Computing with generalized p-boxes: preliminary results

Sébastien Destercke, Didier Dubois, Eric Chojnacki

Computing with generalized p-boxes: preliminary results

Sébastien Destercke
Institut de Radioprotection
et de Sûreté Nucléaire
CE Cadarache
13115 St-Paul-Lez-Durance
sebastien.destercke@irsn.fr

Didier Dubois
IRIT - équipe RPDMP
118 Route de Narbonnes
31000 Toulouse
dubois@irit.fr

Eric Chojnacki
Institut de Radioprotection
et de Sûreté Nucléaire
CE Cadarache
13115 St-Paul-Lez-Durance
eric.chojnacki@irsn.fr

Abstract

The need to propagate uncertainties through a model is present in many applications. In most cases, the nature of this model is either graphical or functional. In this paper, we interest ourselves to the latter case. We consider here that uncertainty on each model input is described either by generalized p-boxes or possibility distributions, two special cases of random sets that can be interpreted in term of confidence bounds over nested sets. We then study their practical propagation for different cases.

Keywords: uncertainty propagation, generalized p-boxes, random sets, possibility distributions.

1 Introduction

The propagation of uncertainty modeled by classical probabilities is an old research topic that still faces a lot of challenges, but past years have also witnessed a growing interest for the problem of propagating uncertainty modeled by means of other theories explicitly coping with imprecision. The main reasons for this interest is that imprecision is a feature of the information that classical probabilities cannot adequately account for, and that the problem of propagating uncertainty is at the core of many practical applications. Nevertheless, explicitly modeling imprecision in an uncertainty model often increases the complexity of the propagation, since propagating imprecision requires propagating sets of values, while classical probabilistic propagation is based on repeated calculations of precise values. Consequently, there is a great need of efficient methods to propagate uncertainty through mathematical models. These models can be either graphical (e.g. extensions of bayesian networks [2]) or analytical (e.g. models of physical phenomena [1]).

Figure 1 gives a synopsis of the general problem of propagating uncertainty models through a function. Given the choice of a theoretical framework, some information on the inputs and on their mutual (in)dependencies, there are mainly three ways of increasing the efficiency of the propagation[1]:

- uncertainty models: as a general rule, more expressive approaches allow to model more complex information, but also implies more computational effort when propagating. By using less expressive models, one can intentionally choose to give away some information in order to gain some efficiency,
- propagation mechanisms: another way of increasing the propagation efficiency is to design more efficient algorithms, possibly using some knowledge we have on the model,
- approximate propagations: as exact propagation can be difficult to achieve in general, one can use alternative propagation methods that will give results approximating the exact propagation. In this last case, it is important to know what is the relationship (guaranteed outer/inner approximation, neither) between this approximated result and the exact result.

[1] We consider here that the model is fixed. Otherwise, another solution is to simplify the model

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 1–8
Torremolinos (Málaga), June 22–27, 2008

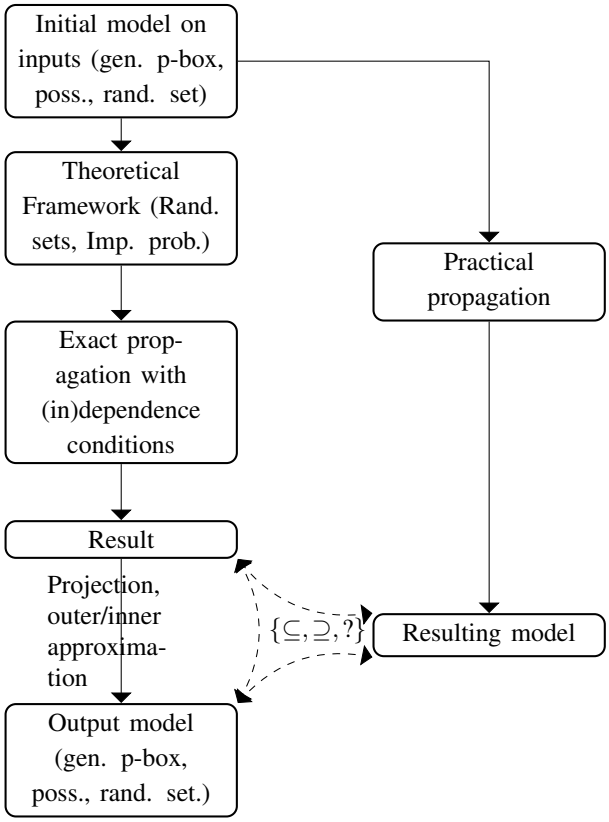

Figure 1: Propagating Uncertainty through function: synopsis

In this paper, we focus on the first and third points, and take random set theory as our basic theoretical framework. The uncertainty models we consider here are so-called generalized p-boxes and/or possibility distributions, that have appealing properties but have limited expressiveness, as they are special cases of random sets. We then use the particularities of generalized p-boxes and/or possibility distributions to propose practical propagation techniques for various situations.

