ARTICLE IN PRESS
FSS:8180

Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems ••• (••••) •••–•••

FUZZY sets and systems

www.elsevier.com/locate/fss

Statistical depth for fuzzy sets

Luis González-De La Fuente [a], Alicia Nieto-Reyes [a,*], Pedro Terán [b]

[a] *Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Spain*
[b] *Departamento de Estadística e Investigación Operativa y Didáctica de las Matemáticas, Universidad de Oviedo, Spain*

Received 4 December 2020; received in revised form 18 September 2021; accepted 18 September 2021

Abstract

Statistical depth functions provide a way to order the elements of a space by their centrality in a probability distribution. That has been very successful for generalizing non-parametric order-based statistical procedures from univariate to multivariate and (more recently) to functional spaces. We introduce two general definitions of statistical depth which are adapted to fuzzy data. For that purpose, two concepts of symmetric fuzzy random variables are introduced and studied. Furthermore, a generalization of Tukey's halfspace depth to the fuzzy setting is presented and proved to satisfy the above notions, through a detailed study of its properties.

Keywords: Fuzzy data; Fuzzy random variable; Nonparametric statistics; Statistical depth; Tukey depth

1. Introduction

L. A. Zadeh [45] introduced in 1965 the concept of a fuzzy set in order to formalize mathematically properties with unsharp boundaries or vaguely defined. Since then, the study of fuzzy sets has grown and is nowadays a field incorporating experts from different areas, including statistics. In 1975, Tukey coined the term statistical depth function [44]. Since $\mathbb{R}$ is totally ordered, one can easily say which of two points lies deeper in a distribution by comparing their quantiles. The point whose quantile is closer to .5 is considered deeper; the median is the deepest point since one must travel through at least half of the probability mass in order to leave the support of the distribution. Statistical depth aims at an analogous center–outward ordering of multivariate data in the space $\mathbb{R}^p$ which does not carry a natural order.

The use of depth functions has substantially grown over the years and it is a popular research topic in nonparametric statistics [28], with applications in regression [6], classification [19] or outlier detection [23], just to name a few examples. There exists a formal definition of statistical depth for multivariate spaces [47] and another for functional spaces [33]. Both are constructed on the basis of a list of desirable properties. In the multivariate case, they are affine equivariance, maximality at the center of symmetry, monotony with respect to the deepest element and vanishing at infinity. In the functional case: isometry equivariance, maximality at center, strict monotony with respect to the

* Corresponding author.
E-mail addresses: gdelafuentel@unican.es (L. González-De La Fuente), alicia.nieto@unican.es (A. Nieto-Reyes), teranpedro@uniovi.es (P. Terán).

https://doi.org/10.1016/j.fss.2021.09.015
0165-0114/

Please cite this article as: L. González-De La Fuente, A. Nieto-Reyes and P. Terán, Statistical depth for fuzzy sets, Fuzzy Sets and Systems, https://doi.org/10.1016/j.fss.2021.09.015

deepest element, upper semicontinuity, receptivity to the convex hull width across the domain, and continuity with respect to the distribution.

Since fuzzy sets are functions, that functional definition can potentially be applied to fuzzy data. However, that is at odds with how fuzzy data are understood in the literature. Functions have the variable under study in the y-axis, while the x-axis typically represents time in the analysis of functional data. For fuzzy sets, the values of the variable are in the x-axis while the y-axis represents the grade of membership. Similarly, the standard operations of fuzzy sets and functions do not match, and the analysis of a fuzzy set via its α-cuts does not have an obvious analog for functions.

An avenue for overcoming the mismatch is to identify each convex fuzzy set with its support function (like in [22]). Some properties demanded in the case of functional depth still do not suit fuzzy sets perfectly. For instance, receptivity (see [33]) does not have any clear meaning for fuzzy data. Also, demanding equivariance under arbitrary isometries has at least two disadvantages:

(1) The isometry groups of metric spaces of fuzzy sets being unknown, currently it is not possible to prove that this property holds for a given function. It may be possible if the definition of the function involves explicitly the set of all isometries, but then it would not be possible to compute the depth values themselves!
(2) Instead of arbitrary isometries, since fuzzy sets are objects in the underlying space $\mathbb{R}^p$ it looks more reasonable to consider equivariance under transformations of $\mathbb{R}^p$ instead of the whole space of fuzzy sets.

In view of that, we found it convenient to study a notion of depth which is specifically taylored to the particularities of fuzzy data (and, in particular, also set-valued data).

A random process generating fuzzy data is formalized with the notion of a *fuzzy random variable* (also called a random fuzzy set). Let us provide some pointers for readers with a background in statistical depth (and all readers). Attempts to combine probability and fuzzy sets started in the 70s. The modern approach to fuzzy random variables originates in [17,35,36]. Some books with a varying degree of mathematical or practical emphasis are [26,2,31,32], unfortunately there seem to be no general books more recent than 2006. There are some journal special issues, however, which include [9,8,16] from that date on. An interesting panorama of the analysis of fuzzy data appears in [3] which surveys the work of the SMIRE group. It describes some of the difficulties in handling fuzzy data, like the absence of a difference operation and the lack of parametric models. A nice guide to the semantics of fuzzy sets is Dubois and Prade's [14].

Our first main objective is to provide a formal definition of statistical depth in the space of fuzzy sets. In doing so, we present two feasible formal definitions. They are based on the multivariate and functional formal definitions: one is a natural extension to fuzzy sets of the multivariate notion and the other is closer to the functional one, in the sense that it uses explicitly a metric in the space of fuzzy sets; while the first one only depends on the operations of sum and product by scalars. We justify our two proposals by comparing and contrasting their defining properties which relate to the existing ones: affine equivariance, maximality at the center of symmetry, monotony with respect to the deepest point and vanishing at infinity. Additionally, we study the relationship between the multivariate notion of depth and our proposed notions of fuzzy depth when particularized to non-fuzzy points.

By analogy with the univariate case, the depth function of a symmetric distribution is expected to be maximized at the center of symmetry [48]. Thus this property requires of a notion of symmetry for fuzzy random variables. It turns out that an immediate extension of central symmetry (with which, e.g., X is symmetric with respect to the point 0 if and only if X and $(-1)\cdot X$ are identically distributed) is not best suited for the task and we need further notions which allow for symmetry with respect to more general fuzzy sets. We propose two such notions, which make use of the support function; one directly and the other through the mid and spread functions (see Section 2 for the required background). We analyze their relevance and contrast them through particular examples.

The most well-known statistical depth function is, in the multivariate case, the Tukey (or halfspace) depth [44], which has a computationally efficient approximation, the random Tukey depth [10]. The two other more well-known multivariate instances are spatial depth [37] and simplicial depth [27], which has in [20] a generalized version convenient for applications. Generally, multivariate instances do not behave properly in functional spaces [15], exceptions are the spatial [5] and the random Tukey [11] depths. Other functional instances are integrated depth [18], h-depth [12], band depth [29], elastic depth [23] and metric depth [34].

We extend Tukey depth to the fuzzy setting, showing that it falls under both notions of depth presented in this paper (semilinear and geometric depth). While it is clearly necessary to consider more depth functions, we found that mathematically studying the properties becomes quite harder due to the greater complexity of spaces of fuzzy sets compared to $\mathbb{R}^p$. For instance, in the case of the Tukey depth, we need a generalization of the notion of a halfspace but some complications appear since the space of fuzzy sets is not even a linear space. Similarly, an extension of simplicial depth would require working around the notion of simplices in the space of fuzzy sets. Hence we content ourselves with studying the Tukey depth in this paper and leave other depth functions to future work.

The structure of the paper is as follows. Section 2 contains the notation and background in fuzzy sets and statistical depth required for a comprehensive understanding of our proposals. The proposed formal definitions of fuzzy depth are in Section 3 and the proposed fuzzy symmetry notions in Section 4. The theoretical study of the properties that constitute the formal definitions is in Section 5. Section 6 presents and studies the generalization of Tukey's depth to the fuzzy setting. All proofs in the paper are deferred to Section 8. The paper concludes with some final remarks in Section 9.

2. Preliminaries and notation

This section contains the necessary definitions, notation and results which are used in the sequel.

A *fuzzy subset* of $\mathbb{R}^p$ is a function $A : \mathbb{R}^p \to [0, 1]$. By $\mathcal{F}_c(\mathbb{R}^p)$ we denote the class of fuzzy sets A on $\mathbb{R}^p$ such that each α-level of A is in the class $\mathcal{K}_c(\mathbb{R}^p)$, of all non-empty, compact and convex subsets of $\mathbb{R}^p$, for $\alpha \in [0, 1]$. The *α-levels* or *α-cuts* are the sets

$$A_\alpha := \{x \in \mathbb{R}^p : A(x) \geq \alpha\}, \quad \alpha \in (0, 1]$$

and the closed support A_0. Since fuzzy sets generalize the indicator function of a set, a fuzzy set having convex α-levels is a generalization of a convex set. It is thus called *convex*, although this property does not coincide with the usual convexity of functions (a fuzzy set is convex if and only if it is quasiconcave as a function). This work focuses on $\mathcal{F}_c(\mathbb{R}^p)$, thus when speaking of fuzzy sets we will implicitly mean elements of $\mathcal{F}_c(\mathbb{R}^p)$.

Note that $\mathcal{K}_c(\mathbb{R})$, in particular, is the class of non-empty compact intervals in $\mathbb{R}$. Let $\mathbb{S}^{p-1} := \{x \in \mathbb{R}^p : \|x\| \leq 1\}$ be the unit sphere on $\mathbb{R}^p$, with $\|.\|$ denoting the Euclidean norm. The *support function* of $A \in \mathcal{F}_c(\mathbb{R}^p)$ is defined as the mapping $s_A : \mathbb{S}^{p-1} \times [0, 1] \to \mathbb{R}$ such that

$$s_A(u, \alpha) := \sup\{\langle u, v\rangle : v \in A_\alpha\}, \tag{1}$$

for every $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0, 1]$, where $\langle \cdot, \cdot\rangle$ is the usual inner product in $\mathbb{R}^p$. By I_A we denote the indicator function of a set $A \subseteq \mathbb{R}^p$; i.e. $\mathrm{I}_A(t)$ is 1 if $t \in A$ and 0 otherwise. The symbol $=^d$ denotes equality in distribution.

In $\mathcal{F}_c(\mathbb{R})$ it is common to consider the *triangular fuzzy numbers* [24, Section 4.1], which will appear in some of our examples. For any $a \leq b \leq c$, the triangular fuzzy number $T(a, b, c)$ is the fuzzy set (tent function) given by

$$T(a,b,c)(x) := \begin{cases} \dfrac{x-a}{b-a} & \text{if } x \in [a,b] \text{ and } a < b, \\ \dfrac{x-c}{b-c} & \text{if } x \in [b,c] \text{ and } b < c \\ 1 & \text{if } x \in [a,b] \text{ and } a = b \text{ or } x \in [b,c] \text{ and } b = x, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

2.1. Arithmetics and Zadeh's extension principle

Operations in $\mathcal{F}_c(\mathbb{R}^p)$ are defined as follows.

Definition 2.1 *([45]).* Let $A, B \in \mathcal{F}_c(\mathbb{R}^p)$ and $\gamma \in \mathbb{R}$. Then, we define

- the *sum* $A + B$ as

$$(A+B)(t) := \sup_{x,y \in \mathbb{R}^p : x+y=t} \min\{A(x), B(y)\}, \text{ with } t \in \mathbb{R}^p$$

or, equivalently,

$$(A+B)_\alpha = \{x+y : x \in A_\alpha, y \in B_\alpha\}$$

for all $\alpha \in [0,1]$.

- the *product* $\gamma \cdot A$ of A *by a scalar* γ, as

$$(\gamma \cdot A)(t) := \sup_{x \in \mathbb{R}^p : t = \gamma \cdot x} A(y) = \begin{cases} A\left(\frac{t}{\gamma}\right), & \text{si } \gamma \neq 0 \\ I_{\{0\}}(t) & \text{si } \gamma = 0, \end{cases}$$

for $t \in \mathbb{R}^p$. Equivalently,

$$(\gamma \cdot A)_\alpha = \{\gamma \cdot x : x \in A_\alpha\}$$

for all $\alpha \in [0,1]$.

While the sum has $I_{\{0\}}$ as its neutral element, it is important to notice that $(-1) \times A$ is not the additive inverse of A in general. Taking the support function is a linear operator in the sense that

$$s_{\gamma \cdot A + \gamma' B} = \gamma \cdot s_A + \gamma' \cdot s_B$$

whenever $\gamma, \gamma' \geq 0$ and $A, B \in \mathcal{F}_c(\mathbb{R}^p)$.

Zadeh's extension principle [46] allows a crisp function (non-fuzzy) to act on fuzzy sets in the following way. Given a function $f : \mathbb{R}^p \to \mathbb{R}^p$ and $A \in \mathcal{F}_c(\mathbb{R}^p)$, the image $f(A) \in \mathcal{F}_c(\mathbb{R}^p)$ is defined to be

$$f(A)(t) := \sup\{A(y) : y \in \mathbb{R}^p, f(y) = t\}$$

for all $t \in \mathbb{R}^p$. For instance, the operations defined above are consistent with Zadeh's extension principle.

For an arbitrary f, it may happen that $A \in \mathcal{F}_c(\mathbb{R}^p)$ does not imply $f(A) \in \mathcal{F}_c(\mathbb{R}^p)$. However, we will only apply Zadeh's extension to continuous functions, where that cannot happen. The main example in this paper is the following.

Let $M \in \mathcal{M}_{p\times p}(\mathbb{R})$ be a non-singular matrix and let $f : \mathbb{R}^p \to \mathbb{R}^p$ be given by $f(x) = M \cdot x$. Then Zadeh's extension yields, for any $A \in \mathcal{F}_c(\mathbb{R}^p)$, a value $f(A) = M \cdot A \in \mathcal{F}_c(\mathbb{R}^p)$ as

$$(M \cdot A)(t) := \sup\{A(y) : y \in \mathbb{R}^p, M \cdot y = t\}.$$

Since f is bijective, actually

$$(M \cdot A)(t) = A(M^{-1} \cdot t).$$

An important tool in order to deal with support functions of fuzzy sets is the *mid/spr* decomposition [43]. Given $A \in \mathcal{F}_c(\mathbb{R}^p)$, the support function of A, s_A, can be expressed as $s_A = \text{mid}(s_A) + \text{spr}(s_A)$, where the mid and spread of a function $f : \mathbb{S}^{p-1} \times [0,1] \to \mathbb{R}$ are defined by

$$\text{mid}(f(u,\alpha)) := (f(u,\alpha) - f(-u,\alpha))/2 \tag{3}$$

$$\text{spr}(f(u,\alpha)) := (f(u,\alpha) + f(-u,\alpha))/2. \tag{4}$$

Analogously, for any $K \in \mathcal{K}_c(\mathbb{R})$, we set

$$\text{mid}(K) = (\sup(K) + \inf(K))/2 \text{ and } \text{spr}(K) = (\sup(K) - \inf(K))/2. \tag{5}$$

2.2. Metrics in $\mathcal{F}_c(\mathbb{R}^p)$

There are several different metrics in $\mathcal{F}_c(\mathbb{R}^p)$. We will use some of the best known: d_r, which use the Hausdorff metric over the α-levels, and ρ_r, which are L^r-type metrics.

First of all, the *Hausdorff metric* in $\mathcal{K}_c(\mathbb{R}^p)$ is given by

$$d_H(S,T) := \max\left\{\sup_{s\in S}\inf_{t\in T} \| s-t \|, \sup_{t\in T}\inf_{s\in S} \| s-t \|\right\}$$

for any $S, T \in \mathcal{K}_c(\mathbb{R}^p)$. Thus, given two fuzzy sets $A, B \in \mathcal{F}_c(\mathbb{R}^p)$, the following family of metrics makes use of the Hausdorff distance between α-levels:

$$d_r(A,B) := \begin{cases} \left(\int_0^1 (d_H(A_\alpha, B_\alpha))^r \, d\alpha\right)^{1/r}, & r \in [1,\infty) \\ \sup_{\alpha\in[0,1]} d_H(A_\alpha, B_\alpha), & r = \infty. \end{cases}$$

While $(\mathcal{F}_c(\mathbb{R}^p), d_r)$ is a non-complete and separable metric space, for any $r \in (1,\infty)$, the metric space $(\mathcal{F}_c(\mathbb{R}^p), d_\infty)$ is non-separable and complete (see [13]).

It is also possible to consider L^r-type metrics (see [13]).

$$\rho_r(A,B) := \left(\int_{\mathbb{S}^{p-1}} \int_{[0,1]} \|s_A - s_B\|^r \, du \, d\alpha\right)^{1/r}, \tag{6}$$

for all $A, B \in \mathcal{F}_c(\mathbb{R}^p)$. The most commonly used metrics are the cases $r = 1$ and $r = 2$.

2.3. Fuzzy random variables

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

Definition 2.2 *([30]).* A *random compact convex set* (or simply a *random set*) is a function $\Gamma : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ such that $\{\omega \in \Omega : \Gamma(\omega) \cap K \neq \emptyset\} \in \mathcal{A}$ for all $K \in \mathcal{K}_c(\mathbb{R}^p)$.

