overline{F}_K(x_i) &= \min\{1, F_{P_0}(x_i) + \delta\} = F_{P_0}(x_i) + \delta \qquad \forall i = 1, \ldots, n-1\end{aligned}$$

since $B_{d_K}^{\delta}(P_0) \subseteq \mathbb{P}^*(P_0)$, and $\underline{F}_K(x_n) = \overline{F}_K(x_n) = 1$. This means that $B_{d_K}^{\delta}(P_0) = \mathcal{M}(\underline{F}_K, \overline{F}_K)$, so it coincides with the credal set of a p-box. We will denote as $\underline{P}_K$

and $\overline{P}_K$ the lower and upper probabilities associated with $B^{\delta}_{d_K}(P_0)$. We refer to [31, Sec. 3] for a study of the Kolmogorov model.

7.1. Merging.

Conjunction

We start studying the behaviour of the Kolmogorov model under conjunction. Next example demonstrates that the conjunction of two K models, while being a p-box, does not necessarily correspond to a p-box induced by a K model.

Example 20. *For the model induced by the Kolmogorov distance and for those events A of the type $\{x_1, x_2, \dots, x_k\}$ such that $0 < \underline{P}_K(A) \leq \overline{P}_K(A) < 1$, we have that $\overline{P}_K(A) - \underline{P}_K(A) = 2\delta$. Let us now show that this is not necessarily the case for their conjunction.*

Consider a three-element space $\mathcal{X} = \{x_1, x_2, x_3\}$, $P_0^1 = (0.25, 0.25, 0.5)$, $\delta_1 = 0.05$ and consider the associated K model $B^{\delta_1}_{d_K}(P_0^1) \subseteq \mathbb{P}^*(\mathcal{X})$. On the other hand, let $B^{\delta_2}_{d_K}(P_0^2) \subseteq \mathbb{P}^*(\mathcal{X})$ be the K model induced by the probability measure $P_0^2 = (0.15, 0.35, 0.5)$ and $\delta_2 = 0.05$. Since the conjunction of two p-boxes $(\underline{F}_1, \overline{F}_1)$ and $(\underline{F}_2, \overline{F}_2)$ is the p-box $\big(\max\{\underline{F}_1, \underline{F}_2\}, \min\{\overline{F}_1, \overline{F}_2\}\big)$, we have that

$$\min\left\{\overline{F}_{K_1}(x_1), \overline{F}_{K_2}(x_1)\right\} - \max\left\{\underline{F}_{K_1}(x_1), \underline{F}_{K_2}(x_1)\right\} = 0.2 - 0.2 = 0, \quad \textit{and}$$
$$\min\left\{\overline{F}_{K_1}(x_2), \overline{F}_{K_2}(x_2)\right\} - \max\left\{\underline{F}_{K_1}(x_2), \underline{F}_{K_2}(x_2)\right\} = 0.55 - 0.45 = 0.1,$$

meaning that this conjunction is not a distortion model induced by the Kolmogorov distance. ♦

However, despite the fact that the K model is not closed under conjunction, this conjunction still remains a p-box [16]. Among other things, this means that we have a straightforward way to check whether the conjunction of two models $\mathcal{M}_1 = B^{\delta_1}_{d_K}(P_0^1)$ and $\mathcal{M}_2 = B^{\delta_2}_{d_K}(P_0^2)$ is non-empty. We have that $\mathcal{M}_1 \cap \mathcal{M}_2 \neq \emptyset$ if and only if

$$\min\left\{\overline{F}_{K_1}(x_i), \overline{F}_{K_2}(x_i)\right\} - \max\left\{\underline{F}_{K_1}(x_i), \underline{F}_{K_2}(x_i)\right\} \geq 0 \qquad \forall x_i \in \mathcal{X},$$

which is very easy to check.

Disjunction

The family of Kolmogorov models is not closed under disjunction:

Example 21. *Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and the probability measures $P_0^1 = (0.5, 0.3, 0.2), P_0^2 = (0.3, 0.5, 0.2)$ and $\delta_1 = \delta_2 = 0.05$. The lower probabilities $\underline{P}_{K_1}, \underline{P}_{K_2}$ associated with the K models $B^{\delta_1}_{d_K}(P_0^1) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta_2}_{d_K}(P_0^2) \subseteq \mathbb{P}^*(\mathcal{X})$ and their disjunction $\underline{P}^{\cup}$, are given in the following table:*

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{K_1}$	0.45	0.2	0.15	0.75	0.6	0.45
$\underline{P}_{K_2}$	0.25	0.4	0.15	0.75	0.4	0.65
$\underline{P}^{\cup}$	0.25	0.2	0.15	0.75	0.4	0.45

If $\underline{P}^{\cup}$ was the K model associated with a distribution function F_0 and a distortion factor δ, its associated p-box $(\underline{F}, \overline{F})$ would be determined by the constraints

$$\big[\underline{F}(x_1), \overline{F}(x_1)\big] = [F_0(x_1) - \delta, F_0(x_1) + \delta] = [0.25, 0.55] \Rightarrow F_0(x_1) = 0.4, \delta = 0.15,$$
$$\big[\underline{F}(x_2), \overline{F}(x_2)\big] = [F_0(x_2) - \delta, F_0(x_2) + \delta] = [0.75, 0.85] \Rightarrow F_0(x_2) = 0.8, \delta = 0.05.$$

Thus, $\underline{P}_K^{\cup}$ is not a K model. To see that moreover it does not have a unique undominated outer approximation in the family of K models, consider the K models $B_{d_K}^{\delta_A}(P_A)$ and $B_{d_K}^{\delta_B}(P_B)$ induced by $P_A = (0.4, 0.33, 0.27), P_B = (0.4, 0.36, 0.24)$ and $\delta_A = \delta_B = 0.16$. These produce the following lower probabilities:

	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$
$\underline{P}_{K_A}$	0.24	0.01	0.11	0.57	0.35	0.44
$\underline{P}_{K_B}$	0.24	0.04	0.08	0.6	0.32	0.44

Thus, $\underline{P}_{K_A}, \underline{P}_{K_B}$ are outer approximations of $\underline{P}_K^{\cup}$. If there was a unique undominated outer approximation, then there should be another probability measure P' and $\delta' > 0$ such that either $\underline{P}_{K_A} \lneq \underline{P}' \le \underline{P}_K^{\cup}$ and $\underline{P}_{K_B} \lneq \underline{P}' \le \underline{P}_K^{\cup}$, where $\underline{P}'$ denotes the lower envelope of $B_{d_K}^{\delta'}(P')$. However, this means that:

- *$\overline{P}'(\{x_1\}) - \underline{P}'(\{x_1\}) \ge \overline{P}^{\cup}(\{x_1\}) - \underline{P}^{\cup}(\{x_1\}) = 0.55 - 0.25 = 0.3$, which implies that $\delta' = \frac{1}{2}\Big(\overline{P}'(\{x_1\}) - \underline{P}'(\{x_1\})\Big) \ge 0.15$.*
- *The same reasoning with the event $\{x_1, x_2\}$ leads to:*

$$\begin{aligned}&\overline{P}'(\{x_1, x_2\}) - \underline{P}'(\{x_1, x_2\}) \\ &\quad \le \min\{\overline{P}_{K_A}(\{x_1, x_2\}), \overline{P}_{K_B}(\{x_1, x_2\})\} - \max\{\underline{P}_{K_A}(\{x_1, x_2\}), \underline{P}_{K_B}(\{x_1, x_2\})\} \\ &\quad \le 0.89 - 0.6 = 0.29,\end{aligned}$$

which implies that $\delta' = \frac{1}{2}\Big(\overline{P}'(\{x_1, x_2\}) - \underline{P}'(\{x_1, x_2\})\Big) \le 0.145$.

