



LUDWIG-  
MAXIMILIANS-  
UNIVERSITÄT  
MÜNCHEN

INSTITUT FÜR STATISTIK



Christoph Jansen, Georg Schollmeyer and Thomas Augustin

# Concepts for Decision Making under Severe Uncertainty with Partial Ordinal and Partial Cardinal Preferences

Technical Report Number 216, 2018  
Department of Statistics  
University of Munich

<http://www.statistik.uni-muenchen.de>



# Concepts for Decision Making under Severe Uncertainty with Partial Ordinal and Partial Cardinal Preferences

C. Jansen<sup>a,\*</sup>, G. Schollmeyer<sup>a</sup>, T. Augustin<sup>a</sup>

<sup>a</sup>*Department of Statistics, LMU Munich, Ludwigsstr. 33, 80539 Munich, Germany*

---

## Abstract

We introduce three different approaches for decision making under uncertainty if (I) there is only partial (both cardinally and ordinally scaled) information on an agent's preferences and (II) the uncertainty about the states of nature is described by a credal set (or some other imprecise probabilistic model). Particularly, situation (I) is modeled by a pair of binary relations, one specifying the partial rank order of the alternatives and the other modeling partial information on the strength of preference. Our first approach relies on decision criteria constructing complete rankings of the available acts that are based on generalized expectation intervals. Subsequently, we introduce different concepts of global admissibility that construct partial orders between the available acts by comparing them all simultaneously. Finally, we define criteria induced by suitable binary relations on the set of acts and, therefore, can be understood as concepts of local admissibility. For certain criteria, we provide linear programming based algorithms for checking optimality/admissibility of acts. Additionally, the paper includes a discussion of a prototypical situation by means of a toy example.

*Keywords:* partial preferences; decision making under uncertainty; ordinality; cardinality; utility representation; imprecise probabilities; credal set; stochastic dominance; admissibility; linear programming.

---

## 1. Introduction

One of the constantly recurring topics discussed in the community of researchers working with imprecise probabilities (and on ISIPTA conferences in particular) is defining meaningful criteria for decision making under complex uncertainty, finding persuading axiomatic justifications for these criteria and providing efficient algorithms capable to deal with them. Examples for such works are ranging from rather early IJAR and ISIPTA contributions by, e.g., Jaffray (1999); Augustin (2001); Schervish et al. (2003); Smets (2005) to more recent ones by, e.g., Troffaes (2007); Utkin and Augustin (2007); Jaffray and Jeleva (2008); Loquin (2013); Bradley (2015).

However, in the vast majority of works in this field, the complexity underlying the decision situation is assumed to solely arise from the fact that the decision maker's beliefs on the mechanism

---

<sup>\*</sup>A preliminary version of this paper was presented at the *Tenth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA)*, 10-14 July 2017 in Lugano, see Jansen et al. (2017b).

\*Corresponding author

*Email address:* christoph.jansen@stat.uni-muenchen.de (C. Jansen)

generating the states of nature are expressed by an imprecise probabilistic model. In contrast, the cardinal utility function adequately describing the decision maker’s preference structure is often unquestioned and assumed to be precisely given in advance.<sup>1</sup> Unfortunately, also this can be problematic. Wrongfully pretending to have perfect information on the level of utilities might lead to bad decision making just as doing the same on the level of beliefs: What’s worth a decision that is derived on the basis of an inadequate utility function?

For this reason, our paper generalizes both the classical setting of decision making under risk as well as the generalized setting of decision making under ambiguity to situations in which the assumption of a known cardinal utility structure is no longer justified. Particularly, we consider the case that the (information on the) decision maker’s preference structure is both of partially ordinal and of partially cardinal scale and, therefore, no longer can be characterized by (a set of positive linear transformations of) *one* cardinal utility function. Instead, we model the decision maker’s utility by the set of all utility representations that are compatible with both the ordinal and the cardinal information concerning her preferences.

The paper is structured as follows: In Section 2, we give a brief overview on the background of our work and show how our approach naturally fits into this picture. Moreover, we discuss related literature and discuss the connections to our work. In Section 3, we introduce the crucial concept of a *preference system* over a set of alternatives that allows for modeling partially ordinal and partially cardinal preference structures. Section 4 introduces three different approaches for decision making with acts taking values in a preference system by proposing decision criteria based on generalized expectation intervals (Section 4.2), on global comparisons of acts (Section 4.3) and on pairwise comparisons of acts (Section 4.4). For certain criteria, we give linear programming driven algorithms for checking feasibility of acts in finite decision settings. Section 5 is devoted to an application of the theory. There, we illustrate all the concepts developed in the paper in an example and thereby also show a class of situations in which our approach seems natural: The case where the consequences that acts can attain belong to some product space with both ordinal and cardinal dimensions. Section 6 concludes the paper.

## 2. Fundamentals underlying our Approach and Related Literature

In classical subjective expected utility theory (SEUT), the decision maker (synonymously called agent in the following) is assumed to be able to specify (I) a real-valued *cardinal* utility function  $u$  (unique up to a positive linear transformation) representing his preferences on a set  $A$  of alternatives and (II) a unique and *precise* subjective probability measure  $\pi$  on the space  $S$  of states of nature adequately specifying his beliefs on the occurrence of the different states  $s \in S$ . Once these two ingredients are specified, according to SEUT, the decision maker should choose any act  $X : S \rightarrow A$  that maximizes the expected utility  $\mathbb{E}_\pi(u \circ X)$  with respect to his utility function  $u$  and his subjective probability measure  $\pi$  among all other available acts.

However, as is well known, in practice both assumptions (I) and (II) often turn out to be systematically too restrictive. In particular, (I) demands the decision maker to act in accordance with the axioms of von Neumann and Morgenstern, i.e. to be able to specify a complete preference

---

<sup>1</sup>Exceptions include Montes (2014, Section 4.2.1), who uses set-valued utility functions, Landes (2014) who axiomatically characterizes preferences over utility intervals and Troffaes and Sahlin (2017), who propose elicitation procedures for partially specified utility functions. These references, among others, are discussed in some more detail at the end of Section 2.

ranking of all simple lotteries on the set  $A$  that is both independent and continuous (see, e.g., Fishburn (1970, Ch. 8) for details), whereas (II) requires that the decision maker can completely order the resulting utility-valued acts by preference in accordance with the axioms of de Finetti, i.e. continuous, additive and monotone (see, e.g., Gilboa (2009, Ch. 9) for details).

Consequently, there exists plenty of literature relaxing these assumptions. If *only* (II) is violated in the sense that there is only *partial* probabilistic information on the occurrence of the states of nature together with a perfectly cardinal preference structure (represented by a cardinal utility function  $u$ ), the common relaxation is to allow for *imprecise* probabilistic models for representing the probabilistic information (for instance one could use the credal set  $\mathcal{M}$  of all probability measures that are compatible with the given probability constraints). In this case, one can define optimality of acts, for instance depending on the attitude of the decision maker towards the ambiguity underlying the situation, in terms of some imprecise decision criterion such as:

- $\Gamma$ -*maximin* ( $\Gamma$ -*maximax*): Choose any arbitrary act  $X$  yielding maximal expected utility with respect to the worst (best) compatible probability measure, i.e. that maximizes the value  $\inf_{\pi \in \mathcal{M}} \mathbb{E}_{\pi}(u \circ X)$  (the value  $\sup_{\pi \in \mathcal{M}} \mathbb{E}_{\pi}(u \circ X)$ ) among all available acts.
- *Maximality*: Dismiss each act  $X$  for which there is available another act  $Y$  that dominates it in expectation with respect to all compatible probability measures, i.e. for which it holds that  $\mathbb{E}_{\pi}(u \circ X) < \mathbb{E}_{\pi}(u \circ Y)$  for all  $\pi \in \mathcal{M}$ .
- *E-admissibility*: Dismiss each act  $X$  that does not maximize expected utility  $\mathbb{E}_{\pi}(u \circ X)$  among the available acts with respect to at least one compatible probability measure  $\pi \in \mathcal{M}$ , i.e. where for all  $\pi \in \mathcal{M}$  there exists an act  $Y_{\pi}$  with  $\mathbb{E}_{\pi}(u \circ X) < \mathbb{E}_{\pi}(u \circ Y_{\pi})$ .

The original sources of the criteria just discussed are given in Kofler and Menges (1976); Levi (1974, 1983); Gilboa and Schmeidler (1989); Walley (1991). Further criteria for the case of cardinal utility and imprecise probabilities, each in its own way taking into account the whole set  $\mathcal{M}$  of compatible probability measures, are reviewed in, e.g., Huntley et al. (2014). Additionally, there exists a variety of efficient and powerful algorithms to deal with this kind of violation of the classical assumptions (see, e.g., Utkin and Augustin (2005); Kikuti et al. (2011); Hable and Troffaes (2014); Jansen et al. (2017a)). However, note that the assumption of a cardinal utility function  $u$  is essential for all these criteria in order to be applicable. The situations where these imprecise decision criteria are appropriate are illustrated as one of the horizontal edges in the schematic cube given in Figure 1.

If it is the case that (I) is violated in the sense that the decision maker has complete *ordinal* preferences but no cardinal information (for instance in form of a complete and transitive binary relation  $R$  on  $A$ ) and (II) is violated in the sense that there is no probabilistic information at all, it is nearly unanimously favored to define optimality of acts based on Wald's classical maximin criterion (see Wald (1949) for details), which reads here as

- *Maximin criterion*: Choose any act  $X$  receiving highest possible rank under the worst possible state of nature, i.e. with  $(\inf_{s \in S} X(s), \inf_{s \in S} Y(s)) \in R$  for every other available act  $Y$ .

However, note that the completeness of the involved ordinal ranking is essential, since, otherwise, the worst consequences of two distinct acts might be incomparable and, therefore, an optimal act with respect to the maximin criterion simply does not exist. Even more severe, also the vacuousness assumption concerning the information on probabilities is crucial: Applying the maximin criterion in the presence of (partial) probabilistic information means willingly ignoring information. This

seems not reasonable at all (cf. also Example 1 for an illustration). Situations in which Wald's maximin criterion appears to be appropriate are illustrated as one of the vertical edges in the schematic cube given in Figure 1.

Finally, if *only* (I) is violated in the sense that there is no cardinal information at all and the available ordinal information is possibly incomplete (meaning that the relation  $R$  from above might not be complete), however, (II) holds true in the sense that beliefs about the states can be described by a precise probability measure  $\pi$ , one common criterion to be applied is the concept of

- (*First order*) *Stochastic dominance*: An act  $Y$  is said to weakly stochastically dominate another act  $X$ , if it holds that  $\mathbb{E}_\pi(u \circ Y) \geq \mathbb{E}_\pi(u \circ X)$  for every measurable utility function  $u$  that is monotone with respect to  $R$  (meaning  $(a, b) \in R$  implies  $u(a) \geq u(b)$  for arbitrary  $a, b \in A$ ). We denote the stochastic dominance relation by  $\geq_{SD}$ . This induces the following choice rule: Dismiss an act  $X$  if it is strictly stochastically dominated by another available act, i.e. if there exists an act  $Y$  such that  $\mathbb{E}_\pi(u \circ Y) \geq \mathbb{E}_\pi(u \circ X)$  for every measurable utility function  $u$  that is monotone with respect to  $R$  and such that  $\mathbb{E}_\pi(u_0 \circ Y) > \mathbb{E}_\pi(u_0 \circ X)$  for at least one such function  $u_0$ .

Also for this case there exist well-established theory as well as efficient algorithms for computation (see, e.g. Lehmann (1955); Kamae et al. (1977); Mosler and Scarsini (1991); Tarp and Osterdal (2007); Range and Osterdal (2013); Schollmeyer et al. (2017)). Situations for which first order stochastic dominance should be the decision criterion of choice are indicated as one of the horizontal edges in the schematic cube drawn in Figure 1.

Further, note that there exists a fair amount of work on generalizing the notion of first order stochastic dominance to situations where the underlying uncertainty is characterized by a credal set of probability measures or situations where the utility function is only partially specified in terms of a multi-valued mapping: In Montes et al. (2014a, Section 5) the authors introduce and study a generalization where an act  $X$  is said to dominate another act  $Y$  whenever it stochastically dominates it with respect to every distribution from the underlying credal set. Moreover, in Montes et al. (2014a, Sections 3 and 4) the authors study different possibilities to extend the notion of stochastic dominance (with respect to a precise probability measure) to suitable binary relations on sets of random variables. These approaches can be viewed as generalizations of stochastic dominance to imprecise utilities. In Montes et al. (2014b) this framework is further investigated and applied to a real world example. The interrelations of these works and the present one are discussed in some more detail at the end of Section 4.4 when the required concepts are formulated.

Beyond the connections to the literature already mentioned, other related work exists: In Denoeux (2009) the author also studies generalizations of stochastic dominance to imprecise probabilistic models, however, for the special case of belief functions on the real line. More precisely, the paper studies how different orderings between intervals on the real line induce different orderings between belief functions and the mass function associated with them. Afterwards, the paper studies how these orderings relate to the notion of stochastic dominance. Some more details on the connection between Denoeux (2009) and the present work are provided in the discussion directly following the proof of Proposition 6 in Section 4.4.

Of course, there is also related work on non-fully specified utilities: In Landes (2014) the author considers the situation of decision making under complete uncertainty (i.e. with a credal set  $\mathcal{M}$  containing all possible probability distributions on the state space) with acts taking values in some *linearly* ordered space. To each such act it is then associated an utility interval. Afterwards, the author axiomatically characterizes desirable properties of binary relation on such utility intervals

and shows that these axioms uniquely determine a particular binary relation, the so-called Min-Max Relation (see Landes (2014, Corollary 3) in particular).

In Troffaes and Sahlin (2017) the authors consider the case of decision making with acts taking values in some multi-attribute space. For such problems they propose a two-step elicitation procedure for utility functions, where the first step consists in precisely eliciting the marginal utility functions on the different attributes and the second step consists in imprecisely eliciting the weights with respect to which these marginal utilities are extended to an utility function on the whole multi-attribute space.

Finally, in Danielson and Ekenberg (1998) and Danielson et al. (2003) and Danielson (2005) the authors study decision making problems where both the utility values assigned to the consequences and the probability values assigned to the states are allowed to be only imprecisely known. In particular, they investigate situations where the statements about probability and utility values can be formed by one of three types of sentences, so-called vague sentences, interval sentences and comparative sentences. The set of available acts together with the sets of probability and utility sentences then forms the so-called information frame. For such decision problems they then propose a decision criterion, the so-called  $t$ -admissibility, which relies on the idea to prefer an act  $X$  to another act  $Y$  whenever it holds that  $\mathbb{E}_\pi(u \circ X) - \mathbb{E}_\pi(u \circ Y) \geq t$  for all pairs  $(\pi, u)$  that are compatible with both the set of utility sentences and the set of probability sentences, that is with the information frame (see Danielson and Ekenberg (1998, Section 2.2)). For evaluating this criterion the authors then propose bilinear as well as linear optimization approaches and apply their theory to a real world problem.

Furthermore, in Danielson (2005) the framework developed in Danielson and Ekenberg (1998) is generalized to the case where the pairwise comparison between the acts  $X$  and  $Y$  is no longer made by considering solely the differences of the expected values, but where it can be made by arbitrary functionals  $f$  depending besides of the acts under consideration also on  $\pi$  and  $u$  as well as additional parameters (see Danielson (2005, Section 3)). Their work is also implemented in the decision user interface *DecideIt*, which is introduced and described in Danielson et al. (2003). Note that our relation  $R_{\forall\forall}$  that is introduced in Equation (9) and our concept of local  $R_{\forall\forall}$ -admissibility from Definition 8 that is based on it are closely related to the concept of  $t$ -admissibility from the works by Danielson et al.. This is discussed in some more detail at the end of Section 4.4.

In the following sections of the paper, we introduce and discuss different concepts for decision making in situations in which simultaneously both assumptions (I) and (II) are violated (i.e. situations corresponding to inner points of the gray-shaded prism from Figure 1) and thus none of the concepts just recalled can be applied. Therefore, the contribution of the present paper consists in filling up the gray-shaded prism in the schematic cube drawn in Figure 1.

