ECAI 2020
G.D. Giacomo et al. (Eds.)

doi:10.3233/FAIA200402

Basic Utility Theory for Belief Functions

Chunlai Zhou[1] and **Biao Qin**[2] and **Deying Li**[3] and **Zhiqiang Xu**[4] and **Xiaoyong Du**[5]

Abstract. In this paper we find common ground for different decision theories in Dempster-Shafer theory by providing an axiomatic utility theory where the completeness requirement is dropped. The resulting preference relation is represented by subjective expectation of *sets* of utilities whose ordering is based on an ordering of outcome sets derived from a *logical* decision theory for complete ignorance. Moreover, we explore the preference aggregation problem within the utility theory and generalize some results by Harsanyi and Mongin to the setting of belief functions.

1 INTRODUCTION

In Bayesian theory, all information is probabilizable. In many situations, however, it is not clear what the probabilities should be. Information may be vague or ambiguous. For that reason, alternative representations are used that do not probabilize all information. Among those alternatives, the belief functions, as introduced in [9] and [30], have proved useful in artificial intelligence. They are suited to represent decision situations where only part of the information can be probabilized and the rest of the information is vague or ambiguous. And the nonprobabilizable information is subject to the principles of complete ignorance. There are two typical situations for decision making with belief functions in AI. The first situation is one in which the decision maker's information concerning the possible states of nature is best described by a belief function [10, 32]. The second situation is that the consequences of each act under each state of nature may not be precisely described [14]. This situation may arise when the decision problem is underspecified: for instance, the set of acts or the state space may be too coarsely defined.

In this paper we construct an axiomatic utility theory for belief functions to accommodate the above two decision situations in a uniform framework. This theory is *basic* because of the following two reasons. First, the partial preference relation in the theory is a common ground for different complete preference relations in other important utility theories for belief functions (Proposition 12 and Eq. (7)). Second, the basic utility theory faithfully observes the most characteristic principle for belief functions:

- (*Least commitment principle*): One should never assign evaluations more than guaranteed by evidence.

Our framework is the well-known Demspter semantics. In a Dempster model, the decision maker (DM for short) or agent uses a two-stage approach to process the information. The first stage deals with the probabilistic information on the knowledge space Ω. A deviation from the Bayesian approach occurs only in the second stage, where the vague information on a different nature space S is processed. We assume that the vague information does not contain any meaningful structure. The Bayesian approach will nevertheless assign probability to describe this information, usually according to the maximum entropy principle or Laplace's principle of insufficient reason. The intent of our approach, on the contrary, is to *preserve full objectivity*; entirely vague information should therefore be processed according to the objective principles of complete ignorance, as laid down in [1] by a multivalued mapping Γ.

Our first contribution is to offer a *logical* theory of complete ignorance which deals with partial acts whose domains are not necessarily the whole nature space but its subsets. The logical approach is to preserve the objectivity of complete ignorance. Complete ignorance is an old topic. There are two main differences from those theories about complete ignorance in the literature. First we drop the completeness assumption from Arrow and Hurwicz's framework. Second and more importantly, we establish the theory in a purely deductive way where preference statements such as $f \succeq g$ are basic propositions. A preference statement is valid if there is a *proof* in our logic. Primitive statements are those satisfying the symmetry and duplication properties characterizing the complete ignorance as well as two rational properties: reflexivity and weak dominance. The transitivity is used as an inference rule instead. With this logic, we show that a preference statement $f \succeq g$ is valid iff the least consequence of f is no less than that of g and the greatest of f is no less than that of g (Proposition 8). Since the ordering of these partial acts depends only on their ranges, we can naturally extend the linear ordering of consequences to a partial ordering of *sets* of consequences (Eq. (1)).

Our second contribution is to establish an Anscombe-Aumann-style utility theory for belief functions. We don't deal directly with complete acts, which are functions from the nature space S to the set $\Delta(\mathcal{C})$ of probability measures over the consequence set $\mathcal{C}$. Instead, we formulate 6 axioms about ordering of *correspondences*, which are compositions of complete acts with multivalued mappings, and prove a representation as subjective expectation of *sets* of utilities (Proposition 11). We also drop the completeness requirement but keep it for the *crisp* correspondences, which associate a single measure to each state, to accommodate the Anscombe-Aumann model for Bayesian theory. The first five axioms are similar to those in the Anscombe-Aumman theory for Bayesian theory. The last one is new and it connects ranking of crisp correspondences (functions) to that of correspondences. Our *new* definition of expectations in terms of belief functions is the same as the corresponding Aumann integrals. We also demonstrate the relationship of the partial preference relation represented by this new expectation to those represented by other different expectations.

Our third contribution is to generalize both Harsanyi's social ag-

[1] Correspondence Author: Renmin University, China and Harvard University, USA, email: czhou@ruc.edu.cn
[2] Correspondence Author: Renmin University, Beijing, China,email: qinbiao@ruc.edu.cn
[3] Renmin University, China,email: deyingli@ruc.edu.cn
[4] BAIDU, China,email: xuzhiqiang04@baidu.com
[5] Renmin University, China,email: duyong@ruc.edu.cn

gregation theorem and Mongin's (im)possibility theorems for social decision making in Bayesian theory to the setting of belief functions. We obtain a social aggregation theorem which is the counterpart of Harsanyi's aggregation theorem for Bayesian theory [20]. Moreover, we prove that, when all the prior beliefs are the same, the Paretian conditions imply that the social observer's utility is a linear combination of the individual's utilities with different coefficient requirements (Corollary 17), and when these belief functions are linearly independent, Paretian conditions imply that there is a utility dictator or inverse utility dictator (Corollary 18).

Our paper proceeds as follows. In Section 2, we set up the decision framework within Dempster models of belief functions by adapting the semantics from Jaffray and Wakker in [21]. In Section 3, we first develop a logical system to derive a decision rule under complete ignorance. Then we propose an Anscombe-Aumann-style axiomatization and prove a representation theorem where the resulting preference relation is represented by subjective expectation of *sets* of utilities whose ordering is based on the *logical* decision theory for complete ignorance. In Section 4, we explore the preference aggregation problem within the utility theory and generalize some results by Harsanyi and Mongin to the setting of belief functions and conclude in Section 5 with related works.

2 DEMPSTER MODEL OF BELIEF FUNCTIONS

In this paper, we consider the Dempster semantics [9]. A *Dempster model* for belief functions is a tuple $M = \langle \Omega, \Gamma, S, Pr\rangle$ where

1. Ω is a finite set of elements called *knowledge states*;
2. S is a finite set of *states of nature*;
3. Γ is a function from Ω to 2^S (called *multi-valued mapping*);
4. Pr is a probability distribution over the space Ω.

We also call Ω the *knowledge space* and S the *nature space*. In a Dempster model, the probabilistic information is processed in the first stage, and the vague information in the second. The information processed in the first stage is modeled through the knowledge space Ω. Exactly one knowledge state is the true one, the others are not true. The decision maker has partial information about which is the true knowledge state. We assume that the information on the set Ω is sufficiently well-structured to be modeled by a probability measure Pr. Concerning the remaining uncertainty, given $\omega \in \Omega$, the decision maker has no information at all, and this uncertainty is not probabilizable. And the nonprobabilizable information is represented by the multi-valued mapping Γ.

