The original publication is available at www.springerlink.com in the book “Soft Methods for Integrated Uncertainty Modeling” in the series “Advances in Soft Computing”. This book contains the proceedings (edited by J. Lawry et al.) of SMPS 2006, which was held in Bristol on 5–7 September 2006.
Cite as: E. Quaeghebeur and G. de Cooman (2006). *Extreme Lower Probabilities*, Advances in Soft Computing **6**, 211–221.

Extreme lower probabilities

Erik Quaeghebeur and Gert de Cooman

Ghent University, EESA Department, SYSTeMS Research Group
Technologiepark-Zwijnaarde 914, 9052 Zwijnaarde, Belgium
Erik.Quaeghebeur@UGent.be, Gert.deCooman@UGent.be

Summary. We consider lower probabilities on finite possibility spaces as models for the uncertainty about the state. These generalizations of classical probabilities can have some interesting properties; for example: k-monotonicity, avoiding sure loss, coherence, permutation invariance. The sets formed by all the lower probabilities satisfying zero or more of these properties are convex. We show how the extreme points and rays of these sets – the extreme lower probabilities – can be calculated and we give an illustration of our results.

Key words: Lower probabilities, extreme points, imprecise probabilities.

1 Introduction

We use and work on theories of imprecise probabilities. This means that we use concepts such as lower (and upper) probabilities to represent uncertainty. Calculating them often entails solving optimization problems. These can be hard, sometimes so hard that approximations seem the only option in practice. We are picky about the kind of approximation, however; it must be conservative. This means that the approximating lower (and upper) probabilities must be less precise than – or dominated by – the ones they approximate.

We were – and still are – aware of very few methods for making such conservative approximations. The useful ones are fewer still. One of the ideas for a new approximation approach is what led to the results communicated in this paper – which is *not* about approximations. The idea was, that perhaps we could write an arbitrary lower probability (that is hard to calculate directly) as a series of some special lower probabilities (that should be easier to calculate; breaking off the series would then constitute an approximation). The germ for this idea entered our minds when we read a paper by Maaß [10], where he mentions what in our terminology became *extreme lower probabilities*.

To get started, let us clear up some terminology and introduce the basic concepts, notation, and assumptions.

A *lower probability* $\underline{P}$ is a concept used in almost all the theories that make up the field of imprecise probabilities; it generalizes the classical concept of a probability P. Important examples of such theories are the ones described by Dempster [6] and Shafer [13], Fine [7], Walley [15], and Weichselberger [16].

Alhough the definition and interpretation of a lower probability differ between the theories, the idea is similar. Like a probability, it is a real-valued set function defined on the set $\wp(\Omega)$ of all *events* (subsets A, B, C) of a *possibility space* Ω of *states* (elements ω).[1] We are uncertain about the state.

A probability P is (i) positive: for any event A, $\underline{P}(A) \geq 0$, (ii) normed: $P(\Omega) = 1$, and (iii) additive: for any disjunct events B and C, $P(B \cup C) = P(B) + P(C)$. Similarly, a lower probability $\underline{P}$ has to satisfy some properties, but these are weaker than those for probabilities. A lower probability is usually required to be (i) normed, and (ii) super-additive: for any disjunct events B and C, $\underline{P}(B \cup C) \geq \underline{P}(B) + \underline{P}(C)$. In this paper, we do not assume a priori that a lower probability satisfies any property (not even positivity).

A probability P *dominates* a lower probability $\underline{P}$ if $\underline{P}(A) \leq P(A)$ for all events A. A probability P is called *degenerate* when it is 1 on a singleton.

The set of all lower probabilities defined on some possibility space is *convex*. This is also the case for the set of all lower probabilities additionally satisfying some interesting properties. Any closed convex set is fully determined by its *extreme points* and *extreme rays*, and vice-versa [12, Thm 18.5]: all elements of the set can be written as a linear combination of (i) the extreme rays, and (ii) a convex combination of extreme points. The extreme points and extreme rays of a convex set of lower probabilities are its *extreme lower probabilities*.

If you can see that a triangle can be described by its three vertices, you have understood the main idea behind extreme lower probabilities. Of course, we will be talking about things that are a bit more complicated than triangles.

