



HAL
open science

Comonotonicity for sets of probabilities

Ignacio Montes, Sébastien Destercke

► **To cite this version:**

Ignacio Montes, Sébastien Destercke. Comonotonicity for sets of probabilities. *Fuzzy Sets and Systems*, 2017, 328, pp.1-34. 10.1016/j.fss.2016.09.012 . hal-01618317

HAL Id: hal-01618317

<https://hal.science/hal-01618317v1>

Submitted on 17 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comonotonicity for sets of probabilities

Ignacio Montes¹ and Sebastien Destercke²

1. *Carlos III University of Madrid, C/ Madrid, 126 - 28903 Getafe, Spain*
2. *Sorbonnes Universite, Universite de Technologie de Compiègne, CNRS, UMR Heudiasyc, 57 Av. de Landshut, 60203 Compiègne*

Abstract

Two variables are called comonotone when there is an increasing relation between them, in the sense that when one of them increases (decreases), so does the other one. This notion has been widely investigated in probability theory, and is related to copulas. This contribution studies how the notion of comonotonicity can be extended to an imprecise setting on discrete spaces, where probabilities are only known to belong to a convex set. We define comonotonicity for such sets and investigate its characterizations in terms of lower probabilities, as well as its connection with copulas. As this theoretical characterization can be tricky to apply to general lower probabilities, we also investigate specific models of practical importance. In particular, we provide some sufficient conditions for a comonotone belief function with fixed marginals to exist, and characterize comonotone bivariate p-boxes.

Keywords: Copula, Sklar's Theorem, Comonotonicity, Lower Probability, Belief Function, P-box.

1. Introduction

When modelling dependencies between two variables in probability theory, the celebrated Sklar's Theorem [20] tells us that we can express any joint distribution by means of a function called copula [17], which contains all the dependency information. For example, the product copula corresponds to independent variables.

Some other copulas also play an important theoretical role: among them are the minimum and the Lukasiewicz operator. Those two copulas correspond to the so-called Fréchet-Hoeffding bounds, and respectively constitute the lower and upper bounds of every possible copula. They also model extreme cases of positive and negative dependencies between variables: the minimum models comonotone variables, i.e., variables that increase simultaneously, while the Lukasiewicz operator models countermonotone random variables, i.e., variables

Email address: `igmontes@est-econ.uc3m.es`, `sebastien.destercke@hds.utc.fr`
(Ignacio Montes¹ and Sebastien Destercke²)

such that when one increases (resp., decreases) the other one decreases (resp., increases). The notions of comonotonicity and countermonotonicity also have a practical importance: they are applied to different fields [4, 6, 11, 12, 24, 25] such as finance, and can be used to model expert information (e.g., “pressure always increases with temperature”) in systems such as Bayesian networks.

While dependency modelling is a well studied matter when using precise probabilities, there are only a few works studying it when considering lower probabilities and sets of probabilities induced by such measures [9, 10, 16, 18, 22], in contrast with the notion of independence [2, 3]. Lower probabilities and sets of probabilities are instrumental models when obtaining a unique probability is difficult or impossible (due to lack of data, imprecise expert opinions, non-probabilistic uncertainty, ...). Formally, they are general enough to include as particular cases many existing models such as belief functions [19], possibility measures [7, 26] and probability boxes [9] (imprecise cumulative distributions), both univariate [21] and bivariate [18]. In this paper, we adopt a view more consistent with the robust interpretation of such models (sets including an ill-known precise probability) than with their subjective one (where lower/upper probabilities are belief measures not assuming the existence of a precise, objective probability [23]).

Given the theoretical and practical importance of comonotonicity and countermonotonicity as dependence models, it seems important to study what they become in an imprecise setting. This is what we do in this paper, where we extend the notion of comonotonicity to lower probabilities, and explore its theoretical properties as well as some of its more interesting practical aspects (all results also apply to countermonotonicity through a simple change of variable).

Our extension of comonotonicity to lower probabilities and probability sets is introduced and studied in Section 3. Two main results of this section are that only a weak form of the link between comonotonicity and Fréchet-Hoeffding bounds holds in the imprecise case, and that there are pairs of marginal probability sets for which no joint comonotonic probability sets exist. This contrasts with precise probabilities, that can always be combined with the minimum copula to produce a comonotonic joint, and implies that situations where a comonotone probability set exists are much harder to characterize.

For this reason, we study in Sections 4 and 5 specific cases of practical interest, for which this characterization is easier. Section 4 focuses on belief functions and uses the fact that they can be seen as probability distributions over subsets to find necessary and sufficient conditions for a joint comonotonic belief function to exist. These conditions will also provide us with practical means to build a joint comonotonic belief function, or to check the existence of such a joint. We then particularize those results to marginal p-boxes in Section 5, since p-boxes are the natural extension of cumulative distributions, and therefore the imprecise probabilistic model best fitted to study (imprecise) copulas. Section 2 reminds the basic notions about lower probabilities, p-boxes and copulas we shall need in the paper. Note that this paper extends significantly (by providing proofs, examples as well as additional discussions and results) a previous short conference version [15].

2. Preliminary notions

This section introduces concepts and results concerning coherent lower probabilities, probability sets and copulas that are necessary in this paper. For an in-depth review of these theories, we refer to [14, 23] and [17], respectively.

2.1. Lower probabilities

A *lower probability* on a (finite) space Ω is a functional $\underline{P} : \mathcal{K} \subseteq \mathcal{P}(\Omega) \rightarrow [0, 1]$ whose domain \mathcal{K} is a subset of the power set $\mathcal{P}(\Omega)$ of Ω . In a subjective setting, the lower probability $\underline{P}(A)$ of event A can be interpreted as the subject's supremum acceptable buying price for the bet that obtains 1 if A happens and 0 otherwise, while in an objective setting, \underline{P} can be interpreted as the lower bound of some ill-known precise probability. Any lower probability is associated, through a conjugacy relation, to an *upper probability* $\overline{P} : \mathcal{K}^c \rightarrow [0, 1]$ such that

$$\overline{P}(A^c) = 1 - \underline{P}(A) \quad \forall A \in \mathcal{K} \quad (1)$$

and where $\mathcal{K}^c = \{A^c : A \in \mathcal{K}\}$. Any lower probability defines a convex set of probabilities, usually called *credal set*, given by the set

$$\mathcal{M}(\underline{P}) = \{P \in \mathbb{P} \mid \underline{P}(A) \leq P(A) \forall A \in \mathcal{K}\} \quad (2)$$

of all probabilities P that dominate \underline{P} , where \mathbb{P} is the set of probability measures defined over $\mathcal{P}(\Omega)$. The upper probability also defines a set of probabilities given by:

$$\mathcal{M}(\overline{P}) = \{P \in \mathbb{P} \mid \overline{P}(A) \geq P(A) \forall A \in \mathcal{K}^c\}. \quad (3)$$

Due to the conjugacy relation of Eq. (1), $\mathcal{M}(\underline{P}) = \mathcal{M}(\overline{P})$ and both sets contain the same information. If $\mathcal{M}(\underline{P}) \neq \emptyset$, then the lower probability is commonly said to *avoid sure loss*, and it then follows that for any $A \in \mathcal{K} \cap \mathcal{K}^c$, we have the inequality $\underline{P}(A) \leq \overline{P}(A)$. Avoiding sure loss can be seen as a minimal consistency requirement, as it ensures that the lower probability is dominated by at least one probability.

Another stronger consistency requirement usually imposed on lower probabilities is *coherence*: a lower probability \underline{P} is coherent when

$$\underline{P}(A) = \min_{P \in \mathcal{M}(\underline{P})} P(A) \quad \forall A \in \mathcal{K}, \quad (4)$$

that is, when \underline{P} is the lower envelope of $\mathcal{M}(\underline{P})$. Due to the conjugacy relation of Eq. (1), when \underline{P} is coherent we also have

$$\overline{P}(A) = \max_{P \in \mathcal{M}(\underline{P})} P(A) \quad \forall A \in \mathcal{K}^c. \quad (5)$$

Some of the most important properties of coherent lower and upper probabilities used in this paper are summarized in the next result.

Proposition 1. [23, Sec. 2.7.4] *The conjugate coherent lower and upper probabilities \underline{P} and \overline{P} satisfy the following properties:*

P.1 $A \subseteq B$ implies both $\underline{P}(A) \leq \underline{P}(B)$ and $\overline{P}(A) \leq \overline{P}(B)$.

P.2 $\overline{P}(A \cup B) \leq \overline{P}(A) + \overline{P}(B)$ and, if $A \cap B = \emptyset$, $\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$.

P.3 $A \cap B = \emptyset \Rightarrow \overline{P}(A \cup B) \geq \underline{P}(A) + \overline{P}(B)$ and $\underline{P}(A \cup B) \leq \underline{P}(A) + \overline{P}(B)$.

Any coherent lower probability defined on \mathcal{K} can be coherently extended to a greater domain $\mathcal{K}' \supseteq \mathcal{K}$. The natural extension, [23] defined as

$$\underline{E}(A) = \min\{P(A) : P \in \mathcal{M}(\underline{P})\} \text{ for all } A \in \mathcal{K}', \quad (6)$$

is the most conservative way to do so, in the sense that $\underline{E}(A) \leq \underline{P}'(A)$ for all $A \in \mathcal{K}'$ and for every other coherent extension \underline{P}' of \underline{P} to \mathcal{K}' . It can be interpreted as the extension that adds no further information to \underline{P} , in contrast with other, less conservative ones.

2.2. Cumulative distributions and p-boxes

Cumulative distributions, both univariate and bivariate, and their imprecise counterparts, p-boxes [9], are strongly connected to copulas and will play an important role in this paper.

From now on, we consider the finite ordered spaces¹ $\mathcal{X} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, \dots, y_m\}$, and their cartesian product $\mathcal{X} \times \mathcal{Y}$. Using the notation introduced in [18], we consider the sets

$$A_x = [x_1, x], \quad A_y = [y_1, y], \quad \text{and} \quad A_{x,y} = A_x \times A_y, \quad (7)$$

where $[x_1, x]$ (respectively, $[y_1, y]$) stands for the set $\{x_1, x_2, \dots, x\}$ (respectively, $\{y_1, y_2, \dots, y\}$). $P_{X,Y}$ will denote a joint probability $P_{X,Y} : \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \rightarrow [0, 1]$, while $P_X : \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$ and $P_Y : \mathcal{P}(\mathcal{Y}) \rightarrow [0, 1]$ denote the marginal probabilities. Also, the joint and marginal cumulative distribution functions $F_{X,Y}$, F_X and F_Y given by:

$$F_{X,Y}(x, y) = P_{X,Y}(A_{x,y}), \quad F_X(x) = P_X(A_x) \quad \text{and} \quad F_Y(y) = P_Y(A_y) \quad (8)$$

for all $x \in \mathcal{X}, y \in \mathcal{Y}$.

Let us now introduce uni- and bivariate p-boxes, which can be understood as bounds of a unknown or imprecisely defined (uni- or bivariate) cumulative distribution function.

Definition 2. A (discrete) univariate p-box defined on an ordered finite set $\mathcal{X} = \{x_1, \dots, x_n\}$ is a pair of increasing functions $\underline{F}, \overline{F} : \mathcal{X} \rightarrow [0, 1]$ such that $\underline{F} \leq \overline{F}$ and $\underline{F}(x_n) = \overline{F}(x_n) = 1$.

Definition 3. A (discrete) bivariate p-box defined on the Cartesian product $\mathcal{X} \times \mathcal{Y} = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$ of two ordered finite sets is a pair of component-wise increasing functions $\underline{F}, \overline{F} : \mathcal{X} \times \mathcal{Y} \rightarrow [0, 1]$ such that $\underline{F} \leq \overline{F}$ and $\underline{F}(x_n, y_m) = \overline{F}(x_n, y_m) = 1$.

¹We assume the elements are indexed according to this order.

Remark 4. In the previous definition we have not required univariate p-boxes $(\underline{F}, \overline{F})$ to satisfy the condition $\underline{F}(x_1) = \overline{F}(x_1) = 0$, as $(\underline{F}, \overline{F})$ models here some imprecise knowledge of a cumulative distribution function. Indeed, consider a precise probability P defined on $\mathcal{P}(\mathcal{X})$, where $\mathcal{X} = \{x_1, \dots, x_n\}$. Denote by F_P its associated cumulative distribution function, which satisfies the following properties:

- F_P is increasing.
- $F_P(x_n) = P(\{x_1, \dots, x_n\}) = P(\mathcal{X}) = 1$.

However, as soon as $P(\{x_1\}) > 0$, $F_P(x_1) > 0$ hence we do not require \underline{F} and \overline{F} to be 0 in x_1 . The same reasoning can be used to justify why we do not require bivariate p-boxes $(\underline{F}, \overline{F})$ to satisfy $\underline{F}(x_1, y_1) = \overline{F}(x_1, y_1) = 0$. ♦

Univariate [21] and bivariate [18] p-boxes are related to coherent and 2-coherent lower probabilities [23], respectively. They can be used to model the imprecise information about (univariate or bivariate) cumulative distribution functions, as we retrieve precise information when $\underline{F} = \overline{F} = F$.

In the bivariate case, a cumulative distribution function F satisfies, in addition to being component-wise increasing and normalized ($F(x_n, y_m) = 1$), the rectangle inequality

$$F(x_i, y_i) + F(x_j, y_j) \geq F(x_j, y_i) + F(x_i, y_j) \quad (9)$$

for any $x_i, x_j \in \mathcal{X}$, $y_i, y_j \in \mathcal{Y}$ such that $x_i \leq x_j$ and $y_i \leq y_j$. Therefore, a bivariate p-box $(\underline{F}, \overline{F})$ with $\underline{F} = \overline{F}$ may not induce a bivariate cumulative distribution function because in Definition 3 we do not require $\underline{F}, \overline{F}$ to satisfy Eq. (9). We refer to [18] for more details.

Definition 5. A univariate p-box $(\underline{F}, \overline{F})$ on \mathcal{X} defines a lower probability \underline{P} on the domain $\mathcal{K}_1 = \{A_x, A_x^c : x \in \mathcal{X}\}$, given by

$$\underline{P}(A_x) = \underline{F}(x) \text{ and } \underline{P}(A_x^c) = 1 - \overline{F}(x) \quad \forall x \in \mathcal{X}. \quad (10)$$

This lower probability is coherent [21, 23], and it can be extended to $\mathcal{P}(\mathcal{X})$ by applying the natural extension defined on Eq. (6).

Definition 6. A bivariate p-box $(\underline{F}, \overline{F})$ defines a lower probability \underline{P} on the domain $\mathcal{K}_2 = \{A_{x,y}, A_{x,y}^c : (x,y) \in \mathcal{X} \times \mathcal{Y}\}$ given by

$$\underline{P}(A_{x,y}) = \underline{F}(x,y) \text{ and } \underline{P}(A_{x,y}^c) = 1 - \overline{F}(x,y) \quad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}. \quad (11)$$

Nevertheless, this lower probability defined from a bivariate p-box may not be coherent. A bivariate p-box will be called *coherent* when its associated lower probability given in Eq. (11) is coherent. In such a case, it can be extended to $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ by using the natural extension in Eq. (6).

If we consider a univariate or bivariate coherent p-box $(\underline{F}, \overline{F})$ and its associated lower probability \underline{P} given in Eq. (10) or Eq. (11), its natural extension to $\mathcal{P}(\mathcal{X})$ or $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ can be equivalently expressed by:

$$\underline{P}(A) = \min\{P(A) \mid \underline{F} \leq F_P \leq \overline{F}\} \quad \forall A, \quad (12)$$

where F_P denotes the cumulative distribution function associated with P . Therefore, the credal set of \underline{P} can be expressed as:

$$\mathcal{M}(\underline{P}) = \{P \in \mathbb{P} \mid \underline{F} \leq F_P \leq \overline{F}\}, \quad (13)$$

where again F_P is the cumulative distribution function associated with P .

Conversely, any coherent lower probability \underline{P}_X and \underline{P} , respectively defined on a sub-domain of $\mathcal{P}(\mathcal{X})$ or $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$, define a univariate p-box

$$\underline{F}_X(x) = \underline{P}(A_x) = \inf\{F_P(x) : P \in \mathcal{M}(\underline{P}_X)\}; \quad (14)$$

$$\overline{F}_X(x) = \overline{P}(A_x) = \sup\{F_P(x) : P \in \mathcal{M}(\underline{P}_X)\}, \quad (15)$$

and a bivariate p-box

$$\underline{F}(x, y) = \underline{E}(A_{x,y}) = \inf\{F_P(x, y) : P \in \mathcal{M}(\underline{P})\}; \quad (16)$$

$$\overline{F}(x, y) = \overline{E}(A_{x,y}) = \sup\{F_P(x, y) : P \in \mathcal{M}(\underline{P})\}; \quad (17)$$

that are obtained by considering the natural extensions of $\underline{P}_X, \underline{P}$ to the domains $\mathcal{K}_1, \mathcal{K}_2$. In general, the lower probabilities \underline{P}'_X and \underline{P}' induced by these p-boxes $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}, \overline{F})$ through Eq. (6) will only outer-approximate the original lower probabilities from which the p-boxes have been built (i.e., $\underline{P}'_X(A) \leq \underline{P}_X(A)$ and $\underline{P}'(A) \leq \underline{P}(A)$ for all A , with the inequality being possibly strict).

2.3. Copulas and Sklar's Theorem

In probability theory, Sklar's Theorem shows that we can express a joint cumulative distribution function in terms of the marginals by means of a function called copula.

Definition 7. [17] *A copula is a binary operator $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ satisfying the following properties:*

1. $C(x, 0) = C(0, x) = 0$ and $C(x, 1) = C(1, x) = x$ for all $x \in [0, 1]$.
2. $C(x_1, y_1) + C(x_2, y_2) \geq C(x_2, y_1) + C(x_1, y_2)$ for all $x_1, x_2, y_1, y_2 \in [0, 1]$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$.

Theorem 8 (Sklar's Theorem [20]). *Let $F_{X,Y}$ be a joint cumulative distribution function with marginals F_X and F_Y . Then, there exists a copula C such that*

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)) \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}. \quad (18)$$

Conversely, given two marginal cumulative distribution functions F_X and F_Y and a copula C , they define a joint cumulative distribution function $F_{X,Y}$ using Eq. (18).