Note that this definition is equivalent to demanding Γ to be measurable with respect to the Borel σ-algebra induced by the Hausdorff metric.

Definition 2.3 *([36]).* A *fuzzy random variable* is a function $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ such that $\mathcal{X}_\alpha(\omega)$ is a random compact set for all $\alpha \in [0,1]$, where $\mathcal{X}_\alpha : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$ is defined as $\mathcal{X}_\alpha(\omega) := (\mathcal{X}(\omega))_\alpha$ for any $\omega \in \Omega$.

It is not explicit in this definition that a fuzzy random variable is a measurable function in the ordinary sense. But clearly, $\mathcal{X}$ is a fuzzy random variable if and only if it is measurable when $\mathcal{F}_c(\mathbb{R}^p)$ is endowed with the σ-algebra generated by the α-cut mappings $L_\alpha : A \in \mathcal{F}_c(\mathbb{R}^p) \mapsto A_\alpha \in \mathcal{K}_c(\mathbb{R}^p)$, namely the smallest σ-algebra which makes each L_α measurable. As shown by Krätschmer [25], that is the Borel σ-algebra generated by any of the metrics d_r or ρ_r for $r \in [1,\infty)$. The Borel σ-algebra of d_∞ is strictly larger.

Given a fuzzy random variable $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$, we define $s_\mathcal{X}$ the support function of $\mathcal{X}$ as the function $s_\mathcal{X} : \mathbb{S}^{p-1} \times [0,1] \times \Omega \to \mathbb{R}$ defined by

$$s_\mathcal{X}(u,\alpha,\omega) := s_{\mathcal{X}(\omega)}(u,\alpha), \tag{7}$$

for all $u \in \mathbb{S}^{p-1}, \alpha \in [0,1]$ and $\omega \in \Omega$. Whenever $\mathcal{X}$ is a fuzzy random variable, each $s_\mathcal{X}(u,\alpha)$ is a real random variable.

Definition 2.4 *([39]).* Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probabilistic space associated with the fuzzy random variable $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$. A *1-median* of the distribution of $\mathcal{X}$ is any fuzzy set $\tilde{M}e(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ such that $\mathrm{E}(\rho_1(\mathcal{X}, \tilde{M}e(\mathcal{X}))) = \min_{U\in\mathcal{F}_c(\mathbb{R})} \mathrm{E}(\rho_1(\mathcal{X}, U))$.

Given a real random variable X, $\mathrm{Med}(X)$ denotes the median of X. $\mathrm{Med}(X)$ does not need to be unique and will be generally assumed to be a set of points. In the particular cases in which $\mathrm{Med}(X)$ will denote a representative of the set instead of the whole set, we will specifically point it out.

Theorem 2.5 *([39]). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probabilistic space associated with the fuzzy random variable $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R})$. We have that for any $\alpha \in [0,1]$, the fuzzy set(s), $\tilde{M}e(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$, satisfies that*

$$(\tilde{M}e(\mathcal{X}))_\alpha = [\mathrm{Med}(\inf(\mathcal{X}_\alpha)), \mathrm{Med}(\sup(\mathcal{X}_\alpha))]. \tag{8}$$

In (8), if $\mathrm{Med}(\inf \mathcal{X}_\alpha)$ and/or $\mathrm{Med}(\sup \mathcal{X}_\alpha)$ are not unique, they are defined as the midpoint of the set of medians.

2.4. Basic definitions about Banach spaces

Definition 2.6. A normed vector space $(\mathbb{E}, \|\cdot\|)$ is called a Banach space if it is complete with respect to the metric derived from its norm, that is, every Cauchy sequence is convergent.

Definition 2.7 *([21]).* A Banach space $(\mathbb{E}, \|\cdot\|)$ is said to be *strictly convex* if $x = y$ whenever $\|(1/2)\cdot(x+y)\| = \|x\| = \|y\|$, for all $x, y \in \mathbb{E}$.

The following result presents two characterizations of strictly convex Banach spaces.

Theorem 2.8 *([21]). Let $(\mathbb{E}, \|\cdot\|)$ be a Banach space. The following properties are equivalent.*

1. $(\mathbb{E}, \|\cdot\|)$ *is strictly convex.*
2. *For all $x, y \in \mathbb{E}$ such that $x \neq y$ and $\|x\| = \|y\| = 1$, we have that $\|x+y\| < 2$.*
3. *For all $x, y \in \mathbb{E}$ such that $x \neq y$ and $\|x\| = \|y\| = 1$, we have that $\|(1-\lambda)\cdot x + \lambda \cdot y\| < 1$ for all $\lambda \in (0, 1)$.*

2.5. Multivariate statistical depth

The definition of a *statistical depth function* is as follows.

Definition 2.9 *([47]).* Let $\mathcal{H}$ be a class of random variables. Let $D(\cdot;\cdot): \mathbb{R}^p \times \mathcal{H} \to [0, \infty)$ be a map satisfying:

(**M1**.) $D(M\cdot x + b; M\cdot X + b) = D(x; X)$, for any $X \in \mathcal{H}$, any non-singular matrix $M \in \mathcal{M}_{p\times p}(\mathbb{R})$ and any $x, b \in \mathbb{R}^p$.
(**M2**.) $D(\theta; X) = \sup_{x\in\mathbb{R}^p} D(x; X)$, for any $X \in \mathcal{H}$ with center θ with respect to some notion of symmetry.
(**M3**.) For any $X \in \mathcal{H}$ having deepest point θ, $D(x; X) \leq D((1-\lambda)\theta + \lambda x; X)$, for any $\lambda \in [0, 1]$.
(**M4**.) $D(x; X) \to 0$ as $\|x\| \to \infty$, for each $X \in \mathcal{H}$.

Then $D(\cdot;\cdot)$ is called a *statistical depth function* or just a depth function. To distinguish the multivariate and fuzzy cases, we also call it a *multivariate depth function*.

Recall the notion of central symmetry of a random variable.

Definition 2.10 *([47]).* Let $X: \Omega \to \mathbb{R}^p$ be a random variable associated with a probabilistic space $(\Omega, \mathcal{A}, \mathbb{P})$ and let $c \in \mathbb{R}^p$. X is *centrally symmetric* with respect to c if $X - c =^d c - X$.

It must also be emphasized that it is customary in the statistical community to regard those properties as favorable features ('desirable properties' [47, Section 2.1]) rather than proper axioms. The intuitive rationale behind properties M1–M4 is the following.

M1. The depth of a point should not depend on the coordinate system we choose to use.
M2. The center of symmetry of a distribution, if it exists, should be the deepest point in the distribution.
M3. Depth should increase when we follow a straight line from any point towards a point of maximal depth.
M4. Depth should decrease to 0 as we move towards infinity.

While these intuitive explanations make M1–M4 plausible requirements, it is hard to view them as forceful axioms since it is unclear, in most cases, that dangerous consequences would necessarily follow from omitting one or another. That explains that some functions which do not satisfy all those properties are generally accepted and routinely called depth functions.

3. Definition of statistical depth function for fuzzy sets

This section is devoted to proposing two notions of statistical depth function for fuzzy data, specifically for fuzzy sets in $\mathcal{F}_c(\mathbb{R}^p)$. They are based on the notions of statistical depth function in $\mathbb{R}^p$ (Section 2.5) and function spaces. However, due to the special properties of spaces of fuzzy sets (in particular, they are not linear spaces and more than one metric is available), some adaptations become necessary.

More precisely, the intuitive properties that depth decreases along any ray departing from a deepest point, and that depth must tend to 0 when moving outwards 'towards infinity', must be reimagined in the context of fuzzy sets. The former, for instance, will be expressed in terms of betweenness: all fuzzy sets which are *between* a given fuzzy set and a maximally deep fuzzy set, must have an intermediate depth. The notion of being between can be formalized by resorting to the algebraic structure (i.e., a convex combination of two fuzzy sets is declared to be between them) or to the geometric structure (i.e., a fuzzy set which splits the distance between other two is declared to be between them). In $\mathbb{R}^p$ or, in the context of functional data, in Hilbert spaces, these two notions are equivalent, but they must be distinguished in our context. Similar considerations affect the notion of 'moving towards infinity'.

That leads us to proposing two different notions or 'styles' of depth, one algebraic and the other geometric. Since the structure of fuzzy set spaces is often called semilinear (as it has a sum and a product by scalars but fails some key axioms of a linear space), we call them semilinear depth and geometric depth.

Let $L^0[\mathcal{F}_c(\mathbb{R}^p)]$ be the class of all fuzzy random variables on the measurable space $(\Omega, \mathcal{A})$, and consider subsets $\mathcal{H} \subseteq L^0[\mathcal{F}_c(\mathbb{R}^p)]$ and $\mathcal{J} \subseteq \mathcal{F}_c(\mathbb{R}^p)$. A *statistical depth function* is a mapping

$$D(\cdot;\cdot) : \mathcal{J} \times \mathcal{H} \to [0, \infty)$$

that satisfies a number of properties listed below. Subset $\mathcal{H}$ accounts for the fact that many methods do not provide a depth function for every single random object (for instance, a definition may involve expectations and so it is valid only provided those expectations exist). Usually $\mathcal{J}$ is the whole $\mathcal{F}_c(\mathbb{R}^p)$, however one may want to consider depth functions restricted to specific subclasses of fuzzy sets, like triangular, trapezoidal and LR fuzzy numbers [24, Section 4.1].

The first two desirable properties are the same for both semilinear depth and geometric depth:

(**P1**.) $D(M \cdot A + B; M \cdot \mathcal{X} + B) = D(A; \mathcal{X})$ for any non-singular matrix $M \in \mathcal{M}_{p\times p}(\mathbb{R})$, any $A, B \in \mathcal{J}$ and any $\mathcal{X} \in \mathcal{H}$.

(**P2**.) For (some notion of symmetry and) any symmetric fuzzy random variable $\mathcal{X} \in \mathcal{H}$,

$$D(U; \mathcal{X}) = \sup_{B\in\mathcal{F}_c(\mathbb{R}^p)} D(B; \mathcal{X}),$$

where $U \in \mathcal{J}$ is a center of symmetry of $\mathcal{X}$.

The product of a matrix and a fuzzy set is performed with Zadeh's extension principle (Section 2.1) and represents the change of the fuzzy set under a coordinate change in $\mathbb{R}^p$ which preserves the origin. Property P1 is a generalization of the affine equivariance property of multivariate depth. Property P2 formally coincides with its analog in the multivariate [47] and functional [33] cases. However, some notion of symmetry suitable for the fuzzy case needs to be considered. In Section 4 below we provide two notions of symmetry for fuzzy random variables in order to make property P2 operative. Symmetry is meant here as a property of the probability distribution of the fuzzy random variable, not of the individual fuzzy sets it takes on as values. Also, in the literature of statistical depth it is not meant that a depth function must satisfy P2 simultaneously for every conceivable definition of symmetry, but only that it is desirable that symmetrically distributed random objects have the center of symmetry as the maximally deep object for some notion of symmetry.

Semilinear depth and geometric depth differ in how further properties are adapted. We complete their definitions as follows.

Definition 3.1. Let $\mathcal{H} \subseteq L^0[\mathcal{F}_c(\mathbb{R}^p)]$ and $\mathcal{J} \subseteq \mathcal{F}_c(\mathbb{R}^p)$. A mapping $D(\cdot;\cdot) : \mathcal{J} \times \mathcal{H} \to [0, \infty)$ is a *semilinear depth function* if it satisfies P1 – P2 and, for each fuzzy random variable $X \in \mathcal{H}$ and each $A \in \mathcal{J}$ such that $D(A; \mathcal{X}) = \sup\{D(B; \mathcal{X}) : B \in \mathcal{J}\}$, the following two properties are met.

(**P3a.**) $D(A;\mathcal{X}) \geq D((1-\lambda)\cdot A + \lambda \cdot B;\mathcal{X}) \geq D(B;\mathcal{X})$ for all $\lambda \in [0,1]$ and all $B \in \mathcal{F}_c(\mathbb{R}^p)$.
(**P4a.**) $\lim_{\lambda\to\infty} D(A+\lambda \cdot B;\mathcal{X}) = 0$ for all $B \in \mathcal{J} \setminus \{\mathrm{I}_{\{0\}}\}$.

Definition 3.2. Let $\mathcal{H} \subseteq L^0[\mathcal{F}_c(\mathbb{R}^p)]$, $\mathcal{J} \subseteq \mathcal{F}_c(\mathbb{R}^p)$ and d a metric in $\mathcal{F}_c(\mathbb{R}^p)$. A mapping $D(\cdot;\cdot): \mathcal{J}\times\mathcal{H} \to [0,\infty)$ is a *geometric depth function* with respect to d if it satisfies P1 – P2 and, for each fuzzy random variable $X \in \mathcal{H}$ and each $A \in \mathcal{J}$ such that $D(A;\mathcal{X}) = \sup\{D(B;\mathcal{X}) : B \in \mathcal{J}\}$, the following two properties are met.

(**P3b**.) $D(A;\mathcal{X}) \geq D(B;\mathcal{X}) \geq D(C;\mathcal{X})$ for all $B, C \in \mathcal{J}$ satisfying $d(A,C) = d(A,B) + d(B,C)$.
(**P4b**.) $\lim_{n\to\infty} D(A_n;\mathcal{X}) = 0$ for every sequence of fuzzy sets $\{A_n\}_n$ such that the $\lim_{n\to\infty} d(A_n, A) = \infty$.

Properties P3b and P4b are based also on the functional case, in the sense that we consider the space of fuzzy sets as a metric space and formulate these properties with respect to a specific metric. Concerning P4b, notice that, by the triangle inequality, A can be replaced by any other fuzzy set like $I_{\{0\}}$, which makes it possible to check P4b without computing the fuzzy sets with maximum depth.

We will study the relationships between these two definitions in Section 5.

The distribution of a fuzzy random variable is the probability measure it induces in $\mathcal{F}_c(\mathbb{R}^p)$. Note that the examples of statistical depth considered in this paper only depend on the fuzzy random variable via its distribution, i.e., identically distributed fuzzy random variables will share the same depth function. Thus there is no essential difference with defining depth as a function of distributions instead of variables, as more commonly found in the statistical literature. Using variables will improve the readability of some proofs.

4. Symmetry for fuzzy random variables

Property P2 requires the use of some notion of symmetry for fuzzy random variables. It is clear that one can define central symmetry (Section 2.5) with respect to a point c by requiring $\mathcal{X}$ to be such that $\mathcal{X} - \mathrm{I}_{\{c\}}$ and its product by -1 are identically distributed (see, e.g., [40] among other papers). However, that would lead to a very weak variant of P2 as it would be void whenever maximally deep fuzzy sets are not crisp points; i.e., whenever they are not points in $\mathbb{R}^p$. Therefore, we need a notion which allows for symmetry with respect to a more general $A \in \mathcal{F}_c(\mathbb{R}^p)$ rather than only crisp points.

We will introduce now two such notions. They make use of the support function, in (1) and (7), as a way of considering the space of fuzzy sets as a space of real functions defined over the compact set $\mathbb{S}^{p-1} \times [0,1]$. Definition 4.1 below provides a notion of symmetry that takes into account measures of shape and location of the symmetric fuzzy set with respect to those of the fuzzy random variable. That is due to the fact that the support function of a fuzzy set provides information about the α-level boundaries of the fuzzy set. Definition 4.2, meanwhile, makes use of the functions mid and spr over the support function, whose definitions are in (3) and (4), respectively.

Definition 4.1. Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ be a fuzzy random variable and $A \in \mathcal{F}_c(\mathbb{R}^p)$ a fuzzy set. $\mathcal{X}$ is *F-symmetric* with respect to A if

$$s_A(u,\alpha) - s_{\mathcal{X}}(u,\alpha) =^d s_{\mathcal{X}}(u,\alpha) - s_A(u,\alpha),$$

for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$.

Equivalently, $\mathcal{X}$ is F-symmetric with respect to A if each $s_{\mathcal{X}}(u,\alpha)$ is centrally symmetric with respect to $s_A(u,\alpha)$.

Definition 4.2. Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ be a fuzzy random variable and $A \in \mathcal{F}_c(\mathbb{R}^p)$ a fuzzy set. $\mathcal{X}$ is (mid, spr)-*symmetric* with respect to A if the following two conditions are satisfied for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$:

$$\mathrm{mid}(s_A(u,\alpha)) - \mathrm{mid}(s_{\mathcal{X}}(u,\alpha)) =^d \mathrm{mid}(s_{\mathcal{X}}(u,\alpha)) - \mathrm{mid}(s_A(u,\alpha)),$$
$$\mathrm{spr}(s_A(u,\alpha)) - \mathrm{spr}(s_{\mathcal{X}}(u,\alpha)) =^d \mathrm{spr}(s_{\mathcal{X}}(u,\alpha)) - \mathrm{spr}(s_A(u,\alpha)).$$

Through the next examples, we show the similarities and differences between Definitions 4.1 and 4.2.

