Coarsening Approximations of Belief Functions

Amel Ben Yaghlane[1], Thierry Denœux[2] and Khaled Mellouli[1]

[1] Institut Supérieur de Gestion de Tunis,
41 Avenue de la liberté, cité Bouchoucha, 2000 Le Bardo - Tunis - Tunisia
byaghlan@planet.tn, khaled.mellouli@ihec.rnu.tn
[2] Université de Technologie de Compiègne
UMR CNRS 6599 Heudiasyc
BP 20529 - F-60205 Compiègne cedex - France
Thierry.Denoeux@hds.utc.fr

Abstract. A method is proposed for reducing the size of a frame of discernment, in such a way that the loss of information content in a set of belief functions is minimized. This approach allows to compute strong inner and outer approximations which can be combined efficiently using the Fast Möbius Transform algorithm.

1 Introduction

The Dempster-Shafer theory of Belief Functions (BF's) is now widely accepted as a rich and flexible framework for representing and reasoning with imperfect information. The concept of belief function subsumes those of probability and possibility measures, making the theory very general. Situations of weak knowledge and heterogeneous information sources are easily modeled within this theory, making it quite suitable in many application domains such as medical diagnosis, sensor fusion and pattern recognition [14].

This generality, however, has a cost in terms of computational complexity. A BF (or, equivalently, a mass function) assigns a number to each of the 2^n subsets of the frame of discernment Ω (with $|\Omega| = n$), with $2^n - 1$ degrees of freedom, which is much larger that what is needed to specify a probability or a possibility measure. Although BF's as elicited from experts or inferred from observation data are usually constrained to be of a simple form, the fusion of several BF's using the Dempster's rule of combination almost inevitably increases the number of focal sets (i.e., subsets of Ω with a positive mass of belief), resulting in high storage and computational requirements for large-scale problems.

The algorithmic complexity of combining several BF's has been studied from a theoretical point of view by Orponen [10], who proved that the problem is $\#P$ complete. Recently, Wilson [16] provided a very complete review of algorithmic issues related to the manipulation of BF's. Currently, two algorithms exist for computing the conjunctive combination $m_1 \cap m_2$ of two mass functions m_1 and m_2 (similar methods hold for the disjunctive combination):

- the mass-based algorithm, initially sketched by Shafer, involves considering each focal set A of m_1, each focal set B of m_2, and assigning the mass

$m_1(A)m_2(B)$ to the set $A \cap B$. Using this method, the combination can be performed in time proportional to $n|\mathcal{F}(m_1)||\mathcal{F}(m_2)|$, where $\mathcal{F}(m_i)$ denotes the number of focal sets of m_i $(i = 1, 2)$. The time needed for the combination of K BF's $m_1, \ldots, m_K$ depends on the particular structure of the mass functions, and is at worst roughly proportional to $n \prod_{i=1}^{K} |\mathcal{F}(m_i)|$, as shown by Wilson [16].

- the Fast Möbius Transform (FMT) method [8] converts each mass function m_i into its associated commonality function q_i; the product of these functions is computed, and the result is converted back into a mass function. The algorithm takes time proportional to $Kn^2 2^n$.

The choice of one of these methods depends on the structure of the mass functions. As remarked by Wilson, if the number of focal sets of the combined belief function is much small than 2^n, then the mass-based method is likely to be faster. However, this is generally not known in advance. If one of the BF's has a number of focal sets close to 2^n, then the FMT method is likely to be better. However, this method becomes impractical when Ω has more than 15 to 20 elements.

When the combination of several BF's cannot be computed exactly, one has to resort to stochastic or deterministic approximation procedures [16]. Since the mass-based method for combining BF's is the most widely used, most deterministic methods (which are exclusively considered here) have been designed with the aim of reducing the number of focal elements. This is true, in particular, for the summarization method initially introduced by Lowrance et al. [6], and for the more sophisticated methods proposed subsequently [15] [1] [5] [11] [2].

In this paper, a different approach is investigated. Instead of reducing the number of focal elements, we propose to reduce the size of the frame of discernment, which can be expected to drastically decrease the computing time of the FMT combination method, and can even make it applicable to find reasonable approximations in the case of large-size problems. Given a set of BF's, we propose to find a coarsening of the frame Ω that will preserve as much as possible of the information content of the belief functions. This approach allows to compute inner and outer approximations, from which lower and upper bounds for the combined belief values can be derived.

The following section summarizes the background definitions and results needed in the sequel. Our approximation method is then described in Section 3, and a simulation example is presented in Section 4.

