A generalization of local divergence measures

Carlo Bertoluzza
Dipartimento de Informatica e Sistemistica
INFM Section of Pavia
University of Pavia, Italy
carlo.bertoluzza@unipv.it

Pedro Miranda
Departamento de Estadística e I.O.
Complutense University of Madrid, Spain
pmiranda@mat.ucm.es

Pedro Gil
Departamento de Estadística, I.O. y D. M.
University of Oviedo, Spain
pedro@pinon.ccu.uniovi.es

Abstract

In this paper we propose a generalization of the concept of the local property for divergence measures. These new measures will be called g-local divergence measures, and we study some of their properties. Once this family is defined, a characterization based on Ling's Theorem is given. From this result, we obtain the general form of g-local divergence measures as a function of the divergence in each element of the reference set; this study is divided in three parts according to the cardinality of the reference set: finite, infinite countable or non-countable. Finally, we study the problem of componible divergence measures as a dual concept of g-local divergence measures.

Keywords: Divergence measures, local property, Ling's Theorem, componibility.

1 Introduction

Recently, Prof. Montes [10, 11] has developed the concept of divergence measures between two fuzzy subsets. Her approach is based on three axioms modelling the minimal requirements for a function that tries to measure the separation or difference between two fuzzy subsets.

From this first definition, it is possible to define measures of divergence between two fuzzy partitions [13] and, at the present time, the applications of these measures to the Theory of fuzzy questionnaires are being studied [14]. On the other hand, divergence measures have revealed themselves as an interesting tool in Decision Making [9].

However, an important problem in the theory of divergence measures derives from the fact that the conditions of divergence measure are too general and thus, the class of all divergence measures is very wide. As a consequence, it is rather hard to find properties beyond the axioms. This leads us to work with different subclasses of divergence measures, considering additional properties.

Among all these subclasses, the most important ones from a mathematical point of view and also in terms of the interpretation, are the so-called local and componible divergence measures.

Unfortunately, the local (resp. componibility) property is too restrictive in many problems. The goal of this paper is to define a new property, that we will call g-local property for generalized local property, extending the local property and increasing the practical situations in which it can be applied, but keeping at the same time its mathematical properties.

The paper is organized as follows: next section is devoted to preliminary concepts and results that we will need in the rest of the paper. Then, in Section 3 we define g-local divergence measures and we study some of their properties; we also give a representation theorem based on Ling's theorem. In Sections 4 and 5 we study the consequences of this result for finite, infinite countable and infinite non-countable references. In Section 6 we deal with the problem of componibility. Sections 7 and 8 are devoted to conclusions, open problems and acknowledgements.

2 Basic concepts

Let us start with the basic concepts and results that will be needed throughout the paper. In the sequel, we will use the following notations: Ω is the reference set; crisp subsets of Ω are denoted by capital letters A, B and so on, while fuzzy subsets are denoted by $\tilde{A}, \tilde{B}$, ...; the set of all crisp subsets of Ω is denoted by $\mathcal{P}(\Omega)$ and the set of all fuzzy subsets by $\tilde{\mathcal{P}}(\Omega)$. The membership function of the fuzzy subset $\tilde{A}$ at the point $x \in \Omega$ will be denoted by $\tilde{A}(x)$. We will also consider the standard fuzzy union, fuzzy intersection and fuzzy complementary, i.e.

$$(\tilde{A} \cup \tilde{B})(x) = \max\{\tilde{A}(x), \tilde{B}(x)\},$$
$$(\tilde{A} \cap \tilde{B})(x) = \min\{\tilde{A}(x), \tilde{B}(x)\},$$
$$\tilde{A}^c(x) = 1 - \tilde{A}(x).$$

We will denote by $\tilde{E}$ the equilibrium subset, i.e. the subset such that $\tilde{E}(x) = \frac{1}{2}, \forall x \in \Omega$. Finally, we will denote by $\tilde{A}_C$ the fuzzy subset defined for any $C \in \mathcal{P}(\Omega)$ by

$$\tilde{A}_C(x) = \left\{ \begin{array}{ll} \tilde{A}(x) & \text{if } \; x \in C \\ 0 & \text{otherwise} \end{array} \right\} = \tilde{A} \cap C(x), \forall x \in \Omega.$$

Definition 1 *Let $\tilde{A}$ be a fuzzy subset of Ω. We will call the* **crisp subset closest** *to $\tilde{A}$ the subset given by:*

$$N_{\tilde{A}}(x) = \begin{cases} 1 & \text{if} \quad \tilde{A}(x) \geq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

Definition 2 *[12] Let Ω be a reference set. A function $D : \tilde{\mathcal{P}}(\Omega) \times \tilde{\mathcal{P}}(\Omega) \mapsto \mathbb{R}$ is called a* **divergence measure between two fuzzy subsets** *(or* **divergence measure***) if and only if it satisfies the following conditions for all $\tilde{A}, \tilde{B}, \tilde{C} \in \tilde{\mathcal{P}}(\Omega)$:*

1. $D(\tilde{A}, \tilde{A}) = 0$.
2. $D(\tilde{A}, \tilde{B}) \geq 0$.
3. $D(\tilde{A}, \tilde{B}) \geq \min\{D(\tilde{A} \cup \tilde{C}, \tilde{B} \cup \tilde{C}), D(\tilde{A} \cap \tilde{C}, \tilde{B} \cap \tilde{C})\}$.

First and second axioms are obvious. The third axiom models the fact that when joining (resp. intersecting) both $\tilde{A}$ and $\tilde{B}$ with another fuzzy subset $\tilde{C}$, the divergence should decrease, as $\tilde{A} \cup \tilde{C}$ and $\tilde{B} \cup \tilde{C}$ (resp. $\tilde{A} \cap \tilde{C}$ and $\tilde{B} \cap \tilde{C}$) are more similar than $\tilde{A}$ and $\tilde{B}$ (see Figure 1 below for a graphical interpretation of this axiom, in which the divergence is defined as the area between the fuzzy subsets).

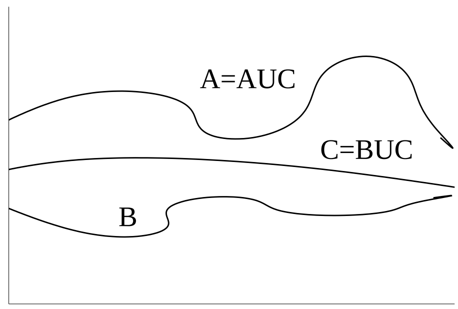

Figure 1: A graphical interpretation of third axiom of divergence measures.

Definition 3 *[12] Let Ω be a finite reference set. A divergence measure D is said to be a* **local divergence measure** *if and only if*

$$D(\tilde{A}, \tilde{B}) - D(\tilde{A} \cup \Omega^i, \tilde{B} \cup \Omega^i) = g(\tilde{A}(x_i), \tilde{B}(x_i)), \ \forall x_i \in \Omega,$$

where

$$\Omega^i(x) = \begin{cases} 1 & \text{if} \quad x = x_i \\ 0 & \text{otherwise} \end{cases}$$

It must be noted that in Definition 3, Ω is supposed to be a finite set. An extension of this definition for infinite reference sets can be found in [6].

