The Annals of Statistics
2014, Vol. 42, No. 4, 1598–1634
DOI: 10.1214/14-AOS1237

WHEN UNIFORM WEAK CONVERGENCE FAILS: EMPIRICAL PROCESSES FOR DEPENDENCE FUNCTIONS AND RESIDUALS VIA EPI- AND HYPOGRAPHS[1]

BY AXEL BÜCHER, JOHAN SEGERS AND STANISLAV VOLGUSHEV

Ruhr-Universität Bochum, Université catholique de Louvain and Ruhr-Universität Bochum

In the past decades, weak convergence theory for stochastic processes has become a standard tool for analyzing the asymptotic properties of various statistics. Routinely, weak convergence is considered in the space of bounded functions equipped with the supremum metric. However, there are cases when weak convergence in those spaces fails to hold. Examples include empirical copula and tail dependence processes and residual empirical processes in linear regression models in case the underlying distributions lack a certain degree of smoothness. To resolve the issue, a new metric for locally bounded functions is introduced and the corresponding weak convergence theory is developed. Convergence with respect to the new metric is related to epi- and hypo-convergence and is weaker than uniform convergence. Still, for continuous limits, it is equivalent to locally uniform convergence, whereas under mild side conditions, it implies L^p convergence. For the examples mentioned above, weak convergence with respect to the new metric is established in situations where it does not occur with respect to the supremum distance. The results are applied to obtain asymptotic properties of resampling procedures and goodness-of-fit tests.

1. Introduction. The Hoffman–Jørgensen weak convergence theory in the space of bounded functions is a great success story in mathematical statistics [Kosorok (2008), van der Vaart and Wellner (1996)]. Measurability assumptions are reduced to a minimum, no smoothness assumptions on the trajectories are needed, it applies in a vast variety of circumstances, and the topology of uniform convergence is fine enough so that, through the continuous mapping theorem and functional delta method, it implies weak convergence of a countless list of interesting statistical functionals.

But precisely because of the strength of uniform convergence, there are circumstances where it does not hold. Weak convergence can fail when the (pointwise)

Received May 2013; revised February 2014.

[1]Supported by the collaborative research center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG), by IAP research network Grant No. P7/06 of the Belgian government (Belgian Science Policy) and by contract "Projet d'Actions de Recherche Concertées" No. 12/17-045 of the "Communauté française de Belgique."

MSC2010 subject classifications. Primary 60F05, 62G30; secondary 62G32, 62M09.

Key words and phrases. Bootstrap, copula, epigraph, hypograph, linear regression, local alternative, power curve, residual empirical process, stable tail dependence function, weak convergence.

candidate limit process has discontinuous trajectories. Think of empirical distributions based on residuals of some sort, that is, observations that are themselves approximations of some latent random variables. Because of the measurement error in the ordinates, jump locations fail to be located exactly, and uniform convergence fails. The examples which interest us in this paper concern empirical copula processes, the empirical process based on residuals in a linear regression setting and empirical tail dependence function processes.

A radical solution to the lack-of-convergence issue is to seek for another metric on an appropriate function space. The metric should be weak enough so that convergence does take place, but still strong enough to enable statistical applications. Ideally, when the limit process has continuous trajectories, it should be equivalent to uniform convergence, so that in standard situations, nothing is lost. This is a difficult task, and it turns out that the various extensions of Skorohod's metrics [Skorohod (1956)] to functions of several variables [Bass and Pyke (1985), Bickel and Wichura (1971), Neuhaus (1971), Straf (1972)] are not suitable for the examples that we consider.

In the present paper, we construct such a metric by building on ideas that originate in variational analysis and optimization theory. In the context of minimization problems, one identifies a real function f on a suitable metric space $\mathbb{T}$ with its *epigraph*, which is the set of all points (x, y) in $\mathbb{T} \times \mathbb{R}$ such that $f(x) \le y$. *Epi-convergence* of functions is then defined as Painlevé–Kuratowski convergence of their epigraphs [Beer (1993), Molchanov (2005), Rockafellar and Wets (1998)]. For maximization problems, *hypographs* and *hypo-convergence* are defined in the same way, the inequality sign pointing in the other direction.

Combining these modes of convergence, we will say that f_n *hypi-converges* to f if the epigraphs of f_n converge to the closure of the epigraph of f and the hypographs of f_n converge to the closure of the hypograph of f. This mode of convergence is to be distinguished from epi/hypo-convergence, a concept arising in connection with saddle points [Attouch and Wets (1983)].

Broadly speaking, hypi-convergence is intermediate between uniform convergence and L^p convergence. Hypi-convergence implies uniform convergence on compact subsets of the domain that are contained in the set of continuity points of the limit function. Hence, for continuous limits, we are back to uniform convergence. But even without continuity, hypi-convergence implies convergence of global extrema. Moreover, for limit functions which are continuous almost everywhere, hypi-convergence implies L^p convergence on compact sets.

In a similar way as one can consider weak epi-convergence of random lower semicontinuous functions [Geyer (1994), Molchanov (2005)], we develop Hoffman–Jørgensen weak convergence theory with respect to the hypi-(semi)metric. Thanks to an extension of the continuous mapping theorem for semimetric spaces, we are able to leverage the above properties of hypi-convergence to yield weak convergence of finite-dimensional distributions, Kolmogorov–Smirnov type statistics, and procedures related to L^p spaces, notably Cramér–von Mises statistics. An extension of the functional delta method is also discussed.

We investigate weak convergence with respect to the hypi-semimetric of empirical copula processes, empirical tail dependence functions, and empirical processes based on regression residuals. Weak hypi-convergence of the empirical copula process is established for copulas whose partial derivatives exist and are continuous everywhere except for an arbitrary Lebesgue null set of the unit cube, which extends results from Rüschendorf (1976) and others (see Section 4 for more references on the empirical copula process). From there, we show validity of the bootstrap [see Fermanian, Radulović and Wegkamp (2004)] and extend results on power curves for goodness-of-fit tests under local alternatives [Genest, Quessy and Remillard (2007)]. Similar results are shown for tail dependence functions, extending Bücher and Dette (2013) and Einmahl, Krajina and Segers (2012). Classical results on the empirical distribution function of regression residuals [Koul (1969), Loynes (1980)] are extended to the case where the true distribution has a discontinuous density.

The structure of the paper is as follows. The hypi-topology is introduced in Section 2. Weak convergence in hypi-space is the topic of Section 3. We provide tools for checking weak hypi-convergence and for exploiting it in a statistical context. The new framework is applied for empirical copula processes in Section 4, for empirical tail dependence function processes in Section 5, and for the empirical process of regression residuals in Section 6. These three sections can be read independently of one another. In order to preserve the flow of the text, a number of auxiliary results and all proofs are deferred to a sequence of Appendices and an online supplement [Bücher, Segers and Volgushev (2014)]. The weak convergence theory for semimetric spaces in Appendix B, including a version of the extended continuous mapping theorem and the functional delta method, is perhaps of independent interest.

2. Hypi-convergence of locally bounded functions.

We introduce a mode of convergence for real-valued, locally bounded functions on a locally compact, separable metric space (Section 2.1). For continuous limits, the metric is equivalent to locally uniform convergence, but for discontinuous limits, it is strictly weaker, while still implying L^p convergence (Section 2.2). The proofs for the results in this section are given in Appendix F.1.

2.1. *The hypi-semimetric.*

Let $(\mathbb{T}, d)$ be a locally compact, separable metric space. The space $\mathbb{T} \times \mathbb{R}$ is a locally compact, separable metric space, too, when equipped, for instance, with the metric $d_{\mathbb{T}\times\mathbb{R}}((x_1, y_1), (x_2, y_2)) = \max\{d(x_1, x_2), |y_1 - y_2|\}$.

Let $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ denote the space of locally bounded functions $f : \mathbb{T} \to \mathbb{R}$, that is, functions that are uniformly bounded on compacta. If $\mathbb{T}$ is itself compact, we will simply write $\ell^\infty(\mathbb{T})$. Functions $f \in \ell^\infty_{\mathrm{loc}}(\mathbb{T})$ will be identified with subsets of $\mathbb{T} \times \mathbb{R}$

by considering their epigraphs and hypographs:

$$\operatorname{epi} f = \{(x, y) \in \mathbb{T} \times \mathbb{R} : f(x) \le y\},$$
$$\operatorname{hypo} f = \{(x, y) \in \mathbb{T} \times \mathbb{R} : y \le f(x)\}.$$

Except for being locally bounded, functions f in $\ell^\infty_{\text{loc}}(\mathbb{T})$ can be arbitrarily rough. A minimal amount of regularity will come from the lower and upper semicontinuous hulls $f_\wedge \le f \le f_\vee$:

$$f_\wedge(x) = \sup_{\varepsilon>0} \inf\{f(x') : d(x', x) < \varepsilon\}, \tag{2.1}$$

$$f_\vee(x) = \inf_{\varepsilon>0} \sup\{f(x') : d(x', x) < \varepsilon\}, \tag{2.2}$$

functions which are elements of $\ell^\infty_{\text{loc}}(\mathbb{T})$, too. Note that $(-f)_\wedge = -f_\vee$. A convenient link between epi- and hypographs on the one hand and lower and upper semicontinuous hulls on the other hand is that

$$\operatorname{cl}(\operatorname{epi} f) = \operatorname{epi} f_\wedge, \qquad \operatorname{cl}(\operatorname{hypo} f) = \operatorname{hypo} f_\vee,$$

where “cl” denotes topological closure, in this case, in the space $\mathbb{T} \times \mathbb{R}$. In particular, a function f is lower (upper) semicontinuous if and only if its epigraph (hypograph) is closed.

Functions being identified with sets, notions of set convergence can be applied to define convergence of functions. We rely on classical theory exposed in, among others, Matheron (1975), Beer (1993), Rockafellar and Wets (1998) and Molchanov (2005). A standard topology on the space of closed subsets of a topological space is the Fell hit-and-miss topology. If the underlying space is locally compact and separable, as in our case, then the Fell topology is metrizable. Moreover, in that case, convergence of a sequence of closed sets in the Fell topology is equivalent to their Painlevé–Kuratowski convergence. Recall that (not necessarily closed) sets A_n of a topological space converge to a set A in the Painlevé–Kuratowski sense if and only if (i) for every $x \in A$ there exists a sequence x_n with $x_n \in A_n$ such that $x_n \to x$ and (ii) whenever $x_{n_k} \in A_{n_k}$ for some subsequence n_k converges to a limit x, we must have $x \in A$. The limit set A is necessarily closed, and Painlevé–Kuratowski convergence of A_n to A is equivalent to Painlevé–Kuratowski convergence of $\operatorname{cl}(A_n)$ to A.

Let $\mathcal{F}(\mathbb{T} \times \mathbb{R})$ be the space of closed subsets of $\mathbb{T} \times \mathbb{R}$ and let $d_{\mathcal{F}}$ be a metric inducing the Fell topology, or equivalently, Painlevé–Kuratowski convergence. Examples of metrics $d_{\mathcal{F}}$ for the Fell topology are to be found in Rockafellar and Wets (1998), Molchanov (2005) and Ogura (2007). A versatile notion of convergence of functions in optimization theory is epi-convergence: a sequence of functions $f_n : \mathbb{T} \to \mathbb{R}$ is said to epi-converge to a function f if and only if the Painlevé–Kuratowski limit of $\operatorname{epi} f_n$ (or equivalently, its closure) is equal to $\operatorname{epi} f$, that is, if $d_{\mathcal{F}}(\operatorname{cl}(\operatorname{epi} f_n), \operatorname{cl}(\operatorname{epi} f)) \to 0$ as $n \to \infty$. Necessarily, the limit set $\operatorname{epi} f$ is closed and, therefore, f must be lower semicontinuous. Similarly, hypo-convergence of functions is defined as Painlevé–Kuratowski convergence of their hypographs (or their closures), and the limit function is necessarily upper semicontinuous.

If f_n both epi-converges to $f_\wedge$ and hypo-converges to $f_\vee$, then we say that f_n *hypi-converges* to f. This mode of convergence is the one that we propose in this paper. According to the following result, hypi-convergence is metrizable and can be checked conveniently by pointwise criteria.

PROPOSITION 2.1 (Hypi-convergence). *Let $f_n, f \in \ell^\infty_{\mathrm{loc}}(\mathbb{T})$. The following statements are equivalent*:

(i) *f_n epi-converges to $f_\wedge$ and hypo-converges to $f_\vee$.*

(ii) *The following pointwise criteria hold*:

$$\begin{cases} \forall x \in \mathbb{T} : \forall x_n \to x : f_\wedge(x) \le \liminf_{n\to\infty} f_n(x_n), \\ \forall x \in \mathbb{T} : \exists x_n \to x : \limsup_{n\to\infty} f_n(x_n) \le f_\wedge(x) \end{cases} \tag{2.3}$$

and

$$\begin{cases} \forall x \in \mathbb{T} : \forall x_n \to x : \limsup_{n\to\infty} f_n(x_n) \le f_\vee(x), \\ \forall x \in \mathbb{T} : \exists x_n \to x : f_\vee(x) \le \liminf_{n\to\infty} f_n(x_n). \end{cases} \tag{2.4}$$

(iii) *The distance $d_{\mathrm{hypi}}(f_n, f)$ converges to 0, where d_{hypi} denotes the* hypi-semimetric *defined as*

$$d_{\mathrm{hypi}}(f, g) = \max\{d_{\mathcal{F}}(\operatorname{epi} f_\wedge, \operatorname{epi} g_\wedge), d_{\mathcal{F}}(\operatorname{hypo} f_\vee, \operatorname{hypo} g_\vee)\},$$

and $d_{\mathcal{F}}$ is a metric on $\mathcal{F}(\mathbb{T} \times \mathbb{R})$ inducing the Fell topology.

(iv) *f_n converges to f in the* hypi-topology, *which is defined as the coarsest topology on $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ for which the map*

$$\ell^\infty_{\mathrm{loc}}(\mathbb{T}) \to \mathcal{F}(\mathbb{T} \times \mathbb{R}) \times \mathcal{F}(\mathbb{T} \times \mathbb{R}) : f \mapsto (\operatorname{cl}(\operatorname{epi} f), \operatorname{cl}(\operatorname{hypo} f))$$

is continuous, *that is*, *the hypi-open sets in $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ are the inverse images of open sets in $\mathcal{F}(\mathbb{T} \times \mathbb{R}) \times \mathcal{F}(\mathbb{T} \times \mathbb{R})$.*

Note that in (2.3) and (2.4), we can replace f_n by $f_{n,\wedge}$ and $f_{n,\vee}$, respectively (Lemma A.1). The equivalence of (i) and (ii) follows from well-known pointwise criteria for epi- and hypo-convergence [Molchanov (2005), Chapter 5, Proposition 3.2(ii)]. Statements (iii) and (iv) are just reformulations of what it means to have both epi- and hypo-convergence in (i).