2 Background

2.1 Basic Concepts

The main concepts of evidence theory are only summarized here. More details can be found in Refs. [12] and [13]. Let Ω denote a finite set called the frame of discernment. A mass function, or *basic belief assignment*(bba) is a function $m : 2^\Omega \to [0, 1]$ verifying:

$$\sum_{A \subseteq \Omega} m(A) = 1. \tag{1}$$

Each mass of belief $m(A)$ measures the amount of belief that is exactly committed to A. A bba m such that $m(\emptyset) = 0$ is said to be normal. This condition will not be imposed here. The subsets A of Ω such that $m(A) > 0$ are called *focal sets* of m. Let $\mathcal{F}(m) \subseteq 2^{\Omega}$ denote the set of focal sets of m.

The *belief function* induced by m is a function bel : $2^{\Omega} \rightarrow [0,1]$, defined as:

$$\text{bel}(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \tag{2}$$

for all $A \subseteq \Omega$. bel(A) represents the amount of support given to A.

The *plausibility function* associated with a bba m is a function pl : $2^{\Omega} \rightarrow [0,1]$, defined as:

$$\text{pl}(A) = \sum_{\emptyset \neq B \cap A} m(B) \qquad \forall A \subseteq \Omega. \tag{3}$$

pl(A) represents the potential amount of support that could be given to A.

Given two bba's m_1 and m_2 defined over the same frame of discernment Ω and induced by two distinct pieces of evidence, we can combine them in two ways using the conjunctive or the disjunctive rules of combination [13] defined, respectively, as:

$$(m_1 \textcircled{\cap} m_2)(A) = \sum_{B \cap C = A} m_1(B) m_2(C) \tag{4}$$

$$(m_1 \textcircled{\cup} m_2)(A) = \sum_{B \cup C = A} m_1(B) m_2(C) \tag{5}$$

for all $A \subseteq \Omega$. The choice of one of these combination rules is related to the reliability of the two sources. In fact, if we know that both sources of information are fully reliable, then we combine them conjunctively. However, if we only know that at least one of the two sources is reliable, then we combine them disjunctively.

The conjunctive and disjunctive rules can be conveniently expressed by means of the commonality function q and the implicability function b, defined, respectively, as

$$q(A) = \sum_{A \subseteq B} m(B) \tag{6}$$

and

$$b(A) = \text{bel}(A) + m(\emptyset) \tag{7}$$

for all $A \subseteq \Omega$. If $q_1 \textcircled{\cap} q_2$ denotes the commonality function associated to $m_1 \textcircled{\cap} m_2$, and $b_1 \textcircled{\cup} b_2$ denotes the implicability function associated to $m \textcircled{\cup} m_2$, we have the following simple relations:

$$q_1 \textcircled{\cap} q_2 = q_1 q_2 \tag{8}$$

$$b_1 \textcircled{\cup} b_2 = b_1 b_2 \tag{9}$$

The importance of this result arises from the fact that the functions m, q and b (as well as bel and pl) are equivalent representations, in the sense that, given any of these functions, it is possible to recover all the others. The conversion between these functions can be efficiently done using the FMT algorithm [8] in time proportional to $n^2 2^n$ [16]. Relations (8) and (9) provide the basis for the FMT-based method for combining BF's, which consists in transforming the BF's or the bba's to q or b, computing the product, and converting back the result into a mass or a belief function. In contrast, the more traditional mass-based approach relies exclusively on Eqs (4) and (5).

2.2 Coarsenings and Refinements

In applying the BF framework to a real-world problem, the definition of the frame of discernment is a crucial step. As remarked by Shafer [12], the degree of "granularity" of the frame is always a matter of convention, as any element ω of Ω representing a "state of nature" could always be split into several possibilities. Hence, it is fundamental to examine how a BF defined on a frame may be expressed in a finer or, conversely, in a coarser frame.

Let Ω and Θ denote two finite sets. A mapping $\rho : 2^\Theta \to 2^\Omega$ is called a *refining* if it verifies the following properties:

1. The set $\{\rho(\{\theta\}), \theta \in \Theta\} \subseteq 2^\Omega$ is a partition of Ω.
2. For all $A \subseteq \Theta$, we have

$$\rho(A) = \bigcup_{\theta \in A} \rho(\{\theta\}) \tag{10}$$

Following the terminology introduced by Shafer, the set Θ is then called a *coarsening* of Ω, and Ω is called a *refinement* of Θ.

Note that defining a coarsening of a frame Ω is formally equivalent to defining a partition of Ω. Let Θ be such a partition. The function $\rho : 2^\Theta \to 2^\Omega$ such that $\rho(\{\theta\}) = \theta$ for all $\theta \in \Theta$, and verifying (10) is a refining of Θ, and Θ is a coarsening of Ω.