The description of the information concerning the state space S combines the probabilizable information concerning Ω and the information given each $\omega \in \Omega$. The information on S is completely described by the mapping: $A \mapsto Pr(\Omega_p^A)$ where $\Omega_p^A = \{\omega \in \Omega : \Gamma(\omega) \subseteq A\}$, which is a *belief function* and denoted by bel. In general, a *belief function* is a function $bel : 2^S \to [0,1]$ satisfying the following conditions: $bel(\emptyset) = 0$, $bel(S) = 1$, and $bel(\bigcup_{i=1}^n A_i) \geq \sum_{\emptyset \neq I \subseteq \{1,\cdots,n\}} (-1)^{|I|+1} bel(\cap_{i\in I} A_i)$ where $A_i \in 2^S$ for all $i \in \{1,\cdots,n\}$. We require that $\Gamma(\omega) \neq \emptyset$ for every ω in order to avoid re-normalization. Note that $bel(A) = Pr(\Omega_p^A) = Pr(\{\omega \in \Omega : \Gamma(\omega) \subseteq A\})$. If we define the function $m : 2^S \to [0,1]$ by $B \mapsto Pr(\{\omega \in \Omega : \Gamma(\omega) = B\})$, then $bel(A) = \sum_{B\subseteq A} m(B)$. Such defined m is a mass function. In general, a *mass function* on S is a mapping $m : 2^S \to [0,1]$ satisfying $\sum_{A\in 2^S} m(A) = 1$. A mass function m is called *normal* if $m(\emptyset) = 0$. Without further notice, all mass functions in this paper are assumed to be normal. A set A is called *focal* if $m(A) > 0$. A mapping $f : 2^S \to [0,1]$ is a belief function if and only if its Möbius transform is a mass assignment [30]. In other words, if $m : 2^S \to [0,1]$ is a mass assignment, then it determines a belief function $bel : 2^S \to [0,1]$ as follows: $bel(A) = \sum_{B\subseteq A} m(B)$ for all $A \in 2^\Omega$. Moreover, given a belief function bel, we can obtain its corresponding mass function m as follows: $m(A) = \sum_{B\subseteq A} (-1)^{|A\setminus B|} bel(B)$ for all $A \in 2^S$. The belief function bel is called *Bayesian* if $m(A) = 0$ for all non-singletons A. The corresponding *plausibility function* $pl : 2^S \to [0,1]$ is dual to bel in the sense that $pl(A) = 1 - bel(\overline{A})$ for all $A \subseteq S$. If m_1 and m_2 are two mass functions on S induced by two *independent* evidential sources, the combined mass function is calculated according to *Dempster's rule of combination*: for any $\mathcal{C} \subseteq S$,

$$(m_1 \oplus m_2)(C) = \frac{\sum_{A\cap B=C} m_1(A)m_2(B)}{\sum_{A\cap B\neq\emptyset} m_1(A)m_2(B)}$$

It is undefined if $\sum_{A\cap B\neq\emptyset} m_1(A)m_2(B) = 0$.

A *decision situation* in the Dempster model $M = \langle \Omega, \Gamma, S, Pr\rangle$ is a tuple $DS = (M, \mathcal{C}, \mathcal{F})$, where $\mathcal{C}$ is the set of *consequences* or outcomes and $\mathcal{F}$ is the set of acts (an *act* is a function f from S to $\mathcal{C}$). We are also interested in *partial acts* whose domains are not necessarily the whole S but subsets of S. Sometimes, in order to emphasize the difference in domains, we also call an act a complete act. Let $\mathcal{D}$ denote the set of all partial acts. A function from Ω to $2^{\mathcal{C}}$ is called a *correspondence*. Our definition is a little different from that in [14] in that the domain is not S but Ω. Let Σ denote the set of all correspondences. We say that a function $h : \Omega \to \mathcal{C}$ *belongs to* a correspondence σ (denoted as $h \in \sigma$) if, for all $\omega \in \Omega$, $h(\omega) \in \sigma(\omega)$. Throughout the paper, we make two technical assumptions that $|\Omega| \geq |2^S|$ and $|S| \geq |\mathcal{C}|$. The first is to guarantee that any belief function on S can be expressed in our Dempster-model $\langle \Omega, S, \Gamma, Pr\rangle$ for some Pr and the second is to guarantee that any *evidential lottery*, which is a belief function on $\mathcal{C}$, can be generated in our decision situation. Note that, if Pr is fixed, then one can generate only a finite subset of belief functions on S and on $\mathcal{C}$ under these two assumptions.

The DM is willing to express his preferences among these elements in a finite set X. A compared pair of the form $x \succeq^X x'$ formalizes the meaning "I prefer x to x' or am indifferent between them". Any form of $x \succeq^X x'$ or any Boolean combination of the forms is called a *proposition*. $x \sim^X x'$ is short for the conjunction of $x \succeq^X x'$ with $x' \succeq^X x$, which expresses indifference relation, and $x \succ^X x'$ denotes $x \succeq^X x'$ but not $x \sim^X x'$. $x \bowtie^X x'$ denotes incomparability. Usually $\succeq^X$ is called *weak preference relation* and $\succ^X$ the strict preference relation. When the context is clear, we usually drop the superscript X in $\succeq^X, \succ^X, \bowtie^X$ and $\sim^X$.

3 UTILITY THEORY

We start in this section with the most fundamental part: a constructive (or logical) system of decision theory under complete ignorance. This system deals with *partial acts*. It turns out that decision theory under complete ignorance with partial acts is a special case of decision-making in Dempster models when the multi-valued mapping Γ is *constant*. Each *partial act* f with the domain $A \subseteq S$ can be identified with its composition with the *constant* multivalued mapping Γ_A defined by $\Gamma_A(\omega) = A$ for all $\omega \in \Omega$. So all partial acts with the same range and the same domain are identified with the same correspondence. In this paper, without further notice, we assume that the ordering on the consequence set $\mathcal{C}$ is *linear*, denoted by $\geq$. In this part, we don't need to assume that the consequence ordering is represented by a utility function.

3.1 Decision theory under complete ignorance

For a partial act d, let $S(d)$ denote its domain and $Rg(d)$ denote its range (sometimes we also denote as $d(S(d))$). Note that $S(d) \subseteq S$ and d may be a partial function. So different acts may have different domains. Denote the set of partial acts with the same domain S' as $\mathcal{D}_{S'}$. Now we are proposing a *new* axiomatization for the weak preference relation $\succeq$ of decision making under complete ignorance. There are *two main differences* of our theory from those about decision making under ignorance in the literature[1, 25, 24, 5, 22]. The first is the absence of the completeness axiom, and the second is our purely *logical* system.

Definition 1 Two partial decisions d_1 and d_2 are *isomorphic* if there is a one-to-one and onto mapping $h : S(d_1) \to S(d_2)$ such that $d_1(s) = d_2(h(s))$ for all $s \in S(d_1)$. The partial act d_2 is said to be *derived from* d_1 if it satisfies the following two conditions:

1. $S(d_2) \subseteq S(d_1)$ and d_1 and d_2 coincide on $S(d_2)$;
2. for each $s \in S(d_1) \setminus S(d_2)$, there is a $s' \in S(d_2)$ such that $d_1(s) = d_1(s')$.

In other words, d_2 is obtained from d_1 by deleting duplicate states. $\lhd$

Just like any standard logical system, our new axiomatic system $\mathcal{Z}$ consists of two parts: the first is *primitive* preferences, and the other *inference rule*:

- Primitive preferences:
  1. (P1: Reflexivity) $d \succeq d$ for any $d \in \mathcal{D}$;
  2. (P2: Weak Dominance Property) For any two partial acts d and d' with the same domain, if $d(s) \geq d'(s)$ for all $s \in S(d)$, then $d \succeq d'$;
  3. (P3: Symmetry) Isomorphic partial acts are preferentially indifferent.
  4. (P4: Duplication) If d' is derived from d, then they are preferentially indifferent.
- Inference Rule:
  - (P5: Transitivity) If $d_1 \succeq d_2$ and $d_2 \succeq d_3$, $d_1 \succeq d_3$.