The product, minimum and Łukasiewicz operator are three of the most important copulas among operators satisfying Definition 7. Indeed, the product copula $\Pi(x, y) = x \cdot y$ applies to independence since $F_{X,Y}(x, y) = F_X(x) \cdot$

$F_Y(y)$. The minimum copula $M(x, y) = \min(x, y)$ and the Łukasiewicz operator $W(x, y) = \max(x + y - 1, 0)$ correspond to extreme assumptions of positive and negative dependence, and are commonly called Fréchet-Hoeffding bounds since any copula satisfies the so-called Fréchet-Hoeffding inequality:

$$M(x, y) \leq C(x, y) \leq W(x, y). \quad (19)$$

In order to explain in more details the dependence assumptions behind the copulas M and W , we first need the following definition.

Definition 9. A subset S of \mathbb{R}^2 is increasing (respectively, decreasing) when for all $(x, y), (u, v) \in S$, $x < u$ (resp. $x < u$) implies $y \leq v$ (resp., $y \geq v$), and $y < v$ (resp. $y < v$) implies $x \leq u$ (resp., $x \geq u$). Similarly, two points in \mathbb{R}^2 , (x, y) and (u, v) , are called increasing (resp., decreasing) when the set $S = \{(x, y), (u, v)\}$ is increasing (resp., decreasing). We can then define the partial orders \preceq and \lesssim given by:

$$(x, y) \preceq (u, v) \Leftrightarrow x \leq u, y \leq v \text{ and } (x, y) \lesssim (u, v) \Leftrightarrow x \leq u, y \geq v, \quad (20)$$

such that all the elements in increasing or decreasing sets are ordered with respect to \preceq and \lesssim , respectively.

We are now ready to define the notions of comonotonicity and countermonotonicity when the probabilities (and associated cumulative distributions) are precisely known.

Definition 10. Consider a joint probability $P_{X,Y}$ on $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ with support $\text{Supp}(P_{X,Y})$ given by:

$$\text{Supp}(P_{X,Y}) = \{(x, y) : P_{X,Y}(\{(x, y)\}) > 0\}. \quad (21)$$

$P_{X,Y}$ is called comonotone when any, and hence all, of the following equivalent conditions are satisfied:

- The copula that links the marginals is the minimum:

$$F_{X,Y}(x, y) = \min(F_X(x), F_Y(y)) \text{ for all } (x, y) \in \mathcal{X} \times \mathcal{Y}. \quad (22)$$

- The support $\text{Supp}(P_{X,Y})$ of (X, Y) is an increasing set.
- For all $(x, y) \in \mathcal{X} \times \mathcal{Y}$, either $P_{X,Y}(\{(u, v) : u \leq x, v > y\}) = 0$ or $P_{X,Y}(\{(u, v) : u > x, v \leq y\}) = 0$.

$P_{X,Y}$ is called countermonotone when any, and hence all, of the following equivalent conditions are satisfied:

- The copula that links the marginals is the Łukasiewicz operator:

$$F_{X,Y}(x, y) = \max(F_X(x) + F_Y(y) - 1, 0) \text{ for all } (x, y) \in \mathcal{X} \times \mathcal{Y}. \quad (23)$$

- The support $\text{Supp}(P_{X,Y})$ of (X,Y) is a decreasing set.
- For all $(x,y) \in \mathcal{X} \times \mathcal{Y}$, either $P_{X,Y}(\{(u,v) : u \leq x, v < y\}) = 0$ or $P_{X,Y}(\{(u,v) : u > x, v \geq y\}) = 0$.

Remark 11. *In practice, any result applying to comonotonicity also applies to countermonotonicity. To see this, it is sufficient to note that if $P_{X,Y}$ is comonotone, the probability $P'_{X,Y}$ given by $P'_{X,Y}(A) = P_{X,Y}(A')$, where $A' = \{(x, y_{m-i+1}) : (x, y_i) \in A\}$, is countermonotone. In practice, this comes down to reverse the order on one of the space \mathcal{X} or \mathcal{Y} . This is why, in the rest of the paper, we focus solely on the notion of comonotonicity. \blacklozenge*

When going from precise to imprecise probabilities, it has been demonstrated that only a weaker version of Sklar’s theorem [16] holds: combining a pair of imprecise cumulative distributions (p-boxes) by a convex set of copulas still results in a bivariate p-box inducing a coherent lower probability, yet there are bivariate coherent p-boxes that cannot be decomposed as a pair of marginal p-boxes combined through a set of copulas.

Therefore, it makes sense to wonder to which extent the equivalences of Definition 10 still hold in an imprecise setting, and if they do not hold anymore, how can we obtain comonotonic models in practice? The rest of the paper addresses these two issues, the latter being considered only for specific models (belief functions) that offer a good compromise between generality and practicality.

3. Comonotone lower probabilities

We first provide a definition of comonotone lower probabilities extending the one of precise comonotone probabilities, and study what becomes of Definition 10 in this setting. We then touch upon the tricky question of practically dealing with comonotone lower probabilities, before focusing on specific models in Sections 4 and 5.

3.1. Definition and characterization

Definition 10 gives three equivalent ways to define comonotonicity in the precise framework: by means of the copula linking the marginals, by means of the increasingness of the support, or by means of the probability of the sets

$$D_1^{(x,y)} = \{(u,v) : u > x, v \leq y\} \text{ and } D_2^{(x,y)} = \{(u,v) : u \leq x, v > y\}. \quad (24)$$

Figure 1 provides an illustration of those sets. We now investigate to what extent these conditions, or similar ones, also hold in the case of comonotone lower probabilities.

In our framework, we consider a lower probability \underline{P} defined on the power set of $\mathcal{X} \times \mathcal{Y}$, where both \mathcal{X} and \mathcal{Y} are finite subsets of \mathbb{R} . We assume that \underline{P} models the imprecise information about a joint probability $P_{X,Y}$. The question is: how can we model the additional information that $P_{X,Y}$ is comonotone?

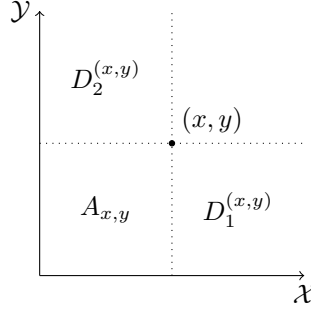


Figure 1: Illustration of $A_{x,y}$, $D_1^{(x,y)}$, $D_2^{(x,y)}$.

	$\{(1, 1)\}$	$\{(1, 2)\}$	$\{(2, 2)\}$	$\{(2, 1)\}$
P_1	α	$1 - \alpha - \beta$	β	0
P_2	α	0	$1 - \alpha$	0
P_3	$1 - \beta$	0	β	0

Table 1: Extreme points for Example 13.

If \underline{P} models the imprecise information about a comonotone probability $P_{X,Y}$, we must impose that all the probabilities compatible with \underline{P} must also be comonotone.

Definition 12. A lower probability \underline{P} defined on the power set of the finite set $\mathcal{X} \times \mathcal{Y}$ is called comonotone when all $P \in \mathcal{M}(\underline{P})$ are comonotone.

Example 13. Consider a lower probability \underline{P} defined on the power set of $\mathcal{X} \times \mathcal{Y} = \{1, 2\} \times \{1, 2\}$ by:

$$\begin{aligned}
\underline{P}(\{(1, 1)\}) &= \underline{P}(\{(1, 1), (1, 2)\}) = \underline{P}(\{(1, 1), (2, 1)\}) \\
&= \underline{P}(\{(1, 1), (1, 2), (2, 1)\}) = \alpha \in (0, 0.5), \\
\underline{P}(\{(2, 2)\}) &= \underline{P}(\{(1, 2), (2, 2)\}) = \underline{P}(\{(2, 1), (2, 2)\}) \\
&= \underline{P}(\{(1, 2), (2, 1), (2, 2)\}) = \beta \in (0, 0.5), \\
\underline{P}(\{(1, 1), (2, 2)\}) &= \underline{P}(\{(1, 1), (2, 1), (2, 2)\}) = \alpha + \beta, \\
\underline{P}(\{(1, 1), (1, 2), (2, 2)\}) &= \underline{P}(\{1, 2\}^2) = 1, \\
\underline{P}(A) &= 0 \text{ otherwise.}
\end{aligned}$$

It can be checked that \underline{P} is coherent. In fact, the extreme points of $\mathcal{M}(\underline{P})$ are the precise probabilities P_1, P_2 and P_3 given by Table 1. Then, $\mathcal{M}(\underline{P})$ is formed by all the convex combinations of P_1, P_2 and P_3 . Consequently, the support of every P in $\mathcal{M}(\underline{P})$ is included in $\{(1, 1), (1, 2), (2, 2)\}$, an increasing set, hence all the probabilities in $\mathcal{M}(\underline{P})$ as well as \underline{P} are comonotone. \blacklozenge

In Example 13, every extreme point of $\mathcal{M}(\underline{P})$ is comonotone, as well as every $P \in \mathcal{M}(\underline{P})$. Therefore, someone may think that it is sufficient to only

require comonotonicity of the extreme points of $\mathcal{M}(\underline{P})$ to satisfy Definition 12. However, we only have the next result:

Theorem 14. *If a coherent lower probability \underline{P} defined on the power set of the finite set $\mathcal{X} \times \mathcal{Y}$ is comonotone, then the extreme points of the set $\mathcal{M}(\underline{P})$ are comonotone.*

Proof. Since if \underline{P} is comonotone, then by definition 12 all precise probabilities within $\mathcal{M}(\underline{P})$ are so, including its extreme points. \square

Indeed, the comonotonicity of the extreme points in $\mathcal{M}(\underline{P})$ does not ensure the comonotonicity of all the probabilities in $\mathcal{M}(\underline{P})$, as next example shows.

Example 15. *Consider a lower probability defined on the power set of $\mathcal{X} \times \mathcal{Y} = \{1, 2\} \times \{1, 2\}$ by*

$$\underline{P}(A) = \begin{cases} 1 & \text{if } A \supseteq \{(1, 2), (2, 1)\}. \\ 0 & \text{otherwise.} \end{cases}$$

This lower probability is coherent, and its credal set is given by:

$$\mathcal{M}(\underline{P}) = \{P \in \mathbb{P} \mid P(\{(1, 2)\}) = \alpha, P(\{(2, 1)\}) = 1 - \alpha, \alpha \in [0, 1]\}.$$

The extreme points of $\mathcal{M}(\underline{P})$ are P_1 and P_2 given by $P_1(\{(1, 2)\}) = 1$ and $P_2(\{(2, 1)\}) = 1$. Both P_1 and P_2 are comonotone, but any convex combination of P_1 and P_2 is not, because for $P_\alpha = \alpha P_1 + (1 - \alpha)P_2$ ($\alpha \in (0, 1)$), the support is $\{(1, 2), (2, 1)\}$, which is not an increasing set. \blacklozenge

As previous example has shown, it is not sufficient to have every extreme point comonotone for every $P \in \mathcal{M}(\underline{P})$ to be comonotone. In fact, in the previous example the only comonotone probabilities in $\mathcal{M}(\underline{P})$ are the extreme points. For this reason, in Definition 12 we have required all the probabilities in the credal set to be comonotone.

We now investigate how comonotone coherent lower probabilities can be equivalently expressed. We first express it by means of the sets $D_1^{(x,y)}$ and $D_2^{(x,y)}$.

Theorem 16. *A coherent lower probability \underline{P} defined on the power set of the finite set $\mathcal{X} \times \mathcal{Y}$ is comonotone if and only if for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ either*

$$\overline{P}(D_1^{(x,y)}) = 0 \text{ or } \overline{P}(D_2^{(x,y)}) = 0, \quad (25)$$

where \overline{P} is the conjugate upper probability of \underline{P} .

Proof. Only if: consider $(x, y) \in \mathcal{X} \times \mathcal{Y}$ and let us see that one of the equalities of Eq. (25) holds. Since \underline{P} is coherent, there are $P_1, P_2 \in \mathcal{M}(\underline{P})$ such that:

$$\overline{P}(D_1^{(x,y)}) = P_1(D_1^{(x,y)}) \text{ and } \overline{P}(D_2^{(x,y)}) = P_2(D_2^{(x,y)}). \quad (26)$$

Define $P = \frac{(P_1+P_2)}{2}$, which belongs to $\mathcal{M}(\underline{P})$ because $\mathcal{M}(\underline{P})$ is convex. P is comonotone because \underline{P} is comonotone, and according to Definition 10, one of the following conditions holds:

$$P(D_1^{(x,y)}) = 0 \text{ or } P(D_2^{(x,y)}) = 0.$$

Assume without loss of generality that the former equality holds. In such a case, since $P = \frac{(P_1+P_2)}{2}$, it also holds that:

$$P_1(D_1^{(x,y)}) = 0 \text{ and } P_2(D_1^{(x,y)}) = 0.$$

Therefore:

$$\bar{P}(D_1^{(x,y)}) = P_1(D_1^{(x,y)}) = 0.$$

If: assume that at least one of $\bar{P}(D_1^{(x,y)})$ or $\bar{P}(D_2^{(x,y)})$ is zero. Then, for all $P \in \mathcal{M}(\underline{P})$:

$$P(D_1^{(x,y)}) \leq \bar{P}(D_1^{(x,y)}) \text{ and } P(D_2^{(x,y)}) \leq \bar{P}(D_2^{(x,y)}).$$

Then, P is comonotone since one of the previous probabilities is zero. \square

Theorem 16 shows that Definition 12 can also be expressed by only using lower and upper probabilities of particular events, meaning that our characterization also makes sense within a subjective interpretation. In particular, Eq. (25) can be interpreted as constraints on betting rates to impose in order to guarantee comonotonicity. Now, we are going to see that, if we define the support $\text{Supp}(\underline{P})$ of a lower probability \underline{P} by:

$$\text{Supp}(\underline{P}) = \bigcup_{P \in \mathcal{M}(\underline{P})} \text{Supp}(P), \quad (27)$$

its comonotonicity can also be equivalently expressed in terms of the increasingness of $\text{Supp}(\underline{P})$.

Theorem 17. *A coherent lower probability \underline{P} defined on the power set of the finite set $\mathcal{X} \times \mathcal{Y}$ is comonotone if and only if its support $\text{Supp}(\underline{P})$ is an increasing set.*

Proof. Only if: First observe that another way to specify $\text{Supp}(\underline{P})$ is

$$\text{Supp}(\underline{P}) = \{(x, y) \in \mathcal{X} \times \mathcal{Y} : \bar{P}(\{(x, y)\}) > 0\} \quad (28)$$

since if $\bar{P}(\{(x, y)\}) > 0$, then there is $P \in \mathcal{M}(\underline{P})$ with $P(\{(x, y)\}) > 0$, and if $\bar{P}(\{(x, y)\}) = 0$, then $P(\{(x, y)\}) = 0$ for all $P \in \mathcal{M}(\underline{P})$. Now, let us consider for each $(x, y) \in \mathcal{X} \times \mathcal{Y}$ a probability $P_{(x,y)} \in \mathcal{M}(\underline{P})$ such that

$$P_{(x,y)}(\{(x, y)\}) = \bar{P}(\{(x, y)\}).$$

Now, we can define

$$P = \frac{1}{|\mathcal{X}| \cdot |\mathcal{Y}|} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{(x,y)},$$

which belongs to $\mathcal{M}(\underline{P})$ because $\mathcal{M}(\underline{P})$ is convex, hence is comonotone. As P is positive if and only if $\overline{P}(\{(x,y)\}) > 0$ by construction, we have $\text{Supp}(P) = \text{Supp}(\underline{P})$ and $\text{Supp}(P)$ is increasing (Definition 10).

If: assume that $\text{Supp}(\underline{P})$ is increasing. Then, for every probability $P \in \mathcal{M}(\underline{P})$, its support $\text{Supp}(P)$ is included in $\text{Supp}(\underline{P})$, and then it is increasing. Then, we conclude that all $P \in \mathcal{M}(\underline{P})$ is comonotone, hence \underline{P} is comonotone. \square

Therefore, this second equivalent expression also holds for coherent lower probabilities. Now, it only remains to check whether or not the comonotonicity of lower probabilities is related to the copula that links the marginals. The next result shows one implication.

Theorem 18. *Let \underline{P} be a coherent comonotone lower probability defined on the power set of $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, \dots, y_m\}$. If $(\underline{F}, \overline{F})$, $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ respectively denote the bivariate and the marginal univariate p -boxes induced by \underline{P} , then for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$:*

$$\underline{F}(x, y) = \min(\underline{F}_X(x), \underline{F}_Y(y)) \text{ and } \overline{F}(x, y) = \min(\overline{F}_X(x), \overline{F}_Y(y)). \quad (29)$$

Proof. Let us prove that $\underline{F}(x, y) = \min(\underline{F}_X(x), \underline{F}_Y(y))$. First of all, using P.1 of Proposition 1, note that

$$\begin{aligned} \underline{F}(x, y) &= \underline{P}(A_{x,y}) \leq \underline{P}(A_{x_n,y}) = \underline{F}_Y(y) \text{ and} \\ \underline{F}(x, y) &= \underline{P}(A_{x,y}) \leq \underline{P}(A_{x,y_m}) = \underline{F}_X(x) \end{aligned}$$

since $\underline{P}(A) \leq \underline{P}(B)$ whenever $A \subseteq B$. It follows that

$$\underline{F}(x, y) \leq \min(\underline{F}_X(x), \underline{F}_Y(y)). \quad (30)$$

Let us now show the converse inequality. Consider $P \in \mathcal{M}(\underline{P})$ such that $\underline{P}(A_{x,y}) = P(A_{x,y})$ and denote by F, F_X and F_Y the joint and marginal distribution functions associated with P . Then, since P is comonotone by assumption:

$$\begin{aligned} \underline{F}(x, y) &= \underline{P}(A_{x,y}) = P(A_{x,y}) = F(x, y) \\ &= \min(F_X(x), F_Y(y)) \geq \min(\underline{F}_X(x), \underline{F}_Y(y)). \end{aligned} \quad (31)$$

Using Eqs. (30) and (31), we conclude that $\underline{F}(x, y) = \min(\underline{F}_X(x), \underline{F}_Y(y))$.

Now let us prove the same equality for $\overline{F}(x, y)$. First consider a probability P in $\mathcal{M}(\underline{P})$ such that $P(A_{x,y}) = \overline{P}(A_{x,y})$ and let F, F_X and F_Y be its joint and marginal cumulative distribution functions. Then, since P is comonotone:

$$\overline{F}(x, y) = \overline{P}(A_{x,y}) = P(A_{x,y}) = \min(F_X(x), F_Y(y)) \leq \min(\overline{F}_X(x), \overline{F}_Y(y)). \quad (32)$$

Let us now show that this inequality is an equality.

