ON A GENERAL DEFINITION OF DEPTH FOR FUNCTIONAL DATA

IRÈNE GIJBELS[1] AND STANISLAV NAGY[1,2]

Abstract. In this paper we provide an elaboration on the desirable properties of statistical depths for functional data. Although a formal definition has been put forward in the literature, there are still several unclarities to be tackled, and further insights to be gained. Herein, a few interesting connections between the wanted properties are found. In particular, it is demonstrated that the conditions needed for some desirable properties to hold are extremely demanding, and virtually impossible to be met for common depths. We establish adaptations of these properties which prove to be still sensible, and more easily met by common functional depths.

1. Introduction

For univariate data the sample median is well-known to be appropriately describing the centre of a data cloud. An extension of this concept for multivariate data (say p-dimensional) is the notion of a point (in $\mathbb{R}^p$) for which a statistical depth function is maximized. A key issue is then to define what is a valid statistical depth function, and what are its desirable properties. Zuo and Serfling (2000a) clearly answer this question by listing four properties that a depth function should satisfy in $\mathbb{R}^p$:

- ⋆ Property ZS-1: *Affine invariance*; the depth of a point should not depend on the coordinate system, or on the measurement scales used.
- ⋆ Property ZS-2: *Maximality at the centre*; the depth function should attain its maximum value at the centre of symmetry of the data cloud, if the data cloud is symmetric.
- ⋆ Property ZS-3: *Monotonicity relative to the deepest point*; as a point moves away from the deepest point (the 'centre'), the depth should decrease monotonically.
- ⋆ Property ZS-4: *Vanishing at infinity*; the depth of a point should tend to zero, as the point moves to an infinite distance from the data cloud.

For *functional* data, i.e. data that are functions, it is far more tricky to determine which properties should a statistical depth function satisfy. Let us first introduce some notation. For a compact set $\mathcal{V} \subset \mathbb{R}^d$, consider the functions $x\colon \mathcal{V} \to \mathbb{R}$, and denote this space of functions by $\mathfrak{F}$. The function space $\mathfrak{F}$ is a normed vector space with norm $\|x\|$ and with the corresponding 'distance' metric $d(x,y) = \|x-y\|$, for $x, y \in \mathfrak{F}$. Some examples are (i) $\mathfrak{F} = \mathcal{C}$, the space of all continuous functions with $\|x\|_\infty = \sup_{v\in\mathcal{V}} |x(v)|$ the uniform norm, or (ii) $\mathfrak{F} = \mathbb{L}_2$, the space of all square integrable functions with

[1] KU Leuven, Department of Mathematics and Leuven Statistics Research Centre (LStat), Belgium

[2] Charles University, Department of Probability and Math. Statistics, Czech Republic

E-mail addresses: irene.gijbels@kuleuven.be, nagy@karlin.mff.cuni.cz.

Date: July 31, 2017.

$\|x\|_{\mathbb{L}_2} = \sqrt{\int_{\mathcal{V}} x(v)^2 \,\mathrm{d}\, v}$. In this paper, spaces $\mathcal{C}$ and $\mathbb{L}_2$ are always assumed to be equipped with the norms $\|\cdot\|_\infty$ and $\|\cdot\|_{\mathbb{L}_2}$, respectively.

The set of all Borel probability measures on $\mathfrak{F}$ is denoted by $\mathcal{P}$. Denoting $(\Omega, \mathcal{B}, \mathsf{P})$ the probability space on which all the random variables are defined, a measurable mapping $X\colon \Omega \to \mathfrak{F}$ is called a random function taking values in $\mathfrak{F}$. A *statistical depth functional* is then a mapping $D\colon \mathfrak{F} \times \mathcal{P} \to \mathbb{R}\colon\ (x, P) \mapsto D(x, P)$.

In a recent paper Nieto-Reyes and Battey (2016) provide a general definition of a statistical depth functional (subsequently called just depth, or depth functional) for data taking the form of random functions. They list six key properties (denoted P-1–P-6) that a reasonable depth functional D must satisfy:

- ⋆ Property P-1: *Distance invariance of D*;
- ⋆ Property P-2: *Maximality of D at the centre*;
- ⋆ Property P-3: *Monotonicity of D relative to the deepest point*;
- ⋆ Property P-4: *Upper semi-continuity of D in any function $x \in \mathfrak{F}$*;
- ⋆ Property P-5: *Receptivity of D to convex hull width across domain*;
- ⋆ Property P-6: *Continuity of D in P*.

See Nieto-Reyes and Battey (2016) for formal definitions of these properties. These properties are aimed to be suitable modifications of Properties ZS-1–ZS-4 (established in the multivariate setting) to the infinite-dimensional functional setting. Finding such modifications is far from evident, though, and simple generalizations need to be investigated with a lot of care. Take, for example, the notion of symmetry of a distribution used in Property ZS-2. Even in the multivariate setting, there is no unique concept of symmetry. This led Nieto-Reyes and Battey (2016) to consider Property P-2G: *Maximality of D at a Gaussian process mean*, instead of Property P-2, which is a straightforward translation of Property ZS-2 towards functional data. In Nieto-Reyes and Battey (2016), an extensive comparative study – with respect to (w.r.t.) these properties – of the most important representatives of existing depth functionals, is given. Their findings are summarized in Table 2 of that paper.

Such a coherent survey of theoretical properties of depth functionals is highly relevant. The body of existing literature on the subject is substantial, though theoretical advances are rather scattered, and lack unity of exposition. The overall aim of this paper is to contribute further to the concept of data depth for general infinite-dimensional, and complex data. Our contribution is two-fold: (i) we broaden the scope of the discussion generated in Nieto-Reyes and Battey (2016), provide a close look at the desirable properties P-1–P-6 from an analytical point of view, and point out that for some properties one needs to look for replacements or adaptations; (ii) we extend the survey with several additional depth functionals, and establish some important theoretical properties of these depth functionals.

