On Extreme Points of p-Boxes and Belief Functions

Ignacio Montes, Sébastien Destercke

On extreme points of p-boxes and belief functions

Ignacio Montes[1] and Sebastien Destercke[2]

Abstract The extreme points of convex probability sets play an important practical role, especially for specific, easier to manipulate sets. Although this problem has been studied for many models (probability intervals, possibility distributions), it remains to be studied for imprecise cumulative distributions (a.k.a. p-boxes). This is what we do in this paper, where we characterize the maximal number of extreme points of a p-box, give a family of p-boxes that attains this number and show an algorithm that allows to compute the extreme points of a given p-box. To achieve all this, we also provide what we think to be a new characterization of extreme points of a belief function.

1 Introduction

Imprecise probability theory [10] is a powerful unifying framework for uncertainty treatment, relying on convex sets of probabilities, or *credal sets*, to model the uncertainty. Formally, they encompass many existing models: belief functions, possibility distributions, probability intervals, ... To apply such models, it is important to study their practical aspects, among which is the characterization of their extreme points. Indeed, these extreme points can be used in many settings, such as graphical models or statistical learning.

Extreme points of many models have already been studied. For instance, Dempster [3] shows that the maximal number of extreme points of a belief function on a n-element space is $n!$. It was later [6] proved that the maximal number of extreme points for possibility distributions in a n-element space is 2^{n-1}, and in [8] an algorithm to extract them was provided. In [2], authors studied the extreme points of probability intervals.

Universidad Carlos III de Madrid igmontes@est-econ.uc3m.es · CNRS, UMR Heudiasyc, Universite de Technologie de Compiegne sebastien.destercke@hds.utc.fr

One practical and popular model for which extreme points have not been characterized are p-boxes [4]. They are special kinds of belief functions whose focal elements are ordered intervals [5, 9, 10], and are quite instrumental in applications such as risk and reliability analysis.

In this paper, we investigate extreme point of p-boxes: we demonstrate that their maximal number is the Pell number, and give the family of p-boxes for which this bound is obtained. To do so, we introduce a new way to characterize the extreme points of a belief function. We also provide an algorithm to compute the extreme points of a given p-box. Section 2 introduces the new characterization, while Section 3.2 studies the extreme points of p-boxes. Due to space restrictions, proofs and side results have been removed.

2 Extreme points of belief functions

Given a space $\mathcal{X} = \{x_1, \ldots, x_n\}$, a *probability mass function* is a function $m : \mathcal{P}(\mathcal{X}) \to [0,1]$ satisfying $m(\emptyset) = 0$ and $\sum_{B \subseteq \mathcal{X}} m(B) = 1$. A probability mass function m defines a *belief* Bel and a *plausibility* Pl function by:

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B) \quad \text{and} \quad \mathrm{Pl}(A) = \sum_{B : A \cap B \neq \emptyset} m(B) \quad \forall A \subseteq \mathcal{X}.$$

These two functions are conjugate since $\mathrm{Bel}(A) = 1 - \mathrm{Pl}(A^c)$, and we can focus on one of them. A *focal set* of the belief function Bel is a set E such that $m(E) > 0$, and $\mathcal{F}$ will denote the set of focal sets. A belief function also induces a credal set

$$\mathcal{M}(\mathrm{Bel}) = \{P \text{ Prob.} \mid \mathrm{Bel}(A) \leq P(A) \ \forall A \subseteq \mathcal{X}\}.$$

Being convex, the set $\mathcal{M}(\mathrm{Bel})$ can be characterized by its extreme points[1], that we will denote $\mathcal{E}xt(\mathrm{Bel})$. It is known [1, 3] that there is a correspondence between the extreme points of a belief function and the permutations of the elements of $\mathcal{X}$. The extreme point $P_\sigma \in \mathcal{E}xt(\mathrm{Bel})$ associated to the permutation σ of $\{1, \ldots, n\}$ is given by

$$P_\sigma(\{x_{\sigma(i)}\}) = \mathrm{Bel}(\{x_{\sigma(i)}, \ldots, x_{\sigma(n)}\}) - \mathrm{Bel}(\{x_{\sigma(i+1)}, \ldots, x_{\sigma(n)}\}) \tag{1a}$$

$$= \sum_{E \subseteq A_i^\sigma} m(E) - \sum_{E \subseteq A_{i+1}^\sigma} m(E) = \sum_{x_{\sigma(i)} \in E, E \cap A_i^{\sigma,C} = \emptyset} m(E) \tag{1b}$$

where $A_i^\sigma = \{x_{\sigma(i)}, \ldots, x_{\sigma(n)}\}$ and $A_i^{\sigma,C} = \{x_{\sigma(1)}, \ldots, x_{\sigma(i-1)}\}$ is its complement, and the convention $A_{n+1}^\sigma = A_1^{\sigma,C} = \emptyset$. However, we may have that

