International Journal on Artificial Intelligence Tools

BRIDGING JEFFREY'S RULE, AGM REVISION AND DEMPSTER CONDITIONING IN THE THEORY OF EVIDENCE

JIANBING MA, WEIRU LIU

School of Electronics, Electrical Engineering and Computer Science,
Queen's University of Belfast, Belfast, UK, BT7 1NN
{jma03,w.liu}@qub.ac.uk

DIDIER DUBOIS, HENRI PRADE

IRIT, Université Paul Sabatier
118 Route de Narbonne 31062 Toulouse, Cedex 9, France
{dubois,prade}@irit.fr

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)

Belief revision characterizes the process of revising an agent's beliefs when receiving new evidence. In the field of artificial intelligence, revision strategies have been extensively studied in the context of logic-based formalisms and probability kinematics. However, so far there is not much literature on this topic in evidence theory. In contrast, combination rules proposed so far in the theory of evidence, especially Dempster rule, are symmetric. They rely on a basic assumption, that is, pieces of evidence being combined are considered to be on a par, i.e. play the same role. When one source of evidence is less reliable than another, it is possible to discount it and then a symmetric combination operation is still used. In the case of revision, the idea is to let prior knowledge of an agent be altered by some input information. The change problem is thus intrinsically asymmetric. Assuming the input information is reliable, it should be retained whilst the prior information should be changed minimally to that effect. To deal with this issue, this paper defines the notion of *revision* for the theory of evidence in such a way as to bring together probabilistic and logical views. Several revision rules previously proposed are reviewed and we advocate one of them as better corresponding to the idea of revision. It is extended to cope with inconsistency between prior and input information. It reduces to Dempster rule of combination, just like revision in the sense of Alchourron, Gärdenfors, and Makinson (AGM) reduces to expansion, when the input is strongly consistent with the prior belief function. Properties of this revision rule are also investigated and it is shown to generalize Jeffrey's rule of updating, Dempster rule of conditioning and a form of AGM revision.

Keywords: Theory of evidence; belief revision; probability; Dempster conditioning.

1. Introduction

Belief[a] revision depicts how an agent revises its prior beliefs when new evidence (called *input*) is received. Often, the input is to some extent conflicting with prior beliefs. There-

[a]This paper is an extended version of a conference paper [21].

fore, belief revision is a framework aiming at characterising the process of belief change, so as to accommodate new evidence and reach a new consistent set of beliefs.

Logic-based belief revision has been intensively studied since the well known AGM postulates [1] were proposed, e.g., by Katsuno and Mendelzon [17], Darwiche and Pearl [6], Nayak et al. [22], Booth and Meyer [4], Jin and Thielscher [16], etc. Belief revision on ranked epistemic states is also receiving more and more attention nowadays, e.g., Benferhat and colleagues [3,2], Ma and colleagues [19,20], etc. For numerical oriented uncertainty formalisms, Jeffrey's rule [15] is frequently referred to in probability theory. It has direct counterparts in possibility theory [6,2] and in the ordinal conditional function framework[29,18].

The theory of evidence, also known as Dempster-Shafer theory of evidence (DS theory) [7,24,32], rapidly gained a widespread interest for modeling and reasoning with uncertain/incomplete information. When two pieces of evidence are collected from two distinct sources, it is necessary to combine them to get an overall result. So far, many combination rules (e.g., Smets [28] for a review) have been proposed in the literature. These rules involve an implicit assumption that all pieces of evidence come from parallel sources that play the same role, hence they are symmetric. Especially, when a source is less important than another, the corresponding piece of information is discounted and the above symmetry assumption still applies. Combination is typically applied to pieces of information received from the "outside". However, an agent may have its own prior opinion (from the inside), and then receive some input information coming from outside. In such a case, the problem is no longer one of combination, it is a matter of revision. Revision is intrinsically asymmetric as it adopts an insider point of view so that the input information and prior knowledge play specific roles, while combination is an essentially symmetric process, up to the proper handling of unequal reliabilities of sources.

Combination and revision share a common feature: they deal with pieces of information of the same nature (for instance pieces of generic knowledge, or uncertain observations). However, two principles that do away with symmetry should guide the revision process:

(1) **Success postulate**: information conveyed by the input evidence should be retained after revision;
(2) **Minimal change**: the prior belief should be altered as little as possible while complying with the former postulate.

It should be noted that new evidence (input information) can be either sure or uncertain. Furthermore, if new evidence is uncertain, uncertainty can either be part of the input information, hence *enforced* as a constraint guiding the belief change operation (this is the idea of the success postulate) or it is meant to qualify the reliability of the (otherwise crisp) input information [10]. In the latter case, the success postulate is questionable. In this paper, we focus on the former case, where new uncertain evidence is accepted and serves as a constraint on the resulting belief state.

In probability theory, revision strategies are studied in the context of probability kinematics [8]. Probability kinematics considers how a prior probability measure should be changed based on a new probability measure, which should be preserved after revision, on a coarser frame (a subalgebra of the probability space). Jeffrey's rule [15] is the most commonly used rule for achieving this objective. It respects a form of minimal change and grants priority to the input information.

However, to the best of our knowledge, there is no widely accepted revision rule in evidence theory, partly because Jeffrey's rule cannot be directly extended to general mass functions. There are several proposals for an extension of Jeffrey's rule for belief functions. Quite often, the input information still takes the form of masses assigned to the elements of a partition of the frame of discernment. In this paper, we first discuss the form a revision rule (or operator) should take in order to comply with the success and the minimal change postulates when the input is a general mass function. We study previous mass-function-based revision rules, and adapt them to the case of inconsistent input. One of these proposals is shown to obey the principles of revision. It is also defined using another equivalent constructive approach that uses a bipartite graph framework. This equivalence is significant since these two approaches start from different perspectives, and in some sense, one can be seen as a justification for the other. Finally, we prove that our revision rule generalizes Jeffrey's rule, Dempster's conditioning, and a form of AGM revision conjointly, and start investigating the issue of iterated revision.

The rest of the paper is organized as follows. We give some preliminaries on DS theory in Section 2. In Section 3, we review revision rules in frameworks less expressive than mass functions. This section also discusses principles a revision rule on mass functions should satisfy. We then review existing revision rules for mass functions in Section 4 and propose one that generalizes Dempster rule of conditioning, Jeffrey's rule, and a crude form of AGM revision. We call it Jeffrey-Dempster revision rule. Section 5 reconsiders this revision rule from a different angle, proposing an algorithm for computing it. Section 6 provides several properties of the revision rule and preliminary results about its iteration.

2. The theory of evidence

Let W be a set, understood as a set of possible worlds, and called the frame of discernment. W contains the solution to a problem, or a description of possible states of affairs. An epistemic state is often either represented as a mere subset of possible worlds (for instance the set of models of a belief set made of propositional formulas), or by a single (subjective) probability distribution (typically in the Bayesian setting, where probabilities represent degrees of belief). The theory of evidence [24] in some sense reconciles the two views by defining a probability distribution over subsets of possible worlds.

Definition 1. A mass function is a mapping $m : 2^W \to [0, 1]$ such that

$$\sum_{E \subseteq W} m(E) = 1 \text{ and } m(\emptyset) = 0.$$

E is called a focal set of m if $m(E) > 0$. The set of focal sets of m is denoted by $\mathcal{F}_m$. Let $S(m)$ denote the support of m, i.e. the union of the focal sets, that is, $S(m) = \bigcup_{E \in \mathcal{F}_m} E$ where the Es are focal sets of m.

A mass $m(E)$ is not the probability of an event E, it is the probability that the epistemic state of an agent is properly represented by the set E. Denoting by $P(A)$ the probability of an event A, the weight $m(E)$ is really construed as the probability $P(\{E\})$ on the power set of W.

A mass function m is called *Bayesian* if and only if all its focal sets are singletons. Then it is a standard probability distribution. A mass function m is called *partitioned* if and only if its focal sets $E_1, \cdots, E_k$ form a partition of W, i.e., $E_1 \cup \cdots \cup E_k = W$ and $E_i \cap E_j = \emptyset$ for $i \neq j$. A mass function m is called *consonant* if and only if its focal sets are nested. In particular, if m has only one focal set E with $m(E) = 1$, it represents an epistemic state whereby all that is known is that the state of the world lies in E.

Given m, its corresponding belief function $Bel : 2^W \rightarrow [0, 1]$ is defined as

$$Bel(A) = \sum_{E \subseteq A} m(E),$$

and its corresponding plausibility function $Pl : 2^W \rightarrow [0, 1]$ is defined as

$$Pl(A) = 1 - Bel(\overline{A}) = \sum_{E \cap A \neq \emptyset} m(E).$$

Plausibility and belief coincide with a probability measure when the mass function is Bayesian. A belief function is a necessity measure, i.e. $Bel(A \cap B) = \min(Bel(A), Bel(B))$, equivalently a plausibility function is a possibility measure, i.e., $Pl(A \cup B) = \max(Pl(A), Pl(B))$, if and only if the mass function is consonant [24].

There are several conditioning methods for belief/plausibility functions [10]. The following, called Dempster conditioning is the most commonly used one [24].

Definition 2. Let B be a subset of W such that $Pl(B) > 0$, then the Dempster conditioning functions on B can be defined as

$$m(A|B) = \frac{\sum_{E: A = B \cap E} m(E)}{Pl(B)},$$

$$Pl(A|B) = \frac{Pl(A \cap B)}{Pl(B)},$$

$$Bel(A|B) = 1 - Pl(\overline{A}|B) = \frac{Pl(B) - Pl(\overline{A} \cap B)}{Pl(B)},$$

where $A \neq \emptyset$ and $m(\emptyset|B)$ is defined as 0.

Dempster conditioning can be viewed as a revision rule that transfers the mass bearing on each subset A to its subset $A \cap B$, thus sanctioning the success postulate. Moreover, resulting masses bearing on non-empty sets are renormalized via simple division by $Pl(B)$, i.e. the mass function inside B does not change in relative value, which expresses a form of minimal change. However the apparent asymmetry of the conditioning rule is due to the assumption that B is a true fact, while m involves uncertainty. It does not presuppose an intrinsic asymmetry, whereby the mass m represents the epistemic state of an agent receiving information B. Dempster conditioning can also be viewed as the *symmetric* fusion of uncertain pieces of evidence represented by m and a sure fact B represented by $m'(B) = 1$. In particular it is a special case of the well-known Dempster rule of combination defined as

$$(m \otimes m')(A) = \frac{\sum_{B,C:B\cap C=A} m(B)m'(C)}{\sum_{B,C:B\cap C\neq\emptyset} m(B)m'(C)}, \qquad (1)$$

which is formally a random set intersection followed by a renormalization. The status of Dempster conditioning with respect to symmetry (i.e., is it revision or combination?) is totally ambiguous. Dempster rule of combination is a symmetric extension of Dempster conditioning. The revision rule we shall advocate is an asymmetric extension thereof.

Another important concept in evidence theory is the one of specialisation. It generalizes a natural notion of comparison of information content. An epistemic state A is less informed than another one B if and only if $B \subseteq A$: the former is a specialisation of the latter.

This notion is naturally extended to mass functions as follows:

Definition 3. (Specialisation [11]) We write $m \sqsubseteq m'$ ($\sqsubseteq$ is typically called *s-ordering*) if and only if there exists a square matrix Σ with general term $\sigma(A, B)$, $A, B \in 2^W$ verifying

$$\sum_{A\subseteq W} \sigma(A, B) = 1, \forall B \subseteq W,$$

$$\sigma(A, B) > 0 \Rightarrow A \subseteq B,\ A, B \subseteq W,$$

such that $m(A) = \sum_{B\subseteq W} \sigma(A, B)m'(B), \forall A \subseteq W$.

