Theory of Stochastic Processes
Vol. 20 (36), no. 2, 2015, pp. 13–41

D. FERGER

ARGINF-SETS OF MULTIVARIATE CADLAG PROCESSES AND THEIR CONVERGENCE IN HYPERSPACE TOPOLOGIES

Let $X_n, n \in \mathbb{N}$, be a sequence of stochastic processes with trajectories in the multivariate Skorokhod-space $D(\mathbb{R}^d)$. If $A(X_n)$ denotes the set of all infimizing points of X_n, then $A(X_n)$ is shown to be a random closed set, i.e. a random variable in the hyperspace $\mathcal{F}$, which consists of all closed subsets of $\mathbb{R}^d$. We prove that if X_n converges to X in $D(\mathbb{R}^d)$ in probability, almost surely or in distribution, then $A(X_n)$ converges in the analogous manner to $A(X)$ in $\mathcal{F}$ endowed with appropriate hyperspace topologies. Our results immediately yield continuous mapping theorems for measurable selections $\xi_n \in A(X_n)$. Here we do not require that $A(X)$ is a singleton as it is usually assumed in the literature. In particular it turns out that ξ_n converges in distribution to a Choquet capacity, namely the capacity functional of $A(X)$. In fact, this motivates us to extend the classical concept of weak convergence. In statistical applications it facilitates the construction of confidence regions based on M-estimators even in the case that the involved limit process has no longer an a.s. unique infimizer as it was necessary so far.

1. Introduction

For some fixed natural number d let $X = \{X(t) : t \in \mathbb{R}^d\}$ be a real-valued stochastic process defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with trajectories in the *multivariate Skorokhod-space* $D = D(\mathbb{R}^d)$ which is defined as follows. If $R = (R_1, \ldots, R_d) \in \{<, \geq\}^d$ is a ordered list of the usual relations $<$ and $\geq$ in $\mathbb{R}$ and $t = (t_1, \ldots, t_d) \in \mathbb{R}^d$ is a point in the euclidean space, then

$$Q_R := Q_R(t) := \{s \in \mathbb{R}^d : s_i R_i t_i, 1 \leq i \leq d\}$$

is the *R-quadrant of t*. Given a function $f : \mathbb{R}^d \to \mathbb{R}$ the quantity

$$f(t + R) := \lim_{s \to t, s \in Q_R(t)} f(s)$$

is called the *R-quadrant-limit of f in t*. Then $D(\mathbb{R}^d)$ consists of all functions f such that for each $t \in \mathbb{R}^d$

(a) $f(t + R)$ exists for all $R \in \{<, \geq\}^d$,
(b) $f(t + R) = f(t)$ for $R = (\geq, \ldots, \geq)$.

Relations (a) and (b) extend the notions „limits from below" and „continuous from above" from the univariate case (d=1) to the multivariate one. Therefore it is convenient to call $f \in D$ a *cadlag function* (continue à droite limite à gauche). $D(\mathbb{R}^d)$ endowed with the *Skorokhod-metric* s is a complete separable metric space, confer [12], p.332. The pertaining Borel-σ-algebra $\mathcal{D}$ is generated by the sets of all cylinders, confer Theorem 2 of [12], whence X can be considered as a random element $X : (\Omega, \mathcal{A}) \to (D, \mathcal{D})$.

2010 *Mathematics Subject Classification.* 60F05, 49J53 ; 60B05, 60B10.
Key words and phrases. Multivariate Skorokhod-space, weak convergence, epi-convergence, sets of infimizing points, random closed sets, hyperspace topologies, Choquet capacity, continuous mapping theorems.

The main object in this paper is the random set Arginf(X) of all *infimizers* of X, where

$$\text{Arginf}(f) \equiv A(f) := \{t \in \mathbb{R}^d : \min_{R \in \{<,\geq\}^d} f(t+R) = \inf_{s \in \mathbb{R}^d} f(s)\}, \quad f \in D. \tag{1}$$

We will see that A(f) is a closed subset of $\mathbb{R}^d$ (possibly empty). Let $\mathcal{F} = \mathcal{F}(\mathbb{R}^d)$ denote the family of all closed subsets of $\mathbb{R}^d$ (including the empty set $\emptyset$) and let $\mathcal{F}$ be equipped with some appropriate topology τ and the pertaining Borel-σ-algebra $\mathcal{B}_\tau$. It turns out that $A(X)$ is a random element in the space $(\mathcal{F}, \mathcal{B}_\tau)$. We focus on the following problem:

If (X_n) is a sequence of random elements in $(D, \mathcal{D})$ converging to some limit X in distribution, in probability or almost surely, respectively, then under which conditions does this entail the corresponding convergence of the arginf-sets?

In other words we want to formulate versions of the Continous Mapping Theorem for the Arginf-functional, which by (1) is a well-defined mapping $A : D \to \mathcal{F}$.

Let D' be the collection of all $f \in D$ with $A(f) \neq \emptyset$. Then by the axiom of choice there exists a function $a : D' \to \mathbb{R}^d$ such that $a(f) \in A(f)$ for every $f \in D'$. This means via a we can chose a single infimizing point of every f and denote it alternatively by arginf$(f) \equiv a(f)$. Once again the question arises which types of continuous mapping theorems are valid for the functional $a : D' \to \mathbb{R}^d$.

Example 1.1. Our motivation for considering these problems stems from statistics. Here, the well-known principle of *M-estimation* yields estimators $\hat{\theta}_n$ defined as infimizing point of some random criterion function $M_n(t), t \in \mathbb{R}^d$, that is $\hat{\theta}_n := a(M_n)$ estimates a (unique) parameter $\theta = a(M)$ where M is some theoretical criterion function. In non-parametric statistics the involved criterion functions typically are cadlag and its tempting to conclude as follows: If, for instance, $M_n \to M$ a.s. then hopefully $a(M_n) \to a(M)$ and $A(M_n) \to A(M)$ a.s. Clearly, we wish the convergence $M_n \to M$ to be as weak as possible and in case of set-convergence the topology on the *hyperspace* $\mathcal{F}$ to be as large as possible. It turns out that epi-convergence seems to be the minimal setup to make the above conclusions to become true as long as a certain compactness condition is fulfilled. In applications to establish epi-convergence of the stochastic processes M_n is hard to prove or even intractable. In contrast, the Skorokhod-convergence is much easier to handle, because there is a nice equivalent characterization. A crucial result of our paper is that convergence in the Skorokhod-metric implies epi-convergence, whence we eventually obtain sufficient and manageable criterions for proving consistency of $(\hat{\theta}_n)$. Moreover we want to derive convergence in distribution, namely $\Gamma_n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \xi$ in $\mathbb{R}^d$ for some positive diagonal $d \times d$-matrices Γ_n and with limit variable ξ we wish to identify. The basic idea here is to introduce *rescaled* M_n-processes defined as

$$X_n(t) = \gamma_n\{M_n(\theta + \Gamma_n^{-1}t) - M_n(\theta)\},\ t \in \mathbb{R}^d,$$

or alternatively as

$$X_n(t) = \gamma_n \frac{M_n(\theta + \Gamma_n^{-1}t)}{M_n(\theta)},\ t \in \mathbb{R}^d,$$

with some *normalizing* sequence (γ_n) of positive real numbers. As a consequence of the transformation in time $t \mapsto \theta + \Gamma_n^{-1}t$ one obtains that $\Gamma_n(\hat{\theta}_n - \theta) = a(X_n)$ and that $\Gamma_n(A(M_n) - \theta) = A(X_n)$. Thus if $X_n \xrightarrow{\mathcal{L}} X$ in (D, s) entailed distributional convergence of $a(X_n)$ and $A(X_n)$ we had a powerful tool to find the limit distributions of M-estimators of Euclidean parameters.

The paper is organized as follows: In section 2 the functional $A : D \to \mathcal{F}$ and the choice function $a : D' \to \mathbb{R}^d$ are investigated as to various continuity properties. For that purpose it is necessary to introduce two topologies on $\mathcal{F}$, namely the *Fell-topology* τ_{Fell} and the *missing-topology* τ_{miss}, and to use some of their properties. In this context $\mathcal{F}$ is called *hyperspace* and these underlying topologies (but also others) are known as *hyperspace topologies*, where the Fell-topology is strictly stronger than the missing-topology. It turns out that convergence of f_n to f in (D, s) ensures that of $A(f_n)$ to $A(f)$ in the hyperspace $(\mathcal{F}, \tau_{miss})$. As a useful consequence measurability of A is guaranteed. If in addition

$$A(f_n) \cap K \neq \emptyset \text{ for eventually all } n \in \mathbb{N}, \tag{2}$$

where $K \subseteq \mathbb{R}^d$ is some compact set, and if $A(f)$ is a singleton then actually $A(f_n) \to A(f)$ in $(\mathcal{F}, \tau_{Fell})$. Once we have continuity of A as described above we obtain several continuity properties for the choice function a. Under the assumption (2) for each selection $a(f_n) \in K$ it follows that $d(a(f_n), A(f)) \to 0$, where $d(x, B)$ is the usual distance of a point $x \in \mathbb{R}^d$ to a subset B of $\mathbb{R}^d$. Thus, if $a(f)$ is the unique infimizer of f, then $a(f_n) \to a(f)$. All of this requires several analytical notions and new results on the Skorokhod-space (D, s) and in particular its relation to the space $SC(\mathbb{R}^d)$ of lower semicontinuous (lsc) functions endowed with the metrik e of *epi-convergence*. Namely, if $\bar{f}$ denotes the *lower-semicontinuous regularization* of f it is shown that the map $f \mapsto \bar{f}$ is a continuous injection from $(D(\mathbb{R}^d), s)$ into $(SC(\mathbb{R}^d), e)$. The analytical results in section 2 are in turn fundamental for the main part of the paper, section 3. Here, we derive a collection of probabilistic limit theorems for $A(X_n)$ and $a(X_n)$. The first part deals with almost sure convergence and in probability. It smoothly follows that if $X_n \to X$ a.s. then $A(X_n) \to A(X)$ in $(\mathcal{F}, \tau_{miss})$ a.s. and if in addition the above conditions analogously hold a.s. for X_n and X then $A(X_n) \to A(X)$ in $(\mathcal{F}, \tau_{Fell})$, $d(a(X_n), A(X)) \to 0$ and $a(X_n) \to a(X)$ a.s., respectively. In the second part of section 3 we focus on what conclusions can be drawn under the assumption that X_n converges in distribution in (D, s) to some limit process X. Firstly, it follows that $A(X_n) \xrightarrow{\mathcal{L}} A(X)$ in $(\mathcal{F}, \tau_{miss})$. Furthermore assume that $(A(X_n))$ is *stochastically bounded*. Roughly speaking this means that with arbitrarily high probability eventually all $A(X_n)$ lie in some compact set $K \subseteq \mathbb{R}^d$. Then we obtain *quasi-convergence in distribution* and give a precise characterization. Naturally the question arises whether there is a hyperspace topology which is related to this new type of convergence and which of course simultaneously is strictly finer than the missing topology. It turns out that the *upper Vietoris** *topology* denoted by τ^*_{uV} meets this requirement, and thus we have that $A(X_n) \xrightarrow{\mathcal{L}} A(X)$ in $(\mathcal{F}, \tau^*_{uV})$. Moreover, if $A(X)$ is a singleton a.s. then actually $A(X_n) \xrightarrow{\mathcal{L}} A(X)$ in $(\mathcal{F}, \tau_{Fell})$ which is the strongest conclusion. In statistics one is interested in *measurable selections* ξ_n, which play the role of estimators for some parameter of interest there. These are random variables with $\xi_n \in A(X_n)$ a.s. Here $X_n \xrightarrow{\mathcal{L}} X$ in (D, s) entails

$$\limsup_{n\to\infty} \mathbb{P}(\xi_n \in K) \leq T(K) := \mathbb{P}(A(X) \cap K \neq \emptyset) \quad \text{for all compact } \ K \subseteq \mathbb{R}^d.$$

Under the additional assumption that (ξ_n) is stochastically bounded (in the classical sense) the above relation can be sharpened to

$$\limsup_{n\to\infty} \mathbb{P}(\xi_n \in F) \leq T(F) \quad \text{for all closed } \ F \subseteq \mathbb{R}^d. \tag{3}$$

Notice that this relation formally is in complete accordance with the equivalent characterization of convergence in distribution stated in the Portmanteau-Theorem. In fact the set-function T can be extended onto the Borel-σ algebra $\mathcal{B}(\mathbb{R}^d)$ on $\mathbb{R}^d$ such that $T(B) = \mathbb{P}(A(X) \cap B \neq \emptyset)$ for all Borel-sets $B \in \mathcal{B}(\mathbb{R}^d)$. Now, the key difference is that $T : \mathcal{B}(\mathbb{R}^d) \to [0, 1]$ in general is not a probability measure on $\mathcal{B}(\mathbb{R}^d)$. On the other hand T is a so-called *Choquet-capacity*, which in a suitable manner generalizes the notion of a probability. It is well-known by a Theorem of Choquet that T uniquely determines the distribution of the random closed set $A(X)$. For that reason we can say that the points

ξ_n *converge in distribution to the* <u>set</u> $A(X)$. Finally, T is in fact a probability distribution if and only if $A(X) = \{\xi\}$ is a singleton a.s. Thus (3) and the Portmanteau-Theorem yield that in case of uniqueness $\xi_n \xrightarrow{\mathcal{L}} \xi$ in $\mathbb{R}^d$.

2. Arginf-sets and infimizing points of cadlag functions

The basic idea is to involve the concept of epi-convergence, which has proven to be the suitable tool for the analysis of deterministic minimization problems. For that purpose let $\bar{\mathbb{R}} := [-\infty, \infty]$ be the extended real line. Recall that a function $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ is *lower semicontinuous at* x if $\liminf_{n\to\infty} f(x_n) \geq f(x)$ for any sequence (x_n) whose limit is x and f is *lower-semicontinuous* (lsc in short), if it is lower semicontinuous at every $x \in \mathbb{R}^d$. Set

$$SC(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \bar{\mathbb{R}};\ f \text{ lsc}\}.$$

For each $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ the *lsc regularization* $\bar{f}$ of f is defined by $\bar{f}(x) := \sup\{h(x) : h \leq f, h \text{ lsc}\}$. According to Proposition 1.8 in [6] $\bar{f}$ is lsc with $\bar{f} \leq f$ and it is the greatest of all lsc functions g such that $g \leq f$. By 1(7) and Lemma 1.7 of [21] it admits the alternative representation

$$\bar{f}(x) = \min\{\alpha \in \bar{\mathbb{R}} : \exists x_n \to x \text{ with } f(x_n) \to \alpha\}. \tag{4}$$

The following lemma yields the lsc regularization for $f \in D$.

Lemma 2.1. *If f is cadlag, then*

$$\bar{f}(x) = \min\{f(x+R) : R \in \{<, \geq\}^d\}. \tag{5}$$

In particular,

$$\bar{f}(x) = f(x) \quad \text{for all } x \in C_f, \tag{6}$$

where $C_f := \{x \in \mathbb{R}^d : f$ is continuous at $x\}$ is the set of all continuity points *of f.*

Proof. We denote the right-hand side of (5) by $\tilde{f}(x)$. By (4) there exists a sequence $x_n \to x$ with $f(x_n) \to \bar{f}(x)$. Since $\{Q_R(x) : R \in \{<, \geq\}^d\}$ is a partition of $\mathbb{R}^d$ there exists some $R \in \{<, \geq\}^d$ and a subsequence $(x_{n_k})_k$ of $(x_n)_n$ with $(x_{n_k})_k \subseteq Q_R(x)$. Thus, $f(x_{n_k}) \to \bar{f}(x)$ as $k \to \infty$, but by definition of the R- quadrant limit $f(x_{n_k}) \to f(x+R)$ as $k \to \infty$ as well, whence $\bar{f}(x) = f(x+R) \geq \tilde{f}(x)$. To see the reverse inequality note that by definition of the R-quadrant limit it follows that $\{f(x+R) : R \in \{<, \geq\}^d\} \subseteq \{\alpha \in \bar{\mathbb{R}} : \exists x_n \to x \text{ with } f(x_n) \to \alpha\}$, whence $f(x+R) \geq \bar{f}(x)$ for all $R \in \{<, \geq\}^d$ and consequently $\tilde{f}(x) \geq \bar{f}(x)$.

The second part (6) of the lemma follows from (5) upon noticing that if x is a continuity point of f then $f(x+R) = f(x)$ for all $R \in \{<, \geq\}^d$. □

For $f : \mathbb{R}^d \to \bar{\mathbb{R}}$ let

$$\text{Argmin}(f) := \{x \in \mathbb{R}^d : f(x) = \inf_{y \in \mathbb{R}^d} f(y)\}$$

be the *set of all minimizing points* of f. If f is lsc then Argmin(f) is closed. Indeed, this well-known fact is easy to see: Let $(x_n) \subseteq \text{Argmin}(f)$ with $x_n \to x$. Then $\inf_{y\in\mathbb{R}^d} f(y) = \liminf_{n\to\infty} f(x_n) \geq f(x)$, whence $x \in \text{Argmin}(f)$.

