An Interval-Valued Utility Theory for Decision Making with Dempster-Shafer Belief Functions

Thierry Denœux\textsuperscript{a} \hspace{1cm} Prakash P. Shenoy\textsuperscript{b}

thierry.deneoeux@utc.fr \hspace{1cm} pshenoy@ku.edu

\textsuperscript{a}Université de Technologie de Compiègne
UMR CNRS 7253 Heudiasyc,
CS 60319 - Rue Roger Couttolenc
60203 Compiègne cedex, France

\textsuperscript{b}University of Kansas School of Business
1654 Naismith Dr, Capitol Federal Hall
Lawrence, KS 66045 USA

July 4, 2019

\textsuperscript{1}Comments and suggestions for improvement are welcome.
Contents

1 Introduction .............................................. 1

2 von Neumann-Morgenstern’s Utility Theory ................. 2

3 Basic Definitions in the D-S Belief Function Theory .... 5
   3.1 Representations of belief functions .................. 5
   3.2 Basic operations in the D-S theory .................. 6
   3.3 Conditional belief functions ....................... 7

4 A Utility Theory for D-S Belief Function Theory .......... 8
   4.1 Belief function lotteries ........................... 8
   4.2 Assumptions of our framework ..................... 10
   4.3 Representation theorem ............................ 14
   4.4 A practical model for assessments of utilities .. 17

5 Comparison with Some Existing Decision Theories ........ 19
   5.1 Comparison with Jaffray’s Axiomatic Theory ....... 19
   5.2 Comparison with Smets’ Decision Theory ........... 20
   5.3 Comparison with Shafer’s Constructive Decision Theory .... 21

6 Summary and Conclusions ................................ 21
List of Tables

1. Four belief function lotteries in Example 2 ........................................... 10
2. Acts in Example 7 .............................................................................. 18
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A two-stage compound lottery reduced to an indifferent simple lottery</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>The substitutability assumption 2.5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Reducing a lottery to an indifferent compound lottery and then to an indifferent reference lottery</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Reducing a bf lottery to a bf reference lottery</td>
<td>13</td>
</tr>
</tbody>
</table>
Abstract

The main goal of this paper is to describe an axiomatic utility theory for Dempster-Shafer belief function lotteries. The axiomatic framework used is analogous to von Neumann-Morgenstern’s utility theory for probabilistic lotteries as described by Luce and Raiffa. Unlike the probabilistic case, our axiomatic framework leads to interval-valued utilities, and therefore, to a partial (incomplete) preference order on the set of all belief function lotteries. If the belief function reference lotteries we use are Bayesian belief functions, then our representation theorem coincides with Jaffray’s representation theorem for his linear utility theory for belief functions. We illustrate our framework using some examples discussed in the literature, and we propose a simple model based on an interval-valued pessimism index representing a decision-maker’s attitude to ambiguity and indeterminacy. Finally, we compare our decision theory with those proposed by Jaffray, Smets, and Shafer.

Keywords: Dempster-Shafer theory of evidence, von Neumann-Morgenstern’s utility theory, interval-valued utility function, Jaffray’s linear utility theory, Smets’ two-level decision theory, Shafer’s constructive decision theory.
1 Introduction

The main goal of this paper is to propose an axiomatic utility theory for lotteries described by belief functions in the Dempster-Shafer (D-S) theory of evidence [5, 26]. The axiomatic theory is constructed similar to von Neumann-Morgenstern’s (vN-M’s) utility theory for probabilistic lotteries [38, 20, 19, 25, 22, 14]. Unlike the probabilistic case, our axiomatic theory leads to interval-valued utilities, and therefore to a partial (incomplete) preference order on the set of all belief function lotteries. Also, we compare our decision theory to those proposed by Jaffray [21], Smets [35], and Shafer [31].

In the foreword to Glenn Shafer’s 1976 monograph [26], Dempster writes: “... I believe that Bayesian inference will always be a basic tool for practical everyday statistics, if only because questions must be answered and decisions must be taken, so that a statistician must always stand ready to upgrade his vaguer forms of belief into precisely additive probabilities.” More than 40 years after these lines were written, a lot of approaches to decision-making have been proposed (see the recent review in [17]). However, most of these methods lack a strong theoretical basis. The most important steps toward a decision theory in the D-S framework have been made by Jaffray [21], Smets [35], and Shafer [31]. However, we argue that these proposals are either not sufficiently justified from the point of view of D-S theory, or not sufficiently developed for practical use. Our goal is to propose and justify a utility theory that is in line with vN-M’s utility theory, but adapted to be used with lotteries whose uncertainty is described by D-S belief functions.

In essence, the D-S theory consists of representations—basic probability assignments (also called mass functions), belief functions, plausibility functions, etc.—together with Dempster’s combination rule, and a rule for marginalizing joint belief functions. The representation part of the D-S theory is also used in various other theories of belief functions. For example, in the imprecise probability community, a belief function is viewed as the lower envelope of a convex set of probability mass functions called a credal set. Using these semantics, it makes more sense to use the Fagin-Halpern combination rule [13] (also proposed by de Campos et al. [4]), rather than Dempster’s combination rule [18, 29, 30]. The utility theory this article proposes is designed specifically for the D-S belief function theory, and not for the other theories of belief functions. This suggests that Dempster’s combination rule should be an integral part of our theory, a property that is not satisfied in the proposals by Jaffray and Smets.

There is a large literature on decision making with a (credal) set of probability mass functions motivated by Ellsberg’s paradox [12]. An influential work in this area is the axiomatic framework by Gilboa-Schmeidler [16], where they use Choquet integration [3, 17] to compute expected utility. A belief function is a special case of a Choquet capacity. Jaffray’s [21] work can also be regarded as belonging to the same line of research, although Jaffray works directly with belief functions without specifying a combination rule. A review of this literature can be found in, e.g., [15], where the authors propose a modification of the Gilboa-Schmeidler [16] axioms. As we said earlier, our focus here is on decision-making with D-S theory of belief functions, and not on decision-making based in belief functions with a credal set interpretation. As we will see, our interval-valued utility functions lead to intervals that are contained in the Choquet lower and upper expected utility intervals.

The remainder of this article is as follows. In Section 2 we sketch vN-M’s axiomatic utility theory for probabilistic lotteries as described by Luce and Raiffa [25]. In Section 3 we summarize the basic definitions in the D-S belief function theory. In Section 4 we describe our adaptation of vN-M’s utility theory for lotteries in which uncertainty is described by D-S belief functions. Our assumptions lead to an interval-valued utility function, and consequently, to a partial (incomplete) preference order on the set of all belief function lotteries. We also describe a model for assessments
of utilities. In Section 5 we compare our utility theory with those described by Jaffray [21], Smets [35], and Shafer [31]. Finally, in Section 6 we summarize and conclude.

2 von Neumann-Morgenstern’s Utility Theory

In this section, we describe vN-M’s utility theory for decision under risk. Most of the material in this section is adapted from [25]. A decision problem can be seen as a situation in which a decision-maker (DM) has to choose a course of action (or act) in some set \( F \). An act may have different outcomes, depending on the state of nature \( X \). Exactly one state of nature will obtain, but this state is unknown. Denoting by \( \Omega_X = \{x_1, \ldots, x_n\} \) the set of states of nature and by \( O = \{O_1, \ldots, O_r\} \) the set of outcomes, an act can thus be formalized as a mapping \( f \) from \( \Omega \) to \( O \).

In this section, we assume that uncertainty about the state of nature is described by a probability mass function (PMF) \( p_X \) on \( \Omega_X \). If the DM select act \( f \), they will get outcome \( O_i \) with probability

\[
p_i = \sum_{\{x \in \Omega_X | f(x) = O_i\}} p_X(x).
\]

To each act \( f \) thus corresponds a PMF \( p = (p_1, \ldots, p_r) \) on \( O \). We call \( L = (O, p) \) a probabilistic lottery. As only one state in \( \Omega_X \) will obtain, a probabilistic lottery will result in exactly one outcome \( O_i \) (with probability \( p_i \)), and we suppose that the lottery will not be repeated. Another natural assumption is that two acts that induce the same lottery are equivalent: the problem of expressing preference between acts then boils down to expressing preference between lotteries.

We are thus concerned with a DM who has preferences on \( L \), the set of all probabilistic lotteries on \( O \), and our task is to find a real-valued utility function \( u : L \to \mathbb{R} \) such that the DM strictly prefers \( L \) to \( L' \) if and only if \( u(L) > u(L') \), and the DM is indifferent between \( L \) and \( L' \) if and only if \( u(L) = u(L') \). We write \( O_i \gg O_j \) if the DM strictly prefers \( O_i \) to \( O_j \), write \( O_i \sim O_j \) if the DM is indifferent between \( O_i \) and \( O_j \), and write \( O_i \succeq O_j \) if the DM either strictly prefers \( O_i \) to \( O_j \) or is indifferent between the two.

Of course, finding such a utility function is not always possible, unless the DM’s preferences satisfy some assumptions. We can then construct a utility function that is linear in the sense that the utility of a lottery \( L = (O, p) \) is equal to its expected utility \( \sum_{i=1}^{r} p_i \cdot u(O_i) \), where \( O_i \) is regarded as a degenerate lottery where the only possible outcome is \( O_i \) with probability 1. In the remainder of this section, we describe the assumptions that lead to the existence of such a linear utility function.

**Assumption 2.1** (Weak ordering of outcomes). For any two outcomes \( O_i \) and \( O_j \), either \( O_i \succeq O_j \) or \( O_j \succeq O_i \). Also, if \( O_i \succeq O_j \) and \( O_j \succeq O_k \), then \( O_i \succeq O_k \). Thus, the preference relation \( \succeq \) over \( O \) is a weak order, i.e., it is complete and transitive.

Given Assumption 2.1 without loss of generality, let us assume that the outcomes are labelled such that \( O_1 \succeq O_2 \succeq \cdots \succeq O_r \), and to avoid trivialities, assume that \( O_1 \gg O_r \).