More specifically, for Properties P-1–P-6 (including P-2G), we establish the following additional insights:

- Property P-1 turns out to be very demanding in functional spaces. In full generality (i.e. without further restrictions) it does not hold for any of the considered depth functionals (Section 2.1);
- Property P-2G is not valid for the band depth (Section 2.2);
- Property P-3 does not represent a counterpart of the monotonicity property ZS-3 for finite-dimensional depths introduced in Zuo and Serfling (2000a). Actually, it

is much stronger and no common finite-dimensional, or functional depth satisfies this property (Section 2.3);

- None of the investigated functional depths satisfies Property P-5. In fact, that property seems to be void in the functional setting (Section 2.5);
- Property P-6 is, arguably, rather weak in its formulation, and in an already rich body of literature a stronger version of it is studied for functional depths. Moreover, P-6 is known not to be satisfied for some functional depths considered in the exposition (Section 2.6).

In view of these findings, we provide an extensive update of the survey presented in Nieto-Reyes and Battey (2016), including a revision of some conclusions. These amended results are presented in Table 1 below. The majority of the new results are in the red bold rows, as well as in the last four columns. For results that are new, proofs are provided in the on-line Supplementary Material document that accompanies this paper. That document also contains several examples that illustrate the points made.

In addition, in Section 2.1 we provide a detailed elaboration on various invariance properties of depths in functional spaces. A novel and appealing, general definition of symmetry for random functions, and Banach space-valued random variables, is presented in Section 2.2.

From the contributions in this paper it be can concluded that: (a) it is possible to achieve a consent on what the desirable properties of functional depths should be; (b) a great caution should be taken when making generalizations to a functional setting, and deriving theoretical results for this.

2. Discussion on statistical depth properties

For convenience of the readers we use the same notations as in Nieto-Reyes and Battey (2016). In particular, we denote (see that paper for formal definitions)

D_h = the h-depth (Cuevas et al., 2007);
D_{RT} = the random Tukey depth (Cuesta-Albertos and Nieto-Reyes, 2008);
D_J = the band depth (López-Pintado and Romo, 2009);
D_{MJ} = the modified band depth (López-Pintado and Romo, 2009);
D_{HR} = the halfregion depth (López-Pintado and Romo, 2011);
D_{MHR} = the modified halfregion depth (López-Pintado and Romo, 2011).

In addition to these depths, included in the initial survey of Nieto-Reyes and Battey (2016), we consider four additional important depth functionals:

D_T = (the functional version of) the Tukey depth (Dutta et al., 2011);
D_S = the spatial depth (Chakraborty and Chaudhuri, 2014b);
D_{L^∞} = the L^∞ depth (Long and Huang, 2016);
D_I = the infimal depth (or Φ-depth) (Mosler, 2013).

For definitions of these functional depths see Appendix A.

We will often consider finite-dimensional Euclidean spaces $\mathbb{R}^p$ as special cases of the functional space $\mathfrak{F}$. This is justified by restricting from $\mathfrak{F}$ to its p-dimensional linear subspace spanned by a sequence of linearly independent (and possibly orthogonal) functions $b_1, \ldots, b_p \in \mathfrak{F}$. A point $x = (x_1, \ldots, x_p) \in \mathbb{R}^p$ can then be identified with a function $\sum_{j=1}^p x_j b_j \in \mathfrak{F}$, and observations made in $\mathbb{R}^p$ can be translated into an appropriate subspace of $\mathfrak{F}$.

TABLE 1. *A revised, expanded table of results summarizing which of the theoretical properties P-1–P-6 from Nieto-Reyes and Battey (2016) are satisfied (indicated by ✓) or not (indicated by ✗) for the considered functional depths (standard rows). In red bold rows we indicate adherence of the depths to the properties P-0–P-6U discussed in the present paper.*

Properties		Depth Functionals									
		D_h	D_{RT}	D_J	D_{MJ}	D_{HR}	D_{MHR}	D_T	D_S	D_{L^∞}	D_I
Functional space	$\mathfrak{F}$	$\mathbb{L}_2$	$\mathbb{L}_2$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathcal{C}$	$\mathbb{L}_2$	$\mathbb{L}_2$	$\mathcal{C}$	$\mathcal{C}$
Non-degeneracy	**P-0**	✓	✓	✗	✓	✗	✓	✗	✓	✓[1]	✗
Invariance	P-1	✓[2]	✗	✗	✗	✗	✗	✗	✓[3]	✓[2]	✗
	P-1S	✓[2]	✓	✓	✓	✓	✓	✓	✓	✓[2]	✓
	P-1F	✗	✗	✓	✓	✓	✓	✗	✗	✗	✓
Maximality at centre	P-2G	✓	✓	✗	✓	✗	✓	✓	✓	✓	✗
	P-2C	✗	✗	✗	✓	✗	✗	✓	✓	✗	✗
	P-2H	✗	✗	✗	✓	✗	✗	✓	✗	✗	✗
Decreasing w.r.t. deepest point	P-3	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
	P-3D	✗	✓	✗	✓	✗	✗	✗	✗	✓[1]	✗
Vanishing at infinity	**P-3V**	✓	✗	✓	✗	✓	✗	✓	✓	✓	✓
Semi-continuity in x	P-4	✓	✓	✓	✓	✓	✓	✓	✓[4]	✓	✓
Domain receptivity	P-5	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Consistency / Robustness	P-6	✓	✓[5]	✗	✓[6]	✗	✓[6]	✗	✓[7]	✓[8]	✓[9]
	P-6U	✓	✓[5]	✗	✓[6]	✗	✓[6]	✗	✓[7]	✓[8]	✓[9]

[1] Only if $\mathbb{E}\, \|X\|_\infty < \infty$ for $X \sim P$.
[2] Only for $a_f = 1$ and $a = 1$ in P-1 and P-1S, respectively.
[3] Only for f surjective in P-1.
[4] For P non-atomic.
[5] For P such that the joint distribution $\left(P_{\langle u_1, X\rangle}, \dots, P_{\langle u_k, X\rangle}\right)$ for $\{u_1, \dots, u_k\} = \mathfrak{U}$ (see Nieto-Reyes and Battey (2016, Section 4.1.2)) is absolutely continuous, or for the sequence of empirical measures.
[6] For $X \sim P$ such that $X(v)$ has no atoms for each $v \in \mathcal{V}$, or for the sequence of empirical measures.
[7] Uniformly over compact sets in $\mathfrak{F}$, for the sequence of empirical measures, and for $X \sim P$ such that $\sup_{\|x\|_{\mathbb{L}_2} \leq C} \mathbb{E}_{P_1} 1/\|x - X\|_{\mathbb{L}_2} < \infty$ for each $C > 0$, where P_1 is the non-atomic part of P.
[8] Uniformly over compact sets in $\mathfrak{F}$, and for the sequence of empirical measures.
[9] For $X \sim P$ such that $X(v)$ has no atoms for each $v \in \mathcal{V}$.