Our next result says that x is an infimizing point of a cadlag function f if and only if x is a minimizing point of its pertaining lsc regularization $\bar{f}$. Moreover, we give some useful properties which hold under rescaling.

Lemma 2.2. *(1) If f is cadlag, then*

$$A(f) = Argmin(\bar{f}). \tag{7}$$

In particular, $A(f)$ is closed.

(2) Let Λ be the group of all transformations $\lambda : \mathbb{R}^d \to \mathbb{R}^d$ of the form $\lambda(t_1,\dots,t_d) = (\lambda_1(t_1),\dots,\lambda_d(t_d))$, where each $\lambda_i : \mathbb{R} \to \mathbb{R}, 1 \le i \le d$, is continuous, strictly increasing with $\lambda_i(-\infty) = -\infty$ and $\lambda_i(\infty) = \infty$. Note that $\lambda^{-1} = (\lambda_1^{-1},\dots,\lambda_d^{-1})$ is the inverse of λ. Consider the following transform of $f \in D$:

$$g(t) = a\{f(\lambda^{-1}(t)) + b\},\ t \in \mathbb{R}^d,$$

where $a > 0$ and $b \in \mathbb{R}$. Then:
(i) $g \in D$.
(ii) $\bar g(t) = a\{\bar f(\lambda^{-1}(t)) + b\}$.
(iii) $x \in A(f) \Leftrightarrow \lambda(x) \in A(g)$.
(iv) $A(g) = \lambda(A(f))$.

Proof. (1) If $x \in A(f)$, then there exists some $R \in \{<,\ge\}^d$ such that $f(x+R) = \inf_{y\in\mathbb{R}^d} f(y) \le f(x+R')$ for all $R' \in \{<,\ge\}^d$. Thus by Lemma 2.1, $\bar f(x) = f(x+R) = \inf_{y\in\mathbb{R}^d} f(y) \le \inf_{y\in\mathbb{R}^d} \bar f(y) \le \bar f(x)$, where the first inequality follows from the definition of $\bar f$, because the constant mapping $h(y) = \inf_{y\in\mathbb{R}^d} f(y)$ is lsc. It follows that $\bar f(x) = \inf_{y\in\mathbb{R}^d} \bar f(y)$, i.e. $x \in \operatorname{Argmin}(\bar f)$.

To see the reverse inclusion let $x \in \operatorname{Argmin}(\bar f)$. By Lemma 2.1 there exists some $R \in \{<,\ge\}^d$ such that $f(x+R) = \bar f(x) = \inf_{y\in\mathbb{R}^d} \bar f(y) \le \inf_{y\in\mathbb{R}^d} f(y) \le f(x+R)$, where the first inequality holds since $\bar f \le f$. Thus $f(x+R) = \inf_{y\in\mathbb{R}^d} f(y)$ resulting in $x \in \mathrm{A}(f)$.

(2) Part (i) follows, because $t_n \to t$ in $Q_R(t)$ implies $\lambda^{-1}(t_n) \to \lambda^{-1}(t)$ in $Q_R(\lambda^{-1}(t))$ for every $R \in \{<,\ge\}^d$. Here, we need the special form of λ and its inverse λ^{-1}. Consequently, $g(t+R) = a\{f(\lambda^{-1}(t)+R)+b\}$, which in turn yields (ii) by (5) of Lemma 2.1. As to the validity of (iii) note that λ is a bijection. Then with the help of (ii) one easily verifies that $\bar f(x) \le \bar f(t)\ \forall\ t \in \mathbb{R}^d$ if and only if $\bar g(\lambda(x)) \le \bar g(u)\ \forall\ u \in \mathbb{R}^d$. This confirms (iii) by part (1) and (i). Finally, (iv) follows from (iii). ☐

Let D_f be the *set of discontinuity points* of $f \in D$. In the univariate case D_f is a most countable, confer [4], p.122, and therefore C_f is dense in $\mathbb{R}$. However, for multivariate f the set D_f in general is uncountable. In fact, consider the simple indicator function $f = 1_{I_1\times\ldots\times I_d}$, where $I_j, 1 \le j \le d$ is a left-closed, right-open interval in $\mathbb{R}$. Here, D_f is equal to the boundary of $I_1 \times \ldots \times I_d$, which for $d \ge 2$ is non-denumerable. Nevertheless, it turns out that C_f is dense in $\mathbb{R}^d$ for higher dimensions d as well. Actually, we have much more.

Lemma 2.3. *C_f lies dense in $\mathbb{R}^d$ for each $f \in D$. In fact, for every sequence $(f_j)_{j\in\mathbb{N}} \subseteq D$ the intersection $\bigcap_{j\in\mathbb{N}} C_{f_j}$ is dense in $\mathbb{R}^d$.*

Proof. We first prove the assertion for a single function $f \in D$. Let us assume that C_f is not dense in $\mathbb{R}^d$. Then there exist $x \in \mathbb{R}^d$ and $\eta > 0$ such that the open ball $B_\eta(x) := \{y \in \mathbb{R}^d : |y-x| < \eta\} \subseteq D_f$. Here, $|\cdot|$ is the maximum-norm on $\mathbb{R}^d$. We choose some $a > 0$ large enough such that $x \in [-a,a]^d$ and $B_\eta(x) \subseteq [-a,a]^d$. According to Lemma 1.5 of [16] for every $\epsilon > 0$ there exist a $\delta = \delta(\epsilon) > 0$ and a (finite) partition $\mathcal{R} = \mathcal{R}(\epsilon)$ of $[-a,a]^d$ into rectangles such that for points $x, x' \in I, I \in \mathcal{R}$, with $|x-x'| < \delta$ the inequality $|f(x) - f(x')| < \epsilon$ holds. If

$$H_f(y) := \max\{|f(y+R) - f(y+R')| : R, R' \in \{<,\ge\}^d\}$$

denotes the *magnitude of jump* at point y, then clearly f is continuous at y if and only if $H_f(y) = 0$. Consequently,

$$B_\eta(x) \subseteq D_f \cap [-a,a]^d \subseteq \bigcup_{k\in\mathbb{N}} \{y \in [-a,a]^d : H_f(y) > 1/k\}. \tag{8}$$

For fixed $k \in \mathbb{N}$ we have that

$$\{y \in [-a,a]^d : H_f(y) > 1/k\} \subseteq \bigcup_{I\in\mathcal{R}(1/k)} \partial I =: \Gamma_k \ , \tag{9}$$

where ∂I denotes the boundary of (any set) I. To see the inclusion, let $y \in [-a,a]^d$ with $H_f(y) > 1/k$. Since $\mathcal{R}(1/k)$ is a partition of $[-a,a]^d$ there exists a rectangle $I \in \mathcal{R}(1/k)$ containing y. Then y cannot be located in the interior of I. In fact, in that case for all $R, R' \in \{<, \geq\}^d$ it follows that

$$|f(y+R) - f(y+R')| = \lim_{n\to\infty} |f(s_n) - f(u_n)|, \tag{10}$$

with sequences $(s_n) \subseteq Q_R(y)$ and $(u_n) \subseteq Q_{R'}(y)$ both converging to y. Since the joint limit y is an interior point of I by assumption we know that s_n and u_n are in I for eventually all $n \in \mathbb{N}$ and in addition $|s_n - u_n| < \delta(1/k)$. Thus the above Lemma 1.5 of [16] yields that $|f(s_n) - f(u_n)| < 1/k$ for eventually all $n \in \mathbb{N}$. Therefore by (10) we obtain that $|f(y+R) - f(y+R')| \leq 1/k$ for all $R, R' \in \{<, \geq\}^d$ and consequently $H_f(y)$ as maximal value is less than or equal to $1/k$ as well, in contradiction to $H_f(y) > 1/k$. For that reason y must lie on the boundary of I, which shows the validity of the inclusion (9).

In view of (8) we can infer that $B_\eta(x) \subseteq \bigcup_{k\in\mathbb{N}} \Gamma_k$, where Γ_k is a finite union of hyperplanes with dimension $d-1$. As it is well-known such hyperplanes are null-sets under the d-dimensional Lebesgue-measure λ_d. It follows that each Γ_k is a null-set as well and so $\lambda_d(B_\eta(x)) \leq \sum_{k\in\mathbb{N}} \lambda_d(\Gamma_k) = 0$ in contradiction to $\lambda_d(B_\eta(x)) = (2\eta)^d > 0$. Thus our lemma is proved in the special case of a single function f.

In the general case let us again assume that $\bigcap_{j\in\mathbb{N}} C_{f_j}$ is not dense in $\mathbb{R}^d$. Then there exist $x \in [-a,a]^d$ for some $a > 0$ large enough and $\eta > 0$ such that the open ball $B_\eta(x) \subseteq \bigcup_{j\in\mathbb{N}} (D_{f_j} \cap [-a,a]^d)$. If for each fixed $j \in \mathbb{N}$, we apply (8) and (9) to $f = f_j$ we obtain that $B_\eta(x) \subseteq \bigcup_{j\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} \Gamma_{jk}$, where each Γ_{jk} is a finite union of hyperplanes with dimension $d-1$. Therefore, by σ-additivity of the Lebesgue-measure we obtain the contradiction that $\lambda_d(B_\eta(x)) \leq 0$. □

Our next result relates *Skorokhod-convergence*, i.e. convergence in $(D(\mathbb{R}^d), s)$ with the so-called *epi-convergence*, which is defined as follows. A sequence (f_n) of functions from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ *epi-converges* to some $f : \mathbb{R}^d \to \bar{\mathbb{R}}$, if at each $x \in \mathbb{R}^d$ the following holds:

(a) For all sequences (x_n) converging to x it is $\liminf_{n\to\infty} f_n(x_n) \geq f(x)$.
(b) There exists at least one sequence (x_n) converging to x such that $\limsup_{n\to\infty} f_n(x_n) \leq f(x)$.

We write $f_n \overset{epi}{\to} f$ for short. According to Theorem 2.78 and Corollary 2.79 of [1] there exists a metric e on the space $SC(\mathbb{R}^d)$ such that

$$f_n \overset{epi}{\to} f \quad \Leftrightarrow \quad e(f_n, f) \to 0, \ n \to \infty \quad (f_n \overset{e}{\to} f)$$

Moreover $(SC(\mathbb{R}^d), e)$ is compact and separable.

So far we did not give the definition of the Skorokhod-metric s. Indeed, for our purposes it suffices to work with the following equivalent characterization of the Skorokhod-convergence as given in Theorem 1 of [12]. Here, recall the definition of the transformation group Λ introduced in Lemma 2.2 (2).

(*Characterization of Skorokhod-convergence*) If f and $f_n, n \in \mathbb{N}$, are functions in D then $s(f_n, f) \to 0$ $(f_n \xrightarrow{s} f)$ if and only if there exists a sequence $(\underline{\lambda}_n) \subseteq \Lambda$ such that
(a) $\sup_{t \in \mathbb{R}^d} |\underline{\lambda}_n(t) - t| \to 0,\ n \to \infty$.
(b) $\sup_{t \in [-a,a]^d} |f_n(\underline{\lambda}_n(t)) - f(t)| \to 0,\ n \to \infty \quad \forall\, a > 0$.

The following result is the key for many proofs in our paper.

Proposition 2.1. *The mapping* $\Phi : (D(\mathbb{R}^d), s) \to (SC(\mathbb{R}^d), e)$ *given by* $\Phi(f) = \bar{f}$ *is continuous, i.e., for each* $f \in D(\mathbb{R}^d)$ *and every sequence* $f_n \xrightarrow{s} f$ *it follows that* $\bar{f}_n \xrightarrow{epi} \bar{f}$. *Moreover,* Φ *is an injection.*

Proof. Let $f \in D$ and (f_n) be a sequence with $f_n \xrightarrow{s} f$. Assume that $\bar{f}_n \overset{epi}{\not\longrightarrow} \bar{f}$. This means that there exists some $x \in \mathbb{R}^d$ such that

$$\liminf_{n\to\infty} \bar{f}_n(x_n) < \bar{f}(x) \quad \text{for some sequence } (x_n) \text{ with } x_n \to x, \tag{11}$$

or

$$\limsup_{n\to\infty} \bar{f}_n(x_n) > \bar{f}(x) \quad \text{for all sequences } (x_n) \text{ with } x_n \to x. \tag{12}$$

In the first case we can deduce from (11) that there exist an $\alpha > 0$ and a subsequence (x_{n_k}) of (x_n) such that

$$\bar{f}_{n_k}(x_{n_k}) < \bar{f}(x) - 3\alpha \quad \forall\, k \in \mathbb{N}. \tag{13}$$

Clearly, $x \in [-a,a]^d$ for some $a > 0$. Since $f_n \xrightarrow{s} f$ we find a sequence $(\underline{\lambda}_n) \in \Lambda$ with

$$\sup_{t\in\mathbb{R}^d} |\underline{\lambda}_n(t) - t| \to 0,\ n \to \infty, \tag{14}$$

and

$$\sup_{t\in[-a,a]^d} |f_n(\underline{\lambda}_n(t)) - f(t)| \to 0,\ n \to \infty. \tag{15}$$

According to Lemma 2.1 for each fixed $n \in \mathbb{N}$ there is a $R_n \in \{<, \geq\}^d$ such that

$$\bar{f}_n(x_n) = f_n(x_n + R_n) = \lim_{y\to x_n, y\in Q_{R_n}(x_n)} f_n(y).$$

Thus for our $\alpha > 0$ there exists a positive $\delta_0 = \delta_0(\alpha, x_n) =: \delta_{0,n}$ such that for all $\delta \in (0, \delta_{0,n}]$ we have that

$$|\bar{f}_n(x_n) - f_n(y)| \leq \alpha \quad \forall\, y \in Q_{R_n}(x_n) \ \text{ with } |y - x_n| \leq \delta. \tag{16}$$

Since C_{f_n} lies dense in $\mathbb{R}^d$ by Lemma 2.3 we find some $y_n \in Q_{R_n}(x_n) \cap C_{f_n}$ such that

$$|y_n - x_n| \leq \min\{n^{-1}, \delta_{0,n}\} \leq \delta_{0,n}, \tag{17}$$

whence by (6) and (16)

$$|\bar{f}_n(x_n) - \bar{f}_n(y_n)| = |\bar{f}_n(x_n) - f_n(y_n)| \leq \alpha.$$

In particular, we have constructed a sequence (y_n) with $y_n \in C_{f_n},\ n \in \mathbb{N}$, satisfying

$$|y_n - x_n| \to 0,\ n \to \infty, \tag{18}$$

and

$$\bar{f}_n(y_n) \leq \bar{f}_n(x_n) + \alpha \quad \forall\, n \in \mathbb{N}. \tag{19}$$

Along the subsequence (n_k) occurring in (13) we define

$$z_{n_k} := \underline{\lambda}_{n_k}^{-1}(y_{n_k}), \quad k \in \mathbb{N}. \tag{20}$$

Since for each fixed $k \in \mathbb{N}$,

$$\begin{aligned} f_{n_k}(\underline{\lambda}_{n_k}(z_{n_k})) &= f_{n_k}(y_{n_k}) && \text{by (20)} \\ &= \bar{f}_{n_k}(y_{n_k}) && \text{by (6)} \\ &\leq \bar{f}_{n_k}(x_{n_k}) + \alpha && \text{by (19)} \\ &< \bar{f}(x) - 3\alpha + \alpha && \text{by (13)} \\ &= \min\{f(x+R) : R \in \{<, \geq\}^d\} - 2\alpha && \text{by (5),} \end{aligned}$$

we arrive at

$$f_{n_k}(\underline{\lambda}_{n_k}(z_{n_k})) \leq f(x+R) - 2\alpha \quad \forall\, R \in \{<, \geq\}^d\} \quad \forall\, k \in \mathbb{N}. \tag{21}$$

Moreover, observe that $|x - z_{n_k}| \leq |x - x_{n_k}| + |x_{n_k} - y_{n_k}| + |y_{n_k} - z_{n_k}|$, where the first two summands converge to zero as $k \to \infty$ in view of (11) and (18). From the definition (20) it follows that the third summand is equal to $|\underline{\lambda}_{n_k}(z_{n_k}) - z_{n_k}|$, which converges to zero by (14). This yields that $z_{n_k} \to x, k \to \infty$. In particular there exists a quadrant $Q_R(x)$ for some $R \in \{<, \geq\}^d$, which contains a subsequence of (z_{n_k}). Therefore w.l.o.g. we have that $z_{n_k} \to x, k \to \infty$ with $z_{n_k} \in Q_R(x)\ \forall\ k \in \mathbb{N}$. This ensures that $f(z_{n_k}) \to f(x+R), k \to \infty$, which in turn guarantees the existence of some natural number $k_0 = k_0(\alpha)$ such that

$$f(x+R) \leq f(z_{n_k}) + \alpha \quad \forall\, k \geq k_0,$$

and so with (21) we can conclude that

$$f_{n_k}(\underline{\lambda}_{n_k}(z_{n_k})) \leq f(x+R) - 2\alpha \leq f(z_{n_k}) - \alpha \quad \forall\, k \geq k_0. \tag{22}$$

As a convergent sequence $(z_{n_k})_{k \geq k_0}$ is bounded and by enlarging $a > 0$ if necessary we may assume that $(z_{n_k})_{k \geq k_0} \subseteq [-a, a]^d$. Thus with (22) one obtains that

$$\begin{aligned} \sup_{t \in [-a,a]^d} |f_{n_k}(\underline{\lambda}_{n_k}(t)) - f(t)| &\geq |f_{n_k}(\underline{\lambda}_{n_k}(z_{n_k})) - f(z_{n_k})| \\ &\geq f(z_{n_k}) - f_{n_k}(\underline{\lambda}_{n_k}(z_{n_k})) \geq \alpha \quad \forall\, k \geq k_0, \end{aligned}$$

which is in contradiction to (15).

