International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems

BIVARIATE *P*-BOXES

RENATO PELESSONI
DEAMS Bruno de Finetti, University of Trieste, Piazzale Europa 1
Trieste, I-34127, Italy
renato.pelessoni@econ.units.it

PAOLO VICIG
DEAMS Bruno de Finetti, University of Trieste, Piazzale Europa 1
Trieste, I-34127, Italy
paolo.vicig@econ.units.it

IGNACIO MONTES
Dept. of Statistics and Operational Research, University of Oviedo, Campus of Llamaquique
Oviedo, E-33007, Spain
imontes@uniovi.es

ENRIQUE MIRANDA
Dept. of Statistics and Operational Research, University of Oviedo, Campus of Llamaquique
Oviedo, E-33007, Spain
mirandaenrique@uniovi.es

Received (received date)
Revised (revised date)

A p-box is a simple generalization of a distribution function, useful to study a random number in the presence of imprecision. We propose an extension of p-boxes to cover imprecise evaluations of pairs of random numbers and term them bivariate p-boxes. We analyze their rather weak consistency properties, since they are at best (but generally not) equivalent to 2-coherence. We therefore focus on the relevant subclass of coherent p-boxes, corresponding to coherent lower probabilities on special domains. Several properties of coherent p-boxes are investigated and compared with those of (one-dimensional) p-boxes or of bivariate distribution functions.

Keywords: p-boxes, coherent and 2-coherent lower probabilities, cumulative distribution functions, rectangle inequalities.

1. Introduction and Preliminary Concepts

Uncertainty modelling with imprecise probabilities includes a variety of simplified representations which are especially fit for reasoning with certain specific situations (see for instance Augustin et al.,[1] Chapter 4).

Among these, a *p-box* is a generalization of the cumulative distribution function (cdf) of a random number X. The idea is relatively simple: a p-box $(\underline{F}, \overline{F})$ is a pair of cdfs $\underline{F}$, $\overline{F}$, such that $\underline{F} \leq \overline{F}$.[2] Recalling that a distribution function F for X, $F(x) = P(X \leq x)$, represents the probability P of the events $(X \leq x)$ for any x, the p-box $(\underline{F}, \overline{F})$ supplies a lower bound $\underline{F}$ and an upper bound $\overline{F}$ to these probabilities. Conceptually, this is a straightforward way of imprecisely describing a random number, which works also when X is unbounded, unlike other more general representations with imprecise probabilities or previsions.[3] Operationally, it helps evaluating questions of the type 'how likely it is that X exceeds a certain threshold?', which are essential, for instance, in (industrial or financial) risk analysis. The literature on p-boxes appears to be limited so far, in spite of their usefulness. P-boxes were discussed in Ferson et al.[2], and later appeared in a few other works, including Ferson and Tucker[4], Troffaes and Destercke[5], Troffaes et al.[6], Utkin and Destercke[7].

In this paper we analyze the still largely unexplored generalization of a p-box in order to jointly describe a couple of random numbers (X, Y), i.e. what we shall call *bivariate p-box.* In addition to the motivations for using (univariate) p-boxes, we meet bivariate p-boxes when coming to a joint evaluation for (X, Y) by combining marginals for X and Y, with the latter given in the form of (univariate) p-boxes. We tackle this aspect of bivariate p-boxes in the companion paper[8], while focusing here on foundational aspects that differentiate bivariate from univariate p-boxes.

A core issue is that the same distribution function F may be obtained *in the univariate case* in a number of ways, which are no longer equivalent in higher dimensions. For instance, F corresponds to the restriction of a probability defined on a suitable set, but may be derived also from a coherent lower probability or even a capacity (i.e. a monotone non-decreasing and normalized measure on an algebra). Technically, this is because several different non-additive measures are characterized in the same way on monotone families (chains), that are the sets of events evaluated by a (univariate) distribution function. Further, the infimum and the supremum of a set of (univariate) cdfs are again cdfs, which is *not* necessarily true in higher dimensions. This suggests that a formal definition of bivariate p-box as an ordered pair of distribution functions would be too restrictive. Actually, the first question is precisely how to define a bivariate p-box; secondly, choosing as we are going to do a broad definition, the question arises of identifying those bivariate p-boxes with satisfactory properties.

Prior to this, some preliminary material is supplied. Special sets of events and distribution functions are recalled in Section 1.1, while Section 1.2 summarizes known facts about lower probabilities and univariate p-boxes. Then, a bivariate p-box is defined in Section 2 as a pair of functions $\underline{F}$, $\overline{F}$ such that $\underline{F} \leq \overline{F}$ and satisfying some minimal properties (normalization, componentwise monotonicity). In Section 3 we show that there is a correspondence between a bivariate p-box and a lower probability (Definition 8), and because of this we investigate thereafter the consistency properties of bivariate p-boxes within the theory of lower probabilities. In Section 3.1 we prove a characterization of 2-coherence (Proposition 4), which is of some interest in itself, as it shows that this little investigated notion corresponds to a sort of strengthened capacity on certain sets of events. It allows to demonstrate that bivariate p-boxes can be associated, at best, with 2-coherent probabilities, but may even fail this correspondence (Proposition 5). In Section 3.2 bivariate p-boxes that avoid sure loss are defined, and characterized in terms of distribution functions (Proposition 6); some further properties of these p-boxes are proved. The notion of coherent bivariate p-box is defined in Section 3.3 in terms of coherence of its associated lower probability. Coherent bivariate p-boxes are characterized as envelopes of distribution functions in Proposition 9. Since coherent bivariate p-boxes are essentially equivalent representations of coherent lower probabilities defined on a suitable set, they appear to be the most prominent class of (bivariate) p-boxes. Their properties are studied in Section 3.4. Extending the classical *rectangle inequality* for distribution functions (equation (RI) in Proposition 2), it is shown that four *imprecise rectangle inequalities* are all necessary for coherence (Proposition 10). In some cases, they are also sufficient for a bivariate p-box to be coherent: the most general result of this kind is here Theorem 3. Further properties related with these inequalities are then discussed. Section 3.5 concerns the relationship between coherent bivariate p-boxes and 2-monotonicity, which is shown to be less tight and more complex than in the one-dimensional environment. Section 4 summarizes and concludes the paper.

1.1. *Distribution functions and related concepts*

The basic families of events we encounter when studying p-boxes are monotone families.

Definition 1. Given a family of events $(A_x)_{x\in I}$ and a strict total order $\prec$ in I, say that $(A_x)_{x\in I}$ is *monotone non-decreasing* (*monotone non-increasing*) if $\forall x, y \in I$, $x \prec y$ implies $A_x \subseteq A_y$ (implies $A_y \subseteq A_x$).

In the sequel, it is always assumed that the impossible event $\varnothing$ and the sure event Ω belong to the monotone families considered. This is not restrictive, as clearly, $\forall A_x$, $\varnothing \subseteq A_x \subseteq \Omega$. A convenient way to ensure this assumption is to include a minimum and a maximum 'value', $-\infty$, $+\infty$ respectively, into I, putting $A_{-\infty} = \varnothing$, $A_{+\infty} = \Omega$.

The family $(A_x^c)_{x\in I}$ *associated* with a monotone family $(A_x)_{x\in I}$ is also monotone, non-decreasing (non-increasing) if $(A_x)_{x\in I}$ is non-increasing (non-decreasing).

We shall have to deal with precise probabilities defined on monotone families or other related sets of events. We shall precisely refer to coherent probabilities in the sense of de Finetti[9], termed here *dF-coherent* probabilities to better distinguish them from coherent imprecise probabilities. Probabilities that are dF-coherent are naturally defined on *arbitrary* sets of events, but may be characterized on special sets, like algebras; a dF-coherent probability on an algebra is a finitely additive probability. On monotone families of events, the following characterization holds (Crisma[10], Thm. 11.1.2; Denneberg[11], Prop. 2.10):

Proposition 1. *Let $(A_x)_{x\in I}$ be a monotone family of events. A map $P : (A_x)_{x\in I} \to [0,1]$ is a dF-coherent probability on $(A_x)_{x\in I}$ if and only if a real function $F : I \to [0,1]$ can be chosen, such that*

(a) F is monotone non-decreasing, i.e., $x \prec y \Rightarrow F(x) \leq F(y)$;
(b) $F(x) = 0$ if $A_x = \varnothing$, $F(x) = 1$ if $A_x = \Omega$;
(c) $\forall x, y \in I, A_x = A_y$ implies $F(x) = F(y)$, or equivalently, $x \prec y, A_y \wedge A_x^c = \varnothing$, implies $\Delta F(x; y) := F(y) - F(x) = 0$,

and such that

$$P(A_x) = F(x), \forall x \in I. \tag{1}$$

Any real function F satisfying conditions (a)÷(c) in Proposition 1 is termed *cumulative distribution function*, or *cdf* in short.

Note that the characterization above, like the subsequent one in Proposition 2, regards dF-coherent, hence not necessarily σ-additive, probabilities. Therefore the concept of cdf in both propositions is larger than the classical one, which corresponds instead to σ-additive measures.

An important consequence of Proposition 1 is the following: if we define $F(x) = \mu(A_x)$, for all $x \in I$, where μ is any monotone and normalized function on a set including $(A_x)_{x\in I}$, then the restriction of μ on $(A_x)_{x\in I}$ is a dF-coherent probability. For this reason, the coherent lower probabilities or capacities we shall consider later on cannot be distinguished from dF-coherent probabilities when we focus on their restrictions to monotone families.

Remark 1. *(Cdf of a random number.)* A very common and important example is the probabilistic description of a real-valued random number X. The domain of X is $\mathcal{X} \subseteq \mathbb{R}$, so that X cannot take the 'values' $-\infty, +\infty$. Yet, here $I = \overline{\mathbb{R}}$, the compact real line, and A_x is the event $(X \leq x)$. The very reason for taking $I = \overline{\mathbb{R}}$ instead of $I = \mathbb{R}$ is the need for guaranteeing, whatever is X, that the family $(A_x)_{x\in\overline{\mathbb{R}}}$ includes $\varnothing = A_{-\infty} = (X \leq -\infty)$ and $\Omega = A_{+\infty} = (X \leq +\infty) = (X < +\infty)$. The cdf of X, by Eq. (1), is then given by $F(x) = P(X \leq x)$, $\forall x \in I$.

In the case that X takes up only finitely many different values, $x_1 < x_2 < \ldots < x_n$, it is customary to define again $I = \overline{\mathbb{R}}$. Yet, it is important to observe that by Proposition 1, (c) we only need to know the values $F(x_i)$, $i = 1, \ldots, n$ to fully describe F. In fact, for $x \in (x_i, x_{i+1})$, $A_x = A_{x_i}$, and with this idea

$$F(x) = \begin{cases} F(x_i) & \text{if } x \in [x_i, x_{i+1}) \\ 1 & \forall x \geq x_n \\ 0 & \forall x < x_1, \end{cases}$$

using (b) and (c) in Proposition 1. ♦

More generally, we have n-tuples of monotone families and related concepts. For what follows it will suffice to consider the case $n = 2$.

Definition 2. Let $({}_1A_x)_{x \in I_1}$, $({}_2A_y)_{y \in I_2}$ be two monotone non-decreasing families of events and define

$$A_{(x,y)} := {}_1A_x \wedge {}_2A_y \ \forall x \in I_1, y \in I_2, \quad (2)$$
$$\mathcal{D} := \{A_{(x,y)} : x \in I_1, y \in I_2\}.$$

Then, $\mathcal{D}$ is a component-wise monotone non-decreasing (bivariate) family of events.

A dF-coherent probability is characterized on $\mathcal{D}$ by the following result, a special case of Thm.11.2.2 in Crisma [10].

Proposition 2. *A map $P : \mathcal{D} \to \mathbb{R}$ is a dF-coherent probability on $\mathcal{D}$ if and only if a real function $F : I_1 \times I_2 \to \mathbb{R}$ can be chosen, such that*

(a) *F is component-wise monotone non-decreasing,*
(b) *$A_{(x,y)} = \varnothing$ implies $F(x, y) = 0$ and $A_{(x,y)} = \Omega$ implies $F(x, y) = 1$,*
(c) (Rectangle inequality) *For every $x_1, x_2 \in I_1, y_1, y_2 \in I_2$ such that $x_1 \prec x_2$, $y_1 \prec y_2$, defining*

$$\Delta F(x_1, x_2; y_1, y_2) := F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1),$$

it is

$$\Delta F(x_1, x_2; y_1, y_2) \geq 0, \quad \text{(RI)}$$

while

$$({}_1A_{x_2} \wedge {}_1A^c_{x_1}) \wedge ({}_2A_{y_2} \wedge {}_2A^c_{y_1}) = \varnothing \text{ implies } \Delta F(x_1, x_2; y_1, y_2) = 0, \quad (3)$$

and P is such that

$$P(A_{(x,y)}) = F(x, y), \ \forall x \in I_1, \forall y \in I_2.$$

Any function F satisfying conditions (a)÷(c) above characterizes a dF-coherent probability on $\mathcal{D}$ by means of the previous proposition, and is thus termed *bivariate (cumulative) distribution function.*

Remark 2. Condition (3) implies

$$A_{(x_1,y_1)} = A_{(x_2,y_2)} \Rightarrow F(x_1, y_1) = F(x_2, y_2). \quad (4)$$

Let us give a sketch of the proof. From $A_{(x_1,y_1)} = A_{(x_2,y_2)}$, we get

$$({}_1A_{x_2} \wedge {}_1A^c_{x_1}) \wedge ({}_2A_{y_2} \wedge {}_2A^c_{y_1}) = {}_1A^c_{x_1} \wedge {}_2A^c_{y_1} \wedge A_{(x_1,y_1)} = \varnothing.$$

Since necessarily $A_{(x_1,y_1)} = A_{(x_1,y_2)} = A_{(x_2,y_1)}$, we get analogously

$$({}_1A_{x_1} \wedge {}_1A^c_{-\infty}) \wedge ({}_2A_{y_2} \wedge {}_2A^c_{y_1}) = A_{(x_1,y_1)} \wedge {}_1A^c_{-\infty} \wedge {}_2A^c_{y_1} = \varnothing,$$

and

$$({}_2A_{y_1} \wedge {}_2A^c_{-\infty}) \wedge ({}_1A_{x_2} \wedge {}_1A^c_{x_1}) = A_{(x_1,y_1)} \wedge {}_2A^c_{-\infty} \wedge {}_1A^c_{x_1} = \varnothing.$$

Hence $\Delta F(x_1, x_2; y_1, y_2) = \Delta F(-\infty, x_1; y_1, y_2) = \Delta F(x_1, x_2; -\infty, y_1) = 0$ by (3). The system of these equalities implies $F(x_1, y_1) = F(x_2, y_2)$.

However, conditions (3) and (4) are not equivalent: it may be that $A_{(x_1,y_1)} \neq A_{(x_2,y_2)}$ but $({}_1A_{x_2} \wedge {}_1A^c_{x_1}) \wedge ({}_2A_{y_2} \wedge {}_2A^c_{y_1}) = \varnothing$. See Example 3 later on and its footnote for an instance.[a] The two conditions are instead equivalent in the univariate case, as stated in Proposition 1, c). There are instances when (3) trivially holds and therefore has not to be checked. An important case is that of X and Y being logically independent, as we discuss next. ♦

Definition 3. Two random numbers X, Y, taking values in $\mathcal{X}$, $\mathcal{Y}$ respectively, are *logically independent* iff $(X = x) \wedge (Y = y) \neq \varnothing\ \forall x \in \mathcal{X}, \forall y \in \mathcal{Y}$.