Section 2 recalls the basics needed in the rest of the paper. Section 3 then concentrates on the case where uncertainty concerns only one input (univariate case). Finally, section 4 deals with the case of uncertainty bearing on multiple inputs (multivariate case) that can be considered mutually independent (in the sense of the so-called random set independence).

2 Preliminaries

We consider in this paper that each input space X^i is a finite space of n^i elements x^i (i.e., upper indexes denote dimension index). When considering one space, we drop the indices.

2.1 Random sets

Formally, a random set [4] is a mapping from a probability space to the power set of another space. In the discrete case [12], a random set can also be termed as a mass distribution $m : \wp(X) \to [0,1]$ s.t. $\sum_{E \in \wp(X)} m(E) = 1$. In this case, subsets E having a strictly positive mass are called focal elements. From a random set, we can define two uncertainty measures, respectively the belief and plausibility functions, which reads, for all $A \subset X$:

$$Bel(A) = \underline{P}(A) = \sum_{E, E \subseteq A} m(E)$$
$$Pl(A) = \overline{P}(A) = \sum_{E, E \cap A \neq \emptyset} m(E)$$

The belief function quantifies our credibility in event A, by summing all the masses that **surely** support A, while the plausibility function measures the maximal confidence that can be given to event A, by summing all masses that **could** support A. They are dual measures, in the sense that for all events A, we have $Bel(A) = 1 - Pl(A^c)$.

In the sequel, random sets will be denoted $(m, \mathscr{F})$, with m the mass distribution and $\mathscr{F}$ the set of focal elements.

2.2 Possibility distributions

A possibility distribution [6] is a mapping $\pi : X \to [0,1]$ from a (here finite) space X to the unit interval such that $\pi(x) = 1$ for at least one element x in X. Formally, a possibility distribution is equivalent to the membership function of a fuzzy set. From this possibility distribution, we can define two uncertainty measures, respectively the belief and plausibility functions, which reads, for all $A \subset X$:

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

Given a possibility distribution π and a degree $\alpha \in [0,1]$, the strong and regular α-cuts are subsets respectively defined as the sets $A_{\overline{\alpha}} = \{x \in X | \pi(x) > \alpha\}$ and $A_\alpha = \{x \in X | \pi(x) \geq \alpha\}$. These α-cuts are nested, since if $\alpha > \beta$, then $A_\alpha \subset A_\beta$. In the finite case, a possibility distribution takes at most n values. Let us note $\alpha_0 = 0 < \alpha_1 < \ldots < \alpha_m = 1$ these m values.

Possibility distributions can also be interpreted as particular random sets. Namely, they are equivalent to random sets whose focal elements are nested: a belief function (resp. a plausibility function) is a necessity measure (resp a possibility measure) if and only if it derives from a mass function with nested focal sets. Given a possibility distribution π, the corresponding random set will have the following focal elements E_i with masses $m(E_i)$, $i = 1, \ldots, m$:

$$\begin{cases} E_i = \{x \in X | \pi(x) \geq \alpha_i\} = A_{\alpha_i} \\ m(E_i) = \alpha_i - \alpha_{i-1} \end{cases} \quad (1)$$

and this random set is called consonant by Shafer [12].

As practical models, possibility distributions can be naturally interpreted as nested sets of confidence intervals (i.e. cuts of level α has confidence $1-\alpha$), and are thus easy to assess. Moreover, their simplicity makes them easy to use. The weak side of possibility distributions is that their expressivity is limited (i.e. for an event A, bounds $[N(A), \Pi(A)]$ are either of the kind $[0, \alpha]$ or $[\beta, 1]$)), thus they can be found insufficient models if available information is more complex.