We now discuss the desirable properties P-1–P-6 as stated in Nieto-Reyes and Battey (2016), and explore their relations with ZS-1–ZS-4. In the exposition, the discussion on Property P-4 (*Semi-continuity of the depth*) is omitted, as this is a rather standard assumption, and this property is satisfied for most depth functionals (Nieto-Reyes and Battey, 2016, Theorem 4.6).

To avoid various technical difficulties, before proceeding to the main conditions we start with an additional property that any reasonable depth function must obey.

P-0 *Non-degeneracy.* For any $P \in \mathcal{P}$ we have $\inf_{x \in \mathfrak{F}} D(x, P) < \sup_{x \in \mathfrak{F}} D(x, P)$.

Technically, P-0 is new compared to ZS-1–ZS-4. Although it can be found as part of P-2G in Nieto-Reyes and Battey (2016), in a more restrictive setting of Gaussian processes, see Section 2.2 below, it is important to consider it separately. In finite-dimensional spaces P-0 is trivially true for all recognized depths. In the functional case, this is surprisingly no longer true, and P-0 must be considered (Chakraborty and Chaudhuri, 2014a, Kuelbs and Zinn, 2013, 2015).

Properties P-2, P-3 and P-5 from Nieto-Reyes and Battey (2016) operate with the concept of the depth-median, i.e. a point (function) $x \in \mathfrak{F}$ at which the value of $D(\cdot, P)$ is maximized over $\mathfrak{F}$. This notion obviously lacks meaning if P-0 is not true for P. Thus, in the sequel, whenever a condition requires depth-medians, P-0 is a mandatory preliminary imposed on D.

2.1. **P-1: Distance invariance.** To understand the importance of Property P-1 and its relation to its intended counterpart ZS-1, four different types of mappings between functional spaces must be distinguished:

M1 *Multiple of an isometry:* A mapping $g\colon \mathfrak{F} \to \mathfrak{F}$ is called an isometry on $\mathfrak{F}$ if

$$d(x, y) = d(g(x), g(y)) \qquad \text{for } x, y \in \mathfrak{F}.$$

If a mapping $f\colon \mathfrak{F} \to \mathfrak{F}$ can be written as $f(x) = c\, g(x)$ for some $c \in \mathbb{R} \setminus \{0\}$ for all $x \in \mathfrak{F}$, then f is called a multiple of the isometry g on $\mathfrak{F}$.

M2 *Affine mapping:* A mapping $f\colon \mathfrak{F} \to \mathfrak{F}$ is called affine if

$$f((1-\alpha)x + \alpha y) = (1-\alpha) f(x) + \alpha f(y) \qquad \text{for } x, y \in \mathfrak{F} \text{ and } \alpha \in [0,1].$$

M3 *Function-affine mapping:* A mapping $f\colon \mathfrak{F} \to \mathfrak{F}$ is called function-affine if

$$f(x) = ax + b \qquad \text{for } x, a, b \in \mathfrak{F}, \text{ where } a(v) \neq 0 \text{ for all } v \in \mathcal{V} \text{ and } ax \in \mathfrak{F}.$$

M4 *Scalar-affine mapping:* A mapping $f\colon \mathfrak{F} \to \mathfrak{F}$ is called scalar-affine if

$$f(x) = ax + b \qquad \text{for } x, b \in \mathfrak{F} \text{ and } a \in \mathbb{R} \setminus \{0\}.$$

In P-1, mappings of type M1 are utilized as equivalents of the affine mappings (type M2 for $\mathfrak{F} = \mathbb{R}^p$) used in ZS-1 for finite-dimensional depths. In the proof of Theorem 4.1 of Nieto-Reyes and Battey (2016) one relies on the equivalence, for functional spaces ($\mathbb{L}_2$ and $\mathcal{C}$), between M1 and M4, and consequently the assertion of Theorem 4.1 is only shown for M4 mappings.

A crucial remark, however, is that this equivalence is not valid — the relations between M1–M4 are more complicated. First of all, recall the Mazur-Ulam theorem (see Väisälä, 2003). This result states that any surjective isometry (but also any surjective mapping of type M1) on a normed vector space over $\mathbb{R}$ is of type M2. This somewhat justifies the use of M1 mappings instead of M2 mappings in functional spaces. Nevertheless, it is not true that M2 mappings are the same as M1. To see this take, for simplicity, $\mathfrak{F} = \mathbb{R}^2$ and $f\colon (x_1, x_2) \mapsto (x_1, 2x_2)$. The mapping f is M2, but not M1 (for d the Euclidean metric on $\mathbb{R}^2$). Therefore, imposing P-1 is in fact, already in finite-dimensional spaces, to some extent weaker than ZS-1. Subsequently, this raises the question of what should then be imposed, as reasonable analogue to ZS-1, in the functional data setting.

