International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems

A CONSONANT APPROXIMATION OF THE PRODUCT OF INDEPENDENT CONSONANT RANDOM SETS

SEBASTIEN DESTERCKE

INRA/CIRAD, UMR1208, 2 place P. Viala, F-34060 Montpellier cedex 1, FRANCE.
sdestercke@gmail.com

DIDIER DUBOIS

Université Paul Sabatier, IRIT/RPDMP, 118 Route de Narbonne, 31062 Toulouse.
dubois@irit.fr

ERIC CHOJNACKI

Institut de Radioprotection et Sûreté nucléaire, Bât 720, 13115 St-Paul lez Durance, FRANCE.
eric.chojnacki@irsn.fr

Received 15 December 2008
Revised 5 August 2009

The belief structure resulting from the combination of consonant and independent marginal random sets is not, in general, consonant. Also, the complexity of such a structure grows exponentially with the number of combined random sets, making it quickly intractable for computations. In this paper, we propose a simple guaranteed consonant outer approximation of this structure. The complexity of this outer approximation does not increase with the number of marginal random sets (i.e., of dimensions), making it easier to handle in uncertainty propagation. Features and advantages of this outer approximation are then discussed, with the help of some illustrative examples.

Keywords: Belief functions, Possibility theory, approximation, independence

1. Introduction

We consider the problem of modeling uncertainty concerning the values that several variables $X_1,\dots,X_N$ can respectively assume on domains $\mathscr{X}_1,\dots,\mathscr{X}_N$ (finite sets, intervals). For a long time, such a task has been handled by the sole means of probability theory. However, many arguments converge to the conclusion that probability distributions alone cannot faithfully model the incompleteness, scarcity or unreliability of information. [1] In this case, other theories explicitly modeling these issues can be advocated. In this paper, we mainly consider two such theories: possibility theory [2] and random set theory. [3]

In practical applications, uncertainty is seldom modeled or elicited directly over the whole Cartesian product $\times_{i=1}^{N}\mathscr{X}_i$. A more common practice is to build or elicit marginal models for each variable $X_1,\dots,X_N$ and then to combine them by taking into account possible dependencies between them, this last step being easier under an independence assumption. However, as the number N of variables increases, the structural complexity resulting from this combination often increases exponentially, making it uneasy to handle computationally. In such cases, simple outer-approximating models are easier to handle when propagating uncertainty and they can guarantee conservative results (i.e., they do not consider more information than available).

In this paper, we consider the case where the marginal uncertainty on each variable $X_1, \ldots, X_N$ is modeled by a consonant random set (i.e., a possibility distribution) and that these random sets can be combined into a joint uncertainty model by assuming random set independence. Since manipulating such a joint structure can be difficult in practice, we provide a joint outer-approximating possibility distribution that can be built by a simple transformation of each marginal random set. This result extends to any number of dimensions a result already given by Dubois and Prade [4] for the 2-dimensional case ($N = 2$). The features and potential advantages of this outer-approximation are then discussed and compared with other methods by means of illustrative examples.

Although the situation considered here (consonant random sets with random set independence) can be viewed as somewhat restrictive, it is likely to occur in many practical situations. First, there are many cases where possibility distributions will adequately model available information: experts expressing their opinion by lower confidence bounds over nested intervals; [5] nested statistical prediction intervals; [6,7] partial probabilistic information; [8] consonant approximation of multinomial sampling. [9,10] Second, random set independence can be interpreted and used in various ways: for example, it can correspond to independence between information sources, [11] or be used as a conservative (but mathematically convenient, as it can be simulated by sampling methods) modeling of stochastic independence between variables whose true probabilities are ill-known. [12,13] .

The paper is organized as follows: basics about possibility theory, random sets and (in)dependence notions in these theories are recalled in Section 2. The possibilistic outer approximation is then introduced and discussed in Section 3. Potential advantages of such an outer approximation when treating information are then illustrated on simple examples in Section 4.

2. Preliminaries

This section provides basics about possibility distributions and random sets needed in the sequel. Also recall that our aim is to describe the joint uncertainty over variables $X_1, \ldots, X_N$ assuming values on some domains $\mathscr{X}_1, \ldots, \mathscr{X}_N$. As we will often work with Cartesian product of spaces, we adopt the following notation: given two values k, ℓ such that $1 \leq k \leq \ell \leq N$, we denote by $\mathscr{X}_{(k:\ell)} := \times_{i=k}^{\ell} \mathscr{X}_i$ the Cartesian product of the $\ell - k + 1$ domains $\mathscr{X}_k, \ldots, \mathscr{X}_\ell$. Similarly, we denote by $X_{(k:\ell)} := (X_k, \ldots, X_\ell)$ a variable assuming values on $\mathscr{X}_{(k:\ell)}$, and $x_{(k:\ell)} := (x_k, \ldots, x_\ell) \in \mathscr{X}_{(k:\ell)}$ a specific element of $\mathscr{X}_{(k:\ell)}$.

2.1. *Random sets*

A discrete normal random set, here denoted by $(m, \mathscr{F})$, over a domain $\mathscr{X}$ is defined as a mapping $m : \wp(\mathscr{X}) \to [0, 1]$ from the power set $\wp(\mathscr{X})$ of $\mathscr{X}$ to the unit interval, with $\sum_{E \subseteq \mathscr{X}} m(E) = 1$ and $m(\emptyset) = 0$. We call m a mass assignment, and a set E that receives strictly positive mass a focal set. The mass $m(E)$ can be interpreted as the probability that the most precise description of what is known about a particular situation is of the form "$x \in E$". Weights $m(E)$ should be shared between elements of E but are not by lack of information. From this mass assignment, Shafer [14] defines two set functions, called *belief and plausibility functions*, for any event $A \subseteq \mathscr{X}$:

$$Bel(A) = \sum_{E, E \subseteq A} m(E); \qquad Pl(A) = 1 - Bel(A^c) = \sum_{E, E \cap A \neq \emptyset} m(E),$$

where the belief function measures the certainty of A (i.e., it sums all masses that cannot be distributed outside A) and the plausibility function measures the plausibility of A (i.e., sums all masses that it is possible to distribute inside A). In this view, sets E are called disjunctive in the sense that they are made of mutually exclusive elements. They represent incomplete information inducing uncertainty [a]. Note that the two functions Bel, Pl are conjugate, in the sense that specifying one of them for all events is enough to characterize the other. Shafer also defines another set-function, the commonality function, which reads, for any event $A \subseteq \mathscr{X}$,

$$q(A) = \sum_{E, A \subseteq E} m(E).$$

This function sums all the masses that could go to any element of A. Since the greater the mass given to larger sets, the higher the values of the commonality function, it can be argued that this function reflects the imprecision of information.

The two functions Bel, Pl can also be interpreted as lower and upper probabilistic bounds describing an imprecise state of knowledge. In this latter case, a random set $(m, \mathscr{F})$ induces a convex set $\mathscr{P}_{(m,\mathscr{F})}$ of probability distributions such that

$$\mathscr{P}_{(m,\mathscr{F})} = \{P \in \mathbb{P}_{\mathscr{X}} | \forall A \subseteq \mathscr{X}, Bel(A) \leq P(A)\},$$

with $\mathbb{P}_{\mathscr{X}}$ the set of all probability distributions on $\mathscr{X}$. This view is closer to the one adopted by Dempster, [15] while Shafer (like Smets [16] later on) does not refer to any underlying standard probabilistic framework.