2.3 Generalized p-boxes

A Generalized p-box is defined as follows:

Definition 1. A generalized p-box $[\underline{F}, \overline{F}]$ over a finite space X is a pair of comonotonic[2] mappings $\underline{F}, \overline{F}, \underline{F} : X \to [0,1]$ and $\overline{F} : X \to [0,1]$ from X to $[0,1]$ such that $\underline{F}$ is point-wise lower than $\overline{F}$ (i.e. $\underline{F} \leq \overline{F}$) and there is at least one element x in X for which $\overline{F}(x) = \underline{F}(x) = 1$.

Given a generalized p-box $[\underline{F}, \overline{F}]$, we can always define a complete pre-ordering $\leq_{[\underline{F},\overline{F}]}$ on elements x of X that is such that $x \leq_{[\underline{F},\overline{F}]} y$ if $\underline{F}(x) \leq \underline{F}(y)$ and $\overline{F}(x) \leq \overline{F}(y)$. The name generalized p-box comes from the fact that if X is the real line and the order is the natural order of numbers, we retrieve the usual notion of p-boxes [9].

To shorten notations, we will consider in the sequel that given a general p-box $[\underline{F}, \overline{F}]$ on X, elements x of X are indexed by natural integers in a way such that $x_i \leq_{[\underline{F},\overline{F}]} x_j$ if and only if $i \leq j$. Let us now denote for all $i = 1, \ldots, n$ by A_i the sets $\{x_j \in X | x_j \leq x_i\}$.

Uncertainty modeled by generalized p-boxes can also be mapped into a set of constraints that are upper and lower confidence bounds on the uncertainty of A_i, namely, for $i = 1, \ldots, n$:

$$\alpha_i \leq P(A_i) \leq \beta_i \quad (2)$$

where $\alpha_i = \underline{F}(x_i)$, $\beta_i = \overline{F}(x_i)$, $P(A_i)$ is the (unknown) probability of event A_i and with $A_0 = \emptyset$, $\alpha_0 = \beta_0 = 0$. We also have, for all i from 0 to $n-1$, $\alpha_i \leq \alpha_{i+1}$,$\beta_i \leq \beta_{i+1}$ and $A_i \subseteq A_{i+1}$.

It can also be shown [5] that the uncertainty modeled by any generalized p-box can be mapped into a particular random set. This random set can be built by the following procedure: consider a threshold $\theta \in [0,1]$. When $\alpha_{i+1} > \theta \geq \alpha_i$ and $\beta_{j+1} > \theta \geq \beta_j$, then, the corresponding focal set is $A_{i+1} \setminus A_j$, with weight

$$m(A_{i+1} \setminus A_j) = \min(\alpha_{i+1}, \beta_{j+1}) - \max(\alpha_i, \beta_j). \quad (3)$$

Generalized p-boxes can also be linked to possibility distributions in the following way [5]: the uncertainty modeled by a generalized p-box $[\underline{F}, \overline{F}]$ is equivalent to the uncertainty modeled by a pair of possibility distributions $\pi_{\overline{F}}, \pi_{\underline{F}}$ that are such that, for $i = 1, \ldots, n$,

$$\pi_{\overline{F}}(x_i) = \beta_i$$

and

$$\pi_{\underline{F}}(x_i) = 1 - \max\{\alpha_j | j = 0, \ldots, i\ \alpha_j < \alpha_i\}.$$

and the random sets $m_{\pi_{\overline{F}}}$ and $m_{\pi_{\underline{F}}}$ modeling the uncertainty of these distributions are such that, for $i = 0, \ldots, n-1$,

$$m_{\pi_{\overline{F}}}(A_i^c) = \beta_i - \beta_{i-1} \text{ and } m_{\pi_{\underline{F}}}(A_{i+1}) = \alpha_{i+1} - \alpha_i$$

Thus, we have three different ways of characterizing a generalized p-box: by a set of lower/upper bounds

[2] Two mappings f and f' on a ranked-space $X = \{x_1, \ldots, x_n\}$ are said comonotonic if there exist a common permutation σ such that $f(x_{\sigma(1)}) \geq f(x_{\sigma(2)}) \geq \ldots \geq f(x_{\sigma(n)})$ and $f'(x_{\sigma(1)}) \geq f'(x_{\sigma(2)}) \geq \ldots \geq f'(x_{\sigma(n)})$

on nested sets, by an equivalent random set or by a pair of possibility distributions. Each of these views suggest different propagation techniques, that will be explored in the next section.

