The Annals of Probability
2013, Vol. 41, No. 2, 785–816
DOI: 10.1214/11-AOP711

A CLT FOR EMPIRICAL PROCESSES INVOLVING TIME-DEPENDENT DATA

BY JAMES KUELBS, THOMAS KURTZ[1] AND JOEL ZINN[2]

University of Wisconsin, University of Wisconsin and Texas A&M University

For stochastic processes $\{X_t : t \in E\}$, we establish sufficient conditions for the empirical process based on $\{I_{X_t \le y} - \Pr(X_t \le y) : t \in E, y \in \mathbb{R}\}$ to satisfy the CLT uniformly in $t \in E, y \in \mathbb{R}$. Corollaries of our main result include examples of classical processes where the CLT holds, and we also show that it fails for Brownian motion tied down at zero and $E = [0, 1]$.

1. Introduction. To form the classical empirical process, one starts with i.i.d. random variables $\{X_j : j \ge 1\}$, with distribution function F, and with

$$\mathbb{P}_n(A) = \frac{1}{n}\sum_{j=1}^{n} I_{X_j \in A}, \tag{1}$$

one considers the process $\mathbb{F}_n(y) = \mathbb{P}_n((-\infty, y])$.

By the classical Glivenko–Cantelli theorem,

$$\sup_{y\in\mathbb{R}} |\mathbb{F}_n((-\infty, y]) - F(y)| \longrightarrow 0 \qquad \text{a.s.}$$

By Donsker's theorem,

$$\{\sqrt{n}(\mathbb{F}_n(y) - F(y)) : y \in \mathbb{R}\}$$

converges in distribution in a sense described more completely below. Hence limit theorems for such processes, such as the law of large numbers and the central limit theorem (CLT), allow one to asymptotically get uniform estimates for the unknown cdf, $F(y) = \Pr(X \le y)$, via the sample data.

A more general version of these processes is to replace the indicators of half-lines in (1) by functions of a "random variable" taking values in some abstract space $(S, \mathcal{S})$. More specifically,

$$\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(f(X_j) - \mathbb{E}f(X)) : f \in \mathcal{F}\right\}, \tag{2}$$

Received August 2010; revised June 2011.

[1]Supported in part by NSF Grant DMS-08-05793.

[2]Supported in part by NSA Grant H98230-08-1-0101.

MSC2010 subject classifications. Primary 60F05; secondary 60F17.

Key words and phrases. Central limit theorems, empirical processes.

where the index set, $\mathcal{F}$, is a subset of $\mathcal{L}_\infty(S, \mathcal{S})$ or an appropriate subset of $\mathcal{L}_2(S, \mathcal{S}, P)$. We use the notation $\mathcal{L}_p(S, \mathcal{S}, P), 0 < p < \infty$, to denote the $\mathcal{S}$-measurable functions on S whose absolute value to the pth power is integrable with respect to P, rather than the equivalence classes of these functions. Of course, when $p = \infty$ the functions are $\mathcal{S}$-measurable and uniformly bounded on S. The standard notation $L_p(S, \mathcal{S}, P)$ is used when we are dealing with equivalence classes.

However, even in the case that the class of functions is a class of indicators, unlike the classical case, it is easy to see there are many classes of sets, $\mathcal{C}$, for which the limit theorem does not hold. As a matter of fact, the limiting Gaussian may not be continuous, for example, if $\mathcal{C} =$ all Borel sets of $\mathbb{R}$ or even $\mathcal{C} =$ all finite sets of $\mathbb{R}$. And further, even if the limiting Gaussian process is continuous, the limit theorem may still fail.

Luckily, in the case of sets, modulo questions of measurability, there are necessary and sufficient conditions for this sequence to converge in distribution to some mean zero Gaussian process. However, all the nasc's are described in terms of the asymptotic behavior of a complicated function of the sample, $\{X_n\}_{n=1}^\infty$. What we attempt to do in this paper is to obtain additional sufficient conditions that are useful when X takes values in some function space S, and the sets in $\mathcal{C}$ involve the time evolution of the stochastic process X. Of course, $\mathcal{C}$ is still a class of sets, but a primary goal that emerges here is to provide sufficient conditions for a uniform CLT in terms of the process $X = \{X(t) : t \in E\}$ that depend as little as possible on the parameter set E. However, classes of sets such as this rarely satisfy the Vapnik–Červonenkis condition. Also, this class of examples arises naturally from the study of the median process for independent Brownian Motions [see Swanson (2007, 2011)], where he studies the limiting quantile process for independent Brownian motions. This was observed by Tom Kurtz, and the follow-up questions led us to start this study. Here we concentrate on empirical process CLTs, and our main result is Theorem 3 below. Another theorem and some examples showing the applicability of Theorem 3 appear in Section 7. In Section 8 there are additional examples which show some obvious conjectures one might be tempted to make, concerning the CLT formulated here, are false. In particular, the examples in Section 8.4 motivate the various assumptions we employ in Theorem 3. As for future work, an upgraded version of Vervaat (1972) would perhaps allow one to relate the results obtained here and those of Swanson, but this is something to be done elsewhere.

2. Previous results and some definitions. Let $(S, \mathcal{S}, P)$ be a probability space, and define (Ω, Σ, Pr) to be the infinite product probability space $(S^N, \mathcal{S}^N, P^N)$. If $X_j : \Omega \to S$ are the natural projections of Ω into the jth copy of S, and $\mathcal{F}$ is a subset of $\mathcal{L}^2(S, \mathcal{S}, P)$ with

$$\sup_{f \in \mathcal{F}} |f(s)| < \infty, \qquad s \in S, \tag{3}$$

then we define

$$\nu_n(f) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \big(f(X_j) - \mathbb{E} f(X)\big), \qquad f \in \mathcal{F}. \tag{4}$$

Let $\ell_\infty(\mathcal{F})$ be the bounded real valued functions on $\mathcal{F}$, with the sup-norm, and recall that a Radon measure μ on $\ell_\infty(\mathcal{F})$ is a finite Borel measure which is inner regular from below by compact sets. Then the functions $f \to f(X_j) - \mathbb{E}(f(X_j))$, $j \geq 1$, are in $\ell_\infty(\mathcal{F})$, and we say $\mathcal{F} \in \mathrm{CLT}(P)$ if the stochastic processes $\{\nu_n(f), f \in \mathcal{F}\}, n \geq 1$, converge weakly to a centered Radon Gaussian measure γ_P on $\ell_\infty(\mathcal{F})$. More precisely, we have the following definition.