In the second case we can deduce from (12) that for every sequence (x_n) with $x_n \to x$ there exist a subsequence (x_{n_k}) of (x_n) and some $\alpha > 0$ such that

$$\bar{f}_{n_k}(x_{n_k}) > \bar{f}(x) + 3\alpha \quad \forall\, k \in \mathbb{N}. \tag{23}$$

By (5) of Lemma 2.1 there is some $R \in \{<, \geq\}^d$ with $\bar{f}(x) = f(x+R)$. Let us introduce

$$\epsilon_n := \sup_{t \in \mathbb{R}^d} |\underline{\lambda}_n(t) - t| = \sup_{t \in \mathbb{R}^d} |\underline{\lambda}_n^{-1}(t) - t|.$$

For our given $R = (R_1, \ldots, R_d)$ and $x = (x_1, \ldots, x_d)$ we define $x_n = (x_{n1}, \ldots, x_{nd})$ for each $n \in \mathbb{N}$ by

$$x_{nj} := \begin{cases} x_j + \epsilon_n + 2/n & , \quad R_j = \ \geq \\ x_j - \epsilon_n - 2/n & , \quad R_j = \ < \end{cases}$$

Using the same arguments as in the derivation of (17) and (19) we find $y_n \in C_{f_n}$ and positive $\delta_{0,n}$ such that

$$|y_n - x_n| \leq \min\{1/n, \delta_{0,n}\} \leq 1/n \quad \forall\, n \in \mathbb{N}, \tag{24}$$

and

$$\bar{f}_n(y_n) \geq \bar{f}_n(x_n) - \alpha \quad \forall\, n \in \mathbb{N}. \tag{25}$$

Given $y_n = (y_{n1}, \ldots, y_{nd})$ and $\underline{\lambda}_n = (\lambda_{n1}, \ldots, \lambda_{nd})$ we put

$$z_n = (z_{n1}, \ldots, z_{nd}) := \underline{\lambda}_n^{-1}(y_n) = (\lambda_{n1}^{-1}(y_{n1}), \ldots, \lambda_{nd}^{-1}(y_{nd})). \tag{26}$$

Observe that $|z_n - y_n| = |\underline{\lambda}_n^{-1}(y_n) - y_n| \leq \epsilon_n$ and recall that $|\cdot|$ is the max-norm on $\mathbb{R}^d$. Therefore we know that

$$-\epsilon_n \leq z_{nj} - y_{nj} \leq \epsilon_n \quad \forall\, 1 \leq j \leq d, \tag{27}$$

and by (24) that

$$-1/n \leq y_{nj} - x_{nj} \leq 1/n \quad \forall\, 1 \leq j \leq d. \tag{28}$$

Conclude from (27) and (28) that for all $1 \leq j \leq d$,

$$x_j + 1/n \leq z_{nj} \leq x_j + 3/n + 2\epsilon_n, \quad \text{ if } R_j = \ \geq$$

and

$$x_j - 3/n - 2\epsilon_n \leq z_{nj} \leq x_j - 1/n, \quad \text{ if } R_j = \ < ,$$

whence $z_{nj}\ R_j\ x_j \quad \forall\, 1 \leq j \leq d$ and so $z_n \in Q_R(x)$. Moreover, since $\epsilon_n \to 0$ by (14), it follows that $z_n \to x$. In particular, $z_{n_k} \to x, k \to \infty$ with $z_{n_k} \subseteq Q_R(x)\ \forall\, k \in \mathbb{N}$ and so $f(z_{n_k}) \to f(x+R) = \bar{f}(x), k \to \infty$. Consequently, there exists some natural number $k_0 = k_0(\alpha)$ such that

$$\bar{f}(x) \geq f(z_{n_k}) - \alpha \quad \forall\, k \geq k_0. \tag{29}$$

As above we find some $a > 0$ such that $x \in [-a,a]^d$ and $(z_{n_k})_{k \geq k_0} \subseteq [-a,a]^d$. Combining our results we obtain:

$$\begin{array}{rcll} \sup_{t\in[-a,a]^d} |f_{n_k}(\underline{\lambda}_{n_k}(t)) - f(t)| & \geq & f_{n_k}(\underline{\lambda}_{n_k}(z_{n_k}) - f(z_{n_k}) & \\ & = & f_{n_k}(y_{n_k}) - f(z_{n_k}) & \text{by (26)} \\ & = & \bar{f}_{n_k}(y_{n_k}) - f(z_{n_k}) & \text{by (6)} \\ & \geq & \bar{f}_{n_k}(x_{n_k}) - \alpha - f(z_{n_k}) & \text{by (25)} \\ & > & \bar{f}(x) + 2\alpha - f(z_{n_k}) & \text{by (23)} \\ & \geq & \alpha \quad \forall\, k \geq k_0 & \text{by (29),} \end{array}$$

which is in contradiction to (15). This finishes our proof of continuity.

To see that Φ is injective, let $\bar{f} = \bar{g}$. It follows from (6) of Lemma 2.1 that $f(x) = g(x)$ for all $x \in C_f \cap C_g =: \Delta$, where by Lemma 2.3 the set Δ is dense in $\mathbb{R}^d$. Thus for an arbitrary point $x \in \mathbb{R}^d$ and $R = (\geq, \ldots, \geq)$ there exists a sequence $(x_n) \subseteq Q_R(x) \cap \Delta$ with limit x, whence by "continuity from above" $f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} g(x_n) = g(x)$. ☐

In the univariate case ($d = 1$) continuity of Φ has been shown by [28] in the proof of her Lemma 8.6 (ii). Here, Vogel uses an equivalent variant of the Skorokhod-metric, which goes back to [20] and which is specifically designed for dimension one.

By Lemma 2.2 the assignment $f \mapsto \mathrm{A}(f), f \in D$ is a map A: $D \to \mathcal{F}$ from the Skorokhod-space into the family of all closed subsets in $\mathbb{R}^d$. We seek for strong (fine) topologies on $\mathcal{F}$ that guarantee continuity of A at as many functions f as possible.

Let $\mathcal{K} = \mathcal{K}(\mathbb{R}^d)$ denote the family of all compact subsets in $\mathbb{R}^d$. The first topology on $\mathcal{F}$ we consider is the so-called *missing-topology* τ_{miss} also known as *upper Fell-topology.* By definition it is generated from a subbase which consists of all *missing-sets* $\mathcal{M}(K) := \{F \in \mathcal{F} : F \cap K = \emptyset\}, K \in \mathcal{K}$. [29] gives a characterization for convergence of a sequence (F_n) in the topology τ_{miss} via the *Kuratowski-outer limit*

$$K - \limsup_{n\to\infty} F_n := \{x \in \mathbb{R}^d : \exists (n_j) \subset \mathbb{N}\ \exists\, x_{n_j} \in F_{n_j} \text{ with } x_{n_j} \to x, j \to \infty\}.$$

Lemma 2.4. ***(Vogel)*** *Let F and $F_n, n \in \mathbb{N}$, be closed subsets in $\mathbb{R}^d$. Then*

$$F_n \to F \quad \text{in } \tau_{miss} \quad \Leftrightarrow \quad K - \limsup_{n\to\infty} F_n \subseteq F. \tag{30}$$

If $\mathcal{F}$ is equipped with the missing-topology then the functional A proves to be continuous on its entire domain. As a consequence of this continuity we also obtain a certain stability for infimizing-points x_n of f_n with s-limit f. Stability is specified in terms of

$$d(x, M) := \inf\{|x - y| : y \in M\}, \quad x \in \mathbb{R}^d,\ M \subseteq \mathbb{R}^d,$$

which defines the *distance of the point x to the set M*. In the following propositions we make these statements precise.

Proposition 2.2. *Assume that $f_n \overset{s}{\to} f$ in D. Then:*
(a) $A(f_n) \to A(f)$ in τ_{miss}.
(b) $K - \limsup_{n\to\infty} A(f_n) \subseteq A(f)$.

Proof. By Lemma 2.4 the relations (a) and (b) are equivalent. From Proposition 2.1 we know that $\bar{f}_n \overset{epi}{\to} \bar{f}$. By Lemma 2.2 it suffices to show that

$$K - \limsup_{n\to\infty} \operatorname{Argmin}(\bar{f}_n) \subseteq \operatorname{Argmin}(\bar{f}). \tag{31}$$

The inclusion (31) follows immediately from Proposition 7.18 of [6] upon noticing that the concept of Γ-convergence as considered there coincides with epi-convergence. However, we prefer to give an elementary proof.

First notice that it is trivially fulfilled, if the left-hand side is empty. So, let $x \in K - \limsup_{n\to\infty} \operatorname{Argmin}(\bar{f}_n)$. By definition there exist a subsequence (n_j) of the natural numbers and $x_{n_j} \in \operatorname{Argmin}(\bar{f}_{n_j}), j \in \mathbb{N}$, such that $x_{n_j} \to x$ as $j \to \infty$. Assume that $x \notin \operatorname{Argmin}(\bar{f})$. Then there exists some $y \in \mathbb{R}^d$ with $\bar{f}(y) < \bar{f}(x)$. By (a) and (b) in the definition of epi-convergence there exists a sequence (y_n) with $y_n \to y$ and $\lim_{n\to\infty} \bar{f}_n(y_n) = \bar{f}(y)$. Thus we can conclude as follows:

$$\bar{f}(x) > \bar{f}(y) = \lim_{n\to\infty} \bar{f}_n(y_n) = \liminf_{j\to\infty} \bar{f}_{n_j}(y_{n_j}) \geq \liminf_{j\to\infty} \bar{f}_{n_j}(x_{n_j}) \geq \bar{f}(x).$$

Here, the second to last inequality holds, because x_{n_j} is a minimizing point of $\bar{f}_{n_j}$ for all $j \in \mathbb{N}$. Observe that in particular $\bar{f}_{n_j} \overset{epi}{\to} \bar{f}$ as $j \to \infty$ and so the last inequality follows from the first part (a) in the definition of epi-convergence. Summing up we arrive at a contradiction, whence x has to lie in $\operatorname{Argmin}(\bar{f})$ as desired. □

If one picks out infimizing points of f_n in a compact set K, then the distances of these points to the set containing all infimizers of f in K converge to zero. More precisely, we have

Proposition 2.3. *Assume that $f_n \overset{s}{\to} f$ in D with $A(f) \neq \emptyset$. If there exist some compact $K \subseteq \mathbb{R}^d$ and $N \in \mathbb{N}$ with $A(f_n) \cap K \neq \emptyset$ for all $n \geq N$, then for each selection $x_n \in A(f_n) \cap K, n \geq N$, it follows that*

$$d(x_n, A(f)) \to 0,\ n \to \infty.$$

If actually $A(f) \cap K \neq \emptyset$ then $d(x_n, A(f) \cap K) \to 0$.

Proof. To prove the first convergence, let us assume that $d(x_n, A(f)) \not\to 0,\ n \to \infty$. Then there exists some $\epsilon > 0$ and a subsequence (x_{n_j}) of (x_n) such that

$$d(x_{n_j}, A(f)) \geq \epsilon \quad \forall\, j \in \mathbb{N}. \tag{32}$$

Since (x_{n_j}) is in the compact set K by assumption it contains a convergent subsequence with limit in K. For the sake of notational simplicity we may assume that $x_{n_j} \to x \in K, j \to \infty$. Furthermore, by definition of the K-outer limit it follows that $x \in K - \limsup_{n\to\infty} A(f_n) \subseteq A(f)$ by part (b) of Proposition 2.2, whence $x \in A(f) \cap K$. On the other hand, the map $x \mapsto d(x, M)$ is continuous for each fixed set $M \neq \emptyset$. Thus by taking the limit $j \to \infty$ in (32) one obtains that $d(x, A(f)) \geq \epsilon > 0$. Since $d(x, M) = 0$

if and only if x lies in the closure of M and $A(f)$ is closed by Lemma 2.2 we arrive at the contradiction $x \notin A(f)$.

The derivation of the second convergence follows the same lines upon noticing that $A(f) \cap K$ is closed as well. □

Intuitively it is clear that one cannot expect a convergence of (x_n) to some x as long as the limit f has more than one infimizing point x. On the other hand if in fact x is unique, i.e. $A(f) = \{x\}$, then the convergence holds indeed. Actually, we have a bit more than that. It is merely needed that x is a unique infimizing point of f *relative to* K, i.e. $A(f) \cap K = \{x\}$, so that the existence of further infimizers outside of K is not forbidden.

Proposition 2.4. *(convergence of relative infimizers) Assume that $f_n \stackrel{s}{\to} f$ in D with $A(f) = \{x\}$. If there exist some compact $K \subseteq \mathbb{R}^d$ and some $N \in \mathbb{N}$ with $A(f_n) \cap K \neq \emptyset$ for all $n \geq N$, then for each selection $x_n \in A(f_n) \cap K, n \geq N$, it follows that*

$$x_n \to x, \; n \to \infty.$$

Under the weaker assumption $A(f) \cap K = \{x\}$ the convergence still holds.

Proof. The assertion follows immediately from Proposition 2.3 upon noticing that $d(x_n, A(f)) = |x_n - x|$ and $d(x_n, A(f) \cap K) = |x_n - x|$ if $A(f) = \{x\}$ and $A(f) \cap K = \{x\}$, respectively. □

In view of applications we like to give a reformulation.

Corollary 2.1. *Let $f_n \stackrel{s}{\to} f$ in D with f possessing a unique infimizing point x. If there exists $N \in \mathbb{N}$ with $A(f_n) \neq \emptyset$ for all $n \geq N$, then for each selection $x_n \in A(f_n), n \geq N$, such that $(x_n)_{n \geq N}$ is* bounded*, we obtain that*

$$x_n \to x.$$

If x is only unique relative to $K_\epsilon := \{t \in \mathbb{R}^d : |t| \leq \limsup_{N \leq n \to \infty} |x_n| + \epsilon\}$ for an arbitrary small $\epsilon > 0$ then $x_n \to x$ still holds.

Proof. By assumption $s := s_N := \sup_{n \geq N} |x_n|$ is finite, whence $K := [-s, s]^d$ is compact with $A(f_n) \cap K \neq \emptyset$ for all $n \geq N$. Thus the first assertion follows from Proposition 2.4. As to the second assertion note that $a := \limsup_{N \leq n \to \infty} |x_n| = \inf_{M \geq N} s_M$. Consequently there exists a natural number $M = M(\epsilon) \geq N$ such that $s_M \leq a + \epsilon$, which in turn ensures that $A(f_n) \cap K_\epsilon \neq \emptyset$ for all $n \geq M$. Since $A(f) \cap K_\epsilon = \{x\}$ by assumption we can apply Proposition 2.4 with M and K_ϵ in place of N and K, respectively. □

Since every convergent sequence is bounded, one might expect that the boundedness assumption could be dropped. The following example shows that this is not true.

Example 2.1. For every $n \in \mathbb{N}$ let $f_n : \mathbb{R}^d \to \mathbb{R}$ be defined by

$$f_n(x) := \begin{cases} |x| & , \quad |x| \leq n \\ 2n - |x| & , \quad n < |x| \leq 2n \\ |x| - 2n & , \quad |x| > 2n \end{cases} ,$$

where $|\cdot|$ is any norm on $\mathbb{R}^d$. Then the f_n are actually continuous and converge uniformly on compacta to f with $f(x) = |x|$, whence $f_n \stackrel{s}{\to} f$ in D. Moreover, $\mathrm{A}(f) = \{0\}$ and $\mathrm{A}(f_n) = \{0\} \cup \{x \in \mathbb{R}^d : |x| = 2n\}$, so that each choice $0 \neq x_n \in A(f_n)$ yields $x_n \to \infty$ and convergence to $x = 0$ fails.

The example shows that even uniform convergence on compacta is not sufficient to get rid of boundedness. The problem lies in whenever an (arbitrary) set $T \subseteq \mathbb{R}^d$ is fixed uniform convergence on that T does not include any kind of control on the f_n outside of T for eventually all $n \in \mathbb{N}$. As a way out we let $T = T_n$ vary with n such that the T_n

exhaust the entire Euclidean space $\mathbb{R}^d$ as n tends to infinity. Moreover, the limit function f must have a *well-separated* infimizing point x as specified in (34) below. This means it is excluded that there is some neighborhood U of x such that f comes arbitraray close to the infimum value at points outside of U.