From a practical point of view, there are various reasons to give attention to generalized p-boxes and to their propagation: similarly to possibility distributions, they can be interpreted in terms of confidence bounds given to nested subsets, making them easy to assess and explain; they have more expressive power than possibility distributions, since lower and upper confidence bounds on an event A can now be of the kind $[\alpha,\beta]$, and since they remain special cases of random sets, we can try to use their specific properties to derive propagation methods more efficient than those used for general random sets.

2.4 Propagation of random sets

Let f be a function from the Cartesian product $\times_{i=1}^{p} X^i$ of input spaces X^i to the output space Y.

If we then consider a joint random set $(m,\mathscr{F})^{1,\ldots,p}$ with n focal elements $E_j \subset \times_{i=1}^{p} X^i$ and weights $m(E_j)$, then, the propagated random set is such that, for $j=1,\ldots,n$:

$$E_j^y = f(E_j) = \{f(\mathbf{x}) \in Y | \mathbf{x} \in E_j\}$$
$$m(E_j^y) = m(E_j)$$

where $\mathbf{x}$ denote a vector of $\times_{i=1}^{p} X^i$. The propagation of the joint random set thus consists of mapping every focal element E_j into $f(E_j)$. Depending on our knowledge of f, this operation can be more or less complex. For instance, if sets E_j are Cartesian products of closed intervals defined on the real line, computing $f(E_j)$ is usually easy when f is isotone, but can become very greedy in computational efforts if the behavior of f is complex and/or ill-known.

In the case where the information is given in terms of p marginal random sets $(m,\mathscr{F})^i$ on each space X^i, a first step before propagating the information is to build the joint random set $(m,\mathscr{F})^{1,\ldots,p}$. We will deal with this step in section 4, since we do not need it in the univariate case.

3 Univariate case

In this section, we consider propagating uncertainty bearing on variable x (which takes values on X) through a function $f(x)=y$ where y is the output variable. Note that f can depend on other parameters and be a complex functional, but we consider that only the value of x is imperfectly known. We thus consider that uncertainty on x is modeled by a generalized p-box $[\underline{F},\overline{F}]$ that we have to propagate.

There is (at least) three ways of doing this propagation: by propagating the nested sets and their lower/upper confidence bounds, by propagating the random set equivalent to this generalized p-box, and by independently propagating the two possibility distributions. After each propagation, we can build the corresponding random set, and then compare these random sets.

The first solution, propagating nested sets and their confidence bounds consists of computing for each set A_i the propagated set $f(A_i)$, and to consider the generalized p-box induced by the constraints:

$$\forall i=1,\ldots,n,\ \alpha_i \leq P(f(A_i)) \leq \beta_i \qquad (4)$$

where α_i,β_i are the confidence bounds originally related to set A_i. Given this propagated generalized p-box (it is still a generalized p-box, since sets $f(A_i)$ are also nested), we can build the counterpart of the random set given by equation (3), which is here:

$$\left.\begin{array}{r}\theta \in [0,1]\\ \alpha_{i+1} > \theta \geq \alpha_i\\ \beta_{j+1} > \theta \geq \beta_j\end{array}\right\} \begin{array}{l} m(f(A_{i+1}) \setminus f(A_j)) = \\ \min(\alpha_{i+1},\beta_{j+1}) - \max(\alpha_i,\beta_j)\end{array}$$

that we note $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$

The second solution, directly propagating through function f the focal elements $A_{i+1} \setminus A_j$ of the random set given by (3), gives:

$$\left.\begin{array}{r}\theta \in [0,1]\\ \alpha_{i+1} > \theta \geq \alpha_i\\ \beta_{j+1} > \theta \geq \beta_j\end{array}\right\} \begin{array}{l} m(f(A_{i+1} \setminus A_j)) = \\ \min(\alpha_{i+1},\beta_{j+1}) - \max(\alpha_i,\beta_j)\end{array}$$

a random set potentially different from the one given by the first propagation. We note this second random set $(m,\mathscr{F})_{f((m,\mathscr{F}))}$.