This is a contradiction, from which we deduce that there is not a unique undominated outer approximation. ♦

In fact, it is worth noting that the disjunction of two Kolmogorov models is not always a p-box, as we show in this example:

Example 22. *Let $\mathcal{X} = \{x_1, x_2, x_3\}$, and consider the distortion models $B_{d_K}^{\delta_1}(P_0^1) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_{d_K}^{\delta_2}(P_0^2) \subseteq \mathbb{P}^*(\mathcal{X})$ given by the probability measures $P_0^1 = (0.3, 0.3, 0.4)$, $P_0^2 = (0.6, 0.3, 0.1)$ and the distortion factors $\delta_1 = \delta_2 = 0.05$. They induce the K models whose associated p-boxes $(\underline{F}_1, \overline{F}_1)$ and $(\underline{F}_2, \overline{F}_2)$ are:*

	x_1	x_2	x_3
$\underline{F}_1$	0.25	0.55	1
$\overline{F}_1$	0.35	0.65	1

	x_1	x_2	x_3
$\underline{F}_2$	0.55	0.85	1
$\overline{F}_2$	0.65	0.95	1

Then the disjunction $B_{d_K}^{\delta_1}(P_0^1) \cup B_{d_K}^{\delta_2}(P_0^2)$ determines the lower and upper cdfs:

	x_1	x_2	x_3
$\underline{F}^{\cup}$	0.25	0.55	1
$\overline{F}^{\cup}$	0.65	0.95	1

Thus, the cdf associated with $P = (0.25, 0.7, 0.05)$ *would belong to* $\mathcal{M}(\underline{F}^{\cup}, \overline{F}^{\cup})$ *but not to* $B_{d_K}^{\delta_1}(P_0^1) \cup B_{d_K}^{\delta_2}(P_0^2)$. ♦

We conclude that the behaviour of the Kolmogorov model is quite inadequate when taking its intersection or union, because it is not preserved by conjunction or disjunction, but also the disjunction is not even a p-box.

Convex mixture

As for the total variation model, it is rather direct to check that the convex mixture of two K models $B_{d_K}^{\delta_1}\left(P_0^1\right)$ and $B_{d_K}^{\delta_2}\left(P_0^2\right)$ is again a K model. Indeed, since the convex combination of two p-boxes is simply the convex combination of the lower and upper cdfs, we have

$$\epsilon \underline{F}_{K_1}(x) + (1-\epsilon)\underline{F}_{K_2}(x) = \epsilon F_{P_0^1}(x) + (1-\epsilon)F_{P_0^2}(x) - \epsilon\delta_1 - (1-\epsilon)\delta_2$$
$$\epsilon \overline{F}_{K_1}(x) + (1-\epsilon)\overline{F}_{K_2}(x) = \epsilon F_{P_0^1}(x) + (1-\epsilon)F_{P_0^2}(x) + \epsilon\delta_1 + (1-\epsilon)\delta_2$$

which are the cdfs induced by the K model $B_{d_K}^{\delta_\epsilon}\left(P_0^\epsilon\right)$ with $\delta_\epsilon = \epsilon\delta_1 + (1-\epsilon)\delta_2$ and

$$F_{P_0^\epsilon}(x) = \epsilon F_{P_0^1}(x) + (1-\epsilon)F_{P_0^2}(x) \quad \forall x \in \mathcal{X}.$$

7.2. **Multivariate setting.**

Marginalisation

In order to study the Kolmogorov model in a multivariate setting, we first need to provide its definition in this context. For this aim, we are now dealing with two ordered spaces $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $\mathcal{Y} = \{y_1, \ldots, y_m\}$. A bivariate p-box [39] is a pair of component-wise increasing functions $\underline{F}_{\mathcal{X},\mathcal{Y}}, \overline{F}_{\mathcal{X},\mathcal{Y}} : \mathcal{X} \times \mathcal{Y} \to [0,1]$ satisfying $\underline{F}_{\mathcal{X},\mathcal{Y}} \leq \overline{F}_{\mathcal{X},\mathcal{Y}}$ and $\underline{F}_{\mathcal{X},\mathcal{Y}}(x_n, y_m) = \overline{F}_{\mathcal{X},\mathcal{Y}}(x_n, y_m) = 1$. A bivariate p-box defines a credal set by:

$$\mathcal{M}\big(\underline{F}_{\mathcal{X},\mathcal{Y}}, \overline{F}_{\mathcal{X},\mathcal{Y}}\big) = \big\{P \in \mathbb{P}(\mathcal{X} \times \mathcal{Y}) \mid \underline{F} \leq F_P \leq \overline{F}\big\}.$$

In contrast with the univariate case, the credal set $\mathcal{M}\big(\underline{F}_{\mathcal{X},\mathcal{Y}}, \overline{F}_{\mathcal{X},\mathcal{Y}}\big)$ may be empty. When it is non-empty, we can define a coherent lower and upper probability by taking lower and upper envelopes.