As one can show, the five properties (P1-P5) are nothing but the four properties proposed by Hurwicz and Arrow in [1] *without* the completeness axiom. Both $P1$ and $P2$ are primitive preferences which are *universally* accepted by a rational agent. The two properties $P3$ and $P4$ are *specifically* used to characterize the complete ignorance in decision making. $P3$ says that no state in the domain is more important than any other state or relabeling acts and states is of no fundamental importance. In other words, the comparison of two partial acts have to be based on comparing the consequences they lead to in different possible states without the knowledge which state of nature is more likely than others. $P4$ requires that merging or splitting states in a domain does not make it better or worse. More than any other property, it captures the idea of complete ignorance, for, in effect, it asserts that dividing a state into several substates should have no effect on decision making. With this property, our characterization of complete ignorance is essentially different from that in subjective probability framework where complete ignorance is expressed by the assignment of equal probabilities to all the states of nature according to Laplace's Principle of Indifference or Insufficient Reason. The transitivity property $P5$ is usually treated as a necessary condition for preference relation. But, in this paper, we regard it as the inference rule.

Definition 2 A *proof* of a proposition ϕ in the system $\mathcal{Z}$ is a finite sequence $\Phi = \phi_1 \cdots \phi_n$ such that $\phi_n = \phi$, and for any $i = 1, \cdots, n$, ϕ_i is either a primitive axiom instance or is obtained from two preceding propositions ϕ_k and $\phi_l (k, l < i)$ by the Inference Rule or propositional reasoning. Proposition ϕ is called *valid* if there is a proof of ϕ in $\mathcal{Z}$. $\lhd$

Lemma 3 *For any two partial acts $d : A \to d(A)$ and $d' : A' \to d'(A')$ (not necessarily with the same domain), if their ranges are the same, i.e., $d(A) = d'(A')$, then $d \sim d'$ is valid in $\mathcal{Z}$.*

Proof. Assume that $d(A) = d'(A')$. For each element $c \in d(A)(= d'(A'))$, pick up one element $a_c \in A$ and $a'_c \in A'$ such that $d(a_c) = c = d'(a'_c)$. It is easy to see that

- $d \restriction_{\{a_c : c \in d(A)\}}$ and $d' \restriction_{\{a'_c : c \in d'(A')\}}$ are isomorphic where $d \restriction_{\{a_c : c \in d(A)\}}$ and $d' \restriction_{\{a'_c : c \in d'(A')\}}$ denote the restrictions of d and d' to $\{a_c : c \in d(A)\}$ and $\{a'_c : c \in d'(A')\}$ respectively;
- $d \restriction_{\{a_c : c \in d(A)\}}$ is derived from d and $d' \restriction_{\{a'_c : c \in d'(A')\}}$ is derived from d'.

So, according to the properties $P3$ and $P4$, the following propositions are valid: $d \sim d \restriction_{\{a_c : c \in d(A)\}}$, $d \restriction_{\{a_c : c \in d(A)\}} \sim d' \restriction_{\{a'_c : c \in d'(A')\}}$ and $d' \restriction_{\{a'_c : c \in d'(A')\}} \sim d'$. It follows immediately from $P5$ that $d \sim d'$ is valid. QED

Lemma 4 *For any two partial acts $d : A \to d(A)$ and $d' : A' \to d'(A')$ (not necessarily with the same domain), if $min(d(A)) = min(d'(A'))$ and $max(d(A)) = max(d'(A'))$, then $d \sim d'$ is valid in $\mathcal{Z}$.*

Proof. By $P4$, we may assume that d is a bijection with the domain $A = \{s_1, \cdots, s_n\}$ with $d(s_i) = c_i (i = 1, \cdots, n)$ where $c_1 > c_2 > \cdots > c_n$. Now we construct three partial acts d_e, g and h as follows:

- $S(d_e) = \{s_1, s_n\}$ and $d_e(s_1) = c_1$ and $d_e(s_n) = c_n$;
- $S(g) = \{s_1, \cdots, s_n\}$ and $g(s_1) = c_1$ and $g(s_k) = c_n$ for $k = 2, \cdots, n$;
- $S(h) = \{s_1, \cdots, s_n\}$ and $h(s_n) = c_n$ and $g(s_k) = c_1$ for $k = 1, \cdots, n-1$;

It is easy to see that $d(s) \geq g(s)$ for all $s \in S(d)(= S(g))$. So, according to P2, $d \succeq g$ is valid. It is easy to see that $h(s) \geq d(s)$ for all $s \in S(d)(= S(g))$. So, according to P2, $h \succeq d$ is valid. Moreover, according to Lemma 3, both $g \sim d_e$ and $d_e \sim h$ are valid. By $P5$, we have that $d_e \sim d$ is valid. Similarly, we can show that d' is preferentially indifferent to the following d'_e:

$$d'_e : \{s'_1, s'_2\} \to \{c_1, c_n\} \text{ s.t. } d'_e(s'_1) = c_1 \text{ and } d'_e(s'_2) = c_n$$

By Lemma 3, we know that $d_e \sim d'_e$ is valid. So we know from $P5$ that $d \sim d'$ is valid.

QED

Lemma 5 *For any two partial acts d_1 and d_2 (not necessarily with the same domain), there are two partial acts d'_1 and d'_2 with the same domain such that both $d_1 \sim d'_1$ and $d_2 \sim d'_2$ are valid.*

Proof. Let $S_\cup$ be the union of $S(d_1)$ and $S(d_2)$. Define a new act d_1' with domain $S_\cup$ as follows:

$$d_1'(S) = \begin{cases} d_1(s) & \text{if } s \in S(d_1), \\ d_1(s_0) & \text{if } s \in S_\cup \setminus S(d_1). \end{cases}$$

where s_0 is some element in $S(d_1)$. According to $P4$, $d_1' \sim d_1$ is valid. Similarly, we can define a new act d_2' with the domain $S_\cup$ such that $d_2' \sim d_2$ is valid.

QED

Lemma 6 *For any two partial acts d and d' with the domains A and B (not necessarily with the same domain) respectively, if $max(d(A)) \geq max(d'(B))$ and $min(d(A)) \geq min(d'(B))$, then $d \succeq d'$ is valid.*

Proof. From Lemma 5, we may assume that d and d' share the same domain $S_0 = \{s_1, \cdots, s_n\}$ and $max(d(S_0)) \geq max(d'(S_0))$ and $min(d(S_0)) \geq min(d'(S_0))$. Define two new acts d_e and d_e' as follows:

- $S(d_e) = \{s_1, s_n\}$, $d_e(s_1) = max(d(S_0))$ and $d_e(s_n) = min(d(S_0))$;
- $S(d_e') = \{s_1, s_n\}$, $d_e'(s_1) = max(d'(S_0))$ and $d_e'(s_n) = min(d'(S_0))$

It is easy to see that propositions $d_e \succeq d_e', d \sim d_e$ and $d' \sim d_e'$ are all valid. So $d \succeq d'$ is valid. QED

The above lemmas are standard in the literature about complete ignorance [1, 22]. One only needs to check that the proof there is within our logic. However, the following proposition, which is converse to Lemma 6, offers a *new* logical perspective.

Lemma 7 *For any two partial acts d and d' with the domains A and B (not necessarily with the same domain) respectively, if $d \succeq d'$ is valid, then $max(d(A)) \geq max(d'(A))$ and $min(d(A)) \geq min(d'(A))$.*

Proof. We employ the proof-by-structural induction. Assume that $\phi := (d \succeq d')$ is valid. According to Definition 2, there is a finite sequence of propositions $\Phi = \phi_1 \cdots \phi_n$ such that $\phi_n = \phi$, and for any $i = 1, \cdots, n$, ϕ_i is either a primitive axiom instance or is obtained from two preceding propositions ϕ_k and $\phi_l (k, l < i)$ by the Inference Rule or propositional reasoning. We prove by structural induction. If ϕ is a primitive axiom instance, it is easy to see that the proposition holds. We assume that ϕ is obtained from two preceding propositions $\phi_k = (d \succeq h)$ and $\phi_l = (h \succeq d')(k, l < i)$ for some partial act h with the same domain S_0 as d and d' (Lemma 5) by the Inference Rule such that

- $max(d(S_0)) \geq max(h(S_0)), min(d(S_0)) \geq min(h(S_0))$, and
- $max(h(S_0)) \geq max(d'(S_0)), min(h(S_0)) \geq min(d'(S_0))$.