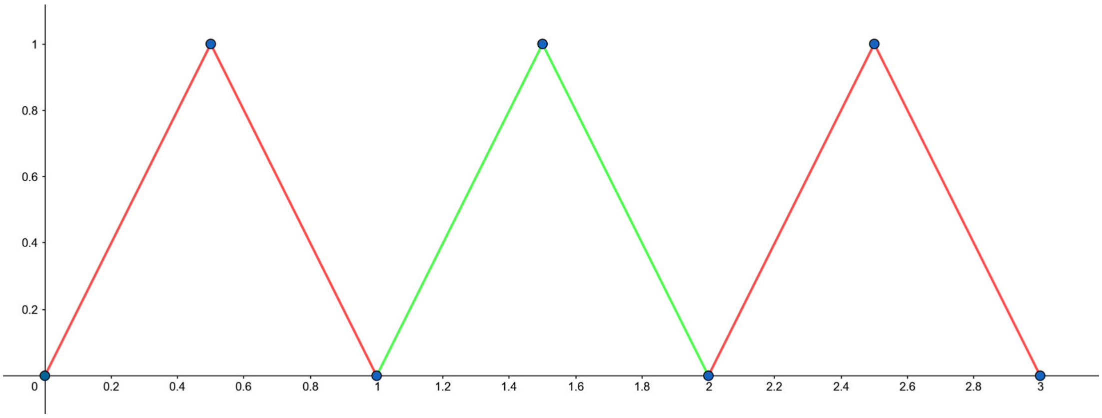

Fig. 1. Illustration of the triangular fuzzy sets T(0,0.5,1) and T(2,2.5,3) in red and A in green. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Example 4.3. Given the probability space $(\{\omega_1, \omega_2\}, \mathcal{P}(\{\omega_1, \omega_2\}), \mathbb{P})$ with $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2)$, let $\mathcal{X} : \{\omega_1, \omega_2\} \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable over it such that $\mathcal{X}(\omega_1)$ and $\mathcal{X}(\omega_2)$ are the triangular fuzzy numbers $T(0{,}0.5{,}1)$ and $T(2{,}2.5{,}3)$, respectively (see Section 2.3 for the definition).

The fuzzy random variable $\mathcal{X}$ is F-symmetric with respect to the triangular fuzzy number $A := T(1, 1.5, 2)$ (see Fig. 1). That is due to the following. For any $\alpha \in [0, 1]$,

$$(\mathcal{X}(\omega_1))_\alpha = [\alpha/2, 1 - \alpha/2], \quad (\mathcal{X}(\omega_2))_\alpha = [2 + \alpha/2, 3 - \alpha/2]$$

and

$$A_\alpha = [1 + \alpha/2, 2 - \alpha/2].$$

Since $\mathbb{S}^0 = \{-1, 1\}$,

$$s_{\mathcal{X}(\omega_i)}(1, \alpha) = \begin{cases} 1 - \alpha/2, & i = 1, \\ 3 - \alpha/2, & i = 2 \end{cases}$$

and

$$s_{\mathcal{X}(\omega_i)}(-1, \alpha) = \begin{cases} -\alpha/2, & i = 1, \\ -2 - \alpha/2, & i = 2. \end{cases}$$

Their centers of symmetry are, respectively, $2 - \alpha/2$ and $-1 - \alpha/2$. Since

$$s_A(1, \alpha) = 2 - \alpha/2, \quad s_A(-1, \alpha) = -1 - \alpha/2,$$

indeed each $s_{\mathcal{X}}(u, \alpha)$ is centrally symmetric with respect to $s_A(u, \alpha)$, namely $\mathcal{X}$ is F-symmetric with respect to A.

One similarly shows that the fuzzy random variable $\mathcal{X}$ is also (mid, spr)-symmetric with respect to A.

Observe that $\mathcal{X}(\omega_1)$ and $\mathcal{X}(\omega_2)$ are centrally symmetric with respect to the crisp point 1.5. But $I_{\{1.5\}}$ is not a reasonable candidate for being maximally deep in the distribution of $\mathcal{X}$. The variable takes on fuzzy values with the same triangle shape, while the supposed center of the distribution, $I_{\{1.5\}}$, is entirely somewhere else in the space $\mathcal{F}_c(\mathbb{R}^p)$ while also not being appropriate from the perspective of Property P4. On the contrary, $A = T(1, 1.5, 2)$, which is the center of F-symmetry and (mid, spr)-symmetry, is the natural candidate.

Example 4.4. Given the probability space $(\{\omega_1, \omega_2, \omega_3\}, \mathcal{P}(\{\omega_1, \omega_2, \omega_3\}), \mathbb{P})$, with $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \mathbb{P}(\omega_3)$, let $\mathcal{X} : \{\omega_1, \omega_2, \omega_3\} \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable over it such that $\mathcal{X}(\omega_1) = \mathrm{I}_{[-2,4]}$, $\mathcal{X}(\omega_2) = \mathrm{I}_{[0,2]}$ and $\mathcal{X}(\omega_3) = \mathrm{I}_{[2,6]}$.

Definitions 4.1 and 4.2 are not equivalent. This can be deduced from $\mathcal{X}$ being F-symmetric with respect to the fuzzy set $\mathrm{I}_{[0,4]}$ but not (mid, spr)-symmetric. One can check

$$s_{\mathcal{X}(\omega_i)}(1,\alpha) = \begin{cases} 4, & i=1 \\ 2, & i=2 \\ 6, & i=3 \end{cases}$$

for all $\alpha \in [0,1]$, whence $s_{\mathcal{X}(\omega_i)}(1,\alpha)$ is centrally symmetric with respect to 4, which is $s_{\mathrm{I}_{[0,4]}}(1,\alpha)$. Similarly,

$$s_{\mathcal{X}(\omega_i)}(-1,\alpha) = \begin{cases} 2, & i=1 \\ 0, & i=2 \\ -2, & i=3 \end{cases}$$

whence $s_{\mathcal{X}(\omega_i)}(-1,\alpha)$ is centrally symmetric with respect to 0, which is $s_{\mathrm{I}_{[0,4]}}(-1,\alpha)$. Thus F-symmetry with respect to $\mathrm{I}_{[0,4]}$ is proved. However,

$$\mathrm{mid}(s_{\mathcal{X}(\omega_i)}(1,\alpha)) = \begin{cases} 1, & i=1 \\ 1, & i=2 \\ 4, & i=3 \end{cases}$$

which shows that $\mathrm{mid}(s_{\mathcal{X}(\omega_i)}(1,\alpha))$ takes on values 1 and 4 with probabilities 2/3 and 1/3, i.e. is not centrally symmetric. Thus $\mathcal{X}$ cannot be (mid, spr)-symmetric.

Example 4.5. Given the probability space $(\{\omega_1,\omega_2,\omega_3,\omega_4\}, \mathcal{P}(\{\omega_1,\omega_2,\omega_3,\omega_4\}), \mathbb{P})$, with $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \mathbb{P}(\omega_3) = \mathbb{P}(\omega_4)$, let $\mathcal{X} : \{\omega_1,\omega_2,\omega_3,\omega_4\} \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable over it such that $\mathcal{X}(\omega_1) = \mathrm{I}_{[1,5]}$, $\mathcal{X}(\omega_2) = \mathrm{I}_{[-2,4]}$, $\mathcal{X}(\omega_3) = \mathrm{I}_{[1,3]}$ and $\mathcal{X}(\omega_4) = \mathrm{I}_{[2,6]}$.

This example presents a (mid, spr)-symmetric fuzzy random variable which is not F-symmetric. Indeed, for any $\alpha \in [0,1]$ and $u \in \{-1,1\}$,

$$\mathrm{mid}(s_{\mathcal{X}(\omega_i)}(u,\alpha)) = \begin{cases} 3u, & i=1 \\ u, & i=2 \\ 2u, & i=3 \\ 4u, & i=4, \end{cases}$$

which is centrally symmetric with respect to 2.5u. Moreover,

$$\mathrm{spr}(s_{\mathcal{X}(\omega_i)}(u,\alpha)) = \begin{cases} 2, & i=1 \\ 3, & i=2 \\ 1, & i=3 \\ 2, & i=4, \end{cases}$$

which is centrally symmetric with respect to 2. The fuzzy set with those values of mid and spread is $\mathrm{I}_{[.5,4.5]}$, hence $\mathcal{X}$ is (mid,spr)-symmetric with respect to it.

However,

$$s_{\mathcal{X}(\omega_i)}(-1,\alpha) = \begin{cases} -1, & i=1 \\ 2, & i=2 \\ -1, & i=3 \\ -2, & i=4. \end{cases}$$

Its distribution takes on values $-2, -1, 2$ with respective probabilities .25, .5, .25 which is not a centrally symmetric distribution: the only possibility is central symmetry with respect to the point with probability .5; but it is not the midpoint between the other two.

Next we provide a result on the median of a function of symmetric fuzzy random variables.

Lemma 4.6. *Let $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ be a fuzzy random variable over a probabilistic space $(\Omega, \mathcal{A}, \mathbb{P})$ and $A \in \mathcal{F}_c(\mathbb{R}^p)$ a fuzzy set.*

- *If $\mathcal{X}$ is F-symmetric with respect to $A \in \mathcal{F}_c(\mathbb{R}^p)$ then $s_A(u, \alpha)$ is a median of $s_{\mathcal{X}}(u, \alpha)$ for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0, 1]$.*
- *If $\mathcal{X}$ is (mid, spr)-symmetric with respect to an $A \in \mathcal{F}_c(\mathbb{R}^p)$ then $mid(s_A(u, \alpha))$ is a median of $mid(s_{\mathcal{X}}(u, \alpha))$ and $spr(s_A(u, \alpha))$ is a median of $spr(s_{\mathcal{X}}(u, \alpha))$ for each $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0, 1]$.*

Let $\Gamma : \Omega \to \mathcal{K}_c(\mathbb{R})$ be a random interval, associated with a probabilistic space $(\Omega, \mathcal{A}, \mathbb{P})$. In [38], a median of a random interval is defined to be a compact interval $\mathrm{Me}[\Gamma] \in \mathcal{K}_c(\mathbb{R})$ such that $\mathrm{mid}(\mathrm{Me}[\Gamma]) = \mathrm{Med}(\mathrm{mid}(\Gamma))$ and $\mathrm{spr}(\mathrm{Me}[\Gamma]) = \mathrm{Med}(\mathrm{spr}(\Gamma))$, see (5). Thus, by Lemma 4.6, if I_Γ is (mid, spr)-symmetric with respect to the indicator function of a compact interval $A \in \mathcal{K}_c(\mathbb{R})$, then A is the median of the random interval Γ.

5. Relationship between depth properties

The aim of this section is to clarify the relationships between both proposed definitions of depth, and between them and multivariate depth. Since geometric depth depends on the choice of a metric, its defining properties cannot be expected to be equivalent, in general, to those of semilinear depth. We will give sufficient conditions for P3a to imply P3b and vice versa, as well as for P4b to imply P4a. Those conditions are satisfied, for example, by the ρ_r-metrics for $r \in (1, \infty)$. Examples will be provided in which the implications fail.

Property P4a does not imply P4b since it fails to imply M4 when restricted to the multivariate case, as discussed below. Hence we will propose an alternative P4a* so that properties P1–P3 and P4a* become equivalent to M1–M4 in the multivariate case.

The subset

$$\mathcal{R}^p = \{\mathrm{I}_{\{x\}} \in \mathcal{F}_c(\mathbb{R}^p) : x \in \mathbb{R}^p\}$$

of $\mathcal{F}_c(\mathbb{R}^p)$ and $\mathbb{R}^p$ can be identified and the operations sum and product by a scalar are preserved, i.e. $\mathrm{I}_{\{x\}} + \mathrm{I}_{\{y\}} = \mathrm{I}_{\{x+y\}}$ and $\gamma \cdot \mathrm{I}_{\{x\}} = \mathrm{I}_{\{\gamma \cdot x\}}$, for all $x, y \in \mathbb{R}^p$ and for all $\gamma \in \mathbb{R}$. It seems, then, natural to study the relation between both notions of fuzzy depth and that of multivariate depth function, reproduced in Section 2.5 together with the central symmetry notion.

Proposition 5.1. *Take F-symmetry to be the notion of symmetry for Definitions 3.1 and 3.2, and take central symmetry for Definition 2.9. Then for any of the metrics ρ_2 and d_r ($r \in [1, \infty]$), Definition 3.2 restricted to $\mathcal{R}^p$ and Definition 2.9 are equivalent. Additionally, Definition 2.9 implies Definition 3.1 restricted to $\mathcal{R}^p$ and the properties P3a and M3 are equivalent.*

The next proposition proves the equivalence of Definition 2.9 and Definition 3.1 restricted to $\mathcal{R}^p$ when P4a is replaced by

(P4a$*$.) $\lim_{n\to\infty} D(A + B_n; \mathcal{X}) = 0$ for all $B_n \in \mathcal{F}_c(\mathbb{R}^p)$ with $n \in \mathbb{N}$ such that there exists $i \in \{1, 2, \cdots, p\}$ and $\alpha \in [0, 1]$ with $\lim_{n\to\infty} |s_{B_n}(e_i, \alpha)| = \infty$.

Here, $\{e_i\}_{i=1}^p$ is the canonical basis of $\mathbb{R}^p$. Clearly, it is equivalent to consider only the case $\alpha = 0$.

Proposition 5.2. *The following hold:*

(a) *Property P4a$*$ restricted to $\mathcal{R}^p$ is equivalent to M4.*
(b) *Property M4 implies P4a restricted to $\mathcal{R}^p$.*
(c) *Property P4b (for the d_∞-metric) restricted to $\mathcal{R}^p$ implies M4.*

Property M4 can be rewritten by replacing the condition $\|x\| \to \infty$ by the equivalent one that $x_i \to \infty$ for some $i = 1, \ldots, p$ (where $x = (x_1, \ldots, x_p) \in \mathbb{R}^p$). By doing this, one omits any reference to the metric (norm) of $\mathbb{R}^p$ and focuses on the vector coordinates, which are part of the algebraic structure. That is the idea behind P4a*, however we do not feel that property P4a* is purely algebraic since it involves support functions, which are still geometric in nature. Thus we chose the weaker P4a for the definition of semilinear depth; but P1–P3 and P4a* are a valid generalization of the notion of multivariate depth to the fuzzy setting.

The rest of the section focuses on comparing the third and fourth properties in Definitions 3.1 and 3.2, since P1–P2 are common. That will give us a better understanding of the relationships between the geometric and semilinear forms of depth.

For that purpose, we consider metrics $d : \mathcal{F}_c(\mathbb{R}^p) \times \mathcal{F}_c(\mathbb{R}^p) \to [0, \infty)$ that satisfy the following assumptions. These are common for many distances between fuzzy sets and are satisfied by each of the metrics in this paper.

(**A1**.) $d(\gamma \cdot A, \gamma \cdot B) = \gamma \cdot d(A, B)$, for all $A, B \in \mathcal{F}_c(\mathbb{R}^p)$ and for all $\gamma \in [0, \infty)$.
(**A2**.) $d(A + W, B + W) = d(A, B)$, for all $A, B, W \in \mathcal{F}_c(\mathbb{R}^p)$.

Proposition 5.3 establishes that, under these assumptions, P3b implies P3a while Theorem 5.4 is for studying the reverse problem. Examples 5.6 and 5.7 show, however, that the equivalence does not hold in general; for instance, for the ρ_1 distance or for d_r, for any $r \in [1, \infty]$.

Proposition 5.3. *Let $\mathcal{X}$ a fuzzy random variable and $D(\cdot; \mathcal{X}) : \mathcal{F}_c(\mathbb{R}^p) \to [0, \infty)$ a function satisfying P3b for a metric that fulfills A1 and A2. Then $D(\cdot; \mathcal{X})$ satisfies P3a.*

Theorem 5.4. *Let $(\mathbb{E}, \|\cdot\|)$ be a strictly convex Banach space, let d be a metric in $\mathcal{F}_c(\mathbb{R}^p)$ fulfilling A1 and A2 and $j : (\mathcal{F}_c(\mathbb{R}^p), d) \to (\mathbb{E}, \|\cdot\|)$ an isometry. Whenever $A, B, C \in \mathcal{F}_c(\mathbb{R}^p)$ are such that*

$$d(A, B) = d(A, C) + d(B, C), \tag{9}$$

the fuzzy set C has the form $(1 - \lambda) \cdot A + \lambda B$ for some $\lambda \in [0, 1]$.

Remark 5.5. In particular, L^p-type norms are strictly convex for $p \in (1, \infty)$. Many commonly used L^p-type metrics for fuzzy sets are thus in the assumptions of Theorem 5.4, like the ρ_r metrics ($r \in (1, \infty)$). Note that if the mapping j preserves the linear operations then A1 and A2 always hold, since every norm has those properties.