### 3. Preference Systems

In this section we start by defining the concept of a preference system, which is essential for what follows throughout the rest of the paper. The intuition behind this concept is very simple: In many practically relevant decision problems, the (available information on the) agent's preferences are (is) incomplete. More precisely, it often will be the case that certain pairs of possible decision outcomes are incomparable for the agent, whereas others can be ordered by preference. Additionally, for some pairs there might even be an idea of the strength of the preference, that is an idea of how much the one outcome is preferred to the other.

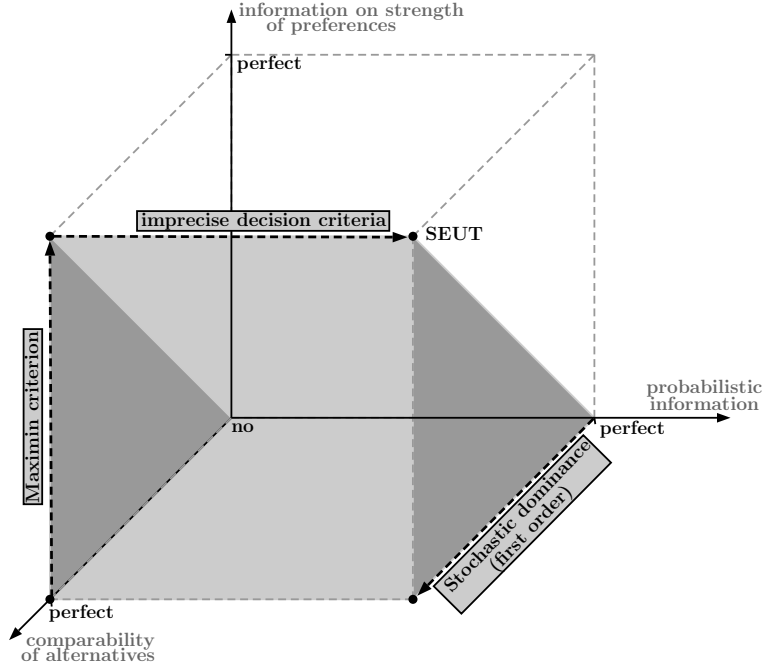


Figure 1: Schematic illustration of the approaches recalled in Section 2: The points on the black dotted arrows indicate situations in which classical criteria exist, whereas the points in the filled gray-shaded prism indicate situations that are captured by our approach.

There are several circumstances that could give rise to this type of incomplete preferences. For example, if a company wants to analyze the choice behavior of their (potential) customers, the information on the customer’s preferences will often be given in form of observed binary choices and/or survey data. Obviously, usually such data won’t be sufficient to specify the full preference structure of the customer, since this requires too many observations. In this case, incompleteness is a missing data problem and originates in lacking information about the choice behavior.

However, also the agent herself might have incomplete preferences. Suppose she knows (e.g. from earlier choice experience) certain potential decision outcomes better than others. Then for pairs involving better known outcomes, she might be able to specify a preference ranking and even some intuition for the strength of the preference, whereas for pairs involving unfamiliar outcomes, she might be able to specify only a ranking or even can’t make a comparison at all. The following definition captures the intuition just described.

**Definition 1.** Let  $A$  be a non-empty set and let  $R_1 \subseteq A \times A$  denote a pre-order (i.e. reflexive and transitive) on  $A$ . Moreover, let  $R_2 \subseteq R_1 \times R_1$  denote a pre-order on  $R_1$ . Then the triplet  $\mathcal{A} = [A, R_1, R_2]$  is called a **preference system** on  $A$ .

The interpretation of the relations  $R_1$  and  $R_2$  contained in a preference system  $\mathcal{A}$  in the sense of Definition 1 is in perfect accordance with the intuition that should be captured by it: If some pair  $(a, b) \in R_1$ , we interpret that as  $a$  is at least as desirable as  $b$ , that is  $a$  and  $b$  can be ordered by preference. If both  $(a, b) \notin R_1$  and  $(b, a) \notin R_1$ , then  $a$  and  $b$  are incomparable. Moreover, if a pair of pairs  $((a, b), (c, d)) \in R_2$ , we interpret this as exchanging alternative  $b$  by alternative  $a$  is at least as desirable as exchanging alternative  $d$  by alternative  $c$ , that is  $a$  is more strongly preferred over

$b$  than  $c$  is over  $d$ . Again, if both  $((a, b), (c, d)) \notin R_2$  and  $((c, d), (a, b)) \notin R_2$ , then the exchange of  $b$  by  $a$  is incomparable with the exchange of  $d$  by  $c$ .

Except from transitivity, Definition 1 makes no rationality and/or compatibility assumption regarding the relations  $R_1$  and  $R_2$ . Accordingly, a preference system in the sense of the above Definition 1 needs by no means to be reasonable or rational. In Krantz et al. (1971, Chapter 4), an axiomatic approach for characterizing consistent preference systems is provided for the case that the involved relations are complete. The corresponding axioms then imply the existence of a real valued function representing both relations simultaneously that is unique up to a positive linear transformation. Another axiomatization that uses quaternary relations instead of pairs of relations is established in Pivato (2013), where it is shown that under some quite strong conditions (like, e.g., solvability) there exists a multi-utility characterization of the corresponding quaternary relation.

A weaker consistency condition that still applies to settings in which conditions like solvability no longer can be expected is given in the following definition, for which we need some further notation: If  $R$  is a pre-order on  $A$ , we denote by  $I_R$  and  $P_R$  its indifference and its strict part, respectively. More precisely, for  $(a, b) \in A \times A$ , we have  $(a, b) \in I_R \Leftrightarrow ((a, b) \in R \wedge (b, a) \in R)$  and  $(a, b) \in P_R \Leftrightarrow ((a, b) \in R \wedge (b, a) \notin R)$ .

**Definition 2.** Let  $\mathcal{A} = [A, R_1, R_2]$  be a preference system. Then  $\mathcal{A}$  is said to be **consistent** if there exists a function  $u : A \rightarrow [0, 1]$  such that for all  $a, b, c, d \in A$  the following two properties hold:

- i) If  $(a, b) \in R_1$ , then  $u(a) \geq u(b)$  with equality iff  $(a, b) \in I_{R_1}$ .
- ii) If  $((a, b), (c, d)) \in R_2$ , then  $u(a) - u(b) \geq u(c) - u(d)$  with equality iff  $((a, b), (c, d)) \in I_{R_2}$ .

Every such function  $u$  is then said to (**weakly**<sup>2</sup>) **represent** the preference system  $\mathcal{A}$ . The set of all (weak) representations  $u$  of  $\mathcal{A}$  is denoted by  $\mathcal{U}_{\mathcal{A}}$ . The set of all  $u \in \mathcal{U}_{\mathcal{A}}$  satisfying  $\inf_{a \in A} u(a) = 0$  and  $\sup_{a \in A} u(a) = 1$  is denoted by  $\mathcal{N}_{\mathcal{A}}$ .

We will call a preference system *non-trivial* if there exists a pair  $(a, b) \in P_{R_1}$ , that is if there is at least one alternative that is strictly preferred to another one. In the rest of the paper we will throughout consider non-trivial preference systems and, therefore, drop the prefix non-trivial from now on. Note that trivial preference systems are represented by arbitrary maps  $c : A \rightarrow [0, 1]$ .

The idea behind the set  $\mathcal{N}_{\mathcal{A}}$  of normalized representations in the above definition is the following: For the special case, that the preference system  $\mathcal{A}$  is in accordance with the axioms in Krantz et al. (1971, Chapter 4), the representation is unique up to a positive linear transformation. Hence, the conditions  $\inf_a u(a) = 0$  and  $\sup_a u(a) = 1$  guarantee a unique representation for that special case. For the general case of a consistent preference system  $\mathcal{A}$  with non complete relations  $R_1$  and  $R_2$ , restricting analysis to the set  $\mathcal{N}_{\mathcal{A}}$  ensures that comparison will not be made with respect to equivalent representations which only measure utility on a different scale. Specifically, if  $u \in \mathcal{U}_{\mathcal{A}}$ , we have also  $\lambda \cdot u \in \mathcal{U}_{\mathcal{A}}$  for arbitrary  $\lambda \in (0, 1)$ . This means that both functions  $u$  and  $\lambda \cdot u$  represent the preference system  $\mathcal{A}$ , however, they measure utility on different scales, namely  $[\min_{a \in A} u(a), \max_{a \in A} u(a)]$  and  $[\min_{a \in A} \lambda \cdot u(a), \max_{a \in A} \lambda \cdot u(a)]$ . By restricting analysis to the set

---

<sup>2</sup>Here, the term *weakly* refers to the fact that the representation is meant in the *if* and not in the *if and only if* (short: *iff*) sense.



$\mathcal{N}_{\mathcal{A}}$  we therefore ensure that all considered representations measure utility on a  $[0, 1]$ -scale. The restriction on  $\mathcal{N}_{\mathcal{A}}$ , together with the concept of *granularity* from Definition 3, will prove crucial when comparing acts by means of the numerical representation in Section 4.2.

Further, note that for finite  $A$ , the boundedness condition on the utility function in Definition 2 implies the existence of alternatives in  $A$  that attain a greatest and a lowest utility value, but not necessarily of worst and best alternatives in  $A$  with respect to the relation  $R_1$ : An element of  $A$  attaining the highest utility value for a certain representation  $u : A \rightarrow [0, 1]$  might indeed be incomparable to all other elements of  $A$ .

Obviously, for a preference system  $\mathcal{A} = [A, R_1, R_2]$  to be consistent, certain compatibility criteria between the relations  $R_1$  and  $R_2$  have to be satisfied. For example it cannot be the case that, for some elements  $a, b, c \in A$ , it simultaneously holds that  $(c, a) \in P_{R_1}$  and  $((a, b), (c, b)) \in R_2$ , since any element  $u \in \mathcal{U}_{\mathcal{A}}$  would have to satisfy  $u(c) > u(a)$  and  $u(a) - u(b) \geq u(c) - u(b)$ . We now provide an algorithm for checking the consistency of a finite preference system.

**Proposition 1.** *Let  $\mathcal{A} = [A, R_1, R_2]$  be a preference system, where  $A = \{a_1, \dots, a_n\}$  is a finite and non-empty set. Consider the linear optimization problem*

$$\varepsilon = \langle (0, \dots, 0, 1)', (u_1, \dots, u_n, \varepsilon)' \rangle \longrightarrow \max_{(u_1, \dots, u_n, \varepsilon) \in \mathbb{R}^{n+1}} \quad (1)$$

with constraints  $0 \leq (u_1, \dots, u_n, \varepsilon) \leq 1$  and

- i)  $u_p = u_q$  for all  $(a_p, a_q) \in I_{R_1} \setminus \{(a, a) : a \in A\}$
- ii)  $u_q + \varepsilon \leq u_p$  for all  $(a_p, a_q) \in P_{R_1}$
- iii)  $u_p - u_q = u_r - u_s$  for all  $((a_p, a_q), (a_r, a_s)) \in I_{R_2} \setminus \{((a, b), (a, b)) : (a, b) \in R_1\}$
- iv)  $u_r - u_s + \varepsilon \leq u_p - u_q$  for all  $((a_p, a_q), (a_r, a_s)) \in P_{R_2}$

Then  $\mathcal{A}$  is consistent if and only if the optimal outcome of (1) is strictly positive.

**Proof.** First, note that  $(0, \dots, 0) \in \mathbb{R}^{n+1}$  defines an admissible solution of (1). Thus, the set of admissible solutions of (1) is non-empty. Since it is also bounded due to  $0 \leq (u_1, \dots, u_n, \varepsilon) \leq 1$ , we can deduce the existence of an optimal solution of (1) by utilizing that linear programming problems with a bounded and non-empty set of admissible solutions always possess an optimal solution. Let  $(u_1^*, \dots, u_n^*, \varepsilon^*)$  denote such an optimal solution.

**If:** Assume  $\varepsilon^* > 0$ . Define  $u : A \rightarrow [0, 1]$  by setting  $u(a_i) := u_i^*$  for all  $i \in \underline{n} := \{1, \dots, n\}$ . One then straightforwardly verifies that conditions i) to iv) imply that  $u \in \mathcal{U}_{\mathcal{A}}$ . Hence,  $\mathcal{A}$  is consistent.

**Only if:** Assume  $\varepsilon^* = 0$  and, for contradiction, that there exists  $v \in \mathcal{U}_{\mathcal{A}}$ . Define the values  $v_i := v(a_i)$  for all  $i \in \underline{n}$  and  $\delta := \min\{m_1, m_2\}$ , where  $m_1 := \min\{v_i - v_j : a_i P_{R_1} a_j\}$  and  $m_2 := \min\{(v_i - v_j) - (v_k - v_l) : (a_i, a_j) P_{R_2} (a_k, a_l)\}$ . Then, since  $v \in \mathcal{U}_{\mathcal{A}}$  and it therefore holds that  $m_1 > 0$  due to Definition 2 i) and that  $m_2 > 0$  due to Definition 2 ii), we have  $\delta > 0$ . One then straightforwardly verifies that  $(v_1, \dots, v_n, \delta)$  is an admissible solution to (1) with  $\delta > \varepsilon^*$ , contradicting the optimality of  $(u_1^*, \dots, u_n^*, \varepsilon^*)$ .  $\square$

The linear programming problem (1) possesses  $|R_2| + n + 2$  constraints. Thus, the number of constraints increases with the preciseness of the available information on the agent's preferences.

In applications, typically the relation  $R_2$  will be rather sparse (i.e. contain few comparable pairs of pairs), whereas the relation  $R_1$  will be rather dense (i.e. contain many comparable pairs). This is intuitive: While  $R_1$  is directly observable in the choice behavior of the agent, edges in  $R_2$  need to be gained by hypothetical comparisons in interviews and polls by asking questions like: “Imagine you have objects  $a$  and  $b$ . Would you rather be willing to accept the exchange of  $a$  by  $c$  or the exchange of  $b$  by  $d$ ?”

In order to reduce the number of constraints of the problem, note that (weak) representability of a preference system  $\mathcal{A} = [A, R_1, R_2]$  automatically implies transitivity of the represented relations. Therefore, in the constraints of the above optimization problem (1) it actually suffices to quantify only over (the corresponding indifference parts  $I_{R_1^*}, I_{R_2^*}$  and strict parts  $P_{R_1^*}, P_{R_2^*}$  of) some transitive reductions  $R_1^*, R_2^*$  of the relations  $R_1$  and  $R_2$ . However, note that this makes necessary to compute the corresponding transitive reductions which, again, raises the complexity of the problem to some extent.

Before turning to decision theory with preference system valued acts, we need one further concept, which will be of particular relevance in Section 4.2.

**Definition 3.** Let  $\mathcal{A} = [A, R_1, R_2]$  be a consistent preference system. Moreover, for  $\delta \in (0, 1)$ , let  $\mathcal{N}_{\mathcal{A}}^{\delta}$  denote the set of all  $u \in \mathcal{N}_{\mathcal{A}}$  satisfying  $u(a) - u(b) \geq \delta$  for all  $(a, b) \in P_{R_1}$  and  $u(a) - u(b) - u(c) + u(d) \geq \delta$  for all  $((a, b), (c, d)) \in P_{R_2}$ . Then,  $\mathcal{N}_{\mathcal{A}}^{\delta}$  is called the **(weak) representation set of granularity** (at least)  $\delta$ . Moreover, the decision system  $\mathcal{A}$  is called  **$\delta$ -consistent** if  $\mathcal{N}_{\mathcal{A}}^{\delta} \neq \emptyset$ .

On the one hand, the granularity  $\delta$  from Definition 3 can be given a similar interpretation as the just noticeable difference in the context of psychophysics (see Luce (1956) for details): It is the minimal difference in utility that the specific decision maker under consideration is able to notice given that utility is measured on a  $[0, 1]$ -scale. On the other hand, the granularity can also be given a more constructive interpretation, namely as a controlling device for the specific decision maker: Choosing a granularity parameter  $\delta > 0$  ensures that an act will not be labeled superior to another based solely on some utility function that involves utility differences that are practically not noticeable at all and, accordingly, should not influence the decision to be made.

The restriction of the analysis to utility functions that reflect the fact that utility differences below some threshold are not distinguishable empirically will play a crucial role when it comes to defining generalized expectations (and the decision criteria based on these) in Section 4.2. For now, it is sufficient to note that the algorithm given in Proposition 1 straightforwardly extends to checking whether a preference system  $\mathcal{A} = [A, R_1, R_2]$   $\delta$ -consistent. This is the statement of the following proposition.