Proposition 2.5. *For each $n \in \mathbb{N}$ let $T_n \subseteq \mathbb{R}^d$ be an open subset and x_n an infimizing point of the restriction of f_n on T_n, i.e., $x_n \in T_n$ and $\min_{R\in\{<,\geq\}^d} f_n(x_n + R) = \inf_{t\in T_n} f_n(t)$. Assume that*

$$\sup_{t\in T_n} |f_n(t) - f(t)| \to 0, \ n \to \infty, \tag{33}$$

and that f has an infimizing point x, which additionally is well-separated *in the following sense:*

$$\inf_{t\in\mathbb{R}^d} f(t) < \inf\{f(t) : |t - x| > \epsilon\} \quad \forall\, \epsilon > 0. \tag{34}$$

Then, if

$$\liminf_{n\to\infty} T_n = \mathbb{R}^d, \tag{35}$$

it follows that

$$x_n \to x, \ n \to \infty.$$

Proof. Let $\epsilon > 0$ and $m(\epsilon) := \inf\{f(t) : |t - x| > \epsilon\}$. Then by (34) the quantity $b(\epsilon) := ((m(\epsilon) - \inf_{t\in\mathbb{R}^d} f(t))/3$ is positive. Thus by (33) we find some $n_0 = n_0(\epsilon) \in \mathbb{N}$ such that

$$\sup_{t\in T_n} |f_n(t) - f(t)| \leq b(\epsilon) \quad \forall\, n \geq n_0. \tag{36}$$

Since x and x_n are infimizing points there exist $R(x)$ and $R(x_n)$ in $\{<,\geq\}^d$ with $f(x + R(x)) = \inf_{t\in\mathbb{R}^d} f(t)$ and $f(x_n + R(x_n)) = \inf_{t\in T_n} f_n(t)$. From (35) we can infer that $x \in T_n$ for all $n \geq n_1 \in \mathbb{N}$.

Next observe that for some sequence $(s_m) \subseteq Q_{R(x)}(x)$ with limit $x \in T_n$ it follows that $|f_n(x + R(x)) - f(x + R(x))| = \lim_{m\to\infty} |f_n(s_m) - f(s_m)| \leq \sup_{t\in T_n} |f_n(t) - f(t)|$. Thus (36) guarantees

$$f_n(x + R(x)) \leq f(x + R(x)) + b(\epsilon) \quad \forall\, n \geq n_2 := \max\{n_0, n_1\}. \tag{37}$$

Furthermore, let $y \in G_n := T_n \cap \{t \in \mathbb{R}^d : |t - x| > \epsilon\}$. Then another application of (36) and the definition of $m(\epsilon)$ yield that

$$f_n(y) \geq f(y) - b(\epsilon) \geq m(\epsilon) - b(\epsilon) \quad \forall\, n \geq n_2. \tag{38}$$

Combining (37) and (38) we arrive at

$$f_n(y) - f_n(x + R(x)) \geq m(\epsilon) - b(\epsilon) - b(\epsilon) - f(x + R(x)) = b(\epsilon) \quad \forall\, n \geq n_2. \tag{39}$$

where the last equality holds, because $m(\epsilon) - f(x + R(x)) = m(\epsilon) - \inf_{t\in\mathbb{R}^d} f(t) = 3b(\epsilon)$ by definition of $b(\epsilon)$. Since G_n is an open set, we can take the quadrant-limits in (39) and obtain $f_n(y + R) \geq f_n(x + R(x)) + b(\epsilon) > f_n(x + R(x))$ for all $R \in \{<,\geq\}^d$. To sum up the following relation holds:

$$f_n(y + R) > f_n(x + R(x)) \quad \forall\, y \in G_n \,\forall\, R \in \{<,\geq\}^d \,\forall\, n \geq n_2. \tag{40}$$

Finally, consider an $n \geq n_2$ and assume that $|x_n - x| > \epsilon$. Then $x_n \in G_n$ and from (40) we can deduce the following contradiction:

$$\inf_{t\in T_n} f_n(t) = f_n(x_n + R(x_n)) > f_n(x + R(x)) \geq \inf_{t\in T_n} f_n(t),$$

where the last inequality holds, because $f_n(x+R(x)) = \lim_{m\to\infty} f_n(t_m)$ for some sequence $(t_m) \subseteq Q_{R(x)}(x)$ with limit x. Since x is an interior point of G_n it is $t_m \in G_n \subseteq T_n$ for eventually all $m \in \mathbb{N}$.

To sum up we have shown that for every $\epsilon > 0$ it is $|x_n - x| \leq \epsilon \; \forall \, n \geq n_2 \in \mathbb{N}$, which finishes our proof. □

Remark 2.1. If x is a well-seperated infimizing point of f, then $A(f) = \{x\}$. Indeed, assume that there is another infimizing point y. Then $\epsilon := |y - x|$ is positive. Since in particular $|y - x| > \epsilon/3$ we obtain with (1) and (34) that

$$\inf_{t \in \mathbb{R}^d} f(t) = \min_{R \in \{<, \geq\}^d} f(y + R) \geq \inf\{f(t) : |t - x| > \epsilon/3\} > \inf_{t \in \mathbb{R}^d} f(t).$$

Under the assumptions of Proposition 2.4 we also obtain convergence of the arginf-sets with respect to a topology on $\mathcal{F}$ which is strictly finer (stronger) than the missing-topology. This so-called *Fell-topology* τ_{Fell} is generated from a subbase $\{\mathcal{M}(K) : K \in \mathcal{K}\} \cup \{\mathcal{H}(G) : G \in \mathcal{G}\}$, where $\mathcal{H}(G) := \{F \in \mathcal{F} : F \cap G \neq \emptyset\}$ is a *hitting-set* and $\mathcal{G} = \mathcal{G}(\mathbb{R}^d)$ denotes the family of all open subsets in $\mathbb{R}^d$. It is well-known that $(\mathcal{F}, \tau_{Fell})$ is compact, second countable and Haussdorff, confer Theorem A2.5 in [11]. In particular, the Fell-topology is induced by the *Kuratowski-metric* δ and it corresponds to the *Painlevé-Kuratowski convergence*. These concepts are defined as follows. For a given sequence (F_n) of closed sets the *Kuratowski-inner limit* is the set

$$K - \liminf_{n \to \infty} F_n := \{x \in \mathbb{R}^d : \exists \, n_0 \in \mathbb{N} \; \forall \, n \geq n_0 \; \exists \, x_n \in F_n \text{ with } x_n \to x\}.$$

Obviously,

$$K - \liminf_{n \to \infty} F_n \subseteq K - \limsup_{n \to \infty} F_n, \tag{41}$$

but if in fact $K - \liminf_{n \to \infty} F_n = K - \limsup_{n \to \infty} F_n =: F$ then F is called the *Painlevé-Kuratowski limit* of (F_n) denoted by $K - \lim_{n \to \infty} F_n$. For the introduction of the Kuratowski-metric δ let $\{x_i : i \in \mathbb{N}\}$ be any countable dense subset in $\mathbb{R}^d$. Then

$$\delta(F, H) := \sum_{i=1}^{\infty} 2^{-i} \big(\min\{d(x_i, F), 1\} - \min\{d(x_i, H), 1\} \big), \quad F, H \in \mathcal{F}.$$

It is known that, confer, e.g., [19] or [21]:

$$F_n \to F \text{ in } \tau_{Fell} \quad \Leftrightarrow \quad K - \lim_{n \to \infty} F_n = F \quad \Leftrightarrow \quad \delta(F_n, F) \to 0. \tag{42}$$

Proposition 2.6. *Assume that* $f_n \xrightarrow{s} f$ *in* D *with* $A(f) = \{x\}$. *If in addition there exists a compact* $K \subseteq \mathbb{R}^d$ *such that*

$$A(f_n) \cap K \neq \emptyset \text{ for eventually all } n \in \mathbb{N}, \tag{43}$$

then the following hold:

(a) $A(f_n) \to \{x\}$ *in* τ_{Fell}.

(b) $K - \lim_{n \to \infty} A(f_n) = \{x\}$.

Proof. First recall that by (42) the statements (a) and (b) are equivalent, whence it remains to prove (b). For that purpose observe that by assumption there exist a natural number N and points $\xi_n \in A(f_n) \cap K$ for all $n \geq N$. Let $(\xi_{n_k})_{k \geq 1}$ be an arbitrary subsequence of $(\xi_n)_{n \geq N}$. Since (ξ_{n_k}) runs through the compact set K it contains a convergent subsequence $(\xi_{n_{k(m)}})_{m \geq 1}$ with $\xi_{n_{k(m)}} \to \xi \in K$, $m \to \infty$. As $\xi_{n_{k(m)}} \in A(f_{n_{k(m)}})$ for all $m \in \mathbb{N}$ we can infer that $\xi \in K - \limsup_{n \to \infty} A(f_n) \subseteq A(f) = \{x\}$, where the inclusion follows from Proposition 2.2 (b). Consequently, $\xi = x$ and the sequence $(\xi_n)_{n \geq N}$ has exactly one cluster point resulting in $\xi_n \to x$, $n \to \infty$, which in turn by definition of the inner limit guarantees that $x \in K - \liminf_{n \to \infty} A(f_n)$. Therefore, by (41) we obtain that

$$\{x\} \subseteq K - \liminf_{n \to \infty} A(f_n) \subseteq K - \limsup_{n \to \infty} A(f_n) \subseteq A(f) = \{x\},$$

which gives the desired result. □

Let $\mathcal{B}_{miss}$ and $\mathcal{B}_{Fell}$ denote the Borel-σ-algebra generated by the missing- or Fell-topology, respectively. Since $\tau_{Fell} \supseteq \tau_{miss}$ we have that $\mathcal{B}_{Fell} \supseteq \mathcal{B}_{miss}$. As it happens both σ-algebras actually coincide:

$$\mathcal{B}_{Fell} = \mathcal{B}_{miss}. \tag{44}$$

To see the remaining reverse inclusion note that [22] prove that

$$\mathcal{B}_{Fell} = \sigma(\{\mathcal{M}(K) : K \in \mathcal{K}\}), \tag{45}$$

whence in fact $\mathcal{B}_{Fell} \subseteq \mathcal{B}_{miss}$, because $\{\mathcal{M}(K) : K \in \mathcal{K}\} \subseteq \tau_{miss}$. As a simple but very useful consequence we obtain

Proposition 2.7. *The map* A: $(D, \mathcal{D}) \to (\mathcal{F}, \mathcal{B}_{Fell})$ *is measurable.*

Proof. By Proposition 2.2 and Theorem 4.10 in [7] the map A: $(D, s) \to (\mathcal{F}, \tau_{miss})$ is continuous and hence $\mathcal{D} - \mathcal{B}_{miss}$ measurable, which in view of (44) immediately gives the result. □

Recall $D' := \{f \in D : A(f) \neq \emptyset\}$. The axiom of choice guarantees the existence of a mapping $a : D' \to \mathbb{R}^d$ such that $a(f) \in \mathrm{A}(f)$ for all $f \in D'$, which is commonly known as *choice function*. If $\mathcal{D}'$ denotes the trace of $\mathcal{D}$ on D', then the question is whether one can find a choice function, which is actually measurable. Notice that $D' = \mathrm{A}^{-1}(\mathcal{H}(\mathbb{R}^d)) \in \mathcal{D}$ by Proposition 2.7, because $\mathcal{H}(\mathbb{R}^d) \in \tau_{Fell}$ is a Borel-set. For this reason $\mathcal{D}' = \{O \in \mathcal{D} : O \subseteq D'\}$. A measurable choice function $a : (D', \mathcal{D}') \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel-σ algebra on $\mathbb{R}^d$ is called *mesaurable selection.*

Proposition 2.8. *(existence of measurable selection) There exists a* $\mathcal{D}' - \mathcal{B}(\mathbb{R}^d)$ *measurable function* $a : D' \to \mathbb{R}^d$ *with* $a(f) \in A(f)$ *for all* $f \in D'$.

Proof. By Proposition 2.7 A is a closed-valued measurable mapping and the assertion follows immediately from Corollary 14.6 of [21]. □

3. Continuous Mapping Theorems for arginf-sets and measurable selections

Let $X = \{X(t) : t \in \mathbb{R}^d\}$ be a stochastic process defined on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with cadlag trajectories. For $k \in \mathbb{N}$ points $t_1, \ldots, t_k$ in $\mathbb{R}^d$ consider the natural projections $\pi_{t_1,\ldots,t_k}$ from D to $\mathbb{R}^k$ defined by $\pi_{t_1,\ldots,t_k}(f) := (f(t_1), \ldots, f(t_k))$. Theorem 2 of [12] says that $\mathcal{D} = \sigma(\pi_{t_1,\ldots,t_k} : t_1, \ldots, t_k \in \mathbb{R}^d, k \in \mathbb{N})$. Since $\pi_{t_1,\ldots,t_k} = (\pi_{t_1}, \ldots, \pi_{t_k})$ is measurable if and only if each component π_{t_j} is measurable we also have that

$$\mathcal{D} = \sigma(\pi_t : t \in \mathbb{R}^d). \tag{46}$$

By (46) each such process can be identified with a measurable map $X : (\Omega, \mathcal{A}) \to (D, \mathcal{D})$. Therefore Proposition 2.7 ensures that $A(X) = A \circ X$ is a *random closed set* in the sense that $A(X) : (\Omega, \mathcal{A}) \to (\mathcal{F}, \mathcal{B}_{Fell})$ is measurable. Similarly, if $A(X)$ is nonempty on Ω it follows by Proposition 2.8 that $a(X) : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is measurable as well. More generally, if $\Omega' := \{A(X) \neq \emptyset\} = \{X \in D'\} \neq \emptyset$, then $\Omega' \in \mathcal{A}$ since $D' \in \mathcal{D}$ as shown above. Consequently, the trace $\mathcal{A}' := \mathcal{A} \cap \Omega'$ of $\mathcal{A}$ in Ω' is given by $\mathcal{A}' = \{A \in \mathcal{A} : A \subseteq \Omega'\}$. Clearly, the restriction $X : (\Omega', \mathcal{A}') \to (D', \mathcal{D}')$ is measurable, whence by Proposition 2.8

$$a(X) : (\Omega', \mathcal{A}') \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \tag{47}$$

is a well-defined measurable map.

3.1. **Convergence in probability and almost surely.** We start with Continuous Mapping Theorems concerning almost sure (a.s.) convergence of the involved random elements. If there is an analogon for convergence in probability, then it is formulated simultaneously in parantheses. In each case these counterparts follow form the subsequence criterion, confer Lemma 3.2 in [11], which is valid for random variables in separable metric spaces. As a consequence results about the missing-topology are excluded, since this topology is not metrizable, confer [9].

Theorem 3.1. *Let X and $X_n, n \in \mathbb{N}$, be cadlag stochastic processes with*

$$X_n \xrightarrow{s} X \quad a.s. \quad \textit{(in probability)}, \; n \to \infty. \tag{48}$$

Then

$$A(X_n) \to A(X) \quad \textit{in } \tau_{Miss} \;\; a.s., \; n \to \infty. \tag{49}$$

Let us further assume that there exists a measurable map $\xi : (\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, which is a.s. the unique infimizing point of X, i.e.

$$A(X) = \{\xi\} \quad a.s. \tag{50}$$

and that with probability one:

$$\textit{There exists a compact } K \textit{ with } A(X_n) \cap K \neq \emptyset \textit{ for eventually all } n \in \mathbb{N}. \tag{51}$$

Then

$$A(X_n) \to A(X) = \{\xi\} \quad \textit{in } \tau_{Fell} \;\; a.s. \quad \textit{(in probability)}. \tag{52}$$

In particular, the events in (48) - (52) are in fact measurable subsets of $(\Omega, \mathcal{A})$.

