# A Savage-style Utility Theory for Belief Functions

**Chunlai Zhou, Biao Qin**[*]**, Xiaoyong Du,**
Computer Science Department, Renmin University of China, Beijing, CHINA
{czhou,qinbiao,duyong}@ruc.edu.cn

## Abstract

In this paper, we provide an axiomatic justification for decision making with belief functions by studying the belief-function counterpart of Savage's Theorem where the state space is finite and the consequence set is a continuum $[l, M](l < M)$. We propose six axioms for a preference relation over acts, and then show that this axiomatization admits a definition of qualitative belief functions comparing preferences over events that guarantees the existence of a belief function on the state space. The key axioms are uniformity and an analogue of the independence axiom. The uniformity axiom is used to ensure that all acts with the same maximal and minimal consequences must be equivalent. And our independence axiom shows the existence of a utility function and implies the uniqueness of the belief function on the state space. Moreover, we prove without the independence axiom the neutrality theorem that two acts are indifferent whenever they generate the same belief functions over consequences. At the end of the paper, we compare our approach with other related decision theories for belief functions.

## 1 Introduction

The problem of decision making under uncertainty is a fundamental issue for Artificial Intelligence. In this problem the decision maker (DM) is concerned with the selection of an appropriate decision alternative, in the face of uncertainty with respect to the environment. The uncertainty manifests itself in that a different payoff is obtained for different states of nature. Characterizing the behavior of DM using *subjective expected utility* was promoted and axiomatized by Savage [Savage, 1954] following previous work in [de Finetti, 1937] and [von Neumann and Morgenstern, 1944]. The theory of subjective expected utility combines two subjective concepts: first, a personal *utility function* (a measure of preferences over outcomes or consequences), and second a *subjective probability* distribution quantifying likelihood judgements about events. Different DMs may make different decisions because

---

[*]Correspondence author.

they may have different utility functions or different beliefs about the probabilities of different outcomes. Despite its achievements in economics and AI, experiments have shown that many DMs do not behave in a manner consistent with Savage's axioms of subjective expected utility for probability functions [Ellsberg, 1961].

In many uncertain situations, the Dempster–Shafer belief structure provides a more suitable representation framework a DM may face in regards to the state of nature [Jaffray, 1989; Strat, 1990; Smets and Kennes, 1994; Ghirardato, 2001; Yager, 2008; Giang and Shenoy, 2003; Smets, 2005; Gul and Pesendorfer, 2014; Shafer, 2016; Giang, 2016; Ma *et al.*, 2017]. Here we provide an axiomatic justification of decision making with belief functions by showing the counterpart of Savage's theorem. Among those semantics for belief functions in the literature, we choose Shafer's interpretation as allocations of probability [Shafer, 1979] or Fagin and Halpern's view of subjective belief functions as *generalized probability* [Fagin and Halpern, 1991], i.e., inner probability.

Our main contribution in this paper is to propose six axioms for a preference relation over acts, which are represented by functions from the state space $S$ to the consequence set $Z$, and prove the Savage-style representation theorem (Theorem 3.12). The theorem shows that the axioms are equivalent to the existence of an *interval* utility function and a belief function (inner probability), such that decisions are being made as if to maximize the expectation of the utility relative to the belief function. Since finite domains are arguably those of most interest in AI applications, we assume that the state space is finite. The first part of our system consists of four axioms (Axioms 1-4), which are adapted directly from those for expected uncertain utility (EUU) theory in [Gul and Pesendorfer, 2014]. The most important concept in this part is called *ideal events*. They are intended to capture aspects of the uncertainty that the DM can quantify without difficulty. The Savage sure-thing principle does not hold generally for any events, but for ideal events. We show that Axiom 3 induces a preference relation on ideal events called a *qualitative probability*, and further derives a preference relation on general events called a *qualitative belief function*. Although a qualitative probability does not imply the existence of a (quantitative) probability function on the state space, a qualitative belief function guarantees the existence of a *quantitative* belief function [Wong *et al.*, 1991].

The key axioms in our theory are the last two: uniformity (Axiom 5) and independence (Axiom 6). Axiom 5 implies that all acts with the same maximal and minimal consequences must be equivalent (Lemma 3.9). It formalizes *the principle of choice under complete ignorance* derived in [Arrow and Hurwicz, 1972; Jaffray and Wakker, 1994]. By using a technique called *ideal splitting* (Def. 3.8), we *prove* the neutrality theorem (Theorem 3.10) which says that two acts are equivalent whenever they generate the same belief functions on the consequences. This axiom also applies to infinite domains. It derives a similar Axiom 3 in [Gul and Pesendorfer, 2014] (Lemma 3.7) where the diffuse sets must be infinite, and hence is more general. Since Axiom 5 (Archimedean) in [Gul and Pesendorfer, 2014] rules out the possibility of a finite state space, we replace it with our independence axiom about acts here. This independence axiom is introduced to ensure the conditions for the mixture space theorem [Herstein and Milnor, 1953] and hence guarantees the existence of a linear function on acts as well as the uniqueness of the quantitative belief function on the state space. By using ideal splitting, we separate this linear function into an interval utility function and a belief function on consequences and show the representation theorem.

