Revision Rules in the Theory of Evidence

Jianbing Ma, Weiru Liu
Queen's University of Belfast,
Belfast, UK, BT7 1NN
Email: {jma03,w.liu}@qub.ac.uk

Didier Dubois, Henri Prade
IRIT, Université Paul Sabatier
118 Route de Narbonne 31062 Toulouse, Cedex 9, France
Email: {dubois,prade}@irit.fr

***Abstract*—Combination rules proposed so far in the Dempster-Shafer theory of evidence, especially Dempster rule, 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 a 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. Although belief revision is already an important subfield of artificial intelligence, so far, it has been little addressed in evidence theory. In this paper, we define the notion of *revision* for the theory of evidence and propose several different revision rules, called the inner and outer revisions, and a modified adaptive outer revision, which better corresponds to the idea of revision. Properties of these revision rules are also investigated.**

I. Introduction

Dempster-Shafer theory of evidence (DS theory) [1], [2], [3], 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., [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], etc.) 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. However, when a source is less important than another, the corresponding piece of information is discounted and the above rules still apply. Usually such reliability information is not contained in the input evidence. So, 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 receives 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 possibility of unequal reliabilities of sources. Let us look at the following example of revision problem (adapted from [14]).

Example 1: 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 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?*

Evidently, the input evidence has priority over the prior belief, hence after revision, we should conclude that the cloth color is more possibly green. However, 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 is the most plausible color. The counterintuitive result produced by the combination rules 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. Two principles 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 (new 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 (as in the example) or it is meant to qualify the reliability of the (otherwise crisp) input information [15]. In the latter case, the success postulate is questionable In this paper, we focus on the former, where new uncertain evidence is accepted and serves as a constraint on the resulting belief state.

In the field of artificial intelligence, revision strategies are extensively studied in the contexts of logical theory revision and probability kinematics. Belief revision [16], [17], [18] is a framework for characterizing the process of belief change in order to revise the agent's current beliefs to accommodate new evidence and to reach a new consistent set of beliefs. Probability kinematics [14] considers how a prior probability measure should be changed based on a new probability measure on a coarser frame, which should be preserved after revision. Jeffrey's rule [14] is the most commonly used rule for achieving this objective.

Within the scope of DS theory, the only revision rule is one first addressed in [19] and later re-formulated in [20], as a counterpart to Jeffrey's rule. However, this rule only accepts 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 $Bel(U_i) = \alpha_i$ and $\sum \alpha_i = 1$. Ideally, a belief function based revision rule should accept evidence in the form of a general belief function definable on 2^W rather than just on a partition of W. Furthermore, like the original Jeffrey's rule, this rule requires that $Pl_0(U_i) > 0$ whenever $\alpha_i > 0$ where Pl_0 is the plausibility function for the prior epistemic state. In other words, this rule requires that new evidence be not *in conflict* with the agent's current beliefs, which restricts the application of this rule. As a belief function is defined by means of a probability distribution on 2^W called a mass function, it is natural to express the revision rule in terms of mass functions.

In [21], [22] (and later in [15]), a revision rule for mass functions was proposed and dubbed *plausible conditioning* in [23]. Two additional revision rules were also proposed in the same paper (re-examined in [24]), dubbed *credible* and *possible conditioning*. However, these revision rules suffer from the same drawback as the one in [20], that is, the new evidence must be consistent with the prior belief in order to be applied. In [25], the need for investigating revision strategies for mass functions was addressed but no concrete revision rule was proposed. In [26], Jeffrey's rule was studied in DS theory which showed that it could be seen as a special case of Dempster's combination rule.

In this paper, we first discuss the form a mass-function-based revision rule (or operator) should take in order to comply with the success and the minimal change postulates. We define a family of mass-function-based revision rules, dubbed *inner* and *outer revision*, *modified outer revision*, and *adaptive revision*. We also prove the equivalence between the modified outer revision and the adaptive revision. This result is significant since these two revision rules start from different perspectives, and in some sense, the adaptive revision can be seen as a justification for the modified outer revision. Finally, we prove that our revision rules generalize both Jeffrey's rule and Halpern's rule.

The rest of the paper is organized as follows. We give some preliminaries in Section II. In Section III, we discuss the principles a revision rule on mass functions shall satisfy. We then propose a set of revision rules in Section IV. Section V contains several rational properties of the revision rules. In Section VI, we conclude the paper.

II. Preliminaries

Let W be a set of possible worlds (or the frame of discernment). A mass function is a mapping $m : 2^W \to [0, 1]$ such that $\sum_{A \subseteq W} m(A) = 1$ and $m(\emptyset) = 0$. A is called a focal set of m if $m(A) > 0$. Let $S(m)$ denote the support of m, i.e. the union of the focal sets, that is, $S(m) = \bigcup_{i=1}^{n} A_i$ where A_is are focal sets of m.

A mass function m is called *Bayesian* iff all its focal sets are singletons. A mass function m is called *Partitioned* iff its focal sets $A_1, \cdots, A_k$ form a partition of W, i.e., $A_1 \cup \cdots \cup A_k = W$ and $A_i \cap A_j = \emptyset$ for $i \neq j$. Given m, its corresponding belief function $Bel : 2^W \to [0, 1]$ is defined as $Bel(B) = \sum_{A \subseteq B} m(A)$ and its corresponding plausibility function $Pl : 2^W \to [0, 1]$ is defined as $Pl(B) = 1 - Bel(\overline{B})$.