Proof. Conclude from Proposition 2.2 that

$$\Omega_0 := \{X_n \xrightarrow{s} X\} \subseteq \{A(X_n) \to A(X) \text{ in } \tau_{Miss}\}$$

and that the first set $\Omega_0 \in \mathcal{A}$ because $s(X_n, X)$ is a real random variable by Proposition 8.1.4 of [10] upon noticing that (D, s) is separable. Since without loss of generality we do assume that $(\Omega, \mathcal{A}, \mathbb{P})$ is complete it follows that the superset is measurable as well and hence (49) follows immediately.
For the proof of the second part (52) let $(C_k)_{k\in\mathbb{N}}$ be a sequence of compact sets such that $C_k \uparrow \mathbb{R}^d$, for instance $C_k = [-k, k]^d$. Observe that

$$\Omega_1 := \{\omega \in \Omega : (51) \text{ holds}\} = \bigcup_{k\in\mathbb{N}} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \{A(X_n) \cap C_k \neq \emptyset\} \in \mathcal{A},$$

because A(X_n) are random closed sets and therefore

$$\{A(X_n) \cap C_k \neq \emptyset\} = (A(X_n))^{-1}((\mathcal{M}(C_k))^c) \in \mathcal{A}.$$

Notice that measurability of ξ entails $\{\{\xi\} \in \mathcal{M}(K)\} = \{\xi \notin K\} \in \mathcal{A}$ for each compact K, whence $\{\xi\}$ is a random closed set by (45). Thus another application of Proposition 8.1.4 of [10] ensures that $\delta(A(X), \{\xi\})$ is a real random variable, for $(\mathcal{F}, \delta)$ is separable, because $(\mathcal{F}, \tau_{Fell})$ is second countable. Consequently, $\Omega_2 := \{A(X) = \{\xi\}\} \in \mathcal{A}$ and so we have that $\Omega_0 \cap \Omega_1 \cap \Omega_2 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0 \cap \Omega_1 \cap \Omega_2) = 1$ by (48), (50) and (51). Now Proposition 2.6 guarantees that

$$\Omega_0 \cap \Omega_1 \cap \Omega_2 \subseteq \{A(X_n) \to \{\xi\} \text{ in } \tau_{Fell}\} = \{\delta(A(X_n), \{\xi\}) \to 0\} \in \mathcal{A}$$

resulting in (52). □

Our next results straightforwardly follow from Propositions 2.3–2.5. and constitute their stochastic counterparts.

Theorem 3.2. *Assume that $X_n \stackrel{s}{\to} X$ a.s. with $A(X) \neq \emptyset$ a.s. If ξ_n are random variables with*

$$\xi_n \in A(X_n) \quad a.s. \quad \forall\, n \geq N \in \mathbb{N} \tag{53}$$

and such that $(\xi_n)_{n \geq N}$ is bounded a.s., then

$$d(\xi_n, A(X)) \to 0 \quad a.s.$$

In case of a.s. uniqueness (50) it follows that

$$\xi_n \to \xi \quad a.s.$$

Proof. Let $\Omega_0 := \{X_n \stackrel{s}{\to} X\}, \Omega_1 := \{A(X) \neq \emptyset\}, \Omega_2 := \bigcap_{n \geq N}\{\xi_n \in A(X_n)\}$ and $\Omega_3 := \{(\xi_n)_{n\geq N}$ is bounded$\}$. We easily infer from the assumptions that $\mathbb{P}(\Omega_0 \cap \ldots \cap \Omega_3) = 1$. On the other hand $\Omega_0 \cap \ldots \cap \Omega_3 \subseteq \{d(\xi_n, A(X)) \to 0\}$ by Proposition 2.3, whence the first assertion follows immediately. The derivation of the second assertion works in the same fashion by using Corollary 2.1. □

As to the existence of the random variables ξ_n recall that $\Omega_n' := \{A(X_n) \neq \emptyset\} \in \mathcal{A}$. If $\mathbb{P}(\Omega_n') = 1$, then we are free to define $\xi_n := a(X_n)$ on Ω_n' and, e.g., $\xi_n := 0$ on $\Omega \setminus \Omega_n'$. Conclude from (47) that ξ_n is a random variable on $(\Omega, \mathcal{A})$ with $\xi_n \in A(X_n)$ on Ω_n', and thus with probability one as desired. Random variables ξ_n with $\xi_n \in A(X_n)$ a.s. are also known as *(measurable) selections* of $A(X_n)$. In statistical applications X_n plays the role of a random criterion function, which typically has certain path properties allowing to find an explicit solution of the pertaining minimization problem. In general this solution is easily seen to be measurable.

Notice that by Example 2.1 the a.s. boundedness requirement on the selections $\xi_n \in A(X_n)$ is inevitable. On the other hand this is hard to prove even in a special framework. The following proposition offers a way out.

Theorem 3.3. *Let $(T_n) \subseteq \mathbb{R}^d$ be a sequence of open sets (possibly random) such that $\liminf_{n\to\infty} T_n = \mathbb{R}^d$ a.s. and assume that*

$$\sup_{t \in T_n} |X_n(t) - X(t)| \to 0 \quad a.s. \qquad (in\ probability),$$

where the limit process X possesses a.s. a well-separeted infimizing point ξ. If $\xi_n \in T_n$ is an infimizing point of the restriction of X_n on T_n a.s. for each $n \geq N \in \mathbb{N}$, i.e.,

$$\min_{R \in \{<,\geq\}^d} X_n(\xi_n + R) = \inf_{t \in T_n} X_n(t) \quad a.s. \quad \forall\, n \geq N \tag{54}$$

then:

$$\xi_n \to \xi \quad a.s. \quad (in\ probability).$$

Proof. Use Proposition 2.5 upon noticing that $(\Omega, \mathcal{A}, \mathbb{P})$ is complete. □

Remark 3.1. All of our results so far remain valid if $\mathbb{R}^d$ is replaced by an open subset $O \subseteq \mathbb{R}^d$. Notice that $\mathcal{K}(O) = \{K \in \mathcal{K}(\mathbb{R}^d) : K \subset O\}$. In particular, boundedness of a sequence $(x_n)_{n\in\mathbb{N}}$ in O means that there exists a compact $K \subset O$ such that $x_n \in K$ for all $n \in \mathbb{N}$. Thus, for instance, $x_n = 1/n$ is not bounded in the open interval $O = (0, 1)$.

3.2. **Convergence in distribution.** In this section we present several Continuous Mapping Theorems for the functionals $A = \text{Arginf}$ and $a = \text{arginf}$ under the assumption that the X_n converge in distribution as random variables in the multivariate Skorokhod-space (D, s). We come up with semi-convergence (or: inner-approximation) in distribution in the sense of [29] and in particular with the new concept of *quasi-distributional convergence* for random closed sets. As a consequence there will also be an extension of the classical notion of weak convergence of probability measures, where now the limit is allowed to be a Choquet-capacity.

Theorem 3.4. *Assume that* $X_n \xrightarrow{\mathcal{L}} X$ *in* (D, s)*. Let* $\varphi_n, n \in \mathbb{N}$*, be random closed sets with* $\varphi_n \subseteq A(X_n)$ *a.s. for every* $n \in \mathbb{N}$*. Then*

$$\varphi_n \xrightarrow{\mathcal{L}} A(X) \quad \text{in } (\mathcal{F}, \tau_{miss}), \tag{55}$$

which is equivalent to semi-convergence in distribution*, i.e.*

$$\limsup_{n\to\infty} \mathbb{P}\Big(\bigcap_{j=1}^{r}\{\varphi_n \cap K_j \neq \emptyset\}\Big) \leq \mathbb{P}\Big(\bigcap_{j=1}^{r}\{A(X) \cap K_j \neq \emptyset\}\Big) \quad \forall\, r \in \mathbb{N} \quad \forall\, K_1, \ldots, K_r \in \mathcal{K}. \tag{56}$$

If in addition (φ_n) is stochastically bounded *in the sense that*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}(\varphi_n \not\subseteq [-k, k]^d) = 0,$$

then we obtain quasi-convergence in distribution*, i.e.*

$$\limsup_{n\to\infty} \mathbb{P}\Big(\bigcap_{j=1}^{r}\{\varphi_n \cap F_j \neq \emptyset\}\Big) \leq \mathbb{P}\Big(\bigcap_{j=1}^{r}\{A(X) \cap F_j \neq \emptyset\}\Big) \quad \forall\, r \in \mathbb{N} \quad \forall\, F_1, \ldots, F_r \in \mathcal{F}. \tag{57}$$

We denote this by

$$\varphi_n \xrightarrow{q-\mathcal{L}} A(X) \quad \text{in } \mathcal{F}.$$

Proof. For every $f \in SC(\mathbb{R}^d)$ the *epigraph* of f is defined by

$$\text{epi}(f) := \{(x, y) \in \mathbb{R}^{d+1} : f(x) \leq y\}.$$

It is easy to see that the correspondence between functions and epigraphs is one-to-one. By Proposition 1.7 in [6] the epigraph of f is closed in $\mathbb{R}^{d+1}$, i.e., $\text{epi}(f) \in \mathcal{F}(\mathbb{R}^{d+1})$. Moreover, by Theorem 4.16 in [6] the following equivalence holds:

$$f_n \to f \text{ in } \text{SC}(\mathbb{R}^d, e) \Leftrightarrow \text{epi}(f_n) \to \text{epi}(f) \text{ in } (\mathcal{F}(\mathbb{R}^{d+1}), \delta) \tag{58}$$

Thus, if $E := \{\text{epi}(f) : f \in SC(\mathbb{R}^d\}$ denotes the subspace of all epigraphs, then

$$\text{epi} : (SC(\mathbb{R}^d), e) \to (E, \delta)$$

is a homeomorphism. Now, $X_n \xrightarrow{\mathcal{L}} X$ in (D, s) and Proposition 2.1 in combination with the (traditional) Continuous Mapping Theorem (CMT) yield:

$$\bar{X}_n \xrightarrow{\mathcal{L}} \bar{X} \quad \text{in } (SC(\mathbb{R}^d), e). \tag{59}$$

Since the map epi is a homeomorhpism, another application of the CMT together with Lemma 3.26 in [11] for subspaces ensures that (59) is equivalent to

$$\text{epi}(\bar{X}_n) \xrightarrow{\mathcal{L}} \text{epi}(\bar{X}) \quad \text{in } (\mathcal{F}(\mathbb{R}^{d+1}), \delta).$$

Therefore Pflug's [19] Theorem 1.3 and our Lemma 2.2 show that (56) holds for the special case $\varphi_n = A(X_n)$. The general case $\varphi_n \subseteq A(X_n)$ a.s. then follows immediately upon noticing that $\{\varphi_n \cap K_j \neq \emptyset, \varphi_n \subseteq A(X_n)\} \subseteq \{A(X_n) \cap K_j \neq \emptyset\}$. Furthermore, a combination of Vogel's [29] Lemma 2.1 with the Portmanteau-Theorem in topological spaces, confer Proposition 8.4.9 in [10], guarantees the equivalence of (55) and (56). Finally, quasi-convergence in distribution (57) follows immediately from Proposition 1.9 of [8]. □

Theorem 3.5. *Assume that* $X_n \xrightarrow{\mathcal{L}} X$ *in* (D, s) *and let* φ_n*,* $n \in \mathbb{N}$*, be random closed sets. If*

$$\emptyset \neq \varphi_n \subseteq A(X_n) \;\; a.s. \quad \forall\, n \geq N \in \mathbb{N},$$

$$(\varphi_n)_{n\geq N} \text{ is stochastically bounded}$$

and

$$A(X) = \{\xi\} \text{ a.s. for some random variable } \xi,$$

then

(60) $$\varphi_n \xrightarrow{\mathcal{L}} A(X) \quad \text{in } (\mathcal{F}, \tau_{Fell}).$$

Proof. From Theorem 3.4 we know that

(61) $$\varphi_n \xrightarrow{\mathcal{L}} A(X) \text{ in } (\mathcal{F}, \tau_{miss}).$$

Put $\Omega_0 := \bigcap_{n \geq N}\{\emptyset \neq \varphi_n \subseteq A(X_n)\} \cap \{A(X) = \{\xi\}\}$. Then $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ and $\mathcal{A}_0 = \{A \in \mathcal{A} : A \subseteq \Omega_0\}$ is the trace of $\mathcal{A}$ in Ω_0. Consequently the semi-convergence (61) also holds for the restrictions of $\varphi_n, n \geq N$, and $A(X)$ on $(\Omega_0, \mathcal{A}_0)$. Therefore an application of Proposition 1.9 of [8] immediately yields the desired result. ☐

Let us recapitulate our last results. In (55) and (60) we obtain weak convergence of φ_n to $A(X)$ in the topological spaces $(\mathcal{F}, \tau_{miss})$ and $(\mathcal{F}, \tau_{Fell})$, respectively. Thus the question arises whether there is a third hyperspace topology on $\mathcal{F}$ such that the induced weak convergence coincides with quasi-convergence in distribution (57). A natural candidate for that is the *upper Vietoris topology* τ_{uV}, which is generated by the family $\{\mathcal{M}(F) : F \in \mathcal{F}\}$. If $\mathcal{B}_{uV} := \sigma(\tau_{uV})$ denotes the pertaining Borel-σ algebra, then $\mathcal{B}_{uV} = \mathcal{B}_{Fell}$. To see this first notice that $\tau_{uV} \supseteq \tau_{miss}$, whence $\mathcal{B}_{uV} = \sigma(\tau_{uV}) \supseteq \sigma(\tau_{miss}) = \mathcal{B}_{miss} = \mathcal{B}_{Fell}$, where the last equality is stated in (44). On the other hand $\mathcal{B}_{Fell} = \sigma(\{\mathcal{M}(F) : F \in \mathcal{F}\})$ by Theorem 14.3(b) of [21] and so $\mathcal{B}_{Fell} \subseteq \mathcal{B}_{uV}$ upon noticing that $\{\mathcal{M}(F) : F \in \mathcal{F}\} \subseteq \tau_{uV}$. Consequently,

(62) $$\mathcal{B}_{uV} = \mathcal{B}_{Fell} = \mathcal{B}_{miss}.$$

In the literature a map $\varphi : (\Omega, \mathcal{A}) \to (\mathcal{F}, \tau_{Fell})$, which is Borel-measurable is called *random closed set*. Thus (62) means that every Borel-measurable map $\varphi : (\Omega, \mathcal{A}) \to (\mathcal{F}, \tau)$ is a random closed set no matter which hyperspace topology $\tau \in \{\tau_{Fell}, \tau_{miss}, \tau_{uV}\}$ is used.

Since for finitely many (arbitrary) subsets $F_1, \ldots, F_m$ of $\mathbb{R}^d$ one has $\bigcap_{i=1}^m \mathcal{M}(F_i) = \mathcal{M}(\bigcup_{i=1}^m F_i)$ and since $\mathcal{F}$ is closed with respect to finite unions it follows that

(63) $$\{\mathcal{M}(F) : F \in \mathcal{F}\}$$

is a base for the upper Vietoris topology τ_{uV}. Unfortunately, as Gerald Beer, California State University, points out (private communication) there is no countable sub-collection of (63) which remains to be a base. This has an effect on the relation between weak convergence in $(\mathcal{F}, \tau_{uV})$ and *quasi-convergence in distribution*, which we define for general random closed sets in (iv) of the following theorem.

Proposition 3.1. *Let $\varphi_n, n \in \mathbb{N}$, and φ be random closed sets defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Consider the following list of statements:*

(i) $\varphi_n \xrightarrow{\mathcal{L}} \varphi$ *in* $(\mathcal{F}, \tau_{uV})$.

(ii) $\liminf_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{O}) \geq \mathbb{P}(\varphi \in \mathcal{O})$ *for all* $\mathcal{O} \in \tau_{uV}$.

(iii) $\limsup_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{C}) \leq \mathbb{P}(\varphi \in \mathcal{C})$ *for all* τ_{uV}*-closed* $\mathcal{C}$.