## 2 Belief Functions

Let $\Omega$ be a frame of discernment and $2^\Omega$ be the powerset of $\Omega$. A *mass assignment* (or *mass function*) is a mapping $m : 2^\Omega \to [0,1]$ satisfying $\sum_{A \in 2^\Omega} m(A) = 1$. 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 *belief function* is a function $bel : 2^\Omega \to [0,1]$ satisfying the following conditions: $bel(\emptyset) = 0$, $bel(\Omega) = 1$; 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 $A_i \in 2^\Omega$ for all $i \in \{1, \cdots, n\}$. A mapping $f : 2^\Omega \to [0,1]$ is a belief function if and only if its Möbius transform is a mass assignment [Shafer, 1976]. In other words, if $m : 2^\Omega \to [0,1]$ is a mass assignment, then it determines a belief function $bel : 2^\Omega \to [0,1]$ 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 $A \in 2^\Omega$. Intuitively, for a subset (or event) $A$, $m(A)$ measures the belief that the DM commits *exactly* to $A$, not the total belief $bel(A)$ that he commits to $A$. The corresponding *plausibility function* $pl : 2^\Omega \to [0,1]$ is dual to $bel$ in the sense that $pl(A) = 1 - bel(\overline{A})$ for all $A \subseteq \Omega$. Suppose that $\Theta$ is a *finer* frame than $\Omega$. This means that the elements $\omega_1, \cdots, \omega_{|\Omega|}$ of $\Omega$ correspond to a partition $\Pi_1, \cdots, \Pi_{|\Omega|}$ of $\Theta$: a subset $\{\omega_{i_1}, \cdots \omega_{i_k}\}$ of $\Omega$ has the same meaning as the subset $\Pi_{i_1} \cup \cdots \cup \Pi_{i_k}$ of $\Theta$. This identification can represented by a mapping $\rho : 2^\Omega \to 2^\Theta$ such that $\rho(\{\omega_i\}) = \Pi_i (1 \leq i \leq |\Omega|)$ and $\rho(\{\omega_{i_1}, \cdots \omega_{i_k}\}) = \cup_{j=1}^k \rho(\omega_{i_j}) = \cup_{j=1}^k \Pi_{i_j}$. The partition $\Pi_1, \cdots, \Pi_{|\Omega|}$ of $\Theta$ as a basis defines a subalgebra $\mathbb{A}^\rho$ of $2^\Theta$ as a Boolean algebra with set operations, which is isomorphic to the set algebra $2^\Omega$. We call $\Theta$ a *refinement* of $\Omega$ and $\Omega$ a coarsening of $\Theta$. Given a belief function $bel$ over $\Theta$

with the refining mapping $\rho : 2^\Omega \to 2^\Theta$, its *marginal bel* $\restriction_\Omega$ over $\Omega$ is defined as follows: $(bel \restriction_\Omega)(\{\omega_{i_1}, \cdots, \omega_{i_k}\}) = bel(\Pi_{i_1} \cup \cdots \cup \Pi_{i_k})$.

A *probability space* $\langle S, \mathbb{A}, pr \rangle$ consists of a set $S$ (called the *sample space*), a $\sigma$-algebra $\mathbb{A}$ of subsets of $S$ and a *probability measure* $pr : \mathbb{A} \to [0,1]$. The *inner probability* $pr_*$ associated with $pr$ is defined as $pr_*(A) = sup\{pr(E) : E \subseteq A, E \in \mathbb{A}\}$ for any $A \subseteq S$. Without further notice, we assume that the space is finite in this paper.

**Proposition 2.1** *(Proposition 3.1 and Corollary 3.6 in [Fagin and Halpern, 1991])Belief functions and inner probabilities are equivalent in the following sense:*

1. *If $\langle S, \mathbb{A}, pr \rangle$ is a probability space, then $pr_*$ is a belief function on $S$;*

2. *Given a belief function $bel$ on a finite set $S$, there is a probability space $\langle S', \mathbb{A}, pr \rangle$ and a surjection $f : S' \to S$ such that for each $A \subseteq S$, $bel(A) = pr_*(f^{-1}(A))$.*

This proposition contains Shafer's interpretation of belief functions as allocations of probability [Shafer, 1979]: for each $A \subseteq S$, $\rho : 2^S \to \mathbb{A}$ defined as $\rho(A)$ being the largest element of $\mathbb{A}$ contained in $f^{-1}(A)$ is an allocation mapping. These semantics are related to but different from the multi-valued-mapping semantics and the coherent-lower-probability interpretation (See Section 6 in [Fagin and Halpern, 1991]).

## 3 Representation Theorem

In this section, we first set up the background for decision-making and propose 6 axioms for the preference relations over acts. Next we prove the Savage-style representation theorem.

### 3.1 Decision-theoretic Setup and Axioms

There is a set of items $X$, and the DM is willing to express his preferences among these items by making paired comparisons of the form: "I strictly prefer $x$ to $y$" which is written $x \succ_X y$. A binary relation on $X$ is called a *preference relation* if it is asymmetric, i.e., $x \succ_X y$ implies not $y \succ_X x$ and negatively transitive, i.e., if not $x \succ_X y$ and not $y \succ_X z$, then not $x \succ_X z$. With the strict preference relation $\succ_X$, we define two other binary relations: $x \succeq_X y$ if $y \not\succ_X x$, and $x \sim_X y$ if $x \succeq_X y$ and $y \succeq_X x$. The first is called *weak preference relation* and the second *indifference relation*. If $\succ_X$ is a preference relation on $X$, then (1) $\succ_X$ is transitive; (2) for all $x$ and $y$ in $X$, exactly one of $x \succ_X y$, $y \succ_X x$ and $x \sim_X y$ holds; (3) $\succeq_X$ is complete and transitive; (4) $\sim_X$ is reflexive, symmetric and transitive; and $x \succeq_X y$ iff $x \succ_X y$ or $x \sim_X y$. In the following, we drop the subscripts $X$ whenever the context is clear.

Just as in Savage's theorem for probability functions, uncertainty is viewed as being subjective in our Savage-style models and belief functions will be supplied by a decision maker (DM) on the basis of his subjective *preferences*. Our model includes two primitive concepts:

- a *finite* set of states of nature denoted by $S = \{s_1, \cdots, s_K\}$ with typical element denoted by $s$.

- a set of outcomes or consequences, denoted by $Z$. Following [Gul and Pesendorfer, 2014], we assume that $Z = [l, M] \subseteq \mathbb{R}$ where $l < M$. For technical reasons, we assume that $l = 0$ in this paper.

From $S$ and $Z$, we construct the choice space, which is denoted by $\mathcal{F}$, as the set of all functions from $S$ to $Z$. Formally, we would write $\mathcal{F} = Z^S = \{f : f \text{ is a function from } S \text{ to } Z\}$. Elements of $\mathcal{F}$ are called *acts*. For any property $P$, let $\{P\}$ denote the set of all states where $P$ holds. For example, $\{f > g\} = \{s \in S : f(s) > g(s)\}$. For $\{P\} = S$, we simply write $P$; that is, $f \in [x, y]$ means $\{s \in S : f(s) \in [x, y]\} = S$. We identify $z \in Z$ with the *constant* act $f = z$. In other words, $z \in Z$ is also understood as $z(s) = z$ for all $s \in S$. The DM cannot fully specify the consequences that ensue from the action chosen. Instead, his choice of action sets up a function $f$ from states of nature to the consequences-he chooses one such action from some set of available actions, and his preferences over actions are modeled by a binary relation $\succ$ $(\subseteq \mathcal{F} \times \mathcal{F})$ defined on $\mathcal{F}$. First we describe axioms and explain the basic logic behind whenever necessary.