Example 5.6. Let $(\{\omega_1, \omega_2\}, \mathcal{P}(\{\omega_1, \omega_2\}), \mathbb{P})$ be a probability space such that $\mathbb{P}(\{\omega_1\}) = 3/4$ and $\mathbb{P}(\{\omega_2\}) = 1/4$ and $\mathcal{X} : \{\omega_1, \omega_2\} \to \mathcal{F}_c(\mathbb{R})$ a fuzzy random variable on that probability space such that $A := \mathcal{X}(\omega_1) = \mathrm{I}_{[1,2]}$ and $B := \mathcal{X}(\omega_2) = \mathrm{I}_{[2,7]}$. Consider $C := \mathrm{I}_{[3,5]}$. Clearly,

$$5 = d_\infty(A, B) = d_\infty(A, C) + d_\infty(B, C) = 2 + 3.$$

Let $D_{FT}(\cdot; \mathcal{X}) : \mathcal{F}_c(\mathbb{R}) \to [0, 1]$ be defined by

$$D_{FT}(U; \mathcal{X}) := \inf_{u \in \mathbb{S}^0, \alpha \in [0,1]} \min\{\mathbb{P}(s_\mathcal{X}(u, \alpha) \le s_A(u, \alpha)), \mathbb{P}(s_\mathcal{X}(u, \alpha) \ge s_A(u, \alpha))\}. \tag{10}$$

We study D_{FT} in the general case in Section 6, and in Theorem 6.6 we will show that it is a semilinear depth function in the sense of Definition 3.1. Thus, D_{FT} satisfies property P3a. It is easy to see that A satisfies

$$D_{FT}(A; \mathcal{X}) = \mathbb{P}(s_\mathcal{X}(-1, 0) \ge s_A(-1, 0)) = 3/4$$

and it is a fuzzy set with maximal depth. Now

$$D_{FT}(B; \mathcal{X}) = \mathbb{P}(s_\mathcal{X}(-1, 0) \le s_B(-1, 0)) = 1/4$$

and

$$D_{FT}(C; \mathcal{X}) = \mathbb{P}(s_\mathcal{X}(-1, 0) \ge s_C(-1, 0)) = 0.$$

Then, we have $A, B, C \in \mathcal{F}_c(\mathbb{R})$ such that

$$d_\infty(A, B) = d_\infty(A, C) + d_\infty(B, C)$$

but $D_{FT}(B;\mathcal{X}) > D_{FT}(C;\mathcal{X})$. Hence P3b does not hold for the metric d_∞. In fact

$$d_r(A,B) = d_r(A,C) + d_r(B,C)$$

for all $r \in [1,\infty)$, so the same counterexample is valid for any $r \in [1,+\infty]$.

Example 5.7. Given the probability space $(\{\omega_1,\omega_2\}, \mathcal{P}(\{\omega_1,\omega_2\}), \mathbb{P})$ with $\mathbb{P}(\omega_1) = 3/4$ and $\mathbb{P}(\omega_2) = 1/4$, let $\mathcal{X}: \{\omega_1,\omega_2\} \to \mathcal{F}_c(\mathbb{R})$ be a fuzzy random variable such that $\mathcal{X}(\omega_1) = T(0,.5,1)$ and $\mathcal{X}(\omega_2) = T(2,2.5,3)$. Let us define the function $D_1(\cdot;\mathcal{X}) : \mathcal{F}_c(\mathbb{R}) \to [0,1]$ as

$$D_1(U;\mathcal{X}) := (1 + \mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_U)\|_1])^{-1},$$

where $\|\cdot\|_1$ is the L^1-norm $\|f\|_1 = \int_{\mathbb{S}^0}\int_{[0,1]} |f(u,\alpha)| d\alpha du$. The value of $D_1(U;\mathcal{X})$ is meant as 0 if the expectation in the denominator is infinite.

Let us show that this function satisfies property P3a but fails P3b.

Observe that $\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_U)\|_1]$ is convex as a function of $U \in \mathcal{F}_c(\mathbb{R})$, that is, for all $U, V \in \mathcal{F}_c(\mathbb{R})$ and $\lambda \in [0,1]$,

$$\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_{\lambda U+(1-\lambda)V})\|_1] \le \lambda \mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_U)\|_1] + (1-\lambda)\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_V)\|_1],$$

because $\|\cdot\|_1$ is a norm and the spread and the expectation are linear. Then, if $A \in \mathcal{F}_c(\mathbb{R})$ maximizes $D_1(\cdot;\mathcal{X})$ we have $\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_A)\|_1] \le \mathrm{E}[\|\mathrm{spr}(s_B) - \mathrm{spr}(s_U)\|_1]$ and there follows, for any B,

$$\begin{aligned}
&D_1((1-\lambda)A + \lambda B;\mathcal{X}) \ge \\
&(1 + (1-\lambda)\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_A)\|_1] + \lambda \mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_B)\|_1])^{-1} \ge \\
&(1 + (1-\lambda)\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_B)\|_1] + \lambda \mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_B)\|_1])^{-1} = D_1(B;\mathcal{X}),
\end{aligned}$$

which establishes P3a.

Now let A, B and C be the fuzzy sets $A := T(0,.5,1) = \mathcal{X}(\omega_1)$, $B := T(2,2.5,3) = \mathcal{X}(\omega_2)$, $C := \mathrm{I}_{\{2\}}$.

First of all, it is clear that $A_\alpha = [\alpha/2, 1-\alpha/2]$ and $B_\alpha = [2+\alpha/2, 3-\alpha/2]$. Then we know that $s_A(-1,\alpha) = -\alpha/2$, $s_A(1,\alpha) = 1-\alpha/2$, $s_B(-1,\alpha) = -2-\alpha/2$ and $s_B(1,\alpha) = 3-\alpha/2$ for all $\alpha \in [0,1]$. It is easy to check $\rho_1(A,B) = 2$, $\rho_1(A,C) = 3/2$, $\rho_1(B,C) = 1/2$ taking into account $s_C(-1,\alpha) = -2$ and $s_C(1,\alpha) = 2$ for all $\alpha \in [0,1]$. Then A, B, C satisfy

$$\rho_1(A,B) = \rho_1(A,C) + \rho_1(B,C).$$

Further, $\mathcal{X}$, A and B have been chosen so that

$$\mathrm{spr}(s_\mathcal{X}) = \mathrm{spr}(s_A) = \mathrm{spr}(s_B),$$

whence

$$\mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_A)\|_1] = \mathrm{E}[\|\mathrm{spr}(s_\mathcal{X}) - \mathrm{spr}(s_B)\|_1] = 0$$

and, accordingly,

$$D_1(A;\mathcal{X}) = D_1(B;\mathcal{X}) = 1.$$

Since, by definition, D_1 is bounded above by 1, both A and B maximize the depth value. If P3b held true, since C is metrically between A and B we would have $D_1(C;\mathcal{X}) = 1$ as well. But $\mathrm{spr}(C) = 0$, whence

$$D_1(C;\mathcal{X}) = (1 + E[\|\mathrm{spr}(s_\mathcal{X})\|_1])^{-1} = \frac{4}{5} < 1$$

and P3b does not hold.

Proposition 5.8 shows that P4b implies P4a under mild conditions. Example 5.9 shows that the converse is not satisfied in general.

Proposition 5.8. *Let $\mathcal{X}$ be a fuzzy random variable and $D(\cdot;\mathcal{X}):\mathcal{F}_c(\mathbb{R}^p)\to[0,\infty)$ a function for which P4b holds with respect to a metric that fulfills A1 and A2. Then $D(\cdot;\mathcal{X})$ satisfies P4a.*

Example 5.9. Let us consider the function $D_1(U;\mathcal{X})=(1+\mathrm{E}[\rho_1(\mathcal{X},U)])^{-1}$ for any fuzzy set U and fuzzy random variable $\mathcal{X}$, where (like in Example 5.7) the value is meant as 0 if the denominator is infinite.

Let $A=\mathrm{I}_{\{0\}}\in\mathcal{F}_c(\mathbb{R})$ and let $\mathcal{X}$ the fuzzy random variable such that $\mathcal{X}(\omega_1)=A$, associated with the probabilistic space $(\{\omega_1\},\mathcal{P}(\{\omega_1\}),\mathbb{P})$. Since the only value of $\mathcal{X}$ is A, clearly $D_1(A;\mathcal{X})=1$ which maximizes $D_1(\cdot;\mathcal{X})$ and is in fact the unique maximizer since $\rho_1(B,A)>0$ for any other B.

Let us show that P4a is satisfied by $D_1(\cdot,\mathcal{X})$ but P4b fails for the metric d_r, for any $r\in(1,\infty)$. Accordingly, it fails as well for the corresponding ρ_r-metrics, since they are equivalent to the d_r.

For any $B\in\mathcal{F}_c(\mathbb{R})$ such that $B\neq\mathrm{I}_{\{0\}}$, and any $\lambda>0$, taking into account that ρ_1 has properties A1 and A2 we obtain

$$D_1(A+\lambda B;\mathcal{X})=(1+\rho_1(A+\lambda B,A))^{-1}=(1+\rho_1(\lambda B,\mathrm{I}_{\{0\}}))^{-1}=(1+\lambda\rho_1(B,\mathrm{I}_{\{0\}}))^{-1}$$

which goes to 0 as $\lambda\to\infty$. Therefore P4a holds.

To prove that P4b fails, consider the sequence $\{A_n\}_n\in\mathcal{F}_c(\mathbb{R})$ given by

$$A_n(x)=\begin{cases}1, & x=0\\ n^{-1}, & x\in(0,n]\\ 0, & \text{otherwise.}\end{cases}$$

Since $(A_n)_\alpha=[0,n]$ if $\alpha\in[0,n^{-1}]$ and $\{0\}$ otherwise,

$$\rho_1(A_n,A)=\frac{1}{2}\cdot 0+\frac{1}{2}\cdot n^{-1}\cdot n=\frac{1}{2}.$$

Accordingly, for each $n\in\mathbb{N}$,

$$D_1(A_n;\mathcal{X})=(1+\rho_1(A_n,A))^{-1}=\frac{2}{3}>0.$$

Thus if P4b held for d_r, that would imply $d_r(A_n,A)\not\to\infty$. But

$$d_r(A_n,A)=(\int_0^1 d_H((A_n)_\alpha,A_\alpha)^r\mathrm{d}\alpha)^{1/r}=(n^{-1}\cdot n^r)^{1/r}=n^{1-1/r}\to\infty$$

since $r>1$. We conclude that property P4b cannot hold.

Remark 5.10. As a consequence of Propositions 5.3 and 5.8, if a function satisfies the properties of geometric depth function for a metric fulfilling A1 and A2, then it is also a semilinear depth function.

6. The Tukey depth

Given a probability space $(\Omega,\mathcal{A},\mathbb{P})$, the multivariate *Tukey depth* [44] (or *halfspace depth*) of a point $x\in\mathbb{R}^p$ with respect to a random vector X is defined as

$$D_T(x;X):=\inf_{u\in\mathbb{S}^{p-1}}\{\mathbb{P}[\omega\in\Omega:X(\omega)\in S^-_{u,x}]\},$$

where $S^-_{u,x}:=\{y\in\mathbb{R}^p:\langle y-x,u\rangle\le 0\}$ is a closed halfspace in $\mathbb{R}^p$. We propose in Definition 6.1 a generalization for fuzzy sets which makes use of the support function of the fuzzy set. As shown in Proposition 6.3, it coincides with the Tukey depth when restricting $\mathcal{F}_c(\mathbb{R}^p)$ to $\mathcal{R}^p$.

Definition 6.1. Let $\mathcal{H}\subseteq L^0[\mathcal{F}_c(\mathbb{R}^p)]$ and $\mathcal{J}\subseteq\mathcal{F}_c(\mathbb{R}^p)$. The *Tukey depth* of a fuzzy set with respect to a fuzzy random variable is given by the function $D_{FT}(\cdot;\cdot):\mathcal{J}\times\mathcal{H}\to[0,1]$ with

$$D_{FT}(U;\mathcal{X}) := \inf_{u\in\mathbb{S}^{p-1},\alpha\in[0,1]} \min\left(\mathbb{P}[\omega\in\Omega:\mathcal{X}(\omega)\in S^-_{u,\alpha}],\mathbb{P}[\omega\in\Omega:\mathcal{X}(\omega)\in S^+_{u,\alpha}]\right),$$

where

$$S^-_{u,\alpha} := \{V\in\mathcal{F}_c(\mathbb{R}^p): s_V(u,\alpha)-s_U(u,\alpha)\le 0\},$$
$$S^+_{u,\alpha} := \{V\in\mathcal{F}_c(\mathbb{R}^p): s_V(u,\alpha)-s_U(u,\alpha)\ge 0\}.$$

Although it is not made explicit, the sets $S^-_{u,\alpha}$ and $S^+_{u,\alpha}$ depend on U. The same will happen with the sets $S^-_{u,\alpha,t}$ and $S^+_{u,\alpha,t}$ in Theorem 6.2 below.

Our definition of Tukey depth is based on using the support functionals $A\mapsto s_A(u,\alpha)$ as a replacement for the products $x\mapsto\langle x,u\rangle$ in $\mathbb{R}^p$. The rationale for doing so is multiple. First, the support functionals reduce to the inner product functionals in $\mathcal{R}^p$, which will ensure this is an actual generalization. Second, each $A\in\mathcal{F}_c(\mathbb{R}^p)$ is identified by the value of all the support functionals on it, like each $x\in\mathbb{R}^p$ is identified by the value of all the products $\langle x,u\rangle$. Third, they are additive and positively homogeneous so they work well with the operations on fuzzy sets. The main difference is that $s_{-A}(u,\alpha)\neq -s_A(u,\alpha)$ in general, whence we need to consider both types of generalized halfspaces $S^-_{u,\alpha}$ and $S^+_{u,\alpha}$. In $\mathbb{R}^p$ one would just have $S^+_{u,x}=S^-_{-u,x}$, which makes it unnecessary to include both halfspaces in the definition.

A known feature of the multivariate Tukey depth of a point $x\in\mathbb{R}^p$ is that it is equivalent to consider all halfspaces containing x or only those which contain x in their boundary. As we see in the following theorem, the situation is analogous for the fuzzy Tukey depth, because $S^-_{u,\alpha}=S^-_{u,\alpha,s_U(u,\alpha)}\subseteq S^-_{u,\alpha,t}$ for any $t\ge s_U(u,\alpha)$, where the second type of halfspace is defined below. The situation is analogous for the generalized halfspaces with the positive sign. Also, the infimum and the minimum can be interchanged.

Theorem 6.2. *It is equivalent to define, for any $U\in\mathcal{F}_c(\mathbb{R}^p)$ and $\mathcal{X}\in L^0[\mathcal{F}_c(\mathbb{R}^p)]$, the Tukey depth of U with respect to $\mathcal{X}$ as*

$$D_{FT}(U;\mathcal{X}) = \min\Big(\inf_{(u,\alpha,t)\in I^-_U}\mathbb{P}[\omega\in\Omega:\mathcal{X}(\omega)\in S^-_{u,\alpha,t}],\ \inf_{(u,\alpha,t)\in I^+_U}\mathbb{P}[\omega\in\Omega:\mathcal{X}(\omega)\in S^+_{u,\alpha,t}]\Big),$$

where

$$I^-_U := \{(u,\alpha,t)\in\mathbb{S}^{p-1}\times[0,1]\times\mathbb{R}: t\ge s_U(u,\alpha)\},$$
$$I^+_U := \{(u,\alpha,t)\in\mathbb{S}^{p-1}\times[0,1]\times\mathbb{R}: t\le s_U(u,\alpha)\},$$
$$S^-_{u,\alpha,t} := \{V\in\mathcal{F}_c(\mathbb{R}^p): \max(s_V(u,\alpha),s_U(u,\alpha))\le t\},$$
$$S^+_{u,\alpha,t} := \{V\in\mathcal{F}_c(\mathbb{R}^p): \max(s_V(u,\alpha),s_U(u,\alpha))\ge t\}.$$

The fuzzy version of the Tukey depth is actually a generalization in the sense that it coincides with the multivariate Tukey depth when applied to the indicator function of a crisp random vector.

Proposition 6.3. *Let $\mathcal{X}$ be a fuzzy random variable taking on values in $\mathcal{R}^p$ and let X be the associated random vector such that $\mathcal{X}=I_{\{X\}}$. Then $D_{FT}(I_{\{x\}};\mathcal{X})=D_T(x;X)$ for all $x\in\mathbb{R}^p$.*

It is well known that many multivariate depth functions provide multivariate generalizations of the univariate median. For instance, the Tukey depth is maximized at the halfspace median, which coincides with the ordinary median for univariate random variables. Sinova et al. provide in [39] a definition of median for fuzzy random variables, which we reproduce in Section 2.3. The next theorem proves that, under mild conditions, the fuzzy Tukey depth is maximized at that median.

Theorem 6.4. *Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probabilistic space associated with the fuzzy random variable $\mathcal{X}:\Omega\to\mathcal{F}_c(\mathbb{R})$ such that* Med(inf $\mathcal{X}_\alpha$) *and* Med(sup $\mathcal{X}_\alpha$) *are unique for each $\alpha\in[0,1]$. Then,*

$$\tilde{M}e(\mathcal{X}) = \arg\sup_{U\in\mathcal{F}_c(\mathbb{R})} D_{FT}(U;\mathcal{X}),$$

where $\tilde{M}e(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$ satisfies

$$(\tilde{M}e(\mathcal{X}))_\alpha = [\mathrm{Med}(\inf \mathcal{X}_\alpha), \mathrm{Med}(\sup \mathcal{X}_\alpha)]$$

for every $\alpha \in [0, 1]$.