For this special class of divergence measures the following result can be proved ([10, 12]):

Proposition 1 *If we have a finite set Ω, then D is a local divergence measure if and only if there exists an application $h : [0,1] \times [0,1] \mapsto \mathbb{R}$ such that*

$$D(A,B) = \sum_{i=1}^{n} h(A(x_i), B(x_i)), \tag{1}$$

satisfying the following conditions:

- $h(x,y) = h(y,x)$, $\forall x,y \in [0,1]$;
- $h(x,x) = 0$, $\forall x \in [0,1]$;
- $h(\cdot,y)$ *is a non-increasing function on [0,y] and non-decreasing on [y,1].*

Remark that for local divergence measures each coordinate is independent of the others and they are all equally important.

It is straightforward to see that the conditions for h in Proposition 1 are the same as those for divergence measures over a single reference set. Then, we conclude that when a divergence measure D presents the local property, it can be decomposed as the sum of the divergences of the coordinates or, in other words, D can be defined from a divergence measure h applied over each coordinate.

Definition 4 *[8] A binary operator $\perp : [0,1] \times [0,1] \mapsto [0,1]$ is called a* **t-conorm** *if and only if it satisfies the following conditions:*

1. $0 \perp x = x$, $\forall x \in [0,1]$, $1 \perp 1 = 1$. *(Boundary conditions).*
2. $u \perp v = v \perp u$
3. *If $u < u'$ then $u \perp v \leq u' \perp v$.*
4. $\perp$ *is associative.*

A binary operator $\top : [0,1] \times [0,1] \mapsto [0,1]$ is called a **t-norm** *if it satisfies conditions 2, 3 and 4, and 1 is substituted by*

1'- $1 \top x = x$, $\forall x \in [0,1]$, $0 \top 0 = 0$.

The following result, based on Ling's theorem [7], characterizes t-conorms:

Theorem 1 *[3] Let $\mathcal{D} = \{(\alpha_i, \beta_i) \,|\, i \in N \subset \mathbb{N}\}$ be a countable family of open disjoint subintervals of [0,1], and let $\mathcal{F} = \{f_i : [\alpha_i, \beta_i] \mapsto \mathbb{R} \,|\, i \in N\}$ be a family of continuous and strictly increasing functions with $f_i(\alpha_i) = 0$. Then, the map $\perp$ defined on $[0,1] \times [0,1] = [0,1]^2$ by*

$$x \perp y = \begin{cases} f_i^{(-1)}[f_i(x) + f_i(y)] & \text{if } \ x,y \in (\alpha_i, \beta_i) \\ \sup(x,y) & \text{otherwise} \end{cases}$$

where $f_i^{(-1)}(z) = f_i^{-1}[\min\{z, f_i(\beta_i)\}]$ is the pseudo-inverse function of f_i, is a t-conorm. Conversely, any continuous t-conorm has this form, with a suitable choice of $\mathcal{D}$ and $\mathcal{F}$. The function f_i is called the additive generator of the t-conorm with respect to the interval $[\alpha_i, \beta_i]$.

A dual theorem can be stated for t-norms.

Consider the subset $\Delta = [0,1] - \cup_{i \in N}(\alpha_i, \beta_i)$. This subset is closed and completely determines the family $\mathcal{D}$.

Definition 5 *[3] Any element in Δ is called* **idempotent***.*

Idempotent elements satisfy the following property:

$$x \in \Delta \Leftrightarrow x \perp x = x.$$

Let us now turn to the concept of fuzziness. We need a previous definition.

Definition 6 *Consider $\tilde{A}, \tilde{B} \in \tilde{\mathcal{P}}(\Omega)$. Subset $\tilde{A}$ is said to be* **sharper** *than $\tilde{B}$, and we will denote it by $\tilde{A} << \tilde{B}$ if and only if $\forall x \in \Omega$, either $\tilde{A}(x) \leq \tilde{B}(x) \leq \frac{1}{2}$ or $\tilde{A}(x) \geq \tilde{B}(x) \geq \frac{1}{2}$.*

Definition 7 *[4] A* **measure of fuzziness** *is a real function $f : \tilde{\mathcal{P}}(\Omega) \mapsto \mathbb{R}$, satisfying for any $\tilde{A}, \tilde{B} \in \tilde{\mathcal{P}}(\Omega)$:*

- $f(\tilde{A}) = 0 \Leftrightarrow \tilde{A} \in \mathcal{P}(X)$.
- *If $\tilde{A} << \tilde{B}$, then $f(\tilde{A}) \leq f(\tilde{B})$.*
- *f attains the maximum value in $\tilde{E}$.*
- $f(\tilde{A}) = f(\tilde{A}^c)$.

Definition 8 *(see e.g. [2]) A* **fuzzy measure** *is a map $\bar{m} : \tilde{\mathcal{P}}(\Omega) \mapsto [0, M]$ with the following properties:*

- $\bar{m}(\emptyset) = 0, \bar{m}(\Omega) = M$.
- $\tilde{A} \subset \tilde{B} \Rightarrow \bar{m}(\tilde{A}) \leq \bar{m}(\tilde{B})$.

This definition extends for fuzzy subsets the one given by Sugeno in [15].

3 g-local divergence measures

Let us now define the concept of *g-local divergence measure*. The concept of locality is based on the notion of diramativity which appears in Information Theory. This property is given by the following definition.

Definition 9 *[5] Given an uncertainty measure H, we say that it verifies the* **diramativity property** *if and only if*

$$H_{n+1}(A_1', A_1'', A_2, ..., A_n) - H_n(A_1' \cup A_1'', A_2, ..., A_n) = G(A_1', A_1'').$$

The following can be proved:

Proposition 2 *An uncertainty measure H has the diramativity property if and only if*

$$H(A_1, ..., A_n, B_1, ..., B_m) = H(A_1, ..., A_n, B) + H(A, B_1, ..., B_m) - H(A, B),$$

in which $A = \cup_{i=1}^{n} A_i, B = \cup_{i=1}^{m} B_i$.

This last formula has been generalized in [1] to

$$H(A_1, ..., A_n, B_1, ..., B_m) = \Phi(H(A_1, ..., A_n, B), H(A, B_1, ..., B_m), H(A, B)). \quad (2)$$

We will follow the same way to generalize the concept of local divergence measure. Let us consider a reference set Ω (finite or infinite) and a local divergence measure D. Let $\tilde{A}, \tilde{B}$ be two fuzzy subsets of Ω and consider a crisp partition in two subsets $\{\Omega_1, \Omega_2\}$ of Ω. Let us define $\tilde{A}_1, \tilde{A}_2, \tilde{B}_1, \tilde{B}_2$ by

$$\tilde{A}_i(x) = \begin{cases} \tilde{A}(x) & \text{if } \; x \in \Omega_i \\ 1 & \text{otherwise} \end{cases}, \; \tilde{B}_i(x) = \begin{cases} \tilde{B}(x) & \text{if } \; x \in \Omega_i \\ 1 & \text{otherwise} \end{cases}, \; i = 1, 2.$$

The property

$$D(\tilde{A}, \tilde{B}) = D(\tilde{A}_1, \tilde{B}_1) + D(\tilde{A}_2, \tilde{B}_2), \quad (3)$$

of the classical locality can be extended to the generalized on Equation 2, thus obtaining:

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2, \tilde{B}_2)). \quad (4)$$

Remark that Φ is associative: to show this, just consider a partition of Ω in three subsets $\{\Omega_1, \Omega_2, \Omega_3\}$. Then,

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2 \cup \tilde{A}_3, \tilde{B}_2 \cup \tilde{B}_3)) = \Phi(D(\tilde{A}_1, \tilde{B}_1), \Phi(D(\tilde{A}_2, \tilde{B}_2), D(\tilde{A}_3, \tilde{B}_3))).$$

On the other hand

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1 \cup \tilde{A}_2, \tilde{B}_1 \cup \tilde{B}_2), D(\tilde{A}_3, \tilde{B}_3)) = \Phi(\Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2, \tilde{B}_2)), D(\tilde{A}_3, \tilde{B}_3)),$$

whence the associativity.