There are several conditioning methods for belief/plausibility functions [15]. The following, called Dempster conditioning is the most commonly used one [20].

Definition 1: Let U be a subset of W such that $Pl(U) > 0$, then conditioning belief/plausibility functions on U can be defined as

$$Pl(V|U) = \frac{Pl(V \cap U)}{Pl(U)},$$

$$Bel(V|U) = 1 - Pl(\overline{V}|U) = \frac{Pl(U) - Pl(\overline{V} \cap U)}{Pl(U)}.$$

It is a revision rule that transfers the mass bearing on each subset V to its subset $V \cap U$, thus sanctioning the success postulate. Moreover, resulting masses bearing on non-empty sets are renormalized via simple division by $Pl(U)$, i.e. do not change in relative value, which expresses minimal change.

Definition 2: (Specialization [22]) We write $m \sqsubseteq m'$ ($\sqsubseteq$ is typically called *s-ordering*) iff 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 specialization matrix*, and m is said to be a specialization of m'. Specialization is an extension of set-inclusion to random sets.

Example 2: 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 specialization of m'. It can be considered as m' *flows a mass value* 0.2 of $\{w_1, w_2\}$ to $\{w_1\}$ (i.e., $\sigma(\{w_1\}, \{w_1, w_2\}) = 0.4$), a mass value 0.2 of $\{w_1, w_2\}$ to $\{w_2\}$ and a mass value 0.3 of $\{w_2, w_3\}$ to $\{w_2\}$.

$m \setminus m'$	{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

Table 1: The Matrix $\sigma(A, B)$.

Notation $\{1\}$ etc stands for subset $\{w_1\}$ etc. Value $2/5$ on the 1st row shows the ratio of the mass on subset $\{w_1, w_2\}$ from m' that will flow down to subset $\{w_1\}$, namely $\sigma(\{w_1\}, \{w_1, w_2\}) = 2/5$.

A. Revision Rules

Jeffrey's probability kinematics rule was introduced as follows.

Definition 3: Let P be a probability measure over W denoting the prior epistemic state and P_I be a probability measure over a partition $\{U_1, \cdots, U_n\}$ of W denoting the input evidence. Let $\circ_{Jef}$ denote Jeffrey's rule. Then :

$P \circ_{Jef} P_I(w) = \sum_{i=1}^{n} P_I(U_i)P(w|U_i)$.

Note that for $w \in U_i$, the above equation can be simplified as $P \circ_{Jef} P_I(w) = P_I(U_i)\frac{P(w)}{P(U_i)}$. That is, the revised probability of element w within each U_i is the same as their prior probability in relative value.

Similarly, Halpern's belief function revision rule is defined as follows [20].

Definition 4: Let Bel be a belief function over 2^W denoting the prior epistemic state and Bel_I be a belief function over a partition $\{U_1, \cdots, U_n\}$ of W denoting the input evidence. Let $\circ_{Hal}$ denote Halpern's revision rule, then we have

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

In [23], alternative revision rules are defined as follows. Let m denote the prior mass function and m_I the input evidence.

- **Credible** $m_{cr}(A|m_I) = \sum_{A \subseteq B} \frac{m(A)m_I(B)}{Bel(B)}$ where for any focal set B of m_I, $Bel(B) > 0$.
- **Possible** $m_{po}(A|m_I) = \sum_{A \cap B \neq \emptyset} \frac{m(A)m_I(B)}{Pl(B)}$ where for any focal set B of m_I, $Pl(B) > 0$.
- **Plausible** $m_{pl}(A|m_I) = \sum_{C \cap B = A} \frac{m(C)m_I(B)}{Pl(B)}$ where m_I must satisfy $m_{pl}(\emptyset|m_I) = 0$ and for any focal set B of m_I, $Pl(B) > 0$.

Obviously, these rules are only applicable when m and m_I are highly consistent.

III. Principles of Mass Function based Revision

Now we discuss how the two general revision principles can be applied to mass-function-based revision. Let $\hat{m} = m \circ m_I$ be the posterior mass function, and $\circ$ be a revision operator which associates a resultant mass function $\hat{m}$ with two given mass functions, one represents the prior belief state (m) and the other new evidence (m_I). Moreover, we focus on rules that generalize Jeffrey's probability kinematics, and Dempster rule of conditioning.

Success postulate through specialization: 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. But how can we define the notion of *implication* between mass functions?

In propositional logics, when we write $\phi \vdash \psi$ (ϕ implies ψ), we in fact state that ϕ is more specialized than ψ, e.g., *a grey bird* (i.e., $\phi = g \wedge b$) is a more specialized concept than *a bird* ($\psi = b$). It corresponds to inclusion between sets of models $A = Mod(\phi)$ and $B = Mod(\psi)$. Hence it is natural for us to use the notion of specialization between two mass functions (Def. 2). In fact, specialization between two mass functions can be equally seen as a generalization of implication in propositional logic. That is, we have

Proposition 1: Let m and m' be two mass functions defined on a set of possible worlds W s.t. $m(Mod(\phi)) = 1$ and $m'(Mod(\psi)) = 1$. If m is a specialization of m', then $\phi \models \psi$ (i.e. $A \subseteq B$).