2.2. *Possibility distributions*

Possibility distributions are the primary mathematical tools of possibility theory. A possibility distribution is a mapping $\pi : \mathscr{X} \to [0,1]$ from a space $\mathscr{X}$ to the unit interval such that $\pi(x) = 1$ for at least one element x in $\mathscr{X}$.

As for random sets, several set functions can be defined from a possibility distribution, among which are the *possibility and necessity* functions [17]:

$$\Pi(A) = \sup_{x \in A} \pi(x); \qquad N(A) = 1 - \Pi(A^c) = \inf_{x \in A^c} (1 - \pi(x)).$$

[a] This is in contrast with other uses of random sets.

Possibility and necessity functions respectively measure the plausibility and certainty of event A. Their characteristic properties are: $N(A \cap B) = \min(N(A), N(B))$ and $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$ for any pair of events A, B of $\mathscr{X}$.

Given a degree $\alpha \in [0,1]$ the strong ($A_{\overline{\alpha}}$) and regular (A_α) α-cuts of a distribution π are subsets respectively defined as

$$A_{\overline{\alpha}} = \{x \in \mathscr{X} | \pi(x) > \alpha\}, \tag{1}$$

$$A_\alpha = \{x \in \mathscr{X} | \pi(x) \geq \alpha\}. \tag{2}$$

These α-cuts are nested, since if $\alpha > \beta$, then $A_\alpha \subseteq A_\beta$. When possibility distributions are discrete, the set of values $\{\pi(x) | x \in \mathscr{X}\}$ is of the form $1 = \alpha_1 > \ldots > \alpha_M > \alpha_{M+1} = 0$, meaning that in this case there are only M distinct α-cuts.

It can be shown that any necessity (resp. possibility) function is a special kind of belief (resp. plausibility) function, [14] whose associated random set has nested focal sets. In this case, the random set is commonly called consonant. Thus, any possibility distribution π, defines a random set $(m_\pi, \mathscr{F}_\pi)$ having, for $i = 1, \ldots, M$, the following focal sets E_i with masses $m(E_i)$: [18]

$$\begin{cases} E_i = \{x \in \mathscr{X} | \pi(x) \geq \alpha_i\} = A_{\alpha_i}, \\ \qquad m(E_i) = \alpha_i - \alpha_{i+1}. \end{cases} \tag{3}$$

Conversely, any random set with nested focal sets can be modeled by a unique possibility distribution in general [b].

Again, necessity and possibility measures of a distribution π can be seen as lower and upper probabilistic bounds, and can be associated to the convex set $\mathscr{P}_\pi$ of probabilities such that

$$\mathscr{P}_\pi = \{P \in \mathbb{P}_{\mathscr{X}} | \forall A \subseteq \mathscr{X}, N(A) \leq P(A)\}. \tag{4}$$

2.3. *Specificity in possibility and random set theory*

The issue of comparing the informative power (or specificity) of random set representations of incomplete information relies on extending the notion of set inclusion. In the case of possibility distributions, fuzzy set inclusion is instrumental.

Definition 1. Let π_1, π_2 be two possibility distributions. π_1 is then said to be included in π_2 if and only if $\pi_1 \leq \pi_2$, and we denote this inclusion by $\pi_1 \sqsubseteq_\pi \pi_2$.

Many different notions extending classical set-inclusion to random sets can be found in the literature: the notions of pl-,q- and s-inclusions are the older ones [c], [20] while Denoeux

[b]The link between nested random sets and possibility measures is less straightforward in more abstract infinite mathematical settings [19].

[c]Notions pl- and s-inclusions are the most commonly used, and are often respectively called weak and strong inclusion between random sets.

recently introduced yet other notions (w- and v-inclusions) based on Smets canonical decomposition of belief functions. [21] [22] Each of these notions induces a different partial order on the set of all random sets.

Definition 2. Let $(m_1,\mathscr{F}_1)$, $(m_2,\mathscr{F}_2)$ be two random sets. $(m_1,\mathscr{F}_1)$ is then said to be *pl*-included in $(m_2,\mathscr{F}_2)$ if and only if, for all $A \subseteq \mathscr{X}$,

$$Pl_1(A) \leq Pl_2(A),$$

and we denote this inclusion by $(m_1,\mathscr{F}_1) \sqsubseteq_{pl} (m_2,\mathscr{F}_2)$.

Note that $(m_1,\mathscr{F}_1)$ is *pl*-included in $(m_2,\mathscr{F}_2)$ if and only if $\mathscr{P}_{(m_1,\mathscr{F}_1)} \subseteq \mathscr{P}_{(m_2,\mathscr{F}_2)}$

Definition 3. Let $(m_1,\mathscr{F}_1)$, $(m_2,\mathscr{F}_2)$ be two random sets. $(m_1,\mathscr{F}_1)$ is then said to be *q*-included in $(m_2,\mathscr{F}_2)$ if and only if, for all $A \subseteq \mathscr{X}$,

$$q_1(A) \leq q_2(A),$$

and we denote this inclusion by $(m_1,\mathscr{F}_1) \sqsubseteq_{q} (m_2,\mathscr{F}_2)$.

And neither of these notions implies the other (that is, two random sets can be *pl*-included in each other and not *q*-included, and vice versa). [20]

Definition 4. Let $(m_1,\mathscr{F}_1)$, $(m_2,\mathscr{F}_2)$ be two random sets and $\mathscr{F}_1 = \{E_1,\dots,E_q\}$, $\mathscr{F}_2 = \{E'_1,\dots,E'_p\}$ be their respective sets of focal elements. Then, $(m_1,\mathscr{F}_1)$ is said to be s-included in $(m_2,\mathscr{F}_2)$, or to be a *specialization* of $(m_2,\mathscr{F}_2)$ if and only if there exists a non-negative matrix G, of generic term g_{ij} and such that

$$\text{for } i = 1,\dots,q, \quad \sum_{j=1}^{p} g_{ij} = 1,$$

$$g_{ij} > 0 \Rightarrow E_i \subseteq E'_j,$$

$$m_2(E'_j) = \sum_{i=1}^{q} m_1(E_i) g_{ij}.$$

The term g_{ij} is the proportion of the mass $m(E'_j)$ that "flows down" to focal set E_i. In other words, $(m_1,\mathscr{F}_1)$ is s-included in $(m_2,\mathscr{F}_2)$ if the mass of any focal set E'_j of $(m_2,\mathscr{F}_2)$ can be redistributed among subsets of E'_j in $(m_1,\mathscr{F}_1)$. When $(m_1,\mathscr{F}_1)$ is s-included in $(m_2,\mathscr{F}_2)$, we denote it by $(m_1,\mathscr{F}_1) \sqsubseteq_s (m_2,\mathscr{F}_2)$. Dubois and Prade [20] have shown that $(m_1,\mathscr{F}_1) \sqsubseteq_s (m_2,\mathscr{F}_2)$ implies both $(m_1,\mathscr{F}_1) \sqsubseteq_{pl} (m_2,\mathscr{F}_2)$ and $(m_1,\mathscr{F}_1) \sqsubseteq_q (m_2,\mathscr{F}_2)$.