Remark 3. *(Cdf of a couple of random numbers.)* As a notable example, consider two random numbers X, Y. Then $I_1 \times I_2 = \overline{\mathbb{R}} \times \overline{\mathbb{R}}$ and $A_{(x,y)}$ is the event $(X \leq x \wedge Y \leq y)$. Hence $F(x, y) = P(X \leq x \wedge Y \leq y)$ and condition (3) reads as

$$x_1 < X \leq x_2 \wedge y_1 < Y \leq y_2 = \varnothing \implies P(x_1 < X \leq x_2 \wedge y_1 < Y \leq y_2) = 0.$$

Often in the sequel X, Y will be discrete random numbers. Similarly to Remark 1, in that case we shall assume that the domain of F is $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$. F is determined by (4) and its values $F(x_i, y_j)$, $\forall (x_i, y_j) \in \mathcal{X} \times \mathcal{Y}$, where $\mathcal{X} = \{x_1, \ldots, x_n\}$, $\mathcal{Y} = \{y_1, \ldots, y_m\}$ are the sets of possible values for X and Y, respectively. In fact, $\forall (x, y) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$, we have

$$A_{(x,y)} = A_{(x',y')},$$

where

$$x' = \max\{x_s \in \mathcal{X} \cup \{-\infty\} : x_s \leq x\}, y' = \max\{y_t \in \mathcal{Y} \cup \{-\infty\} : y_t \leq y\}, \qquad (5)$$

and, recalling (4),

$$F(x, y) = F(x', y').$$

A similar procedure applies when at least one of $\mathcal{X}$, $\mathcal{Y}$ does not coincide with $\mathbb{R}$. ♦

Remark 4. Recalling Remark 3, it is easy to realize that we need not check (3) when X, Y are logically independent. In fact, if $x_1 < x_2 \in \mathcal{X}$, $y_1 < y_2 \in \mathcal{Y}$, the event $(x_1 < X \leq x_2) \wedge (y_1 < Y \leq y_2)$ includes the event $(X = x_2) \wedge (Y = y_2)$, which is non-impossible by logical independence; otherwise, $(x_1 < X \leq x_2) \wedge (y_1 < Y \leq y_2) = (x'_1 < X \leq x'_2) \wedge (y'_1 < Y \leq y'_2) = \varnothing$ iff $x'_1 = x'_2$ or $y'_1 = y'_2$. If this condition occurs, $\Delta F(x_1, x_2; y_1, y_2) = \Delta F(x'_1, x'_2; y'_1, y'_2) = 0$, as can be easily checked. ♦

[a] In the original version of Theorem 2 in Crisma [10], (4) instead of (3) is stated among the hypotheses, although its proof implicitly assumes (3) (L. Crisma, personal communication).

1.2. *Lower probabilities and univariate p-boxes*

The theory of dF-coherent probabilities has been extended to the imprecise case by Williams [12] and Walley [3]. Although in Walley [3] the theory is established in terms of bounded real-valued functions, or gambles, for the purposes of this paper we shall consider only maps defined on events. Several other consistency concepts for non-additive measures have been considered in the literature. We recall now those used, at different levels, in this paper. In the relevant formulas I_E is used to denote the *indicator* function of the event E: the function that takes the value 1 on the elements of E, and 0 otherwise.

Definition 4. (Walley [3]) Let $\mathcal{A}$ be an *arbitrary* set of events and $\underline{P}$ a map, $\underline{P} : \mathcal{A} \to \mathbb{R}$.

(a) $\underline{P}$ is a lower probability that *avoids sure loss* on $\mathcal{A}$ iff, $\forall n > 0, \forall s_1, \ldots, s_n \geq 0, \forall E_1, \ldots, E_n \in \mathcal{A}$, it holds that $\max \sum_{i=1}^{n} s_i(I_{E_i} - \underline{P}(E_i)) \geq 0$.
(b) $\underline{P}$ is a *coherent lower probability* on $\mathcal{A}$ iff $\forall n > 0, \forall s_0, \ldots, s_n \geq 0$, and $\forall E_0, \ldots, E_n \in \mathcal{A}$, it holds that $\max\{\sum_{i=1}^{n} s_i(I_{E_i} - \underline{P}(E_i)) - s_0(I_{E_0} - \underline{P}(E_0))\} \geq 0$.
(c) $\underline{P}$ is a 2-*coherent lower probability* on $\mathcal{A}$ iff, $\forall E_0, E_1 \in \mathcal{A}, \forall s_0 \in \mathbb{R}, \forall s_1 \geq 0$, it holds that $\max\{s_1(I_{E_1} - \underline{P}(E_1)) + s_0(I_{E_0} - \underline{P}(E_0))\} \geq 0$.

The most important of these concepts is that of coherence, which implies the other two.

It is customary to relate lower ($\underline{P}$) and upper ($\overline{P}$) probabilities by the *conjugacy* equality

$$\underline{P}(A) = 1 - \overline{P}(A^c). \tag{6}$$

Because of (6), one may focus on lower probabilities only, as we shall mainly do.

There is an important characterization of Definition 4, (a) and (b), in terms of the *credal set* $\mathcal{M}(\underline{P})$ of a lower probability $\underline{P}$,

$$\mathcal{M}(\underline{P}) := \{P : \mathcal{A} \to \mathbb{R}, \text{dF-coherent} : P(A) \geq \underline{P}(A)\ \forall A \in \mathcal{A}\}. \tag{7}$$

Theorem 1. *(Walley [3], Section 3.3.4) Let* $\underline{P} : \mathcal{A} \to \mathbb{R}$.

- $\underline{P}$ avoids sure loss *if and only if* $\mathcal{M}(\underline{P}) \neq \emptyset$.
- $\underline{P}$ *is* coherent *if and only if* $\underline{P}(A) = \min\{P(A) : P \in \mathcal{M}(\underline{P})\}\ \forall A \in \mathcal{A}$.

Consider now $(A_x)_{x \in \overline{\mathbb{R}}}$. A dF-coherent probability $P : \mathcal{A} \to [0, 1]$, where $\mathcal{A} \supseteq (A_x)_{x \in \overline{\mathbb{R}}}$, induces a cdf $F_P : \overline{\mathbb{R}} \to [0, 1]$ by means of Eq. (1),

$$F_P(x) = P(A_x)\ \forall x \in \overline{\mathbb{R}}. \tag{8}$$

As a consequence, a coherent lower probability $\underline{P}$ on the domain $\mathcal{A}$ induces a set of distribution functions

$$\mathcal{F} := \{F_P : P \in \mathcal{M}(\underline{P})\},$$

with F_P given by Eq. (8). From $\mathcal{F}$ we can derive the functions $\underline{F}, \overline{F} : \overline{\mathbb{R}} \to [0,1]$ by

$$\underline{F}(x) = \inf\{F(x) : F \in \mathcal{F}\}, \ \overline{F}(x) = \sup\{F(x) : F \in \mathcal{F}\}.$$

It is easy to check that both $\underline{F}, \overline{F}$ satisfy conditions (a)÷(c) from Proposition 1, and are therefore distribution functions. We shall refer to $(\underline{F}, \overline{F})$ as the p-box associated with the coherent lower probability $\underline{P}$. More generally, we have the following definition:

Definition 5. (Ferson et al.[2], Ferson and Tucker[4]) A (univariate) *p-box* is a pair $(\underline{F}, \overline{F})$ where $\underline{F}, \overline{F} : \overline{\mathbb{R}} \to [0,1]$ are cumulative distribution functions satisfying $\underline{F}(x) \leq \overline{F}(x)$ for every $x \in \overline{\mathbb{R}}$.

However, two different coherent lower probabilities can be associated with the same p-box. Given the p-box $(\underline{F}, \overline{F})$, $(A_x)_{x\in\overline{\mathbb{R}}}$, define

$$\mathcal{E}_0 := \{A_x, A_x^c : x \in \overline{\mathbb{R}}\}.$$

Then a coherent lower probability $\underline{P}$ with domain $\mathcal{A} \supseteq \mathcal{E}_0$ induces the p-box $(\underline{F}, \overline{F})$ if and only if $\underline{P}(A_x) = \underline{F}(x)$ and $\underline{P}(A_x^c) = 1 - \overline{F}(x)$. There may be more than one coherent lower probability on $\mathcal{A}$ with the same restriction to $\mathcal{E}_0$, and all of them will be associated with the same p-box. This was established in the precise case by Miranda et al.[13]

The following result clarifies the correspondence between coherent lower probabilities and p-boxes in the univariate case. It was stated by Walley[3] and proved by Troffaes and Destercke[5], Troffaes and de Cooman[14].

Theorem 2. *(Troffaes and de Cooman[14], Thm. 7.16; Troffaes and Destercke[5], Sect. 3) Consider two maps* $\underline{F}, \overline{F} : \overline{\mathbb{R}} \to [0,1]$ *and let* $\underline{P}_{(\underline{F},\overline{F})} : \mathcal{E}_0 \to [0,1]$ *be the lower probability they induce by means of*

$$\underline{P}_{(\underline{F},\overline{F})}(A_x) = \underline{F}(x) \text{ and } \underline{P}_{(\underline{F},\overline{F})}(A_x^c) = 1 - \overline{F}(x) \ \forall x \in \overline{\mathbb{R}}.$$

Consider also the restrictions $\underline{P}_{\underline{F}} : (A_x)_{x\in\overline{\mathbb{R}}} \to [0,1]$ *and* $\underline{P}_{\overline{F}} : (A_x^c)_{x\in\overline{\mathbb{R}}} \to [0,1]$ *of* $\underline{P}_{(\underline{F},\overline{F})}$ *given by*

$$\underline{P}_{\underline{F}}(A_x) = \underline{F}(x) \text{ and } \underline{P}_{\overline{F}}(A_x^c) = 1 - \overline{F}(x) \ \forall x \in \overline{\mathbb{R}}.$$

The following are equivalent:

(a) $\underline{P}_{(\underline{F},\overline{F})}$ *is a coherent lower probability on* $\mathcal{E}_0$.
(b) $\underline{F}, \overline{F}$ *are distribution functions and* $\underline{F} \leq \overline{F}$ *(i.e.,* $(\underline{F}, \overline{F})$ *is a p-box).*
(c) $\underline{P}_{\underline{F}}$ *and* $\underline{P}_{\overline{F}}$ *are dF-coherent and* $\underline{F} \leq \overline{F}$.

2. Bivariate *P*-boxes

Univariate p-boxes can be used as a model of uncertainty for a real-valued random number, when there is some imprecision in its associated cumulative distribution function. In this section, we shall investigate how to generalize this model to the case where we consider the joint behaviour of two real random numbers X, Y.

Definition 6. A map $F : \overline{\mathbb{R}}\times\overline{\mathbb{R}} \to [0,1]$ is called *standardized* when it is component-wise non-decreasing and $F(+\infty,+\infty) = 1$, $F(-\infty,\cdot) = F(\cdot,-\infty) = 0$.

In the univariate case, one possible interpretation of a p-box is a model for the imprecise knowledge of a (precise) distribution function F: if we consider a set $\mathcal{F}$ of possible candidates, this set can be summarized by its lower and upper envelopes $\underline{F}, \overline{F}$, which are distribution functions themselves. As a consequence, a univariate p-box can be seen as the set of cdfs bounded between two particular distribution functions that determine the lower and upper bounds for the cumulative probabilities.

Unfortunately, the situation is not so clear-cut in the bivariate case: the envelopes of a set of distribution functions are standardized maps, but not necessarily distribution functions, since they do not necessarily satisfy condition (c) in Proposition 2.

Proposition 3. *Let $\mathcal{F}$ be a family of distribution functions $F : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0,1]$. Their lower and upper envelopes $\underline{F}, \overline{F} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0,1]$, given by*

$$\underline{F}(x,y) = \inf_{F\in\mathcal{F}} F(x,y) \text{ and } \overline{F}(x,y) = \sup_{F\in\mathcal{F}} F(x,y)$$

for every $x, y \in \overline{\mathbb{R}}$, are standardized maps and satisfy the condition

$$A_{(x_1,y_1)} = A_{(x_2,y_2)} \Rightarrow \begin{cases} \underline{F}(x_1,y_1) = \underline{F}(x_2,y_2) \\ \overline{F}(x_1,y_1) = \overline{F}(x_2,y_2) \end{cases}, \forall A_{(x_1,y_1)}, A_{(x_2,y_2)} \in \mathcal{D}. \quad (9)$$

Proof. It suffices to take into account that conditions (a) and (b) in Proposition 2, as well as (4), are preserved by lower and upper envelopes. □

To see that these envelopes are not necessarily distribution functions, consider the following example, where $\underline{F}$, $\overline{F}$ do not satisfy (RI):

Example 1. Let P_1 and P_2 be the probability measures associated with the following mass functions[b]:

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
P_1	0.1	0.1	0	0.4	0.1	0	0	0	0.3
P_2	0.4	0	0.2	0.1	0	0	0.1	0	0.2

[b]It is understood in this and most of the following examples that we consider two random numbers X, Y, taking values, respectively, in $\mathcal{X}$, $\mathcal{Y}$ (here $\mathcal{X} = \mathcal{Y} = \{1,2,3\}$). The values (i,j) in the first row of the tables are those of the product $\mathcal{X} \times \mathcal{Y}$ (here $(1,1),\ldots,(3,3)$).

Their associated distribution functions F_1, F_2 are determined by the following values:

	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$	$(3,1)$	$(3,2)$	$(3,3)$
F_1	0.1	0.2	0.2	0.5	0.7	0.7	0.5	0.7	1
F_2	0.4	0.4	0.6	0.5	0.5	0.7	0.6	0.6	1

and extended to $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ in the manner discussed in Remark 3. Their lower and upper envelopes are

	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$	$(3,1)$	$(3,2)$	$(3,3)$
$\underline{F}$	0.1	0.2	0.2	0.5	0.5	0.7	0.5	0.6	1
$\overline{F}$	0.4	0.4	0.6	0.5	0.7	0.7	0.6	0.7	1

Then

$$\underline{F}(2,2) + \underline{F}(1,1) - \underline{F}(1,2) - \underline{F}(2,1) = 0.5 + 0.1 - 0.2 - 0.5 = -0.1 < 0$$

and

$$\overline{F}(2,3) + \overline{F}(1,2) - \overline{F}(1,3) - \overline{F}(2,2) = 0.7 + 0.4 - 0.6 - 0.7 = -0.2 < 0.$$

As a consequence, neither $\underline{F}$ nor $\overline{F}$ are distribution functions. ♦

Taking this result into account, we give the following definition:

Definition 7. Let $\underline{F}, \overline{F} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0,1]$ be two standardized functions satisfying $\underline{F}(x,y) \leq \overline{F}(x,y)$ for every $x, y \in \overline{\mathbb{R}}$ and (9). Then the pair $(\underline{F}, \overline{F})$ is called a *bivariate p-box.*

Remark 5. Definition 7 generalises Definition 5 of (univariate) p-box. In fact, let us apply Definition 7 to univariate $\underline{F}$, $\overline{F}$. Then (a) and (b) of Proposition 1 hold for them because of standardisation, while (9) ensures (c). Therefore, by Proposition 1, $\underline{F}$ and $\overline{F}$ are cdfs. Since $\underline{F} \leq \overline{F}$, $(\underline{F}, \overline{F})$ is a (univariate) p-box. ♦

Proposition 3 shows that bivariate p-boxes can be obtained in particular by means of a set of distribution functions, taking their lower and upper envelopes. However, not all bivariate p-boxes are of this type: if we consider for instance a map $\underline{F} = \overline{F}$ that is standardized but not a distribution function, then there is no bivariate distribution function between $\underline{F}$ and $\overline{F}$, and as a consequence these cannot be obtained as envelopes of a set of distribution functions.