[1] Recall that an extreme point P of $\mathcal{M}(\mathrm{Bel})$ is a point such that, if $P_1, P_2 \in \mathcal{M}(\mathrm{Bel})$ and $\alpha P_1 + (1-\alpha) P_2 = P$ for some $\alpha \in (0,1)$, then $P_1 = P_2 = P$.

$P_{\sigma_1} = P_{\sigma_2}$, as in general not all permutations give rise to different extreme points, otherwise every belief function would have $n!$ extreme points. Eq. (1b) tells us that an extreme point is built iteratively, according to Algorithm 1.

Algorithm 1: Extreme point computation

Input: $\sigma, (Bel), \mathcal{E} = \mathcal{F}$
Output: P_σ
for *k=1,...,n* **do**
  For all $E \in \mathcal{E}$ s.t. $x_{\sigma(k)} \in E$, assign $m(E)$ to $P_\sigma(\{x_{\sigma(k)}\})$;
  $\mathcal{E} \leftarrow \mathcal{E} \setminus \{E \in \mathcal{E} | x_{\sigma(k)} \in E\}$
end

Let us now introduce another way to characterize this extreme point. To do so, we will denote by $\overline{v}_{i\setminus A} = |\{E \in \mathcal{F} | x_i \in E, E \cap A = \emptyset\}|$ the number of focal sets counting x_i as an element and having an empty intersection with A. Given a permutation σ, let us denote by $\mathbf{v}^\sigma = (v_1^\sigma, \dots, v_n^\sigma)$ the vector such that

$$v_i^\sigma = \overline{v}_{i\setminus A^{\sigma,C}_{\sigma^{-1}(i)}} = |\{E \in \mathcal{F} | x_i \in E, E \cap \{x_{\sigma(1)}, \dots, x_{\sigma(\sigma^{-1}(i)-1)}\} = \emptyset\}| \quad (2)$$

and by $\mathcal{V}(\mathrm{Bel})$ the set of vectors obtained for all permutation. We will also denote $\overline{\mathbf{v}}_A = (\overline{v}_{1\setminus A}, \dots, \overline{v}_{n\setminus A})$. We then have the following result.

Proposition 1. *Given Bel, if two permutations σ_1, σ_2 satisfy $P_{\sigma_1} = P_{\sigma_2}$, then $\mathbf{v}^{\sigma_1} = \mathbf{v}^{\sigma_2}$.*

Also note that any vector $\mathbf{v} \in \mathcal{V}(\mathrm{Bel})$ can be associated to a permutation σ generating an extreme points (to see this, note the link between Eqs. (2) and (1b)), for instance the permutation having generated it. Since by contraposition of Proposition 1, $\mathbf{v}^{\sigma_1} \neq \mathbf{v}^{\sigma_2}$ implies $P_{\sigma_1} \neq P_{\sigma_2}$, $\mathcal{V}(\mathrm{Bel})$ is in bijection with $\mathcal{E}xt(\mathrm{Bel})$ (any vector induces one and only one distinct extreme point). Given a vector $\mathbf{v} \in \mathcal{V}(\mathrm{Bel})$, we can easily find back a permutation generating it by using Algorithm 2

Algorithm 2: Permutation generating algorithm

Input: $\mathbf{v} \in \mathcal{V}(\mathrm{Bel})$, $\mathcal{E} = \mathcal{F}$
Output: One permutation σ generating $\mathbf{v}$
for *k=1,...,n* **do**
  Define $\mathbf{v}$ s.t. $v_i = |\{E \in \mathcal{E} | x_i \in E\}|$;
  Find i s.t. $v_i = \overline{v}_{i\setminus A^{\sigma,C}_k}$ /* Getting $A^{\sigma,C}_k$ only necessitates $\sigma(k-1)$ */
  Define $\sigma(k) = i$;
  $\mathcal{E} \leftarrow \mathcal{E} \setminus \{E \in \mathcal{E} | x_{\sigma(k)} \in E\}$
end