**Proposition 2.** Let  $\mathcal{A} = [A, R_1, R_2]$  be a preference system, where  $A = \{a_1, \dots, a_n\}$  is a finite and non-empty set and let  $\delta \in (0, 1)$ . Then  $\mathcal{A}$  is  $\delta$ -consistent if and only if the optimal outcome of (1) is at least  $\delta$ .

**Proof.** Let  $(u_1^*, \dots, u_n^*, \varepsilon^*)$  denote an optimal solution to problem (1) and define  $u : A \rightarrow [0, 1]$  by setting  $u(a_i) := u_i^*$  for all  $i \in \underline{n}$ . First, note there have to exist elements  $a^+, a^- \in A$  such that  $u(a^-) = 0$  and  $u(a^+) = 1$ , since otherwise normalizing  $u$  to  $[0, 1]$  would induce a solution to (1) with objective value strictly greater than  $\varepsilon^*$ . Thus,  $u \in \mathcal{N}_{\mathcal{A}}$ . If  $\varepsilon^* \geq \delta$ , then the constraints of (1) guarantee that  $u(a) - u(b) \geq \delta$  for all  $(a, b) \in P_{R_1}$  and  $u(a) - u(b) - u(c) + u(d) \geq \delta$  for all  $((a, b), (c, d)) \in P_{R_2}$ . Hence,  $u \in \mathcal{N}_{\mathcal{A}}^{\delta}$  and  $\mathcal{A}$  is  $\delta$ -consistent. If conversely  $\mathcal{A}$  is  $\delta$ -consistent, we can choose  $u^+ \in \mathcal{N}_{\mathcal{A}}^{\delta}$  and we know from the proof of Proposition 1 that  $(u_1^+, \dots, u_n^+, \delta)$  with

$u_i^+ := u^+(a_i)$  for  $i \in \underline{n}$  defines an admissible solution to (1). Since  $\varepsilon^*$  is the optimal outcome of (1), we know that  $\varepsilon^* \geq \delta$ .  $\square$

#### 4. Decision Theory with ps-valued Acts

Differently from axiomatic approaches followed in, e.g., Seidenfeld et al. (1995); Nau (2006); Galaabaatar and Karni (2013), where (multi-)utility and (imprecise) probability representations are obtained by different axiomatic characterizations of preferences over acts, the aim of the present paper is to go the opposite direction and to obtain preferences on acts given a preference system and some additional, commonly partial, probabilistic information about the occurrence of the states of nature.

As already discussed in more detail in Section 2, most existing criteria for decision making under uncertainty are not applicable in such situations, since they require either a perfectly cardinal preference structure (like, e.g., maximizing expected utility or  $\Gamma$ -maximin) or, complementary, a precise probability measure representing the beliefs on the states of nature (like, e.g., first order stochastic dominance<sup>3</sup>). Therefore, we now propose and discuss three different approaches for decision making under uncertainty when the considered acts take values in some arbitrary preference system (abbreviated by *ps-valued* acts in the following) and when there is partial probabilistic information on the occurrence of the states available.

##### 4.1. Basic Setting

We start by defining the central concepts of the theory for the most general case. Let  $S$  denote some non-empty set equipped with some suitable  $\sigma$ -algebra  $\sigma(S)$ . The elements of  $S$  are interpreted as all possible states of nature about whose occurrence the decision maker under consideration is uncertain. Moreover, let  $\mathcal{M}$  denote some credal set on the measurable space  $(S, \sigma(S))$ , which is interpreted as the set of all probability measures on  $(S, \sigma(S))$  that are compatible with the available (partial) probabilistic information and thus describing the uncertainty about the occurrence of the states. For a given consistent preference system  $\mathcal{A}$ , a state space  $S$  and a credal set  $\mathcal{M}$ , a ps-valued act is a mapping  $X : S \rightarrow A$  assigning states of nature to elements of the preference system.

Given this, define the set  $\mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)} \subseteq A^S := \{X | X : S \rightarrow A\}$  by setting

$$\mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)} := \left\{ X \in A^S : u \circ X \text{ is } \sigma(S)\text{-}\mathcal{B}_{\mathbb{R}}\text{-measurable for all } u \in \mathcal{U}_{\mathcal{A}} \right\} \quad (2)$$

where  $\mathcal{B}_{\mathbb{R}}$  denotes the Borel sigma field on  $\mathbb{R}$ . By construction, the space  $\mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  consists of exactly those acts  $X : S \rightarrow A$  whose expectation exists with respect to all pairs  $(u, \pi) \in \mathcal{U}_{\mathcal{A}} \times \mathcal{M}$  of compatible probability measure and utility representation (since bounded and measurable random variables have finite expectation). We can now define our main object of study:

**Definition 4.** *In the situation above, call every subset  $\mathcal{G} \subseteq \mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  a **decision system** (with information base  $(\mathcal{A}, \mathcal{M})$ ). Moreover, call a decision system  $\mathcal{G}$  finite, if both  $|\mathcal{G}| < \infty$  and  $|S| < \infty$ , that is if both the set of states and the set of available acts are finite.*

---

<sup>3</sup>For approaches directly generalizing stochastic dominance to credal sets, see Denoeux (2009); Couso and Dubois (2012); Montes et al. (2014b); Couso and Destercke (2015).

The elements of a decision system  $\mathcal{G}$  are interpreted as those elements of the space  $\mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  that are available in the specific choice situation under consideration. Given a decision system  $\mathcal{G}$ , we are interested in the following question: How can we utilize the information base  $(\mathcal{A}, \mathcal{M})$  best possibly in order to define meaningful and reasonable choice criteria on the set  $\mathcal{G}$ ? In the following sections, we propose three different approaches that address exactly this question.

#### 4.2. Criteria based on Generalized Expectation Intervals

We start by introducing and studying decision criteria that are based on the analysis of generalized expectation intervals of the available acts. Depending on the attitude of the agent of interest towards the ambiguity underlying the situation (for instance she could be ambiguity *seeking* or ambiguity *averse* or something inbetween), such intervals can give rise to a variety of different optimality criteria for decision making. Specifically, for a ps-valued act  $X$  and a decision maker with a granularity parameter  $\delta > 0$ , the corresponding interval will range from the lowest to the highest possible expected value that choosing the act  $X$  can lead to under some pair  $(u, \pi) \in \mathcal{N}_{\mathcal{A}}^{\delta} \times \mathcal{M}$  that is compatible with the preference system  $\mathcal{A}$  and the probabilistic information  $\mathcal{M}$ . This leads to the definition of the basic concept of this section.

**Definition 5.** Let  $X \in \mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  and  $\delta \in (0, 1)$ . With  $\mathcal{D}_{\delta} := \mathcal{N}_{\mathcal{A}}^{\delta} \times \mathcal{M}$ , we call

$$\mathbb{E}_{\mathcal{D}_{\delta}}(X) := \left[ \underline{\mathbb{E}}_{\mathcal{D}_{\delta}}(X), \overline{\mathbb{E}}_{\mathcal{D}_{\delta}}(X) \right] := \left[ \inf_{(u, \pi) \in \mathcal{D}_{\delta}} \mathbb{E}_{\pi}(u \circ X), \sup_{(u, \pi) \in \mathcal{D}_{\delta}} \mathbb{E}_{\pi}(u \circ X) \right] \quad (3)$$

the *generalized interval expectation* of  $X$  with respect to  $\mathcal{A}, \mathcal{M}$  and granularity  $\delta$ .

Examples for how the generalized interval expectation is computed in concrete decision situations are given in the second part of Example 1 in Section 4.3 and in the application example in Section 5.2. Note that, in the spirit of the theory of imprecise probabilities, the set  $\mathbb{E}_{\mathcal{D}_{\delta}}(X)$  can be given an epistemic or an ontologic interpretation (see Walley and Fine (1982) or Augustin et al. (2014, p. 143)): If the imprecision/ambiguity in the sets arises from lack of information in the sense of e.g. partially observed choice behavior and/or partially known precise probabilities, the set  $\mathbb{E}_{\mathcal{D}_{\delta}}(X)$  is the set of all expectations arising in at least one situation that is compatible with the data. In contrast, if both sets  $\mathcal{N}_{\mathcal{A}}^{\delta}$  and  $\mathcal{M}$  have an ontologic interpretation, i.e. are interpreted as holistic entities of their own, the same holds true for the set of expectations  $\mathbb{E}_{\mathcal{D}_{\delta}}(X)$ .

Of course, all decision theory that is based on comparisons of the set  $\mathbb{E}_{\mathcal{D}_{\delta}}(X_i)$  of different acts  $X_i$  should reflect the underlying interpretation. The following definition gives three criteria rather relying on an ontologic interpretation of the set  $\mathcal{D}_{\delta}$ . Note that all of them are straightforward generalizations of the (complete order inducing) decision criteria commonly used in the theory of imprecise probabilities and reviewed, e.g., in Huntley et al. (2014).

**Definition 6.** Let  $\mathcal{G} \subseteq \mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  be a decision system and  $\delta \in (0, 1)$ . An act  $X \in \mathcal{G}$  is called

- i)  $\mathcal{D}_{\delta}$ -*maximin* iff  $\forall Y \in \mathcal{G} : \underline{\mathbb{E}}_{\mathcal{D}_{\delta}}(X) \geq \underline{\mathbb{E}}_{\mathcal{D}_{\delta}}(Y)$
- ii)  $\mathcal{D}_{\delta}$ -*maximax* iff  $\forall Y \in \mathcal{G} : \overline{\mathbb{E}}_{\mathcal{D}_{\delta}}(X) \geq \overline{\mathbb{E}}_{\mathcal{D}_{\delta}}(Y)$
- iii)  $\mathcal{D}_{\delta}^{\alpha}$ -*maximax* iff

$$\forall Y \in \mathcal{G} : \alpha \underline{\mathbb{E}}_{\mathcal{D}_{\delta}}(X) + (1 - \alpha) \overline{\mathbb{E}}_{\mathcal{D}_{\delta}}(X) \geq \alpha \underline{\mathbb{E}}_{\mathcal{D}_{\delta}}(Y) + (1 - \alpha) \overline{\mathbb{E}}_{\mathcal{D}_{\delta}}(Y)$$

where  $\alpha \in [0, 1]$  is some fixed parameter.

We denote by  $\underline{\mathcal{G}}_\delta$ ,  $\overline{\mathcal{G}}_\delta$  and  $\mathcal{G}_\delta^\alpha$  the sets of  $\mathcal{D}_\delta$ -maximin,  $\mathcal{D}_\delta$ -maximax and  $\mathcal{D}_\delta^\alpha$ -maximax acts in  $\mathcal{G}$ .

Independent of its interpretation, we need ways for computing the set  $\mathbb{E}_{\mathcal{D}_\delta}(X)$  in concrete situations. The following proposition gives a linear programming based algorithm for doing so in finite decision systems. However, note that applying the proposition requires the extreme points of the underlying credal set  $\mathcal{M}$  and, therefore, is ideal for situations where the number of extreme points is moderate and where closed formulas for computing the extreme points are available. For credal sets induced by 2-monotone lower/ 2-alternating upper probabilities such formulas exist (cf., Shapley (1971, Theorem 3, p.19)). While generally the number of extreme points could be very high (maximally  $|S|!$  for lower probabilities, cf. Derks and Kuipers (2002) and Wallner (2007)), convenient cases exist where furthermore efficient enumeration procedures are available (such special cases include ordinal probabilities (cf., Kofler (1989, p. 26)), comparative probabilities (cf., Miranda and Destercke (2015)), necessity measures (cf., Schollmeyer (2015)), p-boxes (cf., Montes and Destercke (2017)), probability intervals (cf., Weichselberger and Pöhlmann (1990, Chapter 2) or de Campos et al. (1994)) or pari-mutuel models (cf., Montes et al. (2017))).

**Proposition 3.** *Let  $\mathcal{A} = [A, R_1, R_2]$  be a consistent preference system, where  $A = \{a_1, \dots, a_n\}$  such that  $(a_1, b), (b, a_n) \in R_1$  for all  $b \in A$  and let  $\varepsilon^*$  denote the optimal outcome of problem (1). Moreover, let  $S = \{s_1, \dots, s_m\}$  be finite,  $\mathcal{M}$  be some polyhedral credal set on  $(S, 2^S)$  with extreme points  $\mathcal{E}(\mathcal{M}) := \{\pi^{(1)}, \dots, \pi^{(T)}\}$  and let  $X \in \mathcal{G}$ . For  $\varepsilon^* \geq \delta > 0$ , consider the collection of linear programs  $LP_1^\delta, \dots, LP_T^\delta$  given by:*

$$\sum_{i=1}^n u_i \cdot \pi^{(t)}(X^{-1}(\{a_i\})) \longrightarrow \min_{(u_1, \dots, u_n) \in \mathbb{R}^n} / \max_{(u_1, \dots, u_n) \in \mathbb{R}^n} \quad (LP_t^\delta)$$

with constraints  $0 \leq (u_1, \dots, u_n) \leq 1$ ,  $u_1 = 1$ ,  $u_n = 0$  and *i) to iv) as given in Proposition 1 (with  $\varepsilon := \delta$  fixed). Let  $\underline{v}(t, \delta)$  and  $\overline{v}(t, \delta)$  denote the optimal outcomes of problem  $LP_t^\delta$  in minimum and maximum form. Then, we have  $\mathbb{E}_{\mathcal{D}_\delta}(X) = [\min_t \underline{v}(t, \delta), \max_t \overline{v}(t, \delta)]$ .*

**Proof.** Let  $X \in \mathcal{G}$  and  $\varepsilon^* \geq \delta > 0$ . Then, the set  $\mathcal{N}_{\mathcal{A}}^\delta$  is non-empty and we can define the function  $f : \mathcal{D}_\delta \rightarrow \mathbb{R}$ ,  $(u, \pi) \mapsto \mathbb{E}_\pi(u \circ X)$ . For any representation  $u \in \mathcal{N}_{\mathcal{A}}^\delta$  fixed, the function  $\pi \mapsto f(u, \pi)$  is linear in  $\pi$  and, therefore, both convex and concave. By utilizing the facts that the pointwise infimum of any family of concave functions is a concave function and that the pointwise supremum of any family of convex functions is a convex function, we know that the functions  $\pi \mapsto \inf_u f(u, \pi)$  and  $\pi \mapsto \sup_u f(u, \pi)$  have to be concave and convex, respectively. But concave functions on polyhedral set attain their minimum and convex functions on polyhedral set attain their maximum on the set of extreme points. Hence, in order to find global maximum and minimum of the function  $f$ , it suffices to check for it on the set  $\mathcal{N}_{\mathcal{A}}^\delta \times \mathcal{E}(\mathcal{M})$ .

Now, let  $(u_1^*, \dots, u_n^*)$  denote an optimal solution to problem  $LP_t^\delta$  in maximum form for fixed  $t \in \{1, \dots, T\}$ . One then easily verifies that the constraints imply  $u^* \in \mathcal{N}_{\mathcal{A}}^\delta$ , where  $u^* : A \rightarrow [0, 1]$ ,  $u^*(a_i) := u_i^*$  and  $\overline{v}(t, \delta) = \mathbb{E}_{\pi^{(t)}}(u^* \circ X) = \sup\{\mathbb{E}_{\pi^{(t)}}(u \circ X) : u \in \mathcal{N}_{\mathcal{A}}^\delta\}$ . Analogous reasoning for the problem in minimum form yields  $\underline{v}(t, \delta) = \inf_{u \in \mathcal{N}_{\mathcal{A}}^\delta} \mathbb{E}_{\pi^{(t)}}(u \circ X)$ . Thus, applying our considerations from before yields  $\mathbb{E}_{\mathcal{D}_\delta}(X) = [\min_t \underline{v}(t, \delta), \max_t \overline{v}(t, \delta)]$ .  $\square$

Another way to compute the bounds in (3) in the case of 2-alternating upper probabilities (2-monotone lower probabilities) on a finite space  $A$  is to use the Choquet representation of the upper (lower) expectation (cf., e.g., Denneberg (1994, Proposition 10.3, p. 126)): For a fixed utility  $u$

and a 2-alternating upper probability  $\nu$  with associated credal set  $\mathcal{M}_\nu$  the corresponding expected upper utility can be written as  $\mathbb{E}_{\{u\} \times \mathcal{M}_\nu}(X) = \sum_{i=1}^n (u_{(i)} - u_{(i-1)}) \cdot \nu(\{s \in S \mid u(X(s)) \geq u_{(i)}\})$ , where  $u_{(i)}$  denotes the  $i$ -th value of the increasingly ordered involved utility values  $u_1, \dots, u_n$ .

