Christoph Jansen, Georg Schollmeyer and Thomas Augustin

Quantifying Degrees of E-admissibility in Decision Making with Imprecise Probabilities

Technical Report Number 215, 2018
Department of Statistics
University of Munich

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

Quantifying Degrees of E-admissibility in Decision Making with Imprecise Probabilities

C. Jansen[a,*], G. Schollmeyer[a], T. Augustin[a]

[a]Department of Statistics, LMU Munich, Ludwigsstr. 33, D-80539 Munich, Germany

Abstract

This paper is concerned with decision making using imprecise probabilities. In the first part, we introduce a new decision criterion that allows for explicitly modeling how far decisions that are optimal in terms of Walley's maximality are accepted to deviate from being optimal in the sense of Levi's E-admissibility. For this criterion, we also provide an efficient and simple algorithm based on linear programming theory. In the second part of the paper, we propose two new measures for quantifying the extent of E-admissibility of an E-admissible act, i.e. the size of the set of measures for which the corresponding act maximizes expected utility. The first measure is the maximal diameter of this set, while the second one relates to the maximal barycentric cube that can be inscribed into it. Also here, for both measures, we give linear programming algorithms capable to deal with them. Finally, we discuss some ideas in the context of ordinal decision theory. The paper concludes with a stylized application example illustrating all introduced concepts.

Keywords: Decision Making under Uncertainty; Imprecise Probabilities; E-admissibility; Maximality; Linear Programming; Ordinal Decision Theory; Stochastic Dominance

1. Introduction

A fair amount of the challenges arising in the modern sciences, e.g. parameter estimation and hypothesis testing in statistics, modeling an agent's preferences and choice behavior in philosophy and economics or the formalization of game theoretic problems, can be embedded into the formal framework of decision theory under uncertainty. If moreover the uncertainty underlying the decision situation is describable by some classical probability measure on the space of uncertain states of nature, we find ourselves within the framework of maximizing expected utility and we can draw on the whole toolbox of this well-investigated and elegant mathematical theory.

However, it is well known that in practice the necessity to specify a precise (i.e. classical) probability measure on the space of uncertain states might involve too strong consistency conditions regarding the beliefs of the decision maker of interest: It for instance might be the case that some decision maker finds it highly probable, say at least 0.8, that she will have had dinner in some restaurant by 9 p.m. tonight. However, since she doesn't know at all what the city she is traveling to has to offer, she cannot split this belief among different types of restaurants. That is even so if she made the (rather simplifying) assumption that the above probability exactly equals 0.8, there

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

is too less information for inferring, for instance, the probability of the dinner to take place in a Chinese restaurant. For situations of this kind (and also for less artificial ones), working with imprecise probabilities (Walley (1991); Weichselberger (2001), see also, e.g., Augustin et al. (2014) for an introduction) has become more and more attractive recently, since these allow for utilizing also partial probabilistic knowledge without the necessity of making assumptions that aren't met.

Much work has been undertaken on decision making with imprecise probabilities, and several strategies for optimal decision making have been proposed. Surveys of the theory are given in Seidenfeld (2004b); Troffaes (2007); Etner et al. (2012); Huntley et al. (2014). For original sources see, e.g., Kofler and Menges (1976); Levi (1974); Walley (1991); Gilboa and Schmeidler (1989). In the present paper, we contribute some new insights especially in the context of Levi's E-admissibility.

The paper is organized as follows: In Section 2, we recall the basic model of finite decision theory (Section 2.1) and the most commonly applied decision principles from precise and imprecise decision making (Section 2.2) for reference. Section 3 is divided into two parts: In Section 3.1 we contrast the criteria $\mathcal{M}$-maximality and E-admissibility and introduce a new decision criterion that in some sense lies in between the two. In Section 3.2, we propose two measures, one optimistic and one pessimistic, for quantifying the extent of E-admissibility of some E-admissible acts under consideration. In Section 4 we discuss decision problems in which the utility function is only interpretable in terms of an ordinal utility representation, however, utility differences have no meaning. Again, we recall and discuss criteria for both the precise and the imprecise case. In Section 5, we analyze a stylized application example and apply the theory developed in the paper. Section 6 concludes.

2. The Basic Model

We start our discussion by recalling the classical setup of decision making under uncertainty in Section 2.1 and the most commonly applied decision criteria under different types of uncertainty in Section 2.2 for reference.

2.1. Framework

Throughout most parts of the paper[1], we will consider the common model of finite decision theory: Some agent (or decision maker) is asked to decide which act a_i to choose from a finite set $\mathbb{A} = \{a_1, \ldots, a_n\}$ of available acts. However, the utility of the chosen act is fraught with uncertainty: it depends on which state of nature from a finite set $\Theta = \{\theta_1, \ldots, \theta_m\}$ of possible states corresponds to the true description of reality. Specifically, we assume that the utility of every pair $(a, \theta) \in \mathbb{A} \times \Theta$ can be evaluated by some real-valued cardinal utility function $u : \mathbb{A} \times \Theta \to \mathbb{R}$ that is unique up to a positive linear transformation.[2] We denote by $u_{ij} := u(a_i, \theta_j)$ the utility of choosing a_i given θ_j is the true state. For every act $a \in \mathbb{A}$, the utility function u is naturally associated with a random variable $u_a : (\Theta, 2^\Theta) \to \mathbb{R}$ defined by $u_a(\theta) := u(a, \theta)$ for all $\theta \in \Theta$. Similarly, for every $\theta \in \Theta$, we can define a random variable $u^\theta : (\mathbb{A}, 2^\mathbb{A}) \to \mathbb{R}$ by setting $u^\theta(a) := u(a, \theta)$ for all $a \in \mathbb{A}$. The structure of the basic model is visualized in Table 1.

Depending on the situation, the standard model will sometimes be extended for randomized acts, which are classical probability measures λ on $(\mathbb{A}, 2^\mathbb{A})$. Choosing λ is then interpreted as

[1]The one exception is the discussion in Section 4, where we do not assume a cardinal utility representation.

[2]See Schervish et al. (2013) for the situation where multiple utilities through different currencies are available and exchange rates have to be taken into account.

Table 1: The basic model of finite decision theory.

$u(a_i, \theta_j)$	$\theta_\mathbf{1}$	$\cdots$	$\theta_\mathbf{m}$
$\mathbf{a_1}$	$u(a_1, \theta_1)$	$\cdots$	$u(a_1, \theta_m)$
$\vdots$	$\vdots$	$\cdots$	$\vdots$
$\mathbf{a_n}$	$u(a_n, \theta_1)$	$\cdots$	$u(a_n, \theta_m)$

leaving the final decision to a random experiment which yields act a_i with probability $\lambda(\{a_i\})$. We denote the set of randomized acts on $(\mathbb{A}, 2^{\mathbb{A}})$ by $G(\mathbb{A})$.

If also randomized acts are considered, the original utility function u on $\mathbb{A} \times \Theta$ can straightforwardly be extended to a utility function $G(u)$ on $G(\mathbb{A}) \times \Theta$ by assigning each pair (λ, θ) the expectation of the random variable u^θ under the measure λ, i.e. $G(u)(\lambda, \theta) := \mathbb{E}_\lambda\big[u^\theta\big]$, i.e. the expectation of utility that choosing the randomized act λ will lead to, given θ is the true description of reality. Every act $a \in \mathbb{A}$, sometimes called pure act when the difference to randomized acts needs to be emphasized, then can uniquely be identified with the Dirac-measure $\delta_a \in G(\mathbb{A})$, and we have $u(a, \theta) = G(u)(\delta_a, \theta)$ for all $(a, \theta) \in \mathbb{A} \times \Theta$. Again, for every $\lambda \in G(\mathbb{A})$ fixed, the extended utility function $G(u)$ is associated with a random variable $G(u)_\lambda$ on $(\Theta, 2^\Theta)$ by setting $G(u)_\lambda(\theta) := G(u)(\lambda, \theta)$ for all $\theta \in \Theta$. We refer to the triplet $(\mathbb{A}, \Theta, u)$ as the (finite) decision problem and to the triplet $(G(\mathbb{A}), \Theta, G(u))$ as the corresponding randomized extension.

Finally, note that the standard model of decision theory also contains statistical estimation and hypothesis testing problems as special cases: If we, in addition to the basic problem $(\mathbb{A}, \Theta, u)$, observe some random variable $X : \Omega \to \mathcal{X}$ such that $X \sim P_\theta$ if $\theta \in \Theta$ is the true state of nature, that is we know the distribution of the random experiment if we know the true state, then statistical procedures can be viewed as decision functions $d : \mathcal{X} \to \mathbb{A}$ that map observed data on acts. The utility function u of the original problem then very naturally can be extended to a gain function $U : \mathcal{D} \times \Theta \to \mathbb{R}$ for evaluating decision functions by setting $U(d, \theta) := \mathbb{E}_{P_\theta}[u^\theta \circ d]$. Here, $\mathcal{D}$ denotes some appropriate set of possible statistical procedures. Formally, the resulting triplet $(\mathcal{D}, \Theta, U)$ then again can be viewed as a basic decision problem. Thus, even if we do not explicitly formulate our results for statistical procedures in the following, they always also can be interpreted in a statistical context.[3]

2.2. Criteria for Decision Making

Given a decision model $(\mathbb{A}, \Theta, u)$ of the form just recalled, the challenge is quickly explained: Determine an (in some sense) optimal act $a^* \in \mathbb{A}$ (or, depending on the context, optimal randomized act $\lambda^* \in G(\mathbb{A})$). The subtlety rather comes with the definition of the term optimality, since any

[3]It should, however, be emphasized that in the context of imprecise probabilistic models (like for instance credal sets or interval probabilities) the relationship between optimal decision functions in terms of prior risk and posteriori loss optimal acts may be more subtle than in the context of precise probability: the main theorem of Bayesian decision theory may fail (cf., e.g., Augustin (2003, Section 2.3)). This failure is in essence a variant of the general phenomenon of potential sequential incoherence in decision making and discrepancy between extensive and normal forms, as investigated in depth by Seidenfeld (e.g., Seidenfeld (1988, 1994)). Immediate counter-examples arise from the phenomenon of dilation, which has intensively been studied by Seidenfeld and co-authors (cf., e.g., Seidenfeld (1994), Seidenfeld and Wassermann (1993), Wassermann and Seidenfeld (1994)), see also, e.g., Liu (2015).

meaningful definition necessarily has to depend on (what the decision maker assumes about) the mechanism generating the states of nature. Here, traditional decision theory mainly covers two extreme poles: (I) The generation of the states follows a known classical probability measure π on $(\Theta, 2^{\Theta})$ or (II) it can be compared to a game against an omniscient enemy. In these cases optimality is almost unanimously defined by the following two well-known principles:

(I) *Maximizing Expected Utility:* Label any act $a^* \in \mathbb{A}$ optimal that receives maximal expected utility with respect to π, i.e. for which $\mathbb{E}_\pi[u_{a^*}] \geq \mathbb{E}_\pi[u_a]$ for all other $a \in \mathbb{A}$. We denote by $\mathbb{A}_\pi$ the set of all acts from $\mathbb{A}$ that maximize expected utility with respect to π.