The third solution consists of propagating both possibility distributions by the so-called extension principle. This is equivalent to propagate the respective focal elements of each distribution through f, which gives us the random sets $(m,\mathscr{F})_{f(\pi_{\underline{F}})}$ and

$(m,\mathscr{F})_{f(\pi_{\overline{F}})}$ whose mass functions respectively are, for $i = 0,\ldots,n-1$,

$$m(f(A_i^c)) = \beta_i - \beta_{i-1} \text{ and } m(f(A_{i+1})) = \alpha_{i+1} - \alpha_i$$

and, if we take from these two random sets the counterpart of the random set given by equation (3), we end up with:

$$\left.\begin{array}{r} \theta \in [0,1] \\ \alpha_{i+1} > \theta \geq \alpha_i \\ \beta_{j+1} > \theta \geq \beta_j \end{array}\right\} \begin{array}{l} m(f(A_{i+1}) \setminus f(A_j^c)^c) = \\ \min(\alpha_{i+1},\beta_{j+1}) - \max(\alpha_i,\beta_j) \end{array}$$

that we note $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$.

We can already note that the three random sets $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$, $(m,\mathscr{F})_{f((m,\mathscr{F}))}$, $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$ have the same mass function distributed over different focal elements. To compare the results of the three propagations, we thus have to compare the informative content of their respective focal elements. The following proposition can be used to do such a comparison:

Proposition 1. *Let A and B be two subsets of a space X such that $A \subset B$, and let f be a function from X to another space Y. Then, we have the following inclusion relations:*

$$f(B) \setminus f(A) \subseteq f(B \setminus A) \subseteq f(B) \setminus f(A^c)^c$$

and inclusion relationships become equalities if f is injective

Proof. We will first prove the first inclusion relationship, then the second one, each time showing that we have equality if f is injective.

Let us first prove that any element of $f(B) \setminus f(A)$ is in $f(B \setminus A)$. Let us consider an element y in $f(B) \setminus f(A)$. This implies:

$$\left.\begin{array}{l} y \in f(B) \\ y \notin f(A) \end{array}\right\} \Rightarrow \exists x \in X \left\{\begin{array}{l} f(x) \in f(B) \\ f(x) \notin f(A) \end{array}\right.$$

and this x is in B and not in A (i.e. in $B \setminus A$), which implies that $y = f(x)$ is in $f(B \setminus A)$. This means that $f(B) \setminus f(A) \subseteq f(B \setminus A)$, and we still have to show that this inclusion can be strict. To see it, consider the case where one of the element x in $B \setminus A$ is such that $f(x)$ takes the same value as $f(x')$, where x' is in A, thus this particular $f(x)$ is in $f(B \setminus A)$ and not in $f(B) \setminus f(A)$ (since by assumption it is in $f(A)$), showing that the inclusion can be strict. This case does not happen if f is injective (since if f is injective $f(x) = f(x')$ if and only if $x = x'$).

To prove the second inclusion relation, first note that $f(B \setminus A) = f(B \cap A^c)$ and that $(f(B) \setminus f(A^c)^c) = (f(B) \cap f(A^c))$. Known results immediately give $f(B \cap A^c) \subseteq f(B) \cap f(A^c)$. Strict inclusion happens in the case where we have an element x of X in B and in A, and another element x' not in A and not in B (i.e. x' is in A^c) for which $f(x) = f(x')$, thus we have that x and x' are not in $B \cap A^c$, but are respectively in B and A^c, and thus $f(x)$ is in $f(B) \cap f(A^c)$. Again, this case cannot happen when f is injective (since in this case, $x \neq x'$ implies $f(x) \neq f(x')$). □

What proposition 1 tell us is that, when f is not injective, we have in general

$$(m,\mathscr{F})_{f([\underline{F},\overline{F}])} \subseteq (m,\mathscr{F})_{f((m,\mathscr{F}))} \subseteq (m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$$

thus showing that $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ is more optimistic than $(m,\mathscr{F})_{f((m,\mathscr{F}))}$, which is itself more optimistic than $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$. And in the case where f is injective, all these propagations are equivalent. Note that other recent works underly the particular role of injectivity when propagating p-boxes and corresponding random sets [13].

The question is then, if f is not injective, why should we choose one propagation rather than the other?