3. Lower Probabilities and Bivariate P-boxes

Using Eq. (2), define the sets

$$\mathcal{D} := \{A_{(x,y)} : x, y \in \overline{\mathbb{R}}\},\ \mathcal{D}_c := \{A^c_{(x,y)} : x, y \in \overline{\mathbb{R}}\} \text{ and } \mathcal{E} := \mathcal{D} \cup \mathcal{D}_c, \tag{10}$$

and consider a bivariate p-box $(\underline{F}, \overline{F})$. When an analogue of (4) holds for $\underline{F}$, $\overline{F}$, it determines a lower probability on $\mathcal{E}$ as follows.

Definition 8. Let $(\underline{F}, \overline{F})$ be a bivariate p-box.

The *lower probability induced by* $(\underline{F}, \overline{F})$ is the map $\underline{P}_{(\underline{F},\overline{F})} : \mathcal{E} \to [0,1]$ given by:

$$\underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) = \underline{F}(x,y), \quad \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) = 1 - \overline{F}(x,y) \tag{11}$$

for every $x, y \in \overline{\mathbb{R}}$.

Conversely, a lower probability $\underline{P} : \mathcal{E} \to [0,1]$ determines a couple of functions $\underline{F}_{\underline{P}}$, $\overline{F}_{\underline{P}}$ by

$$\underline{F}_{\underline{P}}(x,y) = \underline{P}(A_{(x,y)}) \text{ and } \overline{F}_{\underline{P}}(x,y) = 1 - \underline{P}(A^c_{(x,y)}) \ \forall x, y \in \overline{\mathbb{R}}. \tag{12}$$

Remark 6. It is interesting to consider the correspondences (11), (12) in the precise case:

- On the one hand, if $\underline{F} = \overline{F} = F$ then the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ determined by means of Eq. (11), that we shall denote P_F, is uniquely determined by additivity from its restriction on $\mathcal{D}$, because $P_F(A^c_{(x,y)}) = 1 - F(x,y) = 1 - P_F(A_{(x,y)})$ for every $x, y \in \overline{\mathbb{R}}$. Moreover, in the particular case where F is a distribution function, we deduce from Proposition 2 that P_F is a dF-coherent probability.
- Conversely, if $P : \mathcal{D} \to [0,1]$ is a dF-coherent probability, it has a unique dF-coherent extension on $\mathcal{E}$, obtained putting
$$P(A^c_{(x,y)}) = 1 - P(A_{(x,y)}).$$
Then P is dF-coherent on $\mathcal{E}$, and its associated lower and upper distribution functions coincide: they are both equal to
$$F_P(x,y) = P(A_{(x,y)}), \forall x, y \in \overline{\mathbb{R}}, \tag{13}$$
which is a bivariate distribution function. ♦

3.1. *Bivariate p-boxes and 2-coherent probabilities*

Next, we investigate the consistency properties of the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ defined by (11), i.e. by means of a bivariate p-box. We shall prove that, unlike the univariate case, where analogous conditions on $(\underline{F}, \overline{F})$ guarantee dF-coherence, here we only come to 2-coherence under an assumption of logical independence.

To see how this comes about, we first establish a characterization of 2-coherence in the next proposition. It extends an analogous result for lower probabilities defined on algebras of events in Walley [3], Appendix B, Theorem B3(b) to sets of events closed under complementation.

Proposition 4. *Let $\mathcal{S}$ be a set of events that is closed under complementation, and let $\underline{P} : \mathcal{S} \to \mathbb{R}^+ \cup \{0\}$. Then, $\underline{P}$ is 2-coherent iff it satisfies the following conditions*

i) $\forall E, F \in \mathcal{S}, E \subseteq F$ *implies* $\underline{P}(E) \le \underline{P}(F)$;
ii) $\underline{P}(E) + \underline{P}(E^c) \le 1, \forall E \in \mathcal{S}$;

iii) if $\emptyset, \Omega \in \mathcal{S}$, $\underline{P}(\emptyset) = 0$, $\underline{P}(\Omega) = 1$.

Proof. Assume $\underline{P}$ satisfies i), ii), iii) above. Note that ii) and non-negativity of $\underline{P}$ imply also $\underline{P}(E) \leq 1, \forall E \in \mathcal{S}$. For any $E_0, E_1 \in \mathcal{A}$, $\forall s_0 \in \mathbb{R}$, $\forall s_1 \geq 0$, let us define $\underline{G}_2 = s_1(I_{E_1} - \underline{P}(E_1)) + s_0(I_{E_0} - \underline{P}(E_0))$. The random number $\underline{G}_2$ is defined on the partition $\{E_0 \wedge E_1, E_0^c \wedge E_1, E_0 \wedge E_1^c, E_0^c \wedge E_1^c\}$. In order to prove, by applying Definition 4, (c), that $\max \underline{G}_2 \geq 0$, we consider several cases:

a) $s_0 \geq 0$.

a1) If $E_0 \wedge E_1 \neq \emptyset$, then $\underline{G}_2(E_0 \wedge E_1) = s_1(1 - \underline{P}(E_1)) + s_0(1 - \underline{P}(E_0)) \geq 0$.

a2) If $E_0 \wedge E_1 = \emptyset$, we have that $E_0 \subseteq E_1^c$ and $E_1 \subseteq E_0^c$. Assume for instance that $s_1 \geq s_0$ (if $s_1 < s_0$, it suffices to exchange the role of E_0 and E_1). By applying i), ii), we get $\underline{P}(E_0) \leq \underline{P}(E_1^c) \leq 1 - \underline{P}(E_1)$. If $E_0^c \wedge E_1 = E_1 \neq \emptyset$, then

$$\begin{aligned}\underline{G}_2(E_1) &= s_1(1 - \underline{P}(E_1)) - s_0\underline{P}(E_0) \\ &\geq s_1\underline{P}(E_0) - s_0\underline{P}(E_0) = (s_1 - s_0)\underline{P}(E_0) \geq 0.\end{aligned}$$

Otherwise, if $E_0^c \wedge E_1 = E_1 = \emptyset$, then $\underline{G}_2 = 0$ when $E_0 = \emptyset$ (by iii)), whilst if $E_0 \neq \emptyset$ then $\underline{G}_2(E_0) = s_0(1 - \underline{P}(E_0)) \geq 0$.

b) $s_0 < 0$.

b1) If $E_0^c \wedge E_1 \neq \emptyset$, then $\underline{G}_2(E_0^c \wedge E_1) = s_1(1 - \underline{P}(E_1)) - s_0\underline{P}(E_0) \geq 0$.

b2) If $E_0^c \wedge E_1 = \emptyset$, we have $E_0^c \subseteq E_1^c$ and $E_1 \subseteq E_0$. By i), we deduce that $\underline{P}(E_1) \leq \underline{P}(E_0)$.
Assume next that $s_1 \geq -s_0$. If $E_1 = E_1 \wedge E_0 \neq \emptyset$, then

$$\begin{aligned}\underline{G}_2(E_1) &= s_1(1 - \underline{P}(E_1)) + s_0(1 - \underline{P}(E_0)) \\ &\geq s_1(1 - \underline{P}(E_0)) + s_0(1 - \underline{P}(E_0)) \\ &= (s_1 + s_0)(1 - \underline{P}(E_0)) \geq 0.\end{aligned}$$

Otherwise, if $E_1 = E_1 \wedge E_0 = \emptyset$, $\underline{G}_2 = 0$ when $E_0^c = \emptyset$, while $\underline{G}_2(E_0^c) = -s_0\underline{P}(E_0) \geq 0$ when $E_0^c = E_0^c \wedge E_1^c \neq \emptyset$.
Finally, assume $s_1 < -s_0$.
If $E_0^c = E_0^c \wedge E_1^c \neq \emptyset$, then

$$\begin{aligned}\underline{G}_2(E_0^c) &= -s_1\underline{P}(E_1) - s_0\underline{P}(E_0) \\ &\geq -s_1\underline{P}(E_0) - s_0\underline{P}(E_0) = -(s_1 + s_0)\underline{P}(E_0) \geq 0.\end{aligned}$$

Otherwise, if $E_0^c = E_0^c \wedge E_1^c = \emptyset$, $\underline{G}_2 = 0$ when $E_1 = \emptyset$, whilst $\underline{G}_2(E_1) = s_1(1 - \underline{P}(E_1)) \geq 0$ when $E_1 \neq \emptyset$.

Conversely, let $\underline{P}$ be 2-coherent. To prove that i), ii), iii) hold, we exploit the condition $\max \underline{G}_2 \geq 0$ in Definition 4, (c) for some suitable $\underline{G}_2$. In detail, we have:

i) Take $s_1 = -s_0 = 1$, $E_1 = E$, $E_0 = F$. Then, since $E \subseteq F$ implies $I_E \leq I_F$, $0 \leq \max \underline{G}_2 = \max[I_E - \underline{P}(E) - I_F + \underline{P}(F))] \leq \underline{P}(F) - \underline{P}(E)$. This implies $\underline{P}(E) \leq \underline{P}(F)$.

ii) Take $s_1 = s_0 = 1$, $E_1 = E$, $E_0 = E^c$ in Definition 4, (c). Then

$$\max \underline{G}_2 = \max[I_E - \underline{P}(E) + I_{E^c} - \underline{P}(E^c)] = 1 - \underline{P}(E) - \underline{P}(E^c) \geq 0$$

iff $\underline{P}(E) + \underline{P}(E^c) \leq 1$.

iii) To prove $\underline{P}(\varnothing) = 0$, take first $s_1 = 1$, $s_0 = 0$, $E_1 = \varnothing$ in Definition 4, (c), to get $\max \underline{G}_2 = -\underline{P}(\varnothing) \geq 0$, i.e. $\underline{P}(\varnothing) \leq 0$. This and non-negativity of $\underline{P}$ imply $\underline{P}(\varnothing) = 0$. Analogously, by taking $s_1 = 0$, $s_0 = -1$, $E_0 = \Omega$, we get $\underline{P}(\Omega) \geq 1$, ensuring with ii) that $\underline{P}(\Omega) = 1$. □ □

Taking account of Proposition 4, we have:

Proposition 5. *Consider two random numbers X, Y, taking values in $\mathcal{X}, \mathcal{Y}$ respectively, and let $\mathcal{D}$, $\mathcal{D}_c$, $\mathcal{E}$ be defined by* (10)*, where $A_{(x,y)} = (X \leq x \wedge Y \leq y)$, $x, y \in \overline{\mathbb{R}}$.*

a) Given a 2*-coherent lower probability $\underline{P} : \mathcal{E} \to \mathbb{R}$, the pair $(\underline{F}_{\underline{P}}, \overline{F}_{\underline{P}})$ associated with $\underline{P}$ by means of Eq.* (12) *is a bivariate p-box.*

b) If X, Y are logically independent, the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ associated with a bivariate p-box $(\underline{F}, \overline{F})$ by means of Eq. (11) *is* 2*-coherent on $\mathcal{E}$.*

Proof.

a) We start proving that $\underline{F}_{\underline{P}}$, $\overline{F}_{\underline{P}}$ are standardized.
Let $x_1 \leq x_2, y_1 \leq y_2 \in \overline{\mathbb{R}}$. Then, $A_{(x_1,y_1)}$ is included in $A_{(x_2,y_2)}$. Hence

$$\underline{F}_{\underline{P}}(x_1, y_1) = \underline{P}(A_{(x_1,y_1)}) \leq \underline{P}(A_{(x_2,y_2)}) = \underline{F}_{\underline{P}}(x_2, y_2) \text{ and}$$
$$\overline{F}_{\underline{P}}(x_1, y_1) = 1 - \underline{P}(A^c_{(x_1,y_1)}) \leq 1 - \underline{P}(A^c_{(x_2,y_2)}) = \overline{F}_{\underline{P}}(x_2, y_2)$$

by Proposition 4, i).
By Proposition 4, iii), we get also $\underline{F}_{\underline{P}}(+\infty, +\infty) = \underline{P}(A_{(+\infty,+\infty)}) = \underline{P}(\Omega) = 1$ and $\underline{F}_{\underline{P}}(-\infty, \cdot) = \underline{P}(A_{(-\infty,\cdot)}) = \underline{F}_{\underline{P}}(\cdot, -\infty) = \underline{P}(A_{(\cdot,-\infty)}) = \underline{P}(\varnothing) = 0$.
To prove that $\underline{F}_{\underline{P}}(x, y) \leq \overline{F}_{\underline{P}}(x, y)$, $\forall x, y \in \overline{\mathbb{R}}$, apply Proposition 4, ii): $\underline{F}_{\underline{P}}(x, y) = \underline{P}(A_{(x,y)}) \leq 1 - \underline{P}(A^c_{(x,y)}) = \overline{F}_{\underline{P}}(x, y)$.
Since (9) holds trivially, the pair $(\underline{F}_{\underline{P}}, \overline{F}_{\underline{P}})$ is a bivariate p-box.

b) We prove that $\underline{P}_{(\underline{F},\overline{F})}$ satisfies i), ii), iii) in Proposition 4.
Proof of i) Consider $E, F \in \mathcal{E}$, $E \subseteq F$. We distinguish four cases.

- If $E = A_{(x_1,y_1)}$, $F = A_{(x_2,y_2)}$, let us define $x' = \min\{x_1, x_2\}$, $y' = \min\{y_1, y_2\}$. Since $E = E \wedge F = A_{(x',y')}$, we get

$$\underline{P}_{(\underline{F},\overline{F})}(E) = \underline{P}_{(\underline{F},\overline{F})}(A_{(x',y')}) = \underline{F}(x', y') \leq \underline{F}(x_2, y_2) = \underline{P}_{(\underline{F},\overline{F})}(F).$$

- Let $E = A^c_{(x_1,y_1)} \subseteq F = A^c_{(x_2,y_2)}$ or, equivalently, $F^c = A_{(x_2,y_2)} \subseteq E^c = A_{(x_1,y_1)}$. Again, as above, we get $F^c = E^c \wedge F^c = A_{(x',y')}$, i.e. $F = A^c_{(x',y')}$. Hence,

$$\underline{P}_{(\underline{F},\overline{F})}(E) = 1 - \overline{F}(x_1, y_1) \leq 1 - \overline{F}(x', y') = \underline{P}_{(\underline{F},\overline{F})}(F).$$

- If $E = A_{(x_1,y_1)} \subseteq F = A^c_{(x_2,y_2)}$ then, taking again $x' = \min\{x_1, x_2\}$, $y' = \min\{y_1, y_2\}$, it follows that $E \wedge F^c = A_{(x',y')} = \varnothing$. Logical independence of X, Y implies $x' < x \ \forall x \in \mathcal{X}$ or $y' < y \ \forall y \in \mathcal{Y}$, since, otherwise, taking $x \in \mathcal{X}, x \leq x'$, $y \in \mathcal{Y}, y \leq y'$, it would be $\varnothing \neq (X = x \wedge Y = y) \subseteq A_{(x',y')}$, a contradiction. It follows easily $E = \varnothing$ or $F = \Omega$. Hence $\underline{P}_{(\underline{F},\overline{F})}(E) \leq \underline{P}_{(\underline{F},\overline{F})}(F)$.
- Let $E = A^c_{(x_1,y_1)} \subseteq F = A_{(x_2,y_2)}$, i.e. $A_{(x_1,y_1)} \vee A_{(x_2,y_2)} = \Omega$. Logical independence of X, Y implies that $x_1 \geq x$ and $y_1 \geq y \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}$, or $x_2 \geq x$ and $y_2 \geq y \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}$. To prove this fact, assume that there exist $(x_1^*, y_1^*) \in \mathcal{X} \times \mathcal{Y}$ such that $x_1^* > x_1$ or $y_1^* > y_1$ and $(x_2^*, y_2^*) \in \mathcal{X} \times \mathcal{Y}$ such that $x_2^* > x_2$ or $y_2^* > y_2$, and define $x'' = \max\{x_1^*, x_2^*\}$, $y'' = \max\{y_1^*, y_2^*\}$. Then $\varnothing \neq (X = x'' \wedge Y = y'') \nsubseteq A_{(x_1,y_1)} \vee A_{(x_2,y_2)} = \Omega$, a contradiction. Again, it follows easily $E = \varnothing$ or $F = \Omega$, hence $\underline{P}_{(\underline{F},\overline{F})}(E) \leq \underline{P}_{(\underline{F},\overline{F})}(F)$.