A bivariate p-box $(\underline{F}_{\mathcal{X},\mathcal{Y}}, \overline{F}_{\mathcal{X},\mathcal{Y}})$ defines marginal (univariate) p-boxes $(\underline{F}_{\mathcal{X}}, \overline{F}_{\mathcal{X}})$ and $(\underline{F}_{\mathcal{Y}}, \overline{F}_{\mathcal{Y}})$, respectively in $\mathcal{X}$ and $\mathcal{Y}$, by:

$$\underline{F}_{\mathcal{X}}(x_i) = \underline{F}_{\mathcal{X},\mathcal{Y}}(x_i, y_m), \quad \overline{F}_{\mathcal{X}}(x_i) = \overline{F}_{\mathcal{X},\mathcal{Y}}(x_i, y_m), \quad \forall i = 1, \ldots, n.$$
$$\underline{F}_{\mathcal{Y}}(y_j) = \underline{F}_{\mathcal{X},\mathcal{Y}}(x_n, y_j), \quad \overline{F}_{\mathcal{Y}}(y_j) = \overline{F}_{\mathcal{X},\mathcal{Y}}(x_n, y_j), \quad \forall j = 1, \ldots, m.$$

In that case, the credal sets $\mathcal{M}\big(\underline{F}_{\mathcal{X}}, \overline{F}_{\mathcal{X}}\big)$ and $\mathcal{M}\big(\underline{F}_{\mathcal{Y}}, \overline{F}_{\mathcal{Y}}\big)$ coincide with the $\mathcal{X}$ and $\mathcal{Y}$ projections, respectively, of the probability measures in $\mathcal{M}\big(\underline{F}_{\mathcal{X},\mathcal{Y}}, \overline{F}_{\mathcal{X},\mathcal{Y}}\big)$. We refer to [28, 39] for some studies about bivariate p-boxes, and to [32] for some comments on the connection between uni- and bivariate p-boxes.

Given the ordered spaces $\mathcal{X}$ and $\mathcal{Y}$, we define the K model for a bivariate cdf $F_{P_0}(x,y)$ as the credal set induced by the bivariate p-box [39]

$$\underline{F}(x,y) = \max\{F_{P_0}(x,y) - \delta, 0\}, \quad \overline{F}(x,y) = \min\{F_{P_0}(x,y) + \delta, 1\} \qquad (29)$$

for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$. We then get the following result regarding the marginals of a bivariate K model:

Proposition 7. *Consider the model* $B^{\delta}_{d_K}\big(P_0^{\mathcal{X},\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{X} \times \mathcal{Y})$ *and its associated lower prevision* $\underline{P}_K$. *Then, the marginal model* $\underline{P}_K^{\mathcal{X}}$ *induces the credal set* $B^{\delta}_{d_K}(P_0^{\mathcal{X}})$ *with* $P_0^{\mathcal{X}}$ *the marginal probability of* $P_0^{\mathcal{X},\mathcal{Y}}$ *on* $\mathcal{X}$.

Proof. Let us first notice that the bivariate p-box defined by Equation (29) induces a coherent lower probability, as it corresponds to the projection over events of the kind $\{x_1, \dots, x_i\} \times \{y_1, \dots, y_j\}$ of the coherent lower probability induced by $B^{\delta}_{d_{TV}}(P_0^{\mathcal{X},\mathcal{Y}})$. We can then easily check that the marginal p-box $\big(\underline{F}_{\mathcal{X}}, \overline{F}_{\mathcal{X}}\big)$ satisfies:

$$\underline{F}_{\mathcal{X}}(x_i) = \underline{F}(x_i, y_m) = F_{P_0^{\mathcal{X},\mathcal{Y}}}(x_i, y_m) - \delta = F_{P_0^{\mathcal{X}}}(x_i) - \delta \quad \forall i = 1, \dots, n-1$$

that is the lower cdf induced by $B^{\delta}_{d_K}(P_0^{\mathcal{X}})$. Similarly,

$$\overline{F}_{\mathcal{X}}(x_i) = \overline{F}(x_i, y_m) = F_{P_0^{\mathcal{X},\mathcal{Y}}}(x_i, y_m) + \delta = F_{P_0^{\mathcal{X}}}(x_i) + \delta \quad \forall i = 1, \dots, n-1.$$

Therefore the marginal model also belongs to the Kolmogorov family. □

Independent products

Consider now two marginal K models $B^{\delta}_{d_K}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta}_{d_K}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$. Regarding the problem of going from marginal models $P_0^{\mathcal{X}}, P_0^{\mathcal{Y}}$ to joint ones, we can first notice that on the Cartesian product of the events $A_i = \{x_1, \dots, x_i\}$ and $B_j = \{y_1, \dots, y_j\}$ (with $i < n$ and $j < m$), we have

$$\begin{aligned}
\underline{P}_K^{\mathcal{X}\times\mathcal{Y}}(A_i \times B_j) &= \underline{F}^{\mathcal{X}\times\mathcal{Y}}(x_i, y_j) = F_{P_0}(x_i, y_j) - \delta, \text{ while} \\
\underline{P}_K^{\mathcal{X}} \boxtimes \underline{P}_K^{\mathcal{Y}}(A_i \times B_j) &= \big(F^{\mathcal{X}}_{P_0}(x_i) - \delta\big)\big(F^{\mathcal{Y}}_{P_0}(y_j) - \delta\big) \\
&= F_{P_0}(x_i, y_j) - \delta\big(F^{\mathcal{X}}_{P_0}(x_i) + F^{\mathcal{Y}}_{P_0}(y_j) - \delta\big)
\end{aligned}$$

where last equation follows from the factorization property in Equation (2). Hence, the equality $\underline{P}_K^{\mathcal{X}\times\mathcal{Y}}(A_i \times B_j) = \underline{P}_K^{\mathcal{X}} \boxtimes \underline{P}_K^{\mathcal{Y}}(A_i \times B_j)$ may not hold. The inequality can go both ways (either $\underline{P}_K^{\mathcal{X}\times\mathcal{Y}}(A \times B) < \underline{P}_K^{\mathcal{X}} \boxtimes \underline{P}_K^{\mathcal{Y}}(A \times B)$ or the reverse) depending on the value of δ, as we show next.

Example 23. *Consider the ordered spaces* $\mathcal{X} = \{x_1, x_2, x_3\}$ *and* $\mathcal{Y} = \{y_1, y_2\}$, *and the probability measures* $P_0^{\mathcal{X}}$ *and* $P_0^{\mathcal{Y}}$ *given by:*

$$P_0^{\mathcal{X}}(\{x_1\}) = 0.3,\ P_0^{\mathcal{X}}(\{x_2\}) = 0.5,\ P_0^{\mathcal{X}}(\{x_3\}) = 0.2,\ P_0^{\mathcal{Y}}(\{y_1\}) = P_0^{\mathcal{Y}}(\{y_2\}) = 0.5.$$

Consider $\delta = 0.1$, *the K models* $B^{\delta}_{d_K}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ *and* $B^{\delta}_{d_K}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$ *and their associated lower probabilities* $\underline{P}_K^{\mathcal{X}}$ *and* $\underline{P}_K^{\mathcal{Y}}$. *Then we obtain on the one hand*

$$\begin{aligned}
\underline{P}_K^{\mathcal{X}\times\mathcal{Y}}\big(\{x_1\} \times \{y_1\}\big) &= 0.15 - \delta = 0.05 < 0.08 = (0.3 - \delta)(0.5 - \delta) \\
&= \underline{P}_K^{\mathcal{X}}(\{x_1\})\underline{P}_K^{\mathcal{Y}}(\{y_1\}) = \underline{P}_K^{\mathcal{X}} \boxtimes \underline{P}_K^{\mathcal{Y}}\big(\{x_1\} \times \{y_1\}\big),
\end{aligned}$$

while on the other hand:

$$\begin{aligned}
\underline{P}_K^{\mathcal{X}\times\mathcal{Y}}\big(\{x_1, x_2\} \times \{y_1\}\big) &= 0.4 - \delta = 0.3 > 0.28 = (0.8 - \delta)(0.5 - \delta) \\
&= \underline{P}_K^{\mathcal{X}}(\{x_1, x_2\})\underline{P}_K^{\mathcal{Y}}(\{y_1\}) = \underline{P}_K^{\mathcal{X}} \boxtimes \underline{P}_K^{\mathcal{Y}}\big(\{x_1, x_2\} \times \{y_1\}\big).
\end{aligned}$$

We conclude that there is no dominance relationship between the coherent lower probabilities obtained by the two approaches. ♦

Let us now show that, as with the other models, K models are not closed under strong products.