(iv) $\varphi_n \xrightarrow{q-\mathcal{L}} \varphi$, *i.e.,* by definition

$$\limsup_{n\to\infty} \mathbb{P}(\bigcap_{j=1}^r \{\varphi_n \cap F_j \neq \emptyset\}) \leq \mathbb{P}(\bigcap_{j=1}^r \{\varphi \cap F_j \neq \emptyset\}) \quad \forall\, r \in \mathbb{N} \quad \forall\, F_1, \ldots, F_r \in \mathcal{F}.$$

(v)

$$\limsup_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{C}) \leq \mathbb{P}(\varphi \in \mathcal{C}) \text{ for all } \mathcal{C} \in \mathfrak{C} := \{\bigcap_{i\in\mathbb{N}} \mathcal{H}(F_i) : (F_i)_{i\in\mathbb{N}} \subseteq \mathcal{F}\}.$$

(vi)

$$\liminf_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{O}) \geq \mathbb{P}(\varphi \in \mathcal{O}) \text{ for all } \mathcal{O} \in \mathfrak{O} := \{\bigcup_{i\in\mathbb{N}} \mathcal{M}(F_i) : (F_i)_{i\in\mathbb{N}} \subseteq \mathcal{F}\}.$$

Then the following relations hold:

$$(i) \Leftrightarrow (ii) \Leftrightarrow (iii) \Rightarrow (iv) \Leftrightarrow (v) \Leftrightarrow (vi)$$

Proof. The first equivalence holds by definition and the second one follows easily by complementation. (iv) is implied by (iii), because $\bigcap_{j=1}^r \{\varphi_n \cap F_j \neq \emptyset\} = \{\varphi_n \in \bigcap_{j=1}^r \mathcal{H}(F_j)\}$ and $\bigcap_{j=1}^r \mathcal{H}(F_j) = \mathcal{F} \setminus \bigcup_{j=1}^r \mathcal{M}(F_j)$ is τ_{uV}-closed. To see $(iv) \Rightarrow (v)$ we adapt the arguments of [29] in the proof of her Lemma 2.1. Suppose that $\limsup_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{C}) > \mathbb{P}(\varphi \in \mathcal{C})$ for at least one $\mathcal{C}$ of the type $\mathcal{C} = \bigcap_{i\in\mathbb{N}} \mathcal{H}(F_i)$ with closed sets F_i. Then $\mathcal{C}_k := \bigcap_{i=1}^k \mathcal{H}(F_i) \downarrow \mathcal{C}, k \to \infty$ and therefore

$$\mathbb{P}(\varphi \in \mathcal{C}_k) \downarrow \mathbb{P}(\varphi \in \mathcal{C}),\ k \to \infty. \tag{64}$$

By assumption $\alpha := (\limsup_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{C}) - \mathbb{P}(\varphi \in \mathcal{C}))/2$ is positive and such that $\limsup_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{C}) > \mathbb{P}(\varphi \in \mathcal{C}) + \alpha$, whence there exists a subsequence $(n_l)_{l\geq 1}$ of the natural numbers with

$$\mathbb{P}(\varphi_{n_l} \in \mathcal{C}) > \mathbb{P}(\varphi \in \mathcal{C}) + \alpha \quad \forall\, l \in \mathbb{N}. \tag{65}$$

According to (64) there is a natural k such that $\mathbb{P}(\varphi \in \mathcal{C}) + \alpha/2 > \mathbb{P}(\varphi \in \mathcal{C}_k)$. Since $\mathcal{C}_k \supseteq \mathcal{C}$ we arrive with (65) at

$$\mathbb{P}(\varphi_{n_l} \in \mathcal{C}_k) \geq \mathbb{P}(\varphi_{n_l} \in \mathcal{C}) > \mathbb{P}(\varphi \in \mathcal{C}) + \alpha/2 + \alpha/2 > \mathbb{P}(\varphi \in \mathcal{C}_k) + \alpha/2 \quad \forall\, l \in \mathbb{N}. \tag{66}$$

Taking the limit $l \to \infty$ in (66) we obtain

$$\begin{aligned}
\mathbb{P}(\varphi \in \mathcal{C}_k) + \alpha/2 &\leq \limsup_{l\to\infty} \mathbb{P}(\varphi_{n_l} \in \mathcal{C}_k) && \text{by (66)}\\
&\leq \limsup_{n\to\infty} \mathbb{P}(\varphi_n \in \mathcal{C}_k)\\
&= \limsup_{n\to\infty} \mathbb{P}(\textstyle\bigcap_{j=1}^k \{\varphi_n \cap F_j \neq \emptyset\})\\
&\leq \mathbb{P}(\textstyle\bigcap_{j=1}^k \{\varphi \cap F_j \neq \emptyset\}) && \text{by (iv)}\\
&= \mathbb{P}(\varphi \in \mathcal{C}_k),
\end{aligned}$$

which is a contradiction to $\alpha > 0$. This shows that (v) holds. Furthermore, $\bigcap_{j=1}^r \mathcal{H}(F_j) = \bigcap_{j=1}^\infty \mathcal{H}(F_j)$, if $F_j := F_r$ for all $j > r$, and thus (iv) follows immediately from (v). Finally, by complementation (v) and (vi) are equivalent. □

Proposition 3.1 tells us that weak convergence in $(\mathcal{F}, \tau_{uV})$ entails quasi-convergence in distribution, but whether the reverse direction holds is questionable. In fact, for instance in 3.1 (vi) we only have that $\mathfrak{O} \subseteq \tau_{uV}$. If we had equality here, than indeed weak convergence would follow. However, recall that there is no countable base for τ_{uV}, which in turn would guarantee this equality. Therefore, we conjecture that in general weak convergence in $(\mathcal{F}, \tau_{uV})$ does not follow from quasi-convergence in distribution. As a way out we go over to a smaller version of the upper Vietoris topology, which on the other side is still finer than the missing-topology. For that purpose let $\mathcal{F}^* \subseteq \mathcal{F} \setminus \mathcal{K}$ be any countable family of unbounded closed sets. By analogy to τ_{uV} we define τ_{uV}^* to be the weakest topology that contains $\{\mathcal{M}(K) : K \in \mathcal{K}\} \cup \{\mathcal{M}(F) : F \in \mathcal{F}^*\}$. Of course we wish $\mathcal{F}^*$ to be as big as possible. For example for every $p \geq 1$ let $B_p(x,r) := \{y \in \mathbb{R}^d : |y - x|_p < r\}$ be the open ball at $x \in \mathbb{R}^d$ with radius r with respect to the L_p-norm $|\cdot|_p$ on $\mathbb{R}^d$. Then

$$\mathcal{F}_p := \{[\bigcup_{i=1}^m B_p(x_i, r_i)]^c : m \in \mathbb{N}, x_i \in \mathbb{Q}^d, 0 < r_i \in \mathbb{Q}\ \ \forall\, 1 \leq i \leq m\}$$

consists of all complements of finite unions of *rational* balls. If $\mathcal{F}_\infty$ combines all complements of finite unions of open rectangles with rational endpoints then $\mathcal{F}^* = \bigcup_{1\leq p\in\mathbb{Q}} \mathcal{F}_p \cup$

$\mathcal{F}_\infty$ is countable. One can enlarge $\mathcal{F}^*$ in many ways. For instance consider a denumerable system containing epigraphs of lsc functions and hypographs of uper-semicontinuous functions, straight lines, planes, affine subspaces or orthants. Then we may add any finite union of these unbounded closed sets. In the sequel we give a countable base for our new topology. Let $\mathcal{K}^*$ be the family of all finite unions of compact rectangles with rational endpoints.

Lemma 3.1. *τ^*_{uV} has a countable base, which is given by*

$$\{\mathcal{M}(K \cup F) : K \in \mathcal{K}^*, F \in \mathcal{F}^*_u\},$$

where $\mathcal{F}^*_u$ denotes the collection of all finite unions of $\mathcal{F}^*$-sets.

Proof. Since $\mathcal{K}^*$ and $\mathcal{F}^*_u$ are countable, so is the given system as well. To see that it is a base first observe that according to the construction a general base-set is given by

$$\bigcap_{i=1}^{r} \mathcal{M}(K_i) \cap \bigcap_{j=1}^{s} \mathcal{M}(F_j) = \mathcal{M}(\bigcup_{i=1}^{r} K_i \cup \bigcup_{j=1}^{s} F_j)$$

with $K_i \in \mathcal{K}$ and $F_j \in \mathcal{F}^*$.

In view of $\bigcup_{i=1}^r K_i =: K \in \mathcal{K}$ and $\bigcup_{j=1}^s F_j =: F \in \mathcal{F}^*_u$ we see that $\{\mathcal{M}(K \cup F) : K \in \mathcal{K}, F \in \mathcal{F}^*_u\}$ is the canonical base of τ^*_{uV}. Therefore each open $\mathcal{O} \in \tau^*_{uV}$ can be represented as

$$\mathcal{O} = \bigcup_{i \in I} \bigcup_{j \in J} \mathcal{M}(K_i \cup F_j) \tag{67}$$

with some $(K_i)_{i\in I} \subseteq \mathcal{K}$ and $(F_j)_{j\in J} \subseteq \mathcal{F}^*_u$ and some index-sets I and J.

Next, we show that for every compact K there exists a sequence $(C_n)_{n\in\mathbb{N}} \subseteq \mathcal{K}^*$ such that $C_n \downarrow K$. In fact, observe that for each $n \in \mathbb{N}$ we have that

$$K \subseteq \bigcup_{x\in K} (x - 1/n, x + 1/n) \subseteq \bigcup_{x\in K} [x - 1/n, x + 1/n].$$

Since K has a finite subcovering, we find $m_n \in \mathbb{N}$ points $x_i(n) \in K, 1 \le i \le m_n$ such that

$$K \subseteq \bigcup_{i=1}^{m_n} [x_i(n) - 1/n, x_i(n) + 1/n] =: K_n.$$

Put $C_n := K_1 \cap \ldots \cap K_n, n \in \mathbb{N}$. By the distributive law $\mathcal{K}^*$ is closed with respect to finite intersections, whence $(C_n) \subseteq \mathcal{K}^*$ and clearly is monotone decreasing. Moreover,

$$\bigcap_{n\ge1} C_n = K. \tag{68}$$

In fact, the inclusion $\supseteq$ is trivial, because $K_j \supseteq K$ for all j. So, let $x \in C_n$ for all $n \in \mathbb{N}$. Since $C_n \subseteq K_n$ we find for each $n \in \mathbb{N}$ an index $i_n \in \{1, \ldots, m_n\}$ such that $|x - x_{i_n}(n)|_\infty \le 1/n$. Consequently, $\xi_n := x_{i_n}(n) \in K$ converges to x, which must lie in K, because K is closed. This shows (68).

It follows that

$$\mathcal{M}(K) = \bigcup_{n\ge1} \mathcal{M}(C_n). \tag{69}$$

Again, the inclusion $\supseteq$ is trivial upon noticing that $C_n \supseteq K$. Therefore, let $F \in \mathcal{M}(K)$, which means that $F \cap K = \emptyset$. Now, assume that $F \notin \bigcup_{n\ge1} \mathcal{M}(C_n)$, i.e., $F \cap C_n \ne \emptyset \; \forall\, n \in \mathbb{N}$. Thus there exists a sequence (x_n) with $x_n \in F \cap C_n$ for all $n \ge 1$. In particular, for each $m \in \mathbb{N}$ the sequence $(x_n)_{n\ge m}$ runs through the compact set $F \cap C_m$, because $C_n \subseteq C_m$. Hence $(x_n)_{n\ge m}$ converges to some $x \in F \cap C_m$ as $n \to \infty$ for all $m \ge 1$, where x does not depend on m. Infer that $x \in F \cap \bigcap_{m\geq 1} C_m = F \cap K$ by (68). This is a contradiction to $F \cap K = \emptyset$, which yields (69).

Finally, let $\mathcal{O} \in \tau^*_{uV}$ be an arbitrary open set, which as we have seen admits the representation (67). According to (69) for every $i \in I$ there exists a sequence $(C_n^{(i)})_{n\geq 1} \subseteq \mathcal{K}^*$ such that $\mathcal{M}(K_i) = \bigcup_{n\geq 1} \mathcal{M}(C_n^{(i)})$. Inserting this in (67) and taking into account that $\mathcal{M}(K_i \cup F_j) = \mathcal{M}(K_i) \cap \mathcal{M}(F_j)$ gives by the distributive law:

$$\begin{aligned}(70)\quad \mathcal{O} &= \bigcup_{i\in I}\bigcup_{j\in J}\bigcup_{n\geq 1}(\mathcal{M}(C_n^{(i)}) \cap \mathcal{M}(F_j)) \\ &= \bigcup_{i\in I}\bigcup_{j\in J}\bigcup_{n\geq 1}\mathcal{M}(C_n^{(i)} \cup F_j) = \bigcup_{j\in J}\bigcup_{(i,n)\in(I\times\mathbb{N})}\mathcal{M}(C_n^{(i)} \cup F_j).\end{aligned}$$

This yields the desired result upon noticing that $\{F_j : j \in J\} \subseteq \mathcal{F}^*_u$ and $\{C_n^{(i)} : (i,n) \in (I \times \mathbb{N})\} \subseteq \mathcal{K}^*$. □

By construction $\tau_{miss} \subseteq \tau^*_{uV} \subseteq \tau_{uV}$, and so by (62) the Borel-σ algebra $\mathcal{B}^*_{uV} := \sigma(\tau^*_{uV})$ also coincides with $\mathcal{B}_{Fell}$. Thus we can extend our conclusion of (62) by saying that every Borel-measurable map $\varphi : (\Omega, \mathcal{A}) \to (\mathcal{F}, \tau)$ is a random closed set whatever hyperspace topology $\tau \in \{\tau_{miss}, \tau^*_{uV}, \tau_{uV}, \tau_{Fell}\}$ is used.

Proposition 3.2. *If $\varphi_n, n \in \mathbb{N}$, and φ are random closed sets, then $\varphi_n \xrightarrow{q-\mathcal{L}} \varphi$ entails $\varphi_n \xrightarrow{\mathcal{L}} \varphi$ in $(\mathcal{F}, \tau^*_{uV})$.*

Proof. Let $\mathcal{O} \in \tau^*_{uV}$. By Lemma 3.1 there exist $(K_i)_{i\geq 1} \subseteq \mathcal{K}^*$ and $(F_i)_{i\geq 1} \subseteq \mathcal{F}^*_u$ such that $\mathcal{O} = \bigcup_{i\geq 1} \mathcal{M}(K_i \cup F_i)$. Since $K_i \cup F_i \in \mathcal{F}$ for all $i \geq 1$, we can infer that $\tau^*_{uV} \subseteq \mathfrak{O}$. An application of Proposition 3.1 yields the desired result. □

Combining Theorem 3.4 and Proposition 3.2 immediately results in the following

Theorem 3.6. *Assume that $X_n \xrightarrow{\mathcal{L}} X$ in (D, s). Let $\varphi_n, n \in \mathbb{N}$, be random closed sets with $\varphi_n \subseteq A(X_n)$ a.s. for every $n \in \mathbb{N}$. If (φ_n) is stochastically bounded then*

$$\varphi_n \xrightarrow{\mathcal{L}} A(X) \quad \text{in } (\mathcal{F}, \tau^*_{uV}).$$

In case of $\varphi_n = A(X_n)$ we obtain that

$$A(X_n) \xrightarrow{\mathcal{L}} A(X) \quad \text{in } (\mathcal{F}, \tau^*_{uV}).$$

We end our investigation of random closed arginf-sets by presenting a counterpart of Pflug's [19] Theorem 1.4 on *confidence sets for infimizing points.* In fact, in section 4 below we will use this result to construct confidence regions for parameters θ in statistics as described in Example 1.1.

Theorem 3.7. *Assume that $X_n \xrightarrow{\mathcal{L}} X$ in (D, s). Let $\varphi_n, n \in \mathbb{N}$, be random closed sets with $\varphi_n \subseteq A(X_n)$ a.s. for every $n \in \mathbb{N}$ and such that (φ_n) is stochastically bounded.*

For a given $\alpha \in (0,1)$ let $G = G_\alpha$ be an open subset of $\mathbb{R}^d$ with $\mathbb{P}(A(X) \subseteq G) \geq 1-\alpha$. Then

$$\liminf_{n\to\infty} \mathbb{P}(\varphi_n \subseteq G) \geq 1 - \alpha.$$

Proof. First note that G^c is closed by assumption. Since $M \subseteq G \Leftrightarrow M \cap G^c = \emptyset$ for every set $M \subseteq \mathbb{R}^d$ it follows from (57) with $r = 1$ by complementation that

$$\liminf_{n\to\infty} \mathbb{P}(\varphi_n \subseteq G) = \liminf_{n\to\infty} \mathbb{P}(\varphi_n \cap G^c = \emptyset) \geq \mathbb{P}(A(X) \cap G^c = \emptyset) = \mathbb{P}(A(X) \subseteq G) \geq 1-\alpha.$$

□

Once we have our results on sets φ_n of infimizing points we easily obtain limit theorems for single selections ξ_n of $A(X_n)$ simply by considering the special case $\varphi_n = \{\xi_n\}$.

Theorem 3.8. *Let ξ_n be random variables with $\xi_n \in A(X_n)$ a.s. $\forall\ n \geq N$. Put $C := A(X)$. If*

$$X_n \stackrel{\mathcal{L}}{\to} X \quad \text{in } (D, s),$$

then it follows that

$$\limsup_{n\to\infty} \mathbb{P}(\xi_n \in K) \leq T_C(K) \quad \forall\ K \in \mathcal{K},$$

where

$$T_C(K) := \mathbb{P}(C \cap K \neq \emptyset) = \mathbb{P}(A(X) \cap K \neq \emptyset).$$

Proof. Set $\varphi_n := \{\xi_n\}, n \geq N$. Then each φ_n is a random closed set with $\varphi_n \subseteq A(X_n)$ a.s. for all $n \geq N$. Since $\{\varphi_n \cap K \neq \emptyset\} = \{\xi_n \in K\}$ for all compact K, the assertion follows from (56) with $r = 1$. ☐

Recall that $C = A(X)$ is a random closed set, that is a measurable mapping $C : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathcal{F}, \mathcal{B}_{Fell})$. In the sequel we may replace $\mathbb{R}^d$ by any locally compact, second countable, Hausdorff space (lcscH) E, confer [14]. Recall that by Uryson's Theorem E is metrizable. Let $\mathcal{F} = \mathcal{F}(E), \mathcal{G} = \mathcal{G}(E), \mathcal{K} = \mathcal{K}(E), \mathcal{B} = \mathcal{B}(E)$ denote the classes of all closed, open, compact, and Borel subsets, respectively. Then the construction of the Fell-topology $\tau_{Fell} = \tau_{Fell}(E)$ and the pertaining Borel-σ algebra $\mathcal{B}_{Fell} = \mathcal{B}_{Fell}(E)$ remains the same including the relation (45), i.e.,

$$\mathcal{B}_{Fell}(E) = \sigma(\{\mathcal{M}(K) : K \in \mathcal{K}(E)\}). \tag{71}$$

In particular every measurable mapping $C : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathcal{F}(E), \mathcal{B}_{Fell}(E))$ is called *random closed set* in E. For any such random closed set C the set-function $T_C : \mathcal{K} \to \mathbb{R}$ defined by $T_C(K) := \mathbb{P}(C \cap K \neq \emptyset)$ is called *capacity functional of C*. The properties of $\mathbb{P}$ lead to the following characteristics of T_C:

(T1) $T_C(\emptyset) = 0$; $0 \leq T_C \leq 1$;

(T2) T_C is *upper-semi-continuous*, i.e.,

$$K_n \downarrow K \text{ in } \mathcal{K} \quad \Rightarrow \quad T_C(K_n) \downarrow T_C(K);$$

(T3) T_C is monotone increasing on $\mathcal{K}$ and for $K_1, K_2, \ldots, K_n \in \mathcal{K}, n \geq 2$,

$$T_C\Big(\bigcap_{i=1}^n K_i\Big) \leq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} T_C\Big(\bigcup_{i\in I} K_i\Big)$$

Every functional $T : \mathcal{K} \to \mathbb{R}$ satisfying (T1)-(T3) is called *Choquet capacity (functional)*. The following two results on Choquet capacities are well-known in the theory of random closed sets, confer [17], [14] or [15].