Given a particular notion of inclusion, we say that a first random set $(m_1,\mathscr{F}_1)$ is an outer-approximation (resp. inner-approximation) of a second random set $(m_2,\mathscr{F}_2)$ when $(m_2,\mathscr{F}_2)$ is included in (resp. includes) $(m_1,\mathscr{F}_1)$. If $(m_2,\mathscr{F}_2)$ is included in $(m_1,\mathscr{F}_1)$, we also say that $(m_2,\mathscr{F}_2)$ is more committed, or is more specific than $(m_1,\mathscr{F}_1)$.

In this paper, we will only use the notion of *s*-inclusion. It it is the most natural inclusion notion to use with random sets, since it is expressed by means of inclusion between focal elements. Also, since *s*-inclusion implies both *pl*- and *q*-inclusion, an outer-approximation with respect to *s*-inclusion is ensured to be an outer-approximation with respect to both *pl*- and *q*-inclusion, while it would not be the case if we focused on one of these two later notions.

When working with possibility distributions and their induced random sets, then the tree notions of inclusion collapse into Definition 1, and results holding for one of them holds for the others. This is not the case when working with the recent notions of w- and v-inclusions introduced by Denoeux [21], which do not reduce to possibilistic inclusion (Definition 1) when particularised to random sets with nested focal sets (i.e., possibility distributions). This is why we do not consider such notions in the present paper.

2.4. *Independence modeling*

Given N marginal random sets $(m_1,\mathscr{F}_1),\dots,(m_N,\mathscr{F}_N)$ respectively modeling uncertainty over variables $X_1,\dots,X_N$, assuming random set independence allows to easily build a joint random set over $\mathscr{X}_{(1:N)}$. Let $E_1,\dots,E_N$ be any collection of focal elements of $(m_1,\mathscr{F}_1),\dots,(m_N,\mathscr{F}_N)$ (i.e., $E_i \in \mathscr{F}_i$), then the joint random set resulting from $(m_1,\mathscr{F}_1),\dots,(m_N,\mathscr{F}_N)$ under the assumption of random set independence, is denoted by $(m_{RSI,X_{(1:N)}},\mathscr{F}_{RSI,X_{(1:N)}})$, such that

$$m_{RSI,X_{(1:N)}}(\times_{i=1}^N E_i) = \prod_{i=1}^N m_i(E_i), \tag{5}$$

that is, the Cartesian product of focal sets receives as joint mass the product of marginal masses of these focal sets. As recalled in the introduction, random set independence is likely to be useful in many practical situations, but we can see from (5) that the number of focal sets will grow exponentially with the number N of dimensions. Such a joint structure is thus likely to become quickly intractable in practice.

When each marginal random set $(m_{\pi_1},\mathscr{F}_{\pi_1}),\dots,(m_{\pi_N},\mathscr{F}_{\pi_N})$ is consonant, that is stems from possibility distributions $\pi_1,\dots,\pi_N$, another way to combine these random sets, originating from possibility theory and first proposed by Zadeh, [23] is to consider the joint possibility distribution denoted by $\pi_{(1:N)}$ and defined, for every $x_{(1:N)} \in \mathscr{X}_{(1:N)}$, by:

$$\pi_{(1:N)}(x_{(1:N)}) = \min_{i=1,\dots,N} \pi_i(x_i). \tag{6}$$

We will denote by $(m_{\pi_{(1:N)}},\mathscr{F}_{\pi_{(1:N)}})$ the corresponding random set. This notion, called *possibilistic non-interaction* by Zadeh, is sometimes also referred to as fuzzy set independence. [13]. In fact it corresponds to an absence of assumption about dependence of underlying variables. Here, we adopt the first terminology, and call *non-interactive* a joint possibility distribution built from marginal distributions by *min*-combination (i.e. using Equation (6)), as well as the induced random set, and the underlying variables. If $\{1 = \alpha_1 > \ldots > \alpha_M > \alpha_{M+1} = 0\}$ is the (finite) set of all distinct values taken by $\pi_1,\dots,\pi_N$ (resp. on $\mathscr{X}_1,\dots,\mathscr{X}_N$), then $(m_{\pi_{(1:N)}},\mathscr{F}_{\pi_{(1:N)}})$ has, for $j = 1,\dots,M$, the following focal elements:

$$\begin{cases} E_{\pi_{(1:N)},j} = \{x_{(1:N)} \in \mathscr{X}_{(1:N)} | \pi_{(1:N)}(x_{(1:N)}) \geq \alpha_i\} = \times_{j=1}^N E_{i,j}, \\ m(E_{\pi_{(1:N)},j}) = \alpha_j - \alpha_{j+1}, \end{cases} \tag{7}$$

with $E_{i,j}$ the α_j-cut of the marginal distribution π_i. In other words, focal elements of $(m_{\pi_1}, \mathscr{F}_{\pi_1}), \ldots, (m_{\pi_N}, \mathscr{F}_{\pi_N})$ are combined level-wise, and correspond to an assumption of complete correlation between α-cuts. It means sources provide cuts with the same confidence levels but variables x_i are otherwise logically independent in $\times_{i=1}^{N} E_{i,j}$. Note that, in the above Equation (7), the number M of focal elements of the joint structure can only increase linearly with the number N of dimensions, thus providing a more manageable joint structure than (5). Also note that (5) does not preserve consonance of joint focal sets, while (7) ensures it by construction.

It is then tempting to use the simpler joint possibility distribution $(m_{\pi_{(1:N)}}, \mathscr{F}_{\pi_{(1:N)}})$ to approximate the more complex belief structure $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$. However, it is well known that given some marginal random sets $(m_{\pi_1}, \mathscr{F}_{\pi_1}), \ldots, (m_{\pi_N}, \mathscr{F}_{\pi_N})$, [24,25] the joint structure $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ neither is s-included nor s-includes the joint structure $(m_{\pi_{(1:N)}}, \mathscr{F}_{\pi_{(1:N)}})$. Hence, using the more manageable $(m_{\pi_{(1:N)}}, \mathscr{F}_{\pi_{(1:N)}})$ to approximate $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ is not without risk, as it does not guarantee any kind of conservatism. In the rest of this paper, we focus on finding a minimal guaranteed outer approximation (in the sense of the s-inclusion) of $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ that has the features of $(m_{\pi_{(1:N)}}, \mathscr{F}_{\pi_{(1:N)}})$.

3. Possibilistic outer-approximation of independent consonant random sets

The question we address in this section is the following: is it possible to transform the marginal distributions $\pi_1, \ldots, \pi_N$ into distributions $\pi'_1, \ldots, \pi'_N$ and then to combine these new distributions into a joint *consonant* random set $(m_{\pi'_{(1:N)}}, \mathscr{F}_{\pi'_{(1:N)}})$ over $\mathscr{X}_{(1:N)}$ using Equation (7), such that $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}}) \sqsubseteq_s (m_{\pi'_{(1:N)}}, \mathscr{F}_{\pi'_{(1:N)}})$ and is minimal with this property? In other words, can we define, from $\pi_1, \ldots, \pi_N$, a joint possibility distribution $\pi'_{(1:N)}$ whose induced random set s-includes $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$?

