ESAIM: PS 20 (2016) 95–130
DOI: 10.1051/ps/2016005

ESAIM: Probability and Statistics
www.esaim-ps.org

INTEGRATED DEPTH FOR FUNCTIONAL DATA: STATISTICAL PROPERTIES AND CONSISTENCY

Stanislav Nagy[1,2], Irène Gijbels[1], Marek Omelka[2] and Daniel Hlubinka[2]

Abstract. Several depths suitable for infinite-dimensional functional data that are available in the literature are of the form of an integral of a finite-dimensional depth function. These functionals are characterized by projecting functions into low-dimensional spaces, taking finite-dimensional depths of the projected quantities, and finally integrating these projected marginal depths over a preset collection of projections. In this paper, a general class of integrated depths for functions is considered. Several depths for functional data proposed in the literature during the last decades are members of this general class. A comprehensive study of its most important theoretical properties, including measurability and consistency, is given. It is shown that many, but not all, properties of the integrated depth are shared with the finite-dimensional depth that constitutes its building block. Some pending measurability issues connected with all integrated depth functionals are resolved, a broad new notion of symmetry for functional data is proposed, and difficulties with respect to consistency results are identified. A general universal consistency result for the sample depth version, and for the generalized median, for integrated depth for functions is derived.

Mathematics Subject Classification. 62G20, 62H99, 60F15.

Received December 22, 2014. Accepted February 24, 2016.

1. Introduction

Nonparametric inference is often based on the notions of ordering, quantiles, or ranking of real-valued observations. These concepts cannot be generalized straightforwardly to multivariate settings. One option for introducing ranking to multivariate data is to utilize the concept of statistical depth function. Originally designed for finite-dimensional random vectors, the method assigns to a given point in the sample space and a probability measure a single non-negative number, the depth of the point, characterizing how much the observation is central with respect to the probability. High depth values mean centrality, low depth values indicate potential outlyingness. Likewise, the point of the sample space at which the highest depth value is attained may be considered as the most central point. This, for example, leads to a generalization of median to finite-dimensional measures. For a general survey on finite-dimensional depths see [22] or [39].

Keywords and phrases. Center of symmetry, functional data, generalized median, integrated depth, measurability, strong consistency, weak consistency.

[1] KU Leuven, Department of Mathematics and Leuven Statistics Research Center (LStat), Celestijnenlaan 200b, box 2400, 3001 Leuven, Belgium.

[2] Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Math. Statistics, Sokolovská 83, Praha 8, 186 75 Czech Republic. hlubinka@karlin.mff.cuni.cz

Article published by EDP Sciences

Since the pioneering paper of Fraiman and Muniz [14], introducing the first depth designed for infinite-dimensional functional data, a great deal of attention has been on statistical depth methodology in connection with spaces of functions as well. The existing approaches towards the assignment of a depth value to a random function can be categorized into two distinct families:

- integrated depth, and
- non-integrated, or geometric depth.

The first family involves taking integrals over given collections of depths of low-dimensional projections of functions. For instance, let x be the function whose depth is to be determined with respect to a given probability on the space of functions. It is typically the functional value of x at t in its domain, that is taken to constitute the elementary one-dimensional projection of x. Then, a univariate depth of $x(t)$ can be computed with respect to the corresponding marginal distribution of the given probability. Equipped with the collection of one-dimensional depths of these projections for all t, one computes an "average" of them as the integrated depth of x, taking the (Lebesgue) integral with respect to t over the common domain of functions.

Fraiman and Muniz [14] proceeded in this way to define their basic notion of integrated depth for scalar-valued functions. Since then, numerous authors generalized or modified this original approach to obtain depth functionals more suitable either theoretically, or from the practical point of view. For example, Cuevas and Fraiman [4] broadened the integrated depth notion to encompass general Banach space-valued random variables, and Claeskens *et al.* [2] elaborated a depth of integrated type tailored for functions with vector-valued response. The modified versions of band depth [23], half-region depth [24], and simplicial band depth [25] also fall logically into the setup of integrated depths.

The second approach of non-integrated type depths is comprehensively covered by the recent notion of Φ-depths introduced by Mosler [30], see also Mosler and Polyakova [31]. Again, in the first step, finite-dimensional depths of low-dimensional functionals of all the quantities involved are computed. The main difference to the first approach is that the "averaging" of low-dimensional depths is now replaced by using the infimum of these values. Since there is no longer an integral involved, we refer to this type of depths as the non-integrated type depths. This class includes the band depth (see [23]), half-region depth (*cf.* [24]), simplicial band depth [25], or functional halfspace depth of Dutta *et al.* [12]. The random Tukey depth of Cuesta-Albertos and Nieto-Reyes [3] is a finite approximation of the functional halfspace depth, and hence is not, strictly speaking, an infinite-dimensional depth. Although none of these functional depths fall directly into the framework of Φ-depths, they are very similar intrinsically and can well be studied simultaneously. See, for instance, [19] and [15] for a discussion on their consistency properties.

In the present article we focus on the first, more popular, approach of integrated depth for functions. The unifying framework of a general family of integrated depths for functional data encompasses all depths of this kind available in the literature. Therefore our study provides a general theoretical treatment of all depths of such type. Some theoretical properties of certain depths can be found in the literature. A thorough theoretical treatment of the class of all integrated depths is however lacking. This is precisely where this paper contributes. The main contributions of this paper consist of

(1) investigating measurability issues of integrated depths;
(2) studying which desirable properties of depth functions hold for integrated depths;
(3) investigating which conditions are needed to ensure weak/strong consistency of the sample versions of integrated depths;
(4) illustrating clearly with well-chosen examples the intrinsic problems with respect to measurability issues, to failure of desirable properties to hold, and to consistency of existing depths for functional data.

Due to the unifying framework our study treats this at once, for all integrated depths. For certain depths we as such also recover results available in the literature, as special cases.

It is immediate that theoretical properties of all integrated depths depend closely on the properties of the finite-dimensional depth used in their construction. In the first section we briefly discuss some finite-dimensional depths and recall the desirable properties that a good finite-dimensional depth should, at least partially, possess (see Sect. 2.1). Then, in Section 2.2 we introduce the most relevant integrated depth functionals known from the literature, and outline the connections between them and the integrated depth as is conceived here. In the next sections we then subsequently deal with items (1), (2) and (3) above, providing also illustrative examples (4).

In Section 3 we address the question of measurability of the integrand function. So far, this measurability issue got very little attention in the literature. We show that under an additional measurability condition imposed on the finite-dimensional evaluating depth, it is possible to guarantee measurability of an integrated data depth. In Section 4 we proceed by studying which desirable properties, imposed on the finite-dimensional depth, are also inherited by its integrated extension for functional data. As a byproduct of this study, we introduce, in Section 4.2, a new general notion of symmetry suitable especially to treat vector-valued functions from the viewpoint of integrated depth functional. The second main part of the paper is Section 5, with a detailed study of consistency results for integrated depth functionals. Our contribution consists of establishing a very general form of (weak and strong) consistency for integrated depths. As corollaries of these results we obtain, in particular, consistency results for the depth functionals of Fraiman and Muniz [14], López-Pintado and Romo [23, 24], López-Pintado *et al.* [25] and Claeskens *et al.* [2]. Moreover, in Section 5.3, the consistency of the depth-based generalized median function is shown, for all above mentioned depth functionals. As usual in the terminology of depth functions, by a (depth-based) generalized median function we understand the function at which the depth (nearly) attains its maximal value over the whole sample space. Finally, in Section A in the Appendix, we apply the developed theory to some of the most important finite-dimensional depth functions, establishing as such basic probabilistic and statistical properties of various integrated depths. Appendix B contains the Proof of Theorem 3.1, the main result on measurability properties for integrated functional depth. Finally, in Appendix C we point out several difficulties in the available proofs of strong consistency of integrated depth for functional data.

2. Finite-dimensional data depths and integrated data depths

2.1. Finite-dimensional data depth

Throughout the paper the following notation is adopted. Let Ω be the sample space and $\mathcal{F}$ a σ-field on Ω, constituting a measurable space $(\Omega, \mathcal{F})$. Denote by P a probability measure on this space, leading to the probability space $(\Omega, \mathcal{F}, \mathsf{P})$ on which all the random variables are defined.

For an arbitrary measurable space S equipped with some fixed σ-algebra of subsets, $\mathcal{P}(S)$ stands for the collection of all probability measures on S. Further, $\mathrm{P}_n \in \mathcal{P}(S)$ stands for the probability measure of the empirical type based on (independent) observations $X_1, \ldots, X_n$ from $\mathrm{P} \in \mathcal{P}(S)$.

Let D be any function

$$D\colon \begin{array}{ll} S \times \mathcal{P}(S) \to [0,1] \\ (s, \mathrm{P}) \qquad \mapsto D(s; \mathrm{P}) \,. \end{array} \tag{2.1}$$

The function D will be used for evaluating the centrality of a point $s \in S$ with respect to a distribution $\mathrm{P} \in \mathcal{P}(S)$ – higher values of D should indicate that s is located centrally with respect to P, lower values stand for its potential outlyingness. Throughout the paper, the function D will be called (generic) depth function.

In reference [39], a statistical depth function is defined as D in (2.1) possessing certain properties (namely (D_1), (D_2), $(\mathrm{D}_3^{\mathrm{W}})$ and (D_4) listed below). However, in our exposition, we a priori do not assume any properties of D as we aim to discuss, in general, the properties of various integrated depth functionals resulting from a given depth function D.

Among the well-known examples of depths are Tukey's halfspace depth of Tukey [37] and the simplicial depth of Liu [21]. For $\mathrm{Q} \in \mathcal{P}(\mathbb{R}^K)$ and $u \in \mathbb{R}^K$, Tukey's halfspace depth is defined as

$$hD\big(u; \mathrm{Q}\big) = \inf_{H \in \mathcal{H}\colon u \in H} \mathrm{Q}(H), \tag{2.2}$$

where $\mathcal{H}$ is the class of all closed halfspaces on $\mathbb{R}^K$. The simplicial depth of Liu [21] is defined as

$$sD\big(u;\mathrm{Q}\big) = \mathsf{P}\big(u \in \overline{co}(U_1, \ldots, U_{K+1})\big), \tag{2.3}$$

where $U_1, \ldots, U_{K+1}$ are independent identically distributed random variables with distribution Q and $\overline{co}(u_1, \ldots, u_{k+1})$ is a closed convex hull whose vertices are $u_1, \ldots, u_{k+1}$.

Let us now discuss the desirable properties of D. Some of these are standardly imposed on a statistical depth function (*cf.* [31, 35, 39]) and some of them are new (more or less technical ones). These new properties will be useful to show the main results for the corresponding integrated depths in Sections 4 and 5.

In what follows let $\mathrm{Q} \in \mathcal{P}\left(\mathbb{R}^K\right)$, U be a random vector with values in $\mathbb{R}^K$ and Q_U the probability measure associated to the vector U. We first recall the definition of halfspace symmetry, the current most general notion of symmetry for finite-dimensional distributions (see Sect. 2.1.3 of [40]).

Definition 2.1. We say that a distribution $\mathrm{Q} \in \mathcal{P}\left(\mathbb{R}^K\right)$ is *halfspace symmetric* around $u^* \in \mathbb{R}^K$, if

$$\mathrm{Q}\,(H) \geq 0.5 \text{ for each closed halfspace } H \text{ with } u^* \text{ on the boundary.}$$

Each one-dimensional distribution $\mathrm{Q} \in \mathcal{P}\,(\mathbb{R})$ having a median $u^* \in \mathbb{R}$ is halfspace symmetric around u^*. Notice, however, that especially in one dimension, the center of halfspace symmetry does not need to be unique if the support of Q is not a connected support.

We are now ready to list the desired properties of a depth function.

(D_1) *Affine invariance:* For any non-singular matrix $A \in \mathbb{R}^{K\times K}$, any $b, u \in \mathbb{R}^K$ and U

$$D\big(u;\mathrm{Q}_U\big) = D\big(Au + b;\mathrm{Q}_{AU+b}\big).$$

(D_2) *Maximality at center:* If the distribution of U is halfspace symmetric around $u^* \in \mathbb{R}^K$, then u^* maximizes $D(u;\mathrm{Q}_U)$.

(D_3) *Quasi-concavity as a function of u:* The level set $\big\{u: D(u;\mathrm{Q}) \geq c\big\}$ is convex for each $c \in \mathbb{R}$.

(D_4) *Vanishing at infinity:* $D(u;\mathrm{Q}) \xrightarrow[\|u\|_{\mathbb{R}^K} \to \infty]{} 0$, where $\|u\|_{\mathbb{R}^K}$ stands for the Euclidean norm on $\mathbb{R}^K$.

(D_5) *Upper semicontinuity of D as a function of u:* For all $u \in \mathbb{R}^K$ and $u_\nu \xrightarrow[\nu\to\infty]{} u$

$$\limsup_{\nu\to\infty} D\left(u_\nu;\mathrm{Q}\right) \leq D\left(u;\mathrm{Q}\right).$$

(D_6) *Weak continuity of D as a functional of* Q*:* For all $\mathrm{Q}_\nu \xrightarrow[\nu\to\infty]{w} \mathrm{Q}$

$$\sup_{u\in\mathbb{R}^K} \left|D\left(u;\mathrm{Q}_\nu\right) - D\left(u;\mathrm{Q}\right)\right| \xrightarrow[\nu\to\infty]{} 0.$$

Note that while it is common to require that (D_1)–(D_5) hold for each $\mathrm{Q} \in \mathcal{P}\left(\mathbb{R}^K\right)$, such a requirement would be too strict for (D_6). Thus in what follows we always specify for which Q we assume (D_6) to hold.

Condition (D_1) can be decomposed into two separate items, as discussed in Mosler and Polyakova [31]. These two items together are equivalent with (D_1).

($\mathrm{D}_1^{\mathrm{A}}$) *Translation invariance:* For any $b, u \in \mathbb{R}^K$ and U

$$D\big(u;\mathrm{Q}_U\big) = D\big(u + b;\mathrm{Q}_{U+b}\big).$$

($\mathrm{D}_1^{\mathrm{B}}$) *Linear invariance:* For any non-singular matrix $A \in \mathbb{R}^{K\times K}$, $u \in \mathbb{R}^K$ and U

$$D\big(u;\mathrm{Q}_U\big) = D\big(Au;\mathrm{Q}_{AU}\big).$$

In (D_2) we could equally well utilize any other notion of symmetry for multivariate data, *e.g.* angular symmetry, central symmetry or elliptical symmetry. One reason why we restricted ourselves to halfspace symmetry is the level of generality provided by the use of halfspaces. Indeed, as Zuo and Serfling [40] discuss, each of the previously mentioned symmetry notions implies halfspace symmetry. Moreover, halfspace symmetry arises very naturally when dealing with depth functions; the center of halfspace symmetry coincides with the point(s) with the highest halfspace depth value hD, and hD is an important representative of finite-dimensional depths.

Condition (D_3) is sometimes in the literature replaced by the weaker

(D_3^W) *Decreasing along rays:* If a maximum depth D is attained at $u^* \in \mathbb{R}^K$, then for every $u \in \mathbb{R}^K$ and $\gamma \in [0,1]$

$$D\big(\gamma u^* + (1-\gamma)u; \mathrm{Q}\big) \geq D\big(u; \mathrm{Q}\big).$$

In Appendix A, a comprehensive summary of theoretical properties (D_1)–(D_6) for the primal representatives of finite-dimensional depths that are available in the literature is given.

2.2. Integrated data depth for functional data

Suppose the observations are independent and identically distributed vector stochastic processes $X_1(t), \ldots, X_n(t)$, *i.e.* X_i is a random variable taking values in

$$\mathcal{C}^K([a,b]) = \left\{ x = (x^{(1)}, \ldots, x^{(K)})^\intercal : x^{(k)} \in \mathcal{C}\big([a,b]\big),\ k = 1, \ldots, K \right\},$$

where $\mathcal{C}\big([a,b]\big)$ stands for the space of continuous functions on $[a,b]$. Without loss of generality we assume in the following that $[a,b] = [0,1]$ and we equip the space $\mathcal{C}^K([0,1])$ with the norm

$$\|x\| = \sum_{k=1}^{K} \sup_{t\in[0,1]} \left| x^{(k)}(t) \right|. \tag{2.4}$$

Further, let $\mathcal{B}^K([0,1])$ stand for the space of all Borel measurable functions $x\colon [0,1] \to \mathbb{R}^K$.

Note that the reason for considering the univariate domain $[0,1]$ is merely notational convenience, as this is a standard choice of the domain in functional data analysis [33]. However, any compact and convex domain T in a Euclidean space could be used in the sequel as well, without any further complications.

For a probability distribution $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and a random function $X \sim \mathrm{P}$, the marginal probability of $X(t) \in \mathbb{R}^K$ for $t \in [0,1]$ will be designated by $\mathrm{P}_t \in \mathcal{P}\left(\mathbb{R}^K\right)$. Furthermore, to avoid any confusion, we use a double subscript for the marginal distribution of a random sample of size $n \in \mathbb{N}$ from P at point $t \in [0,1]$, and denote it by $\mathrm{P}_{n,t} \in \mathcal{P}\left(\mathbb{R}^K\right)$.

For the rest of the paper let D be a fixed, arbitrary depth function on $\mathbb{R}^K$ as in (2.1). At this point we do not presume that any other property listed in Section 2.1 holds. Each time a property is needed, it will specifically be mentioned in the text.

A very general depth for vector-valued functional data was recently proposed by Claeskens *et al.* [2] and López-Pintado *et al.* [25].

