Proceedings of Machine Learning Research 147:122–131, 2021 ISIPTA 2021

# Processing Multiple Distortion Models: a Comparative Study

**Sébastien Destercke** SEBASTIEN.DESTERCKE@HDS.UTC.FR
*UMR CNRS 7253 Heudiasyc, Sorbonne universités, Université de technologie de Compiègne CS - 60203 Compiègne cedex, France*

**Ignacio Montes** IMONTES@UNIOVI.ES
**Enrique Miranda** MIRANDAENRIQUE@UNIOVI.ES
*University of Oviedo, Department of Statistics and Operations Research, Oviedo, Spain*

## Abstract

When dealing with uncertain information, distortion or neighbourhood models are convenient practical tools, as they rely on very few parameters. In this paper, we study their behaviour when such models are combined and processed. 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, natural extension

## 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.

The mathematical properties of such neighbourhood models heavily depend on the chosen distance. In our recent works [16, 17], we analysed the polytopes of probabilities induced by different distances . Yet we did not explore what happens when dealing with multiple neighbourhood models. This is what we do in this paper, where we look at two important tasks: (1) merging models bearing on the same domain [19]. In particular, we focus on the operations of conjunction, disjunction and convex mixtures. And (2) combining models defined on different domains. We analyse the properties of the distortion models when we marginalise a joint model, or when we build a joint model using marginal ones.

The rest of the paper is organised as follows: we first provide necessary notions and notations in Section 2, and then investigate in the following sections the behaviour of the most commonly used models in the literature, when those are merged or combined, reminding the basics of each distortion model in the corresponding section. Section 3 deals with the pari mutuel model [17, 21, 27]; Section 4, with the linear-vacuous model [11, 16, 27]; Section 5 focuses on the constant-odds ratio model [27]; finally Section 6 deals with the model induced by the total variation distance [10, 27]. Some additional comments are provided in Section 7.

## 2. Preliminary Concepts

We consider finite possibility spaces, denoted by $\mathscr{X}$, $\mathscr{Y}$ or their product space $\mathscr{X}\times\mathscr{Y}$. We denote by $\mathscr{P}(\mathscr{X})$ the power set of a space $\mathscr{X}$, by $\mathbb{P}(\mathscr{X})$ the set of probability measures on $\mathscr{X}$, and by $\mathbb{P}^*(\mathscr{X})$ the set of probability measures $P$ satisfying $P(A)\in(0,1)$ for any $A\neq\emptyset,\mathscr{X}$.

### 2.1. Imprecise Probabilities

Let us introduce some basic notions from imprecise probability theory used in this paper; we refer to [1, 24, 27] for details.

A *lower probability* on $\mathscr{X}$ is a function $\underline{P}:\mathscr{P}(\mathscr{X})\to[0,1]$ that is monotone ($A\subseteq B$ implies $\underline{P}(A)\leq\underline{P}(B)$) and normalized ($\underline{P}(\emptyset)=0,\underline{P}(\mathscr{X})=1$). Its conjugate *upper probability* is given by $\overline{P}(A)=1-\underline{P}(A^c)$ for every $A\subseteq\mathscr{X}$.

To any $\underline{P}$, we can associate a closed and convex *credal set*:

$$\mathscr{M}(\underline{P}):=\{P\in\mathbb{P}(\mathscr{X})\mid P(A)\geq\underline{P}(A)\quad\forall A\subseteq\mathscr{X}\},$$

and $\underline{P}$ is called *coherent* when it is the lower envelope of a non-empty $\mathscr{M}(\underline{P})$. All $\underline{P}$ in this paper will be coherent.

A more general notion than lower probability is that of *lower prevision*. A *gamble* on $\mathscr{X}$ is a function $f:\mathscr{X}\to\mathbb{R}$, and the set of all the gambles on $\mathscr{X}$ is denoted by $\mathscr{L}(\mathscr{X})$. A *lower prevision* is a map $\underline{P}:\mathscr{L}(\mathscr{X})\to\mathbb{R}$. The credal set induced by the lower prevision $\underline{P}$ is defined as:

$$\mathscr{M}(\underline{P})=\{P\in\mathbb{P}(\mathscr{X})\mid P(f)\geq\underline{P}(f)\quad\forall f\in\mathscr{L}(\mathscr{X})\}.$$

$\underline{P}$ is called *coherent* when $\underline{P}$ is the lower envelope of $\mathscr{M}(\underline{P})$, meaning that $\underline{P}(f)=\min_{P\in\mathscr{M}(\underline{P})}P(f)$, where $P(f)$ denotes the expectation of the gamble $f$ with respect to the probability measure $P$.

### 2.2. Distortion Models

Our focus is on a family of imprecise probability models usually referred to as *distortion models* [4, 6, 11]. 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 [12, 22, 25]), or making a transformation of a given (lower) probability (as in [5, 7, 23]). We showed [16, Prop.3.2] that the latter approach can be embedded in the former, hence we will only focus on that one. Given a distorting function $d : \mathbb{P}(\mathscr{X}) \times \mathbb{P}(\mathscr{X}) \to [0,\infty)$, a distortion parameter $\delta > 0$ and a fixed probability measure $P_0 \in \mathbb{P}(\mathscr{X})$, we can define the set of probabilities:

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

Whenever $d$ is convex and continuous, $B_d^\delta(P_0)$ is a convex and closed set of probabilities [16, Prop.3.1]. This means that if we consider its lower envelope:

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

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

In [16, 17] we assumed that $P_0 \in \mathbb{P}^*(\mathscr{X})$, i.e. $P_0$ is strictly positive for every non-empty event, and also that $\delta$ is small enough such that $B_d^\delta(P_0) \subseteq \mathbb{P}^*(\mathscr{X})$. In this paper, we shall also assume that this simplifying hypothesis holds throughout, and will only recall it when it is necessary. See [17, Appendix 2] for some additional comments.

### 2.3. Processing Imprecise Probabilistic Models

The variety of distortion models makes it necessary to have tools at our disposal that allow to select the best one for each scenario. In this sense, one desirable property is that the model is closed under a number of operations of interest. The ones analysed in this paper are introduced next.

**Merging** The first operation we shall consider is *merging*. By this, we will refer to the procedure where we aggregate a number of belief models, defined on $\mathscr{X}$, into a unified one. These models may arise from the opinion of different experts or from the use of several data sources, for instance. We refer to [19, 20, 26] for relevant works on this topic.

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 $\mathscr{M}_1, \mathscr{M}_2$, they will produce the sets $\mathscr{M}_1 \cap \mathscr{M}_2$, $\mathscr{M}_1 \cup \mathscr{M}_2$ and $\varepsilon\mathscr{M}_1 + (1-\varepsilon)\mathscr{M}_2 = \{\varepsilon P_1 + (1-\varepsilon)P_2 | P_i \in \mathscr{M}_i\}$ with $\varepsilon \in [0,1]$, respectively.