In this paper, we restrict ourselves to finite possibility spaces Ω, of cardinality $|\Omega| = n \in \mathbb{N}$.[2] We will look at sets of lower probabilities satisfying some interesting properties; for example 2-monotonicity: for any events B and C, $\underline{P}(B \cup C) + \underline{P}(B \cap C) \geq \underline{P}(B) + \underline{P}(C)$. It turns out that for finite possibility spaces, not surprisingly, the number of extreme points is finite. We show how the extreme points can be calculated for these cases and illustrate this.

The rest of this paper is structured as follows. In the next section, we will outline the approach we took to calculating extreme lower probabilities: calculating a set of constraints, and using these to compute extreme points. Then, we look at how we can obtain *manageable* sets of constraints for the properties that interest us. Finally, before concluding, we give an illustration of our results, with comments added.

---

[1] An *upper probability* $\overline{P}$ can be defined using its so-called conjugate lower probability $\underline{P}$: for any event A, $\overline{P}(A) = 1 - \underline{P}(\Omega \setminus A)$. Because $\wp(\Omega)$ is closed under complementation, we can – and do so here – work with lower probabilities only.

[2] Notation for number sets: $\mathbb{N}$, $\mathbb{Q}$, and $\mathbb{R}$ respectively denote the nonnegative integers, the rationals, and the reals. To denote common subsets of these, we use predicates as subscripts; e.g., $\mathbb{R}_{>0} = \{r \in \mathbb{R} | r > 0\}$ denotes the strictly positive reals.

2 On constraints and vertex enumeration

A (lower) probability can satisfy a variety of interesting properties; most can be expressed using sets of linear constraints. (We do not consider things like independence and conditioning.) The introduction featured three simple examples: additivity, super-additivity, and 2-monotonicity. In general, these constraints are linear inequalities; equalities are expressed using two inequalities.

Sometimes, not all constraints are necessary: a constraint can be implied by another constraint, or a set of other constraints. A so-called *redundant* constraint can be removed from the set of constraints. If a constraint makes another one redundant, we call the former more *stringent* than the latter.

It is useful to look at this problem in a geometrical framework. Lower probabilities can be viewed as points in a 2^n-dimensional vector space – one dimension per subset in $\wp(\Omega)$. A linear inequality then generally corresponds to a half-space delimited by a hyperplane. With some property, there corresponds a set of half-spaces, and the intersection of these is the convex set of all lower probabilities satisfying that property.

The geometrical approach is illustrated in Fig. 1 below using a toy example. We consider $n = 1$, so the dimension of the vector space is 2 (this is actually the only one we can draw directly). The constraints are given using a set of hyperplanes $\{h_i | i = 1, \dots, 6\}$, the 'hairs' indicate the half-spaces the constraints correspond to. Constraints h_3 and h_6 are redundant; the former because of h_1 and h_2, the latter because h_5 is more stringent. The set of points satisfying the constraints is filled.

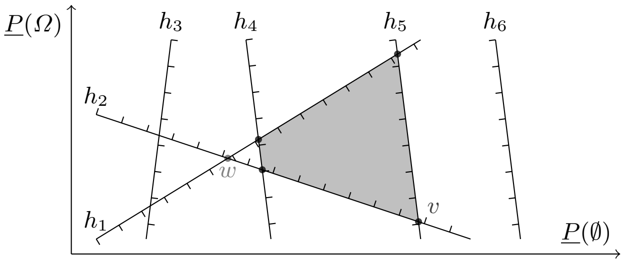

Fig. 1. Illustration of constraints and vertices.

Figure 1 also shows the vertices – v, for example – of the convex set defined by the set of half-spaces (constraints). These are the geometrical equivalent of what we call extreme lower probabilities.

In general, the *vertex enumeration problem* – finding the vertices corresponding to a given set of half-spaces – is hard: no polynomial time algorithm (in dimension, number of constraints, and number of vertices) is known [8]. To get a feeling for the complexity, realize that not only all intersections of nonredundant constraints have to be found, but we must also decide which of these to discard (such as w).

To obtain the set of constraints for different cardinalities and for the various properties we looked at, we have written our own program, `constraints` [11]. The properties in question will be described in the next section, where we will also give the theorems that form the basis of this program.

We have used publicly available programs (`redund`, `lrs`, and `cdd`) to remove redundant constraints and do vertex enumeration. They are maintained by Avis [1] and Fukuda [9].