From these induction hypotheses, it follows immediately that $max(d(S_0)) \geq max(d'(S_0))$ and $min(d(S_0)) \geq min(d'(S_0))$. So we may conclude that $max(d(A)) \geq max(d'(A))$ and $min(d(A)) \geq min(d'(A))$.

QED

The following proposition follows directly from Lemmas 6 and 7.

Proposition 8 *For any two partial acts d and d' with domains A and B respectively, $d \succeq d'$ is valid iff $min(d'(B)) \leq min(d(A))$ and $max(d'(B)) \leq max(d(A))$.*

Corollary 9 *For any two acts d and d' with domains A and B respectively,*

1. *$d \succ d'$ iff $[min(d'(B)) < min(d(A))$ and $max(d'(B)) \leq max(d(A))]$ or $[min(d'(B)) \leq min(d(A))$ and $max(d'(B)) < max(d(A))]$.*
2. *$d \bowtie d'$ iff $[min(d'(B)) < min(d(A))$ and $max(d'(B)) > max(d(A))]$ or $[min(d'(B)) > min(d(A))$ and $max(d'(B)) < max(d(A))]$.*

In other words, the result of our inferential system is the transitive closure of the set of primitive sentences, producing a relation on the set $\mathcal{D}$ of partial acts. For any $\mathcal{C}' \subseteq \mathcal{C}$, let $\Omega_{\mathcal{C}'}$ denote the mapping $\Omega \to 2^{\mathcal{C}}$ defined as $\Omega_{\mathcal{C}'}(\omega) = \mathcal{C}'$ for all $\omega \in \Omega$. Note that each partial act d with the domain A is identified with its composition with the constant multi-valued mapping $\Gamma_A : \Gamma_A(\omega) = A$ for all $\omega \in \Omega$, which is the constant mapping $\Omega_{d(A)}$. Proposition 8 says that this identification is well-defined. Following a similar convention in Bayesian decision theory, we also identify each subset $\mathcal{C}'$ of consequences with the constant correspondence $\Omega_{\mathcal{C}'}$. From these two identifications justified by the principle of complete ignorance, we obtain a natural extension of the ordering of outcomes to that of sets of outcomes [3]: for any two subsets $\mathcal{C}_1, \mathcal{C}_2$ of outcomes,

$$\mathcal{C}_1 \succeq \mathcal{C}_2 \text{ iff } min\mathcal{C}_1 \geq min\mathcal{C}_2 \& max\mathcal{C}_1 \geq max\mathcal{C}_2 \tag{1}$$

3.2 Decision theory in Dempster models

For any two subsets X and Y of real numbers, their (Minkowski) *sum* is defined by $X + Y = \{x + y : x \in X, y \in Y\}$. The *product* aX of a nonnegative real number a and a subset X is defined by $aX = \{ax : x \in X\}$. Here we use the notation $+$ for both the Minkowski sum of sets and standard addition of reals. The context will determine which one we mean.

First we motivate three different but *equivalent* forms of expectation in terms of belief functions in Dempster models and then provide an axiomatization. Assume that the decision situation is $\langle M, \mathcal{C}, \mathcal{F} \rangle$ where $M = \langle \Omega, S, \Gamma, Pr \rangle$. Let f be an (complete) act from S to $\mathcal{C}$, i.e., a function from S to $\mathcal{C}$. For each possible knowledge state ω, the act f will have the set of consequences $f(\Gamma(\omega))$, which is a subset of $\mathcal{C}$. Let $u : \mathcal{C} \to \mathbb{R}$ be a *utility function*. With the prior probability Pr defined on Ω, it is natural to define the expectation of f with respect to the Dempster model as follows:

$$\mathbb{E}^{\Gamma}_{Pr}(f) = \sum_{\omega \in \Omega} Pr(\omega) u(f(\Gamma(\omega))) \tag{2}$$

Note that, in order to define Minkowski sum in the expectation, the utility function is needed. We can also define the expected utility (set) of f with respect to the mass function m on S generated by Pr and Γ. For any subset $A \subseteq S$, the agent's belief on A is represented by $m(A)$. According to the principle of complete ignorance, since the agent could not distinguish the elements in A, the consequence *set* is $f(A)$. So f can be decomposed as a "sum" of partial acts f_A's ($A \subseteq S$) withe domain A defined as $f_A(A) = f(A)$. A *natural* expectation of the act f with respect to the mass function m can be defined as follows:

$$\mathbb{E}^{Mi}_m(f) := \sum_{A \subseteq S} m(A) u(f(A)) \tag{3}$$

It is easy to see that Eq. (3) is derived from Eq. (2). On the other hand, given the Dempster model, each act $f : S \to \mathcal{C}$ is associated with a *correspondence* from Ω to $2^{\mathcal{C}}$ which is the composition $f \circ \Gamma : \Omega \to 2^{\mathcal{C}}$. For the correspondence $f \circ \Gamma$, we define $\mathcal{F}_{f\circ\Gamma} = \{g : g$ is a function from Ω to $\mathcal{C}$ such that $g(\omega) \in g \circ \Gamma(\omega)$ for all $\omega \in \Omega\}$. Now we define a new expectation of the act f in terms of the *Aumann integral* of the correspondence $f \circ \Gamma$ with respect to the probability function Pr on Ω

$$\mathbb{E}^{Au}_{Pr,\Gamma}(f) = \int_\Omega u \circ f \circ \Gamma(\omega) Pr(\omega) := \{\sum_{\omega\in\Omega} u(g(\omega))Pr(\omega) : g \in F_{f\circ\Gamma}\} \quad (4)$$

Both Minkowski sum and Aumann integral are defined over sets of real numbers. Mathematically, sets of real numbers are incomparable. But, if sets of consequences are mapped by utility function into sets of real numbers, they can be compared according to Eq. (1) (the ordering may be incomplete though).

Lemma 10 *For the above Dempster model and any complete act f,*

$$\mathbb{E}^{\Gamma}_{Pr}(f) = \mathbb{E}^{Mi}_{m}(f) = \mathbb{E}^{Au}_{Pr,\Gamma}(f).$$

Lemma 10 supports our claim that our definition of expectations in terms of belief functions in Eq. (3) is natural and robust. Given an act f, its composition with the multivalued mapping Γ is a correspondence from Ω to $\mathcal{C}$. On the other hand, any correspondence $\sigma : \Omega \to 2^{\mathcal{C}}$ can be decomposed into a multivalued mapping $\Gamma_\sigma : \Omega \to 2^S$ and an act $f_\sigma : S \to \mathcal{C}$. Note that Γ_σ may be different from Γ. Although such composition is not unique, all these compositions are equivalent in the sense of Lemma 10. So, in this section, we choose to represent the preference over correspondences instead of that of acts. σ is called a *crisp* correspondence if $\sigma(\omega)$ is a singleton for all $\omega \in \Omega$. It is also regarded as a function from Ω to $\mathcal{C}$. Let Σ_c denote the set of all crisp correspondences. A crisp correspondence σ_S belongs to a correspondence σ (denoted as $\sigma_S \in \sigma$) if $\sigma_S(\omega) \in \sigma(\omega)$ for all $\omega \in \Omega$. For any correspondence σ, σ_m denotes the crisp correspondence which always takes the minimum in each consequence set. In other words, $\sigma_m(\omega) = min(\sigma(\omega))$. Similarly σ_M denote the crisp correspondence that always takes the largest in each correspondence set, i.e., $\sigma_M(\omega) = max(\sigma(\omega))$ for all $\omega \in \Omega$. It is easy to see that $\sigma_m \in \sigma$ and $\sigma_M \in \sigma$.