Note that $m(Mod(\phi)) = 1$ (resp. $m'(Mod(\psi)) = 1$) is a mass function representation stating that *proposition* ϕ *(resp.* ψ*) is true in* m *(resp.* m'*)*. Therefore the success postulate in evidence theory reads : $\hat{m} \sqsubseteq m_I$.

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 specialization 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 specialization of m and m is at minimal distance from itself among all specializations 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 $\triangle(\phi, \phi) = \phi$ ($\triangle$ is a belief merging operator acting on logic formulas). Similarly, in belief revision, $\phi \circ_r \phi = \phi$ (if $\circ_r$ is a belief revision operator), but for mass functions, we do not necessarily expect $m \circ m = m$, instead, 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 this 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.

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 strictly satisfies the minimal change principle in the sense that $P \circ P = P$, which involves no confirmation effect.

Generalization of Dempster conditioning: Finally, if $m_I(A) = 1$ for some subset $U \subseteq W$ such that $Pl(U) > 0$, then $\hat{Pl} = Pl(\cdot|U)$, 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 specializes 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 epistemic state. In the logical setting revision becomes symmetric, and in the evidence setting, revision sounds asymmetric due to the difference of nature of the input information and the prior epistemic state.

IV. Mass Function based Revision operators

A. Inner and 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 belief and plausibility measures [1].

An *inner* revision operator is defined as follows.

Definition 5: Let m and m_I be two mass functions over W and let $\circ_i$ be an inner revision operator that revises m with m_I, then the revision result 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 need to *flow down* some of the mass value of every positive $m_I(B)$ to subsets $A \subseteq B$. Furthermore, the *flowing-down* portion $\sigma_i(A,B)$ of $m_I(B)$ should be proportional to $m(A)$ across all subsets of B (hence $\sigma_i(A,B) = \frac{m(A)}{Bel(B)}$). If m does not consider B possible at all ($Bel(B) = 0$), then value $m_I(B)$ should be totally allocated to B. By construction, the inner revision operator is a specialization of m_I, that preserves as much information from m as possible. It is easy to prove that $m \circ_i m_I$ is a mass function, i.e., $\sum_{A \subseteq W} m \circ_i m_I(A) = 1$ and $m \circ_i m_I(\emptyset) = 0$.

An *outer* revision operator is defined as follows.

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

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

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 $Pl(B)$). Note that for outer revision, the revised result is not necessarily a specialization of m_I, but this change rule naturally appears by duality. Similarly, it is easy to prove that $m \circ_o m_I$ is a mass function.

The inner (resp. outer) revision rule extends the credible (resp. possible) conditioning rules to the revision situation where new evidence totally conflicts with prior beliefs. That is, revision can be done even when $Pl(B) = 0$.

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 $m_{in} = m \circ_i m_I$ where

$$m_{in}(\{w_1\}) = m_I(\{w_1\}) \frac{m(\{w_1\})}{Bel(\{w_1\})} + m_I(\{w_1, w_2\}) \frac{m(\{w_1\})}{Bel(\{w_1, w_2\})} = 0.48,$$

$$m_{in}(\{w_1, w_2\}) = m_I(\{w_1, w_2\}) \frac{m(\{w_1, w_2\})}{Bel(\{w_1, w_2\})} = 0.32,$$

$$m_{in}(\{w_4\}) = 0.2.$$

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

However, these two rules 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 $m(\{w_1\}) = 1$. However, from inner revision, the revised result is $m_{in}(\{w_1, w_3\}) = 1$ whilst from outer revision, the revised result is $m_{out}(\{w_1, w_2\}) = 1$. Both revision results are not fully agreeable with intuitions.

B. A modified outer revision operator

As mentioned earlier, the result of outer revision is not necessarily a specialization of the mass function representing new evidence, hence strictly speaking, from the viewpoint of Section III, the outer revision is in fact not a revision. In this section, we define a modified outer revision that yields a specialization of the new evidence.

Definition 7: Let m and m_I be two mass functions over W. Operator $\circ_m$ is a modified outer revision operator that revises m with m_I s.t. for any $C \neq \emptyset$, $m \circ_m m_I(C) = \sum_{A \cap B = C} \sigma_m(A,B) m_I(B)$ 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 A ($A \cap B \neq \emptyset$), we flow down this portion to $A \cap B$. This modification makes the revision result truly a specialization of m_I. Also, the modified outer revision extends the plausible conditioning rule [21], [23] to situations where $Pl(B) = 0$.

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

C. Adaptive revision

The inner and outer revision rules are described using Bel and Pl functions that are inner and outer measures respectively. We should however describe a mass-function-based revision rule, in terms of the mass function only. In this subsection we propose such an adaptive revision rule for mass functions. It also overcomes the weaknesses of inner and outer revision.