Example 24. *Consider the same setting of Example 23. The K models are the p-boxes* $(\underline{F}_{\mathcal{X}}, \overline{F}_{\mathcal{X}})$, $(\underline{F}_{\mathcal{Y}}, \overline{F}_{\mathcal{Y}})$*:*

$\mathcal{X}$	x_1	x_2	x_3
$\underline{F}_{\mathcal{X}}$	0.2	0.7	1
$\overline{F}_{\mathcal{X}}$	0.4	0.9	1

$\mathcal{Y}$	y_1	y_2
$\underline{F}_{\mathcal{Y}}$	0.4	1
$\overline{F}_{\mathcal{Y}}$	0.6	1

Using the factorisation property of the strong product, these generate the following joint bounds for the events $\{x_1, \ldots, x_i\} \times \{y_1, \ldots, y_j\}$*:*

	$\{x_1\} \times \{y_1\}$	$\{x_1\} \times \mathcal{Y}$	$\{x_1, x_2\} \times \{y_1\}$	$\{x_1, x_2\} \times \mathcal{Y}$	$\mathcal{X} \times \{y_1\}$
$\underline{P}_K^{\mathcal{X}} \boxtimes \underline{P}_K^{\mathcal{Y}}$	0.08	0.2	0.28	0.7	0.4
$\overline{P}_K^{\mathcal{X}} \boxtimes \overline{P}_K^{\mathcal{Y}}$	0.24	0.4	0.54	0.9	0.6

If this was a K model, than the differences between the upper and lower probabilities for these events should be constant, which is not the case here. ♦

Natural extension of marginal models

Consider now two K models $B_{d_K}^{\delta_{\mathcal{X}}}(P_0^{\mathcal{X}})$ and $B_{d_K}^{\delta_{\mathcal{Y}}}(P_0^{\mathcal{Y}})$, and let us focus on the problem of applying to them the natural extension. We already know that these K models are equivalent to the univariate p-boxes $(\underline{F}_{\mathcal{X}}, \overline{F}_{\mathcal{X}})$ and $(\underline{F}_{\mathcal{Y}}, \overline{F}_{\mathcal{Y}})$. Thus, we can use the results from [32, Sec. 3.2], where we studied the natural extension of two p-boxes. In particular, in [32, Prop. 5] we proved that $\mathcal{E}\big((\underline{F}_{\mathcal{X}}, \overline{F}_{\mathcal{X}}), (\underline{F}_{\mathcal{Y}}, \overline{F}_{\mathcal{Y}})\big) = \mathcal{M}\big(\underline{F}, \overline{F}\big)$, where $\big(\underline{F}, \overline{F}\big)$ is the bivariate p-box given, for any $(x, y) \in \mathcal{X} \times \mathcal{Y}$, by:

$$\underline{F}(x, y) = \max\big\{\underline{F}_{\mathcal{X}}(x) + \underline{F}_{\mathcal{Y}}(y) - 1, 0\big\}, \quad \overline{F}(x, y) = \min\big\{\overline{F}_{\mathcal{X}}(x), \overline{F}_{\mathcal{Y}}(y)\big\}. \tag{30}$$

Then, the natural extension of two K models is again a p-box. However, our next example shows that this natural extension is not a K model.

Example 25. *Consider the spaces* $\mathcal{X}$ *and* $\mathcal{Y}$ *and the probabilities* $P_0^{\mathcal{X}}$ *and* $P_0^{\mathcal{Y}}$ *from Example 23 and the associated p-boxes detailed in Example 24. Applying Equation* (30)*, their natural extension is the p-box* $\big(\underline{F}, \overline{F}\big)$*:*

y_2	$[0.2, 0.4]$	$[0.7, 0.9]$	$[1, 1]$
y_1	$[0, 0.4]$	$[0.1, 0.6]$	$[0.4, 0.6]$
$[\underline{F}(x_i, y_j), \overline{F}(x_i, y_j)]$	x_1	x_2	x_3

We can see that:

$$\overline{F}(x_1, y_2) - \underline{F}(x_1, y_2) = 0.2, \quad \overline{F}(x_2, y_1) - \underline{F}(x_2, y_1) = 0.5.$$

Since the differences between $\overline{F}$ *and* $\underline{F}$ *are not constant, we conclude that* $\big(\underline{F}, \overline{F}\big)$ *does not correspond to a K model.* ♦

We conclude that, even if there is a simple formula for computing the natural extension (Equation (30)), the latter is not a Kolmogorov model.

8. L_1 DISTORTION MODEL

We conclude our investigation by considering the distortion model associated with the L_1 distance. Given two probability measures P, Q, their L_1 distance is

$$d_{L_1}(P,Q) = \sum_{A \subseteq \mathcal{X}} |P(A) - Q(A)|.$$

In order to alleviate the notation, we shall also denote it by d_1. This distance has been used in robust statistics in [42]. When $P_0 \in \mathbb{P}^*(\mathcal{X})$ and δ is small enough, it induces the credal set $B^{\delta}_{d_1}(P_0) \subseteq \mathbb{P}^*(\mathcal{X})$ whose associated lower and upper probabilities are [31, Thm. 11]:

$$\underline{P}_{L_1}(A) = P_0(A) - \frac{\delta}{\varphi(|\mathcal{X}|, |A|)} \quad \forall A \neq \emptyset, \mathcal{X}, \tag{31}$$

where

$$\varphi(n,k) = \sum_{l=0}^{k} \binom{k}{l} \sum_{j=0}^{n-k} \binom{n-k}{j} \left| \frac{l}{k} - \frac{j}{n-k} \right| \quad \forall k = 1, \dots, n.$$

8.1. **Merging.** Let us analyse the behaviour of the L_1 model under conjunction, disjunction and convex mixture.

