The Total Belief Theorem

Chunlai Zhou[1,2], Fabio Cuzzolin[3]
[1]DEKE(MOE), Computer Science Dept., Renmin University, Beijing, CHINA
[2]State Key Lab of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, CHINA
[3]Artificial Intelligence and Vision, Department of Computing and Communication Technologies,
Oxford Brookes University, Oxford, UK
czhou@ruc.edu.cn, fabio.cuzzolin@brookes.ac.uk

Abstract

In this paper, motivated by the treatment of conditional constraints in the data association problem, we state and prove the generalisation of the law of total probability to belief functions, as finite random sets. Our results apply to the case in which Dempster's conditioning is employed. We show that the solution to the resulting total belief problem is in general not unique, whereas it is unique when the a-priori belief function is Bayesian. Examples and case studies underpin the theoretical contributions. Finally, our results are compared to previous related work on the generalisation of Jeffrey's rule by Spies and Smets.

1 INTRODUCTION

A number of researchers have been working on the generalisation to belief functions (BFs) of a fundamental result of probability theory: the law of total probability. The latter is sometimes called 'Jeffrey's rule' [14, 15, 26, 13], for it is mathematically equivalent to a rule for updating a probability distribution based on a new distribution with values on a smaller set of events. Both laws concern the way non-conditional, joint models can be soundly built from conditional information and unconditional information.

A motivating example comes from *data association* [27, 22, 3]. There, a number of targets moving in the space are tracked by one or more cameras, appearing in an image sequence as unlabeled feature points [4]. When the targets belong to an object for which a topological model is known, an evidential solution can be proposed by expressing the prior, logical information carried by the body model in term of belief functions on a suitable frame of discernment (see [5], Chapter 7). In particular, a rigid motion constraint from each link in the topological model can be derived and expressed in a *conditional* way – in order to test the rigidity of the motion of two observed feature points at time k, we need to know the correct association between targets and features at time $k-1$. To construct an overall BF describing the association, such conditional pieces of information need to be combined. This requires us to equip the theory of belief functions with a law of total probability.
Another problem in which a total belief problem arises is example-based pose estimation [11], as soon as a regression framework based on belief functions is applied.

The generalisation of Jeffrey's rule to belief functions has been mainly studied by Spies [30] and Smets [28]. Ruspini [23] also reported results on deduction assuming approximate knowledge about the truth of conditional propositions. Spies [30] proved the existence of a solution to the generalisation of Jeffrey's rule to belief functions within his original conditioning framework. Smets also proposed generalisations of Jeffrey's rule based on both geometric and Dempster's conditioning [28].

1.1 CONTRIBUTIONS AND OUTLINE

In this paper we first provide a formal statement of the problem. Namely, we seek to combine conditional belief functions defined over disjoint subsets of a frame of discernment, while simultaneously constraining the resulting total belief function to be compatible with a second BF defined on the coarsening of the original frame.
We then adapt Smets' original proof of his well-known generalized Bayesian Theorem to construct a total belief function and show that it satisfies the prescribed marginalization and conditioning properties. The problem is shown to be equivalent to building a square linear system with positive solution, whose columns are associated with the focal elements (non-zero mass events) of the candidate total BF.

We first recall the necessary definitions from belief theory (Section 2). We then briefly review the belief-theoretical solution to the motivating data association problem (Section 3). In Section 4 we provide the formal statement of the total belief theorem and give a constructive proof for it. We show how to translate a total belief problem to a group of linear equations and analyze its possible solutions. Finally, Section 5 runs a critical comparison between our result and previous relevant work by Spies and Smets on the generalisation of Jeffrey's rule to belief functions. Section 6 concludes the paper.

2 BELIEF FUNCTIONS

2.1 BELIEF MEASURES

A *mass function* [2] over a *frame of discernment* Θ is a set function [8, 7] $m : 2^\Theta \to [0,1]$ defined on the collection 2^Θ of all subsets of Θ such that: $m(\emptyset) = 0$, $\sum_{A \subset \Theta} m(A) = 1$. The quantity $m(A)$ is called the *basic probability number* or 'mass' [17, 16] assigned to A, and measures the belief committed exactly to $A \in 2^\Theta$. The elements of the power set 2^Θ associated with non-zero values of m are called the *focal elements* of m. The *belief function* associated with a mass function $m : 2^\Theta \to [0,1]$ is the set function $b : 2^\Theta \to [0,1]$ defined as: $b(A) = \sum_{B \subseteq A} m(B)$. The domain Θ on which a belief function is defined is usually interpreted as the set of possible answers to a given problem, exactly one of which is the correct one. For each subset ('event') $A \subseteq \Theta$ the quantity $b(A)$ takes on the meaning of 'degree of belief' that the truth lies in A, and represents the total belief committed to a set of possible outcomes A by the available evidence m. Given a belief function b, we can obtain its corresponding mass function m as follows: $m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} b(B)$ for all $A \subseteq \Theta$. The belief function b is called *Bayesian* if $m(A) = 0$ for all non-singletons A. It is called *categorical* if it has only one focal set. And it is called *vacuous* if Θ is the only focal element. A vacuous belief represents a state of total ignorance. The corresponding *plausibility function* $pl : 2^\Omega \to [0,1]$ is defined by $pl(A) = \sum_{E \cap A \neq \emptyset} m(E)$ for all $A \subseteq \Theta$. For m, b and pl, if we know any one of them, then we can determine the other two.

2.2 CONDITIONING

In Bayesian reasoning, where all evidence comes in the form of a proposition A being true, conditioning (as we know) is performed via Bayes' rule. In belief theory, however, the onus is on combining the belief function representing our current knowledge state with a new one encoding the new evidence. After an initial proposal by Dempster, several other aggregation operators have been proposed, based on different assumptions on the nature and properties of the sources of evidence to combine.

Definition 1. *The* orthogonal sum *or* Dempster's combination $b_1 \oplus b_2 : 2^\Theta \to [0,1]$ *of two belief functions* $b_1 : 2^\Theta \to [0,1]$, $b_2 : 2^\Theta \to [0,1]$ *defined on the same frame of discernment* Θ *is the unique BF on* Θ *whose focal elements are all the possible intersections of focal elements of* b_1 *and* b_2, *and whose mass is given by:*

$$(m_1 \oplus m_2)(A) = \frac{\sum_{i,j: A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{i,j: A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)},$$

where m_i *denotes the mass function related to* b_i.

Belief functions can also be conditioned, rather than combined, whenever they need to be updated based on similar hard evidence. However, just as in the case of combination rules, a variety of conditioning operators can be defined for belief functions [9, 12, 10, 6, 33], many of them generalisations of Bayes' rule itself. In particular, Dempster's rule of combination naturally induces a conditioning operator, as follows. Given a conditioning event $A \subseteq \Theta$, the 'logical' (or 'categorical', in Smets' terminology) belief function b_A such that $m(A) = 1$ is combined via Dempster's rule with the a-priori belief function b. The resulting belief function $b \oplus b_A$ is the conditional belief function given A *a la Dempster*, denoted by $b(A|B)$.

Suppose that $\Theta' \supseteq \Theta$ and m is a mass function over Θ. The mass function m can be identified with a mass function $\overrightarrow{m}_{\Theta'}$ over the larger frame Θ': for any $E' \subseteq \Theta'$, $\overrightarrow{m}_{\Theta'}(E') = m(E)$ if $E' = E \cup (\Theta' \setminus \Theta)$ and $\overrightarrow{m}_{\Theta'}(E') = 0$ otherwise. Such $\overrightarrow{m}_{\Theta'}$ is called the *conditional embedding* (or *ballooning extension*) of m into Θ'. When the context is clear, we can drop the subscript Θ'.