Suppose that \( L = \{L^{(1)}, \ldots, L^{(s)}\} \) is a set of \( s \) lotteries, where each of the \( s \) lotteries \( L^{(j)} = (O, p^{(j)}) \) are over outcomes in \( O \), with PMFs \( p^{(j)} \) for \( j = 1, \ldots, s \). Suppose \( q = (q_1, \ldots, q_s) \) is a PMF on \( L \) such that \( q_j > 0 \) for \( j = 1, \ldots, s \), and \( \sum_{j=1}^{s} q_j = 1 \). Then \( [L, q] \) is called a compound lottery whose outcome is exactly one lottery \( L^{(i)} \) (with probability \( q_i \)), and lottery \( L^{(i)} \) will result in one outcome \( O_j \) (with probability \( p^{(j)}_j \)). Notice that the PMF \( p^{(i)} \) is a conditional PMF for \( O \) in

\footnote{The assumption of finiteness of the sets \( \Omega_X \) and \( O \) is only for ease of exposition. It is unnecessary for the proof of the representation theorem in this section.}
the second stage given that lottery \( L^{(i)} \) is realized (with probability \( q_i > 0 \)) in the first stage (see Figure 1). We can compute the joint PMF for \((L, O)\), and then compute the marginal \( p \) of the joint for \( O \). The following assumption states that the resulting lottery \([O, p] \) is indifferent to the compound lottery \([L, q] \).

**Assumption 2.2** (Reduction of compound lotteries). Any compound lottery \([L, q] \), where \( L^{(i)} = [O_i, p_i] \), is indifferent to a simple (non-compound) lottery \([O, p] \), where

\[
p_i = q_1 p_i^{(1)} + \ldots + q_s p_i^{(s)}
\]

for \( i = 1, \ldots, r \). PMF \((p_1, \ldots, p_r)\) is the marginal for \( O \) of the joint PMF of \((L, O)\).

A simple lottery involving only outcomes \( O_1 \) and \( O_r \) with PMF \((u, 1-u)\), where \( 0 \leq u \leq 1 \), is called a reference lottery, and is denoted by \([O_1, O_r, (u, 1-u)]\). Let \( O_2 \) denote the set \([O_1, O_r]\).

**Assumption 2.3** (Continuity). Each outcome \( O_i \) is indifferent to a reference lottery \( \tilde{O}_i = [O_2, (u_i, 1-u_i)] \) for some \( u_i \), where \( 0 \leq u_i \leq 1 \), i.e., \( O_i \sim \tilde{O}_i \).

**Assumption 2.4** (Weak order). The preference relation \( \succsim \) for lotteries in \( L \) is a weak order, i.e., it is complete and transitive. Assumption 2.4 generalizes Assumption 2.1 for outcomes, which can be regarded as degenerate lotteries.

**Assumption 2.5** (Substitutability). In any lottery \( L = [O, p] \), if we substitute an outcome \( O_i \) by the reference lottery \( \tilde{O}_i = [O_2, (u_i, 1-u_i)] \) that is indifferent to \( O_i \), then the result is a compound lottery that is indifferent to \( L \) (see Figure 2), i.e,

\[
[(O_1, \ldots, O_{i-1}, O_i, O_{i+1}, \ldots, O_r), p] \sim [(O_1, \ldots, O_{i-1}, \tilde{O}_i, O_{i+1}, \ldots, O_r), p].
\]

From Assumptions 2.1 and 2.5, given any lottery \( L = [O, p] \), it is possible to find a reference lottery \( \tilde{L} = [O_2, (u, 1-u)] \) that is indifferent to \( L \) (see Figure 3). This is expressed by Theorem 2.1 below.

**Theorem 2.1** (Reducing a lottery to an indifferent reference lottery). Under Assumptions 2.1, 2.5, any lottery \( L = [O, p] \) is indifferent to a reference lottery \( \tilde{L} = [O_2, (u, 1-u)] \) with

\[
u = \sum_{i=1}^{r} p_i u_i.
\]
Proof. (25) First, we replace each \( O_i \) by \( \tilde{O}_i \) for \( i = 1, \ldots, r \). Assumption 2.3 (continuity) states that these indifferent lotteries exist, and Assumption 2.5 (substitutability) says that they are substitutable without changing the preference relation. So by using Assumption 2.4 serially, \([O, p] \sim [\tilde{O}, p]\). Now if we apply Assumption 2.2 (reduction of compound lotteries), then \([O, p] \sim [O_2, (u, 1-u)]\), where \( u \) is given by Eq. (3).

Assumption 2.6 (Monotonicity). A reference lottery \( L = [O_2, (u, 1-u)] \) is preferred or indifferent to reference lottery \( L' = [O_2, (u', 1-u')] \) if and only if \( u \geq u' \).

As \( O_1 \sim \tilde{O}_1 = [O, (u_1, 1-u_1)] \) and \( O_r \sim \tilde{O}_r = [O, (u_r, 1-u_r)] \), Assumptions 2.4 and 2.6 imply that \( u_1 = 1 \) and \( u_r = 0 \). Also, from \( O_1 \succcurlyeq O_2 \succcurlyeq \cdots \succcurlyeq O_r \), we can deduce that \( 1 = u_1 \geq u_2 \geq \cdots \geq u_r = 0 \).

Assumptions 2.1–2.6 allow us to define the utility of a lottery as the probability of the best outcome \( \tilde{O}_1 \) in an indifferent reference lottery, and this utility function for lotteries on \( O \) is linear. This is stated by the following theorem.

Theorem 2.2 (25). If the preference relation \( \succcurlyeq \) on \( L \) satisfies Assumptions 2.1–2.6, then there are numbers \( u_i \) associated with outcomes \( O_i \) for \( i = 1, \ldots, r \), such that for any two lotteries \( L = [O, p] \),
and $L' = [O, p']$, $L \succ L'$ if and only if
\[
\sum_{i=1}^{r} p_i u_i \geq \sum_{i=1}^{r} p'_i u_i. \tag{4}
\]
Thus, we can define the utility of lottery $L = [O, p]$ as $u(L) = \sum_{i=1}^{r} p_i u_i$, where $u_i = u(O_i)$. Also, such a linear utility function is unique up to a strictly increasing affine transformation, i.e., if $u'_i = a u_i + b$, where $a > 0$ and $b$ are real constants, then $u(L) = \sum_{i=1}^{r} p_i u'_i$ also qualifies as a utility function.

3 Basic Definitions in the D-S Belief Function Theory

In this section, we review the basic definitions in the D-S theory of belief functions. Like various uncertainty theories, D-S belief function theory includes functional representations of uncertain knowledge, and basic operations for making inferences from such knowledge. Most of this material is taken from [24].

3.1 Representations of belief functions

Belief functions can be represented in several different ways, including as basic probability assignments, plausibility functions and belief functions. These are briefly discussed below.

Definition 1 (Basic Probability Assignment). Suppose $X$ is an unknown quantity (variable) with possible values (states) in a finite set $\Omega_X$ called the state space of $X$. We assume that $X$ takes one and only one value in $\Omega_X$, but this value is unknown. Let $2^{\Omega_X}$ denote the set of all subsets of $\Omega_X$. A basic probability assignment (BPA) $m_X$ for $X$ is a function $m_X : 2^{\Omega_X} \to [0, 1]$ such that
\[
\sum_{a \subseteq \Omega_X} m_X(a) = 1, \quad \text{and} \quad m_X(\emptyset) = 0. \tag{5}
\]

The subsets $a \subseteq \Omega_X$ such that $m_X(a) > 0$ are called focal sets of $m_X$. An example of a BPA for $X$ is the vacuous BPA for $X$, denoted by $\iota_X$, such that $\iota_X(\Omega_X) = 1$. We say that $m_X$ is deterministic if $m_X$ has a single focal set (with mass 1). Thus, the vacuous BPA for $X$ is deterministic with focal set $\Omega_X$. If all focal sets of $m_X$ are singleton subsets (of $\Omega_X$), then we say that $m_X$ is Bayesian. In this case, $m_X$ is equivalent to the PMF $P_X$ for $X$ such that $P_X(x) = m_X(\{x\})$ for each $x \in \Omega_X$.

Definition 2 (Plausibility Function). The information in a BPA $m_X$ can be represented by a corresponding plausibility function $Pl_{m_X}$ defined as follows:
\[
Pl_{m_X}(a) = \sum_{b \subseteq \Omega_X \mid b \cap a \neq \emptyset} m_X(b) \quad \text{for all} \ a \subseteq \Omega_X. \tag{6}
\]

For an example, suppose $\Omega_X = \{x, \bar{x}\}$. Then, the plausibility function $Pl_{\iota_X}$ corresponding to BPA $\iota_X$ is given by $Pl_{\iota_X}(\emptyset) = 0$, $Pl_{\iota_X}(\{x\}) = 1$, $Pl_{\iota_X}(\{x, \bar{x}\}) = 1$, and $Pl_{\iota_X}(\Omega_X) = 1$.

2Belief functions can also be mathematically represented by a convex set of PMFs called a credal set, but the semantics of such a representation are incompatible with Dempster’s combination rule [27, 29, 30, 18]. For these reasons, we skip a credal set representation of a belief function.
Definition 3 (Belief Function). The information in a BPA $m_X$ can also be represented by a corresponding belief function $Bel_{m_X}$ that is defined as follows:

$$Bel_{m_X}(a) = \sum_{\{b \subseteq \Omega_X | b \subseteq a\}} m_X(b) \text{ for all } a \subseteq \Omega_X.$$  \hspace{1cm} (7)

For the example above with $\Omega_X = \{x, \bar{x}\}$, the belief function $Bel_{\iota_X}$ corresponding to BPA $\iota_X$ is given by $Bel_{\iota_X}(\emptyset) = 0$, $Bel_{\iota_X}(\{x\}) = 0$, $Bel_{\iota_X}(\{\bar{x}\}) = 0$, and $Bel_{\iota_X}(\Omega_X) = 1$. For any proposition $a \in 2^{\Omega_X}$, it is easy to see that $Bel_{m_X}(a) \leq Pl_{m_X}(a)$. Thus, if a DM’s belief in proposition $a$ is an interval, say $[p, p + q]$, where $p, q \geq 0$ and $p + q \leq 1$, then such beliefs can be represented by a BPA $m_X$ such that $m_X(a) = p$, $m_X(\Omega_X \setminus a) = 1 - p - q$, and $m_X(\Omega_X) = q$. For such a BPA, $Bel_{m_X}(a) = p \leq p + q = Pl_{m_X}(a)$.

All three representations—BPA, belief and plausibility functions—have exactly the same information, as any one of them allows us to recover the other two \cite{26}. Next, we describe the two main operations for making inferences.

3.2 Basic operations in the D-S theory

There are two main operations in the D-S theory—Dempster’s combination rule and marginalization.