DEFINITION 1. *Let $\mathcal{F} \subset \mathcal{L}_2(P)$ and satisfy* (3). *Then $\mathcal{F} \in \mathrm{CLT}(P)$, or $\mathcal{F}$ is a P-Donsker class if there exists a centered Radon Gaussian measure γ_P on $\ell_\infty(\mathcal{F})$ such that for every bounded continuous real valued function H on $\ell_\infty(\mathcal{F})$, we have*

$$\lim_{n\to\infty} \mathbb{E}^*(H(\nu_n)) = \int H \, d\gamma_P,$$

where $\mathbb{E}^* H$ is the usual upper integral of H. If $\mathcal{C}$ is a collection of subsets from $\mathcal{S}$, then we say $\mathcal{C} \in \mathrm{CLT}(P)$ if the corresponding indicator functions are a P-Donsker class.

The probability measure γ_P of Definition 1 is obviously the law of the centered Gaussian process G_P, indexed by $\mathcal{F}$ having covariance function

$$\mathbb{E} G_P(f) G_P(g) = \mathbb{E}_P fg - \mathbb{E}_P f \mathbb{E}_P g, \qquad f, g \in \mathcal{F},$$

and L^2 distance

$$\rho_P(f, g) = \mathbb{E}_P\big(\{(f-g) - \mathbb{E}_P(f-g)\}^2\big)^{1/2}, \qquad f, g \in \mathcal{F}.$$

Moreover, if γ_P is as in Definition 1, then it is known that the process G_P admits a version all of whose trajectories are bounded and uniformly ρ_P continuous on $\mathcal{F}$. Hence we also make the following definition.

DEFINITION 2. *A class of functions $\mathcal{F} \subset \mathcal{L}_2(P)$ is said to be P-pre-Gaussian if the mean zero Gaussian process $\{G_P(f) : f \in \mathcal{F}\}$ with covariance and L_2 distance as indicated above has a version with all the sample functions bounded and uniformly continuous on $\mathcal{F}$ with respect to the L_2 distance $\rho_P(f, g)$.*

Now we state some results which are useful for what we prove in this paper. The first is important as it helps us establish counterexamples to natural conjectures one might make in connection to our main result, appearing in Theorem 3 below.

THEOREM 1 [Giné and Zinn (1984) for sufficiency and Talagrand (1988) for necessity of the $\Delta^{\mathcal{C}}$ condition in (ii)]. *Let $\Delta^{\mathcal{C}}(A)$ denote the number of distinct subsets of A obtained when one intersects all sets in $\mathcal{C}$ with A. Then, modulo measurability assumptions, conditions* (i) *and* (ii) *below are equivalent.*

(i) *The central limit theorem holds for the process*

$$\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}[I_{X_j\in C}-\Pr(X\in C)]:C\in\mathcal{C}\right\}$$

or more briefly $\mathcal{C}\in \mathrm{CLT}(P)$.

(ii)

$$\text{(a)}\quad \frac{\ln \Delta^{\mathcal{C}}(\{X_1,\dots,X_n\})}{\sqrt{n}}\to 0 \qquad \textit{in (outer) probability}\quad \textit{and} \tag{5}$$

$$\text{(b)}\quad \mathcal{C}\ \textit{is}\ P\textit{-pre-Gaussian}. \tag{6}$$

A sufficient condition for the empirical CLT, which is used in the proof of our main theorem, is given in Theorem 4.4 of Andersen et al. (1988).

THEOREM 2 [Andersen et al. (1988)]. *Let*

$$\mathcal{F}\subset\mathcal{L}_2(S,\mathcal{S},P),\qquad F=\sup_{f\in\mathcal{F}}|f(X)|$$

and P be the distribution of X with respect to Pr, *that is, $P=\Pr\circ X^{-1}$. Also, let $Pf=\int f(x)P(dx)$. Assume that $\|Pf\|_{\mathcal{F}}\equiv\sup_{f\in\mathcal{F}}|P(f)|<\infty$ and*:

(i) $u^2\Pr^*(F>u)\to 0$ *as* $u\to\infty$;
(ii) *$\mathcal{F}$ is P-pre-Gaussian*;
(iii) *there exists a centered Gaussian process $\{G(f):f\in\mathcal{F}\}$ with L_2 distance d_G such that G is sample bounded and uniformly d_G continuous on $\mathcal{F}$, and for some $K>0$, all $f\in\mathcal{F}$, and all $\varepsilon>0$,*

$$\Big[\sup_{u>0}u^2\Pr^*\Big(\sup_{\{g:d_G(g,f)<\varepsilon\}}|f-g|>u\Big)\Big]^{1/2}\le K\varepsilon.$$

Then $\mathcal{F}\in\mathrm{CLT}(P)$.

In this paper we take i.i.d. copies $\{X_j\}_{j=1}^{\infty}$ of a process $\{X(t):t\in E\}$, and consider

$$\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\big[I_{X_j(t)\le y}-\Pr\big(X(t)\le y\big)\big]:t\in E,\,y\in\mathbb{R}\right\} \tag{7}$$

with the goal of determining when these processes converge in distribution in some uniform sense to a mean zero Gaussian process. For example, if the process X has continuous sample paths on E, then $S = C(E)$ and the class of sets $\mathcal{C}$ in (i) of Theorem 1 consists of the sets $\{f \in C(E) : f(t) \le y\}$ for $t \in E$ and $y \in \mathbb{R}$, and we examine when $\mathcal{C} \in \mathrm{CLT}(P)$. As a result of the previous definitions, the limiting centered Gaussian process has a version with sample paths in a separable subspace of $\ell_\infty(\mathcal{C})$, and as a consequence of the addendum to Theorem 1.5.7 of van der Vaart and Wellner [(1996), page 37], almost all sample paths are uniformly L_2 continuous on $\mathcal{C}$ provided we identify the indicator functions of sets in $\mathcal{C}$ with $\mathcal{C}$ itself. Furthermore, we also have that $\mathcal{C}$ is totally bounded in the L_2 distance with this identification. In addition, the following remark is important in this setting.

REMARK 1. The assumption that a centered process $\{X(t) : t \in T\}$ with L_2 distance d is sample bounded and uniformly continuous on (T, d) is easily seen to follow if (T, d) is totally bounded and the process is uniformly continuous on (T, d). Moreover, if the process is Gaussian, the converse also holds using Sudakov's result as presented in Corollary 3.19 of Ledoux and Talagrand (1991).