3 From properties to constraints

Although for some properties – such as avoiding sure loss (see below) and additivity plus positivity plus normalization – it is possible to obtain the corresponding extreme points directly, we were not able to do this for the most interesting ones (k-monotonicity and coherence). For these, we use the vertex enumeration approach described in the previous section: generate the constraints corresponding to the property of interest and then use vertex enumeration to obtain the corresponding extreme lower probabilities. With this approach it is also easy to combine properties; one just has to combine the corresponding sets of constraints. The (big) downside is that it cannot be used in practice for 'large' possibility spaces – large here means $n \geq 5$.

We will now look at how we can obtain the constraints for some interesting properties. At this point we assume nothing about lower probabilities, not even that they are positive, normed, or super-additive.

Most of the results we mention in this section are either not hard to obtain or not entirely new. The most innovative part of this research was the combination of these results with vertex enumeration.

3.1 *k*-Monotonicity

In the theory of Dempster [6] and Shafer [13] lower probabilities are completely monotone. This is an extreme case of a mathematically interesting type of property: *k-monotonicity*, where $k \in \mathbb{N}_{>0}$.

A formal definition (adapted from De Cooman et al. [5]): a lower probability $\underline{P}$ is k-monotone if and only if for all $\ell = 1, \ldots, k-1$, any event A and any ℓ-tuple of events $(B_i | i \in \mathbb{N}_{<\ell})$, it holds that $\sum_{I \subseteq \mathbb{N}_{<\ell}} (-1)^{|I|} \underline{P}(A \cap \bigcap_{i \in I} B_i) \geq 0$, where the convention $\bigcap_{i \in \emptyset} B_i = \Omega$ is used. You can see that a k-monotone lower probability is also ℓ-monotone, for $\ell = 1, \ldots, k-1$. We have seen the case $k = 2$ in the introduction (in a different, but equivalent form).

Monotonicity is the same as 1-monotonicity: $\underline{P}(B) \leq \underline{P}(A)$ for any event A and all $B \subseteq A$. *Completely monotone* means k-monotone for all $k \in \mathbb{N}_{>0}$.

The above definition gives rise to a lot of redundant constraints. A lot of constraints are equivalent or are trivially satisfied. Removing them allowed us to formulate the following definition, leading to a more efficient program.

Theorem 1 (constraints for k-monotonicity). *A lower probability $\underline{P}$ is k-monotone if and only if it is monotone, and for all nonempty events A and all $\mathcal{A} \subseteq \wp(A) \setminus \{A, \emptyset\}$ such that*
(i) $0 < |\mathcal{A}| \leq k$,
(ii) $\bigcup_{B \in \mathcal{A}} B = A$, and such that
(iii) no event $C \in \mathcal{A}$ exists for which $C = \bigcap_{B \in \mathcal{A}} B$,
it holds that $\underline{P}(A) + \sum_{\mathcal{B} \subseteq \mathcal{A}} (-1)^{|\mathcal{B}|} \underline{P}(\bigcap_{B \in \mathcal{B}} B) \geq 0$.

We close this subsection with some remarks. It is a consequence of a result by Chateauneuf and Jaffray [2, Cor. 1] that k-monotonicity for some $k \geq n$ is equivalent to complete monotonicity. Note that $\underline{P}(\emptyset) = 0$ and $\underline{P}(\Omega) = 1$ are not included in k-monotonicity; they are commonly added, however.

3.2 Avoiding sure loss

Avoiding sure loss is a property that is useful in the behavioral theory of Walley [15, §2.4.1].[3] A lower probability $\underline{P}$ avoids sure loss if and only if for all $\mathcal{B} \subseteq \wp(\Omega)$, and all $\lambda \in (\mathbb{R}_{\geq 0})^{\mathcal{B}}$, it holds that $\sum_{B \in \mathcal{B}} \lambda_B \underline{P}(B) \leq \sup \sum_{B \in \mathcal{B}} \lambda_B I_B$, where I_B is the indicator function of B, which is 1 for $\omega \in B$ and 0 elsewhere.

If we only require a lower probability to avoid sure loss, the extreme lower probabilities can be determined by reasoning. Because a lower probability avoids sure loss if and only if it is dominated by a probability [15, §3.3.3], the set of extreme lower probabilities consists of the degenerate probabilities as extreme points and all negative main directions as extreme rays.

If we want to use vertex enumeration to obtain the extreme lower probabilities, the problem arises that the definition gives an infinite number of constraints (because λ is real-valued). This is of course unmanageable for any computer. This situation is inevitable when the usual assumption of positivity is made and the extreme lower probabilities cannot be determined by reasoning.