Dempster’s Combination Rule In the D-S theory, we can combine two BPAs $m_1$ and $m_2$ representing distinct pieces of evidence by Dempster’s rule \cite{5} and obtain the BPA $m_1 \oplus m_2$, which represents the combined evidence. Dempster refers to this rule as the product-intersection rule, as the product of the BPA values are assigned to the intersection of the focal sets, followed by normalization. Normalization consists of discarding the mass assigned to $\emptyset$, and normalizing the remaining values so that they add to 1. In general, Dempster’s rule of combination can be used to combine two BPAs for arbitrary sets of variables.

Let $\mathcal{X}$ denote a finite set of variables. The state space of $\mathcal{X}$ is $\times_{X \in \mathcal{X}} \Omega_X$. Thus, if $\mathcal{X} = \{X, Y\}$ then the state space of $\{X, Y\}$ is $\Omega_X \times \Omega_Y$. Projection of states simply means dropping extra coordinates; for example, if $(x, y)$ is a state of $\{X, Y\}$, then the projection of $(x, y)$ to $X$, denoted by $(x, y)^{\downarrow X}$, is simply $x$, which is a state of $X$. Projection of subsets of states is achieved by projecting every state in the subset. Suppose $b \in 2^{\Omega_{\{X, Y\}}}$. Then $b^{\downarrow X} = \{x \in \Omega_X : (x, y) \in b\}$. Notice that $b^{\downarrow X} \subseteq 2^{\Omega_X}$.

Vacuous extension of a subset of states of $\mathcal{X}_1$ to a subset of states of $\mathcal{X}_2$, where $\mathcal{X}_2 \supseteq \mathcal{X}_1$, is a cylinder set extension, i.e., if $a \in 2^{\mathcal{X}_1}$, then $a^{\uparrow \mathcal{X}_2} = a \times \Omega_{\mathcal{X}_2 \setminus \mathcal{X}_1}$. Thus, if $a \in 2^{\mathcal{X}_X}$, then $a^{\uparrow \{X, Y\}} = a \times \Omega_Y$.

Definition 4 (Dempster’s rule using BPAs). Suppose $m_1$ and $m_2$ are BPAs for $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. Then $m_1 \oplus m_2$ is a BPA for $\mathcal{X}_1 \cup \mathcal{X}_2 = \mathcal{X}$, say, given by $(m_1 \oplus m_2)(\emptyset) = 0$ and

$$(m_1 \oplus m_2)(a) = K^{-1} \sum_{\{b_1, b_2 \subseteq \Omega_X | b_1 \cap b_2 = a\}} m_1(b_1^{\downarrow \mathcal{X}_1}) m_2(b_2^{\downarrow \mathcal{X}_2}),$$  \hspace{1cm} (8)

for all $a \subseteq \Omega_X$, where $K$ is a normalization constant given by

$$K = 1 - \sum_{\{b_1, b_2 \subseteq \Omega_X | b_1 \cap b_2 = \emptyset\}} m_1(b_1^{\downarrow \mathcal{X}_1}) m_2(b_2^{\downarrow \mathcal{X}_2}).$$  \hspace{1cm} (9)
The definition of Dempster’s rule assumes that the normalization constant $K$ is non-zero. If $K = 0$, then the two BPAs $m_1$ and $m_2$ are said to be in total conflict and cannot be combined. If $K = 1$, we say $m_1$ and $m_2$ are non-conflicting.

**Marginalization**  Marginalization in D-S theory is addition of values of BPAs.

**Definition 5** (Marginalization). Suppose $m$ is a BPA for $\mathcal{X}$. Then, the marginal of $m$ for $\mathcal{X}_1$, where $\mathcal{X}_1 \subset \mathcal{X}$, denoted by $m^{\downarrow \mathcal{X}_1}$, is a BPA for $\mathcal{X}_1$ such that for each $a \subseteq \Omega_{\mathcal{X}_1}$,

$$m^{\downarrow \mathcal{X}_1}(a) = \sum_{\{b \subseteq \Omega_{\mathcal{X}} | b^{\downarrow \mathcal{X}_1} = a\}} m(b).$$

(10)

### 3.3 Conditional belief functions

In probability theory, it is common to construct joint PMFs for a set of discrete variables by using conditional probability distributions. For example, we can construct joint PMF for $(X, Y)$ by first assessing PMF $P_X$ of $X$, and conditional PMFs $P_{Y | x}$ for each $x \in \Omega_X$ such that $P_X(x) > 0$. The pointwise multiplication of $P_{Y | x}$ for all $x \in \Omega_X$ is called a CPT, and denoted by $P_{Y | X}$. Then, $P_{X, Y} = P_X \otimes P_{Y | X}$. We can construct joint BPA for $\{X, Y\}$ in a similar manner.

Suppose that there is a BPA for $Y$ expressing our belief about $Y$ if we know that $X = x$, and denote it by $m_{Y | x}$. Notice that $m_{Y | x} : 2^{\Omega_Y} \rightarrow [0, 1]$ is such that $\sum_{b \in \Omega_Y} m_{Y | x}(b) = 1$. We can embed this conditional BPA for $Y$ into a BPA for $\{X, Y\}$, which is denoted by $m_{X, Y}$, such that the following three conditions hold. First, $m_{x, Y}$ tells us nothing about $X$, i.e., $m_{x, Y}(\Omega_X) = 1$. Second, $m_{x, Y}$ tells us nothing about $Y$, i.e., $m_{X=x}^{\downarrow \mathcal{Y}}(\Omega_Y) = 1$. Third, if we combine $m_{x, Y}$ with the deterministic BPA $m_{X=x}$ for $X$ such $m_{X=x}(\{x\}) = 1$ using Dempster’s rule, and marginalize the result to $Y$ we obtain $m_{Y | x}$, i.e., $(m_{x, Y} \oplus m_{X=x})^{\downarrow \mathcal{Y}} = m_{Y | x}$. The least committed way to obtain such an embedding, called conditional embedding, was derived by Smets [33, 34] (see also [28]). It consists of taking each focal set $b \in 2^{\Omega_Y}$ of $m_{Y | x}$, and converting it to a corresponding focal set of $m_{x, Y}$ (with the same mass) as follows: $(\{x\} \times b) \cup (\{x\} \times \Omega_Y)$, where $\{x\}$ denotes the complement of $\{x\}$ in $\Omega_X$. It is easy to confirm that this method of embedding satisfies the three conditions mentioned above, and $m_{x, Y}$ is the least committed (minimally informative) BPA verifying this property.

**Example 1** (Conditional embedding). Consider discrete variables $X$ and $Y$, with $\Omega_X = \{x, \bar{x}\}$ and $\Omega_Y = \{y, \bar{y}\}$. Suppose that $m_X$ is a BPA for $X$ such that $m_X(x) > 0$. If we have a conditional BPA $m_{Y | x}$ for $Y$ given $X = x$ as follows:

$$m_{Y | x}(y) = 0.8, \quad \text{and} \quad m_{Y | x}(\Omega_Y) = 0.2,$$

(11)

then its conditional embedding into BPA $m_{x, Y}$ for $\{X, Y\}$ is

$$m_{x, Y}(\{(x, y), (\bar{x}, y), (x, \bar{y}), (\bar{x}, \bar{y})\}) = 0.8, \quad \text{and} \quad m_{x, Y}(\Omega_{\{X, Y\}}) = 0.2.$$

(12)

There are some differences with conditional probability distributions. First, in probability theory, $P_{Y | X}$ consists of all conditional distributions $P_{Y | x}$ that are well-defined, i.e., for all $x \in \Omega_X$
such that $P_X(x) > 0$. In D-S belief function theory, we do not have similar constraints. We can include only those non-vacuous conditionals $m_Y|_x$ such that $m_X(\{x\}) > 0$. Also, if we have more than one conditional BPA for $Y$, given, say for $X = x_1$, and $X = x_2$ (assuming $m_X(\{x_1\}) > 0$, and $m_X(\{x_2\}) > 0$), we embed these two conditionals for $Y$ to get BPAs $m_{x_1,Y}$ and $m_{x_2,Y}$ for $\{X,Y\}$, and then combine them using Dempster’s rule of combination to obtain one conditional BPA $m_{Y|X} = m_{x_1,Y} \oplus m_{x_2,Y}$, which corresponds to $P_{Y|X}$ in probability theory.

Second, given any joint PMF $P_{X,Y}$ for $\{X,Y\}$, we can always factor this into $P_{X,Y}^{X} = P_X$ for $X$, and $P_{Y|X}$ for $\{X,Y\}$, such that $P_{X,Y} = P_X \otimes P_{Y|X}$. This is not true in D-S belief function theory. Given a joint BPA $m_{X,Y}$ for $\{X,Y\}$, we cannot always find a BPA $m_{Y|X}$ for $\{X,Y\}$ such that $m_{X,Y} = m_{Y|X} \oplus m_{Y|X}$. However, we can always construct joint BPA $m_{X,Y}$ for $\{X,Y\}$ by first assessing $m_X$ for $X$, and assessing conditionals $m_{Y|x_i}$ for $Y$ for those $x_i$ that we have knowledge about and such that $m_X(\{x_i\}) > 0$, embed these conditionals into BPAs for $\{X,Y\}$, and combine all such BPAs to obtain the BPA $m_{Y|X}$ for $\{X,Y\}$. An implicit assumption here is that BBAs $m_{x_i,Y}$ are distinct, and it is acceptable to combine them using Dempster’s rule. We can then construct $m_{X,Y} = m_X \oplus m_{Y|X}$.

This completes our brief review of the D-S belief function theory. For further details, the reader is referred to [26] and to [8] for a recent review.

4 A Utility Theory for D-S Belief Function Theory

In this section, we describe a new utility theory for lotteries where the uncertainty is described by D-S belief functions. These lotteries, called belief function lotteries\(^3\), will be introduced in Section 4.1. We present and discuss assumptions in Section 4.2 and state a representation theorem in Section 4.3. Finally, we describe a practical model allowing us to compute the utility of a bf lottery based on two parameters in Section 4.4.