Definition 2.2. Let $w\colon [0,1] \to [0,\infty)$ be a weight function that integrates to one over its domain. For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and $x \in \mathcal{C}^K([0,1])$, the multivariate functional depth of x with respect to P is defined as

$$MFD\left(x; \mathrm{P}, D\right) = \int_0^1 D\left(x(t); \mathrm{P}_t\right) w(t)\,\mathrm{d}t. \tag{2.5}$$

In the above definition the weight function w can be an arbitrary non-negative function integrating to one and it may depend on P_t. In particular, in reference [2] w is considered to be either constant, or a function related to the dispersion of the marginal distribution P_t for $t \in [0,1]$. The latter approach however turns out to be difficult to deal with theoretically, as the estimation of the dispersion in the sample case introduces an additional source of randomness into the integral (2.5) (see also Sect. C.2 for some additional discussion). For these reasons we restrict ourselves in what follows to a particular type of MFD with the weight function w to be predefined and deterministic – that is, we assume that w does not depend on P. This choice still enables the incorporation of prior information into (2.5) (*e.g.* by relating w to the population dispersion function), or the introduction of weighting to the domain. Under these assumptions, w can be conceived as a density of a particular measure μ on $[0,1]$, which is absolutely continuous with respect to the Lebesgue measure on $[0,1]$, and can be incorporated into the differential part of the integral. More generally, let μ be an arbitrary probability measure on $[0,1]$, not necessarily absolutely continuous, and define the integrated functional depth as follows.

Definition 2.3. For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and $x \in \mathcal{C}^K([0,1])$, the integrated depth of x with respect to P is defined as

$$FD\left(x;\mathrm{P},D\right) = \int_0^1 D\left(x(t);\mathrm{P}_t\right)\,\mathrm{d}\mu(t). \tag{2.6}$$

Remark 2.4. For the sample version of the integrated depth FD, it is assumed that the random sample functions $X_1,\ldots,X_n \in \mathcal{C}^K([0,1])$ from P are observed completely, that is the whole trajectory of each X_i, $i=1,\ldots,n$ is known. Then, the sample version of FD is defined as a mere FD with respect to the empirical measure P_n of the random sample of functions as

$$FD\left(x;\mathrm{P}_n,D\right) = \int_0^1 D\left(x(t);\mathrm{P}_{n,t}\right)\,\mathrm{d}\mu(t).$$

Claeskens *et al.* [2] work in the framework that the functions of the random sample $X_1,\ldots,X_n$ are observed discretely on a predefined grid of points $0=t_1<t_2<\ldots<t_T=1$ for $T\in\mathbb{N}$, instead of on the whole interval $[0,1]$. Then, for the observed n-by-T matrix of K-dimensional observations $(X_i(t_j))_{\substack{i=1,\ldots,n,\\ j=1,\ldots,T}}$ they define the sample version of MFD, with $t_0=t_1$ and $t_{T+1}=t_T$, as

$$MFD_n\left(x;D\right) = \sum_{j=1}^{T} D\left(x(t_j);\mathrm{P}_{n,t_j}\right)\int_{(t_{j-1}+t_j)/2}^{(t_j+t_{j+1})/2} w(t)\,\mathrm{d}t. \tag{2.7}$$

The consistency results of MFD_n are then meant for $n\to\infty$ and $T\to\infty$ at the same time. See Section C.2 for a further discussion.

A number of other depth concepts proposed in the literature fall directly into the framework of integrated depth. All of these share the common property that the measure μ is the Lebesgue measure on $[0,1]$. The modified simplicial band depth of López-Pintado *et al.* [25] coincides with the general integrated depth for the choice $D=sD$ (see (2.3)). Other integrated depths have been designed specially for $K=1$. We hereby refer to the work of Fraiman and Muniz [14] for the original integrated depth for functions, López-Pintado and Romo [23] for the modified band depth, and López-Pintado and Romo [24] for the modified half-region depth. In all these cases, the authors utilized a particular choice of univariate depth D to define integrated depth as in (2.6). While Fraiman and Muniz [14] focus on establishing strong consistency result for their depth (see Sect. C.1 herein), López-Pintado and Romo [23, 24] and López-Pintado *et al.* [25] already list a few properties that a functional depth should satisfy. See further Sections A.3, A.4 and A.2 for a discussion on respectively modified band depth, modified half-region depth, and modified simplicial band depth.

To provide a comprehensive review of depth functionals relevant to integrated depth FD, we also briefly mention integrated dual depth introduced by Cuevas and Fraiman [4]. This approach is very general in nature and is applicable to general Banach space-valued datasets. In the most general setup of Banach space-valued random variables, assume we are given a Banach space $\mathbb{E}$ with a separable dual $\mathbb{E}^*$ and a fixed probability measure $\mu^* \in \mathcal{P}\left(\mathbb{E}^*\right)$.

Definition 2.5. For $X \sim \mathrm{P} \in \mathcal{P}(\mathbb{E})$ and $x \in \mathbb{E}$, the integrated dual depth of x with respect to P is defined as the integral

$$IDD(x;\mathrm{P}) = \int_{\mathbb{E}^*} D(f(x);\mathrm{P}_f)\,\mathrm{d}\mu^*(f),$$

where P_f stands for the distribution of the random variable $f(X)$ and D is a multiple of the univariate simplicial depth sD, *i.e.*

$$D(u;\mathrm{Q}) = \mathrm{Q}((-\infty,u])\,\mathrm{Q}([u,\infty)) \text{ for } u \in \mathbb{R} \text{ and } \mathrm{Q} \in \mathcal{P}(\mathbb{R}). \tag{2.8}$$

Remark 2.6. For the interesting instance of the infinite-dimensional space of continuous functions (for $K = 1$), the integrated dual depth may take a form of integrated depth (2.6). This can be viewed as follows. Let $\mathbb{E} = \mathcal{C}([0,1])$ and suppose that the support of μ^* is given by the set of coordinate projections

$$\{f : \mathcal{C}([0,1]) \to \mathbb{R} : x \mapsto x(t) : t \in [0,1]\}$$

Then we can canonically identify the projection corresponding to $t \in [0,1]$ with t itself, obtaining a bijection between this set in $\mathbb{E}^*$ and $[0,1]$. Using this construction, for any measure μ^* on $[0,1]$ we can define an integrated dual depth taking the form of a simpler integrated depth FD as understood above.

3. Measurability of the integrand function

The first question that needs to be addressed when studying integrated type depth functionals, is whether the integrand is indeed a measurable function. So far, the issue of possible non-measurability of the function

$$\begin{array}{ccc} [0,1] & \longrightarrow & [0,1] \\ t & \mapsto & D(x(t);\mathrm{P}_t) \end{array} \tag{3.1}$$

has not been addressed in the literature. However, without the guarantee that the integrand is a measurable function, the integral (2.6) is not properly defined, and even more trouble is encountered when trying to show anything about FD. Also most of the proof techniques, as will be seen later, rely upon measure-theoretic tools such as Lebesgue's dominated convergence theorem or Fubini–Tonelli's theorem, and hence it is of vital importance to have the measurability of (3.1) guaranteed.

While the measurability of (3.1) may appear to follow trivially – and it often does – a much more complicated question is the measurability of functions arising when proving consistency results. Here, we often encounter functions like

$$\begin{array}{ccc} [0,1] & \longrightarrow & [0,1] \\ t & \mapsto & \sup_{x\in\mathcal{C}^K([0,1])} |D(x(t);\mathrm{P}_t) - D(x(t);\mathrm{P}_{n,t})| \end{array}$$

as we will see in Section 5. Functions like this, constituting a supremum over an *uncountable* index set, are by no means easy to be handled. It is especially in connection with these functions that we need the strongest notions of measurability for the proofs to hold true.

In this section, we address the measurability issue for the integrated type depth defined in (2.6). By doing this, in particular, we fully recover the results of Fraiman and Muniz [14] and part of the work of López-Pintado and Romo [23,24] on modified versions of depths they proposed. We partially recover also the results of Cuevas and Fraiman [4] for a special choice of dual depth measure IDD and $K = 1$ as outlined in Remark 2.6, and the results of Claeskens *et al.* [2] for multivariate functional depth MFD.

Further complications, not addressed here, include the introduction of a weight function possibly depending on P_t in (2.5), or the measurability of the integrand in Definition 2.5. These cases still deserve a more thorough treatment.

To establish the main result on measurability of all quantities encountered later, an appropriate measurability of the finite-dimensional depth function D used in (2.6) is necessary. This will be referred to later as Condition/Property (D_7) to be satisfied by D. As will be seen in Appendix A several commonly-used depths satisfy this condition.

(D$_7$) *Measurability:* The mapping

$$D\colon \mathbb{R}^K \times \mathcal{P}\left(\mathbb{R}^K\right) \to [0,1]\colon (u,\mathrm{Q}) \mapsto D(u;\mathrm{Q})$$

is jointly Borel measurable and $D(.;\mathrm{Q}) \not\equiv 0$ for all $\mathrm{Q} \in \mathcal{P}\left(\mathbb{R}^K\right)$.

Recall that for a given σ-algebra $\mathcal{S}$ of measurable sets of S, a universally measurable set is measurable for the completion of every probability measure on $\mathcal{S}$ ([9], Sect. 11.5). The universally measurable sets form a σ-algebra, and a function f from $(S,\mathcal{S})$ is called universally measurable if for every Borel set B in its range the set $f^{-1}(B)$ is universally measurable. In particular, a universally measurable function on $[0,1]$ is μ-integrable for any probability measure μ on $[0,1]$.

The main measurability result is formulated in Theorem 3.1, the proof of which is deferred to Appendix B.

Theorem 3.1. *Consider measurable spaces $\mathbb{R}^K$ and $[0,1]$ equipped with the usual Borel σ-algebras and the space of measures $\mathcal{P}\left(\mathbb{R}^K\right)$ equipped with the Borel σ-algebra generated by the metric of the weak convergence of measures. Let the depth function D satisfy (D$_7$). Then the following holds.*

(i) *Functions*

$$h\colon [0,1] \times \mathcal{C}^K([0,1]) \times \mathcal{P}\left(\mathcal{C}^K([0,1])\right) \to [0,1]\colon\ (t,x,\mathrm{P}) \mapsto D\big(x(t);\mathrm{P}_t\big)$$

and

$$g\colon [0,1] \times \mathbb{R}^K \times \mathcal{P}\left(\mathcal{C}^K([0,1])\right) \to [0,1]\colon\ (t,u,\mathrm{P}) \mapsto D\left(u;\mathrm{P}_t\right)$$

are jointly Borel measurable on their domains.

(ii) *Functions*

$$h_2\colon [0,1] \to [0,1]\colon t \mapsto D\big(x(t);\mathrm{P}_t\big), \tag{3.2}$$

$$g_2\colon [0,1] \times \mathbb{R}^K \to [0,1]\colon\ (t,u) \mapsto D\left(u;\mathrm{P}_t\right) \tag{3.3}$$

are Borel measurable for each $x \in \mathcal{C}^K([0,1])$ and $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$.

(iii) *Function*

$$h_3\colon [0,1] \to [0,1]\colon t \mapsto D\big(x(t);\mathrm{P}_t\big)$$

is universally measurable on $[0,1]$ for each $x\colon [0,1] \to \mathbb{R}^K$ Borel measurable and $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$.

(iv) *Function*

$$\widetilde{h}\colon \big(t,\mathrm{P},\widetilde{\mathrm{P}}\big) \mapsto \sup_{x \in \mathcal{C}^K([0,1])} \Big| D\big(x(t);\mathrm{P}_t\big) - D\big(x(t);\widetilde{\mathrm{P}}_t\big) \Big| \tag{3.4}$$

is universally measurable on $[0,1] \times \mathcal{P}\left(\mathcal{C}^K([0,1])\right) \times \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$.

(v) *Function*

$$m\colon [0,1] \to [0,1]\colon t \mapsto \sup_{u \in \mathbb{R}^K} D\big(u;\mathrm{P}_t\big)$$

is universally measurable on $[0,1]$ for each $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$.

Recall that the universal measurability of the integrand functions in Theorem 3.1 is weaker than their Borel measurability. Nevertheless, it is still strong enough for defining and handling integrals as in the definitions of integrated depths. Therefore, the assertion of Theorem 3.1 is sufficient for all our present purposes.

From now on, we shall assume that D satisfies (D$_7$), without specifically mentioning it.

Part (i) of Theorem 3.1 is rather technical, but it is crucial for the proof of the following statements of the theorem. (3.2) in property (ii) ensures that the integrated depth (2.6) is defined correctly for a fixed continuous function. Further, (3.3) in (ii) and (iii) are utilized mainly in Section 4.2 that deals with the existence of a function maximizing FD. Result (iv) ensures good measurability properties of various supremum functions encountered in Section 5 when establishing consistency results for integrated depth. Part (v) is utilized in Section 4.2, where the function m is used to define the maximal value of FD over the sample space.

4. Desirable properties of depth functionals

In this section we address the question of existence of analogues of the desired properties (D$_1$)–(D$_6$) in the setup of depth for vector-valued functional data as introduced in Section 2.2. For some of the conditions the generalization is straightforward – for these we only present and discuss the requirements under which the functional analogies of them hold true. For some of them however the extensions are not so obvious.

In the literature one can find some scattered results regarding (D$_1$)–(D$_6$) for functional depths. A major reference for a more general discussion on functional analogies for (D$_1$)–(D$_6$) is Mosler and Polyakova [31]. Assuming only that the necessary condition (D$_7$) is satisfied for a finite-dimensional depth D, we explore to what extent conditions (D$_1$)–(D$_6$) for D are inherited by an integrated depth FD based on D.

Cuevas and Fraiman [4] show some results related to (D$_1$)–(D$_6$) for $K = 1$, but only in a finite-dimensional case, not covering the more interesting instance of functional data. Claeskens *et al.* [2] discuss a multivariate functional depth, and for general K, the authors also show a range of theoretical properties of MFD, not too different from the kind of theoretical properties we discuss here in a general setting (of any depth function D). We explore the suggested properties methodically and without any additional assumptions imposed on the (generic) depth function D, except from a minimal subset of the usual collection (D$_1$)–(D$_6$).

Let $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$, $X \sim \mathrm{P}$ and P_X a measure corresponding to X, and consider an integrated depth FD for functionals as in (2.6). For convenience of the reader we start by summarizing desirable properties for functional depths, denoted by (DF$_1^{\mathrm{A}}$)–(DF$_6$), which can be seen as analogues of (D$_1$)–(D$_6$) for functional depths. We discuss and motivate these in detail in the following subsections. In particular, the rigorous definition of the notion of halfspace symmetry for functional data utilized in condition (DF$_2$) is postponed to Section 4.2.2. Here, also some discussion on this concept can be found.

- *Invariance Properties:*
  - (DF$_1^{\mathrm{A}}$) *Translation invariance:* For any $b, x \in \mathcal{C}^K([0,1])$ and X
$$FD\left(x; \mathrm{P}_X, D\right) = FD\left(x + b; \mathrm{P}_{X+b}, D\right).$$
  - (DF$_1^{\mathrm{B}}$) *Function-scale invariance:* For any matrix-valued function $a\colon [0,1] \to \mathbb{R}^{K\times K}$ such that $a(t)$ is non-singular for all $t \in [0,1]$ and each $x \in \mathcal{C}^K([0,1])$
$$FD\left(x; \mathrm{P}_X, D\right) = FD\left(a.x; \mathrm{P}_{a.X}, D\right),$$
where $a.x(t) = a(t) \cdot x(t)$ in the sense of matrix multiplication.
  - (DF$_1^{\mathrm{C}}$) *Rearrangement invariance:* For any $\rho\colon [0,1] \to [0,1]$ bijective and $x \in \mathcal{C}^K([0,1])$
$$FD\left(x; \mathrm{P}_X, D\right) = FD\left(x \circ \rho; \mathrm{P}_{X\circ\rho}, D\right).$$
  - (DF$_1^{\mathrm{D}}$) *Measure-preserving rearrangement invariance:* (DF$_1^{\mathrm{C}}$) holds true for any bijective and measure-preserving $\rho\colon [0,1] \to [0,1]$.
- *Maximality at the center:*
  - (DF$_2$) If the distribution of X is halfspace symmetric around function $x^*\colon [0,1] \to \mathbb{R}^K$, then we can define
$$FD\left(x^*; \mathrm{P}_X, D\right) = \int_0^1 D\left(x^*(t); \left(\mathrm{P}_X\right)_t\right) \mathrm{d}\mu(t),$$
and
$$\sup_{x\in\mathcal{C}^K([0,1])} FD\left(x; \mathrm{P}_X, D\right) = \sup_{x\in\mathcal{B}^K([0,1])} FD\left(x; \mathrm{P}_X, D\right) = FD\left(x^*; \mathrm{P}_X, D\right).$$
- *Quasi-concavity as a function of x:*
  - (DF$_3$) The level set $\left\{x : FD(x; \mathrm{P}, D) \geq c\right\}$ is convex for each $c \in \mathbb{R}$.

- *Vanishing at infinity:*
  (DF$_4$) $FD\,(x; \mathrm{P}, D) \xrightarrow[\|x\|\to\infty]{} 0.$
- *Continuity as a function of x:*
  (DF$_5$) *Upper semicontinuity of FD as a function of x:* For all $x \in \mathcal{C}^K([0,1])$ and $x_\nu \xrightarrow[\nu\to\infty]{} x$

$$\limsup_{\nu\to\infty} FD\,(x_\nu; \mathrm{P}, D) \leq FD\,(x; \mathrm{P}, D)\,.$$

- *Weak continuity as a functional of* P*:*
  (DF$_6$) For all $\mathrm{P}_\nu \xrightarrow[\nu\to\infty]{w} \mathrm{P}$

$$\sup_{x\in\mathcal{C}^K([0,1])} |FD\,(x; \mathrm{P}, D) - FD\,(x; \mathrm{P}_\nu, D)| \xrightarrow[\nu\to\infty]{} 0.$$

4.1. Invariance properties

The affine invariance property of depth D can, from the theoretical point of view, very easily be transcribed into the functional setup. For this, we could employ translations by an arbitrary function and isomorphisms on the space $\mathcal{C}^K([0,1])$ as proposed by ([31], Condition FD2L on p. 8). This general notion of linear invariance appears however to be far too restrictive when it comes to functional datasets. The authors therefore suggested to use other, much weaker than FD2L, analogies of (D$_1$) – translation, rearrangement, and function-scale invariance. Mosler and Polyakova [31] consider invariance conditions only for the case $K = 1$. Conditions (DF$_1^{\mathrm{A}}$)–(DF$_1^{\mathrm{D}}$) are a suitable adaptation of these to the framework of integrated depth.