If  $R_1$  is complete then the expectation is a linear form in the utility  $u$  and the maximization  $\max_{u \in \mathcal{N}_A^\delta} \mathbb{E}_{\{u\} \times \mathcal{M}_\nu}(X)$  translates to a simple linear program. If the relation  $R_1$  is not complete then the ordering of the utility values  $u_i$  can change as  $u$  ranges in  $\mathcal{N}_A^\delta$  and one has to compute the expectation separately for every possible ordering of the utility values and then take the maximum. If there are totally comparable values  $u_i$  meaning that for every  $u_j$  either  $u_i \leq u_j$  or  $u_i > u_j$ , independently from the concrete  $u \in \mathcal{N}_A^\delta$  then one can split the sum in a part containing all utility values below  $u_i$  and a part containing all utility values above  $u_i$  and then analyze every sub-sum independently which would help in reducing the combinatorial complexity.

The criteria from Definition 6 allow for comparing acts *given* the granularity  $\delta$  of the specific decision maker of interest. However, note that knowing the granularity might be a strong assumption if  $R_1$  and  $R_2$  are partial orderings, since experimental settings in which this additional parameter could precisely be elicited are not as straightforward as in the case of complete orderings. A natural way for addressing this issue in practical problems is to compute the generalized interval expectation along varying values of  $\delta$ . Clearly, it holds that  $\mathbb{E}_{\mathcal{D}_{\delta_1}}(X) \subseteq \mathbb{E}_{\mathcal{D}_{\delta_2}}(X)$  whenever  $\delta_1 \geq \delta_2$ , since it holds  $\mathcal{D}_{\delta_1} \subseteq \mathcal{D}_{\delta_2}$  and, thus, the inf and the sup in (3) are taken for a smaller set for  $\delta_1$  than for  $\delta_2$ . That is, the generalized interval expectation of an act  $X$  becomes narrower (or more precisely, not gets wider) as the value of  $\delta$  increases. Utilizing this fact, in order to decide between two competing acts  $X$  and  $Y$ , one could proceed as follows: Once having decided for one of the criteria from Definition 6, one can compute the general interval expectation for increasing values of  $\delta$  until the chosen criterion can discriminate between the acts  $X$  and  $Y$  for the first time, for instance in favor of  $X$  (say this happens for the value  $\delta^*$ ). Afterwards, the decision maker is asked whether it is acceptable for her that utility differences below  $\delta^*$  are not taken into account by the decision procedure. If the answer is yes, the decision maker should rank act  $X$  before  $Y$ , otherwise no decision can be made.

Further possibilities to deal with these issues are treated in the next two sections, where we propose two approaches completely overcoming the choice of a granularity parameter.

#### 4.3. Criteria based on Global Comparisons

The decision criteria defined in Section 4.2 all construct complete rankings on the set  $\mathcal{G}$  by comparing numerical representations of parts of the decision system and by somehow ignoring the inherent utility and probability structure. Therefore, when defining optimality of acts in terms of one of the criteria from Definition 6, it makes no difference if the ranking is constructed by pairwise or global comparisons. In the next sections, we turn to two approaches that explicitly take into account a global and local viewpoint for defining optimality of acts, respectively.<sup>4</sup>

We start with defining criteria taking the global perspective: For an act  $X$  in order to be labeled optimal, it is necessary that there exists (depending on the concrete approach at least) one fixed pair  $(u, \pi) \in \mathcal{U}_A \times \mathcal{M}$  for which this act maximizes the expected utility among all other available acts  $Y \in \mathcal{G}$ . In particular, the pair  $(u, \pi)$  for which  $X$  dominates the other acts in expectation must not depend on the concrete competing act under consideration, but has to be constant for all

---

<sup>4</sup>Note that in the context of IP decision theory, fundamental differences between global criteria and criteria based on pairwise comparisons have already been discussed Schervish et al. (2003).

acts from  $\mathcal{G}$ . Optimality criteria for which the pair  $(u, \pi)$  may depend on the concrete competing act are considered in Definition 8 in Section 4.4.

Moreover, note that the concepts from the following definition take a global point of view also from another perspective: The utility component in the required pair(s)  $(u, \pi)$  should not depend on its probability component and also its probability component should not depend on the utility component. This is reflected in the fact that in the admissibility concepts of Definition 7 a  $\forall$  quantifier can follow an  $\exists$  quantifier but not vice versa.

**Definition 7.** Let  $\mathcal{G} \subseteq \mathcal{F}_{(\mathcal{A}, \mathcal{M}, \mathcal{S})}$  denote a decision system. We call an act  $X \in \mathcal{G}$

- i)  $\mathcal{A}|\mathcal{M}$ -admissible iff  $\exists u \in \mathcal{U}_{\mathcal{A}} \exists \pi \in \mathcal{M} \forall Y \in \mathcal{G} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y)$
- ii)  $\mathcal{A}$ -admissible iff  $\exists u \in \mathcal{U}_{\mathcal{A}} \forall \pi \in \mathcal{M} \forall Y \in \mathcal{G} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y)$
- iii)  $\mathcal{M}$ -admissible iff  $\exists \pi \in \mathcal{M} \forall u \in \mathcal{U}_{\mathcal{A}} \forall Y \in \mathcal{G} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y)$
- iv)  $\mathcal{A}|\mathcal{M}$ -dominant iff  $\forall u \in \mathcal{U}_{\mathcal{A}} \forall \pi \in \mathcal{M} \forall Y \in \mathcal{G} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y)$

Denote by  $\mathcal{G}_{\mathcal{A}|\mathcal{M}}$ ,  $\mathcal{G}_{\mathcal{A}}$ ,  $\mathcal{G}_{\mathcal{M}}$  and  $\mathcal{G}_{\mathcal{A}|\mathcal{M}}^d$  the sets of such acts, respectively.

All four act properties just defined rely on the idea that, if there *was* perfect information on both the state probabilities (i.e.  $\mathcal{M} = \{\pi\}$  is a singleton) and the utility values (i.e. the utility representation  $u$  is unique up to a positive linear transformation), then an act  $X$  should be labeled optimal iff  $X$  has greater or equal expected utility than every other act  $Y \in \mathcal{G}$  with respect to  $(u, \pi)$ . However, they differ in the way they handle the ambiguity underlying the involved sets  $\mathcal{M}$  and  $\mathcal{U}_{\mathcal{A}}$ : While  $\mathcal{A}|\mathcal{M}$ -admissibility only demands the existence of at least *one* compatible combination  $(u, \pi)$  with respect to which  $X$  maximizes expected utility,  $\mathcal{A}|\mathcal{M}$ -dominance requires this for *all* compatible combinations.  $\mathcal{M}$ - and  $\mathcal{A}$ -admissibility relax the  $\forall$ -assumption on probability and utility level, respectively. Clearly, it holds that  $\mathcal{G}_{\mathcal{A}}, \mathcal{G}_{\mathcal{M}}, \mathcal{G}_{\mathcal{A}|\mathcal{M}}^d \subseteq \mathcal{G}_{\mathcal{A}|\mathcal{M}}$  and  $\mathcal{G}_{\mathcal{A}|\mathcal{M}}^d \subseteq \mathcal{G}_{\mathcal{A}}$  and  $\mathcal{G}_{\mathcal{A}|\mathcal{M}}^d \subseteq \mathcal{G}_{\mathcal{M}}$ , but in general neither  $\mathcal{G}_{\mathcal{A}} \subseteq \mathcal{G}_{\mathcal{M}}$  nor  $\mathcal{G}_{\mathcal{M}} \subseteq \mathcal{G}_{\mathcal{A}}$ .

Note that Danielson (2005, Section 4) also proposes a decision rule that rather could be viewed from a global perspective. Precisely, in that paper an act  $X$  is labeled simultaneously superior to the remaining acts from  $\mathcal{G} \setminus \{X\}$ , whenever the expectation  $\mathbb{E}_{\pi}(u \circ X)$  of  $X$  is greater or equal as the mean of the expectations  $\frac{1}{|\mathcal{G}|-1} \sum_{Y \in \mathcal{G} \setminus \{X\}} \mathbb{E}_{\pi}(u \circ Y)$  of the remaining acts for every compatible pair  $(u, \pi)$  of utility and probability representation. The author then proves that this criterion induces the same ranking of the acts than a similar criterion applied solely for pairwise comparisons of acts. This is an important difference to the concepts introduced in the present paper: Here, in general, the global admissibility concepts introduced in Definition 7 do indeed induce different orderings of the acts than the concepts of local admissibility introduced in Definition 8 from Section 4.4.

Note also that, if the involved set of utility representations  $\mathcal{U}_{\mathcal{A}}$  is a class of positive linear transformations, i.e. belongs to a perfectly cardinal preference structure, then both  $\mathcal{A}|\mathcal{M}$ -admissibility and  $\mathcal{M}$ -admissibility reduce to E-admissibility as recalled in Section 2. The following example demonstrates that ignoring the available information base and applying the maximin criterion instead leads to counter-intuitive decisions even in very simple situations.

**Example 1.** Let  $A = \{a_1, a_2, a_3, a_4\}$ , the (complete) relation  $R_1$  induced by  $a_2 P_{R_1} a_3 P_{R_1} a_4 P_{R_1} a_1$  and  $P_{R_2} = \{((a_2, a_4), (a_3, a_1))\}$  consists of one single edge. Consider the decision system  $\mathcal{G} = \{X_1, X_2\}$ , where the acts  $X_1, X_2 : \{s_1, s_2\} \rightarrow A$  are defined by  $(X_1(s_1), X_1(s_2)) = (a_1, a_2)$  and  $(X_2(s_1), X_2(s_2)) = (a_3, a_4)$ . An illustration of the decision system is given in Figure 2.

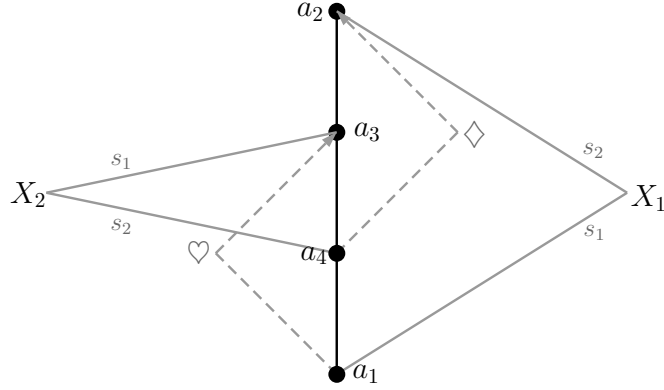


Figure 2: The black part shows the Hasse graph of the relation  $R_1$ , the gray dashed lines show the edges  $\heartsuit = (a_3, a_1)$  and  $\diamond = (a_2, a_4)$  with  $\diamond$  being preferred to  $\heartsuit$ . The solid gray lines show which elements of  $A$  are attained by the acts  $X_1$  and  $X_2$  under the different states of nature.

Moreover, suppose there is additional probabilistic information available which is given by the credal set  $\mathcal{M} := \{\pi : \pi(\{s_1\}) \leq 0.5\}$ . In this case, act  $X_1$  is  $\mathcal{A}|\mathcal{M}$ -dominant, since it maximizes expected utility with respect to every pair  $(u, \pi) \in \mathcal{U}_A \times \mathcal{M}$ . In contrast,  $X_2$  is not even  $\mathcal{A}|\mathcal{M}$ -admissible, although it is the unique optimal act with respect to Wald's maximin criterion (since it holds that  $\inf_{s \in S} X_2(s) = a_4 P_{R_1} a_1 = \inf_{s \in S} X_1(s)$ ). Moreover, we can go on computing the generalized interval expectations of the acts  $X_1$  and  $X_2$  (in the sense of Definition 5) for varying degrees of granularity, say  $\delta = 0, 0.1, 0.15, 0.25$ . The resulting expectation intervals for the acts are visualized in Figure 3.

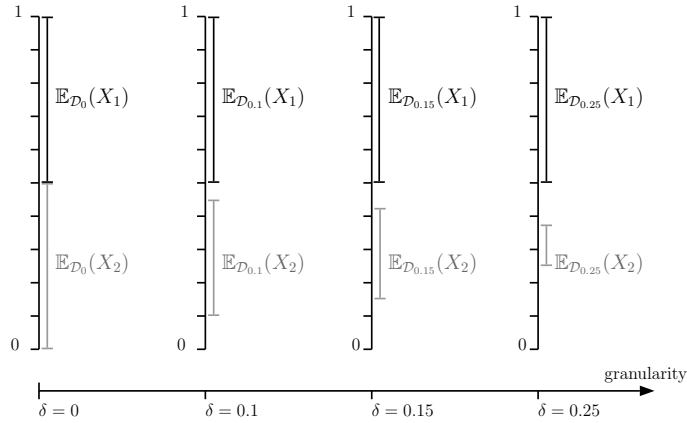


Figure 3: Development of the generalized interval expectations of the acts  $X_1$  and  $X_2$  along increasing granularity. For the case  $\delta = 0$  the lower expectation of act  $X_1$  and the upper expectation of act  $X_2$  coincide with a value of 0.5.

To complete the section, we give a proposition containing a linear programming based approach for checking whether an act  $X$  is  $\mathcal{A}$ -admissible in finite decision settings.

**Proposition 4.** Let  $\mathcal{A} = [A, R_1, R_2]$  be a consistent preference system, where  $A = \{a_1, \dots, a_n\}$ . Moreover, let  $S = \{s_1, \dots, s_m\}$  be finite,  $\mathcal{M}$  be some polyhedral credal set on  $(S, 2^S)$  with extreme



points  $\mathcal{E}(\mathcal{M}) := \{\pi^{(1)}, \dots, \pi^{(T)}\}$  and let  $\mathcal{G} := \{X_1, \dots, X_k\} \subseteq \mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  denote a finite decision system with  $X_z \in \mathcal{G}$ . Consider again the linear optimization problem (1) with additional constraints

$$\sum_{i=1}^n u_i \cdot \pi^{(t)}(X_z^{-1}(\{a_i\})) \geq \sum_{i=1}^n u_i \cdot \pi^{(t)}(X_l^{-1}(\{a_i\})) \quad \text{for all } l = 1, \dots, k \quad (C_t)$$

for every  $t = 1, \dots, T$ . Then  $X_z$  is  $\mathcal{A}$ -admissible if and only if the optimal outcome of this optimization problem is strictly greater than 0.

**Proof.** A similar argument as in the proof of Proposition 1 guarantees the existence of an optimal solution  $(u_1^*, \dots, u_n^*, \varepsilon^*)$  to the optimization problem. If  $\varepsilon^* = 0$ , then there exists no vector  $(u_1, \dots, u_n, \varepsilon)$  with  $\varepsilon > 0$  satisfying the constraints of the optimization problem. Since every function  $u \in \mathcal{U}_{\mathcal{A}}$  with  $\mathbb{E}_{\pi^{(t)}}(u \circ X_z) \geq \mathbb{E}_{\pi^{(t)}}(u \circ X_l)$  for all  $l = 1, \dots, k$  and  $t = 1, \dots, T$  induces such a vector, we conclude that such  $u$  cannot exist. Since  $\mathcal{E}(\mathcal{M}) \subseteq \mathcal{M}$ , we conclude that  $X_z$  is not  $\mathcal{A}$ -admissible.