4.1 Belief function lotteries

We generalize the decision framework outline in Section 2 by assuming that uncertainty about the state of nature $X$ is described by a BPA $m_X$ for $X$. The probabilistic framework is recovered as a special case when $m_X$ is Bayesian. As before, we define an act as a mapping $f$ from $\Omega_X$ to the set $O$ of outcomes. Mapping $f$ pushes $m_X$ forward from $\Omega_X$ to $O$, transferring each mass $m_X(a)$ for $a \in 2^{\Omega_X}$ to $b = \{f(x) : x \in a\}$. The resulting BPA $m$ for $O$ is then defined as

$$m(b) = \sum_{\{a \in 2^{\Omega_X} | f(a) = b\}} m_X(a), \quad (13)$$

for all $b \subseteq O$, where $f[a]$ denotes the image of subset $a$ by $f$. Eq. (13) clearly generalizes Eq. (11). The pair $[O, m]$ will be called a belief function (bf) lottery. As before, we assume that two acts can be compared from what we believe their outcomes will be, irrespective of the evidence on which we base our beliefs. This assumption is a form of what Wakker [39] calls the principle of complete ignorance (PCI). It implies that two acts resulting in the same bf lottery are equivalent.

The problem of expressing preferences between acts becomes that of expressing preferences between bf lotteries.

Remark 1. As a consequence of the PCI, preferences between acts do not depend on the cardinality of the state space $\Omega_X$ in case of complete ignorance. For instance, assume that we define $\Omega_X =$

\(^3\)This notion was previously introduced in [7] under the name “evidential lottery.”
\{x_1, x_2\}, and we are completely ignorant of the state of nature, so that our belief state is described by the vacuous BPA \(m_X(\Omega_X) = 1\). Consider two acts \(f_1\) and \(f_2\) that yield \$100 if, respectively, \(x_1\) or \(x_2\) occurs, and \$0 otherwise. These two acts induce the same vacuous bf lottery \(m(O) = 1\) with \(O = \{\{\text{\$100, \$0}\}\}\): consequently, they are equivalent according to the PCI. Now, assume that we decide to express the states of nature with finer granularity and we refine state \(x_1\) into two states \(x_{11}\) and \(x_{12}\). Let \(\Omega_{X_1} = \{x_{11}, x_{12}, x_2\}\) denote the refined frame. We still have \(m_{X_1}(\Omega_{X_1}) = 1\) and \(m(O) = 1\), so that our preferences between acts \(f_1\) and \(f_2\) are unchanged. We note that a Bayesian DM applying Laplace’s principle of indifference (PI) would reach a different conclusion: before the refinement, the PI implies \(p_X(x_1) = p_X(x_2) = 1/2\), which results in the same probabilistic lottery \(p = (1/2, 1/2)\) on \(O = \{\{\text{\$100, \$0}\}\}\) for the two acts, but after the refinement the same principle gives us \(p_X(x_{11}) = p_X(x_{12}) = p_X(x_2) = 1/3\); this results in two different lotteries \(p_1 = (2/3, 1/3)\) for act \(f_1\) and \(p_2 = (1/3, 2/3)\) for act \(f_2\), which makes \(f_1\) strictly preferable to \(f_2\). Considering that the granularity of the state space is often partly arbitrary (as discussed by Shafer in [20]), we regard this property of invariance to refinement under complete ignorance as a valuable feature of a decision theory based on D-S belief functions.

We are thus concerned with a DM who has preferences on \(L_{bf}\), the set of all bf lotteries. Our task is to find a utility function \(u : L_{bf} \rightarrow [\mathbb{R}]\), where \([\mathbb{R}]\) denotes the set of closed real intervals, such that the \(u(L) = [u, u + w]\) is viewed as an interval-valued utility of \(L\). The interval-valued utility can be interpreted as follows: \(u + w\) are, respectively, the degrees of belief and plausibility of receiving the best outcome in a bf reference lottery equivalent to \(L\). Given two lotteries \(L\) and \(L'\), \(L\) is preferred to \(L'\) if and only if \(u \geq u'\) and \(u + w \geq u' + w'\). This leads to incomplete preferences on the set of all bf lotteries. If we assume \(w = 0\) for all bf lotteries, then we have a real-valued utility function on \(L_{bf}\), and consequently, complete preferences.

**Example 2 (Ellsberg’s Urn).** Ellsberg [12] describes a decision problem that questions the adequacy of the vN-M axiomatic framework. Suppose we have an urn with 90 balls, of which 30 are red, and the remaining 60 are either black or yellow. We draw a ball at random from the urn. Let \(X\) denote the color of the ball drawn, with \(\Omega_X = \{r, b, y\}\). Notice that the uncertainty of \(X\) can be described by a BPA \(m_X\) for \(X\) such that \(m_X(\{r\}) = 1/3\), and \(m_X(\{b, y\}) = 2/3\).

First, we are offered a choice between Lottery \(L_1\): \$100 on red, and Lottery \(L_2\): \$100 on black, i.e., in \(L_1\), you get \$100 if the ball drawn is red, and \$0 if the ball drawn is black or yellow, and in \(L_2\), you get \$100 if the ball drawn is black and \$0 if the ball drawn is red or yellow. Choice of \(L_1\) can be denoted by alternative \(f_1 : \Omega_X \rightarrow \{\{\text{\$100, \$0}\}\}\) such that \(f_1(r) = \$100\), \(f_1(b) = f_1(y) = \$0\). Similarly, choice of \(L_2\) can be denoted by alternative \(f_2 : \Omega_X \rightarrow \{\{\text{\$100, \$0}\}\}\) such that \(f_2(b) = \$100\), \(f_2(r) = f_2(y) = \$0\). \(L_1\) can be represented by the BPA \(m_1\) for \(O = \{\{\text{\$0, \$100}\}\}\) as follows: \(m_1(\{\text{\$100}\}) = 1/3\), \(m_1(\{\text{\$0}\}) = 2/3\). \(L_2\) can be represented by BPA \(m_2\) for \(O\) as follows: \(m_2(\{\text{\$0}\}) = 1/3\), \(m_2(\{\text{\$100}\}) = 2/3\). Notice that \(L_1\) and \(L_2\) are bf lotteries. Ellsberg notes that a frequent pattern of response is \(L_1\) preferred to \(L_2\).

Second, we are offered a choice between \(L_3\): \$100 on red or yellow, and \(L_4\): \$100 on black or yellow, i.e., in \(L_3\) you get \$100 if the ball drawn is red or yellow, and \$0 if the ball drawn is black, and in \(L_4\), you get \$100 if the ball drawn is black or yellow, and \$0 if the ball drawn is red. \(L_3\) can be represented by BPA \(m_3\) as follows: \(m_3(\{\text{\$100}\}) = 1/3\), and \(m_3(\{\text{\$0, \$100}\}) = 2/3\), and \(L_4\) can be represented by the BPA \(m_4\) as follows: \(m_4(\{\text{\$0}\}) = 1/3\), \(m_4(\{\text{\$100}\}) = 2/3\). \(L_3\) and \(L_4\) are also belief function lotteries. Ellsberg notes that \(L_4\) is often strictly preferred to \(L_3\). Also, the same subjects who prefer \(L_1\) to \(L_2\), prefer \(L_4\) to \(L_3\). Table I is a summary of the four bf lotteries.

Thus, if the outcomes of a lottery are based on the states of a random variable \(X\), which is described by a BPA \(m_X\) for \(X\), then we have a belief function lottery. In this example, we have only two outcomes, \$100, and \$0. \(L_1\) and \(L_4\) can also be regarded as probabilistic lotteries as the
Table 1: Four belief function lotteries in Example 2

<table>
<thead>
<tr>
<th>Lottery</th>
<th>$m_i({$100})$</th>
<th>$m_i({$0})$</th>
<th>$m_i({$100, $0})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_1$ ($$100$ on $r$)</td>
<td>$1/3$</td>
<td>$2/3$</td>
<td>$2/3$</td>
</tr>
<tr>
<td>$L_2$ ($$100$ on $b$)</td>
<td>$1/3$</td>
<td>$2/3$</td>
<td>$2/3$</td>
</tr>
<tr>
<td>$L_3$ ($$100$ on $r$ or $y$)</td>
<td>$1/3$</td>
<td>$2/3$</td>
<td>$2/3$</td>
</tr>
<tr>
<td>$L_4$ ($$100$ on $b$ or $y$)</td>
<td>$1/3$</td>
<td>$2/3$</td>
<td>$2/3$</td>
</tr>
</tbody>
</table>

corresponding BPAs are Bayesian. $L_2$ and $L_3$ have BPAs with non-singleton focal sets. Thus, these two lotteries can be considered as involving “ambiguity” as the exact distribution of the probability (of $2/3$) between outcomes $\$100$ and $\$0$ is unknown. Regardless of how the probability of $2/3$ is distributed between $b$ and $y$, the preferences of subjects violate the tenets of vN-M utility theory.

4.2 Assumptions of our framework

As in the probabilistic case, we will assume that a DM’s preferences for bf lotteries are reflexive and transitive. However, unlike the probabilistic case, we do not assume that these preferences are complete. In the probabilistic case, incomplete preferences are studied in [1], and in the case of sets of utility functions, in [10].

Our first assumption is identical to Assumption 2.1.

**Assumption 4.1 (Weak ordering of outcomes).** The DM’s preferences $\succ$ for outcomes in $O = \{O_1, \ldots, O_r\}$ are complete and transitive.

This allows us to label the outcomes such that

$$O_1 \succ O_2 \succ \cdots \succ O_r, \quad \text{and} \quad O_1 > O_r. \quad (14)$$

Let $\mathcal{L}_{bf}$ denote the set of all bf lotteries on $O = \{O_1, \ldots, O_r\}$, where the outcomes satisfy Eq. (14). As every BPA $m$ on $O$ is a bf lottery, $\mathcal{L}_{bf}$ is essentially the set of all BPAs on $O$. As the set of all BPAs include Bayesian BPAs, the set $\mathcal{L}_{bf}$ is a superset of $\mathcal{L}$, i.e., every probabilistic lottery on $O$ can be considered a bf lottery.

Consider a compound lottery $[L, m]$, where $L = \{L_1, \ldots, L_s\}$, $m$ is a BPA for $L$, and $L_j = [O, m_{j|L_j}]$ is a bf lottery on $O$, where $m_{j|L_j}$ is a conditional BPA for $O$ in the second stage given that lottery $L_j$ is realized in the first stage. Assumption 4.2 below posits that we can reduce the compound lottery to a simple bf lottery on $O$ using the D-S calculus, and that the compound lottery is equally preferred to the reduced simple lottery on $O$.

**Assumption 4.2 (Reduction of compound lotteries).** Suppose $[L, m]$ is a compound lottery as described in the previous paragraph. Then, $[L, m] \sim [O, m']$, where

$$m' = \left( m \oplus \left( \bigoplus_{j=1}^{s} m_{j|L_j} \right) \right)_{\downarrow O}, \quad (15)$$

and $m_{j|L_j}$ is a BPA for $(L, O)$ obtained from $m_{j|L_j}$ by conditional embedding, for $j = 1, \ldots, s$.