In terms of the lower probabilities associated with these sets, it should be noted that, while $\mathscr{M}_1 \cup \mathscr{M}_2$ is not convex in general, its lower envelope, that coincides with the lower envelope of its convex hull $ch(\mathscr{M}_1 \cup \mathscr{M}_2)$, is given by $\underline{P} := \min\{\underline{P}_1, \underline{P}_2\}$, where $\underline{P}_1, \underline{P}_2$ denote the lower envelopes of $\mathscr{M}_1, \mathscr{M}_2$, respectively.

In contrast, while $\mathscr{M}_1 \cap \mathscr{M}_2$ is convex, its lower envelope $\underline{P}$ will dominate in general $\max\{\underline{P}_1, \underline{P}_2\}$. A sufficient condition for the equality is precisely the convexity of $\mathscr{M}_1 \cup \mathscr{M}_2$, as shown in [28, Thm.6].

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

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

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

***Independent products*** Conversely, we may start from two marginal models $\underline{P}^{\mathscr{X}}$ and $\underline{P}^{\mathscr{Y}}$ on domains $\mathscr{X}$ and $\mathscr{Y}$, respectively, and build a joint model on $\mathscr{X} \times \mathscr{Y}$ that is compatible with them. When the sources are assumed to be independent, this leads us to consider an *independent product*. Among the many possible choices [8], we consider here the *strong product* of $\underline{P}^{\mathscr{X}}$ and $\underline{P}^{\mathscr{Y}}$, that we shall denote $\underline{P}^{\mathscr{X}} \boxtimes \underline{P}^{\mathscr{Y}}$. It is the lower envelope of the credal set

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

where $P^{\mathscr{X}} \times P^{\mathscr{Y}}$ is the probability obtained from the marginals $P^{\mathscr{X}}$ and $P^{\mathscr{Y}}$ by stochastic independence. The strong product $\underline{P}^{\mathscr{X}} \boxtimes \underline{P}^{\mathscr{Y}}$ and its conjugate $\overline{P}^{\mathscr{X}} \boxtimes \overline{P}^{\mathscr{Y}}$ satisfy the following properties for every $A \subseteq \mathscr{X}, B \subseteq \mathscr{Y}$:

$$\begin{aligned} \underline{P}^{\mathscr{X}} \boxtimes \underline{P}^{\mathscr{Y}}(A \times B) &= \underline{P}^{\mathscr{X}}(A) \cdot \underline{P}^{\mathscr{Y}}(B) \\ \overline{P}^{\mathscr{X}} \boxtimes \overline{P}^{\mathscr{Y}}(A \times B) &= \overline{P}^{\mathscr{X}}(A) \cdot \overline{P}^{\mathscr{Y}}(B). \end{aligned} \quad (1)$$

***Natural extension of marginal models*** We may also consider the most conservative joint model on $\mathscr{X} \times \mathscr{Y}$ with marginals $\underline{P}^{\mathscr{X}}$ and $\underline{P}^{\mathscr{Y}}$, imposing no dependence assumption whatsoever. This corresponds to *natural extension* [13, 27] $\underline{E}$ of the coherent lower probability $\underline{P}$ that is defined on $\{A \times \mathscr{Y} : A \subseteq \mathscr{X}\} \cup \{\mathscr{X} \times B : B \subseteq \mathscr{Y}\}$ by $\underline{P}(A \times \mathscr{Y}) = \underline{P}^{\mathscr{X}}(A)$ and $\underline{P}(\mathscr{X} \times B) = \underline{P}^{\mathscr{Y}}(B)$. It can be equivalently obtained as the lower envelope of the credal set $\mathscr{E}(\underline{P}^{\mathscr{X}}, \underline{P}^{\mathscr{Y}})$ given by those probabilities whose marginals are compatible with $\underline{P}^{\mathscr{X}}$ and $\underline{P}^{\mathscr{Y}}$:

$$\left\{P \in \mathbb{P}(\mathscr{X} \times \mathscr{Y}) \mid P^{\mathscr{X}} \in \mathscr{M}\left(\underline{P}^{\mathscr{X}}\right),\ P^{\mathscr{Y}} \in \mathscr{M}\left(\underline{P}^{\mathscr{Y}}\right)\right\}. \quad (2)$$

The associated coherent lower and upper probabilities on events $C \subseteq \mathscr{X} \times \mathscr{Y}$ are

$$\underline{E}(C) = \inf_{P \in \mathscr{E}(\underline{P}^{\mathscr{X}}, \underline{P}^{\mathscr{Y}})} P(C), \quad \overline{E}(C) = \sup_{P \in \mathscr{E}(\underline{P}^{\mathscr{X}}, \underline{P}^{\mathscr{Y}})} P(C). \quad (3)$$

Our next proposition gives the expression of $\underline{E}, \overline{E}$ on Cartesian products of events:

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

$$\underline{E}(A \times B) = \max\left\{\underline{P}^{\mathscr{X}}(A) + \underline{P}^{\mathscr{Y}}(B) - 1, 0\right\}, \quad (4)$$

$$\overline{E}(A \times B) = \min\left\{\overline{P}^{\mathscr{X}}(A), \overline{P}^{\mathscr{Y}}(B)\right\}. \quad (5)$$

### 2.4. Aim of the Paper

Our goal in this paper is to complement the analysis performed in [17, Sec.5] by investigating the behaviour of different families of distortion models (pari mutuel, linear vacuous, constant odds ratio, total variation) under the procedures described in Sec. 2.3. Specifically, we shall tackle the following problems:

***Merging*** We first consider two distortion models $B_d^{\delta_1}(P_0^1)$ and $B_d^{\delta_2}(P_0^2)$ in some specific family. We analyse whether their conjunction $B_d^{\delta_1}(P_0^1) \cap B_d^{\delta_2}(P_0^2)$, their disjunction $B_d^{\delta_1}(P_0^1) \cup B_d^{\delta_2}(P_0^2)$ or their mixture $\varepsilon B_d^{\delta_1}(P_0^1) + (1-\varepsilon)B_d^{\delta_2}(P_0^2)$ belong to the same family, in the sense that it is equal to $B_d^{\delta^*}(P_0^*)$ for some appropriate $\delta^*$ and $P_0^*$.

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

***Independent products*** Consider two distortion models $B_d^{\delta}(P_0^{\mathscr{X}})$ and $B_d^{\delta}(P_0^{\mathscr{Y}})$ with the same distortion parameter, and an assumption of independence. We want to know whether the joint model that gathers this information belongs to the same family. We may consider two approaches for determining this joint model:

- Combine $P_0^{\mathscr{X}}$ and $P_0^{\mathscr{Y}}$ into a joint and distort it. In this way, we obtain the distortion model $B_d^{\delta}(P_0^{\mathscr{X},\mathscr{Y}})$. We shall denote by $\underline{P}^{\mathscr{X}\times\mathscr{Y}}$ and $\overline{P}^{\mathscr{X}\times\mathscr{Y}}$ the resulting lower and upper probabilities.

- Consider the distortion models $B_d^{\delta}(P_0^{\mathscr{X}})$ and $B_d^{\delta}(P_0^{\mathscr{Y}})$ and combine them using the strong product, leading to the credal set $\mathscr{M}(\underline{P}^{\mathscr{X}}) \boxtimes \mathscr{M}(\underline{P}^{\mathscr{Y}}) \subseteq \mathbb{P}^*(\mathscr{X} \times \mathscr{Y})$ with associated lower and upper probabilities $\underline{P}^{\mathscr{X}} \boxtimes \underline{P}^{\mathscr{Y}}$ and $\overline{P}^{\mathscr{X}} \boxtimes \overline{P}^{\mathscr{Y}}$.