Moreover, Φ is commutative. Let us now impose the following properties:

- $u < u' \Rightarrow \Phi(u, v) \leq \Phi(u', v)$.
- $\Phi(0, v) = v$.

When D is a local divergence measure, then $\Phi(x, y) = x + y$, and it is clear that in this case Φ satisfies these properties.

This leads us to the following definition:

Definition 10 *Given a divergence measure D over a reference set Ω, we say that D is a* **g-local divergence measure** *if and only if for any partition of Ω in two subsets $\{\Omega_1, \Omega_2\}$, it is*

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2, \tilde{B}_2)), \tag{5}$$

where Φ satisfies

1. $\Phi(u, v) = \Phi(v, u)$.
2. $u < u' \Rightarrow \Phi(u, v) \leq \Phi(u', v)$.
3. Φ *is associative.*
4. $\Phi(0, v) = v$.

In this definition we are dividing our reference set Ω in two new reference sets Ω_1, Ω_2, and defining new divergence measures D_1, D_2 from D in each reference set.

Remark that in the definition of $\tilde{A}_i, \tilde{B}_i$, we have chosen the value 1 (the maximal possible value for the membership function) for coordinates $x_j \notin \Omega_i$. Remark that (3) also holds if we had chosen the value 0 instead of 1. The reason is that 0 is a neutral element while 1 is the absorbent. If we had chosen 0, the interpretation should be that we are joining two referential sets, where in each of them a divergence measure has been defined, and we are trying to define a new divergence measure; we will come back to this case in Section 6.

Of course, a local divergence measure is a g-local divergence measure. Let us now give an example of a g-local divergence measure that is not local:

Example 1 *Consider $D(\tilde{A}, \tilde{B}) = \sup_{x\in\Omega} |\tilde{A}(x) - \tilde{B}(x)|$. It is clear that this divergence measure is a g-local divergence measure with $\Phi(x, y) = \sup\{x, y\}$, but it is not a local one.*

In next proposition, we list some properties of g-local divergence measures.

Proposition 3 *Let D be a g-local divergence measure. Then, $\forall \tilde{A}, \tilde{B}, \tilde{C} \in \tilde{\mathcal{P}}(\Omega)$, $\forall Z, V \in \mathcal{P}(\Omega)$:*

- $D(\tilde{A}, \tilde{B}) = D(\tilde{A} \cup \tilde{B}, \tilde{A} \cap \tilde{B})$.
- $D(\tilde{A} \cup Z, \tilde{B} \cup Z) = D(\tilde{A} \cap Z^c, \tilde{B} \cap Z^c)$.
- $D(\tilde{A}, \tilde{B}) \geq \sup_{x\in\Omega} D(\tilde{A}(x), \tilde{B}(x))$.
- *For any permutation σ in Ω, $D(\tilde{A}, \tilde{B}) = D(\tilde{A}_\sigma, \tilde{B}_\sigma)$.*
- $D(Z, Z^c) = D(V, V^c)$.
- $D(\tilde{A}, \tilde{B}) \leq D(Z, Z^c)$.
- $D(\tilde{A}, \tilde{B}) \geq \min\{D(\tilde{A}, \tilde{C}), D(\tilde{C}, \tilde{B})\}$, *whenever $\tilde{A}(x) \leq \tilde{C}(x) \leq \tilde{B}(x)$ or $\tilde{A}(x) \geq \tilde{C}(x) \geq \tilde{B}(x)$, $\forall x \in \Omega$.*

Proof: The proof is straightforward. It just suffices to translate the proofs for the local case that appear in [10] for the g-local case. We prove the first one as an example:

Let us define $\Omega_1 = \{x \in \Omega \,|\, \tilde{A}(x) \geq \tilde{B}(x)\}$ and $\Omega_2 = \{x \in \Omega \,|\, \tilde{A}(x) < \tilde{B}(x)\}$. Then, $\{\Omega_1, \Omega_2\}$ is a partition of Ω and applying the definition of g-locality,

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2, \tilde{B}_2)).$$

Now, by the second axiom of divergence, $D(\tilde{A}_2, \tilde{B}_2) = D(\tilde{B}_2, \tilde{A}_2)$. Thus,

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{B}_2, \tilde{A}_2)) = D(\tilde{A}_1 \cap \tilde{B}_2, \tilde{A}_2 \cap \tilde{B}_1),$$

and now, $\tilde{A}_1 \cap \tilde{B}_2 = \tilde{A} \cup \tilde{B}$, whereas $\tilde{A}_2 \cap \tilde{B}_1 = \tilde{A} \cap \tilde{B}$. ■

As it has been done in [10] for local measures, it is possible to define measures of fuzziness from g-local divergence measures:

Proposition 4 *Let D be a g-local divergence measure. Then, $f_1, f_2 : \tilde{\mathcal{P}}(\Omega) \mapsto \mathbb{R}$ defined by*

$$f_1(\tilde{A}) = D(Z, Z^c) - D(\tilde{A}, \tilde{A}^c),\ Z \in \mathcal{P}(\Omega)$$

$$f_2(\tilde{A}) = D(\tilde{A}, N_{\tilde{A}}),$$

are measures of fuzziness.

We give an example of a divergence measure that it does not present g-locality, even if the divergence can be decomposed as a function of the divergence on each coordinate.

Example 2 *Consider a finite reference set and the divergence measure given by $D(\tilde{A}, \tilde{B}) = \Pi_{x\in\Omega}|\tilde{A}(x) - \tilde{B}(x)|$. This divergence measure does not satisfy the conditions of Definition 10 as it can be easily checked (0 is not the neutral element but an absorbent). We will return to this example in Section 6.*

Let us now study the general form of a g-local divergence measure. Remark that any function $\Phi : [0, M] \times [0, M] \mapsto [0, M]$ in the conditions of Definition 10 satisfies all properties of a t-conorm except 1) that changes into

1') $0 \perp x = x,\ \forall x \in [0, M],\ M \perp M = M.$

Consequently, we can apply Ling's Theorem and Forte,Benvenuti,Kampe de Feriet's result, thus obtaining:

Theorem 2 *For any function Φ in the conditions of Definition 10, it follows that Φ can be written as taking a suitable sequence of disjoint open intervals (α_i, β_i) on [0,M] and strictly increasing functions $f_i : [\alpha_i, \beta_i] \mapsto \mathbb{R}$ such that $f_i(\alpha_i) = 0$, and putting*

$$\Phi(u, v) = \begin{cases} f_i^{(-1)}[f_i(u) + f_i(v)] & \text{if } u, v \in (\alpha_i, \beta_i) \\ \sup(u, v) & \text{otherwise} \end{cases}$$

in which $f_i^{(-1)}$ denotes the pseudo-inverse function of f_i. We will call f_i the additive generator of Φ over $[\alpha_i, \beta_i]$.