The assumption of surjectivity in the statement of the Mazur-Ulam theorem cannot be omitted, as not every mapping of type M1 is M2. To see this, consider $\mathfrak{F} = \mathcal{C}$ for $\mathcal{V} = [0,1]$ and a mapping f assigning to $x \in \mathcal{C}$ the function

$$f(x)(v) = \begin{cases} x(2v) & \text{for } v \in [0, 1/2], \\ (2v-1)\,\|x\|_\infty + (2-2v)x(1) & \text{for } v \in (1/2, 1]. \end{cases} \tag{1}$$

The mapping f "shrinks" x from $[0,1]$ to the interval $[0,1/2]$, and extends it linearly to the non-negative value $\|x\|_\infty$ at $v=1$ on the rest of the domain. It is M1 on $\mathcal{C}$, but is not surjective. To see this, take $x, y \in \mathcal{C}$ and write

$$\|f(x)-f(y)\|_\infty = \max\{\|x-y\|_\infty, |x(1)-y(1)|, |\|x\|_\infty - \|y\|_\infty|\} = \|x-y\|_\infty .$$

Here, the first equality holds true by (1) and the fact that for $v > 1/2$ the function $f(x)$ is linear; the second equality follows from the reverse triangle inequality. Also, f is neither M2, M3, nor M4. To see that f is not M2 write

$$\begin{aligned} f\left(x/2+(-3x)/2\right)(1) &= f(-x)(1) = \|-x\|_\infty = \|x\|_\infty \\ &\neq 2\|x\|_\infty = 1/2\|x\|_\infty + 1/2\|-3x\|_\infty = \left(f(x)/2+f(-3x)/2\right)(1), \end{aligned}$$

provided $\|x\|_\infty \neq 0$, i.e. M2 is not satisfied for $\alpha = 1/2$ for any $x \in \mathcal{C}\setminus\{0\}$, and $y=-3x$.

From what we showed, we see that M1 $\not\Longleftrightarrow$ M2. If $\mathfrak{F}$ contains a constant function, then M4 $\Longrightarrow$ M3. For other relations between the four concepts of affinity it is easy to see that (recall that the metric d is defined by the norm $\|\cdot\|$ on $\mathfrak{F}$) M3 $\Longrightarrow$ M2 and M4 $\Longrightarrow$ M1, but M3 $\not\Longrightarrow$ M1. Consequently, the most general way of rephrasing ZS-1 into the functional setup would be imposing invariance of D w.r.t. all M2 mappings. Such a formulation is, however, very intractable, as the structure of M2 mappings remains unclear in functional spaces. Furthermore, it appears that in functional spaces general affine invariance with respect to M2 mappings is hardly a desirable trait. Therefore, in the literature authors often resort to either type M3, or M4 mappings. These were considered, for instance, by López-Pintado and Romo (2009), and by Claeskens et al. (2014), leading to the properties

P-1S *Scalar-affine invariance.* $D\left(f(x), P_{f(X)}\right) = D(x, P_X)$ for any $P_X \in \mathcal{P}$, $x \in \mathfrak{F}$ and f of type M4.

P-1F *Function-affine invariance.* $D\left(f(x), P_{f(X)}\right) = D(x, P_X)$ for any $P_X \in \mathcal{P}$, $x \in \mathfrak{F}$ and f of type M3.

Another sensible alternative to condition P-1 is the requirement of the depth D being invariant with respect to all surjective mappings f from P-1. As can be seen from Table 1 (and its footnotes), with such a restriction, the spatial depth satisfies Property P-1.

To conclude the discussion on P-1, let us demonstrate that indeed none of the considered depths satisfy P-1 in its full generality. Take $P_X \in \mathcal{P}$ for $\mathfrak{F} = \mathcal{C}$ and $\mathcal{V} = [0,1]$ given by $\mathsf{P}(X \equiv 1) = 1/2$, $\mathsf{P}(X \equiv -1) = 1/2$, and $x \equiv 0$. Then, for instance for D_J (for J=2) we have $D_J(x, P_X) = 1/2$, but $D_J\left(f(x), P_{f(X)}\right) = 0$ for f from (1). For depths defined on $\mathbb{L}_2$, more elaborate isometries (see Fleming and Jamison, 2003) lead to the same negative results.

2.2. **P-2 and P-2G: Maximality at centre.** Condition P-2 is indeed a straightforward translation of ZS-2, and P-2G surely should be valid for any reasonable functional depth. However, the commonly-used band depth D_J fails to meet both P-2 and P-2G, as pointed out by Chakraborty and Chaudhuri (2014a, Theorem 4).

One way of specifying a condition like P-2 arises using the standard technique of projections by means of the functionals from the dual space $\mathfrak{F}^*$ of continuous linear mappings $\varphi\colon \mathfrak{F} \to \mathbb{R}$ (see Rudin, 1991). Consider the following, straightforward notion of symmetry for functional data.

Definition. We say that the distribution $P_X \in \mathcal{P}$ in $\mathfrak{F}$ is symmetric around $\theta \in \mathfrak{F}$ if and only if for any $\varphi \in \mathfrak{F}^*$ the distribution of the random variable $\varphi(X)$ is symmetric around $\varphi(\theta)$.

Note that this definition depends on the univariate notion of symmetry employed for the random variable $\varphi(X)$. This may be chosen w.r.t. the practical problem that one solves. Herein, we follow the approach of Serfling (2006), and in the sequel we consider two concepts of symmetry of random vectors:

- central symmetry, and
- halfspace symmetry.

Recall that a distribution $X \sim P$ on $\mathbb{R}^p$ is centrally symmetric around $\theta \in \mathbb{R}^p$ if and only if $X - \theta$ and $-(X - \theta)$ have the same distribution. X is halfspace symmetric around $\theta \in \mathbb{R}^p$ if and only if $P(H) \geq 1/2$ for all closed halfspaces H in $\mathbb{R}^p$ that contain θ. As argued by Serfling (2006), both central and halfspace symmetry of random vectors are very natural concepts. Halfspace symmetry is very broad — all other standardly used notions of multivariate symmetry (central, elliptical or angular) imply halfspace symmetry.

For $\mathfrak{F} = \mathcal{C}$, the previous definition can be applied to the subset of $\mathcal{C}^*$ consisting of the Dirac functionals

$$\varphi_v \colon \mathcal{C} \to \mathbb{R} \colon x \mapsto x(v) \quad \text{for } v \in \mathcal{V}.$$

Symmetry of P_X then implies the symmetry of the vectors of functional values: for any $p = 1, 2, \ldots$ and $v_1, \ldots, v_p \in \mathcal{V}$, the distribution of the random vector $(X(v_1), \ldots, X(v_p))$ must be symmetric around the vector $(\theta(v_1), \ldots, \theta(v_p))$ (as follows from Zuo and Serfling, 2000b, Lemma 2.1 and Theorem 2.4).