2.3 OUTER REDUCTIONS

Suppose that Θ is a *finer* frame than Ω. This means that the elements $\omega_1, \cdots, \omega_{|\Omega|}$ of Ω correspond to a partition $\Pi_1, \cdots, \Pi_{|\Omega|}$ of Θ: a subset $\{\omega_{i_1}, \cdots \omega_{i_k}\}$ of Ω has the same meaning as the subset $\Pi_{i_1} \cup \cdots \cup \Pi_{i_k}$ of Θ. This identification can represented by a mapping $\rho : 2^\Omega \to 2^\Theta$ such that $\rho(\{\omega_i\}) = \Pi_i (1 \leq i \leq |\Omega|)$ and $\rho(\{\omega_{i_1}, \cdots \omega_{i_k}\}) = \cup_{j=1}^k \rho(\omega_{i_j}) = \cup_{j=1}^k \Pi_{i_j}$. The partition $\Pi_1, \cdots, \Pi_{|\Omega|}$ of Θ as a basis defines a subalgebra $\mathbb{A}^\rho$ of 2^Θ as a Boolean algebra with set operations, which is isomorphic to the set algebra $\mathbb{A} = 2^\Omega$.

In the theory of evidence, two frames are called *compatible* if and only if they concern propositions which can be both expressed in terms of propositions of a common, finer frame [25]. In particular, two compatible frames must admit a common refinement, i.e., a frame which is a refinement of both. Each collection of compatible frames has many common refinements. In particular, if $\Theta_1, ..., \Theta_n$ are elements of a family of compatible frames $\mathcal{F}$, then there exists a unique frame $\Theta \in \mathcal{F}$ such that:

1. $\exists$ a refining $\rho_i : 2^{\Theta_i} \rightarrow 2^{\Omega}$ for all $i = 1, ..., n$;
2. $\forall \theta \in \Theta \; \exists \theta_i \in \Theta_i \; \forall i = 1, ..., n$ such that:

$$\{\theta\} = \rho_1(\{\theta_1\}) \cap ... \cap \rho_n(\{\theta_n\}).$$

This unique frame is called the *minimal refinement* $\Theta_1 \otimes \cdots \otimes \Theta_n$ of the collection $\Theta_1, ..., \Theta_n$, and is the simplest space in which we can compare propositions pertaining to different compatible frames.

If Θ_1 and Θ_2 are two compatible frames, then two belief functions $b_1 : 2^{\Theta_1} \rightarrow [0, 1]$, $b_2 : 2^{\Theta_2} \rightarrow [0, 1]$ can potentially be expression of the same body of evidence. Two belief functions b_1 and b_2 defined over two compatible frames Θ_1 and Θ_2 are said to be *consistent* if $b_1(A_1) = b_2(A_2)$ whenever $\rho_1(A_1) = \rho_2(A_2)$, $A_1 \subseteq \Theta_1, \; A_2 \subseteq \Theta_2$, where ρ_i is the refining between Θ_i and the minimal refinement $\Theta_1 \otimes \Theta_2$ of Θ_1 and Θ_2. When the two belief functions are defined on frames connected by a refining $\rho : 2^{\Theta_1} \rightarrow 2^{\Theta_2}$ (i.e., Θ_2 is a refinement of Θ_1), b_1 and b_2 are consistent iff: $b_1(A) = b_2(\rho(A)), \forall A \subseteq \Theta_1$.

Definition 2. *The* outer reduction *associated with a refining $\rho : 2^{\Omega} \rightarrow 2^{\Theta}$ is the mapping $\overline{\rho} : 2^{\Theta} \rightarrow 2^{\Omega}$ defined as, for any $E \in 2^{\Theta}$,*

$$\overline{\rho}(E) : = \{\omega \in \Omega : \rho(\{\omega\}) \cap E \neq \emptyset\}.$$

Note that, for any $E \subseteq \Theta$, $\rho(\overline{\rho}(E))$ is the smallest element of $\mathbb{A}^{\rho}$ that contains E. So $\rho(\overline{\rho}(E))$ is called the *upper approximations* of E in $\mathbb{A}^{\rho}$, respectively. Given a belief function b over Θ with the refining mapping $\rho : 2^{\Omega} \rightarrow 2^{\Theta}$, its *marginal* $b \restriction_{\Omega}$ over Ω is defined as follows: $(b \restriction_{\Omega})(\{\omega_{i_1}, \cdots, \omega_{i_k}\}) = b(\Pi_{i_1} \cup \cdots \cup \Pi_{i_k})$. The corresponding mass function is:

$$(m \restriction_{\Omega})(\{\omega_{i_1}, \cdots, \omega_{i_k}\}) = \sum_{\substack{E \subseteq \Theta, \\ \overline{\rho}(E) = \{\omega_{i_1}, \cdots, \omega_{i_k}\}}} m(E). \tag{1}$$

A mass function m over Ω can be extended to a mass function $m^{\uparrow\Theta}$ over the finer frame Θ: $m^{\uparrow\Theta}(E) = m(\{\omega_{i_1}, \cdots, \omega_{i_k}\})$ if E is a union of some partition classes $\Pi_{i_1}, \cdots, \Pi_{i_k}$, $m^{\uparrow\Theta}(E) = 0$ otherwise. Such $m^{\uparrow\Theta}$ is called the *vacuous extension* of m. When the context is clear, we will omit the subscript Θ. Trivially, vacuous extension is the inverse of marginalization.

3 DATA ASSOCIATION WITH BFS

In data association we are given a sequence of images $\{I(k), k\}$, each containing a number of feature points $\{t_i(k)\}$ at time k which are projections of real world targets $\{T_1, ..., T_M\}$. We seek the correspondences $t_i(k) \longleftrightarrow t_j(k+1)$ between feature points in consecutive images which are projections of the same target.

3.1 CONDITIONAL CONSTRAINTS

If we assume that targets represent fixed positions on an articulated body connected by a rigid link, we can address the association task in critical situations in which several targets coalesce (*model-based* data association) via a set of logical constraints on the admissible positions of the targets. We can identify, among others [5]: (i) a 'prediction' constraint which encodes the likelihood of a measurement in the current image being associated with a measurement of the past image (e.g. produced by a Kalman filter [27] in joint probabilistic data association [4]); (ii) a *rigid motion* constraint, acting on pairs of targets T_{j_1}, T_{j_2} connected by a rigid link in the model:

$$\|t_i(k) - t_j(k)\| \cong \|t_{i'}(k-1) - t_{j'}(k-1)\|,$$

assuming that $t_i(k), t_{i'}(k-1)$ are both projections of T_{j_1}, and $t_j(k), t_{j'}(k-1)$ are images of T_{j_2}.
All such constraints can be expressed as belief functions over a suitable frame of discernment. However, whereas prediction information inherently concerns associations between feature points belonging to consecutive images, the rigid motion constraint depends on the target-to-measurement associations (e.g. $t_{i'}(k-1) \sim T_{j_1}, t_{j'}(k-1) \sim T_{j_2}$) estimated at the previous step.