(II) *Wald's Maximin Principle:* Label any act $a^* \in \mathbb{A}$ optimal that receives highest possible utility value under that state that is worst possible for this particular act, i.e. for which $\min_{\theta \in \Theta} u(a^*, \theta) \geq \min_{\theta \in \Theta} u(a^*, \theta)$. We denote by $\mathbb{A}_W$ the set of all maximin acts.

Straightforwardly, principles (I) and (II) generalize to randomized acts, and we will denote the corresponding sets of optimal randomized acts by $G(\mathbb{A})_\pi$ and $G(\mathbb{A})_W$, respectively. In contrast, defining optimality of acts becomes less obvious if (A) the probability measure π is only partially known (case of imprecise probabilities) or (B) there is uncertainty about the complete appropriateness of it (case of uncertainty about precise probabilities). In situation (A), one commonly assumes that the available probabilistic information is describable by a polyhedral[4] set $\mathcal{M}$ of probability measures on $(\Theta, 2^{\Theta})$ of the form

$$\mathcal{M} := \left\{\pi \middle| \ \underline{b}_s \leqslant \mathbb{E}_\pi(f_s) \leqslant \overline{b}_s \ \forall s = 1, ..., r\right\} \tag{1}$$

where, for all $s = 1, ..., r$, we have $(\underline{b}_s, \overline{b}_s) \in \mathbb{R}^2$ such that $\underline{b}_s \leqslant \overline{b}_s$[5] and $f_s : \Theta \to \mathbb{R}$, which is an example for a imprecise probabilistic model. Specifically, the available information is assumed to be describable by lower and upper bounds for the expected values of a finite number of random variables on the space of states. Note that this also includes models in which the uncertainty arises from a variety of different (possibly precise) expert opinions: If, for instance, each from a bunch of experts gives precise expected payoff estimates for a number of stocks, we take $\mathcal{M}$ to be the set of probabilities that yield for every stock an expectation that ranges within the lowest and the highest expert guess. Most simply, this includes also the case where every expert specifies a precise probability measure on the state space, since a probability measure is always representable by a family of indicator functions. The picture of $\mathcal{M}$ being the opinions of a committee of experts will be used at different points in the paper (similarly as also done in, e.g., Bradley (2015)).

Under an imprecise probabilistic model of form (1), several optimality criteria for decision making had been proposed (cf., e.g., Troffaes (2007); Etner et al. (2012); Huntley et al. (2014) for general surveys and Utkin and Augustin (2005); Kikuti et al. (2011); Hable and Troffaes (2014); Jansen et al. (2017a) for computational aspects). We now briefly recall the ones among them which are most important for our purposes:

[4]See, however, e.g., Wheeler (2012), Majo-Wilson and Wheeler (2016, Section 2), and the references therein, for arguments to consider also non-convex sets of probabilities.

[5]For technical convenience we assume, wlog, that $0 \in [\underline{b}_s, \overline{b}_s]$ for all $s = 1, ..., r$ in the following. Note that if $\mathcal{M}$ is described by functions $(f_1, \ldots, f_r)$ and bounds $((\underline{b_1}, \overline{b}_1), \ldots, (\underline{b_r}, \overline{b}_r))$ not meeting this assumptions, we can always equivalently characterize it by functions $(f_1 - c_1, \ldots, f_r - c_r)$ and bounds $((\underline{b_1} - c_1, \overline{b}_1 - c_1), \ldots, (\underline{b_r} - c_r, \overline{b}_r - c_r))$, where, for all $s = 1, \ldots, r$, we set $c_s = \underline{b}_s$ if $\underline{b}_s > 0$ and $c_s = -\overline{b}_s$ if $\overline{b}_s < 0$ and $c_s = 0$ if $0 \in [\underline{b}_s, \overline{b}_s]$.

(A_1) $\mathcal{M}$-*Maximin* ($\mathcal{M}$-*Maximax*):[6] Label any act $a^* \in \mathbb{A}$ optimal that maximizes expected utility with respect to the worst (best) compatible probability measure, i.e. for which $\min_{\pi \in \mathcal{M}} \mathbb{E}_\pi(u_{a^*}) \geq \min_{\pi \in \mathcal{M}} \mathbb{E}_\pi(u_a)$ (resp. $\max_{\pi \in \mathcal{M}} \mathbb{E}_\pi(u_{a^*}) \geq \max_{\pi \in \mathcal{M}} \mathbb{E}_\pi(u_a)$) for all $a \in \mathbb{A}$. We denote by $\mathbb{A}_{\underline{\mathcal{M}}}$ (resp. $\mathbb{A}_{\overline{\mathcal{M}}}$) the set of $\mathcal{M}$-maximin (resp. $\mathcal{M}$-maximax) acts.

Clearly, $\mathcal{M}$-maximin is a rather pessimistic criterion that reflects the attitude of decision makers that react averse to the ambiguity between the different compatible probabilities from $\mathcal{M}$. Contrarily, $\mathcal{M}$-maximax reflects the attitude of ambiguity seeking agents. Note also that in the extreme cases where either the credal set $\mathcal{M}$ is the set of all precise probability measures (vacuousness) or it contains only one such measure (ideal stochasticity), the criterion $\mathcal{M}$-maximin reduces to Wald's maximin principle or precise expectation maximization, respectively.

(A_2) $\mathcal{M}$-*Maximality*:[7] Label any act $a^* \in \mathbb{A}$ optimal that dominates every other available act $a \in \mathbb{A}$ in expectation with respect to at least one probability measure $\pi_a \in \mathcal{M}$, i.e. if for every $a \in \mathbb{A}$ there exists $\pi_a \in \mathcal{M}$ such that $\mathbb{E}_{\pi_a}(u_{a^*}) \geq \mathbb{E}_{\pi_a}(u_a)$. We denote the set of all $\mathcal{M}$-maximal acts by $\mathbb{A}_{\max}$.

The idea of $\mathcal{M}$-maximality thus is to exclude every act a_0 from the decision problem for which there exists another act a_1 that dominates it with respect to every compatible probability measure. Note that $\mathcal{M}$-maximality can be viewed as a *local* decision criterion: The preference between the acts a_0 and a_1 is independent of the other available acts in $\mathbb{A} \setminus \{a_0, a_1\}$ or, as Schervish et al. (2003)[8] puts it, $\mathcal{M}$-maximality is induced by pairwise comparisons of acts in $\mathbb{A}$ only. Note further that, in the extreme case of $\mathcal{M}$ being a singleton, the criterion reduces to classical expectation maximization.

(A_3) *E-Admissibility*:[9] Label an act $a^* \in \mathbb{A}$ optimal if it maximizes expected utility among all other available acts with respect to a least one compatible probability measure, i.e. if there exists $\pi^* \in \mathcal{M}$ such that $a^* \in \mathbb{A}_{\pi^*}$. We denote by $\mathbb{A}_{\mathcal{M}}$ the set of all E-admissible acts from $\mathbb{A}$ with respect to the credal set $\mathcal{M}$.

In contrast to $\mathcal{M}$-maximality, the concept of E-admissibility can rather by viewed as a *global* decision criterion: In order to be able to build a preference between two acts a_0 and a_1, utility information for all the other available acts in $\mathbb{A} \setminus \{a_0, a_1\}$ is required. To put it in the words of Schervish et al. (2003) again: E-admissibility, in general, is not induced by pairwise comparisons of acts in $\mathbb{A}$ only. Again, in the case of ideal stochasticity the criterion reduces to classical expectation maximization. Contrarily, in the case of vacuousness every act that is not dominated by another act in every state is E-admissible.[10]

Again, if randomized acts are of interest, we denote the corresponding optimal sets by $G(\mathbb{A})_{\underline{\mathcal{M}}}$, $G(\mathbb{A})_{\overline{\mathcal{M}}}$, $G(\mathbb{A})_{\max}$ and $G(\mathbb{A})_{\mathcal{M}}$. As easy to see, it holds that $G(\mathbb{A})_\pi = \text{conv}(\tilde{\mathbb{A}}_\pi)$, where we have $\tilde{\mathbb{A}}_\pi := \{\delta_a : a \in \mathbb{A}_\pi\}$ and $\text{conv}(S)$ denotes the convex hull of S. Thus, we can easily construct

[6] See, for instance, Kofler and Menges (1976) and Gilboa and Schmeidler (1989). Many authors denote $\mathcal{M}$ by Γ, and thus the name Γ-maximin is common as well.

[7] This criterion is mainly advocated by Walley (1991) and work following him.

[8] Compare also Seidenfeld et al. (2010) and Vicig and Seidenfeld (2012, Section 3).

[9] This criterion is introduced by Levi (1974).

[10] Both criteria just discussed are also of high interest in forecasting with imprecise probabilities. While for imprecise probabilities there is no real-valued strictly proper scoring rule, it is possible to formulate an appropriate lexicographic strictly proper scoring rule with respect to $\mathcal{M}$-maximinity and E-admissibility, supplemented by $\mathcal{M}$-maximinity (Seidenfeld et al., 2012).

the set of randomized actions that maximize expected utility with respect to π by taking all convex combinations of pure acts with the same property. This fact is often used to argue that randomization does not pay out in the context of maximizing expected utility. Moreover, as shown by Walley (cf., Walley (1991, p. 163)) and emphasized in Schervish et al. (2003), we have that $G(\mathbb{A})_{\max} = G(\mathbb{A})_{\mathcal{M}}$, i.e. in the context of randomized acts the criteria $\mathcal{M}$-maximality and E-admissibility coincide in the sense of selecting the same optimal acts.

To complete the section, we now recall one criterion of optimality for situation (B), i.e. the case of an uncertain precise probability π: The criterion of Hodges and Lehmann (cf. Hodges and Lehmann (1952)). One motivation of this decision principle is to model the decision maker's skepticism in the available probability measure more directly. It is defined as follows:

(B_1) *Hodges and Lehmann Optimality:* Label any act $a^* \in \mathbb{A}$ optimal that maximizes the term $\alpha\mathbb{E}_\pi(u_{a^*}) + (1-\alpha)\min_\theta u(a^*, \theta)$ among all other acts $a \in \mathbb{A}$, that is which maximizes a weighted sum of the expected utility and the worst state utility. The value $\alpha \in [0,1]$ expresses the degree of trust that the agent assigns to the probability measure π.