Using P.2 of Proposition 1, the notation of Eq. (24) and the fact that $A_{x,y} \cup D_1^{(x,y)} = A_{x_n,y}$, $A_{x,y} \cap D_1^{(x,y)} = \emptyset$, it holds that:

$$\bar{F}_Y(y) = \bar{P}(A_{x_n,y}) \leq \bar{P}(A_{x,y}) + \bar{P}(D_1^{(x,y)}) = \bar{F}(x,y) + \bar{P}(D_1^{(x,y)}),$$

and similarly

$$\bar{F}_X(x) = \bar{P}(A_{x,y_m}) \leq \bar{P}(A_{x,y}) + \bar{P}(D_2^{(x,y)}) = \bar{F}(x,y) + \bar{P}(D_2^{(x,y)}).$$

Since \underline{P} is comonotone, from Theorem 16 it holds that either $\bar{P}(D_1^{(x,y)}) = 0$, which implies that $\bar{F}_Y(y) \leq \bar{F}(x,y)$, or $\bar{P}(D_2^{(x,y)}) = 0$, which implies that $\bar{F}_X(x) \leq \bar{F}(x,y)$. In both cases it holds that $\bar{F}(x,y) \geq \min\{\bar{F}_X(x), \bar{F}_Y(y)\}$, and therefore using Eq. (32) we conclude $\bar{F}(x,y) = \min\{\bar{F}_X(x), \bar{F}_Y(y)\}$. \square

The next example shows that, unfortunately, the converse implication does not hold in general.

Example 19. Consider the joint coherent lower probability \underline{P} defined on the power set of the finite set $\{1, 2\}^2$ by:

$$\begin{aligned} \underline{P}(\{(1, 1), (1, 2), (2, 2)\}) &= \alpha > 0, \\ \underline{P}(\{(1, 1), (2, 1), (2, 2)\}) &= 1 - \alpha > 0, \\ \underline{P}(\{1, 2\}^2) &= 1, \\ \underline{P}(A) &= 0 \text{ otherwise.} \end{aligned}$$

Then, regardless of α , $\underline{F} = I_{\{(x,y):x,y \geq 2\}}$ and $\bar{F} = I_{\{(x,y):x,y \geq 1\}}$, with I_A the indicator function of event A . Furthermore:

$$\underline{F}_X(x) = \underline{F}_Y(y) = I_{\{x \geq 2\}}(x) \text{ and } \bar{F}_X(x) = \bar{F}_Y(y) = I_{\{x \geq 1\}}(x).$$

Then, for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$, \underline{F} and \bar{F} can be expressed as in Eq. (29). However, the probability P defined by $P(\{(1, 2)\}) = \alpha$ and $P(\{(2, 1)\}) = 1 - \alpha$ belongs to $\mathcal{M}(\underline{P})$. Therefore, $\text{Supp}(P)$ is not increasing and neither is $\text{Supp}(\underline{P})$, hence \underline{P} cannot be comonotone by Theorem 17. \blacklozenge

This means that in an imprecise setting Definition 12 only has two equivalent characterizations: by means of the increasingness of the support or by means of the upper probability of the sets $D_1^{(x,y)}$ and $D_2^{(x,y)}$. Indeed, the bivariate p-box of a comonotone lower probability is the minimum of the marginals, but the minimum of two marginal p-boxes will not necessarily generate a comonotone lower probability (this is in line with the imprecise Sklar Theorem [16]). Figure 2 summarizes the conditions we have seen along this section.

Comonotone lower probabilities

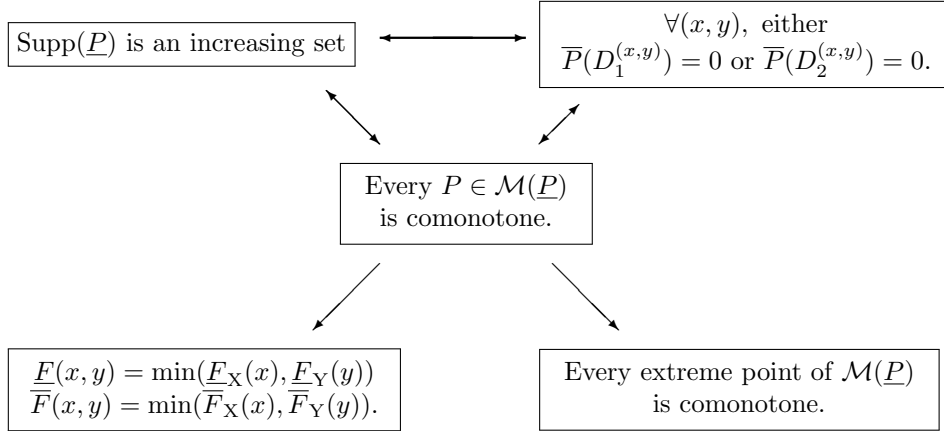


Figure 2: Equivalences and implications concerning the comonotonicity of a coherent lower probability \underline{P} .

Remark 20. As mentioned in Remark 11, results similar to Theorems 14, 16, 17 and 18 can be obtained in the countermonotone case through a simple change of variable:

- If \underline{P} is a countermonotone lower probability, then the extreme points of $\mathcal{M}(\underline{P})$ are also countermonotone, but the converse does not hold in general.
- A lower probability \underline{P} is countermonotone if and only if for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$, either $\overline{P}(\{(u, v) : u \leq x, v < y\}) = 0$ or $\overline{P}(\{(u, v) : u > x, y \geq v\}) = 0$.
- A lower probability \underline{P} is countermonotone if and only if $\text{Supp}(\underline{P})$ is a decreasing set.
- If \underline{P} is a countermonotone lower probability, then its associated p-box $(\underline{F}, \overline{F})$ can be expressed for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ by:

$$\underline{F}(x, y) = \max(\underline{F}_X(x) + \underline{F}_Y(y) - 1, 0). \quad (33)$$

$$\overline{F}(x, y) = \max(\overline{F}_X(x) + \overline{F}_Y(y) - 1, 0). \quad (34)$$

However, the converse does not hold in general.

The proofs are similar to those of comonotonicity. ♦

3.2. Building comonotone lower probabilities

So far, we have provided different theoretical characterizations of comonotone lower probabilities, but have not touch upon more practical questions such as:

- How can we check that a given joint lower probability \underline{P} is indeed comonotone?
- When and how can we build a comonotone lower probability \underline{P} from marginals $\underline{P}_X, \underline{P}_Y$?

Regarding the first question, checking all probabilities within $\mathcal{M}(\underline{P})$ is not doable in practice, and Theorem 14 together with Example 15 indicate that checking the comonotonicity of extreme points of $\mathcal{M}(\underline{P})$ is not sufficient, but only necessary to ensure the comonotonicity of \underline{P} .

The second question may seem more trivial at first sight, as it can always be answered positively in the precise framework: it is sufficient to consider the joint probability induced by the bivariate cumulative distribution $F_{X,Y}$ corresponding to the minimum of the two marginals F_X and F_Y by Eq. (22). Unfortunately, the next example shows that, in the imprecise setting, it is not even true that a comonotone lower probability can be obtained from every pair of marginals.

Example 21. Let \underline{P}_X and \underline{P}_Y be two marginal coherent lower probabilities, defined over $\mathcal{X} = \{1, 2, 3\}$ and $\mathcal{Y} = \{1, 2\}$, and such that

$$\begin{aligned} \underline{P}_X(\{1, 2\}) &= 0.7, & \underline{P}_X(\mathcal{X}) &= 1, & \underline{P}_X(A) &= 0 \text{ otherwise.} \\ \underline{P}_Y(\{1\}) &= 0.3, & \underline{P}_Y(\{2\}) &= 0.7, & \underline{P}_Y(\mathcal{Y}) &= 1. \end{aligned}$$

Let us assume that there is a comonotone joint lower probability \underline{P} with those marginals. If this is the case, then using Theorem 18, the bivariate p-box induced by \underline{P} is the minimum of the marginals $\underline{F}_X, \overline{F}_X$ and $\underline{F}_Y, \overline{F}_Y$. Then:

$$\overline{F}(1, 2) = \overline{P}(\{(1, 1), (1, 2)\}) = \min(\overline{F}_X(1), \overline{F}_Y(2)) = 1$$

and

$$\overline{F}(1, 1) = \overline{P}(\{(1, 1)\}) = \min(\overline{F}_X(1), \overline{F}_Y(1)) = 0.3,$$

meaning that $\exists P \in \mathcal{M}(\underline{P})$ such that $P(\{(1, 1)\}) + P(\{(1, 2)\}) = 1$ (otherwise \underline{P} would not be coherent) and $P(\{(1, 1)\}) \leq 0.3$, hence $P(\{(1, 2)\}) > 0$ and $(1, 2) \in \text{Supp}(P) \subseteq \text{Supp}(\underline{P})$. Similarly, we have that

$$\underline{F}(2, 1) = \underline{P}(\{(1, 1), (2, 1)\}) = \min(\underline{F}_X(2), \underline{F}_Y(1)) = 0.3$$

and

$$\underline{F}(1, 1) = \underline{P}(\{(1, 1)\}) = \min(\underline{F}_X(1), \underline{F}_Y(1)) = 0,$$

meaning that $\exists P \in \mathcal{M}(\underline{P})$ such that $P(\{(1, 1)\}) + P(\{(2, 1)\}) \geq 0.3$ and also $P(\{(1, 1)\}) = 0$, hence $P(\{(2, 1)\}) > 0$ and $(2, 1) \in \text{Supp}(P) \subseteq \text{Supp}(\underline{P})$. Since both $(1, 2)$ and $(2, 1)$ belong to $\text{Supp}(\underline{P})$, $\text{Supp}(\underline{P})$ is not an increasing set and by Theorem 17, \underline{P} cannot be comonotone, a contradiction. \blacklozenge

In the next sections, we shall investigate these two questions for particular cases of lower probabilities, namely belief functions and p-boxes, using their particular features at our advantage to get answers. As we shall see, the conditions to have a comonotone model can already be quite tricky in such contexts, suggesting that getting general but practically useful conditions in the general case could be quite difficult.

4. Comonotone belief functions

In this section, we focus on belief functions, a particular family of lower probabilities. This family is general enough to include many models used in practice (possibility distributions, p-boxes, ...), but specific enough to offer some computational and practical advantages. After reminding some basics, we study first sufficient conditions to ensure that a joint comonotone belief function can be built from given marginals, and then necessary conditions that a joint comonotone belief function must satisfy to be comonotone.

4.1. Basic definitions

Belief functions correspond to coherent lower probabilities satisfying the additional property of complete monotonicity, as defined below.

Definition 22. *A coherent lower probability \underline{P} on $\mathcal{P}(\Omega)$ is called n -monotone if and only if:*

$$\underline{P}(\cup_{i=1}^p A_i) \geq \sum_{\emptyset \neq I \subseteq \{1, \dots, p\}} (-1)^{|I|+1} \underline{P}(\cap_{i \in I} A_i) \quad (35)$$

for all $2 \leq p \leq n$ and all $A_1, \dots, A_p \in \mathcal{P}(\Omega)$. A lower probability that is n -monotone for all n is called a belief function, here denoted Bel , and its conjugate upper probability is called a plausibility function, here denoted Pl .

A belief function and its conjugate plausibility function are coherent lower and upper probabilities and satisfy $\text{Bel}(A) \leq \text{Pl}(A)$ for all $A \subseteq \Omega$. Using the so-called Möbius inverse, they define a *mass distribution* [7] in the following way:

$$m(A) = \sum_{E \subseteq A} (-1)^{|A \setminus E|} \text{Bel}(E) \quad \forall A \subseteq \Omega. \quad (36)$$

A mass distribution $m : \mathcal{P}(\Omega) \rightarrow [0, 1]$ satisfies $m(\emptyset) = 0$ and $\sum_{E \subseteq \Omega} m(E) = 1$. Conversely, any mass distribution defines a pair of belief and plausibility functions by:

$$\text{Bel}(A) = \sum_{E \subseteq A} m(E) \quad \forall A \subseteq \Omega, \quad (37)$$

$$\text{Pl}(A) = \sum_{E \cap A \neq \emptyset} m(E) \quad \forall A \subseteq \Omega. \quad (38)$$

Note that the non-negativity of the mass m is characteristic of belief functions, in the sense that the value of Eq. (36) will be positive if and only if it is applied to a completely monotone lower probability.

Definition 23. [19] *Given a belief function Bel with mass distribution m , the elements $E \subseteq \Omega$ with positive mass, $m(E) > 0$, are called focal elements of Bel , and we will denote by \mathcal{F} the set of focal elements. The union of all focal elements is called the core of Bel :*

$$\text{Core}(\text{Bel}) = \bigcup_{E \in \mathcal{F}} E. \quad (39)$$

As previously, we consider belief functions defined on the power set of the finite set $\mathcal{X} \times \mathcal{Y} \subseteq \mathbb{R}^2$. We will call them *joint belief functions*. Furthermore, every joint belief function Bel with mass distribution m defines two *marginal belief functions* $\text{Bel}_{\mathcal{X}}$ and $\text{Bel}_{\mathcal{Y}}$ on $\mathcal{X} \subseteq \mathbb{R}$ and $\mathcal{Y} \subseteq \mathbb{R}$, respectively, with associated mass distributions $m_{\mathcal{X}}$ and $m_{\mathcal{Y}}$:

$$m_{\mathcal{X}}(A) = \sum_{E: E^{\downarrow \mathcal{X}}=A} m(E) \text{ and } m_{\mathcal{Y}}(B) = \sum_{E: E^{\downarrow \mathcal{Y}}=B} m(E) \quad (40)$$

for every $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$, and where

$$E^{\downarrow \mathcal{X}} = \{x \in \mathcal{X} : \exists y \in \mathcal{Y} \text{ such that } (x, y) \in E\} \quad (41)$$

(resp., $E^{\downarrow \mathcal{Y}}$) denotes the projection of E on the space \mathcal{X} (resp., \mathcal{Y}). Two important models to which we will devote particular attention and that constitute special cases of belief functions are univariate p-boxes and possibility measures.

From now on we will use the following notation: if E is a finite set of an ordered space, \underline{e} and \bar{e} denote its minimum and maximum, that is

$$\underline{e} = \min E \text{ and } \bar{e} = \max E. \quad (42)$$

We will also use the term interval to denote a subset containing all elements between its minimum and maximum, that is E is an interval if $e \in E$ for all $\underline{e} \leq e \leq \bar{e}$.

It is known that, starting with a univariate p-box, the lower probability defined by Eq. (12) is not only coherent but it is also a belief function [5, 13]. This means that the p-box (\underline{E}, \bar{F}) and the belief function Bel defined by Eq. (12) contain the same probabilistic information. We will therefore associate a p-box (\underline{E}, \bar{F}) with the belief function given by Eq. (12). Furthermore, we will use the term *focal element of the p-box* to refer to the focal elements of its associated belief function. It is known [5] that the focal elements of a p-box, named E_1, \dots, E_k , are intervals that can be ordered such that $\underline{e}_i \leq \underline{e}_{i+1}$ and $\bar{e}_i \leq \bar{e}_{i+1}$. This is a characteristic property, in the sense that any set of focal elements E_1, \dots, E_k that are intervals and that can be ordered in this way induce a belief function that can be obtained from a p-box using Eq. (12) (we refer to [5, 13] for detailed proofs and algorithms). Figure 3 provides an illustration of a univariate p-box and its corresponding focal elements. Whenever referring to focal elements of p-boxes, we will consider that they are indexed according to this ordering.

Possibility measures are other particular instances of plausibility functions whose focal sets are nested. Possibility measures are usual tools used in both Evidence Theory [19] and Fuzzy Set Theory [7, 26].

Definition 24. A possibility measure $\Pi : \mathcal{P}(\Omega) \rightarrow [0, 1]$ is a supremum-preserving function: $\Pi(\cup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i)$ for all I , $A_i \subseteq \Omega$.

The conjugate of a possibility measure:

$$N(A) = 1 - \Pi(A^c) \quad \forall A \subseteq \Omega, \quad (43)$$

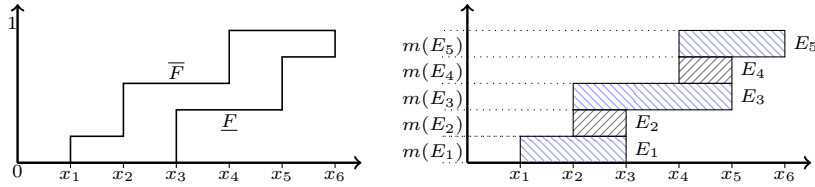


Figure 3: P-box (left) and its associated belief function (right), with focal elements $E_1 = \{x_1, x_2, x_3\}$, $E_2 = \{x_2, x_3\}$, $E_3 = \{x_2, x_3, x_4, x_5\}$, $E_4 = \{x_4, x_5\}$ and $E_5 = \{x_4, x_5, x_6\}$.

called necessity measure, is a belief function. Its focal elements are nested: if E_1 and E_2 are focal elements of N , then either $E_1 \subseteq E_2$ or $E_2 \subseteq E_1$. In fact, since in this work we are dealing with finite sets, there is only a finite number of focal elements E_1, \dots, E_k , and for possibility measures we will consider that they are indexed such that $E_1 \subseteq \dots \subseteq E_k$.

Since any belief function is a coherent lower probability, the definition of comonotonicity for belief functions straightforwardly follows.

Definition 25. *A belief function Bel defined on $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ is comonotone if every $P \in \mathcal{M}(Bel)$ is comonotone. Similarly, a possibility measure Π defined on $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ is comonotone if every $P \in \mathcal{M}(\Pi)$ is comonotone.*

In the case of belief functions, it can be easily proven that $\text{Supp}(Bel) = \text{Core}(Bel)$. The reason is that if $x \in \text{Core}(Bel)$, x belongs to some focal set meaning that the upper probability of x induced by $P \in \mathcal{M}(Bel)$ is positive ($\text{Pl}(\{x\}) > 0$), and therefore there exists some probability $P \in \mathcal{M}(Bel)$ with $P(\{x\}) > 0$. On the other hand, if $P(\{x\}) > 0$ for some $P \in \mathcal{M}(Bel)$, x must belong to some focal set because then $\text{Pl}(\{x\}) > 0$ and therefore it also belongs to $\text{Core}(Bel)$.

Using this fact, Theorem 17 can be directly adapted to the case of belief functions.

Corollary 26. *A belief function Bel defined on $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ is comonotone if and only if $\text{Core}(Bel)$ is an increasing set.*

We may think that the converse of Theorem 18 could hold when dealing with belief functions. However, this is not the case, since the lower probability given in Example 19 is in fact a belief function with focal elements $E_1 = \{(1, 1), (1, 2), (2, 2)\}$ and $E_2 = \{(1, 1), (2, 1), (2, 2)\}$ with $m(E_1) = \alpha$ and $m(E_2) = 1 - \alpha$, respectively.