Intuitively, two functions are close in the hypi-semimetric if both their epigraphs and their hypographs are close. Two functions on $\mathbb{T} = [0, 1]$ whose epigraphs are close but whose hypographs are far away are depicted in the upper part of Figure 1: for instance, the point (0.5, 1) belongs to the hypograph of the dotted-line

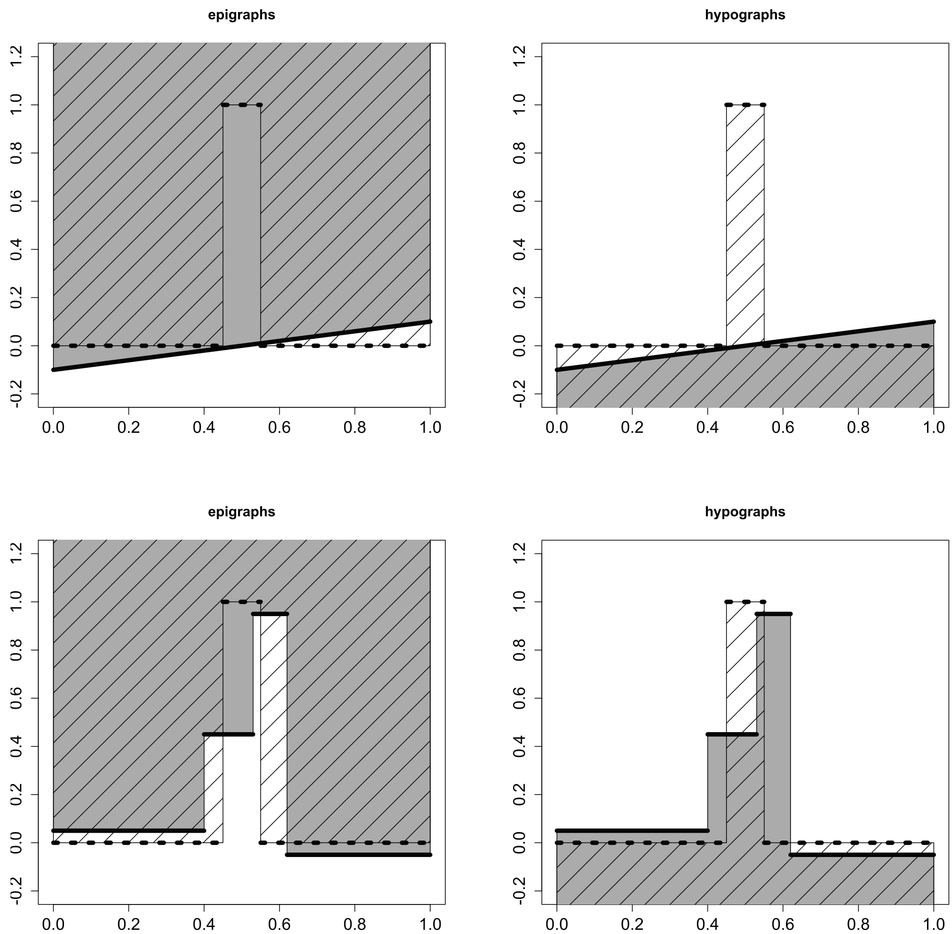

FIG. 1. Top: *two functions whose epigraphs are close but whose hypographs are far away.* Bottom: *two functions of which both the epi- and hypographs are close. The light gray areas represent epigraphs* (*left column*) *and hypographs* (*right column*) *of the functions depicted by solid lines*, *whereas shaded areas represent epigraphs* (*left column*) *and hypographs* (*right column*) *of the functions depicted by dotted lines*.

function but is far away from any point in the hypograph of the solid-line function. As a consequence, these two functions are not close in the hypi-semimetric. For comparison, two functions that are close in the hypi-semimetric are depicted in the lower part of Figure 1.

By Proposition 2.1(ii), if f_n hypi-converges to f and if f is continuous at x, then $f_n(x_n) \to f(x)$ whenever $x_n \to x$. Moreover, it follows that locally uniform convergence of locally bounded functions implies their hypi-convergence. The converse is not true if the hypi-limit is not continuous.

Hypi-convergence of sequences f_n and g_n to f and g, respectively, does in general not imply hypi-convergence of the sequence of sums, $f_n + g_n$, to the sum of the limits, $f + g$. For instance, let x_n converge to x in $\mathbb{T}$ with $x_n \neq x$ and set $f_n = \mathbb{1}_{\{x_n\}}$ and $g_n = -\mathbb{1}_{\{x\}}$. Still, a sufficient condition is that at least one of the limit functions is continuous; see Lemma A.4 for an even more general result.

An alternative view on the hypi-topology can be gained by identifying $f \in \ell^\infty_{\mathrm{loc}}(\mathbb{T})$ with its completed graph $\Gamma(f) = \mathrm{epi}(f_\wedge) \cap \mathrm{hypo}(f_\vee)$ [Vervaat (1981)]. We suspect that for certain domains, hypi-convergence is equivalent to set convergence of completed graphs. For càdlàg functions on $\mathbb{T} = [0, 1]$, the latter convergence can be seen to be equivalent to Skorohod M_2-convergence [Molchanov (2005), page 377], whence hypi-convergence can be regarded as a coordinate-free extension of Skorohod M_2-convergence to nonsmooth functions on rather general domains.

2.2. *Leveraging hypi-convergence.* As mentioned already, uniform convergence implies hypi-convergence but not conversely. Nevertheless, at subsets of the domain where the limit function is continuous, the converse does hold. In this sense, working in hypi-space does not necessarily yield weaker results than in the uniform topology. All proofs for this section are given in Appendix F.1 in the supplement [Bücher, Segers and Volgushev (2014)].

PROPOSITION 2.2. *Let $K \subset \mathbb{T}$ be compact and let $f \in \ell^\infty_{\mathrm{loc}}(\mathbb{T})$ be continuous at every $x \in K$. If f_n hypi-converges to f in $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$, then $\sup_{x\in K} |f_n(x) - f(x)| \to 0$ as $n \to \infty$.*

Being a combination of epi- and hypo-convergence, hypi-convergence preserves convergence of extrema. Later, we will make use of this property when investigating Kolmogorov–Smirnov type test statistics (Section 4.3).

PROPOSITION 2.3. *Let $G \subset \mathbb{T}$ be an open subset with compact closure. If f_n hypi-converges to f in $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ and if f is continuous on the boundary of G, then $\inf f_n(G) \to \inf f(G)$ and $\sup f_n(G) \to \sup f(G)$ as $n \to \infty$. If $G = \mathbb{T}$ is compact, then the boundary of G in $\mathbb{T}$ is empty, and hence the conclusions hold true without imposing any conditions on f.*

Hypi-convergence implies L^p-convergence for finite p, provided that the limit function is not too rough. This is useful, for instance, when studying Cramér–von Mises statistics (Section 4.3) and other statistical procedures based on the L^2-distance, such as minimum distance estimators. Recall that upper and lower semicontinuous functions are necessarily Borel measurable.

PROPOSITION 2.4. *Let μ be a finite Borel measure supported on a compact subset of $\mathbb{T}$. If f_n hypi-converges to f in $\ell^\infty(\mathbb{T})$ and if f is continuous μ-almost* *everywhere, then, for every* $p \in [1, \infty)$, *we have* $\int |f_{n,\vee} - f_{n,\wedge}|^p \, d\mu \to 0$ *and* $\int |f_n^* - f^*|^p \, d\mu \to 0$ *as* $n \to \infty$, *where* f_n^* *and* f^* *represent arbitrary Borel measurable functions on* $\mathbb{T}$ *such that* $f_{n,\wedge} \le f_n^* \le f_{n,\vee}$ *and* $f_\wedge \le f^* \le f_\vee$.

3. Weak hypi-convergence of stochastic processes. When applying Hoffman–Jørgensen weak convergence theory, it is customary to work in a metric space. However, d_{hypi} is a semimetric and not a metric: if functions $f, g \in \ell_{\text{loc}}^\infty(\mathbb{T})$ share the same lower and upper semicontinuous hulls, then $d_{\text{hypi}}(f, g) = 0$ even if f and g are different functions.

To obtain a metric space, we consider equivalences classes of functions at hypi-distance zero. For $f \in \ell_{\text{loc}}^\infty(\mathbb{T})$, let $[f]$ be the set of all $g \in \ell_{\text{loc}}^\infty(\mathbb{T})$ such that $d_{\text{hypi}}(f, g) = 0$. Let $L_{\text{loc}}^\infty(\mathbb{T})$ be the space of all such equivalence classes. Then $L_{\text{loc}}^\infty(\mathbb{T})$ becomes a metric space when equipped with the hypi-metric (abusing notation) $d_{\text{hypi}}([f], [g]) := d_{\text{hypi}}(f, g)$. The map $[\cdot]$ from $\ell_{\text{loc}}^\infty(\mathbb{T})$ into $L_{\text{loc}}^\infty(\mathbb{T})$ sending f to $[f]$ is continuous and it sends open sets to open sets and closed sets to closed sets.

Let X_n and X be maps from probability spaces Ω_n and Ω, respectively, into $\ell_{\text{loc}}^\infty(\mathbb{T})$. Assume that X is hypi Borel measurable, that is, measurable with respect to the σ-field generated by the hypi-open sets of $\ell_{\text{loc}}^\infty(\mathbb{T})$. Then the map $[X] = [\cdot](X)$ into $L_{\text{loc}}^\infty(\mathbb{T})$ is Borel measurable, too. Since $L_{\text{loc}}^\infty(\mathbb{T})$ is a metric space, weak convergence theory as in van der Vaart and Wellner (1996) applies: we say that X_n weakly hypi-converges to X in $\ell_{\text{loc}}^\infty(\mathbb{T})$ if and only if $[X_n] \rightsquigarrow [X]$ in $L_{\text{loc}}^\infty(\mathbb{T})$. Simplifying notation, we sometimes omit brackets and write $X_n \rightsquigarrow X$ in $L_{\text{loc}}^\infty(\mathbb{T})$.

In order to prove weak hypi-convergence, we will usually combine an initial result on weak convergence of some stochastic process, usually some empirical process and with respect to the supremum distance, with the (extended) continuous mapping theorem [van der Vaart and Wellner (1996), Theorems 1.3.6 and 1.11.1]. The task then consists of proving hypi-continuity of the relevant mappings into $\ell_{\text{loc}}^\infty(\mathbb{T})$ on sufficiently large subsets of their domains. Two situations of particular importance are the following:

- convergence of sums, inducing in particular a variant of Slutsky's lemma (Lemma 3.1 and Appendix A.1);
- convergence to a function in $\ell_{\text{loc}}^\infty(\mathbb{T})$ that is defined as the upper or lower semi-continuous hull of some other function that is originally defined on a dense subset of $\mathbb{T}$ only (Appendix A.2).

LEMMA 3.1 (Slutsky). *Let* $X_n, Y_n : \Omega_n \to \ell_{\text{loc}}^\infty(\mathbb{T})$ *be arbitrary maps and let* $X : \Omega \to \ell_{\text{loc}}^\infty(\mathbb{T})$ *be Borel measurable with respect to the hypi-semimetric. If* $[X_n] \rightsquigarrow [X]$ *and* $[Y_n] \rightsquigarrow [0]$ *in* $L_{\text{loc}}^\infty(\mathbb{T})$, *then* $[X_n + Y_n] \rightsquigarrow [X]$ *in* $L_{\text{loc}}^\infty(\mathbb{T})$.

The proof of this and all other results from this section are given in Appendix F.2 in the supplement [Bücher, Segers and Volgushev (2014)].

By Proposition 2.2, the map from $(\ell^\infty_{\mathrm{loc}}(\mathbb{T}), d_{\mathrm{hypi}})$ into $(\ell^\infty(K), \|\cdot\|_\infty)$ sending a function f to its restriction $f|_K$ on a compact subset K of $\mathbb{T}$ is continuous at every function f which is continuous in every point x of K; here $\|\cdot\|_\infty$ denotes the supremum norm. By a generalization of the continuous mapping theorem to semimetric spaces (Theorem B.2), weak hypi-convergence implies weak convergence with respect to the supremum distance insofar the limit process has continuous trajectories. More precisely, we have the following result.

COROLLARY 3.2. *Let X_n and X be maps from probability spaces Ω_n and Ω, respectively, into $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ such that X is hypi Borel measurable. If $[X_n] \rightsquigarrow [X]$ in $L^\infty_{\mathrm{loc}}(\mathbb{T})$ and if $K \subset \mathbb{T}$ is a nonempty, compact set such that, with probability one, X is continuous in every $x \in K$, then $X_n|_K \rightsquigarrow X|_K$ in $(\ell^\infty(K), \|\cdot\|_\infty)$.*

Taking K to be finite, we find that weak hypi-convergence implies weak convergence of finite-dimensional distributions at points where the limit process is continuous almost surely.

For finite Borel measures μ on $\mathbb{T}$ with compact support, Proposition 2.4 states the continuity of the embedding from the set of Borel measurable functions of $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ equipped with the hypi-topology into $L^p(\mu)$, for every $1 \le p < \infty$. Again by the continuous mapping theorem (Theorem B.2), weak hypi-convergence then implies weak L^p-convergence. A technical nuisance is that in order to view $L^p(\mu)$ as a metric space, we have to consider equivalence classes of functions that are equal μ-almost everywhere; notation $[\cdot]_\mu$.

COROLLARY 3.3. *Let X_n and X be maps from probability spaces Ω_n and Ω, respectively, into $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ such that X is hypi Borel measurable. Let μ be a finite Borel measure on $\mathbb{T}$ with compact support. If $[X_n] \rightsquigarrow [X]$ in $L^\infty_{\mathrm{loc}}(\mathbb{T})$ and if X is μ-almost everywhere continuous with probability one, then $\int |X_{n,\vee} - X_{n,\wedge}|^p \, d\mu$ converges to 0 in outer probability and both $[X_{n,\vee}]_\mu$ and $[X_{n,\wedge}]_\mu$ converge weakly in $L^p(\mu)$ to $[X_\vee]_\mu = [X_\wedge]_\mu$, for every $p \in [1, \infty)$.*

Addition not being continuous on $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$, the latter space is not a topological vector space. This prohibits a direct application of the functional delta method [van der Vaart and Wellner (1996), Theorem 3.9.4] to weak hypi-convergence. However, in Appendix B, we provide a variant of the functional delta method (Theorem B.7) that is sufficiently flexible to deal with maps defined on semimetric spaces endowed with an addition operator that is not necessarily continuous.

4. Empirical copula processes.

Usually, empirical copula processes are studied in the space of bounded functions on $[0, 1]^d$ equipped with the supremum distance. Weak convergence then requires existence and continuity of the first-order partial derivatives of the copula on the interior and some subsets of the boundary of $[0, 1]^d$. In this section, we show what can be done in case the latter smoothness condition is not satisfied. Existence and continuity of the partial derivatives almost everywhere is still enough to ensure weak hypi-convergence of the empirical copula process (Section 4.1). The result is strong enough to validate the bootstrap (Section 4.2) and to analyze Kolmogorov–Smirnov and Cramér–von Mises test statistics, even under local alternatives (Section 4.3). The proofs of the results in this section are given in Appendix C.1 and, partially, in Appendix F.3 in the supplement [Bücher, Segers and Volgushev (2014)].

4.1. *Weak hypi-convergence.* Let $\mathbf{X}_i = (X_{i1}, \dots, X_{id})$, with $i \in \mathbb{N}$, be a strictly stationary sequence of d-variate random vectors. (No confusion should arise from the use of the symbol "d" for both the metric on $\mathbb{T}$ above and the dimension of the random vectors here.) Throughout this section, the joint distribution function F of $\mathbf{X}_i$ is assumed to have continuous marginal distributions $F_1, \dots, F_d$ and its copula is denoted by C. Further, for $j = 1, \dots, d$, let $U_{ij} = F_j(X_{ij})$ and set $\mathbf{U}_i = (U_{i1}, \dots, U_{id})$. Note that $\mathbf{U}_i$ is distributed according to C. Consider the empirical distribution functions

$$F_n(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{\mathbf{X}_i \le \mathbf{x}\}, \qquad G_n(\mathbf{u}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{\mathbf{U}_i \le \mathbf{u}\}$$

for $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{u} \in [0,1]^d$. For a distribution function H on the reals, let

$$H^-(u) := \begin{cases} \inf\{x \in \mathbb{R} : H(x) \ge u\}, & 0 < u \le 1, \\ \sup\{x \in \mathbb{R} : H(x) = 0\}, & u = 0, \end{cases}$$

denote the (left-continuous) generalized inverse function of H.

The object of interest is the empirical copula, defined by

$$C_n(\mathbf{u}) = F_n\bigl(F_{n1}^-(u_1), \dots, F_{nd}^-(u_d)\bigr), \qquad \mathbf{u} \in [0,1]^d,$$

where F_{nj} denotes the jth marginal empirical distribution function. For convenience, we will abbreviate the notation for the empirical copula by $C_n(\mathbf{u}) = F_n(\mathbf{F}_n^-(\mathbf{u}))$, with $\mathbf{F}_n^-(\mathbf{u}) = (F_{n1}^-(u_1), \dots, F_{nd}^-(u_d))$.

Often, the empirical copula is defined as the distribution function of the vector of rescaled ranks, and/or it is turned into a genuine copula via linear interpolation. Since these variants often differ from the empirical copula by at most a term of order $o_p(n^{-1/2})$, uniformly over $[0,1]^d$, they do not affect the asymptotic distribution of the empirical copula process, defined by

$$\mathbb{C}_n = \sqrt{n}(C_n - C). \tag{4.1}$$

The asymptotic behavior of $\mathbb{C}_n$, especially its weak convergence in the space $\ell^\infty([0,1]^d)$ equipped with the supremum norm $\|\cdot\|_\infty$, has been investigated by several authors under various conditions [Bücher and Volgushev (2013), Deheuvels (2009), Fermanian, Radulović and Wegkamp (2004), Ghoudi and

Rémillard (2004), Rüschendorf (1976), Segers (2012), Tsukahara (2005), van der Vaart and Wellner (2007)].