The following proposition states that Assumption 4.2 generalizes Assumption 2.2.
Proposition 1. Let \( L = \{L_1, \ldots, L_s\} \) be a set of bf lotteries, with \( L_j = [O, m_{j|L_j}] \), in which \( m_{j|L_j} \) is a Bayesian conditional BPA for \( O \) such that \( m_{j|L_j}(\{O_i\}) = p_{ij} \) and \( \sum_{j=1}^{s} p_{ij} = 1 \) for \( j = 1, \ldots, s \). Let \([L, m] \) be a compound lottery in which \( m \) is a Bayesian BPA for \( L \) such that \( m(\{L_j\}) = q_j \) for \( j = 1, \ldots, s \) with \( \sum_{j=1}^{s} q_j = 1 \). Then BPA \( m' \) defined by (15) is Bayesian and it verifies

\[
m'(\{O_i\}) = \sum_{j=1}^{s} q_j p_{ij}^{(j)}
\]

for \( i = 1, \ldots, r \).

Proof. The conditional embedding of \( m_{j|L_j} \) is given by

\[
m_{L_j,j}(\{(L_j, O_i)\} \cup \{L_j\} \times O) = p_{ij}^{(j)}, \quad i = 1, \ldots, r.
\]

Let \( m_0 = \bigoplus_{j=1}^{s} m_{L_j,j} \). It is a BPA for \( L \times O \) defined by

\[
m_0(\{(L_1, O_{i_1}), \ldots, (L_s, O_{i_s})\}) = p_{i_1}^{(1)} \cdots p_{i_s}^{(s)}
\]

for all \((i_1, \ldots, i_s) \in \{1, \ldots, r\}^s\). Combining \( m_0 \) with \( m \), we get a Bayesian BPA \( m'_0 \) on \( L \times O \) such that

\[
m'_0(\{(L_j, O_i)\}) = q_j p_{ij}^{(j)}.
\]

After marginalizing on \( O \), we finally get Eq. (16). \( \square \)

Next, we define a bf reference lottery \([O_2, m] \) as a bf lottery on \( O_2 = \{O_1, O_r\} \). A bf reference lottery has three parameters \( u = m(\{O_1\}) \), \( v = m(\{O_r\}) \), and \( w = m(O_2) \), which are all non-negative and sum to 1. The following assumption states that any deterministic bf lottery is equally preferred to some bf reference lottery.

Assumption 4.3 (Continuity). Any subset of outcomes \( a \subseteq O \) (considered as a deterministic bf lottery) is indifferent to a bf reference lottery \( \tilde{a} = [O_2, m_a] \) such that

\[
m_a(\{O_1\}) = u_a, \quad (17a)
\]

\[
m_a(\{O_r\}) = v_a, \quad \text{and} \quad (17b)
\]

\[
m_a(O_2) = w_a, \quad (17c)
\]

where \( u_a, v_a, w_a \geq 0 \), and \( u_a + v_a + w_a = 1 \). Furthermore, \( w_a = 0 \) if \( a = \{O_i\} \) is a singleton subset.

Notice that \( Bel_{m_a}(\{O_1\}) = u_a \), and \( Pl_{m_a}(\{O_1\}) = u_a + w_a = 1 - v_a \). For singleton subsets, the equivalent bf reference lottery is Bayesian: this ensures that Assumption 4.3 is a generalization of Assumption 2.3. For non-singleton subsets \( a \) of outcomes, we may have \( w_a > 0 \), i.e., the bf reference lottery may not be Bayesian. In other words, we do not assume that ambiguity can be resolved by selecting an equivalent probabilistic reference lottery.

Example 3. Consider lottery \( L_2 = [\{\$100, \$0\}, m_2] \) in Example 2, where \( m_2(\{\$0\}) = 1/3 \), and \( m_2(\{\$100, \$0\}) = 2/3 \). Suppose we wish to assess the utility of focal set \( \{\$100, \$0\} \) using a probabilistic reference lottery \([\{\$100, \$0\}, (p, 1-p)] \). A DM may have the following preferences. For any \( p \leq 0.2 \) she prefers \( \{\$100, \$0\} \) to the probabilistic reference lottery, and for any \( p \geq 0.3 \), she prefers the probabilistic reference lottery to \( \{\$100, \$0\} \). However, she is unable to give us a precise \( p \) such that \( \{\$100, \$0\} \sim [\{\$100, \$0\}, (p, 1-p)] \). For such a DM, we can assess a bf reference lottery \([\{\$100, \$0\}, m_a] \) such that \( Bel_{m_a}(\{\$100\}) = 0.2 \) and \( Pl_{m_a}(\{\$100\}) = 0.3 \), i.e., \( u_{\{\$100, \$0\}} = 0.2 \), \( v_{\{\$100, \$0\}} = 0.7 \), and \( w_{\{\$100, \$0\}} = 0.1 \).
Assumption 4.4 (Quasi-order). The preference relation $\succsim$ for bf lotteries on $L_{bf}$ is a quasi-order, i.e., it is reflexive and transitive.

In contrast with the probabilistic case (Assumption 2.4), we do not assume that $\succsim$ is complete. There are many reasons we may not wish to assume completeness. It is not descriptive of human behavior. Even from a normative point of view, it is questionable that a DM has complete preferences on all possible lotteries. The assumption of incomplete preferences is consistent with the D-S theory of belief functions where we have non-singleton focal sets. Several authors, such as Aumann [1], and Dubra et al. [10] argue why the assumption of complete preferences may not be realistic in many circumstances.

The substitutability assumption is similar to the probabilistic case (Assumption 2.5)– we replace an outcome in the probabilistic case by a focal set of $m$ in the bf case.

Assumption 4.5 (Substitutability). In any bf lottery $L = [O, m]$, if we substitute a focal set $a$ of $m$ by an equally preferred bf reference lottery $\tilde{a} = [O_2, m_\tilde{a}]$, then the result is a compound lottery that is equally preferred to $L$.

It follows from Assumptions 4.1–4.5 that given any bf lottery, we can reduce it to an equally preferred bf reference lottery. This is stated as Theorem 4.1 below.

Theorem 4.1 (Reducing a bf lottery to an indifferent bf reference lottery). Under Assumptions 4.1–4.5 any bf lottery $L = [O, m]$ with focal sets $a_1, \ldots, a_k$ is indifferent to a bf reference lottery $\tilde{L} = [O_2, \tilde{m}]$, such that

$$
\tilde{m}({\{O_1}\}) = \sum_{i=1}^{k} m(a_i) u_{a_i}, \quad (18a)
$$

$$
\tilde{m}({\{O_r}\}) = \sum_{i=1}^{k} m(a_i) v_{a_i}, \quad \text{and} \quad (18b)
$$

$$
\tilde{m}(O_2) = \sum_{i=1}^{k} m(a_i) w_{a_i}, \quad (18c)
$$

where $u_{a_i}$, $v_{a_i}$ and $w_{a_i}$ are the masses assigned, respectively, to $\{O_1\}$, $\{O_r\}$ and $O_2$ by the bf reference lottery $\tilde{a}_i$ equivalent to $a_i$.

Proof. From Assumption 4.3 (continuity), we can replace each focal set $a_i$ of $m$ one at a time by an indifferent bf reference lottery $\tilde{a}_i = [O_2, m_{i\tilde{a}_i}]$, yielding a sequence of compound lotteries. From Assumptions 4.5 (substitutability) and 4.4 (quasi-order), these compound lotteries are all indifferent to $L$ (see Figure 4). Let $L' = [L, m']$ be the compound lottery obtained after all focal sets $a_i$ has been substituted, with $m'({\{a_i\}}) = m(a_i)$. From Assumption 4.2 (reduction of compound lotteries), $L'$ can be reduced to a reference bf lottery $\tilde{L} = [O_2, \tilde{m}]$, by considering each BBA $m_{i\tilde{a}_i}$ as a conditional BPA and applying the rules of D-S calculus. The reduced bf lottery $\tilde{L} = [O_2, \tilde{m}]$ is then given by

$$
\tilde{m} = \left( m' \oplus \left( \bigoplus_{i=1}^{k} m_{i\tilde{a}_i} \right) \downarrow O_2 \right)
$$

where $m_{i\tilde{a}_i}$ is the BPA for $O_2$ obtained from $m_{i\tilde{a}_i}$ by conditional embedding. If we have $m_{i\tilde{a}_i}({\{O_1}\}) = u_{a_i}$, $m_{i\tilde{a}_i}({\{O_r}\}) = v_{a_i}$ and $m_{i\tilde{a}_i}(O_2) = w_{a_i}$, then after conditional embedding BPA
Similarly, assigned to focal sets of the type \[ \{ \tilde{a}_i \} \] respectively, with mass \( m \)

Let \( m \) be as follows:

\[
\begin{align*}
    m(a_i) = & \{ (a_i, O_1) \cup (a_i, O_2) \} = u_i, \\
    m(a_i) = & \{ (a_i, O_r) \cup (a_i, O_2) \} = v_i, \\
    m(a_i) = & \{ L \times O_2 \} = w_i.
\end{align*}
\]

Next, we combine Bayesian BPA \( m' \) for \( L \) with \( m_0 \). The focal sets of \( m' \oplus m_0 \) are of the form \( \{ \tilde{a}_i \} \times \{ O_1 \}, \{ \tilde{a}_i \} \times \{ O_r \}, \) and \( \{ \tilde{a}_i \} \times O_2 \) depending on whether \( i \in I_1 \) or \( i \in I_2 \), or \( i \in I_3 \), respectively, with mass \( m(a_i)m_0(a_{i1,2,3}) \). Finally we marginalize \( m' \oplus m_0 \) to \( O_2 \). The total mass assigned to focal sets of the type \( \{ \tilde{a}_i \} \times \{ O_1 \} \) is \( \sum_{i=1}^{k} m(a_i)u_i \). Thus, \( \tilde{m} \{ O_1 \} = \sum_{i=1}^{k} m(a_i)u_i \). Similarly, \( \tilde{m} \{ O_r \} = \sum_{i=1}^{k} m(a_i)v_i \), and \( \tilde{m} \{ O_2 \} = \sum_{i=1}^{k} m(a_i)w_i \).