Proof of ii) Let $E \in \{A_{(x,y)}, A^c_{(x,y)}\} \subseteq \mathcal{E}$. It holds that $\underline{P}_{(\underline{F},\overline{F})}(E) + \underline{P}_{(\underline{F},\overline{F})}(E^c) = \underline{F}(x,y) + 1 - \overline{F}(x,y) \leq 1$, taking into account that $\underline{F}(x,y) \leq \overline{F}(x,y)$.
Proof of iii) Note that $\Omega = A_{(+\infty,+\infty)}, \varnothing = A_{(-\infty,\cdot)} = A_{(\cdot,-\infty)} \in \mathcal{E}$. Therefore, we get $\underline{P}_{(\underline{F},\overline{F})}(\Omega) = \underline{F}(+\infty,+\infty) = 1$, $\underline{P}_{(\underline{F},\overline{F})}(\varnothing) = \underline{F}(-\infty,\cdot) = 0$. □

Remark 7. Logical independence is required in Proposition 5 to prove that the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ associated with a given bivariate p-box $(\underline{F}, \overline{F})$ through (11) is 2-coherent. To show that logical independence cannot simply be dropped in Proposition 5, take two binary random numbers X, Y, both assuming their values in $\mathcal{X} = \mathcal{Y} = \{1, 2\}$. Assume $(X = 1 \wedge Y = 1) = \varnothing$, so that X, Y are logically dependent, and consider the bivariate p-box $(\underline{F}, \overline{F})$ determined by the following table:

	$(1,2)$	$(2,1)$	$(2,2)$
$\underline{F}$	α_1	α_2	1
$\overline{F}$	β_1	β_2	1

where $0 \leq \alpha_i \leq \beta_i \leq 1$ $(i = 1, 2)$. On non-trivial events the lower probability associated with this bivariate p-box is given by (11):

	$A_{(1,2)}$	$A_{(2,1)}$	$A^c_{(1,2)}$	$A^c_{(2,1)}$
$\underline{P}_{(\underline{F},\overline{F})}$	α_1	α_2	$1-\beta_1$	$1-\beta_2$

Then $A_{(1,2)} = (X = 1 \wedge Y = 2) \subseteq A^c_{(2,1)} = (X = 1 \wedge Y = 2) \vee (X = 2 \wedge Y = 2)$ and similarly $A_{(2,1)} \subseteq A^c_{(1,2)}$. These are the only non-trivial implications among the events in $\mathcal{E}$. Hence, Proposition 4, i) holds if and only if both $\alpha_1 \leq 1 - \beta_2$ and $\alpha_2 \leq 1 - \beta_1$. ♦

Proposition 5 and Remark 7 show that the consistency properties of bivariate p-boxes are weaker than or at best corresponding to those of 2-coherent lower probabilities. Logical independence is sufficient to ensure the correspondence, but not at all necessary. This is patent from Remark 7, where the lower probability associated with $(\underline{F}, \overline{F})$ may or may not be 2-coherent, depending on the values of α_i, β_i $(i = 1, 2)$. But even ensuring 2-coherence seems unsatisfactory, since the properties of 2-coherent lower probabilities appear rather weak: by Proposition 4, they are those of a capacity with the extra condition ii). We shall therefore investigate hereafter special subsets of bivariate *p*-boxes, with stronger consistency properties.

3.2. *Bivariate p-boxes that avoid sure loss*

In this subsection and the next, we are going to study which consistency properties of the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ determined by a bivariate *p*-box $(\underline{F}, \overline{F})$ by means of Eq. (11) can be characterized by the lower and upper distribution functions $\underline{F}, \overline{F}$.

We begin with the property of avoiding sure loss. Recall that, by Theorem 1, a lower probability $\underline{P}$ with domain $\mathcal{A}$ avoids sure loss if and only if there is a dF-coherent probability that dominates $\underline{P}$ on $\mathcal{A}$.

Lemma 1. *Let $(\underline{F}, \overline{F})$ be a bivariate p-box and $\underline{P}_{(\underline{F},\overline{F})}$ the lower probability it induces on $\mathcal{E}$ by means of Eq.* (11)*.*

(a) Let P be a dF-coherent probability on $\mathcal{E}$, and let F_P be its associated distribution function given by (13)*. Then*

$$P(A) \geq \underline{P}_{(\underline{F},\overline{F})}(A)\ \forall A \in \mathcal{E} \iff \underline{F} \leq F_P \leq \overline{F}.$$

(b) Conversely, let F be a distribution function on $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$, and let P_F be its associated dF-coherent probability on $\mathcal{E}$. Then

$$\underline{F} \leq F \leq \overline{F} \iff P_F(A) \geq \underline{P}_{(\underline{F},\overline{F})}(A)\ \forall A \in \mathcal{E}.$$

Proof. Let us establish the first statement; the proof for the second is analogous. On the one hand, given $A_{(x,y)} \in \mathcal{D} \subset \mathcal{E}$, it holds that

$$P(A_{(x,y)}) \geq \underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) \iff F_P(x, y) \geq \underline{F}(x, y),$$

where F_P is the distribution function associated with P by means of Eq. (13). On the other hand, given $A^c_{(x,y)} \in \mathcal{D}_c \subset \mathcal{E}$, it holds that $P(A^c_{(x,y)}) \geq \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)})$ if and only if $F_P(A_{(x,y)}) = P(A_{(x,y)}) = 1 - P(A^c_{(x,y)}) \leq 1 - \underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) = \overline{F}(x, y)$, where the last equality follows from Eq. (11). □

This allows us to conclude the following:

Proposition 6. *The lower probability $\underline{P}_{(\underline{F},\overline{F})}$ induced by the bivariate p-box $(\underline{F}, \overline{F})$ by means of Eq.* (11) *avoids sure loss if and only if there is a distribution function $F : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0, 1]$ satisfying $\underline{F} \leq F \leq \overline{F}$.*

Proof. $\underline{P}_{(\underline{F},\overline{F})}$ avoids sure loss iff there exists a dF-coherent probability P such that $P \geq \underline{P}_{(\underline{F},\overline{F})}$ on $\mathcal{E}$. By Lemma 1 this is equivalent to the thesis. □

This motivates the following definition:

Definition 9. We shall say that $(\underline{F}, \overline{F})$ *avoids sure loss* when the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ it induces by means of Eq. (11) does.

This notion is a minimal consistency requirement in order to interpret a bivariate p-box as a model for the imprecise knowledge of a bivariate distribution function, as it is equivalent to the existence of some distribution function compatible with the available bounds. Next, we investigate to which extent the notion of avoiding sure loss can be established in terms of $\underline{F}, \overline{F}$. By Proposition 6, a sufficient condition is that either $\underline{F}$ or $\overline{F}$ is a distribution function. It is not difficult to show that this condition is not necessary (simply take a distribution function F and $\underline{F} \leq F \leq \overline{F}$ such that neither $\underline{F}$ nor $\overline{F}$ are distribution functions). Our next result gives a necessary condition:

Proposition 7. *If $(\underline{F}, \overline{F})$ avoids sure loss, then for every $x_1 \leq x_2 \in \overline{\mathbb{R}}$ and every $y_1 \leq y_2 \in \overline{\mathbb{R}}$ it holds that*

$$\overline{F}(x_2, y_2) + \overline{F}(x_1, y_1) - \underline{F}(x_1, y_2) - \underline{F}(x_2, y_1) \geq 0. \tag{I-RI0}$$

Proof. Assume that $(\underline{F}, \overline{F})$ avoids sure loss. By Proposition 6, there is a distribution function F bounded by $\underline{F}, \overline{F}$. Given $x_1 \leq x_2$ and $y_1 \leq y_2 \in \overline{\mathbb{R}}$, it follows from Eq. (RI) that

$$\begin{aligned} 0 &\leq F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1) \\ &\leq \overline{F}(x_2, y_2) + \overline{F}(x_1, y_1) - \underline{F}(x_1, y_2) - \underline{F}(x_2, y_1), \end{aligned}$$

where the second inequality follows from $\underline{F} \leq F \leq \overline{F}$. □

This necessary condition is not sufficient in general. Before showing that, we must remark that although we consider bivariate p-boxes in $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$, we can deal with finite sets as particular cases:

Remark 8. With respect to the verification of the conditions related to avoiding sure loss and coherence for bivariate p-boxes, in many of the results and examples that follow, we shall consider maps $\underline{F}, \overline{F}$ related to discrete random numbers X, Y, whose jointly possible values are included into a finite subset $\mathcal{X} \times \mathcal{Y}$ of $\mathbb{R}^2$. Here the values of $\underline{F}, \overline{F}$ on $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ are determined by their values on $\mathcal{X} \times \mathcal{Y}$ (or by normalization at $(+\infty, +\infty)$, $(-\infty, \cdot)$, $(\cdot, -\infty)$), exactly like those of F in Remark 3.

It is easy to see that it is sufficient to verify that conditions such as (I-RI0) hold for $(x, y) \in \mathcal{X} \times \mathcal{Y}$. In fact, given $x_1 \leq x_2$ and $y_1 \leq y_2 \in \overline{\mathbb{R}}$, they hold trivially if either x_1 or y_1 is $-\infty$; if not, for any $(x_i, y_j) \notin \mathcal{X} \times \mathcal{Y}$, we may define

$$F^*(x_i, y_j) = F^*(x'_i, y'_j) \ (i, j = 1, 2), \tag{14}$$

where F^* is either $\overline{F}$ or alternatively $\underline{F}$ and x'_i, y'_j are given by (5). As a consequence, we obtain for instance that

$$\overline{F}(x_2,y_2)+\overline{F}(x_1,y_1)-\underline{F}(x_1,x_2)-\underline{F}(x_2,y_1)$$
$$=\overline{F}(x'_2,y'_2)+\overline{F}(x'_1,y'_1)-\underline{F}(x'_1,x'_2)-\underline{F}(x'_2,y'_1).$$

♦

Example 2. Let $\mathcal{X}\times\mathcal{Y}=\{1,2,3\}\times\{1,2,3\}$, and $\underline{F}$, $\overline{F}$ be given by:

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
$\underline{F}$	0	0.65	0.7	0.2	0.8	0.8	0.35	0.9	1
$\overline{F}$	0.1	0.7	0.7	0.25	0.8	0.8	0.4	0.9	1

It is immediate to check that both these maps are standardized and that together they satisfy Eq. (I-RI0). However, $(\underline{F},\overline{F})$ does not avoid sure loss: from Proposition 6, it suffices to show that there is no distribution function F satisfying $\underline{F}(x,y)\le F(x,y)\le\overline{F}(x,y)$ for every $x,y\in\{1,2,3\}$. To see that this is indeed the case, note that any distribution function $F\in(\underline{F},\overline{F})$ should satisfy

$$F(1,3)=0.7,\quad F(2,2)=0.8,\quad F(2,3)=0.8,\quad F(3,2)=0.9 \text{ and } F(3,3)=1.$$

Applying Eq. (RI) to $(x_1,y_1)=(1,2)$ and $(x_2,y_2)=(2,3)$, we deduce that $F(1,2)=0.7$. If we apply again the rectangle inequality, now to $(x_1,y_1)=(1,1)$ and $(x_2,y_2)=(2,2)$, we deduce that

$$F(2,2)+F(1,1)-F(1,2)-F(2,1)=0.8+F(1,1)-0.7-F(2,1)\ge 0,$$

i.e. $F(1,1)+0.1\ge F(2,1)$. From this inequality and since $F(1,1)\le\overline{F}(1,1)=0.1$, $F(2,1)\ge\underline{F}(2,1)=0.2$, we get $F(1,1)=0.1$, $F(2,1)=0.2$. If we now apply Eq. (RI) to $(x_1,y_1)=(2,1)$ and $(x_2,y_2)=(3,2)$, we deduce that

$$F(3,2)+F(2,1)-F(2,2)-F(3,1)=0.9+0.2-0.8-F(3,1)\ge 0,$$

i.e. $F(3,1)\le 0.3$. But on the other hand we must have $F(3,1)\ge\underline{F}(3,1)=0.35$, a contradiction. Hence, $(\underline{F},\overline{F})$ does not avoid sure loss. ♦

Interestingly, (I-RI0) is a necessary and sufficient condition when $\underline{F}$ and $\overline{F}$ describe binary and logically independent random numbers:

Proposition 8. *Let X, Y be binary random numbers, with domain, respectively, $\mathcal{X}=\{x_1,x_2\}$ and $\mathcal{Y}=\{y_1,y_2\}$. Let also $(\underline{F},\overline{F})$ be a bivariate p-box determined by its values on $\mathcal{X}\times\mathcal{Y}$. Then, given the following statements*

(a) $(\underline{F},\overline{F})$ *avoids sure loss;*
(b) $\overline{F}(x_2,y_2)+\overline{F}(x_1,y_1)-\underline{F}(x_1,y_2)-\underline{F}(x_2,y_1)\ge 0$;
(c) $\underline{F}(x_2,y_2)+\overline{F}(x_1,y_1)-\underline{F}(x_1,y_2)-\underline{F}(x_2,y_1)\ge 0$;

it holds that (a) ⇒ (b) ⇒ (c). If in addition X, Y are logically independent then (a), (b), (c) are equivalent.

Proof. The first statement implies the second by Proposition 7. To see that the second implies the third note that, since $\underline{F}, \overline{F}$ are standardized maps and they are determined by Eq. (14), it must be $1 = \underline{F}(+\infty, +\infty) = \underline{F}(x_2, y_2)$ and $1 = \overline{F}(+\infty, +\infty) = \overline{F}(x_2, y_2)$.