The main arguments to derive the limit of $\mathbb{C}_n$ are as follows. For the sake of a clear explanation, let us assume for the moment that the random vectors $(\mathbf{X}_i)_{i\in\mathbb{N}}$ form an i.i.d. sequence, even though the same arguments work for many time series models with short-range dependence. Observing that $C_n = F_n(\mathbf{F}_n^-) = G_n(\mathbf{G}_n^-)$, we can decompose $\mathbb{C}_n$ into two terms:

$$\mathbb{C}_n = \sqrt{n}\{G_n(\mathbf{G}_n^-) - C\} = \alpha_n(\mathbf{G}_n^-) + \sqrt{n}\{C(\mathbf{G}_n^-) - C\}, \tag{4.2}$$

where $\alpha_n = \sqrt{n}(G_n - C)$ denotes the usual empirical process associated to the sequence $(\mathbf{U}_i)_{i\in\mathbb{N}}$.

Deriving the limit of the first term in (4.2) is standard: since $\alpha_n \rightsquigarrow \alpha$ in $\ell^\infty([0,1]^d)$ with respect to the supremum norm, for a C-Brownian bridge α, and since $\sup_{0\le u_j\le 1} |G_{nj}^-(u_j) - u_j| = o_p(1)$, we obtain that $\alpha_n(\mathbf{G}_n^-) \rightsquigarrow \alpha$ in $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$, too.

Regarding the second term in (4.2), the argumentation is harder. Set $\beta_n = (\beta_{n1}, \ldots, \beta_{nd})$, where $\beta_{nj} = \sqrt{n}(G_{nj}^- - \mathrm{id}_{[0,1]})$ denotes the quantile process of the jth coordinate and where id_A is the identity map on a set A. It follows from the functional delta method applied to the inverse mapping $H \mapsto H^-$ that $\|\beta_{nj} + \alpha_{nj}\|_\infty = o_p(1)$, where $\alpha_{nj}(u_j) = \alpha_n(1, \ldots, 1, u_j, 1, \ldots, 1)$, with $u_j \in [0,1]$ at the jth position. Therefore, $\beta_{nj} \rightsquigarrow -\alpha_j$ in $(\ell^\infty([0,1]), \|\cdot\|_\infty)$, where, similarly, $\alpha_j(u_j)$ is defined as $\alpha(1, \ldots, 1, u_j, 1, \ldots, 1)$. Now,

$$\sqrt{n}\{C(\mathbf{G}_n^-) - C\} = \sqrt{n}\{C(\mathrm{id}_{[0,1]^d} + \beta_n/\sqrt{n}) - C\}, \tag{4.3}$$

which can be handled under suitable differentiability conditions on C. To conclude upon weak convergence with respect to the supremum distance, the weakest assumption so far has been stated in Segers (2012).

CONDITION 4.1. For $j = 1, \ldots, d$ the partial derivatives $\dot{C}_j(\mathbf{u})$ exist and are continuous on $\{\mathbf{u} \in [0,1]^d : u_j \in (0,1)\}$.

Under Condition 4.1,

$$\sqrt{n}\{C(\mathbf{G}_n^-) - C\}(\mathbf{u}) \rightsquigarrow -\sum_{j=1}^d \dot{C}_j(\mathbf{u})\alpha_j(u_j), \tag{4.4}$$

where $\dot{C}_j(\mathbf{u})$ can be defined, for instance, as 0 if $u_j \in \{0,1\}$. Hence,

$$\mathbb{C}_n(\mathbf{u}) \rightsquigarrow \mathbb{C}(\mathbf{u}) = \alpha(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u})\alpha_j(u_j) \tag{4.5}$$

in $\ell^\infty([0,1]^d)$ with respect to the supremum distance.

Condition 4.1 ensures that the limit process $\mathbb{C}$ in (4.5) has continuous trajectories. Actually, if $\mathbb{C}_n$ is to converge weakly with respect to the supremum distance, then the weak limit must have continuous trajectories with probability one. The reason is that the mapping

$$\Delta : \ell^\infty([0,1]^d) \to [0,\infty) : f \mapsto \sup_{\mathbf{u}\in[0,1]^d} \left| f_\vee(\mathbf{u}) - f_\wedge(\mathbf{u}) \right|$$

is continuous with respect to $\|\cdot\|_\infty$ and that $0 \le \Delta\mathbb{C}_n \le d/\sqrt{n} \to 0$ almost surely. The expression for $\mathbb{C}$ in (4.5) then suggests that $\mathbb{C}_n$ does not converge weakly in $(\ell^\infty([0,1]), \|\cdot\|_\infty)$ if Condition 4.1 does not hold.

EXAMPLE 4.2 (Mixture model). For $\lambda \in (0,1)$, consider the bivariate copula given by

$$C(u_1,u_2) = (1-\lambda)u_1u_2 + \lambda \min(u_1,u_2).$$

For $u_1 \neq u_2$, the partial derivatives are

$$\dot{C}_1(u_1,u_2) = (1-\lambda)u_2 + \lambda\mathbb{1}(u_1 < u_2),$$
$$\dot{C}_2(u_1,u_2) = (1-\lambda)u_1 + \lambda\mathbb{1}(u_2 < u_1).$$

On the diagonal $u_1 = u_2$, the partial derivatives do not exist. Still, by the decomposition in (4.2), the finite-dimensional distributions of $\mathbb{C}_n$ can be seen to converge to the ones of the process $\tilde{\mathbb{C}}$ defined as

$$\tilde{\mathbb{C}}(u_1,u_2) = \alpha(u_1,u_2) - \dot{C}_1(u_1,u_2)\alpha_1(u_1) - \dot{C}_2(u_1,u_2)\alpha_2(u_2),$$

if $u_1 \neq u_2$, whereas, on the diagonal $u_1 = u_2 = u$,

$$\tilde{\mathbb{C}}(u,u) = \alpha(u,u) - (1-\lambda)u\{\alpha_1(u) + \alpha_2(u)\} - \lambda\max(\alpha_1(u), \alpha_2(u)),$$

the distribution of which is non-Gaussian.

Now suppose that $\mathbb{C}_n \rightsquigarrow \mathbb{C}$ in $(\ell^\infty([0,1]^d, \|\cdot\|_\infty)$ for some $\mathbb{C}$. Then the finite-dimensional distributions of $\mathbb{C}$ must be equal to the ones of $\tilde{\mathbb{C}}$. Additionally, the trajectories of $\mathbb{C}$ must be continuous almost surely, and thus the law of the random variable $\mathbb{C}(u_1,u_2)$ must depend continuously on the coordinates (u_1,u_2). However, by the above expressions for $\tilde{\mathbb{C}}$, continuity cannot hold at points on the diagonal. This yields a contradiction and, therefore, $\mathbb{C}_n$ cannot converge weakly with respect to the supremum distance.

By considering weak hypi-convergence, we can go far beyond Condition 4.1. Condition 4.3 imposes the regularity needed to deal with the left-hand side of (4.3) in the hypi-semimetric.

CONDITION 4.3. The set $\mathcal{S}$ of points in $[0,1]^d$ where the partial derivatives of the copula C exist and are continuous has Lebesgue measure 1.

Since a copula is monotone in each of its arguments, its partial derivatives automatically exist almost everywhere. Condition 4.3 then only concerns continuity of these partial derivatives. In practice, Condition 4.3 poses no restriction at all. Still, there do exist copulas that do not satisfy Condition 4.3. It can be shown that a bivariate example is given by the copula with Lebesgue density

$$c(u, v) = \frac{\mathbb{1}_{A\times A}(u, v)}{\lambda_1(A)} + \frac{\mathbb{1}_{B\times B}(u, v)}{\lambda_1(B)},$$

where λ_1 denotes the one-dimensional Lebesgue measure, $A \subset [0, 1]$ is a closed set which is at the same time nowhere dense and satisfies $\lambda_1(A) \in (0, 1)$ and where $B = [0, 1] \setminus A$.

For broad applicability, we relax the assumption of serial independence and replace it by the following condition, which holds for i.i.d. sequences as well as for stationary sequences under various weak dependence conditions [Dehling and Durieu (2011), Doukhan, Fermanian and Lang (2009), Rio (2000)].

CONDITION 4.4. The empirical process $\alpha_n = \sqrt{n}(G_n - C)$ converges weakly in $(\ell^\infty([0, 1]^d), \|\cdot\|_\infty)$ to some limit process α which has continuous sample paths, almost surely.

Under Condition 4.3, the term on the right-hand side of (4.4) is defined only on $\mathcal{S}$. We extend it to the whole of $[0, 1]^d$ by taking lower semicontinuous hulls as in Appendix A.2. Let $|\cdot|$ denote the Euclidean norm and let $\mathcal{C}(A)$ be the set of continuous real-valued functions on a domain A. Recall our convention of omitting the brackets $[\cdot]$ when working in $L^\infty_{\mathrm{loc}}(\mathbb{T})$.

THEOREM 4.5. *Suppose that Condition 4.4 holds and that C satisfies Condition 4.3. Then*

$$\mathbb{C}_n \rightsquigarrow \mathbb{C} = \alpha + dC_{(-\alpha_1,\dots,-\alpha_d)} \tag{4.6}$$

in $(L^\infty([0, 1]^d), d_{\mathrm{hypi}})$, where, for $a = (a_1, \dots, a_d) \in \{\mathcal{C}([0, 1])\}^d$,

$$dC_a(\mathbf{u}) = \sup_{\varepsilon>0} \inf\left\{ \sum_{j=1}^d \dot{C}_j(\mathbf{v}) a_j(v_j) : \mathbf{v} \in \mathcal{S}, |\mathbf{v} - \mathbf{u}| < \varepsilon \right\}.$$

By Section 2, Theorem 4.5 has several useful consequences.

- First, it implies weak convergence with respect to the supremum distance of the restriction of the empirical copula process to compact subsets of the union of $\mathcal{S}$ and the boundary of $[0, 1]^d$; see Corollary 3.2. This is akin to the convergence results for multilinear empirical copulas for count data in Genest, Nešlehová and Rémillard (2014). Note that, in particular, we obtain the weak convergence result in (4.5) under the stronger Condition 4.1.

- Furthermore, we obtain weak convergence of the empirical copula process in $(L^p([0,1]^d), \|\cdot\|_p)$ for any $1 \le p < \infty$. To the best of our knowledge, this result is new and opens the door to L^p-type inference procedures for a broad class of copulas.

Two possible applications are treated in the following subsections.

4.2. *A bootstrap device.* Assume that $\mathbf{X}_1, \dots, \mathbf{X}_n$ are serially independent. We show that the bootstrap based on resampling with replacement [Fermanian, Radulović and Wegkamp (2004)] and the bootstrap based on the multiplier central limit theorem [Bücher and Dette (2010)] provide valid approximations for $\mathbb{C}$ with respect to the hypi-semimetric. Our multiplier bootstrap is different from the approach in Rémillard and Scaillet (2009), which requires estimation of the first-order partial derivatives of C.

Let $M \in \mathbb{N}$ be some large integer and, for each $m \in \{1, \dots, M\}$, let $\mathbf{X}_1^{[m]}, \dots, \mathbf{X}_n^{[m]}$ be drawn with replacement from the sample. The *resampling bootstrap empirical copula process* is defined as

$$\mathbb{C}_n^{[m]} = \sqrt{n}\bigl(C_n^{[m]} - C_n\bigr), \tag{4.7}$$

where $C_n^{[m]}$ denotes the empirical copula calculated from the bootstrap sample $\mathbf{X}_1^{[m]}, \dots, \mathbf{X}_n^{[m]}$. Note that we can represent $C_n^{[m]}$ by $F_n^{[m]}(\mathbf{F}_n^{[m]-})$, where

$$F_n^{[m]}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n W_{ni}^{[m]} \mathbb{1}(\mathbf{X}_i \le \mathbf{x})$$

and where $W_n^{[m]} = (W_{n1}^{[m]}, \dots, W_{nn}^{[m]})$ denotes a multinomial random vector with n trials, n possible outcomes, and success probabilities $(1/n, \dots, 1/n)$, independent of the sample and independent across $m \in \{1, \dots, M\}$.

Regarding the multiplier bootstrap, let $\{\xi_i^{[m]} : i \ge 1, m = 1, \dots, M\}$ be i.i.d. random variables, independent of the sample, with both mean and variance equal to one and such that $\int_0^\infty \sqrt{\mathbb{P}(\xi_i > x)}\,dx < \infty$. Let

$$\tilde{F}_n^{[m]}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n \xi_i^{[m]} \mathbb{1}(\mathbf{X}_i \le \mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^d$$

and define

$$\tilde{\mathbb{C}}_n^{[m]} = \sqrt{n}\bigl\{\tilde{F}_n^{[m]}\bigl(\tilde{\mathbf{F}}_n^{[m]-}\bigr) - C_n\bigr\} \tag{4.8}$$

as the *multiplier bootstrap empirical copula process*. The following proposition shows that both $\mathbb{C}_n^{[1]}, \dots, \mathbb{C}_n^{[M]}$ and $\tilde{\mathbb{C}}_n^{[1]}, \dots, \tilde{\mathbb{C}}_n^{[M]}$ can be regarded as asymptotically independent copies of $\mathbb{C}_n$.

PROPOSITION 4.6. *Let $\mathbf{X}_i$, $i \in \mathbb{N}$, be i.i.d. d-variate random vectors with common distribution function F having continuous margins and a copula C satisfying Condition 4.3. Let $\mathbb{C}_n$, $\mathbb{C}_n^{[m]}$ and $\tilde{\mathbb{C}}_n^{[m]}$ be as in (4.1), (4.7) and (4.8), respectively. Then both $(\mathbb{C}_n, \mathbb{C}_n^{[1]}, \dots, \mathbb{C}_n^{[M]})$ and $(\mathbb{C}_n, \tilde{\mathbb{C}}_n^{[1]}, \dots, \tilde{\mathbb{C}}_n^{[M]})$ weakly converge to $(\mathbb{C}, \mathbb{C}^{[1]}, \dots, \mathbb{C}^{[M]})$ in the space $(L^\infty([0,1]^d), d_{\mathrm{hypi}})^{M+1}$, where $\mathbb{C}^{[1]}, \dots, \mathbb{C}^{[M]}$ denote independent copies of $\mathbb{C}$ in (4.6).*

By hypi-continuity of the supremum and infimum functionals (see Proposition 2.3), the bootstrap approximation can, for instance, be used to construct asymptotic uniform confidence bands for the copula.

4.3. *Power curves of tests for independence.* In the present section, we derive weak hypi-convergence of the empirical copula process for triangular arrays. We apply it to the problem of comparing statistical tests for independence by local power curves. This comparison has been carried out by Genest, Quessy and Remillard (2007) under strong differentiability assumptions on copula densities. By considering hypi-convergence, we can extend their results to copulas that do not have a density with respect to the Lebesgue measure.

We consider a triangular array of random vectors $\mathbf{X}_1^{(n)}, \dots, \mathbf{X}_n^{(n)}$ which are row-wise i.i.d. with continuous marginals and copula $C^{(n)}$. We suppose that there exists a copula C satisfying Condition 4.3 such that

$$\Delta_n = \sqrt{n}\{C^{(n)} - C\} \to \Delta \tag{4.9}$$

uniformly, for some continuous function Δ on $[0,1]^d$. Let $C_n^{(n)}$ denote the empirical copula based on $\mathbf{X}_1^{(n)}, \dots, \mathbf{X}_n^{(n)}$. Let $\mathbf{U}_1^{(n)}, \dots, \mathbf{U}_n^{(n)}$ denote the sample obtained by the marginal probability integral transform and let $G_n^{(n)}$ and $\alpha_n^{(n)}$ denote its empirical distribution function and empirical process, respectively. Similarly as before, we have the decomposition

$$\begin{aligned}
\mathbb{C}_n^{(n)} &= \sqrt{n}\{C_n^{(n)} - C^{(n)}\} \\
&= \sqrt{n}\{G_n^{(n)}(\mathbf{G}_n^{(n)-}) - C^{(n)}(\mathbf{G}_n^{(n)-})\} + \sqrt{n}\{C^{(n)}(\mathbf{G}_n^{(n)-}) - C^{(n)}\} \\
&= \alpha_n^{(n)}(\mathbf{G}_n^{(n)-}) + \sqrt{n}\{C(\mathbf{G}_n^{(n)-}) - C\} + \{\Delta_n(\mathbf{G}_n^{(n)-}) - \Delta_n\}.
\end{aligned}$$

We will show in Appendix F.3 in the supplement [Bücher, Segers and Volgushev (2014)] that $\alpha_n^{(n)} \rightsquigarrow \alpha$ in $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$, where α is a C-Brownian bridge. Therefore, the first summand weakly converges to α with respect to the supremum norm. The second summand weakly converges in the hypi-topology to $dC_{(-\alpha_1,\dots,-\alpha_d)}$, while the last one converges to $\Delta - \Delta \equiv 0$, uniformly. This motivates the following result.