To state and prove our result, we will make use of a distributional transform that appears in a number of places in the literature; see Ferguson (1967). Rüschendorf (2009) provides an excellent introduction to its history, and some uses. In particular, it is used there to obtain an elegant proof of Sklar's theorem [see Sklar (1973)], and also in some related applications.

Given the distribution function F of a real valued random variable Y, let V be a random variable uniformly distributed on $[0, 1]$ and independent of Y. In this paper we use the distributional transform of F defined as

$$\tilde{F}(x, V) = F(x^-) + V\big(F(x) - F(x^-)\big),$$

and Proposition 1 in Rüschendorf (2009) shows that

$$\tilde{F}(Y, V) \text{ is uniform on } [0, 1]. \tag{8}$$

Rüschendorf calls $\tilde{F}(Y, V)$ the distributional transform of Y, and we also note that $\tilde{F}(x, V)$ is nondecreasing in x.

3. The main conditions. Let $\{X(t) : t \in E\}$ be a stochastic process as in (7), and assume

$$\rho(s, t) = \big(\mathbb{E}(H_t - H_s)^2\big)^{1/2}, \qquad s, t \in E,$$

where $\{H(t) : t \in E\}$ is some Gaussian process which is sample bounded and ρ uniformly continuous on E. In our main result (see Theorem 3 below), we hypothesize the relationship between $\{X(t) : t \in E\}$ and $\rho(s, t), s, t \in E$ given in (10). The importance of this condition in the proof of our theorem is 2-fold. First, it allows one to establish the limiting Gaussian process for our CLT actually exists. This verifies condition (ii) in Theorem 2 above, and is accomplished via a subtle application of the necessary and sufficient conditions for the existence and sample function regularity of a Gaussian process given in Talagrand (2005). Second, it also allows us to verify that the remaining nontrivial condition sufficient for our CLT, namely condition (iii) of Theorem 2, applies in this setting. This is useful in applications as condition (iii) is in terms of a single random element involved in our CLT, and hence is far easier to verify than the typical condition which depends on the full sample of the random elements as in (5). The reader should also note that Theorem 2 is a refinement of a number of previous results sufficient for the CLT, and covers a number of important earlier empirical process papers.

The existence of such a $\rho(\cdot,\cdot)$ is obtained for a number of specific processes $\{X(t): t \in E\}$ in Section 7. Nevertheless, given the process X, determining whether a suitable ρ satisfying (10) exists, or does not exist, may be quite difficult. However, using (12) we may limit our choice for ρ in (10) to be such that

$$\rho(s,t) \geq c^{-1} \sup_{x\in\mathbb{R}} \bigl(\mathbb{E}(|I_{X_t \leq x} - I_{X_s \leq x}|^2)\bigr)^{1/2}$$

for all $s,t \in E$ and some constant $c \in (0,\infty)$.

Throughout, $\rho(s,t), s,t \in E$, denotes the L_2 metric of a Gaussian process indexed by E and, to simplify notation, we let

$$\tilde{F}_t(x) \equiv (\tilde{F}_t)(x,V), \qquad x \in \mathbb{R},$$

be the distributional transform of F_t, the distribution function of X_t. Note that this simplification of notation also includes using X_t for $X(t)$ when the extra parenthesis make the latter clumsy or unnecessary. Moreover, this variable notation is also employed for other stochastic processes in the paper. In addition, for each $\varepsilon > 0$, let

$$\sup_{\{s,t\in E:\, \rho(s,t)\leq\varepsilon\}} \Pr{}^*\bigl(|\tilde{F}_t(X_s) - \tilde{F}_t(X_t)| > \varepsilon^2\bigr) \leq L\varepsilon^2 \tag{9}$$

(the weak L condition)

and

$$\sup_{t\in E} \Pr{}^*\Bigl(\sup_{\{s:\, \rho(s,t)\leq\varepsilon\}} |\tilde{F}_t(X_s) - \tilde{F}_t(X_t)| > \varepsilon^2\Bigr) \leq L\varepsilon^2 \tag{10}$$

(the L condition).

REMARK 2. In the L conditions the probabilities involve an ε^2. However, since for any constant $C \in (0,\infty)$, an L_2 metric ρ, is pre-Gaussian if and only if $C\rho$ is pre-Gaussian, WLOG we can change to $C\varepsilon^2$. Moreover, note that any constant L sufficient for (10) will also suffice for (9), and hence to simplify notation we do not distinguish between them.

LEMMA 1. *Let L be as in (9), and take $s, t \in E$. Then, for all $x \in \mathbb{R}$,*

$$\Pr(X_s \le x < X_t) \le (L+1)\rho^2(s,t) \tag{11}$$

and by symmetry,

$$\begin{aligned} \mathbb{E}|I_{X_t \le x} - I_{X_s \le x}| &= \Pr(X_t \le x < X_s) + \Pr(X_s \le x < X_t) \\ &\le 2(L+1)\rho^2(s,t). \end{aligned} \tag{12}$$

Further, we have

$$\sup_x |F_t(x) - F_s(x)| \le 2(L+1)\rho^2(s,t). \tag{13}$$

REMARK 3. As in Lemma 1 and Lemmas 2 and 3 below, use only the weak L condition (9). Actually, for Lemmas 2 and 3, all we need is Lemma 1. However, in Lemma 4 we need the stronger form as stated in (10).

PROOF OF LEMMA 1. Since $\tilde{F}_t$ is nondecreasing and $x < y$ implies $F_t(x) \le \tilde{F}_t(y)$, we have

$$\Pr(X_s \le x < X_t) \le \Pr\bigl(\tilde{F}_t(X_s) \le \tilde{F}_t(x), F_t(x) \le \tilde{F}_t(X_t)\bigr).$$

Thus

$$\begin{aligned} \Pr(X_s \le x < X_t) \le{}& \Pr\bigl(F_t(x) \le \tilde{F}_t(X_t) \le F_t(x) + \rho^2(s,t), \tilde{F}_t(X_s) \le \tilde{F}_t(x)\bigr) \\ &+ \Pr\bigl(\tilde{F}_t(X_t) > F_t(x) + \rho^2(s,t), \tilde{F}_t(X_s) \le \tilde{F}_t(x)\bigr) \end{aligned}$$

and hence

$$\begin{aligned} \Pr(X_s \le x < X_t) \le{}& \Pr\bigl(F_t(x) \le \tilde{F}_t(X_t) \le F_t(x) + \rho^2(s,t)\bigr) \\ &+ \Pr\bigl(|\tilde{F}_t(X_t) - \tilde{F}_t(X_s)| > \rho^2(s,t)\bigr). \end{aligned}$$