3.2 A FAMILY OF ASSOCIATIONS FRAMES

We thus need to introduce a *past target-to-feature* associations frame: $\Theta_M^{k-1} \doteq \big\{t_i(k-1) \leftrightarrow T_j, i = 1, ..., n(k-1), j = 1, ..., M\big\}$, a *feature-to-feature* associations frame: $\Theta_k^{k-1} \doteq \big\{t_i(k-1) \leftrightarrow t_j(k), \forall i = 1, ..., n(k-1) \; \forall j = 1, ..., n(k)\big\}$, and a *current target-to-feature* associations frame: $\Theta_M^k \doteq \big\{t_i(k) \leftrightarrow T_j, \forall i = 1, ..., n(k) \;\; \forall j = 1, ..., M\big\}$. These form a family of compatible frames, as they are all connected by refining maps (see Figure 1). Prediction belief functions (i), for instance, will live on Θ_k^{k-1}. The belief estimate of the associations at time $k-1$ will live on Θ_M^{k-1}. The BFs encoding the various constraints can thus be combined on their minimal refinement $\Theta_M^{k-1} \otimes \Theta_k^{k-1}$. Marginalizing the resulting BF back onto the current target-to-feature association frame Θ_M^k yields the current best estimate.

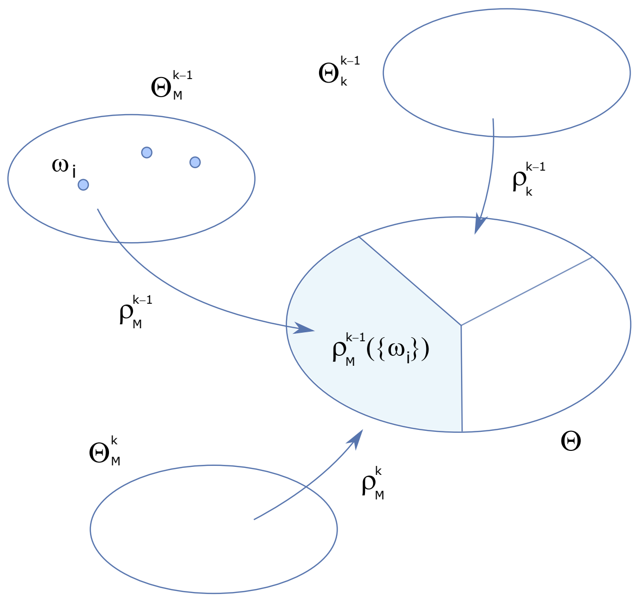

Figure 1: The family of past and present association frames. All the constraints of the model-based association problem are combined over the common refinement Θ and then marginalized onto the current association frame Θ^k_M to yield a belief estimate of the current feature-to-model association.

3.3 TOTAL BELIEF FOR DATA ASSOCIATION

The rigid motion constraint, however, generates an entire set of (conditional) belief functions $b_i : 2^{\rho^{k-1}_M(\{\omega_i\})} \rightarrow [0,1]$, each defined over an element $\rho^{k-1}_M(\{\omega_i\})$ of the disjoint partition of $\Theta = \Theta^{k-1}_M \otimes \Theta^{k-1}_k$ induced there by its coarsening Θ^{k-1}_M (the past target-to-feature frame, see Figure 1 again), where $\omega_i \in \Theta^{k-1}_M$ is the i-th possible association at time $k-1$. Merging all pieces of evidence on Θ thus requires combining these conditional belief functions into a single 'total' BF, which is eventually pooled with those generated by the remaining evidence.

4 THE TOTAL BELIEF THEOREM

4.1 TOTAL PROBABILITY

Suppose P is defined on a σ-algebra $\mathbb{A}$, and that a new probability measure P' on a sub-algebra $\mathbb{B}$ of $\mathbb{A}$. We seek an updated probability P'' which:

- meets the probability values specified by P' for events in the sub-algebra $\mathbb{B}$;
- is such that $\forall\, B \in \mathbb{B}, X, Y \subset B, X, Y \in \mathbb{A}$

$$\frac{P''(X)}{P''(Y)} = \begin{cases} \frac{P(X)}{P(Y)} & if\ P(Y) > 0 \\ 0 & if\ P(Y) = 0. \end{cases}$$

It can be proven that there is a unique solution to the above problem, given by *Jeffrey's rule*, also called the *law of total probability*:

$$P''(A) = \sum_{B \in \mathbb{B}} P(A|B)P'(B). \qquad (2)$$

The initial probability measure 'stands corrected' by the second one on a number of events (but not all). The law of total probability thus generalises standard conditioning, as the special case in which $P'(B) = 1$ for some B and the sub-algebra $\mathbb{B}$ reduced to a single event B.

4.2 CONSTRAINTS

The law of total probability involves, given a subalgebra of events $\mathbb{B}$: (i) a prior probability $P(B)$ on the events of $\mathbb{B}$, and (ii) a family of conditional probabilities $P(A|B)$ for every event in $\mathbb{B}$. In particular, $\mathbb{B}$ can be the subalgebra associated with the power set of a disjoint partition of the original sample space.

Abstracting from the data association problem, we can then state the conditions an overall, total belief function b must obey, given a set of conditional belief functions $b_i : 2^{\Pi_i} \rightarrow [0,1]$ over the elements Π_i of the partition $\Pi = \{\Pi_1, ..., \Pi_{|\Omega|}\}$ of a frame Θ induced by a coarsening Ω.

1. *A-priori constraint*: the marginal on the coarsening Ω of the frame Θ of the candidate total belief function b must coincide with a given *a-priori* b.f. $b_0 : 2^\Omega \rightarrow [0,1]$.

As we showed above, in the data association problem the *a-priori* constraint is represented by the BF encoding the estimate of the past feature-to-model association $M \leftrightarrow m(k-1)$, defined over Θ^{k-1}_k (Figure 1). It ensures that the belief total function is compatible with the last available estimate.

2. *Conditional constraint*: the belief function $b(\cdot|\Pi_i)$ obtained by (Dempster's) conditioning the total belief function b with respect to each element Π_i of the partition Π must coincide with the corresponding given conditional belief function b_i:

$$b(\cdot|\Pi_i) = b \oplus b_{\Pi_i} = b_i \qquad \forall i = 1, ..., N,$$

where $m_{\Pi_i} : 2^\Theta \rightarrow [0,1]$ is such that:

$$m_{\Pi_i}(A) = \begin{cases} 1 & A = \Pi_i \\ 0 & A \subseteq \Theta, A \neq \Pi_i. \end{cases} \qquad (3)$$

4.3 FORMULATION AND PROOF

The generalization of the total probability theorem to the theory of belief functions – the *total belief theorem* – thus reads as follows (Figure 2).

Theorem 1. *Suppose Θ and Ω are two frames of discernment, and $\rho : 2^{\Omega} \to 2^{\Theta}$ a given refining between them. Let b_0 be a belief function defined over $\Omega = \{\omega_1, ..., \omega_{|\Omega|}\}$. Suppose there exists a collection of belief functions $b_i : 2^{\Pi_i} \to [0, 1]$, where $\Pi = \{\Pi_1, ..., \Pi_{|\Omega|}\}$, $\Pi_i = \rho(\{\omega_i\})$, is the partition of Θ induced by its coarsening Ω. Then, there exists a total belief function $b : 2^{\Theta} \to [0, 1]$ such that:*

- *(P1) $b \oplus b_{\Pi_i} = b_i \ \forall i = 1, ..., |\Omega|$, where b_{Π_i} is the categorical belief function with mass m_{Π_i} (3);*
- *(P2) b_0 is the marginal of b on Ω, $b_0 = b \upharpoonright_{\Omega}$.*

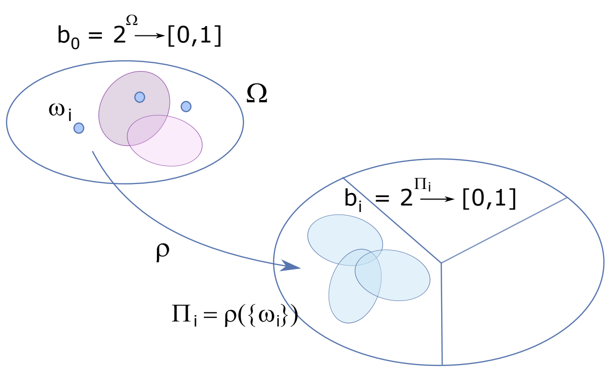

Figure 2: Pictorial representation of the total belief theorem hypotheses.