PROPOSITION 4.7. *Given the above set-up and if* (4.9) *is met with C satisfying Condition* 4.3, *we have* $\mathbb{C}_n^{(n)} \rightsquigarrow \mathbb{C}$ *in* $(L^\infty([0,1]^d, d_{\mathrm{hypi}})$, *where* $\mathbb{C}$ *is the same process as in Theorem* 4.5. *Additionally, in* $(L^\infty([0,1]^d), d_{\mathrm{hypi}})$,

$$\sqrt{n}(C_n^{(n)} - C) \rightsquigarrow \mathbb{C} + \Delta.$$

To illustrate the latter result, we investigate the local efficiency of tests for independence as considered in Genest, Quessy and Remillard (2007). Instead of imposing conditions (i) and (ii) on page 169 in their paper, we only suppose that (4.9) holds with $C = \Pi$, the independence copula, and $\Delta = \delta\Lambda$, where $\Lambda \in \mathcal{C}([0,1]^d)$ and $\delta \geq 0$. For brevity, we only compare the test statistics

$$T_n = n \int_{[0,1]^d} \{C_n^{(n)} - \Pi\}^2 \, d\Pi \quad \text{and} \quad S_n = \sqrt{n} \|C_n^{(n)} - \Pi\|_\infty.$$

From weak hypi-convergence of $\sqrt{n}(C_n^{(n)} - C)$ and Propositions 2.2 and 2.4, we obtain that

$$T_n \rightsquigarrow \mathcal{T}_\delta = \int_{[0,1]^d} (\mathbb{C} + \delta\Lambda)^2 \, d\Pi, \qquad S_n \rightsquigarrow \mathcal{S}_\delta = \|\mathbb{C} + \delta\Lambda\|_\infty.$$

Hence, the local power curves of the tests to the level $\alpha \in (0,1)$ in direction Λ are given by

$$\delta \mapsto \mathbb{P}\{\mathcal{T}_\delta > q_{\mathcal{T}_0}(1-\alpha)\}, \qquad \delta \mapsto \mathbb{P}\{\mathcal{S}_\delta > q_{\mathcal{S}_0}(1-\alpha)\},$$

where $q_{\mathcal{T}_0}(1-\alpha)$ and $q_{\mathcal{S}_0}(1-\alpha)$ denote the $(1-\alpha)$-quantiles of $\mathcal{T}_0$ and $\mathcal{S}_0$, respectively. These curves can be compared by analytical calculations as in Genest, Quessy and Remillard (2007) or by simulation.

5. Stable tail dependence functions. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$, where $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})$, be i.i.d. d-variate random vectors with distribution function F and continuous marginal distribution functions $F_1, \ldots, F_d$. We assume that the following limit, called the stable tail dependence function of F,

$$L(\mathbf{x}) = \lim_{t \downarrow 0} t^{-1} \mathbb{P}\{1 - F_1(X_{11}) \leq t x_1 \text{ or } \cdots \text{ or } 1 - F_d(X_{1d}) \leq t x_d\}, \tag{5.1}$$

exists as a function $L : [0,\infty)^d \to [0,\infty)$.

For $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$, let R_i^j denote the rank of X_{ij} among $X_{1j}, \ldots, X_{nj}$. Replacing all distribution functions in (5.1) by their empirical counterparts and replacing t by k/n where $k = k_n$ is a positive sequence such that $k_n \to \infty$ and $k_n = o(n)$, we obtain the following nonparametric estimator for L_n, called the empirical (stable) tail dependence function:

$$\hat{L}_n(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^n \mathbb{1}\left\{R_i^1 > n + \frac{1}{2} - k x_1 \text{ or } \cdots \text{ or } R_i^d > n + \frac{1}{2} - k x_d / n\right\}$$

[Drees and Huang (1998), Huang (1992)]. The inclusion of the term $1/2$ inside the indicators serves to improve the finite sample behavior of the estimator.

In Einmahl, Krajina and Segers (2012), a functional central limit theorem for $\sqrt{k}(\hat{L}_n - L)$ is given in the topology of uniform convergence on compact subsets of $[0,\infty)^d$. The result requires L to have continuous first-order partial derivatives on sufficiently large subsets of $[0,\infty)^d$, similar to Condition 4.1 for copulas. By switching to weak hypi-convergence, we are able to get rid of smoothness conditions altogether.

Similarly as in Section 4, let $\mathcal{S}$ denote the set of all points $\mathbf{x} \in [0,\infty)^d$ where L is differentiable. The function L being convex, Theorem 25.5 in Rockafellar (1970) implies that the complement of $\mathcal{S}$ is a Lebesgue null set and that the gradient $(\dot{L}_1, \dots, \dot{L}_d)$ of L is continuous on $\mathcal{S}$. Proceeding as in Appendix A.2, we may define, for any $(a_1, \dots, a_d) \in \{C([0,\infty))\}^d$, a function on $[0,\infty)^d$ by

$$dL_{(a_1,\dots,a_d)}(\mathbf{x}) = \sup_{\varepsilon>0} \inf\left\{\sum_{j=1}^d \dot{L}_j(\mathbf{y}) a_j(y_j) : \mathbf{y} \in \mathcal{S}, |\mathbf{x}-\mathbf{y}| < \varepsilon\right\}. \tag{5.2}$$

As in Einmahl, Krajina and Segers (2012), let Λ be the Borel measure on $[0,\infty)^d$ such that $\Lambda(A(\mathbf{x})) = L(x)$ where $A(\mathbf{x}) = \bigcup_{j=1}^d \{y \in [0,\infty)^d : y_j \leq x_j\}$ for $\mathbf{x} \in [0,\infty)^d$. Let $\mathbb{W}$ be a mean-zero Gaussian process on $[0,\infty)^d$ with continuous trajectories and with covariance function $\mathbb{E}[\mathbb{W}(\mathbf{x})\mathbb{W}(\mathbf{y})] = \Lambda(A(\mathbf{x}) \cap A(\mathbf{y}))$. Let $\Delta_{d-1} = \{\mathbf{x} \in [0,1]^d : x_1 + \cdots + x_d = 1\}$ be the unit simplex in $\mathbb{R}^d$. For $f \in \ell^\infty_{\mathrm{loc}}([0,\infty)^d)$ and $j = 1, \dots, d$, define $f_j^0 \in \ell^\infty_{\mathrm{loc}}([0,\infty))$ through $f_j^0(x_j) = f(0, \dots, 0, x_j, 0, \dots, 0)$. Recall our convention of omitting the brackets $[\cdot]$ when working in $L^\infty_{\mathrm{loc}}(\mathbb{T})$.

THEOREM 5.1. *Let $\mathbf{X}_i$, $i \in \mathbb{N}$, be i.i.d. d-dimensional random vectors with common distribution function F with continuous margins $F_1, \dots, F_d$ and stable tail dependence function L. Suppose that the following conditions hold*:

(C1) *For some $\alpha > 0$ we have, uniformly in $\mathbf{x} \in \Delta_{d-1}$,*

$$\begin{aligned} &t^{-1}\mathbb{P}\{1 - F_1(X_{11}) \leq tx_1 \text{ or } \cdots \text{ or } 1 - F_d(X_{1d}) \leq tx_d\} \\ &\quad = L(\mathbf{x}) + O(t^\alpha), \qquad t \downarrow 0. \end{aligned}$$

(C2) *We have $k = o(n^{2\alpha/(1+2\alpha)})$ and $k \to \infty$ as $n \to \infty$.*

Then, in $(L^\infty_{\mathrm{loc}}([0,\infty)^d), d_{\mathrm{hypi}})$,

$$\sqrt{k}(\hat{L}_n - L) \rightsquigarrow \mathbb{W} + dL_{(-\mathbb{W}_1^0, \dots, -\mathbb{W}_d^0)}, \qquad n \to \infty.$$

The proof of Theorem 5.1 is similar to the one of Theorem 4.5 and is deferred to the supplement [Bücher, Segers and Volgushev (2014)].

Conditions (C1) and (C2) also appear in Theorem 4.6 in Einmahl, Krajina and Segers (2012) and are needed to ensure that the estimator is asymptotically unbiased. The difference with their theorem is that we do not need their condition (C3) on the partial derivatives of L. Therefore, Theorem 5.1 also covers piecewise linear stable tail dependence functions arising from max-linear models [Wang and Stoev (2011)].

Weak hypi-convergence of $\sqrt{k}(\hat{L}_n - L)$ can be exploited to validate statistical procedures for tail dependence functions in the same way as was done with weak hypi-convergence of empirical copula processes in Section 4. In contrast to copulas, no smoothness conditions on L are needed at all. Applications include the bootstrap [Peng and Qi (2008)] and minimum L^2-distance estimation [Bücher and Dette (2013)]. Hypi-convergence implying L^2-convergence, Theorem 5.1 also provides another way to prove the asymptotic normality of the M-estimator in Einmahl, Krajina and Segers (2012).

6. Error distributions in regression models. Consider a linear regression model for a sample $(\mathbf{X}_i, Y_i)$, $i \in \{1, \dots, n\}$, in $\mathbb{R}^p \times \mathbb{R}$, of the form

$$Y_i = \mathbf{X}_i' \boldsymbol{\beta} + \varepsilon_i. \tag{6.1}$$

Here, $(\mathbf{X}_i, \varepsilon_i)$, for $i \in \{1, \dots, n\}$, are i.i.d. random vectors in $\mathbb{R}^p \times \mathbb{R}$. It is assumed that $\mathbf{X}_i$ and ε_i are independent and that the distribution of ε_i is constrained in such a way that the vector of regression coefficients $\boldsymbol{\beta}$ is identifiable (provided the support of $\mathbf{X}_i$ is sufficiently large). For instance, the requirement $\mathbb{E}(\varepsilon_i) = 0$ yields a mean regression model, whereas $\operatorname{median}(\varepsilon_i) = 0$ yields a median regression model. For simplicity, we restrict attention to serial independence and to a scalar dependent variable.

The model is semiparametric with parametric component $\boldsymbol{\beta} \in \mathbb{R}^p$ and nonparametric components P^X and P^ε, the distributions of the explanatory variables $\mathbf{X}_i$ and the errors ε_i. We are interested in the estimation of the cumulative distribution function, F, of ε_i:

$$F(z) = \mathbb{P}(\varepsilon_i \le z), \qquad z \in \bar{\mathbb{R}},$$

where $\bar{\mathbb{R}} = [-\infty, \infty]$, a convenient compactification of the real line.

Let $\hat{\boldsymbol{\beta}}_n$ be a consistent estimator for $\boldsymbol{\beta}$. In Theorem 6.1 below, we will be more specific about the asymptotic distribution of $\hat{\boldsymbol{\beta}}_n$. We define estimated residuals as

$$\hat{\varepsilon}_{n,i} = Y_i - \mathbf{X}_i' \hat{\boldsymbol{\beta}}_n = \varepsilon_i - \mathbf{X}_i' (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \tag{6.2}$$

and obtain a simple estimator for F by

$$\hat{F}_n(z) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\hat{\varepsilon}_{n,i} \le z), \qquad z \in \bar{\mathbb{R}}. \tag{6.3}$$

The *empirical residual process* corresponding to $\hat{F}_n$ is

$$\mathbb{F}_n(z) = \sqrt{n}\{\hat{F}_n(z) - F(z)\}, \qquad z \in \bar{\mathbb{R}}. \tag{6.4}$$

Weak convergence results for $\mathbb{F}_n$ play a central role in, for example, testing the goodness-of-fit of error distributions or in the derivation of the asymptotic behavior of more sophisticated estimators for F; see Koul and Qian (2002) for an overview. First results on the asymptotic behavior of $\mathbb{F}_n$ were derived in Koul (1969) and Loynes (1980) (in generalized regression models), and more recently those findings were extended in various directions such as, for instance, time series analysis [see Engler and Nielsen (2009), Koul and Qian (2002), and the references cited therein] or coefficient vectors of growing dimension [see Chen and Lockhart (2001), for an overview]. All of those extensions share the assumption that F has a continuous probability density function f. In that case, weak convergence takes place with respect to the supremum distance and the process admits an expansion of the form

$$\mathbb{F}_n(z) = \left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\mathbb{1}(\varepsilon_i \le z) - F(z)\}\right] + f(z)\mathbb{E}[\mathbf{X}]'\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) + o_p(1) \tag{6.5}$$

uniformly in $z \in \bar{\mathbb{R}}$, where $\mathbf{X}$ denotes a random vector with the same distribution as $\mathbf{X}_i$. In the present section, we will drop the assumption of continuity of f and consider weak hypi-convergence of $\mathbb{F}_n$.

The main arguments underlying the derivation of the limit of $\mathbb{F}_n$ are as follows. Let $\mathbb{P}_n = n^{-1}\sum_{i=1}^n \delta_{\mathbf{X}_i,\varepsilon_i}$ denote the empirical measure of the sample $(\mathbf{X}_i, \varepsilon_i)$, $i \in \{1,\dots,n\}$. For $(z, \boldsymbol{\delta}) \in \bar{\mathbb{R}} \times \mathbb{R}^p$, consider the function

$$\begin{aligned} f_{z,\boldsymbol{\delta}} &: \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}, \\ &(\mathbf{x}, \varepsilon) \mapsto \mathbb{1}(\varepsilon \le z + \mathbf{x}'\boldsymbol{\delta}), \end{aligned} \tag{6.6}$$

and let $\mathcal{F}$ denote the collection of all those functions, that is,

$$\mathcal{F} = \{f_{z,\boldsymbol{\delta}} : z \in \bar{\mathbb{R}}, \boldsymbol{\delta} \in \mathbb{R}^p\}. \tag{6.7}$$

Combining (6.2) and (6.3) on the one hand with (6.6) on the other hand, we find

$$\hat{F}_n(z) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}(\varepsilon_i \le z + \mathbf{X}_i'(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})) = \mathbb{P}_n f_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}},$$

where we use the usual operator notation $Qh = \int h\,dQ$ for a signed measure Q and a measurable function h. Moreover, let P denote the common law of the random vectors $(\mathbf{X}_i, \varepsilon_i)$, yielding $F(z) = \mathbb{E}[f_{z,\mathbf{0}}(\mathbf{X}_i, \varepsilon_i)] = Pf_{z,\mathbf{0}}$ for $z \in \bar{\mathbb{R}}$. Then the empirical process $\mathbb{F}_n$ in (6.4) admits the decomposition

$$\begin{aligned} \mathbb{F}_n(z) &= \sqrt{n}(\mathbb{P}_n f_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}} - Pf_{z,\mathbf{0}}) \\ &= \mathbb{G}_n f_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}} + \sqrt{n}(Pf_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}} - Pf_{z,\mathbf{0}}), \end{aligned} \tag{6.8}$$

where $\mathbb{G}_n$ is shorthand for $\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - P)$. The decomposition in (6.8) is akin to the one in (4.2) for the empirical copula process. If $\hat{\boldsymbol{\beta}}_n$ is consistent for $\boldsymbol{\beta}$, the first term can be shown to be

$$\mathbb{G}_n f_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}} = \mathbb{G}_n f_{z,\mathbf{0}} + o_p(1) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbb{1}(\varepsilon_i \le z) - F(z)\} + o_p(1)$$

uniformly in $z \in \bar{\mathbb{R}}$. The process on the right-hand side is the usual empirical process corresponding to $\varepsilon_1, \ldots, \varepsilon_n$, and its weak convergence is one of the classical results of empirical process theory.

The treatment of the second term in (6.8) will be based on a linear expansion of the map $\boldsymbol{\delta} \mapsto P f_{z,\boldsymbol{\delta}}$ around $\mathbf{0}$. For $(z, \boldsymbol{\delta}) \in \bar{\mathbb{R}} \times \mathbb{R}^p$, we have

$$P f_{z,\boldsymbol{\delta}} - P f_{z,\mathbf{0}} = \int_{\mathbb{R}^p} \{F(z + \mathbf{x}'\boldsymbol{\delta}) - F(z)\} P^X(d\mathbf{x}).$$

Therefore, if F is continuously differentiable with derivative f, we can expect that

$$\begin{aligned} \sqrt{n}\{P f_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}} - P f_{z,\mathbf{0}}\} &= \sqrt{n} \int_{\mathbb{R}^p} f(z)\mathbf{x}'(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) P^X(d\mathbf{x}) + o_p(1) \\ &= f(z)\mathbb{E}[\mathbf{X}]'\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) + o_p(1), \end{aligned} \tag{6.9}$$

which will converge weakly provided $\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})$ converges weakly. However, if F is not differentiable at a point z or if f exists but is not continuous in z, then (6.9) and as a consequence weak convergence with respect to the supremum distance may fail. Still, weak hypi-convergence continues to hold, as the main result of this section shows.