4.2. Sufficient conditions: from marginals to joint

First note that Example 21 also applies to belief functions, as the marginals \underline{P}_X and \underline{P}_Y used in this example correspond to belief functions induced by the mass distributions

$$m_X(\{1, 2\}) = 0.7, \quad m_X(\{1, 2, 3\}) = 0.3, \quad m_Y(\{1\}) = 0.3, \quad m_Y(\{2\}) = 0.7.$$

Hence the question of when and how can we build a comonotone belief function from given marginals is still relevant.

Sufficient conditions to ensure the existence of some joint comonotone belief functions are instrumental to build such joint, in particular when those conditions are related to well-known models. This is why we start from conditions related to such models, i.e., possibility distributions and p-boxes, to finish with more general conditions.

The first case we investigate is when the marginals are possibility measures. To show that a joint comonotone can always be defined from such marginals, we will first build an increasing set, and then we will show that we can define a joint mass whose focal elements are subsets of this increasing set, and whose marginals are the imposed ones.

Before that, first recall that we can always transform two mass functions m_X, m_Y having focal elements A_1, \dots, A_m and B_1, \dots, B_ℓ into two equivalent mass functions m'_X, m'_Y (in terms of induced belief functions, i.e., $\text{Bel}_X = \text{Bel}'_X$ and $\text{Bel}_Y = \text{Bel}'_Y$) having the same number of focal elements C_1, \dots, C_n and D_1, \dots, D_n and such that $m'_X(C_i) = m'_Y(D_i)$. m'_X, m'_Y are called *commensurate mass functions*, and can be obtained by allowing multiple focal elements to be the same and by splitting the mass of the initial focal elements A_1, \dots, A_m and B_1, \dots, B_ℓ . We refer to [8] for details.

We can therefore assume, without loss of generality, that every two possibility measures have the same number of focal sets and that their masses coincide. The next example illustrates this procedure.

Example 27. Consider the possibility measures m_X, m_Y with the following focal elements:

$$A_1 = \{2\}, A_2 = \{1, 2\}, A_3 = \{1, 2, 3\}, B_1 = \{1, 2\}, B_2 = \{1, 2, 3, 4\},$$

with the following masses:

$$m_X(A_1) = 0.3, m_X(A_2) = 0.5, m_X(A_3) = 0.2, m_Y(B_1) = 0.5, m_Y(B_2) = 0.5.$$

Now, we rewrite the focal elements in the following way

$$\begin{array}{cccc} A_1 & A_2 & A_2 & A_3 \\ B_1 & B_1 & B_2 & B_2 \\ \hline C_1 & C_2 & C_3 & C_4 \\ D_1 & D_2 & D_3 & D_4 \\ \hline m & 0.3 & 0.2 & 0.3 & 0.2 & \blacklozenge \end{array}$$

Proposition 28. Let m_X and m_Y be the masses induced by two possibility measures Π_X, Π_Y . Then, there exists a joint comonotone possibility measure Π inducing a joint mass m , whose marginals are the original possibility measures.

Proof. Assume that the focal elements of m_X and m_Y are given in the following way:

$$\begin{aligned} A_1 &\subseteq A_2 \subseteq \dots \subseteq A_n, \\ B_1 &\subseteq B_2 \subseteq \dots \subseteq B_n, \end{aligned}$$

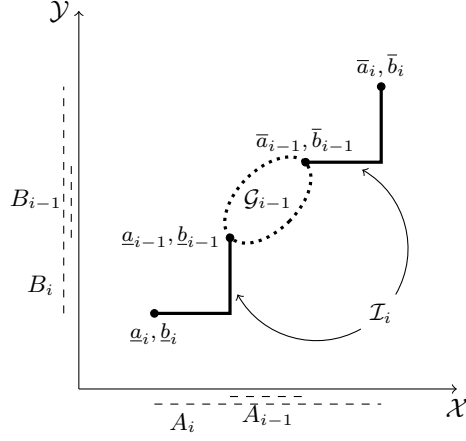


Figure 4: Graphical representation of the construction of the set \mathcal{G}_i , formed by the union of \mathcal{G}_{i-1} and \mathcal{I}_i .

such that for all $i = 1, \dots, n$, $m_X(A_i) = m_Y(B_i)$. We will associate to each pair A_i, B_i a focal element of the joint belief function.

Consider, for $i = 2, \dots, n$ the following increasing set²

$$\begin{aligned} \mathcal{I}_i &= \{(x, \underline{b}_i) : x \in [\underline{a}_i, \underline{a}_{i-1}]\} \cup \{(\underline{a}_{i-1}, y) : y \in [\underline{b}_i, \underline{b}_{i-1}]\} \\ &\quad \cup \{(x, \bar{b}_{i-1}) : x \in [\bar{a}_{i-1}, \bar{a}_i]\} \cup \{(\bar{a}_i, y) : y \in [\bar{b}_{i-1}, \bar{b}_i]\} \\ &= ([\underline{a}_i, \underline{a}_{i-1}] \times \{\underline{b}_i\}) \cup (\{\underline{a}_{i-1}\} \times [\underline{b}_i, \underline{b}_{i-1}]) \\ &\quad \cup ([\bar{a}_{i-1}, \bar{a}_i] \times \{\bar{b}_{i-1}\}) \cup (\{\bar{a}_i\} \times [\bar{b}_{i-1}, \bar{b}_i]) \end{aligned}$$

and the other increasing set³

$$\begin{aligned} \mathcal{G}_1 &= \{(x, \underline{b}_1) : x \in [\underline{a}_1, \bar{a}_1]\} \cup \{(\bar{a}_1, y) : y \in [\underline{b}_1, \bar{b}_1]\} \\ &= ([\underline{a}_1, \bar{a}_1] \times \{\underline{b}_1\}) \cup (\{\bar{a}_1\} \times [\underline{b}_1, \bar{b}_1]). \end{aligned}$$

Then, we define iteratively the sets \mathcal{G}_i , $i = 2, \dots, n$ as $\mathcal{G}_i = \mathcal{I}_i \cup \mathcal{G}_{i-1}$ which by construction are increasing. A graphical illustration of this construction is given in Figure 4.

Now, for all $i = 1, \dots, n$ we define the following set:

$$E_i = \mathcal{G}_i \cap (A_i \times \mathcal{Y}) \cap (\mathcal{X} \times B_i).$$

These sets E_1, \dots, E_n are nested and their union is a subset of \mathcal{G}_n , hence they form an increasing set. By definition, the \mathcal{X} -projection $E_i^{\downarrow \mathcal{X}}$ of E_i is A_i , and its

²Again, we are using the notation introduced in Eq. (42): if A (resp., B) is a finite set, \underline{a}, \bar{a} (resp. \underline{b}, \bar{b}) denote their minimum and maximum elements.

³Similarly, $\underline{a}_i, \bar{a}_i$ and $\underline{b}_i, \bar{b}_i$ denote the minimal and maximal elements of A_i, B_i .

\mathcal{Y} -projection $E_i^{\downarrow\mathcal{Y}}$ is B_i . Finally, we define the joint mass function

$$m(E_i) = m_X(A_i) = m_Y(B_i)$$

that has the desired marginals and which induces, according to Theorem 17, a comonotone joint belief function since its core $\text{Core}(\text{Bel}) \subseteq \mathcal{G}_n$ is increasing. Finally, the focal sets of the constructed joint comonotone belief function are nested, and therefore it is a necessity measure, and the plausibility function is a possibility measure. \square

Example 29. Consider the possibility measures of Example 27. Applying the previous proposition we define the following focal elements for the joint comonotone possibility measure:

$$\begin{aligned} E_1 &= \{(2, 1), (2, 2)\}, & E_2 &= \{(1, 1), (2, 1), (2, 2)\}, \\ E_3 &= \{(1, 1), (2, 1), (2, 2), (2, 3), (2, 4)\}, \\ E_4 &= \{(1, 1), (2, 1), (2, 2), (2, 3), (2, 4), (3, 4)\}, \end{aligned}$$

and the masses are:

$$m(E_1) = 0.3, \quad m(E_2) = 0.2, \quad m(E_3) = 0.3, \quad m(E_4) = 0.2. \blacklozenge$$

Let us now look at the case where marginals are p-boxes. Recall that in this case, the focal elements A_1, \dots, A_m and B_1, \dots, B_ℓ of the two p-boxes $(\underline{F}_X, \overline{F}_X)$ and $(\underline{F}_Y, \overline{F}_Y)$ are ordered, $\underline{a}_i \leq \underline{a}_{i+1}$, $\overline{a}_i \leq \overline{a}_{i+1}$ and $\underline{b}_j \leq \underline{b}_{j+1}$, $\overline{b}_j \leq \overline{b}_{j+1}$ for all $i = 1, \dots, m-1$ and $j = 1, \dots, \ell-1$, and are intervals, in the sense that every A_i, B_j contains all elements in \mathcal{X} and \mathcal{Y} between $\underline{a}_i, \overline{a}_i$ and $\underline{b}_j, \overline{b}_j$, respectively. Similarly to possibility distributions, we can always build two equivalent commensurate mass functions with focal elements C_1, \dots, C_n and D_1, \dots, D_n such that

$$\underline{c}_i \leq \underline{c}_{i+1}, \quad \overline{c}_i \leq \overline{c}_{i+1}, \quad \underline{d}_i \leq \underline{d}_{i+1} \quad \text{and} \quad \overline{d}_i \leq \overline{d}_{i+1} \quad \text{for all } i \in \{1, \dots, n\}. \quad (44)$$

Again, we will assume without loss of generality that every two p-boxes have the same number of focal elements, who are intervals ordered according to their end-points.

Example 30. Consider the belief functions Bel_X and Bel_Y whose focal elements are $A_1 = \{1, 2\}$, $A_2 = \{2, 3\}$, $A_3 = \{3, 4\}$ and $B_1 = \{1, 2\}$, $B_2 = \{2, 3\}$, and whose masses are

$$m_X(A_1) = 0.4, \quad m_X(A_2) = 0.3, \quad m_X(A_3) = 0.3, \quad m_Y(B_1) = 0.6, \quad m_Y(B_2) = 0.4.$$

Then, we can rewrite the focal elements in the following way:

	A_1	A_2	A_2	A_3	
	B_1	B_1	B_2	B_2	
	C_1	C_2	C_3	C_4	
	D_1	D_2	D_3	D_4	
m	0.4	0.2	0.1	0.3	\blacklozenge

The next proposition shows that we can induce a comonotone joint belief function from specific p-boxes.

Proposition 31. *Consider two marginal belief functions Bel_X and Bel_Y with mass distributions m_X, m_Y whose focal elements $\mathcal{A} = \{A_1, \dots, A_n\}$, $\mathcal{B} = \{B_1, \dots, B_n\}$ are such that A_i and B_i are intervals and $m_X(A_i) = m_Y(B_i)$ for all $i = 1, \dots, n$. If \mathcal{A} and \mathcal{B} satisfy the following constraints:*

31.i) $\underline{a}_i \leq \underline{a}_{i+1}$ and $\bar{a}_i \leq \bar{a}_{i+1}$ for every $i = 1, \dots, n$,

31.ii) $\underline{b}_i \leq \underline{b}_{i+1}$ and $\bar{b}_i \leq \bar{b}_{i+1}$ for every $i = 1, \dots, n$,

31.iii) If $\bar{a}_i < \underline{a}_j$, that is if A_j dominates A_i , then $\bar{b}_i \leq \underline{b}_j$, that is B_j also dominates B_i ,

31.iv) If $\bar{b}_i < \underline{b}_j$, that is if B_j dominates B_i , then $\bar{a}_i \leq \underline{a}_j$, that is A_j also dominates A_i ,

then, there exists a joint comonotone belief function Bel such that its marginal masses coincide with m_X and m_Y .

Proof. Like Proposition 28, we will first build an increasing set, and then will define a joint belief function, having the assumed marginals and whose focal elements are subsets of the built increasing set. First, let us consider the set

$$\mathcal{G} = \{(\underline{a}_i, \underline{b}_i), (\bar{a}_i, \bar{b}_i) : i = 1, \dots, n\}$$

and show that it is increasing. First note that for all $i < j$, we have $\underline{a}_i \leq \underline{a}_j, \underline{b}_i \leq \underline{b}_j$ on one hand, and $\bar{a}_i \leq \bar{a}_j, \bar{b}_i \leq \bar{b}_j$ on the other hand, by constraints 31.i) and 31.ii). Hence sets of couples of lower and upper bounds respectively form two increasing sets.

Now let us show by contradiction that their union is also increasing. If \mathcal{G} is not increasing, there is $i > j$ such that $(\underline{a}_i, \underline{b}_i)$ and (\bar{a}_j, \bar{b}_j) are not increasing points. Then, either $\bar{a}_j > \underline{a}_i$ and $\bar{b}_j < \underline{b}_i$ or $\bar{a}_j < \underline{a}_i$ and $\bar{b}_j > \underline{b}_i$. Without loss of generality, assume $\bar{a}_j > \underline{a}_i$ and $\bar{b}_j < \underline{b}_i$ (the reasoning is analogous in the other situation). Then we would have:

$$\underline{a}_j \leq \underline{a}_i < \bar{a}_j \leq \bar{a}_i \tag{45}$$

however, since by hypothesis the inequality $\bar{b}_j < \underline{b}_i$ implies (constraints 31.iv)) $\bar{a}_j \leq \underline{a}_i$, we have a contradiction with Eq. (45). Hence \mathcal{G} is increasing.

Let us now rename by (c_k, d_k) the re-ordered points of \mathcal{G} such that $(c_k, d_k) \leq (c_{k+1}, d_{k+1})$, with an arbitrary indexing for equal points. Clearly, we have $(c_1, d_1) = (\underline{a}_1, \underline{b}_1)$ and $(c_{2n}, d_{2n}) = (\bar{a}_n, \bar{b}_n)$. Let us now consider, for $k = 1, \dots, 2n - 1$, the increasing sets

$$\mathcal{I}_k = \{(x, d_k) : x \in [c_k, c_{k+1}]\} \cup \{(c_{k+1}, y) : y \in [d_k, d_{k+1}]\}.$$

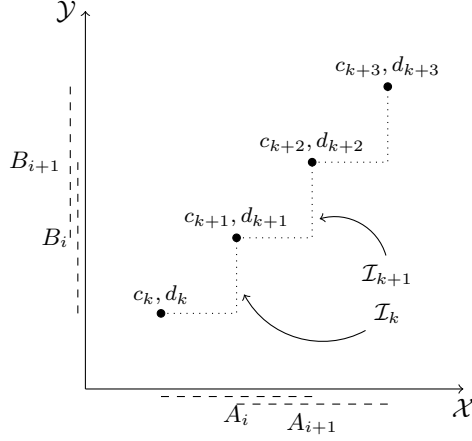


Figure 5: Graphical representation of $E_i = \mathcal{I}_k \cup \mathcal{I}_{k+1}$ and of the increasing set construction in Proposition 31.

Figure 5 illustrates the situation. Clearly, we have that $\cup_{k=1}^{2n-1} \mathcal{I}_k$ is increasing. Now, define the set E_i as

$$E_i = \bigcup_{(\underline{a}_i, \underline{b}_i) \leq (c_k, d_k) < (\bar{a}_i, \bar{b}_i)} \mathcal{I}_k$$

whose \mathcal{X} -projection $E_i^{\downarrow \mathcal{X}}$ and its \mathcal{Y} -projection $E_i^{\downarrow \mathcal{Y}}$ are respectively A_i and B_i , by definition. Finally, we define the mass function:

$$m(E_i) = m_{\mathcal{X}}(A_i) = m_{\mathcal{Y}}(B_i).$$

Then, we have defined a belief function whose core $\text{Core}(\text{Bel}) = \cup_{k=1}^{2n-1} \mathcal{I}_k$ is increasing, and therefore it is comonotone. Moreover, its marginals are $\text{Bel}_{\mathcal{X}}$ and $\text{Bel}_{\mathcal{Y}}$. \square

Example 32. *Let us continue Example 30. Using the previous proposition, we build the following focal sets for the joint belief function:*

$$\begin{aligned} E_1 &= \{(1, 1), (2, 1), (2, 2)\}, & E_2 &= \{(2, 1), (2, 2), (3, 2)\}, \\ E_3 &= \{(2, 2), (3, 2), (3, 3)\}, & E_4 &= \{(3, 2), (3, 3), (3, 4)\}. \end{aligned}$$

Now, we assign the following masses:

$$m(E_1) = 0.4, \quad m(E_2) = 0.2, \quad m(E_3) = 0.1, \quad m(E_4) = 0.3.$$

Then, according to the previous proposition, the joint belief function we have defined is comonotone and its marginals coincide with $\text{Bel}_{\mathcal{X}}$ and $\text{Bel}_{\mathcal{Y}}$. \blacklozenge

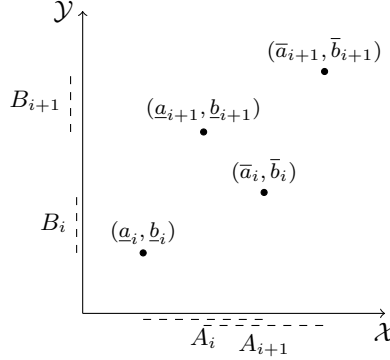


Figure 6: Necessity of Proposition 31 constraints 31.iii) and 31.iv).

Remark 33. *As we have already mentioned, a belief function satisfying constraints 31.i) and 31.ii) in Proposition 31, together with the fact that focal elements should be intervals, is equivalent to a univariate p-box, in the sense that the belief function and its associated p-box contain the same probabilistic information. Constraints 31.iii) and 31.iv) means that we only look at specific p-boxes. ♦*

We know that constraints 31.i) to 31.iv) are sufficient (but not necessary) to ensure the existence of some comonotone joint belief functions, but it would be tempting to drop constraints 31.iii) and 31.iv), in order to retrieve generic p-boxes. However, constraints 31.i) and 31.ii) alone are not sufficient, as illustrated in Figure 6, where the points (a_{i+1}, b_{i+1}) and (\bar{a}_i, \bar{b}_i) not satisfying constraints 31.iii) and 31.iv) are not comonotone.

Similarly, the condition in Proposition 31 that focal elements should be intervals is also essential, as the next example shows.