Conjunction and disjunction

In general, the conjunction of two L_1 models will not lead to a new L_1 model. In fact, observe that in the case of ternary spaces, we can establish a relationship between the total variation and L_1 models. In that case, $\varphi(3,1) = \varphi(3,2) = 4$, whence from Equation (31) $\underline{P}_{L_1}$ is given by:

$$\underline{P}_{L_1}(A) = P_0(A) - \frac{\delta}{4} \quad \forall A \neq \emptyset, \mathcal{X}.$$

Also, from Equation (25), the total variation model generated by the probability measure P_0 and the distortion parameter $\frac{\delta}{4}$ is given by:

$$\underline{P}_{TV}(A) = P_0(A) - \frac{\delta}{4} \quad \forall A \neq \emptyset, \mathcal{X}.$$

Since we are assuming that δ is small enough such that $\underline{P}_{d_1}(A) > 0$ for every $A \neq \emptyset$, we conclude that in cardinality three $B^{\delta}_{d_1}(P_0) = B^{\delta/4}_{d_{TV}}(P_0)$. Thus, if we consider $\mathcal{X} = \{x_1, x_2, x_3\}$, $P_0^1 = (0.41, 0.37, 0.22)$, $P_0^2 = (0.37, 0.41, 0.22)$ and $\delta_1 = \delta_2 = 0.48$, we obtain that

$$B^{0.48}_{d_1}(P_0^1) = B^{0.12}_{d_{TV}}(P_0^1) \text{ and } B^{0.48}_{d_1}(P_0^2) = B^{0.12}_{d_{TV}}(P_0^2).$$

Using Example 15, we conclude that the family of L_1 models is not closed under conjunction nor disjunction.

Moreover, taking into account this same connection between the TV and the L_1 models on ternary spaces as well as Example 16, we conclude that the disjunction of two L_1 models does not possess a unique undominated outer approximation in the L_1 family.

Several facts indicate that checking in which cases the conjunction of two L_1 models is a L_1 model is a difficult task, for a number of reasons: (i) a L_1 model is not determined in general by the lower probability it defines on events, in the sense that a probability measure P may dominate the lower envelope of $B^{\delta}_{d_1}(P_0)$ on any event but still do not belong to $B^{\delta}_{d_1}(P_0)$; (ii) it is not described by an explicit lower prevision; and (iii) even enumerating the extreme points of $B_{d_1}(P_0)$ is to this point an open issue [31].

Convex mixture

Again, it is rather direct to check that, if we consider their restriction on events, the convex mixture of two lower probabilities associated to L_1 models $B^{\delta_1}_{d_1}(P^1_0)$ and $B^{\delta_2}_{d_1}(P^2_0)$ is again a lower probability associated to an L_1 model, as we have

$$\epsilon \underline{P}^{\delta_1}_{L_1}(A) + (1-\epsilon)\underline{P}^{\delta_2}_{L_1}(A) = \epsilon P^1_0(A) + (1-\epsilon)P^2_0(A) - \epsilon\frac{\delta_1}{\varphi(|\mathcal{X}|,|A|)} - (1-\epsilon)\frac{\delta_2}{\varphi(|\mathcal{X}|,|A|)}$$

which are lower probabilities induced by the L_1 model $B^{\delta_\epsilon}_{d_1}(P^{\epsilon}_0)$ with $\delta_\epsilon = \epsilon\delta_1 + (1-\epsilon)\delta_2$ and

$$P^{\epsilon}_0(\{x\}) = \epsilon P^1_0(\{x\}) + (1-\epsilon)P^2_0(\{x\}) \quad \forall x \in \mathcal{X}.$$

Since L_1 models can be described by their lower probabilities whenever $n \leq 11$, this is sufficient to show that in those cases the model is closed under convex mixture. The case $n > 11$, for which we have no simple and explicit description of the lower envelope of the L_1 model, remains an open problem.

8.2. **Multivariate setting.**

Marginalisation

For the multivariate case, let us first look at the marginals of a joint L_1 model $B^{\delta}_{d_1}(P^{\mathcal{X},\mathcal{Y}}_0) \subseteq \mathbb{P}^*(\mathcal{X}\times\mathcal{Y})$. Somewhat surprisingly, the marginal model of a L_1 model is not a L_1 model, as our next example shows.

Example 26. *Let $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$, $\mathcal{Y} = \{y_1, y_2\}$, $P^{\mathcal{X},\mathcal{Y}}_0$ the uniform distribution on $\mathcal{X}\times\mathcal{Y}$ and consider the distortion parameter $\delta = 0.1$. Then:*

$$\begin{aligned}
\underline{P}^{\mathcal{X}}_{L_1}(\{x_i\}) &= \underline{P}_{L_1}(\{x_i\}\times\mathcal{Y}) = P_0(\{x_i\}\times\mathcal{Y}) - \frac{0.1}{\varphi(8,2)} = \frac{1}{4} - \frac{0.1}{84} \\
&= P^{\mathcal{X}}_0(\{x_i\}) - \frac{\delta_{\mathcal{X}}}{\varphi(4,1)} = \frac{1}{4} - \frac{\delta_{\mathcal{X}}}{8} \Rightarrow \delta_{\mathcal{X}} = \frac{0.1}{10.5}. \\
\underline{P}^{\mathcal{X}}_{L_1}(\{x_i, x_j\}) &= \underline{P}_{L_1}(\{x_i, x_j\}\times\mathcal{Y}) = P_0(\{x_i, x_j\}\times\mathcal{Y}) - \frac{0.1}{\varphi(8,4)} = \frac{1}{2} - \frac{0.1}{70} \\
&= P^{\mathcal{X}}_0(\{x_i, x_j\}) - \frac{\delta_{\mathcal{X}}}{\varphi(4,2)} = \frac{1}{2} - \frac{\delta_{\mathcal{X}}}{6} \Rightarrow \delta_{\mathcal{X}} = \frac{0.3}{35}.
\end{aligned}$$

Since $\frac{0.1}{10.5} \neq \frac{0.3}{35}$, we conclude that the marginal model does not belong to the L_1 family. ♦

Independent products

We now investigate what happens when building a joint from marginal L_1 models $B^{\delta}_{d_1}(P^{\mathcal{X}}_0) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B^{\delta}_{d_1}(P^{\mathcal{Y}}_0) \subseteq \mathbb{P}^*(\mathcal{Y})$. Again, on Cartesian products of events, we have

$$\underline{P}^{\mathcal{X}\times\mathcal{Y}}_{L_1}(A\times B) = P^{\mathcal{X}\times\mathcal{Y}}_0(A\times B) - \frac{\delta}{\varphi(|\mathcal{X}\times\mathcal{Y}|, |A\times B|)},$$

while

$$\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B) = \left(P_0^{\mathcal{X}}(A) - \frac{\delta}{\varphi(|\mathcal{X}|,|A|)}\right)\left(P_0^{\mathcal{Y}}(B) - \frac{\delta}{\varphi(|\mathcal{Y}|,|B|)}\right).$$

In that case, it may happen that $\underline{P}_{L_1}^{\mathcal{X}\times\mathcal{Y}}(A\times B) \geq \underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B)$ or $\underline{P}_{L_1}^{\mathcal{X}\times\mathcal{Y}}(A\times B) \leq \underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B)$, even with strict inequality, as we show in next example.