THEOREM 6.1. *Consider a model of the form* (6.1) *such that* $(\mathbf{X}_i, \varepsilon_i)$, $i \in \mathbb{N}$, *are i.i.d. random vectors in* $\mathbb{R}^p \times \mathbb{R}$ *and such that* $\mathbf{X}_i$ *and* ε_i *are independent. Additionally, suppose that the following conditions hold*:

(R1) *The estimator* $\hat{\boldsymbol{\beta}}_n$ *admits a linear expansion of the form*

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = (\mathbb{G}_n \psi_1, \ldots, \mathbb{G}_n \psi_p)' + o_p(1), \qquad n \to \infty,$$

in terms of zero-mean, square-integrable functions $\boldsymbol{\psi}_j : \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$, *for* $j \in \{1, \ldots, p\}$.

(R2) *The distribution* F *is absolutely continuous. There exists a version of its density* f *which is uniformly bounded and which is* làdlàg, *that is*, *which admits right-hand and left-hand limits at every* $z \in \mathbb{R}$:

$$f(z+) = \lim_{0<s\to 0} f(z+s), \qquad f(z-) = \lim_{0<s\to 0} f(z-s).$$

Moreover, $f(\pm\infty) := \lim_{z\to\pm\infty} f(z) = 0$.

(R3) *The norm of* $\mathbf{X}$ *is integrable*, *that is*, $\mathbb{E}[|\mathbf{X}|] < \infty$.

Set $\psi = (\psi_1, \ldots, \psi_p)$, $\mathcal{G} = \mathcal{F} \cup \{\psi_1, \ldots, \psi_p\}$ *and let* $\mathbb{G}$ *denote a* P*-Brownian bridge in* $\ell^\infty(\mathcal{G})$, *that is, a zero-mean Gaussian process on* $\mathcal{G}$ *with covariance function*

$$\operatorname{cov}(\mathbb{G}g_1, \mathbb{G}g_2) = \operatorname{cov}\bigl(g_1(\mathbf{X}, \varepsilon), g_2(\mathbf{X}, \varepsilon)\bigr), \qquad g_1, g_2 \in \mathcal{G}. \tag{6.10}$$

Then, in $(L^\infty(\bar{\mathbb{R}}), d_{\mathrm{hypi}})$, *we have* $\mathbb{F}_n \rightsquigarrow \mathbb{F}$ *as* $n \to \infty$, *where the limiting process* $\mathbb{F}$ *can be written as* $\mathbb{F}(\pm\infty) = 0$ *a.s. and*

$$\begin{aligned} \mathbb{F}(z) = \mathbb{G} f_{z,\mathbf{0}} &- f(z-) \int_{-\infty}^{0} P^X\bigl(\{\mathbf{x} : \mathbf{x}'\mathbb{G}\boldsymbol{\psi} < y\}\bigr)\, dy \\ &+ f(z+) \int_{0}^{+\infty} P^X\bigl(\{\mathbf{x} : \mathbf{x}'\mathbb{G}\boldsymbol{\psi} > y\}\bigr)\, dy, \qquad z \in \mathbb{R}. \end{aligned} \tag{6.11}$$

Note that the limit in (6.11) is not càdlàg, whence the classical Skorohod-topologies cannot be applied in the present context.

The influence function $\boldsymbol{\psi} = (\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_p)$ in (R1) depends on the estimator and on the true model. A classical example is given by the ordinary least squares estimator: if the errors ε_i have mean zero and finite variance and if the components of $\mathbf{X}$ have finite second moments and the $p \times p$ matrix $\mathbb{E}(\mathbf{X}\mathbf{X}')$ is invertible, then

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \bigl\{\mathbb{E}\bigl(\mathbf{X}\mathbf{X}'\bigr)\bigr\}^{-1} \mathbf{X}_i \varepsilon_i + o_p(1).$$

If f happens to be continuous in z, then $f(z-) = f(z+) = f(z)$, and we obtain that $\mathbb{F}(z) = \mathbb{G} f_{z,\mathbf{0}} + f(z)\mathbb{E}[\mathbf{X}']\mathbb{G}\boldsymbol{\psi}$, which, under (R1), coincides with the limit of the classical representation in (6.5). If f is continuous everywhere, then $\mathbb{F}$ is almost surely continuous, and, by Corollary 3.2 with $K = \bar{\mathbb{R}}$, the weak convergence in $\mathbb{F}_n \rightsquigarrow \mathbb{F}$ takes place with respect to the supremum distance.

EXAMPLE 6.2 (Mixtures of exponential distributions). Consider the probability density function

$$f_{\boldsymbol{\theta}}(z) = \begin{cases} \dfrac{1}{\theta_-} e^{z/\theta_-}, & \text{if } z < 0, \\[2ex] (1 - w)\dfrac{1}{\theta_+} e^{-z/\theta_+}, & \text{if } z > 0, \end{cases} \tag{6.12}$$

where $w = \theta_+/(\theta_- + \theta_+)$. This density is a mixture of the exponential distribution on $(-\infty, 0)$ with mean $-\theta_-$ and the exponential distribution on $(0, \infty)$ with mean θ_+, with weights chosen so that the total mean is zero. The left-hand and right-hand limits of $f_{\boldsymbol{\theta}}$ at 0 are

$$f_{\boldsymbol{\theta}}(0-) = \frac{\theta_+}{\theta_-} \frac{1}{\theta_- + \theta_+}, \qquad f_{\boldsymbol{\theta}}(0+) = \frac{\theta_-}{\theta_+} \frac{1}{\theta_- + \theta_+}.$$

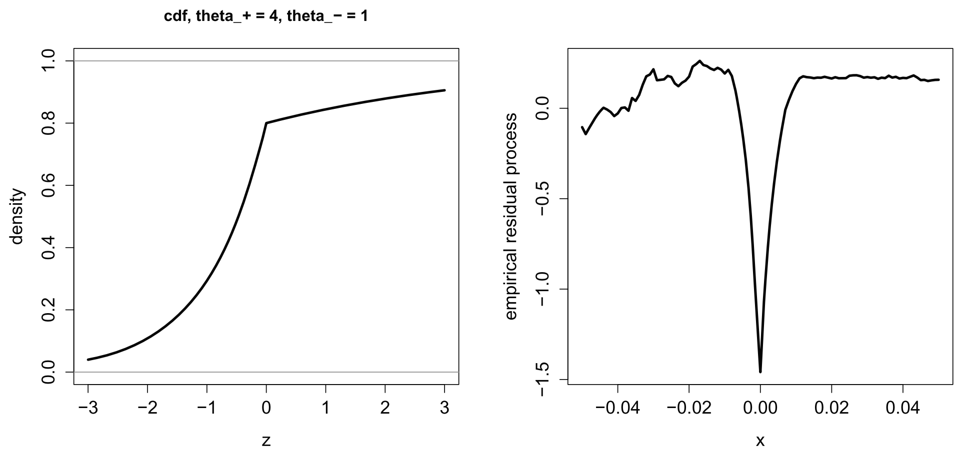

FIG. 2. Left: *cumulative distribution function of the mixed double exponential distribution in* (6.12). Right: *trajectories of the corresponding empirical residual process* $\mathbb{F}_n$ *for* $n = 10^6$. *In both cases*: $(\theta_-, \theta_+) = (1, 4)$.

If θ_- is different from θ_+, these limits are different, and thus the associated distribution function, $F_{\boldsymbol{\theta}}$, is not continuously differentiable at 0. See the left-hand side of Figure 2 for the graph of F_θ when (θ_-, θ_+) is equal to $(1, 4)$.

Now, consider the linear regression model in (6.1) with $p = 1$, $X_i \sim N(0, 1)$ independent of ε_i, and with ε_i distributed according to (6.12). The parameter $\boldsymbol{\beta} \in \mathbb{R}$ is estimated by ordinary least squares, $\hat{\boldsymbol{\beta}}_n$, and the corresponding empirical residual process $\mathbb{F}_n$ is calculated as in (6.4). Theorem 6.1 implies that $\mathbb{F}_n$ converges in $(L^\infty(\bar{\mathbb{R}}), d_{\text{hypi}})$ to the process $\mathbb{F}$ given by (note the simplification arising from $\mathbb{E}[X] = 0$)

$$\mathbb{F}(z) = \mathbb{G} f_{z,\mathbf{0}} + \big(f_{\boldsymbol{\theta}}(z+) - f_{\boldsymbol{\theta}}(z-)\big) \int_0^{+\infty} P^X\big(\{x : x\mathbb{G}\psi > y\}\big)\, dy,$$

where $\mathbb{G}$ is a P-Brownian bridge for $P = P^X \otimes P^\varepsilon$ and where $f_{z,\mathbf{0}}$ and ψ are certain functions in $L^2(P)$. We find that $\mathbb{F}(z) = \mathbb{G} f_{z,\mathbf{0}}$ for $z \neq 0$, a continuous Gaussian process. The only discontinuity occurs at $z = 0$, when the left-hand and right-hand limits of $f_{\boldsymbol{\theta}}$ are different. The "spike" in $\mathbb{F}_n$ then goes upward or downward according to whether $f_{\boldsymbol{\theta}}(z+)$ is larger than or smaller than $f_{\boldsymbol{\theta}}(z-)$. A simulated typical trajectories of $\mathbb{F}_n(z)$ for $n = 10^6$ and $z \in [-0.05, 0.05]$ is shown on the right-hand side of Figure 2 when (θ_-, θ_+) is equal to $(1, 4)$.

APPENDIX A: VERIFYING HYPI-CONVERGENCE

In this appendix, we provide some tools for showing convergence of a sequence of functions with respect to the hypi-semimetric. Proofs are deferred to Appendix D in the supplement [Bücher, Segers and Volgushev (2014)].

A.1. Pointwise convergence and convergence of sums. Let $(\mathbb{T}, d)$ be a metric space. For $f : \mathbb{T} \to \mathbb{R}$, define extended real-valued functions $f_\wedge$ and $f_\vee$ as in (2.1) and (2.2), respectively. Since we do not require f to be locally bounded, $f_\wedge$ and $f_\vee$ can attain $-\infty$ and $+\infty$, respectively.

For $f_n : \mathbb{T} \to \mathbb{R}$, we say that f_n epi-converges to $\alpha \in \mathbb{R}$ at $x \in \mathbb{T}$ if the following two conditions are met:

$$\text{(A.1)} \qquad \begin{cases} \text{(i) } \forall x_n \to x : \liminf_{n\to\infty} f_n(x_n) \geq \alpha, \\ \text{(ii) } \exists x_n \to x : \limsup_{n\to\infty} f_n(x_n) \leq \alpha. \end{cases}$$

Similarly, f_n hypo-converges to α at x if

$$\text{(A.2)} \qquad \begin{cases} \text{(i) } \forall x_n \to x : \limsup_{n\to\infty} f_n(x_n) \leq \alpha, \\ \text{(ii) } \exists x_n \to x : \liminf_{n\to\infty} f_n(x_n) \geq \alpha, \end{cases}$$

which is equivalent to epi-convergence of $-f_n$ to $-\alpha$ at x. If additionally $f : \mathbb{T} \to \mathbb{R}$, then f_n is said to epi- or hypo-converge to f at x if $\alpha = f(x)$ in the preceding conditions. According to Proposition 2.1, f_n hypi-converges to f in $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$ if and only if f_n epi-converges to $f_\wedge$ and hypo-converges to $f_\vee$ at every $x \in \mathbb{T}$. For $x \in \mathbb{T}$ and $\varepsilon > 0$, let $B(x, \varepsilon) = \{y \in \mathbb{T} : d(x, y) < \varepsilon\}$.

LEMMA A.1 (Convergence of hulls). *Let $f_n : \mathbb{T} \to \mathbb{R}$, $x \in \mathbb{T}$ and $\alpha \in \mathbb{R}$. Then f_n epi-converges to α at x if and only if $f_{n,\wedge}$ epi-converges to α at x, and f_n hypo-converges to α at x if and only if $f_{n,\vee}$ hypo-converges to α at x. Moreover,*

$$\text{(A.3)} \qquad f_n(x_n) \to \alpha \qquad \forall x_n \to x$$

is equivalent to

$$\text{(A.4)} \qquad f_{n,\wedge}(x_n) \to \alpha \quad \text{and} \quad f_{n,\vee}(x_n) \to \alpha \qquad \forall x_n \to x.$$

The following three lemmas contain results on hulls of sums and on epi-, hypo- and hypi-convergence of sums.

LEMMA A.2 (On sums of hulls and hulls of sums). *For $f, g : \mathbb{T} \to \mathbb{R}$ such that $g_\wedge$ and $g_\vee$ are both finite, we have*

$$f_\wedge + g_\wedge \leq (f + g)_\wedge \leq f_\wedge + g_\vee,$$
$$f_\vee + g_\wedge \leq (f + g)_\vee \leq f_\vee + g_\vee.$$

In particular, if g is continuous in $x \in \mathbb{T}$, then $(f + g)_\wedge(x) = f_\wedge(x) + g(x)$ and $(f + g)_\vee(x) = f_\vee(x) + g(x)$.

LEMMA A.3 (Epi- and hypo-convergence of hulls of sums). *Let $f_n, g_n : \mathbb{T} \to \mathbb{R}$ and let $x \in \mathbb{T}$ be such that $g_n(x_n) \to \beta \in \mathbb{R}$ for all sequences $x_n \to x$. If $f_{n,\wedge}$ epi-converges to α at x, then $(f_n + g_n)_\wedge$ epi-converges to $\alpha + \beta$ at x. Similarly for upper semicontinuous hulls and hypo-convergence.*

LEMMA A.4 (Hypi-convergence of sums). *Let $\mathbb{T}$ be locally compact and separable. If f_n and g_n hypi-converge to f and g in $\ell^\infty_{\mathrm{loc}}(\mathbb{T})$, respectively, and if at every point $x \in \mathbb{T}$, at least one of the two functions f or g is continuous, then $f_n + g_n$ hypi-converges to $f + g$.*

A.2. Upper and lower semicontinuous extensions. The limit processes in Theorems 4.5 and 5.1 are defined by extending a continuous function defined on a dense subset of a metric space to the whole space. In this section, some useful elementary properties of such extensions are recorded. The main tool is Corollary A.7, giving a criterion for proving hypi-convergence to a function defined by such an extension procedure.

Let $(\mathbb{T}, d)$ be a metric space, let $A \subset \mathbb{T}$ be dense, and let $f : A \to \mathbb{R}$. Extend the domain of f from A to the whole of $\mathbb{T}$ by

$$f_\wedge^{A:\mathbb{T}}(x) = \sup_{\varepsilon>0} \inf f\big(B(x,\varepsilon) \cap A\big) \in [-\infty, \infty], \tag{A.5}$$

$$f_\vee^{A:\mathbb{T}}(x) = \inf_{\varepsilon>0} \sup f\big(B(x,\varepsilon) \cap A\big) \in [-\infty, \infty], \tag{A.6}$$

for $x \in \mathbb{T}$; as before, $B(x, \varepsilon) = \{y \in \mathbb{T} : d(x, y) < \varepsilon\}$ is the open ball centered at x of radius ε. Note that these definitions also make sense if $A = \mathbb{T}$, and that for $f \in \ell^\infty_{\mathrm{loc}}(\mathbb{T})$ we have $f_\wedge^{\mathbb{T}:\mathbb{T}} = f_\wedge$ and $f_\vee^{\mathbb{T}:\mathbb{T}} = f_\vee$; see the definitions in (2.1) and (2.2).

Clearly, $f_\wedge^{A:\mathbb{T}}(x) \le f(x) \le f_\vee^{A:\mathbb{T}}(x)$ for every $x \in A$. For any open set $U \subset \mathbb{T}$, we have

$$\inf f_\wedge^{A:\mathbb{T}}(U) = \inf f(U \cap A), \qquad \sup f_\vee^{A:\mathbb{T}}(U) = \sup f(U \cap A).$$

The functions $f_\wedge^{A:\mathbb{T}}$ and $f_\vee^{A:\mathbb{T}}$ from $\mathbb{T}$ into $[-\infty, +\infty]$ are lower and upper semicontinuous, respectively. If every x in A admits a neighborhood on which f is bounded, then $f_\wedge^{A:\mathbb{T}}$ and $f_\vee^{A:\mathbb{T}}$ are real-valued.