Note that Hodges and Lehmann optimality can be viewed as a special case of $\mathcal{M}$-maximinity (cf., for instance, Jansen et al. (2017a)): If the underlying credal set is chosen to arise from an ε-contamination model (a.k.a. linear-vacuous mixture model) having the form

$$\mathcal{M}_{(\pi_0,\varepsilon)} := \big\{(1-\varepsilon)\pi_0 + \varepsilon\pi : \pi \in \mathcal{P}(\Theta)\big\}$$

where $\mathcal{P}(\Theta)$ is the set of all probabilities on $(\Theta, 2^\Theta)$, $\varepsilon > 0$ is a fixed contamination parameter and $\pi_0 \in \mathcal{P}(\Theta)$ is the central distribution, it holds

$$\begin{aligned} \underline{\mathbb{E}}_{\mathcal{M}_{(\pi_0,\varepsilon)}}(X) &= \min\nolimits_{\pi\in\mathcal{P}(\Theta)}((1-\varepsilon)\mathbb{E}_{\pi_0}(X) + \varepsilon\mathbb{E}_\pi(X)) \\ &= (1-\varepsilon)\mathbb{E}_{\pi_0}(X) + \varepsilon \min\nolimits_{\pi\in\mathcal{P}(\Theta)} \mathbb{E}_\pi(X) \\ &= (1-\varepsilon)\mathbb{E}_{\pi_0}(X) + \varepsilon \min\nolimits_{\theta\in\Theta} X(\theta) \end{aligned}$$

for arbitrary random variables $X : (\Theta, 2^\Theta) \to \mathbb{R}$. Thus, maximizing the lower expectation w.r.t. the ε-contamination model is equivalent to maximizing the Hodges and Lehmann-criterion with trust parameter $(1-\varepsilon)$ and probability π_0. This connection is also of interest for Bayesian statistical inference with imprecise probabilities: As pointed out by Seidenfeld and Wassermann (1996) in the discussion of Walley (1996) and made explicit in Herron et al. (1997), the well-investigated Imprecise Dirichlet Model (IDM) for generalized Bayesian statistical learning is mathematically equivalent to an ε-contamination model with the relative frequencies as the central distribution π_0. Taking into account the above calculation, this also shows a very close relation between decision making in the IDM (e.g., Utkin and Augustin (2007)) and the criterion suggsted by Hodges and Lehmann.

3. *E*-admissibility, Maximality and a Criterion in between

We start our discussion by setting focus on the criteria $\mathcal{M}$-maximality and E-admissibility and develop some new ideas in this context. The discussion is divided in two main parts: In Section 3.1 we briefly compare the two criteria and then propose a new criterion providing an adjustable trade-off between them, for which we also derive a linear programming based algorithm. Afterwards, in Section 3.2, we discuss a new measure for quantifying the extent of E-admissibility of an E-admissible act of interest.

3.1. Comparing E-admissibility and Maximality

As already seen in the previous section, when considering also randomized acts, the concepts of $\mathcal{M}$-maximality and E-admissibility with respect to $\mathcal{M}$ induce the same optimal acts and therefore coincide. However, for a finite (or more general non-convex) set of acts $\mathbb{A}$ the criterion of $\mathcal{M}$-maximality is the strictly weaker condition in the sense that $\mathbb{A}_{\mathcal{M}} \subset \mathbb{A}_{\max}$. Our first result describes how to construct the set $G(\mathbb{A})_{\mathcal{M}}$ of all randomized E-admissible acts (and therefore also the set $G(\mathbb{A})_{\max}$ of randomized $\mathcal{M}$-maximal acts) from the set $\mathbb{A}_{\mathcal{M}}$ of pure E-admissible acts.

Proposition 1. *Let the decision problems* $(\mathbb{A}, \Theta, u)$ *and* $(G(\mathbb{A}), \Theta, G(u))$ *and the sets* $\mathcal{M}$, $\mathbb{A}_\pi$, $\mathbb{A}_{\mathcal{M}}$, $G(\mathbb{A})_\pi$, $G(\mathbb{A})_{\mathcal{M}}$ *be defined as before. The following holds:*

$$G(\mathbb{A})_{\mathcal{M}} = \bigcup\nolimits_{\pi \in \mathcal{M}} \mathrm{conv}(\tilde{\mathbb{A}}_\pi)$$

where $\tilde{\mathbb{A}}_\pi := \{\delta_a : a \in \mathbb{A}_\pi\}$ *and* $\mathrm{conv}(S)$ *denotes the convex hull of a set* S*.*

Proof. $\subset$: Let $\lambda^* \in G(\mathbb{A})_{\mathcal{M}}$. Choose $\pi^* \in \mathcal{M}$ such that $\lambda^* \in G(\mathbb{A})_{\pi^*}$. Assume, for contradiction, there exists $a_0 \in \mathbb{A}$ such that $\lambda^*(\{a_0\}) > 0$ and $a_0 \notin \mathbb{A}_{\pi^*}$. Pick then $a_1 \in \mathbb{A}_{\pi^*}$ and define a randomized act $\lambda_0 \in G(\mathbb{A})$ by setting $\lambda_0(\{a\}) := \lambda^*(\{a\})$ for $a \in \mathbb{A} \setminus \{a_0, a_1\}$, $\lambda_0(\{a\}) := 0$ for $a = a_0$ and $\lambda_0(\{a\}) := \lambda^*(\{a_0, a_1\})$ for $a = a_1$. Then, the following calculation are immediate:

$$\begin{aligned}
\mathbb{E}_{\pi^*}[G(u)_{\lambda_0}] &= \sum_{a \in \mathbb{A}} \lambda_0(\{a\})\mathbb{E}_{\pi^*}(u_a) \\
&= \sum_{a \in \mathbb{A} \setminus \{a_0, a_1\}} \lambda^*(\{a\})\mathbb{E}_{\pi^*}(u_a) \quad + \quad \lambda^*(\{a_0, a_1\})\mathbb{E}_{\pi^*}(u_{a_1}) \\
&> \sum_{a \in \mathbb{A}} \lambda^*(\{a\})\mathbb{E}_{\pi^*}(u_a) \quad = \quad \mathbb{E}_{\pi^*}[G(u)_{\lambda^*}]
\end{aligned}$$

This contradicts $\lambda^* \in G(\mathbb{A})_{\pi^*}$. Therefore, we have $\lambda^* \in \mathrm{conv}(\tilde{\mathbb{A}}_{\pi^*})$.

$\supset$: Let conversely $\lambda^* \in \bigcup_{\pi \in \mathcal{M}} \mathrm{conv}(\tilde{\mathbb{A}}_\pi)$. Then there exists $\pi^* \in \mathcal{M}$ such that $\lambda^* \in \mathrm{conv}(\tilde{\mathbb{A}}_{\pi^*})$ and we have $\mathbb{E}_{\pi^*}[G(u)_{\lambda^*}] = \mathbb{E}_{\pi^*}(u_a)$ for all $a \in \mathbb{A}_{\pi^*}$. Choose $a_0 \in \mathbb{A}_{\pi^*}$ and observe that for arbitrary $\lambda \in G(\mathbb{A})$ it holds that

$$\mathbb{E}_{\pi^*}[G(u)_\lambda] = \sum_{a \in \mathbb{A}} \lambda(\{a\})\mathbb{E}_{\pi^*}(u_a) \ \leq \ \mathbb{E}_{\pi^*}(u_{a_0}) = \mathbb{E}_{\pi^*}[G(u)_{\lambda^*}]$$

Thus there exists $\pi^* \in \mathcal{M}$ with respect to which λ^* maximizes expected utility implying that $\lambda^* \in G(\mathbb{A})_{\mathcal{M}}$. □

Since Proposition 1 allows us to construct both sets $G(\mathbb{A})_{\mathcal{M}}$ and $G(\mathbb{A})_{\max}$ once having computed the set $\mathbb{A}_{\mathcal{M}}$ of pure E-admissible acts, we restrict analysis to non-randomized acts for the rest of the section. For this setting, we now propose a new decision criterion that allows for labeling only such $\mathcal{M}$-maximal acts optimal that are not too far from being E-admissible with respect to $\mathcal{M}$ in the sense that the probabilities for which the corresponding act expectation dominates the other acts differ not too much from each other. The deviation of an act from E-admissibility can be explicitly controlled by an additional parameter ε.

Definition 1. *Let* $(\mathbb{A}, \Theta, u)$ *and* $\mathcal{M}$ *be defined as before and let* $\varepsilon \geq 0$*. An act* $a^* \in \mathbb{A}$ *is called* E_ε*-admissible if there exists a family of probability measures* $(\pi_a)_{a \in \mathbb{A}}$ *from* $\mathcal{M}$ *such that the following two conditions are satisfied:*

i) $\mathbb{E}_{\pi_a}(u_{a^*}) \geq \mathbb{E}_{\pi_a}(u_a)$ *for all* $a \in \mathbb{A}$ *and*

ii) $\| \pi_a - \pi_{a'} \| \leq \varepsilon$ *for all* $a, a' \in \mathbb{A}$*, where* $\| \cdot \|$ *denotes a norm on* $\mathcal{M}$*.*

We denote by $\mathbb{A}_\varepsilon$ *the set of all* E_ε*-admissible acts from* $\mathbb{A}$*.*

Remark 1. *Obviously, the set of* E_0*-admissible acts coincides with the set of* E*-admissible acts with respect to* $\mathcal{M}$*, i.e.* $\mathbb{A}_0 = \mathbb{A}_{\mathcal{M}}$*. Moreover, for* ε^* *chosen sufficiently large, namely* $\varepsilon^* \geq b := \sup_{\pi,\pi' \in \mathcal{M}} \| \pi - \pi' \|$*, the set of* E_{ε^*}*-admissible acts coincides with the set of* $\mathcal{M}$*-maximal acts, i.e.* $\mathbb{A}_{\varepsilon^*} = \mathbb{A}_{max}$*. For* $\varepsilon \in (0, b)$*, it usually will hold that* $\mathbb{A}_{\mathcal{M}} \subsetneq \mathbb{A}_\varepsilon \subsetneq \mathbb{A}_{max}$ *and the set* $\mathbb{A}_\varepsilon$ *then exactly contains those* $\mathcal{M}$*-maximal acts that are not too far (controlled by* ε*) from being* E*-admissible. Moreover, the criterion is monotone in the sense that for* $\varepsilon_1 \geq \varepsilon_2$ *we have that* $\mathbb{A}_{\varepsilon_2} \subset \mathbb{A}_{\varepsilon_1}$*. Thus, it allows for ranking* $\mathcal{M}$*-maximal acts with respect to their deviation from being* E*-admissible.*

If we again take the point of view that $\mathcal{M}$ arises from different expert opinions, it turns out that the criterion of E_ε-admissibility is based on a quite convincing intuition: Consider for instance a political decision maker that consults an advisory body of experts when it comes to facing difficult decisions. In this situation, applying E-admissibility corresponds to only choosing acts which one fixed expert labels optimal among all other options. Contrarily, in terms of $\mathcal{M}$-maximality an act is already optimal if for each other act there is at least one expert preferring the former to the latter, no matter how different the involved experts are in opinion. Here, E_ε-admissibility builds a bridge between these two extremes: While the decision maker can still make use of opinions of different experts, she nevertheless can explicitly control by an additional parameter ε how strong the experts contributing to the decision process are allowed to differ in opinion.