- AXIOM 1 (Weak order): The binary relation $\succ$ is a preference relation.

- AXIOM 2 (Dominance): For any two acts $f$ and $g \in \mathcal{F}$, if $f > g$, then $f \succ g$.

For any $f, g \in \mathcal{F}$ and $A \subseteq S$, let $fAg$ denote the act that agrees with $f$ on $A$ and with $g$ on $A^c$, the complement of $A$. Formally,

$$fAg(s) = \begin{cases} f(s) & \text{if } s \in A, \\ g(s) & \text{otherwise.} \end{cases}$$

The axiom that is most characteristic of subjective probabilities is Savage's sure-thing principle [Savage, 1954]. As pointed out in [Jaffray and Wakker, 1994], the belief-function approach violates this principle. Although the sure thing principle does not hold for all events of $S$, it does for a smaller set of *ideal* events.

**Definition 3.1** An event $E$ is *ideal* if, for all acts $f, g, h$ and $h' \in \mathcal{F}$, $fEh' \succcurlyeq gEh'$ and $h'Ef \succcurlyeq h'Eg$ provided that $fEh \succcurlyeq gEh$ and $hEf \succcurlyeq hEg$. An event $A$ is *null* if $fAh \sim gAh$ for all $f, g, h \in \mathcal{F}$. ◁

Let $\mathcal{E}$ denote the set of all ideal events and $\mathcal{E}_+$ that of ideal events that are not null. So $\mathcal{E}$ is the domain where the sure thing principle holds. We wish to measure not only the ranking of outcomes, but also of the events. Specifically, we wish to find out whether DM thinks that event $A$ is more likely than event $B$. The main hypothesis is that the DM uses elements of $\mathcal{E}_+$ to quantify the uncertainty of all events. For any event $A$, $A^* = \bigcup\{E \in \mathcal{E}_+ : E \cap A \neq \emptyset\}$ and $A_* = \bigcup\{E \in \mathcal{E}_+ : E \subseteq A\}$ are called respectively *the upper and lower approximations* of $A$ w.r.t. $\mathcal{E}_+$. Axiom 3 below is Savage's comparative probability axiom (P4) applied to *ideal* events.

- AXIOM 3 (Comparative Belief): For any $x, y, w, z \in Z$ and any ideal events $E, E' \in \mathcal{E}$, if $y > x$ and $w > z$, then $wEz \succcurlyeq wE'z$ provided that $yEx \succcurlyeq yE'x$.

For any two ideal events $E$ and $E' \in \mathcal{E}$, say that $E$ is *preferred to* $E'$, if, for all $x$ and $y$ from $Z$ such that $x \succ y$, $xEy \succ xE'y$. According to Axiom 3, it doesn't matter which $x$ and $y$ are used to define the preference. Axiom 3 imposes a natural requirement that the betting preferences be independent of the specific consequences that define the bets. From Axioms 1-3, we know that the above relation is a preference relation, which is also denoted by $\succ$. Recall that $S = \{s_1, \cdots, s_K\}$. Since $Z$ is a subset of real numbers, we view each act $f \in \mathcal{F}$ as an element $(f(s_1), \cdots, f(s_K)) \in \mathbb{R}^K$ and hence $\mathcal{F}$ as a subset of $\mathbb{R}^K$. A subset $G$ of acts is called *closed* if it is a closed subset of $\mathbb{R}^K$ in the usual topology induced by the Euclidean metric.

- AXIOM 4 (Continuity): For any act $f$, both $U(f) := \{g \in \mathcal{F} : g \succcurlyeq f\}$ and $L(f) := \{g \in \mathcal{F} : f \succcurlyeq g\}$ are closed.

AXIOM 4 compensates for Savage's (P6) and is needed to ensure the continuity of the von-Neumann-Morgenstern utility index when proving the following Savage-style theorem in a setting with real-valued outcomes.

**Proposition 3.2** *(Lemmas B1 and B2 in [Gul and Pesendorfer, 2014] ) Assume that the collection $\mathcal{F}$ of acts satisfies the above four axioms Axioms 1-4. Let $f$ and $g$ be two acts in $\mathcal{F}$. Then*

1. *$f \geq g$ implies $f \succcurlyeq g$;*

2. *$f \succ g$ implies $f \succ z \succ g$ for some $z \in Z$;*

3. *if $E \in \mathcal{E}_+$ and $y > x$, then $yEh \succ xEh$ for all $h \in \mathcal{F}$;*

4. *if $E \in \mathcal{E}_+$ and $f \in \mathcal{F}$, there exists a unique $c_E(f) \in Z$ such that $c_E(f)Ef \sim f$.*

5. *The set $\mathcal{E}$ of ideal events is a $\sigma$-algebra.*

$E$ is called an *atom* if it is an indecomposable ideal event, i.e., $E \in \mathcal{E}$ and no other nonempty ideal event is contained in $E$. Let $\mathcal{E}_0$ denote the set of atoms in $\mathcal{E}$.

**Lemma 3.3** *The above defined preference relation on $\mathcal{E}$ is a* qualitative probability. *That is to say, $\succ$ is a preference relation; $E \succcurlyeq \emptyset$ for all ideal events $E \in \mathcal{E}$; $S \succ \emptyset$, and, if $E \cap (E_1 \cup E_2) = \emptyset$, then $E_1 \succ E_2$ iff $E_1 \cup E \succ E_2 \cup E$. Moreover,*

1. *If $E_1 \sim E_2, E_3 \sim E_4$ and $E_1 \cap E_3 = E_2 \cap E_4 = \emptyset$, then $E_1 \cup E_3 \sim E_2 \cup E_4$.*

2. *For any ideal events $E_1, E_2, E_3, E_4 \in \mathcal{E}$, if $E_1 \succcurlyeq E_2, E_3 \succ E_4$ and $E_1 \cap E_3 = \emptyset$, then $E_1 \cup E_3 \succ E_2 \cup E_4$.*

3. *For any ideal events $E$ and $E'$, if $E \supseteq E'$, then $S \succcurlyeq E \succcurlyeq E' \succcurlyeq \emptyset$.*

**Proof.** The propositions follow from Axioms 1-3. The proof is similar to that in Savage's theory about qualitative probabilities (refer to Proposition 9.12 (Page 133) and Proposition 8.4 (Page 119) of [Kreps, 1988]). QED

Now we extend the preference relation on ideal events to that on general events. An ideal event $E \in \mathcal{E}$ is called the *core* of event $A$ if $E \subseteq A$ and $E$ is the largest ideal event

contained in $A$. It is easy to see that, for any event $A$, its core always exists and is unique because the core is simply the union of all ideal events contained in $A$. Let $A^\circ$ denote the core. Now we express comparative beliefs in terms of cores. For any two events $A$ and $B$, $A$ is called *preferred to* $B$ (denoted as $A \succ B$) if $A^\circ$ is preferred to $B^\circ$.