Next, we formulate the monotonicity assumption. This is less obvious than it was in the probabilistic case (Assumption 2.6), as there are several ways in which intervals may be ordered. Assumption 4.6 below states that, given two Bf reference lotteries \( L \) and \( L' \), the former will be preferred if and only if it assigns a higher degree of belief to the best consequence \( O_1 \), and a lower degree of belief to the worst consequence \( O_r \).

Assumption 4.6 (Monotonicity). Suppose \( L = [O_2, m] \) and \( L' = [O_2, m'] \) are Bf reference lotteries, with \( m \{ O_1 \} = u \), \( m \{ O \} = w \), \( m' \{ O_1 \} = u' \), \( m' \{ O \} = w' \). Then, \( L \succeq L' \) if and only if \( u \geq u' \) and \( u + w \geq u' + w' \).

Figure 4: Reducing a Bf lottery to a Bf reference lottery
Thus, \( L \succeq L' \) if and only if \( \text{Bel}_m(\{O_1\}) \geq \text{Bel}_m(\{O_1\}) \) and \( \text{Bel}_m(\{O_r\}) \leq \text{Bel}_m(\{O_r\}) \), i.e., if and only if outcome \( O_1 \) is deemed both more credible and more plausible under \( L \) than it is if under \( L' \). The corresponding indifference relation is \( L \sim L' \) if and only if \( u = u' \) and \( w = w' \). It is clear that \( \succeq \) as defined in Assumption 4.6 is reflexive and transitive. Also, the preference relation \( \succeq \) on the set of all bf reference lotteries is obviously incomplete. Thus, two lotteries are incomparable if not \( L \succeq L' \) and not \( L' \succeq L \), i.e., if one of the intervals \([u, u + w]\) and \([u', u' + w']\) is strictly included in the other.

Assumptions 4.1, 4.3 and 4.6 imply the following consistency constraints between the reference bf lotteries equivalent to single outcomes:

\[
1 = u_{\{O_1\}} \geq u_{\{O_2\}} \geq \ldots \geq u_{\{O_r\}} = 0. \tag{19}
\]

### 4.3 Representation theorem

**Theorem 4.2** (Interval-valued utility function). Suppose \( L = [O, m] \) and \( L' = [O, m'] \) are bf lotteries on \( O \). If the preference relation \( \succeq \) on \( \mathcal{L}_{bf} \) satisfies Assumptions 4.1, 4.3 and 4.6, then there are intervals \([u_{a_i}, u_{a_i} + w_{a_i}]\) associated with nonempty subsets \( a_i \subseteq O \) such that \( L \succeq L' \) if and only if

\[
\sum_{a_i \in 2^O} m(a_i) u_{a_i} \geq \sum_{a_i \in 2^O} m'(a_i) u_{a_i} \tag{20a}
\]

and

\[
\sum_{a_i \in 2^O} m(a_i) (u_{a_i} + w_{a_i}) \geq \sum_{a_i \in 2^O} m'(a_i) (u_{a_i} + w_{a_i}). \tag{20b}
\]

Thus, for a bf lottery \( L = [O, m] \), we can define

\[
[u](L) = [u, u + w] \tag{21}
\]

as an interval-valued utility of \( L \), with

\[
u = \sum_{a_i \in 2^O} m(a_i) u_{a_i} \quad \text{and} \quad w = \sum_{a_i \in 2^O} m(a_i) w_{a_i}. \tag{22}
\]

Also, such a utility function is unique up to a strictly increasing affine transformation, i.e., if \( u' = au + b \), and \( w' = aw + b \), where \( a > 0 \), and \( b \) are real constants, then

\[
[u'](L) = [u', u' + w']
\]

also qualifies as an interval-valued utility function.

**Proof.** The proof is immediate from Theorem 4.1 and Assumption 4.6 (monotonicity). \( \square \)

A special case of Theorem 4.2 is if we use Bayesian bf reference lotteries for the continuity assumption, i.e., \( w_{a} = 0 \) for all focal sets \( a \) of \( m \). In this case, Theorem 4.2 implies Corollary 4.1 below where we have a real-valued utility function and a complete ordering on \( \mathcal{L}_{bf} \).

**Corollary 4.1** (Real-valued utility function). Suppose \( L = [O, m] \) and \( L' = [O, m'] \) are bf lotteries on \( O \). If the preference relation \( \succeq \) on \( \mathcal{L}_{bf} \) satisfies Assumptions 4.1, 4.3 and 4.6, and if \( w_{a} = 0 \) for all focal sets \( a \) of \( m \) and \( m' \), then there are numbers \( u_{a} \) associated with nonempty subsets \( a \subseteq O \) such that \( L_1 \succeq L_2 \) if and only if

\[
\sum_{a_i \in 2^O} m(a_i) u_{a} \geq \sum_{a_i \in 2^O} m'(a) u_{a}. \]
is a BPA for \( X \) variable \( X \) where \((y)\). For such a DM, we thus have \( u = \sum_{a \in 2^O} m(a) u_a \)

as the utility of \( L \). Also, such a utility function is unique up to a strictly increasing affine transformation, i.e., if \( u'_a = a u_a + b \), where \( a > 0 \), and \( b \) are real constants, then

\[
u'(L) = \sum_{a \in 2^O} m(a) u'_a
\]

also qualifies as a utility function.

The utility function in Eq. (23) has exactly the same form as Jaffray’s linear utility [21]. This is discussed further in Section 5.1.

Next, we illustrate the application of Theorem 4.2 to some examples: Ellsberg’s urn problem described in Example 2, the one red ball problem described in [23], and the 1,000 balls urns described in [2].

**Example 4** (Ellsberg’s urn). Consider the four bfi lotteries described in Example 2 (also in Table 1). Given a vacuous bfi lottery \([\{100, 0\}, m(\{100, 0\}) = 1]\), what is an indifferent bfi reference lottery? For an ambiguity-averse DM,

\[
([100, 0], [1/2, 1/2]) \succ ([100, 0], m(\{100, 0\}) = 1).
\]

For such a DM, we thus have \( u_{\{100, 0\}} + w_{\{100, 0\}} < 1/2 \).

For the first choice problem between \( L_1 \) (\$100 on \( r \)) and \( L_2 \) (\$100 on \( b \)), using Eq. (23), \( [u](L_1) = [1/3, 1/3] \), and

\[
[u](L_2) = \frac{2}{3} [u_{\{100, 0\}}, u_{\{100, 0\}} + w_{\{100, 0\}}].
\]

Thus, an ambiguity-averse DM would choose \( L_1 \). This result is valid as long as \( u_{\{100, 0\}} + w_{\{100, 0\}} < 1/2 \) and is consistent with Ellsberg’s findings. For the second choice problem between \( L_3 \) (\$100 on \( r \) or \( y \)), and \( L_4 \) (\$100 on \( b \) or \( y \)),

\[
[u](L_3) = \frac{1}{3} (1) + \frac{2}{3} [u_{\{100, 0\}}, u_{\{100, 0\}} + w_{\{100, 0\}}],
\]

and \( [u](L_4) = [2/3, 2/3] \). An ambiguity-averse DM would choose \( L_4 \), as

\[
\frac{1}{3} + \frac{2}{3} u_{\{100, 0\}} + w_{\{100, 0\}} < \frac{2}{3}
\]

as long as \( u_{\{100, 0\}} + w_{\{100, 0\}} < 1/2 \), a result that is also consistent with Ellsberg’s empirical findings.

**Example 5** (One red ball). Consider the following example called ‘one red ball’ in [23]. An urn possibly contains balls of six colors: red (\( r \)), blue (\( b \)), green (\( g \)), orange (\( o \)), white (\( w \)), and yellow (\( y \)). One ball is drawn at random from the urn. We are informed that the urn has a total of \( n \) balls, where \( n \) is a positive integer, and that there is exactly one red ball in the urn. Suppose random variable \( X \) denotes the color of the ball drawn from the urn. Then \( \Omega_X = \{r, b, g, o, w, y\} \), and \( m_X \) is a BPA for \( X \) such that \( m_X(\{r\}) = 1/n \), and \( m_X(\{b, g, o, w, y\}) = (n - 1)/n \). First, you pick
a color, and then you draw a ball at random from the urn. You win $100 if the color of the ball
drawn from the urn matches the color you picked and you win $0 if it doesn’t. What color do you
pick? In [23], the authors describe some informal experiments where all respondents chose red for
n ≤ 7, and for n ≥ 8, several respondents preferred a color different from red.

Suppose you pick r. The bf lottery $L_r$ based on $m_X$ is as follows: \([\{\$100, \$0\}, m_r]\), where
$m_r(\{\$100\}) = 1/n$, and $m_r(\{\$0\}) = (n−1)/n$. If the color you pick is b, then the bf lottery
$L_b$ is \([\{\$100, \$0\}, m_b]\), where $m_b(\{\$0\}) = 1/n$, and $m_b(\{\$100, \$0\}) = (n−1)/n$. Thus, we have
$[u](L_r) = [1/n, 1/n]$, and

$$[u](L_b) = \frac{n−1}{n} [u(\$100, \$0), u(\$100, \$0) + w(\$100, \$0)].$$

So, $L_b$ is strictly preferred to $L_r$ whenever

$$\frac{n−1}{n} u(\$100, \$0) > \frac{1}{n},$$
i.e., whenever $u(\$100, \$0) > 1/(n−1)$, and $L_r$ is strictly preferred to $L_b$ whenever

$$\frac{n−1}{n} (u(\$100, \$0) + w(\$100, \$0)) < \frac{1}{n},$$
i.e., whenever $u(\$100, \$0) + w(\$100, \$0) < 1/(n−1)$. Hence, $L_b$ is increasingly preferred to $L_r$ when $n$
arbitrarily. As the experiment reported in [23] did not allow the respondents to express inability to
choose between the two lotteries, it does not provide any evidence for or against our model.

Example 6 (Urn with 1,000 balls). The following example is discussed in [2], where it is credited
to Ellsberg in an oral conversation (with the authors of [2]). It is also discussed in [17]. There
are two urns, each with 1,000 balls, numbered from 1−1,000. Urn 1 has exactly one ball for each
number, and there is no ambiguity. Urn 2 has unknown number of balls of each number, and there
is much ambiguity. One ball is to be chosen at random from an urn of your choosing. If the number
on the ball matches a specific number, e.g., 687, you win $100, and if not, you win nothing ($0).
Which one of the two urns will you choose? It is reported in [2] that many respondents chose Urn
2. Why? Urn 1 has only one ball numbered 687, and therefore, the probability of winning $100
if the choice is Urn 1 is very small, 0.001. Urn 2 could possibly have anywhere from 0 to 1,000
balls numbered 687. Thus, the choice of Urn 2, although ambiguous, is appealing. Let’s analyze this
problem using Theorem 4.2.