Similar invariance properties of depth for functional data have been studied by other authors as well. López-Pintado and Romo ([23], Thm. 3) (for $K = 1$), and López-Pintado *et al.* ([25], Thm. 1) (for general $K \in \mathbb{N}$) establish for (non-integrated) band depth and simplicial band depth, respectively, invariance properties (DF$_1^{\mathrm{A}}$), (DF$_1^{\mathrm{B}}$), and (DF$_1^{\mathrm{C}}$) as presented above. For $K = 1$, both López-Pintado and Romo [24] for (non-integrated) half-region depth for functions and Cuevas and Fraiman [4] for integrated dual depth for functions mention (DF$_1^{\mathrm{A}}$) and (DF$_1^{\mathrm{B}}$). Finally, Claeskens *et al.* [2] show for general K for MFD defined in Section 2.2 an equivalent of (DF$_1^{\mathrm{A}}$), (DF$_1^{\mathrm{B}}$) for function a constant in t, and (DF$_1^{\mathrm{D}}$) for mappings ρ in a special form. All these invariance results for MFD hold true only if the weight function w is properly invariant as well.

In a general setting, the establishment of the validity of conditions (DF$_1^{\mathrm{A}}$), (DF$_1^{\mathrm{B}}$) and (DF$_1^{\mathrm{D}}$) is straightforward.

Theorem 4.1. *For FD based on D, the following invariance properties hold true:*

- *If D satisfies (D$_1^{\mathrm{A}}$), then FD satisfies (DF$_1^{\mathrm{A}}$).*
- *If D satisfies (D$_1^{\mathrm{B}}$), then FD satisfies (DF$_1^{\mathrm{B}}$).*
- *If μ is the Lebesgue measure on $[0,1]$, then FD satisfies (DF$_1^{\mathrm{D}}$).*

Proof. The assertions follow immediately from the definition. □

It follows immediately from Theorem 4.1 that if D satisfies (D$_1$) and μ is the Lebesgue measure, then the corresponding depth FD satisfies all the invariance properties except from (DF$_1^{\mathrm{C}}$). The following example shows that imposing (DF$_1^{\mathrm{C}}$) on a functional depth is rather questionable.

Example 4.2. Let $K = 1$, μ the Lebesgue measure on $[0,1]$, $D = hD$ and $X \sim \mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ be a Dirac measure concentrated on a constant zero function

$$\mathrm{P}\,(X(t) = 0 \text{ for all } t \in [0,1]) = 1.$$

For $\varepsilon > 0$ small, set

$$x(t) = \begin{cases} 0 & \text{for } t \in [0, 1-\varepsilon], \\ t + \varepsilon - 1 & \text{for } t \in (1-\varepsilon, 1]. \end{cases}$$

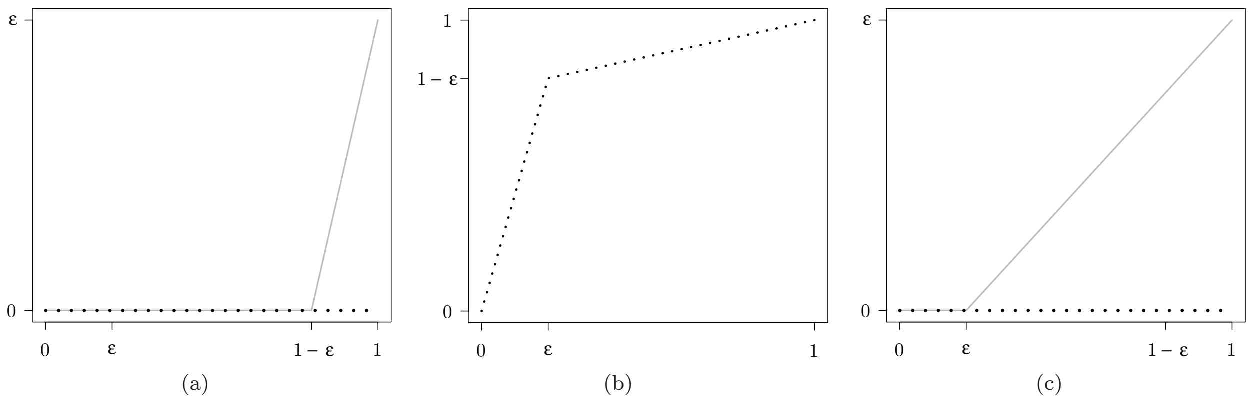

FIGURE 1. (a) Function x (*solid line*) and a constant zero function (*dotted*) on which P is concentrated; (b) bijection ρ not preserving measure; and (c) transformed function $x \circ \rho$ along with (transformed) zero function from Example 4.2 for $\varepsilon = 0.2$.

Then $FD\,(x; \mathrm{P}, hD) = 1 - \varepsilon$, a value very close to 1, the maximal possible one. Consider the transformation

$$\rho(t) = \begin{cases} \frac{t(1-\varepsilon)}{\varepsilon} & \text{for } t \in [0, \varepsilon], \\ \frac{\varepsilon t + 1 - 2\varepsilon}{1-\varepsilon} & \text{for } t \in (\varepsilon, 1]. \end{cases}$$

Then ρ is a bijection as in $(\mathrm{DF}_1^{\mathrm{C}})$ and

$$FD\,(x \circ \rho; \mathrm{P}_{X \circ \rho}, hD) = \varepsilon,$$

a value very close to 0, the smallest possible one. Figure 1 depicts some examples of functions $x(t)$. In this figure, the original function x has with respect to P a high integrated depth value 0.8, while in the transformed setup x attains a very low value 0.2, showing that $(\mathrm{DF}_1^{\mathrm{C}})$ is not satisfied for integrated depths.

Remark 4.3. If the concept of depth for vector-valued functions is used for smooth scalar-valued functions in the way discussed by Mosler and Polyakova ([31], Sect. 5.2), then it is easy to see that if (D_1) holds true for D, then the integrated depth FD based on D is invariant with respect to affine transformations given by $a(t).y(t) + b(t)$. Here, a, b are scalar-valued functions on $[0, 1]$ that are appropriately continuously differentiable and $a(t).y(t)$ is the usual pointwise product of functions.

4.2. Maximality at center

4.2.1. Maximality at center: general case

Consider a set-valued mapping M and a function x^* defined as follows

$$x^*(t) \in M(t) = \left\{ y \in \mathbb{R}^K : D(y; \mathrm{P}_t) = \sup_{u \in \mathbb{R}^K} D(u; \mathrm{P}_t) \right\} \text{ for } t \in [0, 1]. \tag{4.1}$$

The mapping M can be conceived as having values in $2^{\mathbb{R}^K}$ – the collection of all subsets of $\mathbb{R}^K$, and x^* as an arbitrary selection from M. By definition (2.6) of FD it immediately follows that any such function x^* satisfies

$$\sup_{x \in \mathcal{C}^K([0,1])} FD\,(x; \mathrm{P}, D) \leq FD\big(x^*; \mathrm{P}, D\big) = \int_0^1 \sup_{u \in \mathbb{R}^K} D\big(u; \mathrm{P}_t\big)\, \mathrm{d}\mu(t) \tag{4.2}$$

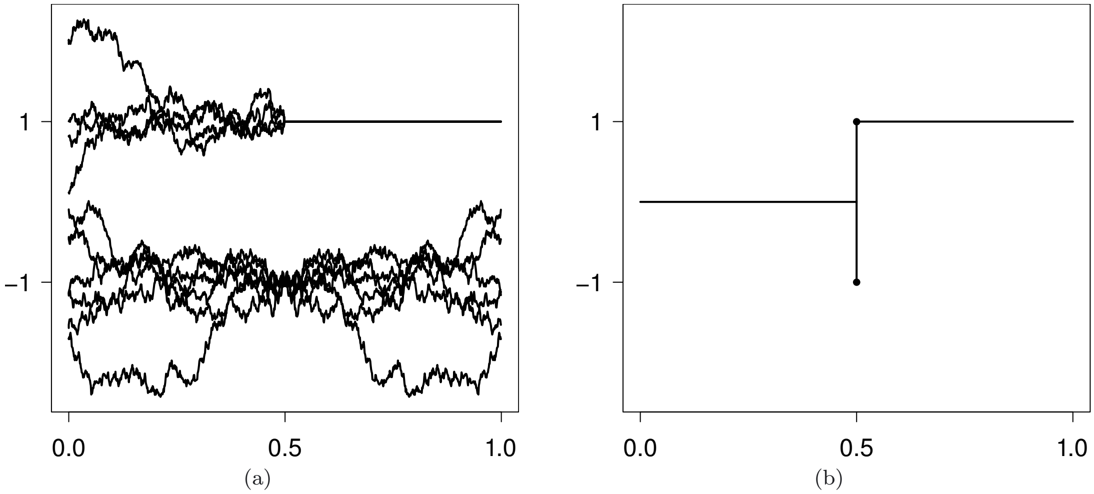

FIGURE 2. (a) A random sample of size 10 from P; and (b) the set-valued mapping M_{hD} from Example 4.4.

where the existence of $FD\big(x^*; \mathrm{P}, D\big)$ is justified by (v) of Theorem 3.1. In this section, we focus on the properties of such maximizers x^*. We investigate when a continuous, or at least measurable selection x^* can be drawn from M. These observations will later in Section 4.2.2 help us to define a reasonable notion of symmetry for distributions in $\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and extend condition (D$_2$) to a vector-valued functional data setup.

The following example illustrates that it is not always possible to find a continuous x^* satisfying (4.2). In this example μ is taken to be the Lebesgue measure on $[0,1]$, without loss of generality.

Example 4.4. Take $K = 1$ and let D be either hD or sD. Consider the two stochastic processes Y_1 and Y_2 on $[0,1]$ defined as

$$Y_1(t) = \begin{cases} 1 + W_{0.5-t}, & \text{if } 0 \le t \le 0.5, \\ 1, & \text{if } 0.5 < t \le 1, \end{cases} \qquad Y_2(t) = \begin{cases} -1 + \widetilde{W}_{0.5-t}, & \text{if } 0 \le t \le 0.5, \\ -1 + \widetilde{W}_{t-0.5}, & \text{if } 0.5 < t \le 1, \end{cases}$$

where W and $\widetilde{W}$ are two independent copies of Wiener processes. The distribution P is given as a mixture of distributions of Y_1 and Y_2 with equal mixing proportions. Then $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}\left([0,1]\right)\right)$ because the processes Y_1 and Y_2 have continuous trajectories P-almost surely. A random sample of size 10 from P is plotted in Figure 2a.

Now, for each $t \in [0, 0.5)$ the marginal distribution P_t is given by a mixture of two independent normal distributions $N(-1, 0.5-t)$ and $N(1, 0.5-t)$, each having weight 0.5. Thus P_t is halfspace symmetric around zero and by (D$_2$) the maximum depth is attained at zero for D (both hD and sD satisfy (D$_2$) for absolutely continuous distributions, see Appendix A).

For $t = 0.5$ both $N(-1, 0.5-t)$ and $N(1, 0.5-t)$ degenerate to Dirac measures at the points -1 and 1, respectively. It is easy to compute that the maximal hD with respect to $\mathrm{P}_{0.5}$ is 0.5 and it is attained at each point of the closed interval $[-1,1]$. The corresponding maximal value of sD is, however, 0.75 and this value is attained at two distinct points -1 and 1.

For $t \in (0.5, 1]$ it is not difficult to see that both hD and sD attain its maximal value at 1. Thus, the set-valued function M introduced in (4.1) is for hD and sD given by, respectively,

$$M_{hD}(t) = \begin{cases} \{0\}\,, & \text{if } 0 \leq t < 0.5, \\ [-1,1], & \text{if } t = 0.5, \\ \{1\}\,, & \text{if } 0.5 < t \leq 1, \end{cases} \qquad M_{sD}(t) = \begin{cases} \{0\}\,, & \text{if } 0 \leq t < 0.5, \\ \{-1,1\}\,, & \text{if } t = 0.5, \\ \{1\}\,, & \text{if } 0.5 < t \leq 1. \end{cases}$$

M_{hD} is plotted in Figure 2b. M_{sD} almost coincides with M_{hD}. The only difference between the two graphs is that the vertical line at point 0.5 for M_{hD} is replaced by two points $(0.5, -1)$ and $(0.5, 1)$ for M_{sD}.

Note that one cannot find a continuous function x on $[0, 1]$ so that $x(t) \in M_{hD}(t)$ or $x(t) \in M_{sD}(t)$ for almost all $t \in [0, 1]$. On the other hand, it is easy to see that one can choose a sequence of continuous functions x_ν such that

$$FD\big(x_\nu; \mathrm{P}, D\big) \xrightarrow[\nu\to\infty]{} \int_0^1 \sup_{u \in \mathbb{R}^K} D\big(u; \mathrm{P}_t\big)\, \mathrm{d}\mu(t) \tag{4.3}$$

The next example shows that even if the marginal distribution P_t is centrally symmetric around $x^*(t)$ for each $t \in [0, 1]$, there might not exist a continuous function x^* on $[0, 1]$ satisfying (4.2).

Example 4.5. Consider the random process Y_1 as in Example 4.4 and define

$$Y_3(t) = \begin{cases} -1 + \widetilde{W}_{0.5-t}, & \text{if } 0 \leq t \leq 0.5, \\ -1, & \text{if } 0.5 < t \leq 1, \end{cases}$$

with the process $\widetilde{W}$ as in Example 4.4. Let the distribution P be again given as a mixture of distributions of Y_1 and Y_3 with equal mixing proportions. Then all the marginals P_t are centrally (thus halfspace) symmetric around $x^*(t) = 0$ for all $t \in [0, 1]$. However, for the simplicial depth sD we have

$$M_{sD}(t) = \begin{cases} \{0\}\,, & \text{if } 0 \leq t < 0.5, \\ \{-1,1\}\,, & \text{if } 0.5 \leq t \leq 1, \end{cases}$$

and again, no continuous function on $[0, 1]$ can be selected from this set. Note that also here it is easy to find a sequence of continuous functions such that (4.3) holds.

Note that all the examples presented in this paper are aimed to be as simple as possible in order to demonstrate the problems one may encounter when dealing with integrated depths for functions. This however certainly does not limit their scope to random functions which are non-differentiable, atomic, or not having absolutely continuous marginals. For instance, one array of models related to Examples 4.4 and 4.5 but providing nicely behaved functional distributions with discontinuous coordinatewise median function can be found in ([32], Sects. 3 and 4).

The following theorem states that (4.3) is not an exceptional feature of the presented examples, but holds generally. Moreover, the theorem also states that there always exists a Borel measurable function x^* on $[0, 1]$ such that the point $x^*(t)$ maximizes $D\,(.; \mathrm{P}_t)$ for all $t \in [0, 1]$.

Theorem 4.6. *Let D satisfy (D_4) and (D_5). Then*

(i) *for any $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ there exists a Borel measurable function $x^*\colon [0,1] \to \mathbb{R}^K$ such that (4.1) holds true, and*
(ii) *a sequence of functions $x_\nu \in \mathcal{C}^K([0,1])$ can be found so that (4.3) holds.*

Proof. By part (ii) of Theorem 3.1 we know that for any $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ the function $g_2(t,u) = D(u;\mathrm{P}_t)$ is jointly Borel measurable on $[0,1]\times\mathbb{R}^K$. For each $t\in[0,1]$ the function $g_2(t,\cdot)$ is by (D_5) upper semicontinuous and by (D_4) and (D_5) it attains its maximal value over the σ-compact space $\mathbb{R}^K$. Using a general measurable selection theorem of Brown and Purves ([1], Cor. 1), there exists a Borel measurable function $x^*\colon[0,1]\to\mathbb{R}^K$ such that (4.1) holds true.

For the second part of the statement, by Lusin's Theorem (*cf.* Thm. 7.5.2 of [9]) we have that for each $\nu\in\mathbb{N}$ there exists a compact set $C_\nu\subset[0,1]$, $\mu(C_\nu)\geq 1-1/\nu$ such that x^* is continuous on C_ν. Finally, by Tietze–Urysohn's Extension Theorem (*cf.* Thm. 2.6.4. of [9]), the continuous restriction of x^* to C_ν can be extended to a continuous function x_ν on $[0,1]$ and $x_\nu(t)=x^*(t)$ for $t\in C_\nu$. Hence,

$$\begin{aligned}FD(x_\nu;\mathrm{P},D) &\leq \sup_{x\in\mathcal{C}^K([0,1])} FD(x;\mathrm{P},D)\leq FD(x^*;\mathrm{P},D)\\ &= \int_{C_\nu} D\big(x_\nu(t);\mathrm{P}_t\big)\,\mathrm{d}\mu(t)+\int_{[0,1]\setminus C_\nu} D\big(x^*(t);\mathrm{P}_t\big)\,\mathrm{d}\mu(t)\leq FD(x_\nu;\mathrm{P},D)+1/\nu\end{aligned}$$

and taking the limit $\nu\to\infty$ yields (4.3). □

The following theorem says that if the set $M(t)$ contains only one point for each $t\in[0,1]$ and the depth function D satisfies (D_6), then the function $x^*(t)=M(t)$ is in $\mathcal{C}^K([0,1])$.

Theorem 4.7. *Let D satisfy assumptions (D_4), (D_5) and (D_6) for each P_t, $t\in[0,1]$, and let D have a unique maximizing point*

$$M(t)=\arg\max_{u\in\mathbb{R}^K} D(u;\mathrm{P}_t) \tag{4.4}$$

for all $t\in[0,1]$. Then the function x^* defined as $x^*(t)=M(t)$, $t\in[0,1]$, is in $\mathcal{C}^K([0,1])$.