Example 27. *Consider $\mathcal{X}$ and $\mathcal{Y}$ such that $|\mathcal{X}| = |\mathcal{Y}| = 2$, P_0 the uniform distribution on $\mathcal{X}\times\mathcal{Y}$ and A, B two singletons. The connection between $\underline{P}_{L_1}^{\mathcal{X}\times\mathcal{Y}}(A\times B)$ and $\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B)$ depends on whether $\delta > \frac{3}{2}$ or $\delta < \frac{3}{2}$. The reason is that if $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$ are singletons, $\varphi(|\mathcal{X}|,|A|) = \varphi(|\mathcal{Y}|,|B|) = 2$, $\varphi(|\mathcal{X}\times\mathcal{Y}|,|A\times B|) = \varphi(4,1) = 8$, $P_0(A\times B) = \frac{1}{4}$ and $P_0^{\mathcal{X}}(A) = P_0^{\mathcal{Y}}(B) = \frac{1}{2}$. Hence:*

$$\underline{P}_{L_1}^{\mathcal{X}\times\mathcal{Y}}(A\times B) = \frac{1}{4} - \frac{\delta}{8}, \quad \underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B) = \left(\frac{1}{2} - \frac{\delta}{2}\right)\cdot\left(\frac{1}{2} - \frac{\delta}{2}\right).$$

Operating, we obtain that $\underline{P}_{L_1}^{\mathcal{X}\times\mathcal{Y}}(A\times B) > \underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B)$ if and only if $\delta < \frac{3}{2}$ and $\underline{P}_{L_1}^{\mathcal{X}\times\mathcal{Y}}(A\times B) < \underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(A\times B)$ if and only if $\delta > \frac{3}{2}$. Therefore, in general there is no dominance relation between the two approaches. ♦

Let us now show through an example that the L_1 model is also not closed under strong products.

Example 28. *Let us consider the case where $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2\}$ with $P_0^{\mathcal{X}}$ and $P_0^{\mathcal{Y}}$ uniform with some δ. Let us now assume that $\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}$ is a L_1 model with some δ^*. Due to the factorization property, we have*

$$\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}\big(\{(x_i, y_j)\}\big) = \left(0.5 - \frac{\delta}{2}\right)\left(0.5 - \frac{\delta}{2}\right) \text{ for } i, j \in \{1, 2\}$$

and since by assumption this should also be a L_1 model with a uniform distribution, we should also have

$$\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}\big(\{(x_i, y_j)\}\big) = \frac{1}{4} - \frac{\delta^*}{\varphi(4,1)} = \frac{1}{4} - \frac{\delta^*}{8}.$$

Fixing $\delta = 0.1$, the two equalities lead to

$$\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}\big(\{(x_i, y_j)\}\big) = 0.45^2 = \frac{1}{4} - \frac{\delta^*}{8},$$

which gives us $\delta^* = 0.38$ for the joint model $\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}$. However, on the event $\{x_1\}\times\mathcal{Y}$ this gives $\underline{P}_{L_1}^{\mathcal{X}} \boxtimes \underline{P}_{L_1}^{\mathcal{Y}}(\{x_1\}\times\mathcal{Y}) = 0.45$, which is different from $1/2 - \delta^*/\varphi(4,2) = 0.437$. ♦

Natural extension of marginal models

Consider now two L_1 models $B_{d_1}^{\delta_{\mathcal{X}}}\big(P_0^{\mathcal{X}}\big) \subseteq \mathbb{P}^*(\mathcal{X})$ and $B_{d_1}^{\delta_{\mathcal{Y}}}\big(P_0^{\mathcal{Y}}\big) \subseteq \mathbb{P}^*(\mathcal{Y})$ in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Using Equations (5) and (6), we can give the form of the lower and upper bounds of their natural extension in the events $A\times B$:

$$\underline{E}_{L_1}(A\times B) = \max\left\{P_0^{\mathcal{X}}(A) - \frac{\delta_{\mathcal{X}}}{\varphi(|\mathcal{X}|,|A|)} + P_0^{\mathcal{Y}}(B) - \frac{\delta_{\mathcal{Y}}}{\varphi(|\mathcal{Y}|,|B|)} - 1, 0\right\}.$$

$$\overline{E}_{L_1}(A\times B) = \min\left\{P_0^{\mathcal{X}}(A) + \frac{\delta_{\mathcal{X}}}{\varphi(|\mathcal{X}|,|A|)}, P_0^{\mathcal{Y}}(B) + \frac{\delta_{\mathcal{Y}}}{\varphi(|\mathcal{Y}|,|B|)}\right\}.$$

When the distortion parameters coincide, these expressions simplify to:

$$\underline{E}_{L_1}(A\times B)=\max\left\{P_0^{\mathcal{X}}(A)+P_0^{\mathcal{Y}}(B)-1-\delta\left(\frac{1}{\varphi(|\mathcal{X}|,|A|)}+\frac{1}{\varphi(|\mathcal{Y}|,|B|)}\right),0\right\}.$$

$$\overline{E}_{L_1}(A\times B)=\min\left\{P_0^{\mathcal{X}}(A)+\frac{\delta}{\varphi(|\mathcal{X}|,|A|)},P_0^{\mathcal{Y}}(B)+\frac{\delta}{\varphi(|\mathcal{Y}|,|B|)}\right\}.$$

Note that none of the equations can be simplified, because even if the distortion parameters coincide, the expressions depend on the values $\varphi(|\mathcal{X}|,|A|)$ and $\varphi(|\mathcal{Y}|,|B|)$.

Taking this into account, we deduce that in general $\mathcal{E}\big(\underline{P}_{L_1}^{\mathcal{X}},\underline{P}_{L_1}^{\mathcal{Y}}\big)$ cannot be expressed as $B_{d_1}^{\delta}\big(\underline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}\big)$: it suffices to note that, since $\mathcal{X},\mathcal{Y}$ are binary spaces, $B_{d_1}^{\delta}(P_0)=B_{d_{TV}}^{\delta/2}(P_0)$ (see for instance [31, p.646]), so we can use the same Example 19 for the TV model to deduce that we do not have the equality between $\mathcal{E}\big(\underline{P}_{L_1}^{\mathcal{X}},\underline{P}_{L_1}^{\mathcal{Y}}\big)$ and $B_{d_1}^{\delta}\big(\underline{E}_{P_0^{\mathcal{X}},P_0^{\mathcal{Y}}}\big)$.