The term $\sigma(A, B)$ may be seen as the proportion of the mass $m'(B)$ which is transferred (*flows down*) to A. Matrix Σ is called *a specialisation matrix*, and m is said to be a specialisation of m'. Specialisation is an extension of set-inclusion to random sets. In fact, specialisation between two mass functions can be equally seen as a generalization of implication in propositional logic. That is, we have

Proposition 1. *Let ϕ and ψ be two propositions in a language such that W is its set of interpretations. Denote $Mod(\phi)$ (resp. $Mod(\psi)$) the set of models of ϕ (resp. ψ) and let m and m' be two mass functions on W s.t. $m(Mod(\phi)) = 1$ and $m'(Mod(\psi)) = 1$ stating that proposition ϕ (resp. ψ) is true in m (resp. m'). If m is a specialisation of m', then $\phi \vdash \psi$.*

Proof: From Def. 3 and the fact that each of m and m' only has one focal set, i.e., $Mod(\phi)$ and $Mod(\psi)$, respectively, it must be $\sigma(Mod(\phi), Mod(\psi)) > 0$ (in fact $\sigma(Mod(\phi), Mod(\psi)) = 1$) which implies $Mod(\phi) \subseteq Mod(\psi)$ which is equivalent to $\phi \vdash \psi$. □

Example 1. Let $W = \{w_1, w_2, w_3\}$, and let m and m' be two mass functions such that $m(\{w_1\}) = 0.3$, $m(\{w_2\}) = 0.5$, $m(\{w_1, w_2\}) = 0.1$, $m(\{w_2, w_3\}) = 0.1$, and

$m'(\{w_1\}) = 0.1$, $m'(\{w_1, w_2\}) = 0.5$, $m'(\{w_2, w_3\}) = 0.4$. Then m is a specialisation of m', since m' *flows a mass value* 0.2 from $\{w_1, w_2\}$ to $\{w_1\}$ (i.e., $\sigma(\{w_1\}, \{w_1, w_2\}) = 0.4$), a mass value 0.2 from $\{w_1, w_2\}$ to $\{w_2\}$ and a mass value 0.3 from $\{w_2, w_3\}$ to $\{w_2\}$.

Table 1: The Specialisation Matrix Σ.

$m \setminus m'$	{1}	{2}	{3}	{1,2}	{1,3}	{2,3}	W
{1}	1	0	0	$\frac{2}{5}$	0	0	0
{2}	0	0	0	$\frac{2}{5}$	0	$\frac{3}{4}$	0
{3}	0	0	0	0	0	0	0
{1,2}	0	0	0	$\frac{1}{5}$	0	0	0
{1,3}	0	0	0	0	0	0	0
{2,3}	0	0	0	0	0	$\frac{1}{4}$	0
W	0	0	0	0	0	0	0

Notation $\{i\}$ stands for subset $\{w_i\}$. Value $\frac{2}{5}$ on the 1st row shows the fraction of the $m'(\{w_1, w_2\})$ that will flow down to subset $\{w_1\}$, namely $\sigma(\{w_1\}, \{w_1, w_2\}) = \frac{2}{5}$.

Note that for probability distributions specialisation is an empty concept: a probability measure is a specialisation of another if and only if they are equal. This is because a probability measure cannot convey incomplete information, contrary to an epistemic state represented by a set. Usually, informational comparison between probability measures is done via their respective entropy measures.

3. Revision rules

The major approach to belief revision is the one described by Gärdenfors and his colleagues[1]. Although couched in logical terms using a propositional language, its thrust is basically semantic, a closed belief set being equated to its subset of models. The axiomatic framework for the AGM revision explicitly uses the success postulate as a means of making the revision process asymmetric. The idea of minimal change is expressed in a more distributed way via the other AGM axioms. On the other hand, in probability kinematics, Jeffrey's rule [15] was also devised with a view to respect the success postulate, and the minimal change is stated as an inertia axiom (what need not change will not change). The minimal change principle has also been expressed by showing the result of Jeffrey's rule minimizes an informational distance to the prior probability [31]. In this section, we recall the AGM style revision in terms of sets of possible worlds, and describe the basics of Jeffrey's rule. Then we provide a rationale for a revision rule in evidence theory, that may, in some sense extend both types of revision.

3.1. *Revising sets of possible worlds*

Let $E \subseteq W, E \neq \emptyset$ be the prior epistemic state of an agent who knows that the actual world lies in E. Let $E_I \subseteq W, E_I \neq \emptyset$ be the input information stating that the actual world actually lies in E_I, that the agent supposedly accepts as being true. There are two cases distinguished by the AGM theory:

(1) Either $E \cap E_I$ is not empty (the two pieces of information are consistent), and the resulting epistemic state is $E \circ E_I = E \cap E_I$.
(2) Or $E \cap E_I = \emptyset$, (the two pieces of information are inconsistent), and the resulting epistemic state is some subset $E \circ E_I \subseteq E_I$.

The first situation is called *expansion* in the AGM theory. It is clearly symmetric and coincides with a fusion rule. The minimal change principle is at work since E is indeed minimally modified to accept E_I. The second situation corresponds to what Gärdenfors and his colleagues calls genuine revision, in that the resulting epistemic state denies the prior one. The AGM axioms enforce a certain method for the choice of a suitable subset $E \circ E_I \subseteq E_I$. Namely, they come down to assuming that the prior epistemic state E is the visible part of a deeper knowledge made of a weak ordering $\succeq$ on W, where by $w_1 \succeq w_2$ means that w_1 is more plausible than w_2, such that E contains all the maximal elements (we say that $\succeq$ is E-faithful). The weak order $\succeq$ defines a well-ordered partition $\{U_1, \cdots, U_n\}$ of W, where $\forall w_i \in U_i, w_j \in U_j, i > j$ if and only if $w_i \succ w_j$ and $E = U_1$. Let $i = \min\{j : U_j \cap E_I \neq \emptyset\}$. Then $E \circ E_I = \{w \in E_I, \nexists w' \in E_I, w' \succ w\}$ is the set $E_I \cap U_i$ of most plausible worlds in E_I.

It has been argued [9] that the existence of this plausibility ordering makes this construction ambiguous:

- One possible interpretation of the weak ordering $\succeq$ is that it represents the agent's background knowledge expressing that in general if what is known is E_I then the most plausible worlds are $E \circ E_I \subseteq E_I$. Then the agent possesses an explicit representation of this background knowledge in the form of rules $A \to B$ from which the weak ordering $\succeq$ can be derived. In that case, the revision process comes down to a form of non-monotonic inference whereby the background knowledge is queried, based on the incomplete observation E, and the plausible conclusion is $E \circ E_I$. But clearly the plausibility relation is not revised.
- Another interpretation is that of a fusion rule on a par between E_I that is a crisp information (whereby the elements $w \notin E_I$ are deemed impossible) and a *ranked* epistemic state $\succeq$ conveying some uncertainty, resulting in the restriction of $\succeq$ to E_I, that is, not really a revision process (in fact a kind of ranked expansion).

However, the genuine revision problem is an asymmetric process of minimally changing a *prior* plausibility ordering $\succeq$, viewed as a refined description of an epistemic state by means of another input plausibility ordering $\succeq_I$, that must be respected. In this setting the AGM axioms must be lifted to the revision of ranked epistemic states. There are works on the latter topic (see Dubois [9] for some bibliography). We adopt this view in this paper. However, in this subsection, we assume that the well-ordered partitions reduce to $(E, \overline{E})$ for $\succeq$, and $(E_I, \overline{E}_I)$ for the input weak order. Then the AGM revision of E by E_I reduces to:

$$\begin{aligned} E \circ E_I &= E \cap E_I \text{ if } E \cap E_I \neq \emptyset \\ &= E_I \text{ otherwise} \end{aligned} \tag{2}$$

Note that this revision rule also underlies a principle of minimal commitment: Any result of $E \circ E_I$ of the form $A \subset E_I$ makes sense when input and prior are inconsistent, and is of the AGM style: the ordering of possible worlds induced by the AGM axioms explains how to select the subset A. However if this ordering is not available then it is cautious to consider $E \circ E_I = E_I$. While we will be using the above revision rule in the sequel, all definitions of revision rules below can be adapted to the case where $E \circ E_I$ is the subset of most plausible worlds in E_I dictated by an E-faithful plausibility relation $\succeq$.

3.2. *Revising probability distributions*

Jeffrey's rule applies when the two weak orders $\succeq$ and $\succeq_I$ are encoded as probability distributions P (the prior epistemic state) and P_I respectively. Jeffrey's rule was introduced as follows.

Definition 4. Let P be a probability measure over W denoting and P_I be a probability measure over a partition $\{U_1, \cdots, U_n\}$ of W denoting the input evidence. Let $\circ_J$ denote Jeffrey's rule. Then, $\forall A \subseteq W$:

$$P \circ_J P_I(A) = \sum_{i=1}^{n} P_I(U_i)P(A|U_i). \tag{3}$$

Some axioms guide the revision process carried out by Jeffrey's rule. The success postulate writes:

$$P \circ_J P_I(U_i) = P_I(U_i). \tag{4}$$

It is clear that the coefficient $P_I(U_i)$ represents what the probability of U_i should be, and not (for instance) uncertainty about the reliability of piece of information $P_I(U_i)$.

Note that for $w \in U_i$, Jeffrey's rule (3) can be simplified as

$$P \circ_J P_I(\{w\}) = P_I(U_i)\frac{P(\{w\})}{P(U_i)}. \tag{5}$$

That is, the revised probability of element w within each U_i is the same as its prior probability in relative value. More generally, the probability of any event A conditional to any uncertain event U_i in this subalgebra is the same in the original and the revised epistemic states. Namely,

$$\forall U_i, \forall A, P(A|U_i) = P \circ_J P_I(A|U_i). \tag{6}$$

The minimal change principle expressed by the constraint of Eq. (6) says that the revised probability measure $P \circ_J P_I$ must preserve the conditional probability degree of any event A given uncertain event U_i. Jeffrey's rule of conditioning yields the unique distribution that satisfies Eqs. (4) and (6) (see, e.g., Chan and Darwiche[5]). The posterior probability also minimizes relative entropy with respect to the original distribution under the probabilistic constraints defined by the input P_I [31].

Shafer [25] noticed that one can see the input probability function P_I on the partition as a belief function on W, where the mass function is m_I such that $m_I(U_i) = P_I(U_i), \forall i$. Then he proved that for each input m_I, there exists a belief function with mass function m_D, such that Jeffrey's rule can be written as $P \circ_J P_I = m_P \otimes m_D$ applying Dempster rule of combination (1) between the Bayesian mass m_P associated to P and the mass function m_D that can be built from m_I. Note that this remark brings the obviously asymmetric Jeffrey's rule back to the symmetric camp of fusion rules just like it does for Dempster conditioning. Following this line, only Dempster rule of combination is needed for revision.

However, we claim that this formal result is misleading: it does not rule out the possibility of an asymmetric generalisation of Jeffrey's rule that is is more in the spirit of its inventor, that is, one that obeys the success postulate and a rule of minimal change with respect to the prior mass function, while extending Dempster conditioning.

3.3. *The need for an asymmetric revision rule in DS theory*

Let us look at the following example of revision problem (adapted from Jeffrey [15]).