For $\mathfrak{F} = \mathbb{L}_2$ (where we denote the inner product of $x, y \in \mathbb{L}_2$ by $\langle x, y\rangle = \int_{\mathcal{V}} x(v)y(v)\,\mathrm{d}\,v$), P_X is symmetric around θ if and only if the random vector $(\langle X, u_1\rangle, \ldots, \langle X, u_p\rangle)$ is symmetric around $(\langle \theta, u_1\rangle, \ldots, \langle \theta, u_p\rangle)$ for all $u_1, \ldots, u_p \in \mathbb{L}_2$ and $p = 1, 2, \ldots$.

Our notion of symmetry for random functions is certainly natural. Any Gaussian process in $\mathcal{C}$ or $\mathbb{L}_2$ is obviously centrally, and halfspace symmetric around its mean function. For centrally symmetric functional distributions, the definition coincides with the usual definition of central symmetry in general spaces, as stated in the following lemma, whose proof is provided in the on-line Supplementary Material document, Section S.1.

Lemma 1. *For a random function $X \sim P$ in $\mathfrak{F} = \mathcal{C}$ or $\mathfrak{F} = \mathbb{L}_2$, X is centrally symmetric around $\theta \in \mathfrak{F}$ if and only if $X - \theta = -(X - \theta)$ in distribution.*

Having at hand a reasonable notion of symmetry for random functions, we can specify two alternatives for P-2, stronger than P-2G.

P-2C *Maximality at centre (central symmetry).* For any centrally symmetric $P \in \mathcal{P}$ we have that $D(\theta, P) = \sup_{x \in \mathfrak{F}} D(x, P)$ if and only if P is centrally symmetric around $\theta \in \mathfrak{F}$.

P-2H *Maximality at centre (halfspace symmetry).* For any halfspace symmetric $P \in \mathcal{P}$ we have that $D(\theta, P) = \sup_{x \in \mathfrak{F}} D(x, P)$ if and only if P is halfspace symmetric around $\theta \in \mathfrak{F}$.

Obviously, P-2H $\Longrightarrow$ P-2C $\Longrightarrow$ P-2G. Note that P-2H and P-2C are formulated in terms of equivalence: the depth D must be maximized at, and only at the centre of symmetry, as opposed to P-2G, where just the "at" part of the statement is required.

The above approach towards functional symmetry in the space $\mathcal{C}$ is adopted in Nagy et al. (2016), where some discussion on uniqueness and other properties of the centre of symmetry function can be found. In that paper, it is also shown that, even though P-2H is a stronger condition than both P-2C and P-2G, several functional depths satisfy it, see also Table 1. Therefore, both P-2C and P-2H constitute attainable refinements of the weak condition P-2G, and represent sensible analogues of ZS-2 to the functional setting.

2.3. **P-3: Strictly decreasing w.r.t. the deepest point.** This condition is proposed as an extension of ZS-3 to the functional data setting. To gain geometric intuition for this concept, consider P and two points $x, z \in \mathfrak{F}$ such that only z maximizes $D(\cdot, P)$ over $\mathfrak{F}$. The set

$$R(x,z) = \{y \in \mathfrak{F} \colon \max\{d(y,z), d(y,x)\} < d(x,z)\}$$

then constitutes the intersection of two balls of fixed diameter $d(x,z)$ centred at x and z, respectively. Property P-3 requires that the depth at any point in $R(x,z) \setminus \{x\}$ attains a depth value strictly larger than $D(x,P)$, regardless of the form of P.

This condition is surely stronger than the straightforward extension of ZS-3 (see P-3D below). However, already for finite-dimensional depths Property P-3 appears to be too strict. To see this, consider $\mathfrak{F} = \mathbb{R}^2$ equipped with the Euclidean norm, and take for D the Tukey depth D_T (Tukey, 1975). This depth, being the prime example of a depth in $\mathbb{R}^p$, satisfies all properties ZS-1–ZS-4. Now, take P to be a centred bivariate Gaussian distribution with independent marginals X_1 and X_2, with $\mathrm{Var}(X_1) = 1$, $\mathrm{Var}(X_2) = 2$, and set $x = (0,1)$. The centre of (elliptical, central and halfspace) symmetry is the unique point $z = (0,0)$ maximizing D_T, and the contours of D_T coincide with the density contours of P. As can be seen in Figure 1, P-3 is violated already in this, simplest non-trivial example, because the set $R(x,z)$ contains points with depth values lower than $D(x,P)$.

In a functional setting, despite Lemma 4.3 in Nieto-Reyes and Battey (2016), the h-depth D_h does not satisfy P-3. For a counterexample, define $P_X \in \mathcal{P}$ for $\mathfrak{F} = \mathbb{L}_2$ and $\mathcal{V} = [0,1]$ so that $\mathsf{P}(X \equiv 0) = 1/2$, $\mathsf{P}(X \equiv 1) = \mathsf{P}(X \equiv -1) = 1/4$ and take $h = 1/4$. Recall the definition of the h-depth: $D_h(x,P) = \mathbb{E}\, K_h\left(\|x - X\|_{\mathbb{L}_2}\right)$, where $h > 0$ is a fixed constant, and $K_h(\cdot) = K(\cdot/h)/h$ with $K(\cdot)$ the Gaussian kernel. Then D_h is maximized at a single function $\mathbb{E}\, X = z \equiv 0$, $D_h(z,P) = 1/2K_h(0) + 1/2K_h(1) \approx 0.798$, but for $x \equiv 1$ we have $D_h(x/2,P) = 3/4K_h(1/2) + 1/4K_h(3/2) \approx 0.162 < D_h(x,P) = 1/4K_h(0) + 1/2K_h(1) + 1/4K_h(2) \approx 0.399$. This is due to the fact that already in $\mathbb{R}^p$ the h-depth (for finite samples equivalent with kernel density estimation) does not satisfy ZS-3. See also Figure 1.