We now provide an algorithm for checking whether an act in a given decision problem is E_ε-admissible for a fixed value ε. It turns out that this, provided the L_1-norm is used for measuring the distances between the elements of $\mathcal{M}$, can be done by solving one single, relatively simple, linear programming problem. We arrive at the following proposition.

Proposition 2. *Let* $(\mathbb{A}, \Theta, u)$ *and* $\mathcal{M}$ *be defined as before and let* $\varepsilon \geq 0$*. For some act* $a_z \in \mathbb{A}$*, consider the following linear programming problem:*

$$\sum_{i=1}^{n} \Big(\sum_{j=1}^{m} \gamma_{ij} \Big) \longrightarrow \max_{(\gamma_{11},\ldots,\gamma_{nm})} \tag{2}$$

with constraints $(\gamma_{11}, \ldots, \gamma_{nm}) \geqslant 0$ *and*

- $\sum_{j=1}^{m} \gamma_{ij} \leqslant 1$ *for all* $i = 1, \ldots, n$
- $\underline{b}_s \leqslant \sum_{j=1}^{m} f_s(\theta_j) \cdot \gamma_{ij} \leqslant \overline{b}_s$ *for all* $s = 1, ..., r$, $i = 1, \ldots, n$
- $\sum_{j=1}^{m} (u_{ij} - u_{zj}) \cdot \gamma_{ij} \leqslant 0$ *for all* $i = 1, \ldots, n$
- $\sum_{j=1}^{m} |\gamma_{i_1 j} - \gamma_{i_2 j}| \leq \varepsilon$ *for all* $i_1 > i_2 \in \{1, \ldots, n\}$

Then $a_z \in \mathbb{A}$ *is* E_ε*-admissible iff the optimal outcome of (2) equals* n*.*

Proof. Clearly, if (2) possesses an optimal solution $(\gamma_{11}^*, \ldots, \gamma_{nm}^*)$ yielding an objective value of n, then the constraints guarantee that setting $\pi_{a_i}(\{\theta_j\}) := \gamma_{ij}^*$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$ defines a family of probability measures $(\pi_{a_i})_{i=1,\ldots n,}$ from $\mathcal{M}$ satisfying the properties from Definition 1. Thus, $a_z \in \mathbb{A}_\varepsilon$.

If conversely $a_z \in \mathbb{A}_\varepsilon$, we can choose a family of probability measures $(\pi_{a_i})_{i=1,\ldots n,}$ from $\mathcal{M}$ satisfying the properties from Definition 1. One then easily verifies that setting $\gamma_{ij}^* := \pi_{a_i}(\{\theta_j\})$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$ defines an admissible solution $(\gamma_{11}^*, \ldots, \gamma_{nm}^*)$ to (2) that yields an objective value of n. □

Remark 2. *To see the linearity of the constraint $\sum_{j=1}^m |\gamma_{i_1 j} - \gamma_{i_2 j}| \leq \varepsilon$ for all $i_1 > i_2 \in \{1, \ldots, n\}$ in the above linear programming problem, one can proceed as follows: Add $2m$ decision variables $l_1, \ldots, l_m$ and $o_1, \ldots, o_m$ and replace the above constraints equivalently by the constraints $l_j \leq \gamma_{ij} \leq o_j$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$ as well as $\sum_{j=1}^m (o_j - l_j) \leq \varepsilon$. In sum, the programming problem (2) thus possesses $n+rn+n+nm+1 = n(1+r+m)+1$ constraints and $nm+m = m(n+1)$ decision variables.*

We conclude the section by illustrating the results so far by a brief toy example, which in parts is also discussed in Seidenfeld (2004a, p. 2) in order to demonstrate that $\mathcal{M}$-maximinity does not imply E-admissibility with respect to $\mathcal{M}$ and vice versa. We additionally show how the proposed concept of E_ε-admissibility can help to clarify analysis in such situations. The example reads as follows:

Example 1. *Consider the basic decision problem $(\mathbb{A}, \Theta, u)$ that is defined by the following utility table*

u_{ij}	θ_1	θ_2
a_1	1	0
a_2	0	1
a_3	4/10	4/10
a_4	6/10	11/35

Moreover, suppose the uncertainty on the states is modeled by the credal set

$$\mathcal{M} = \Big\{ \pi : 0.3 \leq \pi(\{\theta_1\}) \leq 0.8 \Big\}$$

In this case, we have $\mathbb{A}_{\mathcal{M}} = \{a_1, a_2\}$, $\mathbb{A}_{\max} = \mathbb{A}$ and $\mathbb{A}_{\underline{\mathcal{M}}} = \{a_3, a_4\}$. Thus, we have a situation with two different $\mathcal{M}$-maximin acts, which are both not E-admissible. In order to make a decision between the acts a_3 and a_4, we can apply the E_ε-criterion to see which of the two is closer to being E-admissible. We receive the following results:

outcome of (2) for	$\varepsilon = 0.1$	$\varepsilon = 0.2$	$\varepsilon = 0.3$	$\varepsilon = 0.6$
a_3	≈ 0.76	≈ 1.51	≈ 2.27	4
a_4	2.3	4	4	4

The results show that act a_4 is that act among the $\mathcal{M}$-maximin acts that is closer to E-admissibility, since it is E_ε-admissible already for an ε-level of 0.2 whereas act a_3 is not. Thus, it could be argued that it a_4 is preferable. Finally, if we additionally consider randomized acts, then Proposition 1 and the discussions in Section 2.2 imply that it holds

$$G(\mathbb{A})_{\mathcal{M}} = G(\mathbb{A})_{\max} = G(\mathbb{A})_\varepsilon = \mathrm{conv}(\{\delta_{a_1}, \delta_{a_2}\})$$

for arbitrary values of $\varepsilon \geq 0$.

3.2. The Extents of E-admissible acts

In the previous section, we considered acts optimal that are not too far from being E-admissible. We accordingly weakened the concept of E-admissibility towards acts that are in some sense almost E-admissible. In this section we take rather the opposed direction and address the following question: Given an E-admissible act $a \in \mathbb{A}_{\mathcal{M}}$ with respect to some credal set $\mathcal{M}$, how large is the set of compatible probability measures from $\mathcal{M}$ for which act a maximizes expected utility? If we again use the picture of $\mathcal{M}$ modeling the opinions of some committee of experts, the question translates as follows: How diverse can these experts be in opinion while still all sharing the view that act a is optimal?

In order to answer this question, we propose two measures for the extent of E-admissibility of acts in the following: The maximal extent and the uniform extent. While the first concept measures the maximal diameter of the set of measures for which the considered act maximizes expected utility, the latter one searches for a maximal set that can be inscribed into this set. Together, the two measures will be shown to give a pretty good impression about the extent of E-admissibility. We start by defining the concept of maximal extent.

Definition 2. *Let $(\mathbb{A}, \Theta, u)$ and $\mathcal{M}$ be defined as before and let $\| \cdot \|$ denote some norm on $\mathcal{M}$. Moreover, let $a \in \mathbb{A}_{\mathcal{M}}$ be an E-admissible act with respect to $\mathcal{M}$ and denote by $\mathcal{M}_a$ the set $\{\pi \in \mathcal{M} : a \in \mathbb{A}_\pi\}$. We define the (maximal) extent $ext_{\mathcal{M}}(a)$ of act a as the number*

$$ext_{\mathcal{M}}(a) := \sup_{\pi, \pi' \in \mathcal{M}_a} \| \pi - \pi' \|$$

i.e. as the maximum distance of probability measures $\pi, \pi' \in \mathcal{M}_a$ with respect to $\| \cdot \|$ for which act a maximizes expected utility.

Why is the measure $ext_{\mathcal{M}}(\cdot)$ sensible for the question motivating the section? To see that, first note that intuitively if $ext_{\mathcal{M}}(a)$ is large, then act a maximizes expected utility with respect to very different (in the sense of highly distant) probability measures from $\mathcal{M}$. To directly connect this observation to the size of the set $\mathcal{M}_a$, it is important to mention that $\mathcal{M}_a$ is a convex set and therefore all measure lying on the "line" between the two maximum distance measures again have to be contained in $\mathcal{M}_a$. Thus, $ext_{\mathcal{M}}(a)$ indeed can be viewed as a measure of the size of the set of probabilities for which act a is optimal and therefore is sensible for the above questions. The following proposition gives an algorithm for computing $ext_{\mathcal{M}}(a)$ by solving a series of linear programming problems for the case that $\| \cdot \| := \| \cdot \|_\infty$.

Proposition 3. *Let $(\mathbb{A}, \Theta, u)$ and $\mathcal{M}$ be defined as before and let $a_z \in \mathbb{A}_{\mathcal{M}}$. Consider, for every $j = 1, \dots, m$, the linear programming problem*

$$\gamma_{1j} - \gamma_{2j} \longrightarrow \max_{(\gamma_{11}, \dots, \gamma_{1m}, \gamma_{21}, \dots, \gamma_{2m})} \tag{P_j}$$

with constraints $(\gamma_{11}, \dots, \gamma_{1m}, \gamma_{21}, \dots, \gamma_{2m}) \geq 0$ and

- $\sum_{j=1}^m \gamma_{ij} = 1$ *for all $i = 1, 2$*
- $\underline{b}_s \leqslant \sum_{j=1}^m f_s(\theta_j) \cdot \gamma_{ij} \leqslant \overline{b}_s$ *for all $s = 1, ..., r$ and $i = 1, 2$*
- $\sum_{j=1}^m (u_{\ell j} - u_{zj}) \cdot \gamma_{ij} \leqslant 0$ *for all $i = 1, 2$ and $\ell = 1, \dots, n$*

Denote by $g(j)$ the optimal objective of problem (P_j). Then the maximal extent of act a with respect to $\|\cdot\|_\infty$ is given by $ext_{\mathcal{M}}(a) = \max_j g(j)$.