Elements that do not belong to $\bigcup_{i\in\mathbb{N}}(\alpha_i, \beta_i)$ are therefore idempotent. If the range of D is [0,M], we will denote by Δ the set of idempotent elements. It is clear that Δ is a closed set and that $0, M \in \Delta$.

At this point, it must be remarked that the only t-conorm satisfying the local property is the sum. Hence, g-locality provides a wide generalization of locality.

When D is a local divergence measure, then $\Delta = \{0, M\}$. However, the converse is not true, as next example shows:

Example 3 *Consider $|\Omega| = n$ and the divergence measure defined by*

$$D(\tilde{A}, \tilde{B}) = \max\{\sum_{x\in\Omega} |\tilde{A}(x) - \tilde{B}(x)|, 1\}.$$

It is straightforward to check that D is a continuous g-local divergence measure for which Φ is given by

$$\Phi(x, y) = \begin{cases} 1 & \text{if } \ x = 1 \ \textit{or } y = 1 \\ \sup\{x + y, 1\} & \text{otherwise} \end{cases}$$

Therefore, Φ satisfies the conditions of a t-conorm. However, 0 and $M = 1$ are the only idempotent elements.

In next sections we will study some consequences of this representation theorem for different choices of the cardinality of Ω.

4 The case of finite or countable referential set

Let us start with the finite case. We have the following:

Proposition 5 *Assume $\Omega = \{x_1, ..., x_n\}$ and consider a divergence measure D. If D is a g-local divergence measure, then $\exists g_n : \mathbb{R}^n \mapsto \mathbb{R}$ such that*

$$D(\tilde{A}, \tilde{B}) = g_n(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), ..., h_n(\tilde{A}(x_n), \tilde{B}(x_n))),$$

where, for all $i = 1, ..., n$, $h_i : [0, 1] \times [0, 1] \mapsto \mathbb{R}$ satisfies the following conditions:

- *$h_i(x, y) = h_i(y, x)$, $\forall x, y \in [0, 1]$;*
- *$h_i(x, x) = 0$, $\forall x \in [0, 1]$;*
- *$h_i(\cdot, y)$ is a non-increasing function on [0,y] and non-decreasing on [y,1].*

Proof: By the definition of the g-local property applied to $\Omega_1 = \{x_1\}, \Omega_2 = \Omega - \Omega_1$, it follows

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2, \tilde{B}_2)) = \Phi(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), D(\tilde{A}_2, \tilde{B}_2)).$$

Applying again the definition for $D(\tilde{A}_2, \tilde{B}_2)$ with $\{x_2\}$and $\Omega_2 - \{x_2\}$ we have

$$D(\tilde{A}_2, \tilde{B}_2) = \Phi(h_2(\tilde{A}(x_2), \tilde{B}(x_2)), D(\tilde{A}_3, \tilde{B}_3)).$$

Hence,

$$\begin{aligned} D(\tilde{A}, \tilde{B}) &= \Phi(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), \Phi(h_2(\tilde{A}(x_2), \tilde{B}(x_2)), D(\tilde{A}_3, \tilde{B}_3))) \\ &=: g_3(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), h_2(\tilde{A}(x_2), \tilde{B}(x_2)), D(\tilde{A}_3, \tilde{B}_3)). \end{aligned}$$

Iterating this process, we obtain that D can be written as

$$D(\tilde{A}, \tilde{B}) = g_n(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), h_2(\tilde{A}(x_2), \tilde{B}(x_2)), ..., h_n(\tilde{A}(x_n), \tilde{B}(x_n))).$$

Finally, remark that all functions h_i are indeed divergence measures over single references. Therefore, they must satisfy the conditions of h in Proposition 1. ■

It must be noticed that function g_n is the iterated t-conorm Φ, where the iteration is justified by associativity.

As a consequence of this result, we will find the value of $D(\tilde{A}, \tilde{B})$ for $\tilde{A}, \tilde{B} \in \tilde{\mathcal{P}}(\Omega)$. We have two different situations:

Lemma 1 *Assume $\Omega = \{x_1, ..., x_n\}$ and consider a g-local divergence measure D. Let us denote by Δ the set of idempotent elements. If $\max_i\{h_i(\tilde{A}(x_i), \tilde{B}(x_i))\} \in \Delta$, then $D(\tilde{A}, \tilde{B}) = \max_i\{h_i(\tilde{A}(x_i), \tilde{B}(x_i))\}$.*

Proof: We will make the proof by induction on the cardinality of Ω.
If $n = 2$, then $g_2 = \Phi$ and the result holds by Theorem 2.
Suppose $n > 2$ and that the result holds until $n - 1$. Then,

$$g_n(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), ..., h_n(\tilde{A}(x_n), \tilde{B}(x_n))) =$$

$$g_{n-1}(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), ..., \Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n)))).$$

Now, we have two possible cases:

1. If $\max\{h_n(\tilde{A}(x_n), \tilde{B}(x_n)), h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1}))\} = \max\{h_i(\tilde{A}(x_i), \tilde{B}(x_i))\}$, then

$$\Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))) = \max\{h_i(\tilde{A}(x_i), \tilde{B}(x_i))\}$$

   and applying the induction hypothesis, the result holds.

2. Otherwise, $\Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))) < \max\{h_i(\tilde{A}(x_i), \tilde{B}(x_i))\}$, as this last value is idempotent. Applying again the induction hypothesis, the result holds. ■

Lemma 2 *Assume $\Omega = \{x_1, ..., x_n\}$ and let us consider a g-local divergence measure D. Let us suppose that $\max_i\{h_i(\tilde{A}(x_i), \tilde{B}(x_i))\} \in (\alpha_i, \beta_i)$ and let us denote by f_i the additive generator of Φ in (α_i, β_i). Then, $D(\tilde{A}, \tilde{B})$ can be written as*

$$D(\tilde{A}, \tilde{B}) = f_i^{(-1)}[\sum_{x_j | h_j(\tilde{A}(x_j), \tilde{B}(x_j)) \in (\alpha_i, \beta_i)} f_i(h_j(\tilde{A}(x_j), \tilde{B}(x_j)))].$$

Proof: We will make again the proof by induction on the cardinality of Ω.
For $n = 2$ it is $g_2 = \Phi$ and the result holds.
Suppose $n > 2$ and assume the result holds until $n - 1$. Then,

$$g_n(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), ..., h_n(\tilde{A}(x_n), \tilde{B}(x_n))) =$$

$$g_{n-1}(h_1(\tilde{A}(x_1), \tilde{B}(x_1)), ..., \Phi[h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))]).$$

Now, we have three possibilities:

1. If both $x = h_n(\tilde{A}(x_n), \tilde{B}(x_n))$ and $y = h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1}))$ are in (α_i, β_i), then

$$\Phi(x, y) = f_i^{(-1)}(f_i(x) + f_i(y)) \in (\alpha_i, \beta_i]. \quad (6)$$

   In such a case, we have again two possible cases:

   (a) If $\Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))) = \beta_i$, then $D(\tilde{A}, \tilde{B}) = \beta_i$ by Lemma 1. On the other hand,

$$\beta_i \geq f_i^{(-1)}(\sum_{x_j | h_j(\tilde{A}(x_j), \tilde{B}(x_j)) \in (\alpha_i, \beta_i)} f_i(h_j(\tilde{A}(x_j), \tilde{B}(x_j)))) \geq f_i^{(-1)}(f_i(x) + f_i(y)) = \beta_i,$$

   whence the result.