Example 1. Consider a belief function Bel defined on $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ such that

$$m(E_1 = \{x_1, x_2\}) = 0.2, \quad m(E_2 = \{x_2, x_3, x_4\}) = 0.5, \quad m(E_3 = \{x_3\}) = 0.3$$

Consider for example the permutation $\sigma = (1, 2, 3, 4)$. It generates the extreme point $P_\sigma = (0.2, 0.5, 0.3, 0)$. Indeed, according to Alg. 1, $m(E_1)$ is assigned to x_1, $m(E_2)$ to x_2 and $m(E_3)$ to x_3. Then, σ generates the vector $\mathbf{v} = (1, 1, 1, 0)$. Algorithm 2 can then generate permutations $(1, 2, 4, 3)$ or $(1, 2, 3, 4)$, as in the first iteration we only have $v_1 = \overline{v}_{1 \setminus A_1^{\sigma,C}} = 1$, meaning $\sigma(1) = 1$, and in the second iteration we only have $v_2 = \overline{v}_{2 \setminus A_2^{\sigma,C}} = 1$, and so on ....

The extreme points of the belief function in this example, as well as the permutations that generate them, can be seen in Table 1.

Permutation	Probability	$(v_1^\sigma, v_2^\sigma, v_3^\sigma, v_4^\sigma)$
(1,2,3,4) (1,2,4,3)	$P_{\sigma_1} = (0.2, 0.5, 0.3, 0)$	(1,1,1,0)
(1,3,2,4) (1,3,4,2) (3,4,1,2) (3,1,2,4) (3,1,4,2)	$P_{\sigma_2} = (0.2, 0, 0.8, 0)$	$(1, 0, 2, 0)$
(1,4,3,2) (1,4,2,3) (4,3,1,2) (4,1,3,2) (4,1,2,3)	$P_{\sigma_3} = (0.2, 0, 0.3, 0.5)$	$(1, 0, 1, 1)$
(2,3,1,4) (2,3,4,1) (2,4,1,3) (2,1,3,4) (2,1,4,3) (2,4,3,1)	$P_{\sigma_4} = (0, 0.7, 0.3, 0)$	$(0, 2, 1, 0)$
(3,2,1,4) (3,2,4,1) (3,4,2,1)	$P_{\sigma_5} = (0, 0.2, 0.8, 0)$	(1,0,2,0)
(4,3,2,1) (4,2,3,1) (4,2,1,3)	$P_{\sigma_6} = (0, 0.2, 0.3, 0.5)$	(0,1,1,1)

Table 1 Extreme points of the belief function of Example 1.

Note that this new characterization in terms of "counting" vectors allows us to derive new results about the extreme points of belief functions.

Proposition 2. *Let Bel be a belief function on $\mathcal{X} = \{x_1, \ldots, x_n\}$. The number of extreme points of Bel is $n!$ if and only if $\{x_i, x_j\}$ is a focal set for any $i, j \in \{1, \ldots, n\}$ such that $i \neq j$.*

Proposition 3. *Let Bel be a belief function on $\mathcal{X} = \{x_1, \ldots, x_n\}$. Denote by $\mathcal{F}$ the set of focal sets of Bel. Let Bel′ be another belief function and let $\mathcal{F}' = \mathcal{F} \cup \{E\}$ be the set of focal sets of Bel′, where $E \notin \mathcal{F}$. Then, Bel′ has at least as many extreme points as Bel.*

3 Extreme points of p-boxes

Before studying the extreme points of p-boxes, we need to make a small, useful digression about a specific number sequence: the Pell numbers. Quite like the

Fibonacci sequence, Pell numbers form a sequence that follows a recursive relation $\mathcal{P}_0 = 0, \quad \mathcal{P}_1 = 1, \quad \mathcal{P}_n = \mathcal{P}_{n-2} + 2\mathcal{P}_{n-1}$. The first numbers are: $0, 1, 2, 5, 12, 29, 70, \ldots$. It is known that $2^{n-1} \leq \mathcal{P}_n \leq n!$ for any $n \geq 1$. As we shall see, it turns out that the maximal number of extreme points of p-boxes on $\mathcal{X}$ is $\mathcal{P}_n$