We argue that the inclusion of P-3 to the list of desirable properties of depths is questionable. The following straightforward extension of ZS-3 is a more plausible requirement.

P-3D *Decreasing w.r.t. the deepest point.* For any $P \in \mathcal{P}$ such that $D(z,P) = \sup_{x \in \mathfrak{F}} D(x,P)$ we have that $D(z,P) > \inf_{x \in \mathfrak{F}} D(x,P)$ and $D(x,P) \leq D(z + \alpha(x - z), P)$ holds for all $x \in \mathfrak{F}$ and $\alpha \in [0,1]$.

As can be seen in Table 1, there are functional depths that satisfy P-3D. As such, this condition presents the minimal criterion distinguishing (global) depth functionals as unimodal estimators, from other measures of centrality, such as local depths (Paindaveine and Van Bever, 2013), or (functional) density-like estimators (Ferraty and Vieu, 2006). Geometrically, P-3D means that the upper levels sets of the depth D are star-shaped in $\mathfrak{F}$, relative to the depth-median (see, for instance, Valentine, 1964). In particular, they

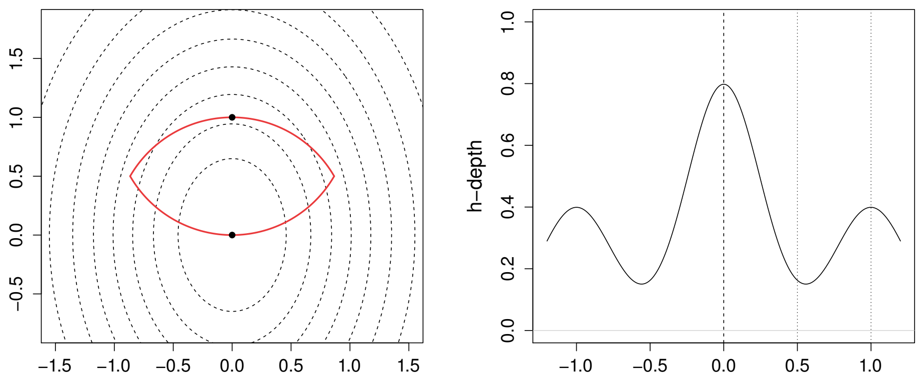

FIGURE 1. *Left: Some contours of D_T of a centred bivariate Gaussian distribution P (dashed), the point maximizing the depth $z = (0,0)$, and $x = (0,1)$. The set $R(x,z)$ is the area inside the solid curve passing through x and z. The points in $R(x,z)$ far from x and z attain lower depth values than $D(x,P)$, meaning that P-3 is not satisfied for D_T. Right: Depth D_h of an atomic distribution P in $\mathbb{R}$ supported in three points $z = 0$, $x_1 = 1$ and $x_2 = -1$. Here, z is the expectation of $X \sim P$, and the sole maximizer of D_h over $\mathbb{R}$. The set $R(x_1,z)$ is the interval (z,x_1). We see that $D_h(1/2,P) < D_h(x_1,P)$, meaning that D_h satisfies neither ZS-3, nor P-3.*

are always connected, and the depth induces a reasonable centre-outwards ordering of the data, as required in applications.

2.4. **Vanishing at infinity.** Unfortunately, the vanishing at infinity property ZS-4 does not follow from P-3, despite (3.2) and Lemma 3.4 in Nieto-Reyes and Battey (2016). To see this, assume for simplicity $\mathfrak{F} = \mathbb{R}$, and for $x \in \mathbb{R}$ and $P \in \mathcal{P}$ define

$$D(x,P) = \begin{cases} 1/(1+x) & \text{for } x \geq 0, \\ 1/(2-x) + 1/2 & \text{for } x < 0. \end{cases} \tag{2}$$

Note that D does not depend on P, though still, quite formally, it is a depth function on $\mathbb{R}$. It satisfies P-3 with $z = 0$, but $\lim_{x\to-\infty} D(x,P) = 1/2 > 0 = \lim_{x\to\infty} D(x,P) = \inf_{x\in\mathfrak{F}} D(x,P)$. Thus, ZS-4 is not guaranteed by P-3 (or P-3D), and needs to be established separately.

P-3V *Vanishing at infinity.* For any $P \in \mathcal{P}$ we have $\lim_{\|x\|\to\infty} D(x,P) = \inf_{x\in\mathfrak{F}} D(x,P)$.

2.5. **P-5: Receptivity to convex hull width across the domain.** Nieto-Reyes and Battey (2016) stated that Property P-5 holds for the h-depth D_h defined on $\mathbb{L}_2$. Surprisingly, it turns out that Condition P-5 is void. To see this, assume that $\mathfrak{F} = \mathbb{L}_2$ for $\mathcal{V} = [0,1]$ and define $P_X \in \mathcal{P}$ by $\mathsf{P}(X \equiv z) = \mathsf{P}(X \equiv -z) = 1/2$, where

$$z(v) = \begin{cases} 1 & \text{for } v \in [0,1/2) \\ 0 & \text{for } v \in [1/2,1]. \end{cases}$$

For $\delta = 0$ we have $L_\delta = [1/2, 1]$. For any function α as in the statement of P-5 and

$$x(v) = \begin{cases} 1/2 & \text{for } v \in [0, 1/2) \\ 0 & \text{for } v \in [1/2, 1], \end{cases}$$

and P-5 implies $D(x, P_X) < D(\alpha x, P_{\alpha X}) = D(x, P_X)$, a contradiction.

In addition, the finite-dimensional counterpart of P-5 conflicts with the basic affine invariance property ZS-1. We verify this by considering P-5 in $\mathfrak{F} = \mathbb{R}^p$ for $p \geq 2$. It can be seen that P-5 follows if, for any $x \in \mathbb{R}^p$, $P_X \in \mathcal{P}$, and a $p \times p$ diagonal matrix A given by its diagonal $(\alpha_1, \ldots, \alpha_p) \in (0, 1]^p$ it is true that $D(x, P_X) < D(Ax, P_{AX})$. This contradicts with ZS-1 that requires equality in the last formula.