Now (9) implies for all $s, t \in E$ that

$$\Pr\bigl(|\tilde{F}_t(X_t) - \tilde{F}_t(X_s)| > \rho^2(s,t)\bigr) \le L\rho^2(s,t),$$

since its failure for $s_0, t_0 \in E$ and $\varepsilon = \rho(s_0, t_0)$ in (9) implies a contradiction. Therefore, since $\tilde{F}_t(X_t)$ is uniform on [0, 1], we have

$$\Pr(X_s \le x < X_t) \le \rho^2(s,t) + L\rho^2(s,t).$$

The last conclusion follows by moving the absolute values outside the expectation. $\square$

4. The main result. Recall the relationship of the X process and ρ as described at the beginning of Section 3. Then we have:

THEOREM 3. *Let ρ be given by $\rho^2(s,t) = \mathbb{E}(H(s) - H(t))^2$, for some centered Gaussian process H that is sample bounded and uniformly continuous on (E, ρ) with probability one. Furthermore, assume that for some $L < \infty$, and all $\varepsilon > 0$, the L condition* (10) *holds, and $D(E)$ is a collection of real valued functions on E such that* $\Pr(X(\cdot) \in D(E)) = 1$. *If*

$$\mathcal{C} = \{C_{s,x} : s \in E, x \in \mathbb{R}\},$$

where

$$C_{s,x} = \{z \in D(E) : z(s) \leq x\}$$

for $s \in E, x \in \mathbb{R}$, then $\mathcal{C} \in \mathrm{CLT}(P)$.

REMARK 4. Note that a sample function of the X-process is in the set $C_{s,x}$ iff $X_s \leq x$. Hence, if we identify a point $(s,x) \in E \times \mathbb{R}$ with the set $C_{s,x}$, then instead of saying $\mathcal{C} \in \mathrm{CLT}(P)$, we will often say

$$\{I_{X_s \leq x} - \Pr(X_s \leq x) : s \in E, x \in \mathbb{R}\}$$

satisfies the CLT in $\ell_\infty(\mathcal{C})$ [or in $\ell_\infty(E \times \mathbb{R})$].

REMARK 5. At this point one might guess that the reader is questioning the various assumptions in Theorem 3. First we mention that $D(E)$ is some convenient function space. For example, typically the process X has continuous sample paths on E, so $D(E) = C(E)$ in these situations. More perplexing, at least for most readers, is probably the appearance of the distributional transforms $\{\tilde{F}_t : t \in E\}$ in the L condition (10). If the distribution functions F_t are all continuous, then $F_t = \tilde{F}_t, t \in E$, and our proof obviously holds with F_t replacing $\tilde{F}_t$ in the L condition. However, without all the distribution functions F_t assumed continuous, the methods required in our proof fail with this substitution. An interesting case where the distributional transforms are useful occurs when one has a point $t_0 \in E$ such that $\Pr(X(t) = X(t_0)$ for all t in $E) = 1$, and F_{t_0} is possibly discontinuous. In this situation, the L condition (10) holds for the Gaussian process $H(t) = g$ for all $t \in E$, g a standard Gaussian random variable and $X(t_0)$ having any distribution function F_{t_0}. Thus Theorem 3 applies and yields the classical empirical CLT when the set S is the real line, and the class of sets consists of half-lines for all laws F_{t_0}. A similar result also applies if E is a finite disjoint union of nonempty sets, and the process $\{X_t : t \in E\}$ is constant on each of the disjoint pieces of E regardless of the distribution functions $F_t, t \in E$. More importantly, however, allowing even a single discontinuous distribution F_t may invalidate the empirical CLT on $\mathcal{C}$. For example, if $\{X(t) : t \in [0,1]\}$ is standard Brownian motion with $P(X(0) = 0) = 1$, then in Section 8.1 we show the empirical CLT fails, but Corollary 2 shows that it holds if we allow the distribution at time zero to have a bounded density. Furthermore, in Section 8.4 we provide some additional examples where the empirical process is pre-Gaussian, and the input process $\{X(t) : t \in E\}$ satisfies a modified L condition, that is, for all $\varepsilon > 0$, there is an $L < \infty$ such that

$$\sup_{t\in E} \Pr{}^*\Bigl(\sup_{\{s : \rho(s,t)\le\varepsilon\}} |F_t(X_s) - F_t(X_t)| > \varepsilon^2\Bigr) \le \varepsilon^2,$$

yet the empirical CLT we seek fails. Hence one needs to assume something more, and our results show that the L condition given in (10) is sufficient for the empirical CLT.

In Section 7 we will provide another theorem showing how Theorem 3 can be applied, and hence the examples obtained there are motivation for its formulation in terms of the L condition (10). The following remark also motivates the presence of the process $\{\tilde{F}_t(X_s) : s \in E\}$ and the L condition in our CLT. In particular, we sketch an argument that for each $t \in E$ a symmetric version of this process satisfies a CLT in $\ell_\infty(E)$. This remark is meant only for motivation, and in its presentation we are unconcerned with a number of details.

REMARK 6. Let

$$\{I_{X_s\le x} - \Pr(X_s \le x) : s \in E, x \in \mathbb{R}\}$$

satisfy the central limit theorem in the closed subspace of $\ell_\infty(E \times \mathbb{R})$ consisting of functions whose s-sections are Borel measurable on $\mathbb{R}$. We denote this subspace by $\ell_{\infty,m}(E \times \mathbb{R})$, and also assume the distribution functions F_t are all continuous. Then, for each fixed $t \in E$, we define the bounded linear operator $\phi : \ell_{\infty,m}(E \times \mathbb{R}) \longrightarrow \ell_\infty(E)$ given by

$$\phi(f)(s) = \int f(s,x) F_t(dx).$$

Now by the symmetrization lemma [Lemma 2.7 in Giné and Zinn (1984)], we have for a Rademacher random variable ε independent of the empirical process variables that $\{\varepsilon I_{C_{s,x}} : s \in E, x \in \mathbb{R}\}$ satisfy the CLT in $\ell_\infty(E \times \mathbb{R})$. Taking $f(s,x) = I_{C_{s,x}}$, we have for all $t \in E$ fixed that