Proof. For $j \in \{1, \dots, m\}$, let $(\gamma_{11}^j, \dots, \gamma_{1m}^j, \gamma_{21}^j, \dots, \gamma_{2m}^j)$ denote an optimal solution to problem (P_j).[11] Then the constraints guarantee that by setting $\pi_1^j(\{\theta_t\}) := \gamma_{1t}^j$ and $\pi_2^j(\{\theta_t\}) := \gamma_{2t}^j$ for all $t = 1, \dots, m$ we define two measures $\pi_1^j, \pi_2^j \in \mathcal{M}_a$ with the property

i) $g(j) = |\pi_1^j(\{\theta_j\}) - \pi_2^j(\{\theta_j\})| \geq |\pi_1(\{\theta_j\}) - \pi_2(\{\theta_j\})|$ for all $\pi_1, \pi_2 \in \mathcal{M}_a$

Let $j^* \in \{1, \dots, m\}$ with $g(j^*) = |\pi_1^{j^*}(\{\theta_{j^*}\}) - \pi_2^{j^*}(\{\theta_{j^*}\})| = \max_j g(j)$. Due to i), for all $j \in \{1, \dots, m\}$ arbitrary, it then holds that:

ii) $|\pi_1^{j^*}(\{\theta_{j^*}\}) - \pi_2^{j^*}(\{\theta_{j^*}\})| \geq |\pi_1(\{\theta_j\}) - \pi_2(\{\theta_j\})|$ for all $\pi_1, \pi_2 \in \mathcal{M}_a$

This implies that $\| \pi_1^{j^*} - \pi_2^{j^*} \|_\infty \geq \| \pi_1 - \pi_2 \|_\infty$ for all $\pi_1, \pi_2 \in \mathcal{M}_a$, which then implies that $ext_{\mathcal{M}}(a_z) = g(j^*) = \max_j g(j)$. □

Remark 3. *Note that, instead of solving m linear programming problems for computing the value $ext_{\mathcal{M}}(a)$ as proposed in Proposition 3, one alternatively could solve one bilinear programming problem with objective function*

$$\sum_{j=1}^{m} \xi_j(\gamma_{1j} - \gamma_{2j}) + \sum_{j=1}^{m} \xi_{m+j}(\gamma_{2j} - \gamma_{1j}) \longrightarrow \max_{(\gamma_{11}, \dots, \gamma_{2m}, \xi_1, \dots \xi_{2m})} \tag{3}$$

with the same constraints as above and additional constraints $\xi_1, \dots \xi_{2m} \geq 0$ and $\sum_{j=1}^{2m} \xi_j = 1$. This approach has the advantage that the value $ext_{\mathcal{M}}(a)$ can also be computed with respect to $\|\cdot\|_1$ instead of $\|\cdot\|_\infty$. To see that, simply replace the constraint $\sum_{j=1}^{2m} \xi_j = 1$ by the set of constraints $\sum_{j=1}^{2m} \xi_j = m$ and $\xi_j + \xi_{m+j} = 1$ for all $j = 1, \dots, m$. However, note that the resulting bilinear programming problem then no longer is decomposable into m linear programming problems, since the solutions of the single problems can no longer be treated independently of each other as in the case of the $\|\cdot\|_\infty$-norm.

Despite its intuitiveness, the extent $ext_{\mathcal{M}}(\cdot)$ of an E-admissible act has a drawback in certain situations: It measure the size of set $\mathcal{M}_a$ only in one direction, namely the most extreme one. Therefore, the maximal extent alone might be not capable of distinguishing situations that definitely are worth to be distinguished in this context. This drawback is most easily explained by the schematic picture in Figure 1.

In order to react the problem that might arise when only considering the extent $ext_{\mathcal{M}}(a)$ of an E-admissible act $a \in \mathbb{A}_{\mathcal{M}}$ for measuring the size of the set $\mathcal{M}_a$, we now introduce another concept for addressing this question, and call it uniform extent. This measure relates to the diameter of the largest barycentric ε-cube that can be inscribed into $\mathcal{M}_a$.

Definition 3. *Let $(\mathbb{A}, \Theta, u)$ and $\mathcal{M}$ be defined as before and let $a \in \mathbb{A}_{\mathcal{M}}$. We define the uniform extent $uxt_{\mathcal{M}}(a)$ of act a with respect to $\mathcal{M}$ as the number*

$$uxt_{\mathcal{M}}(a) := \max\Big\{\varepsilon : \exists \pi \in \mathcal{M}_a \text{ s.t. } B_\varepsilon(\pi) \subset \mathcal{M}_a\Big\} \tag{4}$$

[11] Note that, due to standard results from linear programming theory, such an optimal solution always exists since the constraint set is bounded and there always exists an admissible solution since $a_z \in \mathbb{A}_{\mathcal{M}}$.

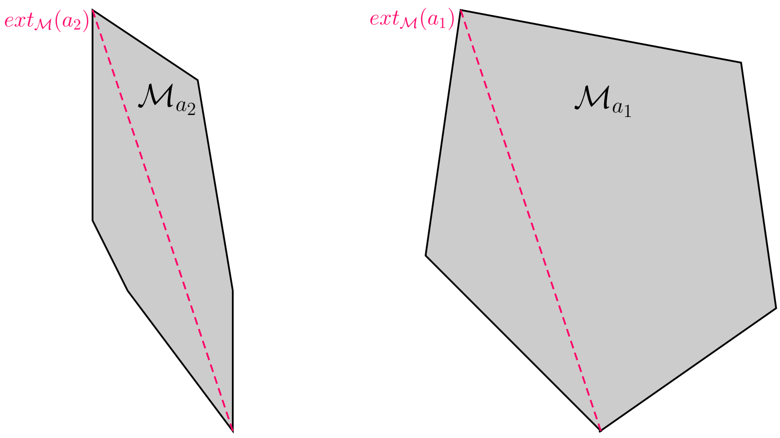

Figure 1: Two sets $\mathcal{M}_{a_1}$ and $\mathcal{M}_{a_2}$ with the same extent, however, quite different size.

where $B_\varepsilon(\pi) := \mathrm{conv}(\{\pi_\varepsilon^{1+}, \pi_\varepsilon^{1-}, \pi_\varepsilon^{2+}, \pi_\varepsilon^{2-}, \dots, \pi_\varepsilon^{m+}, \pi_\varepsilon^{m-}\})$ *with*

$$\pi_\varepsilon^{j^*+}(\{\theta_j\}) = \begin{cases} \pi(\{\theta_j\}) + \varepsilon & \text{if } j = j^* \\ \pi(\{\theta_j\}) - \frac{\varepsilon}{m-1} & \text{if } j \neq j^* \end{cases}$$

$$\pi_\varepsilon^{j^*-}(\{\theta_j\}) = \begin{cases} \pi(\{\theta_j\}) - \varepsilon & \text{if } j = j^* \\ \pi(\{\theta_j\}) + \frac{\varepsilon}{m-1} & \text{if } j \neq j^* \end{cases}$$

is the barycentric ε*-cube around* π*. Thus, the uniform extent of act* a *is half the diameter of the largest barycentric* ε*-cube that can be inscribed into* $\mathcal{M}_a$*.*

The uniform extent for the schematic picture discussed earlier is illustrated in Figure 2. We see that we now can distinguish between the two situations. However, as easy to imagine, also the uniform extent might sometimes be too pessimistic just as the maximal extent is too optimistic. Hence, a good approach is to consider both measures *ext* and *uxt* simultaneously. Together, they will give a pretty good impression of the extent of E-admissibility.

We now propose an algorithm for computing the uniform extent of some fixed E-admissible act with respect to $\mathcal{M}$ under consideration. Again, it shows that this can be done by solving one single, relatively simple, linear programming problem. Here, the main idea is to explicitly model the distributions $\pi_\varepsilon^{1+}, \pi_\varepsilon^{1-}, \dots, \pi_\varepsilon^{m+}, \pi_\varepsilon^{m-}$ from Definition 3 by decision variables and utilizing the fact that $\mathcal{M}_a$ is a convex set. The uniform extent is then computed by maximizing over the value of ε. Precisely, we arrive at the following proposition.

Proposition 4. *Let* $(\mathbb{A}, \Theta, u)$ *and* $\mathcal{M}$ *be defined as before and let* $a_z \in \mathbb{A}_\mathcal{M}$*. Consider the linear programming problem*

$$\varepsilon \longrightarrow \max_{(\gamma_1,\dots,\gamma_m,\varepsilon)} \tag{5}$$

with constraints $(\gamma_1, \dots, \gamma_m, \varepsilon) \geq 0$ *and*

- $\sum_{j=1}^m \gamma_j = 1$
- $\gamma_j \geq \varepsilon$ *for all* $j = 1, \dots, m$

- $\underline{b}_s \leqslant \sum_{j=1}^m f_s(\theta_j) \cdot \gamma_j + \varepsilon \cdot d(j^*, s) \leqslant \overline{b}_s$ *for all* $s = 1, ..., r, j^* = 1, \dots, m$
- $\underline{b}_s \leqslant \sum_{j=1}^m f_s(\theta_j) \cdot \gamma_j - \varepsilon \cdot d(j^*, s) \leqslant \overline{b}_s$ *for all* $s = 1, ..., r, j^* = 1, \dots, m$
- $\sum_{j=1}^m (u_{\ell j} - u_{zj}) \cdot \gamma_j + \varepsilon \cdot c(j^*, \ell) \leqslant 0$ *for* $\ell = 1, \dots, n, j^* = 1, \dots, m$
- $\sum_{j=1}^m (u_{\ell j} - u_{zj}) \cdot \gamma_j - \varepsilon \cdot c(j^*, \ell) \leqslant 0$ *for* $\ell = 1, \dots, n, j^* = 1, \dots, m$

where

$$c(j^*, \ell) = (u_{\ell j^*} - u_{zj^*}) - \frac{1}{(m-1)} \sum_{j \neq j^*} (u_{\ell j} - u_{zj})$$

and

$$d(j^*, s) = f_s(\theta_{j^*}) - \frac{1}{(m-1)} \sum_{j \neq j^*} f_s(\theta_j)$$

Then the uniform extent $uxt_{\mathcal{M}}(a_z)$ *of* a_z *is given by the optimal value of problem (5).*

Proof. First, note that every pair $(\pi, \varepsilon) \in \mathcal{M}_{a_z} \times \mathbb{R}_0^+$ with $B_\varepsilon(\pi) \subset \mathcal{M}_{a_z}$ induces an admissible solution to (5) with objective value ε by setting $\gamma_j := \pi(\{\theta_j\})$, since we have

$$\sum_{j=1}^m f_s(\theta_j) \cdot \gamma_j + \varepsilon \cdot d(j^*, s) = \mathbb{E}_{\pi_\varepsilon^{j^*+}}(f_s) \in (\underline{b}_s, \overline{b}_s)$$

$$\sum_{j=1}^m f_s(\theta_j) \cdot \gamma_j - \varepsilon \cdot d(j^*, s) = \mathbb{E}_{\pi_\varepsilon^{j^*-}}(f_s) \in (\underline{b}_s, \overline{b}_s)$$

for all $s = 1, ..., r, j^* = 1, \dots, m$ and it then holds that

$$\sum_{j=1}^m (u_{\ell j} - u_{zj}) \cdot \gamma_j + \varepsilon \cdot c(j^*, \ell) = \mathbb{E}_{\pi_\varepsilon^{j^*+}}(u_\ell - u_z) \leqslant 0$$

$$\sum_{j=1}^m (u_{\ell j} - u_{zj}) \cdot \gamma_j - \varepsilon \cdot c(j^*, \ell) = \mathbb{E}_{\pi_\varepsilon^{j^*-}}(u_\ell - u_z) \leqslant 0$$

for $\ell = 1, \dots, n, j^* = 1, \dots, m$ due to the constraints. Since $a_z \in \mathbb{A}_{\mathcal{M}}$ and, therefore, there exists $\pi_0 \in \mathcal{M}_{a_z}$, it is then guaranteed that problem (5) always possesses an admissible solution (just take the one induced by $(\pi_0, 0)$). Since the set of admissible solutions is obviously bounded, it also possesses an optimal solution by standard results on linear programming theory.