Intuitively, for mass-function-based revision, only *correlated* information needs to be taken into account. By correlated information, we mean focal sets of m_I that are consistent with $S(m)$. That is, if A is a focal set of m_I and $A \cap S(m) \neq \emptyset$, then the new mass value on A after revision should reflect both $m_I(A)$ and the mass $m(A)$; otherwise, $m_I(A)$ should be retained after revision.

Example 6: 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.

From Example 6, we also observe that

A. Focal element $\{w_1, w_2\}$ of m_I is correlated 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 be involved in the revised value of $\{w_1, w_2\}$. Similarly, $\{w_6, w_7\}$ of m and $\{w_6\}$ of m_I are correlated. 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.
B. $\{w_3\}$ is not contained in $S(m_I)$, hence should not be contained in $S(\hat{m})$.

These observations show that we can partition W on the basis of correlated focal sets of m and m_I as follows. Let $S_1 \cup \cdots \cup S_k \cup S_{uncor} = S(m) \cup S(m_I)$ where

1) 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 $S(m)$.
2) Each S_i is the union of correlated focal sets. That is, for a focal set A of m (resp. m_I) s.t. $A \subseteq S_i$, then for any focal set B of m_I (resp. m), we have $B \subseteq S_i$ whenever $A \cap B \neq \emptyset$. In addition, if A is a focal set of m (resp. m_I) s.t. $A \subseteq S_i$ and B is a focal set of m_I (resp. m) s.t., $B \subseteq S_{uncor}$ or $B \subseteq S_j$ for $j \neq i$, then we have $A \cap B = \emptyset$. For instance, in Example 6, we can either have $k = 2$, s.t. $S_1 = \{w_1, w_2, w_8\}, S_2 = \{w_6, w_7\}$, or $k = 1$, s.t. $S_1 = \{w_1, w_2, w_6, w_7, w_8\}$.
3) k is the maximum number of correlated groups which satisfies the above two properties. Hence in Example 6, we should have $k = 2$ and $S_1 = \{w_1, w_2, w_8\}, S_2 = \{w_6, w_7\}$.

Each element S_i of the partition corresponds to a subset $\mathcal{F}_i$ of focal sets. A partition of the set of focal sets $\mathcal{F} \cup \mathcal{F}_I$ containing those of m and m_I is thus obtained. Given two mass functions, the partition of W into union of correlated focal sets can be obtained by Algorithm 1. The algorithm comes down to computing maximal connected components in a certain non-directed bipartite graph induced by the sets of focal sets $\mathcal{F}$ and $\mathcal{F}_I$. Namely, consider the bipartite graph whose nodes consist in focal sets in $\mathcal{F}$ and $\mathcal{F}_I$ (if the same set appears in $\mathcal{F}$ and $\mathcal{F}_I$ it produces two nodes). Arcs connect one focal set $A \in \mathcal{F}$ 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[1] graph.

Proposition 2: Let m and m_I be two mass functions, and S_1 and S_2 be the sets of focal sets for them respectively, Algorithm 1 produces a unique partition of W containing the maximum number of correlated groups.

Proof: The proof is easy after the following two results.

- Let A be a focal set of m (resp. m_I) s.t. $A \subseteq S_i$ and B is a focal set of m_I (resp. m) s.t., $B \subseteq S_{uncor}$ or $B \subseteq S_j$ for $j \neq i$, then we have $A \cap B = \emptyset$.

[1]A bipartite graph is a set of graph vertices decomposed into two disjoint sets such that no two graph vertices within the same set are adjacent.

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.

1: Set $S_{uncor} = \emptyset$, $k = 0$;
2: **while** $\mathcal{F}_1 \neq \emptyset$ **do**
3: Select a focal set A in $\mathcal{F}_1$;
4: **if** A does not overlap with any focal sets in $\mathcal{F}_2$ **then**
5: $S_{uncor} = S_{uncor} \cup A$; $\mathcal{F}_1 = \mathcal{F}_1 \setminus \{A\}$;
6: **else**
7: $k = k + 1$, $i = 2$, $S_k = A$, $\mathcal{F}_1 = \mathcal{F}_1 \setminus \{A\}$, $pre_B = S_k$;
8: **repeat**
9: 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;
10: Let $S_k = \bigcup_{B \in \mathcal{B}} B \cup S_k$; $pre_B = \bigcup_{B \in \mathcal{B}} B$;
11: $\mathcal{F}_i = \mathcal{F}_i \setminus \mathcal{B}$;
12: $i = 3 - i$; (repeatedly checking elements in $\mathcal{F}_1$ and $\mathcal{F}_2$)
13: **until** S_k can not be changed any further;
14: **end if**
15: **end while**
16: $S_{uncor} = \bigcup_{A \in \mathcal{F}_2} A \cup S_{uncor}$;
17: **return** $\{S_1, \cdots, S_k, S_{uncor}\}$;

This is clear from the procedure of Algorithm 1 (lines 8-13).