If $\mathrm{Med}(\inf \mathcal{X}_\alpha)$ and/or $\mathrm{Med}(\sup \mathcal{X}_\alpha)$ are not unique, we take the usual convention of defining them as the midpoint of the set of medians. The next example shows that without the uniqueness assumption, the result in Theorem 6.4 is not satisfied.

Example 6.5. Let $(\{\omega_1, \omega_2\}, \mathcal{P}(\{\omega_1, \omega_2\}), \mathbb{P})$ be a probability space with $\mathbb{P}(\omega_1) = 1/4$ and $\mathbb{P}(\omega_2) = 3/4$ and $\mathcal{X} : \{\omega_1, \omega_2\} \to \mathcal{F}_c(\mathbb{R})$ a fuzzy random variable over it such that $\mathcal{X}(\omega_1) = \mathrm{I}_{[1,2]}$ and $\mathcal{X}(\omega_2) = \mathrm{I}_{[3,4]}$.

Now $\tilde{M}e(\mathcal{X}) = \mathrm{I}_{[2,3]}$ as $\mathrm{Med}(\inf \mathcal{X}_\alpha)$ is the midpoint of the interval $[1, 3]$ and $\mathrm{Med}(\sup \mathcal{X}_\alpha)$ of the interval $[2, 4]$. One easily checks

$$D_{FT}(\tilde{M}e(\mathcal{X}); \mathcal{X}) = 1/4 < 3/4 = D_{FT}(\mathcal{X}(\omega_2); \mathcal{X}).$$

In Theorem 6.6 we show that D_{FT} is a depth function in the sense of both Definitions 3.1 and 3.2.

Theorem 6.6. *When symmetry is understood in the sense of F-symmetry, D_{FT} is a semilinear depth and a geometric depth with respect to the metric ρ_r for any $r \in (1, \infty)$. Further, it satisfies property P4b for the metric d_∞ and property P4a*.*

Proposition 6.7. *D_{FT} is not a geometric depth function with respect to d_r for any $r \in [1, \infty]$.*

Remark 6.8. The difference between ρ_r and d_r in this regard can be tracked down to Theorem 5.4, which applies to ρ_r while d_r admits an embedding into a Banach space but not a strictly convex one.

As shown in Section 4, F-symmetry and (mid, spr)-symmetry are not equivalent concepts. We provide an example below where a fuzzy random variable is (mid, spr)-symmetric with respect to a fuzzy set at which the maximum Tukey depth is not reached.

Example 6.9. Consider the fuzzy random variable $\mathcal{X}$ in Example 4.5, which takes on four values with equal probabilities. We showed there that $\mathcal{X}$ is (mid, spr)-symmetric with respect to $A = \mathrm{I}_{[1/2,9/2]}$.

One can straightforwardly check that, for any $\alpha \in [0, 1]$,

$$\begin{aligned}
P(s_{\mathcal{X}}(1, \alpha) \le s_A(1, \alpha)) &= .5 \\
P(s_{\mathcal{X}}(1, \alpha) \ge s_A(1, \alpha)) &= .5 \\
P(s_{\mathcal{X}}(-1, \alpha) \le s_A(-1, \alpha)) &= .75 \\
P(s_{\mathcal{X}}(-1, \alpha) \ge s_A(-1, \alpha)) &= .25
\end{aligned}$$

whence there follows $D_{FT}(A; \mathcal{X}) = .25$.

However, taking $B = \mathrm{I}_{[1,9/2]}$ we have

$$\begin{aligned}
P(s_{\mathcal{X}}(1, \alpha) \le s_B(1, \alpha)) &= .5 \\
P(s_{\mathcal{X}}(1, \alpha) \ge s_B(1, \alpha)) &= .5 \\
P(s_{\mathcal{X}}(-1, \alpha) \le s_B(-1, \alpha)) &= .75 \\
P(s_{\mathcal{X}}(-1, \alpha) \ge s_B(-1, \alpha)) &= .75
\end{aligned}$$

and therefore $D_{FT}(B; \mathcal{X}) = .5$. That shows that A is not maximally deep and thus D_{FT} violates P2 when symmetry is understood in the sense of (mid, spr)-symmetry.

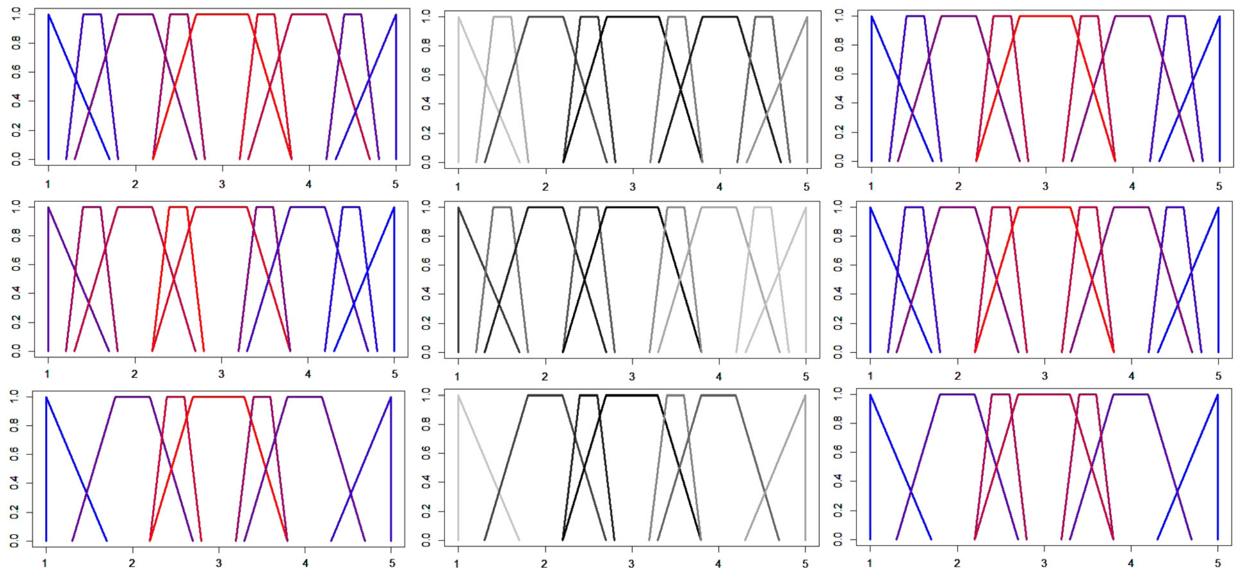

Fig. 2. Display of the fuzzy sets in the three species in the *Trees* dataset: birch (top row), sessile oak (middle row) and rowan (bottom row). In the first column, color is assigned based on the Tukey depth of each fuzzy set in the empirical distribution of the corresponding species' sample. Colors range from red (high depth) to blue (low depth) with shades of purple for intermediate levels of depth. In the second column, the weights of the elements in each of the three samples (i.e., their relative frequencies in the sample) are codified in a grey scale, from light grey (light weight) to black (heavy weight). The third column applies the procedure in the first column but discarding the weights of the fuzzy sets in the samples (i.e., this is what would appear if all quality types were distributed uniformly in the population of trees).

7. Real data illustration

To illustrate the performance of the Tukey depth of fuzzy sets on a real dataset we use the *Trees* dataset in the `SAFD` (Statistical Analysis of Fuzzy Data) R package [7]. It comes from a study carried out by the INDUROT forest institute of the University of Oviedo (Spain) on a reforestation project in the Northern Spanish region of Asturias. It contains random samples of three species of trees, namely birch (*Betula celtiberica*), sessile oak (*Quercus petraea*) and rowan (*Sorbus aucuparia*). There are $n_1 = 133$ birches, $n_2 = 109$ sessile oaks and $n_3 = 37$ rowans, for a total sample size of $n = 279$. An important variable considered in the study is the *quality* of trees, which is not measured as a real number but rather on the basis of expert subjective opinions on the leaf structure, height-diameter ratio, and other factors. Each tree's *quality* is represented by a trapezoidal fuzzy set, see Fig. 2. The x-axis represents the tree quality on a scale from 0 to 5, with 0 meaning a total absence of quality and 5 a perfect quality. The y-axis represents the trapezoidal membership function, with the 0-level being the interval where the experts are absolutely sure that the quality is contained and the 1-level being the interval where the experts think that the quality is contained.

For each species, the Tukey depth of each fuzzy set in the sample is computed. First, we discuss the adequacy of the Tukey depth for this dataset. The data are trapezoidal fuzzy sets, which are commonly used in practice for their simplicity. The defining feature of a trapezoidal fuzzy set is that it is determined by four values, the maximum and minimum values of its 0-level (support) and 1-level (kernel). Moreover, the operations between trapezoidal fuzzy sets correspond to the addition and product by a scalar of those four values. From an algebraic point of view, their operations are isomorphic to those of 4-dimensional vectors (elements of $\mathbb{R}^4$). Hence one can reason that using a semilinear depth is appropriate, since its requirements describe the notion of statistical depth making use only of the algebraic structure.

Before discussing the possible adequacy of geometric depth, observe that one could identify each datum with a vector in $\mathbb{R}^4$ and apply a multivariate depth function to assign each trapezoidal fuzzy set a depth value. But the defining properties of multivariate depth are not the best fit for these data. To understand this without an elaborate discussion, just compare properties P1 and M1. Property M1 requires invariance with respect to all affine transformations in $\mathbb{R}^4$, many of which will transform the vector into another vector which does not correspond to any fuzzy set at all. However, Property P1 deals with affine transformations of the underlying space $\mathbb{R}$, i.e., each of the four vector components identifying the fuzzy datum must be affected by the same transformation in a coherent way. This is why we cannot simply reduce fuzzy data depth to multivariate data depth.

Concerning geometric depth, we should ask whether some metric is particularly appropriate for these data. Since they have a simple form which only depends on the 0-level and the 1-level, it is hard to see practical reasons to discard any metric *a priori*. They are not a special type of fuzzy data with specificities which should be taken into account with an *ad hoc* metric. One may reasonably consider that, in order to assess the distance between two quality values, it may be disadvantageous to use metrics like d_∞: since it is a supremum-type metric, it can produce the same distance value in situations where visually it looks like one distance should definitely be larger than the other. From a theoretical perspective, a metric such as ρ_2 which embeds the space of fuzzy data into a Hilbert space has the best geometric properties (closest to those of the Euclidean space). Moreover, for that very reason ρ_2 does not have the above mentioned shortcoming of d_∞. This analysis suggests that a depth function which were both a semilinear depth and a geometric depth with respect to the metric ρ_2 (or another metric embedding fuzzy data into a Hilbert space) would be the optimal choice. As shown in Theorem 6.6, Tukey depth satisfies both criteria.

Now let us discuss the computation of the depth values. Let $\{S_i^j\}_{i=1}^{n_j}$ be the sample associated to species j, where $j = 1$ for birch, $j = 2$ for sessile oak, and $j = 3$ for rowan. For each $j = 1, 2, 3$, let us also denote by $\mathcal{X}_j$ a fuzzy random variable corresponding to the empirical distribution associated to $\{S_i^j\}_{i=1}^{n_j}$ (namely, each fuzzy value has the probability given by its relative frequency in the sample). Thus, for each $i = 1, \ldots, n_j$ and $j = 1, 2, 3$, our objective is to compute $D_{FT}(S_i^j, \mathcal{X}_j)$, providing an order of the fuzzy sets in $\{S_i^j\}_{i=1}^{n_j}$ by their depth values. Since the fuzzy sets are trapezoidal, each S_i^j, $i = 1, \ldots, n_j$ and $j = 1, 2, 3$, is determined by the following four values of its support function: $s_{S_i^j}(1, 0)$, $s_{S_i^j}(-1, 0)$, $s_{S_i^j}(1, 1)$ and $s_{S_i^j}(-1, 1)$. Taking this into account in the expression of D_{FT} in Definition 6.1, we obtain

$$D_{FT}(S_i^j, \mathcal{X}_j) = \min_{u \in \{-1,1\}, \alpha \in \{0,1\}} D_T(s_{S_i^j}(u, \alpha); s_{\mathcal{X}_j}(u, \alpha)), \tag{11}$$

for each $i = 1, \ldots, n_j$ and $j = 1, 2, 3$, where D_T denotes the Tukey depth for real-valued data. This provides a simplified formula which avoids considering all $\alpha \in [0, 1]$ and only uses $\alpha \in \{0, 1\}$; moreover, one can use existing algorithms for the calculation of the univariate Tukey depth.

Making use of Equation (11), the first column of Fig. 2 displays the center-outward ordered sample based on the Tukey depth for each species. Blue color represents the lowest depth and red the highest, with purple shades in between. Thus the most reddish fuzzy value in each sample is the analog of the median (the innermost value). The most blueish fuzzy values are the analogs of the maximum and the minimum. Since the outward directions in the real line are obvious, the interpretation of the plots in the example is very clear.

The fuzzy sets representing quality make up a fuzzy scale (with nine values) which was agreed upon by the researchers and the experts. Hence, each fuzzy quality value appears with a different multiplicity in each sample as shown in the second column of Fig. 2. The third column of Fig. 2 is suitable for comparison with the color-coded depth values that appear in the first column, as it represents what the depth values would be if each quality value appeared in the sample equally many times. In that case, the colors in the plot are symmetric and the shift from red to blue happens at a constant rate in each direction. Also note that, in the third species, only 7 values appear (the rowan sample is quite smaller and not all possible quality values were used).

In the actual depth plots (first column of Fig. 2) that color symmetry is broken. This is clearest in the middle row (sessile oak sample). The deepest value (corresponding to the median) is below the middle of the quality range, and lower quality values consistently have higher depth than their symmetric values. Even the lowest quality value in the fuzzy scale is clearly non-blue (corresponding to it having a relatively high relative frequency in the second column). That shows visually that the quality of sessile oaks is lower than that of the other two species. In contrast, the center of the quality distribution (the area with the highest depth values) for birches appears at higher (center to relatively high) values.

The rowan plot looks quite symmetric but it illustrates that the Tukey depth of a fuzzy set tries to measure how much of the distribution is 'from that fuzzy set outwards'. The depth of a datum in the sample, then, depends critically on the position of the other data besides its own frequency. That is shown here as the second and sixth values (counting from the left) have approximately the same depth, while the grey-scale plot shows that they have rather different frequencies. What the plot shows is that approximately the same amount of data are at the second quality value or worse, than at the sixth quality value or better. The plot also shows that there are more highest-quality than lowest-quality rowans; this is consistent with the frequencies in the rowan sample's grey-scale plot.

8. Proofs

This section includes the proofs of our results, as well as some necessary lemmas.

Proof of Lemma 4.6. We prove the first part, the second is analogous. Since A is F-symmetric with respect to $\mathcal{X}$, by Definition 4.1, we have

$$s_{\mathcal{X}}(u,\alpha) - s_A(u,\alpha) =^d s_A(u,\alpha) - s_{\mathcal{X}}(u,\alpha)$$

for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$. For every $x \in \mathbb{R}$,

$$\mathbb{P}\left(s_{\mathcal{X}}(u,\alpha) - s_A(u,\alpha) \le x\right) = \mathbb{P}\left(s_A(u,\alpha) - s_{\mathcal{X}}(u,\alpha) \le x\right).$$

Fixing $x = 0$ we have

$$\mathbb{P}\left(s_X(u,\alpha) \le s_A(u,\alpha)\right) = \mathbb{P}\left(s_{\mathcal{X}}(u,\alpha) \ge s_A(u,\alpha)\right).$$

Then, $s_A(u,\alpha)$ is a median of the random variable $s_{\mathcal{X}}(u,\alpha)$ for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$. □

Proof of Proposition 5.1. Let $\mathcal{H}$ be the set of all fuzzy random variables taking on values in $\mathcal{R}^p$ and let $D : \mathcal{R}^p \times \mathcal{H}$ be a depth function as in Definition 3.1.

Proof for P1. Set $A = \mathrm{I}_{\{a\}}$, $B = \mathrm{I}_{\{b\}} \in \mathcal{R}^p$ and let $M \in \mathcal{M}_{p\times p}(\mathbb{R})$ a non-singular matrix. Then

$$(M \cdot A)(t) = \mathrm{I}_{\{a\}}(M^{-1} \cdot t) = \mathrm{I}_{\{M\cdot a\}}(t).$$

Clearly, $M \cdot A + B = \mathrm{I}_{\{M\cdot a+b\}}$. Identifying $\mathrm{I}_{\{x\}} \in \mathcal{R}^p$ with $x \in \mathbb{R}^p$ proves M1. The converse is shown analogously.