3.1. *Main result*

First, note that when constructing a non-interactive possibility distribution $\pi'_{X_{(1:N)}}$ (and the induced joint random set) from a transformation of $\pi_1, \ldots, \pi_N$, the focal elements of $(m_{\pi'_{X_{(1:N)}}}, \mathscr{F}_{\pi'_{X_{(1:N)}}})$ will be of type $\times_{i=1}^{N} E_{i,j}$. That is, they must still be Cartesian products of α-cuts of distributions $\pi_1, \ldots, \pi_N$. We can then answer the above question by the following proposition:

Proposition 1. *The most specific non-interactive possibility distribution $\pi'_{X_{(1:N)}}$ inducing a random set $(m_{\pi'_{X_{(1:N)}}}, \mathscr{F}_{\pi'_{X_{(1:N)}}})$ outer approximating (in the sense of s-inclusion) $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ is such that, for any $x_{(1:N)} \in \mathscr{X}_{(1:N)}$,*

$$\pi'_{X_{(1:N)}}(x_{(1:N)}) = \min_{i=1,\ldots,N} \{(-1)^{N+1}(\pi_i(x_i) - 1)^N + 1\}, \tag{8}$$

The detailed proof, which can be found in Appendix A, consists in showing that by applying Equation (8), the mass allocated to focal sets of type $\times_{i=1}^{N} E_{i,j}$ is the sum of the masses of all focal elements of $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ included in $\times_{i=1}^{N} E_{i,j}$ but not in $\times_{i=1}^{N} E_{i,j-1}$.

Proposition 1 extends to any number N of dimensions the result provided by Dubois and Prade [4] for the 2-dimensional case. It indicates that if one transforms each distribution π_i into

$$\pi_i' = (-1)^{N+1}(\pi_i - 1)^N + 1, \tag{9}$$

then consider the associated joint non-interactive possibility distribution $\min_{i=1}^{N} \pi_i'$ is a guaranteed outer approximation of the random set $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$. We thus cut down the size of the representation, from a structure whose complexity grows exponentially with the number of dimensions, to one that has a linear complexity in the number of dimensions.

Also note that we could use a t-norm other than the minimum as a combination operator in Equation (8) (dropping the assumption of non-interactivity [26]), and search for suitable transforms $\pi_1'', \dots, \pi_N''$ of $\pi_1, \dots, \pi_N$ providing the most specific outer-approximation. However, since the minimum is the most conservative of all t-norms, any joint possibility distribution outer-approximating $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ and combined by means of another t-norm would imply a transformation of marginal distributions such that $\pi_i'' \geq \pi_i'$ for any $i \in \{1, \dots, N\}$, thus losing even more information on each variable.

Our approximation intends to provide a conservative structure that directly approximates a complex joint structure $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$. It is straightforward to build and remains easy to manipulate within the family of consonant random sets. There are other approaches allowing to outer approximate some given random set. For example, given a random set modelling information on a single variable X, Denoeux [27] proposes to assign weights of some focal elements to coarser focal elements, building an outer-approximation s-including the original random set, without making any assumption about the structure of focal elements.

3.2. *Evaluating the loss of information*

The price to pay for moving from exponential to linear complexity in the number of dimensions is a loss of information in the obtained result. This loss must be measured. In particular, we can see that the value of π_i' in Equation (9) converges to 1 if $\pi_i(x_i) > 0$ as N increases, and is 0 if $\pi_i(x_i) = 0$. This means that, as N increases, the outer-approximation π' converges to a Boolean possibility distribution such that $\pi'_{X_{(1:N)}}(x_{(1:N)}) = 1$ if $x_{(1:N)} \in \times_{j=1}^{N} \pi_{i,\bar{0}}$, zero otherwise (i.e., towards the Cartesian product of supports of distributions $\pi_i, i = 1, \dots, N$). Both Figures 1 and 2 provide some intuition about the rate of convergence.

Before commenting these figures, recall a known result concerning the best inner consonant approximation of independent random set: [4]

Proposition 2. *The most specific joint possibility distribution $\pi^{\prod}_{X_{(1:N)}}$ whose induced random set inner-approximate $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ (in the sense of s-inclusion) is such that*

$$\pi^{\prod}_{X_{(1:N)}}(x_{(1:N)}) = \prod_{i=1}^{N} \pi_i(x_i)$$

Fig. 1. Evolution of outer-approximating (—) and inner-approximating (- - -) distribution degree (α) versus input space dimension (N), for a given starting $\pi(x)$ ($N = 1$).

Fig. 2. Possibility distributions π'_i obtained from a marginal triangular possibility distribution π_i for different input space dimensions (1,2,3,4,5,10,15,20)

For this inner approximation, all values strictly lower than one converge to zero as the number of dimension increases, indicating that the inner approximation converges towards the Cartesian product of cores of distributions π_i. Note that there does not currently exist any easy means to express $\pi^{\prod}_{X_{(1:N)}}$ as a non-interactive joint possibility distribution (i.e., as a min-based combination of transformed marginal possibility distributions). This makes the inner approximation less attractive from a computational perspectives, as one will have to consider the joint model as a whole and will not be able to make level-wise computations on marginal distributions.

Figure 1 plots the evolution of fixed initial possibility degree values against the number of dimensions. That is, each full line represent $\pi'_i(x)$ versus the number of dimensions, for a given $\pi_i(x)$, while dotted lines represent the same information for $\pi^{\prod}$. It shows that the information loss induced by the adoption of the possibilistic outer-approximation can be important, since it converges quickly to one. Still, the approximation is potentially useful when dealing with a reasonable number of variables (i.e., less than 10).

Figure 2 then sketches some distributions π'_i for different input space dimensions, starting from a triangular possibility distribution π_i on the real line, with center 0 and support $[-1,1]$. We can see on this figure that, even if the loss of information is important (and thus the approximation likely to be coarse), part of this information remains, even for high dimensions.

4. Comparisons with other approaches on illustrative examples

As we have seen in the previous section, the proposed outer approximation allows for a significant decrease of complexity of the resulting joint structure, but it also implies an important loss of information. After this study of the approximation itself, it is legitimate to wonder if, in applications, this approximation could be useful compared to other ones, and in which specific cases is it better to use it?

In this section, we bring some insight by focusing on the problem of uncertainty propagation. We consider that $X_1,\dots,X_N$ take their values on closed intervals of the real line, that uncertainty about these values are modeled by (discrete or discretized) possibility distributions $\pi_1,\dots,\pi_N$ and that either the variables or the sources providing information about them can be judged independent. We then consider the problem of propagating uncertainty on input variables $X_1,\dots,X_N$ through a (functional) model $T : \mathscr{X}_{(1:N)} \to \mathscr{Y}$ in order to evaluate the resulting uncertainty on Y.

Propagating uncertainty with random sets is, from a mathematical standpoint, easy.