- Let $\{S_1, \cdots, S_k, S_{uncor}\}$ be the output of Algorithm 1. For any two focal sets $B, B' \subseteq S_i$ (regardless of B, B' from m or m_I), $1 \leq i \leq k$, then for any other partition result $\{S'_1, \cdots, S'_{k'}, S'_{uncor}\}$ satisfying conditions 1-3, there should be a S'_t such that $B, B' \subseteq S'_t$, hence $S_i \subseteq S'_t$. Now we prove the 2nd result. Let the first focal set being included in S_i be A (Algorithm line 4), then from the procedure of Algorithm 1 (lines 8-13), there exists a series of focal sets $C_0, \cdots, C_m$ such that $C_0 = A, C_m = B, C_j \cap C_{j+1} \neq \emptyset$ (e.g., $\{w_1\}, \{w_1, w_2\}, \{w_2\}$) and also a series of focal sets $D_0, \cdots, D_n$ such that $D_0 = A, D_n = B', D_l \cap D_{l+1} \neq \emptyset$ (e.g., $\{w_1\}, \{w_1, w_3\}, \{w_3\}$). Since the partition $\{S'_1, \cdots, S'_{k'}, S'_{uncor}\}$ satisfies condition 1 and A intersects with C_1 and D_1 which are focal sets of the other mass function, we have $A \not\subseteq S'_{uncor}$. So there exists S'_t such that $A \subseteq S'_t$, hence by condition 2, we have $C_1, D_1 \subseteq S'_t$, $C_2, D_2 \subseteq S'_t$, and so on. Finally we must have $B = C_m, B' = D_n \subseteq S'_t$.

From the first result, we have that the partition result $\{S_1, \cdots, S_k, S_{uncor}\}$ satisfies conditions 1 and 2. From the second result, we have that for any partition result $\{S'_1, \cdots, S'_{k'}, S'_{uncor}\}$ satisfying conditions 1-3, $k \geq k'$ must hold. Otherwise if $k < k'$, then there exists S'_j such that $\forall i$, $S_i \not\subseteq S'_j$. However, for any focal set $A \subseteq S'_j$ ($A \not\subseteq S'_{uncor}$), A must intersect at least one focal set A' of the other mass function, hence $A \not\subseteq S_{uncor}$. Let $A \subseteq S_i$, then from the second result, we can easily infer that $S_i \subseteq S'_j$ which leads to a contradiction. Therefore, $k \geq k'$ holds. Since $\{S'_1, \cdots, S'_{k'}, S'_{uncor}\}$ satisfies condition 3, we should have $k = k'$ which shows that $S_1, \cdots, S_k, S_{uncor}$ also satisfies condition 3. Furthermore, from the second result it is not difficult to prove that $\{S'_1, \cdots, S'_k, S'_{uncor}\}$ and $\{S_1, \cdots, S_k, S_{uncor}\}$ form a bijection. Therefore, $\{S_1, \cdots, S_k, S_{uncor}\}$ is the unique result satisfying conditions 1-3. Q.E.D.

Based on the obtained partition, we only need to consider revision inside S_i. For convenience, let focal sets of m included in S_i be $A_i^1, \cdots, A_i^s$ and let $S_i^A = \cup_{k=1}^s A_i^k$. Similarly, let focal sets of m_I included in S_i be $B_i^1, \cdots, B_i^t$ 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 of $D \subseteq B_i^j$. More precisely, given a subset S_i in the partition, flowing down the mass of B_i^j can be performed in the following procedure.

1) For each 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.
2) For C, the flown down value from B_i^j is $\hat{m}^j(C) = m_I(B_i^j)\frac{supp(C)}{\sum_{C \subseteq B_i^j} supp(C)}$.

This technique can be seen as a kind of conditioning on B_i^j, i.e., $\frac{\hat{m}^j(C)}{m_I(B_i^j)} = \frac{supp(C)}{\sum_{C \subseteq B_i^j} supp(C)}$. Evidently, this equation is to some extent similar to the form of Jeffrey's rule.

Example 7: (Ex. 6 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 $m(\{w_1, w_2\}) = 0.4$ can be seen as a positive support for giving the revised 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, $m(\{w_1, w_8\}) = 0.2$ supports $\{w_1\}$ from $m_I(\{w_1, w_2\})$ 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 an adaptive revision operator.

Definition 8: Let m and m_I be two mass functions and $\{S_1, \cdots, S_k, S_{uncor}\}$ be the partition of W obtained from Algorithm 1. Let $\hat{m}^j(C)$ and $supp(C)$ be defined as above. Then an adaptive revision operator for mass functions $\circ_a$ is defined as $\hat{m} = m \circ_a m_I$ such that $\hat{m}(C) = \sum_{j=1}^t \hat{m}^j(C)$.

From Algorithm 1, 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. 8 is indeed the final result for C.

Proposition 3: $\hat{m}$ is a specialization of m_I.

Example 8: (Ex. 6 Cont') Let m and m_I be as defined in Ex. 6, then we have $\hat{m} = m \circ_a m_I$ s.t. $\hat{m}(\{w_1\}) = 1/6$, $\hat{m}(\{w_1, w_2\}) = 1/3$, $\hat{m}(\{w_4, w_5\}) = 0.3$, and $\hat{m}(\{w_6\}) = 0.2$.

V. Properties of Mass Function Based Revision

We prove the equivalence between the modified outer revision and the adaptive revision. This finding is significant since these two revision strategies are from different perspectives and the proof of equivalence shows that the modified outer revision is well justified.