To see that the third statement implies the first when X, Y are logically independent, define $F : \{x_1, x_2\} \times \{y_1, y_2\} \to [0, 1]$ by

$$F(x_1, y_1) = \overline{F}(x_1, y_1) \qquad F(x_1, y_2) = \max\{\overline{F}(x_1, y_1), \underline{F}(x_1, y_2)\}$$
$$F(x_2, y_1) = \max\{\overline{F}(x_1, y_1), \underline{F}(x_2, y_1)\} \qquad F(x_2, y_2) = 1,$$

and let us extend it to $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ by Eq. (14). By construction, F is a standardized map bounded by $\underline{F}, \overline{F}$. To see that it is a distribution function, use Remark 8 and note that if either $F(x_1, y_2)$ or $F(x_2, y_1)$ is equal to $\overline{F}(x_1, y_1) = F(x_1, y_1)$, then it follows from the componentwise monotonicity of $\underline{F}, \overline{F}$ that

$$\Delta F(x_1, x_2; y_1, y_2) = F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1) \geq 0.$$

Otherwise, if $F(x_1, y_2) = \underline{F}(x_1, y_2)$ and $F(x_2, y_1) = \underline{F}(x_2, y_1)$, then

$$\Delta F(x_1, x_2; y_1, y_2) = \underline{F}(x_2, y_2) + \overline{F}(x_1, y_1) - \underline{F}(x_1, y_2) - \underline{F}(x_2, y_1) \geq 0.$$

This proves that F satisfies (RI). By applying Proposition 2 and recalling also Remark 4 we conclude that F is a distribution function. Then, $(\underline{F}, \overline{F})$ avoids sure loss by Proposition 6. □

3.3. *Coherent bivariate p-boxes*

Let us turn now to coherence. We begin by establishing a result akin to Proposition 6:

Proposition 9. *The lower probability $\underline{P}_{(\underline{F},\overline{F})}$ induced by the bivariate p-box $(\underline{F}, \overline{F})$ by means of Eq. (11) is coherent if and only if $\underline{F}$ (resp., $\overline{F}$) is the lower (resp., upper) envelope of the set*

$$\mathcal{F} = \{F : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0, 1] \text{ distribution function} : \underline{F} \leq F \leq \overline{F}\}. \tag{15}$$

Proof. Denote by $\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})$ the set of dF-coherent probabilities associated with $\underline{P}_{(\underline{F},\overline{F})}$ by means of Eq. (7). Assume first that $\underline{F}$ and $\overline{F}$ are the lower and upper envelopes of $\mathcal{F}$. We get

$$\begin{aligned}\underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) = \underline{F}(x, y) &= \inf_{F \in \mathcal{F}} F(x, y) \\ &= \inf_{F \in \mathcal{F}} P_F(A_{(x,y)}) = \inf_{P \in \mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} P(A_{(x,y)})\end{aligned}$$

and

$$\begin{aligned}\underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) &= 1 - \overline{F}(x, y) = 1 - \sup_{F \in \mathcal{F}} F(x, y) = 1 - \sup_{F \in \mathcal{F}} P_F(A_{(x,y)}) \\ &= \inf_{F \in \mathcal{F}} P_F(A^c_{(x,y)}) = \inf_{P \in \mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} P(A^c_{(x,y)}),\end{aligned}$$

where the last equality follows in both derivations by Lemma 1. We conclude from this that $\underline{P}_{(\underline{F},\overline{F})}$ is coherent. Conversely, if $\underline{P}_{(\underline{F},\overline{F})}$ is coherent, we get

$$\begin{aligned}\underline{F}(x,y) = \underline{P}_{(\underline{F},\overline{F})}(A_{(x,y)}) &= \inf_{P\in\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} P(A_{(x,y)}) \\ &= \inf_{P\in\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} F_P(x,y) = \inf_{F\in\mathcal{F}} F(x,y)\end{aligned}$$

and

$$\begin{aligned}\overline{F}(x,y) &= 1-\underline{P}_{(\underline{F},\overline{F})}(A^c_{(x,y)}) = 1-\inf_{P\in\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} P(A^c_{(x,y)}) = \sup_{P\in\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} P(A_{(x,y)}) \\ &= \sup_{P\in\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})} F_P(x,y) = \sup_{F\in\mathcal{F}} F(x,y),\end{aligned}$$

again using Lemma 1 for the last equalities. □

This motivates the following definition:

Definition 10. A bivariate p-box $(\underline{F},\overline{F})$ is *coherent* iff its associated lower probability $\underline{P}_{(\underline{F},\overline{F})}$ is.

Thus, from a sensitivity analysis point of view, only coherent bivariate p-boxes make sense, since they are the ones that can be regarded as equivalent to a set of bivariate distribution functions. One interesting difference with the univariate case is that the bounds $\underline{F},\overline{F}$ of the bivariate p-box need not be distribution functions for $(\underline{F},\overline{F})$ to be coherent (although if $\underline{F},\overline{F}$ are distribution functions then trivially $(\underline{F},\overline{F})$ is coherent by Proposition 9, since both $\underline{F}$, $\overline{F}$ belong to $\mathcal{F}$). This can be seen for instance with Example 1, where the lower envelope of a set of distribution functions (which determines the lower distribution function of a coherent bivariate p-box) is not a distribution function itself.

As for the condition of avoiding sure loss, if a bivariate p-box $(\underline{F},\overline{F})$ avoids sure loss but is not coherent, then it is necessary that at least one of $\underline{F}$, $\overline{F}$ is not a distribution function. One instance of this is provided in Example 4 in the next subsection.

The above remark suggests a further difference with the univariate case. Let the p-box $(\underline{F},\overline{F})$ avoid sure loss, and $\underline{F}$, $\overline{F}$ be univariate. Then, recalling Remark 5, $\underline{F}$, $\overline{F}$ are cdfs and, by Theorem 2, $\underline{P}_{\underline{F},\overline{F}}$ is a coherent lower probability. Hence, by Definition 10, $(\underline{F},\overline{F})$ is a coherent p-box. The interesting conclusion is that the concepts of coherence and of avoiding sure loss are distinct with bivariate p-boxes, but undistinguishable with univariate ones.

3.4. *Properties of coherent bivariate p-boxes*

Comparing Propositions 1 and 2 we realize that adding one dimension, from $n=1$ to $n=2$, requires the additional conditions in Proposition 2 (c) in the characterization of dF-coherent probabilities on (componentwise) monotone families of events.

The other conditions in Proposition 2 are simple generalizations of the corresponding ones of Proposition 1. When we consider lower probabilities on $\mathcal{E}$, we might wonder whether there exists some analogue of condition (c), and in particular of the rectangle inequality (RI). As we shall see in what follows, the situation is more complex, involving various imprecise rectangle inequalities and related properties.

As a preliminary step, we recall the following properties of coherent upper and lower probabilities (see Walley [3], Section 2.4.7):

$$A \wedge B = \varnothing \Rightarrow \underline{P}(A \vee B) \geq \underline{P}(A) + \underline{P}(B). \tag{16}$$
$$A \wedge B = \varnothing \Rightarrow \overline{P}(A \vee B) \geq \underline{P}(A) + \overline{P}(B). \tag{17}$$
$$\underline{P}(A \vee B) + \overline{P}(A \wedge B) \geq \underline{P}(A) + \underline{P}(B). \tag{18}$$
$$\overline{P}(A \vee B) + \underline{P}(A \wedge B) \geq \underline{P}(A) + \underline{P}(B). \tag{19}$$
$$\overline{P}(A \vee B) + \overline{P}(A \wedge B) \geq \overline{P}(A) + \underline{P}(B). \tag{20}$$

These properties are useful in obtaining imprecise rectangle inequalities as necessary conditions for coherence of a bivariate p-box:

Proposition 10. (Imprecise Rectangle Inequalities) *Let $(\underline{F}, \overline{F})$ be a bivariate p-box on $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$. If it is coherent, then the following conditions hold for every $x_1 \leq x_2 \in \overline{\mathbb{R}}$ and $y_1 \leq y_2 \in \overline{\mathbb{R}}$:*

$$\underline{F}(x_2, y_2) + \overline{F}(x_1, y_1) - \underline{F}(x_1, y_2) - \underline{F}(x_2, y_1) \geq 0. \tag{I-RI1}$$
$$\overline{F}(x_2, y_2) + \underline{F}(x_1, y_1) - \underline{F}(x_1, y_2) - \underline{F}(x_2, y_1) \geq 0. \tag{I-RI2}$$
$$\overline{F}(x_2, y_2) + \overline{F}(x_1, y_1) - \overline{F}(x_1, y_2) - \underline{F}(x_2, y_1) \geq 0. \tag{I-RI3}$$
$$\overline{F}(x_2, y_2) + \overline{F}(x_1, y_1) - \underline{F}(x_1, y_2) - \overline{F}(x_2, y_1) \geq 0. \tag{I-RI4}$$

Proof. Consider (x_1, y_1) and (x_2, y_2) such that $x_1 \leq x_2$, $y_1 \leq y_2$ and let $\underline{P}_{(\underline{F},\overline{F})}$ be the coherent lower probability induced by $(\underline{F}, \overline{F})$ by means of Eq. (11). Recall also in the following derivations that $A_{(x_1,y_2)} \wedge A_{(x_2,y_1)} = A_{(x_1,y_1)}$.

[Proof of (I-RI1)**].** It holds that:

$$\begin{aligned}\underline{P}(A_{(x_2,y_2)}) &\overset{Eq.\ (16)}{\geq} \underline{P}(A_{(x_1,y_2)} \vee A_{(x_2,y_1)}) + \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) \\ &\overset{Eq.\ (18)}{\geq} \underline{P}(A_{(x_1,y_2)}) + \underline{P}(A_{(x_2,y_1)}) - \overline{P}(A_{(x_1,y_1)}) \\ &\quad + \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c).\end{aligned}$$

Thus:

$$\begin{aligned}&\underline{P}(A_{(x_2,y_2)}) - \underline{P}(A_{(x_1,y_2)}) - \underline{P}(A_{(x_2,y_1)}) + \overline{P}(A_{(x_1,y_1)}) \\ &\qquad \geq \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) \geq 0.\end{aligned}$$

If we write the previous equation in terms of the maps $\underline{F}, \overline{F}$, we obtain that:

$$\underline{F}(x_2, y_2) - \underline{F}(x_1, y_2) - \underline{F}(x_2, y_1) + \overline{F}(x_1, y_1) \geq 0.$$

[Proof of (I-RI2)**].** It holds that:

$$\begin{aligned}\overline{P}(A_{(x_2,y_2)}) &\overset{Eq.\ (17)}{\geq} \overline{P}(A_{(x_1,y_2)} \vee A_{(x_2,y_1)}) + \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) \\ &\overset{Eq.\ (19)}{\geq} \underline{P}(A_{(x_1,y_2)}) + \underline{P}(A_{(x_2,y_1)}) - \underline{P}(A_{(x_1,y_1)}) \\ &\quad + \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c).\end{aligned}$$

Then:

$$\begin{aligned}&\overline{P}(A_{(x_2,y_2)}) + \underline{P}(A_{(x_1,y_1)}) - \underline{P}(A_{(x_1,y_2)}) - \underline{P}(A_{(x_2,y_1)}) \\ &\qquad\qquad \geq \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) \geq 0.\end{aligned}$$

In terms of $\underline{F}, \overline{F}$, this means that

$$\overline{F}(x_2,y_2) + \underline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.$$

[Proof of (I-RI3) **and** (I-RI4)**].** Analogously:

$$\overline{P}(A_{(x_2,y_2)}) \overset{Eq.\ (17)}{\geq} \overline{P}(A_{(x_1,y_2)} \vee A_{(x_2,y_1)}) + \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c)$$

and, from Eq. (20), this is greater than or equal to both

$$\underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) + \underline{P}(A_{(x_1,y_2)}) + \overline{P}(A_{(x_2,y_1)}) - \overline{P}(A_{(x_1,y_1)})$$

and

$$\underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) + \overline{P}(A_{(x_1,y_2)}) + \underline{P}(A_{(x_2,y_1)}) - \overline{P}(A_{(x_1,y_1)}).$$

Then:

$$\begin{aligned}0 &\leq \underline{P}(A_{(x_2,y_2)} \wedge (A_{(x_1,y_2)} \vee A_{(x_2,y_1)})^c) \\ &\qquad \leq \begin{cases}\overline{P}(A_{(x_2,y_2)}) - \underline{P}(A_{(x_1,y_2)}) - \overline{P}(A_{(x_2,y_1)}) + \overline{P}(A_{(x_1,y_1)}). \\ \overline{P}(A_{(x_2,y_2)}) - \overline{P}(A_{(x_1,y_2)}) - \underline{P}(A_{(x_2,y_1)}) + \overline{P}(A_{(x_1,y_1)}).\end{cases}\end{aligned}$$

In terms of $\underline{F}, \overline{F}$, this means that:

$$\begin{aligned}\overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \overline{F}(x_2,y_1) \geq 0; \\ \overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \overline{F}(x_1,y_2) - \underline{F}(x_2,y_1) \geq 0.\end{aligned}$$

This completes the proof. □

Inequality (RI) is stated in terms of $\Delta F(x_1,x_2;y_1,y_2)$, the second order mixed difference. It is interesting to explicit the corresponding differences

$$\Delta\underline{F}(x_1,x_2;y_1,y_2) = \underline{F}(x_2,y_2) + \underline{F}(x_1,y_1) - \underline{F}(x_1,y_2) - \underline{F}(x_2,y_1)$$

and

$$\Delta\overline{F}(x_1,x_2;y_1,y_2) = \overline{F}(x_2,y_2) + \overline{F}(x_1,y_1) - \overline{F}(x_1,y_2) - \overline{F}(x_2,y_1)$$

in the inequalities (I-RI1)÷(I-RI4). With simple algebraic computations, we obtain that (I-RI1) and (I-RI2) are equivalent to

$$\Delta\underline{F}(x_1, x_2; y_1, y_2) + \min_{i=1,2}\{\overline{F}(x_i, y_i) - \underline{F}(x_i, y_i)\} \geq 0 \tag{21}$$

while (I-RI3) and (I-RI4) are equivalent to

$$\Delta\overline{F}(x_1, x_2; y_1, y_2) + \min_{i,j=1,2,i\neq j}\{\overline{F}(x_i, y_j) - \underline{F}(x_i, y_j)\} \geq 0. \tag{22}$$

We deduce from (21) that $\Delta\underline{F}$, unlike ΔF, may possibly be negative. In this case however, the uncertainty evaluation is necessarily imprecise at the points (x_1, y_1), (x_2, y_2). Equation (22) provides the same insight: when $\Delta\overline{F}$ is negative, $\overline{F}$ is necessarily greater than $\underline{F}$ at (x_1, y_2) and at (x_2, y_1).

Another difference between ΔF and $\Delta\underline{F}$ or $\Delta\overline{F}$ is that ΔF is the probability of a half-open 'rectangle' with vertices (x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2), and as such must be 0 when there is no admissible value of X, Y in the rectangle (this is condition (3) in Proposition 2, see also Remark 3). On the contrary, $\Delta\underline{F}$ and $\Delta\overline{F}$ have no analogue meaning, and may be non-zero even when ΔF must be zero. One key issue here is that $\underline{F}, \overline{F}$ need not be distribution functions themselves for the bivariate p-box $(\underline{F}, \overline{F})$ to be coherent (recall again Example 1). The following simple example illustrates this further, making use of an assumption of logical dependence:

Example 3. Let X, Y be binary random numbers, both with domain $\mathcal{X} = \mathcal{Y} = \{1, 2\}$. Assume that X and Y are logically dependent, and cannot simultaneously be equal to 2, i.e. $(X = 2) \wedge (Y = 2) = \varnothing$.

Given the following probabilities P_1, P_2 on $\mathcal{X} \times \mathcal{Y}$:

	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
P_1	0.25	0.50	0.25	0
P_2	0.50	0.10	0.40	0

we obtain the corresponding cdfs F_1, F_2 and their lower (upper) envelope $\underline{F}$ ($\overline{F}$):

	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
F_1	0.25	0.75	0.50	1
F_2	0.50	0.60	0.90	1
$\underline{F}$	0.25	0.60	0.50	1
$\overline{F}$	0.50	0.75	0.90	1

Clearly, $(\underline{F}, \overline{F})$ is a coherent bivariate p-box. Note that $(X = 2) \wedge (Y = 2) = (1 < X \leq 2) \wedge (1 < Y \leq 2) = \varnothing$, and, as required by (3), we have in fact that $\Delta F_1 = \Delta F_2 = 0$.[c] As for $\underline{F}$, $\overline{F}$, we have instead $\Delta\underline{F} = 1 + 0.25 - 0.60 - 0.50 = 0.15$, $\Delta\overline{F} = 1 + 0.50 - 0.75 - 0.90 = -0.15$. ♦

[c]Referring to Remark 2, here $A_{(1,1)} = (X \leq 1) \wedge (Y \leq 1) \neq A_{(2,2)} = \Omega$, but $({}_1A_2 \wedge {}_1A_1^c) \wedge ({}_2A_2 \wedge {}_2A_1^c) = (1 < X \leq 2) \wedge (1 < Y \leq 2) = \varnothing$.