Let $(\gamma_1^*, \dots, \gamma_m^*, \varepsilon^*)$ denote such an optimal solution to (5). Utilizing again the above identities (in the opposite way), we see that setting $\pi^*(\{\theta_j\}) := \gamma_j^*$ defines a probability measure $\pi^* \in \mathcal{M}_{a_z}$ such that $B_{\varepsilon^*}(\pi^*) \subset \mathcal{M}_{a_z}$ with $\varepsilon^* = uxt_{\mathcal{M}}(a_z)$. $\square$

Remark 4. *The linear programming problem (5) possesses* $m+1$ *decision variables and* $1+2mr+2mn$ *constraints. It therefore might become computationally expensive for very large problems.*

We conclude the section by applying the proposed measures of the extent of E-admissibility to the toy example that was already introduced at the end of Section 3.1.

Example 2. *Consider again the situation of Example 1. We want to compute the extent* $ext_{\mathcal{M}}(\cdot)$ *of both* E*-admissible acts* a_1 *and* a_2*. Solving the series of linear programming problems from Proposition 3 for both acts gives* $ext_{\mathcal{M}}(a_1) = 0.3$ *and* $ext_{\mathcal{M}}(a_2) = 0.2$ *with respect to the* $\|\cdot\|_\infty$*-norm. Therefore, it could be argued that* a_1 *is the most preferable among the* E*-admissible acts with respect to* $\mathcal{M}$*. Additionally, we are interested in the uniform extent* $uxt_{\mathcal{M}}(\cdot)$ *of the acts* a_1 *and* a_2*. Solving the linear programming problem introduced in Proposition 4 gives the results* $uxt_{\mathcal{M}}(a_1) = 0.15$ *and* $uxt_{\mathcal{M}}(a_2) = 0.1$*, even strengthening the argument that act* a_1 *is the most preferable among the* E*-admissible acts.*

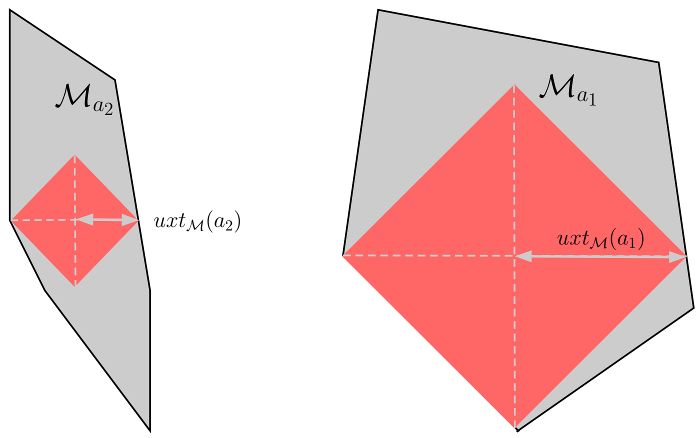

Figure 2: The measure $uxt_{\mathcal{M}}(\cdot)$ indeed gives different values to the sets $\mathcal{M}_{a_1}$ and $\mathcal{M}_{a_2}$ and therefore resolves the drawback of the measure $ext_{\mathcal{M}}(\cdot)$.

4. The Ordinal Case

Up to this point, all decision criteria discussed, with the exception of Wald's maximin principle, made explicit use of the cardinality of the utility function u involved in the basic decision problem $(\mathbb{A}, \Theta, u)$. However, as widely known, assuming cardinal utility implicitly demands the decision maker's preferences to satisfy pretty strong axiomatic assumptions which are often not met in practice. If the deviation from these axioms is too strong, it often makes sense to work with decision criteria that can cope with purely ordinal preferences.[12] For this reason, in this section the utility function u in the decision problem $(\mathbb{A}, \Theta, u)$ is solely interpreted as an ordinal utility representation. Particularly, utility differences with respect to u have no meaningful interpretation apart from their sign in what follows.

We again start by briefly summarizing some criteria that still make sense in the presence of purely ordinal preferences. If, additional to the ordinal utility information, a precise probability measure π on the state space is available, again several different criteria appear natural:

(C_1) *Pairwise Stochastic Dominance:* Label any act $a_0 \in \mathbb{A}$ optimal for which there does not exist another act $a_1 \in \mathbb{A} \setminus \{a_0\}$ such that $\mathbb{E}_\pi(t \circ u_{a_1}) \geq \mathbb{E}_\pi(t \circ u_{a_0})$ for every non-decreasing function $t : \mathbb{R} \to \mathbb{R}$. If, contrarily, it is the case that $\mathbb{E}_\pi(t \circ u_{a_1}) \geq \mathbb{E}_\pi(t \circ u_{a_0})$ for every non-decreasing function $t : \mathbb{R} \to \mathbb{R}$, we say that a_1 stochastically dominates a_0 (cf., e.g., Lehmann (1955); Mosler and Scarsini (1991)).

Clearly, pairwise stochastic dominance can rather be viewed as a local decision criterion, since the preference between two acts $a_0, a_1 \in \mathbb{A}$ does not depend on which other acts from $\mathbb{A} \setminus \{a_0, a_1\}$ are also available to the decision maker. Moreover, it also possesses a very natural interpretation: Act a_1 is preferred to act a_0 if every expectation maximizing decision maker with the same ordinal utility function would have the same preference between the two acts. Note that often acts will be incomparable with respect to stochastic dominance, since it will hold $\mathbb{E}_\pi(t_1 \circ u_{a_1}) > \mathbb{E}_\pi(t_1 \circ u_{a_0})$

[12]Another, very prominent, way for proceeding in such situations is working with partially cardinal preference relations as done in Seidenfeld et al. (1995).

for one function t_1 and $\mathbb{E}_\pi(t_2 \circ u_{a_1}) < \mathbb{E}_\pi(t_2 \circ u_{a_0})$ for another function t_2.

(C_2) *Joint Stochastic Dominance:* Label every act $a_0 \in \mathbb{A}$ optimal for which there exists a strictly increasing function $t^* : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}_\pi(t^* \circ u_{a_0}) \geq \mathbb{E}_\pi(t^* \circ u_a)$ for all $a \in \mathbb{A}$, i.e. if there exists one expectation maximizing agent with the same ordinal utility function for which a_0 maximizes expected utility among all other available acts (cf. Jansen et al. (2017b)).

Obviously, this is an example for a global criterion: If there exists a function t^* with the desired properties for all $a \in \mathbb{A}$, this does not necessarily imply the existence of such a function for $\mathbb{A}^* := \mathbb{A} \cup \{a^*\}$ (simply choose a^* to have higher utility that every act in $\mathbb{A}$ in every state of nature).

(C_3) *Pairwise Statistical Preference:* Label every act $a_0 \in \mathbb{A}$ optimal for which there exists no other act $a_1 \in \mathbb{A} \setminus \{a_0\}$ such that

$$\pi(\{\theta : u_{a_1}(\theta) \geq u_{a_0}(\theta)\}) > \pi(\{\theta : u_{a_0}(\theta) \geq u_{a_1}(\theta)\})$$

i.e. if there is no other act a_1 which has higher probability of yielding a higher utility value than a_0. If contrarily there exists such an act a_1, we say that a_1 statistically dominates a_0 (cf., e.g., Montes (2014, Section 2.2.1)).

Clearly, statistical preference can rather be viewed as a local decision criterion, since the preference between two acts a_0 and a_1 does not depend on acts from $\mathbb{A} \setminus \{a_0, a_1\}$.

(C_4) *Joint Statistical Preference:* Label every act $a_0 \in \mathbb{A}$ optimal for which it holds that $D_\pi(a_0) \geq D_\pi(a)$ for all $a \in \mathbb{A}$, where

$$D_\pi(a) := \pi(\{\theta : u(a, \theta) \geq u(a', \theta) \text{ for all } a' \in \mathbb{A}\})$$

that is if a_0 has the highest probability to be utility dominant among all other available acts.

This criterion is clearly global: Enlarging the set of acts $\mathbb{A}$ to a new set of acts $\mathbb{A}^* := \mathbb{A} \cup \{a^*\}$ might completely change the preference between two acts acts $a_0, a_1 \in \mathbb{A}$ in the sense that $D_\pi^{\mathbb{A}}(a_0) > D_\pi^{\mathbb{A}}(a_1)$ but $D_\pi^{\mathbb{A}^*}(a_0) < D_\pi^{\mathbb{A}^*}(a_1)$.[13]

If no precise probability measure π is available and the uncertainty on the state space is again characterized by a credal set $\mathcal{M}$ of the form defined in (1), then there are several possibilities to generalize the decision criteria (C_1), (C_2). (C_3) and (C_4). A detailed study of these different possibilities as well as algorithmic approaches that are capable to deal with the resulting criteria is given in Montes (2014, Sections 4.1.1 and 4.1.2) and Jansen et al. (2017b). An algorithm for detecting stochastic dominance for the case that the different decision consequences are only partially ordered is introduced in Schollmeyer et al. (2017). Here, we only give a small selection of

[13]For a simple example consider the decision problems $(\mathbb{A}, \Theta, u)$ and $(\mathbb{A} \cup \{a^*\}, \Theta, \tilde{u})$ given by

	θ_3	θ_2	θ_3
a_1	2	2	5
a_2	3	3	3

	θ_3	θ_2	θ_3
a_1	2	2	5
a_2	3	3	3
a^*	1	2	6

and the prior π on Θ induced by $(\pi(\{\theta_1\}), \pi(\{\theta_2\}), \pi(\{\theta_3\})) = (0.2, 0.2, 0.6)$. Here we have that $D_\pi^{\mathbb{A}}(a_1) = 0.6 > 0.4 = D_\pi^{\mathbb{A}}(a_2)$ but $D_\pi^{\mathbb{A}^*}(a_2) = 0.4 > 0 = D_\pi^{\mathbb{A}^*}(a_1)$.

the criteria:

(D_1) *Joint Statistical Preference (Imprecise Version)*: Label any act $a_0 \in \mathbb{A}$ optimal for which it holds that $\min_{\pi \in \mathcal{M}} D_\pi(a_0) \geq \min_{\pi \in \mathcal{M}} D_\pi(a)$ for all other $a \in \mathbb{A}$, i.e. which maximizes the lower probability of the act to be dominant to all other available acts.