Proof for P2. Let $A = \mathrm{I}_{\{a\}} \in \mathcal{R}^p$ and $\mathcal{X} : \Omega \to \mathcal{R}^p$ a fuzzy random variable with $\mathcal{X}(\omega) = \mathrm{I}_{\{X(\omega)\}} \in \mathcal{R}^p$ such that $\mathcal{X}$ is F-symmetric with respect to A. By Definition 4.1,

$$s_A(u,\alpha) - s_{\mathcal{X}}(u,\alpha) =^d s_{\mathcal{X}}(u,\alpha) - s_A(u,\alpha)$$

for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$. It is clear that $s_A(u,\alpha) = \langle u, a\rangle$. Writing $a = (a_1,\ldots,a_p)$ and $X(\omega) = (X_1(\omega),\ldots,X_p(\omega))$, it is equivalent to say

$$\sum_{i=1}^{p} u_i \cdot a_i - \sum_{i=1}^{p} u_i \cdot X_i(\omega) =^d \sum_{i=1}^{p} u_i \cdot X_i(\omega) - \sum_{i=1}^{p} u_i \cdot a_i, \tag{12}$$

for all $u = (u_1,\ldots,u_p) \in \mathbb{S}^{p-1}$. If we fix $t = (t_1,\ldots,t_p) \in \mathbb{R}^p$, there exists some $u_0 = (u_{0,1},\ldots,u_{0,p}) \in \mathbb{S}^{p-1}$ such that $u_0 = (1/\|t\|) \cdot t$, which is equivalent to $t = \|t\| \cdot u_0$. By Equation (12)

$$\sum_{i=1}^{p} \|t\| \cdot u_{0,i} \cdot a_i - \sum_{i=1}^{p} \|t\| \cdot u_{0,i} \cdot X_i(\omega) =^d \sum_{i=1}^{p} \|t\| \cdot u_{0,i} \cdot X_i(\omega) - \sum_{i=1}^{p} \|t\| \cdot u_{0,i} \cdot a_i$$

Equivalently, it is easy to see that,

$$\sum_{i=1}^{p} t_i \cdot a_i - \sum_{i=1}^{p} t_i \cdot X_i(\omega) =^d \sum_{i=1}^{p} t_i \cdot X_i(\omega) - \sum_{i=1}^{p} t_i \cdot a_i, \tag{13}$$

for all $t = (t_1,\ldots,t_p) \in \mathbb{R}^p$. Note that by the Cramér–Wold theorem [4, Theorem 29.4], two random vectors $Y = (Y_1,\cdots,Y_p)$ and $Z = (Z_1,\cdots,Z_p)$ are identically distributed if and only if the random variables $\sum_{i=1}^k t_i Y_i$ and $\sum_{i=1}^k t_i Z_i$ are identically distributed for each $(t_1,\cdots,t_k) \in \mathbb{R}^p$. Thus, Equation (13) is equivalent to $a - X =^d X - a$, i.e. X is centrally symmetric with respect to a.

Since the Cramér–Wold theorem is an equivalence, the converse is proved the same way.

Proof for P3a. It results from the equality $(1-\lambda)\cdot A+\lambda\cdot B=\mathrm{I}_{\{(1-\lambda)\cdot a+\lambda\cdot b\}}$, when taking $A=\mathrm{I}_{\{a\}}\in\mathcal{R}^p$, $B=\mathrm{I}_{\{b\}}\in\mathcal{R}^p$ and $\lambda\in[0,1]$; with $a,b\in\mathbb{R}^p$, and the possibility of taking $x=[a-(1-\lambda)\theta]/\lambda$.

Proof for P3b. Let $A=\mathrm{I}_{\{a\}}$ and $B=\mathrm{I}_{\{b\}}$. Thus,

$$d_\infty(A,B)=d_H(\{x\},\{y\})=\|x-y\|.$$

If we take $C=\mathrm{I}_{\{c\}}\in\mathcal{R}^p$ such that $d_\infty(A,B)=d_\infty(A,C)+d_\infty(B,C)$. It is equivalent to say $\|a-b\|=\|a-c\|+\|c-b\|$. As $(\mathbb{R}^p,\|\cdot\|)$ is a strictly convex normed space, then $c=(1-\lambda)\cdot a+\lambda\cdot b$ for some $\lambda\in[0,1]$. Thus, $C=(1-\lambda)\cdot A+\lambda\cdot B$. As we have said $d_H(\{a\},\{b\})=\|a-b\|$, then the result is valid for any d_r distance with $r\in[1,\infty)$.

For L^2-type fuzzy distances (such as ρ_2) the two properties are also equivalent, by taking into account Theorem 5.4 and the fact that $(\mathbb{R}^p,\|\cdot\|)$, being a Hilbert space, is strictly convex.

Proof for P4a. If M4 is satisfied, for all sequences $\{x_n\}_n\subseteq\mathbb{R}^p$ such that $\lim_n\|x_n\|=\infty$ we have $\lim_n D(x_n;X)=0$. Let $a,b\in\mathbb{R}^p$ be such that a maximizes $D(\cdot;X)$. Then, $\lim_n D(a+nb;X)=0$. Then, $\lim_n(\mathrm{I}_{a+nb};\mathcal{X})=0$ and P4a holds.

Proof for P4b. **Case of d_r distance.** Let $\mathrm{I}_{\{a\}}\in\mathcal{R}^p$ be a fuzzy set which maximizes $D(\cdot;\mathcal{X})$. Let $\{\mathrm{I}_{\{a_n\}}\}_n$ be a sequence of elements of $\mathcal{R}^p$ such that $\lim_n d_r(\mathrm{I}_{\{a\}},\mathrm{I}_{\{a_n\}})=\infty$. By definition of d_r distance, it is clear that $d_r(\mathrm{I}_{\{x\}},\mathrm{I}_{\{y\}})=\|x-y\|$, where $\|\cdot\|$ is the Euclidean norm. Thus, property P4b can be viewed as

$$\lim_{\|x\|\to\infty} D(\mathrm{I}_{\{x\}};\mathcal{X})=0,$$

which is equivalent to M4.

Case of ρ_r distance. Since ρ_r and d_r are equivalent metrics and property P4b is stated in terms of convergence, it holds for ρ_r if and only if it holds for d_r. □

Proof of Proposition 5.2. *Proof of part (a).* P4a* restricted to $\mathcal{R}^p$ reads:

> $\lim_{n\to\infty} D(\mathrm{I}_{\{a\}}+\mathrm{I}_{\{b_n\}};\mathcal{X})=0$ for a sequence $\{b_n\}_n$ of elements of $\mathbb{R}^p$ such that there exists $i\in\{1,2,\cdots,p\}$ such that $\lim_{n\to\infty}|b_{n,i}|=\infty$, where $b_{n,i}$ is the i-th component of the vector b_n for all $n\in\mathbb{N}$.

In order to prove the equivalence of the two properties, we show that the two following sets of sequences of $\mathbb{R}^p$ are equal:

$$C_1:=\{\{x_n\}_n:\lim_n\|x_n\|=\infty\}$$
$$C_2:=\{\{a+b_n\}_n:\text{ there exists } i\in\{1,2,\ldots,p\}\text{ such that }\lim_n|b_{n,i}|=\infty\}.$$

Let $(x_n)_n\in C_1$. It is clear that there exists some component of vector x_n going to infinity, that is, there exists $i\in\{1,2,\ldots,p\}$ such that $\lim_n x_{n,i}=\infty$, because $\lim_n\|x_n\|=\infty$. Then, if we take $b_n=x_n-a$, it is clear that $\{x_n\}_n\in C_2$.

Let $\{a+b_n\}_n\in C_2$. By the triangle inequality,

$$\|a+b_n\|\geq|\|b_n\|-\|a\||$$

for all $n\in\mathbb{N}$. Without loss of generality, we suppose that $\lim_n|b_{n,1}|=\infty$. Thus

$$\|b_n\|=\sqrt{\sum_{i=1}^p b_{n,i}^2}\geq|b_{n,1}|,$$

for all $n\in\mathbb{N}$. If we take limits, we have $\lim_{n\to\infty}\|b_n\|\geq\lim_{n\to\infty}|b_{n,1}|=\infty$. Then, $\lim_{n\to\infty}\|b_n\|=\infty$ and $\lim_{n\to\infty}\|a+b_n\|=\infty$. We have that $\{a+b_n\}_n\in C_1$.

Proof of part (b). By part (a), it is enough to show M4 ⇒ P4a. Let $A=\mathrm{I}_{\{a\}}\in\mathcal{R}^p$ maximize depth in $\mathcal{R}^p$, and let $B=\mathrm{I}_{\{b\}}\in\mathcal{R}^p$. Then

$$D(A+\lambda B;\mathcal{X})=D(\mathrm{I}_{\{a+\lambda b\}};\mathcal{X})\to 0$$

as $\lambda \to \infty$ by the fact that $\|a+\lambda b\| \to \infty$ and property M4.

Proof of part (c). Let A maximize depth and B_n be such that $|s_{B_n}(e_i,\alpha)| \to \infty$ for some $\alpha \in [0,1]$ and $i = 1, \ldots, p$. Since

$$d_\infty(A+B_n, A) = d_\infty(B_n, \mathrm{I}_{\{0\}}) = \sup_{u\in S^{d-1}, \alpha\in[0,1]} |s_{B_n}(u,\alpha)| \geq |s_{B_n}(e_i,\alpha)| \to \infty,$$

by property P4b for the d_∞-metric we indeed have

$$D(A+B_n; \mathcal{X}) \to 0$$

as $n \to \infty$. □

Proof of Proposition 5.3. Let $A, B \in \mathcal{F}_c(\mathbb{R}^p)$ be fuzzy sets, with A attaining maximal depth. It suffices to prove that for every $\lambda \in [0,1]$, the fuzzy set $C = (1-\lambda)\cdot A + \lambda \cdot B$ satisfies $d(A,B) = d(A,C) + d(C,B)$. This follows from the fact that

$$\begin{aligned} d(A,C) &= d(A, (1-\lambda)\cdot A + \lambda\cdot B) = \\ &= d((1-\lambda)\cdot A + \lambda A, (1-\lambda)\cdot A + (\lambda-1)\cdot A + \lambda\cdot B) = \\ &= d(\lambda\cdot A, \lambda\cdot B) = \lambda \cdot d(A,B), \end{aligned} \tag{14}$$

(where the second equality is due to A1 and the third to A2) and that, analogously to (14), we have $d(C,B) = (1-\lambda)\cdot d(A,B)$. □

Proof of Theorem 5.4. The proof is split into two steps. In the first step we prove that $j(C)$ is in the sphere $S(j(A), d(A,C))$ having center $j(A)$ and radius $d(A,C)$, as well as in $S(j(B), d(B,C))$. In the second step, we prove that if non-empty, the set $S(j(A), d(A,C)) \cap S(j(B), d(B,C))$ is a singleton.

Step I. By (9), it is clear that $d(A,C) \leq d(A,B)$. Then there exists $\lambda_0 \in [0,1]$ such that

$$d(A,C) = \lambda_0 \cdot d(A,B). \tag{15}$$

By A1, A2, and the convexity of A,

$$d(A,C) = d(\lambda_0\cdot A, \lambda_0 \cdot B) = d(A, (1-\lambda_0)\cdot A + \lambda_0\cdot B). \tag{16}$$

Additionally,

$$\begin{aligned} d(B,C) &= d(A,B) - d(A,C) = d(A,B) - \lambda_0\cdot d(A,B) = \\ &= (1-\lambda_0)\cdot d(A,B) = d((1-\lambda_0)\cdot A, (1-\lambda_0)\cdot B) = \\ &= d(B, (1-\lambda_0)\cdot A + \lambda_0 \cdot B). \end{aligned} \tag{17}$$

where the first equality is due to (9), the second to (15), the fourth to A2 and the fifth to A1 by adding up $\lambda_0 \cdot B$. By (16) and (17),

$$d(A,B) = d(A, (1-\lambda_0)\cdot A + \lambda_0\cdot B) + d(B, (1-\lambda_0)\cdot A + \lambda_0\cdot B) \tag{18}$$

Since j is an isometry,

$$\begin{aligned} \|j(A) - j(B)\| = \|j(A) - j((1-\lambda_0)\cdot A + \lambda_0\cdot B)\| + \\ + \|j(B) - j(\lambda_0\cdot A + \lambda_0\cdot B)\|. \end{aligned} \tag{19}$$

Thus, by equations (15), (17) and (19),

$$\begin{aligned} j(C), j((1-\lambda_0)\cdot A + \lambda_0\cdot B) \in S(j(A), d(A,C)), \\ j(C), j((1-\lambda_0)\cdot A + \lambda_0\cdot B) \in S(j(B), d(B,C)). \end{aligned} \tag{20}$$

Step II. Reasoning by contradiction, assume that there exist $x, y \in \mathbb{E}$ with $x \neq y$ such that $x, y \in S(j(A), d(A,C)) \cap S(j(B), d(B,C))$. Necessarily, $d(A,C) > 0$ and $d(B,C) > 0$. Then

$$\|d(A,C)^{-1}(j(A)-x)\| = \|d(A,C)^{-1}(j(A)-y)\| = 1$$

whence, by the strict convexity of $\mathbb{E}$,

$$\|\frac{1}{2}d(A,C)^{-1}(j(A)-x) + \frac{1}{2}d(A,C)^{-1}(j(A)-y)\| < 1,$$

equivalently

$$\|j(A) - \frac{1}{2}(x+y)\| < d(A,C).$$

We analogously prove

$$\|j(B) - \frac{1}{2}(x+y)\| < d(B,C).$$

Therefore

$$\|j(A) - \frac{1}{2}\cdot(x+y)\| + \|j(B) - \frac{1}{2}\cdot(x+y)\| < d(A,C) + d(B,C) = d(A,B). \tag{21}$$

Using the triangle inequality and (21),

$$d(A,B) = \|j(A) - j(B)\| \leq \|j(A) - \frac{1}{2}\cdot(x+y)\| + \|j(B) - \frac{1}{2}\cdot(x+y)\| < d(A,B),$$

which leads to a contradiction. We deduce that such $x \neq y$ cannot exist and, hence, the set $S(j(A), d(A,C)) \cap S(j(B), d(B,C))$ is a singleton.

To conclude the proof note that, by (20),

$$j(C) = j((1-\lambda_0)\cdot A + \lambda_0 \cdot B)$$

whence

$$C = (1-\lambda_0)\cdot A + \lambda_0 \cdot B$$

because j is an isometry and so $d(C, (1-\lambda_0)\cdot A + \lambda_0 \cdot B) = 0$. □

Proof of Proposition 5.8. Let $A \in \mathcal{F}_c(\mathbb{R}^p)$ be a fuzzy set with maximal depth and let $B \in \mathcal{F}_c(\mathbb{R}^p) \setminus \mathrm{I}_{\{0\}}$. By the assumptions on the metric d,

$$d(A + \lambda\cdot B, A) = d(\lambda\cdot B, \mathrm{I}_{\{0\}}) = \lambda\cdot d(B, \mathrm{I}_{\{0\}}) \to \infty \text{ as } \lambda \to \infty.$$

Then the sequence of fuzzy sets $\{A + \lambda\cdot B\}_{\lambda\in\mathbb{N}}$ satisfies the hypothesis of P4b and, consequently, $\lim_{\lambda\to\infty} D(A + \lambda\cdot B; \mathcal{X}) = 0$. □

The next result is used in the proof of Theorem 6.2.