Proposition 4: For any two mass functions m and m_I over W, we have $m \circ_m m_I = m \circ_a m_I$.

Proof. It can be shown that

$$\begin{aligned}\sum_{C \subseteq B_i^j} supp(C) &= \sum_{C \subseteq B_i^j} \sum_{A_i^k \cap B_i^j = C} m(A_i^k) \\ &= \sum_{A_i^k \cap B_i^j \neq \emptyset} m(A_i^k) = Pl(B_i^j).\end{aligned}$$

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 $\circ_m = \circ_a$. Q.E.D.

Now we prove that our adaptive revision rule (also the modified outer revision rule) generalizes both Jeffrey's rule and Halpern's rule.

Proposition 5: If m is a Bayesian mass function and m_I is a partitioned mass function, then $m \circ_a m_I = m \circ_{Jef} m_I$. If m is a mass function and m_I is a partitioned mass function, then we have $Bel(m \circ_a m_I) = Bel(m) \circ_{Hal} Bel(m_I)$.

We show that the vacuous mass function plays no role in revision. It can also be seen as a reflection of *minimal change*.

Proposition 6: Let m be a mass function and m_W be such that $m_W(W) = 1$, then we have $m \circ_a m_W = m_W \circ_a m = m$.

Furthermore, we can also prove that if prior beliefs and new evidence are in total conflict, then the revision result is simply the latter.

Proposition 7: Let m and m_I be two mass functions such that $S(m) \cap S(m_I) = \emptyset$, then we have $m \circ_a m_I = m_I$.

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 $m \circ_a m_I = m_I$.

Proposition 8: Let m and m_I be two mass functions such that $m_I(U) = 1$, then $m \circ_a m_I$ comes down to Dempster conditioning if $Pl(U) > 0$ and $m \circ_a m_I = m_I$ otherwise.

Proposition 9: Let m and m_I be two mass functions such that $\forall A \in \mathcal{F}, B \in \mathcal{F}_I, A \cap B \neq \emptyset$, then $m \circ_a m_I$ comes down to Dempster rule of combination.

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.

For iterated revision, we have the following result.

Proposition 10: [2] Let $m, m_I, m_{I'}$ be three mass functions on W. If $m_{I'}$ yields a finer partition induced by correlated focal sets than m_I, then we have $m \circ_a m_I \circ_a m_{I'} = m \circ_a m_{I'}$.

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 $\hat{\hat{m}} = m \circ_a m_I \circ_a m_{I'}$ with $\hat{\hat{m}}(\{w_1\}) = \frac{6}{35}, \hat{\hat{m}}(\{w_2\}) = 0.3, \hat{\hat{m}}(\{w_3\}) = \frac{1}{35}, \hat{\hat{m}}(\{w_4\}) = 0.5$. $\check{m} = m \circ_a m_{I'}$ has the same set of focal sets and corresponding mass values.

In [18], four postulates on iterated belief revision were proposed, i.e., C1-C4. C1 and C2 are described as follows.

C1 If $\alpha \models \mu$, then $(\Phi \circ \mu) \circ \alpha \equiv \Phi \circ \alpha$.

C2 If $\alpha \models \neg\mu$, then $(\Phi \circ \mu) \circ \alpha \equiv \Phi \circ \alpha$.

$\Phi \circ \alpha \models \beta$ here stands for $BS(\Phi \circ \alpha) \models \beta$ where $BS(\Psi)$ represents the belief set of epistemic state Ψ.

Proposition 10 can be seen as a generalization of the above two iterated belief revision postulates. More precisely, since α, μ, and $\neg\mu$ can be represented in terms of mass functions as $m_\alpha(Mod(\alpha)) = 1$, $m_\mu(Mod(\mu)) = 1$ and $m_{\neg\mu}(Mod(\neg\mu)) = 1$, obviously if $\alpha \models \mu$ (resp. $\alpha \models \neg\mu$), then m_α has a finer partition induced by correlated focal sets than m_μ (resp. $m_{\neg\mu}$), hence C1 and C2 can be seen as special cases of Proposition 10. Due to the limitation of space, here we omit the discussion on relationships between our revision rules and the other revision postulates in [18].

VI. 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. We have proposed a set of revision rules to revise prior beliefs with new evidence. These revision rules are proved to satisfy some useful properties, such as the iteration property (Prop. 10). These rules are also proved to generalize the well known Jeffrey's rule and Halpern's belief function revision rule.

Our modified outer revision rule coincides with the adaptive revision rule that is proposed from a totally different perspective from that of the inner and outer revision rules. This result also demonstrates that a rational, yet mathematically simple, revision of mass functions can be achieved. Further work should strive to simplify the presentation of the revision rule in order to better lay bare its significance and further simplify its computation. Moreover, a precise formulation of the minimal change principle in the presence of reinforcement effect due to independence between the prior and the input is also needed.

[2] For convenience, proofs for this and some other propositions are put in the Appendix Section.

References

[1] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *The Annals of Statistics*, vol. 28, pp. 325–339, 1967.

[2] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[3] R. Yager and L. Liping, Eds., *Classic works of the Dempster-Shafer theory of belief functions*. Springer, 2008.