Proof. Put

$$M^c_\varepsilon(t)=\Big\{u\in\mathbb{R}^K:\inf_{y\in M(t)}\|y-u\|_{\mathbb{R}^K}>\varepsilon\Big\}. \tag{4.5}$$

By the same reasoning as in the proof of (iii) of Lemma 5.5.1 in [6] one can show that thanks to (D_4) and (D_5) it holds that for all $\varepsilon>0$ and $t\in[0,1]$

$$\sup_{u\in\mathbb{R}^K} D(u;\mathrm{P}_t)-\sup_{u\in M^c_\varepsilon(t)} D(u;\mathrm{P}_t)>0. \tag{4.6}$$

Let $t\in[0,1]$ be fixed and $\{t_\nu\}$ be a sequence in $[0,1]$ such that $t_\nu \xrightarrow[\nu\to\infty]{} t$. Note that one can bound

$$\begin{aligned}0&\leq D\big(M(t);\mathrm{P}_t\big)-D\big(M(t_\nu);\mathrm{P}_t\big)\\ &=\Big[D\big(M(t);\mathrm{P}_t\big)-D\big(M(t);\mathrm{P}_{t_\nu}\big)\Big]+\Big[D\big(M(t);\mathrm{P}_{t_\nu}\big)-D\big(M(t_\nu);\mathrm{P}_{t_\nu}\big)\Big]\\ &\quad+\Big[D\big(M(t_\nu);\mathrm{P}_{t_\nu}\big)-D\big(M(t_\nu);\mathrm{P}_t\big)\Big].\end{aligned} \tag{4.7}$$

Similarly as in the Proof of Theorem 3.1 (i) one can show that $\mathrm{P}\in\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ implies that $\mathrm{P}_{t_\nu}\to\mathrm{P}_t$ weakly. This together with assumption (D_6) (that holds for each P_t) yields that the first and the third term on the right-hand side of (4.7) converge to zero. Further, the second term on the right-hand side of (4.7) is non-positive. Thus for each $\varepsilon>0$ and for all sufficiently large $\nu\in\mathbb{N}$ the right-hand side of (4.7) is less than ε, implying that $D(M(t_\nu);\mathrm{P}_t)\xrightarrow[\nu\to\infty]{} D(M(t);\mathrm{P}_t)$. Now (4.6) yields that $M(t_\nu)\xrightarrow[\nu\to\infty]{} M(t)$. As the point t was taken arbitrarily in $[0,1]$, this implies that $M(t)$ is continuous for $t\in[0,1]$. □

4.2.2. Maximality at center: symmetric distributions

There are various definitions of symmetry for finite-dimensional distributions available in the literature [36]. In the setup of functional, or even vector-valued functional data as considered here, the task of determining which probability distribution is symmetric around a function and which function should serve as the center of this symmetry did not get much attention so far.

For the sake of having a reasonable definition of symmetry for such complicated structures at hand, we define a new, natural notion of symmetry for functional data. Circumventing the confusion raised by the ambiguity in $\mathbb{R}^K$ by acceding to the notion of halfspace symmetry in Section 2.1, it is fairly straightforward to see how a symmetric distribution of a random function in $\mathcal{C}^K([0,1])$ could be defined.

Definition 4.8. Let $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$. We say that P is halfspace symmetric around a function $x^*\colon [0,1] \to \mathbb{R}^K$, if there exists x^* Borel measurable such that the marginal distributions $\mathrm{P}_t \in \mathcal{P}\left(\mathbb{R}^K\right)$ are halfspace symmetric around $x^*(t)$ for each $t \in [0,1]$.

A primal question is then what are the properties of the center of symmetry function x^*. One complication, already mentioned in Section 4.2.1, is that the center of symmetry x^* is not necessarily a unique function, and under certain circumstances there can exist a whole set of functions which are centers of symmetry

$$\Theta\left(\mathrm{P}\right) = \left\{x^*\colon [0,1] \to \mathbb{R}^K : x^* \text{ is the center of halfspace symmetry of P}\right\} \text{ for } \mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right).$$

Distributions P halfspace symmetric around a whole set of functions were constructed in Examples 4.4 and 4.5. There, it was shown (as each distribution $\mathrm{Q} \in \mathcal{P}(\mathbb{R})$ is halfspace symmetric around its median) that even for halfspace symmetric probability distributions P attaining values in $\mathcal{C}^K([0,1])$ only, the set $\Theta\left(\mathrm{P}\right)$ may have empty intersection with the space $\mathcal{C}^K([0,1])$.

This however does not prevent us from reasonably extending property (D_2) to the setup of vector-valued functional data as stated in (DF_2). The condition asserts that

- for any $x \in \mathcal{C}^K([0,1])$
$$FD\left(x; \mathrm{P}, D\right) \leq FD\left(x^*; \mathrm{P}, D\right),$$
- and that there exists a sequence of continuous functions $\{x_\nu\} \subset \mathcal{C}^K([0,1])$ so that
$$FD\big(x_\nu; \mathrm{P}, D\big) \xrightarrow[\nu\to\infty]{} FD\left(x^*; \mathrm{P}, D\right).$$

The first assertion is trivially satisfied provided condition (D_2) holds true for D and the second assertion is a direct corollary of part (ii) of Theorem 4.6. This then leads to the following main result.

Theorem 4.9. *If D satisfies (D_2), (D_4) and (D_5), then FD based on D satisfies (DF_2).*

Proof. Follows immediately from the considerations preceding the statement of the theorem. □

To conclude the section on maximality properties of integrated depth we add one additional theorem discussing the existence of the deepest function similar to Theorem 4.7. Here we state that one can be more specific about the uniqueness of the maximizing function x^* if the distribution $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ is halfspace symmetric. To do this, recall from Zuo and Serfling ([40], Thm. 2.1) that the center of the halfspace symmetry of $\mathrm{Q} \in \mathcal{P}\left(\mathbb{R}^K\right)$ is unique, if it exists and

(U) the support of Q is not concentrated on a subset of a line $L \subset \mathbb{R}^K$, on which the (univariate) distribution of Q has more than one median.

Assumption (U) is evidently satisfied if, for instance, the support of Q is connected or contiguous [18], or if Q is absolutely continuous and $K > 1$.

Theorem 4.10. *Let* $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ *be halfspace symmetric and let* P_t *satisfy* (U) *for all* $t \in [0,1]$*. Then* P *is halfspace symmetric around a unique, continuous function* $x^* \in \mathcal{C}^K([0,1])$.

Proof. The existence and uniqueness of the function x^* follow immediately from Zuo and Serfling ([40], Thm. 2.1). It remains to show that $x^* \in \mathcal{C}^K([0,1])$.

Fix t and a sequence $\{t_\nu\}$ in $[0,1]$, $t_\nu \to t$. Suppose now that $x^*(t_\nu) \not\to x^*(t)$. By the tightness of the measure $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ there exists a subsequence $\{t_\eta\}$ and $\tilde{\theta} \neq x^*(t)$ such that $x^*(t_\eta) \to \tilde{\theta}$. Let X be a random function with distribution P. As $x^*(t)$ is the unique center of the symmetry, there exists a closed halfspace $\tilde{H}$ passing through the origin such that

$$\mathsf{P}\big(X(t) - \tilde{\theta} \in \tilde{H}\big) < \frac{1}{2}. \tag{4.8}$$

Note that $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ implies that $X(t_\eta) \to X(t)$ weakly, which further yields that $X(t_\eta) - x^*(t_\eta)$ converges to $X(t) - \tilde{\theta}$ weakly. Now by the Portmanteau theorem ([9], Thm. 11.1.1) used for the closed halfspace $\tilde{H}$ it follows that

$$\tfrac{1}{2} \leq \limsup_{\eta\to\infty} \mathsf{P}\big(X(t_\eta) - x^*(t_\eta) \in \tilde{H}\big) \leq \mathsf{P}\big(X(t) - \tilde{\theta} \in \tilde{H}\big),$$

contradicting (4.8) and finishing the proof of the theorem. □

Note that in Example 4.4 the distribution P satisfies all conditions of Theorem 4.10 except for condition (U) at a single marginal $t = 0.5$. That is why the function of halfspace symmetry cannot be chosen to be continuous there.

To end this section, we discuss already available results related to the maximality of integrated depth. First we note that most of the problems connected with the definition of symmetry in functional spaces have not been dealt with so far. Cuevas and Fraiman [4], López-Pintado and Romo [23], López-Pintado *et al.* [25] and Claeskens *et al.* [2] consider at some point an extension of condition (D_2) to the functional setup, but each deals with a simpler (more specific) framework. López-Pintado and Romo ([23], Prop. 2) and Cuevas and Fraiman ([4], Sect. 2, Prop. (ii)) restrict to the case of central symmetry of P. In the latter paper, even though the considered depth D is basically a multiple of sD used in Example 4.5, the emerging problems outlined in Example 4.5 are overcome by assuming absolutely continuous marginals P_t. Finally, both Claeskens *et al.* ([2], Thm. 1 (ii)) and López-Pintado *et al.* ([25], Thm. 2) assume that the center of symmetry function x^* must be unique and continuous itself.

Concerning existence and continuity of the function maximizing integrated depth, results of a similar fashion have been stated by Claeskens *et al.* ([2], Thm. 2). In this section we study the problem from a more general point of view and moreover (i) address the measurability issues; (ii) provide stronger results in Theorems 4.6, 4.7 and 4.10; (iii) avoid making technical assumptions concerning semicontinuity of some set-valued mappings (such as Assumptions (a) or (b) in Theorem 2 of Claeskens *et al.* [2]) since their validity is not easy to verify for any D, or probability $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$.

In this section we explored which of the firm given conditions (D_1)–(D_6) are sufficient for the existence of a measurable function maximizing FD in Theorem 4.6, and in Theorems 4.7 and 4.10 we have directly specified which conditions ensure the continuity of such generalized median functions.

4.3. Quasi-concavity as a function of x

The extension of condition (D_3) to vector-valued functions and $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ is straightforward, as convexity is well established in any vector space. This leads to (DF_3).

An alternative definition of quasi-concavity of an arbitrary function $f\colon S \to \mathbb{R}$ ([34], Sect. 81) for a vector space S is that for each $\gamma \in [0,1]$ and $u, v \in S$

$$f(\gamma u + (1-\gamma)v) \geq \min\{f(u), f(v)\}. \tag{4.9}$$

As the following simple counter-example shows, quasi-concavity seems to be too restrictive for an integrated depth of functional data, even if D satisfies (D_3).

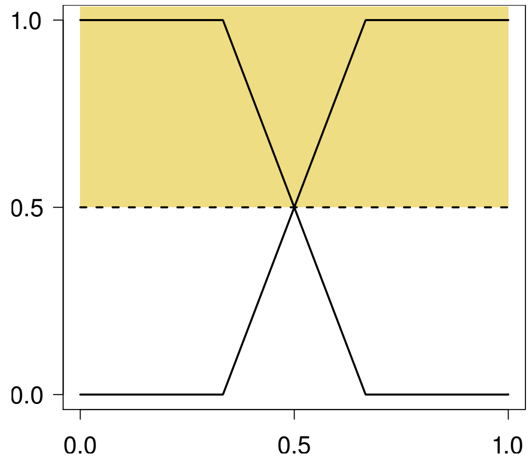

FIGURE 3. Example 4.11. Functions x, y (*solid lines*), their convex combination $0.5x + 0.5y$ (*dashed line*) and a part of the support of the distribution P (*colored region*). The functions x and y have positive integrated depth, but their convex combination has zero depth.

Example 4.11. Let μ be the Lebesgue measure on $[0, 1]$, $K = 1$ and $D = hD$. By Appendix A D satisfies (D_3). Let for all $t \in [0, 1]$ be the marginal distribution P_t distributed uniformly on $[0.5, 1.5]$. Let the functions x and y be defined as follows

$$x(t) = \begin{cases} 0, & t \leq \frac{1}{3}, \\ 3t - 1, & t \in \left(\frac{1}{3}, \frac{2}{3}\right), \\ 1, & t \geq \frac{2}{3}, \end{cases} \qquad y(t) = \begin{cases} 1, & t \leq \frac{1}{3}, \\ 2 - 3t, & t \in \left(\frac{1}{3}, \frac{2}{3}\right), \\ 0, & t \geq \frac{2}{3}. \end{cases}$$

Note that $y(t) = 1 - x(t)$. The functions x and y are depicted in Figure 3. Further it is easy to see that

$$FD\big(\tfrac{1}{2}\,x + \tfrac{1}{2}\,y; \mathrm{P}, hD\big) = 0 < \tfrac{5}{24} = FD\big(x; \mathrm{P}, hD\big) = FD\big(y; \mathrm{P}, hD\big)$$

and hence (4.9) is violated for FD.

Although (DF_3) is not satisfied by integrated depth, a weaker version of it – property $(\mathrm{DF}_3^{\mathrm{W}})$ can be shown to be valid. Mosler and Polyakova ([31], Condition FD4 on p. 7) offer a possible generalization of $(\mathrm{D}_3^{\mathrm{W}})$ into the setup of vector-valued functional data.

$(\mathrm{DF}_3^{\mathrm{W}})$ *Decreasing along rays:* if a maximum depth FD over $\mathcal{C}^K([0, 1])$ is attained at $x^* \colon [0, 1] \to \mathbb{R}^K$ Borel measurable, then for every $x \colon [0, 1] \to \mathbb{R}^K$ Borel measurable and $\gamma \in [0, 1]$

$$FD\big(\gamma x^* + (1 - \gamma)x; \mathrm{P}, D\big) \geq FD\big(x; \mathrm{P}, D\big).$$

Notice that the correctness of the previous condition is justified by Theorem 3.1 (iii).

Observation (4.2) implies that

$$FD\big(x^*; \mathrm{P}, D\big) = \int_0^1 \sup_{u \in \mathbb{R}^K} D\big(u; \mathrm{P}_t\big)\, \mathrm{d}\mu(t),$$

which further yields that $x^*(t)$ maximizes $D(\cdot\,; \mathrm{P}_t)$ for almost all $t \in [0, 1]$. Thus $(\mathrm{DF}_3^{\mathrm{W}})$ is certainly satisfied provided that $(\mathrm{D}_3^{\mathrm{W}})$ holds for D. This was already observed by Claeskens *et al.* ([2], Thm. 1 (iii)) as well.

Theorem 4.12. *If D satisfies $(\mathrm{D}_3^{\mathrm{W}})$, then FD based on D satisfies $(\mathrm{DF}_3^{\mathrm{W}})$.*

Proof. Straightforward. □

4.4. Vanishing at infinity

Transcription of the vanishing at infinity property (D$_4$) of D into functional data is also straightforward, and leads to (DF$_4$). It is easy to see that (DF$_4$) cannot hold true in general for integrated depth based on D satisfying (D$_4$). We illustrate this with an example.

Example 4.13. Let μ be the Lebesgue measure on $[0,1]$, $K=1$ and $D = hD$. By Appendix A, D satisfies (D$_4$). Consider the sequence

$$x_\nu(t) = \begin{cases} \nu(1-\nu t), & \text{if } t < 1/\nu, \\ 0, & \text{otherwise,} \end{cases}$$

for $\nu > 0$. Now $\|x_\nu\| = \nu \to \infty$, but $FD(x_\nu; \mathrm{P}, hD)$ tends to the depth of a zero function. Property (DF$_4$) is then evidently not satisfied for those P for which $FD(0; \mathrm{P}, hD) > 0$. Moreover, if the distribution P is halfspace symmetric around 0 and (DF$_2$) holds for FD, then $FD(x_\nu; \mathrm{P}, hD)$ converges even to $\sup_{x\in\mathcal{C}([0,1])} FD(x; \mathrm{P}, hD)$.

Although the lack of property (DF$_4$) seems to be troubling, one has to keep in mind that the sequence of functions $\{x_\nu\}$ used in the counter-example is rather tricky and out of interest in applications. One can recover (DF$_4$) provided that one can restrict to the functions whose growth is in some way uniformly bounded. For instance let $\widetilde{\mathcal{C}}^K([0,1])$ be a subset of $\mathcal{C}^K([0,1])$ such that

$$\sup_{x\in\widetilde{\mathcal{C}}^K([0,1])} \sum_{k=1}^{K} \left(\sup_{t\in[0,1]} x^{(k)}(t) - \inf_{t\in[0,1]} x^{(k)}(t) \right) < \infty$$

and suppose only $x \in \widetilde{\mathcal{C}}^K([0,1])$.

Another option is to consider a more stringent version of the convergence to infinity criterion, as assumed in Claeskens *et al.* ([2], Thm. 1 (iv)) and López-Pintado *et al.* ([25], Thm. 2). There, the convergence in norm condition $\|x_\nu\| \xrightarrow[\nu\to\infty]{} \infty$ was bent to the form

$$\left| {x_\nu}^{(k)}(t) \right| \xrightarrow[\nu\to\infty]{} \infty \text{ for } \mu\text{-almost all } t \in [0,1] \text{ and some } k = 1,\dots,K.$$

In both cases mentioned, property (DF$_4$) is trivially satisfied.

4.5. Continuity as a function of *x*

It is easy to formulate (D$_5$) for functions (*cf.* [4]), which results into (DF$_5$). The next theorem assures that (DF$_5$) holds when the depth D satisfies (D$_5$).

Theorem 4.14. *If D satisfies (D$_5$), then FD based on D satisfies (DF$_5$).*

Proof. Suppose that $x_\nu, x \in \mathcal{C}^K([0,1])$ and $x_\nu \xrightarrow[\nu\to\infty]{} x$. Put $f_\nu(t) = D(x_\nu(t); \mathrm{P}_t)$ and $f(t) = D(x(t); \mathrm{P}_t)$. As $x_\nu(t) \to x(t)$ for each $t \in [0,1]$, (D$_5$) implies that $f(t) \geq \limsup_{\nu\to\infty} f_\nu(t)$ for every $t \in [0,1]$. The statement then follows by application of Fatou's lemma (see *e.g.* Thm. 4.3.3 of [9]) to the sequence of functions $\{1 - f_\nu\}$, since the depth D in (2.1) is assumed to be bounded by 1 from above. $\square$

4.6. Weak continuity as a functional of P

Condition (D$_6$) can be generalized easily. See (DF$_6$). A slightly different version (for given x, without sup) of this condition was already stated and shown for the integrated dual depth Cuevas and Fraiman [4].

Theorem 4.15. *If D satisfies (D$_6$) for P_t, for almost all $t \in [0,1]$, then FD based on D satisfies (DF$_6$) at* P.