If  $\varepsilon^* > 0$ , constraints i) to iv) guarantee that  $u : A \rightarrow \mathbb{R}$ ,  $u(a_i) := u_i^*$  for all  $i \in \underline{n}$  (weakly) represents the preference system  $\mathcal{A}$ . Now, let  $\pi \in \mathcal{M}$  be arbitrary. Choose  $\alpha \in \Delta_{T-1}$  such that  $\pi(\cdot) = \sum_{t=1}^T \alpha_t \cdot \pi^{(t)}(\cdot)$ . Then, condition  $(C_t)$  additionally guarantees that for all  $l = 1, \dots, k$  it holds

$$\begin{aligned} \mathbb{E}_{\pi}(u \circ X_z) &= \sum_{i=1}^n u_i^* \cdot \pi(X_z^{-1}(\{a_i\})) \\ &= \sum_{i=1}^n u_i^* \cdot \left( \sum_{t=1}^T \alpha_t \cdot \pi^{(t)}(X_z^{-1}(\{a_i\})) \right) \\ &= \sum_{t=1}^T \alpha_t \left( \sum_{i=1}^n u_i^* \cdot \pi^{(t)}(X_z^{-1}(\{a_i\})) \right) \\ &\geq \sum_{t=1}^T \alpha_t \left( \sum_{i=1}^n u_i^* \cdot \pi^{(t)}(X_l^{-1}(\{a_i\})) \right) \\ &= \sum_{i=1}^n u_i^* \cdot \left( \sum_{t=1}^T \alpha_t \cdot \pi^{(t)}(X_l^{-1}(\{a_i\})) \right) \\ &= \mathbb{E}_{\pi}(u \circ X_l) \end{aligned}$$

Hence,  $X_z$  maximizes expected utility with respect to  $(u, \pi)$ . Since  $\pi \in \mathcal{M}$  was chosen arbitrarily, this implies that  $X_z$  is  $\mathcal{A}$ -admissible.  $\square$

Note that a similar algorithm as given in Proposition 4 could be used for checking  $\mathcal{M}$ -admissibility of acts. However, this would require the set  $\mathcal{E}(\mathcal{U}_{\mathcal{A}})$  of extreme points of the representation set to be known, which is way less straightforward than assuming  $\mathcal{E}(\mathcal{M})$  to be known.

#### 4.4. Criteria based on Pairwise Comparisons

While the criteria defined in Section 4.3 rather relied on global comparisons of acts in the sense that an act, in order to be labeled admissible, has to dominate all other available acts from  $\mathcal{G}$  in expectation for (at least one) *fixed* pair  $(\pi, u)$  simultaneously, we now turn to criteria induced by pairwise expectation comparisons of acts (i.e. binary relations on the set of acts). There, roughly

spoken, the idea is to first compare the expectation of a fixed act  $X$  of interest to the expectation of every other available act  $Y \in \mathcal{G}$  *separately* and, afterwards, to label this act admissible if and only if none of the other available acts from  $\mathcal{G}$  dominates it. In particular, the pair  $(u_Y, \pi_Y)$  for which the expectation of act  $X$  is compared to the expectation of act  $Y$  might now depend on  $Y$ , for which reason the following criteria could rather be viewed from a local perspective.

Similarly as already seen in the global case, there are several different ways to define such relations each of which reflecting a different attitude towards the underlying ambiguity between the different compatible probability measures and/or the indeterminacy on the utility level. In particular, we define six binary relations  $R_{\exists\exists}, R_{\exists\forall}^1, R_{\exists\forall}^2, R_{\forall\exists}^1, R_{\forall\exists}^2$  and  $R_{\forall\forall}$  on  $\mathcal{F}_{(\mathcal{A}, \mathcal{M}, \mathcal{S})}$  by setting for all  $X, Y \in \mathcal{F}_{(\mathcal{A}, \mathcal{M}, \mathcal{S})}$ :

$$(X, Y) \in R_{\exists\exists} \quad \text{iff} \quad \exists u \in \mathcal{U}_{\mathcal{A}} \exists \pi \in \mathcal{M} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y) \quad (4)$$

$$(X, Y) \in R_{\exists\forall}^1 \quad \text{iff} \quad \exists u \in \mathcal{U}_{\mathcal{A}} \forall \pi \in \mathcal{M} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y) \quad (5)$$

$$(X, Y) \in R_{\exists\forall}^2 \quad \text{iff} \quad \exists \pi \in \mathcal{M} \forall u \in \mathcal{U}_{\mathcal{A}} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y) \quad (6)$$

$$(X, Y) \in R_{\forall\exists}^1 \quad \text{iff} \quad \forall u \in \mathcal{U}_{\mathcal{A}} \exists \pi \in \mathcal{M} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y) \quad (7)$$

$$(X, Y) \in R_{\forall\exists}^2 \quad \text{iff} \quad \forall \pi \in \mathcal{M} \exists u \in \mathcal{U}_{\mathcal{A}} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y) \quad (8)$$

$$(X, Y) \in R_{\forall\forall} \quad \text{iff} \quad \forall \pi \in \mathcal{M} \forall u \in \mathcal{U}_{\mathcal{A}} : \mathbb{E}_{\pi}(u \circ X) \geq \mathbb{E}_{\pi}(u \circ Y) \quad (9)$$

Obviously, it holds that  $R_{\forall\forall}$  is subset of all other relation, whereas  $R_{\exists\exists}$  is a superset of them. For the remaining relations, in general, no sub- or superset relation has to be satisfied. Furthermore, transitivity is only guaranteed for  $R_{\forall\forall}$  in general. Similarly as already discussed in the global case, each of the desirability relations just defined relies on the idea that, given perfect information on utilities and probabilities, maximizing expected utility should be the criterion of choice. Again, the relations differ only in the way they handle the ambiguity on the involved sets  $\mathcal{U}_{\mathcal{A}}$  and  $\mathcal{M}$ . Naturally, each of the relations defined above induces a different criterion of (local) admissibility. These criteria are summarized in the following definition.

**Definition 8.** Let  $R \in \{R_{\exists\exists}, R_{\exists\forall}^1, R_{\exists\forall}^2, R_{\forall\exists}^1, R_{\forall\exists}^2, R_{\forall\forall}\} =: \mathcal{R}_p$ . We call an act  $X \in \mathcal{G}$  **locally admissible** with respect to  $R$ , if it is an element of the set

$$\max_R(\mathcal{G}) := \{Y \in \mathcal{G} : \nexists Z \in \mathcal{G} \text{ s.t. } (Z, Y) \in P_R\}$$

that is if it is a maximal element in  $\mathcal{G}$  with respect to the relation  $R \cap (\mathcal{G} \times \mathcal{G})$ .

So, which of the relations in  $\mathcal{R}_p$  defined above are most important in our context? To address this question, it certainly makes sense to start by discussing some special cases of them: If the credal set  $\mathcal{M}$  is a singleton  $\mathcal{M} = \{\pi\}$ , that is a precise probability available, and if the set of compatible utility representations  $\mathcal{U}_{\mathcal{A}} = \{a \cdot u_0 + b \mid a > 0, b \in \mathbb{R}\}$  is unique up to a positive linear transformation of one utility function  $u_0$ , that is a perfectly cardinal utility, then all relations  $R \in \mathcal{R}_p$  coincide with the classical expected utility criterion, i.e. with choosing an act that maximizes the expectation with respect to  $\pi$  and one arbitrary chosen utility representation from  $\mathcal{U}_{\mathcal{A}}$ .

If  $\mathcal{M}$  still is a singleton, however,  $\mathcal{U}_{\mathcal{A}}$  is the class of all non-decreasing functions with respect to  $R_1$  (this essentially corresponds to the case where the relation  $R_2$  of the underlying preference system is empty), then the relations  $R_{\exists\forall}^2, R_{\forall\exists}^1$  and  $R_{\forall\forall}$  essentially coincide with the classical concept of first order stochastic dominance (cf., e.g., Lehmann (1955); Kamae et al. (1977); Mosler and Scarsini (1991)), while second order stochastic dominance is obtained if  $\mathcal{U}_{\mathcal{A}}$  is the set of all

continuous concave non-decreasing utility functions that are related to the concept of decreasing returns to scale. An intermediate case would arise if one has information about decreasing returns to scale only for parts of the preference system.

Finally, if the involved credal set  $\mathcal{M}$  is no longer a singleton and utility is given perfectly cardinal again, then the relations  $R_{\exists\exists}$ ,  $R_{\exists\forall}^2$  and  $R_{\forall\exists}^1$  all coincide and are exactly the ones corresponding to the criterion of maximality as recalled in Section 2. More precisely, the acts that are locally admissible with respect to one of the relations  $R_{\exists\exists}$ ,  $R_{\exists\forall}^2$  and  $R_{\forall\exists}^1$  in that special case, are exactly the acts that are not dismissed when applying maximality. Additionally, the relations  $R_{\exists\forall}^1$ ,  $R_{\forall\exists}^2$  and  $R_{\forall\forall}$  for that case reduce to Bewley's structural dominance (see, e.g., Bewley (2002) or Etner et al. (2012, p. 243)).

To check whether an act  $X$  dominates another acts  $Y$  with respect to one of the relations  $R_{\exists\exists}$  and  $R_{\forall\forall}$  in the general (yet finite) case, one can apply a similar technique as described in Proposition 3 by noting that  $\mathbb{E}_\pi(u \circ X) \geq \mathbb{E}_\pi(u \circ Y)$  is equivalent to  $\mathbb{E}_\pi(u \circ X - u \circ Y) \geq 0$ . Utilizing this fact leads us to the following proposition.

**Proposition 5.** *Let  $\mathcal{A} = [A, R_1, R_2]$  be a consistent preference system, where  $A = \{a_1, \dots, a_n\}$ . Moreover, let  $S = \{s_1, \dots, s_m\}$  be finite,  $\mathcal{M}$  be some polyhedral credal set on  $(S, 2^S)$  with extreme points  $\mathcal{E}(\mathcal{M}) := \{\pi^{(1)}, \dots, \pi^{(T)}\}$  and let  $X, Y \in \mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$ . Consider the collection of linear programs  $LO_1, \dots, LO_T$  given by:*

$$\sum_{i=1}^n u_i \cdot [\pi^{(t)}(X^{-1}(\{a_i\})) - \pi^{(t)}(Y^{-1}(\{a_i\}))] \longrightarrow \min_{(u_1, \dots, u_n) \in \mathbb{R}^n} / \max_{(u_1, \dots, u_n) \in \mathbb{R}^n} \quad (LO_t)$$

with constraint i)-iv) from Proposition 1 where  $\varepsilon$  is set to 0. For  $t = 1, \dots, T$ , denote by  $\underline{v}(t)$  and  $\bar{v}(t)$  the optimal value of problem  $LO_t$  in minimum and maximum form, respectively. Then, the following holds:

i)  $(X, Y) \in R_{\forall\forall}$  if and only if  $\min_t \underline{v}(t) \geq 0$

ii)  $(X, Y) \in R_{\exists\exists}$  if  $\max_t \bar{v}(t) > 0$

**Proof.** Define the function  $g : \tilde{\mathcal{U}}_A \times \mathcal{M} \rightarrow \mathbb{R}$  by setting  $g(u, \pi) = \mathbb{E}_\pi(u \circ X - u \circ Y)$ , where  $\tilde{\mathcal{U}}_A$  denotes the set of all functions  $u : A \rightarrow [0, 1]$  which are monotone (but not necessarily strictly monotone) with respect to the relations  $R_1$  and  $R_2$ . Note that  $\mathcal{U}_A \subseteq \tilde{\mathcal{U}}_A$ . A similar argument as performed in the proof of Proposition 3 shows that  $\min_t \underline{v}(t) = \min\{g(u, \pi) : (u, \pi) \in \tilde{\mathcal{U}}_A \times \mathcal{M}\}$  and  $\max_t \bar{v}(t) = \max\{g(u, \pi) : (u, \pi) \in \tilde{\mathcal{U}}_A \times \mathcal{M}\}$ .

Part i): If  $\min_t \underline{v}(t) \geq 0$  (note that this actually means  $\min_t \underline{v}(t) = 0$  since the vector  $(0, \dots, 0)$  is an admissible solution of  $LO_t$  for all  $t = 1, \dots, T$ ), then, according to the above identity, it holds that  $\mathbb{E}_\pi(u \circ X - u \circ Y) \geq 0$  for all  $(u, \pi) \in \tilde{\mathcal{U}}_A \times \mathcal{M}$ . Since  $\mathcal{U}_A \subseteq \tilde{\mathcal{U}}_A$ , this implies  $\mathbb{E}_\pi(u \circ X) \geq \mathbb{E}_\pi(u \circ Y)$  for all  $(u, \pi) \in \mathcal{U}_A \times \mathcal{M}$ . If contrarily  $\min_t \underline{v}(t) < 0$ , let  $(u_1^*, \dots, u_n^*)$  denote a solution yielding  $\min_t \underline{v}(t)$  and define  $u^* \in \tilde{\mathcal{U}}_A$  by setting  $u^*(a_i) := u_i^*$ . If  $u^* \in \mathcal{U}_A$  we are done. If  $u^* \in \tilde{\mathcal{U}}_A \setminus \mathcal{U}_A$ , choose  $u_0 \in \mathcal{U}_A \neq \emptyset$  (this is possible since  $\mathcal{A}$  is assumed to be consistent) such that  $\mathbb{E}_\pi(u_0 \circ X - u_0 \circ Y) < |\min_t \underline{v}(t)|$  for all  $\pi \in \mathcal{M}$  (this is possible since, with any  $u \in \mathcal{U}_A$ , we have also  $\lambda \cdot u \in \mathcal{U}_A$  for arbitrary  $\lambda \in (0, 1)$ ). One then easily verifies that  $u^+ := \frac{u^* + u_0}{2} \in \mathcal{U}_A$ . One also easily verifies that, if  $\pi^+ \in \mathcal{M}$  is chosen to be a credal element yielding outcome  $\min_t \underline{v}(t)$  in combination with  $u^*$ , then it holds  $\mathbb{E}_{\pi^+}(u^+ \circ X) < \mathbb{E}_{\pi^+}(u^+ \circ Y)$ . This completes the proof of i).

Part ii): If  $\max_t \bar{v}(t) > 0$ , then, due to  $\max_t \bar{v}(t) = \max\{g(u, \pi) : (u, \pi) \in \tilde{\mathcal{U}}_{\mathcal{A}} \times \mathcal{M}\}$ , there exists a pair  $(u^*, \pi^+) \in \tilde{\mathcal{U}}_{\mathcal{A}} \times \mathcal{M}$  such that  $\mathbb{E}_{\pi^+}(u^* \circ X) \geq \mathbb{E}_{\pi^+}(u^* \circ Y)$ . If  $u^* \in \mathcal{U}_{\mathcal{A}}$  we are done. If  $u^* \in \tilde{\mathcal{U}}_{\mathcal{A}} \setminus \mathcal{U}_{\mathcal{A}}$ , choose  $u_0 \in \mathcal{U}_{\mathcal{A}} \neq \emptyset$  (again utilizing the consistency of the preference system  $\mathcal{A}$ ) such that  $\mathbb{E}_{\pi}(u_0 \circ X - u_0 \circ Y) > -\max_t \bar{v}(t)$  (again utilizing the fact that, with any  $u \in \mathcal{U}_{\mathcal{A}}$ , we have also  $\lambda \cdot u \in \mathcal{U}_{\mathcal{A}}$  for arbitrary  $\lambda \in (0, 1)$ ). Analogously as in part i), we have that  $u^+ := \frac{u^* + u_0}{2} \in \mathcal{U}_{\mathcal{A}}$ . Moreover, one easily verifies that it holds that  $\mathbb{E}_{\pi^+}(u^+ \circ X) > \mathbb{E}_{\pi^+}(u^+ \circ Y)$ . Thus, there exists a pair  $(u^+, \pi^+) \in \mathcal{U}_{\mathcal{A}} \times \mathcal{M}$  with the desired property and, therefore, it holds that  $(X, Y) \in R_{\exists\exists}$ . This completes the proof of part ii).  $\square$

Note that the converse implication in part ii) of Proposition 5 is not necessarily true (for a trivial example consider the pair  $(X, X) \in R_{\exists\exists}$ ). A non-trivial situation where the opposite direction fails to hold is illustrated by the following toy example:

**Example 2.** Let  $A = \{a_1, a_2, a_3, a_4\}$ , the (complete) relation  $R_1$  induced by  $a_1 I_{R_1} a_4 P_{R_1} a_2 I_{R_1} a_3$  and the relation  $R_2 = \emptyset$ . Consider the decision system  $\mathcal{G} = \{X_1, X_2\}$  consisting of two acts  $X_1, X_2 : \{s_1, s_2\} \rightarrow A$  defined by  $(X_1(s_1), X_1(s_2)) = (a_1, a_2)$  and  $(X_2(s_1), X_2(s_2)) = (a_3, a_4)$ . Suppose the uncertainty about the states is characterized by the credal set  $\mathcal{M} = \{\pi : \pi(\{s_1\}) \leq 0.5\}$ . Then we have  $(X, Y) \in R_{\exists\exists}$ , since for  $\pi$  the uniform distribution and  $u$  defined by  $u(a_1) = u(a_4) = 0.75$  and  $u(a_2) = u(a_3) = 0.25$  we have  $\mathbb{E}_{\pi}(u \circ X_1) = \mathbb{E}_{\pi}(u \circ X_2)$ . However, it holds that

$$\begin{aligned} \max_t \bar{v}(t) &= \sup_{(u, \pi) \in \tilde{\mathcal{U}}_{\mathcal{A}} \times \mathcal{M}} \mathbb{E}_{\pi}(u \circ X_1 - u \circ X_2) \\ &= \sup_{(u, \pi) \in \tilde{\mathcal{U}}_{\mathcal{A}} \times \mathcal{M}} u(a_1)(\pi(\{s_1\}) - \pi(\{s_2\})) + u(a_2)(\pi(\{s_2\}) - \pi(\{s_1\})) \\ &= \sup_{(u, \pi) \in \tilde{\mathcal{U}}_{\mathcal{A}} \times \mathcal{M}} (u(a_1) - u(a_2))(\pi(\{s_1\}) - \pi(\{s_2\})) \\ &= 0 \end{aligned}$$

where the last equality holds since  $u(a_1) - u(a_2) \geq 0$  due to  $a_1 P_{R_1} a_2$  and, therefore, the product is maximal when  $\pi(\{s_1\})$  is, which is the case for  $\pi(\{s_1\}) = 0.5$ . Hence, we have constructed a situation where  $(X, Y) \in R_{\exists\exists}$  but not  $\max_t \bar{v}(t) > 0$ .