Example 2. An agent inspects a piece of cloth by candlelight, gets the impression it is green ($m_I(\{g\}) = 0.7$) but concedes it might be blue or violet ($m_I(\{b, v\}) = 0.3$). However the agent's prior belief (before using a candle) about the piece of cloth (we have no information about how this opinion was formed) was that it was violet ($m(v) = 0.8$) without totally ruling out the blue and the green ($m(b, g) = 0.2$). How can she modify her prior belief so as to acknowledge the observation?

Clearly, the input evidence has priority over the prior belief, hence after revision, we should conclude that the cloth color is more plausibly green. However, symmetric combination rules in DS theory may fail to produce this result. For example, if we apply Dempster's rule of combination, then we get $m(\{v\}) = \frac{6}{11}, m(\{g\}) = \frac{7}{22}, m(\{b\}) = \frac{3}{22}$ which shows violet, in contrast with what the observation suggests, is the most plausible color, and that the situation is more confused than before getting new information. Indeed, one might expect the direct observation of the piece of cloth should result in being more informed than previously. The counterintuitive result produced by the combination rule here stems from the underlying assumption that we treat the prior belief and the input evidence on a par. Therefore, to solve the above belief change problem, the correct action is to perform revision instead of combination.

3.4. *Principles of mass function based revision*

Now we discuss how the two general principles of revision can be applied to mass-functions. Let $\circ$ be a revision operator which computes a posterior mass function $\hat{m} = m \circ m_I$ from two given other mass functions, one representing the prior belief state (m) and the other representing new evidence (m_I). Moreover, we consider revision rules that generalize Jeffrey's probability kinematics, Dempster rule of conditioning and the AGM-style revision rule (2) of epistemic states.

There are some difficulties to do so. First, we must adapt Dempster rule of conditioning to the AGM-style revision rule; the latter prescribes that if $m(E_I) = 0$, then $\hat{m}(E_I) = 1$, while it generalizes (2) whenever $m(E_I) > 0$. Next, it may be impossible to straightforwardly reconduct the basic axioms of Jeffrey's rule (4) and (6) to belief functions, because the expansion effect of the AGM-style revision rule does not apply to probability distributions. We propose the following basic principles for the revision of belief functions:

(1) **Success postulate through specialisation:** The first fundamental principle of revision is to preserve new evidence. Translated into the language of DS theory, this principle states that for $\hat{m} = m \circ m_I$, $\hat{m}$ should in some sense *imply* m_I. The notion of *implication* between mass functions is naturally extended by the notion of specialisation between two mass functions (Def. 3). Therefore the success postulate in evidence theory reads:

    Success Postulate for Belief Functions (SPBF): $\hat{m} \sqsubseteq m_I$.

    Note that Jeffrey's rule satisfies this property, stricto sensu, if the input information is interpreted as a partitioned mass function.

    A straightforward rendering of the probabilistic success postulate for belief functions does not make sense if no restriction is put on the form of m_I. This is because the case of expansion never occurs for a probability function, as argued in Dubois *et al.* [10]: you cannot make the focal elements of a probability measure more precise via revision. On the contrary, when revising a mass function, in general, focal sets of m_I will be made more precise by focal sets of m. So we cannot reconduct the success postulate (4) $\hat{Bel}(E) = Bel_I(E)$ when E is any focal element of m_I. One may expect that in some cases, $\hat{Bel}(E) > Bel_I(E)$ (confirming E). But this is implied by the above success postulate, since specialisation is stronger than the inequality between belief functions [11]. Nevertheless, the probabilistic form (4) of the success postulate can be recovered by a suitable restriction of the form of the input belief function [26], as seen later on.

(2) **Minimal change principle:** The issue is to define what *minimal change* means in DS theory in terms of mass functions. Intuitively, it suggests using informational distance functions d between two mass functions, m and m_I. Namely one can use d to look for a specialisation of m_I at minimal distance from m. However, under this approach, $d(m, m) = 0$ for any distance function d, hence we ought to have $m \circ m = m$ (since m itself is a specialisation of m and m is at minimal distance from itself among all specialisations of m). However, the combination of *independent* mass functions exhibits a *reinforcement* effect which cannot occur in logic-based belief merging. That is, $m \oplus m \neq m$ ($\oplus$ is a mass function combination operator) whilst applied to standard epistemic states, revision is idempotent: $E \circ E_I = E$ if $E_I = E$.

    For mass functions, instead of $m \circ m = m$, we may expect some reinforcement effect if the new evidence is identical to, but considered independent from, the prior beliefs. For instance we may believe to some degree that Toulouse rugby team won the European championship last year. If some friend coming from abroad says he believes it likewise, this piece of information confirms our prior belief, so that even if our opinion remains the same, we become more confident in it.

The bottom line is that there is a certain conflict between the minimal change principle and the confirmation effect when revising uncertain information. However, in the case of revising a belief function by another one that is consistent with it, one should expect that the prior information is still implied by the result of the revision. Two mass functions m_1 and m_2 are said to be *strongly consistent* if and only if $\forall A \in \mathcal{F}_{m_1}, B \in \mathcal{F}_{m_2}, A \cap B \neq \emptyset$. A minimal requirement for minimal change is then as follows

Expansion Minimal Change (EMC): If m and m_I are strongly consistent then
$$m \circ m_I \sqsubseteq m.$$

Note that this property also holds for Dempster rule of combination. It is the extension of one of the AGM postulates [1] requesting identity between expansion and revision in the case of receiving an input that is consistent with the original belief set.

A stronger form of minimal change, mimicking (6) could be the request that $[m \circ m_I](A|B) = m(A|B)$ for some focal subsets A not affected by the input information m_I where $m_I(B) > 0$. In particular, if m_I has focal sets forming a partition $\mathcal{F}_I = \{U_1, \cdots, U_n\}$, then one may request that $\forall i, [m \circ m_I](A|U_i) = m(A|U_i)$ for $A \subseteq U_i$. However this kind of postulate is specific to probability distributions: any focal set of the Bayesian mass m is included into a single focal set of the partitioned mass m_I. So we may adopt a counterpart to the minimal change principle (6) only in very drastic conditions:

Probabilistic Minimal Change (PMC): If $E_I \in \mathcal{F}_I$ is such that $\forall A \in \mathcal{F}_m$, either $A \subseteq E_I$ or $A \subseteq \overline{E}_I$, and moreover $\forall F \neq E_I \in \mathcal{F}_I, F \cap E_I = \emptyset$, then
$$\forall A \subseteq E_I, [m \circ m_I](A|E_I) = m(A|E_I)$$

This property cannot be used in place of EMC because its scope is very narrow, and in many situations it will hold by default because the preconditions are too drastic. On the other hand, any revision rule that violates it can be questioned. In this form it is satisfied by the AGM rule (2). Indeed, if $E \subseteq E_I$ then $E \circ E_I = E$, otherwise PMC does not apply. It is also respected by Dempster conditioning. Note that the EMC postulate also has a limited scope since most of the time m and m_I will not be strongly consistent.

(3) **Generalization of Jeffrey's rule:** Since a Bayesian mass function can be seen as a probability distribution, we would expect that a mass-function-based revision rule should generalize Jeffrey's rule. The latter is idempotent since $P \circ P = P$, which involves no confirmation effect, and also satisfies minimal change in the above sense (this is Eq. (6)). There is no paradox here: a probability P is never strongly consistent with itself and its focal sets cannot be precisiated. So EMC does not apply to this case.

(4) **Generalization of Dempster conditioning:** If $m_I(E_I) = 1$ for some subset $E_I \subseteq W$ such that $Pl(E_I) > 0$, then $\hat{Pl} = Pl(\cdot|E_I)$, in the sense of Dempster conditioning (this is true for Jeffrey's rule that reduces to conditioning when the input information is a sure fact). Note that Dempster rule of combination also specialises to such conditioning in this case. This is because combination and revision collapse to what Gärdenfors calls expansion when the input information is a sure fact consistent with the prior epis-

temic state. In the logical setting, AGM revision becomes symmetric; however, in the evidence setting, Dempster conditioning sounds asymmetric because the input information is not uncertain, contrary to the prior epistemic state.

(5) **Generalization of the AGM revision** Note that the Success Postulate SPBF and the Expansion Minimal Change postulate EMC for mass functions are respected by the revision (2) for epistemic states, considering them as mass functions. This rule yields the least committed result obeying both principles for epistemic states. It is thus natural to require that a revision rule in evidence theory should generalize (2).

As Dempster rule of conditioning is a special case of a revision rule, we must extend Definition 2 as follows in order to recover the AGM revision of epistemic states:

$$m_U(A|E_I) = \begin{cases} \frac{\sum_{E:A=E_I\cap E} m(E)}{Pl(E_I)} & \text{for } Pl(E_I) > 0, \\ 0 & \text{for } Pl(E_I) = 0 \text{ and } A \neq E_I, \\ 1 & \text{for } Pl(E_I) = 0 \text{ and } A = E_I. \end{cases}$$

Let us call this extended conditioning *universal Dempster conditioning*.

The corresponding conditional plausibility and belief functions are denoted by $Pl_U(A|E_I), Bel_U(A|E_I)$ as follows:

$$Pl_U(A|E_I) = \sum_{C\cap A\neq\emptyset} m_U(C|B)$$

and

$$Bel_U(A|E_I) = \sum_{C\subseteq A} m_U(C|E_I).$$

Compared with Definition 2, this definition is extended to the situation where $Pl(E_I) = 0$, in which case the above universal conditioning gives an intuitive result.

It is easy to check that $m_U(\cdot|E_I)$ is a mass function and it has an equivalent form as

$$\begin{aligned} m_U(A|E_I) &= I_{\{Pl(E_I)>0\}} \frac{\sum_{E:A=E_I\cap E} m(E)}{Pl(E_I)} + I_{\{Pl(E_I)=0,A=E_I\}} \\ &= I_{\{Pl(E_I)>0\}} m(A|E_I) + I_{\{Pl(E_I)=0,A=E_I\}}. \end{aligned} \quad (7)$$

where $I_{\{\phi\}}$ is an indicator function such that $I_{\{\phi\}} = 1$ if ϕ is true and $I_{\{\phi\}} = 0$ otherwise.

Similarly, $Pl_U(A|E_I)$ and $Bel_U(A|E_I)$ also have equivalent forms as

$$Pl_U(A|E_I) = I_{\{Pl(E_I)>0\}} Pl(A|E_I) + I_{\{Pl(E_I)=0,A=E_I\}}.$$
$$Bel_U(A|E_I) = I_{\{Pl(E_I)>0\}} Bel(A|E_I) + I_{\{Pl(E_I)=0,A=E_I\}}.$$

Then the following is clear

$$Pl_U(A|E_I) = \begin{cases} \frac{Pl(A\cap E_I)}{Pl(E_I)} & \text{for } Pl(E_I) > 0, \\ 0 & \text{for } Pl(E_I) = 0 \text{ and } A \neq E_I, \\ 1 & \text{for } Pl(E_I) = 0 \text{ and } A = E_I. \end{cases}$$

$$Bel_U(A|E_I) = \begin{cases} \frac{Pl(E_I)-Pl(\overline{A}\cap E_I)}{Pl(E_I)} & \text{for } Pl(E_I) > 0, \\ 0 & \text{for } Pl(E_I) = 0 \text{ and } A \neq E_I, \\ 1 & \text{for } Pl(E_I) = 0 \text{ and } A = E_I. \end{cases}$$

where $A \neq \emptyset$ and $m(\emptyset|E_I) = 0$.