9. Conclusions

The variety of distortion models present in the literature makes it interesting to develop tools to compare their behaviour in a number of situations, so as to be able to choose the most appropriate model in each scenario. In this paper, we have complemented our earlier work in [30, 31] and compared six different distortion models by determining (i) if they are closed under conjunction, disjunction or convex mixture; (ii) whether there is a unique procedure to build an independent product; (iii) if they are closed under marginalisation; and (iv) whether the procedures of distortion and natural extension commute. Tables 1 and 2 summarise our results. From these tables, it is clear that the PMM and LV models are the most stable, and that the least stables are the COR and L_1 models, respectively from the merging and multi-variate point of view.

	Conjunction	Disjunction	Mixture	Unique OA?
PMM	YES [29, Prop.12]	NO (Ex.2)	YES	YES [29, Prop.12]
LV	YES (Prop.3)	NO (Ex.5)	YES	YES [33, Prop.8]
COR	NO (Ex.9)	NO (Ex.10)	NO (Ex.11)	NO (Ex.10)
TV	NO (Ex.15)	NO (Ex.15)	YES	NO (Ex.16)
K	NO (Ex.20)	NO (Ex.21)	YES	NO (Ex.21)
L_1	NO (Ex.15)	NO (Ex.15)	YES	NO (Ex.16)

Table 1. Behaviour of the neighbourhood models under conjunction, disjunction and convex mixture.

In the case of the natural extension, we should remark that, strictly speaking, the natural extension of two marginal pari mutuel models is only a PMM if we regard it as a PMM-distortion like of a lower probability, but not in the sense of Definition 2; see Theorem 2 for more details.

Beyond these comparisons, there are a few global remarks that we find interesting:

- Regarding merging, the union of two convex sets $\mathcal{M}_1$ and $\mathcal{M}_2$ will not be convex in general [54, Thm. 6], and it is common to consider the convex hull $ch(\mathcal{M}_1\cup\mathcal{M}_2)$ of the union. None of the distortion models considered

	Marginalising	Strong Product	Natural Extension
PMM	YES [29, Sec.6.2]	NO (Ex.4)	YES (Thm.2)
LV	YES (Prop.4)	NO (Ex.7)	NO (Ex.8)
COR	YES (Prop.5)	NO (Ex.13)	NO (Ex.14)
TV	YES (Prop.6)	NO (Ex.18)	NO (Ex.19)
K	YES (Prop.7)	NO (Ex.24)	NO (Ex.25)
L_1	NO (Ex.26)	NO (Ex.28)	NO (Ex.19)

TABLE 2. Behaviour of the neighbourhood models under different operations.

in this study is closed under disjunction; we may then consider the problem of outer approximating this disjunction within that family. However, this outer approximation is not unique [33]. The PMM and LV models, being special instances of probability intervals, are remarkable exceptions, as for any set $\mathcal{M}$, its greatest outer-approximation in terms of the pari mutuel or the linear vacuous is unique [33].

- The results regarding the problem of constructing a joint model, and the relation between the two approaches considered here (combine then distort vs. distort then combine), are valid for other independence notions from imprecise probability [11]. Indeed, most of them (including epistemic independence and random set independence, for instance) also satisfy Equation (2), hence the inequality concerning events of the kind $A \times B$ remains true for them. As this factorisation property is also true for lower probabilities, the various examples and discussion given for the different models also apply to them. This may be an important issue when having to choose whether one should first combine then discount, or discount then combine. We can nevertheless observe that there is essentially one way to apply the first option (using stochastic independence), and many to apply the second (as one has to choose an adequate notion of independence).

Taking these results into account, it may be interesting to look at the problem from a different angle: to characterise those distortion models that are closed under merging operations in terms of the properties of the distorting function d. While this is left as future research, we can give a number of preliminary comments. On the one hand, it may be useful to consider some of the properties from [30, 31], where we characterised those distorting functions d determining probability intervals, for which it may be easier to analyse their conjunction and disjunction, using results from de Campos et al. [13]. Note nevertheless that whether those probability intervals will remain distortion models is not guaranteed in general. With respect to the natural extension, we think that Proposition 1 should be useful in this regard; and concerning independent products, we conjecture that the equality between the two approaches considered in the paper will only hold in very particular cases.

Our work in this paper may be extended in a number of ways: on the one hand, we may consider other distortion models, such as those based on divergences such as Kullback-Leibler [12, 35] or the recently introduced nearly-linear models [10]; on the other hand, we may consider other models of merging [52] or of independence [11]; and we may take the approach one step further and consider distorted credal sets, considering the ideas put forward by Moral in [35].

Acknowledgements

This work was carried out in the framework of the Labex MS2T, funded by the French Government, through the National Agency for Research (Reference ANR-11-IDEX-0004-02). We also acknowledge the financial support of project PGC2018-098623-B-I00 from the Ministry of Science, Innovation and Universities of Spain. A preliminary version of this paper was presented at the ISIPTA'2021 conference [18]. The comments and suggestions from the reviewers and participants of the conference are much appreciated. We would also like to thank the anonymous reviewers for their very careful reading of the paper that led to a number of improvements.

References

[1] B.C. Arnold, E. Castillo, and J.M. Sarabia. Compatibility of partial or complete conditional probabilities specifications. *Journal of Statistical Planning and Infernce*, 123:133–159, 2004.

[2] T. Augustin, F. Coolen, G. de Cooman, and M. Troffaes, editors. *Introduction to Imprecise Probabilities*. Wiley Series in Probability and Statistics. Wiley, 2014.

[3] C. Baudrit, I. Couso, and D. Dubois. Joint propagation of probability and possibility in risk analysis: Towards a formal framework. *International Journal of Approximate Reasoning*, 45(1):82–105, 2007.

[4] A. Benavoli and M. Zaffalon. Density-ratio robustness in dynamic state estimation. *Mechanical Systems and Signal Processing*, 37(1–2):54–75, 2013.

[5] J. Berger. Robust Bayesian analysis: sensitivity to the prior. *Journal of Statistical Planning and Inference*, 25:303–328, 1990.

[6] A. Bronevich. On the clousure of families of fuzzy measures under eventwise aggregations. *Fuzzy Sets and Systems*, 153:45–70, 2005.

[7] A. Bronevich. Necessary and sufficient consensus conditions for the eventwise aggregation of lower probabilities. *Fuzzy Sets and Systems*, 158:881–894, 2007.

[8] A. Bronevich and T. Augustin. Approximation of coherent lower probabilities by 2-monotone measures. In *Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'2009)*, pages 61–70, 2009.

[9] A. Chateauneuf. Decomposable capacities, distorted probabilities and concave capacities. *Mathematical Social Sciences*, 31:19–37, 1996.