From a computational complexity standpoint, $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ seems more convenient than $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$, which in turn seems more convenient than $(m,\mathscr{F})_{f((m,\mathscr{F}))}$. The main reason is that, to compute $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ and $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$, we have to compute mappings of focal elements that are collections of nested sets (one collection in the first case, two in the second), allowing us to use this nestedness to cut down the number of required computations, while focal elements of $(m,\mathscr{F})_{f((m,\mathscr{F}))}$ are not nested. To illustrate this, let us consider that f is a complex non-monotonic mapping from $\mathbb{R}$ to $\mathbb{R}$, where $\mathbb{R}$ is the real line. Given the sets $A_0 \subset A_1 \subseteq \ldots \subseteq A_n$, let us consider that the global minimum and maximum of f are respectively in $A_i \setminus A_{i-1}$, and in $A_j \setminus A_{j-1}$, and that we know their values. This means that in the propagation, we no longer have to compute the lower bounds of all $f(A_k), f(A_l^c)$ such that $k > i > l$ nor the upper bounds of all $f(A_{k'}), f(A_{l'}^c)$ such that $k' > j > l'$.

Also, the maximal number of sets that have to be propagated for computing $(m,\mathscr{F})_{f((m,\mathscr{F}))}$ is $(n+1)n/2$, while it is $2n$ for $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$ and only n for $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$.

If we now take theoretical aspects into account, the only exact result is provided by the random set $(m,\mathscr{F})_{f((m,\mathscr{F}))}$ resulting from the third propagation. Moreover, the information modeled by $(m,\mathscr{F})_{f((m,\mathscr{F}))}$ is consistent, in the sense that no mass is affected to the empty set. Also note that it is also coherent with imprecise probability theory, which is not considered here, due to lack of space.

We may then find $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$ useful, because it is conservative when compared to $(m,\mathscr{F})_{f((m,\mathscr{F}))}$, ensuring us that we are cautious and that the resulting information will be consistent. Moreover, this propagation is consistent with the extension principle of possibility theory.

Finally, although $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ is surely the easiest propagation to compute, it is more optimistic than $(m,\mathscr{F})_{f((m,\mathscr{F}))}$, implying that, compared to $(m,\mathscr{F})_{f((m,\mathscr{F}))}$, we could dangerously reduce our uncertainty on y by adding unwanted assumptions. Moreover, the random set $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ can be such that some mass is allocated to the empty set, thus indicating some inconsistency in the information modeled by $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$.

Finally, if faced with a practical problem, the best solution is to compute $(m,\mathscr{F})_{f((m,\mathscr{F}))}$ if possible. If not possible, computing $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$, yields $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ for free (since for computing the former we need to propagate sets A_i). So, another solution is to bracket the information contained in $(m,\mathscr{F})_{f((m,\mathscr{F}))}$ using $(m,\mathscr{F})_{f(\pi_{\underline{F}},\pi_{\overline{F}})}$ and $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$. Computing $(m,\mathscr{F})_{f([\underline{F},\overline{F}])}$ only is not cautious.

Again, if f is injective, such analysis is pointless since the three propagations give the same results. Note that from a practical point of view, sticking to injective functions can be very restrictive. For instance, if X is a subset of $\mathbb{R}$, requiring injectivity of f is equivalent to limiting ourselves to strictly monotone functions on X.

4 Multivariate case

We now consider that our knowledge on multiple parameters $x^1,\ldots,x^p$ respectively taking values on $X^1,\ldots,X^p$ is tainted with uncertainty and that we must propagate this uncertainty through a function $y=f(x^1,\ldots,x^p)$ where y takes values on a space Y. Note that the results of the previous section hardly apply, because such functions are generally not injective when useful (e.g. monotonic ones)[3]. To simplify the problem, we here consider that our uncertainty on each variable x^i is described by a possibility distribution π^i to which correspond a random set $(m,\mathscr{F})^i$.

Before doing anything else, we must first specify how we build the joint random set $(m,\mathscr{F})^{1,\ldots,p}$ that we are going to consider. To do this, we assume here that the random sets are independent between them, in the sense that, for every subset E of $\times_{i=1}^{p} X^i$, the joint mass $m(E)$ is such that:

$$m(E)=\sum_{\substack{\times_{i=1}^{p} E_j^i=E \\ E_j^i\in\mathscr{F}^i}} \left(\prod_{i=1}^{p} m(E_j^i)\right)$$

where $\times_{i=1}^{p} E_j^i$ is the Cartesian product of the focal elements E_j^i. This assumption is commonly called random set independence. The assumption of random set independence can be interpreted as the assumption that the sources of information of each variables x^i are independent (e.g. different sensors measure each variable x^i). Also, an assumption of random set independence is conservative when compared to other notions of independence [3], and can thus be used as a conservative tool to approximate such assumptions (which are often difficult to handle in practice). Although all variables are not often mutually independent, there is often subgroups of variables that can be judged independent, and to which the following results apply.