Luckily – by removing redundant constraints –, we can reduce the set of constraints for avoiding sure loss to a finite set. It was shown by Walley [15, §A.3] that in the definition, to get the most stringent constraints, we only need to consider (i) $\mathcal{B}$ such that $\{I_B | B \in \mathcal{B}\}$ is a linearly independent set and $\bigcup_{B \in \mathcal{B}} B = \Omega$, and (ii) λ such that the function on right hand side is the constant function 1.

Walley [15, §A.3] assumes positivity ($\underline{P}$ is in the first orthant). When we want to do the vertex enumeration approach without this assumption (as we do here), we need to add extra constraints for the cases where some (or all) of the components of $\underline{P}$ are negative. This can be done by taking every original constraint, and creating a new one for each of the 2^{2^n} possible orthants $\underline{P}$ can be located in. We do this by setting the λ_B for which $\underline{P}(B) < 0$ to 0.

The above and some other, minor, changes result in the following theorem.

[3] In Walley's theory [15], *lower previsions* – expectation operators – play a central role. Here, we restrict ourselves to the less general lower probabilities.

Theorem 2 (constraints for avoiding sure loss). *A lower probability* $\underline{P}$ *avoids sure loss if and only if* $\underline{P}(\emptyset) \leq 0$ *and for all*

(*i*) $\mathcal{B}$ *such that* $\bigcup_{B\in\mathcal{B}} B = \Omega$ *and* $\{I_B | B \in \mathcal{B}\}$ *is a linearly independent set,*
(*ii*) $\lambda \in (\mathbb{Q} \cap]0,1])^{\mathcal{B}}$ *such that* $\sum_{B\in\mathcal{B}} \lambda_B I_B = 1$, *and*
(*iii*) *binary masks* $\beta \in \{0,1\}^{\mathcal{B}}$,

it holds that $\sum_{B\in\mathcal{B}} \beta_B \lambda_B \underline{P}(B) \leq 1$.

3.3 Coherence

Coherent lower probabilities are at the core of Walley's theory [15, §2.6.4]. Weichselberger [16] uses the term F-probability for the same concept. A lower probability $\underline{P}$ is coherent if and only if for all events A, all $\mathcal{B} \subseteq \wp(\Omega) \setminus \{A\}$, and all $\lambda \in (\mathbb{R}_{\geq 0})^{\{A\}\cup\mathcal{B}}$, it holds that $\sum_{B\in\mathcal{B}} \lambda_B \underline{P}(B) - \lambda_A \underline{P}(A) \leq \sup(\sum_{B\in\mathcal{B}} \lambda_B I_B - \lambda_A I_A)$. Coherence implies avoiding sure loss and monotonicity, but not 2-monotonicity.

The above definition is similar enough to the one for avoiding sure loss to allow the same techniques for the removal of redundant constraints to be applied, up to some technicalities. Because coherence implies positivity, we need not use binary masks. Working this out results in the following theorem.

Theorem 3 (constraints for coherence). *A lower probability* $\underline{P}$ *is coherent if and only if* $\underline{P}(\emptyset) = 0$, $\underline{P}(\Omega) = 1$, *and*

(*a*) *for all events* A *it holds that* $0 \leq \underline{P}(A) \leq 1$;
(*b*) *for all*
- (*i*) *events* A *and* $\mathcal{B} \subseteq \wp(\Omega) \setminus \{A\}$ *such that* $\{I_B | B \in \mathcal{B}\}$ *is a linearly independent set and* $\bigcup_{B\in\mathcal{B}} B = A$, *and*
- (*ii*) $\lambda \in (\mathbb{Q} \cap]0,1])^{\mathcal{B}}$ *such that* $\sum_{B\in\mathcal{B}} \lambda_B I_B = I_A$,

it must hold that $\sum_{B\in\mathcal{B}} \lambda_B \underline{P}(B) \leq \underline{P}(A)$;

(*c*) *for all*
- (*i*) *events* A *and* $\mathcal{B} \subseteq \wp(\Omega) \setminus \{A\}$ *such that* $\{I_A\} \cup \{I_B | B \in \mathcal{B}\}$ *is a linearly independent set and* $\bigcup_{B\in\mathcal{B}} B = \Omega$, *and*
- (*ii*) $\lambda \in (\mathbb{Q}_{>0})^{\{A\}\cup\mathcal{B}}$ *such that* $\sum_{B\in\mathcal{B}} \lambda_B I_B - \lambda_A I_A = 1$,

it must hold that $\sum_{B\in\mathcal{B}} \lambda_B \underline{P}(B) - \lambda_A \underline{P}(A) \leq 1$.