A bba m^Θ defined on a frame Θ may easily be carried to a refinement Ω by means of the vacuous extension, which transfers the mass $m^\Theta(A)$ to $\rho(A)$, for all $A \subseteq \Theta$ (in the following, the superscript of a bba will always indicate its domain). The resulting bba m^Ω on Ω is then defined as

$$m^\Omega(B) = \begin{cases} m^\Theta(A), & \text{if } B = \rho(A) \quad \text{for some } A \subseteq \Theta \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

The inverse operation, i.e., carrying a bba m^Ω to a coarsening Θ of Ω is not so easy because a refining $\rho : 2^\Theta \to 2^\Omega$ is not, in general, onto; there are usually subsets A of Ω which are not "discerned" by Θ and, hence, are not equal to $\rho(B)$ for any $B \subseteq \Theta$ [12]. In order to associate a subset of Θ with each subset A of Ω, an *inner reduction* $\underline{\theta}$ and an *outer reduction* $\overline{\theta}$ may be defined, respectively, as functions from 2^Ω to 2^Θ, such that:

$$\underline{\theta}(A) = \{\theta \in \Theta | \rho(\{\theta\}) \subseteq A\} \tag{12}$$

$$\overline{\theta}(A) = \{\theta \in \Theta | \rho(\{\theta\}) \cap A \neq \emptyset\} \tag{13}$$

for all $A \subseteq \Omega$. Hence, the mass $m^\Omega(A)$ given to $A \subseteq \Omega$ by a bba m^Ω can be transferred either to $\underline{\theta}(A)$, or to $\overline{\theta}(A)$. This leads to the following definitions:

$$\underline{m}^\Theta(B) = \sum_{\{A\subseteq\Omega, B=\underline{\theta}(A)\}} m^\Omega(A) \quad \forall B \subseteq \Theta \tag{14}$$

$$\overline{m}^\Theta(B) = \sum_{\{A\subseteq\Omega, B=\overline{\theta}(A)\}} m^\Omega(A) \quad \forall B \subseteq \Theta. \tag{15}$$

The bba's $\underline{m}^\Theta$ and $\overline{m}^\Theta$ will be called, respectively, the inner and the outer reduction of m^Ω ($\overline{m}^\Theta$ is called the restriction of m^Ω par Shafer [12, p. 126]; the definition of $\underline{m}^\Theta$ is, to our knowledge, new).

To simplify the manipulation of expressions when changing frames, let us introduce the following definition.

Definition 1 *Let Ω_1 and Ω_2 be two finite sets, φ an application from 2^{Ω_1} to 2^{Ω_2}, m^{Ω_1} a bba on Ω_1, and m^{Ω_2} a bba on Ω_2. We say that m^{Ω_2} is the image of m^{Ω_1} by φ, and we note $m^{\Omega_2} = \varphi(m^{\Omega_1})$, if*

$$m^{\Omega_2}(A) = \sum_{\{B\subseteq\Omega_1, \varphi(B)=A\}} m^{\Omega_1}(B)$$

for all $A \subseteq \Omega_2$.

According to Def. 1, the vacuous extension of m^Θ in Ω may be noted $m^\Omega = \rho(m^\Theta)$, and Eqs (14) and (15) may be rewritten as $\underline{m}^\Theta = \underline{\theta}(m^\Omega)$ and $\overline{m}^\Theta = \overline{\theta}(m^\Omega)$.

2.3 Inclusion of Belief Functions

Another notion of interest is that of strong inclusion of bba's [3]. Let m and m' be two BS's with focal elements $\mathcal{F}(m) = \{F_1, \ldots, F_p\}$ and $\mathcal{F}(m') = \{F'_1, \ldots, F'_{p'}\}$. Then m is said to be strongly included in m', or to be a *specialization* of m' (noted $m \subseteq m'$), iff there exists a non-negative matrix W with entries w_{ij} ($i = 1, \ldots, p; j = 1, \ldots, p'$) such that

$$\sum_{j=1}^{p'} w_{ij} = m(F_i), \quad i = 1, \ldots, p, \tag{16}$$

$$\sum_{i=1}^{p} w_{ij} = m'(F'_j), \quad j = 1, \ldots, p' \tag{17}$$

and $w_{ij} > 0 \Rightarrow F_i \subseteq F'_j$. The relationship between m and m' may be seen as a transfer of mass from each focal element F_i of m to supersets $F'_j \supseteq F_i$, the quantity w_{ij} denoting the part of $m(F_i)$ transferred to F'_j. If $m \subseteq m'$, then we have (with obvious notations) $\mathrm{pl} \leq \mathrm{pl}'$ and $b' \leq b$, but the reverse is not true.