$$\phi(f)(s) = 1 - F_t(z(s)^-) = 1 - F_t(z(s))$$

as we are assuming the F_t are continuous. Therefore the continuous mapping theorem [see, e.g., Theorem 1.3.6 in van der Vaart and Wellner (1996)] implies that for each $t \in E$,

$$Z_n(s) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j \bigl(1 - F_t(X_s)\bigr), \qquad s \in E,$$

satisfies the CLT in $\ell_\infty(E)$. In addition, since we are assuming $F_t = \tilde{F}_t$, we should then have "asymptotically small oscillations;" namely, for every $\delta > 0$ there exists $\varepsilon > 0$ such that

$$\Pr{}^*\left(\sup_{\rho(s,t)\le\varepsilon} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varepsilon_j \bigl(\tilde{F}_t(X_s) - \tilde{F}_t(X_t)\bigr) > \delta \right) \le \delta.$$

By using standard symmetry arguments this last probability dominates

$$\frac{1}{2} \Pr{}^*\Bigl(\max_{j\le n} \sup_{\rho(s,t)\le\varepsilon} |\tilde{F}_t(X_s) - \tilde{F}_t(X_t)| > \sqrt{n}\delta\Bigr) \le \delta,$$

which (again by standard arguments) implies (modulo multiplicative constants)

$$n \Pr{}^*\Bigl(\sup_{\rho(s,t)\le\varepsilon} |\tilde{F}_t(X_s) - \tilde{F}_t(X_t)| > \sqrt{n}\delta\Bigr) \le \delta.$$

While this is different from the hypotheses in our theorem, it indicates that the quantity $\sup_{\rho(s,t)\le\varepsilon} |\tilde{F}_t(X_s) - \tilde{F}_t(X_t)|$ is relevant to any such theorem.

5. Preliminaries for generic chaining. Let T be an arbitrary countable set. Then, following Talagrand (2005) we have:

DEFINITION 3. An admissible sequence is an increasing sequence $(\mathcal{A}_n)$ of partitions of T such that

$$\operatorname{Card} \mathcal{A}_n \le N_n,$$

where $N_0 = 1$, and for $n \ge 1$, $N_n = 2^{2^n}$. The partitions $(\mathcal{A}_n)$ are increasing if every set in $\mathcal{A}_{n+1}$ is a subset of some set of $\mathcal{A}_n$.

We also have:

DEFINITION 4. If $t \in T$, we denote by $A_n(t)$ the unique element of $\mathcal{A}_n$ that contains t. For a psuedo-metric e on T, and $A \subseteq E$, we write $\Delta_e(A)$ to denote the diameter of A with respect to e.

Using generic chaining and the previous definitions, Theorem 1.4.1 of Talagrand (2005) is essentially the following result. Its statement there contains a curious wording at the end of the first sentence, which suggests that cutting and pasting led to something being omitted. After closer inspection we observed that the necessary assumption of total boundedness of the parameter space was required, and it now appears in the statement of the theorem below. Since Theorem 1.4.1 appears without proof, for completeness the proof can be found in the Appendix.

THEOREM 4. *Let $\{X_t : t \in T\}$ be a centered Gaussian process with L_2 distance $d(s,t)$, $s,t \in T$, where T is countable, and (T,d) is totally bounded. Then, the following are equivalent*:

(i) *X_t is uniformly continuous on (T, d) with probability one.*

(ii) *We have*

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big(\sup_{d(s,t) \le \varepsilon} (X_s - X_t)\Big) = 0.$$

(iii) *There exists an admissible sequence of partitions of T such that*

$$\lim_{k \to \infty} \sup_{t \in T} \sum_{n \ge k} 2^{n/2} \Delta(A_n(t)) = 0. \tag{14}$$

Under the assumption that H is centered Gaussian and uniformly continuous on (T, e), then, recalling Remark 1, it follows that H being sample bounded on T is equivalent to (T, e) being totally bounded. Also, an immediate corollary of this result used below is as follows.

PROPOSITION 1. *Let H_1 and H_2 be mean zero Gaussian processes with L_2 distances e_1, e_2, respectively, on T. Furthermore, assume T is countable, and $e_1(s,t) \le e_2(s,t)$ for all $s, t \in T$. Then, H_2 sample bounded and uniformly continuous on (T, e_2) with probability one, implies H_1 is sample bounded and uniformly continuous on (T, e_1) with probability one.*

REMARK 7. One can prove this using Slepian's lemma [see, e.g., Fernique (1975)]. However, the immediate conclusion is that H_1 is sample bounded and uniformly continuous on (T, e_2). Then, a separate argument is needed to show the statement in this proposition. Using the more classical formulation for continuity of Gaussian processes involving majorizing measures [see, e.g., Theorem 12.9 of Ledoux and Talagrand (1991)], the result also follows similarly to what is explained below.

PROOF OF PROPOSITION 1. By the previous theorem $\{H_2(t) : t \in T\}$ is sample bounded and uniformly continuous on (T, e_2) with probability one if and only if there exists an admissible sequence of partitions of T such that

$$\lim_{r \to \infty} \sup_{t \in T} \sum_{n \ge r} 2^{n/2} \Delta_{e_2}(A_n(t)) = 0.$$

Since $\Delta_{e_1}(A_n(t)) \le \Delta_{e_2}(A_n(t))$, we have

$$\lim_{r \to \infty} \sup_{t \in T} \sum_{n \ge r} 2^{n/2} \Delta_{e_1}(A_n(t)) = 0,$$

and hence Theorem 4 implies that H_1 is sample bounded and uniformly continuous on (T, e_1) with probability one. Thus the proposition is proven. □

6. Proof of Theorem 3. First we establish some necessary lemmas, and the section ends with the proof of Theorem 3. Throughout we take as given the assumptions and notation of that theorem.

6.1. *Some additional lemmas*. In order to simplify notation, we denote the L_2 distance on the class of indicator functions

$$\mathcal{F} = \{I_{X_s \le x} : s \in E, x \in \mathbb{R}\}$$

by writing

$$\tau((s,x),(t,y)) = \{E((I_{X_s \le x} - I_{X_t \le y})^2)\}^{1/2}$$

and identifying $\mathcal{F}$ with $E \times \mathbb{R}$. Our next lemma relates the τ-distance and the ρ-distance. It upgrades (12) when $x \neq y$.