   (b) If $\Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))) \in (\alpha_i, \beta_i)$, then $f_i^{(-1)} = f_i^{-1}$ in Equation 6, and the result holds by the induction hypothesis.

2. If only one of them, say $h_n(\tilde{A}(x_n), \tilde{B}(x_n))$, is in (α_i, β_i), then

$$\Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))) = h_n(\tilde{A}(x_n), \tilde{B}(x_n)),$$

   and the result holds applying the induction hypothesis.

3. Otherwise, $\Phi(h_{n-1}(\tilde{A}(x_{n-1}), \tilde{B}(x_{n-1})), h_n(\tilde{A}(x_n), \tilde{B}(x_n))) < \alpha_i$, and applying again the induction hypothesis, the result holds. ■

Lemma 1 and Lemma 2 show that, in order to compute $D(\tilde{A}, \tilde{B})$, some coordinates may not affect the final value; in particular, if we are in the conditions of Lemma 1, it suffices to know the value of function h_x corresponding the coordinate x where the maximum is reached. Let us see an example.

Example 4 *Suppose* $n = 3$ *and* $\Delta = [0, 0.25] \cup [0.75, 1]$. *Let us consider* $f_1 : [0.25, 0.75] \times [0.25, 0.75] \mapsto \mathbb{R}$ *given by* $f_1(x) = x^2 - 0.0625$; *finally,* $h_i(x, y) = |x - y|, \forall i$.

Now, define the fuzzy subsets $\tilde{A} \equiv (0.4, 0.5, 0.1)$ *and* $\tilde{B} = \emptyset$. *Then, by Lemma 2,*

$$D(\tilde{A}, \tilde{B}) = f_1^{(-1)}(f_1(\tilde{A}(x_1)) + f_1(\tilde{A}(x_2))) = f_1^{-1}(0.285) = 0.5895.$$

On the other hand, if we consider $\tilde{C} \equiv (0.5, 0.8, 0)$, *then*

$$D(\tilde{C}, \tilde{B}) = 0.8,$$

by Lemma 1.

Let us now turn to the infinite countable case. First, note that, as in the finite case, we can define a function $h_x(\tilde{A}(x), \tilde{B}(x)), \forall x \in \Omega$ (consider $h_x(\tilde{A}(x), \tilde{B}(x)) = D(\tilde{A}_1, \tilde{B}_1)$ with $\Omega_1 = \{x\}$). Now, the following can be proved:

Theorem 3 *Assume* Ω *is an infinite countable reference set. Consider* $\tilde{A}, \tilde{B} \in \tilde{\mathcal{P}}(\Omega)$ *and a continuous g-local divergence measure* D.

1. *If* $\exists \max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\}$, *then:*

   (a) *If* $\max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \in \Delta$, *then*

   $$D(\tilde{A}, \tilde{B}) = \max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\}.$$

   (b) *Suppose* $\max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \notin \Delta$ *and assume* $\max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \in (\alpha_i, \beta_i)$. *Then,*

   $$D(\tilde{A}, \tilde{B}) = f_i^{(-1)}(\sum_{x|h_x(\tilde{A}(x), \tilde{B}(x)) \in (\alpha_i, \beta_i)} f_i(h_x(\tilde{A}(x), \tilde{B}(x)))).$$

2. *If* $\nexists \max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\}$, *then:*

   (a) *If* $\sup_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \in \Delta - [\cup_{i\in N}\beta_i]$, *then*

   $$D(\tilde{A}, \tilde{B}) = \sup_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\}.$$

   (b) *If* $\sup_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \in (\alpha_i, \beta_i]$, *for some* $i \in N$, *then*

   $$D(\tilde{A}, \tilde{B}) = f_i^{(-1)}(\sum_{x|h_x(\tilde{A}(x), \tilde{B}(x)) \in (\alpha_i, \beta_i)} f_i(h_x(\tilde{A}(x), \tilde{B}(x)))).$$

Proof: Given $\tilde{A}, \tilde{B}$, let us define the families of fuzzy subsets $\{\tilde{A}_n\}_{n\in\mathbb{N}}, \{\tilde{B}_n\}_{n\in\mathbb{N}}$ by

$$\tilde{A}_n(x_i) = \begin{cases} \tilde{A}(x_i) & \text{if } \; i \leq n \\ 0 & \text{otherwise} \end{cases}, \; \tilde{B}_n(x_i) = \begin{cases} \tilde{B}(x_i) & \text{if } \; i \leq n \\ 0 & \text{otherwise} \end{cases}$$

Then, as $\tilde{A}_n \to \tilde{A}$ and $\tilde{B}_n \to \tilde{B}$, it follows that $D(\tilde{A}_n, \tilde{B}_n) \to D(\tilde{A}, \tilde{B})$, by the continuity of D.

Suppose $\exists \max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \in \Delta$. Suppose indeed that this maximum is reached in x_i. Then, for $n \geq i$,

$$D(\tilde{A}_n, \tilde{B}_n) = g_{n+1}(h_{x_1}(\tilde{A}_n(x_1), \tilde{B}_n(x_1)), ..., h_{x_n}(\tilde{A}_n(x_n), \tilde{B}_n(x_n)), D(\emptyset, \emptyset)) = h_{x_i}(\tilde{A}_n(x_i), \tilde{B}_n(x_i)),$$

by Lemma 1 , whence

$$D(\tilde{A}, \tilde{B}) = h_{x_i}(\tilde{A}(x_i), \tilde{B}(x_i)).$$

On the other hand, suppose $\exists \max_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\} \in (\alpha_i, \beta_i)$. Suppose indeed that the maximum is reached in x_j. Then, for $n \geq j$, by Lemma 2,

$$D(\tilde{A}_n, \tilde{B}_n) = g_{n+1}(h_{x_1}(\tilde{A}_n(x_1), \tilde{B}_n(x_1)), ..., h_{x_n}(\tilde{A}_n(x_n), \tilde{B}_n(x_n)), D(\emptyset, \emptyset))$$
$$= f_i^{(-1)}(\sum_{*} f_i(h_{x_k}(\tilde{A}(x_k), \tilde{B}(x_k)))),$$

where the sum applies to

$$* = \{x_k | h_{x_k}(\tilde{A}(x_k), \tilde{B}(x_k)) \in (\alpha_i, \beta_i), k \leq n\}.$$

Now, taking limits, we obtain that

$$D(\tilde{A}, \tilde{B}) = f_i^{(-1)}(\sum_{**} f_i(h_x(\tilde{A}(x), \tilde{B}(x)))),$$

where the sum applies to

$$** = \{x_k | h_{x_k}(\tilde{A}(x_k), \tilde{B}(x_k)) \in (\alpha_i, \beta_i)\}.$$

Suppose on the other hand that the supremum is not reached. In order to simplify the formulas, let us denote $P = \sup_{x\in\Omega}\{h_x(\tilde{A}(x), \tilde{B}(x))\}$.