(D_2) *Joint Stochastic Dominance (Imprecise Version):* Label act $a_0 \in \mathbb{A}$ optimal if there exists a strictly increasing function $t^* : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}_\pi(t^* \circ u_{a_0}) \geq \mathbb{E}_\pi(t^* \circ u_a)$ for all $a \in \mathbb{A}$ and all $\pi \in \mathcal{M}$.

All the ordinal decision criteria just discussed can be handled either by hand or by utilizing linear programming techniques similar as seen in detail for the criteria discussed in Section 3 (see Jansen et al. (2017b) for details). Here, we only give an impression of how this could be done for the example of the imprecise version of joint stochastic dominance: To check whether an act $a_z \in \mathbb{A}$ is optimal in the sense of joint stochastic dominance in the imprecise version, we explicitly model the transformation function t^* by decision variables. Additionally, we require the extreme points $\pi^{(1)}, \dots, \pi^{(T)}$ of the underlying credal set $\mathcal{M}$. We then consider the linear programming problem with the objective function

$$\varepsilon \longrightarrow \max_{(\varepsilon, t_{11}, \dots, t_{nm})} \tag{6}$$

and constraints $(\varepsilon, t_{11}, \dots, t_{nm}) \geq 0$ and

- $t_{11}, \dots, t_{nm} \leq 1$
- $\sum_{j=1}^{m} (u_{zj} t_{zj} - u_{ij} t_{ij}) \cdot \pi^{(t)}(\{\theta_j\}) \geq 0$ for all $t = 1, \dots, T,\ i = 1, \dots, n$
- For $i, i^{'} \in \{1, \dots, n\}$, $j, j^{'} \in \{1, \dots, m\}$: $u_{ij} = u_{i^{'} j^{'}} \Rightarrow t_{ij} = t_{i^{'} j^{'}}$
- For $i, i^{'} \in \{1, \dots, n\}$, $j, j^{'} \in \{1, \dots, m\}$: $u_{ij} < u_{i^{'} j^{'}} \Rightarrow t_{ij} + \varepsilon \leq t_{i^{'} j^{'}}$

One then can show that act a_z is optimal in the sense of joint stochastic dominance in the imprecise version if and only if the optimal objective of the above program is strictly greater than 0. The idea here is that if there exists an optimal solution $(\varepsilon^*, t_{11}^*, \dots, t_{nm}^*)$ with $\varepsilon^* > 0$, then the solution t_{ij}^* describes the necessary strictly increasing transformation of u, and we receive a desired function by choosing any increasing function $t^* : \mathbb{R} \to \mathbb{R}$ satisfying that $t^*(u_{ij}) = t_{ij}^* \cdot u_{ij}$ for all $i = 1, \dots n$ and $j = 1, \dots, m$.

Of course solving this linear program might become computationally very expensive and cumbersome as the number of extreme points of the set $\mathcal{M}$ might become as large as $m!$ (cf., Derks and Kuipers (2002); Wallner (2007)). However, convenient classes of credal sets 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-mutual models (cf., Montes et al. (2017))).

Another easy to handle case appears if the credal set $\mathcal{M}$ under consideration directly arises as the convex hull of a finite number of precise probability estimates $\pi_{E_1}, \dots, \pi_{E_K}$ of a committee of experts $E_1, \dots, E_K$. In such cases the extreme points of the credal set $\mathcal{M}$ are always among the experts guesses $\pi_{E_1}, \ldots, \pi_{E_K}$ for the probabilities, and the algorithm described above can directly be applied without the need for any previous computation of the extreme points. We conclude the section by a small example that continues Examples 1 and 2.

Example 3. *Consider again the situation of Examples 1 and 2. Here, the unique optimal act with respect to joint statistical preference in the imprecise version is a_1 with a value of* 0.3*. If we consider joint stochastic dominance in the imprecise version, we first need to compute the extreme points of $\mathcal{M}$, which are here given by the measures $\pi^{(1)}, \pi^{(2)}$ induced by $\pi^{(1)}(\{\theta_1\}) = 0.3$ and $\pi^{(2)}(\{\theta_1\}) = 0.8$. Solving the above linear programming problem (6) for all acts gives that acts a_3 and a_4 are optimal in terms of joint stochastic dominance in the imprecise version whereas acts a_1 and a_2 are not.*

5. A Stylized Application Example

In this section, we discuss a more realistic, yet stylized, application example in some more detail: Consider the situation where the decision maker wants to invest money in stocks of some company. The acts then correspond to the stocks of the different companies. Say the agent compares ten different stocks collected in $\mathbb{A} = \{a_1 \ldots, a_{10}\}$. Moreover, the states of nature then correspond to different economic scenarios which might or might not occur and which, each differently, would influence the payoffs of the stocks of the different companies. Say the agent incorporates the scenarios collected in $\Theta = \{\theta_1, \ldots, \theta_5\}$ in her market analysis. She summarizes the payoffs of the different stocks under the different scenarios in the following utility table:

$u(a_i, \theta_j)$	$\boldsymbol{\theta_1}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\theta_3}$	$\boldsymbol{\theta_4}$	$\boldsymbol{\theta_5}$
$\mathbf{a_1}$	37	25	23	73	91
$\mathbf{a_2}$	50	67	2	44	94
$\mathbf{a_3}$	60	4	96	1	83
$\mathbf{a_4}$	16	24	31	26	100
$\mathbf{a_5}$	3	86	76	85	11
$\mathbf{a_6}$	12	49	66	56	14
$\mathbf{a_7}$	39	10	92	88	57
$\mathbf{a_8}$	62	52	80	71	42
$\mathbf{a_9}$	90	8	74	70	38
$\mathbf{a_{10}}$	63	68	36	69	9

Moreover, suppose the decision maker has observed the market development for quite a while, so that she can specify bounds for the probabilities of the different economic scenarios to occur (alternatively, the bounds for the scenario probabilities could also come from opinions of different expert the agent has consulted). Precisely, she specifies the uncertainty underlying the situation by the credal set

$$\mathcal{M} = \Big\{\pi : \underline{b}_s \leq \mathbb{E}_\pi[f_s] \leq \overline{b}_s \text{ for } s = 1, \ldots, 5\Big\}$$

where

- $f_s : \Theta \to \mathbb{R}$ is given by $f_s(\theta) := \mathbb{1}_{\{\theta_s\}}(\theta)$ for $s = 1, \ldots, 5$ and

- $\begin{pmatrix} \underline{b}_1 & \overline{b}_1 \\ \underline{b}_2 & \overline{b}_2 \\ \underline{b}_3 & \overline{b}_3 \\ \underline{b}_4 & \overline{b}_4 \\ \underline{b}_5 & \overline{b}_5 \end{pmatrix} = \begin{pmatrix} 0.1 & 0.3 \\ 0.05 & 0.1 \\ 0.1 & 0.2 \\ 0.2 & 0.4 \\ 0.15 & 0.2 \end{pmatrix}$

Applying the different decision criteria and the other concepts discussed in the paper, the decision maker arrives at the following results:

- Stock a_8 is the unique non-randomized $\mathcal{M}$-maximin act, i.e. $\mathbb{A}_{\underline{\mathcal{M}}} = \{a_8\}$. Thus, for a very pessimistic and ambiguity averse agent, act a_8 is the appropriate investment.

- Solving the programming problem from Proposition 2 for ε set to 0 for each act, we find the set of E-admissible acts with respect to $\mathcal{M}$ is given by $\mathbb{A}_{\mathcal{M}} = \{a_7, a_9\}$ (since the optimal value of the program is 10 for both acts). Hence, the $\mathcal{M}$-maximin act is not E-admissible with respect to $\mathcal{M}$. In order to further compare the E-admissible acts a_7 and a_9, we first compute the extent $ext_{\mathcal{M}}(\cdot)$ from Definition 2 for both of them. Solving the series of linear programming problems described in Proposition 3 gives the results $ext_{\mathcal{M}}(a_7) \approx 0.152$ and $ext_{\mathcal{M}}(a_9) = 0.2$, for which reason it could be argued that a_9 is the most preferable among the E-admissible acts. To see how informative the extent of the acts is, we are additionally interested in their uniform extents $uxt_{\mathcal{M}}(\cdot)$ in the sense of Definition 3. Solving the linear programming problem introduced in Proposition 4 gives $uxt_{\mathcal{M}}(a_7) = 0.025$ as well as $uxt_{\mathcal{M}}(a_9) = 0.025$. Thus, if we consider the uniform extent in order to measure the amount of E-admissibility of acts, it could be argued that the decision maker should be indifferent between the E-admissible acts a_7 and a_9.

- Solving the programming problem from Proposition 2 for ε set to 100 for each act, we find the set of $\mathcal{M}$-maximal acts is given by $\mathbb{A}_{\max} = \{a_7, a_8, a_9\}$. In order to make a decision between the $\mathcal{M}$-maximin act a_8 and the E-admissible acts a_7 and a_9, it is of interest how far a_8 is from being E-admissible. Solving the linear program from Proposition 2 for varying value of ε gives that a_8 is E_ε-admissible in the sense of Definition 1 already for a value of $\varepsilon = 0.01$. Hence, a_8 is very close to being E-admissible and, therefore, could be argued to be preferable to a_7 and a_9.

- The unique optimal act with respect to joint statistical preference in the imprecise version is a_7 with a value of 0.2. In order to see which of the acts are optimal in the sense of joint stochastic dominance in the imprecise version as discussed in the previous section, we first need to compute the extreme points of $\mathcal{M}$. There are 15 such extreme points.[14] They are given in the Table 2:

  Having obtained the extreme points, we can use algorithm (6) from Section 4 for every act in $\mathbb{A} = \{a_1, \dots a_{10}\}$. We find that the acts a_6 and a_{10} are not optimal in the sense of joint stochastic dominance in the imprecise version, whereas the acts in $\mathbb{A} \setminus \{a_6, a_{10}\}$ are.

[14]The calculation was performed with the `rcdd package` (see Geyer and Meeden (2017)), which provides an interface for using Fukuda (2017)'s `cdd library` in the R statistical computing environment (see R Core Team (2017)).