In the following proposition, we formulate an additional assumption under which also the opposite implication of Proposition 5 ii) is valid.

**Proposition 6.** Consider again the situation of Proposition 5. Additionally, assume that both  $R_1$  and  $R_2$  are antisymmetric relations and that there exists an element  $a_0 \in A$  such that

$$\pi^{(t)}(X^{-1}(\{a_0\})) - \pi^{(t)}(Y^{-1}(\{a_0\})) > 0$$

for all  $t = 1, \dots, T$ . Then, we have that  $(X, Y) \in R_{\exists\exists}$  if and only if  $\max_t \bar{v}(t) > 0$ .

**Proof.** It follows from Proposition 5 ii) that  $\max_t \bar{v}(t) > 0$  implies  $(X, Y) \in R_{\exists\exists}$ . For the converse implication assume that  $(X, Y) \in R_{\exists\exists}$ . We then can choose  $u^* \in \mathcal{U}_{\mathcal{A}}$  such that  $\mathbb{E}_{\pi^+}(u^* \circ X) \geq \mathbb{E}_{\pi^+}(u^* \circ Y)$  for some  $\pi^+ \in \mathcal{M}$ . If the inequality holds strictly we are done, since then setting  $u_i^* := u^*(a_i)$  for all  $i \in \underline{n}$  induces a solution with an objective value strictly greater than 0. Thus, assume  $\mathbb{E}_{\pi^+}(u^* \circ X) = \mathbb{E}_{\pi^+}(u^* \circ Y)$ . Choose  $\alpha \in \Delta_{T-1}$  such that  $\pi^+(\cdot) = \sum_{t=1}^T \alpha_t \cdot \pi^{(t)}(\cdot)$  and let  $\delta > 0$  be defined as in the proof of Proposition 1, i.e. as the minimal difference with respect to  $u$  over all elements of  $P_{R_1}$  and  $P_{R_2}$ . Define the function  $u^+ : A \rightarrow [0, 1]$  by  $u^+(a) := u^*(a)$  for  $a \neq a_0$

and  $u^+(a_0) := u^*(a_0) + \frac{\delta}{4}$ . Then one can show that  $u^+ \in \mathcal{U}_A$  (utilizing that both relations  $R_1$  and  $R_2$  are antisymmetric).<sup>5</sup> Set  $z(a, t) := \pi^{(t)}(X^{-1}(\{a\})) - \pi^{(t)}(Y^{-1}(\{a\}))$  for  $t = 1, \dots, T$  and  $a \in A$ . Then, we can compute

$$\begin{aligned}
\max_t \bar{v}(t) &\geq \mathbb{E}_{\pi^+}(u^+ \circ X - u^+ \circ Y) \\
&= \sum_{t=1}^T \alpha_t \cdot \mathbb{E}_{\pi^{(t)}}(u^+ \circ X - u^+ \circ Y) \\
&= \sum_{t=1}^T \alpha_t \cdot \left( \sum_{a \in A} u^+(a) \cdot z(a, t) \right) \\
&= \sum_{t=1}^T \alpha_t \cdot \left( \left( \sum_{a \in A \setminus \{a_0\}} u^+(a) \cdot z(a, t) \right) + u^+(a_0) \cdot z(a_0, t) \right) \\
&= \sum_{t=1}^T \alpha_t \cdot \left( \left( \sum_{a \in A \setminus \{a_0\}} u^*(a) \cdot z(a, t) \right) + (u^*(a_0) + \frac{\delta}{4}) \cdot z(a_0, t) \right) \\
&= \sum_{t=1}^T \alpha_t \cdot \left( \left( \sum_{a \in A} u^*(a) \cdot z(a, t) \right) + \frac{\delta}{4} \cdot z(a_0, t) \right) \\
&= \sum_{t=1}^T \alpha_t \cdot \sum_{a \in A} u^*(a) \cdot z(a, t) + \frac{\delta}{4} \cdot \sum_{t=1}^T \alpha_t \cdot z(a_0, t) \\
&= \sum_{t=1}^T \alpha_t \cdot \mathbb{E}_{\pi^{(t)}}(u^* \circ X - u^* \circ Y) + \frac{\delta}{4} \cdot \sum_{t=1}^T \alpha_t \cdot z(a_0, t) \\
&= \mathbb{E}_{\pi^+}(u^* \circ X - u^* \circ Y) + \frac{\delta}{4} \cdot \sum_{t=1}^T \alpha_t \cdot z(a_0, t) \\
&= 0 + \frac{\delta}{4} \cdot \sum_{t=1}^T \alpha_t \cdot (\pi^{(t)}(X^{-1}(\{a_0\})) - \pi^{(t)}(Y^{-1}(\{a_0\}))) \\
&> 0
\end{aligned}$$

where the first inequality sign is valid since, as seen in the proof of Proposition 5, we have that  $\max_t \bar{v}(t) = \max\{g(u, \pi) : (u, \pi) \in \tilde{\mathcal{U}}_A \times \mathcal{M}\}$  and where the last strict inequality holds since we have

---

<sup>5</sup>Antisymmetry is required since otherwise changing  $u^*$  only on the element  $a_0$  would mean that  $u^+$  cannot represent the relations  $I_{R_1}$  and  $I_{R_2}$  on pairs of the form  $(a_0, a_1) \in I_{R_1}$  with  $a_0 \neq a_1$  and pairs of the form  $((a_0, a_1), (a_2, a_3)) \in I_{R_2}$  with  $a_1, a_2, a_3 \in A \setminus \{a_0\}$ .

Given antisymmetry, proving that  $u^+ \in \mathcal{U}_A$  is then straightforward, however, involves some tedious arithmetic exercises. One has to show that  $u^+$  represents both relations  $R_1$  and  $R_2$ . Therefore, one first has to note that by definition of  $\delta$  it holds that  $u^*(a) - u^*(b) > \frac{\delta}{2}$  for all  $(a, b) \in P_{R_1}$  and that  $u^*(a) - u^*(b) - (u^*(c) - u^*(d)) > \frac{\delta}{2}$  for all  $((a, b), (c, d)) \in P_{R_2}$ . It is then immediate that  $u^+$  represents  $R_1$  and  $R_2$  for pairs not containing  $a_0$ , since for such pairs  $u^+$  equals  $u^*$  and  $u^* \in \mathcal{U}_A$ . Thus, we need only care about pairs containing  $a_0$ . There are several cases to distinguish. We only show the most complicated one. Assume for  $a_1, a_2 \in A \setminus \{a_0\}$  it holds that  $((a_1, a_0), (a_0, a_2)) \in P_{R_2}$ . Then, due to the second of the above identities, it holds that  $u^*(a_1) - u^*(a_0) - \frac{\delta}{2} > u^*(a_0) - u^*(a_2)$ . This implies  $u^*(a_1) - (u^*(a_0) + \frac{\delta}{4}) > u^*(a_0) + \frac{\delta}{4} - u^*(a_2)$ , which implies  $u^+(a_1) - u^+(a_0) > u^+(a_0) - u^+(a_2)$ . Thus,  $u^+$  represents  $R_2$  on pairs of this form. The remaining cases are similar.

$\pi^{(t)}(X^{-1}(\{a_0\})) - \pi^{(t)}(Y^{-1}(\{a_0\})) > 0$  for all  $t = 1, \dots, T$  by assumption. This gives  $\max_t \bar{v}(t) > 0$  and completes the proof of the proposition.  $\square$

Note that the other relations  $R \in \mathcal{R}_p \setminus \{R_{\forall\forall}, R_{\exists\exists}\}$  do not appear to be manageable in such a straightforward manner. However, in the special case that  $\mathcal{M}$  is the core of a belief function, all  $\pi \in \mathcal{M}$  can be understood as obtained from a mass transfer of probability mass to singleton sets of  $S$  (cf., e.g., Chateauneuf and Jaffray (1989, Corollary 3, p.273) or Derks et al. (2000, Theorem 2, p.29) in the context of game theory). Since classical first order stochastic dominance can be alternatively checked via the solution of a suitable mass transportation problem (cf., Mosler and Scarsini (1991, p. 269)), the computation of  $R_{\exists\forall}^2$  can be done by solving a composite mass transportation problem.

On the other hand, if  $R_1$  is totally ordered and  $R_2$  is empty, then first order stochastic dominance can be characterized for a precise probability  $\pi$  as  $Y \geq_{SD} X \iff \forall c \in A : \pi(Y \geq c) \geq \pi(X \geq c)$ . This can be generalized to imprecise probabilities by replacing the probability of the events  $X \geq c$  and  $Y \geq c$ , respectively by the lower or the upper probabilities associated with the credal set  $\mathcal{M}$ . This would lead to four other generalizations of stochastic dominance for imprecise probabilities and was studied in detail in Denoeux (2009) for the case of belief functions. In the special case of some notion of “independence” of  $X$  and  $Y$  the relation  $Y \geq X \iff \forall c \in A : \underline{P}(Y \geq c) \geq \bar{P}(X \geq c)$  would be equivalent to our relation  $R_{\forall\forall}$ . Here, we used the notations  $\underline{P}(B) = \inf_{\pi \in \mathcal{M}} \pi(B)$  and  $\bar{P}(B) = \sup_{\pi \in \mathcal{M}} \pi(B)$  for  $B \subseteq A$  and the term “independence” means that there always exists some  $\pi \in \mathcal{M}$  that attains at the same time  $\pi(X \geq c) = \bar{P}(X \geq c)$  and  $\pi(Y \geq c) = \underline{P}(Y \geq c)$ .

Note that the characterization of stochastic dominance via the probability of the events of the form  $X \geq c$  (and  $Y \geq c$ ) becomes far more complicated when dealing with a relation  $R_1$  that is only partial. In this case one has to consider the probabilities of all events of the form  $X \in U$  (and  $Y \in U$ ), where  $U$  is an arbitrary upset<sup>6</sup>, and the number of upsets can become extremely large such that explicitly checking all upsets becomes intractable. (But note that for the case of a precise probability, checking all upsets can be done by using linear programming techniques described in (Schollmeyer et al., 2017).)

It should be further mentioned that in Montes et al. (2014a, Section 3) and in Montes et al. (2014b, Section 3) the authors introduce six binary relations  $\succeq_1, \dots, \succeq_6$  relying on a quite similar idea as the relations in  $\mathcal{R}_p$ , however, in a slightly different context. Here, the authors explore six ways of extending a binary relation  $\succeq$  between random variables to binary relations  $\succeq_1, \dots, \succeq_6$  between sets of random variables that are based on the same construction principle as the ones collected in the set  $\mathcal{R}_p$  (i.e. considering all variants of placing the  $\exists$  and the  $\forall$  quantifier). The authors then propose to apply these relations for decision making with acts attaining uncertain rewards: By considering more general acts  $\Gamma : S \rightarrow 2^A \setminus \{\emptyset\}$  yielding set-valued outcomes (i.e. acts that are random sets), and their associated sets of random variables  $S(\Gamma) = \{X : S \rightarrow A : X(s) \in \Gamma(s)\}$ , they propose to prefer act  $\Gamma_1$  before  $\Gamma_2$  whenever  $S(\Gamma_1) \succeq_{i_0} S(\Gamma_2)$ , where  $\succeq_{i_0} \in \{\succeq_1, \dots, \succeq_6\}$  is the extension of choice. The main difference to the relations proposed in the present paper is that we do not consider acts with uncertain reward, but certain rewards with uncertain utility assignment. More precisely, we do exactly know which consequence from  $A$  is attained by which act under which state of nature, however, we do not know which is the concrete utility assignment. Consequently, we find ourselves in a more structured setting than the authors in Montes et al. (2014a,b) and we would ignore information by solely considering the relations  $\succeq_1, \dots, \succeq_6$ .

<sup>6</sup>A set  $U \subseteq A$  is called an upset if  $\forall x, y \in A : x \in U \ \& \ y \geq x \implies y \in U$ .

Furthermore, in Montes et al. (2014a, Section 5) the authors propose two binary relations  $\succeq_s^{\mathcal{M}}$  and  $\succeq_w^{\mathcal{M}}$  between acts if the uncertainty on the states is characterized by a credal set  $\mathcal{M}$ . Here, they first assume a family  $(\succeq_\pi)_{\pi \in \mathcal{M}}$  on the states each representing the ordering of the acts given  $\pi$  was the true distribution and, afterwards, define act  $X$  to be preferable to act  $Y$ , that is  $(X, Y) \in \succeq_s^{\mathcal{M}}$  or  $(X, Y) \in \succeq_w^{\mathcal{M}}$ , if it holds that  $(X, Y) \in \succeq_\pi$  for all  $\pi \in \mathcal{M}$  or it holds that  $(X, Y) \in \succeq_\pi$  for some  $\pi \in \mathcal{M}$ , respectively. Here, there are some close connections to the relations from the set  $\mathcal{R}_p$ : If we assume a preference system  $\mathcal{A} = [A, R_1, R_2]$  with a complete ordinal relation  $R_1$  and an empty cardinal relation  $R_2$  and we additionally choose  $\succeq_\pi$  to be defined as first order stochastic dominance with respect to  $\pi$  for every  $\pi \in \mathcal{M}$ , then it holds that  $(X, Y) \in R_{\forall\forall}$  if and only if  $(X, Y) \in \succeq_s^{\mathcal{M}}$  as well as  $(X, Y) \in R_{\exists\forall}^2$  if and only if  $(X, Y) \in \succeq_w^{\mathcal{M}}$ .

Finally, note that the relation  $R_{\forall\forall}$  is also discussed in Danielson and Ekenberg (1998, Section 2): If one considers the concept of  $t$ -admissibility proposed in that paper for the special case that  $t = 0$  and one additionally assumes the sets  $V$  and  $P$  from Danielson and Ekenberg (1998) to consist of exactly those utility and probability sentences that characterize the sets  $\mathcal{U}_{\mathcal{A}}$  and  $\mathcal{M}$  respectively, then the set of  $t$ -admissible acts coincides with the set of acts that are locally admissible with respect to  $R_{\forall\forall}$  in the sense of Definition 8. However, in general, it will not always be possible to describe the utility constraints induced by the relation  $R_2$  and the probability constraints induced by the credal set  $\mathcal{M}$  by one of the three types of utility and probability sentences that are considered in that paper. Of course, these constraints could straightforwardly be incorporated in the framework developed in the paper by allowing for larger classes of such sentences, since they are still linear in the corresponding values.