[4] D. Dubois and H. Prade, "Representation and combination of uncertainty with belief functions and possibility measures," *Computational Intelligence*, vol. 4, pp. 244–264, 1988.

[5] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66(2), pp. 191–234, 1994.

[6] R. Yager, "Hedging in the combination of evidence," *Journal of Information and Optimization Science*, vol. 4(1), pp. 73–81, 1983.

[7] L. Zhang, *Representation, independence, and combination of evidence in the Dempster-Shafer theory*. New York: John Wiley & Sons, Inc., 1994, pp. 51–69.

[8] A. Josang, "The consensus operator for combining beliefs," *Artificial Intelligence*, vol. 141(1-2), pp. 157–170, 2002.

[9] T. Denœux, "Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence," *Artifical Intelligence*, vol. 172, no. 2-3, pp. 234–264, 2008.

[10] K. Yamada, "A new combination of evidence based on compromise," *Fuzzy sets and Systems*, vol. 159, pp. 1689–1708, 2008.

[11] F. Smarandache and J. Dezert, "An introduction to the dsm theory for the combination of paradoxical, uncertain, and imprecise sources of information," http://arxiv.org/abs/cs/0608002.

[12] D. Dubois, H. Prade, and P. Smets, "New semantics for quantitative possibility theory," in *ECSQARU 2001*, 2001, pp. 410–421.

[13] S. Destercke, D. Dubois, and E. Chojnacki, "Cautious conjunctive merging of belief functions," in *Proceedings of ECSQARU 2007*, 2007, pp. 332–343.

[14] R. Jeffrey, *The logic of decision*. 2nd ed. Chicago University Press, 1983.

[15] D. Dubois, S. Moral, and H. Prade, *Belief change rules in ordinal and numerical uncertainty theories*. Kluwer Academic Pub., 1998, vol. 3, pp. 311–392.

[16] C. E. Alchourrón, P. Gärdenfors, and D. Makinson, "On the logic of theory change: Partial meet functions for contraction and revision," *Symbolic Logic*, vol. 50, pp. 510–530, 1985.

[17] H. Katsuno and A. O. Mendelzon, "Propositional knowledge base revision and minimal change," *Artificial Intelligence*, vol. 52, pp. 263–294, 1991.

[18] A. Darwiche and J. Pearl, "On the logic of iterated belief revision," *Artificial Intelligence*, vol. 89, pp. 1–29, 1997.

[19] D. Dubois and H. Prade, "Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge," in *Proc. of IJCQQPR*, 1997, pp. 96–107.

[20] J. Y. Halpern, *Reasoning about Uncertainty*. The MIT Press, Cambridge, Massachusetts, London, England, 2003.

[21] D. Dubois and H. Prade, "On the unicity of dempster rule of combination," *Int. J. of Intel. Sys.*, vol. 1, pp. 133–142, 1986.
[22] D. Dubois and H. Prade, "A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets," *Int. J. Gener. Sys.*, vol. 12(3), pp. 193–226, 1986.
[23] H. Ichihashi and H. Tanaka, "Jeffrey-like rules of conditioning for the dempster-shafer theory of evidence," *Int. J. of Approximate Reasoning*, vol. 3, pp. 143–156, 1989.
[24] P. Smets, "Jeffrey's rule of conditioning generalized to belief functions," in *Proc. of UAI*, 1993, pp. 500–505.
[25] P. Smets, "The application of the matrix calculus to belief functions," *International Journal of Approximate Reasoning*, vol. 31(1-2), pp. 1–30, 2002.
[26] G. Shafer, "Jeffrey's rule of conditioning," *Philosophy of Science*, vol. 48(3), pp. 337–362, 1981.

APPENDIX

Proof of Proposition 1: The proof is straightforward and omitted.

Proof of Proposition 3: The proof is straightforward and omitted. In addition, it could be seen as a corollary of Proposition 4.

Proofs of Propositions 5,6,7: The proofs are straightforward and omitted.

Proof of Proposition 10: Let $\hat{m} = m \circ_a m_I$, $\hat{\hat{m}} = m \circ_a m_I \circ_a m_{I'}$ and $\check{m} = m \circ_a m_{I'}$. Here we need show $\forall V$, $\check{m}(V) = \hat{\hat{m}}(V)$.

For any $V \subseteq W$, if $\check{m}(V) = \sum_{A\cap B=V} \check{\sigma}_m(A,B) m_{I'}(B) > 0$, then we obviously have $V \subseteq B$, as B should be a focal set of $m_{I'}$ hence V can only be included in a particular focal set of $m_{I'}$, say B_V, hence we have $\check{m}(V) = m_{I'}(B_V) \sum_{A\cap B_V=V} \check{\sigma}_m(A, B_V)$. Now we discuss two subcases.