**Lemma 3.4** *For any event $A, B, C \subseteq S$,*

1. *$\succ$ is a preference relation,*

2. *(dominance) $A \supseteq B$ implies $A \succeq B$,*

3. *(partial monotonicity) if $A \supseteq B$ and $A \cap C = \emptyset$, then $A \succ B$ implies $A \cup C \succ B \cup C$, and*

4. *(nontriviality) $S \succ \emptyset$.*

**Proof.** Here we mainly prove partial monotonicity (Part 3). All the other parts are straightforward. Assume that $A \supseteq B, A \succ B$ and $A \cap C = \emptyset$. It follows that $A^\circ \supseteq B^\circ, A^\circ \setminus B^\circ \neq \emptyset$ and $A^\circ \succ B^\circ$. From Lemma 3.3, we know that $A^\circ \setminus B^\circ \succ \emptyset$. Note $(B \cup C)^\circ \cap (A^\circ \setminus B^\circ) = \emptyset$. Since $A \cup C \supseteq B \cup C, (A \cup C)^\circ \supseteq (B \cup C)^\circ$. In addition, $(A \cup C)^\circ \supseteq (A^\circ \setminus B^\circ)$. So $(A \cup C)^\circ \supseteq [(B \cup C)^\circ \cup (A^\circ \setminus B^\circ)]$ and hence $(A \cup C)^\circ \succeq [(B \cup C)^\circ \cup (A^\circ \setminus B^\circ)]$ (Lemma 3.3). We already know that $(A^\circ \setminus B^\circ) \succ \emptyset$ and $(B \cup C)^\circ \cap (A^\circ \setminus B^\circ) = \emptyset$. It follows from Lemma 3.3 that $(B \cup C)^\circ \cup (A^\circ \setminus B^\circ) \succ (B \cup C)^\circ$. So we have shown that $(A \cup C)^\circ \succ (B \cup C)^\circ$ and hence $(A \cup C) \succ (B \cup C)$.                QED

A binary relation $R \subseteq S \times S$ is called a *qualitative belief function* if it satisfies the four conditions in Lemma 3.4. The following theorem follows directly from Lemma 3.4 and Theorem 4 in [Wong *et al.*, 1991].

**Theorem 3.5** *(Qualitative Belief Theorem) There exists a belief function $Bel : 2^S \to [0, 1]$ such that, for any two events $A, B \in 2^S$, $A \succ B \Leftrightarrow Bel(A) > Bel(B)$.*

Such a belief function $Bel$ is called *prior* and is not unique. The existence of such a quantitative belief function as $Bel$ is equivalent to the solvability of a group of linear (in)equalities. From the Fourier-Motzkin Lemma, we know that a solution, if it exists, is not unique and hence the quantitative belief function is not unique. We need Axiom 6 to ensure the uniqueness (Theorem 3.12). Note that the belief function $Bel \restriction_\mathcal{E}$ restricted to the $\sigma$-algebra $\mathcal{E}$ is a probability measure, which is not unique. So $Bel$ is actually an inner probability associated with the probability measure $Bel \restriction_\mathcal{E}$.

**Definition 3.6** Let $E$ be an ideal event in $\mathcal{E}$. An event $D$ is called *$E$-diffuse* if it satisfies the following conditions:

1. $D \subseteq E$;

2. $D \cap E' \neq \emptyset$ and $D^c \cap E' \neq \emptyset$ for all $E'$ such that $E' \in \mathcal{E}_+$ and $E' \subseteq E$.

Let $\mathcal{D}_E$ denote the set of all $E$-diffuse sets. Note that $\mathcal{D}_E$ may be empty.                ◁

- AXIOM 5 (Uniformity): For any $x, y \in Z, E, E' \in \mathcal{E}$, $D \in \mathcal{D}_E$ and $D' \in \mathcal{D}_{E'}$, if $(xDy)Ef \sim cEf$ and $(xD'y)E'f \sim c'E'f$, then $c = c'$.

Since $E$ and $E'$ are ideal, the constant $c$ depends only on $x$ and $y$. The upper and lower approximations of any $E$-diffuse $D$ are $E$ and $\emptyset$, respectively and hence each $E$-diffuse event represents a state of complete ignorance regarding $E$. Axiom 5 formalizes the principle of complete ignorance regarding betting on $D$.

**Lemma 3.7** *For any $x, y \in Z, D \in \mathcal{D}_E$ and $D' \in \mathcal{D}_{E'}$,*

1. *if $E \sim E'$, then $xDy \sim xD'y$ for any $x, y \in Z$.*

2. *$(xDy)Ef \sim (xD'y)E'f$.*

**Proof.** This lemma follows from Prop. 3.2 and Axiom 4. QED

Since the consequence set $Z \subseteq \mathbb{R}$, we can define convex combinations of acts. Let $f$ and $g$ be two acts in $\mathcal{F}$. For $\alpha \in [0, 1]$, the *convex combination* $\alpha f + (1 - \alpha)g$ of $f$ and $g$ is defined as: for any $s \in S, (\alpha f + (1 - \alpha)g)(s) = \alpha f(s) + (1 - \alpha)g(s) \in Z$.

- AXIOM 6 (Independence): For any $\alpha \in (0, 1]$ and any acts $f, g, h \in \mathcal{F}$, $f \succ g$ implies $\alpha f + (1 - \alpha)h \succ \alpha g + (1 - \alpha)h$.

It follows that $f_1 \sim f_2$ and $g_1 \sim g_2$ implies $\alpha f_1 + (1 - \alpha)g_1 \sim \alpha f_2 + (1 - \alpha)g_2$.