If f is continuous at $x \in A$, then $f_\wedge^{A:\mathbb{T}}(x) = f_\vee^{A:\mathbb{T}}(x) = f(x)$, and $f_\wedge^{A:\mathbb{T}}$ and $f_\vee^{A:\mathbb{T}}$, seen as functions on $\mathbb{T}$, are continuous at x, too. The following lemma shows that, if f is continuous on the whole of A, then its domain does not really matter insofar as the extension is concerned.

LEMMA A.5. *Let $E \subset A \subset \mathbb{T}$ be such that E is dense in $\mathbb{T}$. Let $f : A \to \mathbb{R}$ and consider the restriction $f|_E : E \to \mathbb{R}$ of f to E and the extensions $(f|_E)_\wedge^{E:\mathbb{T}}$ and $(f|_E)_\vee^{E:\mathbb{T}}$ of $f|_E$ to $\mathbb{T}$. If f is continuous, then $(f|_E)_\wedge^{E:\mathbb{T}} = f_\wedge^{A:\mathbb{T}}$ and $(f|_E)_\vee^{E:\mathbb{T}} = f_\vee^{A:\mathbb{T}}$.*

The following two results provide criterions for proving epi-, hypo- or hypi-convergence to a semicontinuous extension.

PROPOSITION A.6. *Let $A \subset \mathbb{T}$ be dense and let $f : A \to \mathbb{R}$ be continuous. Assume that $f_\wedge^{A:\mathbb{T}}$ is real-valued. If the functions $f_n : \mathbb{T} \to \mathbb{R}$ converge pointwise on A to f and if $\liminf_n f_n(x_n) \ge f_\wedge^{A:\mathbb{T}}(x)$ whenever $x_n \in \mathbb{T}$ converges to $x \in \mathbb{T}$, then f_n epi-converges to $f_\wedge^{A:\mathbb{T}}$. Similarly for hypo-convergence to $f_\vee^{A:\mathbb{T}}$.*

COROLLARY A.7. *Let $A \subset \mathbb{T}$ be dense. Let $f : A \to \mathbb{R}$ be continuous and suppose that its lower and upper semicontinuous extensions $f_\wedge^{A:\mathbb{T}}$ and $f_\vee^{A:\mathbb{T}}$ are real-valued. Let $f^* : \mathbb{T} \to \mathbb{R}$ be such that $f_\wedge^{A:\mathbb{T}} \le f^* \le f_\vee^{A:\mathbb{T}}$. Then $f_\wedge^{A:\mathbb{T}} = (f^*)_\wedge$ and $f_\vee^{A:\mathbb{T}} = (f^*)_\vee$ on $\mathbb{T}$. Moreover, if the functions $f_n : \mathbb{T} \to \mathbb{R}$ are locally bounded and verify*

$$\forall x \in \mathbb{T} : \forall x_n \to x : f_\wedge^{A:\mathbb{T}}(x) \le \liminf_{n\to\infty} f_n(x_n) \le \limsup_{n\to\infty} f_n(x_n) \le f_\vee^{A:\mathbb{T}}(x),$$

then f_n hypi-converges to f^*.

APPENDIX B: WEAK CONVERGENCE AND SEMIMETRIC SPACES

The workhorses of the theory of weak convergence in metric spaces are the continuous mapping theorem, the extended continuous mapping theorem, and for normed vector spaces, the functional delta method; see, for instance, Theorems 1.3.6, 1.11.1 and 3.9.4 in van der Vaart and Wellner (1996). Thanks to these theorems, weak convergence of many empirical processes can be shown and can be exploited to conclude weak convergence of sequences of appropriately normalized estimators and test statistics. However, the space of interest in this paper, $(\ell_{\mathrm{loc}}^\infty(\mathbb{T}), d_{\mathrm{hypi}})$; see Proposition 2.1, is not a metric space but rather a semimetric space. Moreover, addition of functions is ill-compatible with the hypi-semimetric: if $f \in \ell_{\mathrm{loc}}^\infty(\mathbb{T})$ is not continuous, then $d_{\mathrm{hypi}}(f+g, 0)$ need not be equal to zero even if $d_{\mathrm{hypi}}(g, -f) = 0$. Hence, addition is not well defined on the space of equivalence classes of functions at hypi-distance zero.

In this appendix, versions of the (extended) continuous mapping theorem and the functional delta method are given that are adapted to semimetric spaces. In particular, the maps under consideration are not required to be defined on equivalence classes of points at distance zero but rather on the original semimetric space itself. Proofs are deferred to Appendix E in the supplement [Bücher, Segers and Volgushev (2014)].

Let $(\mathbb{D}, d)$ be a semimetric space. For $x \in \mathbb{D}$, put $[x] = \mathrm{cl}\{x\}$, the set of $y \in \mathbb{D}$ such that $d(x, y) = 0$. Since $d(x', y') = d(x, y)$ whenever $x' \in [x]$ and $y' \in [y]$, we can, abusing notation, define a metric $d([x],[y]) := d(x, y)$ on the quotient space $[\mathbb{D}] = \{[x] : x \in \mathbb{D}\}$. Let $[\cdot]$ denote the map $\mathbb{D} \to [\mathbb{D}] : x \mapsto [x]$. Obviously, $[\cdot]$ is continuous. The image of an open (closed) subset of $\mathbb{D}$ under $[\cdot]$ is open (closed) in $[\mathbb{D}]$.

Let $\mathcal{B}(\mathbb{D})$ and $\mathcal{B}([\mathbb{D}])$ be the Borel σ-fields on $(\mathbb{D}, d)$ and $([\mathbb{D}], d)$, respectively, that is, the smallest σ-fields containing the open sets. There is a one-to-one correspondence between both σ-fields: for $B \in \mathcal{B}(\mathbb{D})$, the set $[B] = \{[x] : x \in B\}$ is a Borel set in $\mathbb{D}$, and conversely, every Borel set B of $\mathbb{D}$ can be written as $\bigcup_{[x]\in[B]}[x]$; in particular $x \in B$ if and only if $[x] \in [B]$. A Borel law L on $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ induces a Borel law $L \circ [\cdot]^{-1}$ on $([\mathbb{D}], \mathcal{B}([\mathbb{D}]))$ and vice versa.

One of the merits of Hoffman–Jørgensen weak convergence is that measurability requirements are relaxed. In the context of semimetric spaces, measurability issues require, perhaps, some extra care.

LEMMA B.1 (Measurability). *Let* $(\mathbb{D}, d)$ *and* $(\mathbb{E}, e)$ *be semimetric spaces. Let* $g : \mathbb{D} \to \mathbb{E}$ *be arbitrary. Then the set* D_g *of* $x \in \mathbb{D}$ *such that* g *is not continuous in* x *is Borel measurable. More generally,* $g^{-1}(B) \setminus D_g$ *is a Borel set in* $\mathbb{D}$ *for every Borel set* B *in* $\mathbb{E}$.

In our version of the continuous mapping theorem, the map g is defined on $\mathbb{D}$ and not on $[\mathbb{D}]$, that is, even if $d(x, y) = 0$, it may occur that $g(x) \neq g(y)$. Therefore, we cannot directly apply Theorem 1.3.6 in van der Vaart and Wellner (1996). Nevertheless, the proof is inspired from the proof of that theorem.

THEOREM B.2 (Continuous mapping). *Let* $(\mathbb{D}, d)$ *be a semimetric space and let* $(\mathbb{E}, e)$ *be a metric space. Let* $g : \mathbb{D} \to \mathbb{E}$ *be arbitrary and let* D_g *be the set of* $x \in \mathbb{D}$ *such that* g *is not continuous in* x. *Let* $(\Omega_\alpha, \mathcal{A}_\alpha, P_\alpha)$, $\alpha \in A$, *be a net of probability spaces and let* $X_\alpha : \Omega_\alpha \to \mathbb{D}$ *be arbitrary maps; let* $(\Omega, \mathcal{A}, P)$ *be a probability space and let* $X : \Omega \to \mathbb{D}$ *be Borel measurable. If* $[X_\alpha] \rightsquigarrow [X]$ *in* $[\mathbb{D}]$ *and if* $X(\Omega) \subset \mathbb{D} \setminus D_g$, *then* $g(X) : \Omega \to \mathbb{E}$ *is Borel measurable and* $g(X_\alpha) \rightsquigarrow g(X)$ *in* $\mathbb{E}$.

In many circumstances, one needs a refined version of the continuous mapping theorem that covers maps $g_n(X_n)$, rather than $g(X_n)$ for a fixed g. The following statement and proof are inspired from Theorem 1.11.1(i) and Problem 1.11.1 in van der Vaart and Wellner (1996) and Theorem 18.11(i) in van der Vaart (1998).

THEOREM B.3 (Extended continuous mapping). *Let* $(\mathbb{D}, d)$ *be a semimetric space and let* $(\mathbb{E}, e)$ *be a metric space. For integer* $n \geq 0$, *let there be probability spaces* Ω_n, *subsets* $\mathbb{D}_n \subset \mathbb{D}$, *maps* $X_n : \Omega_n \to \mathbb{D}_n$ *and maps* $g_n : \mathbb{D}_n \to \mathbb{E}$. *Assume the following two conditions*:

- *For every* $x_0 \in \mathbb{D}_0$ *and for every subsequence* $(x_{n_k})_k$ *with* $x_{n_k} \in \mathbb{D}_{n_k}$ *for all* k *and such that* $x_{n_k} \to x_0$ *as* $k \to \infty$, *we have* $g_{n_k}(x_{n_k}) \to g_0(x_0)$.
- *The map* X_0 *is Borel measurable and* $[X_n] \rightsquigarrow [X_0]$ *in* $([\mathbb{D}], d)$.

If $g_0(X_0)$ *is Borel measurable, then* $g_n(X_n) \rightsquigarrow g_0(X_0)$ *in* $(\mathbb{E}, e)$. *If* $g_0(X_0)$ *is not Borel measurable, there still exists a version* X_0' *of* X_0 *such that* $g_0(X_0')$ *is Borel measurable, and thus* $g_n(X_n) \rightsquigarrow g_0(X_0')$ *in* $(\mathbb{E}, e)$.

ADDENDUM B.4. *The law of* X_0 *is concentrated on the set*

$$\mathbb{D}_\infty = \bigcap_{k \geq 1} \operatorname{cl}\Bigl(\bigcup_{m \geq k} \mathbb{D}_m\Bigr) = \limsup_{n \to \infty} \mathbb{D}_n,$$

which is closed in $\mathbb{D}$. *The restriction of the map* g_0 *to* $\mathbb{D}_0 \cap \mathbb{D}_\infty$ *is continuous. Whether* $g_0(X_0)$ *is measurable or not, there always exists a version* X_0' *of* X_0 *which takes values in* $\mathbb{D}_0 \cap \mathbb{D}_\infty$ *and for which* $g_0(X_0')$ *is Borel measurable.*

COROLLARY B.5. *If in Theorem* B.3, $(\mathbb{E}, e)$ *is a semimetric space rather than a metric space, the conclusion still holds with* $[g_n(X_n)]$ *converging weakly to* $[g_0(X_0)]$ *or* $[g_0(X_0')]$, *respectively, in* $([\mathbb{E}], e)$.

The formulation of Theorem B.3 has been chosen to make it suitable for establishing a variant of the functional delta method in semimetric vector spaces. In the remaining part of this section, let $\mathbb{D}$ and $\mathbb{E}$ be real vector spaces and let d and e be semimetrics on $\mathbb{D}$ and $\mathbb{E}$, respectively. Addition is not required to be continuous. Worse still, addition need not even be well defined on equivalence classes.

DEFINITION B.6 (Semi-Hadamard differentiability). Let $\Psi : \mathbb{D}_\Psi \to \mathbb{E}$ with $\mathbb{D}_\Psi \subset \mathbb{D}$. Let $x \in \mathbb{D}_\Psi$ and $\mathbb{W} \subset \mathbb{D}$. Then Ψ is said to be semi-Hadamard differentiable at x tangentially to $\mathbb{W}$ if there exists a map $d\Psi_x : \mathbb{W} \to \mathbb{E}$, called the (semi-)derivative of Ψ at x, with the following property: for every $w \in \mathbb{W}$, every sequence $(t_n)_n$ in $(0, \infty)$ such that $t_n \to 0$ and every sequence $(w_n)_n$ in $\mathbb{D}$ such that $x + t_n w_n \in \mathbb{D}_\Psi$ for all n and $w_n \to w$ as $n \to \infty$, we have

$$t_n^{-1}\bigl(\Psi(x + t_n w_n) - \Psi(x)\bigr) \to d\Psi_x(w), \qquad n \to \infty.$$

The derivative $d\Psi_x$ is not assumed to be linear or continuous. Still, in Addendum B.8 below, we will see that $d\Psi_x$ does enjoy some kind of continuity property. An extension of the chain rule similar to Lemma 3.9.3 in van der Vaart and Wellner (1996) is straightforward and is therefore omitted.

THEOREM B.7. *Let* $\mathbb{D}_\Psi \subset \mathbb{D}$ *and let* $\Psi : \mathbb{D}_\Psi \to \mathbb{E}$ *be semi-Hadamard differentiable at* $x \in \mathbb{D}_\Psi$ *tangentially to* $\mathbb{W} \subset \mathbb{D}$ *with derivative* $d\Psi_x$. *Let* Y_n, $n \geq 1$, *and* X *be maps from probability spaces into* $\mathbb{D}$ *such that* Y_n *takes values in* $\mathbb{D}_\Psi$ *and* X *takes values in* $\mathbb{W}$. *Assume that* X *is Borel measurable and that there exists a positive sequence* r_n *tending to infinity such that, in* $([\mathbb{D}], d)$,

$$\bigl[r_n(Y_n - x)\bigr] \rightsquigarrow [X], \qquad n \to \infty.$$

Passing to a suitable version of X *if necessary to ensure measurability of* $d\Psi_x(X)$, *we then have, in* $([\mathbb{E}], e)$,

$$\bigl[r_n\bigl(\Psi(Y_n) - \Psi(x)\bigr)\bigr] \rightsquigarrow \bigl[d\Psi_x(X)\bigr], \qquad n \to \infty.$$

ADDENDUM B.8. *There exists a subset* $\mathbb{W}_\infty$ *of* $\mathbb{W}$ *and a version* X' *of* X *such that the restriction of* $d\Psi_x$ *to* $\mathbb{W}_\infty$ *is continuous,* X' *takes values in* $\mathbb{W}_\infty$ *only, and* $d\Psi_x(X')$ *is Borel measurable.*

APPENDIX C: PROOFS

This appendix contains the most important proofs for the main part of the paper, namely those for Sections 4 and 6. The remaining proofs are collected in Appendix F of the supplement [Bücher, Segers and Volgushev (2014)].

C.1. Proofs for Section 4.

PROOF OF THEOREM 4.5. For the sake of a clear exposition, we split the proof into two propositions. First, Proposition C.2 shows that Condition 4.3 implies a certain abstract hypi-differentiability property stated in Condition C.1. Then, Proposition C.3 shows that the latter condition suffices to obtain the completion of Theorem 4.5. □

CONDITION C.1 (Hypi-differentiability of C). Define the set

$$\mathcal{W}(t) := \{a \in \{\ell^\infty([0,1])\}^d : \mathbf{u} + ta(\mathbf{u}) \in [0,1]^d \ \forall \mathbf{u} \in [0,1]^d\},$$

where $a(\mathbf{u}) = (a_1(u_1), \dots, a_d(u_d))$. Whenever $t_n \searrow 0$, $t_n \neq 0$, and $a_n = (a_{n1}, \dots, a_{nd}) \in \{\ell^\infty([0,1])\}^d$ converges uniformly to $a \in \mathcal{W} := \{\mathcal{C}([0,1])\}^d$ (i.e., $\|a_{nj} - a_j\|_\infty \to 0$ for all $j = 1, \dots, d$) such that $a_n \in \mathcal{W}(t_n)$ for all $n \in \mathbb{N}$, the functions

$$[0,1]^d \to \mathbb{R} : u \mapsto t_n^{-1}\{C(\mathbf{u} + t_n a_n(\mathbf{u})) - C(\mathbf{u})\}$$

converge in $(\ell^\infty([0,1]^d), d_{\mathrm{hypi}})$ to some limit dC_a.

PROPOSITION C.2. *A copula C satisfying Condition 4.3 also satisfies the hypi-differentiability Condition C.1 with derivative*

$$dC_a(\mathbf{u}) = \sup_{\varepsilon > 0} \inf \left\{ \sum_{j=1}^d \dot{C}_j(\mathbf{v}) a_j(v_j) : \mathbf{v} \in \mathcal{S}, |\mathbf{v} - \mathbf{u}| < \varepsilon \right\}, \qquad \mathbf{u} \in [0,1]^d.$$

Conversely, it is an open problem whether there exists a copula that satisfies Condition C.1 but violates Condition 4.3. According to the next proposition, Condition C.1 can replace Condition 4.3 in Theorem 4.5.