In this section, we establish an Anscombe-Aumann-style decision theory for belief functions. Anscombe-Aumman model has $\mathcal{C}$ and Ω just as does Savage's. However, correspondences do not map states directly into sets of outcomes, but into *finite sets* of von-Neumann-Morgenstern (vNM for short) lotteries over outcomes [29]. Formally, the vNM lotteries are probability distributions over $\mathcal{C}$. Let $\Delta(\mathcal{C})$ denote the set of all probability distributions over $\mathcal{C}$. It is endowed with the usual mixing operation. We will endow the set Σ^Δ of correspondences from Ω to $\{R : R \subseteq \Delta(\mathcal{C}), R \text{ is finite}\}$ with a mixture operation as well, performed pointwise. That is, for every $\sigma, \sigma' \in \Sigma^\Delta$ and every $\alpha \in [0,1], \alpha\sigma + (1-\alpha)\sigma' \in \Sigma^\Delta$ is given by $(\alpha\sigma + (1-\alpha)\sigma')(\omega) = \alpha\sigma(\omega) + (1-\alpha)\sigma'(\omega)$ for any $\omega \in \Omega$. As for the decision-theoretic interpretation of elements of Σ^Δ, we actually adapted a similar interpretation from David Schmeidler in [29] as horse lotteries. Elements of $\Delta(\mathcal{C})$ are as random outcomes or (roulette) lotteries. Σ^Δ_c reduces to the set of classical Anscombe-Aumann acts. We will denote the decision maker's preference order also by $\succcurlyeq \subseteq \Sigma^\Delta \times \Sigma^\Delta$ and we abuse this notation as usual. In particular, we can write, for $C', D' \subseteq \Delta(\mathcal{C}), C' \succcurlyeq D'$ understood as $\sigma_{C'} \succcurlyeq \sigma_{D'}$ where, for every finite $R \subseteq \Delta(\mathcal{C}), \sigma_R$ is the constant correspondence given by $\sigma_R(\omega) = R$ for all $\omega \in \Omega$. For a function $u : \mathcal{C} \to \mathbb{R}$ and $P \in \Delta(\mathcal{C})$, we will use the notation $u(P)$ for $\sum_{c\in\mathcal{C}} P(c)u(c)$.

The following are our Anscombe-Aumann-style axiomatization for belief functions. In order to obtain the representation theorem in Prop.11, we have to consider all elements in Σ^Δ. In other words, we have to consider all Pr on Ω. The first three are the counterparts to the von-Neumann-Morgenstern axioms without the completeness property. The next two axioms are similar to de Finetti's last two axioms, guaranteeing monotonicity and non-triviality. And the last axiom is characteristic: it connects our axiomatization to the classic Anscombe-Aumman axiomatization for Bayesian theory and it takes the same role as Axiom 8 in [14].

- AA1 (Restricted weak order): $\succcurlyeq$ is reflexive, transitive on Σ^Δ and it is complete on the set Σ^Δ_c of crisp ones.
- AA2 (Continuity): For every $\sigma_1, \sigma_2, \sigma_3 \in \Sigma^\Delta$, if $\sigma_1 \succ \sigma_2 \succ \sigma_3$, there exist $\alpha, \beta \in (0,1)$ such that $\alpha\sigma_1 + (1-\alpha)\sigma_3 \succ \sigma_2 \succ \beta\sigma_1 + (1-\beta)\sigma_3$.
- AA3 (Independence): For $\sigma_1, \sigma_2, \sigma_3 \in \Sigma^\Delta$, and every $\alpha \in (0,1)$, $\sigma_1 \succeq \sigma_2$ iff $\alpha\sigma_1 + (1-\alpha)\sigma_3 \succeq \alpha\sigma_2 + (1-\alpha)\sigma_3$.
- AA4 (Monotonicity): For $\sigma, \sigma' \in \Sigma^\Delta$, $\sigma(\omega) \succeq \sigma'(\omega)$ for all $\omega \in \Omega$ implies $\sigma \succeq \sigma'$.
- AA5 (Nontriviality): There exist $\sigma, \sigma' \in \Sigma^\Delta$ such that $\sigma \succ \sigma'$.
- AA6 (Consistency): For any correspondence g and two crisp correspondences σ_S, σ'_S, $\sigma_S \succeq g \succeq \sigma'_S$ iff $\sigma_S \succeq g_S \succeq \sigma'_S$ for all crisp correspondence $g_S \in g$.

Proposition 11 *$\succeq$ satisfies the Axioms AA1-AA6 if and only if there exists a probability measure pr over Ω and a non-constant function $u : \mathcal{C} \to \mathbb{R}$ such that, for every $\sigma, \sigma' \in \Sigma^\Delta$,*

$$\sigma \succeq \sigma' \textit{ iff } \sum_{\omega\in\Omega} pr(\omega)u(\sigma(\omega)) \succeq \sum_{\omega\in\Omega} pr(\omega)u(\sigma'(\omega)) \quad (5)$$

Proof. It is easy to check the sufficiency. Now we only need to show the necessity. First we consider the preference ordering $\succeq$ restricted to the set of crisp correspondences, which is denoted as $\succeq_c$. It is easy to see that $\succeq_c$ satisfies all the the Anscombe-Aumann axioms for Bayesian theory. So there is a probability measure over Ω and a utility function $u : \mathcal{C} \to \mathbb{R}$ such that, for any crisp σ_c, σ'_c,

$$\sigma_c \succeq \sigma'_c \text{ iff } \sum_{\omega\in\Omega} pr(\omega)u(\sigma_c(\omega)) \geq \sum_{\omega\in\Omega} pr(\omega)u(\sigma'_c(\omega))$$

Now it remains to show that the probability measure pr and utility function u work for Σ^Δ, i.e., Eq. (5) holds. Assume that $\sigma \succeq \sigma'$. Let $\mathbf{0}$ and $\mathbf{1}$ denote the two constant correspondences that map to the smallest and largest outcomes respectively. It is easy to see that $\mathbf{1} \succ \mathbf{0}$. Let $\alpha_\sigma = sup\{\alpha : \sigma \succeq \alpha\mathbf{1} + (1-\alpha)\mathbf{0}\}$ and $\beta_\sigma = inf\{\beta : \beta\mathbf{1} + (1-\beta)\mathbf{0} \succeq \sigma\}$. Now we need to show that $\sigma_m \sim \alpha_\sigma$. According to AA6, we know that $\sigma_m \succeq \alpha_\sigma$. Since, for all $\omega \in \Omega$, $\sigma(\omega) \succeq \sigma_m(\omega)$, $\alpha_\sigma \succeq \sigma_m$ according to (AA4). So we have shown that $\sigma_m \sim \alpha_\sigma$. Similarly, we can show that $\sigma_M \sim \beta_\sigma$. So $\sigma \succeq \sigma'$ iff $\sigma_m \succeq \sigma'_m$ and $\sigma_M \succeq \sigma_M$.

QED

The preference ordering of *acts* in the decision situation $\langle M, \mathcal{C}, \mathcal{F}\rangle$ where $M = \langle \Omega, S, \Gamma, Pr\rangle$ is defined as follows: for any two acts $f, g \in \mathcal{F}$,

$$f \succeq g \text{ if } f \circ \Gamma \succeq g \circ \Gamma \quad (6)$$

which depends on the specific multivalued mapping Γ.