Proof. We have

$$\begin{aligned}\sup_{x\in\mathcal{C}^K([0,1])}\left|FD\left(x;\mathrm{P}_\nu,D\right)-FD\left(x;\mathrm{P},D\right)\right| &\leq \int_0^1 \sup_{x\in\mathcal{C}^K([0,1])}\left|D\left(x(t);\mathrm{P}_{\nu,t}\right)-D\left(x(t);\mathrm{P}_t\right)\right|\,\mathrm{d}\mu(t)\\ &= \int_0^1 \sup_{u\in\mathbb{R}^K}\left|D\left(u;\mathrm{P}_{\nu,t}\right)-D\left(u;\mathrm{P}_t\right)\right|\,\mathrm{d}\mu(t),\end{aligned}$$

and the last term vanishes by (D_6) and the dominated convergence theorem (see *e.g.* Thm. 4.3.5 of [9]). Notice that the measurability of the integrand function above is assured by part (iv) of Theorem 3.1. □

Note that (DF_6) implies the qualitative robustness property [16] of FD at given Q (see *e.g.* Thm. 1 of [27]). Further, as observed by Mizera ([27], Prop. 1) condition (DF_6) implies also the strong consistency of the sample depth (see (sC_1) in Sect. 5). However, condition (D_6) is very strict and strong consistency can often be proved under weaker assumptions. In Section 5 we discuss sufficient conditions under which (weak and strong) consistency of a generic functional depth is guaranteed.

5. Consistency

In this section the consistency of the integrated depth functional FD given in (2.6) is investigated. Roughly speaking, we discuss when consistency properties of the underlying finite-dimensional depth function D are inherited by its integral FD.

Recall that for a measurable space S and $\mathrm{P}\in\mathcal{P}(S)$, $\mathrm{P}_n\in\mathcal{P}(S)$ stands for an empirical measure of n independent observations from P. Following the terminology of Dudley *et al.* [10] we say that an estimator $T\colon S\times\mathcal{P}(S)\to\mathbb{R}$ satisfies

(sC_1) *strong consistency at* P if

$$\sup_{x\in S}\left|T\left(x;\mathrm{P}\right)-T\left(x;\mathrm{P}_n\right)\right| \xrightarrow[n\to\infty]{\text{a.s.}} 0;$$

(sC_2) *strong universal consistency* if (sC_1) holds for each $\mathrm{P}\in\mathcal{P}(S)$;

(sC_3) *strong uniform consistency* if for each $\varepsilon>0$

$$\sup_{\mathrm{P}\in\mathcal{P}(S)}\mathsf{P}\left(\sup_{m\geq n}\sup_{x\in S}\left|T\left(x;\mathrm{P}\right)-T\left(x;\mathrm{P}_m\right)\right|>\varepsilon\right)\xrightarrow[n\to\infty]{} 0;$$

(wC_1) *weak consistency at* P if

$$\sup_{x\in S}\left|T\left(x;\mathrm{P}\right)-T\left(x;\mathrm{P}_n\right)\right| \xrightarrow[n\to\infty]{\mathrm{P}} 0;$$

(wC_2) *weak universal consistency* if (wC_1) holds for each $\mathrm{P}\in\mathcal{P}(S)$, and finally;

(wC_3) *weak uniform consistency* if for each $\varepsilon>0$

$$\sup_{\mathrm{P}\in\mathcal{P}(S)}\mathsf{P}\left(\sup_{x\in S}\left|T\left(x;\mathrm{P}\right)-T\left(x;\mathrm{P}_n\right)\right|>\varepsilon\right)\xrightarrow[n\to\infty]{} 0.$$

Note that the following chain of implications holds:

$$\begin{array}{ccccc}(sC_3) & \Rightarrow & (sC_2) & \Rightarrow & (sC_1)\\ \Downarrow & & \Downarrow & & \Downarrow\\ (wC_3) & \Rightarrow & (wC_2) & \Rightarrow & (wC_1)\end{array}$$

Herein, the mildest notion of consistency is weak consistency (wC$_1$), in other words, convergence in probability uniformly in $x \in S$. In general, one can also consider consistency at $x \in S$, for x fixed. If one speaks about consistency in the context of data depth functions however, one most of the times means consistency to hold uniformly in $x \in S$, as this is necessary for the vast majority of applications. For instance, to show the consistency of any of the generalized medians (the point with the highest depth value), depth contours [17], or depth-based L-statistics [13], it is necessary to have consistency of depth to be guaranteed uniformly in $x \in S$. Without this concept of uniform consistency, a depth itself is of very limited practical, and theoretical, importance.

Remark 5.1. If the difference $\left|T\left(x;\mathrm{P}\right) - T\left(x;\mathrm{P}_n\right)\right|$ is bounded uniformly in $n \in \mathbb{N}$, then the convergence in probability is equivalent to L_1-convergence. In particular, condition (wC$_3$) is equivalent to

$$\sup_{\mathrm{P}\in\mathcal{P}(S)} \mathrm{E}\left[\sup_{x\in S}\left|T\left(x;\mathrm{P}\right) - T\left(x;\mathrm{P}_n\right)\right|\right] \xrightarrow[n\to\infty]{} 0.$$

This type of results was discussed in ([4], Eq. (14) in Thm. 2).

In Section 5.1 we establish conditions under which weak consistency results for generic depth functions hold. Section 5.2 is devoted to establishing sufficient conditions under which strong consistency results are guaranteed to hold. We refer the readers interested in the full details of the proofs to Appendix C where we discuss results regarding consistency properties of depths of the integrated type provided in the literature and where we point out some problematic issues in the proofs of the published results.

5.1. Weak consistency results

The following theorem states that weak consistency results for D are taken over by FD.

Theorem 5.2. *Let FD be based on a depth function D and $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$. Then the following holds:*

(i) *If (wC$_1$) is satisfied by D at P_t for all $t \in [0,1]$, then (wC$_1$) is satisfied by FD at* P.
(ii) *If (wC$_2$) is satisfied by D, then (wC$_2$) is satisfied by FD.*
(iii) *If (wC$_3$) is satisfied by D, then (wC$_3$) is satisfied by FD.*

Before we proof the theorem, note that using Remark 2.6 one can deduce that for $K = 1$ and D given by (2.8) the weak consistency result in (wC$_3$) can be derived from ([4], Thm. 2). We extend this result to $K > 1$ and an arbitrary depth measure satisfying an appropriate weak consistency property. We also pay attention to non-trivial measurability issues arising in the proof.

Proof. Similarly as in the proof of Theorem 2 of Cuevas and Fraiman [4], the dominated convergence theorem (see *e.g.* Thm. 4.3.5 of [9]) is utilized in what follows. In comparison to the original proof, appropriate attention is paid to the measurability of the involved functions as the measurability is an important assumption of the dominated convergence theorem.

Let $X_1, X_2, \dots$ be random vector-valued functions defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. We prove only parts (ii) and (iii) of the theorem. The proof of (i) is completely analogous to the proof of (ii).

Part (ii): Weak universal consistency
For a given $\omega \in \Omega$ define a function

$$S_n(t,\omega) = \sup_{u\in\mathbb{R}^K}\left|D\left(u;\mathrm{P}_t\right) - D\left(u;\mathrm{P}_{n,t}(\omega)\right)\right|. \tag{5.1}$$

By Theorem 3.1 this function is jointly universally measurable in ω and t, which further implies that for each $n \in \mathbb{N}$ the function

$$s_n(t) = \mathrm{E}\left[S_n(t,\omega)\right] = \int_\Omega \sup_{u\in\mathbb{R}^K}\left|D\left(u;\mathrm{P}_t\right) - D\left(u;\mathrm{P}_{n,t}(\omega)\right)\right| \,\mathrm{d}\mathsf{P}(\omega)$$

is universally measurable as a function of $t \in [0,1]$. It is easy to see that $s_n \colon [0,1] \to [0,1]$, $|s_n| \le 1$ and by (wC$_2$) of D on $\mathbb{R}^K$ one gets that $s_n(t) \xrightarrow[n\to\infty]{} 0$ for all $t \in [0,1]$. Now, by Fubini–Tonelli's theorem ([9], Thm. 4.4.5)

$$\begin{aligned}
\mathrm{E}\left[\sup_{x\in\mathcal{C}^K([0,1])}\left|FD\big(x;\mathrm{P},D\big)\;-FD\big(x;\mathrm{P}_n,D\big)\right|\right] &\le \mathrm{E}\left[\sup_{x\in\mathcal{C}^K([0,1])}\int_0^1\left|D\big(x(t);\mathrm{P}_t\big)-D\big(x(t);\mathrm{P}_{n,t}\big)\right|\,\mathrm{d}\mu(t)\right]\\
&\le \int_\Omega\int_0^1\sup_{u\in\mathbb{R}^K}\left|D\big(u;\mathrm{P}_t\big)-D\big(u;\mathrm{P}_{n,t}(\omega)\big)\right|\,\mathrm{d}\mu(t)\;\mathrm{d}\mathsf{P}(\omega)\\
&= \int_0^1\int_\Omega\sup_{u\in\mathbb{R}^K}\left|D\big(u;\mathrm{P}_t\big)-D\big(u;\mathrm{P}_{n,t}(\omega)\big)\right|\,\mathrm{d}\mathsf{P}(\omega)\;\mathrm{d}\mu(t)\\
&= \int_0^1 s_n(t)\;\mathrm{d}\mu(t), \qquad (5.2)
\end{aligned}$$

where the last integral tends to zero thanks to the dominated convergence theorem.

Part (iii): Weak uniform consistency

By the assumptions of the theorem and Remark 5.1

$$\sup_{\mathrm{Q}\in\mathcal{P}(\mathbb{R}^K)}\mathrm{E}\left[\sup_{u\in\mathbb{R}^K}\left|D(u;\mathrm{Q})-D(u;\mathrm{Q}_n)\right|\right]\xrightarrow[n\to\infty]{} 0. \qquad (5.3)$$

Analogously as for the proof of weak universal consistency one can make use of Fubini–Tonelli's theorem and bound

$$\begin{aligned}
&\sup_{\mathrm{P}\in\mathcal{P}(\mathcal{C}^K([0,1]))}\mathrm{E}\left[\sup_{x\in\mathcal{C}^K([0,1])}\left|FD\big(x;\mathrm{P},D\big)-FD\big(x;\mathrm{P}_n,D\big)\right|\right]\\
&\le \int_0^1\sup_{\mathrm{P}\in\mathcal{P}(\mathcal{C}^K([0,1]))}\int_\Omega\sup_{u\in\mathbb{R}^K}\left|D\big(u;\mathrm{P}_t\big)-D\big(u;\mathrm{P}_{t,n}(\omega)\big)\right|\,\mathrm{d}\mathsf{P}(\omega)\;\mathrm{d}\mu(t)\\
&\le \int_0^1\sup_{\mathrm{Q}\in\mathcal{P}(\mathbb{R}^K)}\mathrm{E}\left[\sup_{u\in\mathbb{R}^K}\left|D\big(u;\mathrm{Q}\big)-D\big(u;\mathrm{Q}_n\big)\right|\right]\,\mathrm{d}\mu(t)\\
&= \sup_{\mathrm{Q}\in\mathcal{P}(\mathbb{R}^K)}\mathrm{E}\left[\sup_{u\in\mathbb{R}^K}\left|D\big(u;\mathrm{Q}\big)-D\big(u;\mathrm{Q}_n\big)\right|\right],
\end{aligned}$$

which converges to zero by (5.3). □

When trying to show the strong (universal) consistency (sC$_1$) or (sC$_2$), it is tempting to proceed just as in part (ii) of the previous proof and simply consider the function S_n from (5.1), without taking its expectation, in the chain of inequalities (5.2). Then, the resulting integral would vanish as $n \to \infty$ just as in (5.2) and we would seemingly have shown the strong consistency without any additional effort. This is, nevertheless, a step where one needs to proceed very carefully as we are trying to establish the almost sure convergence in $\mathcal{C}^K([0,1])$ – an infinite-dimensional space. The strong consistency (sC$_1$) for $S = \mathcal{C}^K([0,1])$ and $T = FD$ must be understood in the sense of almost sure convergence on the space $\mathcal{C}^K([0,1])$, that is

$$\lim_{n\to\infty}\sup_{x\in\mathcal{C}^K([0,1])}\left|FD\big(x;\mathrm{P},D\big)-FD\big(x;\mathrm{P}_n(\omega),D\big)\right| = 0 \text{ for all } \omega\in\Omega\setminus N, \mathsf{P}(N)=0. \qquad (5.4)$$

On the other hand, if we try to substitute S_n into (5.2), the strong (universal) consistency of D gives us only that for all $t \in [0,1]$

$$\lim_{n\to\infty}\sup_{u\in\mathbb{R}^K}\left|D\big(u;\mathrm{P}_t\big)-D\big(u;\mathrm{P}_{n,t}(\omega)\big)\right| = 0 \text{ for all } \omega\in\Omega\setminus N_t, \mathsf{P}(N_t)=0.$$

One then can try to explore the structure of the set $N \subset \Omega$ of elementary events for which the pointwise convergence of the integrand function in an analogue of (5.2) is violated. Certainly the set N is a subset of a union of P_t-null sets

$$N \subset \bigcup_{t\in[0,1]} N_t. \tag{5.5}$$

However, this relation is far too weak to establish the desired property $\mathsf{P}(N) = 0$ that is necessary for the almost sure convergence in (5.4) to hold true. This important problem has not been tackled in the literature so far, and its resolution plays a crucial role in the theory of almost sure convergence on infinite-dimensional spaces. In Section 5.2 below, we show that for the particular case of the integral functional as the mapping transforming the infinite-dimensional structure $\mathcal{C}^K([0,1])$ to $\mathbb{R}$, it is possible to show that $\mathsf{P}(N) = 0$. However, in general this appears to be a challenging problem of considerable complexity.

5.2. Strong consistency results

In this section we prove strong consistency results for integrated depth (2.6), that maps an infinite-dimensional random function (3.1) into the target space $\mathbb{R}$. For this type of functionals it is possible to overcome the difficulties with the almost sure convergence outlined at the end of Section 5.1. As a basic measure-theoretic argument we use (in the proof of Thm. 5.3) the universal measurability of the function $\widetilde{h}$ from part (iv) of Theorem 3.1. This argument, however, cannot be used straightforwardly when dealing with functionals different from the integral functional in (2.6). Our main strong consistency result is provided in the following theorem.

Theorem 5.3. *Let FD be based on D and $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$.*

(i) *If* (sC$_1$) *is satisfied by D at P_t for all $t \in [0,1]$, then* (sC$_1$) *is satisfied by FD at* P.
(ii) *If* (sC$_2$) *is satisfied by D, then* (sC$_2$) *is satisfied by FD.*

Proof. Just as for Theorem 5.2, the proof of (i) follows easily from the proof of (ii). Therefore, we prove only (ii).

For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ define a random process on $[0,1] \times \Omega$

$$X(t,\omega) = \limsup_{n\to\infty} \sup_{x\in\mathcal{C}^K([0,1])} \left| D\left(x(t); \mathrm{P}_t\right) - D\left(x(t); \mathrm{P}_{n,t}(\omega)\right)\right|$$

and a corresponding collection of random variables $X(t) \equiv X(t,\cdot)\colon \Omega \to [0,1]$.

The process X is positive, bounded by 1 and universally measurable in (t,ω). The measurability follows from part (iv) of Theorem 3.1 and the measurability of the countable lim sup functional. If D satisfies (sC$_2$), then for any $t \in [0,1]$ is the set

$$N_t = \{\omega \in \Omega\colon X(t,\omega) \neq 0\}$$

of zero P-measure in Ω. Therefore, we can write by Fubini–Tonelli's theorem

$$\begin{aligned}
\mathrm{E}\left[\left|\int_0^1 X(t)\,\mathrm{d}\mu(t)\right|\right] &= \int_\Omega \left|\int_0^1 X(t,\omega)\,\mathrm{d}\mu(t)\right| \mathrm{d}\mathsf{P}(\omega) \leq \int_\Omega \int_0^1 |X(t,\omega)|\,\mathrm{d}\mu(t)\,\mathrm{d}\mathsf{P}(\omega)\\
&= \int_0^1 \int_\Omega |X(t,\omega)|\,\mathrm{d}\mathsf{P}(\omega)\,\mathrm{d}\mu(t)\\
&= \int_0^1 \left(\int_{N_t} |X(t,\omega)|\,\mathrm{d}\mathsf{P}(\omega) + \int_{\Omega\setminus N_t} |X(t,\omega)|\,\mathrm{d}\mathsf{P}(\omega)\right)\mathrm{d}\mu(t)\\
&= 0
\end{aligned}$$

and hence

$$\int_0^1 X(t)\,\mathrm{d}\mu(t) = 0 \text{ for } \mathsf{P}\text{-almost all } \omega \in \Omega.$$

Now a chain of inequalities analogous to (5.2) yields the result easily. □

5.3. Consistency of the generalized median function

With the consistency of the integrated depth FD at hand, a further interest in statistical inference goes to quantities derived from the integrated depth. Often interest goes to the maximizer of the sample depth $FD(\cdot; \mathrm{P}_n, D)$, and a major question is whether for large sample sizes n, this maximizer is close to the set of maximizers of the population version of the depth function $FD(\cdot; \mathrm{P}, D)$. A full and correct answer to this question, requires a study of the existence, the measurability and the continuity of the deepest point (function) with respect to the depth function FD defined and discussed in Section 4.2.

For $\mathcal{B}^K([0,1])$ the set of Borel measurable functions on $[0,1]$ define the set

$$FM = \Big\{ x \in \mathcal{B}^K([0,1]) \colon\ FD(x; \mathrm{P}, D) = \int_0^1 \sup_{u \in \mathbb{R}^K} D\big(u; \mathrm{P}_t\big)\, \mathrm{d}\mu(t). \Big\}$$

By Theorem 4.6 in Section 4.2 the set FM is non-empty, provided (D_4) and (D_5) hold true for D.

Note that also some care is needed when defining the sample generalized median, as the set of maximizers of $FD(\cdot; \mathrm{P}_n, D)$ does not need to contain a continuous function. But thanks again to Theorem 4.6 there exists, for each $\eta > 0$, a function $x_\eta \in \mathcal{C}^K([0,1])$ such that

$$\sup_{x \in \mathcal{B}^K([0,1])} \left\{ FD\left(x; \mathrm{P}_n, D\right) - FD\left(x_\eta; \mathrm{P}_n, D\right) \right\} < \eta. \tag{5.6}$$

Finally, to be able to formulate the result about the convergence we need to specify a distance between two vector functions on $[0,1]$. As the distance given by the norm (2.4) of the difference of the functions is not appropriate (see Example 4.13), we use here the distance

$$\mathrm{d}_\mu(x, y) = \inf \Big\{ \varepsilon > 0 : \mu\big\{ t : \|x(t) - y(t)\|_{\mathbb{R}^K} \leq \varepsilon \big\} \geq 1 - \varepsilon \Big\}. \tag{5.7}$$

In the following theorem we utilize the notions of outer convergence (see Chap. 1.9 of [38]) due to possible measurability issues.