4. Revision operators for mass functions

Within the scope of DS theory, the first attempt to actually generalize Jeffrey's rule to mass functions was made by Ichihashi and Tanaka [14]. They proposed three variants of this extended Jeffrey's rule they called *plausible*, *credible* and *possible conditioning*, respectively. Dubois and Prade [12] argued in favor of the first one as being a natural asymmetric variant of Dempster rule of combination (see also Dubois *et al.* [10]). These approaches are criticized by Smets [26], and he makes yet another proposal where the input evidence comes in the form of a belief function on a subalgebra of W. Smets [27] emphasises the need for mass function revision strategies but proposes no concrete revision rule. Halpern [13] makes the same proposal as the one advocated by Dubois and Prade. However, he restricts the rule to input evidence in the form of a belief function that can be represented as a probability measure on a partition $\{U_1, \cdots, U_n\}$ of W, i.e, such that $m(U_i) = \alpha_i$ and $\sum \alpha_i = 1$.

Ideally, a mass function based revision rule should accept evidence in the form of a general mass function definable on 2^W rather than just on a partition of W. Furthermore, like the original Jeffrey's rule, all above proposals require that the prior plausibilities $Pl(A)$ be positive whenever A is a focal set of the input mass function. In other words, this rule requires that no part of the new evidence be *in total conflict* with the agent's current beliefs, which restricts the application of this rule. This requirement is much more drastic than the minimal consistency requirement of Dempster rule of combination for which at least one focal set of the first mass function must be consistent with another one of the other mass function. In many cases, the revision rules of Ichihashi and Tanaka cannot be applied.

In this section we generalize previously proposed extensions of Jeffrey's rule by doing away with the latter limitation, and we show that only one of them generalizes Dempster rule of conditioning and at the same time admits any kind of input.

4.1. *Revision operators for belief functions: a general form*

As a belief function is defined by a mass function, it is natural to express revision rules in terms of mass functions. The idea of all revision operations is to share the input masses $m_I(B)$ among a family of focal subsets $\mathcal{F}_m^B$ of m related to B in some way. This sharing of the mass $m_I(B)$ is made proportionally to the weights $m(A), A \in \mathcal{F}_m^B$, in conformity with the idea of relative minimal change at work in Jeffrey's rule. Namely, the portion $\sigma_i(A, B)$ of $m_I(B)$ allocated to a prior focal set A should be proportional to $m(A)$ across all sets in $\mathcal{F}_m^B$. So these revision rules take the form

$$\hat{m}(A) = [m \circ m_I](A) = \sum_{B: A \in \mathcal{F}_m^B} \sigma(A, B) m_I(B) \tag{8}$$

where $\sigma(A,B) = \frac{m(A)}{\sum_{C \in \mathcal{F}_m^B} m(C)}$ expresses proportional sharing, so that $m \circ m_I$ is obviously a normalized mass function. A natural requirement is that $\forall B \in \mathcal{F}_I, \mathcal{F}_m^B \subseteq \{A \in \mathcal{F}_m : A \cap B \neq \emptyset\}$, since if $A \cap B = \emptyset$ there is no bridge between A and B.

Moreover, in order to deal with inconsistency in the style of AGM revision, i.e. generalizing the rule (2), the extreme case where $\sum_{C \in \mathcal{F}_m^B} m(C) = 0$ is handled as follows:

$$\sigma(A,B) = \begin{cases} 0 \text{ for } A \neq B, \\ 1 \text{ for} A = B. \end{cases}$$

If no focal set of the prior mass function m is in the scope $\mathcal{F}_m^B$ of the input focal set B, then value $m_I(B)$ should be totally allocated to B. It is easy to prove that $m \circ m_I$ is a mass function, i.e., $\sum_{A \subseteq W} (m \circ m_I)(A) = 1$ and $(m \circ m_I)(\emptyset) = 0$.

Such a form for a revision operation already ensures some properties will be fulfilled. First, we do have an extension of Jeffrey's rule.

Proposition 2. *If m is a Bayesian mass function defined by a positive probability distribution p, and m_I is a partitioned mass function on partition $\{U_1, \cdots, U_n\}$ of W, i.e, such that $m(U_i) = \alpha_i > 0$ and $\sum \alpha_i = 1$, then $m \circ m_I = p \circ_J m_I$ coincides with Jeffrey's rule.*

Proof: For a Bayesian prior mass function, and partitioned input mass function, $\mathcal{F}_m^{U_i}$ reduces then to the set of singletons with positive probability in U_i and it is then obvious that $\sigma(\{s\}, U_i) = \frac{P(\{s\})}{P(U_i)}$. Then $[m \circ m_I](\{s\}) = \sum_{s' \in U_i} \frac{p(s))\alpha_i}{P(U_i)} = [P \circ_J m_I](\{s\})$. □

Moreover, the general revision scheme (8) partially preserves the minimal change property of Jeffrey's rule:

Proposition 3. *If the preconditions of the PMC postulate are verified for a focal set B of m_I, then $m \circ m_I$ satisfies PMC for B.*

Proof: Let $B \in \mathcal{F}_I$ is such that $\forall A \in \mathcal{F}_m$, either $A \subseteq B$ or $A \subseteq \overline{B}$, and assume $\forall F \neq B \in \mathcal{F}_I, F \cap B = \emptyset$, then if $A \subseteq B$, (8) reduces to $\hat{m}(A) = \frac{m(A)m_I(B)}{\sum_{C \subseteq B} m(C)}$, and $Pl(B) = Bel(B) = \sum_{C \subseteq B} m(C)$; moreover $\hat{P}l(B) = \sum_{F \cap B \neq \emptyset} \frac{m(F)m_I(B)}{Pl(B)} = m_I(B)$ so that $\hat{m}(A|B) = \sum_{C = A \cap B} \frac{m(C)m_I(B)}{\hat{P}l(B)Pl(B)} = m(A|B)$ and PMC holds. □

4.2. *Inner vs. outer revision operators*

We first propose *inner* and *outer revision* operators which are named after the concepts of inner and outer probability measures, both of which are closely related to upper and lower probabilities in the sense of Dempster [7]. We define an *inner* revision operator denoted by $\circ_i$, as follows.

Definition 5. Let m and m_I be two mass functions over W, then the inner revision of m by m_I is defined as $(m \circ_i m_I)(A) = \sum_{A \subseteq B} \sigma_i(A,B) m_I(B)$ where

$$\sigma_i(A,B) = \begin{cases} \frac{m(A)}{Bel(B)} & \text{for } Bel(B) > 0, \\ 0 & \text{for } Bel(B) = 0 \text{ and } A \neq B, \\ 1 & \text{for } Bel(B) = 0 \text{ and } A = B. \end{cases}$$

The intuition behind inner revision can be illustrated as follows. To obtain the revised mass value for A, we *flow down* some of the mass value of every positive $m_I(B)$ to subsets $A \subseteq B$ proportionally to $m(A)$. The inner revision is (8) with $\mathcal{F}_m^B = \{A : A \subseteq B, m(A) > 0\}$. If event B is not certain at all a priori ($Bel(B) = 0$), then the value $m_I(B)$ is fully allocated to B. By construction, the inner revision operator is a specialisation of m_I, so that it satisfies our success postulate. In some sense, it preserves as much information from m as possible. The latter point is ensured by the proportional reallocation. However, in the case when m and m_I are strongly consistent, the resulting mass function is not necessarily a specialisation of the prior mass function. The reason is that $Bel(B) = 0$ is possible even if each focal set of m_I intersect all focal sets of m. So there may be an input focal set B such that $m(B) = 0, Pl(B) > 0, \hat{m}(B) = m_I(B) > 0$, and B is a subset of no focal set of m. So this rule does not obey the EMC principle. Moreover it does not generalize Dempster conditioning.

The *outer* revision operator denoted by $\circ_o$ is defined as follows.

Definition 6. Let m and m_I be two mass functions over W, then the outer revision operator that revises m with m_I is defined as $[m \circ_o m_I](A) = \sum_{A \cap B \neq \emptyset} \sigma_o(A, B) m_I(B)$ where

$$\sigma_o(A, B) = \begin{cases} \frac{m(A)}{Pl(B)} & \text{for } Pl(B) > 0, \\ 0 & \text{for } Pl(B) = 0 \text{ and } A \neq B, \\ 1 & \text{for } Pl(B) = 0 \text{ and } A = B. \end{cases}$$

This change rule naturally appears as the dual of the former one. The intuition of outer revision is similar to that of inner revision except that here for any A, we flow down portions of mass values of Bs to subsets A such that $A \cap B \neq \emptyset$ preserving the masses $m(A)$ in relative value across the concerned A sets (dividing them by $Pl(B)$). So, the outer revision is (8) with $\mathcal{F}_m^B = \{A : A \cap B \neq \emptyset\}$. Note that for outer revision, the revised result is not necessarily a specialisation of m_I, since sets not included in B receive portions of mass $m_I(B)$.

Our inner (resp. outer) revision rule coincides with Ichihashi and Tanaka's credible (resp. possible) conditioning rule when $Bel(B) > 0$ (resp. $Pl(B) > 0$), and extends them by addressing the revision situation where new evidence totally conflicts with prior beliefs. That is, contrary to Ichihashi and Tanaka [14], revision can be done even when $Bel(B) = 0$ (resp. $Pl(B) = 0$).

Example 3. Let m and m_I be two mass functions on W, such that $m(\{w_1\}) = 0.2$, $m(\{w_1, w_2\}) = 0.8$, and $m_I(\{w_1\}) = 0.4, m_I(\{w_1, w_2\}) = 0.4$, $m_I(\{w_4\}) = 0.2$.

Applying inner revision operator $\circ_i$, we get $\hat{m}_i = m \circ_i m_I$ where
$\hat{m}_i(\{w_1\}) = m_I(\{w_1\}) \frac{m(\{w_1\})}{Bel(\{w_1\})} + m_I(\{w_1, w_2\}) \frac{m(\{w_1\})}{Bel(\{w_1, w_2\})} = 0.48$,
$\hat{m}_i(\{w_1, w_2\}) = m_I(\{w_1, w_2\}) \frac{m(\{w_1, w_2\})}{Bel(\{w_1, w_2\})} = 0.32$,
$\hat{m}_i(\{w_4\}) = 0.2$.

Similarly, applying outer revision operator $\circ_o$, we get $\hat{m}_o = m \circ_o m_I$ s.t. $\hat{m}_o(\{w_1\}) = 0.16$, $\hat{m}_o(\{w_1, w_2\}) = 0.64$, and $\hat{m}_o(\{w_4\}) = 0.2$.

However, none of these change rules respect the two basic revision postulates in the spirit of Jeffrey's rule. As a consequence they do suffer from some drawbacks.

Example 4. Let $m(\{w_1, w_2\}) = 1$ and $m_I(\{w_1, w_3\}) = 1$, then intuitively m supports w_1 while rejects w_3, and hence we expect the revision result to be $\hat{m}(\{w_1\}) = 1$. However, from inner revision, the revised result is $\hat{m}_i(\{w_1, w_3\}) = 1$ whilst from outer revision, the revised result is $\hat{m}_o(\{w_1, w_2\}) = 1$. None of the results are fully agreeable with intuitions. In particular, they are not in agreement with the AGM revision principles (that enforce expansion in this case). They do not extend the revision rule (2) for epistemic states.