Theorem 3.9. *(Choquet) Every probability measure Q on $(\mathcal{F}(E), \mathcal{B}_{Fell}(E))$ determines a Choquet capacity functional T on $\mathcal{K}(E)$ through the correspondence*

$$T(K) = Q(\mathcal{H}(K)) \quad \forall\ K \in \mathcal{K}(E). \tag{72}$$

Conversely, every Choquet capacity functional T on $\mathcal{K}(E)$ determines a unique probability measure Q on $(\mathcal{F}(E), \mathcal{B}_{Fell}(E))$ that satisfies the relation (72)

Each Choquet capacity T can be extended onto the Borel-σ algebra $\mathcal{B}(E)$ by

$$T(B) := \sup\{T(K) : K \in \mathcal{K}, K \subseteq B\}, \quad B \in \mathcal{B}(E).$$

Theorem 3.10. *(Matheron) The extension $T : \mathcal{B}(E) \to E$ is consistent in the sense that*

$$T(B) = Q(\mathcal{H}(B)) \quad \forall\, B \in \mathcal{B}(E),$$

where the hitting sets $\mathcal{H}(B) = \{F \in \mathcal{F} : F \cap B \neq \emptyset\}$ in fact belong to $\mathcal{B}_{Fell}(E)$.

As an immediate consequence the extension T has the following properties in addition to (T1) and (T2):

$(T3)^*$ T is monotone increasing on $\mathcal{B}$ and for $B_1, B_2, \ldots, B_n \in \mathcal{B}, n \geq 2$,

$$T\Big(\bigcap_{i=1}^{n} B_i\Big) \leq \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{|I|+1} T\Big(\bigcup_{i \in I} B_i\Big)$$

(T4) T is σ-continuous from below:

$$B_n \uparrow B \ \text{ in } \mathcal{B} \quad \Rightarrow \quad T(B_n) \uparrow T(B).$$

For a better understanding of our following results we like to discuss briefly some facts on capacity functionals. In view of Choquet's Theorem we see that capacity functionals play the same role as distribution functions do in case of random vectors in the euclidian space. In fact there is another analogy, namely one can also characterize weak convergence of random closed sets in terms of capacity functionals, confer [13], [15] or [22]. In contrast to distribution functions a capacity functional T is a set-function defined on the Borel-sets of E with properties (T1), (T2), $(T3^*)$ and (T4), which are satisfied for every probability measure on $\mathcal{B}(E)$. On the other hand T is merely sub-additive and in general lacks additivity. Thus capacity functionals usually are not probability measures, but appropriate generalizations thereof.

Theorem 3.11. *(convergence in distribution to a random closed set) Let ξ_n be random variables with $\xi_n \in A(X_n)$ a.s. $\forall\, n \geq N$. Put $C := A(X)$. If*

$$X_n \xrightarrow{\mathcal{L}} X \qquad \text{in } (D, s),$$

and $(\xi_n)_{n \geq N}$ is stochastically bounded*, i.e.*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}(|\xi_n| > k) = 0,$$

then it follows:

$$\limsup_{n\to\infty} \mathbb{P}(\xi_n \in F) \leq T_C(F) \quad \forall\, F \in \mathcal{F}(\mathbb{R}^d). \tag{73}$$

Proof. As in the proof of Theorem 3.8 we consider the special random closed sets $\varphi_n = \{\xi_n\}$. Recall that $|\cdot|$ denotes the maximum-norm on $\mathbb{R}^d$. Thus $\{|\xi_n| > k\} = \{\varphi_n \nsubseteq [-k,k]^d\}$, whence (φ_n) is stochastically bounded and the assertion follows from (57) with $r = 1$. □

Inequality (73) formally looks exactly like the equivalent characterization of convergence in distribution (weak convergence) given by the Portmanteau-Theorem, confer [4]. However, the main difference between (73) and the Portmanteau-Theorem lies in that here T_C needs not to be a probability measure as pointed out above. Nevertheless, T_C has many features in common with probability measures. This leads us to introduce a new concept of convergence. In our definition below notice that by Choquet's Theorem in combination with the canonical construction every Choquet capacity T is the capacity functional of some random closed set C in E, that means $T = T_C$. (Put $(\Omega, \mathcal{A}, \mathbb{P}) := (\mathcal{F}, \mathcal{B}_{Fell}, Q)$, where Q is the probability measure determined by T, and let $C : \Omega \to \mathcal{F}$ be the identity map.) We say that C *is associated with* T.

Definition 3.1. (a) Let $P_n, n \in \mathbb{N}$, be probability measures on $(E, \mathcal{B}(E))$ and T be a Choquet capacity. If

$$\limsup_{n\to\infty} P_n(F) \leq T(F) \quad \forall\, F \in \mathcal{F}(E),$$

then we say that P_n *converges weakly to* T and denote this by $P_n \stackrel{w-C}{\longrightarrow} T$.

(b) Let $\xi_n : (\Omega_n, \mathcal{A}_n, \mathbb{P}_n) \to (E, \mathcal{B}(E)), n \in \mathbb{N}$, be random variables. If the distributions $\mathbb{P}_n \circ \xi_n^{-1}$ of ξ_n converge weakly to T, then we say that ξ_n *converges in distribution to* C, where $C : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathcal{F}(E), \mathcal{B}_{Fell}(E))$ is the random closed set associated with T. Therefore, we alternatively write:

$$\xi_n \stackrel{\mathcal{L}}{\to} C.$$

Notice that this is equivalent to:

$$\limsup_{n\to\infty} \mathbb{P}_n(\xi_n \in F) \leq \mathbb{P}(C \cap F \neq \emptyset) \quad \forall\, F \in \mathcal{F}(E). \tag{74}$$

With this new definition we can now say that under the assumptions of Theorem 3.11 the sequence (ξ_n) of infimizing points converges in distribution to the random closed set $A(X)$. Of course for an application of the Portmanteau-Theorem the question remains how one can decide whether T_C actually is a probability measure or not. So, let us consider an arbitrary random closed set C whose capacity functional T_C is a probability measure. Then (e.g. by the canonical construction) we find a random variable ξ on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that ξ has distribution T_C, i.e., $\mathbb{P} \circ \xi^{-1} = T_C$. Conclude that $D := \{\xi\}$ is a random closed set with

$$T_D(K) = \mathbb{P}(\{\xi\} \cap K \neq \emptyset) = \mathbb{P}(\xi \in K) = \mathbb{P} \circ \xi^{-1}(K) = T_C(K) \quad \forall\, K \in \mathcal{K}.$$

Thus $T_C = T_D$ on $\mathcal{K}$ and from Choquet's Theorem we can infer that $C \stackrel{\mathcal{L}}{=} D = \{\xi\}$. Inversely, if $C \stackrel{\mathcal{L}}{=} \{\xi\}$ then $T_C = \mathbb{P} \circ \xi^{-1}$. Hence we have shown a first answer to our question:

$$T_C \text{ is a probability measure} \quad \Leftrightarrow \quad C \stackrel{\mathcal{L}}{=} \{\xi\} \text{ for some random variable } \xi. \tag{75}$$

Actually, we can improve this result as follows.

Proposition 3.3. *Let T_C be the capacity functional of a random closed set C defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then the following statements are equivalent:*

(i) T_C is a probability measure;

(ii) $C = \{\xi\}$ $\mathbb{P}$ - a.s. for some random variable $\xi : (\Omega, \mathcal{A}, \mathbb{P}) \to E$.

Proof. Since equality almost surely entails equality in distribution it follows from (75) that we only have to show the direction $(i) \Rightarrow (ii)$. For that purpose first observe that

$$\{\text{card}(C) \neq 1\} = \{\text{card}(C) = 0\} \cup \{\text{card}(C) \geq 2\}, \tag{76}$$

where

$$\mathbb{P}(\text{card}(C) = 0) = \mathbb{P}(C = \emptyset) = \mathbb{P}(C \cap E = \emptyset) = 1 - T_C(E) = T_C(\emptyset) = 0 \tag{77}$$

by using the complement rule for T_C. Next, let us introduce the countable family

$$\mathcal{U} := \{B_\epsilon(x) : 0 < \epsilon \in \mathbb{Q},\ x \in \Delta\}$$

of open balls, where Δ is countable and dense in E, which in fact is separable by assumption of second countability. Then

$$\{\text{card}(C) \geq 2\} = \bigcup_{U,V \in \mathcal{U}, U \cap V = \emptyset} \{C \cap U \neq \emptyset, C \cap V \neq \emptyset\}.$$

Since, e.g., $\{C \cap U \neq \emptyset\} = C^{-1}(\mathcal{H}(U))$ and $\mathcal{H}(U) \in \mathcal{B}_{Fell}$ by Matheron's Theorem we can infer that $\{\text{card}(C) \geq 2\} \in \mathcal{A}$. Moreover,

$$\mathbb{P}(\text{card}(C) \geq 2) \leq \sum_{U,V \in \mathcal{U}, U \cap V = \emptyset} \mathbb{P}(C \cap U \neq \emptyset, C \cap V \neq \emptyset). \tag{78}$$

Here, each summand pertaining to the index (U, V) vanishes, because by the inclusion-exclusion formula we have that

$$\begin{aligned} &\mathbb{P}(C \cap U \neq \emptyset, C \cap V \neq \emptyset) \\ &= \mathbb{P}(C \cap U \neq \emptyset) + \mathbb{P}(C \cap V \neq \emptyset) - \mathbb{P}(\{C \cap U \neq \emptyset\} \cup \{C \cap V \neq \emptyset\} \\ &= T_C(U) + T_C(V) - \mathbb{P}(C \cap (U \cup V) \neq \emptyset) \\ &= T_C(U) + T_C(V) - T_C(U \cup V) \\ &= 0, \end{aligned} \tag{79}$$

where the last equality follows from the additivity of T_C upon noticing that U and V are disjoint. Combining (76)-(79) gives $\mathbb{P}(\text{card}(C) \neq 1) = 0$. Put $\Omega_0 := \{\text{card}(C) = 1\}$. Then $\mathbb{P}(\Omega_0) = 1$ and for every $\omega \in \Omega_0$ there exists exactly one point $\xi(\omega) \in E$ such that $C(\omega) = \{\xi(\omega)\}$. By defining $\xi(\omega) := 0$ for $\omega \notin \Omega_0$ we obtain a map $\xi : \Omega \to E$, which is measurable. Indeed for every Borel-set B of E we have that

$$\{\xi \in B\} = (C^{-1}(\mathcal{H}(B)) \cap \Omega_0) \cup (\Omega_0^c \cap \{0 \in B\}),$$

where Ω_0^c denotes the complement of Ω_0 in Ω. Since C is a random closed set and $\Omega_0 \in \mathcal{A}$ this equality ensures $\{\xi \in B\} \in \mathcal{A}$ and hence measurability of ξ. To sum up we have shown that $C = \{\xi\}$ on Ω_0 with $\mathbb{P}(\Omega_0) = 1$ and measurable map ξ as desired for (ii). □

Now we are in the position to give a sufficient and necessary condition for distributional convergence. Here recall that in the classical definition of weak convergence $P_n \stackrel{w}{\to} P$ the limit P a priori is a probability measure and not only a Choquet-capacity as in our Definition 3.1.

Theorem 3.12. *(convergence in distribution) Let ξ_n be random variables with $\xi_n \in A(X_n)$ a.s. $\forall\, n \geq N$. If*

$$X_n \stackrel{\mathcal{L}}{\to} X \quad \text{in } (D, s)$$

and $(\xi_n)_{n \geq N}$ is stochastically bounded then

$$\mathbb{P} \circ \xi_n^{-1} \stackrel{w}{\to} T_{A(X)}$$

if and only if $A(X) = \{\xi\}$ a.s. for some random variable ξ. In this case:

$$\xi_n \stackrel{\mathcal{L}}{\to} \xi \quad \text{in } \mathbb{R}^d.$$

Proof. If $A(X) = \{\xi\}$ a.s. then $T_{A(X)} = \mathbb{P} \circ \xi^{-1}$ is a probability measure and weak convergence follows from (73) in combination with the Portmanteau-Theorem. Conversely, if weak convergence holds then $T_{A(X)}$ is a probability measure by definition and we may apply Proposition 3.3. □

For convenience let us restate our conclusions by means of the following short diagram:

$$\mathbb{P} \circ \xi_n^{-1} \stackrel{w}{\to} T_{A(X)} \quad \Leftrightarrow \quad A(X) = \{\xi\} \text{ a.s.} \quad \Rightarrow \quad \xi_n \stackrel{\mathcal{L}}{\to} \xi \quad \text{in } \mathbb{R}^d.$$

It illustrates that we do not claim the validity of the reverse conclusion in the second part. To be specific, if ξ_n converges in distribution to some ξ then one can not infer that the limit process X has an a.s. unique infimizing point, which then necessarily would coincide with ξ. We demonstrate this by the following counter-example. Consider the continuous functions $X_n(t) := X(t) := 1_{\{|t|>1\}}(|t| - 1), t \in \mathbb{R}^d$, which can be taken as stochastic processes with constant value, and therefore $X_n \stackrel{\mathcal{L}}{\to} X$ in (D, s). Observe that

$A(X_n)$ and $A(X)$ are equal to the closed ball $\{t \in \mathbb{R}^d : |t| \le 1\}$. If $\xi_n := t_0$ with any fixed point t_0 chosen from the closed ball, then $\xi_n \xrightarrow{\mathcal{L}} \xi := t_0$, but $A(X) \neq \{\xi\}$.

It is possible that for some specifically chosen infimizing points $\xi_n = a(X_n)$ convergence in distribution $\xi_n = a(X_n) \xrightarrow{\mathcal{L}} \xi = a(X)$ holds in the non-unique case ($A(X)$ is not a singleton with positive probability) as long as additional requirements are fulfilled. Theorem 3.2 of Seijo and Sen [23] goes exactly in this direction. They introduce the functionals a=sargmax and a=largmax, which give in a well-defined sense the smallest and largest infimizing point of a function $f \in D(\mathbb{R}^d)$. For X_n and X with trajectories of some special type and so-called associated jump processes Γ_n and Γ it is shown that if (X_n, Γ_n) converges in distribution to (X, Γ) and (sargmax(X_n)) is stochastically bounded, then in fact sargmax$(X_n) \xrightarrow{\mathcal{L}}$ sargmax(X). The same assertion holds for the selection largmax. Moreover, they give an example which demonstrates that weak convergence of X_n alone is not enough to guarantee weak convergence of the pertaining infimizers. More precisely, let a be any measurable selection. Then if (a) $X_n \xrightarrow{\mathcal{L}} X$ in $(D(\mathbb{R}^d), s)$ and (b) $a(X_n) = O_{\mathbb{P}}(1)$, one **cannot** conclude that $a(X_n) \xrightarrow{\mathcal{L}} a(X)$. But Theorem 3.11 says that $a(X_n) \xrightarrow{\mathcal{L}} A(X)$ and we conjecture that under (a) and (b) this is the furthest reaching conclusion one can draw. On the other hand convergence in distribution to $A(X)$ suffices for the construction of asymptotic confidence intervals in statistics. We will show this in the next section by using the following Continuous Mapping Theorem (CMT). Here, $\overline{M}$ denotes the closure of the set M.