Example 34. *Consider two mass functions m_X and m_Y with $\mathcal{A} = \{A_1, A_2\}$ and $\mathcal{B} = \{B\}$, where:*

$$A_1 = \{1, 3\}, \quad A_2 = \{2, 4\}, \quad B = \{1, 2, \dots, n-1, n\}, \quad n > 3.$$

\mathcal{A} and \mathcal{B} satisfy all the conditions of Proposition 31, except for being intervals. However, there is no joint comonotone belief functions having those marginals. Indeed, following Proposition 31 proof, such a joint would have two focal elements E_1, E_2 with projections A_1, B and A_2, B , respectively, and such that $E_1 \cup E_2$ is increasing. Now, for all $x \in \{1, 1, \dots, n-1, n\}$, $E_1 \cup E_2$ must contain, at least for one x , one of the pair

- $(x, 1)$ and $(x, 2)$,
- $(x, 1)$ and $(x, 4)$,

- $(x, 3)$ and $(x, 2)$,
- $(x, 3)$ and $(x, 4)$

for E_1, E_2 to have the required projections. If we take every two of those pairs for two different $x < y$ in $\{1, 2, \dots, n-1, n\}$, then they form a non-increasing set. For example, take $(x, 1), (x, 4)$ and $(y, 3), (y, 4)$, we have $(x, 4) \not\leq (y, 3)$. Hence it is not possible to build a comonotone joint belief from m_X and m_Y . ♦

The next proposition somehow puts Propositions 28 and 31 together, showing that a sufficient condition is that we can partition focal elements so that one of the situation (possibility measures or specific p-boxes) is retrieved for each partition.

Proposition 35. *Let Bel_X and Bel_Y be two marginal belief functions with masses m_X and m_Y with focal sets in $\mathcal{F}_X, \mathcal{F}_Y$. Assume that there exists a partition of $\mathcal{F}_X, \mathcal{F}_Y$ into n components $\mathcal{A}_i = \{A_1^i, \dots, A_{n_i}^i\}$ and $\mathcal{B}_i = \{B_1^i, \dots, B_{n_i}^i\}$ such that $A_i^j \in \mathcal{F}_X$, $B_i^j \in \mathcal{F}_Y$, $m_X(A_i^j) = m_Y(B_i^j)$ and satisfying the following constraints:*

35.i) For all $i \neq j$, $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$ and $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$.

35.ii) Components $\mathcal{A}_i, \mathcal{B}_i$ can be ordered so that

$$\overline{\mathcal{A}}_i = \max_{a \in A_j^i} a \leq \underline{\mathcal{A}}_{i+1} = \min_{a \in A_j^{i+1}} a \quad (46)$$

and

$$\overline{\mathcal{B}}_i = \max_{b \in B_j^i} b \leq \underline{\mathcal{B}}_{i+1} = \min_{b \in B_j^{i+1}} b. \quad (47)$$

35.iii) For all i , $\mathcal{A}_i, \mathcal{B}_i$ satisfy either $A_1^i \subseteq \dots \subseteq A_{n_i}^i$ and $B_1^i \subseteq \dots \subseteq B_{n_i}^i$ or conditions 31.i)-31.iv) of Proposition 31.

Then, there exists a joint comonotone belief function Bel such that its marginal masses coincide with m_X and m_Y .

Proof. First, note that due to constraint 35.iii), it is possible to build for every $\mathcal{A}_i, \mathcal{B}_i$ an increasing set \mathcal{G}_i included in $\bigcup_{j=1}^{n_i} A_j^i \times \bigcup_{j=1}^{n_i} B_j^i$. For this, we just have to apply either the procedure of Proposition 28 or 31 to $\mathcal{A}_i, \mathcal{B}_i$.

Now, for all $i < j$, we have that any $(a, b) \leq (c, d)$ for any $(a, b) \in \mathcal{G}_i$ and $(c, d) \in \mathcal{G}_j$, since constraint 35.ii) and the definition of \mathcal{G}_i tell us that

$$(a, b) \leq (\overline{\mathcal{A}}_i, \overline{\mathcal{B}}_i) \leq (\underline{\mathcal{A}}_j, \underline{\mathcal{B}}_j) \leq (c, d).$$

This is sufficient to show that if sets \mathcal{G}_i are increasing, so is $\bigcup_{i=1}^n \mathcal{G}_i$.

Using either the techniques of Proposition 28 or 31 within \mathcal{G}_i , we can then build a focal element $E_{i,j}$ included in \mathcal{G}_i , such that its projections are A_i^j, B_i^j and give him the mass

$$m(E_{i,j}) = m_X(A_i^j) = m_Y(B_i^j).$$

As sets $E_{i,j}$ are subsets of $\bigcup_{i=1}^n \mathcal{G}_i$, they form an increasing set, and have m_X, m_Y for marginals. □

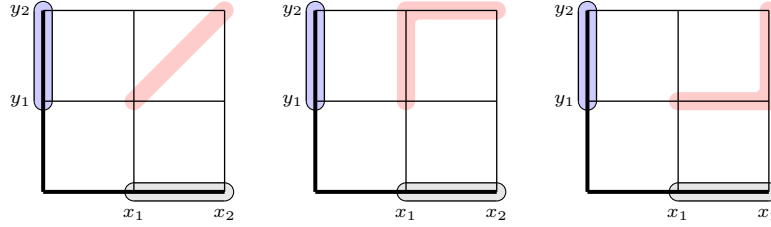


Figure 7: Non-uniqueness of joint comonotone belief function.

Although in this subsection we have some situations in which a joint comonotone belief function exists with fixed marginals, such joint comonotone belief function may not be unique, as next example shows.

Example 36. Consider the marginal belief functions Bel_X and Bel_Y with mass distributions m_X and m_Y , given by:

$$m_X(\{1, 2\}) = m_Y(\{1, 2\}) = 1.$$

In this case, we can define at least three joint belief functions⁴ that are comonotone. If we denote their masses by m , m' and m'' , they are given by:

$$m(\{(1, 1), (2, 2)\}) = m'(\{(1, 1), (2, 2), (1, 2)\}) = m''(\{(1, 1), (2, 2), (2, 1)\}) = 1.$$

This shows that the joint comonotone belief function is not unique. Figure 7 illustrates the three mentioned solutions. \blacklozenge

4.3. Necessary conditions: from joints to marginals

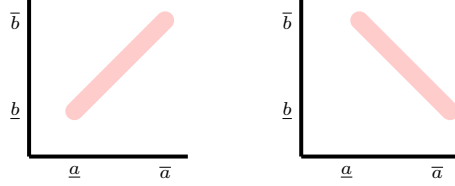
We now examine the properties satisfied by comonotone belief functions, and their impacts on the associated marginal belief functions. This will give us necessary conditions that marginals must satisfy if we want to build a comonotone joint. Of course, all those necessary conditions are satisfied in Propositions 28, 31 and 35. Again, as one of the goal of this section is to offer practical tools to check comonotonicity, we will start with easier conditions and will proceed towards more complex ones.

A first condition is given by the following lemma.

Lemma 37. Let Bel be a joint comonotone belief function, and denote by $\mathcal{E} \subseteq \mathcal{F}$ an arbitrary subset of its focal elements. Using notations

$$\begin{aligned} \bar{a} &= \max\{x : \exists y \text{ s.t. } \exists E \in \mathcal{E} \text{ with } (x, y) \in E\}, \\ \underline{a} &= \min\{x : \exists y \text{ s.t. } \exists E \in \mathcal{E} \text{ with } (x, y) \in E\}, \\ \bar{b} &= \max\{y : \exists x \text{ s.t. } \exists E \in \mathcal{E} \text{ with } (x, y) \in E\}, \\ \underline{b} &= \min\{y : \exists x \text{ s.t. } \exists E \in \mathcal{E} \text{ with } (x, y) \in E\}, \end{aligned}$$

⁴There is actually an infinity of possible solutions.



Lemma 37 satisfied. Lemma 37 not satisfied.

Figure 8: Illustration of Lemma 37 for one focal element.

then there are two (not-necessarily distinct) focal elements $\underline{E}, \overline{E} \in \mathcal{E}$ such that $(\underline{a}, \underline{b}) \in \underline{E}$ and $(\overline{a}, \overline{b}) \in \overline{E}$.

Proof. Let us proceed by contradiction, and assume that there is no focal set in \mathcal{E} containing the element $(\underline{a}, \underline{b})$. This means that there are two elements x^* and y^* such that

$$\begin{aligned} \exists E_1 \in \mathcal{E} \text{ such that } (x^*, \underline{b}) \in E_1, \\ \exists E_2 \in \mathcal{E} \text{ such that } (\underline{a}, y^*) \in E_2, \end{aligned}$$

with E_1, E_2 distinct. By definition, $\underline{a} < x^*$ and $\underline{b} < y^*$. Yet, if Bel is comonotone, $(x^*, \underline{b}), (\underline{a}, y^*)$ cannot both belong to $\text{Core}(\text{Bel})$ since $(\underline{a}, y^*) \not\leq (x^*, \underline{b})$ and $(x^*, \underline{b}) \not\leq (\underline{a}, y^*)$, hence a contradiction.

Analogously, we can prove that there is a focal set in \mathcal{E} containing $(\overline{a}, \overline{b})$. \square

Lemma 37 implies that when \mathcal{E} reduces to one single element E , then E must contain its two “extremal” elements. Hence the need to consider elements $(\underline{a}_i, \underline{b}_i)$ and $(\overline{a}_i, \overline{b}_i)$ in Section 4.2. In practice, this also gives us an easy mean to test comonotonicity of a joint belief function Bel: if the extremal points of its focal elements do not form an increasing set, then Bel is not comonotone. Figure 8 illustrates the simple situation considering one focal element E .

We now introduce a technical result, useful to prove the next ones.

Lemma 38. *Let Bel be a joint belief function whose marginals are m_X and m_Y , and $\mathcal{F}, \mathcal{A}, \mathcal{B}$ their respective sets of focal elements. Then, it is always possible to define partitions*

- $\mathcal{F}_1, \dots, \mathcal{F}_n$ of \mathcal{F} ,
- $\mathcal{A}_1, \dots, \mathcal{A}_n$ of \mathcal{A} ,
- $\mathcal{B}_1, \dots, \mathcal{B}_n$ of \mathcal{B} ,

such that for all $i \in \{1, \dots, n\}$ we have

$$\sum_{A \in \mathcal{A}_i} m_X(A) = \sum_{B \in \mathcal{B}_i} m_Y(B) = \sum_{E \in \mathcal{F}_i} m(E). \quad (48)$$

Proof. We will first build the partitions on \mathcal{A} , \mathcal{B} , and then derive from it the partition on $\mathcal{F} = \{E_1, \dots, E_m\}$. Take $n = 1$ and define $\mathcal{A}_1 = \{A_1\}$, $\mathcal{B} = \{B_1\}$. Now, for every $i = 2, \dots, m$, consider the joint focal element E_i and its projections A_i, B_i on \mathcal{X}, \mathcal{Y} sequentially, and apply the following procedure

- if there is $1 \leq j \leq n$ such that $A_i \in \mathcal{A}_j$, then add B_i to \mathcal{B}_j ,
- if there is $1 \leq j \leq n$ such that $B_i \in \mathcal{B}_j$, then add A_i to \mathcal{A}_j ,
- else, instantiate $\mathcal{A}_{n+1} = \{A_i\}$, $\mathcal{B}_{n+1} = \{B_i\}$ and increase n by one,

until reaching E_m . Indeed, for a given E_i new sets $\mathcal{A}_j, \mathcal{B}_j$ are created only if neither A_i nor B_i can be found in some other sets. Conversely, a set E_i will have its projections in elements of the partitions \mathcal{A}, \mathcal{B} having the same index, by construction. Hence, we can build the partition element \mathcal{F}_j such that

$$\mathcal{F}_j = \{E \in \mathcal{F} \mid E^{\downarrow \mathcal{X}} \in \mathcal{A}_j, E^{\downarrow \mathcal{Y}} \in \mathcal{B}_j\}.$$

Then, these partitions satisfy the following property:

$$\sum_{A \in \mathcal{A}_j} m_X(A) = \sum_{E \in \mathcal{F}_j} m(E) = \sum_{B \in \mathcal{B}_j} m_Y(B).$$

□

Example 39. Consider the joint belief function whose focal elements are:

$$E_1 = \{(1, 1), (2, 2)\}, \quad E_2 = \{(1, 1), (1, 2), (2, 2), (3, 2)\}, \quad E_3 = \{(1, 3), (2, 4)\}, \\ E_4 = \{(3, 4), (4, 5)\}, \quad E_5 = \{(3, 5), (4, 5)\}.$$

Considering the following sets

$$A_1 = \{1, 2\}, \quad A_2 = \{1, 2, 3\}, \quad A_3 = \{3, 4\}, \\ B_1 = \{1, 2\}, \quad B_2 = \{3, 4\}, \quad B_3 = \{4, 5\}, \quad B_4 = \{5\},$$

we have that

Focal set	X projection	Y projection
E_1	A_1	B_1
E_2	A_2	B_1
E_3	A_1	B_2
E_4	A_3	B_3
E_5	A_3	B_4

and $\mathcal{A} = \{A_1, A_2, A_3\}$, $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ with masses

$$m_X(A_1) = m(E_1) + m(E_3), \quad m_X(A_2) = m(E_2), \quad m_X(A_3) = m(E_4) + m(E_5), \\ m_Y(B_1) = m(E_1) + m(E_2), \quad m_Y(B_2) = m(E_3), \quad m_Y(B_3) = m(E_4), \\ m_Y(B_4) = m(E_5).$$

Then, applying Lemma 38, we obtain the following partitions of \mathcal{A}, \mathcal{B} :

$$\begin{aligned}\mathcal{A}_1 &= \{A_1, A_2\} \text{ and } \mathcal{B}_1 = \{B_1, B_2\}. \\ \mathcal{A}_2 &= \{A_3\} \text{ and } \mathcal{B}_2 = \{B_3, B_4\}.\end{aligned}$$

We can easily verify that

$$\begin{aligned}m_X(\mathcal{A}_1) + m_X(\mathcal{A}_2) &= m(E_1) + m(E_3) + m(E_2) = m_Y(\mathcal{B}_1) + m_Y(\mathcal{B}_2). \\ m_Y(\mathcal{A}_3) &= m(E_4) + m(E_5) = m_Y(\mathcal{B}_3) + m_Y(\mathcal{B}_4). \blacklozenge\end{aligned}$$

Now, we investigate the properties of the partition defined in Lemma 38 when the joint belief function is comonotone.

Proposition 40. *Let Bel be a joint belief function with marginals m_X and m_Y . Using Lemma 38, it is possible to make a partition of their sets of focal sets \mathcal{F} , \mathcal{A} and \mathcal{B} : $\mathcal{F}_1, \dots, \mathcal{F}_n$, $\mathcal{A}_1, \dots, \mathcal{A}_n$ and $\mathcal{B}_1, \dots, \mathcal{B}_n$ satisfying Eq. (48).*

Let us use the following notation:

$$\begin{aligned}\bar{a}_i &= \max\{x : \exists y \text{ s.t. } \exists E \in \mathcal{F}_i \text{ with } (x, y) \in E\} = \max \mathcal{A}_i. \\ \underline{a}_i &= \min\{x : \exists y \text{ s.t. } \exists E \in \mathcal{F}_i \text{ with } (x, y) \in E\} = \min \mathcal{A}_i. \\ \bar{b}_i &= \max\{y : \exists x \text{ s.t. } \exists E \in \mathcal{F}_i \text{ with } (x, y) \in E\} = \max \mathcal{B}_i. \\ \underline{b}_i &= \min\{y : \exists x \text{ s.t. } \exists E \in \mathcal{F}_i \text{ with } (x, y) \in E\} = \min \mathcal{B}_i.\end{aligned}$$

Then, when the joint belief function is comonotone, it fulfills the following properties:

$$\begin{aligned}\bar{a}_i < \bar{a}_j &\Rightarrow \bar{b}_i \leq \bar{b}_j. \\ \underline{a}_i < \underline{a}_j &\Rightarrow \underline{b}_i \leq \underline{b}_j. \\ \bar{b}_i < \bar{b}_j &\Rightarrow \bar{a}_i \leq \bar{a}_j. \\ \underline{b}_i < \underline{b}_j &\Rightarrow \underline{a}_i \leq \underline{a}_j.\end{aligned}$$

Proof. Applying Lemma 37 to \mathcal{F}_i and \mathcal{F}_j , we know that there are two focal sets $E_i \in \mathcal{F}_i$ and $E_j \in \mathcal{F}_j$ such that $(\bar{a}_i, \bar{b}_i) \in E_i$ and $(\bar{a}_j, \bar{b}_j) \in E_j$, therefore $(\bar{a}_i, \bar{b}_i), (\bar{a}_j, \bar{b}_j)$ are in the core of Bel. Since Bel is comonotone, its core is increasing, hence $\bar{a}_i < \bar{a}_j$ implies $\bar{b}_i \leq \bar{b}_j$. Other implications can be proved similarly. \square

The result gives necessary conditions on the marginals to ensure the existence of a joint comonotonic belief function with fixed marginals. It is at work in both Propositions 31 and 35. However, this condition is necessary but not sufficient, as next example shows.

Example 41. *Consider two marginal belief functions Bel_X and Bel_Y whose mass distributions are defined by:*

$$m_X(\{1, 2\}) = 1; \quad m_Y(\{1\}) = m_Y(\{2\}) = 0.5.$$

Then, we can consider the partitions $\mathcal{A}_1 = \{\{1, 2\}\}$ and $\mathcal{B}_2 = \{\{1\}, \{2\}\}$ that satisfy the necessary conditions given so far, including the one of Proposition 40.

If Bel is a joint belief function with marginals Bel_X and Bel_Y , then both sets $E_1 = \{1, 2\} \times \{1\}$ and $E_2 = \{1, 2\} \times \{2\}$ are necessarily focal elements. Then, both $(2, 1)$ and $(1, 2)$ belong to $\text{Core}(\text{Bel})$, yet their union do not form an increasing set, hence Bel is not comonotone. \blacklozenge

Proposition 42. *Let Bel be a joint comonotone belief function. For every focal element B of Bel_Y , consider the set:*

$$\mathcal{A}_B := \{A \in \mathcal{A} : \exists E \in \mathcal{F} \text{ with projections } A, B\}. \quad (49)$$

Then, for every $A_1, A_2 \in \mathcal{A}_B$ such that $A_1 \cap A_2 = \emptyset$, it holds that:

- $[\underline{a}_1, \bar{a}_1] \cap [\underline{a}_2, \bar{a}_2] \neq \emptyset$,
- or $|B| = 1$.

Proof. Assume that none of the previous conditions hold. Since $|B| > 1$, it satisfies $\underline{b} < \bar{b}$. Without loss of generality, assume that $\bar{a}_1 < \underline{a}_2$. Then, by Lemma 37, (\bar{a}_1, \bar{b}) and $(\underline{a}_2, \underline{b})$ belong to $\text{Core}(\text{Bel})$. However, this contradicts the comonotonicity of the joint belief function.

Then, we conclude that at least one of the properties must be satisfied. \square

The same result holds when exchanging the role of the projections. We omit the proof as it is analogous to the previous one.