### 3.2 Representation Theorem

With the prior $Bel$, for any act $h : S \to Z$, we define an associated belief function $Bel_h$ on the range $R(h)$ of $h$. For any $A \subseteq R(h)$,

$$Bel_h(A) := Bel(h^{-1}(A)). \tag{1}$$

Since $\{h^{-1}(z) : z \in R(h)\}$ defines a partition of $S, h^{-1}$ can be viewed as refining and $Bel_h$ is the marginal of $Bel$ over $R(h)$. In other words, $Bel_h = Bel \restriction_{R(h)}$. $Bel_h$ may also be identified with the belief function $Bel$ restricted to the subalgebra generated by the sets $\{h^{-1}(z) : z \in R(h)\}$ of preimages of $h$. Now we introduce a technique to *qualitatively* represent a mass function over outcomes in terms of inner probabilities.

**Definition 3.8** (Ideal Splitting) Let $E \in \mathcal{E}, N = \{1, \cdots, n\}$ and $\{A_1, \cdots, A_n\}$ is a finite partition of $E$. The *ideal split* $\{E^J : \emptyset \neq J \subseteq N\}$ of $\{A_1, \cdots, A_n\}$ is inductively defined as follows:

- for all $i \in N, E^{\{i\}} := (A_i)^\circ$;

- for all $J \subseteq N$ such that $J \neq \emptyset$ and $|J| > 1, E^J := (\bigcup_{j \in J} A_j)^\circ \setminus \bigcup_{\emptyset \neq L \subsetneq J} E^L$.

◁

For any act $f$ with the range $\{z_1, \cdots, z_n\}(z_1 < \cdots < z_n)$, let $\{S_f^J : \emptyset \neq J \subseteq N\}$ denote the ideal split of the partition $\{A_i = f^{-1}(z_i) : 1 \leq i \leq n\}$ of $S$. Let $D_f^{(J,j)}$ denote $S_f^J \cap f^{-1}(z_j)$ where $j \in J$. Note that, for any nonempty $E \subseteq S_f^J, f(E) = \{z_i : i \in J\}$ and $D_f^{(J,j)} \in \mathcal{D}_{S_f^J}$. For any $A \subseteq S$, define

$$f \restriction_A (s) = \begin{cases} f(s) & \text{if } s \in A, \\ 0 & \text{otherwise.} \end{cases}$$

So $f = \sum_{\emptyset \neq J \subseteq N} f \upharpoonright_{S_f^J}$.

**Lemma 3.9** *Let $|J| > 1$, $max_J$ and $min_J$ be the maximal and minimal numbers in $J$. $f \upharpoonright_{S_f^J} \sim c \upharpoonright_{S_f^J}$ for some constant $c \in Z$ which depends only on $z_{max_J}$ and $z_{min_J}$.*

**Proof.** By Proposition 3.2, $f \upharpoonright_{S_f^J} \sim c \upharpoonright_{S_f^J}$ for some constant $c \in Z$. Since $z_{min_J}(\bigcup_{j \in J \setminus \{min_J\}} D_f^{(J,j)}) z_{max_J} \upharpoonright_{S_f^J} \leq f \upharpoonright_{S_f^J} \leq (z_{min_J} D_f^{(J,min_J)} z_{max_J}) \upharpoonright_{S_f^J}$, $f \upharpoonright_{S_f^J} \sim (z_{min_J} D_f^{(J,min_J)} z_{max_J}) \upharpoonright_{S_f^J} \sim c \upharpoonright_{S_f^J}$. From Lemma 3.7, we know that $c$ depends only on $z_{min_J}$ and $z_{max_J}$. QED

So we write this $c$ as $c(z_{min_J}, z_{max_J})$ to emphasize its dependency. Moreover, a similar argument on $S_f^{J \cup J'}$ shows that, for another $J'$ such that $\emptyset \neq J' \subseteq N$, if $min_J < min_{J'}$ and $max_J < max_{J'}$, then $c(z_{min_J}, z_{max_J}) < c(z_{min_{J'}}, z_{max_{J'}})$. Now we prove the following important neutrality theorem *without* Axiom 6.

**Theorem 3.10** *(Neutrality Theorem) For any two acts $f, g \in \mathcal{F}$, if $Bel_f = Bel_g$, then $f \sim g$.*

**Proof.** Without loss of generality, we assume that both ranges $R(f) = R(g) = \{z_1, z_2\}(z_1 < z_2)$ and $Bel_f = Bel_g$. This implies that $Bel(f^{-1}(z_1)) = Bel_f(z_1) = Bel_g(z_1) = Bel_g(g^{-1}(z_1))$ and $Bel(f^{-1}(z_2)) = Bel_f(z_2) = Bel_g(z_2) = Bel_g(g^{-1}(z_2))$. From Theorem 3.5, we know that $f^{-1}(z_1) \sim g^{-1}(z_1)$ and $f^{-1}(z_2) \sim g^{-1}(z_2)$. So we need to show that $f = z_1 f^{-1}(z_1) z_2 \sim z_1 g^{-1}(z_1) z_2 = g$. Note that it is not necessarily the case that $f^{-1}(z_1), g^{-1}(z_1) \in \mathcal{E}$. Let $\{S_f^{\{1\}}, S_f^{\{2\}}, S_f^{\{1,2\}}\}$ and $\{S_g^{\{1\}}, S_g^{\{2\}}, S_g^{\{1,2\}}\}$ be the ideal splits of the partitions $\{f^{-1}(z_i)\}$ and $\{g^{-1}(z_i)\}$, respectively. Without loss of generality, we assume that all these ideal events are not null. It immediately follows that $S_f^{\{1\}} \sim S_g^{\{1\}}, S_f^{\{2\}} \sim S_g^{\{2\}}$ and $S_f^{\{1,2\}} \sim S_g^{\{1,2\}}$. Note that, for any nonempty ideal event $E \subseteq S_f^{\{1,2\}}$, $f(E) = \{z_1, z_2\}$. So $D_f^1 := S_f^{\{1,2\}} \cap f^{-1}(1)$ and $D_f^2 := S_f^{\{1,2\}} \cap f^{-1}(2)$ are both $S_f^{\{1,2\}}$-diffuse. Similarly, $D_g^1 := S_g^{\{1,2\}} \cap g^{-1}(1)$ and $D_g^2 := S_g^{\{1,2\}} \cap g^{-1}(2)$ are both $S_g^{\{1,2\}}$-diffuse. It is easy to see that $g = (z_1 D_g^1 z_2) S_g^{\{1,2\}} (z_1 S_g^{\{1\}} z_2)$, and $f = (z_1 D_f^1 z_2) S_f^{\{1,2\}} (z_1 S_f^{\{1\}} z_2)$. According to Lemma 3.7, $(z_1 D_g^1 z_2) S_g^{\{1,2\}} z_2 = z_1 D_g^1 z_2 \sim z_1 D_f^1 z_2 = (z_1 D_f^1 z_2) S_f^{\{1,2\}} z_2$. It follows from Lemma 3.2 that there are constants $c_g^{\{1,2\}}$ and $c_f^{\{1,2\}}$ such that $(z_1 D_g^1 z_2) S_g^{\{1,2\}} z_2 \sim c_g^{\{1,2\}} S_g^{\{1,2\}} z_2$ and $(z_1 D_f^1 z_2) S_f^{\{1,2\}} z_2 \sim c_f^{\{1,2\}} S_f^{\{1,2\}} z_2$. Axiom 5 tells us that $c_g^{\{1,2\}} = c_f^{\{1,2\}}$. For simplicity, we use $c^{\{1,2\}}$ denote this constant. So $g \sim c^{\{1,2\}} S_g^{\{1,2\}} (z_1 S_g^{\{1\}} z_2)$ and $f \sim c^{\{1,2\}} S_f^{\{1,2\}} (z_1 S_f^{\{1\}} z_2)$. Since $S_g^{\{1,2\}} \cup S_g^{\{1\}} \sim S_f^{\{1,2\}} \cup S_f^{\{1\}}$, $z_1 S_g^{\{1,2\}} (z_1 S_g^{\{1\}} z_2) \sim z_1 S_f^{\{1,2\}} (z_1 S_f^{\{1\}} z_2)$.