None of the four rectangle inequalities in Proposition 10 is implied by the remaining ones. We show this for (I-RI3) and (I-RI4) in Example 4; similar examples can be devised for the remaining inequalities. Since each inequality alone is necessary for coherence, no subset of three or fewer such inequalities is in general sufficient for the coherence of a bivariate p-box.

Example 4. Consider the functions $\underline{F}$ and $\overline{F}$ defined by:

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
$\underline{F}$	0	0.3	0.45	0.3	0.6	0.75	0.45	0.8	1
$\overline{F}$	0	0.3	0.5	0.3	0.6	0.85	0.5	0.85	1

Both $\underline{F}$ and $\overline{F}$ are standardized maps. In addition, $\underline{F}$ is a distribution function, and consequently $\underline{F}$ and $\overline{F}$ satisfy Eqs. (I-RI1) and (I-RI2). It can be checked that Eq. (I-RI4) is also satisfied. However, Eq. (I-RI3) is not: $\overline{F}(3,3)+\overline{F}(2,2)-\overline{F}(2,3)-\underline{F}(3,2) = 1+0.6-0.85-0.8 = -0.05 < 0$.

Similarly, if we define $\underline{F}^*$ and $\overline{F}^*$ by $\underline{F}^*(x,y) = \underline{F}(y,x)$ and $\overline{F}^*(x,y) = \overline{F}(y,x)$, we obtain an example where Eqs. (I-RI1), (I-RI2) and (I-RI3) are satisfied but (I-RI4) is not. ♦

A natural question arising at this stage is whether the rectangle inequalities (I-RI1)÷(I-RI4) might have a role similar to inequality (RI) in Proposition 2. In the precise case, (RI) together with standardisation is sufficient for characterizing dF-coherence on $\mathcal{D}$ whenever (3) need not be checked, which happens in several instances, including logical independence of X and Y (cf. Remark 3). We might expect that (I-RI1)÷(I-RI4) are sufficient for the coherence of a bivariate p-box under similar conditions, for instance if X and Y are logically independent. Although it is at present unclear whether this is true, we provide in Theorem 3 an affirmative answer when one of the random numbers is binary. The following proposition is preliminary to this.

Proposition 11. *Let X, Y be two logically independent random numbers, jointly taking all values in $\mathcal{X}\times\mathcal{Y} = \{x_1,\ldots,x_n\}\times\{y_1,y_2\}$. If their bivariate p-box $(\underline{F},\overline{F})$ satisfies* (I-RI1)÷(I-RI4)*, then $(\underline{F},\overline{F})$ is coherent.*

Proof. The proof is based on the following four statements:

(a) If (I-RI1) holds, there exists a cdf F_1 such that $F_1(x_i,y_2) = \underline{F}(x_i,y_2)$ $(i = 1,\ldots,n)$, and $\underline{F} \le F_1 \le \overline{F}$.
(b) If (I-RI2) holds, there exists a cdf F_2 such that $F_2(x_i,y_1) = \underline{F}(x_i,y_1)$ $(i = 1,\ldots,n)$, and $\underline{F} \le F_2 \le \overline{F}$.
(c) If (I-RI3) holds, there exists a cdf F_3 such that $F_3(x_i,y_1) = \overline{F}(x_i,y_1)$ $(i = 1,\ldots,n)$, and $\underline{F} \le F_3 \le \overline{F}$.
(d) If (I-RI4) holds, there exists a cdf F_4 such that $F_4(x_i,y_2) = \overline{F}(x_i,y_2)$ $(i = 1,\ldots,n)$, and $\underline{F} \le F_4 \le \overline{F}$.

In fact, assuming for the moment that (a)÷(d) hold, $\underline{F}$ ($\overline{F}$) is the lower (resp., upper) envelope of $\{F_1, F_2\}$ (resp., of $\{F_3, F_4\}$), while $\underline{F} \leq F_i \leq \overline{F}$, for $i = 1, \ldots, 4$. Therefore, recalling Definition 10 and Proposition 9, the thesis follows.

As for (a)÷(d), each of these statements can be proven with a constructive procedure: a function F_i is obtained sequentially and it is verified that F_i is a cdf, such that $\underline{F} \leq F_i \leq \overline{F}$.

The four proofs are similar in their structure, but lengthy, hence we shall report the details for one of them.

Pseudo-Algorithm 1 Procedure defining F_1

---

if $\Delta_{\underline{F}}^{R_i} \geq 0$ for all i **then**
  $F_1 = \underline{F}$ is a cdf;
else
  Let h be the smallest positive integer such that $\Delta_{\underline{F}}^{R_{n-1}} \geq 0, \ldots, \Delta_{\underline{F}}^{R_{n-h+1}} \geq 0$, $\Delta_{\underline{F}}^{R_{n-h}} < 0$;
  Define $F_1 = \underline{F}$ at all vertices of $R_{n-1}, \ldots, R_{n-h+1}$.
  Let $k \geq h$ be such that

$$\Delta_{\underline{F}}^{R_{n-h}} < 0, \ldots, \Delta_{\underline{F}}^{R_{n-h}} + \ldots + \Delta_{\underline{F}}^{R_{n-k}} < 0 \tag{23}$$

and

$$\text{either } n - k = 1 \text{ or } \Delta_{\underline{F}}^{R_{n-h}} + \ldots + \Delta_{\underline{F}}^{R_{n-k}} + \Delta_{\underline{F}}^{R_{n-k-1}} \geq 0 \tag{24}$$

  for $j = 0, \ldots, k - h$ **do**
    Define

$$F_1(x_{n-h-j}, y_1) = \underline{F}(x_{n-h-j}, y_1) - \sum_{s=0}^{j} \Delta_{\underline{F}}^{R_{n-h-s}} \tag{25}$$

and $F_1(\cdot, y_2) = \underline{F}(\cdot, y_2)$ for the $(\cdot, y_2)$ vertices of $R_{n-h}, \ldots, R_{n-k}$.
  end for
  if $n - k = 1$ **then**
    F_1 is a cdf
  else
    Note that necessarily $\Delta_{\underline{F}}^{R_{n-k-1}} > 0$ (since $\Delta_{\underline{F}}^{R_{n-h}} + \ldots + \Delta_{\underline{F}}^{R_{n-k}} < 0$, $(\Delta_{\underline{F}}^{R_{n-h}} + \ldots + \Delta_{\underline{F}}^{R_{n-k}}) + \Delta_{\underline{F}}^{R_{n-k-1}} \geq 0$)
    Iterate the procedure with starting rectangle R_{n-k-1} at the next iteration, until R_1 is reached.
  end if
end if

In all four cases, the procedure operates with the *elementary rectangles* R_i, $i = 1, \ldots, n-1$, with vertices (x_i, y_1), (x_i, y_2), (x_{i+1}, y_1), (x_{i+1}, y_2). We will refer to (x_i, y_1) as the *South-West* (SW) vertex of the rectangle R_i (and will describe analogously the other vertices).

To shorten the notation, relabel the second order difference $\Delta\underline{F}(x_i, x_{i+1}; y_1, y_2)$ in terms of the rectangle R_i it refers to:

$$\Delta_{\underline{F}}^{R_i} \coloneqq \underline{F}(x_{i+1}, y_2) + \underline{F}(x_i, y_1) - \underline{F}(x_{i+1}, y_1) - \underline{F}(x_i, y_2), \tag{26}$$

while $\Delta_{\overline{F}}^{R_i}$ (or more generally $\Delta_F^{R_i}$, F generic function) is defined replacing $\underline{F}$ with $\overline{F}$ (with F) in (26). The following fact will be exploited later on:

$$\Delta_F^{R_i} + \Delta_F^{R_{i+1}} + \ldots + \Delta_F^{R_{i+r}} = \Delta F(x_i, x_{i+r+1}; y_1, y_2) \text{ for any given function } F. \tag{27}$$

Checking that (27) holds is immediate, as non-extreme terms cancel pairwise and appear in exactly two contiguous rectangles in the left-hand summation.

Let us now prove statement (a).

Proof of (a). The idea is to obtain F_1 by putting $F_1 = \underline{F}$ when possible, and modifying this assessment on (a subset of) the points (x_i, y_1) according to the values of $\Delta_{\underline{F}}^{R_i}$, for each i. This is to ensure $\Delta_{F_1}^{R_i} \geq 0$, for all i, which guarantees by (27) that any $\Delta F_1(x_i, x_{i+h+1}; y_1, y_2)$ is non-negative, one of the conditions for F_1 to be a cdf. We check the values $\Delta_{\underline{F}}^{R_i}$ sequentially, from $i = n-1$ to $i = 1$.

Basically, this procedure (cf. Pseudo-Algorithm 1) identifies at each iteration a sequence of *non-regular* rectangles (corresponding to differences $\Delta_{\underline{F}}^{R_i} < 0$) and defines F_1 by means of (25) so that $F_1 \neq \underline{F}$ at the SW vertices of the non-regular rectangles.

Let us check that F_1 is then a cdf.

By construction $F_1 \geq \underline{F}$. Then, it only remains to show that at a generic (x_i, y_1) it holds that:

$$F_1(x_i, y_1) \leq \min\{F_1(x_{i+1}, y_1), F_1(x_i, y_2), \overline{F}(x_i, y_1)\} \text{ and } \Delta_{F_1}^{R_i} \geq 0. \tag{28}$$

If $F_1 = \underline{F}$ at the vertices of rectangle R_i, there is nothing to prove ((28) holds for the SW vertex (x_i, y_1) of R_i). We only have to check (28) for the SW vertices (x_i, y_1) of each non-regular rectangle of each iteration, and (partly) for the SW vertex of the first 'regular' rectangle after each sequence (R_{n-k-1} at the first iteration). The latter check is necessary because $F_1 \neq \underline{F}$ after the first iteration at the South-East vertex of this rectangle (at (x_{n-k}, y_1) in the first iteration), which is the SW vertex of its right-contiguous, and non-regular, rectangle. $F_1 \neq \underline{F}$ only at this vertex of the regular rectangle. Therefore, the check is partial: $F_1(x_i, y_1) \leq \min\{F_1(x_i, y_2), \overline{F}(x_i, y_2)\}$ holds trivially. To keep the notation simpler, we shall illustrate this partial check referring to R_{n-k-1}.

Let $F_1(x_i, y_1) \neq \underline{F}(x_i, y_1)$. Note that the procedure followed in Pseudo-Algorithm 1, Eq. (25), ensures that $F_1(x_i, y_1) = \underline{F}(x_i, y_1) - (\Delta_{\underline{F}}^{R_i} + \ldots + \Delta_{\underline{F}}^{R_{i+r}})$ for some $r \in \mathbb{N}^+$ (for instance, when $i = n-k$, $r = k-h$). Applying (27) we obtain

$$F_1(x_i, y_1) = -\underline{F}(x_{i+r+1}, y_2) + \underline{F}(x_{i+r+1}, y_1) + \underline{F}(x_i, y_2). \tag{29}$$

We obtain further that:

- $F_1(x_i, y_1) \leq \overline{F}(x_i, y_1)$: use for this inequality (I-RI1) to majorize the right-hand term of (29).
- $F_1(x_i, y_2) - F_1(x_i, y_1) = \underline{F}(x_{i+r+1}, y_2) - \underline{F}(x_{i+r+1}, y_1) \geq 0$ (using (29) at the equality and $F_1(x_i, y_2) = \underline{F}(x_i, y_2)$), i.e. $F_1(x_i, y_1) \leq F_1(x_i, y_2)$.
- $F_1(x_i, y_1) \leq F_1(x_{i+1}, y_1)$. To prove this, we distinguish two cases:
  1) If (x_i, y_1) is the SW vertex of a non-regular rectangle R_i,[d] then

$$\begin{aligned}F_1(x_{i+1}, y_1) - F_1(x_i, y_1) &= \underline{F}(x_{i+1}, y_1) - (\Delta_{\underline{F}}^{R_{i+1}} + \ldots + \Delta_{\underline{F}}^{R_{i+r}}) \\ &- \underline{F}(x_i, y_1) + \Delta_{\underline{F}}^{R_i} + \Delta_{\underline{F}}^{R_{i+1}} + \ldots + \Delta_{\underline{F}}^{R_{i+r}} \\ &= \underline{F}(x_{i+1}, y_1) - \underline{F}(x_i, y_1) + \Delta_{\underline{F}}^{R_i} = \underline{F}(x_{i+1}, y_2) - \underline{F}(x_i, y_2) \geq 0;\end{aligned}$$

  2) If (x_i, y_1) is the SW vertex of R_{n-k-1} (in general, of the first regular rectangle after an iteration), then $F_1(x_{n-k}, y_1) - F_1(x_{n-k-1}, y_1) = \underline{F}(x_{n-k}, y_1) - (\Delta_{\underline{F}}^{R_{n-h}} + \ldots + \Delta_{\underline{F}}^{R_{n-k}}) - \underline{F}(x_{n-k-1}, x_1) \geq 0$, using the monotonicity of $\underline{F}$ and (23) at the inequality.
- $\Delta_{F_1}^{R_i} \geq 0$. We distinguish two cases, like the preceding bullet.
  1) If R_i is a non-regular rectangle[e], then

$$\begin{aligned}\Delta_{F_1}^{R_i} &= \underline{F}(x_{i+1}, y_2) - \underline{F}(x_i, y_2) - \underline{F}(x_{i+1}, y_1) + \Delta_{\underline{F}}^{R_{i+1}} + \ldots \\ &+ \Delta_{\underline{F}}^{R_{i+r}} + \underline{F}(x_i, y_1) - (\Delta_{\underline{F}}^{R_i} + \Delta_{\underline{F}}^{R_{i+1}} + \ldots + \Delta_{\underline{F}}^{R_{i+r}}) = \Delta_{\underline{F}}^{R_i} - \Delta_{\underline{F}}^{R_i} = 0;\end{aligned}$$

  2) If $R_i = R_{n-k-1}$, then

$$\begin{aligned}\Delta_{F_1}^{R_{n-k-1}} &= \underline{F}(x_{n-k}, y_2) - \underline{F}(x_{n-k-1}, y_2) - \underline{F}(x_{n-k}, y_1) + \Delta_{\underline{F}}^{R_{n-h}} + \ldots \\ &+ \Delta_{\underline{F}}^{R_{n-k}} + \underline{F}(x_{n-k-1}, y_1) = \Delta_{\underline{F}}^{R_{n-h}} + \ldots + \Delta_{\underline{F}}^{R_{n-k}} + \Delta_{\underline{F}}^{R_{n-k-1}} \geq 0,\end{aligned}$$

  using the second condition in (24) at the inequality.

Proof of (b), (c), (d).

Similar to the proof of (a). The procedures for defining F_2, F_3, F_4 again modify either $\underline{F}$ or $\overline{F}$ at one vertex of each non-regular rectangle. A major difference is whether a West or an East vertex is modified. For East vertices, the procedure starts with R_1, for West vertices with R_{n-1} (like (a) - this is only the case of (c), whose procedure modifies the North-West vertices).

The procedures starting with R_1 are those for proving (b) and (d). We report the procedure defining F_2 in the proof of (b) (cf. Pseudo-Algorithm 2). It modifies, at each iteration, the North-East vertex of each non-regular rectangle. □

Proposition 11 allows us to establish the following:

[d] If R_i is the first non-regular rectangle of a sequence, $\Delta_{\underline{F}}^{R_{i+1}} + \ldots + \Delta_{\underline{F}}^{R_{i+r}} = 0$ in the following computation.
[e] See footnote d.