3.1 Basic definitions

From now on we consider an ordered space[2] $\mathcal{X} = \{x_1, \ldots, x_n\}$ such that $x_1 < \ldots < x_n$. A *probability box* or p-box [4] $(\underline{F}, \overline{F})$ is a pair of cumulative distribution functions $\underline{F}, \overline{F} : \mathcal{X} \to [0, 1]$ such that $\underline{F} \leq \overline{F}$. Here we interpret p-boxes as lower and upper bounds of an ill-known cumulative distribution, that induce a credal set

$$\mathcal{M}(\underline{F}, \overline{F}) = \{P \text{ Prob.} \mid \underline{F}(x) \leq F_P(x) \leq \overline{F}(x) \ \forall x \in \mathcal{X}\},$$

where F_P denotes the cumulative distribution function associated with the probability P.

It is known that p-boxes are particular instances of belief functions (see [9, 10] for details). That is, to any p-box we can associate a belief function such that $\mathcal{M}(\text{Bel}) = \mathcal{M}(\underline{F}, \overline{F})$. The focal sets $E_1, \ldots, E_k$ of this belief functions are known to be intervals[3] ordered with respect to the order $\preceq$ between intervals such that

$$[a_1, a_2] \preceq [b_1, b_2] \Leftrightarrow a_1 \leq b_1, a_2 \leq b_2.$$

That is, $E_1 \prec E_2 \prec \ldots \prec E_k$. For the reader interested in the way such focal sets can be built, we refer to [5]. This is also a characteristic property, as any belief functions whose focal sets are ordered intervals will be equivalent to a p-box.

3.2 Extreme points of a p-box

We can easily provide first bounds over the number of extreme points of p-boxes

Proposition 4. *The maximal number of extreme points of a p-box on $\mathcal{X} = \{x_1, \ldots, x_n\}$ ($n > 2$) lies in the interval $[2^{n-1}, n!)$.*

The exact maximal number of extreme points of a p-box is reached for the following family of p-boxes: the *Pell* p-boxes on $\mathcal{X} = \{x_1, \ldots, x_n\}$ are those

[2] It should be noted that this order is not necessarily the one of real numbers, it can be defined according to the application.

[3] By interval, we mean that all elements between $\min E$ and $\max E$ are contained in E.

whose focal sets are

$$\{x_1\},\ \{x_n\},\ \{x_1,x_2\},\ \{x_{n-1},x_n\},$$
$$\forall i=2,\dots,n-1,\ \{x_{i-1},x_i,x_{i+1}\},\ \text{and either } [x_{i-1},x_{i+2}] \text{ or } [x_i,x_{i+1}].$$

Theorem 1. *If $(\underline{F},\overline{F})$ is a p-box of the Pell family on $\mathcal{X}=\{x_1,\dots,x_n\}$, its number of extreme points is the Pell number $\mathcal{P}_n$.*

Theorem 2. *The maximal number of extreme points of a p-box defined on $\mathcal{X}=\{x_1,\dots,x_n\}$ is the Pell number $\mathcal{P}_n$, and is reached if and only if the p-box is of the Pell family.*

3.3 Counting the number of extreme points of a p-box

In this section, we provide an algorithm to enumerate the extreme points of a given p-box. This algorithm builds up a tree by incrementally assigning values v_i to vectors $\mathbf{v}\in\mathcal{V}(\text{Bel})$ as well as corresponding probability values. The ith level of the tree corresponds to v_i values, and each leaf then corresponds to a distinct extreme point (whose values can be found back by going from the leaf to the root). Pseudo-Algorithm 4 describes how children are created from a node having depth $d<n$. At a given depth d, a node is created (Loop 4-14 of Algorithm 4) for each possible number of focal elements that affect their masses to x_{d+1} (including 0), and the created node receive the corresponding probability $P(x_{d+1})$, the value v_{d+1} of the corresponding permutation vector in $\mathcal{V}$, and the update set of focal elements determining which mass remains to be distributed to which elements. The whole tree can then be built by applying this method recursively, until a depth n is reached. The root node (level 0) simply starts with $\mathcal{E}=\mathcal{F}$.