Under independence assumption, propagating $(m,\mathscr{F})^{1,\ldots,p}$ is not without difficulty, since the number of focal elements to propagate grows exponentially with the number of input variables tainted with uncertainty. It is thus important to give practical approximation methods allowing to reduce the computational cost of the propagation, especially when

[3] Nevertheless, results of Section 3 still apply when lower and upper confidence bounds are given for a collection of nested multi-dimensional sets

f is complex and the available resources are limited. In the sequel, we provide a technique to get an outer approximation of $(m,\mathscr{F})^{1,\dots,p}$ by means of a joint possibility distribution, that can then be propagated more easily than the general structure $(m,\mathscr{F})^{1,\dots,p}$. In other words, what we want to do is to find a joint possibility distribution $\pi'^{1,\dots,p}$ such that the consonant random set $(m,\mathscr{F})^{\pi'^{1,\dots,p}}$ induced by this possibility distribution is an outer approximation of $(m,\mathscr{F})^{1,\dots,p}$. Such an outer approximation is given by the following property, which extends a result found by Dubois and Prade [7] for the 2-dimensional case:

Proposition 2. *For $i=1,\dots,p$, given the marginal distributions π^i and the joint random set $(m,\mathscr{F})^{1,\dots,p}$, the minimal possibility distribution $\pi'^{1,\dots,p}$ inducing a random set $(m,\mathscr{F})^{\pi^{1,\dots,p}}$ that is an outer approximation of $(m,\mathscr{F})^{1,\dots,p}$ is such that*

$$\pi'^{1,\dots,p}(x^1,\dots,x^p) = \min_{i=1,..,p}\{(-1)^{p+1}(\pi^i(x^i)-1)^p+1\}$$

The proof consists in a generalization of the proof given by Dubois and Prade [7] for the 2-dimensional case, and is omitted here due to lack of space.

In other words, we can transform each distribution π^i into $\pi'^i=(-1)^{p+1}(\pi^i-1)^p+1$ and then, propagating them by means of the possibilistic extension principle to yield an outer approximation of the exact propagation of $(m,\mathscr{F})^{1,\dots,p}$. Let us recall that propagating distributions π'^i through f comes down to compute π'^f such that

$$\pi'^f(y) = \sup_{\substack{x^1,\dots,x^p\in\times_{i=1}^p X^i \\ f(x^1,\dots,x^p)=y}} \min_{i=1,\dots,p} \pi'^i(x^i)$$

And, since this extension principle is equivalent to performing a set propagation of each α-cut [8], it allows us to drastically reduce the computational effort. To illustrate this, let us consider that every marginal possibility distribution π_i takes the same q different values on $[0,1]$, then exactly propagating $(m,\mathscr{F})^{1,\dots,p}$ would require q^p set propagations, while computing the guaranteed outer approximation $(m,\mathscr{F})^{\pi'^{1,\dots,p}}$ by using proposition 2 would only require q set propagations, whatever the dimension of the input space.

Nevertheless, the input space dimension does have an effect on our approximation, since we can see that, for a particular π_i, the transformation $(-1)^{p+1}(\pi_i(x)-1)^p+1$ converges to 1 if $\pi_i(x)>0$ as p increases, and is 0 if $\pi_i(x)=0$. This means that, as p increases, the outer approximation converges to the Cartesian product of the supports of the π_i's. This loss of information is the price to pay for passing from an exponential to a linear complexity while having a guaranteed outer approximation (which ensures a cautious approximation). Moreover, the nestedness of α-cuts of $\pi'^{1,\dots,p}$ can again be used to make the propagation more efficient [10].