We wonder whether the credal sets $B_d^{\delta}(P_0^{\mathscr{X},\mathscr{Y}})$ and the convex hull of $\mathscr{M}(\underline{P}^{\mathscr{X}}) \boxtimes \mathscr{M}(\underline{P}^{\mathscr{Y}})$ coincide, or in case they do not, if there is an inclusion relationship between them.

***Natural extension*** We consider two marginal distortion models $B_d^{\delta_{\mathscr{X}}}(P_0^{\mathscr{X}})$ and $B_d^{\delta_{\mathscr{Y}}}(P_0^{\mathscr{Y}})$, and wonder whether we can give a simple expression of $\mathscr{E}(\underline{P}^{\mathscr{X}}, \underline{P}^{\mathscr{Y}})$, $\underline{E}$ and $\overline{E}$ and also whether $\mathscr{E}(\underline{P}^{\mathscr{X}}, \underline{P}^{\mathscr{Y}})$ is also a distortion model of the same family. For the sake of simplicity, in this part we assume that $\delta_{\mathscr{X}} = \delta_{\mathscr{Y}} := \delta$.

Note that, although $B_d^{\delta}(P_0^{\mathscr{X}})$ and $B_d^{\delta}(P_0^{\mathscr{Y}})$ are included in $\mathbb{P}^*(\mathscr{X})$ and $\mathbb{P}^*(\mathscr{Y})$, respectively, we cannot guarantee that $\mathscr{E}(\underline{P}^{\mathscr{X}}, \underline{P}^{\mathscr{Y}})$ is included in $\mathbb{P}^*(\mathscr{X} \times \mathscr{Y})$.

In the sections that follow, we consider a number of distortion models and analyse their behaviour under the previous operations.

## 3. Pari Mutuel Model

The first model 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:*

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

Since by assumption $P_0 \in \mathbb{P}^*(\mathscr{X})$ and $\underline{P}_{PMM}(A) > 0$ for all $A \neq \emptyset$, the previous expressions simplify to:

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

for every $A \neq \emptyset, \mathscr{X}$, and taking the trivial values 0 and 1 for $\emptyset$ and $\mathscr{X}$, respectively.

The pari mutuel model is equivalent [16, Thm.4.1] to the credal set $B_{d_{PMM}}^{\delta}(P_0)$ where $d_{PMM} : \mathbb{P}^*(\mathscr{X}) \times \mathbb{P}^*(\mathscr{X}) \to [0, \infty)$ is the distorting function given by

$$d_{PMM}(P, P_0) = \max_{A \subset \mathscr{X}} \frac{P_0(A) - P(A)}{1 - P_0(A)}.$$

### 3.1. Merging

Let us first study how the PMM behaves under merging.

***Conjunction*** Given two models $B_{d_{PMM}}^{\delta_1}(P_0^1)$ and $B_{d_{PMM}}^{\delta_2}(P_0^2)$, it was established in [15, Prop.12] that their intersection is non-empty iff

$$\sum_{x \in \mathscr{X}} \min\left\{(1+\delta_1)P_0^1(\{x\}), (1+\delta_2)P_0^2(\{x\}), 1\right\} \geq 1.$$

In that case, the intersection is given by the model $B_{d_{PMM}}^{\delta_\cap}(P_0^\cap)$, where

$$\delta^\cap = \left(\sum_{x \in \mathscr{X}} \min\left\{(1+\delta_1)P_0^1(\{x\}), (1+\delta_2)P_0^2(\{x\})\right\}\right) - 1,$$

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

***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 1** *Consider $P^1_0 = (0.5, 0.3, 0.2), P^2_0 = (0.3, 0.5, 0.2)$ and $\delta_1 = \delta_2 = 0.1$. Then 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 $\mathscr{M}(\underline{P}^{\cup}) = B^{\delta}_{d_{PMM}}(P_0)$ for some $P_0, \delta$, then we would obtain*

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

*on the other hand, the equalities*

$$0.45 = \underline{P}^{\cup}(\{x_2,x_3\}) = (1+\delta)P_0(\{x_2,x_3\}) - \delta$$
$$0.35 = \underline{P}^{\cup}(\{x_2\}) + \underline{P}^{\cup}(\{x_3\}) = (1+\delta)P_0(\{x_2,x_3\}) - 2\delta$$

*mean 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 (see [14, Prop.7]). It is given by the model $B^{\delta_{\cup}}_{d_{PMM}}(P^{\cup}_0)$ such that:

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

This is the most informative PMM including the credal set $B^{\delta_1}_{d_{PMM}}(P^1_0) \cup B^{\delta_2}_{d_{PMM}}(P^2_0)$.

***Convex mixture*** The mixture operation was studied in [15, Sec.6.1], where it was shown that the convex mixture of two PMMs is again a PMM $B^{\delta_\varepsilon}_{d_{PMM}}(P^{\varepsilon}_0)$, where $1+\delta_\varepsilon = \varepsilon(1+\delta_1) + (1-\varepsilon)(1+\delta_2)$ and $\forall x \in \mathscr{X}$:

$$P^{\varepsilon}_0(\{x\}) = \frac{\varepsilon(1+\delta_1)P^1_0(\{x\}) + (1-\varepsilon)(1+\delta_2)P^2_0(\{x\})}{1+\delta_\varepsilon}.$$

### 3.2. Multivariate Setting

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

***Marginalisation*** In [15, Sec.6.2], it was shown that the marginal lower probability $\underline{P}^{\mathscr{X}}$ obtained from a joint PMM $B^{\delta}_{d_{PMM}}(P^{\mathscr{X},\mathscr{Y}}_0)$ is again a PMM $B^{\delta}_{d_{PMM}}(P^{\mathscr{X}}_0)$ with $P^{\mathscr{X}}_0$ the marginal probability of $P^{\mathscr{X},\mathscr{Y}}_0$ on $\mathscr{X}$ and the same distortion factor.