The theorem's proof makes use of the two lemmas that follow. We adapt Smets' original proof of the generalized Bayesian Theorem [29] to the refinement framework presented here, which is more general than his multivariate setting in which only Cartesian products of frames are considered. As usual, we denote by m_0 and m_i the mass functions of b_0 and b_i, respectively. Their sets of focal elements are denoted by $\mathcal{E}_{\Omega}$ and $\mathcal{E}_i$, respectively. For the collection of belief functions $b_i : 2^{\Pi_i} \to [0, 1]$, let $\vec{b_i}$ be the conditional embedding of b_i into Θ and $\vec{b}$ denote the Dempster combination of all $\vec{b_i}$, i.e., $\vec{b} = \vec{b_1} \oplus \cdots \oplus \vec{b_{|\Omega|}}$, with mass function $\vec{m}$.

Lemma 1. *The belief function $\vec{b}$ over Θ satisfies the following two properties: (1) each focal element $\vec{e}$ of $\vec{b}$ is the union of exactly one focal element e_i of each of the conditional belief function b_i; (2) the marginal $\vec{b} \upharpoonright_{\Omega}$ on Ω is the vacuous belief function over Ω.*

Proof. Each focal element $\vec{e_i}$ of $\vec{b_i}$ is of the form $(\bigcup_{j \neq i} \Pi_j) \cup e_i$ where e_i is some focal element of b_i. In other words, $\vec{e_i} = (\Theta \setminus \Pi_i) \cup e_i$. Since $\vec{b}$ is the Dempster combination of $\vec{b_i}$s, it is easy to see that each focal element $\vec{e}$ of $\vec{b}$ is the union of exactly one focal element e_i from each conditional belief function b_i. In other words, $\vec{e} = \bigcup_{i=1}^{|\Omega|} e_i$ where $e_i \in \mathcal{E}_i$, and condition (1) is proven.

Let $\vec{\mathcal{E}}$ denote the set of all focal elements of $\vec{b}$, namely:

$$\vec{\mathcal{E}} = \{\vec{e} \subseteq \Theta : \vec{e} = \textstyle\bigcup_{i=1}^{|\Omega|} e_i \text{ where } e_i \text{ is a focal element of } b_i\}.$$

Note that e_i's coming from different conditional belief functions b_i's are disjoint. For each $\vec{e} \in \vec{\mathcal{E}}$, $\bar{\rho}(\vec{e}) = \Omega$. It follows from Eq. (1) that $\vec{m} \upharpoonright_{\Omega} (\Omega) = 1$ and hence the marginal of $\vec{b}$ on Ω is the vacuous belief function there.

□

Let $b_0^{\uparrow\Theta}$ be the vacuous extension of b_0 from Ω to Θ. We define the desired total belief function b to be the Dempster combination of $b_0^{\uparrow\Theta}$ and $\vec{b}$, namely:

$$b := b_0^{\uparrow\Theta} \oplus \vec{b}. \quad (4)$$

Lemma 2. *The belief function b defined in (4) over Θ satisfies the following two properties:*

1. *$b \oplus b_{\Pi_i} = b_i$ for all $i = 1, \cdots, |\Omega|$ where b_{Π_i} is the categorical belief function with the mass function: $m_{\Pi_i}(A) = 1$ if $A = \Pi_i$, and is 0, otherwise.;*
2. *b_0 is the marginal of b on Ω,i.e., $b_0 = b \upharpoonright_{\Omega}$,*

i.e., is a valid total belief function.

Proof. Let $\vec{m}$ and m_i be the mass functions corresponding to $\vec{b}$ and b_i, respectively. For each $\vec{e} = \bigcup_{i=1}^{|\Omega|} e_i \in \vec{\mathcal{E}}$ where $e_i \in \mathcal{E}_i$, $\vec{m}(\vec{e}) = \Pi_{i=1}^{|\Omega|} m_i(e_i)$. Let $\mathcal{E}_{\Omega}^{\uparrow\Theta}$ denote the set of focal elements $b_0^{\uparrow\Theta}$. Since $b_0^{\uparrow\Theta}$ is the vacuous extension of b_0, $\mathcal{E}_{\Omega}^{\uparrow\Theta} = \{\rho(e_{\Omega}) : e_{\Omega} \in \mathcal{E}_{\Omega}\}$. Each element of $\mathcal{E}_{\Omega}^{\uparrow\Theta}$ is actually the union of some equivalence classes Π_i of the partition Π. Since each focal element of $b_0^{\uparrow\Theta}$ intersects with all focal elements $\vec{e} \in \vec{\mathcal{E}}$,

$$\sum_{e_{\Omega} \in \mathcal{E}_{\Omega}, \vec{e} \in \vec{\mathcal{E}}, \rho(e_{\Omega}) \cap \vec{e} \neq \emptyset} m_0^{\uparrow\Theta}(\rho(e_{\Omega}))\vec{m}(\vec{e}) = 1. \quad (5)$$

Thus, the normalization factor in the Dempster combination $b_0^{\uparrow\Theta} \oplus \vec{b}$ is equal to 1.

Now, let $\mathcal{E}$ denote the set of focal elements of the belief function $b = b_0^{\uparrow\Theta} \oplus \vec{b}$. By Dempster's sum (1) each element e of $\mathcal{E}$ is the union of focal elements of some conditional belief functions b_i, i.e., $e = e_{j_1} \cup e_{j_2} \cup \cdots \cup e_{j_K}$ for some K such that $\{j_1, \cdots, j_K\} \subseteq \{1, \cdots, |\Omega|\}$ and e_{j_l} is a focal element of $b_{j_l}(1 \leq l \leq K)$. Let m denote the mass function for b.

For each such $e \in \mathcal{E}$, $e = \rho(e_\Omega) \cap \overrightarrow{e}$ for some $e_\Omega \in \mathcal{E}_\Omega$ and $\overrightarrow{e} \in \overrightarrow{\mathcal{E}}$, so that $e_\Omega = \bar{\rho}(e)$. Thus we have

$$\begin{aligned}
(m_0^{\uparrow\Theta} \oplus \overrightarrow{m})(e) &= \sum_{\substack{e_\Omega \in \mathcal{E}_\Omega, \overrightarrow{e} \in \overrightarrow{\mathcal{E}}, \\ \rho(e_\Omega) \cap \overrightarrow{e} = e}} m_0^{\uparrow\Theta}(\rho(e_\Omega)) \overrightarrow{m}(\overrightarrow{e}) \\
&= \sum_{e_\Omega \in \mathcal{E}_\Omega, \overrightarrow{e} \in \overrightarrow{\mathcal{E}}, \rho(e_\Omega) \cap \overrightarrow{e} = e} m_0(e_\Omega) \overrightarrow{m}(\overrightarrow{e}) \\
&= m_0(\bar{\rho}(e)) \sum_{\overrightarrow{e} \in \overrightarrow{\mathcal{E}}, \rho(\bar{\rho}(e)) \cap \overrightarrow{e} = e} \overrightarrow{m}(\overrightarrow{e}) \\
&= m_0(\bar{\rho}(e)) m_{j_1}(e_{j_1}) \cdots m_{j_K}(e_{j_K}) \prod_{j \notin \{j_1, \cdots, j_K\}} \sum_{e \in \mathcal{E}_j} m_j(e) \\
&= m_0(\bar{\rho}(e)) m_{j_1}(e_{j_1}) \cdots m_{j_K}(e_{j_K}), \qquad (6)
\end{aligned}$$

as $\overrightarrow{m}(\overrightarrow{e}) = \prod_{i=1}^n m_i(e_i)$ whenever $\overrightarrow{e} = \cup_{i=1}^n e_i$.