Pseudo-Algorithm 2 Procedure defining F_2

if $\Delta_{\underline{F}}^{R_i} \geq 0$ for all i **then**
  $F_2 = \underline{F}$ is a cdf;
else
  Let $h \in \mathbb{N}^+$ be such that $\Delta_{\underline{F}}^{R_1} \geq 0, \ldots, \Delta_{\underline{F}}^{R_{h-1}} \geq 0$, $\Delta_{\underline{F}}^{R_h} < 0$;
  Define $F_2 = \underline{F}$ at all vertices of $R_1, \ldots, R_{h-1}$.
  Let $k \geq h$ be the smallest integer such that

$$\Delta_{\underline{F}}^{R_h} < 0, \ldots, \Delta_{\underline{F}}^{R_h} + \ldots + \Delta_{\underline{F}}^{R_k} < 0$$

and

$$\text{either } k = n-1 \text{ or } \Delta_{\underline{F}}^{R_h} + \ldots + \Delta_{\underline{F}}^{R_k} + \Delta_{\underline{F}}^{R_{k+1}} \geq 0$$

  for $j = 0, \ldots, k-h$ **do**
    Define

$$F_2(x_{h+1+j}, y_2) = \underline{F}(x_{h+1+j}, y_2) - \sum_{s=0}^{j} \Delta_{\underline{F}}^{R_{h+s}}$$

and $F_2(\cdot, y_1) = \underline{F}(\cdot, y_1)$ for the $(\cdot, y_1)$ vertices of $R_h, \ldots, R_k$.
  end for
  if $k = n-1$ **then**
    F_2 is a cdf
  else
    Note that necessarily $\Delta_{\underline{F}}^{R_{k+1}} > 0$
    Iterate the procedure with starting rectangle (at the second run) R_{k+1}, until R_{n-1} is reached.
  end if
end if

Theorem 3. *Let $\underline{F}, \overline{F} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0,1]$ be the bivariate p-box of a couple of logically independent random numbers (X, Y), with X arbitrary, i.e. $\mathcal{X} \subseteq \mathbb{R}$, and Y binary ($\mathcal{Y} = \{y_1, y_2\}$). Then,*

$$(\underline{F}, \overline{F}) \text{ is coherent } \iff \underline{F}, \overline{F} \text{ satisfy (I-RI1)} \div \text{(I-RI4)}.$$

Proof. The direct implication has been established in Proposition 10. To prove the converse, let us show that the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ defined on $\mathcal{E}$ by Eq. (11) is coherent. From Definition 4 (b), $\underline{P}_{(\underline{F},\overline{F})}$ is coherent iff $\forall n$, $\forall s_0, \ldots, s_n \geq 0$, $\forall E_0, \ldots, E_n \in \mathcal{E}$, it holds that

$$\max \left[\sum_{i=1}^{n} s_i (I_{E_i} - \underline{P}_{(\underline{F},\overline{F})}(E_i)) - s_0 (I_{E_0} - \underline{P}_{(\underline{F},\overline{F})}(E_0)) \right] \geq 0. \tag{30}$$

Now, for every $E_i \in \mathcal{E}$ $(i = 0, \ldots, n)$, there exist $x_i \in \overline{\mathbb{R}}$, $y_j \in \mathcal{Y}$ such that $E_i \in \{A_{(x_i,y_j)}, A^c_{(x_i,y_j)}\}$.

Define now $\mathcal{X}' = \{x_0, \ldots, x_n\}$ and consider an auxiliary random number X', taking values in $\mathcal{X}' \cap \mathbb{R}$ (we assume $\mathcal{X}' \cap \mathbb{R} \neq \emptyset$ to avoid trivialities) and such that the bivariate p-box $(\underline{F}', \overline{F}')$ of the couple (X', Y) coincides with $(\underline{F}, \overline{F})$ on $\mathcal{X}' \times \mathcal{Y}$. Then $\underline{F}', \overline{F}'$ satisfy conditions (I-RI1)÷(I-RI4), because they coincide with $\underline{F}, \overline{F}$ on $\mathcal{X}' \times \mathcal{Y}$ and by Remark 3. By Proposition 11, $(\underline{F}', \overline{F}')$ is coherent, i.e. the lower probability $\underline{P}_{(\underline{F}', \overline{F}')}$ associated with $(\underline{F}', \overline{F}')$ by (11) is coherent. But $\underline{P}_{(\underline{F}', \overline{F}')} = \underline{P}_{(\underline{F}, \overline{F})}$ on the events $E_0, E_1, \ldots, E_n$. Because of this (30) may be interpreted as a coherence condition for $\underline{P}_{(\underline{F}', \overline{F}')}$ too, and as such it holds. □

In particular, Proposition 11 also implies that conditions (I-RI1)÷(I-RI4) are equivalent to the coherence of $(\underline{F}, \overline{F})$ when the bivariate p-box describes a couple of binary, logically independent random numbers, like for instance indicators of events. With similar techniques it can be shown that the four conditions together are equivalent to coherence when $\mathcal{X} = \{x_1, x_2, x_3\}$ and $\mathcal{Y} = \{y_1, y_2, y_3\}$; the lengthy proof is somewhat similar to that of Proposition 11.

Interestingly, we deduce from the proof of Proposition 11 that, under the conditions of the result, each of (I-RI1)÷(I-RI4) is sufficient for the bivariate p-box $(\underline{F}, \overline{F})$ to avoid sure loss. This follows applying Proposition 6 since under each of the four conditions we have established the existence of a distribution function F bounded between $\underline{F}$ and $\overline{F}$. To see that this result does not hold in general, not even under logical independence of X and Y, note that the bivariate p-box in Example 2 satisfies (I-RI1) although it does not avoid sure loss. It is easy to show with appropriate examples that the other conditions are not sufficient either.

The further consistency properties of bivariate p-boxes we are about to discuss concern also the restrictions of $\underline{P}_{(\underline{F}, \overline{F})}$ to $\mathcal{D}$ and $\mathcal{D}_c$, denoted $\underline{P}_{\underline{F}}$ and $\underline{P}_{\overline{F}}$. In terms of a p-box $(\underline{F}, \overline{F})$ they are defined, for all x, y, by:

$$\underline{P}_{\underline{F}}(A_{(x,y)}) = \underline{F}(x, y), \quad \underline{P}_{\overline{F}}(A^c_{(x,y)}) = 1 - \overline{F}(x, y). \tag{31}$$

We shall also consider two extreme bivariate p-boxes, where the information supplied by either $\underline{F}$ or $\overline{F}$ is *vacuous*.

Definition 11. *(Vacuous p-boxes)* Define $\underline{F}_0, \overline{F}_1$ by

$$\underline{F}_0(x, y) = \begin{cases} 0 & \text{if } A_{(x,y)} \neq \Omega, \\ 1 & \text{if } A_{(x,y)} = \Omega \end{cases} \text{ and } \overline{F}_1(x, y) = \begin{cases} 1 & \text{if } A_{(x,y)} \neq \varnothing \\ 0 & \text{if } A_{(x,y)} = \varnothing. \end{cases}$$

We say that a bivariate p-box $(\underline{F}, \overline{F})$ is *lower vacuous* if $\underline{F} = \underline{F}_0$, and *upper vacuous* if $\overline{F} = \overline{F}_1$.

Note that by (31) $\underline{F}_0$, $\overline{F}_1$ correspond to the vacuous lower probabilities on $\mathcal{D}$ and $\mathcal{D}_c$, which are well known to be coherent (see Walley [3], Section 2.9.1). It is easily seen that $\underline{F}_0$ is also a distribution function, unlike $\overline{F}_1$, which may not be a cdf in certain special cases, like the following one:

Example 5. Let $\mathcal{X} = \mathcal{Y} = \{1, 2\}$, and assume that X and Y cannot take simultaneously the value 1. Then $\overline{F}_1$ is not a distribution function: the rectangle inequality (RI) gives $\overline{F}_1(2,2) + \overline{F}_1(1,1) - \overline{F}_1(1,2) - \overline{F}_1(2,1) = -1 < 0$. ♦

However, $\overline{F}_1$ is a cdf in several common instances, in particular when it refers to two logically independent random numbers.

The next proposition regards the consistency of the previously defined concepts.

Proposition 12. *Let $(\underline{F}, \overline{F})$ be the bivariate p-box describing two random numbers X, Y and $\underline{P}_{\underline{F}}$, $\underline{P}_{\overline{F}}$ the lower probabilities (on $\mathcal{D}$, $\mathcal{D}_c$, respectively) given by* (31)*. Then*

- *(a) $\underline{P}_{\overline{F}}$ avoids sure loss.*
- *(b) $\underline{P}_{\overline{F}}$ is coherent iff $\underline{P}_{(\underline{F}_0, \overline{F})}$ is coherent.*

If in addition X and Y are logically independent, then

- *(c) $\underline{P}_{\underline{F}}$ avoids sure loss.*
- *(d) Any lower vacuous p-box $(\underline{F}_0, \overline{F})$ is coherent.*

Proof.

- (a) It suffices to take into account that $\underline{F}_0$ is a distribution function that is dominated by $\overline{F}$.
- (b) $\underline{P}_{\overline{F}}$ is coherent if and only if for every (x, y) there is a distribution function $F \leq \overline{F}$ such that $F(x, y) = \overline{F}(x, y)$. The condition $F \leq \overline{F}$ is equivalent to $\underline{F}_0 \leq F \leq \overline{F}$, and on the other hand $\underline{F}_0$ is trivially a distribution function. We deduce from Proposition 9 that $\underline{P}_{(\underline{F}_0, \overline{F})}$ is coherent if and only if $\overline{F}$ is the upper envelope of the distribution functions in $(\underline{F}_0, \overline{F})$, and as a consequence we have the equivalence in (b).
- (c) Under logical independence, $\underline{F}$ is dominated by the distribution function $\overline{F}_1$.
- (d) Consider $x, y \in \overline{\mathbb{R}}$, and let us prove the existence of a distribution function F such that $\underline{F}_0 \leq F \leq \overline{F}$ and that $F(x, y) = \overline{F}(x, y)$. Let $F : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0, 1]$ be defined by

$$F(x', y') = \begin{cases} \overline{F}(x, y) & \text{if } x' \geq x, y' \geq y, (x', y') \neq (+\infty, +\infty) \\ 1 & \text{if } (x', y') = (+\infty, +\infty) \\ 0 & \text{otherwise.} \end{cases}$$

Then $F \geq \underline{F}_0$ trivially, and since $\overline{F}$ is standardized, $F \leq \overline{F}$. By construction F is monotone non-decreasing and $F(x, y) = \overline{F}(x, y)$. By Proposition 2 and Remark 4, only the rectangle inequality (RI) still has to be checked to state that F is a distribution function. Consider for this $x_1 < x_2, y_1 < y_2 \in \overline{\mathbb{R}}$. There are three possibilities:

– If $x_1 < x$, then

$$F(x_2, y_2)+F(x_1, y_1)-F(x_1, y_2)-F(x_2, y_1) = F(x_2, y_2)-F(x_2, y_1) \geq 0.$$

– Similarly, if $y_1 < y$,

$$F(x_2, y_2)+F(x_1, y_1)-F(x_1, y_2)-F(x_2, y_1) = F(x_2, y_2)-F(x_1, y_2) \geq 0.$$

– Finally, if $x_1 \geq x$ and $y_1 \geq y$,

$$F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1) \\ = F(x_2, y_2) + F(x, y) - F(x, y) - F(x, y) = F(x_2, y_2) - F(x, y) \geq 0.$$

Thus, F satisfies (RI) and therefore it is a distribution function. □

From Proposition 12, we see that $\underline{P}_{\underline{F}}$ is not guaranteed to always avoid sure loss, unlike $\underline{P}_{\overline{F}}$. There is an asymmetry also between lower vacuous bivariate p-boxes, always coherent under logical independence, and upper vacuous bivariate p-boxes, which may not be coherent even when X and Y are logically independent. To see this, let $\underline{F}$ be determined by the values

	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
$\underline{F}$	0	0.6	0.6	1

Then $(\underline{F}, \overline{F}_1)$ is not coherent, because it does not satisfy the necessary condition (I-RI2): $\overline{F}_1(2,2) + \underline{F}(1,1) - \underline{F}(1,2) - \underline{F}(2,1) = -0.2 < 0$.

Although Proposition 12 makes clear that there are some relationships between bivariate p-boxes or their corresponding lower probabilities, on one hand, and the restrictions of these lower probabilities on $\mathcal{D}$ or $\mathcal{D}_c$, on the other, these connections are not as tight as in the univariate case. In fact, Theorem 2 does not quite extend to the bivariate case. Indeed, our results imply that, given a bivariate p-box $(\underline{F}, \overline{F})$,

$$\underline{F}, \overline{F} \text{ distribution functions } \Rightarrow \underline{P}_{(\underline{F},\overline{F})} \text{ coherent } \Rightarrow \underline{P}_{\underline{F}}, \underline{P}_{\overline{F}} \text{ coherent;}$$

the second implication holds trivially, because restrictions (here $\underline{P}_{\underline{F}}$, $\underline{P}_{\overline{F}}$) of a coherent lower probability are coherent themselves. However, the converses of these two implications do not hold in general: for the first one, it suffices to recall Example 1, where the envelopes of a set of distribution functions are not distribution functions themselves; for the second, use Example 4: there $\underline{F}$ is a distribution function and we have logical independence, so both $\underline{P}_{\underline{F}}, \underline{P}_{\overline{F}}$ are coherent, by Proposition 12. However, we showed in the example that the bivariate p-box $(\underline{F}, \overline{F})$ is not coherent.

It is also interesting to remark that in the univariate case vacuous p-boxes are quite related to maxitive and possibility measures, as established by Troffaes et al. [6] In particular, Corollary 3.3 there establishes that the upper probability induced by a univariate p-box is maxitive if and only if either $\underline{F}$ or $\overline{F}$ is 0–1-valued. Our results show that such an equivalence does not hold in the bivariate case.

3.5. *Bivariate p-boxes and 2-monotonicity*

In this section we explore the relationships between bivariate p-boxes and 2-monotonicity. Recall that an uncertainty measure μ defined on a *lattice* of events $\mathcal{C}$ is 2-monotone[11] iff

$$\mu(A \vee B) \geq \mu(A) + \mu(B) - \mu(A \wedge B) \;\; \forall A, B \in \mathcal{C},$$

whilst μ is termed 2-*alternating* if the previous inequality is reversed.

The lower probability $\underline{P}_{(\underline{F},\overline{F})}$ induced by a bivariate p-box is defined on $\mathcal{E}$, which is not a lattice. Hence, in order to discuss 2-monotonicity of $\underline{P}_{(\underline{F},\overline{F})}$ we must refer to an extension of $\underline{P}_{(\underline{F},\overline{F})}$ to some larger domain. In this section we shall consider the *natural extension* $\underline{E}$ of $\underline{P}_{(\underline{F},\overline{F})}$ to the algebra $\mathcal{Q}$ generated by $\mathcal{E}$. Since we have established in Proposition 9 a correspondence between the sets $\mathcal{M}(\underline{P}_{(\underline{F},\overline{F})})$ and $\mathcal{F}$ given by Eqs. (7) and (15), we can apply Theorem 3.4.1 in Walley[3] to conclude that

$$\underline{E} := \min\{P_F^* : F \text{ distribution function}, \underline{F} \leq F \leq \overline{F}\}$$

where P_F^* denotes the unique extension from $\mathcal{E}$ to $\mathcal{Q}$ of the dF-coherent probability P_F defined by Eq. (13). Note that this extension is indeed unique by Thm. 11.2.2 in Crisma[10] (cf. also Denneberg[11], Troffaes and Destercke[5] for similar results for the univariate case), but this uniqueness does not hold beyond the algebra $\mathcal{Q}$ (see Miranda et al.[13], Note 4).