By a similar argument, it is easy to see that in purely functional settings P-5 contradicts with P-1F. Moreover, if P-1S is true for a depth D, then by taking $\delta \to d(L, U)$ and $\alpha(v) = 1/2$ for $v \in L_\delta$ in P-5 necessarily $\lim_{\delta \to d(L,U)} D(f(x), P_{f(X)}) = D(x, P_X)$, if D is appropriately (semi-)continuous. Thus, P-5 counteracts with the affine invariance properties of D.

2.6. **P-6: Continuity in P.** It is important to require continuity of D in the distributional argument. However, depth values at particular sample points are usually of little importance. From inference point of view, it is more relevant that the whole set of depth values (the depth surface $\{D(x, P) \colon x \in \mathfrak{F}\}$) is well approximated by its finite sample version. Only this allows the study of depth-medians and sets of depth contours, and enables correct addressing of the problem of distribution-by-depth characterization, necessary for successful construction of nonparametric tests. Therefore, we propose an extension of P-6 to the uniform setting as follows.

P-6U *Uniform continuity in P.* For $\varepsilon > 0$ there exists $\delta > 0$ such that for any $P, Q \in \mathcal{P}$, $d_{\mathcal{P}}(P, Q) < \delta$, it is true that $\sup_{x \in \mathfrak{F}} |D(x, P) - D(x, Q)| < \epsilon$. Here, $d_{\mathcal{P}}$ metricises the topology of weak convergence in $\mathcal{P}$.

Already in $\mathbb{R}^p$ condition P-6U is demanding in full generality. This has to do with the fact that most depths are only semi-continuous for discrete distributions (P-4), and some assumptions on absolute continuity of the measure P must be included. Another complication arising when extending P-6 to P-6U follows from the proposed uniformity aspect, as in functional spaces it is often too strict to assume uniform convergence over the whole space $\mathfrak{F}$. Thus, possible restriction to local uniformity (where convergence is taken over compact sets in $\mathfrak{F}$) might be appropriate in P-6U.

Finally we provide an example demonstrating that P-6 is not valid for most of the depths considered in Nieto-Reyes and Battey (2016), unless appropriate continuity of the distributions is assumed. Take $P \in \mathcal{P}$ concentrated in the constant zero function, and a sequence $\{P_\nu\}_{\nu=1}^{\infty} \subset \mathcal{P}$, where each P_ν is concentrated in the constant $1/\nu$ function. Then P_ν converges weakly to P as $\nu \to \infty$ and $D(0, P) > 0$ for all studied functional depths. Now for any D, with the exception of D_h and D_{L^∞}, $D(0, P_\nu) = 0$ for all $\nu = 1, 2, \ldots$; and P-6 is true only for D_h (for D_{L^∞} see Section S.10 in the Supplementary Material). This disagrees with Theorem 4.8 in Nieto-Reyes and Battey (2016), as we show that P-6 is not true in general for none of the depths D_{MJ}, D_{HR} and D_{MHR}, and also not for D_J and $\mathfrak{F}$ the space of equicontinuous functions. The continuity assumptions that need to be made in order to retrieve P-6 for these depths are indicated in Table 1.

3. Conclusions

In Table 1 we summarize which properties listed in Section 2 can be obtained for the considered depth functionals. Proofs of the results not available in the literature are given in the Supplementary Material document, where also references to known results can be found. Below we comment on overall findings for each of the depths involved in the survey.

As can be seen, the h-depth D_h lacks unimodality. This is a consequence of the fact that D_h takes the form of a more general, kernel estimator for functional data (see Ferraty and Vieu, 2006). On the other hand, D_h pertains excellent continuity and robustness properties. Thus, it appears that D_h is more of a good density-like concept for random functions than a functional depth, because of its localized behaviour.

When interpreting the results for the random Tukey depth D_{RT}, it is necessary to keep in mind that D_{RT} is inherently random, and the depth values for a single distribution P may vary substantially with different sets of projections $\mathfrak{U}$. Furthermore, note that the initial step in the computation of D_{RT} lies in projecting all the elements of the Hilbert space $\mathbb{L}_2$ onto its finite-dimensional subspace spanned by $\mathfrak{U}$. This makes D_{RT} a random finite-dimensional depth, rather than a functional depth. Furthermore, D_T as the infinite-dimensional analogue of D_{RT} (obtained by drawing an infinite number of projections), can be seen to be burdened with substantial theoretical drawbacks. All this makes it very difficult to compare the properties of D_{RT} with those of other, truly infinite-dimensional depth functionals.

The band depth D_J, the halfregion depth D_{HR}, and the infimal depth D_I share very similar features. For complex datasets they tend to degenerate, and they exhibit difficulties with their robustness, and consistency properties. The modified counterparts of D_J and D_{HR} appear to be better behaved. While the use of D_{MHR} is still limited by contrived theoretical issues (see also Kuelbs and Zinn, 2015), D_{MJ} satisfies all the (newly) imposed conditions except for P-3V, vanishing at infinity. Nevertheless, D_{MJ} is only one representative of a larger class of functionals called integrated depths (Fraiman and Muniz, 2001, Cuevas and Fraiman, 2009, Claeskens et al., 2014), and all the properties proved here are true not only for D_{MJ}, but rather simultaneously for the whole class of depth functionals of integrated type (Nagy et al., 2016), under mild conditions.

The frequently overlooked depths D_S and D_{L^∞} also prove to be highly competitive, in comparison to the established depths. Additional research into the theory and practice of these functional depths might be beneficial.

Overall, it can be concluded that, apart from D_S and D_{L^∞}, some integrated depths (and D_{MJ}) appear superior to other studied functional depth concepts. This claim can be supported also from the empirical point of view by observing the great inclination of practitioners performing data analysis using depths of the integrated type. However, a proper choice of a functional depth from the family of integrated depths still poses an interesting problem worth investigation.