Proposition 43. *Let Bel be a joint comonotone belief function. For every focal element A of Bel_X , consider the set:*

$$\mathcal{B}_A := \{B \in \mathcal{B} : \exists E \in \mathcal{F} \text{ with projections } A, B\}. \quad (50)$$

Then, for all $B_1, B_2 \in \mathcal{B}_A$ such that $B_1 \cap B_2 = \emptyset$, it holds that:

- $[\underline{b}_1, \bar{b}_1] \cap [\underline{b}_2, \bar{b}_2] \neq \emptyset$.
- or $|A| = 1$.

These propositions show that if a joint focal element have identical projections on a dimension and disjoint ones on the other dimensions, then either the identical projections should be singleton, or the end-points of the disjoint projections should be well-ordered. Figure 9 illustrates the situation, showing that when Proposition 42 is not satisfied, then extreme points of focal elements are not comonotone.

For the next result, we introduce the notations

$$A^{\uparrow \mathcal{X} \times \mathcal{Y}} = \{E \in \mathcal{F} : E^{\downarrow \mathcal{X}} = A\} \text{ and } B^{\uparrow \mathcal{X} \times \mathcal{Y}} = \{E \in \mathcal{F} : E^{\downarrow \mathcal{Y}} = B\}. \quad (51)$$

We then have the following result.

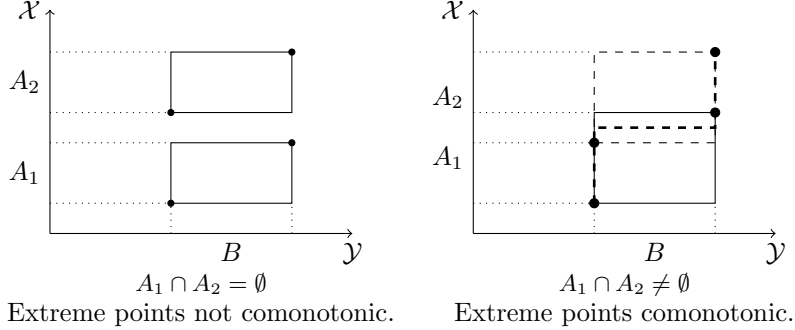


Figure 9: Illustration of Proposition 42 when $|B| > 1$.

Proposition 44. *Let Bel be a joint comonotone belief function, and consider the partitions introduced in Lemma 38. Let $A_1, A_2 \in \mathcal{A}_i$, and assume that $\bar{a}_1 < \underline{a}_2$ and $(A_1^{\uparrow X \times Y})^{\downarrow Y} \cap (A_2^{\uparrow X \times Y})^{\downarrow Y} = \emptyset$. Then, for all $B_1, B_2 \in \mathcal{B}$ such that $B_1 \subseteq (A_1^{\uparrow X \times Y})^{\downarrow Y}$ and all $B_2 \subseteq (A_2^{\uparrow X \times Y})^{\downarrow Y}$ it holds that $\bar{b}_1 \leq \underline{b}_2$.*

Proof. Consider $B_1 \in (A_1^{\uparrow X \times Y})^{\downarrow Y}$ and $B_2 \in (A_2^{\uparrow X \times Y})^{\downarrow Y}$. Then, there are focal elements E_1 and E_2 such that $E_1^{\downarrow X} = A_1$, $E_1^{\downarrow Y} = B_1$, $E_2^{\downarrow X} = A_2$ and $E_2^{\downarrow Y} = B_2$. Now, assume ex-absurdo that $\bar{b}_1 > \underline{b}_2$. Then, since the joint belief function is comonotone, both (\bar{a}_1, \bar{b}_1) and $(\underline{a}_2, \underline{b}_2)$ cannot belong to $\text{Core}(\text{Bel})$ at the same time. However, using Lemma 37:

$$\begin{aligned} (x, y) &= (\bar{a}_1, \bar{b}_1) \in E_1 \subseteq \text{Core}(\text{Bel}). \\ (u, v) &= (\underline{a}_2, \underline{b}_2) \in E_2 \subseteq \text{Core}(\text{Bel}). \end{aligned}$$

This contradicts comonotonicity. \square

This necessary condition is imposed in constraints 31.iii) and 31.iv) of Proposition 31. A similar result holds if we reverse the roles of the X and Y projections.

Proposition 45. *Let Bel be a joint comonotone belief function, and consider the partitions introduced in Lemma 38. Let $B_1, B_2 \in \mathcal{B}_i$, and assume that $\bar{b}_1 < \underline{b}_2$ and $(B_1^{\uparrow X \times Y})^{\downarrow X} \cap (B_2^{\uparrow X \times Y})^{\downarrow X} = \emptyset$. Then, for all $A_1, A_2 \in \mathcal{A}$ such that $A_1 \in (B_1^{\uparrow X \times Y})^{\downarrow X}$ and all $A_2 \in (B_2^{\uparrow X \times Y})^{\downarrow X}$ it holds that $\bar{a}_1 \leq \underline{a}_2$.*

Figure 10 illustrates the conditions required by Proposition 44. We now have both sufficient and necessary conditions for comonotonicity to hold when studying belief functions. However, finding sufficient and necessary conditions easily expressible appears difficult.

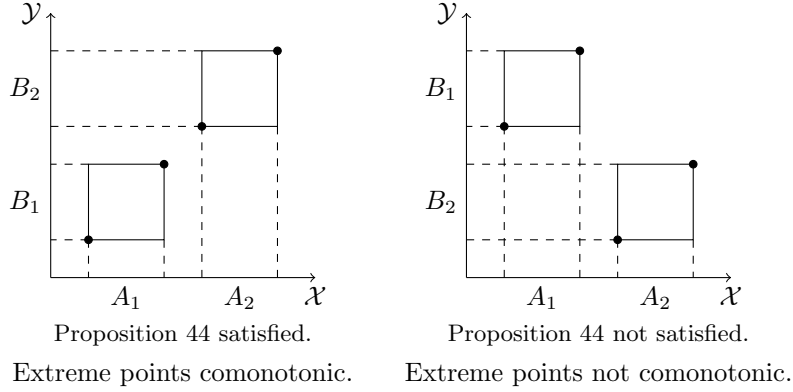


Figure 10: Illustration of Proposition 44.

5. Comonotone p-boxes

We now study the case of comonotone p-boxes, as p-boxes are the straightforward extension of cumulative distributions, for which it is easy to build a comonotone model from given marginals. We may therefore wonder if the same holds for p-boxes. We already know from Proposition 31 that it is not always possible to build a comonotone joint from marginals that are p-boxes, yet we do not know how joint comonotone p-boxes are characterized. This is what we study in this section.

Let us consider a bivariate p-box $(\underline{F}, \overline{F})$ defined on $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} = \{y_1, \dots, y_m\}$. We have already said that a bivariate p-box induces a lower probability on the set \mathcal{K}_2 defined on Eq. (11) by:

$$\underline{P}(A_{x,y}) = \underline{F}(x,y) \text{ and } \underline{P}(A_{x,y}^c) = 1 - \overline{F}(x,y) \quad (52)$$

for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$, and when \underline{P} is coherent, it can be extended to $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ by using the natural extension as in Eq. (12). From now on, we investigate under which condition the natural extension of \underline{P} to $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ is comonotone.

Definition 46. Let $(\underline{F}, \overline{F})$ be a coherent bivariate p-box on $\mathcal{X} \times \mathcal{Y}$. Denote by \underline{P} the lower probability associated with $(\underline{F}, \overline{F})$ by means of Eq. (12). Then, $(\underline{F}, \overline{F})$ is called comonotone whenever \underline{P} is comonotone.

According to the previous definition, $(\underline{F}, \overline{F})$ is comonotone whenever \underline{P} is, or equivalently, if every $P \in \mathcal{M}(\underline{P})$ is comonotone. During this section, and for the sake of simplicity, we shall assume that $n, m > 1$, and

$$\overline{P}(\{x_i\}) > 0 \text{ and } \overline{P}(\{y_j\}) > 0 \text{ for all } i = 1, \dots, n \text{ and } j = 1, \dots, m. \quad (53)$$

From a practical point of view, this assumption has no important consequences, as any element x or y having upper probability zero is deemed impossible, hence

could in principle be removed from \mathcal{X} or \mathcal{Y} . Using property P.3 of Proposition 1 and the previous equation, we obtain that for any $i = 2, \dots, n$:

$$\overline{P}(A_{x_i, y_m}) \geq \overline{P}(\{x_i\}) + \underline{P}(A_{x_{i-1}, y_m}) > \underline{P}(A_{x_{i-1}, y_m}), \quad (54)$$

which is equivalent to $\overline{F}_X(x_i) > \underline{F}_X(x_{i-1})$. Similarly, for every $j = 2, \dots, m$ $\overline{F}_Y(y_j) > \underline{F}_Y(y_{j-1})$.

5.1. Characterization of comonotone bivariate p-boxes

Let us now characterize the conditions under which a bivariate p-box is comonotone. To make the exposition easier, we will gradually introduce sets of necessary conditions, and will then show that they form together a set of sufficient conditions.

Given an increasing set S , we will also often use the fact that its elements can be re-indexed according to the ordering \preceq introduced in Eq. (20) of Definition 9, that is

$$S = \{(u_1, v_1), \dots, (u_k, v_k)\} \quad (55)$$

satisfying $u_1 \leq \dots \leq u_k, v_1 \leq \dots \leq v_k$.

Let us now investigate the properties satisfied by the support of the lower probability associated with a comonotone bivariate p-box.

Proposition 47. *Let $(\underline{F}, \overline{F})$ be a comonotone bivariate p-box defined on $\mathcal{X} \times \mathcal{Y} = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$ satisfying Eq. (53). Denote by \underline{P} the lower probability associated with $(\underline{F}, \overline{F})$ by Eq. (12). If we denote by S the increasing support of \underline{P} , S satisfies the following properties:*

S.1 *The X and Y projections of S are \mathcal{X} and \mathcal{Y} , that is $S^{\downarrow \mathcal{X}} = \mathcal{X}$ and $S^{\downarrow \mathcal{Y}} = \mathcal{Y}$.*

S.2 *If $(u_i, v_i), (u_{i+1}, v_{i+1}) \in S$, then $\underline{F}(u_i, v_i) < \overline{F}(u_{i+1}, v_{i+1})$.*

Proof. Let us start proving property S.1. From Eq. (53), for every $x \in \mathcal{X}$ $\overline{P}(\{x\}) > 0$, which means that there exists $P \in \mathcal{M}(\underline{P})$ and $y \in \mathcal{Y}$ such that $P(\{(x, y)\}) > 0$. Then, $(x, y) \in \text{Supp}(P) \subseteq \text{Supp}(\underline{P})$. We conclude that $S^{\downarrow \mathcal{X}} = \mathcal{X}$. Similarly, $S^{\downarrow \mathcal{Y}} = \mathcal{Y}$.

Let us consider now property S.2. Let $(u_i, v_i), (u_{i+1}, v_{i+1}) \in S$. Since $(u_{i+1}, v_{i+1}) \in S$, $\overline{P}(\{(u_{i+1}, v_{i+1})\}) > 0$, and using P.3 and the monotonicity property P.1 of Proposition 1:

$$\begin{aligned} \underline{F}(u_i, v_i) &< \underline{F}(u_i, v_i) + \overline{P}(\{(u_{i+1}, v_{i+1})\}) = \underline{P}(A_{u_i, v_i}) + \overline{P}(\{(u_{i+1}, v_{i+1})\}) \\ &\leq \overline{P}(A_{(u_i, v_i)} \cup \{(u_{i+1}, v_{i+1})\}) \leq \overline{P}(A_{u_{i+1}, v_{i+1}}) = \overline{F}(u_{i+1}, v_{i+1}). \end{aligned}$$

□

This proposition also tells us that if $(x_i, y_j) \in S$, the support of \underline{P} , then at least one of the element $(x_{i+1}, y_j), (x_i, y_{j+1}), (x_{i+1}, y_{j+1})$ must also be in S . The reason is that, if neither (x_i, y_{j+1}) nor (x_{i+1}, y_j) belong to S , then Condition S.1 implies that there exist $x^* \geq x_{i+1}, y^* \geq y_{i+1}$ satisfying $(x^*, y_{j+1}), (x_{i+1}, y^*) \in S$.

Therefore, since S is increasing and $x_{i+1} \leq x^*$, it must hold $y^* \leq y_{j+1}$, so we conclude that $y^* = y_{j+1}$ and therefore $(x_{i+1}, y_{j+1}) \in S$.

The next proposal indicates some constraints a comonotone bivariate p-box must satisfy when $(x_i, y_j) \in S$ but (x_{i+1}, y_j) or (x_i, y_{j+1}) are not.

Proposition 48. *Let $(\underline{F}, \overline{F})$ be a comonotone bivariate p-box defined on $\mathcal{X} \times \mathcal{Y} = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$ satisfying Eq. (53). Denote by \underline{P} the lower probability associated with $(\underline{F}, \overline{F})$ by Eq. (12). If we denote by S the support of \underline{P} , \underline{F} and \overline{F} satisfy the following conditions:*

S.3 *If $(x_i, y_j) \in S$ and $(x_{i+1}, y_j) \notin S$, then $\underline{F}(x_i, y_j) = \underline{F}_Y(y_j)$ and $\overline{F}(x_i, y_j) = \overline{F}_Y(y_j)$.*

S.4 *If $(x_i, y_j) \in S$ and $(x_i, y_{j+1}) \notin S$, then $\underline{F}(x_i, y_j) = \underline{F}_X(x_i)$ and $\overline{F}(x_i, y_j) = \overline{F}_X(x_i)$.*

Proof. We will only prove the necessary condition S.3, as the proof for S.4 is analogous. Assume that $(x_i, y_j) \in S$, $(x_{i+1}, y_j) \notin S$. By S.1 either $(x_{i+1}, y_{j+1}) \in S$ or $(x_i, y_{j+1}) \in S$.

Let us now prove that $S \cap D_1^{(x_i, y_j)} = \emptyset$. On the one hand, by assumption $(x_{i+1}, y_{j+1}) \notin S$. On the other hand, since $(x_i, y_j) \in S$ and S is increasing, it holds that $(x_{i+1}, y_1), \dots, (x_{i+1}, y_j) \notin S$. Finally, for any (x, y) such that $x > x_i, y \leq j$, it holds that $y < y_{j+1}$ and $x \geq x_{i+1} > x_i$. Therefore, neither $(x, y), (x_i, y_{j+1})$ nor $(x, y), (x_{i+1}, y_{j+1})$ are comonotone, and no element in $D_1^{(x_i, y_j)}$ can belong to S .

Hence for every $(x, y) \in D_1^{(x_i, y_j)}$, $\overline{P}(\{(x, y)\}) = 0$. Using P.2, we conclude that:

$$\overline{P}\left(D_1^{(x_i, y_j)}\right) \leq \sum_{(x, y) \in D_1^{(x_i, y_j)}} \overline{P}(\{(x, y)\}) = 0.$$

Now, let us see that $\overline{P}(D_1^{(x_i, y_j)}) = 0$ implies that $\overline{F}(x_i, y_j) = \overline{F}_Y(y_j)$ and $\underline{F}(x_i, y_j) = \underline{F}_Y(y_j)$. Using this and P.2 of Proposition 1, we have:

$$\overline{P}\left(D_1^{(x_i, y_j)}\right) + \overline{P}(A_{x_i, y_j}) \geq \overline{P}\left(D_1^{(x_i, y_j)} \cup A_{x_i, y_j}\right) \Leftrightarrow \overline{F}(x_i, y_j) \geq \overline{F}_Y(y_j)$$

and since $\overline{F}(x_n, y_j) = \overline{F}_Y(y_j) \geq \overline{F}(x_i, y_j)$ by increasing monotonicity of \overline{F} (Definition 3), we get the desired equality.

Let us prove now the equality between $\underline{F}(x_i, y_j)$ and $\underline{F}_Y(y_j)$. Since $D_1^{(x_i, y_j)} \cap A_{x_i, y_j} = \emptyset$ by construction, P.3 of Proposition 1 implies:

$$\overline{P}\left(D_1^{(x_i, y_j)}\right) + \underline{P}(A_{x_i, y_j}) \geq \underline{P}\left(D_1^{(x_i, y_j)} \cup A_{x_i, y_j}\right) \Leftrightarrow \underline{F}(x_i, y_j) \geq \underline{F}_Y(y_j)$$

and again, we have $\underline{F}(x_n, y_j) = \underline{F}_Y(y_j) \geq \underline{F}(x_i, y_j)$ since \underline{F} is increasing monotone. \square

Note that, as $\underline{F}(x_n, y_j) = \underline{F}_Y(y_j)$ and $\overline{F}(x_n, y_j) = \overline{F}_Y(y_j)$, Proposition 48 also implies that if $(x_i, y_j) \in S$ and $(x_{i+1}, y_j) \notin S$, then

- $\underline{F}(x_i, y_j) = \underline{F}(x_{i+1}, y_j) = \dots = \underline{F}(x_n, y_j)$ and
- $\overline{F}(x_i, y_j) = \overline{F}(x_{i+1}, y_j) = \dots = \overline{F}(x_n, y_j)$,

due to the increasing monotonicity of the bivariate p-box. The next proposition shows that when both (x_i, y_j) and $(x_{i+1}, y_{j+1}) \in S$, then a bivariate p-box has to be precise on (x_i, y_j) to be comonotone.

Proposition 49. *Let $(\underline{F}, \overline{F})$ be a comonotone bivariate p-box defined on $\mathcal{X} \times \mathcal{Y} = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$ satisfying Eq. (53). Denote by \underline{P} the lower probability associated with $(\underline{F}, \overline{F})$ by Eq. (12). If we denote by S the support of \underline{P} , then \underline{F} and \overline{F} satisfy the following condition:*

S.5 If $(x_i, y_j) \in S$ and $(x_{i+1}, y_{j+1}) \in S$, then $\underline{F}(x_i, y_j) = \overline{F}(x_i, y_j)$.

Proof. We will work by contradiction, showing that if $\underline{F}(x_i, y_j) < \overline{F}(x_i, y_j)$, we can build a non-comonotone precise probability within the set $\mathcal{M}(\underline{P})$ induced by the bivariate p-box.