An approximation $\widehat{m}^-$ (resp. $\widehat{m}^+$) of a bba m is called a strong inner (resp. outer) approximation if $\widehat{m}^- \subseteq m$ (resp. $m \subseteq \widehat{m}^+$). Given strong inner and outer approximations of several BF's, it is possible to obtain lower and upper bounds for the belief and the plausibility values of the combined BF [3][2]. Methods for constructing such approximations were propose by Dubois and Prade [4] in a possibilistic setting, and by Denœux [2] using an approach based on the clustering of focal sets.

3 Coarsening Approximations of Belief Functions

In this section, we propose a new heuristic method for constructing strong inner and outer approximations of BF's. Our method consists in finding a coarsening Θ of the initial frame Ω such that the approximating BF can be represented exactly in Θ. We first present the basic principle and the algorithm in the case of a single BF, and then extend the method to the simultaneous approximation of several BF's.

3.1 Basic Principle

Main result Let m^Ω denote a bba on Ω, Θ a coarsening of Ω, ρ the refining from 2^Θ to 2^Ω, and $\underline{\theta}$ and $\overline{\theta}$ the associated inner and outer reduction functions. Let $\underline{m}^\Theta$ and $\overline{m}^\Theta$ denote the inner and outer reductions of m^Ω as defined by Eqs (14) and (15), and let $\underline{m}^\Omega$ and $\overline{m}^\Omega$ be the vacuous extensions of $\underline{m}^\Theta$ and $\overline{m}^\Theta$, respectively, on Ω. We thus have

$$\underline{m}^\Omega = \rho(\underline{m}^\Theta) = \rho \circ \underline{\theta}(m^\Omega) \tag{18}$$

$$\overline{m}^\Omega = \rho(\overline{m}^\Theta) = \rho \circ \overline{\theta}(m^\Omega) \tag{19}$$

Theorem 1 *$\underline{m}^\Omega$ and $\overline{m}^\Omega$ are, respectively, strong inner and outer approximations of m^Ω: $\underline{m}^\Omega \subseteq m^\Omega \subseteq \overline{m}^\Omega$*

Proof: We have, by construction,

$$\underline{m}^\Omega(A) = \sum_{\{B\subseteq\Omega, A=\rho\circ\underline{\theta}(B)\}} m^\Omega(B) \quad \forall A \subseteq \Omega \tag{20}$$

$$\overline{m}^\Omega(A) = \sum_{\{B\subseteq\Omega, A=\rho\circ\overline{\theta}(B)\}} m^\Omega(B) \quad \forall A \subset \Omega \tag{21}$$

From Theorem 6.3 in [12, p.118], we have $\rho(\underline{\theta}(B)) \subseteq B$ for all $B \subseteq \Omega$. Hence, the mass $\underline{m}^\Omega(A)$ is the sum of masses $m^\Omega(B)$ initially attached to supersets of A, which implies that $\underline{m}^\Omega \subseteq m^\Omega$.

Similarly, $B \subseteq \rho(\overline{\theta}(B))$ for all $B \subset \Omega$, which implies that the mass $\overline{m}^\Omega(A)$ is the sum of masses $m^\Omega(B)$ initially attached to subsets of A, which implies that $m^\Omega \subseteq \overline{m}^\Omega$. QED

Matrix representation of bba's A very simple construction of $\underline{m}^{\Omega}$ and $\overline{m}^{\Omega}$ for a given coarsening Θ can be obtained using the following representation. Let us assume that the frame $\Omega = \{\omega_1, \ldots \omega_n\}$ has n elements, and the bba m^{Ω} under consideration has p focal sets: $\mathcal{F}(m^{\Omega}) = \{A_1, \ldots, A_p\}$. One can represent the bba m^{Ω} by a pair $(\mathbf{m}^{\Omega}, \mathbf{F}^{\Omega})$ where $\mathbf{m}^{\Omega}$ is the p-dimensional vector of masses $\mathbf{m}^{\Omega} = (m^{\Omega}(A_1), \ldots, (m^{\Omega}(A_p))$ and $\mathbf{F}^{\Omega}$ is a $p \times n$ binary matrix such that

$$\mathbf{F}^{\Omega}_{ij} = A_i(\omega_j) = \begin{cases} 1, & \text{if } \omega_j \in A_i \\ 0, & \text{otherwise.} \end{cases}$$

where $A_i(\cdot)$ denotes the indicator function of focal set A_i.