Let $X_1$ denote the number on the ball chosen Urn 1, and let $X_2$ denote the number on the
ball chosen Urn 2. $\Omega_{X_1} = \Omega_{X_2} = \{1, \ldots, 1000\}$. Function $m_{X_1}$ is a BPA for $X_1$ as follows:
$m_{X_1}(\{1\}) = \ldots = m_{X_1}(\{1000\}) = 0.001$. BPA $m_{X_2}$ is vacuous, i.e., $m_{X_2}(\Omega_{X_2}) = 1$.

Lottery $L_1$ corresponding to choice of Urn 1 (say, alternative $f_1$) is \([\{\$100, \$0\}, m_1]\), where
$m_1$ is a BPA for \{\$100, \$0\} such that $m_1(\{\$100\}) = 0.001$, and $m_1(\{\$0\}) = 0.999$. $L_1$ is a bf
reference lottery, and thus, $[u](L_1) = [0.001, 0.001]$. Lottery $L_2$ corresponding to choice of Urn 2
(say, alternative $f_2$) is \([\{\$100, \$0\}, m_2]\), where $m_2$ is a vacuous BPA for \{\$100, \$0\}. The utility of
$L_2$ is

$$[u](L_2) = [u(\$100, \$0), u(\$100, \$0) + w(\$100, \$0)].$$
Consequently, $L_2$ is preferred to $L_1$ as long as

$$u_{\{100, 0\}} \geq 0.001,$$

a condition that is easily satisfied. This may explain why many DMs prefer to be ambiguity-seeking in this context, i.e., prefer $L_2$ to $L_1$.

### 4.4 A practical model for assessments of utilities

Whereas Theorem 4.2 guarantees the existence of an interval valued utility function for bf lotteries, there remains the problem of practical elicitation of utilities. The number of utilities to be elicited for each bf lottery equals twice the number of its focal sets, which can be as high as $2|\Omega| - 1$ in the worst case. By making one more reasonable assumption, we can drastically limit the number of parameters to be elicited.

Assumption 4.7 below has no counterpart in the vN-M theory, but it is rooted in decision-making under ignorance [21, 39]. For any nonempty subset of consequences $a \subseteq O$, let $Q_a$ and $\overline{Q}_a$ denote, respectively, the worst and the best outcome in $a$. To simplify the notations, we assimilate a deterministic BPA with its focal set, and we write $a \succeq b$ to mean that the deterministic lottery with focal set $a$ is preferred or indifferent to the deterministic lottery with focal set $b$. Then our last assumption can be stated as follows.

**Assumption 4.7 (Dominance).** For all nonempty subsets $a$ and $b$ of $O$, if $Q_a \supseteq Q_b$ and $\overline{Q}_a \supseteq \overline{Q}_b$, then $a \succeq b$.

Assumption 4.7 implies that the preference between two deterministic lotteries with focal sets $a$ and $b$ is determined only by the best and worst outcomes in $a$ and $b$. In particular, when $Q_a = Q_b$ and $\overline{Q}_a = \overline{Q}_b$, then $a \sim b$. Although counterintuitive at first glance, this assumption cannot be avoided if we accept the PCI, as shown by the following example.

**Example 7.** Assume that the set of outcomes is $O = \{\$0, \$1, \$999, \$1,000\}$. According to Assumption 3.7, a DM would be indifferent between receiving one of the prizes in $a = \{\$0, \$1, \$1,000\}$, without any further information, and receiving one of the prizes in $b = \{\$0, \$999, \$1,000\}$. It may be argued that most DMs would strictly prefer $b$ to $a$. Yet, assume that the state space is $\Omega_X = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, we are in a state of complete ignorance, i.e., $m_X(\Omega_X) = 1$, and the deterministic lotteries $a$ and $b$ are generated by the acts $f_1$ and $f_2$ shown in Table 2. It is clear that $f_1$ dominates $f_2$ (it yields at least as desirable consequences for all states of nature, and strictly preferred consequences for some states of nature), so it would be paradoxical to strictly prefer $f_2$ over $f_1$, i.e., to strictly prefer $b$ over $a$. But $a$ and $b$ might also have been generated by acts $f_1'$ and $f_2'$ in Table 2 and, as $f_2'$ dominates $f_1'$, it would also be paradoxical to strictly prefer $f_1'$ over $f_2'$, i.e., to strictly prefer $a$ over $b$. Consequently, indifference between $a$ and $b$ seems to be the only rational option in this case.

Jaffray [21] shows that, whenever $Q_a \supseteq Q_b$ and $\overline{Q}_a \supseteq \overline{Q}_b$, we can always construct a state space $\Omega_X$ and two acts $f_1$ and $f_2$ such that $f_1(\Omega_X) = a$, $f_2(\Omega_X) = b$ and, for any $\omega \in \Omega_X$, $f_1(\omega) \succeq f_2(\omega)$. As $f_1$ yields at least as desirable outcomes as $f_2$ under any state of nature, it should be preferred whatever our beliefs on $\Omega_X$, and in particular when $m_X(\Omega_X) = 1$. Hence, we should have $a \succeq b$.

Assumption 4.7 implies that $a \succeq Q_a$ and $\overline{Q}_a \succeq a$. From Assumption 4.6, we thus have $u_a \geq u_{Q_a}$ and $u_a + w_a \leq u_{\overline{Q}_a}$. Consequently, the utility bounds $u_a$ and $u_a + w_a$ of subset $a$ can be written as convex combinations of the utilities of its worst and best outcomes:

$$u_a = \alpha(Q_a, \overline{Q}_a) u_{\{Q_a\}} + \left(1 - \alpha(Q_a, \overline{Q}_a)\right) u_{\{\overline{Q}_a\}},$$

$$u_a + w_a = \beta(Q_a, \overline{Q}_a) u_{\{Q_a\}} + \left(1 - \beta(Q_a, \overline{Q}_a)\right) u_{\{\overline{Q}_a\}},$$

(24a)

(24b)
Table 2: Acts in Example 7

<table>
<thead>
<tr>
<th></th>
<th>(\omega_1)</th>
<th>(\omega_2)</th>
<th>(\omega_3)</th>
<th>(\omega_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_1)</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>(f_2)</td>
<td>0</td>
<td>0</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td>(f'_1)</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1000</td>
</tr>
<tr>
<td>(f'_2)</td>
<td>0</td>
<td>999</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

where \(\alpha(O_a, \overline{O}_a)\) and \(\beta(O_a, \overline{O}_a)\) are two coefficients depending only on the best and worst outcomes in \(a\), such that \(0 \leq \alpha(O_a, \overline{O}_a) \leq \beta(O_a, \overline{O}_a) \leq 1\). In Jaffray’s framework [21], \(w_a = 0\) (see Section 5.1) and \(\alpha(O_a, \overline{O}_a)\) is called a local pessimism index. In our framework, we can see the interval \([\alpha(O_a, \overline{O}_a), \beta(O_a, \overline{O}_a)]\) as an interval-valued local pessimism index reflecting both the DM’s attitude to ambiguity and indeterminacy.

Strat [37] proposes independently, but without any axiomatic justification, a similar criterion but with a constant parameter \(\alpha(O_a, \overline{O}_a) = \beta(O_a, \overline{O}_a) = \alpha\) that does not depend on the subset \(a\). Similarly, we can assume that the lower and upper pessimism indices take on constant values: \(\alpha(O_a, \overline{O}_a) = \alpha\) and \(\beta(O_a, \overline{O}_a) = \beta\), with \(0 \leq \alpha \leq \beta \leq 1\). This model allows us to recover some existing decision criteria as special cases:

- When \(\alpha = \beta\), the utility interval \([u](L)\) is reduced to a point \(u(L)\) and we get the real-valued utility (23) in Corollary 4.1.
- In particular, when \(\alpha = \beta = 0\), then

\[
u(L) = \sum_{a \subseteq O} m(a) \min_{O \in a} u_O,
\]

which is the lower expected utility \(\underline{u}_m\) with respect to \(m\) [5, 27, 6]. As shown by Gilboa and Schmeidler [17], \(\underline{u}_m\) is also the Choquet expected utility \(\underline{u}_m\) with respect to the belief function \(Bel_m\) corresponding to \(m\). The preference relation between bf lotteries then corresponds to the maximin criterion, which reflects a pessimistic attitude of the DM.

- Similarly, when \(\alpha = \beta = 1\), we get

\[
u(L) = \sum_{a \subseteq O} m(a) \max_{O \in a} u_O,
\]

which is the upper expected utility \(\overline{u}_m\), or the Choquet expected utility with respect to the plausibility function \(Pl_m\) corresponding to \(m\) [17]. The corresponding decision strategy corresponds to the maximax criterion, which models an optimistic attitude of the DM.

- When \(\alpha = 0\) and \(\beta = 1\), then the interval-valued utility is equal to the lower-upper expected utility interval

\([u](L) = [\underline{u}_m, \overline{u}_m]\).

The corresponding preference relation is then the interval bound dominance relation [9, 7], defined by

\(L \succ L' \Leftrightarrow (\underline{u}_m \geq \underline{u}_{m'} \text{ and } \overline{u}_m \geq \overline{u}_{m'})\). (27)
In the general case, we have
\[ u_m \leq u \leq u + w \leq u_m, \]  
(28)
where \( u \) and \( w \) are as in Eq. (22). Thus, the interval-valued utility \( [u](L) \) of lottery \([O, m]\) as defined in Theorem 4.2 is always included in the lower-upper expected utility interval, and the preference relation induced by our interval-valued utilities compares more lotteries than the interval dominance relation (27). The lower and upper expectations defined by Eqs. (25)-(26) can thus be seen as lower and upper bounds of the interval utility of a lottery \( L = [O, m] \) and could be used as conservative estimates if the equivalent reference lotteries \( \tilde{a}_i \) cannot be elicited.

5 Comparison with Some Existing Decision Theories

In this section, we compare our utility theory to Jaffray’s linear utility theory [21], Smets’ two-level decision theory [35], and Shafer’s constructive decision theory [31].