Theorem 5.4. *Let* $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ *and* FD *be based on* D *that satisfies* (D_4) *and* (D_5). *Let* $\eta_n \searrow 0$ *and* $x_{\eta_n} \in \mathcal{C}^K([0,1])$ *be a sequence such that* (5.6) *holds (with* η *replaced with* η_n*). Put* $\mathrm{d}_\mu\left(x, FM\right) = \inf_{y \in FM} \mathrm{d}_\mu\left(x, y\right)$. *Then the following holds:*

(i) *If* D *satisfies* (wC_1) *at* P, *then*

$$\mathrm{d}_\mu\left(x_{\eta_n}, FM\right) \xrightarrow[n \to \infty]{\mathsf{P}^*} 0,$$

where P^* *denotes the convergence in outer probability.*

(ii) *If* D *satisfies* (sC_1) *at* P, *then*

$$\mathrm{d}_\mu\left(x_{\eta_n}, FM\right) \xrightarrow[n \to \infty]{\text{a.s.}*} 0,$$

where a.s.$*$ *denotes the convergence outer almost surely.*

Proof. From the proof of Theorem 4.6 it follows that the set FM is nonempty.

Fix $\varepsilon > 0$ and suppose for a moment that

$$\delta(\varepsilon) = \sup_{x \in \mathcal{C}^K([0,1])} FD(x; \mathrm{P}, D) - \sup_{x \in FM_\varepsilon^c} FD(x; \mathrm{P}, D) > 0, \tag{5.8}$$

where $FM_\varepsilon^c = \left\{ x \in \mathcal{C}^K([0,1]) : \mathrm{d}_\mu(x, FM) > \varepsilon \right\}$. Note that

$$\Big[\mathrm{d}_\mu(x_{\eta_n}, FM) > \varepsilon \Big] \subset \Big[\sup_{x \in FM_\varepsilon^c} FD(x; \mathrm{P}_n, D) \geq \inf_{x \in FM} FD(x; \mathrm{P}_n, D) - \eta_n \Big] \subset A_n \cup B_n,$$

where

$$A_n = \left[\sup_{x \in \mathcal{C}^K([0,1])} \left| FD(x; \mathrm{P}, D) - FD(x; \mathrm{P}_n, D) \right| > \tfrac{\delta(\varepsilon)}{3} \right],$$

$$B_n = \left[\tfrac{\delta(\varepsilon)}{3} + \sup_{x \in FM_\varepsilon^c} FD(x; \mathrm{P}, D) \geq -\tfrac{\delta(\varepsilon)}{3} + \inf_{x \in FM} FD(x; \mathrm{P}, D) - \eta_n \right].$$

With the help of (5.8) and the weak (strong) consistency of FD one gets $\mathsf{P}\big(A_n\big) \xrightarrow[n\to\infty]{} 0$ or $\mathsf{P}\big(\bigcup_{N=n}^{\infty} A_N\big) \xrightarrow[n\to\infty]{} 0$, respectively. For all sufficiently large n the set B_n (as well as $\bigcup_{N=n}^{\infty} B_N$) is an empty set. Thus to finish the proof it remains to show that (5.8) holds for each $\varepsilon > 0$.

Fix $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$. Let $M(t)$ be the set of points in $\mathbb{R}^K$ that for a given t maximize $D(\cdot, \mathrm{P}_t)$ (see also (4.4)) and $M_\varepsilon^c(t)$ be given by (4.5). Then by the proof of Theorem 4.7 the identifiability mapping $\nu\colon (0,\infty) \times [0,1] \to (0,\infty)$ of the depth function D defined as

$$\nu(\varepsilon, t) = \sup_{u \in \mathbb{R}^K} D\left(u; \mathrm{P}_t\right) - \sup_{u \in M_\varepsilon^c(t)} D\left(u; \mathrm{P}_t\right),$$

is positive for each $\varepsilon > 0$ and $t \in [0,1]$. Further, analogously as in the proof of the measurability of the function (3.4) in Theorem 3.1 one can argue that ν is universally measurable as a function of t for each fixed $\varepsilon > 0$.

Fix $x \in \mathcal{C}^K([0,1])$ such that $\mathrm{d}_\mu(x, FM) > \varepsilon$. Further, for $x^* \in FM$ define

$$N_{x,x^*} = \left\{ t \in [0,1] \colon \left\| x(t) - x^*(t) \right\|_{\mathbb{R}^K} > \varepsilon \right\}$$

and note that $\mu(N_{x,x^*}) > \varepsilon$ by the definition of d_μ.

Now one can bound

$$\begin{aligned} FD\left(x^*; \mathrm{P}, D\right) - FD\left(x; \mathrm{P}, D\right) &= \int_0^1 D\left(x^*(t); \mathrm{P}_t\right) \,\mathrm{d}\mu(t) - \int_0^1 D\left(x(t); \mathrm{P}_t\right) \,\mathrm{d}\mu(t) \\ &> \int_{N_{x,x^*}} \nu(\varepsilon, t) \,\mathrm{d}\mu(t) \geq \inf_{N \in \mathcal{N}_\varepsilon} \left\{ \int_N \nu(\varepsilon, t) \,\mathrm{d}\mu(t) \right\}, \end{aligned} \tag{5.9}$$

where $\mathcal{N}_\varepsilon$ is a set of universally measurable subsets of $[0,1]$ of μ-measure greater than ε. Note that the last term in (5.9) is a lower bound for $\delta(\varepsilon)$ from (5.8). Thus to finish the proof it remains to show that this term is positive.

Denote

$$N_1 = \left\{ t \in [0,1] \colon \nu(\varepsilon, t) > 1 \right\}$$

and

$$N_j = \left\{ t \in [0,1] \colon \nu(\varepsilon, t) \in \left(\tfrac{1}{j}, \tfrac{1}{j-1} \right] \right\} \quad \text{for} \quad j = 2, 3, \ldots$$

Then obviously for each $j \in \mathbb{N}$

$$\int_{N_j} \nu(\varepsilon, t) \,\mathrm{d}\mu(t) > \tfrac{1}{j}\, \mu\left(N_j\right)$$

and by positivity of $\nu(\varepsilon, t)$ the sets N_j form a pairwise disjoint partition of $[0,1]$, which further implies that $\sum_{j=1}^{\infty} \mu\left(N_j\right) = 1$. Thus it is possible to find a $J \in \mathbb{N}$ such that

$$\sum_{j=J}^{\infty} \mu(N_j) = \mu\left(\bigcup_{j=J}^{\infty} N_j \right) < \frac{\varepsilon}{2}.$$

Now one can conclude that for any $N \subset [0,1]$ such that $\mu(N) > \varepsilon$

$$\int_N \nu(\varepsilon, t)\,\mathrm{d}\mu(t) = \int_{\bigcup_{j=J}^{\infty} N_j \cap N} \nu(\varepsilon, t)\,\mathrm{d}\mu(t) + \int_{\bigcup_{j=1}^{J-1} N_j \cap N} \nu(\varepsilon, t)\,\mathrm{d}\mu(t) > 0 + \tfrac{\varepsilon}{2}\tfrac{1}{J-1} > 0. \tag{5.10}$$

Finally (5.10) implies that the right hand side of (5.9) is positive, which finishes the proof. □

By a straightforward modification of the proof one can show that if the functions in FM are uniformly bounded and also the functions $\{x_{\eta_n}\}_{n=1}^{\infty}$ are uniformly bounded, then the statement of Theorem 5.4 holds true when d_μ defined by (5.7) is replaced by the $L_1(\mu)$-distance

$$\mathrm{d}_{L^1(\mu)}(x, y) = \int_0^1 \left\| x(t) - y(t) \right\|_{\mathbb{R}^K}\,\mathrm{d}\mu(t).$$

As far as we know Theorem 5.4 is the only result regarding depth-based generalized medians in an infinite-dimensional setup. The result of Cuevas and Fraiman [4] is restricted to the case of a finite-dimensional distribution $\mathrm{P} \in \mathcal{P}\left(\mathbb{R}^K\right)$, where the compactness argument of the unit ball can be used.

Appendix A. Properties of finite-dimensional depths

In this section, we investigate the validity of the results provided in the paper for some special choices of depth:

- in Section A.1 for Tukey's halfspace depth hD as defined in (2.2);
- in Section A.2 for the simplicial depth sD defined in (2.3);
- in Section A.3 for the one-dimensional band depth defined in (A.5);
- in Section A.4 for the modified half-region depth defined in (A.8).

The main results of this section concern the validity of the measurability condition (D_7) of a given finite-dimensional depth D in (2.1). In most cases, this will be achieved by showing the joint upper semicontinuity (u.s.c.) of D in both its arguments

$$\limsup_{\nu\to\infty} D\left(u_\nu; \mathrm{Q}_\nu\right) \le D\left(u; \mathrm{Q}\right) \qquad \text{for } \mathrm{Q}_\nu \xrightarrow[\nu\to\infty]{w} \mathrm{Q} \text{ and } u_\nu \xrightarrow[\nu\to\infty]{} u. \tag{A.1}$$

As an u.s.c. function is always Borel measurable (see *e.g.* Rem. on pp. 12–13 in [20]), (A.1) implies condition (D_7). Since the non-zero condition from (D_7) is usually trivially satisfied, proving (D_7) is then reduced to establishing (A.1).

A.1. Halfspace Depth

Remark A.1. For the halfspace depth function Properties (D_1), (D_2), (D_3) and (D_4) hold true for each $\mathrm{Q} \in \mathcal{P}\left(\mathbb{R}^K\right)$, as was shown by Zuo and Serfling ([39], Thm. 2.1) and Donoho and Gasko ([7], Lem. 2.2). Property (D_5) is shown by Donoho and Gasko ([7], Lem. 6.1). If Q satisfies condition

$$\mathrm{Q}(L) = 0 \quad \text{for all hyperplanes } L \subset \mathbb{R}^K, \tag{A.2}$$

which is ensured by the absolute continuity of Q, then hD satisfies also (D_6) (see Thm. A.3) below. In conclusion, if Q is absolutely continuous, then hD has all the desired properties (D_1)–(D_6).

Mizera and Volauf ([28], Prop. 1) established (A.1) for hD, which straightforwardly implies that Property (D_7) holds for hD. This is stated formally in Theorem A.2.

Theorem A.2. *Property* (D_7) *holds for the halfspace depth* hD.

Theorem A.3. *The halfspace depth hD satisfies* (sC$_3$)*. If* (A.2) *holds, then Property* (D$_6$) *is satisfied for hD as well.*

Proof. Let us first prove that (sC$_3$) holds for the halfspace depth hD. Using definition (2.2) of hD one can bound

$$\begin{aligned}\sup_{u\in\mathbb{R}^K}\left|hD\left(u;\mathrm{Q}_n\right)-hD\left(u;\mathrm{Q}\right)\right| &= \sup_{u\in\mathbb{R}^K}\left|\inf_{H\in\mathcal{H},u\in H}\mathrm{Q}_n(H)-\inf_{H\in\mathcal{H},u\in H}\mathrm{Q}(H)\right|\\ &\leq \sup_{u\in\mathbb{R}^K}\sup_{H\in\mathcal{H},u\in H}\left|\mathrm{Q}_n(H)-\mathrm{Q}(H)\right|\\ &= \sup_{H\in\mathcal{H}}\left|\mathrm{Q}_n(H)-\mathrm{Q}(H)\right|, \end{aligned}\tag{A.3}$$

where $\mathcal{H}$ is the class of closed halfspaces on $\mathbb{R}^K$.

Now, as the class of sets $\mathcal{H}$ in $\mathbb{R}^K$ is an image admissible Suslin VC class of sets, by Dudley *et al.* ([10], Prop. 11) it is also a strong uniform Glivenko–Cantelli class. This together with (A.3) finishes this part of the proof.

Property (D$_6$) follows analogously to the same result for sD presented in ([11], Cor. 2), if one realizes that $\mathcal{H}\subset\mathcal{A}$ for $\mathcal{A}$ being a collection of sets defined in the original proof. □

Remark A.4. By applying the previously mentioned properties of hD, the integrated depth based on hD satisfies the following properties:

- For all $\mathrm{P}\in\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$: (DF$_1^{\mathrm{A}}$), (DF$_1^{\mathrm{B}}$), (DF$_2$), (DF$_3^{\mathrm{W}}$), (DF$_5$), (wC$_3$), (sC$_2$), Theorem 4.6 and part (ii) of Theorem 5.4. Property (DF$_1^{\mathrm{D}}$) is satisfied if μ is the Lebesgue measure on $[0,1]$.
- For $\mathrm{P}\in\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t satisfying (A.2): (DF$_6$).
- For $\mathrm{P}\in\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t absolutely continuous with connected support: Theorem 4.7 (see Prop. 3.5. [26]).
- For $\mathrm{P}\in\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t satisfying (U): Theorem 4.10.

A.2. Simplicial depth

Remark A.5. It is easy to see that Property (D$_1$) holds for the simplicial depth function defined in (2.3) for each $\mathrm{Q}\in\mathcal{P}\left(\mathbb{R}^K\right)$. By Liu ([21], Thm. 1) also Property (D$_4$) holds for each Q. If Q is absolutely continuous then (D$_2$) holds true ([21], Thm. 3). If Q is absolutely continuous and halfspace symmetric, then (D$_3^{\mathrm{W}}$) holds ([21], Thm. 4). However, for discrete probability distributions these latter properties may fail, as noticed by Zuo and Serfling [39]. By the derivation of Liu ([21], Thm. 2) Property (D$_5$) is satisfied for all Q. Like for the halfspace depth hD, (D$_6$) is satisfied for sD if (A.2) holds true, as follows from Dümbgen ([11], Cor. 2). In general, (D$_3$) is not satisfied for sD as noted by Mosler ([29], p. 120) and the sufficient conditions on Q for the quasi-concavity to hold for are, to the best of our knowledge, unknown.

The following theorem ensures us about the desired measurability property for the simplicial depth.

Theorem A.6. (D$_7$) *is satisfied for the simplicial depth sD.*

Proof. We show only (A.1) for sD.

Assume that $\mathrm{Q}_\nu \xrightarrow[\nu\to\infty]{w} \mathrm{Q}$ and $u_\nu\to u$ in $\mathbb{R}^K$. Let X_ν and X be random variables having the distribution Q_ν and Q respectively. Denote $\widetilde{\mathrm{Q}}_\nu$ and $\widetilde{\mathrm{Q}}$ the distribution of $X_\nu-u_\nu$ and $X-u$. Then $\widetilde{\mathrm{Q}}_\nu\to\widetilde{\mathrm{Q}}$ weakly (see *e.g.* Thm. 11.3.3(b) [9]), which further implies that the sequence of product measures $\widetilde{\mathrm{Q}}_\nu^{K+1}=\bigotimes_{k=1}^{K+1}\widetilde{\mathrm{Q}}_\nu$ converges weakly to $\widetilde{\mathrm{Q}}^{K+1}=\bigotimes_{k=1}^{K+1}\widetilde{\mathrm{Q}}$.

Put

$$F=\left\{(u_1,\ldots,u_{K+1})\in\left(\mathbb{R}^K\right)^{K+1}: 0\in\overline{co}(u_1,\ldots,u_{K+1})\right\}.$$

and note that it is a closed set. Now using the translation invariance (D_1^A) implied by (D_1) of sD (see Rem. A.5), and the Portmanteau theorem ([9], Thm. 11.1.1), one gets

$$\begin{aligned}\limsup_{\nu\to\infty} sD\big(u_\nu;\mathrm{Q}_\nu\big) &= \limsup_{\nu\to\infty} sD\big(0;\widetilde{\mathrm{Q}}_\nu\big) = \limsup_{\nu\to\infty} \widetilde{\mathrm{Q}}_\nu^{K+1}(F)\\ &\leq \widetilde{\mathrm{Q}}^{K+1}(F) = sD\big(0;\widetilde{\mathrm{Q}}\big) = sD\big(u;\mathrm{Q}\big),\end{aligned}$$

which concludes the proof. □

The sample version of the simplicial depth is strongly consistent, as is stated in Theorem A.7.

Theorem A.7. *sD satisfies* (sC_3).

Proof. One can follow the derivation of Dümbgen ([11], Thm. 1) and see that

$$\sup_{u\in\mathbb{R}^K}\big|sD(u;\mathrm{Q}_n) - sD(u;\mathrm{Q})\big| \leq (K+1)\sup_{A\in\mathcal{A}}\big|\mathrm{Q}_n(A) - \mathrm{Q}(A)\big|, \tag{A.4}$$

where $\mathcal{A}$ is the class of intersections of K closed halfspaces in $\mathbb{R}^K$.

Now, just as in the proof of Theorem A.3, the class of sets $\mathcal{A}$ in $\mathbb{R}^K$ is an image admissible Suslin VC class, which together with (A.4) finishes the proof. □

Remark A.8. By applying the previously mentioned properties of the simplicial depth sD, the integrated depth based on sD satisfies the following properties:

- For all $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$: ($DF_1^A$), ($DF_1^B$), ($DF_5$), ($wC_3$), ($sC_2$), Theorem 4.6 and part (ii) of Theorem 5.4. Condition (DF_1^D) is satisfied if μ is the Lebesgue measure on $[0,1]$.
- For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t absolutely continuous: (DF_2).
- For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t absolutely continuous and halfspace symmetric: (DF_3^W).
- For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t satisfying (A.2): (DF_6).
- For $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ with marginals P_t absolutely continuous and satisfying (U): Theorem 4.10.

A.3. (Modified) band depth

Another depth suitable for real-valued ($K = 1$) functional data was presented in reference [23]. The authors study band depth, a depth for functional data that is not of the form of an integrated depth (2.6). A modification of it, the so-called modified band depth, fits into the pattern of integrated depths. We recall the definitions of univariate band depth and modified band depth, as introduced by López-Pintado and Romo ([23], Sect. 5).