PROPOSITION C.3. *Suppose that Condition 4.4 holds and that C satisfies Condition C.1. Then, in $(L^\infty([0,1]^d), d_{\mathrm{hypi}})$,*

$$\mathbb{C}_n \rightsquigarrow \mathbb{C} = \alpha + dC_{(-\alpha_1, \dots, -\alpha_d)}.$$

PROOF OF PROPOSITION C.2. Let $t_n \searrow 0$ and let $a_n \in \mathcal{W}(t_n)$ converge uniformly to $a \in \mathcal{W}$. As in Condition C.1, we use the notation $a_n(\mathbf{u}) = (a_{n1}(u_1), \dots, a_{nd}(u_d))$. We have to prove epi- and hypo-convergence of

$$\mathbf{u} \mapsto F_n(\mathbf{u}) := t_n^{-1}\{C(\mathbf{u} + t_n a_n(\mathbf{u})) - C(\mathbf{u})\}$$

to $F_\wedge$ and $F_\vee$, respectively, where $F = dC_a$. Note that, in the notation of Appendix A.2, we have $F = G_\wedge^{\mathcal{S}:[0,1]^d}$, where $G : \mathcal{S} \to \mathbb{R}$ is defined through

$$G(\mathbf{u}) = \sum_{j=1}^d \dot{C}_j(\mathbf{u}) a_j(u_j).$$

By an application of Corollary A.7, and since $F_\wedge = G_\wedge^{\mathcal{S}:[0,1]^d}$ and $F_\vee = G_\vee^{\mathcal{S}:[0,1]^d}$, it suffices to show that:

(i) $\forall \mathbf{u} \in [0,1]^d : \forall \mathbf{u}_n \to \mathbf{u} : \liminf_{n\to\infty} F_n(\mathbf{u}_n) \geq F_\wedge(\mathbf{u})$,
(ii) $\forall \mathbf{u} \in [0,1]^d : \forall \mathbf{u}_n \to \mathbf{u} : \limsup_{n\to\infty} F_n(\mathbf{u}_n) \leq F_\vee(\mathbf{u})$.

We begin with the proof of (i) and fix a point $\mathbf{u} \in [0,1]^d$ and a sequence $\mathbf{u}_n \to \mathbf{u}$. Choose $\varepsilon > 0$ and let $|\cdot|_1$ denote the L_1-norm on $\mathbb{R}^d$. Due to Lemma C.4, we may choose

$$\mathbf{u}_n^\star \in \{\mathbf{v} \in [0,1]^d : |\mathbf{u}_n - \mathbf{v}|_1 \leq \varepsilon t_n/2\}$$

and

$$\mathbf{u}_n^\circ \in \{\mathbf{v} \in [0,1]^d : |\mathbf{u}_n + t_n a_n(\mathbf{u}_n) - \mathbf{v}|_1 \leq \varepsilon t_n/2\}$$

such that, for the path

$$\gamma_n(s) = (1-s)\mathbf{u}_n^\star + s\mathbf{u}_n^\circ, \qquad s \in [0,1],$$

the set $\{s \in [0,1] : \gamma_n(s) \notin \mathcal{S}\}$ has Lebesgue-measure zero. Define $f_n(s) = t_n^{-1} C(\gamma_n(s))$, $s \in [0,1]$, and note that

$$\begin{aligned} |\{f_n(1) - f_n(0)\} - F_n(\mathbf{u}_n)| &= t_n^{-1} |C(\mathbf{u}_n^\circ) - C(\mathbf{u}_n + t_n a_n(\mathbf{u}_n)) - C(\mathbf{u}_n^\star) + C(\mathbf{u}_n)| \\ &\leq t_n^{-1}\{|\mathbf{u}_n + t_n a_n(\mathbf{u}_n) - \mathbf{u}_n^\circ|_1 + |\mathbf{u}_n^\star - \mathbf{u}_n|_1\} \leq \varepsilon \end{aligned}$$

by Lipschitz-continuity of C. Lipschitz-continuity of C also implies absolute continuity of f_n, which allows us to choose $\mathbf{v}_n \in \gamma_n([0,1]) \cap \mathcal{S}$ such that

$$\begin{aligned} \varepsilon + F_n(u_n) &\geq f_n(1) - f_n(0) \\ &= \int_0^1 f_n'(s)\,ds \\ &= \sum_{j=1}^d t_n^{-1}(u_{nj}^\circ - u_{nj}^\star) \int_0^1 \dot{C}_j(\gamma_n(s))\,ds \\ &= \sum_{j=1}^d [a_{nj}(u_{nj}) + t_n^{-1}\{u_{nj}^\circ - u_{nj}^\star - t_n a_{nj}(u_{nj})\}] \int_0^1 \dot{C}_j(\gamma_n(s))\,ds \\ &\geq \inf_{s:\gamma_n(s)\in\mathcal{S}} \sum_{j=1}^d a_{nj}(u_{nj}) \dot{C}_j(\gamma_n(s)) - \varepsilon \end{aligned}$$

$$\geq \sum_{j=1}^{d} a_{nj}(v_{nj})\dot{C}_j(\mathbf{v}_n) + \sum_{j=1}^{d} \{a_{nj}(u_{nj}) - a_{nj}(v_{nj})\}\dot{C}_j(\mathbf{v}_n) - 2\varepsilon$$

$$\geq \sum_{j=1}^{d} a_{nj}(v_{nj})\dot{C}_j(\mathbf{v}_n) - 3\varepsilon = F(\mathbf{v}_n) - 3\varepsilon = F_\wedge(\mathbf{v}_n) - 3\varepsilon$$

for sufficiently large n, where we have used the bounds $0 \leq \dot{C}_j \leq 1$, uniform convergence of a_{nj} to a_j, uniform continuity of a_j and the fact that F is continuous in $\mathbf{v}_n$. Hence, by lower semicontinuity of $F_\wedge$,

$$\liminf_{n\to\infty} F_n(\mathbf{u}_n) \geq F_\wedge(\mathbf{u}) - 4\varepsilon.$$

As $\varepsilon > 0$ was arbitrary the assertion in (i) follows.

The proof of (ii) is analogous. In the main inequality chain, all signs can be reversed if the infimum is replaced by a supremum and upon noting that on $\mathcal{S}$, the functions F, $F_\wedge$ and $F_\vee$ are equal. □

PROOF OF PROPOSITION C.3. Recall $\beta_n = (\beta_{n1}, \dots, \beta_{nj})$, with $\beta_{nj} = \sqrt{n}(G_{nj}^{-} - \mathrm{id}_{[0,1]})$. It follows from Condition 4.4 and the functional delta method for the inverse mapping, also known as Vervaat's lemma, that

$$(\alpha_n, \beta_n) = (\alpha_n, \beta_{n1}, \dots, \beta_{nd}) \rightsquigarrow (\alpha, -\alpha_1, \dots, -\alpha_d)$$

in $\ell^\infty([0,1]^d) \times \{\ell^\infty([0,1])\}^d$, with respect to the supremum distance in each coordinate. Note that we can write $\mathbb{C}_n = g_n(\alpha_n, \beta_n)$, where $g_n : \ell^\infty([0,1]^d) \times \mathcal{W}(1/\sqrt{n}) \to (L^\infty([0,1]^d), d_{\mathrm{hypi}})$ is defined as

$$g_n(a,b) = a(\mathrm{id}_{[0,1]^d} + b/\sqrt{n}) + \sqrt{n}\{C(\mathrm{id}_{[0,1]^d} + b/\sqrt{n}) - C\}. \tag{C.1}$$

Exploiting Condition C.1 and Lemma A.4 (recall that α is continuous almost surely), the assertion follows from the extended continuous mapping theorem, see Theorem 1.11.1 in van der Vaart and Wellner (1996). □

LEMMA C.4. *Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$ be two distinct points and denote by $H_\mathbf{u}$ and $H_\mathbf{v}$ the hyperplanes being orthogonal to $\mathbf{u} - \mathbf{v}$ and passing through $\mathbf{u}$ and $\mathbf{v}$, respectively. For $\delta > 0$, set $H_\mathbf{u}^\delta = H_\mathbf{u} \cap B_1(\mathbf{u}, \delta)$ and $H_\mathbf{v}^\delta = H_\mathbf{v} \cap B_1(\mathbf{v}, \delta)$, where $B_1(\mathbf{u}, \delta)$ denotes the unit ball of radius δ centered at $\mathbf{u}$ with respect to the $\|\cdot\|_1$-norm. Finally, let $\mathcal{Z}$ denote the cylinder with top area equal to $H_\mathbf{u}^\delta$ and bottom area equal to $H_\mathbf{v}^\delta$, that is,*

$$\mathcal{Z} = \{\mathbf{y} + s(\mathbf{v} - \mathbf{u}) : \mathbf{y} \in H_\mathbf{u}^\delta, s \in [0,1]\}.$$

Let $\mathcal{D}$ be a Lebesgue-null set in $\mathbb{R}^d$ and define, for any $\mathbf{y} \in H_\mathbf{u}^\delta$,

$$\mathcal{Z}_\mathbf{y}^\mathcal{D} = \{s \in \mathbb{R} : \mathbf{y} + s(\mathbf{v} - \mathbf{u}) \in \mathcal{Z} \cap \mathcal{D}\}.$$

Then $\mathcal{Z}_\mathbf{y}^\mathcal{D}$ is a one-dimensional Lebesgue-null set for almost all $\mathbf{y} \in H_\mathbf{u}^\delta$.

The proofs of Lemma C.4, Propositions 4.6 and 4.7 are given in Appendix F.3 in the supplement [Bücher, Segers and Volgushev (2014)].

C.2. Proofs for Section 6.

PROOF OF THEOREM 6.1. The proof consists of two main steps. In the first step, consider $\ell^\infty(\bar{\mathbb{R}}) \times \mathbb{R}^p$ equipped with the metric

$$\rho\big((h_1, \mathbf{y}_1), (h_2, \mathbf{y}_2)\big) = \|h_1 - h_2\|_\infty + |\mathbf{y}_1 - \mathbf{y}_2|.$$

As shown in Appendix F.5 in the supplement [Bücher, Segers and Volgushev (2014)], we have, in $(\ell^\infty(\bar{\mathbb{R}}) \times \mathbb{R}^p, \rho)$, as $n \to \infty$,

$$\big(\mathbb{G}_n f_{\cdot,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}}, \sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})\big) = (\mathbb{G}_n f_{\cdot,\mathbf{0}}, \mathbb{G}_n \boldsymbol{\psi}) + o_p(1) \rightsquigarrow (\mathbb{G} f_{\cdot,\mathbf{0}}, \mathbb{G}\boldsymbol{\psi}) \tag{C.2}$$

where $\mathbb{G}$ denotes a zero-mean Gaussian process on $\mathcal{G} = \mathcal{F} \cup \{\psi_1, \dots, \psi_p\}$ with covariance given in (6.10). Define $T_n : \ell^\infty(\bar{\mathbb{R}}) \times \mathbb{R}^p \to \ell^\infty(\bar{\mathbb{R}})$ by

$$T_n(G, \boldsymbol{\gamma}) = G + g_n(\boldsymbol{\gamma}),$$

where the map $g_n(\gamma) \in \ell^\infty(\bar{\mathbb{R}})$ is defined by $(g_n(\gamma))(\pm\infty) = 0$ and

$$\big(g_n(\boldsymbol{\gamma})\big)(z) = t_n^{-1} \int_{\mathbb{R}^p} \big\{F\big(z + t_n \mathbf{x}'\boldsymbol{\gamma}\big) - F(z)\big\} P^X(d\mathbf{x}), \qquad z \in \mathbb{R}.$$

Note that we can write the second term in (6.8) as

$$\sqrt{n}\{P f_{z,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}} - P f_{z,\mathbf{0}}\} = \big(g_n\big(\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})\big)\big)(z) \tag{C.3}$$

with $t_n = 1/\sqrt{n}$. This also allows to write

$$\mathbb{F}_n = T_n\big(\mathbb{G}_n f_{\cdot,\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}}, \sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})\big).$$

The assertion of Theorem 6.1 will then follow by an application of the extended continuous mapping theorem. More precisely, if $G_n, G \in \ell^\infty(\bar{\mathbb{R}})$ are such that G is continuous and $\|G_n - G\|_\infty \to 0$, and if moreover $\boldsymbol{\gamma}_n \to \boldsymbol{\gamma}$ in $\mathbb{R}^p$, then, in $(\ell^\infty(\bar{\mathbb{R}}), d_{\mathrm{hypi}})$,

$$T_n(G_n, \boldsymbol{\gamma}_n) \to T(G, \boldsymbol{\gamma}) := G + g(\boldsymbol{\gamma}), \tag{C.4}$$

by Lemma C.7 below and Lemma A.4 on weak hypi-convergence of sums. Here, the map $g(\boldsymbol{\gamma}) \in \ell^\infty(\bar{\mathbb{R}})$ is defined by $g(\boldsymbol{\gamma})(\pm\infty) = 0$ and, for $z \in \mathbb{R}$,

$$\begin{aligned}\big(g(\boldsymbol{\gamma})\big)(z) = {}& -f(z-) \int_{-\infty}^{0} \mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} < y\big)\, dy \\ & + f(z+) \int_{0}^{+\infty} \mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} > y\big)\, dy.\end{aligned} \tag{C.5}$$

Note that the integrals on the right-hand side of the last display exist as a consequence of condition (R3) and Fubini's theorem, which implies that

$$\int_0^{+\infty} \mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} > y\big)\, dy = \mathbb{E}\big[\max(\mathbf{X}'\boldsymbol{\gamma}, 0)\big] < \infty, \tag{C.6}$$

$$\int_{-\infty}^{0} \mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} < y\big)\, dy = \mathbb{E}\big[\max(-\mathbf{X}'\boldsymbol{\gamma}, 0)\big] < \infty. \tag{C.7}$$

Finally, as a consequence of (C.4) and since $\mathbb{G}f_{\cdot,\mathbf{0}}$ is continuous almost surely by Lemma F.5 in the supplementary material, the assertion follows from (C.2) and an application of the extended continuous mapping theorem [van der Vaart and Wellner (1996), Theorem 1.11.1]. □

The preceding proof made use of Lemma C.7 below. For its formulation, we need two additional lemmas. The proof of the first one is trivial and, therefore, omitted.

LEMMA C.5. *If f is làdlàg, then both functions $z \mapsto f(z+)$ and $z \mapsto f(z-)$ defined in* (R2) *of Theorem* 6.1 *are làdlàg, too. Their right-hand limits at z are both equal to $f(z+)$ and their left-hand limits at z are both equal to $f(z-)$.*

LEMMA C.6. *If conditions* (R2) *and* (R3) *hold, then for every $\gamma \in \mathbb{R}^p$, the function $g(\boldsymbol{\gamma})$ in* (C.5) *is uniformly bounded and làdlàg, with right- and left-hand limits at $z \in \mathbb{R}$ given by*

$$\big(g(\boldsymbol{\gamma})\big)(z\pm) = f(z\pm)\mathbb{E}\big[\mathbf{X}'\boldsymbol{\gamma}\big]. \tag{C.8}$$

The upper and lower semicontinuous hulls of $g(\boldsymbol{\gamma})$ at $z \in \mathbb{R}$ are

$$\big(g(\boldsymbol{\gamma})\big)_\vee(z) = \max\big\{\big(g(\boldsymbol{\gamma})\big)(z-), \big(g(\boldsymbol{\gamma})\big)(z), \big(g(\boldsymbol{\gamma})\big)(z+)\big\}, \tag{C.9}$$

$$\big(g(\boldsymbol{\gamma})\big)_\wedge(z) = \min\big\{\big(g(\boldsymbol{\gamma})\big)(z-), \big(g(\boldsymbol{\gamma})\big)(z), \big(g(\boldsymbol{\gamma})\big)(z+)\big\}. \tag{C.10}$$

Moreover, $(g(\boldsymbol{\gamma}))_\wedge(\pm\infty) = (g(\boldsymbol{\gamma}))_\vee(\pm\infty) = 0$.