This representation is similar to an (objects $\times$ attributes) binary data matrix as commonly encountered in data analysis. Here, each focal set corresponds to an object, and each element of the frame corresponds to an attribute. Each object A_i has a weight $m^{\Omega}(A_i)$. Since a coarsening is inherently equivalent to a partition of Ω, finding a suitable coarsening is actually a problem of classifying the columns of data matrix $\mathbf{F}$, which is a classical clustering problem (see, e.g. [7]). Note that, in contrast, the clustering approximation method introduced by Denœux [2] is based on the classification of the lines of $\mathbf{F}$.

To see how the bba's $\underline{m}^{\Theta}$, $\overline{m}^{\Theta}$, $\underline{m}^{\Omega}$, $\overline{m}^{\Omega}$ can be constructed from $\mathbf{F}$, let us denote by $P = \{I_1, \ldots, I_c\}$ the partition of $N_n = \{1, \ldots, n\}$ corresponding to the coarsening $\Theta = \{\theta_1, \ldots, \theta_c\}$, i.e.,

$$\theta_r = \{\omega_j, j \in I_r\} \qquad r = 1, \ldots, c.$$

Let $(\underline{\mathbf{m}}^{\Theta}, \underline{\mathbf{F}}^{\Theta})$ denote the matrix representation of $\underline{m}^{\Theta}$. Matrix $\underline{\mathbf{F}}^{\Theta}$ may be obtained from $\mathbf{F}^{\Omega}$ by merging the columns $\mathbf{F}^{\Omega}_{\cdot j}$ for $j \in I_r$, and replacing them by their minimum:

$$\underline{\mathbf{F}}^{\Theta}_{i,r} = \min_{j \in I_r} \mathbf{F}^{\Omega}_{i,j} \quad \forall i, r \tag{22}$$

and we have $\underline{\mathbf{m}}^{\Theta} = \mathbf{m}^{\Omega}$. The justification for this is that the focal elements of $\underline{m}^{\Theta}$ are the sets $\underline{\theta}(A_i)$, and $\theta \in \underline{\theta}(A_i)$ iff $\rho(\theta) \subseteq A_i$, where ρ is the refining associated to Θ.

similarly, if $(\overline{\mathbf{m}}^{\Theta}, \overline{\mathbf{F}}^{\Theta})$ denotes the matrix representation of $\overline{m}^{\Theta}$, we have

$$\overline{\mathbf{F}}^{\Theta}_{i,r} = \max_{j \in I_r} \mathbf{F}^{\Omega}_{i,j} \quad \forall i, r \tag{23}$$

and $\overline{\mathbf{m}}^{\Theta} = \mathbf{m}^{\Omega}$.

The matrix representations of $\underline{m}^{\Omega}$ and $\overline{m}^{\Omega}$, the vacuous extensions of $\underline{m}^{\Theta}$ and $\overline{m}^{\Theta}$, are then obtained as:

$$\underline{\mathbf{F}}^{\Omega}_{i,j} = \underline{\mathbf{F}}^{\Theta}_{i,r} \quad \forall j \in I_r \tag{24}$$

$$\overline{\mathbf{F}}^{\Omega}_{i,j} = \overline{\mathbf{F}}^{\Theta}_{i,r} \quad \forall j \in I_r \tag{25}$$

and $\underline{\mathbf{m}}^{\Omega} = \overline{\mathbf{m}}^{\Omega} = \mathbf{m}^{\Omega}$.

3.2 Clustering Algorithm

As shown above, given a coarsening Θ of a frame of discernment Ω and a basic belief assignment m^Ω, we can define strong inner and outer approximations $\underline{m}^\Omega$ and $\overline{m}^\Omega$. It is clear that the quality of these approximations depends on the coarsenings considered, then how to choose these coarsenings so as to obtain good approximations of m^Ω?

To answer this question, we propose to use a measure of information allowing us to reduce the size of the frame of discernment while *retaining as much information as possible* from the original belief function. Several approaches have been proposed to measure the information contained in a piece of evidence [9]. Among these approaches, we will use the *generalized cardinality* [4, 2] defined as:

$$|m| = \sum_{i=1}^{p} m(A_i)|A_i|, \tag{26}$$

where A_i, $i = 1, \ldots, p$ are the focal sets of m. The bba m is all the more imprecise (and contains all the less information) that $|m|$ is large.

It follows from Theorem 1 and the definition of strong inclusion that

$$|\underline{m}^\Omega| \leq |m^\Omega| \leq |\overline{m}^\Omega|$$

Hence, a way to keep $\underline{m}^\Omega$ and $\underline{m}^\Omega$ as "close" as possible to m^Ω is to minimize the increase of cardinality from m^Ω to $\overline{m}^\Omega$ (which correspond to a loss of information), and to minimize the decrease of cardinality from m^Ω to $\underline{m}^\Omega$ (corresponding to meaningless information).