***Independent products*** When building a joint model from marginal ones $P^{\mathscr{X}}_0$ and $P^{\mathscr{Y}}_0$ under the assumption of independence, it can be seen that there is no dominance relationship between $\underline{P}^{\mathscr{X}\times\mathscr{Y}}_{PMM}$ (combine through stochastic independence and then distort) and $\underline{P}^{\mathscr{X}}_{PMM} \boxtimes \underline{P}^{\mathscr{Y}}_{PMM}$ (distort then combine through strong independence). To see this, note that on the one hand for the Cartesian product of events $A\times B$, it holds that:

$$\overline{P}^{\mathscr{X}\times\mathscr{Y}}_{PMM}(A\times B) = (1+\delta)P^{\mathscr{X}}_0(A)P^{\mathscr{Y}}_0(B) \leq$$
$$(1+\delta)P^{\mathscr{X}}_0(A)(1+\delta)P^{\mathscr{Y}}_0(B) = \overline{P}^{\mathscr{X}}_{PMM} \boxtimes \overline{P}^{\mathscr{Y}}_{PMM}(A\times B),$$

where the inequality is strict whenever we consider non-trivial events $A, B$, i.e. $P^{\mathscr{X}}_0(A), P^{\mathscr{Y}}_0(B) \in (0,1)$. On the other hand, for events $E$ that are not products, the relationship between $\overline{P}^{\mathscr{X}\times\mathscr{Y}}_{PMM}(E)$ and $\overline{P}^{\mathscr{X}}_{PMM} \boxtimes \overline{P}^{\mathscr{Y}}_{PMM}(E)$ can be the reverse one, as we show in the next example:

**Example 2** *Let $\mathscr{X} = \{x_1, x_2\}$, $\mathscr{Y} = \{y_1, y_2\}$, the probability measures $P^{\mathscr{X}}_0$ and $P^{\mathscr{Y}}_0$ given by: $P^{\mathscr{X}}_0(\{x_1\}) = 0.3, P^{\mathscr{X}}_0(\{x_2\}) = 0.7, P^{\mathscr{Y}}_0(\{y_1\}) = P^{\mathscr{Y}}_0(\{y_2\}) = 0.5$ and let $\delta = 0.1$. Given $E_1 = \{(x_2,y_2)\}^c$, it holds that:*

$$\overline{P}^{\mathscr{X}\times\mathscr{Y}}_{PMM}(E_1) = 0.715 > \overline{P}^{\mathscr{X}}_{PMM} \boxtimes \overline{P}^{\mathscr{Y}}_{PMM}(E_1) = 0.6985.$$

*Therefore, there is not a dominance relationship between $\overline{P}^{\mathscr{X}\times\mathscr{Y}}_{PMM}$ and $\overline{P}^{\mathscr{X}}_{PMM} \boxtimes \overline{P}^{\mathscr{Y}}_{PMM}$.* ♦

***Natural extension of marginal models*** Consider the lower and upper probabilities that are the lower and upper envelopes of $B^{\delta}_{d_{PMM}}(P^{\mathscr{X}}_0)$ and $B^{\delta}_{d_{PMM}}(P^{\mathscr{Y}}_0)$. Using Eqs. (4) and (5), we obtain

$$\overline{E}_{PMM}(A\times B) = \min\left\{1, (1+\delta)\min\left\{P^{\mathscr{X}}_0(A), P^{\mathscr{Y}}_0(B)\right\}\right\}, \quad (6)$$
$$\underline{E}_{PMM}(A\times B) = \max\left\{(1+\delta)\left(P^{\mathscr{X}}_0(A) + P^{\mathscr{Y}}_0(B) - 1\right) - \delta, 0\right\}. \quad (7)$$

These are similar to the expressions of Def. 2. Even if Eqs. (6) and (7) are only valid for events of the type $A\times B$, one may think that the natural extension is related to a PMM. Our next result shows that such a connection exists.

**Theorem 3** *Let $B^{\delta}_{d_{PMM}}(P^{\mathscr{X}}_0)$ and $B^{\delta}_{d_{PMM}}(P^{\mathscr{Y}}_0)$ be two PMM with associated lower probabilities $\underline{P}^{\mathscr{X}}_{PMM}$ and $\underline{P}^{\mathscr{Y}}_{PMM}$. Then, the credal set of the natural extension defined in Eq. (2) can be expressed as:*

$$\mathscr{E}\left(\underline{P}^{\mathscr{X}}_{PMM}, \underline{P}^{\mathscr{Y}}_{PMM}\right) = \left\{P \in \mathbb{P}(\mathscr{X}\times\mathscr{Y}) \mid P \leq (1+\delta)\overline{E}_{P^{\mathscr{X}}_0, P^{\mathscr{Y}}_0}\right\};$$

*equivalently, for every $C \subseteq \mathscr{X}\times\mathscr{Y}$,*

$$\overline{E}_{PMM}(C) = \min\left\{(1+\delta)\overline{E}_{P^{\mathscr{X}}_0, P^{\mathscr{Y}}_0}(C), 1\right\},$$

*where $\overline{E}_{P^{\mathscr{X}}_0, P^{\mathscr{Y}}_0}$ corresponds to the upper envelope of the credal set in Eq. (2) applied to the particular case of precise marginals $P^{\mathscr{X}}_0, P^{\mathscr{Y}}_0$.*

This result shows that the procedures of natural extension and the distortion produced by the PMM commute, in the sense illustrated in Fig. 1.

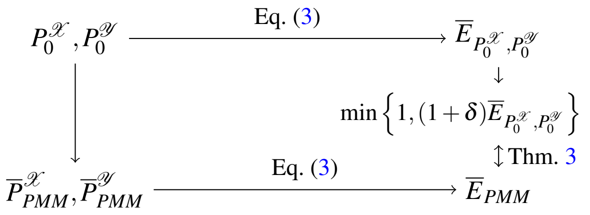


Figure 1: Graphical representation of the computation of the natural extension of two PMMs.

## 4. Linear Vacuous Mixtures

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

**Definition 4** *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 $\forall A \neq \emptyset, \mathcal{X}$:*

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

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

This model, studied in [27] and [16, Sec.5], has been used for instance in robust statistics [11]. The credal set $\mathcal{M}(\underline{P}_{LV})$ coincides [16, Thm.5.1] 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 [16, Thm.5.1]:

$$d_{LV}(P,P_0) = \max_{A \neq \emptyset} \frac{P_0(A) - P(A)}{P_0(A)}.$$

Let us analyse the behaviour of the LV model under the different operations introduced in Sec. 2.3.

### 4.1. Merging

We first look at the behaviour of LV models under merging.

***Conjunction*** Similarly to the PMM, the intersection of two LV models (when non-empty) is again a LV model.

**Proposition 5** *Given two distortion models $B^{\delta_1}_{d_{LV}}(P_0^1)$ and $B^{\delta_2}_{d_{LV}}(P_0^2)$, the set $B^{\delta_1}_{d_{LV}}(P_0^1) \cap B^{\delta_2}_{d_{LV}}(P_0^2)$ is non-empty iff*

$$\sum_{x \in \mathcal{X}} \max\left\{(1-\delta_1)P_0^1(\{x\}), (1-\delta_2)P_0^2(\{x\})\right\} \leq 1.$$

*In that case, this conjunction is the LV model generated by*

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

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

***Disjunction*** Regarding the disjunction, the convex hull of $B^{\delta_1}_{d_{LV}}(P_0^1) \cup B^{\delta_2}_{d_{LV}}(P_0^2)$ will in general not be a LV model, not even when $\delta_1 = \delta_2$ as we show in the next example.