4.3. *A modified outer revision operator*

As mentioned earlier, neither the inner nor the outer revision qualifies as revision operations. In particular the result of outer revision is not necessarily a specialisation of the mass function representing new evidence, but it obeys minimal change beyond what is requested, even by the AGM postulates. In this section, we modify outer revision so as to make it a specialisation of the new evidence, while respecting the EMC postulate. We define operator $\circ_{om}$ as a modified outer revision operator, that extends the third revision rule of Ichihashi and Tanaka:

Definition 7. Let m and m_I be two mass functions over W. The modified outer revision of m by m_I s.t. for any $C \neq \emptyset$, is defined by

$$(m \circ_{om} m_I)(C) = \sum_{A \cap B = C} \sigma_m(A, B) m_I(B) \tag{9}$$

where $\sigma_m(A, B) = \begin{cases} \frac{m(A)}{Pl(B)} & \text{for } Pl(B) > 0, \\ 0 & \text{for } Pl(B) = 0 \text{ and } A \neq B, \\ 1 & \text{for } Pl(B) = 0 \text{ and } A = B. \end{cases}$

Note that $\sigma_m(A, B)$ is exactly the same as $\sigma_o(A, B)$. The only difference between the modified revision rule and its predecessor is that instead of flowing down a portion of $m_I(B)$ to the prior focal set A when $A \cap B \neq \emptyset$, we flow down this portion to $A \cap B$. Here, $\mathcal{F}_m^B = \{A : A \cap B \neq \emptyset\}$, and one could also write

$$[m \circ_{om} m_I](C) = \sum_{A,B: B \cap A = C, A \in \mathcal{F}_m^B} \sigma_m(A, B) m_I(B).$$

This modification makes the revision result truly a specialisation of m_I, hence it satisfies the success postulate, but it is not exactly of the form (8). Likewise, if m and m_I are strongly consistent then it reduces to Dempster rule of combination where normalisation is not needed. So the EMC rule is also satisfied. Also, the modified outer revision extends the plausible conditioning rule [14,12] to situations where $Pl(B) = 0$. Note that conditions $Pl(B) = 0$, $A = B$, and $A \cap B = C$ together imply $Pl(C) = 0$, hence Eq. (9) can also be equivalently written as

$$(m \circ_{om} m_I)(C) = \sum_{A \cap B = C, Pl(B) > 0} \frac{m(A) m_I(B)}{Pl(B)} + I_{\{Pl(C)=0\}} m_I(C) \tag{10}$$

Example 5. (Ex. 4 cont') Let $m(\{w_1, w_2\}) = 1$ and $m_I(\{w_1, w_3\}) = 1$. Applying $\circ_{om}$ we get a revision result $\hat{m}$ such that $\hat{m}(\{w_1\}) = 1$, which is exactly what is expected.

In fact this modified outer rule coincides with the AGM revision rule (2) for flat epistemic states described by sets. When $Pl(B) > 0, \forall B \in \mathcal{F}_I$, Dubois and Prade [12] show that the modified outer revision rule can be expressed using Jeffrey-like equations and Dempster conditioning (Def. 2).

$$\hat{m}(A) = m \circ_{om} m_I(A) = \sum_{B \subseteq W} m(A|B) m_I(B), \tag{11}$$

$$\hat{Pl}(A) = \sum_{B \subseteq W} Pl(A|B) m_I(B), \tag{12}$$

$$\hat{Bel}(A) = \sum_{B \subseteq W} Bel(A|B) m_I(B), \tag{13}$$

These results still hold if universal Dempster conditioning is used to cope with the case where $Pl(B) = 0$ for some $B \in \mathcal{F}_I$, as proposed here. As pointed out by Dubois and Prade [12], these Jeffrey-like equations are also formally justified by Wagner's statement [30] that the only event-wise combination valid for both plausibility and belief functions is the linear convex combination, as it is the case for probability measures.

The above properties suggest a natural name for this change rule, namely, *Jeffrey-Dempster revision rule* and we can denote it by $\circ_{JD}$.

Similarly, Halpern's belief function revision rule is defined as follows [13].

Definition 8. Let m be a belief function over 2^W denoting the prior epistemic state and m_I be a belief function such that $\mathcal{F}_I = \{U_1, \cdots, U_n\}$ is a partition of W denoting the input evidence. Let $\circ_H$ denote Halpern's revision rule, then we have

$$Bel \circ_H Bel_I(V) = \sum_{i=1}^{n} m(U_i) Bel(V|U_i).$$

This is clearly a special case of Jeffrey-Dempster revision rule, since it is of the form (13).

Def. 7 gives a succinct representation of the Jeffrey-Dempster revision operator, but the above expressions suggest a calculation method. It uses the same table as the one for showing how to compute Dempster rule of combination.

We build a table M in which each column stands for a focal element of m_I and each row stands for a focal element of m. In each entry (A, B) of M (A is a row focal element and B is a column focal element), we put $A \cap B$ and a weight $w(A \cap B)$ computed as follows.

For each column B, if there are non-empty $A \cap B$ in the column, then we define $w(A \cap B) = \frac{m_I(B) m(A)}{Pl(B)}$ (and zero if $A \cap B$ is empty). If all sets in the column are empty then we set $w(B) = m_I(B)$. Finally we sum over prior focal sets A in column B all the weights bearing on the same set $C = A \cap B$, to get the resulting posterior mass $\hat{m}(C)$ of that set.

Example 6. Let m and m_I be two mass functions on W, such that $m(\{w_1\}) = 0.3$, $m(\{w_1, w_2\}) = 0.7$, and $m_I(\{w_1\}) = 0.3, m_I(\{w_1, w_2\}) = 0.5, m_I(\{w_3\}) = 0.2$.

The matrix M is as follows.

Table 2: The Matrix M.

$m \setminus m_I$	$\{w_1\}, 0.3$	$\{w_1, w_2\}, 0.5$	$\{w_3\}, 0.2$
$\{w_1\}, 0.3$	$\{w_1\}$, 0.09	$\{w_1\}$, 0.15	$\emptyset, 0$
$\{w_1, w_2\}, 0.7$	$\{w_1\}$, 0.21	$\{w_1, w_2\}$, 0.35	$\emptyset, 0$

Note that in the above matrix, since all the intersected sets in the fourth column are empty, we get $m(\{w_3\}) = 0.2$, so finally we get $\hat{m} = m \circ_{om} m_I$ such that $\hat{m}(\{w_1\}) = 0.45, \hat{m}(\{w_1, w_2\}) = 0.35, \hat{m}(\{w_3\}) = 0.2$.

4.4. *Smets revision rule*

Smets [26] considers the revision of a mass function m by an input m_I having a particular structure. Namely, consider a partition $\mathcal{P} = \{U_1, U_2, \ldots, U_n\}$ of W and let $\mathcal{B}$ the Boolean algebra of subsets induced by $\mathcal{P}$ (unions of elements in $\mathcal{P}$). Assume the set $\mathcal{F}_I$ of focal sets of m_I only contains elements in $\mathcal{B}$.

This revision rule is again of the form (8), but with a different $\mathcal{F}_m^B$. Namely, for each subset $A \subseteq W$, define A^*, the upper approximation of A with respect to $\mathcal{P}$, as:

$$A^* = \cup_{i:U_i \cap A \neq \emptyset} U_i = \cap\{B : A \subseteq B, B \in \mathcal{B}\}.$$

This is the upper approximation in the sense of rough sets [23]. Now let $\mathcal{F}_m^B = \{A, m(A) > 0, [A \cap B]^* = B\}, \forall B \in \mathcal{B}$. Then revision operator $\circ_S$ is such that for any $C \neq \emptyset$,

$$(m \circ_S m_I)(C) = \sum_{A \cap B = C, [A \cap B]^* = B} \sigma_S(A, B) m_I(B) \tag{14}$$

where $\sigma_m(A, B) = \begin{cases} \frac{m(A)}{\sum_{D:[D \cap B]^* = B} m(D)} & \text{if the denominator is positive,} \\ 0 & \text{otherwise and } A \neq B, \\ 1 & \text{otherwise and } A = B. \end{cases}$

The specific feature of this revision rule is that any focal subset C of $m \circ_S m_I$ receives a portion of weight from a single focal set of m_I, namely its upper approximation. This is one feature, apart from the specific form of the input, that makes it differ from the Jeffrey-Dempster revision rule, for which $m_I(B)$ is shared among all possible subsets of B that are the intersection of B and a focal set of the prior m. It enables a stronger form of success to be satisfied, namely

$$Bel_{m \circ_S m_I}(B) = Bel_I(B), \forall B \in \mathcal{B}.$$

This is easy to figure out: For $B \in \mathcal{B}, \sum_{C \subseteq B} \hat{m}(C) = \sum_{C:C^* \subseteq B} k_C m_I(C^*)$, where k_C is the portion of $m_I(C^*)$ allocated to C in the final result. Now by construction $\sum_{D, D^* = C^*} k_D = 1$ since $m_I(C^*)$ is shared among those of its subsets of which it is the upper approximation. So,

$$\sum_{C:C^*\subseteq B} k_C m_I(C^*) = \sum_{B'\subseteq B} m_I(B') \sum_{C:C^*=B'} k_C = \sum_{B'\subseteq B} m_I(B').$$

The price paid is a restriction on the form of the input that is questionable, namely, if two input focal sets E_1 and E_2 intersect, then their intersection is a focal set of m_I. Suppose m_I possesses two intersecting focal sets with $m_I(E_1 \cap E_2) = 0$. The Smets framework requires to define a partition of W such that E_1 and E_2 belong to the corresponding algebra. There is a degree of freedom here as all elements of the corresponding partition will have mass 0; the simplest is the 4-element partition induced by E_1 and E_2. In any case, all focal sets of m included in $E_1 \cap E_2$ will get zero posterior mass since $m_I(E_1 \cap E_2) = 0$ and they can receive no mass from $m_I(E_1)$ nor $m_I(E_2)$. In particular if $E \subseteq E_1 \cap E_2, \forall E \in \mathcal{F}_m$, then $E^* = E_1 \cap E_2, \forall E \in \mathcal{F}_m$ and $m \circ_S m_I = m_I$, even though m is strongly consistent with m_I. This goes against the EMC rule.

Smets [26] also considers a counterpart of the inner revision rule that replaces $A \cap B = C$ by $A \subseteq B$ in (14). But again, it violates the minimal change principle and does not extend Dempster conditioning.

5. An algorithmic view of the revision rule for belief functions.

The above discussion should have convinced the reader that the most reasonable extension of Jeffrey's revision rule to belief functions among existing proposals is the Jeffrey-Dempster revision rule, also retained by Dubois and Prade [12] in their older comparative study. In this section we provide a different view that leads to the same revision rule.

Intuitively, for mass-function-based revision, prior information that is consistent with the new information is refined by expansion. By consistent, we mean focal sets of m_I that overlap the support $S(m)$. That is, if A is a focal set of m_I and $A \cap S(m) \neq \emptyset$, then both $m_I(A)$ and the mass $m(A)$, if positive, should affect the posterior mass function; otherwise, $m_I(A)$ should be retained after revision. Symmetrically if $m(A) > 0$, $A \cap S(m_I) = \emptyset$, then $m(A)$ should be removed from consideration.