Definition A.9. Let $J = 2, 3, \ldots$. Define a univariate depth for $u \in \mathbb{R}$, $\mathrm{Q} \in \mathcal{P}(\mathbb{R})$, and $U_1, \ldots, U_J \sim \mathrm{Q}$ independent of each other, as

$$D_B^J(u;\mathrm{Q}) = \frac{1}{J-1}\sum_{j=2}^{J} \mathsf{P}\left(u \in \left[\min_{i=1,\ldots,j} U_i, \max_{i=1,\ldots,j} U_i\right]\right). \tag{A.5}$$

For $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ and $x \in \mathcal{C}([0,1])$, the modified band depth of J-th order of x with respect to P is defined as

$$BD_J(x;\mathrm{P}) = FD\left(x;\mathrm{P}, D_B^J\right),$$

with μ the Lebesgue measure on $[0,1]$.

First we show that Property (D_7) is satisfied for D_B^J, justifying the results given below.

Theorem A.10. (D$_7$) *is satisfied for univariate depth D_B^J for any $J = 2, 3, \ldots$*

Proof. Following the lines of the proof of Theorem A.6 and exploiting evident translation invariance (D$_1^{\text{A}}$) of D_B^J, define for each $j = 2, \ldots, J$ a set

$$F_j = \left\{ (u_1, \ldots, u_j) \in \mathbb{R}^j : 0 \in \left[\min_{i=1,\ldots,j} u_i, \max_{i=1,\ldots,j} u_i \right] \right\}$$

and notice that this set is closed in $\mathbb{R}^j$. The rest of the proof is a straightforward modification of the proof of Theorem A.6. $\square$

Theorem A.11. *If* Q *is absolutely continuous, then Property* (D$_2$) *is satisfied for the univariate depth D_B^J for any $J = 2, 3, \ldots$*

Proof. Denote by F the continuous distribution function of Q. Then, adhering to the notation from (A.5), we can, for any $u \in \mathbb{R}$, write

$$\begin{aligned} \mathsf{P}\left(u \in \left[\min_{i=1,\ldots,j} U_i, \max_{i=1,\ldots,j} U_i \right] \right) &= 1 - \mathsf{P}\left(u \notin \left[\min_{i=1,\ldots,j} U_i, \max_{i=1,\ldots,j} U_i \right] \right) \\ &= 1 - F(u)^j - (1 - F(u))^j, \end{aligned}$$

and

$$D_B^J(u; \mathrm{Q}) = \frac{1}{J-1} \sum_{j=2}^{J} \left(1 - F(u)^j - (1 - F(u))^j \right). \tag{A.6}$$

As the last expression attains its maximal value at $F(u) = 0.5$, the highest value of D_B^J is attained at the median of Q. In this case, the median is the only center of halfspace symmetry of Q, as already discussed in Section 4.2.2. $\square$

Theorem A.12. *If* Q *is absolutely continuous, then Property* (D$_6$) *is satisfied for the univariate depth D_B^J for any $J = 2, 3, \ldots$.*

Proof. Exploiting (A.6), for $\mathrm{Q}_\nu \xrightarrow[\nu\to\infty]{w} \mathrm{Q}$, Q_ν and Q having distribution functions F_ν and F, respectively, we may write

$$\begin{aligned} \sup_{u\in\mathbb{R}} \left| D_B^J(u; \mathrm{Q}) - D_B^J(u; \mathrm{Q}_\nu) \right| &\leq \sup_{u\in\mathbb{R}} \frac{1}{J-1} \sum_{j=2}^{J} \left| F_\nu(u)^j - F(u)^j + (1 - F_\nu(u))^j - (1 - F(u))^j \right| \\ &\leq \frac{1}{J-1} \sum_{j=2}^{J} \left(\sup_{u\in\mathbb{R}} \left| F_\nu(u)^j - F(u)^j \right| + \sup_{u\in\mathbb{R}} \left| (1 - F_\nu(u))^j - (1 - F(u))^j \right| \right) \\ &\leq \frac{1}{J-1} \sum_{j=2}^{J} \left(j \sup_{u\in\mathbb{R}} \left| F_\nu(u) - F(u) \right| + j \sup_{u\in\mathbb{R}} \left| (1 - F_\nu(u)) - (1 - F(u)) \right| \right), \end{aligned}$$

and the last term vanishes as $\nu \to \infty$ by Pólya's theorem (*cf.* Thm. 7.14 of [5]). $\square$

Remark A.13. As we restrict ourselves only to the case of $K = 1$ for band depth, Property (D$_1$) is trivial to be verified. By López-Pintado and Romo ([23], Thm. 1) Properties (D$_4$) and (D$_5$) hold true for any $\mathrm{Q} \in \mathcal{P}(\mathbb{R})$, whereas (D$_3^{\text{W}}$) is satisfied provided Q is absolutely continuous. By Theorem A.11 above also (D$_2$) is satisfied provided Q is absolutely continuous. (D$_6$) was verified in Theorem A.12 for Q absolutely continuous. (sC$_2$) has been established by López-Pintado and Romo ([23], Thm. 2).

Remark A.14. Given the previously mentioned properties of D_B^J, the modified band depth satisfies:

- For all $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$: ($\mathrm{DF}_1^{\mathrm{A}}$), ($\mathrm{DF}_1^{\mathrm{B}}$), ($\mathrm{DF}_1^{\mathrm{D}}$), ($\mathrm{DF}_5$), ($\mathrm{sC}_2$), Theorem 4.6 and part (ii) of Theorem 5.4.
- For $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ with marginals P_t absolutely continuous: (DF_2), ($\mathrm{DF}_3^{\mathrm{W}}$), ($\mathrm{DF}_6$).
- For $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ with marginals P_t absolutely continuous and satisfying (U): Theorem 4.10.

A.4. (Modified) half-region depth

In López-Pintado and Romo [24] it is suggested that the depth of real-valued ($K = 1$) random functions should be measured by a half-region depth, similar in nature to the band depth. The half-region depth itself is not an integrated depth as defined in (2.6) in Section 2.2. However López-Pintado and Romo [24] also propose a modification of the half-region depth that can be rewritten as a minimum of two members of the family of integrated depth functionals. The definition of modified half-region depth of López-Pintado and Romo [24] reads as follows.

Definition A.15. Define two univariate depths for $u \in \mathbb{R}$ and $\mathrm{Q} \in \mathcal{P}(\mathbb{R})$ as

$$\begin{aligned} D_U(u;\mathrm{Q}) &= \mathrm{Q}([u,\infty)), \\ D_L(u;\mathrm{Q}) &= \mathrm{Q}((-\infty,u]). \end{aligned} \tag{A.7}$$

Then for $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ and $x \in \mathcal{C}([0,1])$, the modified half-region depth of x with respect to P is defined as

$$MS_H(x;\mathrm{P}) = \min\left\{FD(x;\mathrm{P},D_U), FD(x;\mathrm{P},D_L)\right\}, \tag{A.8}$$

with μ the Lebesgue measure on $[0,1]$.

López-Pintado and Romo [24] state that under certain conditions imposed on P the modified half-region depth satisfies (sC_1) for P and the proofs for that are analogous of the ones for the (not modified) half-region depth. However, as shown in ([15], Sect. 3) there are some problematic issues regarding the latter proofs, and they cannot simply be adapted to the setup of the modified half-region depth. That is why in this section we prove the strong universal consistency result (sC_2) for MS_H.

To justify the reasoning in the proof of the consistency result in Theorem A.17 below, we first establish the measurability property (D_7) for the univariate depths D_U and D_L. As the couple of depths D_U and D_L evidently shares the same properties, we consider only D_U here.

Theorem A.16. *(D_7) is satisfied for the univariate depth D_U.*

Proof. Once again, we show that (A.1) holds for D_U. Let $\mathrm{Q}_\nu \xrightarrow[\nu\to\infty]{w} \mathrm{Q}$ in $\mathcal{P}(\mathbb{R})$ and $u_\nu \xrightarrow[\nu\to\infty]{} u$ in $\mathbb{R}$. As ($\mathrm{D}_1^{\mathrm{A}}$) is obviously satisfied by D_U, we may proceed again just as in the proof of Theorem A.6 for sD. The only difference is that we define F to be a set

$$F = \{u \in \mathbb{R}\colon 0 \in [u,\infty)\} = [0,\infty),$$

which is closed in $\mathbb{R}$. Statement (A.1) then follows again from the Portmanteau theorem. □

Theorem A.17. *The modified half-region depth for functions MS_H satisfies (sC_2).*

Proof. By the usual Glivenko–Cantelli's theorem ([38], Chap. 2.4) both univariate depths D_U and D_L from (A.7) satisfy (sC_2). By part (ii) of Theorem 5.3, both integrated depths based on D_U and D_L are strongly universally consistent as well. Since the modified half-space depth is defined as the minimum of the resulting integrated depths, we can write for any $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ as done in López-Pintado and Romo ([24], Prop. 4)

$$\begin{aligned} \sup_{x\in\mathcal{C}([0,1])} \left|MS_H(x;\mathrm{P}) - MS_H(x;\mathrm{P}_n)\right| \leq & \sup_{x\in\mathcal{C}([0,1])} \left|FD(x;\mathrm{P},D_U) - FD(x;\mathrm{P}_n,D_U)\right| \\ & + \sup_{x\in\mathcal{C}([0,1])} \left|FD(x;\mathrm{P},D_L) - FD(x;\mathrm{P}_n,D_L)\right|, \end{aligned}$$

with the right hand side converging to zero P-almost surely. □

As a final result of this section, we establish the weak continuity of the modified half-region depth functional.

Theorem A.18. *For* $\mathrm{P} \in \mathcal{P}(\mathcal{C}([0,1]))$ *with absolutely continuous marginals* P_t *for all* $t \in [0,1]$*,* MS_H *satisfies* (DF$_6$).

Proof. The proof is a direct consequence of Pólya's theorem (*cf*. Thm. 7.14 of [5]), the proof of Theorem 4.15 and the upper bound from the proof of Theorem A.17. □

Appendix B. Proof of Theorem 3.1

Part (i)

The mappings h and g can be written as compositions of the outer mapping D with projections

$$\mathit{\Pi}\colon [0,1] \times \mathcal{C}^K([0,1]) \times \mathcal{P}\left(\mathcal{C}^K([0,1])\right) \to \mathbb{R}^K \times \mathcal{P}\left(\mathbb{R}^K\right) : \left(t, x, \mathrm{P}\right) \to \left(x(t), \mathrm{P}_t\right),$$

and

$$\mathit{\Pi}_2\colon [0,1] \times \mathbb{R}^K \times \mathcal{P}\left(\mathcal{C}^K([0,1])\right) \to \mathbb{R}^K \times \mathcal{P}\left(\mathbb{R}^K\right) : \left(t, u, \mathrm{P}\right) \to \left(u, \mathrm{P}_t\right),$$

respectively. Suppose for a moment that both $\mathit{\Pi}$ and $\mathit{\Pi}_2$ are continuous. Then (D$_7$) yields that also h and g are jointly Borel measurable. As the proofs for both projections are analogous, we demonstrate the Borel measurability only for $\mathit{\Pi}$.

It is not obvious to see that $\mathit{\Pi}$ is continuous as the target space of $\mathit{\Pi}$ is not finite-dimensional. Assume that $t_\nu \to t$ in $[0,1]$, $x_\nu \to x$ uniformly in $\mathcal{C}^K([0,1])$ and $\mathrm{P}_\nu \to \mathrm{P}$ weakly in $\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$. Obviously $x_\nu(t_\nu) \to x(t)$ and it remains to show that also $\mathrm{P}_{t_\nu,\nu} \to \mathrm{P}_t$ weakly.

By the Portmanteau theorem ([9], Thm. 11.1.1) it is sufficient to verify that for each open subset G of $\mathbb{R}^K$:

$$\liminf_{\nu\to\infty} \mathrm{P}_{t_\nu,\nu}(G) = \liminf_{\nu\to\infty} \mathsf{P}\left(X_\nu(t_\nu) \in G\right) \geq \mathsf{P}\left(X(t) \in G\right) = \mathrm{P}_t(G), \tag{B.1}$$

where $X_\nu, X \in \mathcal{C}^K([0,1])$ are random functions with distributions P_ν and P, respectively. To verify (B.1) one needs to show that for each $\eta > 0$ there exists $\nu_\eta \in \mathbb{N}$, depending on G, such that for all $\nu > \nu_\eta$

$$\mathsf{P}\left(X_\nu(t_\nu) \in G\right) \geq \mathsf{P}\left(X(t) \in G\right) - \eta. \tag{B.2}$$

Fix $G \subset \mathbb{R}^K$ open and $\eta > 0$. Further, let $\varepsilon > 0$ and note that the following subset of $\mathcal{C}^K([0,1])$

$$\mathcal{G}_\varepsilon = \left\{y \in \mathcal{C}^K([0,1]) : \, y(t') \in G \text{ for all } |t'-t| \leq \varepsilon\right\}$$

is open. Thus by $t_\nu \to t$, $\mathrm{P}_\nu \xrightarrow{w} \mathrm{P}$ and the Portmanteau theorem there exists $\nu_{\varepsilon,\eta} \in \mathbb{N}$ such that for all $\nu \geq \nu_{\varepsilon,\eta}$ it holds

$$\mathsf{P}\left(X_\nu(t_\nu) \in G\right) \geq \mathsf{P}\left(X_\nu \in \mathcal{G}_\varepsilon\right) \geq \mathsf{P}\left(X \in \mathcal{G}_\varepsilon\right) - \tfrac{\eta}{2}. \tag{B.3}$$

Further, as $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$

$$\lim_{\varepsilon\to 0+} \mathsf{P}\left(X \in \mathcal{G}_\varepsilon\right) = \mathsf{P}\left(X(t) \in G\right),$$

which together with (B.3) implies (B.2) and finishes the proof.

Part (ii)

Follows immediately from (i), as all the marginal functions of a jointly Borel measurable function on a product space are Borel measurable (*cf*. pp. 118–119 in [9]).

Part (iii)

By part (ii) the set

$$\{(t,u) : D(u;\mathrm{P}_t) \leq a\} \subset [0,1]\times\mathbb{R}^K$$

is for any $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and $a \in [0,1]$ Borel measurable. Intersecting this set with the graph $\{(t,x(t)) : t\in[0,1]\}$ of a Borel measurable function $x\colon [0,1] \to \mathbb{R}^K$, which is a Borel measurable set in $[0,1]\times\mathbb{R}^K$ ([9], Lem. 13.2.2) we obtain that the set

$$\{(t,x(t)) : D(x(t);\mathrm{P}_t) \leq a\} \subset [0,1]\times\mathbb{R}^K$$

is a Borel measurable set. The projection of the latter set into the first coordinate

$$\{t\in[0,1] \colon D(x(t);\mathrm{P}_t) \leq a\} \subset [0,1]$$

is analytic ([9], Thm. 13.2.1), thus universally measurable ([9], Lem. 13.2.6), in $[0,1]$. As we have just shown that all lower level sets of the function

$$D(x(\cdot);\mathrm{P}_\cdot) : [0,1] \to [0,1]$$

are universally measurable sets, the assertion follows.

Part (iv)

Adopting the notation of Chapter 5.3 of [8] one can write the function $h(t,x,\mathrm{P}) = X(\omega,f)$, where $\omega \in \Omega = [0,1]\times\mathcal{P}\left(\mathcal{C}^K([0,1])\right)\times\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and $f\in\mathcal{F} = \mathcal{C}^K([0,1])$. As $\mathcal{C}^K([0,1])$ with its Borel σ-algebra is a Polish space, then by *(i)* the function h is an image admissible Suslin (*via* an identity mapping). Analogously, the function $g(t,x,\mathrm{P},\widetilde{\mathrm{P}}) = h(t,x,\mathrm{P}) - h(t,x,\widetilde{\mathrm{P}})$ defined on $[0,1]\times\mathcal{C}^K([0,1])\times\mathcal{P}\left(\mathcal{C}^K([0,1])\right)\times\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ is image admissible Suslin, as g can be written as $g(t,x,\mathrm{P},\widetilde{\mathrm{P}}) = X(\omega,f)$ with $\Omega = [0,1]\times\mathcal{P}\left(\mathcal{C}^K([0,1])\right)\times\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ and $\mathcal{F} = \mathcal{C}^K([0,1])$. By Corollary 5.3.5 of Dudley [8] the function

$$\sup_{f\in\mathcal{F}}\left|X(\omega,f)\right| = \sup_{x\in\mathcal{C}^K([0,1])}\left|g(t,x,\mathrm{P},\widetilde{\mathrm{P}})\right| = \sup_{x\in\mathcal{C}^K([0,1])}\left|D\big(x(t);\mathrm{P}_t\big) - D\big(x(t);\widetilde{\mathrm{P}}_t\big)\right|$$

is universally measurable as a function of ω, which finishes the proof of (iv).

Part (v)

By (ii) g_2 is jointly Borel measurable for any $\mathrm{P}\in\mathcal{P}\left(\mathcal{C}^K([0,1])\right)$. The proof is completed by following the reasoning of the proof of (iv), with $g_2(t,u) = X(\omega,f)$, where $\Omega = [0,1]$ and $\mathcal{F} = \mathbb{R}^K$. □

Appendix C. Discussion of the available consistency results

Here, we briefly discuss the strong consistency results of Fraiman and Muniz [14] and Claeskens *et al.* [2]. We point out some problematic issues, illustrating some of these with examples.

C.1. Integrated depth of Fraiman and Muniz [14]

The consistency result of ([14], Thm. 3.1) reads as follows.

Result 1. *Let μ be the Lebesgue measure on $[0,1]$, $K=1$ and $X\sim\mathrm{P}\in\mathcal{P}(\mathcal{C}([0,1]))$ satisfy the conditions*

(H1) *For a constant $A>0$ let the set of uniformly Lipschitz continuous functions*

$$\mathrm{Lip}_A = \{x\in\mathcal{C}([0,1]) : |x(t)-x(s)| \leq A\,|t-s| \text{ for } s,t\in[0,1]\}.$$

be the space of functions where the stochastic process X takes its values.

(H2) *For any* $\varepsilon > 0$ *there exists* $c > 0$ *so that for all* $u \in \mathrm{Lip}_A$

$$\mathrm{E}\left[\lambda\left(\{t \in [0,1] \colon X(t) \in [u(t), u(t) + c\varepsilon]\}\right)\right] < \varepsilon/2,$$

where λ *stands for the Lebesgue measure on* $\mathbb{R}$.