Without loss of generality, we consider the conditional mass function $m(e_1|\Pi_1)$ where e_1 is a focal element of b_1 and Π_1 is the first partition class associated with the partition Π, and show that $m(e_1|\Pi_1) = m_1(e_1)$. In order to obtain $m(e_1|\Pi_1)$, which is equal to $\frac{\sum_{e \in \mathcal{E}, e \cap \Pi_1 = e_1} m(e)}{pl(\Pi_1)}$, in the following we separately compute $\sum_{e \in \mathcal{E}, e \cap \Pi_1 = e_1} m(e)$ and $pl(\Pi_1)$. For any $e \in \mathcal{E}$, if $e \cap \Pi_1 \neq \emptyset$, $\bar{\rho}(e)$ is a subset of Ω including ω_1.

$$\begin{aligned}
pl(\Pi_1) &= \sum_{e \in \mathcal{E}, e \cap \Pi_1 \neq \emptyset} m(e) \\
&= \sum_{\mathcal{C} \subseteq \{\Pi_2, \cdots, \Pi_{|\Omega|}\}} \left(\sum_{\rho(\bar{\rho}(e)) = \Pi_1 \cup (\bigcup_{E \in \mathcal{C}} E)} m(e) \right) \\
&= \sum_{\mathcal{C} \subseteq \{\Pi_2, \cdots, \Pi_{|\Omega|}\}} m_0^{\uparrow\Theta} \left(\Pi_1 \cup \bigcup_{E \in \mathcal{C}} E \right) \\
&\left(\sum_{e_1 \in \mathcal{E}_1} m_1(e_1) \prod_{\Pi_l \in \mathcal{C}} \sum_{e_l \in \mathcal{E}_l} m_l(e_l) \right) \\
&= \sum_{\mathcal{C} \subseteq \{\Pi_2, \cdots, \Pi_{|\Omega|}\}} m_0^{\uparrow\Theta} \left(\Pi_1 \cup \bigcup_{E \in \mathcal{C}} E \right) \\
&= \sum_{e_\Omega \in \mathcal{E}_\Omega, \omega_1 \in e_\Omega} m_0(e_\Omega) = pl_0(\{\omega_1\}). \qquad (7)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\sum_{e \in \mathcal{E}, e \cap \Pi_1 = e_1} m(e) &= \sum_{\mathcal{C} \subseteq \{\Pi_2, \cdots, \Pi_{|\Omega|}\}} \sum_{\rho(\bar{\rho}(e)) = \bigcup_{E \in \mathcal{C}} E} m(e_1 \cup e) \\
&= m_1(e_1) \sum_{\mathcal{C} \subseteq \{\Pi_2, \cdots, \Pi_{|\Omega|}\}} m_0^{\uparrow\Theta} \Big(\Pi_1 \cup \bigcup_{E \in \mathcal{C}} E \Big) \prod_{\Pi_l \in \mathcal{C}} \sum_{e_l \in \mathcal{E}_l} m_l(e_l) \\
&= m_1(e_1) \sum_{\mathcal{C} \subseteq \{\Pi_2, \cdots, \Pi_{|\Omega|}\}} m_0^{\uparrow\Theta} \Big(\Pi_1 \cup \bigcup_{E \in \mathcal{C}} E \Big) \\
&= m_1(e_1) \sum_{e_\Omega \in \mathcal{E}_\Omega, \omega_1 \in e_\Omega} m_0(e_\Omega) = m_1(e_1) pl_0(\{\omega_1\}). \qquad (8)
\end{aligned}$$

From Eqs. (7) and (8) it follows that $m(e_1|\Pi_1) = \frac{\sum_{e \in \mathcal{E}, e \cap \Pi_1 = e_1} m(e)}{pl(\Pi_1)} = m_1(e_1)$. This proves property 1. Proving 2. is much easier. For any $e_\Omega := \{\omega_{j_1}, \cdots, \omega_{j_K}\} \in \mathcal{E}_\Omega$,

$$\begin{aligned}
m \upharpoonright_\Omega (e_\Omega) &= \sum_{\bar{\rho}(e) = e_\Omega} m(e) \\
&= m_0^{\uparrow\Theta}(\rho(e_\Omega)) \prod_{l=1}^{K} \sum_{e \in \mathcal{E}_{j_l}} m_{j_l}(e) = m_0^{\uparrow\Theta}(\rho(e_\Omega)) = \qquad (9)
\end{aligned}$$

$= m_0(e_\Omega)$. It follows that $b \upharpoonright_\Omega = b_0$, hence the thesis.

□

The proof of the main Theorem 1 immediately follows from Lemmas 1 and 2.

Example 1. *Suppose that the considered coarsening $\Omega := \{\omega_1, \omega_2, \omega_3\}$ induces a partion Π of Θ: $\{\Pi_1, \Pi_2, \Pi_3\}$. Also suppose that the considered conditional belief function b_1 defined on Π_1 has two focal elements e_1^1 and e_1^2; the conditional belief function b_2 defined on Π_2 has a single focal element e_2^1; b_3 defined on Π_3 has two focal elements e_3^1 and e_3^2 (See Figure 3).*

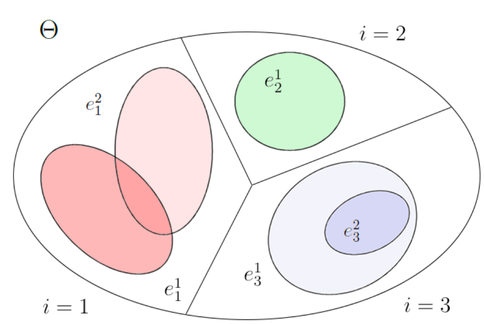

Figure 3: The conditional belief functions considered in our case study. The set-theoretical relations between their focal elements are immaterial to the solution.

According to Lemma 1, Dempster's combination $\overrightarrow{b}$ of *the conditional embeddings of the b_i's has 4 focal elements, which are listed as follows:*

$$e_1 = e_1^1 \cup e_2^1 \cup e_3^1, \quad e_2 = e_1^1 \cup e_2^1 \cup e_3^2$$
$$e_3 = e_1^2 \cup e_2^1 \cup e_3^1, \quad e_4 = e_1^2 \cup e_2^1 \cup e_3^2$$

The four focal total elements can be represented as "elastic bands" as in Figure 4.

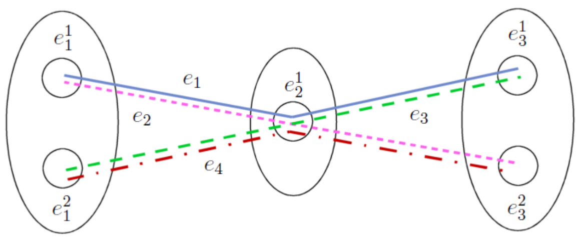

Figure 4: Graphical representation of the four possible focal elements of $\overrightarrow{b}$ in our case study.