3.4 Permutation invariance

As a last property, we look at (weak) *permutation invariance* [4]. It is the odd duck of the lot; whereas all the previous properties allow for lower probabilities that express a quite broad a range of uncertainty models, permutation invariance restricts them to some very specific ones. We mention it to show how easy it can be to add the constraints for another property.

A lower probability is invariant under permutations of the elements of the possibility space if and only if, for any event A and all events B resulting from some permutation, $\underline{P}(A) = \underline{P}(B)$. Let us give an example for $n = 3$: consider the permutation $(1 \to 3, 2 \to 1, 3 \to 2)$, then $A = \{1, 2\}$ becomes $B = \{1, 3\}$.

We can characterize permutation invariance as follows [16, §4.3.1].

Theorem 4 (constraints for permutation invariance). *A lower probability $\underline{P}$ is permutation invariant if and only if for all $k = 1, \ldots, n-1$, any one event A such that $|A| = k$, and for all other events B with $|B| = k$, it holds that $\underline{P}(B) = \underline{P}(A)$.*

4 Results

We are not the first to hunt for extreme lower probabilities. However, as far as we know, we are as of yet the most systematic. ([11] contains a list of results.)

For $n = 4$ Shapley [14] gives a list with 37 of the 41 – he omits the 4 degenerate probabilities – 2-monotone extreme lower probabilities. For $n = 5$, we have found all 117 983 for this case. For $n \leq 4$, we have found the extreme (permutation invariant) k-monotone lower probabilities for all k.

In an example, Maaß [10] mentions the 8 extreme coherent lower probabilities for $n = 3$. We give a graphical representation of them in Fig. 2 using their corresponding credal sets. Let us clarify: The set of all probabilities that dominate some lower probability is called its *credal set* (cfr. *core* in game theory [14]). All probabilities can be represented as a point of the $(n-1)$-dimensional unit simplex – a regular triangle for $n = 3$ – and so coherent lower probabilities can be represented by their credal set, which is a convex subset of this unit simplex [15, §3.6.1]. The vertices of the simplex correspond to the degenerate probabilities, so to the states – here a, b, and c.

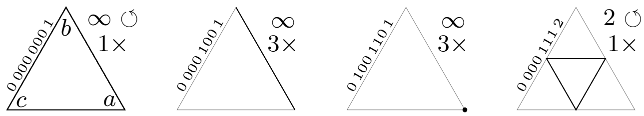

Fig. 2. The credal sets of the 8 extreme coherent lower probabilities for $n = 3$.

In Fig. 2, we only give the border of the credal sets and show only one element of each permutation class. At the top right, we indicate the following: the total number of elements of the permutation class (e.g., 4×), if it is permutation invariant (↻), and k-monotonicity for $k \in \mathbb{N}_{>1}$ (∞ for complete monotonicity, 2 for 2-monotonicity). Along the simplex's left edge, we give a vector that is proportional to the extreme coherent lower probability (component order: $\emptyset$ $\{a\}\{b\}\{c\}$ $\{a, b\}\{a, c\}\{b, c\}$ Ω). Remember that $\underline{P}(\Omega) = 1$.

As convex combinations preserve monotonicity, we can immediately see from Fig. 2 that for $n = 3$, all coherent lower probabilities are 2-monotone. This was already known, but it illustrates how these computational results can help in finding theoretical results. (For $n = 2$, all are completely monotone.)

Once we implemented our program, finding all 402 extreme coherent lower probabilities for $n = 4$ was easy.[4] Figure 3 shows the corresponding credal sets, as well as those for the 16 extreme 3-monotone lower probabilities, in the same way we did for $n = 3$ in Fig. 2. The unit simplex is now a regular tetrahedron, so we had to use a projection from three dimensions to two.

Take $\Omega = \{a, b, c, d\}$. In this case, the component order is $\emptyset$ $\{a\}\{b\}\{c\}\{d\}$ $\{a, b\}\{a, c\}\{a, d\}\{b, c\}\{b, d\}\{c, d\}$ $\{a, b, c\}\{a, b, d\}\{a, c, d\}\{b, c, d\}$ Ω.