For the above Anscombe-Aumann-style setting, given a belief function bel on S, let $\mathcal{P}_{bel}$ denote the set of probability distributions dominating bel, i.e., $\mathcal{P}_{bel} = \{pr : pr$ is a probability distribution on S such that, for any $A \subseteq S, pr(A) \geq bel(A)\}$. Now we can define for any $f \in \mathcal{F}$,

$$\mathbb{E}^{im}_{bel}(f) = \{\textstyle\sum_{s\in S} pr(s)u(f(s)) : pr \in \mathcal{P}_{bel}\}.$$

It is a closed interval $\mathbb{E}^{im}_{bel}(f) = [\underline{\mathbb{E}}_{bel}(f), \bar{\mathbb{E}}_{bel}(f)]$ where $\underline{\mathbb{E}}_{bel}(f) = \sum_{A\subseteq S} m(A)min(u(f(A)))$ and $\bar{\mathbb{E}}_{bel}(f) = \sum_{A\subseteq S} m(A)max(u((f(A))))$ [11]. Let $\succeq^{im}$ defined as:

$$f \succeq^{im} g \text{ iff } \underline{\mathbb{E}}_{bel}(f) \geq \underline{\mathbb{E}}_{bel}(g) \& \bar{\mathbb{E}}_{bel}(f) \geq \bar{\mathbb{E}}_{bel}(g).$$

The following expectation is called *Choquet expected utility* of an act f with respect to the mass function m [29],

$$\mathbb{E}^{Ch}_m(f) := \textstyle\sum_{A\subseteq S} m(A)u(min(f(A))).$$

For the *expected uncertainty utility* [18], we require that the utility function $u_{\mathbb{I}}$ is defined on the set of intervals $\mathbb{I} := \{(c, c') : c \leq c'\}$ satisfying the conditions: it is continuous and strictly increasing,i.e., $u_{\mathbb{I}}(c_1, c_2) > u_{\mathbb{I}}(c'_1, c'_2)$ whenever $c_1 > c'_1$ and $c_2 > c'_2$. Then the expectation $\mathbb{E}^{u_{\mathbb{I}}}_m(f)$ of an act f with respect to mass function m is defined as

$$\mathbb{E}^{u_{\mathbb{I}}}_m(f) = \textstyle\sum_{A\subseteq S} m(A)u_{\mathbb{I}}(min(f(A)), max(f(A))).$$

Let α be a function from 2^S to [0,1]. $\alpha(A)$ is a pessimimism index depending on A. For any act f, the Hurwicz expected utility $\mathbb{E}^{H,\alpha}_m(f)$ is defined as [14]:

$$\textstyle\sum_{A\subseteq S}[\alpha(A)min(u(f(A))) + (1 - \alpha(A))max(u(f(A)))]m(A).$$

Note that the Hurwicz expected utility $\mathbb{E}^{H,\alpha}_m$ is a special case of expected uncertain utility $\mathbb{E}^{u_{\mathbb{I}}}_m$. The last α-Maxmin expected utility is defined in the *imprecise-probability* semantics for belief functions. For an index $\alpha \in [0, 1]$, the α-maxmin expected utility $\mathbb{E}^{\alpha,M}_m(f)$ is defined as

$$\alpha \cdot min_{pr\in\mathcal{P}_{bel}}(\mathbb{E}_{pr}(u \circ f) + (1 - \alpha)max_{pr\in\mathcal{P}_{bel}}(\mathbb{E}_{pr}(u \circ f)).$$

Proposition 12 *Let $\succeq^{Ch}, \succeq^{u_{\mathbb{I}}}, \succeq^{H,\alpha}$ and $\succeq^{\alpha,M}$ denote the corresponding preference ordering represented by these four kinds of expectations. Then the following equalities hold: $\succeq = \succeq^{im} = \bigcap_{u_{\mathbb{I}}} \succeq^{u_{\mathbb{I}}} = \bigcap_\alpha \succeq^{H,\alpha} = \bigcap_\alpha \succeq^{\alpha,M}$.*

Proof. The first part of the proposition that $\succeq = \succeq^{im}$ is known in the literature of imprecise probabilities. Here we give a proof idea of the equality that $\succeq = \bigcap_\alpha \succeq^{H,\alpha}$. It is easy to see that $\succeq$ is a subset of $\bigcap_\alpha \succeq^{H,\alpha}$. For the other direction, it suffices to show that, if it is not true that $f \succeq g$, there are α and α' such that $f \succeq^{H,\alpha} g$ but not $f \succeq^{H,\alpha'} g$. This follows from Corollary 9. QED

To emphasize the relations among different orderings, we make explicit the domains. Let $\geq_{\mathcal{C}}, \succeq_{2^{\mathcal{C}}}$ and $\succeq_{\mathcal{F}}$ denote the linear ordering on $\mathcal{C}$, the partial ordering on $2^{\mathcal{C}}$ in Eq. (1) and the partial ordering on $\mathcal{F}$ in Eq. (6) respectively. The extensions of the above preference relations can be illustrated as follows:

$$\geq_{\mathcal{C}} \Rightarrow \succeq_{2^{\mathcal{C}}} \Rightarrow \succeq_{\mathcal{F}} \Rightarrow \succeq^{u_{\mathbb{I}}} \begin{cases} \succeq^{H,\alpha} \\ \succeq^{\alpha,M} \\ \succeq^{Ch}. \end{cases} \tag{7}$$

The first extension is shown in Proposition 8 and the second is justified in Proposition 11. Proposition 12 and Eq. (7) are two reasons why we call our theory here *basic*.

4 GROUP DECISION MAKING

In this section we extend the above Anscombe-Aumann model for belief functions to the multiagent setting. Individual agents are indexed with $i = 1, \cdots, n$, and a social observer is represented by index $i = 0$. Both individual and social agents express their subjective beliefs indirectly, i.e. by stating their preferences $\succeq^i$ over uncertain prospects ($i = 0, 1, \cdots, n$). We assume that $\succeq^i$ ($i = 0, 1, \cdots, n$) satisfy the subjective interval utility theory, i.e., the AA1-AA6 axioms of last section, for all $i = 0, 1, \cdots, n$ and they are represented by

$$U^i(\cdot) := \sum_\omega p^i(\omega)u^i(\cdot), \tag{8}$$

We consider the following Paretian conditions: for all $\sigma, \sigma' \in \Sigma^\Delta$,

- (C) $\sigma \sim^i \sigma'$ for all $i = 1, \cdots, n$ implies $\sigma \sim^0 \sigma'$
- $(C_1)\sigma \succeq^i \sigma'$ for all $i = 1, \cdots, n$ implies $\sigma \succeq^0 \sigma'$
- $(C_2)\sigma \succ^i \sigma'$ for all $i = 1, \cdots, n$ implies $\sigma \succ^0 \sigma'$
- $(C_3)\sigma \succeq^i \sigma'$ for all $i = 1, \cdots, n$ and $\exists j : \sigma \succ^j \sigma'$ implies $\sigma \succ^0 \sigma'$

In social choice theory, these are the conditions of Pareto-Indifference, Pareto-Weak Preference, Weak Pareto, Strict Pareto, respectively. The Strong Pareto C^+ is the combination of (C) and (C_3).

Lemma 13 *The Paretian indifference (C) holds if and only if there are real numbers $a_1, \cdots, a_n, b$ such that $U^0(\sigma) \sim \sum_{i=1}^n a_iU^i(\sigma)+ b$ for all correspondence σ. And (C_1) (resp. (C^+)) holds if and only if this equation is satisfied for some choice of non-negative (resp. positive) numbers $a_1, \cdots, a_n$.*

Proof. Assume that $\succeq^0, \succeq^1, \cdots, \succeq^n$ satisfy the AA1-AA6 axioms. Now we restrict the considerations to the set of crisp correspondences Σ^Δ_c. Let $\succeq^0_S, \succeq^1_S, \cdots, \succeq^n_S$ denote these restrictions. It is easy to see that they satisfy the AA1-AA5 axioms for Bayesian Anscombe-Aumann models. So, according to Lemma 2 in [27], the corresponding restriction $U^0_S = \sum_{i=1}^n a_iU^i_S + b$ holds. Now consider σ_m and σ_M. It is easy to see that $U^0_S(\sigma_m) = \sum_{i=1}^n a_iU^i_S(\sigma_m) + b$ and $U^0_S(\sigma_M) = \sum_{i=1}^n a_iU^i_S(\sigma_M) + b$. QED

Note that it is generally not necessarily true that $U^0_S(\sigma) = \sum_{i=1}^n a_iU^i_S(\sigma) + b$. From this lemma, we can easily obtain the following theorem.