Without loss of generality, we assume that a prior b_0 on Ω has each subset of Ω as a focal element, i.e., $\mathcal{E}_\Omega = 2^\Omega$. It follows that each focal element e of the total belief function $b := \overrightarrow{b} \oplus b_0^{\uparrow\Theta}$ is the union of some focal elements from different *conditional belief functions b_is. So the set $\mathcal{E}$ of the focal elements of b is $\{e = \bigcup_{1 \le i \le I} e_i : 1 \le I \le 3, e_i \in \mathcal{E}_i\}$ and is the union of the following three sets:*

$$\mathcal{E}_{I=1} := \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3$$
$$\mathcal{E}_{I=2} := \{e \cup e' : (e, e') \in \mathcal{E}_i \times \mathcal{E}_j, 1 \le i, j \le 3, i \ne j\}$$
$$\mathcal{E}_{I=3} := \{e_1 \cup e_2 \cup e_3 : e_i \in \mathcal{E}_i, 1 \le i \le 3\}.$$

So the cardinality $|\mathcal{E}| = 5 + 8 + 4 = 17$. According to Eq. (6), it is very easy to compute the corresponding total mass function m. For example, for the two focal elements $e_1^1 \cup e_3^2$ and $e_1^1 \cup e_2^1 \cup e_3^2$ we have:

$$m(e_1^1 \cup e_3^2) = m_0(\{w_1, w_3\}) m_1(e_1^1) m_3(e_3^2)$$
$$m(e_1^1 \cup e_2^1 \cup e_3^2) = m_0(\Omega) m_1(e_1^1) m_2(e_2^1) m_3(e_3^2).$$

Since we assume that a topological model is known (Section 3), in this paper we take a closed-world approach and hence adopt normalized mass functions. Under an open-world assumption in which mass functions are not necessarily normalized, Theorem 1 does not hold any more. In the unnormalized case, however, a weaker form of (P1) exists. Let $m \bigcirc\!\!\!\!\cap\, m_{\Pi_i}$ denote Smets' conjunctive combination of m and the categorical mass function m_{Π_i} [29]. For any two focal elements e_1 and e_2 of m_i, $\frac{(m \bigcirc\!\!\!\!\cap\, m_{\Pi_i})(e_1)}{(m \bigcirc\!\!\!\!\cap\, m_{\Pi_i})(e_2)} = \frac{m_i(e_1)}{m_i(e_2)}$. Constraint (P2), instead, does not seem to be easily generalizable to unnormalized mass functions.

4.4 NUMBER OF SOLUTIONS

The total belief function b obtained in Theorem 1 is *not unique*. Assume that b^* is a total belief function satisfies the two properties in Theorem 1. Let m^* and $\mathcal{E}^*$ denote its mass function and the set of its focal elements, respectively. Without loss of generality, we still assume that the prior b_0 has every subset of Ω as its focal element, i.e., $\mathcal{E}_\Omega = 2^\Omega$. From the second property that $b^* \oplus b_{\Pi_i} = b_i (1 \le i \le |\Omega|)$, we derive that each focal element of b^* must be a union of focal elements of some conditional belief functions b_is. For, if e^* is a focal element of b^* and $e^* = e_l \cup e'$ where $\emptyset \ne e_l \subseteq \Pi_l$ and $e' \subseteq \Theta \setminus \Pi_l$ for some $1 \le l \le |\Omega|$, then $m_l(e_l) = (m^* \oplus m_{\Pi_l})(e_l) > 0$ and hence $e_l \in \mathcal{E}_l$. So we must have that $\mathcal{E}^* \subseteq \mathcal{E}$, where $\mathcal{E}$ is the set of focal elements of the total belief function b obtained in Theorem 1: $\mathcal{E} = \{\bigcup_{j \in J} e_j : J \subseteq \{1, \cdots, |\Omega|\}, e_j \in \mathcal{E}_j\}$.

In order to find b^* (or m^*), we need to solve a group of linear equations which correspond to the constraints dictated in the two properties. We specify the mass $m^*(e)$ of each focal element $e \in \mathcal{E}$ of the total solution (4) as an unknown variable. There are $|\mathcal{E}|$ variables in the group.

From Properties (P1) and (P2) we know that $pl_0(\omega_i) = pl^*(\Pi_i)(1 \le i \le |\Omega|)$ where pl_0 and pl^* are the corresponding plausibility functions of b_0 and b^*, respectively. In addition, Property (P1) implies the system of linear constraints:

$$\left\{ \sum_{e \cap \Pi_i = e_i, e \in \mathcal{E}} m^*(e) = m_i(e_i) pl_0(\omega_i), \quad \forall i \forall e_i \in \mathcal{E}_i. \right. \tag{10}$$

The total number of such equations is $\sum_{j=1}^{|\Omega|} |\mathcal{E}_j|$. Since, for each $1 \le i \le |\Omega|$, $\sum_{e \in \mathcal{E}_i} m_i(e) = 1$, system (10) include a group of $\sum_{j=1}^{|\Omega|} |\mathcal{E}_j| - |\Omega|$ *independent* linear equations, which we denote as G_1. From property (P2) (the marginal of b^* on Ω is b_0) follow the constaints:

$$\left\{ \sum_{e \in \mathcal{E}, \overline{\rho}(e) = C} m^*(e) = m_0(C), \quad \forall \emptyset \ne C \subseteq \Omega. \right. \tag{11}$$

The total number of linear equations in (11) is the number of all nonempty subsets of Ω. Since $\sum_{C \subseteq \Omega} m_0(C) = 1$, there is a subset of $|2^\Omega| - 2$ independent linear equations in (11), denoted by G_2.

The groups of constraints G_1 and G_2 are independent for, although for each $1 \le i \le |\Omega|$ $\sum_{w_i \in C, C \subseteq \Omega} m_0(C) = pl_0(w_i), m_i(e_i) \sum_{w_i \in C, C \subseteq \Omega} \sum_{e \in \mathcal{E}, \overline{\rho}(e) = C} m^*(e)$ is not generally identical to $\sum_{e \cap \Pi_i = e_i, e \in \mathcal{E}} m^*(e)$. Therefore, the union $G := G_1 \cup G_2$ completely specifies Properties (P1) and (P2) in Theorem 1. Since $|G_1| = \sum_{j=1}^{|\Omega|} |\mathcal{E}_j| -$

$|\Omega|$ and $|G_2| = |2^{\Omega}| - 2$, the cardinality of the union group G is $|G| = \sum_{j=1}^{|\Omega|} |\mathcal{E}_j| - |\Omega| + |2^{\Omega}| - 2$.

From Theorem 1, we know that the system of equations G is solvable and has at least a positive solution (in which each variable has a positive value). This implies that $|\mathcal{E}| \geq |G|$, i.e., the number of variables must be no less than that of the indepedent linear equations in G. If $|\mathcal{E}| > |G|$, in particular, we can apply the Fourier-Motzkin elimination method [24] to show that G has *another distinct positive* solution b^* (i.e., such that $m^*(e) \geq 0 (e \in \mathcal{E})$).