First note that if $(x_i, y_j), (x_{i+1}, y_{j+1}) \in S$, either (x_i, y_{j+1}) or (x_{i+1}, y_j) do not belong to S . Assume without loss of generality that $(x_{i+1}, y_j) \notin S$. By S.3, it holds that

$$\underline{F}(x_i, y_j) = \underline{F}(x_{i+1}, y_j) \text{ and } \overline{F}(x_i, y_j) = \overline{F}(x_{i+1}, y_j). \quad (56)$$

Since $\overline{P}(D_1^{(x_i, y_j)}) = 0$ (see proof of Proposition 48), any probability $P \in \mathcal{M}(\underline{P})$ must have a cumulative distribution F_P such that

$$F_P(x_n, y_j) = \dots = F_P(x_{i+1}, y_j) = F_P(x_i, y_j). \quad (57)$$

Since \underline{P} is coherent, there exist $P_1, P_2 \in \mathcal{M}(\underline{P})$ satisfying:

$$P_1(\{(x_{i+1}, y_{j+1})\}) = \overline{P}(\{(x_{i+1}, y_{j+1})\}) > 0 \text{ and } F_{P_2}(x_i, y_j) = \underline{F}(x_i, y_j).$$

If we take $P = \frac{P_1 + P_2}{2} \in \mathcal{M}(\underline{P})$, it holds that:

$$P(\{(x_{i+1}, y_{j+1})\}) \geq \frac{1}{2} P_1(\{(x_{i+1}, y_{j+1})\}) > 0,$$

If we now assume that $\underline{F}(x_i, y_j) < \overline{F}(x_i, y_j)$, it also holds that:

$$\begin{aligned} F_P(x_i, y_j) &= \frac{F_{P_1}(x_i, y_j) + F_{P_2}(x_i, y_j)}{2} = \frac{F_{P_1}(x_i, y_j) + \underline{F}(x_i, y_j)}{2} \\ &< \frac{F_{P_1}(x_i, y_j) + \overline{F}(x_i, y_j)}{2} \leq \overline{F}(x_i, y_j). \end{aligned}$$

Let us now build, from F_P , a new cumulative distribution F^* such that $\underline{F} \leq F^* \leq \overline{F}$, and giving a non-null mass to (x_{i+1}, y_j) . Take ε such that:

$$0 < \varepsilon < \min\{P(\{(x_{i+1}, y_{j+1})\}), \overline{F}(x_i, y_j) - F_P(x_i, y_j)\}, \quad (58)$$

and define F^* by

$$F^*(x, y) = \begin{cases} F_P(x, y) + \varepsilon & \text{if } y = y_j, x > x_i, \\ F_P(x, y) & \text{otherwise.} \end{cases}$$

F^* satisfies the following properties:

- a) $F^*(x, y) \geq F_P(x, y) \geq \underline{F}(x, y)$ and by Eqs. (56), (57) and (58):

$$F^*(x_{i+k}, y_j) = F_P(x_{i+k}, y_j) + \varepsilon = F_P(x_i, y_j) + \varepsilon < \overline{F}(x_i, y_j) = \overline{F}(x_{i+k}, y_j),$$

and therefore $F^* \leq \overline{F}$. Then, if we denote by P_{F^*} the probability associated with F^* , it holds that $P_{F^*} \in \mathcal{M}(\underline{P})$.

- b) Let us now compute the probability P_{F^*} associated with F^* . For this aim, note that:

$$P_{F^*}(\{(x_i, y_j)\}) = F^*(x_i, y_j) + F^*(x_{i-1}, y_{j-1}) - F^*(x_{i-1}, y_j) - F^*(x_i, y_{j-1}),$$

where $F^*(x_i, y_0) = F^*(x_0, y_j) = 0$.

$$\begin{aligned} P_{F^*}(\{(x_{i+1}, y_{j+1})\}) &= F^*(x_{i+1}, y_{j+1}) + F^*(x_i, y_j) - F^*(x_{i+1}, y_j) - F^*(x_i, y_{j+1}) \\ &= F_P(x_{i+1}, y_{j+1}) + F_P(x_i, y_j) - F_P(x_{i+1}, y_j) - (F_P(x_i, y_{j+1}) + \varepsilon) \\ &= P(\{(x_{i+1}, y_{j+1})\}) - \varepsilon. \\ P_{F^*}(\{(x_{i+1}, y_j)\}) &= F^*(x_{i+1}, y_j) + F^*(x_i, y_{j-1}) - F^*(x_{i+1}, y_{j-1}) - F^*(x_i, y_j) \\ &= (F_P(x_{i+1}, y_j) + \varepsilon) + F_P(x_i, y_{j-1}) - F_P(x_{i+1}, y_{j-1}) - F_P(x_i, y_j) \\ &= P(\{(x_{i+1}, y_j)\}) + \varepsilon. \end{aligned}$$

Note that $P_{F^*}(\{(x_{i+1}, y_{j+1})\}) = P(\{(x_{i+1}, y_{j+1})\}) - \varepsilon > 0$ from Eq. (58). Finally, for any other (x, y) different than $(x_{i+1}, y_{j+1}), (x_{i+1}, y_j)$, it can be easily proven that $P_{F^*}(\{(x, y)\}) = P(\{(x, y)\})$.

From Eq. (58) it holds that $\varepsilon < P(\{(x_{i+1}, y_{j+1})\})$, hence $P_{F^*}(\{(x_{i+1}, y_j)\}) > 0$ and $P_{F^*}(\{(x_{i+1}, y_{j+1})\}) > 0$, thus we reach a contradiction. \square

Finally, we present the characterization of comonotone bivariate p-boxes.

Theorem 50. *Let $(\underline{F}, \overline{F})$ be a coherent bivariate p-box defined on $\mathcal{X} \times \mathcal{Y} = \{x_1, \dots, x_n\} \times \{y_1, \dots, y_m\}$ satisfying Eq. (53). Then, it is comonotone if and only if there is an increasing set $S \subseteq \mathcal{X} \times \mathcal{Y}$ satisfying properties S.1–S.5.*

Proof. Only if: Taking $S = \text{Supp}(\underline{P})$, S is an increasing set and, by Propositions 47, 48 and 49, it satisfies properties S.1 to S.5.

If: assume that there exists an increasing set S satisfying the required properties, and let us prove that $(\underline{F}, \overline{F})$ is comonotone. To this aim, we shall use the characterization of Theorem 17, and we will prove that $\text{supp}(\underline{P})$ is an increasing set because $\text{supp}(\underline{P}) \subseteq S$ by showing that $\overline{P}(\{(x, y)\}) = 0$ for every $(x, y) \notin S$.

Take $(x_i, y_j) \notin S$. Then, using S.1, there exists either $x < x_i$ or $y < y_j$ such that $(x, y_j) \in S$ or $(x_i, y) \in S$. Assume that we are in the first case; the second case follows by analogy. Denote by:

$$x^* = \max\{x \in \mathcal{X} : (x, y_j) \in S\}.$$

We know that $x^* \neq x_i$, and it holds that $x^* < x_i$. The reason is that, since $(x^*, y_j) \in S$ and $(x_i, y_j) \notin S$, from S.1 there exists $y^* > y_j$ such that $(x_i, y^*) \in S$. However, if $x^* > x_i$, then (x_i, y^*) and (x^*, y_j) would not be increasing, but both belong to S , a contradiction.

There are two cases:

1. $\underline{F}(x^*, y_j) = \overline{F}(x^*, y_j)$. By the definition of x^* , we can apply Condition S.3, which implies also that:

$$\underline{F}(x^*, y_j) = \overline{F}(x^*, y_j) = \underline{F}(x_i, y_j) = \overline{F}(x_i, y_j) = \underline{F}_Y(y_j) = \overline{F}_Y(y_j). \quad (59)$$

For any $P \in \mathcal{M}(\underline{P})$, it holds that

$$F_P(x^*, y_j) = \dots = F_P(x_{i-1}, y_j) = F_P(x_i, y_j) = \dots = F_P(x_n, y_j), \quad (60)$$

and this value is also equal to those of Eq. (59). Consequently, since F_P is componentwise increasing and needs to satisfy the rectangle inequality, we have $F_P(x^*, y_{j-1}) = F_P(x_i, y_{j-1})$. Indeed, applying the rectangle inequality to $(x_n, y_j), (x^*, y_{j-1})$, we have⁵:

$$F_P(x_n, y_j) + F_P(x^*, y_{j-1}) - F_P(x^*, y_j) - F_P(x_n, y_{j-1}) \geq 0,$$

but since $F_P(x_n, y_j) = F_P(x^*, y_j)$, we get $F_P(x^*, y_{j-1}) \geq F_P(x_n, y_{j-1})$. However, by monotonicity, we conclude that:

$$F_P(x^*, y_{j-1}) = \dots = F_P(x_i, y_{j-1}) = \dots = F_P(x_n, y_{j-1}). \quad (61)$$

Let us now compute $P(\{(x_i, y_j)\})$, for which we only need to apply the rectangle inequality to the elements $(x_i, y_j), (x_{i-1}, y_{j-1})$ and use Eqs. (60) and (61):

$$\begin{aligned} P(\{(x_i, y_j)\}) &= F_P(x_i, y_j) + F_P(x_{i-1}, y_{j-1}) - F_P(x_{i-1}, y_j) - F_P(x_i, y_{j-1}) \\ &= (F_P(x_i, y_j) - F_P(x_i, y_{j-1})) + (F_P(x_{i-1}, y_{j-1}) - F_P(x_{i-1}, y_j)) \\ &= 0. \end{aligned}$$

Since this happens for all $P \in \mathcal{M}(\underline{P})$,

$$\overline{P}(\{(x_i, y_j)\}) = \sup_{P \in \mathcal{M}(\underline{P})} P(\{(x_i, y_j)\}) = 0.$$

⁵If $j = 1$, we trivially consider $F_P(x, y_0) = 0$ for any $x \in \mathcal{X}$, and the similar reasoning can be applied.

2. $\underline{F}(x^*, y_j) < \overline{F}(x^*, y_j)$. Since S is increasing and using S.1, there exists $y \in \mathcal{Y}$, satisfying $y > y_j$ such that $(x_i, y) \in S$. Take:

$$y_l = \min\{y \in \mathcal{Y} : (x_i, y) \in S\}. \quad (62)$$

Note that $y_l > y_{l-1} \geq y_j$ because $(x^*, y_j) \in S$, $x^* < x_i$ and S is increasing. Also, again using S.1, denote by:

$$x_k = \max\{x \in \mathcal{X} : (x, y_{l-1}) \in S\}. \quad (63)$$

Let us now prove that $x^* \leq x_k \leq x_i$:

- On the one hand, we know that $(x^*, y_j), (x_k, y_{l-1}) \in S$. Since S is increasing and $y_j \leq y_{l-1}$, it must hold $x^* \leq x_k$.
- On the other hand, we know that $(x_k, y_{l-1}), (x_i, y_l) \in S$. Since S is increasing and $y_{l-1} < y_l$, it must hold $x_k \leq x_i$.

Using the definition of x_k in Eq. (63), we have $(x_{k+1}, y_{l-1}) \notin S$. Using S.5, it holds that $\underline{F}(x_k, y_{l-1}) = \overline{F}(x_k, y_{l-1})$, and using S.3 it holds that:

$$\begin{aligned} \underline{F}(x_k, y_{l-1}) &= \overline{F}(x_k, y_{l-1}) = \underline{F}(x_{k+1}, y_{l-1}) \\ &= \overline{F}(x_{k+1}, y_{l-1}) = \underline{F}(x_i, y_{l-1}) = \overline{F}(x_i, y_{l-1}). \end{aligned} \quad (64)$$

From this, every $P \in \mathcal{M}(\underline{P})$ satisfies

$$F_P(x_k, y_{l-1}) = F_P(x_{k+1}, y_{l-1}) = \dots = F_P(x_i, y_{l-1})$$

with values coinciding with those of Eq. (64). Now, let us apply the rectangle inequality to the elements $(x_k, y_{l-2}), (x_i, y_{l-1})$:

$$\begin{aligned} 0 &\leq F_P(x_k, y_{l-2}) + F_P(x_i, y_{l-1}) - F_P(x_k, y_{l-1}) - F_P(x_i, y_{l-2}) \\ &= F_P(x_k, y_{l-2}) - F_P(x_i, y_{l-2}), \end{aligned}$$

but since $x_k \leq x_i$ and F_P is monotone, $F_P(x_k, y_{l-2}) = F_P(x_i, y_{l-2})$ holds. Also, by monotonicity:

$$F_P(x_k, y_{l-2}) = F_P(x_{k+1}, y_{l-2}) = \dots = F_P(x_i, y_{l-2}).$$

Iterating this procedure, we obtain that:

$$\begin{aligned} F_P(x_k, y_{l-3}) &= F_P(x_{k+1}, y_{l-3}) = \dots = F_P(x_i, y_{l-3}). \\ &\dots \\ F_P(x_k, y_2) &= F_P(x_{k+1}, y_2) = \dots = F_P(x_i, y_2). \\ F_P(x_k, y_1) &= F_P(x_{k+1}, y_1) = \dots = F_P(x_i, y_1). \end{aligned}$$

Finally, we compute $P_F(\{(x_i, y_j)\})$:

$$\begin{aligned} P_F(\{(x_i, y_j)\}) &= F(x_i, y_j) + F(x_{i-1}, y_{j-1}) - F(x_{i-1}, y_j) - F(x_i, y_{j-1}) \\ &= (F(x_i, y_j) - F(x_i, y_{j-1})) + (F(x_{i-1}, y_{j-1}) - F(x_{i-1}, y_j)) \\ &= 0. \end{aligned}$$

We conclude that for any $P \in \mathcal{M}(\underline{P})$, $P(\{(x_i, y_j)\}) = 0$, so $\overline{P}(\{(x_i, y_j)\}) = 0$.

Thus, for any $(x, y) \notin S$, $\overline{P}(\{(x, y)\}) = 0$, hence $(x, y) \notin \text{Supp}(\underline{P})$. Therefore, $\text{Supp}(\underline{P}) \subseteq S$, hence $\text{Supp}(\underline{P})$ is an increasing set and by Theorem 17, \underline{P} is comonotone. \square

Example 51. Consider the 5-element spaces \mathcal{X} and \mathcal{Y} and the coherent bivariate p -box $(\underline{F}, \overline{F})$ given by:

y_5	[0.2, 0.3]	[0.2, 0.5]	[0.5, 0.5]	[0.7, 0.7]	[1, 1]
y_4	[0.2, 0.3]	[0.2, 0.5]	[0.5, 0.5]	[0.7, 0.7]	[0.7, 0.9]
y_3	[0.2, 0.3]	[0.2, 0.5]	[0.5, 0.5]	[0.6, 0.6]	[0.6, 0.6]
y_2	[0.2, 0.3]	[0.2, 0.5]	[0.5, 0.5]	[0.5, 0.5]	[0.5, 0.5]
y_1	[0.2, 0.2]	[0.2, 0.2]	[0.2, 0.2]	[0.2, 0.2]	[0.2, 0.2]
$[\underline{F}(x_i, y_j), \overline{F}(x_i, y_j)]$	x_1	x_2	x_3	x_4	x_5

It can be seen that the set S given by:

$$S = \{(x_1, y_1), (x_1, y_2), (x_2, y_2), (x_3, y_2), (x_4, y_3), (x_4, y_4), (x_5, y_4), (x_5, y_5)\} \quad (65)$$

is an increasing set and it satisfies conditions S.1 to S.5, so $(\underline{F}, \overline{F})$ is a comonotone bivariate p -box whose associated lower probability \underline{P} is comonotone. Indeed, $\text{Supp}(\underline{P}) = S$. \blacklozenge

Using the previous theorem, we can characterize the form of the probabilities on the credal set of the lower probability associated with a comonotone bivariate p -box.

Corollary 52. Let $(\underline{F}, \overline{F})$ be a comonotone bivariate p -box satisfying Eq. (53) with associated lower probability \underline{P} , and denote by S the support of \underline{P} . Then, every bivariate cumulative distribution function $F \in (\underline{F}, \overline{F})$ satisfies the following conditions:

1. If $(x_i, y_j) \in S$ and $(x_{i+1}, y_j) \notin S$, $F(x_{i+1}, y_j) = \dots = F(x_n, y_j)$.
2. If $(x_i, y_j) \in S$ and $(x_i, y_{j+1}) \notin S$, $F(x_i, y_{j+1}) = \dots = F(x_i, y_m)$.

Furthermore, the support of \underline{P} can be expressed by $S = \{(u_1, v_1), \dots, (u_l, v_l)\}$ such that $u_k \leq u_{k+1}, v_k \leq v_{k+1}$ for every $k = 1, \dots, l-1$, $(u_1, v_1) = (x_1, y_1)$ and $(u_l, v_l) = (x_n, y_m)$. It holds that:

$$\begin{aligned} \mathcal{M}(\underline{P}) &= \{P \in \mathbb{P} \mid \exists \alpha_k \in [\underline{F}(u_k, v_k), \overline{F}(u_k, v_k)] \text{ such that } \alpha_1 \leq \dots \leq \alpha_l = 1, \\ &\quad P(\{(u_1, v_1)\}) = \alpha_1, P(\{(u_k, v_k)\}) = \alpha_k - \alpha_{k-1} \forall k = 2, \dots, l\}. \end{aligned} \quad (66)$$

Proof. Take $F \in (\underline{F}, \overline{F})$. Then, if $(x_i, y_j) \in S$, consider the following cases:

1. $(x_{i+1}, y_j) \notin S$. Since S is increasing, $D_1^{(x_i, y_j)} \cap S = \emptyset$, and therefore $P_F(D_1^{(x_i, y_j)}) = 0$. We conclude that

$$\begin{aligned} F(x_n, y_j) &= P_F(D_1^{(x_i, y_j)} \cup A_{x_i, y_j}) \\ &= P_F(D_1^{(x_i, y_j)}) + P_F(A_{x_i, y_j}) = P_F(A_{x_i, y_j}) = F(x_i, y_j). \end{aligned}$$

Therefore, since F is component-wise increasing,

$$F(x_i, y_j) = F(x_{i+1}, y_j) = \dots = F(x_n, y_j).$$

2. $(x_i, y_{j+1}) \notin S$. With a similar reasoning, it holds that

$$F(x_i, y_j) = F(x_i, y_{j+1}) = \dots = F(x_i, y_m).$$

Now, take $\alpha_k \in [\underline{F}(u_k, v_k), \overline{F}(u_k, v_k)]$ for every $k = 1, \dots, l$ such that $\alpha_1 \leq \dots \leq \alpha_l = 1$, and consider F such that $F(u_k, v_k) = \alpha_k$ for every $k = 1, \dots, l$. Using cases 1 and 2, F defines a unique cumulative distribution function on $\mathcal{X} \times \mathcal{Y}$. Let us compute its associated probability P_F . First of all, it trivially holds that $P_F(\{(u_1, v_1)\}) = P_F(\{(x_1, y_1)\}) = \alpha_1$. For any $(u_k, v_k) = (x_i, y_j) \in S$ with $k \geq 2$, we have the following cases:

- $(u_{k-1}, v_{k-1}) = (x_{i-1}, y_{j-1})$. From 1 and 2, we know that

$$\alpha_{k-1} = F(x_{i-1}, y_{j-1}) = F(x_{i-1}, y_j) = F(x_i, y_{j-1}).$$

Therefore:

$$\begin{aligned} P_F(\{(u_k, v_k)\}) &= F(x_i, y_j) + F(x_{i-1}, y_{j-1}) - F(x_{i-1}, y_j) - F(x_i, y_{j-1}) \\ &= \alpha_k + \alpha_{k-1} - \alpha_{k-1} - \alpha_{k-1} = \alpha_k - \alpha_{k-1}. \end{aligned}$$

- $(u_{k-1}, v_{k-1}) = (x_{i-1}, y_j)$. Since S is increasing, (x_i, y_{j-1}) cannot belong to S . However, by Condition S.1 there exists $x \in \mathcal{X}$, that must satisfy $x < x_i$, such that $(x, y_{j-1}) \in S$. Denote by $x^* = \max\{x \in \mathcal{X} \mid (x, y_{j-1}) \in S\}$, and denote by $\alpha^* = F(x^*, y_{j-1})$. From condition 1 of this corollary, it holds that:

$$\alpha^* = F(x^*, y_{j-1}) = F(x_{i-1}, y_{j-1}) = F(x_i, y_{j-1}).$$

Finally, we compute $P_F(\{(u_k, v_k)\})$:

$$\begin{aligned} P_F(\{(u_k, v_k)\}) &= F(x_i, y_j) + F(x_{i-1}, y_{j-1}) - F(x_{i-1}, y_j) - F(x_i, y_{j-1}) \\ &= \alpha_k + \alpha^* - \alpha_{k-1} - \alpha^* = \alpha_k - \alpha_{k-1}. \end{aligned}$$

- $(u_{k-1}, v_{k-1}) = (x_i, y_{j-1})$. Following the same steps than in the previous case, we obtain $P_F(\{(x_k, u_k)\}) = \alpha_k - \alpha_{k-1}$.