**Example 3** *Consider the same probabilities and distortion factors as in Ex. 1. The associated LV models $\underline{P}_{LV_1}, \underline{P}_{LV_2}$ and their disjunction $\underline{P}^{\cup}_{LV}$ are:*

| | $\{x_1\}$ | $\{x_2\}$ | $\{x_3\}$ | $\{x_1,x_2\}$ | $\{x_1,x_3\}$ | $\{x_2,x_3\}$ |
|---|---|---|---|---|---|---|
| $\underline{P}_{LV_1}$ | 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 it would be*

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

*which does not hold. As a consequence, the disjunction $B^{\delta_1}_{d_{LV}}(P_0^1) \cup B^{\delta_2}_{d_{LV}}(P_0^2)$ does not produce a LV model.* ♦

This disjunction has a unique undominated LV outer approximation, since by [14, Prop.8] this holds for any given credal set. It is given by the model $B^{\delta_{\cup}}_{d_{LV}}(P_0^{\cup})$ where

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

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

***Convex mixture*** The mixture of two LV models, that is, the credal set $B^{\delta_\varepsilon}_{d_{LV}}(P_0^{\varepsilon})$ for a given $\varepsilon \in (0,1)$ can be established through a reasoning similar to the one made for the PMM in [14, Sec.5.1]. In particular, using in a straightforward way results established in [19] for probability intervals, $B^{\delta_\varepsilon}_{d_{LV}}(P_0^{\varepsilon})$ is described by the constraints

$$\varepsilon(1-\delta_1)P_0^1(\{x\}) + (1-\varepsilon)(1-\delta_2)P_0^2(\{x\}) \leq P(\{x\}) \quad \forall x \in \mathcal{X}.$$

We deduce that $1-\delta_\varepsilon = \varepsilon(1-\delta_1) + (1-\varepsilon)(1-\delta_2)$ and

$$P_0^{\varepsilon}(\{x\}) = \frac{\varepsilon(1-\delta_1)P_0^1(\{x\}) + (1-\varepsilon)(1-\delta_2)P_0^2(\{x\})}{1-\delta_\varepsilon}.$$

### 4.2. Multivariate Setting

Let us now look at the multivariate setting.

***Marginalisation*** It is easy to prove that the marginal model of a joint LV is again a LV model:

**Proposition 6** *Consider the distortion model $B^{\delta}_{d_{LV}}\left(P_0^{\mathcal{X},\mathcal{Y}}\right)$ and its induced lower prevision $\underline{P}_{LV}$. Then, the marginal model $\underline{P}^{\mathcal{X}}_{LV}$ induces the model $B^{\delta}_{d_{LV}}\left(P_0^{\mathcal{X}}\right)$ with $P_0^{\mathcal{X}}$ the marginal of $P_0^{\mathcal{X},\mathcal{Y}}$ on $\mathcal{X}$.*

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

$$\underline{P}_{LV}^{\mathscr{X}\times\mathscr{Y}}(A\times B) = (1-\delta)P_0^{\mathscr{X}}(A)P_0^{\mathscr{Y}}(B) \geq (1-\delta)P_0^{\mathscr{X}}(A)(1-\delta)P_0^{\mathscr{Y}}(B) = \underline{P}_{LV}^{\mathscr{X}} \boxtimes \underline{P}_{LV}^{\mathscr{Y}}(A\times B),$$

where last equality follows from the factorization property in Eq. (1). We may then wonder if $\underline{P}_{LV}^{\mathscr{X}\times\mathscr{Y}} \geq \underline{P}_{LV}^{\mathscr{X}} \boxtimes \underline{P}_{LV}^{\mathscr{Y}}$ in general. The next example shows that this is not the case, hence that we have no dominance relation between the joint models $\underline{P}_{LV}^{\mathscr{X}\times\mathscr{Y}}$ and $\underline{P}_{LV}^{\mathscr{X}} \boxtimes \underline{P}_{LV}^{\mathscr{Y}}$.

**Example 4** *Let us continue with Ex. 2. Given the event $E_1 = \{(x_2,y_2)\}^c$, we obtain*

$$\underline{P}_{LV}^{\mathscr{X}\times\mathscr{Y}}(E_1) = 0.585 < \underline{P}_{LV}^{\mathscr{X}} \boxtimes \underline{P}_{LV}^{\mathscr{Y}}(E_1) = 0.5985,$$

*and therefore $B_{d_{LV}}^{\delta}\left(P_0^{\mathscr{X},\mathscr{Y}}\right)$ is not included in the convex hull of $\mathscr{M}\left(\underline{P}_{LV}^{\mathscr{X}}\right) \boxtimes \mathscr{M}\left(\underline{P}_{LV}^{\mathscr{Y}}\right)$.* ♦

***Natural extension of marginal models*** Using Eqs. (4) and (5) when $\delta_{\mathscr{X}} = \delta_{\mathscr{Y}} = \delta$, we get

$$\underline{E}_{LV}(A\times B) = (1-\delta)\max\left\{P_0^{\mathscr{X}}(A) + P_0^{\mathscr{Y}}(B) - \frac{1}{1-\delta}, 0\right\}, \quad (8)$$

$$\overline{E}_{LV}(A\times B) = (1-\delta)\min\left\{P_0^{\mathscr{X}}(A), P_0^{\mathscr{Y}}(B)\right\} + \delta. \quad (9)$$

The expressions in Eqs. (8) and (9) 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 $\underline{E}_{LV} = (1-\delta)\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}$ does not hold:

**Example 5** *Consider the same spaces and probability measures as in Ex. 2, and take $\delta = 0.2$. Then Eq. (8) gives*

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

*while $\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(\{x_2\}\times\{y_2\}) = \max\{0.7+0.5-1, 0\} = 0.2$, meaning that $(1-\delta)\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(\{x_2\}\times\{y_2\}) = 0.16$.* ♦

## 5. Constant Odds Ratios

We next consider the constant odds ratio model (COR, for short):

**Definition 7** *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 of the equation:*

$$(1-\delta)P_0\big((f-\underline{P}_{COR}(f))^+\big) = P_0\big((f-P_0(f))^-\big), \quad (10)$$

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

While Eq. (10) does not have a explicit expression, the restriction to (indicators of) events of the constant odds ratio can be more easily computed as:

$$\underline{P}_{COR}(A) = \frac{(1-\delta)P_0(A)}{1-\delta P_0(A)} \quad \forall A \subseteq \mathscr{X}. \quad (11)$$

The constant odds ratio was given a behavioural interpretation in [27, Sec.2.9.4]. We refer to [2, 3, 22, 25] for some applications of this model, and to [16, Sec.6] for a detailed study. When $P_0 \in \mathbb{P}^*(\mathscr{X})$ and $\delta$ is small enough, the credal set $\mathscr{M}(\underline{P}_{COR})$ coincides with [16, Thm.6.1] $B_{d_{COR}}^{\delta}(P_0)$, where $d_{COR} : \mathbb{P}^*(\mathscr{X}) \times \mathbb{P}^*(\mathscr{X}) \to [0,\infty)$ is the distorting function given by [16, Thm.6.1]:

$$d_{COR}(P,P_0) = \max_{A,B\neq\emptyset}\left\{1 - \frac{P(A)\cdot P_0(B)}{P(B)\cdot P_0(A)}\right\}.$$

Also, the credal set $\mathscr{M}(\underline{P}_{COR})$ can be expressed as [27, Sec.3.3.5]:

$$\mathscr{M}(\underline{P}_{COR}) = \left\{P \in \mathbb{P}(\mathscr{X}) \mid \frac{P(A)P_0(B)}{P_0(A)P(B)} \geq (1-\delta) \;\; \forall A,B \subseteq \mathscr{X}\right\}. \quad (12)$$

### 5.1. Merging

***Conjunction*** Unlike the PMM and LV models, the intersection of two COR models cannot be expected to be a COR model in general, as next example shows.