PROOF. The existence and the expressions of the right-hand and left-hand limits of $g(\boldsymbol{\gamma})$ at $z \in \mathbb{R}$ are a consequence of Lemma C.5 and the fact that

$$-\int_{-\infty}^{0} \mathbb{P}(\mathbf{X}'\boldsymbol{\gamma} < y)\,dy + \int_{0}^{+\infty} \mathbb{P}(\mathbf{X}'\boldsymbol{\gamma} > y)\,dy = \mathbb{E}\big[\mathbf{X}'\boldsymbol{\gamma}\big],$$

which follows in turn from (C.6) and (C.7). The statement about the upper (lower) semicontinuous hull follows from the fact that for a làdlàg function, the supremum (infimum) over a shrinking neighborhood around a point converges to the maximum (minimum) of the function value at the point itself and the right-hand and left-hand limits at that point. □

LEMMA C.7. *Assume conditions* (R2) *and* (R3) *in Theorem* 6.1. *If $\boldsymbol{\gamma}_n \to \boldsymbol{\gamma}$ in $\mathbb{R}^p$, then*

$$d_{\mathrm{hypi}}\big(g_n(\boldsymbol{\gamma}_n), g(\boldsymbol{\gamma})\big) \to 0, \qquad n \to \infty.$$

PROOF. First of all, for $z \in \mathbb{R}$, we can write $(g_n(\boldsymbol{\gamma}))(z)$ as

$$\int_{\mathbb{R}^p} t_n^{-1} \int_0^{t_n \mathbf{x}'\boldsymbol{\gamma}} f(z+y)\,dy\,P^X(d\mathbf{x}) = \int_{\mathbb{R}^p} \int_0^{\mathbf{x}'\boldsymbol{\gamma}} f(z+t_n y)\,dy\,P^X(d\mathbf{x}).$$

It follows that $g_n(\boldsymbol{\gamma}_n)$ is uniformly close to $g_n(\boldsymbol{\gamma})$: we have

$$\big|\big(g_n(\boldsymbol{\gamma}_n)\big)(z) - \big(g_n(\boldsymbol{\gamma})\big)(z)\big| \le \|f\|_\infty \int_{\mathbb{R}^p} |\mathbf{x}| P^X(d\mathbf{x}) |\boldsymbol{\gamma}_n - \boldsymbol{\gamma}|, \qquad z \in \mathbb{R},$$

and thus, noting that $(g_n(\boldsymbol{\gamma}_n))(\pm\infty) = 0 = (g_n(\boldsymbol{\gamma}))(\pm\infty)$,

$$\big\|g_n(\boldsymbol{\gamma}_n) - g_n(\boldsymbol{\gamma})\big\|_\infty \to 0, \qquad n \to \infty. \tag{C.11}$$

Hence, without loss of generality, we can assume that $\boldsymbol{\gamma}_n = \boldsymbol{\gamma}$ for all n.

Fix $z \in \bar{\mathbb{R}}$. We will prove hypi-convergence of $g_n(\boldsymbol{\gamma})$ to $g(\boldsymbol{\gamma})$ by using the pointwise criteria in (A.1) and (A.2). First, consider the case $z_n \to z = +\infty$. Observe that for any fixed $\mathbf{x} \in \mathbb{R}^p$

$$t_n^{-1}\big|F\big(z_n + t_n \mathbf{x}'\boldsymbol{\gamma}\big) - F(z_n)\big| \le \big|\mathbf{x}'\boldsymbol{\gamma}\big| \sup_{y \ge z_n - t_n |\mathbf{x}'\boldsymbol{\gamma}|} f(y) \to 0$$

since $\lim_{z\to+\infty} f(z) = 0$ by assumption. Hence, by dominated convergence, $(g_n(\boldsymbol{\gamma}))(z_n) \to 0$, which is equal to $(g(\boldsymbol{\gamma}))_\wedge(+\infty) = (g(\boldsymbol{\gamma}))_\vee(+\infty)$ in view of Lemma C.6. The limit $z_n \to -\infty$ can be handled similarly.

It thus remains to consider $z_n \to z \in \mathbb{R}$. By Fubini's theorem, we have

$$\begin{aligned}
\big(g_n(\boldsymbol{\gamma})\big)(z_n) &= -\int\int_{\mathbf{x}'\boldsymbol{\gamma} < y < 0} f(z_n + t_n y)\, dy\, P^X(d\mathbf{x}) \\
&\quad + \int\int_{\mathbf{x}'\boldsymbol{\gamma} > y > 0} f(z_n + t_n y)\, dy\, P^X(d\mathbf{x}) \\
&= -\int_{-\infty}^0 f(z_n + t_n y)\mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} < y\big)\, dy \\
&\quad + \int_0^{+\infty} f(z_n + t_n y)\mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} > y\big)\, dy.
\end{aligned}$$

The idea is now to replace $f(z_n + t_n y)$ by $f(z-)$ or $f(z+)$ according to whether $z_n + t_n y$ is smaller or larger than z. To this end, define the auxiliary functions

$$w(y) = \begin{cases} -\mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} < y\big), & \text{if } y < 0, \\ \mathbb{P}\big(\mathbf{X}'\boldsymbol{\gamma} > y\big), & \text{if } y > 0; \end{cases}$$

$$\eta(a) = f(z-) \int_{-\infty}^a w(y)\, dy + f(z+) \int_a^{+\infty} w(y)\, dy, \qquad a \in \mathbb{R}.$$

Further, put $a_n = (z - z_n)/t_n$ and observe that $z_n + t_n y < z$ if $y < a_n$ while $z_n + t_n y > z$ if $y > a_n$. We have

$$\big(g_n(\boldsymbol{\gamma})\big)(z_n) = \int_{-\infty}^{\infty} f(z_n + t_n y) w(y)\, dy.$$

By the dominated convergence theorem, as $\int_{\mathbb{R}} |w(y)|\,dy = \mathbb{E}[|\mathbf{X}'\boldsymbol{\gamma}|] < \infty$,

$$\begin{aligned}
&\big(g_n(\boldsymbol{\gamma})\big)(z_n) - \eta(a_n) \\
&\quad = \int_{-\infty}^{+\infty} \{f(z_n + t_n y) - f(z-)\}\mathbb{1}(y < a_n) w(y)\,dy \\
&\qquad + \int_{-\infty}^{+\infty} \{f(z_n + t_n y) - f(z+)\}\mathbb{1}(y > a_n) w(y)\,dy = o(1).
\end{aligned} \tag{C.12}$$

Consider the extrema of the function η. The function η can be written as

$$\eta(a) = f(z+) \int_{-\infty}^{+\infty} w(y)\,dy + \{f(z-) - f(z+)\} \int_{-\infty}^{a} w(y)\,dy.$$

It follows that η is absolutely continuous with Radon–Nikodym derivative $\dot{\eta}(a) = \{f(z-) - f(z+)\}w(a)$. Since $w(y) \le 0$ for $y < 0$ and $w(y) \ge 0$ for $y > 0$, we find that η is monotone on $(-\infty, 0)$ and on $(0, \infty)$. Hence, η attains its extrema at either $a \to -\infty$, $a = 0$, or $a \to +\infty$. But for $a \to \pm\infty$, we find from (C.8) that

$$\eta(\mp\infty) = f(z\pm) \int_{-\infty}^{+\infty} w(y)\,dy = f(z\pm)\mathbb{E}\big[\mathbf{X}'\boldsymbol{\gamma}\big] = \big(g(\boldsymbol{\gamma})\big)(z\pm),$$

while for $a = 0$, we find

$$\eta(0) = f(z-) \int_{-\infty}^{0} w(y)\,dy + f(z+) \int_{0}^{+\infty} w(y)\,dy = \big(g(\boldsymbol{\gamma})\big)(z).$$

As a consequence, using (C.9),

$$\begin{aligned}
\sup_{a\in\mathbb{R}} \eta(a) &= \max\{\eta(-\infty), \eta(0), \eta(+\infty)\} \\
&= \max\{\big(g(\boldsymbol{\gamma})\big)(z-), \big(g(\boldsymbol{\gamma})\big)(z), \big(g(\boldsymbol{\gamma})\big)(z+)\} = \big(g(\boldsymbol{\gamma})\big)_{\vee}(z),
\end{aligned}$$

and similarly, by (C.10), $\inf_{a\in\mathbb{R}} \eta(a) = (g(\boldsymbol{\gamma}))_{\wedge}(z)$. In combination with (C.12), we obtain that

$$\begin{aligned}
\big(g(\boldsymbol{\gamma})\big)_{\wedge}(z) &= \inf_{a\in\mathbb{R}} \eta(a) \\
&\overset{(I1)}{\le} \liminf_{n\to\infty} \big(g_n(\boldsymbol{\gamma})\big)(z_n) \le \limsup_{n\to\infty} \big(g_n(\boldsymbol{\gamma})\big)(z_n) \\
&\overset{(I2)}{\le} \sup_{a\in\mathbb{R}} \eta(a) = \big(g(\boldsymbol{\gamma})\big)_{\vee}(z).
\end{aligned}$$

Moreover, the inequalities $(I1)$ and $(I2)$ in the above display become equalities if we choose z_n in such a way that $a_n = (z - z_n)/t_n$ converges to $-\infty$, 0, or ∞, according to where the infinimum and supremum of η are attained.

The above paragraph shows that $g_n(\boldsymbol{\gamma})$ epi-converges to $(g(\boldsymbol{\gamma}))_{\wedge}$ and hypo-converges to $(g(\boldsymbol{\gamma}))_{\vee}$ pointwise at every $z \in \mathbb{R}$. As a consequence, $g_n(\boldsymbol{\gamma})$ hypi-converges to $g(\boldsymbol{\gamma})$. This completes the proof. □

Acknowledgments. The authors would like to thank the Associate Editor and two referees for their constructive comments on an earlier version of this manuscript.

Parts of this paper were written when A. Bücher was a post-doctoral researcher at Université catholique de Louvain, Belgium.

SUPPLEMENTARY MATERIAL

Supplement to: "When uniform weak convergence fails: Empirical processes for dependence functions and residuals via epi- and hypographs" (DOI: 10.1214/14-AOS1237SUPP; .pdf). In the supplement, missing proofs for the results in this paper are given.

REFERENCES

ATTOUCH, H. and WETS, R. J.-B. (1983). A convergence theory for saddle functions. *Trans. Amer. Math. Soc.* **280** 1–41. MR0712247

BASS, R. F. and PYKE, R. (1985). The space $D(A)$ and weak convergence for set-indexed processes. *Ann. Probab.* **13** 860–884. MR0799425

BEER, G. (1993). *Topologies on Closed and Closed Convex Sets. Mathematics and Its Applications* **268**. Kluwer Academic, Dordrecht. MR1269778

BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1670. MR0383482

BÜCHER, A. and DETTE, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statist. Probab. Lett.* **80** 1925–1932. MR2734261

BÜCHER, A. and DETTE, H. (2013). Multiplier bootstrap of tail copulas with applications. *Bernoulli* **19** 1655–1687. MR3129029

BÜCHER, A., SEGERS, J. and VOLGUSHEV, S. (2014). Supplement to "When uniform weak convergence fails: Empirical processes for dependence functions and residuals via epi- and hypographs." DOI:10.1214/14-AOS1237SUPP.

BÜCHER, A. and VOLGUSHEV, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *J. Multivariate Anal.* **119** 61–70. MR3061415

CHEN, G. and LOCKHART, R. A. (2001). Weak convergence of the empirical process of residuals in linear models with many parameters. *Ann. Statist.* **29** 748–762. MR1865339

DEHEUVELS, P. (2009). A multivariate Bahadur–Kiefer representation for the empirical copula process. *J. Math. Sci.* **163** 382–398.

DEHLING, H. and DURIEU, O. (2011). Empirical processes of multidimensional systems with multiple mixing properties. *Stochastic Process. Appl.* **121** 1076–1096. MR2775107

DOUKHAN, P., FERMANIAN, J.-D. and LANG, G. (2009). An empirical central limit theorem with applications to copulas under weak dependence. *Stat. Inference Stoch. Process.* **12** 65–87. MR2486117

DREES, H. and HUANG, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal.* **64** 25–47. MR1619974

EINMAHL, J. H. J., KRAJINA, A. and SEGERS, J. (2012). An M-estimator for tail dependence in arbitrary dimensions. *Ann. Statist.* **40** 1764–1793. MR3015043

ENGLER, E. and NIELSEN, B. (2009). The empirical process of autoregressive residuals. *Econom. J.* **12** 367–381. MR2562392

FERMANIAN, J.-D., RADULOVIĆ, D. and WEGKAMP, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10** 847–860. MR2093613

GENEST, C., NEŠLEHOVÁ, J. and RÉMILLARD, B. (2014). On the empirical multilinear copula process for count data. *Bernoulli* **20** 1344–1371.

GENEST, C., QUESSY, J.-F. and REMILLARD, B. (2007). Asymptotic local efficiency of Cramér–von Mises tests for multivariate independence. *Ann. Statist.* **35** 166–191. MR2332273

GEYER, C. J. (1994). On the asymptotics of constrained M-estimation. *Ann. Statist.* **22** 1993–2010. MR1329179

GHOUDI, K. and RÉMILLARD, B. (2004). Empirical processes based on pseudo-observations. II. The multivariate case. In *Asymptotic Methods in Stochastics. Fields Inst. Commun.* **44** 381–406. Amer. Math. Soc., Providence, RI. MR2106867

HUANG, X. (1992). Statistics of bivariate extreme values. Ph.D. thesis, Tinbergen Institute Research Series, Netherlands.

KOSOROK, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York. MR2724368

KOUL, H. L. (1969). Asymptotic behavior of Wilcoxon type confidence regions in multiple linear regression. *Ann. Math. Statist.* **40** 1950–1979. MR0260126

KOUL, H. L. and QIAN, L. (2002). Asymptotics of maximum likelihood estimator in a two-phase linear regression model. *J. Statist. Plann. Inference* **108** 99–119. C. R. Rao 80th birthday felicitation volume, Part II. MR1947394

LOYNES, R. M. (1980). The empirical distribution function of residuals from generalised regression. *Ann. Statist.* **8** 285–298. MR0560730

MATHERON, G. (1975). *Random Sets and Integral Geometry*. Wiley, New York. MR0385969

MOLCHANOV, I. (2005). *Theory of Random Sets*. Springer, London. MR2132405

NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295. MR0293706

OGURA, Y. (2007). On some metrics compatible with the Fell–Matheron topology. *Internat. J. Approx. Reason.* **46** 65–73. MR2362225

PENG, L. and QI, Y. (2008). Bootstrap approximation of tail dependence function. *J. Multivariate Anal.* **99** 1807–1824. MR2444820

RÉMILLARD, B. and SCAILLET, O. (2009). Testing for equality between two copulas. *J. Multivariate Anal.* **100** 377–386. MR2483426

RIO, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. Mathématiques & Applications (Berlin) [Mathematics & Applications]* **31**. Springer, Berlin. MR2117923

ROCKAFELLAR, R. T. (1970). *Convex Analysis. Princeton Mathematical Series* **28**. Princeton Univ. Press, Princeton, NJ. MR0274683

ROCKAFELLAR, R. T. and WETS, R. J.-B. (1998). *Variational Analysis. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **317**. Springer, Berlin. MR1491362

RÜSCHENDORF, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist.* **4** 912–923. MR0420794

SEGERS, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18** 764–782. MR2948900

SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Teor. Veroyatn. Primen.* **1** 289–319. MR0084897

STRAF, M. L. (1972). Weak convergence of stochastic processes with several parameters. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971). Probability Theory* **II** 187–221. Univ. California Press, Berkeley, CA. MR0402847

TSUKAHARA, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* **33** 357–375. MR2193980

VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics*. Springer, New York. MR1385671

VAN DER VAART, A. W. and WELLNER, J. A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. IMS, Beachwood, OH. MR2459942

VERVAAT, W. (1981). Une compactification des espaces fonctionnels C et D; une alternative pour la démonstration de théorèmes limites fonctionnels. *C. R. Acad. Sci. Paris Sér. I Math.* **292** 441–444. MR0611412

WANG, Y. and STOEV, S. A. (2011). Conditional sampling for spectrally discrete max-stable random fields. *Adv. in Appl. Probab.* **43** 461–483. MR2848386

A. BÜCHER
S. VOLGUSHEV
FAKULTÄT FÜR MATHEMATIK
RUHR-UNIVERSITÄT BOCHUM
UNIVERSITÄTSSTRASSE 150
44780 BOCHUM
GERMANY
E-MAIL: axel.buecher@rub.de
stanislav.volgushev@rub.de

J. SEGERS
INSTITUT DE STATISTIQUE
UNIVERSITÉ CATHOLIQUE DE LOUVAIN
VOIE DU ROMAN PAYS, 20
B-1348 LOUVAIN-LA-NEUVE
BELGIUM
E-MAIL: johan.segers@uclouvain.be