- If $Pl(B_V) = 0$, we then have for any $V \subseteq B_V$, $\check{m}(V) = 0$ if $V \neq B_V$ and $\check{m}(V) = m_{I'}(B_V)$ if $V = B_V$. As $\hat{\hat{m}}(V) = \sum_{A\cap B=V} \hat{\hat{\sigma}}_m(A,B) m_{I'}(B)$, similar to the above, we also have $\hat{\hat{m}}(V) = m_{I'}(B_V) \sum_{A\cap B_V=V} \hat{\hat{\sigma}}_m(A, B_V)$. Now again we have two subcases.
  - If $\hat{Pl}(B_V) = 0$, then based on Def. 7, immediately we have $\hat{\hat{m}}(V) = 0$ if $V \neq B_V$ and $\hat{\hat{m}}(V) = m_{I'}(B_V)$ if $V = B_V$ which implies $\hat{\hat{m}}(V) = \check{m}(V)$.
  - If $\hat{Pl}(B_V) > 0$, then we have $\hat{\hat{\sigma}}_m(A, B_V) = \frac{\hat{m}(A)}{\hat{Pl}(B_V)}$ where $\hat{m}(A) = \sum_{C\cap D=A} \hat{\sigma}_m(C,D) m_I(D)$. Since m_I is a mass function, similarly we have that D can only be a particular D_A of m_I and $\hat{m}(A) = m_I(D_A) \sum_{C\cap D_A=A} \hat{\sigma}_m(C, D_A)$. From $V \subseteq D_A \cap B_V$ and $m_{I'}$ is a refined mass function of m_I, we must have $B_V \subseteq D_A$. Hence from $C \cap D_A \cap B_V = A \cap B_V = V$, we get $C \cap B_V = V$. As $Pl(B_V) = 0$, we get $m(C) = 0$. Hence from Def. 7, we have
  $$\hat{\sigma}_m(C, D_A) = \begin{cases} 1 & \text{for } Pl(D_A) = 0 \wedge C = D_A, \\ 0 & \text{otherwise.} \end{cases}$$
  If $V \neq B_V$, then as $C \cap B_V = V$ and $D_A \cap B_V = B_V$, we get $C \neq D_A$. Hence we have $\hat{\sigma}_m(C, D_A) = 0$, hence $\hat{m}(A) = 0$, hence $\hat{\hat{m}}(V) = 0 = \check{m}(V)$. If $V = B_V$, then we have $\hat{\hat{m}}(B_V) = \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} m_I(D_A) \sum_{C\cap B_V=B_V} \hat{\sigma}_m(C, D_A)$. Now if $Pl(D_A) > 0$, then we get $\forall U \cap B_V \neq \emptyset$, $\hat{m}(U) = \sum_{C'\cap D_A=U} \frac{m(C')}{Pl(D_A)} m_I(D_A) = 0$ as $m(C') = 0$ (obtained by $C' \cap B_V \neq \emptyset$ and $Pl(B_V) = 0$). Hence $\hat{Pl}(B_V) = 0$ which contradicts with $\hat{Pl}(B_V) > 0$. Hence we should have $Pl(D_A) = 0$, then based on Def. 7, we have $\hat{m}(D_A) = m_I(D_A)$ and for any $E \cap D_A \neq \emptyset, E \neq D_A$, $\hat{m}(E) = 0$, hence we have $\hat{Pl}(B_V) = m_I(D_A)$. As $\hat{\sigma}_m(C, D_A) = 1$ for $C = D_A$, we get
  $$\begin{aligned} &\hat{\hat{m}}(B_V) \\ =\ & \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} m_I(D_A) \sum_{C\cap B_V=B_V} \hat{\sigma}_m(C, D_A) \\ =\ & \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} m_I(D_A) = m_{I'}(B_V) = \check{m}(B_V). \end{aligned}$$
- 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 $V \subseteq D$, similarly we must have 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
  $$\begin{aligned} &\hat{\hat{m}}(V) \\ =\ & \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} \sum_{A\cap B_V=V} \sum_{C\cap D_V=A} \hat{\sigma}_m(C, D_V) m_I(D_V) \\ =\ & \frac{m_{I'}(B_V)}{\hat{Pl}(B_V)} \sum_{A\cap B_V=V} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C\cap D_V=A} m(C) \\ =\ & m_{I'}(B_V) \frac{\sum_{A\cap B_V=V} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C\cap D_V=A} m(C)}{\sum_{A'\cap B_V\neq\emptyset} \frac{m_I(D_{A'})}{Pl(D_{A'})} \sum_{C'\cap D_{A'}=A'} m(C')}, \end{aligned}$$
  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 $D_{A'} = D_V$. So we have
  $$\begin{aligned} &\hat{\hat{m}}(V) \\ =\ & m_{I'}(B_V) \frac{\sum_{A\cap B_V=V} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C\cap D_V=A} m(C)}{\sum_{A'\cap B_V\neq\emptyset} \frac{m_I(D_V)}{Pl(D_V)} \sum_{C'\cap D_V=A'} m(C')} \\ =\ & m_{I'}(B_V) \frac{\sum_{A\cap B_V=V} \sum_{C\cap D_V=A} m(C)}{\sum_{A'\cap B_V\neq\emptyset} \sum_{C'\cap D_V=A'} m(C')} \\ =\ & m_{I'}(B_V) \frac{\sum_{C\cap B_V=V} m(C)}{\sum_{C'\cap B_V\neq\emptyset} m(C')} \\ =\ & m_{I'}(B_V) \frac{\sum_{C\cap B_V=V} m(C)}{Pl(B_V)} \\ =\ & \check{m}(V). \end{aligned}$$