**Example 6** *Consider the model $\mathscr{M}_1 = B_{d_{COR}}^{\delta_1}(P_0^1)$ with $P_0^1 = (0.5, 0.3, 0.2)$ and $\delta_1 = 0.2$, and $\mathscr{M}_2 = B_{d_{COR}}^{\delta_2}(P_0^2)$ such that $P_0^2 = (0.35, 0.3, 0.35)$ with $\delta_2 = 0.5$. From Eq. (12), 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 $\mathscr{M}_1$ and $\mathscr{M}_2$. From this, it follows that any $P \in \mathscr{M}_1 \cap \mathscr{M}_2$ must satisfy the constraint $\frac{P(\{x_1\})}{P(\{x_3\})} = 2$. As a consequence, the credal set $\mathscr{M}_1 \cap \mathscr{M}_2$ has at most two extreme points. Since in [16, Prop.6.2] it was proved that a COR model has $2^n - 2$ extreme points, where $n$ is the cardinality of $\mathscr{X}$, it follows that $\mathscr{M}_1 \cap \mathscr{M}_2$ is not a COR model, as it has less than $2^n - 2 = 6$ extreme points.* ♦

***Disjunction*** Similarly, the disjunction of two COR models will not produce a COR model in general, not even when $\delta_1 = \delta_2$, as we show in our next example.

**Example 7** *Consider $P_0^1 = (0.4, 0.3, 0.3)$, $P_0^2 = (0.3, 0.4, 0.3)$ and $\delta_1 = \delta_2 = 0.1$. Using Eq. (11), the associated COR models $\underline{P}_{COR_1}, \underline{P}_{COR_2}$ and their disjunction $\underline{P}^{\cup}$ are given in the following table:*

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

*recalling also that* $\underline{P}^{\cup} = \min\{\underline{P}_{COR_1}, \underline{P}_{COR_2}\}$. *If* $\underline{P}^{\cup}$ *was a COR model, i.e.* $\mathscr{M}(\underline{P}^{\cup}) = B^{\delta}_{d_{COR}}(P_0)$ *for some* $P_0$ *and* $\delta$, *since* $\underline{P}^{\cup}(\{x_1\}) = \underline{P}^{\cup}(\{x_2\}) = \underline{P}^{\cup}(\{x_3\})$, *it must hold that* $P_0(\{x_1\}) = P_0(\{x_2\}) = P_0(\{x_3\}) = \frac{1}{3}$. *But in that case, regardless of the value of* $\delta$, $\underline{P}^{\cup}$ *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. For instance, both COR models induced by* $P_A = (31/80, 31/80, 18/80)$, $P_B = (35/124, 35/124, 27/62)$, $\delta_A = 121/310$ *and* $\delta_B = \frac{1}{2}$ *outer approximate* $\underline{P}^{\cup}$, *and it can be checked that no COR model is both included in* $B^{\delta_A}_{d_{COR}}(P_A)$ *and* $B^{\delta_B}_{d_{COR}}(P_B)$ *and outer approximates* $\underline{P}^{\cup}$. ♦

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 $B^{\delta_1}_{d_{COR}}(P_0^1)$ and $B^{\delta_2}_{d_{COR}}(P_0^2)$ are described by the same set of constraints over $P(A)/P(B)$, their convex mixture is a credal set described by the constraints

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

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

**Example 8** *Consider* $P_0^1 = (1/4, 1/4, 1/2), P_0^2 = (1/2, 1/4, 1/4)$ *and* $\delta_1 = \delta_2 = 0.5$. *Using Eq. (11), the associated COR models* $\underline{P}_{COR_1}, \underline{P}_{COR_2}$ *and their average* $\underline{P}^{0.5}$ *are:*

| | $\{x_1\}$ | $\{x_2\}$ | $\{x_3\}$ | $\{x_1,x_2\}$ | $\{x_1,x_3\}$ | $\{x_2,x_3\}$ |
|---|---|---|---|---|---|---|
| $\underline{P}_{COR_1}$ | 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* $\underline{P}^{0.5}$ *be the lower probability of a COR model* $B^{\delta_{0.5}}_{d_{COR}}(P_0^{0.5})$, *it would be* $P_0^{0.5}(\{x_1\}) = P_0^{0.5}(\{x_3\}) = p$, *hence* $P_0^{0.5}(\{x_2\}) = 1-2p$. *Using this observation and Eq. (11) on events* $\{x_1\}$ *and* $\{x_1,x_3\}$, *we should have* $\delta_{0.5} = 13/28$ *and* $p = 7/19$, *and applying again Eq. (11) with these values on* $\{x_2\}$ *would give* $\underline{P}(\{x_2\}) = 75/467$, *which is close but not equal to the value* $1/7$ *reported in the table above.* ♦

### 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 8** *Consider the distortion model* $B^{\delta}_{d_{COR}}(P_0^{\mathscr{X},\mathscr{Y}})$ *and its induced lower prevision* $\underline{P}_{COR}$. *Then, the marginal model* $\underline{P}^{\mathscr{X}}_{COR}$ *induces the credal set* $B^{\delta}_{d_{COR}}(P_0^{\mathscr{X}})$ *with* $P_0^{\mathscr{X}}$ *the marginal of* $P_0^{\mathscr{X},\mathscr{Y}}$ *on* $\mathscr{X}$.

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

$$\underline{P}^{\mathscr{X}\times\mathscr{Y}}_{COR}(A\times B) \geq \underline{P}^{\mathscr{X}}_{COR} \boxtimes \underline{P}^{\mathscr{Y}}_{COR}(A\times B).$$

We can then wonder if $\underline{P}^{\mathscr{X}\times\mathscr{Y}}_{COR}(C) \geq \underline{P}^{\mathscr{X}}_{COR} \boxtimes \underline{P}^{\mathscr{Y}}_{COR}(C)$ for any event $C \subseteq \mathscr{X}\times\mathscr{Y}$. The next example shows that this is not the case, and therefore that there is no dominance relation between $\underline{P}^{\mathscr{X}\times\mathscr{Y}}_{COR}$ and $\underline{P}^{\mathscr{X}}_{COR} \boxtimes \underline{P}^{\mathscr{Y}}_{COR}$.

**Example 9** *Consider our running Ex. 2. Given* $E_2 = \{(x_1,y_1),(x_2,y_2)\}$, *we obtain* $\underline{P}^{\mathscr{X}\times\mathscr{Y}}_{COR}(E_2) = 0.4737 < \underline{P}^{\mathscr{X}}_{COR} \boxtimes \underline{P}^{\mathscr{Y}}_{COR}(E_2) = 0.4883$. *Therefore, the convex hull of* $\mathscr{M}(\underline{P}^{\mathscr{X}}_{COR}) \boxtimes \mathscr{M}(\underline{P}^{\mathscr{Y}}_{COR})$ *is not included in* $B^{\delta}_{d_{COR}}(P_0^{\mathscr{X},\mathscr{Y}})$.