Example 2. Consider a p-box $(\underline{F},\overline{F})$ on $\{x_1,x_2,x_3,x_4\}$ whose focal sets are given by:

	$E_1=\{x_1\}$	$E_2=\{x_1,x_2,x_3\}$	$E_3=\{x_1,x_2,x_3,x_4\}$	$E_4=\{x_3,x_4\}$
m	0.2	0.1	0.4	0.3

Following Algorithm 4 and starting at the root (level 0), at the first step we have $\underline{Nb}=1,\overline{Nb}=3$, therefore the first level of the tree has three nodes (the root has three children). For $v_1=3$, $P(\{x_1\})=0.7$, the update gives $\mathcal{E}^*=E_4=\{x_3,x_4\}=0.3$, which is used to generate the node children. At the next level, only one node is generated with $v_2=0, P(\{x_2\})=0$, as $\underline{Nb}=\overline{Nb}=0$ (no focal set contains x_2), with $\mathcal{E}^*=E_4=\{x_3,x_4\}=0.3$. This node in turns generate two nodes, as $\underline{Nb}=0$ and $\overline{Nb}=1$, and so on.

Figure 1 illustrates the process in a synthetic way (as not all details are given, due to lack of space), as well as the extreme points corresponding to

Algorithm 3: Tree building algorithm

Input: Tree node with depth $d < n$ and associated set $\mathcal{E}$ of focal elements
Output: Children of node

$\underline{Nb} \leftarrow \begin{cases} 0 & \text{if } \{x_{d+1}\} \notin \mathcal{E}, \\ 1 & \text{else.} \end{cases}$;
$\overline{Nb} \leftarrow |\{E_k \in \mathcal{E} | x_{d+1} \in E_k\}|$ `/* Number of focal sets containing` x_{d+1} `*/` ;
$\underline{k} \leftarrow \inf_{E_k \in \mathcal{E}} k$;
for $i = \underline{Nb}, \ldots, \overline{Nb}$ **do**
    $P(x_{d+1}) \leftarrow \sum_{j=\underline{Nb}}^{i} m(E_{j+\underline{k}-1})$ `/*` $m(E_0) = 0$ `*/` ;
    $v_{d+1} \leftarrow i$;
    $\ell^* \leftarrow \max_\ell \{x_\ell \in E_{i+k-1}\}$ `/*` $\ell^* = d+1$ `if` E_0 `*/`;
    $\mathcal{E}^* \leftarrow \mathcal{E}$;
    foreach $E \in \mathcal{E}^*$ *such that* $x_{d+1} \in E$ **do**
        $E \leftarrow E \setminus \{x_1, \ldots, x_{\ell^*}\}$;
        if $E = \emptyset$ **then** Remove E from $\mathcal{E}^*$;
    end
    foreach $E \in \mathcal{E}^*$ *such that* $x_{d+1} \notin E$ **do**
        if $E \setminus \{x_1, \ldots, x_{\ell^*}\} \neq \emptyset$ **then** $E \leftarrow E \setminus \{x_1, \ldots, x_{\ell^*}\}$;
    end
    Create children of depth $d+1$ and associate $P(x_{d+1}), v_{d+1}, \mathcal{E}^*$ to it. ;
end

Algorithm 4: Tree building algorithm

Input: $\mathcal{E} = \{E_j^* | \text{for } j = 1, \ldots, K, E_j^* = E_j \in \mathcal{F}\}$, $d_0 = 0$, $P_0 = 0$, $v_0 = 0$
Output: Tree of n levels
`/* Tree initilization (root node) */`
$m_1 = \mathbb{1}_{\{x_1\} \in \mathcal{E}}$, $M_1 = \max\{j \in \{1, \ldots, k\} \mid x_1 \in E_j\}$;
Create Node$(m_1, M_1, \mathcal{E}, d_0, v_0, P_0, v_0, Parent = NULL)$;
while $\exists$ *Left-most unvisited node (=Lnode)* **do**
    if $d_{Lnode} = n$ **then** Mark Lnode visited `/* Leaf reached */`;
    else
        for $k = m_{Lnode}, \ldots, M_{Lnode}$ **do**
            Create Childnode(Lnode,#Lnode,$Val = k$,$\mathcal{F}$) ;
            Mark Childnode unvisited
        end
    end
    Mark Lnode visited
end

leaves of the trees. The development of the second level of the three is given only for $v_1 = 1$, to illustrate the update of $\mathcal{E}$ (Line 7 of Alg. 5).