Example 7. Let $W = \{w_1, w_2, ..., w_8\}$, define m such that
$m(\{w_1, w_8\}) = 0.2, m(\{w_1, w_2\}) = 0.4, m(\{w_3\}) = 0.3, m(\{w_6, w_7\}) = 0.1,$
and m_I such that $m_I(\{w_1, w_2\}) = 0.5, m_I(\{w_4, w_5\}) = 0.3, m_I(\{w_6\}) = 0.2,$
then $\hat{m} = m \circ m_I$ should *imply* m_I. Observe that the prior m rules out $\{w_4, w_5\}$. Hence $m_I(\{w_4, w_5\}) = 0.3$ should be retained after revision, i.e.,$\hat{m}(\{w_4, w_5\}) = 0.3$, and no other focal sets of the posterior $\hat{m}$ shall contain w_4 or w_5. On the other hand, one can expect that also $\hat{m}(\{w_3\}) = 0$, since $w_3 \notin S(m_I)$, even though $m(\{w_3\}) = 0.3$.

From Example 7, we also observe: Focal element $\{w_1, w_2\}$ of m_I is consistent with focal sets of $\{w_1, w_8\}$ and $\{w_1, w_2\}$ of m, so $m(\{w_1, w_8\}), m(\{w_1, w_2\})$ should affect the revised mass of $\{w_1, w_2\}$. Similarly, $\{w_6, w_7\}$ of m and $\{w_6\}$ of m_I are consistent. But $\{w_1, w_8\}, \{w_1, w_2\}$ of m and $\{w_6\}$ of m_I are not, and likewise for focal sets $\{w_6, w_7\}$ of m and $\{w_1, w_2\}$ of m_I.

These observations show that we can partition $S(m) \cup S(m_I)$ on the basis of mutually consistent focal sets of m and m_I as follows. Define a reflexive symmetric relation $\sim$ on $\mathcal{F}_m \cup \mathcal{F}_I$ by $A \sim B$ and $B \sim A$ if and only if either $A = B$, or $A \in \mathcal{F}_m, B \in \mathcal{F}_I$ and $A \cap B \neq \emptyset$. Consider its transitive closure $\sim^*$. It is an equivalence relation on $\mathcal{F}_m \cup \mathcal{F}_I$, the equivalence classes of which form a partition of the set of focal sets $\mathcal{F}_m \cup \mathcal{F}_I$. Denote by $\mathcal{F}_1, \ldots \mathcal{F}_k$ its equivalence classes when they contain more than a single focal set, and by $\mathcal{F}_{uncor}$ the set of focal sets that form singleton equivalence classes. Each family $\mathcal{F}_i$ is called a correlated group. Its focal sets can be arranged in a sequence where any two adjacent sets overlap. Clearly, any focal subset in $\mathcal{F}_i$ is disjoint from any other subset in $\mathcal{F}_j$ if $i \neq j$.

Let $S_1 \cup \cdots \cup S_k \cup S_{uncor} = S(m) \cup S(m_I)$, where each element S_i of the partition of $S(m) \cup S(m_I)$ corresponds to the union of focal sets in a family $\mathcal{F}_i \subset \mathcal{F}_m \cup \mathcal{F}_I$ of focal sets of both m and m_I. S_{uncor} is the union of focal sets of m which have no intersection with $S(m_I)$ and focal sets of m_I which have no intersection with $S(m)$. In Example 7, we should have $k = 2, S_1 = \{w_1, w_2, w_8\}$, and $S_2 = \{w_6, w_7\}$ and $S_{uncor} = \{w_3\} \cup \{w_4, w_5\} = \{w_3, w_4, w_5\}$.

The above construction is in fact equivalent to computing maximal connected components in a certain non-directed bipartite graph induced by the sets of focal sets $\mathcal{F}_m$ and those in $\mathcal{F}_I$. A bipartite graph is a set of graph vertices V decomposed into two disjoint sets V_1 and V_2, such that no two graph vertices (or nodes) within the same set are adjacent. Namely, consider the bipartite graph with two sets of nodes: focal sets in $V_1 = \mathcal{F}_m$ and those in $V_2 = \mathcal{F}_I$ (if the same focal set appears in $\mathcal{F}_m$ and $\mathcal{F}_I$ it produces two nodes, one in each node set). Arcs connect one focal set $A \in \mathcal{F}_m$ to one focal set $B \in \mathcal{F}_I$ if and only if $A \cap B \neq \emptyset$. Each S_i is the union of focal sets corresponding to a maximal connected component in the graph. The set S_{uncor} is the union of focal sets corresponding to isolated nodes in the bipartite graph. The bipartite graph for Example 7 is shown in Fig. 1.

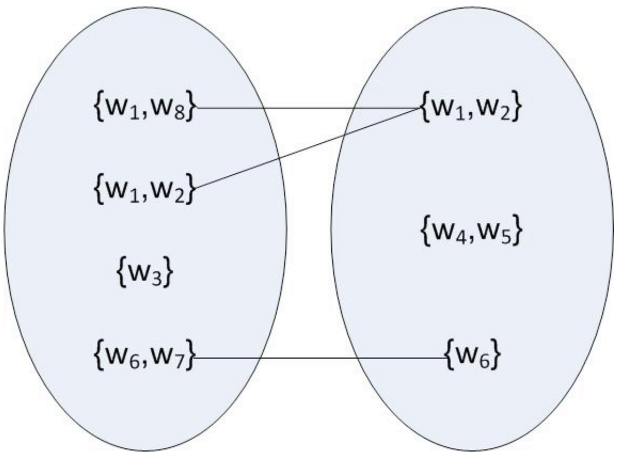

Fig. 1. The bipartite graph for Example 7.

From Fig. 1, it is straightforward to get $S_1 = \{w_1, w_2, w_8\}, S_2 = \{w_6, w_7\}$ and $S_{uncor} = \{w_3, w_4, w_5\}$.

Based on the graph perspective, given two mass functions, the partition of $S(m) \cup S(m_I)$ into the union of mutually consistent focal sets can be obtained by Algorithm 1 with a breadth-first search algorithm used to obtain maximal connected components in the bipartite graph produced by the above procedure.

Algorithm 1 Partitioning into Correlated Groups

Require: $\mathcal{F}_1$: the set of focal sets of m, $\mathcal{F}_2$: the set of focal sets of m_I.
Ensure: A maximum number of correlated groups consisting of focal sets.
Set $S_{uncor} = \emptyset$, $k = 0$;
while $\mathcal{F}_1 \neq \emptyset$ **do**
  Select a focal set A in $\mathcal{F}_1$;
  if A does not overlap with any focal sets in $\mathcal{F}_2$ **then**
    $S_{uncor} = S_{uncor} \cup A$; $\mathcal{F}_1 = \mathcal{F}_1 \setminus \{A\}$;
  else
    $k = k + 1, i = 2, S_k = A, \mathcal{F}_1 = \mathcal{F}_1 \setminus \{A\}, pre_B = S_k$;
    repeat
      Let $\mathcal{B} = \{B : B \in \mathcal{F}_i$ **and** $B \cap pre_B \neq \emptyset\}$ be the set of focal sets of m_i that intersect pre_B;
      Let $S_k = \bigcup_{B \in \mathcal{B}} B \cup S_k$; $pre_B = \bigcup_{B \in \mathcal{B}} B$;
      $\mathcal{F}_i = \mathcal{F}_i \setminus \mathcal{B}$;
      $i = 3 - i$; (repeatedly checking elements in $\mathcal{F}_1$ and $\mathcal{F}_2$)
    until S_k can not be changed any further;
  end if
end while
$S_{uncor} = \bigcup_{A \in \mathcal{F}_2} A \cup S_{uncor}$;
return $\{S_1, \cdots, S_k, S_{uncor}\}$;

Proposition 4. *Let $\mathcal{F}_m$ and $\mathcal{F}_I$ be two families of subsets of W. Algorithm 1 produces a unique partition of $S(m) \cup S(m_I)$ containing the maximum number of correlated groups.*

Proof: The proof is straightforward from the graph-theoretic point of view since the algorithm finds maximal connected components using a standard breadth-first search algorithm. □

Based on the obtained partition, we only need to consider revision inside each subset S_i. For convenience, let focal sets of m included in S_i be $A_i^1, \cdots, A_i^s \in \mathcal{F}_i$ and let $S_i^A = \cup_{k=1}^s A_i^k$. Similarly, let focal sets of m_I included in S_i be $B_i^1, \cdots, B_i^t \in \mathcal{F}_i$ and $S_i^B = \cup_{j=1}^t B_i^j$. Let $S_i^{AB} = S_i^A \cap S_i^B$. Now we aim to *flow down* the masses of B_is to their subsets based on $m(A_i)$ values.

For each B_i^j, different portions of $m_I(B_i^j)$ should flow down to its subsets. Based on the idea of Jeffrey's rule, for each subset C of B_i^j, its share of the mass value $m_I(B_i^j)$

should be proportional to the ratio of its mass value $m(C)$ to the sum of masses $m(D)$ of all the subsets $D \subseteq B_i^j$. More precisely, given a subset S_i in the partition, flowing down the mass of B_i^j can be performed by means of the following procedure:

(1) For each non-empty subset C of B_i^j, we calculate $supp(C) = \sum_{A_i^k \cap B_i^j = C} m(A_i^k)$ which is the measure of support for subset C based on A_i^ks in S_i from the viewpoint of B_i^j. That is, in Example 7, from the viewpoint of $\{w_1, w_2\}$ of m_I, it supports $\{w_1\}$ based on $\{w_1, w_8\}$ of m.
(2) For C, the flown down value from B_i^j is

$$\hat{m}^j(C) = m_I(B_i^j)\frac{supp(C)}{\sum_{\emptyset \neq D \subseteq B_i^j} supp(D)}. \tag{15}$$

Note that the denominator is clearly $Pl(B_i^j)$. This technique can be seen as a kind of conditioning on B_i^j, i.e.,

$$\frac{\hat{m}^j(C)}{m_I(B_i^j)} = \frac{supp(C)}{\sum_{D \subseteq B_i^j} supp(D)}.$$

Evidently, this equation is to some extent similar to the form of Jeffrey's rule.

Example 8. (Ex. 7 Cont') In subset S_1 in the partition: $m(\{w_1, w_8\}) = 0.2$, $m(\{w_1, w_2\}) = 0.4$ and $m_I(\{w_1, w_2\}) = 0.5$, we need to flow down $m_I(\{w_1, w_2\}) = 0.5$ to the subsets of $\{w_1, w_2\}$, i.e., $\{w_1\}$, $\{w_2\}$ and $\{w_1, w_2\}$. Here the existence of $m(\{w_1, w_2\}) = 0.4$ can be seen as a positive support for flowing down the mass value of $m_I(\{w_1, w_2\})$ to $\{w_1, w_2\}$ since $\{w_1, w_2\} \cap \{w_1, w_2\} = \{w_1, w_2\}$. Similarly, the existence of $m(\{w_1, w_8\}) = 0.2$ supports $\{w_1\}$ from the viewpoint of $m_I(\{w_1, w_2\})$ since $\{w_1, w_2\} \cap \{w_1, w_8\} = \{w_1\}$. Therefore, we get $supp(\{w_1, w_2\}) = 0.4$ and $supp(\{w_1\}) = 0.2$, and hence $\hat{m}(\{w_1, w_2\}) = \frac{1}{3}$ and $\hat{m}(\{w_1\}) = 1/6$.

After allocating all fractions of $m_I(B_i^j)$, $1 \leq j \leq t$, we are able to sum up all the masses that each subset C receives. This leads to the following definition of a revision operator.