LEMMA 2. *Assume that* (9) *holds. Then*

$$\tau^2((s,x),(t,y)) \le \min_{u \in \{s,t\}} |F_u(y) - F_u(x)| + (2L+2)\rho^2(t,s). \tag{15}$$

Moreover, if Q denotes the rational numbers, there is a countable dense set E_0 of (E,ρ) such that $\mathcal{F}_0 = \{I_{X_s \le x} : (s,x) \in E_0 \times Q\}$ is dense in $(\mathcal{F},\tau)$.

PROOF. First observe that by using the symmetry in s and t of the right-hand term of (15), we have, by applying (11) in the second inequality below, that

$$\begin{aligned}\tau^2((s,x),(t,y)) &= \mathbb{E}|I_{X_t \le y} - I_{X_s \le x}| \le \mathbb{E}|I_{X_t \le y} - I_{X_t \le x}| + \mathbb{E}|I_{X_t \le x} - I_{X_s \le x}| \\ &= |F_t(y) - F_t(x)| + \Pr(X_s \le x < X_t) + \Pr(X_t \le x < X_s) \\ &\le |F_t(y) - F_t(x)| + (2+2L)\rho^2(s,t).\end{aligned}$$

Similarly, we also have by applying (11) again that

$$\tau^2((s,x),(t,y)) \le |F_s(y) - F_s(x)| + (2+2L)\rho^2(s,t).$$

Combining these two inequalities for τ, the proof of (15) holds. Since (E,ρ) is assumed totally bounded, there is a countable dense set E_0 of (E,ρ), and hence the right continuity of the distribution functions and (15) then imply the final statement in Lemma 2. $\square$

Using Lemma 2 and the triangle inequality, we can estimate the τ-diameter of sets as follows.

COROLLARY 1. *If $t_B \in B \subseteq E$ and $D \subseteq \mathbb{R}$, then*

$$\operatorname{diam}_\tau(B \times D) \le 2\Big\{(2L+2)^{1/2} \operatorname{diam}_\rho(B) + \sup_{x,y \in D} |F_{t_B}(y) - F_{t_B}(x)|^{1/2}\Big\}.$$

LEMMA 3. *Assume that (s,x) and (t,y) satisfy*

$$\tau((s,x),(t,y)) = \|I_{X_s \le x} - I_{X_t \le y}\|_2 \le \varepsilon,$$

$\rho(s,t) \le \varepsilon$, *and* (9) *holds*. *Then*, *for* $c = (2L+2)^{1/2}+1$,

$$|F_t(x) - F_t(y)| \le (c\varepsilon)^2$$

or, *in other words*,

$$|F_t(x) - F_t(y)| \le (c \max\{\tau((s,x),(t,y)), \rho(s,t)\})^2.$$

PROOF. Using (11) in the second inequality below, we have

$$\begin{aligned}|F_t(y) - F_t(x)|^{1/2} &= \|I_{X_t \le x} - I_{X_t \le y}\|_2 \le \|I_{X_s \le x} - I_{X_t \le y}\|_2 + \|I_{X_s \le x} - I_{X_t \le x}\|_2 \\ &\le \varepsilon + \big(\Pr(X_s \le x < X_t) + \Pr(X_t \le x < X_s)\big)^{1/2} \\ &\le \varepsilon + (2L\varepsilon^2 + 2\varepsilon^2)^{1/2} = [(2L+2)^{1/2}+1]\varepsilon \equiv c\varepsilon.\end{aligned}$$

Hence the lemma is proven. $\square$

The next lemma is an important step in verifying the weak-L_2 condition in item (iii) of Theorem 2 above; for example, see Theorem 4.4 of Andersen et al. (1988).

LEMMA 4. *If* (10) *holds*, c *is as in Lemma* 3, *and*

$$\lambda((s,x),(t,y)) = \max\{\tau((s,x),(t,y)), \rho(s,t)\};$$

then for all (t,y) *and* $\varepsilon > 0$,

$$\Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} |I_{X_t \le y} - I_{X_s \le x}| > 0\Big) \le 2(c^2+L+1)\varepsilon^2.$$

PROOF. First we observe that

$$\begin{aligned}&\Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} |I_{X_t \le y} - I_{X_s \le x}| > 0\Big) \\ &\quad = \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{X_t \le y, X_s > x} + I_{X_s \le x, X_t > y} > 0\Big).\end{aligned}$$

Again, using the fact that $x < y$ implies $F_t(x) \le \tilde{F}_t(y)$, we have

$$\begin{aligned}&\Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} |I_{X_t \le y} - I_{X_s \le x}| > 0\Big) \\ &\quad \le \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde{F}_t(X_t) \le \tilde{F}_t(y), F_t(x) \le \tilde{F}_t(X_s)} > 0\Big) \\ &\qquad + \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde{F}_t(X_s) \le \tilde{F}_t(x), F_t(y) \le \tilde{F}_t(X_t)} > 0\Big) = I + II,\end{aligned}$$

where

$$I = \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde{F}_t(X_t) \le \tilde{F}_t(y), F_t(x) \le \tilde{F}_t(X_s)} > 0\Big) \tag{16}$$

and

$$II = \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde F_t(X_s)\le \tilde F_t(x),\, F_t(y)\le \tilde F_t(X_t)} > 0\Big). \tag{17}$$

At this point we use Lemma 3 to see that in (16) we can use

$$\inf_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} F_t(x) \ge F_t(y) - (c\varepsilon)^2.$$

Therefore, since $\tilde F_t(x) \le F_t(x)$ for all x and again using (10)

$$\begin{aligned}
I &\le \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde F_t(X_t)\le F_t(y),\, F_t(y)-(c\varepsilon)^2\le \tilde F_t(X_s)} > 0\Big)\\
&\le \Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde F_t(X_t)\le F_t(y),\, F_t(y)-(c\varepsilon)^2\le \tilde F_t(X_t)+\varepsilon^2} > 0\Big) + L\varepsilon^2\\
&\le \Pr\big(F_t(y) - (c\varepsilon)^2 - \varepsilon^2 \le \tilde F_t(X_t) \le F_t(y)\big) + L\varepsilon^2\\
&\le (c^2 + L + 1)\varepsilon^2 \qquad \text{by (8)}.
\end{aligned}$$