Lemma 14 *If all individual prior beliefs are the same, then (C) holds if and only if there are real numbers $a_1, \cdots, a_n, b$ such that $u_0 = \sum_{i=1}^n a_iu_i + b$. And (C_1) (resp. (C^+)) holds if and only if this equation is satisfied for some choice of non-negative (resp. positive) numbers $a_1, \cdots, a_n$.*

Let $\mathcal{B}$ be the class of all belief functions on the consequence set $\mathcal{C}$. For each DM $i = 0, 1, \cdots, n$, the preference ordering $\succeq^i$ on $\mathcal{B}$ is defined according to the following representation: for any two $bel, bel' \in \mathcal{B}$,

$$bel \succeq^i bel' \text{ iff } \sum_{C\subseteq\mathcal{C}} m(C)u^i(C) \succeq^i \sum_{C\subseteq\mathcal{C}} m'(C)u^i(C)$$

where u^i comes from Eqs. (8). Similarly we can define different Paretian conditions for these preference orderings on $\mathcal{B}$. Since

$\Delta(\mathcal{C}) \subseteq \mathcal{B}$, the following version of aggregation theorem can be derived directly from Harsanyi's aggregation theorem for vNM lotteries [20].

Lemma 15 *(Aggregation Theorem) For the above defined preference ordering on $\mathcal{B}$, if they satisfy the Paretian indifference condition, then there are real numbers $a_1, \cdots, a_n, b$ such that $u^0 = \sum_{i=1}^n a_i u^i + b$. And (C_1) (resp. (C^+)) holds if and only if this equation is satisfied for some choice of non-negative (resp. positive) numbers $a_1, \cdots, a_n$.*

Given the representations in Eq. (8), we say that agent i is a *probability dictator* if $p^0 = p^i$ fo some i, that i is a *utility dictator* if $u^0 = u^i$ up to positive affine transformation, and that i is an *overall dictator* if he is both a probability and a utility dictator. We define i to be an *inverse utility dictator* or an *inverse overall dictator* by changing the clause that $u^0 = u^i$ into $u^0 = -u^i$. In general both probabilities and utilities should be expected to vary from one individual to another. A set of elements $\{\phi_1, \cdots, \phi_k\}$ of a vector space is *affinely independent* if for any set of real numbers $a_1, \cdots, a_k, b$, $a_1\phi_1 + \cdots + a_k\phi_k + b = 0$ implies $a_1 = \cdots = a_k = b = 0$. This concept provides the notion of algebraic independence in the case of utility functions while linear independence characterizes the diversity in individual beliefs [27].

Lemma 16 *Denoting by $p^1, \cdots, p^n$ the probabilities and by $u^1, \cdots, u^n$ the utility functions on consequences in Eq. (5) in Proposition 11. Then, if (C) holds,*

- *there is either a utility or an inverse utility dictator in Case (*): $p^1, \cdots, p^n$ are linearly independent*
- *there is a probability dictator in Case (**): $u^1, \cdots, u^n$ are affinely independent.*

There is an overall or an inverse overall dictator when both (*) and (**) apply. If either (C_1) or (C_2) holds, the same results follow, except that there is always a utility dictator in Case (*).

Proof. Assume that (C) holds and $p^1, \cdots, p^n$ are linearly independent. Just consider the set Σ_c^Δ of all crisp correspondences. According to Proposition 4 in [27], there is a utility dictator or an inverse utility dictator. The second part can be shown similarly. QED

All the above Paretian conditions are defined for correspondences. We may also define the corresponding Paretian conditions $(C'), (C_1'), (C_2')$ and (C_3') for acts instead. Although $f \sim^i g$ for acts f, g and all $i = 1, \cdots, n$, it is generally not true that correspondences $f \circ \Gamma_i \sim^i g \circ \Gamma_i$ for acts f, g and all $i = 1, \cdots, n$ because Γ_i's may be different. So the above theorem generally doesn't hold under these Paretian conditions for acts. However, they are true for the following special cases where we may assume that all individuals have the same multivalued mappings.

Corollary 17 *If all individual prior mass functions are the same, then (C') holds if and only if there are real numbers $a_1, \cdots, a_n, b$ such that $u_0 = \sum_{i=1}^n a_i u_i + b$. And (C_1') holds if and only if this equation is satisfied for some choice of non-negative numbers $a_1, \cdots, a_n$.*

Corollary 18 *If (C') holds, then there is either a utility or an inverse utility dictator provided $m^1, \cdots, m^n$ are linearly independent.*

5 RELATED WORKS AND CONCLUSION

The most comprehensive and up-to-date survey about decision theory for belief functions is [11]. To the best of our knowledge, we are the first to establish a theory for the representation in Eq. (5). It was first proposed in [34]. Recently it was mentioned in [11] and called weak dominance relation. But neither of them justified the representation of this preference relation. Our results here are quite different from those by Denoeux and Shenoy in [12]. Firstly, our axiomatic system in Section 3.2 is of Anscombe-Aumman style which derives both representations of beliefs and tastes. In contrast, their axiomatization is analogous to von Neumann-Morgensterns utility theory for probabilistic lotteries as described by Luce and Raiffa which assumes belief function lotteries first and induces only utilities. Secondly, our representation differs from theirs in that our upper and lower bounds are the expectations of maximal and minimal utilities respectively while the upper and lower bounds in their representation are not dependent on but consistent with maximal and minimal utilities (pointed out by them in the paragraph following Def. 6). Thirdly, our axiomatization is more fundamental and fits better with the interpretation of belief functions as probabilities of sets, in particular with Dempster model. In accordance with this semantics, we obtain the representation (Th.3.9) by combining the logical system for complete ignorance (represented by set-valued mappings) and the Anscombe-Aumann system for probabilities. Both systems follow from the well-established principles. But one key assumption, Assumption 3.3 (Continuity), in their system is not intuitive. In order to make this assumption reasonable, we are led to develop a logical system for complete ignorance in Section 3.1.

In the imprecise probability literature, [2] also removes the completeness assumption from the Anscombe-Aumann formulation of Savage's theory and introduces an inertia assumption. Moreover the lower and upper Choquet integrals in [8] give the same lower and upper bounds as in our representation. But they are not constructed from complete ignorance. The belief-function framework in this paper is adapted from [21]. Jaffray and Wakker justified an interval-utility representation, which is similar to $\succeq^{u_{\mathbb{I}}}$, by means of a neutrality axiom and a weakened form of the sure thing principle. [18] provides a Savage-style utility theory for the same representation but in a general case. [37] proposes a similar theory for Halpern and Fagin's inner measure semantics but in a finite-state setting. [14] also introduces a similar axiomatization not for acts but for correspondences representing unforeseen contingencies. The representations in those papers do depend on the maximal and minimal elements of sets of outcomes but the represented preference relation is complete. [15] considers a similar but more general decision making under uncertainty comprising complete ignorance and probability where decision making with belief functions is a special case. But the issue there is not about the preference representation but the reversibility of an ignorant variable and a probability variable. [33] defends a two-level mental model, composed of a credal level, where an agent's beliefs are represented by belief functions, and the pignistic level, where decisions are made by maximizing expected utility with respect to a probability measure derived from a belief function through the so-called pignistic transformation. [31] suggests that a constructive decision theory should be based not on utilities, but on goals. [19] provides an axiomatization of expectation in terms of belief functions. A preference representation in terms of belief functions may be derived from Choquet integrals [29]. Related literature also includes [36, 16, 23, 13, 6]. Generalizations of the Anscombe-Aumann model in the frameworks of sets of probability measures or belief functions on consequences are studied in [35, 7]. None of them proposes a representation as in this paper.