With Fig. 3 as a guide, we can give some observations and results.

Using results from Choquet [3, Ch. 7], it can be proven that the extreme completely monotone lower probabilities are the *vacuous* lower probabilities (cfr. *unanimity games* in game theory) with respect to events: $\underline{P}$ is vacuous with respect to A if $\underline{P}(B) = 1$ for $A \subseteq B$ and 0 otherwise.[5] These correspond to the first three (Fig. 2) and first four (Fig. 3) permutation classes shown. The last of these classes corresponds to the degenerate probabilities, which are all the extreme (classical) probabilities.

We have observed that the extreme completely monotone probabilities are always included in the extreme coherent proabilities. This is not the case for all of the extreme k-monotone and permutation invariant lower probabilities. An example for $n = 4$ is the only non-completely monotone lower probability of the extreme 3-monotone lower probabilities (shown in Fig. 3).

Notice that, except for the degenerate probabilities, all credal sets touch all tetrahedron faces. This is so because it can be shown that the degenerate probabilities are the only extreme coherent lower probabilities that can be nonzero in singletons.

5 Conclusions

Although we have not intensively pursued our initial goal – finding conservative approximations for lower and upper probabilities –, it did lead us to this interesting research. Obtaining sets of extreme lower probabilities for many cases and formulating a systematic approach to calculating them are the main results of the research presented in this paper.

Apart from these results, this topic also allows one to become familiar with different ways of looking at lower probabilities and their properties. A lower probability can be seen as a set function satisfying some properties, as a convex combination of some special set functions, as a point of a convex subset of a vector space, and – for coherent ones – as a credal set. They can be k-monotone, avoid sure loss, be coherent, be permutation invariant, etc.

And last but not least, this topic can lead to beautiful figures.

[4] For $n = 5$, we know 1 743 093 of the extreme coherent lower probabilities. These have been found by a computer in our lab, after some months of vertex enumeration. A hardware failure cut this gargantuan task short.

[5] In general (Ω can be infinite), the extreme completely monotone lower probabilities are those that take the value 1 on a proper filter of sets and are 0 elsewhere.