Definition 9. Let m and m_I be two mass functions and $\{S_1, \cdots, S_k, S_{uncor}\}$ be the partition of $S(m) \cup S(m_I)$ obtained from Algorithm 1. For any subset C, if $C \subseteq S_i$, then let t be the number of focal sets of m_I contained in S_i and $\hat{m}^j(C)$ be defined in Eq. (15), $1 \leq j \leq t$. Then a revision operator for mass functions is defined as $\hat{m} = m \circ_a m_I$ such that

$$\hat{m}(C) = \begin{cases} \sum_{j=1}^{t} \hat{m}^j(C) & \text{for } C \subseteq S_i, \\ m_I(C) & \text{for } C \subseteq S_{uncor}. \end{cases} \tag{16}$$

From Algorithm 1, set C does not intersect any focal set included in another element of the partition, so the flowing down process for other elements of the partition does not affect the revised mass value of C. Hence $\hat{m}(C)$ obtained in Def. 9 is indeed the final result for C. It should be clear that the Jeffrey-Dempster revision and the revision operator of Def. 9

are the same thing. The latter is just another way of computing the former. This finding is significant since these two revision strategies are from different perspectives.

Proposition 5. *For any two mass functions m and m_I over W, we have $m \circ_{JD} m_I = m \circ_a m_I$.*

Proof: As said earlier,
$\sum_{C \subseteq B_i^j} supp(C) = \sum_{C \subseteq B_i^j} \sum_{A_i^k \cap B_i^j = C} m(A_i^k) = \sum_{A_i^k \cap B_i^j \neq \emptyset} m(A_i^k) = Pl(B_i^j)$.

If $Pl(B_i^j) = 0$, then B_i^j does not intersect any focal set of m, based on Algorithm 1, B_i^j is in S_{uncor} (note that the converse is also right, i.e., if a focal set B of m_I is in S_{uncor}, then $Pl(B) = 0$), hence the mass value of B_i^j remains unchanged after revision. This is equivalent to the following condition in Def. 7:

$$Pl(B) = 0 \Longrightarrow \begin{cases} \sigma_m(A, B) = 0 \text{ for } A \neq B, \\ \sigma_m(A, B) = 1 \text{ for } A = B. \end{cases}$$

If $Pl(B_i^j) > 0$, then B_i^j is not in S_{uncor}. Hence $\forall l$, B_i^l is not in S_{uncor}, we have $Pl(B_i^l) > 0$.

$$\begin{aligned}
\hat{m}(C) &= \sum_{j=1}^{t} \hat{m}^j(C) = \sum_{j=1}^{t} m_I(B_i^j) \frac{supp(C)}{Pl(B_i^j)} \\
&= \sum_{j=1}^{t} m_I(B_i^j) \frac{\sum_{\forall A_i^k, A_i^k \cap B_i^j = C} m(A_i^k)}{Pl(B_i^j)} \\
&= \sum_{j=1}^{t} \sum_{\forall A_i^k, A_i^k \cap B_i^j = C} \frac{m(A_i^k)}{Pl(B_i^j)} m_I(B_i^j) \\
&= \sum_{\forall A, B, A \cap B = C} \frac{m(A)}{Pl(B)} m_I(B)
\end{aligned}$$

Therefore, we have $\circ_{JD} = \circ_a$. □

6. Properties of the revision rule for mass functions

In this section, we present some properties of the Jeffrey-Dempster revision. First, we discuss the relationship between our revision rule and Dempster's rule of conditioning.

Proposition 6. *Let m_I be a mass function such that $m_I(E_I) = 1$, then $m \circ_{JD} m_I$ comes down to Dempster conditioning if $Pl(E_I) > 0$ and $m \circ_{JD} m_I = m_I$ otherwise.*

Proof: If $Pl(E_I) = 0$, then it is obvious. If $Pl(E_I) > 0$, then $\forall A \neq \emptyset, \hat{m}(A) = \sum_{B \subseteq W} m(A|B) m_I(B) = m(A|E_I) \cdot 1 = m(A|E_I)$. □

We can show that the vacuous mass function plays no role in revision: bringing empty information leads to no change. Likewise, if there is no prior information, the input information is accepted as such.

Proposition 7. *Let m be a mass function and m_W be such that $m_W(W) = 1$, then we have $m \circ_{JD} m_W = m_W \circ_{JD} m = m$.*

Proof: $m \circ_{JD} m_W$ coincides with Dempster conditioning on W and $m(\cdot|W) = m$. For $m_W \circ_{JD} m$, we have $\forall A, W \cap A = A$ and $Pl_W(A) = 1$, hence from Def. 7, we get $\forall C,\ (m_W \circ_{JD} m)(C) = \sum_{W \cap C = C} m_W(W)m(C) = m(C)$. Hence $m_W \circ_{JD} m = m$. □

Furthermore, we can also prove that if prior beliefs and new evidence are in total conflict, then the revision result is simply the latter, thus generalizing the AGM revision rule for epistemic states.

Proposition 8. *Let m and m_I be two mass functions such that $S(m) \cap S(m_I) = \emptyset$, then we have $m \circ_{JD} m_I = m_I$.*

Proof Since $S(m) \cap S(m_I) = \emptyset$, we have for any focal set B of m_I, $Pl(B) = 0$, and if $Pl(B) > 0$, then $m_I(B) = 0$, hence from Eq. (10), we have $\forall C,\ (m \circ_{JD} m_I)(C) = \sum_{A \cap B = C, Pl(B) > 0} \frac{m(A)m_I(B)}{Pl(B)} + I_{\{Pl(C)=0\}} m_I(C) = m_I(C)$. □

Example 9. Let $W = \{w_1, ..., w_5\}$ and m be such that $m(\{w_1\}) = 0.4, m(\{w_1, w_2\}) = 0.6$, m_I be such that $m_I(\{w_3, w_4\}) = 0.2, m_I(\{w_3, w_5\}) = 0.4, m_I(\{w_5\}) = 0.4$, then we have $m \circ_{JD} m_I = m_I$.

Proposition 9. *Let m and m_I be two strongly consistent mass functions, i.e., such that $\forall A \in \mathcal{F}_m, B \in \mathcal{F}_I, A \cap B \neq \emptyset$, then $m \circ_{JD} m_I$ comes down to Dempster rule of combination.*

Proof Since $\forall A \in \mathcal{F}_m, B \in \mathcal{F}_I, A \cap B \neq \emptyset$, it must be $\forall B \in \mathcal{F}_I$, $Pl(B) = 1$. Hence, $\forall C \neq \emptyset$, we have

$$\begin{aligned}(m \circ_{Demp} m_I)(C) &= \frac{\sum_{A \cap B = C} m(A)m_I(B)}{\sum_{A' \cap B' \neq \emptyset} m(A')m_I(B')} \\ &= \sum_{A \cap B = C} m(A)m_I(B) \\ &= \sum_{A \cap B = C} \frac{m(A)m_I(B)}{Pl(B)} \\ &= \hat{m}(C).\end{aligned}$$

□

There is no need for renormalization factor in Dempster rule then. It corresponds to an expansion as the input information does not contradict the output.

Next we show that our revision rule generalizes Jeffrey's rule of combination:

Proposition 10. *If m is a Bayesian mass function and m_I is a partitioned mass function, then $m \circ_{JD} m_I = m \circ_J m_I$.*

Proof Let the focal sets of m_I be $U_1, \cdots, U_n$ which is a partition of W. For any w, let $w \in U_i$.

If $Pl(U_i) = 0$, then both $\circ_{JD}$ and Jeffrey's rule can only give $\hat{m}(U_i) = m_I(U_i)$.

If $Pl(U_i) > 0$, then we need to prove that $(m \circ_{JD} m_I)(w) = \frac{m(w)}{Pl(U_i)} m_I(U_i)$ which is exactly what is defined in Def. 7. □

For iterated revision, we have the following result.

Proposition 11. *Let $m, m_I, m_{I'}$ be three mass functions on W. If both m_I and $m_{I'}$ are partitioned mass functions while $m_{I'}$ corresponds to a finer partition than m_I, then we have $(m \circ_{JD} m_I) \circ_{JD} m_{I'} = m \circ_{JD} m_{I'}$.*

Proof: Let $\hat{m} = m \circ_{JD} m_I$, $\hat{\hat{m}} = (m \circ_{JD} m_I) \circ_{JD} m_{I'}$ and $\check{m} = m \circ_{JD} m_{I'}$. Here we need show $\forall V$, $\check{m}(V) = \hat{\hat{m}}(V)$. Note that we only need to consider V which is a subset of a focal set of $m_{I'}$, otherwise from Def. 7, we have $\check{m}(V) = \hat{\hat{m}}(V) = 0$.

For any subset V of a focal set of $m_{I'}$, $\check{m}(V) = \sum_{A \cap B = V} \check{\sigma}_m(A, B) m_{I'}(B)$ from Def. 7. Obviously, for each B in the former equation, $V \subseteq B$. Then since the focal sets of $m_{I'}$ form a partition, there cannot be two different focal sets of $m_{I'}$ that both contain V (as different focal sets of $m_{I'}$ are disjoint). So there is exactly one focal set of $m_{I'}$ containing V, call this focal set B_V. So, we have

$$\check{m}(V) = m_{I'}(B_V) \sum_{A \cap B_V = V} \check{\sigma}_m(A, B_V).$$

Likewise, $\hat{\hat{m}}(V) = m_{I'}(B_V) \sum_{A \cap B_V = V} \hat{\hat{\sigma}}_{\hat{m}}(A, B_V)$. Now we discuss two subcases.

- If $Pl(B_V) = 0$, based on Def. 7, then for any $V \subseteq B_V$, $\check{m}(V) = 0$ if $V \neq B_V$ and $\check{m}(B_V) = m_{I'}(B_V)$ if $V = B_V$.

  Let D_V be the unique focal set of m_I, such that $V \subseteq D_V$; by construction $V \subseteq B_V \subseteq D_V$.

  - If $Pl(D_V) = 0$, then for any $V \subseteq D_V$, $\hat{m}(V) = 0$ if $V \neq D_V$ and $\hat{m}(D_V) = m_I(D_V)$ if $V = D_V$. It is clear that there is no focal set of $\hat{m}$ inside D_V, but for D_V itself. Then $\hat{\hat{m}}(B_V) = m_{I'}(B_V)$, while $\hat{\hat{m}}(V) = 0$ if $V \subset B_V$, i.e., $\hat{\hat{m}} = \check{m}$ inside B_V (including B_V itself).
  - If $Pl(D_V) > 0$, then $\forall A, C, A = C \cap D_V$ and $m(C) > 0$ (such A, C indeed exist), it should be:

    $$\hat{m}(A) = m_I(D_V) \sum_{C : C \cap D_V = A} \frac{m(C)}{Pl(D_V)}.$$

    Note that while $A \subseteq D_V$, it must hold that $A \cap B_V = \emptyset$ since $Pl(B_V) = 0$. So there cannot be, via this construction, a focal set of $\hat{m}$ that intersects B_V. So, $\hat{Pl}(B_V) = 0$ in any case. Then, based on Def. 7, immediately we have $\hat{\hat{m}}(V) = 0$ if $V \neq B_V$ and $\hat{\hat{m}}(B_V) = m_{I'}(B_V)$, which implies $\hat{\hat{m}}(V) = \check{m}(V)$.