5.1 Comparison with Jaffray’s Axiomatic Theory

Jaffray’s axiomatic theory is based on considering the set of all belief functions on \( O \) as a mixture set as follows. Suppose \( m_1 \) and \( m_2 \) are BPAs for \( O \), and suppose \( \lambda \in (0, 1) \). Then \( m \) defined as:
\[ m(a) = \lambda m_1(a) + (1 - \lambda)m_2(a) \]  
(29)
for all \( a \in 2^O \), is a BPA for \( O \). BPA \( m \) can be written as \( m = \lambda m_1 + (1 - \lambda)m_2 \), and called a mixture of \( m_1 \) and \( m_2 \). Using the Jensen-version [22] of vN-M axiom system, Jaffray uses the following axioms, all of which are expressed using mixture BPA functions:

**Assumption 5.1** (Completeness and transitivity). The relation \( \succsim \) is complete and transitive over \( L_{bf} \).

**Assumption 5.2** (Independence). For all \( L_1 = [O, m_1] \) and \( L_2 = [O, m_2] \) in \( L_{bf} \), and \( \lambda \in (0, 1) \), \( L_1 \succ L_2 \) implies \([O, \lambda m_1 + (1 - \lambda)m_2] \rangle [O, \lambda m_2 + (1 - \lambda)m_1] \).

**Assumption 5.3** (Continuity). For all \( L_1 = [O, m_1] \), \( L_2 = [O, m_2] \), and \( L_3 = [O, m_3] \) in \( L_{bf} \) such that \( L_1 \succ L_2 \succ L_3 \), there exists \( \lambda \) and \( \mu \) in \( (0, 1) \) such that
\[ [O, \lambda m_1 + (1 - \lambda)m_3] \succ [O, m_2] \succ [O, \mu m_1 + (1 - \mu)m_3]. \]

**Theorem 5.1** (Jaffray’s representation theorem [21]). The preference relation \( \succsim \) on \( L_{bf} \) satisfies Assumptions 5.1, 5.2, and 5.3 if and only if there exists a utility function \( u : L_{bf} \rightarrow \mathbb{R} \) such that for any lottery \( L = [O, m] \) in \( L_{bf} \),
\[ u(L) = \sum_{a \in 2^O} m(a) u_a \]  
(30)
where \( u_a = u([O, m^d_a]) \), and \( m^d_a \) is a deterministic BPA for \( O \) such that \( m^d_a(a) = 1 \).

Thus, Jaffray’s axioms result in the same solution as that of Corollary 4.1, which is a special case of Theorem 4.2. As Jaffray’s axioms do not use Dempster’s rule explicitly, it is not clear whether Eq. (30) applies to the D-S framework or not. The mixture BPA \( m \) derived from BPAs \( m_1 \) and \( m_2 \) using Eq. (29) is not Dempster’s combination rule, although Eq. (29) can be derived from a belief function model using Dempster’s rule. By deriving this solution from a set of axioms making use of the basic constructs of DS theory (namely, Dempster’s combination rule, marginalization,
and conditional embedding), we provide additional arguments supporting Eq. (30) as a natural definition of the real-valued utility of a bf lottery in the D-S theory.

Also, there is no explicit notion of a bf reference lottery in Jaffray’s framework. Thanks to our continuity axiom (Assumption 4.3), the interval-valued utility \[ u_a, u_a + w_a \] in our framework receives a simple interpretation as an interval-valued probability of a best outcome \( O_1 \), in a bf reference lottery \[ O_2, m_a \] that is indifferent to \( a \) and such that \( m_a(\{O_1\}) = u_a, m_a(\{O_r\}) = 1 - (u_a + w_a) \), and \( m_a(O_2) = w_a \). We believe that this simple interpretation can be very helpful when eliciting utilities from DMs.

5.2 Comparison with Smets’ Decision Theory

Smets’ decision theory [35] is a two-level framework where beliefs, represented by belief functions, are held at a credal level. When a DM has to make a decision, the marginal belief function for a variable of interest is transformed into a PMF, and the Bayesian expected utility framework is then used to make a decision.

Smets uses a transformation called the pignistic transform to transform belief functions into PMFs. This transform is justified in [36] using a mixture property as follows. Let \( T \) denote the belief-PMF transformation. Smets [36] argues that this transformation should be linear, i.e., we should have, for any \( \lambda \in [0,1] \),

\[
T(\lambda m_1 + (1 - \lambda) m_2) = \lambda T(m_1) + (1 - \lambda)T(m_2).
\]

(31)

The unique transformation \( T \) verifying (31) is the pignistic transformation defined as \( T(m) = BetP_m \) with

\[
BetP_{m}(O) = \sum_{a \subseteq O} \frac{m(a)}{|a|} I(O \in a)
\]

(32)

for all \( O \in \mathcal{O} \), where \( I(\cdot) \) is the indicator function. The pignistic PMF \( BetP_m \) is mathematically identical to the Shapley value in cooperative game theory [32]. In [36], Smets attempts to derive Eq. (31) from the maximum expected utility principle. The argument, however, is quite technical and not very compelling.

Given the definition in Eq. (32), the expected utility of a bf lottery \( L = [O, m] \) according to the pignistic PMF is

\[
u_{BetP}(L) = \sum_{O \in \mathcal{O}} BetP_{m}(O) u_{\{O\}}
\]

(33a)

\[= \sum_{O \in \mathcal{O}} \left( \sum_{a \subseteq O} \frac{m(a)}{|a|} I(O \in a) \right) u_{\{O\}}
\]

(33b)

\[= \sum_{a \subseteq \mathcal{O}} m(a) \left( \frac{1}{|a|} \sum_{O \in a} u_{\{O\}} \right).
\]

(33c)

It is a special case of Eq. (23), with

\[
u_a = \frac{1}{|a|} \sum_{O \in a} u_{\{O\}}.
\]

Smets’ decision theory thus amounts to assuming that a DM is indifferent between a bf lottery that gives them an outcome in \( a \) for sure, and a bf reference lottery in which the probability of
the best outcome is equal to the average utilities of the outcomes in \( a \). This is consistent with our Assumptions 4.1–4.6 but it is inconsistent with Assumption 4.7. For instance, in Example 7, a DM using the pignistic criterion would strictly prefer \( b \) to \( a \), even though the act \( f_1 \) generating \( a \) dominates the act \( f_2 \) generating \( b \). Moreover, this restricted model does not have any parameter to represent a DM’s attitude toward ambiguity. As a result, it is unable to explain Ellsberg’s paradox and the ambiguity aversion of human DMs as described, e.g., in the examples in Section 4.3.

5.3 Comparison with Shafer’s Constructive Decision Theory

Shafer [31] argues for a decision theory that allows us to construct both goals and beliefs in response to a decision. In the vN-M utility theory, we start with a probabilistic lottery, and construct a utility function that reflects a DM’s risk attitude. Thus, probabilities and utilities are separate constructs that are then combined for the computation of expected utility. In many situations, we have neither objective nor subjective probabilities. For such situations, Shafer argues for constructing belief functions from available evidence, and constructing a set of consistent and monotonic goals. Given a set of actions, we examine which goals each of the actions will achieve. We use belief functions to make judgments based on evidence about what will happen if an action is taken. We then use these belief functions to compute the expected number of goals that an action will satisfy, and pick an action that satisfies the most goals. This can be generalized to the case where not all goals are equally weighted, some are weighted more than others.

Our utility theory is more in line with vN-M utility theory than Shafer’s constructive decision theory. There is considerable literature in many domains about the use of utility theory for decision making. While Shafer’s constructive decision theory is intriguing and may indicate an interesting direction to explore, there is much to be done before we can apply it in many domains for which we have a decision theory in the vN-M style.

6 Summary and Conclusions

In this section, we summarize our proposal and sketch some future work. We start with Luce and Raiffa’s version of the vN-M utility theory for probabilistic lotteries. We then consider bf lotteries, lotteries when our beliefs about the state of the world is described by DS belief functions. We use a similar set of axioms as vN-M, but first we replace each singleton outcome in a probabilistic lottery by a focal set of a BPA. Second, we replace the reduction of compound lotteries with a corresponding axiom that uses Dempster’s combination rule and belief function marginalization in place of probabilistic combination (pointwise multiplication followed by normalization) and probabilistic marginalization (addition). Third, we use a bf reference lottery with two independent parameters. The axioms lead to a decision theory that involves assessing the utility of each focal element of a BPA as an interval-valued utility. Interval-valued utilities lead to a partial preference relation on the set \( L_{bf} \) of all bf lotteries. If we use Bayesian bf reference lotteries with a single parameter, then our axiomatic framework leads to a real-valued utility function that is exactly the same as in Jaffray’s linear utility theory [21].

The decision theory that results from our axioms is more general than that proposed by Jaffray [21], which can be construed as a decision theory for belief functions interpreted as generalized probabilities. Jaffray’s axiomatic theory is based on a set of mixture BPAs. A mixture of two BPAs is not the same as a Dempster’s combination of two BPAs, although we could construct a belief function model where the mixture BPA is obtained by Dempster’s rule. Thus, it is not clear if Jaffray’s linear utility theory is applicable to D-S belief function lotteries or not. Our utility theory
confirms that this is indeed the case. Our bf reference lotteries lead to interval-valued utilities, and consequently, a partial preference relation on the set of all bf lotteries.

We also compare our axiomatic theory to Smets’ two-level framework [35, 36], and note that his framework is too constrained to explain ambiguity-aversion or ambiguity-seeking behavior of human DMs. Shafer [31] has recently published his constructive decision theory where he rejects the separation of beliefs and utilities. He proposes, instead, constructing a set of consistent and monotonic goals, and measuring the utility of each choice by the number of goals (or weighted goals) achieved by the choice. Shafer’s constructive decision theory needs to be fleshed out before it can be applied to practical decision-making situations.

In practice, implementing the most general form of our axiomatic theory may need assessment of $2^k$ parameters, where $k$ is the number of focal sets of a bf lottery. In the worst case, $k$ can be as large as $2^{|O|} - 1$. Based on additional assumptions, we propose a model based on only two parameters, which can be interpreted as reflecting both the DM’s attitude to ambiguity and their indeterminacy. This model, as well as others, will have to be further studied and developed. More generally, a rigorous methodology to elicit interval-valued utilities remains to be designed.

**Acknowledgments**

This research was done during a visit by the second author to Université de Technologie de Compiègne (UTC) in Spring 2019. This research was funded partially by a sabbatical from the University of Kansas, and supported by the Labex MS2T, which is funded by the French Government through the program “Investments for the future” by the National Agency for Research (reference ANR-11-IDEX-0004-02). Thanks to Sébastien Destercke for many discussions on the topic of this manuscript, and to Suzanna Emelio for proofreading this manuscript.

**References**