We conclude that both approaches for building a joint independent model are not related in general.

***Natural extension of marginal models*** We have already mentioned that there is not an explicit expression for the lower/upper prevision of the COR model in gambles (see Eq. (10)), and it can only be given for events (see Eq. (11)). This complicates 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.

Applying Eqs. (4) and (5) to the lower and upper envelopes of $B^{\delta}_{d_{COR}}(P_0^{\mathscr{X}})$ and $B^{\delta}_{d_{COR}}(P_0^{\mathscr{Y}})$, we get the following forms on Cartesian products $A\times B$, for $A \subseteq \mathscr{X}$ and $B \subseteq \mathscr{Y}$:

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

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

Although these expressions do not seem to resemble a COR model, we may wonder if, similarly to what happened with the PMM (see Thm. 3), the equality $\mathscr{E}(\underline{P}^{\mathscr{X}}_{COR}, \underline{P}^{\mathscr{Y}}_{COR}) = B_{d_{COR}}(\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}})$ holds. As we show next, this is not the case.

**Example 10** *Consider our running Ex. 2. Given* $E_3 = \{(x_2,y_2)\}$, *we obtain* $\underline{E}_{COR}(E_3) = \max\{0.6774 + 0.4737 - 1, 0\} = 0.1511$. *On the other hand, from Eq. (4) we have that* $\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(E_3) = 0.2$, *whence* $(1-\delta)\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(E_3)/1-\delta\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(E_3) = 0.1836$. *Thus, the two values do not coincide.* ♦

## 6. Total Variation Model

Given two probability measures $P, Q$, their total variation distance is given by

$$d_{TV}(P, P_0) = \max_{A\subseteq\mathscr{X}} |P(A) - P_0(A)|.$$

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

**Definition 9** *Given a probability measure $P_0$ and a distortion factor $\delta \in (0,1)$, the* total variation model *(TV, for short) is given by the following lower and upper probabilities:*

$$\underline{P}_{TV}(A) = \max\{P_0(A) - \delta, 0\} \text{ for every } A \neq \mathscr{X}$$
$$\overline{P}_{TV}(A) = \min\{P_0(A) + \delta, 1\} \text{ for every } A \neq \emptyset,$$

*and the trivial values $\underline{P}_{TV}(\mathscr{X}) = 1$ and $\overline{P}_{TV}(\emptyset) = 0$.*

Since we are assuming that $P_0 \in \mathbb{P}^*(\mathscr{X})$ and that $\delta$ is small enough so that $B^{\delta}_{d_{TV}}(P_0) \subseteq \mathbb{P}^*(\mathscr{X})$, the above equations simplify for every $A \neq \emptyset, \mathscr{X}$ to:

$$\underline{P}_{TV}(A) = P_0(A) - \delta, \quad \overline{P}_{TV}(A) = P_0(A) + \delta. \qquad (13)$$

We refer to [10], [17, Sec.2], [21, Sec.3.2] and [27, Sec.3.2.4] for some works on this distortion model.

### 6.1. Merging

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

**Example 11** *From Eq. (13), the lower probability of a TV-model satisfies, for any event $A$ such that $0 < \underline{P}_{TV}(A) \leq \overline{P}_{TV}(A) < 1$, 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 $A \neq \emptyset, \mathscr{X}$. Let us use this to derive that the family of TV models is not closed under conjunction or disjunction.*

*Let $B^{\delta_1}_{d_{TV}}(P^1_0)$ be induced by $P^1_0 = (0.41, 0.37, 0.22)$ and $\delta_1 = 0.12$, and $B^{\delta_2}_{d_{TV}}(P^2_0)$ be determined by $P^2_0 = (0.37, 0.41, 0.22)$ 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 in the following table:*

| | $\{x_1\}$ | $\{x_2\}$ | $\{x_3\}$ | $\{x_1,x_2\}$ | $\{x_1,x_3\}$ | $\{x_2,x_3\}$ |
|---|---|---|---|---|---|---|
| $\underline{P}_{TV_1}$ | 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 $\max\{\underline{P}_{TV_1}, \underline{P}_{TV_2}\}$ is coherent and therefore coincides with $\underline{P}^{\cap}$. We observe that*

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

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

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

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

**Example 12** *Consider the model $\underline{P}^{\cup}$ from the previous example, and let us consider the TV models $B^{\delta_A}_{d_{TV}}(P^A_0)$ and $B^{\delta_B}_{d_{TV}}(P^B_0)$, 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 | 0.34 | 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:*

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

*a contradiction.* ♦

***Convex mixture*** It is rather direct to check that the convex mixture of two TV models $B^{\delta_1}_{d_{TV}}(P^1_0)$ and $B^{\delta_2}_{d_{TV}}(P^2_0)$ is again a TV model, since $\varepsilon \underline{P}^1_{TV}(A) + (1-\varepsilon)\underline{P}^2_{TV}(A) = \varepsilon P^1_0(A) + (1-\varepsilon)P^2_0(A) - \varepsilon\delta_1 - (1-\varepsilon)\delta_2$ are lower probabilities induced by the TV model $B^{\delta_\varepsilon}_{d_{TV}}(P^\varepsilon_0)$ with

$$\delta_\varepsilon = \varepsilon\delta_1 + (1-\varepsilon)\delta_2 \quad \text{and}$$
$$P^\varepsilon_0(\{x\}) = \varepsilon P^1_0(\{x\}) + (1-\varepsilon)P^2_0(\{x\}).$$

### 6.2. Multivariate Setting

***Marginalisation*** It is not difficult to prove that the marginal of a joint TV model is again a TV model.

**Proposition 10** *Consider the distortion model $B^{\delta}_{d_{TV}}(P^{\mathscr{X},\mathscr{Y}}_0)$ and its associated lower prevision $\underline{P}_{TV}$. Then, the marginal model $\underline{P}^{\mathscr{X}}_{TV}$ induces the credal set $B^{\delta}_{d_{TV}}(P^{\mathscr{X}}_0)$ with $P^{\mathscr{X}}_0$ the marginal probability of $P^{\mathscr{X},\mathscr{Y}}_0$ on $\mathscr{X}$.*

***Independent products*** Consider now two marginal models $B^{\delta}_{d_{TV}}(P^{\mathscr{X}}_0)$ and $B^{\delta}_{d_{TV}}(P^{\mathscr{Y}}_0)$. On Cartesian products of events, we have

$$\underline{P}^{\mathscr{X}\times\mathscr{Y}}_{TV}(A \times B) = P^{\mathscr{X}}_0(A)P^{\mathscr{Y}}_0(B) - \delta \neq$$
$$\big(P^{\mathscr{X}}_0(A) - \delta\big)\big(P^{\mathscr{Y}}_0(B) - \delta\big) = \underline{P}^{\mathscr{X}}_{TV} \boxtimes \underline{P}^{\mathscr{Y}}_{TV}(A \times B)$$

where the last equality follows from the factorization property in Eq. (1). We show in our next example that this inequality can go both ways.