We therefore conclude that:

$$P_F(\{(u_1, v_1)\}) = \alpha_1, \quad P_F(\{(u_k, v_k)\}) = \alpha_k - \alpha_{k-1} \quad \forall k = 2, \dots, l. \quad (67)$$

On the other hand, if we take a probability P satisfying Eq. (67), it defines a cumulative distribution function F_P bounded by \underline{F} and \overline{F} , so $P \in \mathcal{M}(\underline{P})$. \square

Example 53. Let us continue with Example 51. Taking into account the previous result, any cumulative distribution function $F \in (\underline{F}, \overline{F})$ has the following form:

y_5	β	δ	0.5	0.7	1
y_4	β	δ	0.5	0.7	γ
y_3	β	δ	0.5	0.6	0.6
y_2	β	δ	0.5	0.5	0.5
y_1	0.2	0.2	0.2	0.2	0.2
$F(x_i, y_j)$	x_1	x_2	x_3	x_4	x_5

Where $\beta \in [0.2, 0.3]$, $\delta \in [0.2, 0.5]$, $\gamma \in [0.7, 0.9]$ and $\beta \leq \delta \leq \gamma$. Then, the probabilities $P \in \mathcal{M}(\underline{P})$ are given by:

y_5	0	0	0	0	$1 - \gamma$
y_4	0	0	0	0.1	$\gamma - 0.7$
y_3	0	0	0	0.1	0
y_2	$\beta - 0.2$	$\delta - \beta$	$0.5 - \delta$	0	0
y_1	0.2	0	0	0	0
$P(\{(x_i, y_j)\})$	x_1	x_2	x_3	x_4	x_5

We can see how Corollary 52 allows us to compute in a very simple way the probabilities in the credal set of the p-box. \blacklozenge

5.2. Comonotone bivariate p-boxes and belief functions

So far we have characterized the comonotone bivariate p-boxes. We now investigate the lower probability they induce, proving that it is a belief function.

Theorem 54. Let $(\underline{F}, \overline{F})$ be a coherent bivariate p-box defined on $\mathcal{X} \times \mathcal{Y}$. If $(\underline{F}, \overline{F})$ is comonotone, then its associated lower probability \underline{P} by means of Eq. (12) is a belief function.

Proof. Denote the support of \underline{P} by $S = \text{Supp}(\underline{P})$. Since it is increasing, we can express it by $S = \{(u_1, v_1), \dots, (u_l, v_l)\}$ such that $u_k \leq u_{k+1}, v_k \leq v_{k+1}$ for every $k = 1, \dots, l-1$, $(u_1, v_1) = (x_1, y_1)$ and $(u_l, v_l) = (x_n, y_m)$.

If we use the notation $z_k = (u_k, v_k)$ for $k = 1, \dots, l$, and using the partial order defined in Eq. (20) we can define an ordered space (S, \preceq) , where $S = \{z_1, \dots, z_l\}$ and $z_1 \preceq \dots \preceq z_l$. Also, we can define the (univariate) p-box $(\underline{F}_S, \overline{F}_S)$ by:

$$\underline{F}_S(z_k) = \underline{F}(u_k, v_k), \quad \overline{F}_S(z_k) = \overline{F}(u_k, v_k) \quad \forall k = 1, \dots, l.$$

Therefore, $(\underline{F}_S, \overline{F}_S)$ defines a belief function \underline{P}_S using Eq. (12). Let us denote by $E_1, \dots, E_r \subseteq S$ the focal sets of \underline{P}_S , and by m_S its associated mass distribution.

Now let us see that there is a one-to-one correspondence between $\mathcal{M}(\underline{P})$ and $\mathcal{M}(\underline{P}_S)$. On the one hand, take $P \in \mathcal{M}(\underline{P})$, which using Corollary 52 satisfies $P(\{(u_k, v_k)\}) = \alpha_k - \alpha_{k-1}$ for some $\alpha_k \in [\underline{F}(u_k, v_k), \overline{F}(u_k, v_k)] \forall k = 1, \dots, l$, $\alpha_1 \leq \dots \leq \alpha_l$ and $\alpha_0 = 0$. Thus for any such P , we can define a probability

P_S by $P_S(\{z_k\}) = \alpha_k - \alpha_{k-1}$ for every $k = 1, \dots, l$ and associate it to P . This probability satisfies

$$F_{P_S}(z_k) = \alpha_k \in [\underline{F}(u_k, v_k), \overline{F}(u_k, v_k)] = [\underline{F}_S(z_k), \overline{F}_S(z_k)] \quad \forall k = 1, \dots, l.$$

Hence, $P_S \in \mathcal{M}(\underline{P}_S)$.

Conversely, take $P_S \in \mathcal{M}(\underline{P}_S)$, which satisfies $F_{P_S}(z_k) = \alpha_k \in [\underline{F}_S(z_k), \overline{F}_S(z_k)]$ for every $k = 1, \dots, l$. Therefore, we can define F by $F(u_k, v_k) = F(z_k) = \alpha_k$ for every $k = 1, \dots, l$. Using the previous corollary, F is a cumulative distribution function bounded by $(\underline{F}, \overline{F})$, so $P_F \in \mathcal{M}(\underline{P})$.

Now, take $A \in \mathcal{X} \times \mathcal{Y}$. This set can be expressed by $A = (A \cap S) \cup (A \cap S^c)$. Therefore, using monotonicity P.1 and P.3 of Proposition 1, we have:

$$\underline{P}(A \cap S) \leq \underline{P}(A) \leq \underline{P}(A \cap S) + \overline{P}(A \cap S^c) \leq \underline{P}(A \cap S) + \overline{P}(S^c) = \underline{P}(A \cap S),$$

where the last inequality follows because S is the support of \underline{P} , and therefore $\overline{P}(\{(x, y)\}) = 0$ for every $(x, y) \notin S$. Hence, $\underline{P}(A \cap S) = \underline{P}(A)$ for any $A \subseteq \mathcal{X} \times \mathcal{Y}$.

Let us see now that $\underline{P}(A \cap S) = \underline{P}_S(A)$ for every $A \subseteq \mathcal{X} \times \mathcal{Y}$. On the one hand, since \underline{P} is coherent, there exists $P \in \mathcal{M}(\underline{P})$ such that $\underline{P}(A \cap S) = P(A \cap S)$. Also, take $P_S \in \mathcal{M}(\underline{P}_S)$ the probability associated with P . It holds that:

$$\underline{P}(A \cap S) = P(A \cap S) = P_S(A \cap S) \geq \underline{P}_S(A \cap S).$$

On the other hand, there exists $P_S \in \mathcal{M}(\underline{P}_S)$ such that $P_S(A \cap S) = \underline{P}_S(A \cap S)$. If we denote by $P \in \mathcal{M}(\underline{P})$ the probability associated with P_S , it holds that:

$$\underline{P}_S(A \cap S) = P_S(A \cap S) = P(A \cap S) \geq \underline{P}(A \cap S).$$

We conclude that $\underline{P}(A \cap S) = \underline{P}_S(A \cap S)$, and therefore $\underline{P}(A) = \underline{P}_S(A \cap S)$ for every $A \subseteq \mathcal{X} \times \mathcal{Y}$. Finally, for any $A \subseteq \mathcal{X} \times \mathcal{Y}$, since \underline{P}_S is a belief function, it holds that:

$$\underline{P}(A) = \underline{P}_S(A \cap S) = \sum_{\substack{E_k \text{ focal of } \underline{P}_S \\ E_k \subseteq A \cap S}} m_S(E_k) = \sum_{\substack{E_k \text{ focal of } \underline{P}_S \\ E_k \subseteq A}} m_S(E_k),$$

where the last equality follows because every focal set E_k satisfies $E_k \subseteq S$, so $E_k \subseteq A \cap S$ is equivalent to $E_k \subseteq A$. Therefore, \underline{P} is a belief function with focal sets E_1, \dots, E_r and mass distribution m such that $m(E_k) = m_S(E_k)$. \square

Example 55. *Let us continue with Example 51. From the previous theorem we know that \underline{P} is a belief function. Also, following the proof, we can define a totally ordered univariate space (S, \preceq) , where S was given in Eq. (65). Therefore, S can be expressed as $S = \{z_1, \dots, z_8\}$ where $z_1 \prec \dots \prec z_8$. We can define the p -box $(\underline{F}_S, \overline{F}_S)$ on S , that is just the restriction of $(\underline{F}, \overline{F})$ to S , given by:*

S	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8
\underline{F}_S	0.2	0.2	0.2	0.5	0.6	0.7	0.7	1
\overline{F}_S	0.2	0.3	0.5	0.5	0.6	0.7	0.9	1

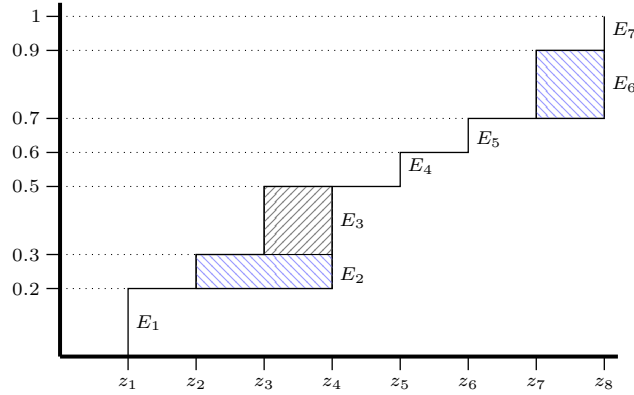


Figure 11: Representation of the p-box $(\underline{E}_S, \overline{F}_S)$ and its focal sets.

The lower probability \underline{P}_S associated with $(\underline{E}_S, \overline{F}_S)$ is a belief function, and its focal sets are given by:

$$\begin{aligned} E_1 &= \{z_1\}, & E_2 &= \{z_2, z_3, z_4\}, & E_3 &= \{z_3, z_4\}, & E_4 &= \{z_5\}, \\ E_5 &= \{z_6\}, & E_6 &= \{z_7, z_8\}, & E_7 &= \{z_9\}, \end{aligned}$$

with masses

E_i	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8
$m(E_i)$	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.1

The graphical representation of $(\underline{E}_S, \overline{F}_S)$ and the focal sets of \underline{P} can be seen in Figure 11.

According to the proof of the previous theorem, the focal sets of \underline{P} coincide with those of \underline{E}_S , but written in terms of (x_i, y_j) :

$$\begin{aligned} E_1 &= \{(x_1, y_1)\}, & E_2 &= \{(x_1, y_2), (x_2, y_2), (x_3, y_2)\}, & E_3 &= \{(x_2, y_2), (x_3, y_2)\}, \\ E_4 &= \{(x_4, y_3)\}, & E_5 &= \{(x_4, y_4)\}, & E_6 &= \{(x_5, y_4), (x_5, y_5)\}, & E_7 &= \{(x_5, y_5)\}, \end{aligned}$$

with masses $m(E_i) = m_S(E_i)$. The focal sets are graphically depicted in Figure 12. ♦

6. Conclusions

Comonotonicity (and countermonotonicity) are specific dependency assumptions that have important practical and theoretical roles when using precise probabilities. They correspond to lower/upper bounds of any other joint distributions, and can be used for instance in sensitivity analysis.

This paper investigates the notion of comonotonicity for lower probabilities and sets of probabilities, models that are used when a unique probability cannot

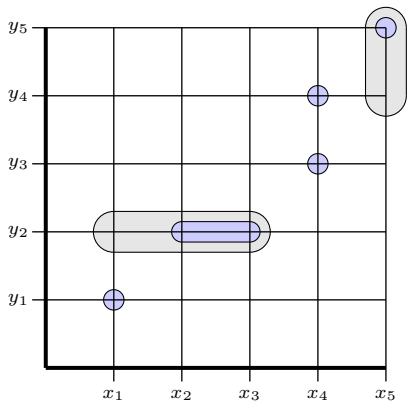


Figure 12: Core of the belief function defined from the bivariate p-box of Example 51.

be identified (due to limited information, etc). Our results show that most equivalent characterizations (increasingness of the support, counter-monotonic events having zero probability) of comonotonicity true in the precise case still hold in the imprecise one. However, only a weaker form of the characterization in terms of cumulative distributions holds: the bivariate p-box associated with a comonotone lower probability can be expressed as the minimum of the marginal p-boxes, yet the converse does not hold in general (the minimum of marginal p-boxes does not produce a comonotonic model).

We have also observed that comonotonicity of a joint model is much harder to obtain and to check in practice when dealing with imprecise probabilities. In particular, it is not always possible to build a comonotonic joint lower probability from given marginals, in contrast with the precise setting. For this reason, we have investigated conditions for which such constructions are possible for specific yet useful models: belief functions and sub-cases of them (p-boxes, possibility distributions). Focusing on particular models allowed us to get practical conditions usable in applications.

In previous works [2, 3] the notion of independence in an imprecise framework was investigated, not even for events but also for gambles. It remains as an open problem to study how and to what extent comonotonicity can be expressed for gambles, and in particular if there is some kind of connection with the Choquet integral [1].

Acknowledgements

We would like to thank the anonymous reviewers for their careful reading of earlier versions of this paper, as well as for their insightful comments that have helped to improve its readability and in one case to correct a result.

This work was carried out and funded in the framework of the Labex MS2T. It was supported by the French Government, through the program “Investments

for the future” managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02). We also acknowledge the Spanish Ministry of Economy and Competitiveness financial support by project TIN2014-59543-P.

References

- [1] G. Choquet. Theory of capacities. *Annales de l’Institut Fourier*, 5:131–295, 1953–1954.
- [2] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy*, 5:165–181, 2000.
- [3] G. de Cooman, E. Miranda, and M. Zaffalon. Independent natural extension. *Artificial*, 175(12–13):1911–1950, 2011.
- [4] D. Denneberg and N. Leufer. Dual volatility and dependence parameters and the copula. *International Journal of Approximate Reasoning*, 48:697–708, 2008.
- [5] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: I generalised p-boxes. *Int. J. of Approximate Reasoning*, 49:649–663, 2008.
- [6] J. Dhaene, M. Denuir, M. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, 31(1):3–33, 2002.
- [7] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.
- [8] D. Dubois and R. Yager. Fuzzy set connectives as combinations of belief structures. *Information Sciences*, 66:245–275, 1992.
- [9] 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.
- [10] S. Ferson, R. B. Nelsen, J. Hajagos, D. J. Berleant, J. Zhang, W. T. Tucker, L. R. Ginzburg, and W. L. Oberkampf. Dependence in probabilistic modeling, dempster-shafer theory, and probability bounds analysis.
- [11] M. J. Goovaerts R. Kaas J. Dhaene, M. Denuit and D. Vyncke. The concept of comonotonicity in actuarial science and finance: applications. *Insurance: Mathematics and Economics*, 31:133–161, 2002.
- [12] E. Jouini and C. Napp. Comonotonic processes. *Insurance: Mathematics and Economics*, 32(2):255–265, 2003.
- [13] E. Kriegler. Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning*, 39(2–3):185–209, 2005.

- [14] E. Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.
- [15] I. Montes and S. Destercke. Comonotone lower probabilities for bivariate and discrete structures. In *ISIPTA '15: Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 207–217, Pescara, 2013. SIPTA.
- [16] I. Montes, E. Miranda, R. Pelessoni, and P. Vicig. Sklar’s Theorem in an imprecise setting. *Fuzzy Sets and Systems*, (278C):48–66, 2015.
- [17] R. Nelsen. *An introduction to Copulas*. Lecture Notes in Statistics, Vol. 139, New York, 2006.
- [18] R. Pelessoni, P. Vicig, I. Montes, and E. Miranda. Bivariate p-boxes. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, to appear, 24(2):229–263, 2016.
- [19] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
- [20] A. Sklar. Fonctions de répartition à n-dimensions et leurs marges. *Publications de l’Institut de Statistique de l’Université de Paris*, 8:229–231, 1959.
- [21] M. Troffaes and S. Destercke. Probability boxes on totally preordered spaces for multivariate modell. *International Journal of Approximate Reasoning*, 52(6):767–791, 2011.
- [22] M.C.M. Troffaes, F.P.A. Coolen, and S. Destercke. A note on learning dependence under severe uncertainty. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 498–507. Springer, 2014.
- [23] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
- [24] Z. Q. Xu. A new characterization of comonotonicity and its application in behavioral finance. *Journal of Mathematical Analysis and Applications*, 418(2):612–625, 2014.
- [25] J. Yang, S. Cheng, and L. Zhang. Bivariate copula decomposition in terms of comonotonicity, coucountermonotonic and independence. *Insurance: Mathematics and Economics*, 39:267–284, 2006.
- [26] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.