[10] C. Corsato, R. Pelessoni, and P. Vicig. Nearly-linear uncertainty measures. *International Journal of Approximate Reasoning*, 114:1–28, 2019.

[11] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5:165–181, 2000.

[12] I. Csiszár. Axiomatic characterization of information measures. *Entropy*, 10:261–273, 2008.

[13] L.M. de Campos, J.F. Huete, and S. Moral. Probability intervals: a tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2:167–196, 1994.

[14] G. de Cooman, M. C. M. Troffaes, and E. Miranda. n-Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347:143–156, 2008.

[15] S. Destercke and V. Antoine. Combining imprecise probability masses with maximal coherent subsets: application to ensemble classification. In *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, pages 27–35. Springer, 2013.

[16] S. Destercke and D. Dubois. The role of generalised p-boxes in imprecise probability models. In *Proceedings of the 6th International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'2009)*, pages 179–188, 2009.

[17] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: I. Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.

[18] S. Destercke, I. Montes, and E. Miranda. Processing multiple distortion models: a comparative study. In *Proceedings of the 12th International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'2021)*, pages 126–135, 2021.

[19] S. Ferson, V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002–4015, Sandia National Laboratories, January 2003.

[20] T. Herron, T. Seidenfeld, and L. Wasserman. Divisive conditioning: further results on dilation. *Philosophy of Science*, 64:411–444, 1997.
[21] P.J. Huber. *Robust Statistics*. Wiley, New York, 1981.
[22] P.J. Huber and V. Strassen. Minimax tests and the Neyman–Pearson lemma for capacities. *The Annals of Statistics*, 1:251–263, 1973.
[23] H.G. Kellerer. Verteilungsfunktionen mit gegebenen marginalverteilungen. *Z. Wahrscheinlichkeitstheorie*, 3:247–270, 1964.
[24] E. Miranda, I. Montes, and P. Vicig. On the selection of an optimal outer approximation of a coherent lower probability. *Fuzzy Sets and Systems*, 424C:1–36, 2021.
[25] E. Miranda, M. Troffaes, and S. Destercke. A geometric and game theoretic study of the conjunction of possibility measures. *Information Sciences*, 298:373–389, 2015.
[26] E. Miranda and M. Zaffalon. Compatibility, desirability and the running intersection property. *Artificial Intelligence*, 283, 2020.
[27] I. Montes and S. Destercke. On extreme points of p-boxes and belief functions. *Annals of Mathematics and Artificial Intelligence*, 81:405–428, 2017.
[28] I. Montes and E. Miranda. Bivariate p-boxes and maxitive functions. *International Journal of General Systems*, 46(4):354–385, 2017.
[29] I. Montes, E. Miranda, and S. Destercke. Pari-mutuel probabilities as an uncertainty model. *Information Sciences*, 481:550 – 573, 2019.
[30] 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(6):605–635, 2020.
[31] I. Montes, E. Miranda, and S. Destercke. Unifying neighbourhood and distortion models: Part II- New models and synthesis. *International Journal of General Systems*, 49(6):636–674, 2020.
[32] I. Montes, E. Miranda, R. Pelessoni, and P. Vicig. Sklar's theorem in an imprecise setting. *Fuzzy Sets and Systems*, 278C:48–66, 2015.
[33] I. Montes, E. Miranda, and P. Vicig. 2-monotone outer approximations of coherent lower probabilities. *International Journal of Approximate Reasoning*, 101:181–205, 2018.
[34] I. Montes, E. Miranda, and P. Vicig. Outer approximations of coherent lower probabilities with belief functions. *International Journal of Approximate Reasoning*, 110:1–30, 2019.
[35] S. Moral. Discounting imprecise probabilities. In E. Gil, E. Gil, J. Gil, and M. Gil, editors, *The Mathematics of the Uncertain*, volume 142 of *Studies in Systems, Decision and Control*. Springer, 2018.
[36] S. Moral and J. del Sagrado. Aggregation of imprecise probabilities. In *Aggregation and fusion of imperfect information*, pages 162–188. Springer, 1998.
[37] R. Nau. The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference*, 105:265–282, 2002.
[38] W.L. Oberkampf, J.C. Helton, S.F Wojtkiewicz, C.A. Joslyn, and S. Ferson. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*, 85(1-3):11–20, 2004.
[39] R. Pelessoni, P. Vicig, I. Montes, and E. Miranda. Bivariate p-boxes. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 24(2):229–263, 2016.
[40] R. Pelessoni, P. Vicig, and M. Zaffalon. Inference and risk measurement with the pari-mutuel model. *International Journal of Approximate Reasoning*, 51:1145–1158, 2010.
[41] L.R. Pericchi and P. Walley. Robust Bayesian credible intervals and prior ignorance. *International Statistical Review*, 59:1–23, 1991.
[42] P.J. Rousseeuw and A.M. Leroy. *Robust regression and outlier detection*. John Wiley & Sons, 2005.
[43] M. Schervish and T. Seidenfeld. An approach to consensus and certainty with increasing evidence. *Journal of Statistical Planning and Inference*, 25(3):401–414, 1990.
[44] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.
[45] T. Seidenfeld, M. Schervish, and J. Kadane. Decisions without ordering. In *Acting and reflecting*, pages 143–170. Springer, 1990.
[46] V. Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36:423–429, 1965.

[47] M.C.M. Troffaes and G. de Cooman. *Lower previsions*. Wiley Series in Probability and Statistics. Wiley, 2014.

[48] M.C.M. Troffaes and S. Destercke. Probability boxes on totally preordered spaces for multivariate modelling. *International Journal of Approximate Reasoning*, 52(6):767–791, 2011.

[49] M.C.M. Troffaes, E. Miranda, and S. Destercke. On the connection between probability boxes and possibility measures. *Information Sciences*, 224:88–108, 2013.

[50] L. Utkin. A framework for imprecise robust one-class classification models. *Journal of Machine Learning Research and Cybernetics*, 5(3):379–393, 2014.

[51] P. Walley. Coherent lower (and upper) probabilities. Statistics Research Report 22, University of Warwick, 1981.

[52] P. Walley. The elicitation and aggregation of beliefs. Statistics Research Report 23, University of Warwick, 1982.

[53] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

[54] M. Zaffalon and E. Miranda. Probability and time. *Artificial Intelligence*, 198:1–51, 2013.

UMR CNRS 7253 Heudiasyc, Sorbonne universités, Université de technologie de Compiègne CS 60319 - 60203 Compiègne cedex, France
Email address: `sebastien.destercke@hds.utc.fr`

University of Oviedo, Department of Statistics and Operations Research
Email address: `imontes@uniovi.es`

University of Oviedo, Department of Statistics and Operations Research
Email address: `mirandaenrique@uniovi.es`