Now, we estimate II in (17). Again using the fact that $\tilde F_t(x) \le F_t(x)$ for all x, Lemma 3, and our definition of L, we therefore have

$$\begin{aligned}
&\Pr^*\Big(\sup_{\{(s,x):\lambda((t,y),(s,x))\le\varepsilon\}} I_{\tilde F_t(X_s)\le \tilde F_t(x),\, F_t(y)\le \tilde F_t(X_t)} > 0\Big)\\
&\qquad\le \Pr\big(\tilde F_t(X_t) - \varepsilon^2 \le F_t(y) + (c\varepsilon)^2,\ F_t(y) \le \tilde F_t(X_t)\big) + L\varepsilon^2\\
&\qquad\le (c^2 + L + 1)\varepsilon^2.
\end{aligned}$$

□

6.2. *The construction and the proof of Theorem* 3. Since (E,ρ) is totally bounded by Remark 1, take E_0 to be any countable dense subset of E in the ρ distance. Then by Theorem 4, Talagrand's continuity theorem, there exists an admissible sequence of partitions, $\mathcal{B}_n$ of E_0, for which

$$\lim_{r\to\infty} \sup_{t\in E_0} \sum_{n\ge r} 2^{n/2} \Delta_\rho(B_n(t)) = 0. \tag{18}$$

Fix n. Then, for each $B \in \mathcal{B}_{n-1}$ choose $t_B \in B$. Fix the distribution function $F_B := F_{t_B}$ and μ_B the associated probability measure. Put $\alpha = (\Delta_\rho(B) + 2^{-n})^2$ and set $z_1 = \sup\{x \in \mathbb{R} : F_B(x) < \alpha\}$. We consider two cases:

- $F_B(z_1) \le \alpha$ and
- $F_B(z_1) > \alpha$.

In the first case $F_B(z_1) = \alpha$. If $F_B(z_1) < \alpha$, then by right continuity there exist $w > z_1$ such that $F_B(w) < \alpha$, which contradicts the definition of z_1. In this case we consider $C_1 = (-\infty, z_1]$ and $D_1 = \varnothing$.

In the second case we let $C_1 = (-\infty, z_1)$ and $D_1 = \{z_1\}$. In either case $\mu_B(C_1 \cup D_1) \ge \alpha$.

If $\mu_B((z_1,\infty)) \ge \alpha$, let $z_2 = \sup\{x > z_k : F_B(x) - F_B(z_1) < \alpha\}$. If $z_2 = \infty$, we set $C_2 = (z_1,\infty)$ and $D_2 = \varnothing$. Otherwise, if $z_2 < \infty$, there are two cases. That is, we have:

- $F_B(z_2) - F_B(z_1) \le \alpha$ and
- $F_B(z_2) - F_B(z_1) > \alpha$.

In the first case we consider $C_2 = (z_1, z_2]$ and $D_2 = \varnothing$. In the second case we let $C_2 = (z_1, z_2)$ and $D_2 = \{z_2\}$. As before, $\mu_B(C_2 \cup D_2) \ge \alpha$.

Now assume that we have constructed $C_1, \dots, C_k$ and $D_1, \dots, D_k$ in this manner. Therefore we have z_k. If $\mu_B((z_k,\infty)) \ge \alpha$, let $z_{k+1} = \sup\{x > z_k : F_B(x) - F_B(z_k) < \alpha\}$. If $z_{k+1} = \infty$, we set $C_{k+1} = (z_k,\infty)$ and $D_{k+1} = \varnothing$. Otherwise, if $z_{k+1} < \infty$, there are two cases. That is, we have:

- $F_B(z_{k+1}) - F_B(z_k) \le \alpha$ and
- $F_B(z_{k+1}) - F_B(z_k) > \alpha$.

In the first case we consider $C_{k+1} = (z_k, z_{k+1}]$ and $D_{k+1} = \varnothing$. In the second case we let $C_{k+1} = (z_k, z_{k+1})$ and $D_{k+1} = \{z_{k+1}\}$. As before, $\mu_B(C_{k+1} \cup D_{k+1}) \ge \alpha$. Hence, there can be at most $\frac{1}{\alpha} + 1$ steps before $\{C_k, D_k\}_k$ cover $\mathbb{R}$. Therefore, after eliminating any empty set, we have a cover of $\mathbb{R}$ with at most $\frac{2}{\alpha} + 2$ sets. By our choice of α the cover has at most $2^{2n+1} + 2$ sets. Hence since we have $B \in \mathcal{B}_{n-1}$, the number of sets used to cover $E_0 \times \mathbb{R}$ of the form $B \times C_k$ or $B \times D_k$ is less than or equal to $2^{2^{n-1}}(2^{2n+1} + 2)$. The reader should note that the points $\{z_k\}$ depend on the set B, but we have suppressed that to simplify notation. We now check the τ-diameters of the nonempty $B \times C_k$ and $B \times D_k$.

Estimating these diameters by doubling the radius of the sets, the triangle inequality allows us to upper bound their radius using one of s and t to be t_B. Also note that in Lemma 2, or Corollary 1, the term which contains $|F_{t_B}(y) - F_{t_B}(x)|$ would cause trouble in the case $D_k \ne \varnothing$, since this is only known to be $\ge \alpha$. Luckily it does not appear when $D_k \ne \varnothing$.

First we consider the τ-diameter of sets of the form $B \times C_k$ when $D_k = \varnothing$. Then $C_k = (z_{k-1,B}, z_{k,B}]$. Hence for $(s,x), (t,y) \in B \times C_k$, Corollary 1 implies

$$\Delta_\tau\bigl(B \times (z_{B,k-1}, z_{B,k}]\bigr) \le 2\biggl((2L+2)^{1/2}\Delta_\rho(B) + \Delta_\rho(B) + \frac{1}{2^n}\biggr).$$

When $D_k \ne \varnothing$, then $C_k = (z_{k-1,B}, z_{k,B})$, so again by Corollary 1 the τ-diameter of $B \times C_k$ has an upper bound as in the previous case.