Lemma 8.1. *Let $f, g : A \to [0,1]$ be two functions such that $A \subseteq \mathbb{R}^p$. Then*

$$\min\{\inf_{a\in A} f(a), \inf_{a\in A} g(a)\} = \inf_{a\in A} \min\{f(a), g(a)\}.$$

Proof of Theorem 6.2. Let $\mathcal{X}$ be a fuzzy random variable. We denote by $D^*_{FT}(\cdot;\mathcal{X})$ the expression in statement of the theorem and by $D_{FT}(\cdot;\mathcal{X})$ the fuzzy Tukey depth from Definition 6.1. By Lemma 8.1, $D_{FT}(U;\mathcal{X})$ equals

$$\min\left\{\inf_{(u,\alpha)\in\mathbb{S}^{p-1}\times[0,1]} \mathbb{P}[\omega\in\Omega : \mathcal{X}(\omega)\in S^-_{u,\alpha}], \inf_{(u,\alpha)\in\mathbb{S}^{p-1}\times[0,1]} \mathbb{P}[\omega\in\Omega : \mathcal{X}(\omega)\in S^+_{u,\alpha}]\right\}, \tag{22}$$

for all $U \in \mathcal{F}_c(\mathbb{R}^p)$. Note that for any $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$, we have

$$S^-_{u,\alpha,s_U(u,\alpha)} = S^-_{u,\alpha} \text{ and } S^+_{u,\alpha,s_U(u,\alpha)} = S^+_{u,\alpha}. \tag{23}$$

Furthermore, making use of (23), it is clear that, for each fixed $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$,

$$\inf_{(u,\alpha,t)\in I_U^-} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,t}] \le \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,s_U(u,\alpha)}]$$
$$= \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha}].$$

Thus

$$\inf_{(u,\alpha,t)\in I_U^-} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,t}] \le \inf_{(u,\alpha)\in \mathbb{S}^{p-1}\times[0,1]} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha}]. \tag{24}$$

For any given $u \in \mathbb{S}^{p-1}$, $\alpha \in [0,1]$ and $t \ge s_U(u,\alpha)$, we have that $S^-_{u,\alpha,s_U(u,\alpha)} \subseteq S^-_{u,\alpha,t}$, and consequently

$$\mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,t}] \ge \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,s_U(u,\alpha)}].$$

Thus, using (23) again,

$$\inf_{(u,\alpha,t)\in I_U^-} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,t}] \ge \inf_{(u,\alpha)\in \mathbb{S}^{p-1}\times[0,1]} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha}]. \tag{25}$$

By (24) and (25) we have the following equality:

$$\inf_{(u,\alpha,t)\in I_U^-} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha,t}] = \inf_{(u,\alpha)\in \mathbb{S}^{p-1}\times[0,1]} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^-_{u,\alpha}].$$

Analogously,

$$\inf_{(u,\alpha,t)\in I_U^+} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^+_{u,\alpha,t}] = \inf_{(u,\alpha)\in \mathbb{S}^{p-1}\times[0,1]} \mathbb{P}[\omega \in \Omega : \mathcal{X}(\omega) \in S^+_{u,\alpha}].$$

From the expression in (22), $D^*_{FT}(U;\mathcal{X}) = D_{FT}(U;\mathcal{X})$ for all $U \in \mathcal{F}_c(\mathbb{R}^p)$. □

Proof of Proposition 6.3. Since $\mathcal{X}(\omega) = \mathrm{I}_{\{X(\omega)\}}$ for all $\omega \in \Omega$ and $s_{\mathrm{I}_{\{a\}}}(u,\alpha) = \langle u, a\rangle$ for all $a \in \mathbb{R}^p$, $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$,

$$D_{FT}(\mathrm{I}_{\{a\}}, \mathcal{X}) = \min\{\inf_{u\in\mathbb{S}^{p-1}} \mathbb{P}[\omega \in \Omega : \langle u, X(\omega)\rangle \le \langle u, a\rangle], \inf_{u\in\mathbb{S}^{p-1}} \mathbb{P}[\omega \in \Omega : \langle u, X(\omega)\rangle \ge \langle u, a\rangle]\}. \tag{26}$$

Besides, for any $u \in \mathbb{S}^{p-1}$,

$$\mathbb{P}[\omega \in \Omega : \langle u, X(\omega)\rangle \ge \langle u, a\rangle] = \mathbb{P}[\omega \in \Omega : \langle -u, X(\omega)\rangle \le \langle -u, a\rangle].$$

Thus,

$$D_{FT}(\mathrm{I}_{\{a\}}; \mathcal{X}) = \inf_{u\in\mathbb{S}^{p-1}} \mathbb{P}[\omega \in \Omega : \langle u, X(\omega)\rangle \le \langle u, a\rangle] = D_T(a, X). \quad \square$$

Proof of Theorem 6.4. For any $A \in \mathcal{F}_c(\mathbb{R})$ we have $s_A(1,\alpha) = \sup A_\alpha$ and $s_A(-1,\alpha) = -\inf A_\alpha$ and, as $\mathbb{S}^0 = \{1,-1\}$,

$$D_{FT}(A;\mathcal{X}) = \inf_{\alpha\in[0,1]} \min\{\mathbb{P}(\sup \mathcal{X}_\alpha \le \sup A_\alpha), \mathbb{P}(\sup \mathcal{X}_\alpha \ge \sup A_\alpha),$$
$$\mathbb{P}(\inf \mathcal{X}_\alpha \le \inf A_\alpha), \mathbb{P}(\inf \mathcal{X}_\alpha \ge \inf A_\alpha)\}.$$

Since, by Theorem 2.5, $s_{\tilde{M}e(\mathcal{X})}(u,\alpha) = \mathrm{Med}(s_{\mathcal{X}}(u,\alpha))$ for all $u \in \mathbb{S}^0$ and $\alpha \in [0,1]$,

$$D_{FT}(\tilde{M}e(\mathcal{X}); \mathcal{X}) = 1/2.$$

Let $U \in \mathcal{F}_c(\mathbb{R}) \backslash \{\tilde{M}e(\mathcal{X})\}$. Then there exist $u_0 \in \mathbb{S}^0$ and $\alpha_0 \in [0,1]$ such that

$$s_U(u_0,\alpha_0) \ne s_{Me(\mathcal{X})}(u_0,\alpha_0) = \mathrm{Med}(s_{\mathcal{X}}(u_0,\alpha_0)).$$

As the medians of $\sup \mathcal{X}_\alpha$ and $\inf \mathcal{X}_\alpha$ are unique for each $\alpha \in [0,1]$, we have either

$$\mathbb{P}(s_{\mathcal{X}}(u_0,\alpha_0) \le s_U(u_0,\alpha_0)) < 1/2$$

or

$$\mathbb{P}(s_{\mathcal{X}}(u_0,\alpha_0) \ge s_U(u_0,\alpha_0)) < 1/2$$

and, consequently, $D_{FT}(U;\mathcal{X}) < 1/2$. □

The following result is for later use; for instance, to prove that D_{FT} is affine equivariant.

Proposition 8.2. *Let $A \in \mathcal{F}_c(\mathbb{R}^p)$ and let $M \in \mathcal{M}_{p\times p}(\mathbb{R})$ be a non-singular matrix. Then,*

$$s_{M\cdot A}(u,\alpha) = \|M^T \cdot u\| \cdot s_A\left(\frac{1}{\|M^T \cdot u\|} \cdot M^T \cdot u, \alpha\right),$$

for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$.

Proof. As $A \in \mathcal{F}_c(\mathbb{R}^p)$ is a fuzzy set and $M \in \mathcal{M}_{p\times p}(\mathbb{R})$ a non-singular matrix, we have that $M \cdot A \in \mathcal{F}_c(\mathbb{R}^p)$ and it is defined as the fuzzy set with α-levels

$$(M\cdot A)_\alpha = \{M \cdot x : x \in A_\alpha\} = M \cdot A_\alpha.$$

Taking into account the definition of the support function of fuzzy sets and Claim 1 below,

$$s_{M\cdot A}(u,\alpha) = \sup_{v\in A_\alpha} \langle u, M\cdot v\rangle = \sup_{v\in A_\alpha} \langle M^T \cdot u, v\rangle, \tag{27}$$

which yields the result. □

Claim 1. *$\langle x, M\cdot y\rangle = \langle M^T \cdot y, x\rangle$ for any $x, y \in \mathbb{R}^p$ and $M \in \mathcal{M}_{p\times p}(\mathbb{R})$.*

Proof. Let $x = (x_1,\dots,x_p)$, $y = (y_1,\dots,y_p) \in \mathbb{R}^p$. Let $M = (m_{i,j})_{i,j} \in \mathcal{M}_{p\times p}(\mathbb{R})$ be a non-singular matrix. Then,

$$\begin{aligned}\langle x, M\cdot y\rangle &= \langle x, (\sum_{i=1}^p m_{1,i}\cdot y_i,\dots,\sum_{i=1}^p m_{p,i}\cdot y_i)\rangle = \\ &= x_1 \cdot \sum_{i=1}^p m_{1,i}\cdot y_i + \dots + x_p \cdot \sum_{i=1}^p m_{p,i}\cdot y_i = \\ &= y_1 \cdot \sum_{i=1}^p m_{i,1}\cdot x_i + \dots + y_p \cdot \sum_{i=1}^p m_{i,p}\cdot x_i = \\ &= \langle M^T \cdot x, y\rangle,\end{aligned}$$

where the second and fourth equalities are by definition of the inner product and third equality is due to a reordering of the summands. □

The following proposition is used in the proof of property 4b for the Tukey depth.

Proposition 8.3. *Let $\{A_n\}_n$ be a sequence of fuzzy sets and let $A \in \mathcal{F}_c(\mathbb{R}^p)$ such that $\lim_n d_\infty(A_n, A) = \infty$. Then, there exists $u \in \mathbb{S}^{p-1}$ such that $\lim_n s_{A_n}(u,0) = \infty$.*

Proof. By the assumption, $\lim_n d_\infty(A, A_n) = \infty$. As A is a fixed fuzzy set, it is equivalent to say that $\lim_n d_\infty(A_n, \mathrm{I}_{\{0\}}) = \infty$, thus

$$\lim_{n\to\infty} \sup_{\alpha\in[0,1]} d_H(A_{n,\alpha}, \{0\}) = \infty,$$

where d_H denotes the Hausdorff distance and $A_{n,\alpha}$ is the α-level of A_n for all $\alpha \in [0,1]$ and $n \in \mathbb{N}$. It is a basic property of the Hausdorff distance that $d_H(A_{n,\alpha}, \{0\}) = \sup\{\|x\| : x \in A_{n,\alpha}\}$ for all $\alpha \in [0,1]$. As $A_{n,\alpha} \subseteq A_{n,0}$ for all $\alpha \in [0,1]$ and $n \in \mathbb{N}$, we have that

$$d_\infty(A_n, \mathrm{I}_{\{0\}}) = d_H(A_{n,0}, \{0\}) = \sup\{\|x\| : x \in A_{n,0}\}.$$

The function $f_n : A_{n,0} \to \mathbb{R}$ defined by $f_n(x) = \|x\|$ is a continuous function defined over a compact set, thus we have that f_n is bounded and attains its supremum on $A_{n,0}$ for all $n \in \mathbb{N}$. Let x_n be an element of $A_{n,0}$ at which f_n is maximized. By hypothesis, we have that $\lim_n \|x_n\| = \infty$. In particular, there exists $u \in \mathbb{S}^{p-1}$ such that $\lim_n \langle u, x_n\rangle = \infty$ (we can take u to be one of the vectors in the standard basis of $\mathbb{R}^p$). By definition of the support function, we have that $\langle u, x_n\rangle \le s_{A_n}(u,0)$ and $\lim_n s_{A_n}(u,0) = \infty$. □

Proof of Theorem 6.6. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the probabilistic space associated with a fuzzy random variable $\mathcal{X}$.

Proof of P1. Due to the linearity of the support functions of fuzzy sets, for any non-singular matrix $M \in \mathcal{M}_{p\times p}(\mathbb{R})$ and $A, B \in \mathcal{F}_c(\mathbb{R}^p)$,

$$\begin{aligned}&\mathbb{P}(s_{M\cdot\mathcal{X}+B}(u,\alpha) \le s_{M\cdot A+B}(u,\alpha)) = \mathbb{P}(s_{M\cdot\mathcal{X}}(u,\alpha) + s_B(u,\alpha) \le s_{M\cdot A}(u,\alpha) + s_B(u,\alpha))\\ &= \mathbb{P}(s_{M\cdot\mathcal{X}}(u,\alpha) \le s_{M\cdot A}(u,\alpha))\end{aligned}$$

for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$. Then, it suffices to prove $D_{FT}(A;\mathcal{X}) = D_{FT}(M\cdot A; M\cdot\mathcal{X})$. By Proposition 8.2,

$$\begin{aligned}&\inf_{(u,\alpha)\in\mathbb{S}^{p-1}\times[0,1]} \mathbb{P}(s_{M\cdot\mathcal{X}}(u,\alpha) \le s_{M\cdot A}(u,\alpha)) =\\ &= \inf_{(u,\alpha)\in\mathbb{S}^{p-1}\times[0,1]} \mathbb{P}(s_{\mathcal{X}}(\frac{1}{\|M^T u\|}\cdot M^T u, \alpha) \le s_A(\frac{1}{\|M^T u\|}\cdot M^T u, \alpha)).\end{aligned} \tag{28}$$

Let us consider the map $f : \mathbb{S}^{p-1} \to \mathbb{S}^{p-1}$ defined by $f(u) := (1/\|M^T u\|)M^T u$ for all $u \in \mathbb{S}^{p-1}$. Its inverse map is $f^{-1}(u) = (1/\|(M^T)^{-1}u\|)(M^T)^{-1}u$, which is well defined, thus f is bijective. Then

$$\begin{aligned}&\inf_{u,\alpha} \mathbb{P}(s_{\mathcal{X}}(\frac{1}{\|M^T u\|}\cdot M^T u, \alpha) \le s_A(\frac{1}{\|M^T u\|}\cdot M^T u, \alpha)) =\\ &= \inf_{u,\alpha} \mathbb{P}(s_{\mathcal{X}}(f(u),\alpha) \le s_A(f(u),\alpha))\\ &= \inf_{u,\alpha} \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \le s_A(u,\alpha)).\end{aligned} \tag{29}$$

Analogously,

$$\begin{aligned}&\inf_{u,\alpha} \mathbb{P}(s_{M\cdot\mathcal{X}}(\frac{1}{\|M^T u\|}\cdot M^T u, \alpha) \ge s_{M\cdot A}(\frac{1}{\|M^T u\|}\cdot M^T u, \alpha)) =\\ &= \inf_{u,\alpha} \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \ge s_A(u,\alpha)).\end{aligned} \tag{30}$$

Now (28), (29) and (30) imply together $D_{FT}(A;\mathcal{X}) = D_{FT}(M\cdot A; M\cdot\mathcal{X})$.

Proof of P2. If $\mathcal{X}$ is F-symmetric with respect to a fuzzy set A then

$$\begin{aligned}\mathbb{P}(s_{\mathcal{X}}(u,\alpha) - s_A(u,\alpha) \le t) &= \mathbb{P}(s_A(u,\alpha) - s_{\mathcal{X}}(u,\alpha) \le t)\\ &= \mathbb{P}(s_{\mathcal{X}}(u,\alpha) - s_A(u,\alpha) \ge -t),\end{aligned}$$

for all $u \in \mathbb{S}^{p-1}$, $\alpha \in [0,1]$ and $t \in \mathbb{R}$. In particular, for $t = 0$,

$$\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \le s_A(u,\alpha)) = \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \ge s_A(u,\alpha)) \tag{31}$$

for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$.

Let $B \in \mathcal{F}_c(\mathbb{R}^p)$. We first study the case in which $s_B(u,\alpha) \in \mathrm{Med}(s_{\mathcal{X}}(u,\alpha))$ for all $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$. Let us fix $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$. Without loss of generality, let us assume that $s_A(u,\alpha) \le s_B(u,\alpha)$ (the other case follows analogously). In this case, $(-\infty, s_A(u,\alpha)] \subset (-\infty, s_B(u,\alpha)]$ and $[s_B(u,\alpha), \infty) \subset [s_A(u,\alpha), \infty)$, which implies

$$\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_B(u,\alpha)) \geq \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_A(u,\alpha))$$

and

$$\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_A(u,\alpha)) \geq \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_B(u,\alpha)).$$

By this and Equation (31),

$$\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_B(u,\alpha)) \leq \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_A(u,\alpha)) =$$
$$\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_A(u,\alpha)) \leq \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_B(u,\alpha)).$$

Consequently,

$$\min\{\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_B(u,\alpha)), \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_B(u,\alpha))\} \leq$$
$$\min\{\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_A(u,\alpha)), \mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_A(u,\alpha))\}$$

and, as $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$ have been taken arbitrarily, we have that

$$D_{FT}(A;\mathcal{X}) \geq D_{FT}(B;\mathcal{X}). \tag{32}$$

Let us study the case in which there exists $u \in \mathbb{S}^{p-1}$ and $\alpha \in [0,1]$ such that $s_B(u,\alpha) \notin \mathrm{Med}(s_{\mathcal{X}}(u,\alpha))$. Then, $\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \leq s_B(u,\alpha)) < 1/2$ or $\mathbb{P}(s_{\mathcal{X}}(u,\alpha) \geq s_B(u,\alpha)) < 1/2$. Then, making use of the Definition 6.1, we have that $D_{FT}(A;\mathcal{X}) > D_{FT}(B;\mathcal{X})$. This together with (32) implies $D_{FT}(A;\mathcal{X}) = \sup_{U \in \mathcal{F}_c(\mathbb{R}^p)} D_{FT}(U;\mathcal{X})$.