Appendix A. Definitions of newly included functional depths

A.1. **Tukey depth.** For $\mathfrak{F} = \mathbb{L}_2$, (the functional version of) the Tukey depth (Dutta et al., 2011) of $x \in \mathfrak{F}$ w.r.t. $P \in \mathcal{P}$, $X \sim P$ is defined as

$$D_T(x, P) = \inf_{u \in \mathfrak{F}} \left(\min \left\{ P_{\langle u, X \rangle}\left((-\infty, \langle u, x \rangle] \right), P_{\langle u, X \rangle}\left([\langle u, x \rangle, \infty) \right) \right\} \right).$$

A.2. **Spatial depth.** For $\mathfrak{F} = \mathbb{L}_2$, the spatial depth (Chakraborty and Chaudhuri, 2014b) of $x \in \mathfrak{F}$ w.r.t. $P \in \mathcal{P}$, $X \sim P$ is defined as

$$D_S(x, P) = 1 - \left\| \mathbb{E} \frac{x - X}{\|x - X\|_{\mathbb{L}_2}} \right\|_{\mathbb{L}_2}.$$

The expectation of the $\mathbb{L}_2$-valued random variable in the definition is meant in the Bochner sense. In the expectation, the convention $0/0 = 0$ is used.

A.3. L^∞ **depth.** For $\mathfrak{F} = \mathcal{C}$, the L^∞ depth (Long and Huang, 2016) of $x \in \mathfrak{F}$ w.r.t. $P \in \mathcal{P}$, $X \sim P$ is defined as

$$D_{L^\infty}(x, P) = \left(1 + \mathbb{E}\, \|x - X\|_\infty\right)^{-1}.$$

In this definition we adhere to the convention $1/\infty = 0$.

A.4. **Infimal depth.** For $\mathfrak{F} = \mathcal{C}$, the infimal depth (or Φ-depth) (Mosler, 2013) of $x \in \mathfrak{F}$ w.r.t. $P \in \mathcal{P}$, $X \sim P$ is defined as

$$D_I(x, P) = \inf_{v \in \mathcal{V}} \min \left\{ P_{X(v)}\left((-\infty, x(v)]\right), P_{X(v)}\left([x(v), \infty)\right) \right\}.$$

Acknowledgement. Helpful remarks of three anonymous referees on an earlier version of the manuscript are greatly appreciated. We are grateful to Daniel Hlubinka for reading the manuscript, and providing useful comments on it. This research is supported by the IAP research network no. P7/06 of the Federal Science Policy (Belgium). The first author gratefully acknowledges support from the GOA/12/014-project of the Research Fund KU Leuven. The second author is a Research Assistant of the Research Foundation - Flanders, and acknowledges support from this foundation.

References

Chakraborty, A. and Chaudhuri, P. (2014a). On data depth in infinite dimensional spaces. *Ann. Inst. Statist. Math.*, 66(2):303–324.

Chakraborty, A. and Chaudhuri, P. (2014b). The spatial distribution in infinite dimensional spaces and related quantiles and depths. *Ann. Statist.*, 42(3):1203–1231.

Claeskens, G., Hubert, M., Slaets, L., and Vakili, K. (2014). Multivariate functional halfspace depth. *J. Amer. Statist. Assoc.*, 109(505):411–423.

Cuesta-Albertos, J. A. and Nieto-Reyes, A. (2008). The random Tukey depth. *Comput. Statist. Data Anal.*, 52(11):4979–4988.

Cuevas, A., Febrero, M., and Fraiman, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Comput. Statist.*, 22(3):481–496.

Cuevas, A. and Fraiman, R. (2009). On depth measures and dual statistics. A methodology for dealing with general data. *J. Multivariate Anal.*, 100(4):753–766.

Dutta, S., Ghosh, A. K., and Chaudhuri, P. (2011). Some intriguing properties of Tukey's half-space depth. *Bernoulli*, 17(4):1420–1434.

Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis. Theory and practice.* Springer Series in Statistics. Springer, New York.

Fleming, R. J. and Jamison, J. E. (2003). *Isometries on Banach spaces: function spaces*, volume 129 of *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*. Chapman & Hall/CRC, Boca Raton, FL.

Fraiman, R. and Muniz, G. (2001). Trimmed means for functional data. *Test*, 10(2):419–440.

Kuelbs, J. and Zinn, J. (2013). Concerns with functional depth. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(2):831–855.

Kuelbs, J. and Zinn, J. (2015). Half-region depth for stochastic processes. *J. Multivariate Anal.*, 142:86–105.

Long, J. P. and Huang, J. Z. (2016). A study of functional depths. *arXiv preprint arXiv:1506.01332*.

López-Pintado, S. and Romo, J. (2009). On the concept of depth for functional data. *J. Amer. Statist. Assoc.*, 104(486):718–734.

López-Pintado, S. and Romo, J. (2011). A half-region depth for functional data. *Comput. Statist. Data Anal.*, 55(4):1679–1695.

Mosler, K. (2013). Depth statistics. In *Robustness and complex data structures*, pages 17–34. Springer, Heidelberg.

Nagy, S., Gijbels, I., Omelka, M., and Hlubinka, D. (2016). Integrated depth for functional data: statistical properties and consistency. *ESAIM Probab. Stat.*, 20:95–130.

Nieto-Reyes, A. and Battey, H. (2016). A topologically valid definition of depth for functional data. *Statist. Sci.*, 31(1):61–79.

Paindaveine, D. and Van Bever, G. (2013). From depth to local depth: A focus on centrality. *J. Amer. Statist. Assoc.*, 108(503):1105–1119.

Rudin, W. (1991). *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition.

Serfling, R. (2006). Multivariate symmetry and asymmetry. *Encyclopedia of Statistical Sciences, Second Edition*, 8:5338–5345.

Tukey, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 523–531. Canad. Math. Congress, Montreal, Que.

Väisälä, J. (2003). A proof of the Mazur-Ulam theorem. *Amer. Math. Monthly*, 110(7):633–635.

Valentine, F. A. (1964). *Convex sets*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Toronto-London.

Zuo, Y. and Serfling, R. (2000a). General notions of statistical depth function. *Ann. Statist.*, 28(2):461–482.

Zuo, Y. and Serfling, R. (2000b). On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry. *J. Stat. Plan. Inference*, 84(1-2):55–79.