Given a random set $(m,\mathscr{F})$ defined over the Cartesian product $\mathscr{X}_{(1:N)}$, then the propagated random set $(m_Y,\mathscr{F}_Y)$ is such that, to any focal set $E \in \mathscr{F}$ ($E \subseteq \mathscr{X}_{(1:N)}$), corresponds the propagated focal set E_Y

$$\begin{cases} E_Y = T(E) = \{T(x_{(1:N)})|x_{(1:N)} \in E\}, \\ m(E_Y) = m(E). \end{cases}$$

Propagating a random set simply consists in mapping every focal set to a set through T. The most difficult parts are (i) in the assessment and construction of the joint random set over $\mathscr{X}_{(1:N)}$ and (ii) in the propagation through T, which can be computationally very demanding, especially when $(m,\mathscr{F})$ has a high number of focal sets and/or when evaluations of T are costly. In this case, it can be useful to relax some assumptions about the dependence structure or to consider some suitable outer approximation in order to cut down the complexity of the propagation. In the following, we will compare two such approaches:

(1) The relaxation of the random set independence assumption by considering all possible dependence structures. The resulting propagation is indeed conservative and allows the use of so-called probabilistic arithmetic, [28] a well-known efficient tool to propagate uncertainties.
(2) The propagation of our proposed outer approximation $(m_{\pi'_{X_{(1:N)}}},\mathscr{F}_{\pi'_{X_{(1:N)}}})$ by means of the extension principle, [29] that is the computation of π'_Y such that, or any $y \in \mathscr{Y}$,

$$\pi'_Y(y) = \sup_{T(x_1,\dots,x_N)=y} \min(\pi'_{X_1}(x_1),\dots,\pi'_{X_N}(x_N)) \tag{10}$$

with, for $i=1,\dots,N$, π'_{X_i} given by Eq. (9) and $x_i \in \mathscr{X}_i$. This amounts to propagating the random set $(m_{\pi'_{X_{(1:N)}}},\mathscr{F}_{\pi'_{X_{(1:N)}}})$ rather than $(m_{RSI,X_{(1:N)}},\mathscr{F}_{RSI,X_{(1:N)}})$.

4.1. *Probabilistic arithmetic*

Let us first recall the basics about probabilistic arithmetic and the uncertainty model it uses, i.e., p-boxes. A p-box is a pair of (discrete) cumulative distributions $[\underline{F},\overline{F}]$ defined on a closed interval of the real line $\mathbb{R}$ that induces a probability family such that $\mathscr{P}_{[\underline{F},\overline{F}]} = \{P \in \mathbb{P}_X | \forall r \in \mathbb{R}, \underline{F}(r) \leq P([-\infty,r) \leq \overline{F}(r)\}$. It is known that a p-box is also a special kind of random set. [30,31,32]

To any possibility distribution π defined on the real line, we can associate a p-box $[\underline{F},\overline{F}]_\pi$ such that, for any $r\in\mathbb{R}$, $\underline{F}_\pi(r) = N((-\infty,r])$ and $\overline{F}_\pi(r) = \Pi((-\infty,r])$, with N,Π the necessity and possibility measures based on π. We also have that the random set $(m_\pi,\mathscr{F}_\pi)$ induced by π is pl-included in the random set $(m_{[\underline{F},\overline{F}]_\pi},\mathscr{F}_{[\underline{F},\overline{F}]_\pi})$ induced by $[\underline{F},\overline{F}]_\pi$(i.e., $\mathscr{P}_\pi \subseteq \mathscr{P}_{[\underline{F},\overline{F}]_\pi}$, hence $[\underline{F},\overline{F}]_\pi$ outer-approximates π). [8]

Now, let $[\underline{F},\overline{F}]_{\pi_1},\dots,[\underline{F},\overline{F}]_{\pi_N}$ be p-boxes deriving from distributions $\pi_1,\dots,\pi_N$. When, and only when T is expressible as a combination of arithmetic operations (or, more generally, of monotonic functions of two variables, e.g., $\log,\exp,\dots$), probabilistic arithmetic provides an efficient tool to propagate all these p-boxes (or their equivalent random set) while assuming unknown dependencies between them. That is, it considers every possible kind of dependencies between $[\underline{F},\overline{F}]_{\pi_1},\ldots,[\underline{F},\overline{F}]_{\pi_N}$ (among which is random set independence). The result is thus an outer approximation of the propagation of $(m_{RSI,X_{(1:N)}},\mathscr{F}_{RSI,X_{(1:N)}})$.

Given two real-valued variables X,Y and some p-boxes $[\underline{F},\overline{F}]_X,[\underline{F},\overline{F}]_Y$ describing uncertainty pervading them, the result of applying each arithmetic operations $\{+,-,\times,\div\}$ reads, for any $z\in\mathbb{R}$:

$$\begin{aligned}
\underline{F}_{X+Y}(z) &= \sup_{x,y\in\mathbb{R},x+y=z}\{\max(\underline{F}_X(x)+\underline{F}_Y(y)-1,0)\}\\
\overline{F}_{X+Y}(z) &= \inf_{x,y\in\mathbb{R},x+y=z}\{\min(\overline{F}_X(x)+\overline{F}_Y(y),1)\}\\
\underline{F}_{X-Y}(z) &= \sup_{x,y\in\mathbb{R},x+y=z}\{\max(\underline{F}_X(x)+\overline{F}_Y(-y),0)\}\\
\overline{F}_{X-Y}(z) &= \inf_{x,y\in\mathbb{R},x+y=z}\{\min(\overline{F}_X(x)+1-\underline{F}_Y(-y),1)\}\\
\underline{F}_{X\times Y}(z) &= \sup_{x,y\in\mathbb{R},x\times y=z}\{\max(\underline{F}_X(x)+\underline{F}_Y(y)-1,0)\}\\
\overline{F}_{X\times Y}(z) &= \inf_{x,y\in\mathbb{R},x\times y=z}\{\min(\overline{F}_X(x)+\overline{F}_Y(y),1)\}\\
\underline{F}_{X\div Y}(z) &= \sup_{x,y\in\mathbb{R},x\times y=z}\{\max(\underline{F}_X(x)+\overline{F}_Y(1/y),0)\}\\
\overline{F}_{X\div Y}(z) &= \inf_{x,y\in\mathbb{R},x\times y=z}\{\min(\overline{F}_X(x)+1-\underline{F}_Y(1/y),1)\}
\end{aligned}$$

Remark Note that the expressions for computing lower cumulative functions are the same as those for computing fuzzy interval arithmetic computations under the extension principle where the minimum is changed into a t-norm (here the Łukasiewicz t-norm [33] [34] $\max(a+b-1,0)$). Likewise, propagating the optimal inner approximation of Proposition 2 comes down to computing with fuzzy intervals using a sup-product extension principle.

4.2. *Comparison on illustrative examples*

Let us now compare, on some illustrative examples, the propagation of p-boxes $[\underline{F},\overline{F}]_{\pi_1},\ldots,[\underline{F},\overline{F}]_{\pi_N}$ by probabilistic arithmetic with the exact propagation of the outer approximation $(m_{\pi'_{X_{(1:N)}}},\mathscr{F}_{\pi'_{X_{(1:N)}}})$. To make this comparison, we will transform the p-box resulting from probabilistic arithmetic, denoted $[\underline{F},\overline{F}]_Y$, into the possibility distribution $\pi_{[\underline{F},\overline{F}]_Y}$ from which it could stem. That is, for any value $r\in\mathbb{R}$,

$$\pi_{[\underline{F},\overline{F}]_Y}(r)=\begin{cases}\overline{F}_Y(r) \text{ if } \overline{F}_Y(r)<1\\ 1-\underline{F}_Y(r) \text{ if } \underline{F}_Y(r)>0\end{cases} \tag{11}$$

Example 1. First consider the simple function $Y=X_1+X_2-X_3$, with X_1,X_2,X_3 positive real-valued variables whose uncertainty is modeled by the same possibility distribution, summarized in Table 1 , together with their transformation (9) and the distribution resulting from the application of Eq (10).