Then for the choice

$$D(u; \mathrm{Q}) = 1 - \left|\frac{1}{2} - \mathrm{Q}\big((-\infty, u]\big)\right| \text{ for } u \in \mathbb{R} \text{ and } \mathrm{Q} \in \mathcal{P}(\mathbb{R})$$

FD *satisfies* (sC$_1$) *with the supremum taken over* Lip_A.

We now briefly indicate some key argumentations in the original proof of Result 1.

Argumentation in the proof of Result 1.
Initially, the authors show that under conditions (H1) and (H2) the statistic

$$F(x, \mathrm{P}) = \int_0^1 \mathrm{P}_t\big((-\infty, x(t)]\big)\,\mathrm{d}t \text{ for } x \in \mathrm{Lip}_A, \mathrm{P} \in \mathcal{P}(\mathrm{Lip}_A) \tag{C.1}$$

is strongly universally consistent over the set Lip_A. The strong universal consistency of FD is then claimed to follows immediately. □

Looking carefully into the argumentation, it is seen though that by the reverse triangle inequality one only obtains the following

$$\begin{aligned}
&\sup_{x \in \mathrm{Lip}_A} |FD(x; \mathrm{P}, D) - FD(x; \mathrm{P}_n, D)| \\
&\quad = \sup_{x \in \mathrm{Lip}_A} \left| \int_0^1 \left\{1 - \left|\frac{1}{2} - \mathrm{P}_t\big((-\infty, x(t)]\big)\right|\right\} \mathrm{d}t - \int_0^1 \left\{1 - \left|\frac{1}{2} - \mathrm{P}_{n,t}\big((-\infty, x(t)]\big)\right|\right\} \mathrm{d}t \right| \\
&\quad \leq \sup_{x \in \mathrm{Lip}_A} \int_0^1 \left| \left|\frac{1}{2} - \mathrm{P}_{n,t}\big((-\infty, x(t)]\big)\right| - \left|\frac{1}{2} - \mathrm{P}_t\big((-\infty, x(t)]\big)\right| \right| \mathrm{d}t \\
&\quad \leq \sup_{x \in \mathrm{Lip}_A} \int_0^1 \left|\mathrm{P}_t\big((-\infty, x(t)]\big) - \mathrm{P}_{n,t}\big((-\infty, x(t)]\big)\right| \mathrm{d}t,
\end{aligned}$$

where it cannot be said that the last term necessarily vanishes for $n \to \infty$, because of the fact that the strong universal consistency of F defined in (C.1) only means that

$$\sup_{x \in \mathrm{Lip}_A} \left| \int_0^1 \mathrm{P}_t\big((-\infty, x(t)]\big) - \mathrm{P}_{n,t}\big((-\infty, x(t)]\big)\,\mathrm{d}t \right| \xrightarrow[n \to \infty]{\text{a.s.}} 0$$

and the absolute value cannot, in general, be put into the integral.

C.2. Multivariate functional depth of Claeskens *et al.* [2]

In this section, we shall return to the original multivariate functional depth as defined in (2.5) and its sample version based on discretely observed random samples of functions, as described in (2.7) in Remark 2.4. Only in this section w, the weight function used in (2.5), is allowed to depend on P and the random sample P_n. This dependency will be highlighted by the use of additional arguments of w. In Claeskens *et al.* ([2], Thm. 3) the following result can be found.

Result 2. *Let* $X_1, \dots, X_n \in \mathcal{C}^K([0,1])$ *be a random sample from* $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$, $X \sim \mathrm{P}$, *with* $\mathrm{E}[X(t)]$ *finite for each* $t \in [0,1]$. *The curves* $X_1, \dots, X_n$ *are observed only at points* $t_1 < t_2 < \dots < t_T$, *which for a fixed distribution function* $G\colon [0,1] \to [0,1]$ *are given by*

$$t_j = G^{-1}\left(\frac{j-1}{T-1}\right) \tag{C.2}$$

and $G^{-1}(0) = 0$, $G^{-1}(1) = 1$. *Assume that* G *is twice differentiable and its first derivative is bounded away from zero on* $[0,1]$. *Then it holds that*

$$\sup_{x \in \mathcal{C}^K([0,1])} \left| MFD\left(x; \mathrm{P}, D\right) - MFD_n\left(x; D\right) \right| \xrightarrow[n\to\infty, T\to\infty]{\text{a.s.}} 0,$$

provided that D *satisfies* (D_1), (sC_1) *and*

$$\int_0^1 \left| w\left(t; \widetilde{\mathrm{P}}_{n,t}\right) - w\left(t; \mathrm{P}_t\right) \right| \,\mathrm{d}t \xrightarrow[n\to\infty]{\text{a.s.}} 0,$$

where $\widetilde{\mathrm{P}}_n \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ *is an empirical measure of random functions defined for* $j = 1, \dots, T-1$, $k = 1, \dots, K$, *as*

$$\widetilde{X}_i^{(k)}(t) = \begin{cases} X_i{}^{(k)}(t_j)\frac{t_j + t_{j+1} - 2t}{t_{j+1} - t_j} + \bar{X}^{(k)}(t_j)\frac{2(t - t_j)}{t_{j+1} - t_j} & \text{for } t \in [t_j, (t_j + t_{j+1})/2] \\ -X_i{}^{(k)}(t_{j+1})\frac{t_j + t_{j+1} - 2t}{t_{j+1} - t_j} - \bar{X}^{(k)}(t_j)\frac{2(t - t_{j+1})}{t_{j+1} - t_j} & \text{for } t \in [(t_j + t_{j+1})/2, t_{j+1}]\,. \end{cases} \tag{C.3}$$

Here $\bar{X}^{(k)}(t)$ *denotes the* k*th coordinate of the average of the functional values of the random sample* $\bar{X}(t) = \frac{1}{n}\sum_{i=1}^n X_i(t) \in \mathbb{R}^K$.

Note that the design setup used by Claeskens *et al.* [2] enables to circumvent the problem described at the end of Section 5.1 as it is sufficient to consider only a countable union of null sets in (5.5). On the other hand it raises several other questions which are not easy to answer as the proof of Result 2 is not written in full detail in [2]. In what follows we would like to point out a potential problem with the interpolation processes $\widetilde{X}_i^{(k)}$ as introduced in (C.3).

A key ingredient in the proof of Result 2 is the statement that the empirical measure of the interpolated functional variable $\widetilde{\mathrm{P}}_n \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ as defined in (C.3) converges weakly to the original measure $\mathrm{P} \in \mathcal{P}\left(\mathcal{C}^K([0,1])\right)$ as n and T tend simultaneously to infinity. As an instrument to show this, it was necessary to establish for fixed $c > 0$ the inequality

$$\sup_{t\in[0,1]} \left\| \widetilde{X}_i(t) - X_i(t) \right\|_{\mathbb{R}^K} \le 2 \sup_{|s-t| \le c/(T-1)} \left\| X_i(t) - X_i(s) \right\|_{\mathbb{R}^K}. \tag{C.4}$$

Via the following example we illustrate that an inequality like this does however not hold true in general, revealing as such a potential problem.

Example C.1. Let the design distribution function G from (C.2) be a distribution function of a random variable uniformly distributed on $[0,1]$ and let $K = 1$. Define $X \sim \mathrm{P} \in \mathcal{P}\left(\mathcal{C}\left([0,1]\right)\right)$ as

$$\begin{aligned} \mathsf{P}\big(X(t) = 1 \text{ for all } t \in [0,1]\big) &= 0.5, \\ \mathsf{P}\big(X(t) = 0 \text{ for all } t \in [0,1]\big) &= 0.5. \end{aligned}$$

Take a random sample of size n from P and denote the number of constant one functions in the sample by $m = m(n) \le n$. Without loss of generality, we may assume that precisely the first $n - m$ functions of the random sample are constant zeroes.

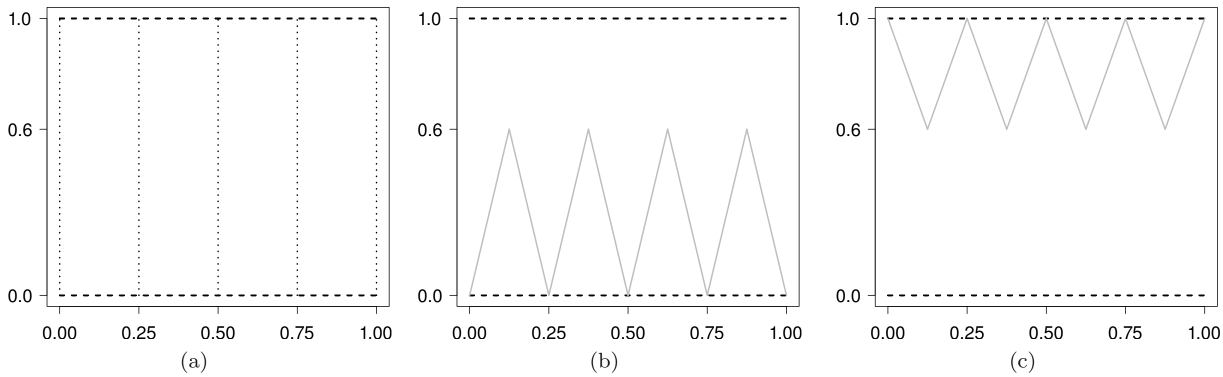

FIGURE C.1. (a) A random sample of $n = 10$ functions from distribution P (it is not visible on the plot, but $m = 6$) from Example C.1 in dashed line along with points t_j, $j = 1, \dots, T$, $T = 5$; (b) approximated random function $\widetilde{X}_i$ for $i = 1, \dots, 4$; and (c) for $i = 5, \dots, 10$, in solid gray lines.

For $T \in \mathbb{N}$ arbitrary, the design points t_j, $j = 1, \dots, T$, at which the sample functions $X_1, \dots, X_n$ are observed are equidistantly spaced with $t_1 = 0$, $t_T = 1$. For any t_j, the average of observed functional values of the random sample is

$$\bar{X}(t_j) = \frac{m}{n} \text{ for all } j = 1, \dots, T.$$

Thus, the approximating functions (C.3) take the form for each $j = 1, \dots, T-1$

$$\widetilde{X}_i(t) = \begin{cases} \frac{m}{n}\frac{2(t-t_j)}{t_{j+1}-t_j} & \text{for } t \in [t_j, (t_j + t_{j+1})/2] \\ -\frac{m}{n}\frac{2(t-t_{j+1})}{t_{j+1}-t_j} & \text{for } t \in [(t_j + t_{j+1})/2, t_{j+1}] \end{cases} \text{ for } i = 1, \dots, m,$$

$$\widetilde{X}_i(t) = \begin{cases} \frac{t_j+t_{j+1}-2t}{t_{j+1}-t_j} + \frac{m}{n}\frac{2(t-t_j)}{t_{j+1}-t_j} & \text{for } t \in [t_j, (t_j + t_{j+1})/2] \\ -\frac{t_j+t_{j+1}-2t}{t_{j+1}-t_j} - \frac{m}{n}\frac{2(t-t_{j+1})}{t_{j+1}-t_j} & \text{for } t \in [(t_j + t_{j+1})/2, t_{j+1}] \end{cases} \text{ for } i = m+1, \dots, n.$$

A random sample of size $n = 10$ generated from P is depicted in Figure C.1a, together with the corresponding functions $\widetilde{X}_i$ (in Figs. C.1b and C.1c).

Take now an arbitrary function from the random sample, for instance X_1. Then it can be easily verified that the left hand side of (C.4) is

$$\sup_{t\in[0,1]} \left\| \widetilde{X}_1(t) - X_1(t) \right\|_{\mathbb{R}^1} = \left| \widetilde{X}_1\left(\frac{t_1+t_2}{2}\right) - X_1\left(\frac{t_1+t_2}{2}\right) \right| = \frac{m}{n}.$$

On the other hand, as P takes values only in constant functions, for the right hand side for any $c > 0$

$$2 \sup_{|s-t|\leq c/(T-1)} \|X_i(t) - X_i(s)\|_{\mathbb{R}^1} = 0$$

and inequality (C.4) does not hold true for any random sample such that $m > 0$.

Moreover, (C.4) is not even valid asymptotically, since the preceding formulae do not depend on the choice of T, and as $n \to \infty$ one can only conclude that the left hand side tends P-almost surely to 0.5, while the right hand side is still almost surely zero.

The setup in Example C.1 can be further utilized to show that it does not generally hold that

$$\widetilde{\mathrm{P}}_n \xrightarrow[n,T\to\infty]{w} \mathrm{P} \quad \text{in } \mathcal{P}\left(\mathcal{C}^K([0,1])\right) , \tag{C.5}$$

a key argument used in the proof of the consistency of MFD of Claeskens *et al.* ([2], Thm. 3).

Acknowledgements. The authors thank the Editor, an Associate Editor and two anonymous referees for their valuable comments on the original version of the manuscript. This research is supported by the IAP research network nr. P7/06 of the Federal Science Policy (Belgium). The work of the first author was partially supported by the Czech Science Foundation Project No. P402/12/G097 “DYME – Dynamic Models in Economics”. Currently he is a Research Assistant of the Research Foundation – Flanders, and acknowledges support from this foundation. The second author gratefully acknowledges support from the GOA/12/014 – project of the Research Fund KU Leuven. The third author gratefully acknowledges support from the Grant GACR 15-04774Y.

References

[1] L.D. Brown and R. Purves, Measurable selections of extrema. *Ann. Statist.* **1** (1973) 902–912.

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

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

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

[5] A. DasGupta, Probability for statistics and machine learning: Fundamentals and advanced topics. *Springer Texts in Statistics*. Springer, New York (2011).

[6] V.H. de la Peña and E. Giné, Decoupling. From dependence to independence. *Probability and its Applications*. Springer-Verlag, New York (1999).

[7] D.L. Donoho and M. Gasko, Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann. Statist.* **20** (1992) 1803–1827.

[8] R.M. Dudley, Uniform central limit theorems. Vol. 63 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (1999).

[9] R.M. Dudley, Real analysis and probability. Vol. 74 of *Cambridge Studies in Advanced Mathematics*. Revised reprint of the 1989 original. Cambridge University Press, Cambridge (2002).

[10] R.M. Dudley, E. Giné and J. Zinn, Uniform and universal Glivenko-Cantelli classes. *J. Theoret. Probab.* **4** (1991) 485–510.

[11] L. Dümbgen, Limit theorems for the simplicial depth. *Statist. Probab. Lett.* **14** (1992) 119–128.

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

[13] R. Fraiman and J. Meloche, Multivariate L-estimation. With comments and a rejoinder by the authors. *Test* **8** (1999) 255–317.

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

[15] I. Gijbels and S. Nagy, Consistency of non-integrated depths for functional data. *J. Multivariate Anal.* **140** (2015) 259–282.

[16] F.R. Hampel, A general qualitative definition of robustness. *Ann. Math. Statist.* **42** (1971) 1887–1896.

[17] X. He and G. Wang, Convergence of depth contours for multivariate datasets. *Ann. Statist.* **25** (1997) 495–504.

[18] L. Kong and Y. Zuo, Smooth depth contours characterize the underlying distribution. *J. Multivariate Anal.* **101** (2010) 2222–2226.

[19] J. Kuelbs and J. Zinn, Concerns with functional depth. *ALEA Latin Am. J. Prob. Math. Statist.* **10** (2013) 831–855.

[20] E.H. Lieb and M. Loss, Analysis. Vol. 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI. 2nd edition (2001).

[21] R.Y. Liu, On a notion of data depth based on random simplices. *Ann. Statist.* **18** (1990) 405–414.

[22] R.Y. Liu, J.M. Parelius and K. Singh, Multivariate analysis by data depth: descriptive statistics, graphics and inference. With discussion and a rejoinder by Liu and Singh. *Ann. Statist.* **27** (1999) 783–858.

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

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

[25] S. López-Pintado, Y. Sun, J. Lin and M.G. Genton, Simplicial band depth for multivariate functional data. *Adv. Data Anal. Classif.* **8** (2014) 321–338.

[26] J.-C. Massé and R. Theodorescu, Halfplane trimming for bivariate distributions. *J. Multivariate Anal.* **48** (1994) 188–202.

[27] I. Mizera, Qualitative robustness and weak continuity: the extreme unction? In Nonparametrics and robustness in modern statistical inference and time series analysis: a Festschrift in honor of Professor Jana Jurečková. Vol. 7 of *Inst. Math. Stat. Collect.* Inst. Math. Statist., Beachwood, OH (2010) 169–181.

[28] I. Mizera and M. Volauf, Continuity of halfspace depth contours and maximum depth estimators: diagnostics of depth-related methods. *J. Multivariate Anal.* **83** (2002) 365–388.

[29] K. Mosler, Multivariate dispersion, central regions and depth: The lift zonoid approach. Vol. 165 of *Lect. Notes Stat.* Springer-Verlag, Berlin (2002).

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

[31] K. Mosler and Y. Polyakova, General notions of depth for functional data. arXiv:1208.1981 (2012).

[32] S. Nagy, Coordinatewise characteristics of functional data. In *Proc. 31th Int. Conf. Mathematical Methods in Economics 2013, Jihlava, Czech Republic*, edited by H. Vojáčková. (Part II). College of Polytechnics Jihlava, September (2013) 655–660.

[33] J.O. Ramsay and B.W. Silverman, Functional data analysis. *Springer Series in Statistics*, 2nd edition. Springer, New York (2005).

[34] A.W. Roberts and D.E. Varberg, Convex functions. Vol. 57 of *Pure and Applied Mathematics*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London (1973).

[35] R. Serfling, Depth functions in nonparametric multivariate inference. In Data depth: robust multivariate analysis, computational geometry and applications. Vol. 72 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* Amer. Math. Soc., Providence, RI (2006) 1–16.

[36] R. Serfling, Multivariate symmetry and asymmetry. Vol. 8 of *Encyclopedia of Statistical Sciences*, 2nd edition (2006) 5338–5345.

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

[38] A.W. van der Vaart and J.A. Wellner, Weak convergence and empirical processes. *Springer Series in Statistics*. Springer-Verlag, New York (1996).

[39] Y. Zuo and R. Serfling, General notions of statistical depth function. *Ann. Statist.* **28** (2000) 461–482.

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