Algorithm 5: Tree node creation algorithm

Input: Node($m, M, \mathcal{E}, d, v, P$, #ParentNode),#Node,Val,$\mathcal{F}$

Output: Childnode of Node

$d_n = d + 1, v_n =$Val, $P_n = \sum_{j=m}^{Val} m(E_j^*)$ /* n =depth of Node */ ;

$Ind = \max_i \{x_i \in E_{Val}^*\}$;

for $j = Val + 1, \ldots, K$ **do** /* Update focal sets of $\mathcal{E}$ */

  if $j \leq M$ **then** $E_j^* = E_j^* \setminus \{x_1, \ldots, x_{Ind}\}$ /* Else E_j^* unchanged */;

end

$m_n = \mathbb{1}_{\{x_{d_n}\} \in \mathcal{E}}, \quad M_n = \max_{j=Val+1,\ldots,K} \{x_{d_n} \in E_j\}$;

Create Childnode($m_n, M_n, \mathcal{E}, d_n, v_n, P_n$, #Node)

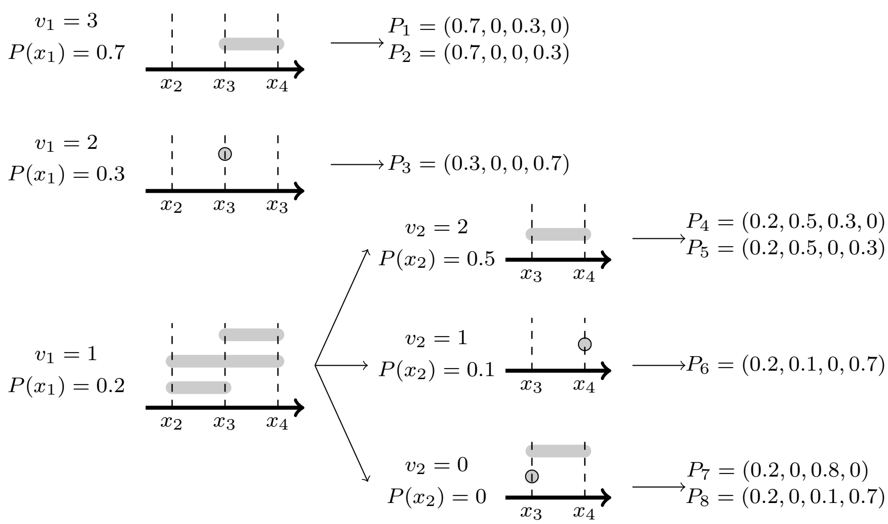

Fig. 1 Algorithm for extracting the extreme points of Example 2.

4 Conclusions

In this paper, we have characterized the maximal number of extreme points and have provided an algorithm to enumerate them through the construction of a tree structure.

There are still some interesting open problems, for instance we could try to extend our present results to the multivariate case (bivariate p-boxes) [7]. Nevertheless, this seems to be a hard problem because the connection between (univariate) p-boxes and belief functions no longer holds for bivariate p-boxes.

Acknowledgements We thank Georg Schollmeyer for insightful exchanges and discussion about this problem. The research reported in this paper has been supported by project TIN2014-59543-P and by LABEX MS2T (ANR-11-IDEX-0004-02).

References

1. Chateauneuf, A., Jaffray, J. Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of Mbius inversion. Mathematical social sciences, 17(3):263–283.
2. De Campos, L. M., Huete, J. F., Moral, S. (1994). Probability intervals: a tool for uncertain reasoning. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2(2):167–196.
3. Dempster AP (1967) Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistic 38:325–339
4. Ferson S, Kreinovich V, Ginzburg L, Myers DS, Sentz K (2003) Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015, Sandia National Laboratories
5. Kriegler E (2005) Utilizaing belief functions for the estimation of future climate change. International Journal of Approximate Reasoning 39(2-3):185–209
6. Miranda E, Couso I, Gil P (2003) Extreme points of credal sets generated by 2-alternating capacities. International Journal of Approximate Reasoning 33(1):95–115
7. Pelessoni R, Vicig P, Montes I, Miranda E (2015) Bivariate p-boxes. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, accepted
8. Schollmeyer G (2015) On the number and characterization of the extreme points of the core of necessity measures on finite spaces. In: Augustin T, Doria S, Miranda E, Quaeghebeur E (eds) Proceedings of the 9th ISIPTA
9. Troffaes MCM, Destercke S (2011) Probability boxes on totally preordered spaces for multivariate modelling. International Journal of Approximate Reasoning 52(6):767–791
10. Walley P (1991) Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London