$\pi_{X_1}, \pi_{X_2}, \pi_{X_3}$	$\Rightarrow_{(8)}$	$\pi'_{X_1}, \pi'_{X_2}, \pi'_{X_3}$	π'_Y
Masses (m)	Focal Sets	Trans. masses (m')	Focal Sets
0.1	$[1,2]$	0.01	[0,3]
0.7	$[0.5,3]$	0.511	[-2,5.5]
0.2	$[0.1,5]$	0.488	[-4.8,9.9]

Table 1. Distributions of Example 1

Figure 3 shows results stemming from the propagation of $(m_{\pi'_{X_{(1:3)}}}, \mathscr{F}_{\pi'_{X_{(1:3)}}})$, from the application of probabilistic arithmetic as well as the possibility distribution covering the propagation of $(m_{RSI,X_{(1:3)}}, \mathscr{F}_{RSI,X_{(1:3)}})$ and centered around $[0,3]$. ♦

Fig. 3. Result comparison for the model $Y = X_1 + X_2 - X_3$

All results from Example 1 are similar and provide reasonably good approximations. Probabilistic arithmetic performs even better in this specific case. Thus, when T is an analytical model in which each variable appears once, the approximation $(m_{\pi'_{X_{(1:N)}}}, \mathscr{F}_{\pi'_{X_{(1:N)}}})$ is likely to be not really useful, as other techniques will have comparable efficiency and performance.

The next example shows that, in more complex cases, using $(m_{\pi'_{X_{(1:N)}}}, \mathscr{F}_{\pi'_{X_{(1:N)}}})$ and exactly propagating its focal sets can be of some usefulness.

Example 2. We now consider a model T where Y is a function of two positive real-valued variables X_1, X_2:

$$Y = T(X_{(1:2)}) = \frac{(X_1^2 + X_2^2)}{(2X_1 + 1)(X_2^3 - 1.9)}.$$

Figure 4 shows the behaviour of the function $T(X_{(1:2)})$. We can see that, while the function is non-decreasing in X_2, it is non-monotonic in X_1 (for example, if we fix $X_2 = 2$). Table 2 describes the possibility distributions describing the uncertainty on X_1, X_2.

The random set $(m_{\pi'_{X_{(1:2)}}}, \mathscr{F}_{\pi'_{X_{(1:2)}}})$ induced by the joint distribution $\pi'_{X_{(1:2)}}$ outer-approximating $(m_{RSI,X_{(1:2)}}, \mathscr{F}_{RSI,X_{(1:2)}})$ is summarized in Table 3, as well as the result of propagating each of its focal elements through T by using Eq.(10).

The resulting distribution has the interval $[0.0113, 0.5478]$ as support (i.e., α-cut of level 0) and $[0.1036, 0.2732]$ as mode (i.e., α-cut of level 1).

π_{X_1}		π_{X_2}	
m_{X_1}	$\mathscr{F}_{X_1}$	m_{X_2}	$\mathscr{F}_{X_2}$
0.1	$[1,2]$	0.5	$[2,3]$
0.7	$[0.5,3]$	0.4	$[2,5]$
0.2	$[0.1,5.1]$	0.1	$[2,10]$

Table 2. Distributions of Example 2

$\pi'_{X_{(1:2)}}$			π'_Y
$m'_{X_{(1:2)}}$	$\mathscr{F}'_{X_{(1:2)}}$		$\mathscr{F}'_Y$
$[1,2]\times[2,3]$	0.01		$[0.1036, 0.2732]$
$[0.5,3]\times[2,3]$	0.24	T	$[0.1013, 0.3484]$
$[0.5,3]\times[2,5]$	0.39	$\Rightarrow$	$[0.0395, 0.3484]$
$[0.1,5.1]\times[2,5]$	0.17		$[0.0368, 0.5478]$
$[0.1,5.1]\times[2,10]$	0.19		$[0.0113, 0.5478]$

Table 3. $(m_{\pi'_{X_{(1:2)}}}, \mathscr{F}_{\pi'_{X_{(1:2)}}})$ and propagation result.

Applying probabilistic arithmetic to the above example and then Eq.(11) results in a possibility distribution having interval $[0.0003, 17.08]$ as support and $[0.007, 2.7868]$ as

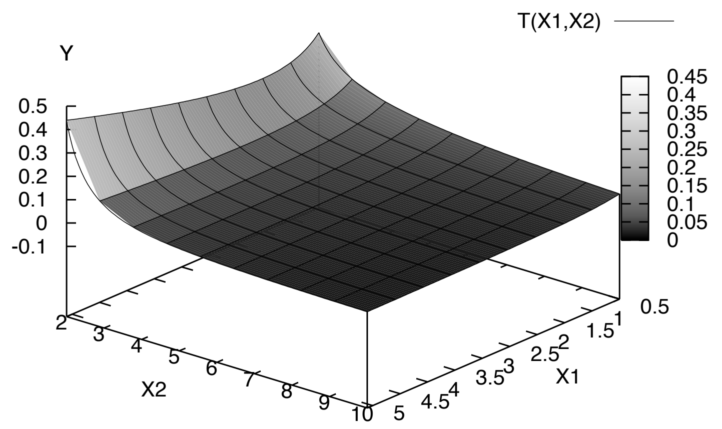

Fig. 4. Function $Y = T(X_1, X_2)$ of Example 2

core. We can see that these intervals are far more conservative than the one obtained with the outer-approximation $\pi'_{X_{(1:2)}}$ (the support of the result obtained by probabilistic arithmetic method is more than 30 times larger than the one produced by our approximation). This is mainly due to the fact that both X_1 and X_2 appears more than once in the analytic expression of the model, and that, in such cases, applying interval arithmetic operations (that are special cases of p-box arithmetic operations) to compute the uncertain output of a model like T does not provide best-possible bounds. Had we applied fuzzy arithmetic to propagate π'_{X_1}, π'_{X_2} through T, we would have obtained a distribution having $[0.00035, 17.21]$ as support and $[0.04, 0.71]$, that is somewhat closer to the result obtained by using probabilistic arithmetic.

Note that it is possible to use methods proposed in Baudrit et al., [35] that make the same dependence assumption as probabilistic arithmetic (i.e., unknown independence) but provide best-possible bounds (i.e., avoid the problem of repeated variables) and can deal with general functions. Such methods would have given us yet another outer approximation, probably closer to the one obtained in Table 3. However, such approaches require, to calculate probabilistic bounds on each event, the resolution of a particular linear programming problem, and have computational complexities even higher than computing the exact propagation of each marginal random sets with an assumption of independence (that is, working directly with $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$). Using such methods is therefore not relevant in this work.

This indicates that the proposed approximate representation and the use of interval analysis methods for implementing the extension principle Eq. (10) is likely to be useful in those cases where the use of probabilistic arithmetic is known to perform poorly, namely when:

- T is locally monotonic (that is, monotonic in each variable when fixing the values of other ones), but its analytical formula, expressed as a combination of arithmetic operations, contains multiple instances of the same variable, and cannot be reduced to a form where each variable appears once.
- The model T is not isotone, that is extrema are not reached on boundaries of intervals, but evaluating extrema of T remains feasible (either by heuristic searches or analytical derivation).

Also, the use of $\pi'_{X_{(1:2)}}$ as an outer-approximation does not constrain in any way the nature of the model T (which can be a complex and non-linear model), while probabilistic arithmetic can only be used within a restricted selection of functions.

5. Conclusions

When working in multiple dimensions, handling the combination of marginal and independent random sets can be a tedious task, especially since the resulting joint structure has an exponentially growing complexity. A way to reduce the complexity of this structure is to work with an approximation that benefits from the computational advantages of a simplified framework.