More precisely, let us denote by $\mathcal{P}_c$ the set of all partitions of N_n in c classes ($c < n$). As shown above, each element of $\mathcal{P}_c$ corresponds to a coarsening of Ω with c elements. The coarsening yielding the "best" (least specific) inner approximation corresponds to the partition $\underline{P}_c$ defined as:

$$\underline{P}_c = \arg\min_{P \in \mathcal{P}_c} \Delta(m^\Omega, \underline{m}^\Omega)$$

with $\Delta(m^\Omega, \underline{m}^\Omega) = |m^\Omega| - |\underline{m}^\Omega|$. Similarly, the partition $\overline{P}_c$ yielding the best (most specific) outer approximation is defined as

$$\overline{P}_c = \arg\min_{P \in \mathcal{P}_c} \Delta(\overline{m}^\Omega, m^\Omega).$$

We are thus searching for the best coarsening over all possible partitions of Ω into c clusters. Unfortunately, the number of possible partitions is huge, and exploring all of them is not computationally tractable. Hierarchical clustering [7] is a heuristic approach for constructing a sequence of nested partitions of a given set. In our case, this approach will consist in aggregating sequentially pairs of elements of Ω until the desired size of the coarsened frame of discernment is reached. At each step, the two elements whose aggregation results in the best value of the criterion will be selected.

More precisely, let $(\mathbf{m}^\Omega, \mathbf{F}^\Omega)$ denote the matrix representation of m^Ω, and suppose that we are looking for the coarsening with $n-1$ elements corresponding to the "best" inner approximation. The aggregation of elements ω_j and ω_k of the frame corresponds to the fusion of columns j and k of $\mathbf{F}^\Omega$ using the minimum operator. In this process, the number of 1's in each line i of matrix $\mathbf{F}^\Omega$ is decreased by one if either $\omega_j \in A_i$ and $\omega_k \notin A_i$, or $\omega_k \in A_i$ and $\omega_j \notin A_i$. Hence, the decrease of cardinality is

$$\delta(\omega_k, \omega_l) = \Delta(m^\Omega, \underline{m}^\Omega) = \sum_{i=1}^{p} \mathbf{m}_i |\mathbf{F}^\Omega_{ij} - \mathbf{F}^\Omega_{il}| \tag{27}$$

Note that $\delta(\omega_k, \omega_l)$ can be interpreted as a degree of dissimilarity between ω_j and ω_l. The hierarchical clustering algorithm can then be described as follows:

- Given: the bba $(\mathbf{m}^\Omega, \mathbf{F}^\Omega)$
- Compute the dissimilarity matrix $D = (\delta(\omega_k, \omega_l)), k, l \in \{1, \ldots, n\}$
- $c \leftarrow n$
- Repeat
  - $c \leftarrow c - 1$
  - find k^* and l^* such that $\delta(\omega_{k^*}, \omega_{l^*}) = \min_{k,l} \delta(\omega_k, \omega_l)$
  - construct $\underline{\mathbf{F}}^\Theta$ with c columns by aggregating columns k^* and l^* using the minimum operator
  - update dissimilarity matrix D
- Until c has the desired value
- Compute $(\underline{\mathbf{m}}^\Omega, \underline{\mathbf{F}}^\Omega)$, the vacuous extension of $(\underline{\mathbf{m}}^\Theta, \underline{\mathbf{F}}^\Theta)$

The computation of outer approximations can be performed in exactly the same way, except that the minimum operator is replaced by the maximum operator. After aggregating columns k and l of matrix $\mathbf{F}^\Omega$, the number of 1's in each line i of matrix $\mathbf{F}^\Omega$ is now increased by one if either $\omega_j \in A_i$ and $\omega_k \notin A_i$, or $\omega_k \in A_i$ and $\omega_j \notin A_i$. Hence, the increase of cardinality is

$$\Delta(\overline{m}^\Omega, m^\Omega) = \sum_{i=1}^{n} \mathbf{m}_i |\mathbf{F}^\Omega_{ij} - \mathbf{F}^\Omega_{il}| = \delta(\omega_k, \omega_l) \tag{28}$$

We thus arrive at the same dissimilarity measure as in the previous case, although the resulting coarsening is, in general, different.