Using a similar argument, we show that there is a constant $c$ such that $z_1 < c < z_2$ and $z_1(S_g^{\{1,2\}})c \sim z_1(S_g^{\{1,2\}})(z_1 S_g^{\{1\}} z_2)$, $z_1(S_f^{\{1,2\}})c \sim z_1(S_f^{\{1,2\}})(z_1 S_f^{\{1\}} z_2)$. This implies $c^{\{1,2\}}(S_g^{\{1,2\}})c \sim c^{\{1,2\}}(S_g^{\{1,2\}})(z_1 S_g^{\{1\}} z_2)$, and $c^{\{1,2\}}(S_f^{\{1,2\}})c \sim c^{\{1,2\}}(S_f^{\{1,2\}})(z_1 S_g^{\{1\}} z_2)$. So we have shown that $f \sim g$. The general case when $|R(f)| > 2$ can be proved similarly by an induction on the ideal split. QED

Let $\mathcal{B}_\mathcal{F}$ denote the set $\{Bel_f : f \in \mathcal{F}\}$ of belief functions associated with all acts. It is not necessarily true that $\alpha Bel_f + (1 - \alpha) Bel_g \in \mathcal{B}_\mathcal{F}$. In other words, $\alpha Bel_f + (1 - \alpha) Bel_g$ is not necessarily generated by an act. For any $\alpha \in [0, 1]$, define $h_\alpha : \mathcal{B}_\mathcal{F} \times \mathcal{B}_\mathcal{F} \to \mathcal{B}_\mathcal{F}$ as $h_\alpha(Bel_f, Bel_g) = Bel_{+_\alpha(f,g)}$ where $+_\alpha$ denotes the convex-combination function mapping $(f, g)$ to $\alpha f + (1 - \alpha) g$. Since $\sim$ is an equivalence relation on acts, convex combination $+_\alpha$ is well defined on the equivalence classes according to Axiom 6. Let $f_\sim$ and $g_\sim$ denote the equivalence classes including acts $f$ and $g$, respectively. $\alpha f_\sim + (1 - \alpha) g_\sim := (\alpha f + (1 - \alpha) g)_\sim$. It is easy to check that $Bel_{(\cdot)} : \mathcal{F} \to \mathcal{B}_\mathcal{F}$ defined as $Bel_{(\cdot)}(f) = Bel_f$ is a homomorphism from $\langle \mathcal{F}_\sim, (+_\alpha)_{\alpha \in [0,1]} \rangle$ to $\langle \mathcal{B}_\mathcal{F}, (h_\alpha)_{\alpha \in [0,1]} \rangle$ where $\mathcal{F}_\sim = \{f_\sim : f \in \mathcal{F}\}$. It is easy to show that $\langle \mathcal{F}_\sim, (+_\alpha)_{\alpha \in [0,1]} \rangle$ is a mixture space, and so is $\langle \mathcal{B}_\mathcal{F}, (h_\alpha)_{\alpha \in [0,1]} \rangle$[Herstein and Milnor, 1953]. For simplicity, we will omit the subscript $\sim$ for equivalence classes in the following whenever no confusion arises. For any two $Bel_f, Bel_g \in \mathcal{B}_\mathcal{F}$, $Bel_f \succ Bel_g$ if $f \succ g$. Assume that $f \succ g \succ h$. According to Lemma 3.2, $f \sim z_f, g \sim z_g, h \sim z_h$ for some constants $z_f, z_g, z_h \in Z$ such that $z_f > z_g > z_h$. There exists $\alpha \in (0, 1)$ such that $\alpha z_f + (1 - \alpha) z_h > z_g$. From Axiom 6, we know that $+_\alpha(f, h) \sim (\alpha z_f + (1 - \alpha) z_h) > z_g$. It follows that $+_\alpha(f, h) \succ g$. Similarly, we can show that there exists $\beta \in (0, 1)$ such that $+_\beta(f, h) \prec g$. From Axioms 1 and 6, we know that both $\langle \mathcal{F}_\sim, (+_\alpha)_{\alpha \in [0,1]} \rangle$ and $\langle \mathcal{B}_\mathcal{F}, (h_\alpha)_{\alpha \in [0,1]} \rangle$ satisfy the conditions for the mixture space theorem ([Herstein and Milnor, 1953] or Theorem 5.11 p. 54 in [Kreps, 1988]).