Here, we have looked at the case where marginal random sets are consonant (note that these marginal random sets can themselves be consonant approximations of non-consonant ones [4]) and are assumed to be random set independent. Even if this is a restricted framework, it is likely to occur in many practical applications (as advocated in the introduction), and can already be hard to deal with.

Consequently, we have proposed a transformation of marginal random sets that allows to build a joint possibility distribution outer-approximating the exact joint structure resulting from an assumption of independence. This outer-approximation cuts down the complexity from exponential to linear in the number of dimensions. This drastic reduction, which significantly alleviates the computational burden of subsequent treatments, is paid by a potentially important loss of information, of which the user must be aware.

However, we have shown that our outer-approximation can provide good results in some situations where other quick approximations perform poorly. Also, there will be some situations where the use of the approximation will be sufficient to give a satisfying answer (e.g., risk analysis), and will therefore avoid the use of computationally more demanding methods.

Finally, since within the setting of imprecise probabilities, many other different notions of independence have emerged, [36] it would be desirable to define possibilistic approximations of independence assumptions similarly to the view developed here, simply because possibilistic approximations are computationally convenient. Some results concerning such an approximation for the notion of epistemic independence can be found. [37]

References

1. P. Walley, *Statistical reasoning with imprecise Probabilities* (Chapman and Hall, New York, 1991).
2. D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty* (Plenum Press, New York, 1988).
3. I. Molchanov, *Theory of Random Sets* (Springer, London, 2005).
4. D. Dubois and H. Prade, “Consonant approximations of belief functions”, *Int. J. of Approximate Reasoning* **4** (1990) 419–449.
5. S. Sandri, D. Dubois and H. Kalfsbeek, “Elicitation, assessment and pooling of expert judgments using possibility theory”, *IEEE Trans. on Fuzzy Systems* **3** (1995) 313–335.
6. A. Birnbaum, “Confidence curves: an omnibus technique for estimation and testing statistical hypothesis”, *Journal of American Statistical Association* **56** (1961) 246–249.
7. D. Dubois, L. Foulloy, G. Mauris and H. Prade, “Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities”, *Reliable Computing* **10** (2004) 273–297.
8. C. Baudrit and D. Dubois, “Practical representations of incomplete probabilistic knowledge”, *Computational Statistics and Data Analysis* **51** (2006) 86–108.
9. M. Masson and T. Denoeux, “Inferring a possibility distribution from empirical data”, *Fuzzy Sets and Systems* **157** (2006) 319–340.
10. A. Aregui and T. Denoeux, “Constructing consonant belief functions from sample data using confidence sets of pignistic probabilities”, *Int. J. of Approximate Reasoning* **49** (2008) 575–594.
11. C. Baudrit, I. Couso and D. Dubois, “Joint propagation of probability and possibility in risk analysis: towards a formal framework”, *Int. J. of Approximate Reasoning* **45** (2007) 82–105.
12. I. Couso, “Independence concepts in evidence theory”, in *Proc. of the 5th Int. Symp. on Imprecise Probability: Theories and Applications* (2007).

13. T. Fetz, "Sets of joint probability measures generated by weighted marginal focal sets", in *Proc. 2st International Symposium on Imprecise Probabilities and Their Applications*, eds. F. Cozman, R. Nau and T. Seidenfeld (2001).
14. G. Shafer, *A Mathematical Theory of Evidence* (Princeton University Press, New Jersey, 1976).
15. A. Dempster, "Upper and lower probabilities induced by a multivalued mapping", *Annals of Mathematical Statistics* **38** (1967) 325–339.
16. P. Smets and R. Kennes, "The transferable belief model", *Artificial Intelligence* **66** (1994) 191–234.
17. D. Dubois, P. Hajek and H. Prade, "Knowledge-driven versus data-driven logics", *J. of Logic, Language and Information* **9** (2000) 65–89.
18. D. Dubois and H. Prade, "On several representations of an uncertain body of evidence", in *Fuzzy Information and Decision Processes*, eds. M. Gupta and E. Sanchez (North-Holland, 1982), pp. 167–181.
19. E. Miranda, I. Couso and P. Gil, "Relationships between possibility measures and nested random sets", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **10** (2002) 1–15.
20. D. Dubois and H. Prade, "A set-theoretic view on belief functions: logical operations and approximations by fuzzy sets", *Int. J. of General Systems* **12** (1986) 193–226.
21. T. Denoeux, "Conjunctive and disjunctive combination of belief functions induced by non-distinct bodies of evidence", *Artificial Intelligence* **172** (2008) 234–264.
22. P. Smets, "The canonical decomposition of a weighted belief", in *Proc. Int. Joint. Conf. on Artificial Intelligence* (Montreal, 1995), pp. 1896–1901.
23. L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning I", *Information Sciences* **8** (1975) 199–249.
24. F. Tonon and S. Chen, "Inclusion properties for random relations under the hypotheses of stochastic independence and non-interactivity", *Int. J. of General Systems* **34** (2005) 615–624.
25. C. Baudrit, D. Guyonnet and D. Dubois, "Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment", *IEEE Trans. Fuzzy Systems* **14** (2006) 593–608.
26. E. Klement, R. Mesiar and E. Pap, *Triangular Norms* (Kluwer Academic Publisher, Dordrecht, 2000).
27. T. Denoeux, "Inner and outer approximation of belief structures using a hierarchical clustering approach", *Int. J. on Uncertainty, Fuzziness and Knowledge-Based Systems* **9** (2001) 437–460.
28. R. Williamson and T. Downs, "Probabilistic arithmetic I : Numerical methods for calculating convolutions and dependency bounds", *Int. J. of Approximate Reasoning* **4** (1990) 8–158.
29. D. Dubois, E. Kerre, R. Mesiar and H. Prade, *Fundamentals of fuzzy sets* (Kluwer, Boston, 2000), chapter Fuzzy interval analysis, pp. 483–581.
30. S. Destercke, D. Dubois and E. Chojnacki, "Unifying practical uncertainty representations: I generalized p-boxes", *Int. J. of Approximate Reasoning* **49** (2008) 649–663.
31. S. Destercke, D. Dubois and E. Chojnacki, "Unifying practical uncertainty representations: II clouds", *Int. J. of Approximate Reasoning* **49** (2008) 664–677.
32. E. Kriegler and H. Held, "Utilizing random sets for the estimation of future climate change", *Int. J. of Approximate Reasoning* **39** (2005) 185–209.
33. D. Dubois and H. Prade, "Additions of interactive fuzzy numbers", *IEEE Trans. on Automatic Control* **26** (1981) 926–936.
34. M. Wagenknecht, R. Hampel and V. Schneider, "Computational aspects of fuzzy arithmetics based on archimedean t-norms", *Fuzzy Sets and Systems* **123** (2001) 49–62.
35. C. Baudrit and D. Dubois, "Comparing methods for joint objective and subjective uncertainty propagation with an example in a risk assessment", in *Proc. Fourth International Symposium on Imprecise Probabilities and Their Application (ISIPTA'05)* (Pittsburg (USA, Pennsylvanie), 2005), pp. 31–40.

36. I. Couso, S. Moral and P. Walley, "A survey of concepts of independence for imprecise probabilities", *Risk Decision and Policy* **5** (2000) 165–181.
37. E. Miranda and G. de Cooman, "Epistemic independence in numerical possibility distribution", *Int. J. of Approximate Reasoning* **32** (2003) 23–42.