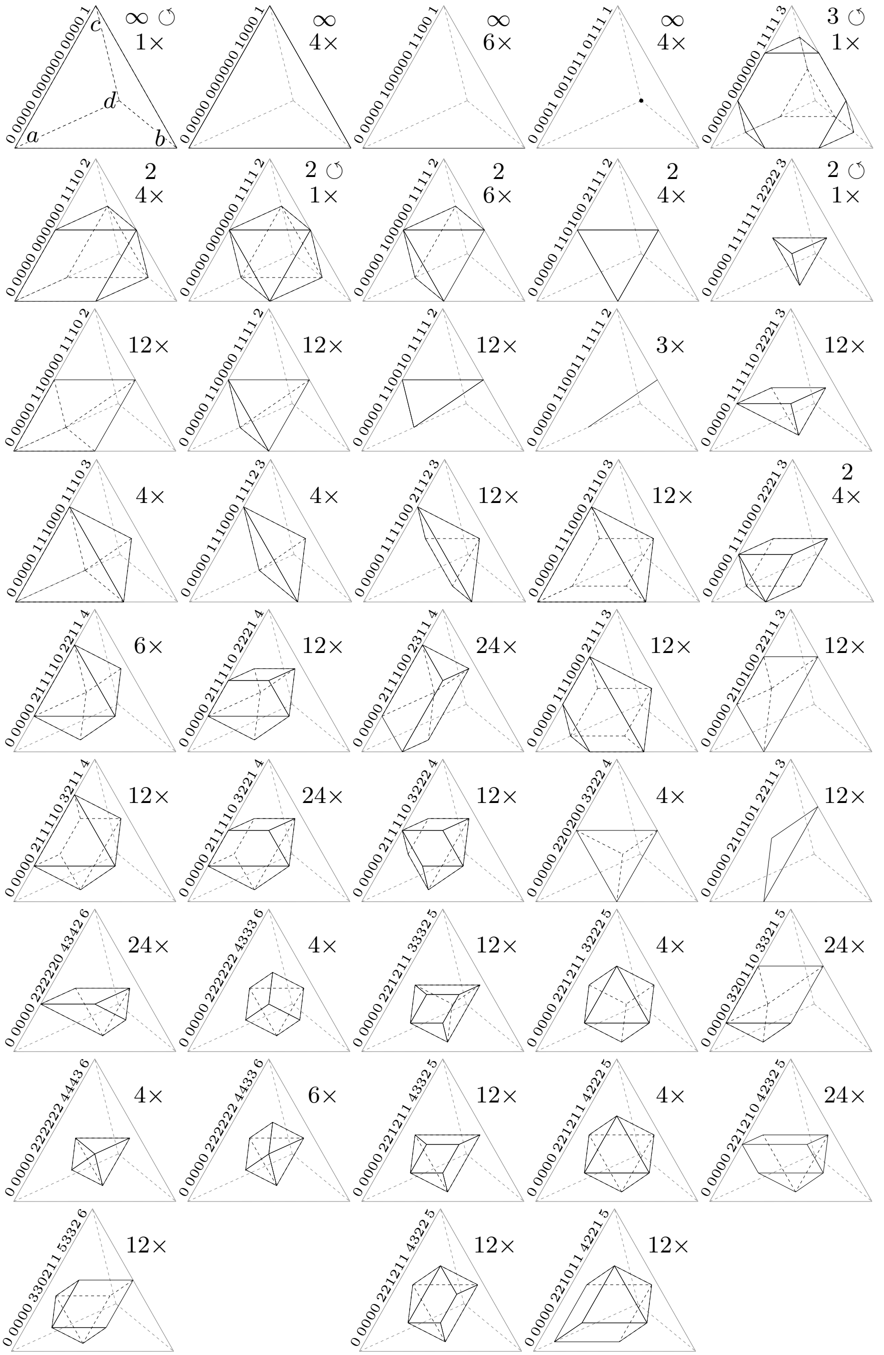

Fig. 3. The credal sets for $n = 4$ of the extreme coherent lower probabilities (all except last of top row) and extreme 3-monotone lower probabilities (top row).

Acknowledgements

We wish to thank M. Troffaes, S. Maaß, A. Wallner, and E. Miranda for many a useful discussion and two referees for useful suggestions. These led to a much clearer view on the subject.

This paper presents research results of the Belgian Program on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors. Erik Quaeghebeur's research is financed by a Ph.D. grant from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

References

1. Avis, D. (2005). `lrs` homepage. `http://cgm.cs.mcgill.ca/~avis/C/lrs.html`.
2. Chateauneuf, A. and Jaffray, J.-Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Math. Social Sci.* **17**, 263–283.
3. Choquet, G. (1954). Theory of Capacities. *Ann. Inst. Fourier* **5**, 131–295.
4. De Cooman, G. and Miranda, E. (2006). Symmetry of models versus models of symmetry. In: Harper, W. L. and Wheeler, G. R. (Editors) *Probability and Inference: Essays in Honor of Henry E. Keyburg, Jr.* King's College Publications.
5. De Cooman, G., Troffaes, M. and Miranda, E. (2005). n-Monotone lower previsions and lower integrals. In: Cozman, F. G., Nau, R. and Seidenfeld, T. (Editors) *ISIPTA '05: Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications*, pp. 145–154.
6. Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.* **38**, 325–339.
7. Fine, T. L. (1973). *Theories of probability: An examination of foundations.* Academic Press.
8. Fukuda, K. (2004). Frequently Asked Questions in Polyhedral Computation. `http://www.ifor.math.ethz.ch/~fukuda/polyfaq/`.
9. Fukuda, K. (2005). `cdd` and `cddplus` homepage. `http://www.ifor.math.ethz.ch/~fukuda/cdd_home/cdd.html`.
10. Maaß, S. (2003). Continuous Linear Representations of Coherent Lower Previsions. In: Bernard, J.-M., Seidenfeld, T. and Zaffalon, M. (Editors) *ISIPTA '03: Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, pp. 372–382.
11. Quaeghebeur, E. (2006). `constraints` homepage. `http://users.UGent.be/~equaeghe/constraints.php`.
12. Rockafellar, R. T. (1970). *Convex Analysis.* Princeton Univ. Press.
13. Shafer, G. (1976). *A mathematical theory of evidence.* Princeton Univ. Press.
14. Shapley, L. S. (1971). Cores of Convex Games. *Int. J. Game Theory* **1**, 11–26.
15. Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall.
16. Weichselberger, K. (2001). *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung*, vol. I: Intervallwarscheinlichkeit als umfassendes Konzept. Physica-Verlag.