If $P \in \Delta - [\cup_{i\in N}\beta_i]$, then, $\exists j$ such that $P \in (\beta_j, \alpha_{j+1}]$. Besides, there exists i such that $h_{x_i}(\tilde{A}(x_i), \tilde{B}(x_i)) > \beta_j$. Thus, for $n \geq i$, by Lemma 1,

$$D(\tilde{A}_n, \tilde{B}_n) = g_{n+1}(h_{x_1}(\tilde{A}_n(x_1), \tilde{B}_n(x_1)), ..., h_{x_n}(\tilde{A}_n(x_n), \tilde{B}_n(x_n)), D(\emptyset, \emptyset)) = \sup_k h_{x_k}(\tilde{A}_n(x_k), \tilde{B}_n(x_k)),$$

and then

$$D(\tilde{A}, \tilde{B}) = \sup_k \{h_{x_k}(\tilde{A}(x_k), \tilde{B}(x_k))\} = P.$$

Finally, assume $P \in (\alpha_j, \beta_j]$. Then, there exists i such that $h_{x_i}(\tilde{A}(x_i), \tilde{B}(x_i)) > \alpha_j$. Thus, for $n \geq i$, by Lemma 2,

$$D(\tilde{A}_n, \tilde{B}_n) = g_{n+1}(h_{x_1}(\tilde{A}_n(x_1), \tilde{B}_n(x_1)), ..., h_{x_n}(\tilde{A}_n(x_n), \tilde{B}_n(x_n)), D(\emptyset, \emptyset))$$

$$= f_i^{(-1)}(\sum_{*} f_i(h_{x_k}(\tilde{A}(x_k), \tilde{B}(x_k)))).$$

Now, taking limits, we obtain that

$$D(\tilde{A}, \tilde{B}) = f_i^{(-1)}(\sum_{**} f_i(h_x(\tilde{A}(x), \tilde{B}(x)))).$$

This completes the proof. ∎

5 The non-countable case

Suppose now that an Ω is a non-countable reference set. For this case, we will follow the same approach of Bertoluzza and Cariolaro ([2]) in the case of fuzzy measures of fuzzy subsets. Let us start with a definition:

Definition 11 *We define the* **discriminant element** *of a pair of fuzzy subsets* $\tilde{A}, \tilde{B}$ *by*

$$\Gamma(\tilde{A}, \tilde{B}) = \sup_{\alpha, \beta, C} D(\alpha C, \beta C),$$

where C *are the crisp subsets of* Ω *and* $\alpha \geq \beta$ *are such that*

$$(\tilde{A} \cap \tilde{B})_C \subset \beta C \subset \alpha C \subset (\tilde{A} \cup \tilde{B})_C.$$

An example of suitable values of α, β for a given C can be seen in Figure 2.

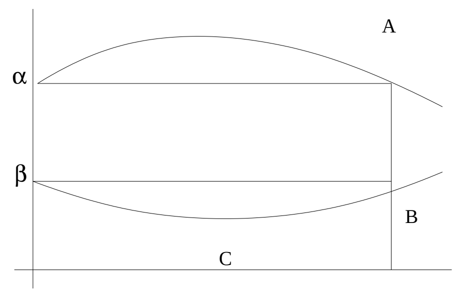

Figure 2: A graphical interpretation of $\Gamma(\tilde{A}, \tilde{B})$

The value of $\Gamma(\tilde{A}, \tilde{B})$ can be computed using next proposition:

Proposition 6 *Let D be a divergence measure and for fixed $\alpha, \beta \in [0,1]$ let us define*

$$C_{\alpha,\beta} = \{x \in \Omega |\, (\tilde{A} \cup \tilde{B})(x) \geq \alpha, (\tilde{A} \cap \tilde{B})(x) \leq \beta\}.$$

Then, $\Gamma(\tilde{A}, \tilde{B}) = \sup_{\alpha,\beta} D(\alpha C_{\alpha,\beta}, \beta C_{\alpha,\beta})$.

Proof: Let us fix α, β. By definition, $(\tilde{A} \cap \tilde{B})_{C_{\alpha,\beta}} \subset \beta C_{\alpha,\beta} \subset \alpha C_{\alpha,\beta} \subset (\tilde{A} \cup \tilde{B})_{C_{\alpha,\beta}}$, whence

$$\Gamma(\tilde{A}, \tilde{B}) \geq \sup D(\alpha C_{\alpha,\beta}, \beta C_{\alpha,\beta}).$$

Now, for any $\epsilon > 0$, there exist $\alpha_0 \geq \beta_0$ and $C \in \mathcal{P}(X)$ such that $(\tilde{A} \cap \tilde{B})_C \subset \beta_0 C \subset \alpha_0 C \subset (\tilde{A} \cup \tilde{B})_C$ and $D(\alpha_0 C, \beta_0 C) \geq \Gamma(\tilde{A}, \tilde{B}) - \epsilon$.

On the other hand, $C \subset C_{\alpha_0,\beta_0}$, and hence $D(\alpha_0 C, \beta_0 C) \leq D(\alpha_0 C_{\alpha_0,\beta_0}, \beta_0 C_{\alpha_0,\beta_0})$, whence

$$\Gamma(\tilde{A}, \tilde{B}) \leq \sup D(\alpha C_{\alpha,\beta}, \beta C_{\alpha,\beta}).$$

This finishes the proof. ■

Note that $D(\alpha C, \beta C) \leq D(\tilde{A}, \tilde{B})$ for any crisp subset C satisfying the conditions of Definition 11 and then, $\Gamma(\tilde{A}, \tilde{B})$ is a lower approximation of $D(\tilde{A}, \tilde{B})$.