Appendix A. Proof of Proposition 1

Proof. We consider the finite set

$$\{\alpha \in [0,1] | i = 1,\dots,N,\ \exists x \in \mathscr{X}_i \text{ s.t. } \pi_i(x) = \alpha\}$$

of M distinct values taken by all distributions $\pi_1,\dots,\pi_N$. We consider that these values are indexed such that $1 = \alpha_1 > \dots > \alpha_M > \alpha_{M+1} = 0$, and we denote by $E_{i,j}$ the α_j-cut of distribution π_i. Note that the masses of each random set $(m_{\pi_i}, \mathscr{F}_{\pi_i})$, $i = 1,\dots,N$ form the same vector $(m_{i,1},\dots,m_{i,M})$, and to simplify notations, we will adopt the notation $m_j := m_{i,j}$ for some i. To prove Proposition 1, let us first express the values that should be assigned to elements $x_{(1:N)}$ of $\mathscr{X}_{(1:N)}$, so as to define a possibility distribution outer-approximating $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$. Let us do it in terms of masses m_j, $j = 1,\dots,M$, and then we will show that this expression is equivalent to the distribution $\pi'_{X_{(1:N)}}$ given by Equation (8).

Let us express the value of the outer approximation in terms of masses $m_{i,j}$. First, note that focal sets of $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ have the general form $\times_{i=1}^N E_{i,j_i}$, with mass $\prod_{i=1}^N m_{j_i}$.

For a given value $j \in \{1,\dots,M\}$, the focal sets of $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ that are included in $\times_{i=1}^N E_{i,j}$ but not in $\times_{i=1}^N E_{i,j-1}$ are of the form

$$\{ \bigotimes_{|I|=k, i\in I} E_{i,j} \times \bigotimes_{i\in\{1,\dots,N\}\setminus I} E_{i,j_i} | k = 1,\dots,N; I \subseteq \{1,\dots,N\}; j_i < j\} \quad \text{(A.1)}$$

with $\bigotimes$ standing for Cartesian product, and $|I|$ for the cardinality of I. For a fixed value k, there are $\binom{N}{k}$ different subset of $\{1,\dots,N\}$ having cardinality k. Following Dubois and Prade, [4] we can define a mass function defined on focal sets that are Cartesian products of type $\times_{i=1}^N E_{i,j}$ (i.e., α_j-cuts of distributions π_i) by

$$m^*(\times_{i=1}^N E_{i,j}) = \sum_{k=1}^N \binom{N}{k} m_j^k \sum_{j_1,\dots,j_{n-k}<j} m_{j_1}\dots m_{j_{n-k}}.$$

The above equation simply being the sum of masses of all elements described by Eq.(A.1). As all the vectors of weights are the same, we can factorize out the polynomial expression $\sum_{j_1,\dots,j_{n-k}<j} m_{j_1}\dots m_{j_{n-k}}$ and get

$$m^*(\times_{i=1}^N E_{i,j}) = \sum_{k=1}^N \binom{N}{k} m_j^k \left(\sum_{l<j} m_l\right)^{N-k}.$$

These masses sum up to one. It corresponds to a possibility distribution with focal elements $\times_{i=1}^N E_{i,j}$. It minimally outer-approximates $(m_{RSI,X_{(1:N)}}, \mathscr{F}_{RSI,X_{(1:N)}})$ in the sense of Proposition 1 by construction. Now, let us consider (as done by Dubois and Prade [4]) an element $x_{(1:N)} \in (\times_{i=1}^{N} E_{i,j}) \setminus (\times_{i=1}^{N} E_{i,j-1})$ (recall that $E_{i,j} \subseteq E_{i,j-1}$ for any $i \in \{1,\dots,N\}$ and $j \in \{2,\dots,M\}$), that is an element $x_{(1:N)}$ that is in the Cartesian product of α_j-cuts, but not α_{j-1}-cuts. Note that only these elements have to be considered, since the outer-approximation is consonant with focal sets of type $\times_{i=1}^{N} E_{i,j}$. Given the outer-approximating mass m^* on sets $\times_{i=1}^{N} E_{i,j}$, we get

$$\pi'_{X_{(1:N)}}(x_{(1:N)}) = \sum_{i \geq j} m^*(\times_{k=1}^{N} E_{k,i}) = \sum_{i \geq j} \left(\sum_{k=1}^{N} \binom{N}{k} m_i^k \left(\sum_{l<i} m_l \right)^{N-k} \right) = \alpha'_j.$$

Given our choice of $x_{(1:N)}$, we also have that $\min_{i=1,\dots,N} \pi_i(x_i) = \alpha_j$. What we want to check is whether, by applying Equation (8), we do have $\min_{i=1,\dots,N} \pi'_i(x_i) = \alpha'_j$. To answer this, first notice that $\pi_i(x_i) = \alpha_j = \sum_{i \geq j} m_i$, and that Equation (8) can be rewritten $(-1)^{N+1}(\sum_{i \geq j} m_i - 1)^N + 1$. Checking that $\min_{i=1,\dots,N}(\pi'_i(x_i)) = \alpha'_j$ then amounts to proving the following equality:

$$\sum_{i \geq j} \left(\sum_{k=1}^{N} \binom{N}{k} m_i^k \left(\sum_{l<i} m_l \right)^{N-k} \right) = (-1)^{N+1}(\sum_{i \geq j} m_i - 1)^N + 1. \tag{A.2}$$

The case $N = 1$ is trivial, and, for $N = 2$, it has been originally checked by Dubois and Prade. [4] We now prove its validity for any N. First, note that

$$\left(\sum_{l \leq i} m_l \right)^N = \left(m_j + \sum_{l<i} m_l \right)^N = \sum_{k=0}^{N} \binom{N}{k} (m_i)^k \left(\sum_{l<i} m_l \right)^{N-k} = \left(\sum_{l<i} m_l \right)^N + \sum_{k=1}^{N} \binom{N}{k} (m_i)^k \left(\sum_{l<i} m_l \right)^{N-k}$$

and thus

$$\sum_{k=1}^{N} \binom{N}{k} (m_i)^k \left(\sum_{l<i} m_l \right)^{N-k} = \left(\sum_{l \leq i} m_l \right)^N - \left(\sum_{l<i} m_l \right)^N.$$

Let $C_i := \sum_{l \leq i} m_l$ the sum of masses $m_1,\dots,m_i$, the left-hand side of equation (A.2) can be rewritten

$$\sum_{i \geq j} \left(\left(\sum_{l \leq i} m_l \right)^N - \left(\sum_{l<i} m_l \right)^N \right) = \sum_{i \geq j} \left((C_i)^N - (C_{i-1})^N \right) = (C_M)^N - (C_{j-1})^N$$

and, likewise, the right-hand side can be rewritten

$$\begin{aligned} (-1)^{N+1}(\sum_{i \geq j} m_i - 1)^N + 1 = &\quad (-1)^{N+1}(-\sum_{i<j} m_i)^N + 1 = (-1)^{N+1}(-1)^N (C_{j-1})^N + 1 \\ = &\quad (-1)^{2N+1}(C_{j-1})^N + 1 = (C_M)^N - (C_{j-1})^N, \end{aligned}$$

since $C_M = 1$ by definition. This completes the proof. □