Proposition 3.4. *(Extended CMT) Let $\xi_n : (\Omega_n, \mathcal{A}_n, \mathbb{P}_n) \to (E, \mathcal{B}(E)), n \in \mathbb{N}$, be random variables and let $C : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathcal{F}(E), \mathcal{B}_{Fell}(E))$ be a random closed set. Assume that*

$$\xi_n \xrightarrow{\mathcal{L}} C.$$

Consider a further lcscH space U and a measurable mapping $h : E \to U$ such that $T_C(D_h) = 0$ with D_h the set of all discontinuity points of h, which is well-known to lie in $\mathcal{B}(E)$, confer [4]*, p.243. Then*

$$\limsup_{n\to\infty} \mathbb{P}_n(h(\xi_n) \in F) \le T_C \circ h^{-1}(F) \quad \forall\, F \in \mathcal{F}(U), \tag{80}$$

and

$$h(\xi_n) \xrightarrow{\mathcal{L}} \overline{h(C)}. \tag{81}$$

If C is compact $\mathbb{P}$-a.s. then $h(C)$ is compact a.s. as well and

$$h(\xi_n) \xrightarrow{\mathcal{L}} h(C). \tag{82}$$

Proof. First observe that $\overline{h^{-1}(F)} \subseteq h^{-1}(F) \cup D_h$ for each closed F in U. Since T_C is monotone and sub-additive it therefore follows that

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}_n(h(\xi_n) \in F) &\le \limsup_{n\to\infty} \mathbb{P}_n(\xi_n \in h^{-1}(F)) \le \limsup_{n\to\infty} \mathbb{P}_n(\xi_n \in \overline{h^{-1}(F)}\ ) \\ &\le T_C(\overline{h^{-1}(F)}) \le T_C(h^{-1}(F) + T_C(D_h) = T_C \circ h^{-1}(F).\end{aligned}$$

Since $\{C \cap h^{-1}(F) \neq \emptyset\} = \{h(C) \cap F \neq \emptyset\} \subseteq \{\overline{h(C)} \cap F \neq \emptyset\}$ we can infer that

$$T_C \circ h^{-1}(F) = T_C(h^{-1}(F)) = \mathbb{P}(C \cap h^{-1}(F) \neq \emptyset) \le \mathbb{P}(\overline{h(C)} \cap F \neq \emptyset) = T_{\overline{h(C)}}(F),$$

and (81) follows from (80). For the proof of (82) notice that by assumption $\mathbb{P}(C \cap D_h = \emptyset) = 1$. So, if $C_h := E \setminus D_h$ is the set of all continuity points of h, then $\mathbb{P}(C \subseteq C_h) = 1$, because $\{C \cap D_h = \emptyset\} = \{C \subseteq C_h\}$. Consequently, by assumption on C the event $\Omega_0 := \{C \subseteq C_h\} \cap \{C \in \mathcal{K}(E)\}$ has probability one. Moreover, on Ω_0 our map h is continuous on the compact set C. Recall from Analysis that the continuous image of a compact set is compact, whence $\Omega_0 \subseteq \{h(C) \in \mathcal{K}(U)\}$ and thus $h(C)$ is compact in U a.s. Let $\mathcal{A}_0 := \mathcal{A} \cap \Omega_0 = \{A \in \mathcal{A} : A \subseteq \Omega_0\}$ and $\mathbb{P}_0$ be the restriction of $\mathbb{P}$ on $\mathcal{A}_0$. Since the restriction $C : (\Omega_0, \mathcal{A}_0, \mathbb{P}_0) \to (\mathcal{F}(E), \mathcal{B}_{Fell}(E))$ is measurable, $h(C) : (\Omega_0, \mathcal{A}_0, \mathbb{P}_0) \to (\mathcal{F}(U), \mathcal{B}_{Fell}(U))$ is a random closed set as well. This follows from (71) and $\{h(C) \cap K = \emptyset\} = \{C \cap h^{-1}(K) = \emptyset\} = C^{-1}(\mathcal{M}(h^{-1}(K))) \in \mathcal{A}_0$ upon noticing the second part of Matheron's Theorem and $h^{-1}(K) \in \mathcal{B}(E)$ for all $K \in \mathcal{K}(U)$ by measurability of h.

Finally, $T_C \circ h^{-1}(F) = \mathbb{P}_0(C \cap h^{-1}(F) \neq \emptyset) = \mathbb{P}_0(h(C) \cap F \neq \emptyset)$ and (82) follows from (80) and the defining relation (74). ☐

Notice that $\xi_n \xrightarrow{\mathcal{L}} \{\xi\}$ is equivalent to $\xi_n \xrightarrow{\mathcal{L}} \xi$ so that the above proposition is a generalization of the classical CMT, confer, e.g., Theorem 1.3 in chapter 19 of [24].

As omnipresent requirement in our last theorems we have to show distributional convergence $X_n \xrightarrow{\mathcal{L}} X$ in $(D(\mathbb{R}^d), s)$. Therefore we end this section with a necessary and sufficient condition given by [12]. It reduces the problem to the Skorokhod space $(D([-a, a]), s)$ of [16] with compact rectangle $[-a, a] := [-a_1, a_1] \times \ldots \times [-a_d, a_d]$, $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$ with $a_i > 0$ for all $1 \leq i \leq d$ $(a > 0)$. Observe that for each cadlag stochastic process $X = \{X(t) : t \in \mathbb{R}^d\}$ the pertaining restriction $X^{(a)} := \{X(t) : t \in [-a, a]\}$ can be considered as random element in $(D([-a, a]), s)$.

Theorem 3.13. *(Lagodowski and Rychlik)* $X_n \xrightarrow{\mathcal{L}} X$ *in* $(D(\mathbb{R}^d), s)$ *if and only if*

$$X_n^{(a)} \xrightarrow{\mathcal{L}} X^{(a)} \quad \text{in } (D([-a, a]), s) \quad \forall\, 0 < a = (a_1, \ldots, a_d) \in T \tag{83}$$

where $T := \{t \in \mathbb{R}^d : \pi_t \text{ continuous } \mathbb{P} \circ X^{-1} \text{ a.e.}\}$.

This result has counterparts for continuous stochastic processes, confer Proposition 14.6 in [11], and for those whose trajectories are locally bounded, confer Theorem 1.6.1 in [27].

The crucial point is that for the proof of (83) there exist manageable sufficient criteria (e.g. moment-criteria), confer [3], [4] or [18] in case $d = 1$ and [16], [2], [25] or [26] for $d \geq 1$.

4. Applications to M-estimators

Recall the M-estimator of Example 1.1 in the introduction:

$$\hat{\theta}_n = (\hat{\theta}_{n,1}, \ldots, \hat{\theta}_{n,d}) = a(M_n) = \operatorname{arginf}(M_n) \tag{84}$$

with criterion function $M_n \in D(\mathbb{R}^d)$ and its theoretical counterpart $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$, which as parameter of interest is to be estimated.

Let Γ_n be a $(d \times d)$ diagonal matrix with positive entries $\alpha_{n,1}, \ldots, \alpha_{n,d}$ on the diagonal. Define

$$X_n(t) = \gamma_n\{M_n(\theta + \Gamma_n^{-1} t) - M_n(\theta)\},\ t \in \mathbb{R}^d,$$

with some positive sequence $(\gamma_n) \subseteq \mathbb{R}$. In Lemma 2.2 put $f = M_n, a = \gamma_n, b = M_n(\theta)$ and $\lambda(t) = \Gamma_n(t - \theta)$. Since $\lambda^{-1}(t) = \theta + \Gamma_n^{-1} t$, we first observe that $g = X_n$. It follows from Lemma 2.2 (i) and (iii) that $\xi_n := \Gamma_n(\hat{\theta}_n - \theta)$ is a random variable with $\xi_n \in A(X_n)$, where X_n is cadlag. Now, assume that we are able to prove that

$$X_n \xrightarrow{\mathcal{L}} X \quad \text{in } (D, s) \tag{85}$$

and that (ξ_n) is stochastically bounded, i.e.,

$$\xi_n = \Gamma_n(\hat{\theta}_n - \theta) = O_{\mathbb{P}}(1). \tag{86}$$

Then Theorem 3.11 yields that

$$\xi_n \xrightarrow{\mathcal{L}} A(X) =: C.$$

Recall that $|\cdot|$ denotes the maximum-norm on $\mathbb{R}^d$, which is continuous on its entire domain. If $C = A(X)$ is compact a.s. then our extended CMT stated in Proposition 3.4 ensures that

$$|\xi_n| \xrightarrow{\mathcal{L}} |C|. \tag{87}$$

This limit theorem (87) can be used for the construction of an asymptotic confidence region as follows. For every $r > 0$ consider the random open intervals

$$I_n(r) := (\hat{\theta}_{n,1} - \frac{r}{\alpha_{n,1}}, \hat{\theta}_{n,1} + \frac{r}{\alpha_{n,1}}) \times \cdots \times (\hat{\theta}_{n,d} - \frac{r}{\alpha_{n,d}}, \hat{\theta}_{n,d} + \frac{r}{\alpha_{n,d}}).$$

We observe that

$$\begin{aligned}
&\liminf_{n\to\infty} \mathbb{P}(\theta \in I_n(r)) \\
&= \liminf_{n\to\infty} \mathbb{P}(|\xi_n| < r) && \text{by definitions} \\
&= 1 - \limsup_{n\to\infty} \mathbb{P}(|\xi_n| \geq r) \\
&= 1 - \limsup_{n\to\infty} \mathbb{P}(|\xi_n| \in [r, \infty)) \\
&\geq 1 - \mathbb{P}(|C| \cap [r, \infty) \neq \emptyset) && \text{by (87), (74), } [r,\infty) \in \mathcal{F}(\mathbb{R}) \\
&= \mathbb{P}(|C| \cap [r, \infty) = \emptyset) \\
&= \mathbb{P}(|C| \subseteq (-\infty, r)) \\
&= \mathbb{P}(\sup_{x \in C} |x| < r) && \text{since } C \text{ is compact a.s.} \\
&= F(r-),
\end{aligned}$$

where F denotes the distribution function of the real random variable $\sigma := \sup_{x\in C} |x|$. Thus we can conclude as follows: If $q_\alpha := F^{-1}(1-\alpha)$ is the $(1-\alpha)$-quantile of F, then $I_n(r)$ is an asymptotic confidence interval for θ at level $1-\alpha$ whenever $r > q_\alpha$, that is

$$\liminf_{n\to\infty} \mathbb{P}(\theta \in I_n(r)) > 1 - \alpha \quad \forall\, r > q_\alpha,$$

and

$$\liminf_{n\to\infty} \mathbb{P}(\theta \in I_n(q_\alpha)) \geq 1 - \alpha \quad \text{if } q_\alpha \text{ is a continuity-point of } F.$$

In general q_α is an unknown quantity for the statistician, but it can be approximated by a Monte-Carlo method. For that purpose first notice that the distribution Q of the limit process X in (85) usually depends on the unknown parameter θ, that is $Q = Q_\theta$. Therefore we generate $m \in \mathbb{N}$ processes $X^{(1)}, \ldots, X^{(m)}$ i.i.d. with common distribution $Q_{\hat{\theta}_n}$. For each $1 \leq i \leq m$ compute $C^{(i)} := A(X^{(i)})$ and $\sigma^{(i)} := \sup_{x\in C^{(i)}} |x|$. Let F_m be the empirical distribution function pertaining to $\sigma^{(1)}, \ldots, \sigma^{(m)}$. Then $q_{m,\alpha} := F_m^{-1}(1-\alpha)$ is a reasonable estimate for q_α.

A more general construction goes as follows. Let $\emptyset \neq \phi_n \subseteq A(M_n)$. Then

$$\varphi_n := \Gamma_n(\phi_n - \theta) = \lambda(\phi_n) \subseteq A(X_n),$$

where the inclusion holds by Lemma 2.2 (iii). Now, if (φ_n) is stochastically bounded, then (85) and Theorem 3.4 ensure that

$$\Gamma_n(\phi_n - \theta) \xrightarrow{q-\mathcal{L}} A(X).$$

Choose an open $G \subseteq \mathbb{R}^d$ with $\mathbb{P}(A(X) \subseteq G) \geq 1-\alpha$. Observe that $\{\Gamma_n(\phi_n - \theta) \subseteq G\} \subseteq \{\theta \in \phi_n - \Gamma_n^{-1}G\}$. So, if $C_n := \phi_n - \Gamma_n^{-1}G$, then

$$\liminf_{n\to\infty} \mathbb{P}(\theta \in C_n) \geq \liminf_{n\to\infty} \mathbb{P}(\Gamma_n(\phi_n - \theta) \subseteq G) \geq 1 - \alpha,$$

where the inequality follows from Theorem 3.7. Thus C_n is an (asymptotic) confidence region for θ at level $1-\alpha$. In any case C_n becomes smaller (with respect to the relation $\subseteq$) for the special choice $\phi_n = \{\hat{\theta}_n\}$. The pertaining confidence region $\hat{C}_n := \hat{\theta}_n - \Gamma_n^{-1}G$ has (d-dimensional) volume $\lambda_d(\hat{C}_n) = \{\prod_{i=1}^d \alpha_{n,i}\}^{-1}\lambda_d(G)$ by Proposition 6.1.2 in [5]. Consequently the optimal confidence region is given by $\hat{C}_{n,opt} = \hat{\theta}_n - \Gamma_n^{-1}G_{opt}$ with

$$G_{opt} = \operatorname{argmin}\{\lambda_d(G) : G \in \mathcal{G}, \mathbb{P}(A(X) \subseteq G) \geq 1-\alpha)\}.$$

The analytical solution of this minimization problem is -if at all- hard to find. The special choice $G := (-r, r)^d$ yields $\hat{C}_n = I_n(r)$ from above.

References

1. H. Attouch, *Variational Convergence for Functions and Operators*, Pitman Advanced Publishing Program, Boston, London, Melbourne, 1984.
2. P. J. Bickel and M. J. Wichura, *Convergence criteria for multiparameter stochastic processes and some applications*, Ann. Math. Statist. **42** (1971), no. 5, 1656–1670.
3. P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.
4. P. Billingsley, *Convergence of Probability Measures, Second Edition*, John Wiley & Sons, New York, 1999.
5. D. L. Cohn, *Measure Theory*, Birkhäuser, Boston, Basel, Berlin, 1997.
6. G. Dal Maso, *An Introduction to* Γ*-Convergence*, Birkhäuser, Boston, Basel, Berlin, 1993.
7. J. H. Dshalalow, *Real Analysis*, Chapman & Hall/CRC, Boca Raton, London, New York, Washington D.C., 2001.
8. D. Ferger, *On the argmin-sets of stochastic processes and their distributional convergence in Fell-type topologies*, Kybernetika **47** (2011), no. 6, 955–968.
9. O. Gersch, *Convergence in Distribution of Random Closed Sets and Applications in Stability*, Ph.D Thesis, Technische Universität Ilmenau, 2007.
10. P. Gänssler and W. Stute, *Wahrscheinlichkeitstheorie*, Springer-Verlag, Berlin, Heidelberg, 1977.
11. O. Kallenberg, *Foundations of Modern Probability*, Springer-Verlag, Berlin, Heidelberg, 1997.
12. A. Lagodowski and Z. Rychlik, *Weak convergence of probability meaures on the function space* $D_d[0, \infty)$, Bull. Polish Acad. Sci. Math. **34** (1986), no. 5-6, 329–335.
13. T. Norberg, *Convergence and existence of random set distributions*, Ann. Probab. **12** (1984), no. 3, 726–732.
14. G. Matheron, *Random Sets and Integral Geometry*, John Wiley & Sons, New York, London, Sydney, 1975.
15. I. Molchanov, *Theory of Random Sets*, Springer-Verlag, London, 2005.
16. G. Neuhaus, *On weak convergence of stochastic processes with multidimensional time parameter*, Ann. Math. Statist. **42** (1971), no. 4, 1285–1295.
17. H. T. Nguyen, *An Introduction to Random Sets*, Chapman & Hall/CRC, Boca Raton, London, New York, 2006.
18. K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.
19. G. Ch. Pflug, *Asymptotic dominance and confidence for solutions of stochastic programs*, Czechoslovak J. Oper. Res. **1** (1992), no. 1, 21–30.
20. D. Pollard, *Convergence of Stochastic Processes*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.
21. R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, Heidelberg, New York, 1998.
22. G. Salinetti and R. J-B. Wets, *On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima*, Math. Oper. Res. **11** (1986), no. 3, 385–419.
23. E. Seijo and B. A. Sen, *A continuous mapping theorem for the smallest argmax functional*, Elektron. J. Stat. **5** (2011), 421–439.
24. G. Shorack, *Probability for Statisticians*, Springer-Verlag, New York, 2000.
25. M. L. Straf, *A general Skorokhod space and its applications to the weak convergence of stochastic processes with several parameters*, Ph.D. dissertation, University of Chicago, 1969.
26. M. L. Straf, *Weak convergence of stochastic processes with several parameters* Proc. Sixth Berkley Symp. Math. Statist. Prob. (1970), 187–221.
27. A. W. van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes*, Springer-Verlag, New York, Berlin, Heidelberg, 1996.
28. S. Vogel, *Qualitative stability of stochastic programs with applications in asymptotic statistics*, Statist. Decisions **23** (2005), 219–248.
29. S. Vogel, *Semiconvergence in distribution of random closed sets with applications to random optimization problems*, Ann. Oper. Res. **142** (2006), 269–282.

Technische Universität Dresden, Department of Mathematics, Zellescher Weg 12-14, D-01069 Dresden, Germany

E-mail address: `dietmar.ferger@tu-dresden.de`