Example 2. *We employ Example 1 to illustrate the whole process to find the total belief function b^*. We further assume that m_0 and $m_i (1 \leq i \leq 3)$ take numerical values:*

- $m_1(e_1^1) = \frac{1}{2} = m_1(e_1^2); m_2(e_2^1) = 1; m_3(e_3^1) = \frac{1}{3}, m_3(e_3^2) = \frac{2}{3};$
- $m_0(\{\omega_1\}) = m_0(\{\omega_2\}) = m_0(\{\omega_3\}) = \frac{1}{16}; \ m_0(\{\omega_1, \omega_2\}) = \frac{2}{16}, m_0(\{\omega_2, \omega_3\}) = \frac{4}{16}, m_0(\{\omega_1, \omega_3\}) = \frac{3}{16}; m_0(\Omega) = \frac{1}{4}.$

If we follow the above prescribed process to translate the two properties into a group G of linear equations, we obtain 17 unknown variables $m^*(e) (e \in \mathcal{E})$ ($|\mathcal{E}| = 17$) and 8 indepedent linear equations ($|G| = 8$). From Theorem 1, we can construct a positive solution m defined according to Eq. (6). For this example:

$$m(\{e_1^1, e_2^1, e_3^1\}) = m_0(\Omega) m_1(e_1^1) m_2(e_2^1) m_3(e_3^1) = \frac{1}{24},$$
$$m(\{e_1^2, e_2^1, e_3^2\}) = m_0(\Omega) m_1(e_1^2) m_2(e_2^1) m_3(e_3^2) = \frac{1}{12}.$$

When solving the equation group G via the Fourier-Motzkin elimination method, we choose $m^*(\{e_1^1, e_2^1, e_3^1\})$ and $m^*(\{e_1^2, e_2^1, e_3^2\})$ to be the last two variables to be eliminated. Moreover, there is a sufficiently small postive number ϵ such that $m^*(\{e_1^1, e_2^1, e_3^1\}) = \frac{1}{24} - \epsilon > 0$, $m^*(\{e_1^2, e_2^1, e_3^2\}) = \frac{1}{12} + \epsilon$, and all other varaibles also take positive values. It is easy to see that such obtained m^* is different from m obtained in Theorem 1.

However, when the prior b_0 is Bayesian, the total belief function obtained according to Eq. (6) is the unique one satisfying the two properties in Theorem 1.

Corollary 1. *For the belief function b_0 over Ω and conditional belief functions b_i over Π_i in Theorem 1, if b_0 is Bayesian (a probability function) such that $b_0(\omega_i) > 0$ for all $1 \leq i \leq |\Omega|$, then there is a unique total belief function $b : 2^{\Theta} \to [0, 1]$ such that:*

1. *$b \oplus b_{\Pi_i} = b_i$ for all $i = 1, \cdots, |\Omega|$ where b_{Π_i} is the categorical belief function with $m_{\Pi_i}(A) = 1$ if $A = \Pi_i$, and is 0, o.w.;*
2. *b_0 is the marginal of b on Ω, i.e., $b_0 = b \restriction_{\Omega}$.*

Moreover, the total mass function m of b is:

$$m(e) = \begin{cases} m_i(e) b_0(\omega_i) & \textit{if } e \in \mathcal{E}_i \textit{ for some } i, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. It is easy to check that the total mass function m defined above satisfies the two properties. Now we need to show that it is unique. Assume that b_0 is Bayesian such that $b_0(\omega_i) > 0$ for all i and m is the mass function of a total belief function b satisfying the two properties. Since $b_0 = b \restriction_{\Omega}$, each focal element e of b is a subset of some equivalence class Π_k for some $1 \leq k \leq |\Omega|$. Together with the first requirement that $b \oplus b_{\Pi_i} = b_i$, it implies that $pl(\Pi_i) = b(\Pi_i) = b_0(\omega_i)$ for all $1 \leq i \leq |\Omega|$ and $\mathcal{E} = \bigcup_{j=1}^{|\Omega|} \mathcal{E}_j$. For any $e \in \mathcal{E}$, $e \in \mathcal{E}_k$ for some $1 \leq k \leq |\Omega|$. It follows that $(m \oplus m_{\Pi_k})(e) = \frac{m(e)}{pl(\Pi_k)} = m_k(e)$ and hence $m(e) = m_k(e) pl(\Pi_k) = m_k(e) b_0(\omega_k)$. If $e \notin \bigcup_{i=1}^{|\Omega|} \mathcal{E}_i$, $e \notin \mathcal{E}$ and hence $m(e) = 0$. So we have shown that the total mass function m is the unique one satisfying the two properties. $\square$

Corollary 1 recalls a more general characterization of uniqueness [1]: if $Bel(A, B) = Bel(A|B) \oplus Bel(B)$, then it will be unique when $Bel(A|B)$ is defined for all primitive elements of B and $Bel(B)$ is Bayesian, or $Bel(A|B)$ is defined for all subsets of B.

4.5 GENERALISATION

If the first requirement (P1) is modified to include conditional constraints with respect to unions of equivalence classes, the approach used to prove Theorem 1 does not work.

For each nonempty subset $\{i_1, \cdots, i_J\} \subseteq \{1, \cdots, |\Omega|\}$, let $b_{\bigcup_{j=1}^J \Pi_{i_j}}$ and $b_{i_1 \cdots i_J}$ denote the categorical belief function with as only focal element $\bigcup_{j=1}^J \Pi_{i_j}$ and a conditional belief function on $\bigcup_{j=1}^J \Pi_{i_j}$, respectively.

We can introduce a new requirement by generalising Property (P1):

- $(P1') : b \oplus b_{\bigcup_{j=1}^J \Pi_{i_j}} = b_{i_1 \cdots i_J}$ for every nonempty subset $\{i_1, \cdots, i_J\} \subseteq \{1, \cdots, |\Omega|\}$.

Let $\overrightarrow{b}_{i_1 \cdots i_J}$ denote the conditional embeddings of $b_{i_1 \cdots i_J}$ and $\overrightarrow{b}_{new}$ the Dempster combination of all these conditional embeddings. Let $b_{new} = \overrightarrow{b}_{new} \oplus b_0^{\uparrow \Theta}$. In the following example, we show that b_{new} satisfies neither $(P1')$ nor $(P2)$.

Example 3. *We continue Example 2 with an additional categorical conditional belief function b_{12} on $\Pi_1 \cup \Pi_2$ with the only focal element e_{new} where $e_{new} = e_1^3 \cup e_2^2$ for some $e_1^3 \subseteq \Pi_1$ and $e_2^2 \subseteq \Pi_2$ such that $e_1^3 \cap e_1^1 \neq \emptyset, e_1^3 \cap e_1^2 = \emptyset$ and $e_2^2 \cap e_2^1 = \emptyset$. It is easy to check that $b_{new} = (\overrightarrow{b} \oplus \overrightarrow{b}_{12}) \oplus b_0^{\uparrow\Theta}$ satisfies neither $P1'$ nor $P2$.*

5 RELATION TO GENERALISED JEFFREY'S RULES

In spirit, our approach in this paper is similar to Spies' Jeffrey's rule for belief functions in [30]. His total belief function is also the Dempster combination of the prior on the subalgebra generated by the partion Π of Θ with conditional belief functions on each equivalence class Π_i. Moreover, he showed that this total belief function satisfies the two properties in Theorem 1. However, his definition of conditional belief function is different from the one used in this paper, derived from Dempster's rule of combination. His definition falls within the framework of random sets, so that a conditional belief function there is a *second-order* belief function whose focal elements are conditional events which are sets of subsets of the underlying frame of discernment. The biggest difference between Spies' approach and ours is thus that his framework depends on probabilities while ours doesn't. It would be interesting to explore the connection of our total belief theorem to his Jeffrey's rule for BFs.