Clearly, the set of locally  $R_{\forall\forall}$ -admissible acts coincides with the set of  $\mathcal{A}|\mathcal{M}$ -dominant acts. Thus, part i) of Proposition 5 can also be used for checking whether an act  $X$  is  $\mathcal{A}|\mathcal{M}$ -dominant by solving the problems  $(LO_t)_{t=1, \dots, T}$  for every pair  $((X, Y))_{Y \in \mathcal{G}}$ , where  $\mathcal{G} \subseteq \mathcal{F}_{(\mathcal{A}, \mathcal{M}, S)}$  once again denotes a finite set of available concurring acts. This is a unique feature offered by  $R_{\forall\forall}$ : In general, the other global concepts of admissibility from Definition 4.3 cannot be expressed as induced by one of the local criteria from Definition 4.4 (for the special case of a cardinal  $u$  this is discussed and shown in Schervish et al. (2003)).

## 5. A Real World Toy Example

In this section, we apply certain aspects of the proposed framework for decision making under uncertainty by computing selected decision criteria for a prototypical toy example. Particularly, we thereby demonstrate, firstly, that our framework is computationally feasible and, secondly, show a class of situations in which ps-valued acts naturally appear in practical applications, namely situations where the orderings  $R_1$  and  $R_2$  arise from the fact that the acts map into some bivariate product space with one cardinal and one (potentially partial) ordinal dimension.

### 5.1. Setup of the Example

The example reads as follows: Suppose the agent under consideration is currently looking for a new job. As she has very high qualification in her field, she immediately receives three different job offers, say  $J_1, J_2$  and  $J_3$ , each of which appears to be of high interest for her at first sight. In order to come to a decision between the different job offers, she decides to address the situation systematically by comparing the jobs with respect to the offered monthly salary after tax and the offered additional benefits. She specifies the additional benefits collected in  $\mathcal{X} = \{b_1, \dots, b_5\}$  to be important for her, where

$b_1$	$b_2$	$b_3$	$b_4$	$b_5$
overtime premium	child care	advanced training	promotion prospects	flexible hours

Under the assumption that the different additional benefits are incomparable for our agent, the situation just described very naturally induces the following preference system  $\mathcal{A}' = [A', R'_1, R'_2]$ , where  $A' = \mathbb{R}^+ \times 2^{\mathcal{X}}$  is the set of possible decision outcomes (each of which consists of a potential salary offer and a set of additional benefits), the relation  $R'_1 \subseteq A' \times A'$  is the component-wise ordering given by

$$R'_1 = \{((y_1, B_1), (y_2, B_2)) : y_1 \geq y_2 \wedge B_1 \supseteq B_2\} \quad (10)$$

and the relation  $R'_2 \subseteq R'_1 \times R'_1$  partially specifying the strength of preferences is given by

$$R'_2 = \{(((y_1, B_1), (y_2, B_2)), ((y_3, B_3), (y_4, B_4))) : y_1 - y_2 \geq y_3 - y_4 \wedge B_1 \supseteq B_3 \supseteq B_4 \supseteq B_2\} \quad (11)$$

The relation  $R'_1$  is interpretable in a pretty straightforward manner: An element  $(y_1, B_1)$  of  $A'$  is preferred to another element  $(y_2, B_2)$ , whenever it is preferable in terms of salary, i.e.  $y_1 \geq y_2$  and offers a super-set of additional benefits, i.e.  $B_1 \supseteq B_2$ . Otherwise, the elements are incomparable with respect to  $R'_1$ . Moreover, also the relation  $R'_2$  possesses a very natural interpretation: Whenever, for elements  $((y_1, B_1), (y_2, B_2)), ((y_3, B_3), (y_4, B_4)) \in R'_1$ , it is clear that exchanging  $B_2$  by  $B_1$  is preferable to exchanging  $B_4$  by  $B_3$  since it holds that  $B_1 \supseteq B_3 \supseteq B_4 \supseteq B_2$ , one can compare the exchanges of elements from  $A$  by simply checking whether the difference  $y_1 - y_2$  is greater than the difference  $y_3 - y_4$  in the salaries.

Finally, the agent specifies a set  $S = \{s_1, \dots, s_4\}$  of four different economic scenarios which might affect the offers of the companies in different ways (for example, here,  $\{s_4\}$  might be some event having very negative influence on the stock price of the company offering job  $J_1$ , whereas  $\{s_1\}$  might be an event causing the opposite). Particularly, the agent can specify the following decision system describing her situation:

	$s_1$	$s_2$	$s_3$	$s_4$
$J_1$	$(5000, \mathcal{X})$ =: $a_1$	$(2700, \{b_1, b_2\})$ =: $a_2$	$(2300, \{b_1, b_2, b_3\})$ =: $a_3$	$(1000, \emptyset)$ =: $a_4$
$J_2$	$(3500, \{b_1, b_5\})$ =: $a_5$	$(2400, \{b_1, b_2\})$ =: $a_6$	$(1700, \{b_1, b_2\})$ =: $a_7$	$(2500, \{b_1\})$ =: $a_8$
$J_3$	$(3000, \{b_1, b_2, b_3\})$ =: $a_9$	$(1000, \{b_1\})$ =: $a_{10}$	$(2000, \{b_1\})$ =: $a_{11}$	$(3000, \{b_1, b_4, b_5\})$ =: $a_{12}$

Once having set up the decision system, the agent can also determine the relevant preference system  $\mathcal{A} = [A, R_1, R_2]$  by setting  $A := \{a_1, \dots, a_{12}\}$ ,  $R_1 := R'_1 \cap (A \times A)$  and  $R_2 := R'_2 \cap (R_1 \times R_1)$ , i.e. by restricting all sets contained in the triplet  $\mathcal{A}'$  to the relevant ones. The Hasse graph of the order  $R_1$  (which is clearly anti-symmetric here, since all elements of  $A$  are distinct) then can be visualized as in Figure 4. Note that  $R_2$  is not anti-symmetric, since for instance it holds that  $((a_3, a_7), (a_9, a_6)) \in I_{R_2}$  while  $(a_3, a_7) \neq (a_9, a_6)$  and therefore distinct equivalent elements with respect to  $R_2$  exist.



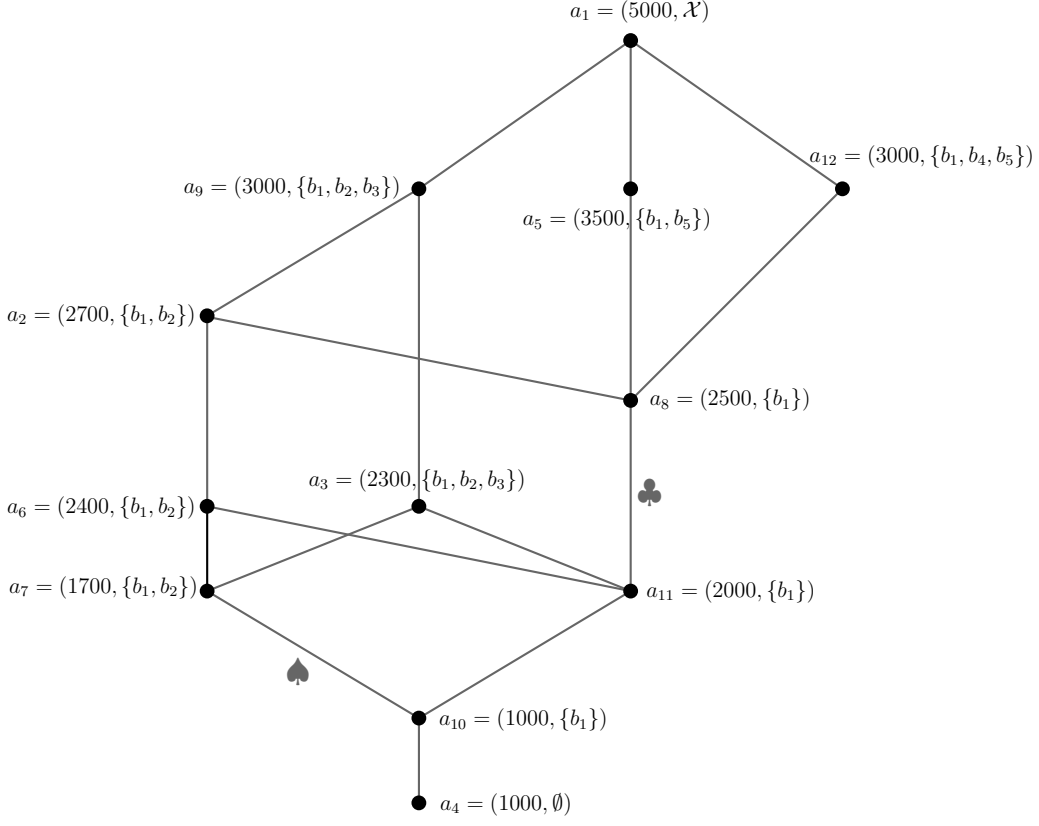


Figure 4: Hasse graph of the relation  $R_1$  of the example. The symbols ♠ and ♣ mark examples of elements of  $R_1 \times R_1$  that are comparable with respect to the strict relation  $P_{R_2}$ .

### 5.2. Checking Consistency and Applying the Decision Criteria

First, we want to check whether the preference system  $\mathcal{A} = [A, R_1, R_2]$  of the considered agent is consistent in the sense of Definition 2. Therefore, we apply the algorithm described in Proposition 1, whose objective function translates as

$$\varepsilon = \langle (0, \dots, 0, 1)', (u_1, \dots, u_{12}, \varepsilon)' \rangle \longrightarrow \max_{(u_1, \dots, u_{12}, \varepsilon) \in \mathbb{R}^{13}} \quad (12)$$

and whose constraints are determined by the relations  $R_1$  and  $R_2$  from the preference system under consideration as described in Proposition 1. For example, since  $(a_9, a_2) \in P_{R_1}$ , we add the constraint  $u_2 + \varepsilon \leq u_9$  and, since  $((a_7, a_{10}), (a_8, a_{11})) \in P_{R_2}$  (the edges  $(a_7, a_{10})$  and  $(a_8, a_{11})$  are indicated with ♠ and ♣ in the figure), we add the constraint  $u_8 - u_{11} + \varepsilon \leq u_7 - u_{10}$ . Solving the resulting linear programming problem (12) gives an optimal objective of  $0.\overline{037}$ . Hence, according to Proposition 1, the preference system  $\mathcal{A}$  is consistent. An optimal solution to the problem is given by  $(u_1^*, \dots, u_{12}^*, \varepsilon^*) \approx (1, 0.\overline{4}, 0.\overline{370}, 0, 0.\overline{629}, 0.\overline{370}, 0.\overline{2}, 0.\overline{370}, 0.\overline{518}, 0.\overline{037}, 0.\overline{259}, 0.\overline{5}, 0.\overline{037})$ , which induces an element  $u^* \in \mathcal{U}_{\mathcal{A}}$  by setting  $u^*(a_i) := u_i^*$  for  $i = 1, \dots, 12$ .

Now, suppose our agent collects some more information that allows her to order the different economic scenarios  $s_1, \dots, s_4$  by their probability to occur, i.e. by an ordinal probability specified by the credal set  $\mathcal{M} = \{\pi : \pi(\{s_1\}) \geq \pi(\{s_2\}) \geq \pi(\{s_3\}) \geq \pi(\{s_4\})\}$ . In this situation, the set

of extreme point of  $\mathcal{M}$  possesses exactly four elements and is given by  $\mathcal{E}(\mathcal{M}) = \{\pi^{(1)}, \dots, \pi^{(4)}\}$ , where we have  $\pi^{(t)}(\{s_j\}) = \mathbb{1}_{\{1, \dots, t\}}(j) \cdot \frac{1}{t}$  for  $j, t \in \{1, 2, 3, 4\}$  (cf., Kofler (1989, p. 26) or Miranda and Destercke (2015, Proposition 5 and Algorithm 1)).

Then, we want to check which of the jobs  $J_1, J_2, J_3$  are  $\mathcal{A}$ -admissible in the sense of Definition 7, part ii). The linear optimization problem described in Proposition 4, for instance applied for job  $J_1$ , then possesses the same objective function as the program for checking consistency, namely (12). Moreover, it also includes all the constraints of problem (12), however, additionally involves the constraints (as described in Proposition 4)

$$(C_1) \quad \begin{aligned} u_1 &\geq u_5 \\ u_1 &\geq u_9 \end{aligned}$$

$$(C_2) \quad \begin{aligned} \frac{1}{2}(u_1 + u_2) &\geq \frac{1}{2}(u_5 + u_6) \\ \frac{1}{2}(u_1 + u_2) &\geq \frac{1}{2}(u_9 + u_{10}) \end{aligned}$$

$$(C_3) \quad \begin{aligned} \frac{1}{3}(u_1 + u_2 + u_3) &\geq \frac{1}{3}(u_5 + u_6 + u_7) \\ \frac{1}{3}(u_1 + u_2 + u_3) &\geq \frac{1}{3}(u_9 + u_{10} + u_{11}) \end{aligned}$$

$$(C_4) \quad \begin{aligned} \frac{1}{4}(u_1 + u_2 + u_3 + u_4) &\geq \frac{1}{4}(u_5 + u_6 + u_7 + u_8) \\ \frac{1}{4}(u_1 + u_2 + u_3 + u_4) &\geq \frac{1}{4}(u_9 + u_{10} + u_{11} + u_{12}) \end{aligned}$$

that are due to the information about the uncertainty that is given by the credal set  $\mathcal{M}$  (where  $C_i$ , for  $i = 1, \dots, 4$ , here describes the constraint induced by the  $i$ th extreme point). Again, solving the resulting linear programming problem gives an optimal objective of  $0.\overline{037}$  and, again, an optimal solution to the problem is given by  $(u_1^*, \dots, u_{12}^*, \varepsilon^*)$  from above. However, the interpretation of the optimal solution is quite different: If we define  $u^* \in \mathcal{U}_{\mathcal{A}}$  as above, then job  $J_1$  maximizes expected utility with respect to  $(u^*, \pi)$  for every  $\pi \in \mathcal{M}$  compatible with the agent's probabilistic information. In contrast, solving the same linear programming problem from Proposition 4 for the jobs  $J_2$  and  $J_3$  gives an optimal objective of 0 each time indicating that both jobs are not  $\mathcal{A}$ -admissible. According to  $\mathcal{A}$ -admissibility, therefore, our agent should decide for job  $J_1$ .

Next, we want to compute the generalized interval expectations from Definition 5 of choosing one of the three jobs under consideration along varying granularity value  $\delta = 0, 0.01, 0.02, 0.03$ .<sup>7</sup> According to Proposition 3, for job  $J_k$  ( $k=1,2,3$ ) and fixed value of the granularity parameter  $\delta$ , this makes necessary solving the optimization problems

$$\sum_{i=1}^{12} u_i \cdot \pi^{(t)}(J_k^{-1}(\{a_i\})) \longrightarrow \min_{(u_1, \dots, u_{12}) \in \mathbb{R}^{12}} / \max_{(u_1, \dots, u_{12}) \in \mathbb{R}^{12}} \quad (13)$$

for  $t = 1, \dots, 4$ , with constraints as described in Proposition 1, with the difference that  $\varepsilon$  here is not one of the variables, but is set to  $\delta$ . Thus, eight linear programs have to be solved for each choice of job  $J_k$  and  $\delta$  and the generalized expectation interval can be computed as described in the proposition (i.e. by taking for each job the minimum of the minima as lower bound and the maximum of the maxima as upper bound). Solving the corresponding optimization problems gives the following results:

<sup>7</sup>Note that, since the optimal objective value of program (12) equals  $0.\overline{037}$ , it makes no sense to consider values of  $\delta$  any greater than that.

	$\delta = 0$	$\delta = 0.01$	$\delta = 0.02$	$\delta = 0.03$
$\mathbb{E}_{\mathcal{D}_\delta}(J_1)$	[0.25,1]	[0.305,1]	[0.36,1]	[0.415,1]
$\mathbb{E}_{\mathcal{D}_\delta}(J_2)$	[0,1]	[0.1075,0.9]	[0.215,0.8]	[0.3225,0.7]
$\mathbb{E}_{\mathcal{D}_\delta}(J_3)$	[0,1]	[0.07 $\bar{3}$ , 0.9]	[0.14 $\bar{6}$ ,0.8]	[0.22,0.7]

Since both lower and upper bound of the interval of  $J_1$  are greater than the respective bounds of the intervals of  $J_2$  and  $J_3$  (independent of which granularity value is chosen), job  $J_1$  is also optimal with respect to all criteria introduced in Definition 6. The generalized expectation intervals of the different jobs along increasing value of granularity are visualized in Figure 5.