**Theorem 3.11** *(Mixture Space Theorem) For the above mixture space $\langle \mathcal{F}, (+_\alpha)_{\alpha \in [0,1]} \rangle$, there exists a function $F : \mathcal{F} \to \mathbb{R}$ such that (1) $f \succ g$ iff $F(f) > F(g)$; and (2) $F(\alpha f + (1 - \alpha) g)) = \alpha F(f) + (1 - \alpha) F(g)$. We say that $F$ represents $\succ$ on $\mathcal{F}$. Moreover, $F$ is unique up to a positive affine transformation: $F' : \mathcal{F} \to \mathbb{R}$ is a representation of $\succ$ iff $F' = aF + b$ for constants $a > 0$ and $b$.*

The above theorem has a more general form: $F(\sum_{i=1}^n a_i f_i) = \sum_{i=1}^n a_i F(f_i)$ for any acts $f_1, \cdots, f_n \in \mathcal{F}$ and nonnegative $a_1, \cdots, a_n$ such that $\sum_{i=1}^n a_i = 1$. Since $F$ is unique up to positive affine transformation, we assume that $F(0) = 0$ and $F(M) = M$ for simplicity. With this assumption, the commutativity of $F$ with convex combinations implies that $F(z) = z$ for any $z \in Z$ (this is the reason why we assume $l = 0$). Note that any constant in $Z$ is viewed as a constant act here. A function $u : \mathbb{I} = \{(l, m) : 0 \leq l < m \leq M\} \to \mathbb{R}$ is called an *interval utility* if it is continuous and strictly increasing,i.e., $u(z_1, z_2) > u(z_1', z_2')$ whenever $z_1 > z_1'$ and $z_2 > z_2'$.

**Theorem 3.12** *(Representation Theorem) If the preference relation $\succ$ on $\mathcal{F}$ satisfies Axioms 1-6, then there exists a belief function $Bel$ on the state space $S$ and an interval utility $u$ such that $f \succ g$ iff $\sum_{A \subseteq S} u(min f(A), max f(A))m_f(A) > \sum_{A \subseteq S} u(min(g(A)), max(g(A)))m_g(A)$ where $m_f$ and $m_g$ are the corresponding mass functions of $Bel_f$ and $Bel_g$ generated by $Bel$ w.r.t. $f$ and $g$ (Eq. 1), respectively. Moreover, $Bel$ is unique and $u$ is unique up to a positive affine transformation.*

**Proof.** For any $\alpha \in [0,1]$ and $f \in \mathcal{F}$, $F(\alpha f) = F(\alpha f + (1-\alpha)0) = \alpha F(f)$. Consider the ideal partition $\mathcal{E}_0 := \{E_i : i \leq i \leq n\}$ of $S$, each element of which is an atom. Since $\frac{f}{n} = \frac{f \upharpoonright_{E_i}}{n} + \cdots + \frac{f \upharpoonright_{E_n}}{n}$, $F(f) = F(f \upharpoonright_{E_i}) + \cdots + F(f \upharpoonright_{E_n})$. In particular, $M = \sum_{i=1}^{n} F(M \upharpoonright_{E_i})$. From Axiom 6, we know that $\frac{F(z \upharpoonright_{E_i})}{z} = \frac{F(z' \upharpoonright_{E_i})}{z'}$ for any $z, z' \in Z$ and hence $\frac{F(M \upharpoonright_{E_i})}{M}$ is independent of the consequence $M \in Z$. Let $Pr(E_i) := \frac{F(M \upharpoonright_{E_i})}{M}$. So $F(z \upharpoonright_{E_i}) = z Pr(E_i)$ for any $z \in Z$. Such defined $Pr$ is a probability measure on $\mathcal{E}_0$ and represents the preference relation $\succ$ on the ideal events. It is easy to show that, for any two ideal events $E$ and $E'$, $E \succ E'$ iff $M \upharpoonright_E \succ M \upharpoonright_{E'}$. This equivalence implies that the combination of Axiom 6 and those properties in Lemma 3.3 meets the necessary and sufficient conditions for existence of a unique measure strictly agreeing with a qualitative probability ordering in [Suppes and Zanotti, 1976]. So $Pr$ is the *unique* probability measure that represents the preference relation $\succ$ on ideal events and $Bel := (Pr)_*$ is also the unique belief function that is derived from qualitative belief function in Lemma 3.4 and satisfies Axiom 6. Let $R(f) = \{z_1, \cdots, z_n\}$ be the range of $f$ such that $z_1 < \cdots < z_n$ and $A_i := f^{-1}(z_i)(1 \leq i \leq n)$. It is easy to see that $\{A_i : 1 \leq i \leq n\}$ is a partition of $S$. Let $N = \{1, \cdots, n\}$ and $\{S_f^J : \emptyset \neq J \subseteq N\}$ be the ideal split (Definition 3.8) of the partition $\{A_i : 1 \leq i \leq n\}$. Recall that, for any nonempty ideal $E \subseteq S_f^J$, $f(E) = \{j : j \in J\}$. It follows from Lemma 3.9 that $F(f) = \sum_{i=1}^{k} F(f \upharpoonright_{S_f^J}) = \sum_{i=1}^{k} F(c(z_{min_J}, z_{max_J}) \upharpoonright_{S_f^J}) = \sum_{i=1}^{k} c(z_{min_J}, z_{max_J})Pr(S_f^J)$. Note that $Pr(S_f^J) = m_f(\{z_j : j \in J\})$. Define $u : \{(l,m) : 0 \leq l < m \leq M\} \to \mathbb{R}$ as $u(x,y) = c(x,y)$. From Axiom 4 and Lemma 3.9, we have that $u$ is continuous and strictly increasing. Also we obtain a version of von Neumann-Morgenstern theorem for the mixture space $\langle \mathcal{B}_{\mathcal{F}}, (h_\alpha)_{\alpha \in [0,1]} \rangle$ with $Bel = Pr_*$.

QED

It is easy to check that the converse to Representation Theorem also holds.