Remark 1 *Several lines of $\underline{\mathbf{F}}^\Omega$ or $\overline{\mathbf{F}}^\Omega$ computed by the above algorithm may be identical, which means that the number of focal sets has decreased. In this case, the binary matrix of focal sets and the mass vector have to be rearranged so that the line dimension becomes equal to the number of focal sets.*

Remark 2 *As remarked by Wilson [16], coarsening a frame may sometimes result in no loss of information. Two elements ω_j and ω_k can be merged without losing information if $\delta(\omega_j, \omega_k) = 0$. Hence, "lossless coarsenings" (using Wilson's terminology) will be found in the first steps of our algorithm, if such solutions exist. Our algorithm will even find the "coarsest lossless coarsening" as defined by Wilson [16].*

Remark 3 *Our algorithm is basically the classical hierarchical clustering algorithm applied to the binary matrix of focal sets. Hence, the time needed to compute an inner or outer coarsening approximation by this method is proportional to* n^3.

3.3 Inner and Outer Approximations of Combined Belief Functions

The approximation method proposed in the previous section can be generalized to compute inner and outer approximations of combined belief functions. Rather than computing the combination of the original belief functions defined on Ω, we will compute the combination of their approximations defined over a common coarsened frame of Ω using the FMT algorithm [8]. Then the vacuous extension defined above will be used to recover the combined belief function on the original frame Ω from its approximations defined over the coarsened frames.

Let $m_1^\Omega, ..., m_K^\Omega$ be K bba's defined over a frame of discernment Ω to be combined using either the conjunctive or the disjunctive rules of combination. Let $(\mathbf{m}_k^\Omega, \mathbf{F}_k^\Omega)$, $k = 1, \ldots, K$ denote their matrix representations. We wish to find a common coarsening $\Theta = \{\theta_1, ..., \theta_c\}$ of Ω that will preserve as much as possible of the information contained in each of the K bba's. For that purpose, let us define the following criterion to be minimized for the construction of an inner approximation: $\sum_{k=1}^K \Delta(m_k^\Omega, \underline{m}_k^\Omega)$, and for the construction of an outer approximation: $\sum_{k=1}^K \Delta(\overline{m}_k^\Omega, m_k^\Omega)$. To minimize these criteria, we may simply apply the same hierarchical clustering approach as above, to the matrix

$$\mathbf{F}^\Omega = \begin{bmatrix} \mathbf{F}_1^\Omega \\ \vdots \\ \mathbf{F}_K^\Omega \end{bmatrix}$$

and the weight vector $\mathbf{m}^\Omega = [\mathbf{m}_1^\Omega, \ldots, \mathbf{m}_K^\Omega]'$ (prime denotes transposition).

Determining Inner and Outer approximations of the Combined Belief Function. Given K bba's $\underline{m}_1^\Theta, \ldots, \underline{m}_K^\Theta$ and $\overline{m}_1^\Theta, \ldots, \overline{m}_K^\Theta$ defined over the common coarsened frame Θ of Ω, we shall proceed as follows to determine strong inner and outer approximations of their combination:

1. use the FMT algorithm to convert these approximated bba's to their related inner and outer commonality or implicability functions.
2. compute the approximated inner and outer combined commonality or implicability functions over the coarsened frame Θ. In the case of inner approximation they are given by: $\underline{q}^\Theta = \prod_{i=1}^K \underline{q}_i^\Theta$ and $\underline{b}^\Theta = \prod_{i=1}^K \underline{b}_i^\Theta$, and similarly for the outer approximations $\overline{q}^\Theta$ and $\overline{b}^\Theta$.
3. convert back these approximated combined commonality or implicability functions to their related inner and outer combined bba's $\underline{m}^\Theta$ and $\overline{m}^\Theta$ using the FMT algorithm.
4. use the vacuous extension to recover the inner and outer approximated combined belief function $\underline{m}^\Omega$ and $\overline{m}^\Omega$ from $\underline{m}^\Theta$ and $\overline{m}^\Theta$.

4 Simulations

As an example, we simulated the conjunctive combination of 3 bba's on a frame Ω with $n = |\Omega| = 30$, with 500 focal sets each. The focal sets were generated randomly in such a way that element ω_i of the frame had probability $(i/(n+1))^2$ to belong to each focal set. Hence, we simulate the realistic situation in which some single hypotheses are more plausible than others. The masses were assigned to focal sets as proposed by Tessem [15]: the mass given to the first one was taken from a uniform distribution on $[0, 1]$, then a random fraction of the rest was given to the second one, etc. The remaining part of the unit mass was finally allocated to the last focal set. The conjunctive sum of the 3 bba's was approximated using the method described above, using a coarsening of size $c = 10$.