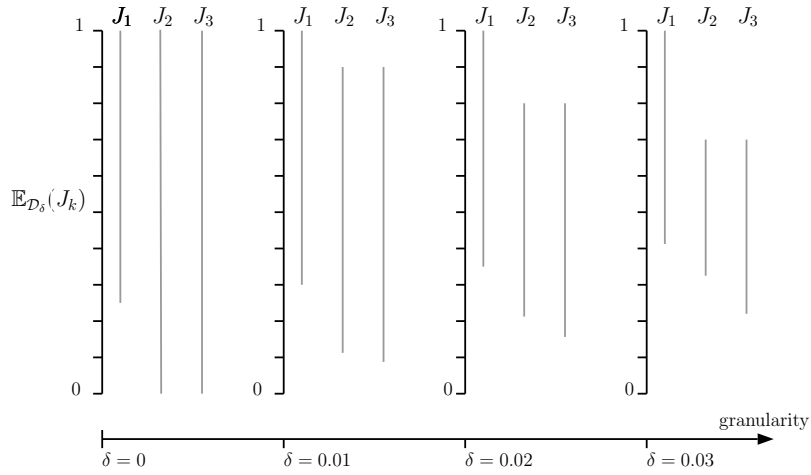


Figure 5: Generalized expectation intervals of the different jobs along increasing value of granularity.

Finally, we apply Proposition 5 in order to investigate how the different job offers  $J_1$ ,  $J_2$  and  $J_3$  relate to each other with respect to the relations  $R_{\exists\exists}$  and  $R_{\forall\forall}$ . The results are summarized in the following table:

	$(J_1, J_2)$	$(J_1, J_3)$	$(J_2, J_1)$	$(J_2, J_3)$	$(J_3, J_1)$	$(J_3, J_2)$
$R_{\exists\exists}$	$\in$	$\in$	$\in$	$\in$	$\in$	$\in$
$R_{\forall\forall}$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$

As discussed in the second paragraph after the proof of Proposition 5, from the fact that  $(J_1, J_2) \notin R_{\forall\forall}$  one can also conclude that job  $J_1$  is not  $\mathcal{A}|\mathcal{M}$ -dominant in the sense of Definition 7.

Concluding the example, we have seen that the agent under consideration should most likely decide for job  $J_1$ , since it is the only offer which is  $\mathcal{A}$ -admissible in the sense of Definition 7. Moreover,  $J_1$  is the unique optimal offer with respect to the criteria based on generalized expectation intervals as introduced in Definition 6. In contrast, when preferring pairwise comparison of the different job offers with respect to the binary relations  $R_{\exists\exists}$  and  $R_{\forall\forall}$ , no clear decision can be made: While every job offer  $J_i$  dominates any other job offer  $J_j$ , where  $i, j \in \{1, 2, 3\}$ , with respect to  $R_{\exists\exists}$ , none of the job offers are comparable with respect to  $R_{\forall\forall}$ .

## 6. Summary and Outlook

In this paper, we proposed three approaches for decision making under severe uncertainty if the acts under consideration take values in some preference system, i.e. can be understood as partial cardinal and partial ordinal valued. Our first approach is based on comparing granularity-dependent expectation intervals. Specifically, we proposed three decision criteria based on these intervals that are direct generalizations of the decision criteria known from the theory of imprecise probabilities. The other two approaches for decision making discussed in the paper rely on local and global comparisons of specific compatible expectations of the considered acts, respectively: For the former approach one searches for compatible pairs  $(u, \pi)$  of utility and probability representations with respect to which the act  $X$  of interest simultaneously dominates all the other available acts  $Y \in \mathcal{G}$  in expectation. For the latter approach, it suffices if for each other available act  $Y \in \mathcal{G}$  there exists a pair  $(u_Y, \pi_Y)$  such that  $X$  dominates  $Y$  with respect to this specific pair. At several points, we discussed how special cases of our criteria relate to concepts from the classical theory like for instance stochastic dominance or the criteria from decision theory using imprecise probabilities. For certain decision criteria proposed in the paper, we moreover provided linear programming algorithms to evaluate them. Finally, we illustrated a class of situations where our framework appears natural by means of a prototypical toy examples.

There are, of course, several challenges that could be addressed in future research. Clearly, further algorithms for evaluating the remaining criteria that were proposed in the paper need to be explored in order to make the theory computationally more tractable and, therefore, applicable in practice (compare, in particular, the discussion directly following the proof of Proposition 6 of Section 4.4). Further, it is certainly worth investigating in more detail how the criteria from the different approaches relate to each other and what can be learned about them by considering special cases of imprecise probabilistic models. Finally, designing experimental settings for eliciting the parameter  $\delta$  could help to receive a more canonical interpretation of granularity (compare, in particular, the discussions directly following Definition 3 of Section 3 as well as the discussions at the end of Section 4.2).

## Acknowledgements

We are most grateful to the anonymous referees, from whose valuable and insightful comments as well as hints to further relevant references the paper benefitted very much. We also thank the referees of the ISIPTA '17 contribution underlying this paper (see Jansen et al. (2017b)) and participants of ProGic 2017 and ISIPTA '17 for helpful comments. Further, we want to thank Jean Baccelli for ongoing stimulating discussions on the topic.

## References

- Augustin, T. (2001). On decision making under ambiguous prior and sampling information. In de Cooman, G., Fine, T., Moral, S., and Seidenfeld, T., editors, *Proceedings of the Second International Symposium on Imprecise Probability: Theories and Applications*, pages 9–16. Shaker.
- Augustin, T., Walter, G., and Coolen, F. (2014). Statistical inference. In Augustin, T., Coolen, Frank, de Cooman, G., and Matthias Troffaes, editors, *Introduction to Imprecise Probabilities*, pages 143–189. Wiley, Chichester.
- Bewley, T. (2002). Knightian decision theory. part i. *Decisions in Economics and Finance*, 25:79–110.
- Bradley, S. (2015). How to choose among choice functions. In Augustin, T., Doria, S., Miranda, E., and Quaeghebeur, E., editors, *Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 57–66. Aracne.

- Chateauneuf, A. and 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.
- Couso, I. and Destercke, S. (2015). Ranking of fuzzy intervals seen through the imprecise probabilistic lens. *Fuzzy Set Syst*, 278:20–39.
- Couso, I. and Dubois, D. (2012). An imprecise probability approach to joint extensions of stochastic and interval orderings. *Comm Com Inf Sc*, 299:388–399.
- Danielson, M. (2005). Generalized evaluation in decision analysis. *Eur J Oper Res*, 162:442–449.
- Danielson, M. and Ekenberg, L. (1998). A framework for analyzing decision under risk. *Eur J Oper Res*, 104:474–484.
- Danielson, M., Ekenberg, L., Johansson, J., and Larsson, A. (2003). The DecideIT decision tool. In Bernard, J.-M., Seidenfeld, T., and Zaffalon, M., editors, *Proceedings of the Third International Symposium on Imprecise Probability: Theories and Applications*, pages 204–217. Carleton Scientific.
- de Campos, L., Huete, J., and Moral, S. (1994). Probability interval: A tool for uncertain reasoning. *Int J Uncertain Fuzz*, 2:167–196.
- Denneberg, D. (1994). *Non-additive Measure and Integral*. Kluwer Academic Publishers, Dordrecht, Boston and London.
- Denoeux, T. (2009). Extending stochastic ordering to belief functions on the real line. *Inform Sciences*, 179:1362–1376.
- Derks, J., Haller, H., and Peters, H. (2000). The selectope for cooperative games. *Int J Game Theory*, 29(1):23–38.
- Derks, J. and Kuipers, J. (2002). On the number of extreme points of the core of a transferable utility game. In Borm, P. and Peters, H., editors, *Chapters in Game Theory in Honour of Stef Tijs*, pages 83–97. Kluwer.
- Etner, J., Jeleva, M., and Tallon, J.-M. (2012). Decision theory under ambiguity. *J Econ Surv*, 26:234–270.
- Fishburn, P. (1970). *Utility Theory for Decision Making*. Wiley, London and New York.
- Galaabaatar, T. and Karni, E. (2013). Subjective expected utility with incomplete preferences. *Econometrica*, 81:255–284.
- Gilboa, I. (2009). *Theory of Decision under Uncertainty*. Cambridge University Press, New York.
- Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *J Math Econ*, 18:141153.
- Hable, R. and Troffaes, M. (2014). Computation. In Augustin, T., Coolen, F., de Cooman, G., and Troffaes, M., editors, *Introduction to Imprecise Probabilities*, pages 329–337. Wiley, Chichester.
- Huntley, N., Hable, R., and Troffaes, M. (2014). Decision making. In Augustin, T., Coolen, Frank, de Cooman, G., and Matthias Troffaes, editors, *Introduction to Imprecise Probabilities*, pages 190–206. Wiley, Chichester.
- Jaffray, J.-Y. (1999). Rational decision making with imprecise probabilities. In de Cooman, G., Cozman, F., Moral, S., and Walley, P., editors, *Proceedings of the First International Symposium on Imprecise Probability: Theories and Applications*, pages 183–188. IPP.
- Jaffray, J.-Y. and Jeleva, M. (2008). Information processing under imprecise risk with an insurance demand illustration. *Int J Approx Reason*, 49:117–129.
- Jansen, C., Augustin, T., and Schollmeyer, G. (2017a). Decision theory meets linear optimization beyond computation. In Antonucci, A., Cholvy, L., and Papini, O., editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty. ECSQARU 2017. Lecture Notes in Computer Science, vol 10369.*, pages 329–339. Springer, Cham.
- Jansen, C., Schollmeyer, G., and Augustin, T. (2017b). Concepts for decision making under severe uncertainty with partial ordinal and partial cardinal preferences. In Antonucci, A., Corani, G., Couso, I., and Destercke, S., editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 181–192. PMLR.
- Kamae, T., Krengel, U., and O’Brien, G. (1977). Stochastic inequalities on partially ordered spaces. *Ann Probab*, pages 899–912.
- Kikuti, D., Cozman, F., and Filho, R. (2011). Sequential decision making with partially ordered preferences. *Artif Intel*, 175:1346 – 1365.
- Kofler, E. (1989). *Prognosen und Stabilität bei unvollständiger Information*. Campus, Frankfurt.
- Kofler, E. and Menges, G. (1976). *Entscheiden bei unvollständiger Information*. Springer, Berlin.
- Krantz, D., Luce, R., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement. Volume I: Additive and Polynomial Representations*. Academic Press, San Diego and London.
- Landes, J. (2014). Minmax decision rules for choice under complete uncertainty: Axiomatic characterizations for preferences over utility intervals. *Int J Approx Reason*, 55:1301–1317.
- Lehmann, E. (1955). Ordered families of distributions. *Ann Math Stat*, 26:399–419.
- Levi, I. (1974). On indeterminate probabilities. *J Philos*, 71:391–418.
- Levi, I. (1983). *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*. MIT Press,

- Cambridge, Mass.
- Loquin, K. (2013). Significance of a decision making problem under uncertainty. In Cozman, F., Denoeux, T., Destercke, S., and Seidenfeld, T., editors, *Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications*, pages 203–212. SIPTA.
- Luce, R. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, 24:178–191.
- Miranda, E. and Destercke, S. (2015). Extreme points of the credal sets generated by comparative probabilities. *J Math Psychol*, 64:44–57.
- Montes, I. (2014). *Comparison of Alternatives under Uncertainty and Imprecision*. PhD thesis, Oviedo. [digibuo.uniovi.es/dspace/bitstream/10651/28953/6/TD\\_IgnacioMontesGutierrez.pdf](http://digibuo.uniovi.es/dspace/bitstream/10651/28953/6/TD_IgnacioMontesGutierrez.pdf).
- Montes, I. and Destercke, S. (2017). Extreme points of p-boxes and belief functions. *Ann Math Artif Intel*, 81:405–428.
- Montes, I., Miranda, E., and Destercke, S. (2017). A study of the pari-mutuel model from the point of view of imprecise probabilities. In Antonucci, A., Corani, G., Couso, I., and Destercke, S., editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 229–240. PMLR.
- Montes, I., Miranda, E., and Montes, S. (2014a). Decision making with imprecise probabilities and utilities by means of statistical preference and stochastic dominance. *Eur J Oper Res*, 234:209–220.
- Montes, I., Miranda, E., and Montes, S. (2014b). Stochastic dominance with imprecise information. *Comput Stat Data An*, 71:867–885.
- Mosler, K. and Scarsini, M. (1991). Some theory of stochastic dominance. In Mosler, K. and Scarsini, M., editors, *Stochastic Orders and Decision under Risk*, pages 203–212. Institute of Mathematical Statistics, Hayward, CA.
- Nau, R. (2006). The shape of incomplete preferences. *Ann Stat*, 34:2430–2448.
- Pivato, M. (2013). Multiutility representations for incomplete difference preorders. *Math Soc Sci*, 66:196–220.
- Range, T. M. and Osterdal, L. (2013). Checking bivariate first order dominance. *Technical report, Discussion Papers on Business and Economics (2013)*. Available under: [http://static.sdu.dk/mediafiles/0/C/2/%7B0C290F51-8643-421F-A9EF-F38D17A7F570%7Ddpbe9\\_2013.pdf](http://static.sdu.dk/mediafiles/0/C/2/%7B0C290F51-8643-421F-A9EF-F38D17A7F570%7Ddpbe9_2013.pdf).
- Schervish, M., Seidenfeld, T., Kadane, J., and Levi, I. (2003). Extensions of expected utility theory and some limitations of pairwise comparisons. In Bernard, J.-M., Seidenfeld, T., and Zaffalon, M., editors, *Proceedings of the Third International Symposium on Imprecise Probability: Theories and Applications*, pages 496–510. Carleton Scientific.
- 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., and Quaeghebeur, E., editors, *Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 277–286. Aracne.
- Schollmeyer, G., Jansen, C., and Augustin, T. (2017). Detecting stochastic dominance for poset-valued random variables as an example of linear programming on closure systems. *Technical Report 209, Department of Statistics, LMU Munich (2017)*. Available under: <https://epub.ub.uni-muenchen.de/40450/>.
- Seidenfeld, T., Kadane, J., and Schervish, M. (1995). A representation of partially ordered preferences. *Ann Stat*, 23:2168–2217.
- Shapley, L. (1971). Cores of convex games. *Int J Game Theory*, 1:11–26.
- Smets, P. (2005). Decision making in the tbm: the necessity of the pignistic transformation. *Int J Approx Reason*, 38:133–147.
- Tarp, F. and Osterdal, L. (2007). Multivariate discrete first order stochastic dominance. *Discussion Papers 07-23, University of Copenhagen. Department of Economics (2007)*. Available under: <http://www.economics.ku.dk/research/publications/wp/2007/0723.pdf>.
- Troffaes, M. (2007). Decision making under uncertainty using imprecise probabilities. *Int J Approx Reason*, 45:17–29.
- Troffaes, M. C. M. and Sahlin, U. (2017). Imprecise swing weighting for multi-attribute utility elicitation based on partial preferences. In Antonucci, A., Corani, G., Couso, I., and Destercke, S., editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 333–345. PMLR.
- Utkin, L. and Augustin, T. (2005). Powerful algorithms for decision making under partial prior information and general ambiguity attitudes. In Cozman, F., Nau, R., and Seidenfeld, T., editors, *Proceedings of the Fourth International Symposium on Imprecise Probability: Theories and Applications*, pages 349–358. SIPTA.
- Utkin, L. and Augustin, T. (2007). Decision making under incomplete data using the imprecise dirichlet model. *Int J Approx Reason*, 44:322–338.
- Wald, A. (1949). Statistical decision functions. *Ann Math Stat*, 20:165–205.
- Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.
- Walley, P. and Fine, T. (1982). Towards a frequentist theory of upper and lower probability. *Ann Stat*, 10:741–761.

- Wallner, A. (2007). Extreme points of coherent probabilities on finite spaces. *Int J Approx Reason*, 44:339–357.
- Weichselberger, K. and Pöhlmann, S. (1990). *A Methodology for Uncertainty in Knowledge-based Systems*. Springer, Heidelberg.