## 4 Related Works and Conclusion

To the best of our knowledge, we are *the first* to provide the belief-function counterpart of Savage's theorem in a classical decision model (the set-up in Section 3.1) where the state space is finite. A decision foundation for belief functions was given by Ghirardato [Ghirardato, 2001], however using a different framework from the classical decision model. He assumed that acts are correspondences, assigning *sets* of outcomes rather than one outcome, to states. With the expected uncertain utility theory from [Gul and Pesendorfer, 2014], we advance Jaffray's theory about decision making with belief functions [Jaffray, 1989; Jaffray and Wakker, 1994]. The conception and structure of our theory mainly comes from [Jaffray and Wakker, 1994]. Jaffray's model is also finite but his semantics is Dempster's model of multi-valued mappings which is different from our generalized probability semantics. In his set-up, the DM uses a two-stage approach to process the information. The first stage deals with the probabilistic information. A deviation from the Bayesian approach occurs in the second stage where the principle of complete ignorance applies. In order to achieve the von Neumann-Morgenstern theorem for belief functions, he did not show but *postulated* the "neutrality axiom" that acts are indifferent whenever they generate the same belief functions over the consequence space. Jaffray demonstrated the incompatibility of Savage's sure thing principle with belief functions and assumed a weakened version by restricting Savage's principle to *unambiguous events*. But his unambiguous events are too sparse for decision making because a belief function with all nonempty events being focal possesses *only* two *trivial* unambiguous events: $\emptyset$ and the state space $S$. This is the main reason why we choose to restrict the sure-thing principle to *ideal* events in a *larger* state space with an inner-probability interpretation (Part 2 of Proposition 2.1). The state space in [Gul and Pesendorfer, 2014] must be infinite because of the Archimedean axiom (Axiom 5 there). The development of their theory is quite different from the familiar EU theory for probability functions. They first proved Savage's theorem on ideal events and then defined the interval utility function without considering the two important theorems in the classical EU theory: qualitative belief theorem and the neutrality theorem as we do in this paper. The extensive literature about expected utilities includes many related works including [Schmeidler, 1989; Gilboa and Schmeidler, 1989; Zhang, 2002]. There is also a rich literature about decision making with belief functions (See Introduction). But none of them provides an axiomatization based on preferences over *acts* in a finite classical decision model as in this paper.

Our axiomatic system (Axiom 6) extends the language of expected (uncertain) utility theories with a *definable* convex-combination operator on acts. The extension is motivated by both our neutrality theorem and Jaffray's idea of equating acts with the belief functions that they generate over outcomes [Jaffray, 1991]. One may replace Axiom 6 here with an axiom which is similar to the state-independent preference intensity axiom in Wakker's model [Wakker, 1989] or to Assumption 3 in [Gul, 1992] *without* the convex-combination operator. We would like to achieve a Savage-style theorem for belief functions *without* any probabilistic interpretation [Smets and Kennes, 1994] which incorporates recent results on de-Finetti-style and von-Neumann-Morgenstern-style methods [Flamino et al., 2015; Coletti et al., 2015].

## Acknowledgments

# References

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

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

[de Finetti, 1937] Bruno de Finetti. La prevision: ses lois logiques, ses sources subjectives. *Annales de Institut Henri Poincare*, 7(1):1–68, 1937.

[Ellsberg, 1961] Daniel Ellsberg. Risk, ambiguity and the savage axioms. *Quarterly J. Economics*, 75:643–669, 1961.

[Fagin and Halpern, 1991] Ronald Fagin and Joseph Y. Halpern. Uncertainty, belief, and probability. *Computational Intelligence*, 7:160–173, 1991.

[Flaminio *et al.*, 2015] 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.

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

[Giang and Shenoy, 2003] Phan Hong Giang and Prakash P. Shenoy. Decision making with partially consonant belief functions. In Christopher Meek and Uffe Kjærulff, editors, *UAI '03,*, pages 272–280. Morgan Kaufmann, 2003.

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

[Gilboa and Schmeidler, 1989] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *J. Math. Economics*, 18:141–153, 1989.

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

[Gul, 1992] Faruk Gul. Savage's theorem with a finite number of states. *J. Economic Theory*, 57(1):99–110, 1992.

[Herstein and Milnor, 1953] Israel N. Herstein and John Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21:291–297, 1953.

[Jaffray and Wakker, 1994] 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.

[Jaffray, 1989] Jean-Yves Jaffray. Linear utility theory for belief functions. *Operations Research Letters*, 8:107–112, 1989.

[Jaffray, 1991] Jean-Yves Jaffray. Linear utility theory for belief functions: A discussion. In A. Chikan, editor, *Progress in decision,utility and risk theory*. Kluwer, Dordrecht, 1991.

[Kreps, 1988] David M. Kreps. *Notes on the Theory of Choice*. Westview Press/Boulder and London, 1988.

[Ma *et al.*, 2017] Wenjun Ma, Xudong Luo, and Yuncheng Jiang. An ambiguity aversion model for decision making under ambiguity. In Satinder P. Singh and Shaul Markovitch, editors, *AAAI'17*, pages 614–621. AAAI Press, 2017.

[Savage, 1954] Leonard Savage. *The Foundations of Statistics*. New York: Wiley, 1954.

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

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

[Shafer, 1979] Glenn Shafer. Allocations of probability. *Annals of probability*, 7(5):827–839, 1979.

[Shafer, 2016] Glenn Shafer. Constructive decision theory. *Int. J. Approx. Reasoning*, 79:45–62, 2016.

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

[Smets, 2005] Philippe Smets. Decision making in the tbm: the necessity of the pignistic transformation. *Int. J. Approx. Reasoning*, 38(2):133–147, 2005.

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

[Suppes and Zanotti, 1976] Patrick Suppes and Mario Zanotti. Necessary and sufficient conditions for existence of a unique measure strictly agreeing with a qualitative probability ordering. *J. Philosophical Logic*, 5(3):431–438, 1976.

[von Neumann and Morgenstern, 1944] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[Wakker, 1989] Peter Wakker. *Additive Representation of Preferences*. Kluwer Academic, Dordrecht, 1989.

[Wong *et al.*, 1991] S. K. Michael Wong, Yiyu Yao, Peter Bollmann, and H. C. Burger. Axiomatization of qualitative belief structure. *IEEE Trans. Systems, Man, and Cybernetics*, 21(4):726–734, 1991.

[Yager, 2008] Ronald R. Yager. Decision making under dempster-shafer uncertainties. In Ronald R. Yager and Liping Liu, editors, *Classic Works of the Dempster-Shafer Theory of Belief Functions*, pages 619–632. Springer, 2008.

[Zhang, 2002] Jiankang Zhang. Subjective ambiguity, expected utility and choquet expected utility. *Economic Theory*, 20(1):159–181, 2002.