Note that if our marginal uncertainty models are generalized p-boxes $[\underline{F},\overline{F}]^i$, we can still use proposition 2 to get an outer approximation of $(m,\mathscr{F})^{1,\dots,p}$, where $(m,\mathscr{F})^{1,\dots,p}$ is the joint random set resulting from assuming random set independence between the marginal random sets induced by generalized p-boxes $[\underline{F},\overline{F}]^i$. To do this, it is sufficient to apply the transformation of proposition 2 to each possibility distributions $\pi^i_{\underline{F}},\pi^i_{\overline{F}}$, and then propagate all possible combination of these transformed possibility distributions by the extension principle. If we still assume that each possibility distribution $\pi^i_{\underline{F}},\pi^i_{\overline{F}}$ takes the same q values, then propagating all combinations by the extension principle will require $2^p\cdot q$ computations, which is generally lower[4] than q^p, and thus remains simpler to compute than $(m,\mathscr{F})^{1,\dots,p}$. But it may be that, due to the lack of injectivity, the result $f(\pi^1_{\underline{F}}\cdots\pi^p_{\underline{F}})$ is not informative. For instance suppose $f(x_1,x_2)=x_1+x_2$, and $x_i\in[a_i,b_i]\setminus[c_i,d_i]$ with $[c_i,d_i]\subset[a_i,b_i], i=1,2$. Then $\pi^i_{\underline{F}}=(-\infty,c_i],\cup[d_i,+\infty)$, but the sum of two such subsets of reals is the whole real line.

5 Conclusion

Propagating uncertainty through a model is a complex problem, and one of the main difficulty encountered by such a propagation is the high computational effort it requires. When the model is simple or the available resources sufficient enough, this computational effort can be supported, but it is no longer the case when resources are limited or when the model is complex (e.g. nuclear computer codes).

In this paper, we've studied the propagation of generalized p-boxes. This model is more general than

[4] It is lower when $2^p<q^{p-1}$, so q must be at least 4 for $p=2$ and 3 for $p=3$, a constraint often satisfied

possibility distributions, but remains a special case of general random sets. It also has a nice interpretation in terms of lower and upper confidence bounds over a collection of nested sets. This model is thus particularly interesting, since it is more expressive than single possibility distributions and more tractable then general random sets.

Both for the univariate and multivariate cases, we have proposed practical methods to propagate generalized p-boxes. In particular, we have provided methodologies that makes the propagation easier while ensuring that the result will encompass the exact propagation (this cautiousness principle is particularly important in safety studies).

Perspectives include (but are not limited to) the comparison of our approximation method in the multivariate case to other conservative propagation methods (e.g. use of so-called probabilistic arithmetic [14]), the psychological evaluation of generalized p-boxes in elicitation process [11], and the use of the presented methods in practical applications.

References

[1] C. Baudrit, D. Guyonnet, and D. Dubois. Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment. *IEEE Trans. Fuzzy Systems*, 14:593–608, 2006.

[2] A. Cano, M. Gomez, S. Moral, and J. Abellan. Hill-climbing and branch-and-bound algorithms for exact and approximate inference in credal networks. *Int. J. of Approximate Reasoning*, 44:261–280, 2007.

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

[4] A.P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[5] S. Destercke, D. Dubois, and E. Chojnacki. Relating practical representations of imprecise probabilities. In *Proc. 5th Int. Symp. on Imprecise Probabilities: Theories and Applications*, 2007.

[6] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

[7] D. Dubois and H. Prade. Consonant approximations of belief functions. *I.J. of Approximate reasoning*, 4:419–449, 1990.

[8] D. Dubois and H. Prade. Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems*, (42):87–101, 1992.

[9] S. Ferson, L. Ginzburg, V. Kreinovich, D.M. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Laboratories, 2003.

[10] M. Hanss. *Applied Fuzzy Arithmetic: An Introduction with Engineering Applications*. Springer, 2004.

[11] E. Raufaste, R.S. Neves, and C. Mariné. Testing the descriptive validity of possibility theory in human judgments of uncertainty. *Artificial Intelligence*, 148:197–218, 2003.

[12] G. Shafer. *A mathematical Theory of Evidence*. Princeton University Press, New Jersey, 1976.

[13] F. Tonon. Some properties of a random set approximation to upper and lower distribution functions. *I.J. of Approximate reasoning*, 48:174–184, 2008.

[14] R.C. Williamson and T. Downs. Probabilistic arithmetic i : Numerical methods for calculating convolutions and dependency bounds. *I. J. of Approximate Reasoning*, 4:8–158, 1990.