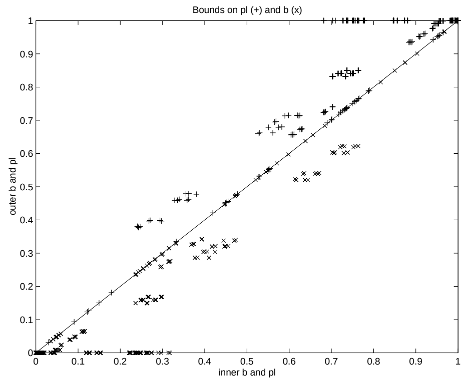

Fig. 1. Simulation results

A part of the results is shown in Fig. 1. The plausibilities and implicabilities $\underline{\mathrm{pl}}^{\Omega}(A)$ and $\underline{b}^{\Omega}(A)$ are plotted on the x axis against $\overline{\mathrm{pl}}^{\Omega}(A)$ and $\overline{b}^{\Omega}(A)$, for 1000 randomly selected subsets of Ω. As expected, we obtain a bracketing of the true plausibilities and implicabilities for any A, since $\underline{\mathrm{pl}}^{\Omega}(A) \leq \mathrm{pl}^{\Omega}(A) \leq \overline{\mathrm{pl}}^{\Omega}(A)$ and $\underline{b}^{\Omega}(A) \geq b^{\Omega}(A) \geq \overline{b}^{\Omega}(A)$. A bracketing of $\mathrm{bel}^{\Omega}(A)$ could also be obtained, as shown by Denœux [2].

5 Conclusion

A new method for computing inner and outer approximations of BF's has been defined. Unlike previous approaches, this method does not rely on the reduction of the number of focal sets, but on the construction of a coarsened frame in which combination can be performed efficiently using the FMT algorithm. Joint strategies aiming at reducing the number of focal sets or the size of the frame, depending on the problem at hand, could be considered as well, and are left for further study.

References

1. M. Bauer. Approximation algorithms and decision making in the Dempster-Shafer theory of evidence – an empirical study. *International Journal of Approximate Reasoning*, 17:217–237, 1997.
2. T. Denœux. Inner and outer clustering approximations of belief structures. In *Proceedings of IPMU'2000*, volume 1, pages 125–132, Madrid, July 2000.
3. D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Comput. Intell.*, 4:244–264, 1988.
4. D. Dubois and H. Prade. Consonant approximations of belief measures. *International Journal of Approximate Reasoning*, 4:419–449, 1990.
5. D. Harmanec. Faithful approximations of belief functions. In K. B. Laskey and H. Prade, editors, *Uncertainty in Artificial Intelligence 15 (UAI99)*, Stockholm, Sweden, 1999.
6. Lowrance J. D, T. D. Garvey, and T. M. Strat. A framework for evidential-reasoning systems. In T. Kehler et al., editor, *Proceedings of AAAI'86*, volume 2, pages 896–903, Philadelphia, August 1986. AAAI.
7. A. K. Jain and R. C. Dubes. *Algorithms for clustering data.* Prentice-Hall, Englewood Cliffs, NJ., 1988.
8. R. Kennes. Computational aspects of the Möbius transform of graphs. *IEEE Trans. SMC*, 22:201–223, 1992.
9. G. J. Klir and M. J. Wierman. *Uncertainty-Based Information. Elements of Generalized Information Theory.* Springer-Verlag, New-York, 1998.
10. P. Orponen. Dempster's rule of combination is # p-complete. *Artificial Intelligence*, 44:245–253, 1990.
11. S. Petit-Renaud and T. Denœux. Handling different forms of uncertainty in regression analysis: a fuzzy belief structure approach. In A. Hunter and S. Pearsons, editors, *Symbolic and quantitative approaches to reasoning and uncertainty (ECSQARU'99)*, pages 340–351, London, June 1999. Springer Verlag.
12. G. Shafer. *A mathematical theory of evidence.* Princeton University Press, Princeton, N.J., 1976.
13. P. Smets. The Transferable Belief Model for quantified belief representation. In D. M. Gabbay and P. Smets, editors, *Handbook of Defeasible reasoning and uncertainty management systems*, volume 1, pages 267–301. Kluwer Academic Publishers, Dordrecht, 1998.
14. P. Smets. Practical uses of belief functions. In K. B. Laskey and H. Prade, editors, *Uncertainty in Artificial Intelligence 15 (UAI99)*, pages 612–621, Stockholm, Sweden, 1999.
15. B. Tessem. Approximations for efficient computation in the theory of evidence. *Artificial Intelligence*, 61:315–329, 1993.
16. N. Wilson. Algorithms for dempster-shafer theory. In D. M. Gabbay and P. Smets, editors, *Hanbook of defeasible reasoning and uncertainty management. Volume 5: Algorithms for Uncertainty and Defeasible Reasoning*, pages 421–475. Kluwer Academic Publishers, Boston, 2000.