Proof of P3a. Let $A \in \mathcal{F}_c(\mathbb{R}^p)$ have maximal Tukey depth. For any $U \in \mathcal{F}_c(\mathbb{R}^p)$, set

$$D_{T_1}(U;\mathcal{X}) := \inf_{(u,\alpha,t)\in I_U^-} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}) \text{ and } D_{T_2}(U;\mathcal{X}) := \inf_{(u,\alpha,t)\in I_U^+} \mathbb{P}(\mathcal{X} \in S^+_{u,\alpha,t});$$

thus

$$D_{FT}(U;\mathcal{X}) = \min\{D_{T_1}(U;\mathcal{X}), D_{T_2}(U;\mathcal{X})\}.$$

Then, to prove $D_{FT}((1-\lambda)\cdot A + \lambda \cdot B;\mathcal{X}) \geq D_{FT}(B;\mathcal{X})$, it suffices to show

$$D_{T_1}((1-\lambda)\cdot A + \lambda \cdot B;\mathcal{X}) \geq D_{FT}(B;\mathcal{X})$$

and

$$D_{T_2}((1-\lambda)\cdot A + \lambda \cdot B;\mathcal{X}) \geq D_{FT}(B;\mathcal{X}).$$

Let us focus on D_{T_1}, as the process for D_{T_2} is analogous. The set

$$\mathbb{A} = \{(u,\alpha,t) \in \mathbb{S}^{p-1} \times [0,1] \times \mathbb{R} : (1-\lambda)\cdot A + \lambda \cdot B \in S^-_{u,\alpha,t}\}$$

can be written as the union $\mathbb{A}_1 \cup \mathbb{A}_2 \cup \mathbb{A}_3$ of

$$\mathbb{A}_1 = \{(u,\alpha,t) \in \mathbb{S}^{p-1} \times [0,1] \times \mathbb{R} : A \in S^-_{u,\alpha,t}, B \in S^-_{u,\alpha,t}\},$$
$$\mathbb{A}_2 = \{(u,\alpha,t) \in \mathbb{S}^{p-1} \times [0,1] \times \mathbb{R} : A \in S^-_{u,\alpha,t}, B \in S^+_{u,\alpha,t}, (1-\lambda)\cdot A + \lambda \cdot B \in S^-_{u,\alpha,t}\}$$

and

$$\mathbb{A}_3 = \{(u,\alpha,t) \in \mathbb{S}^{p-1} \times [0,1] \times \mathbb{R} : A \in S^+_{u,\alpha,t}, B \in S^-_{u,\alpha,t}, (1-\lambda)\cdot A + \lambda \cdot B \in S^-_{u,\alpha,t}\}.$$

Then

$$D_{T_1}((1-\lambda)\cdot A + \lambda \cdot B;\mathcal{X}) = \min\{\inf_{(u,\alpha,t)\in\mathbb{A}_1} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}), \inf_{(u,\alpha,t)\in\mathbb{A}_2} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}), \inf_{(u,\alpha,t)\in\mathbb{A}_3} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t})\}. \tag{33}$$

Since $\mathbb{A}_1 \subseteq \{(u,\alpha,t) \in \mathbb{S}^{p-1} \times [0,1] \times \mathbb{R} : A \in S^-_{u,\alpha,t}\}$,

$$\inf_{(u,\alpha,t)\in\mathbb{A}_1} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}) \geq \inf_{(u,\alpha,t)\in I_U^-} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}) \geq D_{TF}(A;\mathcal{X}) \geq D_{TF}(B;\mathcal{X}), \tag{34}$$

where the second inequality is due to Theorem 6.2. Analogously,

$$\mathbb{A}_2 \subseteq \{(u, \alpha, t) \in \mathbb{S}^{p-1} \times [0, 1] \times \mathbb{R} : A \in S^-_{u,\alpha,t}\}$$

which implies

$$\inf_{(u,\alpha,t)\in\mathbb{A}_2} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}) \geq D_{FT}(B; \mathcal{X}). \tag{35}$$

Finally, $\mathbb{A}_3 \subseteq \{(u, \alpha, t) \in \mathbb{S}^{p-1} \times [0, 1] \times \mathbb{R} : B \in S^-_{u,\alpha,t}\}$ and then

$$\inf_{(u,\alpha,t)\in\mathbb{A}_3} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}) \geq \inf_{(u,\alpha,t)\in I^-_B} \mathbb{P}(\mathcal{X} \in S^-_{u,\alpha,t}) \geq D_{TF}(B; \mathcal{X}). \tag{36}$$

Thus, recalling (33), we have

$$D_{T_1}((1-\lambda) \cdot A + \lambda \cdot B; \mathcal{X}) \geq D_{FT}(B; \mathcal{X})$$

as wished.

Proof of P3b. Using the ρ_2 distance in Theorem 5.4, we have that P3a and P3b are equivalent.

Proof of P4b and P4a.* We begin by proving property P4b for the metric d_∞.

Let $\{A_n\}_n$ be a sequence of fuzzy sets and let $A \in \mathcal{F}_c(\mathbb{R}^p)$ such that $\lim_n d_\infty(A, A_n) = \infty$. It is equivalent to $\lim_n d_\infty(A_n, \mathrm{I}_{\{0\}}) = \infty$. By Proposition 8.3, there exists $u_0 \in \mathbb{S}^{p-1}$ such that $\lim_n s_{A_n}(u_0, 0) = \infty$. By the definition of the Tukey depth,

$$\begin{aligned} 0 \leq D_{FT}(A_n; \mathcal{X}) \leq \mathbb{P}(s_{\mathcal{X}}(u_0, 0) \geq s_{A_n}(u_0, 0)) = \\ 1 - F_{u_0,0}(s_{A_n}(u_0, 0)) + \mathbb{P}(s_{\mathcal{X}}(u_0, 0) = s_{A_n}(u_0, 0)) \end{aligned} \tag{37}$$

where $F_{u,\alpha}$ denotes the cumulative distribution function of the real random variable $s_{\mathcal{X}}(u, \alpha)$. Clearly,

$$\lim_n \mathbb{P}(s_{\mathcal{X}}(u_0, 0) = s_{A_n}(u_0, 0)) = 0,$$
$$\lim_n F_{u_0,0}(s_{A_n}(u_0, 0)) = 1,$$

both from the fact that $\lim_n s_{A_n}(u_0, 0) = \infty$. Taking limits in (37), we obtain $\lim_n D_{FT}(A_n; \mathcal{X}) = 0$ as wished.

By Proposition 5.2.(c), D_{FT} satisfies P4a* as well. Moreover, P4b for d_∞ implies P4b for ρ_r, as follows from the statement of property P4b and the fact that $\rho_r \leq d_\infty$.

Proof of P4a. It is fulfilled by virtue of Proposition 5.8, as ρ_r satisfies A1 and A2 and the fuzzy Tukey depth has just been shown to satisfy P4b. □

Proof of Proposition 6.7. It is a direct consequence of Example 5.6, which shows that the D_{FT} function violates P3b when considering d_r metrics. □

9. Concluding remarks

We proposed two different notions of depth functions for fuzzy sets: semilinear depth and geometric depth. The former generalizes in a natural way the notion of depth proposed for the multivariate case. The latter takes into account that fuzzy sets can be considered as a metric space and we consider some of the most well-known metrics (ρ_r and d_r metrics). We showed the viability of these definitions by giving an instance of depth function which fulfills both and generalizes the corresponding instance of multivariate depth function. Tukey's halfspace depth was the first proper instance of depth function proposed and remains one of the best understood and most useful. We gave an example which shows that for d_r distances, the fuzzy Tukey depth is not a geometric depth function due to the failure of property P3b.

The notion of a semilinear depth function is easy to understand and it is easier to prove whether a certain function satisfies it, because it only depends on the arithmetics of fuzzy sets. Another conclusion is that geometric depth is similar to semilinear depth when the metric ρ_2 is considered (see Theorem 5.4), but metrics based on the Hausdorff distance generate some different situations. Note that ρ_2 embeds $\mathcal{F}_c(\mathbb{R}^p)$ into a separable Hilbert space and hence has the most similar geometry to that of the Euclidean space.

Semilinear depth is specially interesting when one does not have salient reasons for a specific metric to be preferred. However, if some metric is considered most appropriate one can resort to the notion of geometric depth for that metric, taking into account that properties A1 and A2 are enough to ensure that the geometric depth function is also a semilinear depth function.

For future work, it is interesting to adapt more notions of multivariate depth to the fuzzy setting. Notice that the effort needed to prove the depth properties for a specific function is harder than in the multivariate case. It would also be interesting to consider a wider family of metrics in the properties of geometric depth.

It should be remarked that the computation of the depth function induced by a data sample is related to computational geometry problems which admit faster algorithms (see [1] for details). Those ideas and random projection techniques may be useful to create algorithms in the fuzzy case.

Some connections between statistical depth and fuzzy sets have been discussed in the literature [41]. Conceptually, a statistical depth function is similar to a fuzzy set where the degree of membership indicates how deep in the distribution a given point is. This is related, semantically, to the notion of a fuzzy set of central points in [42], where the degree of membership tries to capture the extent to which a given point is the center of the distribution. Some multivariate depth functions were shown to be fuzzy sets of central points, in particular Tukey depth. But depth in a space where each point is a fuzzy set was not consider in those papers.

The effect of replacing the transformation $A \mapsto M \cdot A + B$ by $A \mapsto M \cdot A + I_{\{b\}}$ in property P1 (where $B \in \mathcal{F}_c(\mathbb{R}^p)$ but $b \in \mathbb{R}^p$) is not clear yet. As we have shown, Tukey depth satisfies P1 in its current (stronger) form. So will depth functions whose definition involves a metric satisfying property A2. It remains plausible that some functions may satisfy this weaker version but not property P1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

A. Nieto-Reyes and L. Gonzalez are supported by the Spanish Ministerio de Economía, Industria y Competitividad grant MTM2017-86061-C2-2-P. P. Terán is supported by the Ministerio de Economía y Competitividad grant MTM2015-63971-P, the Ministerio de Ciencia, Innovación y Universidades grant PID2019-104486GB-I00 and the Consejería de Empleo, Industria y Turismo del Principado de Asturias grant GRUPIN-IDI2018-000132.

References

[1] G. Aloupis, Geometric measures of data depth, in: R.Y. Liu, R. Serfling, D.L. Souvaine (Eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 72, American Mathematical Society, Providence, 2006, pp. 147–158.
[2] C. Bertoluzza, M.A. Gil, D.A. Ralescu (Eds.), Statistical Modeling, Analysis and Management for Fuzzy Data, 1st edition, Physica, Heidelberg, 2002.
[3] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo, S. de la Rosa de Sáa, B. Sinova, Random fuzzy sets: a mathematical tool to develop statistical fuzzy data analysis, Iran. J. Fuzzy Syst. 10 (2013) 1–28.
[4] P. Billingsley, Probability and Measure, 3rd edition, Wiley, New York, 1995.
[5] A. Chakraborty, P. Chaudhuri, The spatial distribution in infinite dimensional spaces and related quantiles and depths, Ann. Stat. 42 (2014) 1203–1231.
[6] J. Chowdhury, P. Chaudhuri, Nonparametric depth and quantile regression for functional data, Bernoulli 25 (2019) 395–423.
[7] A. Colubi, Statistical inference about the means of fuzzy random variables. Applications to the analysis of fuzzy- and real-valued data, Fuzzy Sets Syst. 160 (3) (2009) 344–356.
[8] A. Colubi, D. Dubois, Fuzzy sets in statistics, Comput. Stat. Data Anal. 56 (4) (2012) 892–994 (special issue).
[9] R. Coppi, M.A. Gil, H.A.L. Kiers, The fuzzy approach to statistical analysis, Comput. Stat. Data Anal. 51 (1) (2006) 1–452 (special issue).
[10] J.A. Cuesta-Albertos, A. Nieto-Reyes, The random Tukey depth, Comput. Stat. Data Anal. 52 (2008) 4979–4988.
[11] J.A. Cuesta-Albertos, A. Nieto-Reyes, Functional classification and the random Tukey depth. Practical issues, in: C. Borgelt, G. González-Rodríguez, W. Trutsching, M.A. Lubiano, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Combining Soft Computing and Statistical Methods in Data Analysis, vol. 77, Springer, Berlin, 2010, pp. 123–130.

[12] A. Cuevas, M. Febrero, R. Fraiman, Robust estimation and classification for functional data via projection-based depth notions, Comput. Stat. 22 (2007) 481–496.

[13] P. Diamond, P. Kloeden, Metric spaces of fuzzy sets, Fuzzy Sets Syst. 35 (1990) 241–249.

[14] D. Dubois, H. Prade, Gradualness, uncertainty and bipolarity: making sense of fuzzy sets, Fuzzy Sets Syst. 192 (2012) 3–24.

[15] S. Dutta, A-K. Ghosh, P. Chaudhuri, Some intriguing properties of Tukey's halfspace depth, Bernoulli 17 (2011) 1420–1434.

[16] P. D'Urso, M.A. Gil, Fuzzy data analysis and classification, Adv. Data Anal. Classif. 11 (4) (2017) 645–808 (special issue).

[17] R. Féron, Sur les notions de distance et d'écart dans une structure floue et leurs applications aux ensembles aléatoires flous, C. R. Acad. Sci. Paris A 289 (1979) 35–38.

[18] R. Fraiman, G. Muniz, Trimmed means for functional data, Test 10 (2001) 419–440.

[19] G. Francisci, C. Agostinelli, A. Nieto-Reyes, A.N. Vidyashankar, Analytic and statistical properties of local depth functions motivated by clustering applications, Preprint, https://arxiv.org/abs/2008.11957, 2021.

[20] G. Francisci, A. Nieto-Reyes, C. Agostinelli, Generalization of the simplicial depth: no vanishment outside the convex hull of the distribution support, Preprint, https://arxiv.org/abs/1909.02739, 2021.

[21] K. Goebel, Convexity of balls and fixed-point theorems for mappings with nonexpansive square, Compos. Math. 22 (1970) 269–274.

[22] G. Gónzalez-Rodríguez, A. Colubi, M.A. Gil, Fuzzy data treated as functional data: a one-way ANOVA test approach, Comput. Stat. Data Anal. 56 (2010) 943–955.

[23] T. Harris, D. Tucker, B. Li, L. Shand, Elastic depths for detecting shape anomalies in functional data, Technometrics (2020), https://doi.org/10.1080/00401706.2020.1811156.

[24] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic. Theory and Applications, Prentice Hall, Upper Saddle River, 1993.

[25] V. Krätschmer, A unified approach to fuzzy random variables, Fuzzy Sets Syst. 123 (2001) 1–9.

[26] S. Li, Y. Ogura, V. Kreinovich, Limit Theorems and Applications of Set-Valued and Fuzzy Set-Valued Random Variables, 1st edition, Springer, Netherlands, 2002.

[27] R.Y. Liu, On a notion of data depth based upon random simplices, Ann. Stat. 18 (1990) 405–414.

[28] R.Y. Liu, J.M. Parelius, K. Singh, Multivariate analysis by data depth: descriptive statistics, graphics and inference, Ann. Stat. 27 (3) (1999) 783–858, with discussion.

[29] S. López-Pintado, J. Romo, On the concept of depth for functional data, J. Am. Stat. Assoc. 104 (2009) 718–734.

[30] I. Molchanov, Theory of Random Sets, 2nd edition, Springer, London, 2017.

[31] B. Möller, M. Beer, Fuzzy Randomness: Uncertainty in Civil Engineering and Computational Mechanics, 1st edition, Springer, Berlin, 2004.

[32] H.T. Nguyen, B. Wu, Fundamentals of Statistics with Fuzzy Data, 1st edition, Springer, Berlin, 2006.

[33] A. Nieto-Reyes, H. Battey, A topologically valid definition of depth for functional data, Stat. Sci. 31 (2016) 61–79.

[34] A. Nieto-Reyes, H. Battey, A topologically valid construction of depth for functional data, J. Multivar. Anal. 184 (2021), https://doi.org/10.1016/j.jmva.2021.104738.

[35] M.L. Puri, D.A. Ralescu, The concept of normality for fuzzy random variables, Ann. Probab. 11 (1985) 1373–1379.

[36] M.L. Puri, D.A. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. 114 (1986) 409–422.

[37] R. Serfling, A depth function and a scale curve based on spatial quantiles, in: Y. Dodge (Ed.), Statistical Data Analysis Based on L_1-Norm and Related Methods, Birkhäuser, Basel, 2002, pp. 25–38.

[38] B. Sinova, M.R. Casals, A. Colubi, M.A. Gil, The median of a random interval, in: C. Borgelt, G. González-Rodríguez, W. Trutsching, M.A. Lubiano, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Combining Soft Computing and Satistical Methods in Data Analysis, vol. 77, Springer, Heidelberg, 2010, pp. 575–583.

[39] B. Sinova, M.A. Gil, A. Colubi, E. Van Aelst, The median of a random fuzzy number. The 1-norm distance approach, Fuzzy Sets Syst. 200 (2012) 99–115.

[40] B. Sinova, E. Van Aelst, P. Terán, M-estimators and trimmed means: from Hilbert-valued to fuzzy set-valued data, Adv. Data Anal. Classif. 15 (2021), https://doi.org/10.1007/s11634-020-00402-x.

[41] P. Terán, Connections between statistical depth functions and fuzzy sets, in: C. Borgelt, G. González-Rodríguez, W. Trutsching, M.A. Lubiano, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Combining Soft Computing and Statistical Methods in Data Analysis, vol. 77, Springer, Berlin, 2010, pp. 611–618.

[42] P. Terán, Centrality as a gradual notion: a new bridge between fuzzy sets and statistics, Int. J. Approx. Reason. 52 (2011) 1243–1256.

[43] W. Trutschnig, G. González-Rodríguez, A. Colubi, M.A. Gil, A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread, Inf. Sci. 179 (2009) 3964–3972.

[44] J.W. Tukey, Mathematics and picturing data, in: R.D. James (Ed.), Proceedings of the International Congress of Mathematicians, vol. 2, Canadian Mathematical Congress, Montreal QC, 1975, pp. 523–531.

[45] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.

[46] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Part 1, Inf. Sci. 8 (1975) 199–249, Part 2, Inf. Sci. 8 (1975) 301–353, Part 3, Inf. Sci. 8 (1975) 43–80.

[47] Y. Zuo, R. Serfling, General notions of statistical depth function, Ann. Stat. 28 (2000) 461–482.

[48] Y. Zuo, R. Serfling, On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry, J. Stat. Plan. Inference 84 (2000) 55–79.