**Example 13** *Consider the distortion models $B^{\delta}_{d_{TV}}(P_0^{\mathscr{X}})$ and $B^{\delta}_{d_{TV}}(P_0^{\mathscr{Y}})$ and the distortion factor $\delta = 0.1$. On the one hand, assume that there are $A \subseteq \mathscr{X}$ and $B \subseteq \mathscr{Y}$ such that $P_0^{\mathscr{X}}(A) = P_0^{\mathscr{Y}}(B) = 0.5$. We obtain that:*

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

*On the other hand, if there are $A \subseteq \mathscr{X}$ and $B \subseteq \mathscr{Y}$ satisfying $P_0^{\mathscr{X}}(A) = P_0^{\mathscr{Y}}(B) = 0.6$, we obtain that:*

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

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

***Natural extension of marginal models*** Consider now two TV-models $B^{\delta}_{d_{TV}}(P_0^{\mathscr{X}})$ and $B^{\delta}_{d_{TV}}(P_0^{\mathscr{Y}})$. Using Eqs. (4) and (5), we obtain the following expressions on Cartesian products $A \times B$, for $A \subseteq \mathscr{X}$ and $B \subseteq \mathscr{Y}$:

$$\underline{E}_{TV}(A\times B) = \max\left\{P_0^{\mathscr{X}}(A) + P_0^{\mathscr{Y}}(B) - 1, 2\delta\right\} - 2\delta, \quad (14)$$
$$\overline{E}_{TV}(A\times B) = \min\left\{P_0^{\mathscr{X}}(A), P_0^{\mathscr{Y}}(B)\right\} + \delta. \quad (15)$$

The lower and upper natural extension have a similar form as a TV-model. However, the lower and upper bounds have different distortion parameters, respectively $2\delta$ and $\delta$. This suggests that, as with the LV and COR models, the natural extension of the TV model can neither be expressed as $\mathscr{E}(\underline{P}_{TV}^{\mathscr{X}}, \underline{P}_{TV}^{\mathscr{Y}}) = B^{\delta}_{d_{TV}}(\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}})$ nor as $\mathscr{E}(\underline{P}_{TV}^{\mathscr{X}}, \underline{P}_{TV}^{\mathscr{Y}}) = B^{2\delta}_{d_{TV}}(\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}})$:

**Example 14** *We consider the same possibility spaces $\mathscr{X}$ and $\mathscr{Y}$ and precise probabilities $P_0^{\mathscr{X}}$ and $P_0^{\mathscr{Y}}$ as in Ex. 2. If we consider for instance the event $\{(x_2,y_2)\}$, we deduce from Eqs. (14) and (4) that for every $\delta \in (0, 0.1)$*

$$\underline{E}_{TV}(\{(x_2,y_2)\}) = \underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(\{(x_2,y_2)\}) - 2\delta.$$

*On the other hand, if we consider the event $\{(x_2,y_2)\}^c$, we deduce from Eqs. (15) and (5) that*

$$\underline{E}_{TV}(\{(x_2,y_2)\}^c) = 1 - \overline{E}_{TV}(\{(x_2,y_2)\}) = 0.5 - \delta \textit{ while}$$
$$\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(\{(x_2,y_2)\}^c) = 1 - \overline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}(\{(x_2,y_2)\}) = 0.5$$

*meaning that we should distort $\underline{E}_{P_0^{\mathscr{X}}, P_0^{\mathscr{Y}}}$ by $\delta$. We conclude from this that $\underline{E}_{TV}$ is not a TV-model.* ♦

## 7. Conclusions

The variety of distortion models present in the literature makes it interesting to compare their behaviour under a number of different perspectives, so as to be able to choose the most appropriate model in each situation. In this paper, we have complemented our earlier work in [16, 17] and compared four different distortion models by determining (i) if they are closed under conjunction, disjunction or convex mixture; (ii) if they are closed under marginalisation; (iii) whether there is a unique procedure to build an independent product; and (iv) whether the procedures of distortion and natural extension commute. Tables 1 and 2 summarise our results. We see that the PMM and LV models are the most stable, followed by the TV and COR models.

| | Conjunction | Disjunction | Unique OA? |
|---|---|---|---|
| PMM | YES [15, Prop.12] | NO (Ex.1) | YES [15, Prop.12] |
| LV | YES (Prop. 5) | NO (Ex.3) | YES [14, Prop.8] |
| COR | NO (Ex. 6) | NO (Ex.7) | NO (Ex.7) |
| TV | NO (Ex. 11) | NO (Ex.11) | NO (Ex. 12) |

Table 1: Behaviour under conjunction and disjunction.

| | Mixture | Marginalising | Natural extension |
|---|---|---|---|
| PMM | YES | YES [15, Sec. 6.2.] | YES (Thm. 3) |
| LV | YES | YES (Prop. 6) | NO (Ex. 5) |
| COR | NO (Ex. 8) | YES (Prop. 8) | NO (Ex. 10) |
| TV | YES | YES (Prop. 10) | NO (Ex. 14) |

Table 2: Behaviour of the neighbourhood models under mixture, marginalisation and natural extension.

In the case of the natural extension, we should remark that, strictly speaking, the natural extension of two marginal PMMs is only a PMM if we regard it as a PMM-distortion analogue of a lower probability, but not in the sense of Def. 2; see Thm. 3 for more details.

Our work in this paper may be extended in a number of ways: on the one hand, we may analyse other distortion models, such as those based on the Kolmogorov or $L_1$ distances [17], divergences such as Kullback-Leibler [9, 18] or nearly-linear models [7]; we may consider other models of merging [26] or of independence [8]; and we may take the approach one step further and investigate distorted credal sets, considering the ideas put forward by Moral in [18]. More generally, we could also analyse the processing of generic distortion models defined through some function $d$.

## 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, as well as the reviewers for their careful reading.

## References

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

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

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

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

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

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

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

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

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

[10] T. Herron, T. Seidenfeld, and L. Wasserman. Divisive conditioning: further results on dilation. *Philosophy of Science*, 64:411–444, 1997.

[11] P. J. Huber. *Robust Statistics*. Wiley, New York, 1981.

[12] P. J. Huber and V. Strassen. Minimax tests and the Neyman–Pearson lemma for capacities. *The Annals of Statistics*, 1:251–263, 1973.

[13] I. Montes, E. Miranda, R. Pelessoni, and P. Vicig. Sklar's theorem in an imprecise setting. *Fuzzy Sets and Systems*, 278C:48–66, 2015.

[14] I. Montes, E. Miranda, and P. Vicig. 2-monotone outer approximations of coherent lower probabilities. *International Journal of Approximate Reasoning*, 101: 181–205, 2018.

[15] I. Montes, E. Miranda, and S. Destercke. Pari-mutuel probabilities as an uncertainty model. *Information Sciences*, 481:550–573, 2019.

[16] 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):602–635, 2020.

[17] 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.

[18] 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.

[19] S. Moral and J. del Sagrado. Aggregation of imprecise probabilities. In *Aggregation and fusion of imperfect information*, pages 162–188. Springer, 1998.

[20] R. Nau. The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference*, 105: 265–282, 2002.

[21] 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.

[22] L. R. Pericchi and P. Walley. Robust Bayesian credible intervals and prior ignorance. *International Statistical Review*, 59:1–23, 1991.

[23] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.

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

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

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

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

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