Smets [28] also generalized Jeffrey's rule within the framework of models based on belief functions, without relying on probabilities. Recall that ρ is a refining mapping from 2^Ω to 2^Θ and $\mathbb{A}^\rho$ is the Boolean algebra generated by the set of equivalence classes Π_i associated with the refining mapping ρ. Contrary to our total belief theorem which assumes conditional constraints only with respect to the equivalence classes Π_i (the atoms of $\mathbb{A}^\rho$), Smets' generalized Jeffrey's rule considers constraints with respect to unions of equivalence classes, i.e., arbitrary elements of $\mathbb{A}^\rho$. Given two belief functions b_1 and b_2 over Θ, his general idea is to find a BF b_3 there such that:

- (Q1) its marginal on Ω is the same as that of b_1, i.e., $b_3 \upharpoonright_\Omega = b_1 \upharpoonright_\Omega$;
- (Q2) its conditional constraints w.r.t. elements of $\mathbb{A}^\rho$ are the same as those of b_2.

Let m be a mass function over Θ. Smets defines two kinds of conditioning for conditional constraints: for any $E \in \mathbb{A}^\rho$ and $e \subseteq E$ such that $\rho(\bar{\rho}(e)) = E$,

- $m^{in}(e|E) := \frac{m(e)}{\sum_{\rho(\bar{\rho}(e')) = E} m(e')}$;
- $m^{out}(e|E) := \frac{m(e|E)}{\sum_{\rho(\bar{\rho}(e')) = E} m(e'|E)}$.

The first one is the well-known geometric conditioning, whereas the second one is called 'outer conditioning'. Both are distinct from Dempster's rule of conditioning used in this paper. From these two conditioning rules, he obtaines two different forms of generalized Jeffrey's rule: for any $e \subseteq \Theta$,

- $m_3^{in}(e) = m_1^{in}(e|E) m_2(E)$ where $E = \rho(\bar{\rho}(e))$;
- $m_3^{out}(e) = m_1^{out}(e|E) m_2(E)$.

Both m_3^{in} and m_3^{out} satisfy (Q1). As for (Q2), m_3^{in} applies whereas m_3^{out} only *partially* does, since $(m_3^{out})^{in}(e|E) = m_1^{out}(e|E)$ [34].

Finally, in [18] Ma *et al* define a new Jeffrey's rule where the conditional constraints are indeed defined according to Dempster's rule of combination, and w.r.t. the whole power set of the frame instead of a subalgebra as in Smets' framework. In their rule, however, the conditional constraints are not preserved by their total belief functions.

6 CONCLUSIONS

In this paper we stated and proved the generalisation of the law of total probability to belief measures, for the case in which Dempster's conditioning is employed. We showed that the solution is not unique, whereas it is unique when the a-priori belief function is Bayesian. A critical comparison with Spies' and Smets' results on generalised Jeffrey's rules was also conducted.

These results can be further extended in a number of ways. For instance, distinct versions of the law of total belief may arise by replacing Dempster's conditioning with other accepted forms of conditioning for belief functions, such as credal [9], geometric [31], conjunctive and disjunctive [28] conditioning. As belief functions are a special type of coherent lower probabilities, which in turn can be seen as a special class of lower previsions (consult [32], Section 5.13), marginal extension [19] can be applied to them to obtain a total lower prevision. The relationship between marginal extension and the law of total belief needs therefore to be understood.
Finally, fascinating relationships exist between the total belief problem and transversal matroids [20], on one hand, and the theory of positive linear systems [21], on the other, as hinted at in this paper, which will be investigated in the near future.

References

[1] B. Almond, *Graphical Belief Modeling*, Chapman and Hall/CRC, 1995.

[2] T. Augustin, *Modeling weak information with generalized basic probability assignments*, Data Analysis and Information Systems, pp. 101–113, 1996.

[3] A. Ayoun & Ph. Smets., *Data association in multi-target detection using the transferable belief model*, Int.. J. Intell. Systems **16**(10), 1167–82, 2001.

[4] Y. Bar-Shalom & T. E. Fortmann, *Tracking and data association*, Academic Press, Inc., 1988.

[5] F. Cuzzolin, *Visions of a generalized probability theory*, Lambert Academic Publishing, 2014.

[6] D. Denneberg, *Conditioning (updating) non-additive probabilities*, AOR **52**, 21–42, 1994.

[7] D. Denneberg & M. Grabisch, *Interaction transform of set functions over a finite set*, Information Sciences **121**, 149–170, 1999.

[8] D. Dubois & H. Prade, *A set theoretical view of belief functions*, International Journal of Intelligent Systems **12** (1986), 193–226.

[9] R. Fagin & J.Y. Halpern, *A new approach to updating beliefs*, Proc. of UAI, pp. 347–374, 1991.

[10] I. Gilboa & D. Schmeidler, *Updating ambiguous beliefs*, J. Economic Theory **59**, 33–49, 1993.

[11] W. Gong & F. Cuzzolin, *A belief-theoretical approach to example-based pose estimation*, IEEE Transactions on Fuzzy Systems, 2017.

[12] J. Y. Jaffray, *Bayesian updating and belief functions*, IEEE Tr. SMC **22**, 1144–1152, 1992.

[13] R. Jeffrey, *Conditioning, kinematics, and exchangeability*, Causation, chance, and credence **1**, 221–255, 1988.

[14] ______, *The logic of decision*, Mc Graw, 1965.

[15] ______, *Dracula meets wolfman: Acceptance vs. partial belief*, pp. 157–185, Springer, 1970.

[16] R. Kruse, D. Nauck & F. Klawonn, *Reasoning with mass*, Proc. of UAI, 182–187, 1991.

[17] R. Kruse, E. Schwecke & F. Klawonn, *On a tool for reasoning with mass distribution*, Proc. of IJCAI, vol. 2, pp. 1190–1195, 1991.

[18] J. Ma, W. Liu, D. Dubois & H. Prade, *Bridging Jeffrey's rule, AGM revision and Dempster conditioning in the theory of evidence*, International Journal on AI Tools **20 (4)**, 691–720, 2011.

[19] E. Miranda & G. de Cooman, *Marginal extension in the theory of coherent lower previsions*, Int. J. of Approximate Reasoning **46** (1), 188–225, 2007.

[20] J. G. Oxley, *Matroid theory*, Oxford University Press, 1992.

[21] S. Rinaldi & L. Farina, *I sistemi lineari positivi: teoria e applicazioni*, Cittá Studi Edizioni, 1998.

[22] B. Ristic & Ph. Smets, *The TBM global distance measure for the association of uncertain combat ID declarations*, Inform. Fusion **7**(3), 276–284, 2006.

[23] E. H. Ruspini, *Approximate deduction in single evidential bodies*, CoRR **abs/1304.3104**, 2011.

[24] A. Schrijver, *Theory of Linear and Integer Progamming*, John Wiley & Sons, 1986.

[25] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[26] ______, *Jeffrey's rule of conditioning*, Philosophy of Sciences **48**, 337–362, 1981.

[27] P. Smets and B. Ristic, *Kalman filter and joint tracking and classification in the TBM framework*, Proc. of ICIF, vol. 1, pp. 46–53, 2004.

[28] P. Smets, *Jeffrey's rule of conditioning generalized to belief functions*, Proc. of UAI, 500–505, 1993.

[29] ______, *Belief functions : The disjunctive rule of combination and the generalized Bayesian theorem*, IJAR **9**, 1–35, 1993.

[30] M. Spies, *Conditional events, conditioning, and random sets*, IEEE Tr. SMC **24**, 1755–1763, 1994.

[31] P. Suppes & M. Zanotti, *On using random relations to generate upper and lower probabilities*, Synthese **36**, 427–440, 1977.

[32] P. Walley, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, 1991.

[33] C. Yu & F. Arasta, *On conditional belief functions*, IJAR **10**, 155–172, 1994.

[34] C. Zhou, M. Wang, and B. Qin, *Belief-kinematics Jeffrey's rules in the theory of evidence*, Proc. of UAI, 917–926, 2014.