As far as we know, we are the first to investigate social choice theory from the perspective of Dempster-Shafer theory. Here we mainly follow the classical works from Harsanyi [20] and Mongin [26, 27]. In this paper, we have shown that, in many decision problems of interest, preferences that can be *solidly* justified are incomplete. Our choice theory may be regarded to be "rationally objective" in the sense of [17]. Yet decisions eventually have to be made. We would like to investigate other *complete* "subjectively rational choices" and provide their connection to our theory in this paper. We also would like to investigate how social aggregation interacts with the characteristic Dempster's rule of combination. Moreover, we want to explore aggregating partially ordered preferences [28] and, more broadly, social choice theory in AI [4] from the Dempster-Shafer perspective. It is also desirable to provide tools to elicit/operationalize the notions of belief functions and/or utilities.

ACKNOWLEDGEMENTS

We would like to thank the referees for their constructive comments. The first author is partly supported by NSFC No.61732006 and would like to thank Professors Arthur Dempster, Xiao-Li Meng and Ruobin Gong for their helpful discussions during his visiting scholarhsip at Harvard Statistics Department. The second author is partly supported by NSFC No. 61772534.

REFERENCES

[1] Kenneth Arrow and Leonid Hurwicz, 'An optimality criterion for decision making under ignorance.', in *Uncertainty and Expectations in Economics*, eds., C. F. Carter and J. I. Ford., Oxford: Basil Blackwell and Mott Ltd., (1972).

[2] Truman F Bewley, 'Knightian decision theory. part i', *Decisions in economics and finance*, **25**(2), 79–110, (2002).

[3] Walter Bossert, 'On the extension of preferences over a set to the power set: an axiomatic characterization of a quasi-ordering', *Journal of Economic Theory*, **49**(1), 84–92, (1989).

[4] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia, *Handbook of computational social choice*, Cambridge University Press, 2016.

[5] Michele Cohen and Jean-Yves Jaffray, 'Rational behavior under complete ignorance', *Econometrica*, **48**(5), 1281–1299, (1980).

[6] Giulianella Coletti, Davide Petturiti, and Barbara Vantaggi, 'Rationality principles for preferences on belief functions', *Kybernetika*, **51**(3), 486–507, (2015).

[7] Giulianella Coletti, Davide Petturiti, and Barbara Vantaggi, 'Decisions on generalized anscombe-aumann acts under possibly unexpected scenarios', in *Proc. WUPES 2018*, 49–60, (2018).

[8] Giulianella Coletti, Davide Petturiti, and Barbara Vantaggi, 'Models for pessimistic or optimistic decisions under different uncertain scenarios', *International Journal of Approximate Reasoning*, **105**, 305–326, (2019).

[9] Arthur Dempster, 'Upper and lower probabilities induced by a multivalued mapping', *Annals of Math. Stat.*, **38**, 325–339, (1967).

[10] Thierry Denoeux, 'Analysis of evidence-theoretic decision rules for pattern classification', *Pattern recognition*, **30**(7), 1095–1107, (1997).

[11] Thierry Denoeux, 'Decision making with belief functions: a review', *arXiv preprint arXiv:1808.05322*, (2018).

[12] Thierry Denoeux and Prakash P. Shenoy, 'An axiomatic utility theory for dempster-shafer belief functions', in *International Symposium on Imprecise Probabilities: Theories and Applications, ISIPTA 2019, 3-6 July 2019, Thagaste, Ghent, Belgium*, pp. 145–155, (2019).

[13] Tommaso Flaminio, Lluis Godo, and Hykel Hosni, 'Coherence in the aggregate: A betting method for belief functions on many-valued events', *Int. J. Approx. Reasoning*, **58**, 71–86, (2015).

[14] Paolo Ghirardato, 'Coping with ignorance: unforeseen contingencies and non-additive uncertainty', *Economic Theory*, **17**(2), 247–276, (2001).

[15] Phan Giang, 'Decision with dempster-shafer belief functions:decision under ignorance and sequential consistency', *Int. J. Approximate Reasoning*, **53**(1), 38–53, (2016).

[16] Phan Hong Giang and Prakash P. Shenoy, 'Decision making with partially consonant belief functions', in *UAI '03,*, eds., Christopher Meek and Uffe Kjærulff, pp. 272–280. Morgan Kaufmann, (2003).

[17] Itzhak Gilboa, Fabio Maccheroni, Massimo Marinacci, and David Schmeidler, 'Objective and subjective rationality in a multiple prior model', *Econometrica*, **78**(2), 755–770, (2010).

[18] Faruk Gul and Wolfgang Pesendorfer, 'Expected uncertain utility theory', *Econometrica*, **82**(1), 1–39, (2014).

[19] Joseph Y Halpern and Riccardo Pucella, 'Characterizing and reasoning about probabilistic and non-probabilistic expectation', *Journal of the ACM (JACM)*, **54**(3), 15, (2007).

[20] John Harsanyi, 'Cardinal welfare, individual ethics, and interpersonal comparison of utility', *J. of Political Economics*, **63**, 309–321, (1955).

[21] Jean-Yves Jaffray and Peter Wakker, 'Decision making with belief functions: Compatibility and incompatibility with the sure thing principle', *J. Risk and Uncertainty*, **8**, 255–271, (1994).

[22] Ramzi Ben Larbi, Sébastien Konieczny, and Pierre Marquis, 'A characterization of optimality criteria for decision making under complete ignorance', in *KR*, (2010).

[23] Wenjun Ma, Xudong Luo, and Yuncheng Jiang, 'An ambiguity aversion model for decision making under ambiguity', in *AAAI'17*, eds., Satinder P. Singh and Shaul Markovitch, pp. 614–621. AAAI Press, (2017).

[24] Eric Maskin, 'Decision-making under ignorance with implication for social choice', *Theory and Decision*, **11**, 319–337, (1979).

[25] John Milnor, 'Games against nature', in *Decision Processes*, eds., C.H. Coombs and R.L. Davis, Wiley and Chapman & Hall, New York and London, (1954).

[26] Phillipe Mongin, 'Consistent bayesian aggregation', *J. Economic Theory*, **66(2)**, 313–351, (1995).

[27] Phillipe Mongin, 'The paradox of the bayesian experts and state dependent utility theory', *J. of Math Economics*, **29(3)**, 331–361, (1998).

[28] Maria Silvia Pini, Francesca Rossi, Kristen Brent Venable, and Toby Walsh, 'Aggregating partially ordered preferences', *J. Log. Comput.*, **19**(3), 475–502, (2009).

[29] David Schmeidler, 'Subjective probability and expected utility without additivity', *Econometrica*, **57**, 571–587, (1989).

[30] Glenn Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, N.J., 1976.

[31] Glenn Shafer, 'Constructive decision theory', *Int. J. Approx. Reasoning*, **79**, 45–62, (2016).

[32] Philippe Smets, 'Practical uses of belief functions', in *UAI '99, Stockholm, Sweden, July 30 - August 1, 1999*, eds., Kathryn B. Laskey and Henri Prade, pp. 612–621. Morgan Kaufmann, (1999).

[33] Philippe Smets and Robert Kennes, 'The transferable belief model', *Artif. Intell.*, **66**(2), 191–234, (1994).

[34] Thomas M. Strat, 'Decision analysis using belief functions', *Int. J. Approx. Reasoning*, **4**(5-6), 391–417, (1990).

[35] Marie-Louise Vierø, 'Exactly what happens after the anscombe–aumann race?', *Economic Theory*, **41**(2), 175–212, (2009).

[36] Ronald R. Yager, 'Decision making under Dempster-Shafer uncertainties', in *Classic Works of the Dempster-Shafer Theory of Belief Functions*, eds., Ronald R. Yager and Liping Liu, 619–632, Springer, (2008).

[37] Chunlai Zhou, Biao Qin, and Xiaoyong Du, 'A savage-style utility theory for belief functions', in *IJCAI 2018, July 13-19, 2018, Stockholm, Sweden.*, ed., Jérôme Lang, pp. 5135–5141. ijcai.org, (2018).