In the univariate case p-boxes are tightly connected to 2-monotonicity: the natural extension of $\underline{P}_{(\underline{F},\overline{F})}$ on $\mathcal{Q}$ is completely monotone, and hence also 2-monotone (see Troffaes and Destercke[5], Thm. 17). Actually, as we know from Theorem 2, $P_{\underline{F}}$, $P_{\overline{F}}$ are even dF-coherent probabilities and $\underline{F}$, $\overline{F}$ distribution functions. This property is not necessarily ensured in higher dimensions, as we have seen.

In the bivariate case the natural extension of the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ associated with the p-box $(\underline{F}, \overline{F})$ may not be 2-monotone, even if both $\underline{F}, \overline{F}$ are distribution functions:

Example 6. Let $\underline{F}, \overline{F}$ be the standardized maps for X, Y given by:

	$(1,1)$	$(1,2)$	$(2,1)$	$(2,2)$
$\underline{F}$	0	0	0.5	1
$\overline{F}$	0.25	0.25	0.5	1

To see that both $\underline{F}$ and $\overline{F}$ are distribution functions, note that

$$\underline{F}(2,2) + \underline{F}(1,1) - \underline{F}(1,2) - \underline{F}(2,1) = 0.5 > 0;$$
$$\overline{F}(2,2) + \overline{F}(1,1) - \overline{F}(1,2) - \overline{F}(2,1) = 0.5 > 0,$$

and that all the other possible comparisons are trivial. As a consequence, the lower probability $\underline{P}_{(\underline{F},\overline{F})}$ they induce on $\mathcal{E}$ by Eq. (11) is coherent, and from Theorem 3.1.2 in Walley[3], so is its natural extension $\underline{E}$ to $\mathcal{Q}$. Moreover, $\underline{E}(C) = \underline{P}_{(\underline{F},\overline{F})}(C)$ for every $C \in \mathcal{E}$.

Let us show that $\underline{E}$ is not 2-monotone. Denote $a = (X = 1 \wedge Y = 1)$, $b = (X = 1 \wedge Y = 2)$, $c = (X = 2 \wedge Y = 1)$, $d = (X = 2 \wedge Y = 2)$ and take $A = (Y = 1) = a \vee c$ and $B = (X = 2) = c \vee d$. Then, using that $A, B, A \wedge B \in \mathcal{E}$,

- $\underline{E}(A) = \underline{E}(a \vee c) = \underline{P}_{(\underline{F},\overline{F})}(a \vee c) = \underline{F}(2,1) = 0.5$.
- $\underline{E}(B) = \underline{E}(c \vee d) = \underline{P}_{(\underline{F},\overline{F})}(c \vee d)$, which by conjugacy is equal to $1 - \overline{P}_{(\underline{F},\overline{F})}(a \vee b) = 1 - \overline{F}(1,2) = 0.75$.

On the other hand,

- $\underline{E}(A \wedge B) = \underline{E}(c) \leq P_{\overline{F}}(c) = 0.25$, where $P_{\overline{F}}$ is the probability induced by the distribution function $\overline{F}$.
- $\underline{E}(A \vee B) = \underline{E}(a \vee c \vee d) \leq P_F(a \vee c \vee d) = 0.75$, where $F \in \mathcal{F}$ is the distribution function given by $F(1,1) = 0, F(1,2) = 0.25, F(2,1) = 0.5, F(2,2) = 1$ and P_F is the probability it induces by means of Eq. (11).

As a consequence, we conclude that

$$\underline{E}(A \vee B) + \underline{E}(A \wedge B) \leq 1 < 1.25 = \underline{E}(A) + \underline{E}(B),$$

whence the lower probability induced by $(\underline{F}, \overline{F})$ is not 2-monotone. ♦

Interestingly, in this example the natural extension $\underline{E}$ does not coincide with the lower envelope of $\{P_{\underline{F}}, P_{\overline{F}}\}$: these are associated with the mass functions $P_{\underline{F}} = (0, 0, 0.5, 0.5)$ and $P_{\overline{F}} = (0.25, 0, 0.25, 0.5)$, on $\{a, b, c, d\}$, whence

$$\min\{P_{\underline{F}}(A \vee B), P_{\overline{F}}(A \vee B)\} = 1 > 0.75 = \underline{E}(A \vee B).$$

This means that even if the bivariate p-box is determined by the distribution functions $\underline{F}, \overline{F}$, the natural extension of its associated lower probability cannot always be computed as the lower envelope of the probabilities $P_{\underline{F}}$, $P_{\overline{F}}$ corresponding to $\underline{F}$, $\overline{F}$.

It is also interesting to proceed in the converse manner: instead of investigating whether assessing a lower probability $\underline{P}_{(\underline{F},\overline{F})}$ on $\mathcal{E}$, or equivalently a bivariate p-box $(\underline{F}, \overline{F})$, induces 2-monotonicity properties in a larger environment $\mathcal{Q}$, we can consider the effects of 2-monotonicity of a lower probability $\underline{P}$ defined in some $\mathcal{Q} \supset \mathcal{E}$ on its associated $(\underline{F}, \overline{F})$. The following proposition answers this problem.

Proposition 13. *Let $\mathcal{Q}$ be an algebra of events, $\mathcal{Q} \supset \mathcal{E}$. Then,*

a) if $\underline{P}$ is a 2-monotone lower probability on $\mathcal{Q}$, its restriction to $\mathcal{D}$ is a dF-coherent probability;
b) if $\overline{P}$ is a 2-alternating upper probability on $\mathcal{Q}$, its restriction to $\mathcal{D}_c$ is a dF-coherent probability.

Proof. The proof of a) follows applying Corollary 7 in Scarsini [15] to our framework. As for the proof of b), let now $\underline{P}$ be the lower probability conjugate of $\overline{P}$, i.e. $\underline{P}(E) = 1 - \overline{P}(E^c)$, $\forall E \in \mathcal{Q}$. Since $\overline{P}$ is 2-alternating, $\underline{P}$ is 2-monotone by Prop. 2.3,

(iii) in Denneberg [11]. Hence, the restriction of $\underline{P}$ to $\mathcal{D}$ is a dF-coherent probability by a). Since $\overline{P}$ is its unique dF-coherent extension to $\mathcal{D}_c$, the thesis follows. □

Remark 9. When the lower (upper) probability in Proposition 13 is 2-monotone (2-alternating), its restriction to $\mathcal{D}_c$ ($\mathcal{D}$) is not necessarily a dF-coherent probability. We prove this fact in the case of a 2-alternating upper probability with a counterexample. Consider two random numbers X, Y, with $\mathcal{X} = \mathcal{Y} = \{1, 2, 3\}$. Define first a dF-coherent P on $(X = i \wedge Y = j)$, $i, j = 1, 2, 3$, and the related $F(i, j) = P(X \leq i \wedge Y \leq j)$:

	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$	$(3,1)$	$(3,2)$	$(3,3)$
P	0.1	0	0.15	0.2	0.2	0.05	0.15	0.1	0.05
F	0.1	0.1	0.25	0.3	0.5	0.7	0.45	0.75	1

Clearly, P is defined by additivity also on the powerset $\mathcal{Q}$ of $\{X = i \wedge Y = j : i, j = 1, 2, 3\}$. Define now $\overline{P}(A) \coloneqq \min\{1.25 \cdot P(A), 1\}$, $\forall A \in \mathcal{Q}$. $\overline{P}$ is an instance of *pari-mutuel* upper probability, which is well-known to be coherent and 2-alternating (see Walley [3], Sec. 2.9.3; Pelessoni et al. [16]).

The restriction of $\overline{P}$ on $\mathcal{D}$ is not a dF-coherent probability: if it were so, $\overline{F}(i, j) = \overline{P}(X \leq i \wedge Y \leq j) = \min\{1.25 \cdot F(i, j), 1\}$ would be a cdf, which is not. In fact $\overline{F}$ does not satisfy the rectangle inequality (RI). Take for instance $(2, 2)$ and $(3, 3)$: $\overline{F}(3, 3) + \overline{F}(2, 2) - \overline{F}(2, 3) - \overline{F}(3, 2) = 1 + 0.625 - 0.875 - 0.9375 < 0$. ♦

We are now in a position to comment a further difference between univariate and bivariate p-boxes, regarding the way they may be obtained from restrictions of functions defined on larger domains.

In the univariate case, let $\underline{P}$ be a capacity on $\mathcal{A} \supset \mathcal{E}_0 = \{A_x, A_x^c : x \in \overline{\mathbb{R}}\}$ such that $\underline{P}(A_x) + \underline{P}(A_x^c) \leq 1$ $\forall x \in \overline{\mathbb{R}}$. We get a couple of distribution functions by applying Proposition 1 to $\underline{F}(x) = \underline{P}(A_x)$, $\overline{F}(x) = 1 - \underline{P}(A_x^c)$, while condition $\underline{P}(A_x) + \underline{P}(A_x^c) \leq 1$ $\forall x \in \overline{\mathbb{R}}$ ensures that $\underline{F} \leq \overline{F}$. Hence, $(\underline{F}, \overline{F})$ is a p-box and the restriction of $\underline{P}$ to $\mathcal{E}_0$ is a coherent lower probability by Theorem 2.

In the bivariate case, given a 2-coherent lower probability $\underline{P}$ on $\mathcal{Q} \supset \mathcal{E}$, its restriction on $\mathcal{E}$ corresponds to a p-box $(\underline{F}_{\underline{P}}, \overline{F}_{\underline{P}})$, using Eq. (12) and Proposition 5, a). It is coherent if $\underline{P}$ is. However, even the stronger condition that $\underline{P}$ is 2-monotone and $\mathcal{Q}$ is an algebra does not ensure that $\overline{F}_{\underline{P}}$ is a cdf, but only that $\underline{F}_{\underline{P}}$ is. For this, notice that $\overline{P}(A_{(x,y)}) = \overline{F}_{\underline{P}}(x, y)$ is 2-alternating and apply Proposition 13 and Remark 9. Alternatively, a p-box can be obtained from the restriction to $\mathcal{E}$ of a 2-alternating upper probability $\overline{P}$, putting $\overline{F}_{\overline{P}}(x, y) \coloneqq \overline{P}(A_{(x,y)})$, $\underline{F}_{\overline{P}}(x, y) \coloneqq 1 - \overline{P}(A^c_{(x,y)})$. In this way, $\overline{F}_{\overline{P}}$ is a cdf, $\underline{F}_{\overline{P}}$ not necessarily.

4. Conclusions

In this paper we have explored the extension of p-boxes to the bivariate environment, as a way to describe couples of random numbers in the presence of imprecision in their uncertainty evaluation.

We have defined bivariate p-boxes $(\underline{F}, \overline{F})$ by requiring some minimal conditions to $\underline{F}$, $\overline{F}$, and have shown that this only guarantees a poor degree of consistency, which may be equivalent to (or even weaker than) the still weak consistency notion of 2-coherence of a lower probability. Because of this, we have focused on the more restrictive notion of coherent bivariate p-boxes, for which we have established their main properties. Although we exploited the correspondence with a coherent lower probability in most of the derivations, we should remark that a p-box $(\underline{F}, \overline{F})$ corresponds also to a pair $(\underline{P}, \overline{P})$ of lower and upper coherent probabilities. This can be easily seen from Eq. (11) and the conjugacy relation $\underline{P}(A) = 1 - \overline{P}(A^c)$, and holds for univariate p-boxes as well. One difference here is that in the univariate case $\underline{P}$ and $\overline{P}$ are also precise dF-coherent probabilities (cf. the comment after Proposition 1 and Theorem 2, (c)). Hence, the notion of coherent bivariate p-box highlights a consistency property which trivially holds with univariate p-boxes.

It is clear that adding a second dimension and/or allowing imprecise rather than only precise assessments considerably raises the complexity of the properties of coherent p-boxes. This appears in particular from the variety of results involving the rectangle inequalities, and from the relationships with 2-monotonicity. It is unclear at present whether a general characterization theorem for coherent bivariate p-boxes can be found. In the paper, we prove one such theorem under the assumption of logical independence, and when at least one of the random numbers is binary. In this case the four imprecise rectangle inequalities, together with some mild conditions, are sufficient for coherence. It is however patent that generally they are not. Just think of the special case $\underline{F} = \overline{F}$: then these inequalities are all equal to the rectangle inequality (RI) of Proposition 2, which is generally not sufficient for dF-coherence, in this case equivalent to coherence. It seems also unlikely that an additional condition similar to Eq. (3) may be found, given that the interpretation of $\Delta\underline{F}$ or $\Delta\overline{F}$ is not alike to that of ΔF (cf. the discussion following Proposition 10). This question remains an open problem at this stage. Other topics for future work, besides those investigated in Montes et al. [8], are the extension of the concepts and properties investigated here to the n-dimensional environment, and the use of bivariate p-boxes in applications.

Acknowledgements

We acknowledge financial support by the FRA2013 grant ‘Models for Risk Evaluation, Uncertainty Measurement and Non-Life Insurance Applications’ (R. Pelessoni and P. Vicig) and by the project TIN2014-59543-P (I. Montes and E. Miranda).

References

1. T. Augustin, F. Coolen, G. de Cooman, and M. Troffaes, editors. *Introduction to Imprecise Probabilities*. Wiley, 2014.
2. S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Con-

structing probability boxes and Dempster-Shafer structures. Technical Report SAND2002–4015, Sandia National Laboratories, 2003.

3. P. Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London, 1991.
4. S. Ferson and W. Tucker. Sensitivity analysis using probability bounding. *Reliability Engineering and System Safety*, 91(10–1):1435–1442, 2006.
5. M. C. M. Troffaes and S. Destercke. Probability boxes on totally preordered spaces for multivariate modelling. *International Journal of Approximate Reasoning*, 52(6):767–791, 2011.
6. M. C. M. Troffaes, E. Miranda, and S. Destercke. On the connection between probability boxes and possibility measures. *Information Sciences*, 224:88–108, 2013.
7. L. Utkin and S. Destercke. Computing expectations with continuous p-boxes: univariate case. *International Journal of Approximate Reasoning*, 50(5):778 – 798, 2009.
8. I. Montes, E. Miranda, R. Pelessoni, and P. Vicig. Sklar's theorem in an imprecise setting. *Fuzzy Sets and Systems*, 2015. doi: 10.1016/j.fss.2014.10.007. In press.
9. B. de Finetti. *Theory of Probability: A Critical Introductory Treatment.* John Wiley & Sons, Chichester, 1974–1975. English translation of de Finetti,[17] two volumes.
10. L. Crisma. *Introduzione alla teoria delle probabilità coerenti.* Edizioni Università di Trieste, Trieste, 2006.
11. D. Denneberg. *Non-Additive Measure and Integral.* Kluwer Academic, Dordrecht, 1994.
12. P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Reprinted in Williams[18].
13. E. Miranda, G. de Cooman, and E. Quaeghebeur. Finitely additive extensions of distribution functions and moment sequences: The coherent lower prevision approach. *International Journal of Approximate Reasoning*, 48(1):132–155, 2008.
14. M.C.M. Troffaes and G. de Cooman. *Lower previsions.* Wiley Series in Probability and Statistics. Wiley, 2014.
15. M. Scarsini. Copulae of capacities on product spaces. In L. Rüschendorf, B. Schweizer, and M.D. Taylor, editors, *Distributions with Fixed Marginals and Related Topics*, volume 28 of *IMS Lecture Notes-Monograph Series*, pages 307–318. Institute of Mathematical Statistics, 1996.
16. R. Pelessoni, P. Vicig, and M. Zaffalon. Inference and risk measurement with the pari-mutuel model. *International Journal of Approximate Reasoning*, 51 (9):1145–1158, 2010.
17. B. de Finetti. *Teoria delle Probabilità.* Einaudi, Turin, 1970.
18. P. M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of Williams[12].