	$\{\theta_1\}$	$\{\theta_2\}$	$\{\theta_3\}$	$\{\theta_4\}$	$\{\theta_5\}$
$\pi^{(1)}(\cdot)$	0.30	0.10	0.20	0.20	0.20
$\pi^{(2)}(\cdot)$	0.30	0.05	0.20	0.25	0.20
$\pi^{(3)}(\cdot)$	0.30	0.10	0.20	0.25	0.15
$\pi^{(4)}(\cdot)$	0.30	0.05	0.20	0.30	0.15
$\pi^{(5)}(\cdot)$	0.30	0.10	0.10	0.30	0.20
$\pi^{(6)}(\cdot)$	0.30	0.05	0.10	0.35	0.20
$\pi^{(7)}(\cdot)$	0.30	0.10	0.10	0.35	0.15
$\pi^{(8)}(\cdot)$	0.15	0.05	0.20	0.40	0.20
$\pi^{(9)}(\cdot)$	0.15	0.10	0.20	0.40	0.15
$\pi^{(10)}(\cdot)$	0.20	0.05	0.20	0.40	0.15
$\pi^{(11)}(\cdot)$	0.20	0.10	0.10	0.40	0.20
$\pi^{(12)}(\cdot)$	0.25	0.05	0.10	0.40	0.20
$\pi^{(13)}(\cdot)$	0.25	0.10	0.10	0.40	0.15
$\pi^{(14)}(\cdot)$	0.10	0.10	0.20	0.40	0.20
$\pi^{(15)}(\cdot)$	0.30	0.05	0.10	0.40	0.15

Table 2: Extreme points in the application example.

6. Summary and Concluding Remarks

In this paper we introduced and discussed some ideas in the context of decision theory using imprecise probabilistic model. Here, we first introduced a new decision criterion, E_ε-admissibility, that selects acts that are not too far from E-admissibility, where the accepted deviation from E-admissibility can be explicitly controlled by an additional parameter ε. Subsequently, we investigated how to measure the extent of E-admissibility of an E-admissible act of interest. Precisely, we introduced two different measures for this purpose: the maximal extent $ext_{\mathcal{M}}(a)$ and the uniform extent $uxt_{\mathcal{M}}(a)$ of an E-admissible act a. While the former corresponds to the maximal diameter of the set $\mathcal{M}_a$, the latter is related to the side length of the maximal barycentric ε-cube that can be inscribed into $\mathcal{M}_a$. For all concepts discussed we proposed (bi-)linear programming driven algorithms for computation. In the second part of the paper we recalled some concepts for decision making if a cardinal utility function is no longer available and there is (potentially) only imprecise probabilistic information. For the concept of imprecise joint stochastic dominance, we also discussed some details about computation.

There are several interesting directions that could be followed in future research of which we only want to briefly mention one: Consider again the viewpoint that the credal set $\mathcal{M}$ arises from the opinions of a committee of experts. In the discussion directly following Definition 1, we argued in favor of the concept of E_ε-admissibility, since it allows to take into account more than only one expert opinion while simultaneously allowing to control how far the involved experts may differ in opinion. This idea could easily be extended: Instead of only controlling how far the involved experts may differ from each other in terms of opinion, one could also control how far their opinions differ from some externally given criterion. If we take again our example of some politician with an advisory body of experts, the external criterion could for instance be the opinion of the politician herself, so that she only takes expert opinions into account that do not differ too much from her own one. Of course other examples for external criteria are imaginable.

Acknowledgments

We want to express our gratitude to Jean Baccelli, Seamus Bradley, Arthur Merin, Jürgen Landes, Gregory Wheeler and many other members of the Munich Center for Mathematical Philosophy for stimulating talks and discussion on decision theory and imprecise probabilities during the last few years, and to Teddy Seidenfeld for inspiring lectures and discussions not only during his visits to Munich.

References

Augustin, T. (2003). On the suboptimality of the generalized Bayes rule and robust Bayesian procedures from the decision theoretic point of view: A cautionary note on updating imprecise priors. In Bernard, J.-M., Seidenfeld, T., and Zaffalon, M., editors, *Proceedings of the Third International Symposium on Imprecise Probability: Theories and Applications (ISIPTA '03, Lugano)*, pages 31–45. Carleton Scientific.

Augustin, T., Coolen, F., de Cooman, G., and Troffaes, M., editors (2014). *Introduction to Imprecise Probabilities*. Wiley, Chichester.

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 (ISIPTA '15, Pescara)*, pages 57–66. Aracne.

de Campos, L., Huete, J., and Moral, S. (1994). Probability interval: A tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2:167–196.

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. *Journal of Economic Surveys*, 26:234–270.

Fukuda, K. (2017). The cdd and cddplus Homepage. `https://www.inf.ethz.ch/personal/fukudak/cdd_home/` Last accessed Dec 22, 2017.

Geyer, C. J. and Meeden, G. D. (2017). *Package rcdd*. `https://CRAN.R-project.org/package=rcdd` and `www.stat.umn.edu/geyer/rcdd/`, Last accessed Dec 22, 2017.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153.

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.

Herron, T., Seidenfeld, T., and Wasserman, L. (1997). Divisive conditioning: Further results on dilation. *Philosophy of Science*, 64:411–444.

Hodges, J. and Lehmann, E. (1952). The use of previous experience in reaching statistical decisions. *Annals of Mathematical Statistics*, 23:396–407.

Huntley, N., Hable, R., and Troffaes, M. (2014). Decision making. In Augustin, T., Coolen, F., de Cooman, G., and Troffaes, M., editors, *Introduction to Imprecise Probabilities*, pages 190–206. Wiley, Chichester.

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.

Kikuti, D., Cozman, F., and Filho, R. (2011). Sequential decision making with partially ordered preferences. *Artificial Intelligence*, 175:1346–1365.

Kofler, E. (1989). *Prognosen und Stabilität bei unvollständiger Information [Forecasts and Stability under Incomplete Information]*. Campus, Frankfurt.

Kofler, E. and Menges, G. (1976). *Entscheiden bei unvollständiger Information [Decision Making under Incomplete Information]*. Springer, Berlin.

Lehmann, E. (1955). Ordered families of distributions. *Annals of Mathematical Statistics*, 26:399–419.

Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy*, 71:391–418.

Liu, H. (2015). Common knowledge, ambiguity, and the value of information in games. In Augustin, T., Doria, S., Miranda, E., and Quaeghebeur, E., editors, *Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA '15, Pescara)*, pages 167–176. Aracne.

Majo-Wilson, C. and Wheeler, G. (2016). Scoring imprecise credences: A mildly immodest proposal. *Philosophy and Phenomenological Research*, 93:55–78.

Miranda, E. and Destercke, S. (2015). Extreme points of the credal sets generated by comparative probabilities. *Journal of Mathematical Psychology*, 64:44–57.

Montes, I. (2014). *Comparison of Alternatives under Uncertainty and Imprecision.* PhD thesis, Oviedo. `http://digibuo.uniovi.es/dspace/bitstream/10651/28953/6/TD_IgnacioMontesGutierrez.pdf`. Last accessed December 22, 2017.

Montes, I. and Destercke, S. (2017). On extreme points of p-Boxes and belief functions. In Ferraro, M., Giordani, P., Vantaggi, B., Gagolewski, M., Gil, M., Grzegorzewski, P., and Hryniewicz, O., editors, *Soft Methods for Data Science*, pages 363–371. Springer.

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 (ISIPTA '17, Lugano)*, volume 62 of *Proceedings of Machine Learning Research*, pages 229–240. PMLR.

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.

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

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 (ISIPTA '03, Lugano)*, pages 496–510. Carleton Scientific.

Schervish, M. J., Seidenfeld, T., and Kadane, J. B. (2013). The effect of exchange rates on statistical decisions. *Philosophy of Science*, 80:504–532.

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 (ISIPTA '15, Pescara)*, 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/`. Last accessed December 22, 2017.

Seidenfeld, T. (1988). Decision theory without independence or without ordering, What is the difference? (with discussion). *Economics and Philosophy*, 4:267–315.

Seidenfeld, T. (1994). When normal form and extensive form solutions differ. In Prawitz, D., Skyrms, B., and Westerstahl, D., editors, *Logic, Methodology and Philosophy of Science IX (Uppsala, 1991)*, pages 451–463. Elsevier.

Seidenfeld, T. (2004a). A contrast between two decision rules for use with (convex) sets of probabilities: Γ-maximin versus E-admissibility. *Synthese*, 140:69–88.

Seidenfeld, T. (2004b). Some decision theoretic issues for imprecise probability theory. Slides of a lecture given on the *First SIPTA School on Imprecise Probabilities.* Available under: `http://www.sipta.org/ssipta04/topics.htm`. Last accessed December 22, 2017.

Seidenfeld, T., Kadane, J., and Schervish, M. (1995). A representation of partially ordered preferences. *Annals of Statistics*, 23:2168–2217.

Seidenfeld, T., Schervish, M. J., and Kadane, J. B. (2010). Coherent choice functions under uncertainty. *Synthese*, 172:157–176.

Seidenfeld, T., Schervish, M. J., and Kadane, J. B. (2012). Forecasting with imprecise probabilities. *International Journal of Approximate Reasoning*, 53:1248–1261.

Seidenfeld, T. and Wassermann, L. (1993). Dilation for sets of probabilities. *Annals of Statistics*, 21:1139–1154.

Seidenfeld, T. and Wassermann, L. (1996). Discussion of "Inferences from multinomial data: Learning about a bag of marbles" by P. Walley. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58:49.

Troffaes, M. (2007). Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45:17–29.

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 Forth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA '05, Pittsburgh)*, pages 349–358. SIPTA.

Utkin, L. V. and Augustin, T. (2007). Decision making under imperfect measurement using the imprecise Dirichlet model. *International Journal of Approximate Reasoning*, 44:322–338.

Vicig, P. and Seidenfeld, T. (2012). Bruno de Finetti and imprecision: Imprecise probability does not exist! *International Journal of Approximate Reasoning*, 53:1115–1123.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

Walley, P. (1996). Inferences from multinomial data: Learning about a bag of marbles (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)*, 58:3–57.

Wallner, A. (2007). Extreme points of coherent probabilities on finite spaces. *International Journal of Approximate Reasoning*, 44:339–357.

Wassermann, L. and Seidenfeld, T. (1994). The dilation phenomenon in robust Bayesian inference. (with discussion). *Journal of Statistical Planning and Inference*, 40:345–356.

Weichselberger, K. (2001). *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I: Intervallwahrscheinlichkeit als umfassendes Konzept [Elementary Foundations of a More General Calculus of Probability I: Interval Probability as a Comprehensive Concept]*. Physica, Heidelberg.

Weichselberger, K. and Pöhlmann, S. (1990). *A Methodology for Uncertainty in Knowledge-based Systems*. Springer, Heidelberg.

Wheeler, G. (2012). Objective Bayesianism and the problem of non-convex evidence. *British Journal for the Philosophy of Science*, 63:841–850.