In order to simplify the results below, we will suppose $\tilde{A}(x) \geq \tilde{B}(x)$, $\forall x \in \Omega$. This is not restrictive for g-local divergence measures, as we have already proved that $D(\tilde{A}, \tilde{B}) = D(\tilde{A} \cup \tilde{B}, \tilde{A} \cap \tilde{B})$ (Proposition 3). Now, the following can be shown:

Proposition 7 *Let D be a continuous g-local divergence measure. If $\Gamma(\tilde{A}, \tilde{B}) \in \Delta$, then $D(\tilde{A}, \tilde{B}) = \Gamma(\tilde{A}, \tilde{B})$.*

Proof: Let us consider the crisp subset of Ω defined by

$$\{x \in \Omega \,|\, \tilde{A}(x) - \tilde{B}(x) \geq \frac{1}{n}\}.$$

This subset can be partitioned by the family $\{C^n_{i,j}\}_{i,j}$, where

$$C^n_{i,j} = \{x \in \Omega \,|\, \frac{i+1}{n} \geq \tilde{A}(x) > \frac{i}{n}, \frac{j+1}{n} \geq \tilde{B}(x) > \frac{j}{n}\}, \; i > j, \; i,j = 0, ..., n-1.$$

Let us now define:

$$\tilde{A}_n(x) = \begin{cases} \frac{i}{n} & \text{if } x \in C^n_{i,j} \\ \frac{\tilde{A}(x)+\tilde{B}(x)}{2} & \text{otherwise} \end{cases} \quad \tilde{B}_n(x) = \begin{cases} \frac{j+1}{n} & \text{if } x \in C^n_{i,j} \\ \frac{\tilde{A}(x)+\tilde{B}(x)}{2} & \text{otherwise} \end{cases}$$

It is clear that $\tilde{B} \subset \tilde{B}_n \subset \tilde{A}_n \subset \tilde{A}$. Moreover, $\tilde{A}_n \uparrow \tilde{A}$ and $\tilde{B}_n \downarrow \tilde{B}$; then, by the continuity of D, it is $D(\tilde{A}_n, \tilde{B}_n) \uparrow D(\tilde{A}, \tilde{B})$.

If we apply now the definition of g-local property we obtain

$$D(\tilde{A}_n, \tilde{B}_n) = g_m(\{D(\frac{i}{n} C^n_{i,j}, \frac{j+1}{n} C^n_{i,j})\}_{i,j}, 0) \leq g_m(\{\Gamma(\tilde{A}, \tilde{B})\}_{i,j}, 0) = \Gamma(\tilde{A}, \tilde{B}),$$

as $\Gamma(\tilde{A}, \tilde{B}) \in \Delta$.

Then, taking limits $D(\tilde{A}, \tilde{B}) \leq \Gamma(\tilde{A}, \tilde{B})$ and the result is proved. ■

What happens when $\Gamma(\tilde{A}, \tilde{B}) \notin \Delta$?

Lemma 3 *Let D be a continuous g-local divergence measure. If $\Gamma(\tilde{A},\tilde{B}) \in (\alpha_i, \beta_i)$, then $D(\tilde{A},\tilde{B}) \in [\Gamma(\tilde{A},\tilde{B}), \beta_i]$.*

Proof: We already know that $\Gamma(\tilde{A},\tilde{B}) \leq D(\tilde{A},\tilde{B})$.

Let us consider $\tilde{A}_n, \tilde{B}_n$ as in Proposition 7. Then, we know that $D(\tilde{A}_n,\tilde{B}_n) \uparrow D(\tilde{A},\tilde{B})$.

If we apply the definition of g-local property we obtain

$$D(\tilde{A}_n,\tilde{B}_n) = g_m(\{D(\frac{i}{n}C^n_{i,j}, \frac{j+1}{n}C^n_{i,j})\}_{i,j}, 0) \leq g_m(\{\Gamma(\tilde{A},\tilde{B})\}_{i,j}, 0) = f_i^{(-1)}(m\Gamma(\tilde{A},\tilde{B})) \leq \beta_i,$$

as stated. ■

Now, let us give a method for obtaining the exact value of $D(\tilde{A},\tilde{B})$. This result is based on the research of Bertoluzza and Cariolaro [2]. They work with fuzzy measures of fuzzy subsets and they obtain an explicit formula for $\bar{m}(\tilde{C})$. In the following, we will assume that $\Gamma(\tilde{A},\tilde{B}) \in (\alpha_i, \beta_i)$.

If D is continuous and $D(\emptyset,\emptyset) = 0, D(\Omega,\emptyset) = M$, then we have that for any $\tilde{A},\tilde{B} \in \tilde{\mathcal{P}}(\Omega)$, there exists $\tilde{C} \in \tilde{\mathcal{P}}(\Omega)$ depending on $\tilde{A},\tilde{B}$, i.e. $\tilde{C}_{\tilde{A},\tilde{B}}$ such that $D(\tilde{A},\tilde{B}) = D(\tilde{C},\emptyset)$. Consequently, we can define an equivalence relation over the set of all pairs of fuzzy subsets and therefore, it suffices to obtain the value of $D(\tilde{C}_{\tilde{A},\tilde{B}},\emptyset)$. Now, it is easy to prove (see, for example [16]) that

Lemma 4 *$\bar{m}(\tilde{C}) := D(\tilde{C},\emptyset)$ is a fuzzy measure over fuzzy subsets.*

We are now in conditions to apply the results of Bertoluzza and Cariolaro. They find the value of $\bar{m}(\tilde{C})$ from the values of the measure over crisp subsets. Let us denote by m this restriction. Analogously, we will denote by $\bar{\Phi}$ the t-conorm defining $\bar{m}$ and by Φ the t-conorm defining m.

First, for a fixed $\alpha \in [0,1]$, we define

$$\varphi_\alpha(x) := \bar{m}(\alpha C),$$

where $C \in \mathcal{P}(\Omega)$ such that $m(C) = x$. Now, consider

$$T_0 := \{\alpha \,|\, \varphi_\alpha(M) \leq \alpha_i\}.$$

The idea of this definition is the following: We consider the α-levels of $\tilde{C}$. For some values of α, we have that $\bar{m}(\alpha F) < \alpha_i, \forall F \in \mathcal{P}(\Omega)$. These values of α are then excluded of the evaluation of $\bar{m}(\tilde{C})$ as the corresponding α-levels have no influence in the final value of $\bar{m}(\tilde{C})$. The set of these values of α is T_0.

For the other α-levels, $\alpha \notin T_0$, it follows that there exists a minimal measure of the crisp subset such that $\bar{m}(\alpha F) \in (\alpha_i, \beta_i)$. This value is given by

$$\xi(\alpha) := \sup\{x \in [0,M] \,|\, \varphi_\alpha(x) \leq \alpha_i\}.$$

In [2], it is shown that the minimal value of $m(F)$ in this conditions is an α_j. This leads us to the following definition:

$$T_j := \{\delta \in [0,1] \,|\, \xi(\delta) = \alpha_j\}.$$

However, some of these subsets T_j must be excluded, too. The set of excluded T_j is given by

$$T'_j := \{\delta \in T_j \mid \lim_{\epsilon \to 0} m(\tilde{C}_\alpha^{\alpha+\epsilon}) < \alpha_j\}$$

where $\tilde{C}_\alpha^{\alpha+\epsilon} = \{x \in \Omega \,|\, \alpha \leq \tilde{C}(x) \leq \alpha + \epsilon\}$. The sets T'_j are indeed the α-levels whose measure do not reach the minimal value for being in the interval (α_i, β_i); therefore, these α-cuts are removed. The remaining α-levels define the subset $J_j = T_j - T'_j$, from which we define

$$\Omega_j := \{x \in \Omega \,|\, \tilde{C}(x) \in J_j\}.$$

Then, we have:

Theorem 4 *[2] Let $\bar{f}$ be the additive generator of the restriction of $\bar{\Phi}$ to the interval (α_i, β_i), and let f_j be the additive generator of the restriction of Φ to the interval (α_j, β_j). Let us define*

$$\tilde{C}_j(x) = \begin{cases} \tilde{C}(x) & \text{if} \quad x \in \Omega_j \\ 0 & \text{otherwise} \end{cases}$$

Then, we have

$$\bar{m}(\tilde{C}_j) = \bar{f}_j^{(-1)}[\int_{\Omega_j} k(\tilde{C}_j) d(f_j \circ m)],$$

where k is defined by the solution of Equation

$$\bar{m}(\alpha C) = \bar{f}_j^{(-1)}(k(\alpha) f_j(m(C))).$$

If $f_j \circ m$ is additive, then we can decompose the integral:

$$\bar{m}(\tilde{C}_j) = \bar{f}_j^{(-1)}[\sum_r \int_{\Omega_j \cap \Omega^r} k(\tilde{C}_j) d(f_j \circ m)],$$

where $\{\Omega^r\}_r$ is a collection of disjoint sets such that $m(\Omega^r) < m(\Omega), \sum_r m(\Omega^r) > m(\Omega)$. Finally,

$$\bar{m}(\tilde{C}) = \bar{f}_j^{(-1)}[\sum_{j \in I} \bar{f} \circ m(\tilde{C}_j)].$$

6 The problem of componibility

In this section we study briefly the problem of componibility. Let us return to Example 2. This example is very special because indeed all coordinates are independent of each other and are all equally important. The condition that this divergence measure does not satisfy is that 0 (i.e. when coordinates are the same) is not the neutral element but the absorbent of function Φ.

Suppose now normalized divergence measures over two reference sets. Then, if we join these reference sets and we want to define a new divergence measure from the initial ones, it makes sense to consider 0 as absorbent. This translates to the condition $\Phi(u, 1) = u$. Then, we obtain the following definition:

Definition 12 *Given a normalize divergence measure D over a reference set Ω, we say that D is a* **componible divergence measure** *if and only if*

$$D(\tilde{A}, \tilde{B}) = \Phi(D(\tilde{A}_1, \tilde{B}_1), D(\tilde{A}_2, \tilde{B}_2)), \tag{7}$$

where Φ satisfies

- *1.- $\Phi(u, v) = \Phi(v, u)$.*
- *2.- $u < u' \Rightarrow \Phi(u, v) \leq \Phi(u', v)$.*
- *3.- Φ is associative.*
- *4'.- $\Phi(1, v) = v$.*

It comes out that Φ is a t-norm, whence we obtain the following representation theorem:

Theorem 5 *For any function Φ in the conditions of Definition 12, it follows that Φ can be written as taking a suitable sequence of disjoint open intervals (α_i, β_i) on [0,M] and strictly decreasing functions $f_i : [\alpha_i, \beta_i] \mapsto \mathbb{R}$ such that $f_i(\beta_i) = 0$, and putting*

$$\Phi(u, v) = \begin{cases} f_i^{(-1)}[f_i(u) + f_i(v)] & \text{if } \; u, v \in (\alpha_i, \beta_i) \\ \inf(u, v) & \text{otherwise} \end{cases}$$

in which $f_i^{(-1)}$ denotes the pseudo-inverse function of f_i.

Then, all the results we have obtained for local divergence measures have their dual results for componible divergence measures. However, the interpretation is completely different. In the first case, we have one divergence measure and we want to define two new divergence measures satisfying some properties; in the second case, we have two divergence measures and we want to define a new divergence measure satisfying some conditions over the joint reference set.

7 Conclusions

We have given a generalization of local divergence measures based on the relationship between local divergence measures and the property of diramativity for uncertainty measures. This new divergence measures keep most of the properties of local divergence measures. Also a representation theorem of generalized local divergence measures is provided based on Ling's Theorem. Starting from this result, we have studied in detail the finite, infinite countable or infinite non-countable cases, using for the first two the Montes approach and for the last one the Bertoluzza and Cariolaro's. Finally, in Section 6 we have introduced the dual concept of componible divergence measures.

8 Acknowledgements

This research in this communication has been supported in part by DGES Grant No. PB97-1286, a fellowship of Fundación Banco Herrero and INFM Section of Pavia; their financial support is gratefully acknowledged.

References

[1] C. Bertoluzza and I. Bonzani. Su una generalizzazione della nozione di diramativitá in teoria dell'informazione. *Stochastica*, VII(1):29–46, 1983. In Italian.

[2] C. Bertoluzza and D. Cariolaro. On the measure of a fuzzy set based on continuous t-conorms. *Fuzzy Sets and Systems*, (88):355–362, 1997.

[3] J. Kampé de Fériet, B. Forte, and P. Benvenuti. Forme générale de l'opération de composition continue d'une information. *C. R. Acad. Sci. Paris Sér. A-B*, 269:529–534, 1969. In French.

[4] D. Dubois and H. Prade. *Fuzzy Sets and Systems. Theory and Applications.* Academic Press Inc., 1980.

[5] B. Forte and N. Pintacuda. Sull'informazione associata alle esperienze incomplete. *Ann.Matematica pure e applicata*, Serie IV, Tomo LXXX:215–234, 1968. In Italian.

[6] J. Jiménez, S. Montes, and C. Bertoluzza. On the measure of a fuzzy set based on continuous t-conorms. In *Proceedings of EUSFLAT-ESTYLF Joint Conference*, pages 231–234, Palma de Mallorca (Spain), 1999.

[7] C.-H. Ling. Representation of associative functions. *Publications Math. Debrecen*, 12:189–212, 1965.

[8] K. Menger. Statistical metrics. *Proc. Nat. Acad. Sci.*, (28):537–537, 1942.

[9] P. Miranda, E. Torres, and P. Gil. Decision making and divergence measures: An application. *Belgian Journal of Operations Research, Statistics and Computer Science*, 38(2–3):51–66, 1998.

[10] S. Montes. *Particiones y medidas de divergencia en modelos difusos.* PhD thesis, University of Oviedo, 1998. In Spanish.

[11] S. Montes, I. Couso, P. Gil, and C. Bertoluzza. Divergence measure between fuzzy sets. *International Journal of Approximate Reasoning*, 30(2):91–105, August 2002.

[12] S. Montes, P. Gil, and C. Bertoluzza. Divergence between fuzzy sets and fuzziness. In *Proceedings of Seventh International Conference of Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU)*, volume II, pages 1030–1035, Paris, 1998.

[13] S. Montes, P. Miranda, J. Jiménez, and P. Gil. A measure of the difference between two fuzzy partitions. *Tatra Mountains Mathematical Publications*, 21:133–152, 2001.

[14] A. Salas, V. Doldi, and C. Bertoluzza. Pseudoquestionnaires in a deterministic environment. In *Proceedings of Eighth International Conference of Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU)*, volume II, pages 1177–1184, Madrid, 2000.

[15] M. Sugeno. *Theory of fuzzy integrals and its applications.* PhD thesis, Tokyo Institute of Technology, 1974.

[16] E. Torres, P. Miranda, and P. Gil. Relations among divergence measures and fuzzy measures. In *Proceedings of EUROFUSE-SIC'99*, pages 359–364, Budapest, 1999.