- If $Pl(B_V) > 0$, then we have

$$\begin{aligned} \hat{\hat{m}}(V) &= \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} \sum_{A \cap B_V = V} \hat{m}(A) \\ &= \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} \sum_{A \cap B_V = V} \sum_{C \cap D = A} \hat{\sigma}_m(C, D) m_I(D), \end{aligned}$$

since $V \subseteq A \subseteq D$, likewise it must be that D can only be a particular focal set of m_I s.t. $B_V \subseteq D$. We denote it as D_V, and we have $Pl(D_V) \geq Pl(B_V) > 0$. Hence we have

$$\begin{aligned}\hat{\hat{m}}(V) &= \frac{m_{I'}(B_V)}{\hat{P}l(B_V)} \sum_{A \cap B_V = V} \sum_{C \cap D_V = A} \hat{\sigma}_m(C, D_V) m_I(D_V) \\ &= \frac{m_{I'}(B_V)}{\hat{P}l(B_V)} \sum_{A \cap B_V = V} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C \cap D_V = A} m(C) \\ &= m_{I'}(B_V) \frac{\sum_{A \cap B_V = V} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C \cap D_V = A} m(C)}{\sum_{A' \cap B_V \neq \emptyset} \frac{m_I(D_{A'})}{Pl(D_{A'})} \sum_{C' \cap D_{A'} = A'} m(C')},\end{aligned}$$

where we rewrite $\hat{P}l(B_V)$ in the last step and $D_{A'}$ is the unique focal set of m_I which contains A'. As $A' \cap B_V \neq \emptyset$ and $A' \subseteq D_{A'}$, we similarly have $B_V \subseteq D_{A'}$, but only one focal set of m_I contains B_V, hence it should be $D_{A'} = D_V$ for any A'. So we have

$$\begin{aligned}\hat{\hat{m}}(V) &= m_{I'}(B_V) \frac{\sum_{A \cap B_V = V} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C \cap D_V = A} m(C)}{\sum_{A' \cap B_V \neq \emptyset} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C' \cap D_V = A'} m(C')} \\ &= m_{I'}(B_V) \frac{\sum_{A \cap B_V = V} \sum_{C \cap D_V = A} m(C)}{\sum_{A' \cap B_V \neq \emptyset} \sum_{C' \cap D_V = A'} m(C')} \\ &= m_{I'}(B_V) \frac{\sum_{C \cap B_V = V} m(C)}{\sum_{C' \cap B_V \neq \emptyset} m(C')} \\ &= m_{I'}(B_V) \frac{\sum_{C \cap B_V = V} m(C)}{Pl(B_V)} \\ &= \check{m}(V).\end{aligned}$$

□

Example 10. Let m be such that $m(\{w_1\}) = 0.3, m(\{w_1, w_2\}) = 0.3, m(\{w_3\}) = 0.1$,$m(\{w_4\}) = 0.3$, m_I be such that $m_I(\{w_1, w_3\}) = 0.6, m_I(\{w_2, w_4\}) = 0.4$ and $m_{I'}$ be such that $m_{I'}(\{w_1, w_3\}) = 0.2, m_{I'}(\{w_2\} = 0.3, m_{I'}(\{w_4\}) = 0.5$, then we have $\hat{\hat{m}} = m \circ_{JD} m_I \circ_{JD} m_{I'}$ with $\hat{\hat{m}}(\{w_1\}) = \frac{6}{35}, \hat{\hat{m}}(\{w_2\}) = 0.3, \hat{\hat{m}}(\{w_3\}) = \frac{1}{35}, \hat{\hat{m}}(\{w_4\}) = 0.5$. $\check{m} = m \circ_{JD} m_{I'}$ has the same set of focal sets and corresponding mass values.

Darwiche and Pearl [6] proposed four postulates on iterated belief revision, the first of which (written in set-theoretic notation) is

C1: If $E_I \subseteq F_I$, then $(E \circ F_I) \circ E_I \equiv E \circ E_I$

Proposition 11 can be seen as a generalization of the above iterated belief revision postulate, where the condition $E_I \subseteq F_I$ is changed into a simple refinement condition on input partitions. This result includes the case when the partitioned mass function $m_{I'}$ is a specialisation of the partitioned mass m_I, but it does not require it.

The second postulate reads:

C2: $(E \circ \overline{E}_I) \circ E_I \equiv E \circ E_I$

The complement $\overline{m}$ of a mass function m is obtained as $\forall E \subseteq W, \overline{m}(E) = m(\overline{E})$. If $\mathcal{F}_I$ is a partition made of E_I and its complement $\overline{E}_I$, $\overline{m}_I$ just exchanges the weights of m_I, keeping the same partition. Hence if we substitute m_I for $\overline{m}_{I'}$, we are still in the conditions requested by Proposition 11 which thus applies, and generalizes **C2** as well. Note that this remark does not apply beyond binary partitions. If m_I is a partitioned mass, $\overline{m}_I$ is not (the complements of elements of a partition do not make a partition, in general). So we cannot say that $(m \circ_{JD} \overline{m}_{I'}) \circ_{JD} m_{I'} = m \circ_{JD} m_{I'}$ using this proposition, beyond binary partitions, since Proposition 11 has a restricted scope to inputs of the form of partitioned mass functions. Extending this result beyond this case is a matter of further research.

7. Conclusion

Although belief revision in probability theory is fully studied, revision strategies in evidence theory have seldom been addressed. In this paper, we have investigated the issue of revision strategies for mass functions so as to generalize Jeffrey's rule from probability to belief functions. We have tried to express two basic requirements for the revision of belief functions, that seem to be universally valid as revision principles, and hold for Jeffrey's rule: success postulate and minimal change. Moreover, since the natural revision rule of a belief function by a sure fact is Dempster rule of conditioning, we required that the appropriate revision rule also extends Dempster rule of conditioning. Moreover, in case of inconsistency we have proposed using a revision method in the spirit of the AGM theory, respecting the success postulate. We have been led to first extend Dempster rule of conditioning to this case.

We have surveyed the existing literature, where several extensions of Jeffrey's revision rule have been proposed. Only one of them due to Ichiashi and Tanaka, generalizes Dempster conditioning. But none of them can bear total conflict with any focal set of the input information. As a consequence we have extended the definition in order to cover this case and called the obtained revision rule Jeffrey-Dempster revision. Its peculiarity is to accept any kind of belief functions as input, contrary to other past proposals by Smets and Halpern. The paid price is that the Jeffrey-style minimal change postulate holds in a strong form only for special kinds of inputs. However the success postulate is respected as the posterior mass is a specialisation of the input. In case of strong consistency with the prior belief function, the revision rule coincides with Dempster rule of combination, which generalizes the reduction of the AGM revision to a symmetric conjunctive merging called expansion when the input is consistent with the prior information. We have proposed an algorithm to compute the result of a Jeffrey-Dempster revision.

We have seen that when the inputs are successive partitioned mass functions, two known postulates of iterated belief revision due to Darwiche and Pearl still hold. This point suggests further research should be carried out to study more properties of iterating the Jeffrey-Dempster revision rule. Moreover, various proposals exist for the revision of ranked epistemic states [20,2], following Darwiche and Pearl. In these papers, revision patterns similar to Jeffrey's rule are considered, in relation to existing iterated revision postulates and existing revision rules in the ordinal setting. Benferhat et al. [2] consider counterparts to Jeffrey's rule in qualitative and quantitative possibility theory. Ma et al. [20] try to find a revision framework common to possibility and probability theories, bridging the gap between Darwiche-Pearl ordinal framework and probability kinematics, using the idea of partial epistemic state. Studying the relevance of these postulates for the belief function setting is also a line of further research. Finally, we may try to find revision postulates that fully characterise the Jeffrey-Dempster revision rule advocated here.

References

1. C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50:510–530, 1985.
2. S. Benferhat, D. Dubois, H. Prade, and M. A. Williams. A framework for iterated belief revision using possibilistic counterparts to Jeffrey's rule. *Fundamenta Informaticae*, 99(2):147–168, 2010.
3. S. Benferhat, S. Konieczny, O. Papini, and R. P. Pérez. Iterated revision by epistemic states: Axioms, semantics and syntax. In *Proc. of ECAI'00*, pages 13–17, 2000.
4. R. Booth and T. Meyer. Admissible and restrained revision. *Journal of Artificial Intelligence Research*, 26:127–151, 2006.
5. H. Chan and A. Darwiche. On the revision of probabilistic beliefs using uncertain evidence. *Artificial Intelligence*, 163:67–90, 2005.
6. A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.
7. A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Statistics*, 28:325–339, 1967.
8. Domotor, Z.: Probability kinematics and representation of belief change. *Philosophy of Science* **47** (1980) 284–403
9. D. Dubois. Three scenarios for the revision of epistemic states. *Journal of Logic and Computation*, 18(5):721–738, 2008.
10. D. Dubois, S. Moral, and H. Prade. Belief change rules in ordinal and numerical uncertainty theories. In D. Gabbay and P. Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 3, pages 311–392. Kluwer Academic Pub., 1998.
11. D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *Int. J. Gener. Sys.*, 12(3):193–226, 1986.
12. D. Dubois and H. Prade. Updating with belief functions, ordinal conditional functions and possibility measures. In P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence*, volume 6, pages 311–329. Elsevier, Amsterdam, 1991.
13. J. Y. Halpern. *Reasoning about Uncertainty*. The MIT Press, Cambridge, Massachusetts, London, England, 2003.
14. H. Ichihashi and H. Tanaka. Jeffrey-like rules of conditioning for the Dempster-Shafer theory of evidence. *Int. J. of Approximate Reasoning*, 3:143–156, 1989.
15. R. C. Jeffrey. *The Logic of Decision*. McGraw-Hill, New York, 1965. 2nd edition: University of Chicago Press, Chicago, IL, 1983. Paperback Edition, 1990.
16. Y. Jin and M. Thielscher. Iterated belief revision, revised. *Artificial Intelligence*, 171:1–18, 2007.
17. H. Katsuno and A. O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.

18. G. Kern-Isberner. *Conditionals in nonmonotonic reasoning and belief revision*, volume 2087 of *Lecture Notes in Artificial Intelligence*. Springer, Berlin, 2001.
19. J. Ma and W. Liu. Modeling belief change on epistemic states. In *Proc. of 22th FLAIRS*, pages 553–558. AAAI Press, 2009.
20. J. Ma, W. Liu, and S. Benferhat. A belief revision framework for revising epistemic states with partial epistemic states. In *Proc. of AAAI'10*, pages 333–338, 2010.
21. J. Ma, W. Liu, D. Dubois, and H. Prade. Revision rules in the theory of evidence. In *Proc. of ICTAI'10*, pages 295–302, 2010.
22. A. C. Nayak, M. Pagnucco, and P. Peppas. Dynamic belief revision operators. *Artificial Intelligence*, 146:193–228, 2003.
23. Z. Pawlak. *Rough Sets - Theoretical Aspects of Reasning about Data*. Kluwer Academic Publ., Dordrecht, 1991.
24. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
25. G. Shafer. Jeffrey's rule of conditioning. *Philosophy of Science*, 48(3):337–362, 1981.
26. P. Smets. Jeffrey's rule of conditioning generalized to belief functions. In *Proc. of UAI*, pages 500–505, 1993.
27. P. Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31(1-2):1–30, 2002.
28. P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4):387–412, 2007.
29. W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In W. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume 2, pages 105–134. Kluwer Academic Publishers, 1988.
30. C. Wagner. Consensus for belief functions and related uncertainty measures. *Theory and Decision*, 26:295–304, 1989.
31. P.M. Williams. Bayesian conditionalization and the principle of minimum information. *British J. for the Philosophy of Sciences*, 31:131–144, 1980.
32. R. Yager and L. Liping, editors. *Classic works of the Dempster-Shafer theory of belief functions*. Springer, Berlin, 2008.