If $D_k \ne \varnothing$, then the only element of D_k is $z_{k,B}$, and by Corollary 1 we have

$$\Delta_\tau(B \times D_k) \le 2(2L+2)^{1/2}\,\mathrm{diam}_\rho(B).$$

So, in either case,

$$\Delta_\tau(B \times C_{B,k} \text{ or } D_{B,k}) \le 2\biggl((2L+2)^{1/2}\Delta_\rho(B) + \Delta_\rho(B) + \frac{1}{2^n}\biggr). \tag{19}$$

LEMMA 5. *Let $\mathcal{G}_n$ be a sequence of partitions of an arbitrary parameter set T with pseudo metric e on T satisfying both*:

(1) $\mathrm{Card}(\mathcal{G}_n) \le 2^{2^n}$ *and*
(2) $\lim_{r\to\infty} \sup_{t\in T} \sum_{n\ge r} 2^{n/2}\Delta_e(G_n(t)) = 0$,

and set $\mathcal{H}_n := \mathcal{P}(\bigcup_{1\le k\le n-1} \mathcal{G}_k)$, where $\mathcal{P}(\mathcal{D})$ denotes the minimal partition generated by the sets in $\mathcal{D}$. Then the sequence $\mathcal{H}_n$ (notice the $n-1$ in the union) *also satisfies those conditions*.

PROOF. The first condition holds since a simple induction on n implies the minimal partition

$$\mathcal{H}_n = \mathcal{P}\Bigl(\bigcup_{1\le k\le n-1} \mathcal{G}_k\Bigr)$$

has cardinality at most $\prod_{k=1}^{n-1} 2^{2^k} \le 2^{2^n}$. The second condition holds since the partitions are increasing collections of sets, and hence $\mathrm{diam}_e(H_n(t)) \le \mathrm{diam}_e(G_{n-1}(t))$. $\square$

LEMMA 6. *Let E_0 be a countable dense subset of (E, ρ). Then there exists an admissible sequence of partitions $\{\mathcal{A}_n : n \ge 0\}$ of $E_0 \times \mathbb{R}$ such that*

$$\lim_{r\to\infty} \sup_{(t,y)\in E_0\times\mathbb{R}} \sum_{n\ge r} 2^{n/2}\Delta_\tau(A_n((t, y))) = 0. \tag{20}$$

PROOF. We construct the admissible sequence of partitions $\mathcal{A}_n$ as above. More precisely, let $\{\mathcal{B}_n : n \ge 0\}$ be an increasing sequence of partitions of E_0 such that (18) holds, and after the construction above we also have (19). That is, for $k \ge 1$ let

$$\mathcal{G}_k = \{B \times F : B \in \mathcal{B}_{k-1}, F \in \mathcal{E}_{\mathcal{B}}\},$$

where

$$\mathcal{E}_B = \{C_{j,B}, D_{j,B} \text{ all sets nonempty}\}$$

and $C_{j,B}, D_{j,B}$ are constructed from $B \in \mathcal{B}_{k-1}$ as above. Then, for $n \ge 4$ set

$$\mathcal{A}_n = \mathcal{P}\Bigl(\bigcup_{3\le k\le n-1} \mathcal{G}_k\Bigr),$$

where $\mathcal{P}(\mathcal{D})$ is the minimal partition generated by the sets in $\mathcal{D}$, and for $n = 1, 2, 3$ we take $\mathcal{A}_n$ to be the single set $E_0 \times \mathbb{R}$. Since the cardinality of the partitions $\mathcal{G}_k$ defined above is less than or equal to $2^{2^{k-1}}(2^{2k+1} + 2)$, then for $n \ge 4$ a simple computation implies the minimal partition

$$\mathcal{A}_n = \mathcal{P}\Bigl(\bigcup_{3\le k\le n-1} \mathcal{G}_k\Bigr)$$

has cardinality at most $\prod_{k=3}^{n-1} 2^{2^{k-1}}(2^{2k+1}+2) \le \prod_{k=3}^{n-1} 2^{2^{k-1}} 2^{2k+2} \le 2^{2^n}$. By (19) and Lemma 5 we have

$$\sup_{(t,y)} \sum_{n\ge r} 2^{n/2}\Delta_\tau(A_n(t,y)) \le C\left\{\sup_t \sum_{n\ge r} 2^{n/2}\Delta_\rho(B_n(t)) + \sum_{n\ge r} 2^{n/2}2^{-n}\right\}.$$

Thus (18) implies that τ satisfies (20) with respect to the sequence of admissible partitions $\mathcal{A}_n$ on $E_0 \times \mathbb{R}$. $\square$

PROOF OF THEOREM 3. Let Q denote the rational numbers. Then, if we restrict the partitions $\mathcal{A}_n$ of $E_0 \times \mathbb{R}$ in Lemma 6 to $E_0 \times Q$, we immediately have

$$\lim_{r\to\infty} \sup_{(t,y)\in E_0\times Q} \sum_{n\ge r} 2^{n/2}\Delta_\tau(A_n((t,y))) = 0, \tag{21}$$

and $(E_0 \times Q, \tau)$ is totally bounded. Now let $\{G_{(s,x)} : (s,x) \in E \times \mathbb{R}\}$ be a centered Gaussian process with $\mathbb{E}(G_{(s,x)}G_{(t,y)}) = \Pr(X_s \le x, X_t \le y)$. Then, G has L_2 distance τ, and by (21) and Theorem 4, it is uniformly continuous on $(E_0 \times Q, \tau)$. Hence if $\{H_{(s,x)} : (s,x) \in E_0 \times Q\}$ is a centered Gaussian process with

$$\mathbb{E}\big(H_{(s,x)}H_{(t,y)}\big) = \Pr(X_s \le x, X_t \le y) - \Pr(X_s \le x)\Pr(X_t \le y),$$

then

$$\mathbb{E}\big(\big(H_{(s,x)} - H_{(t,y)}\big)^2\big) = \tau^2((s,x),(t,y)) - \big(\Pr(X_s \le x) - \Pr(X_t \le y)\big)^2.$$

Hence the L_2 distance of H is smaller than that of G, and therefore Proposition 1 implies the process H is uniformly continuous on $(E_0 \times Q, d_H)$. By Lemma 2 the set $E_0 \times Q$ is dense in $(E \times \mathbb{R}, \tau)$, and since

$$d_H((s,x),(t,y)) \le \tau((s,x),(t,y)),$$

we also have that $E_0 \times Q$ is dense in $(E \times \mathbb{R}, d_H)$. Thus the Gaussian process $\{H_{(s,x)} : (s,x) \in E \times \mathbb{R}\}$ has a uniformly continuous version, which we also denote by H, and since (E, d_H) is totally bounded, the sample functions are bounded on E with probability one.

If $\mathcal{F} = \{I_{X_s\le x} : (s,x) \in E \times \mathbb{R})\}$, then since

$$d_H((s,x),(t,y)) = \rho_P(I_{X_s\le x}, I_{X_t\le y}),$$

the continuity of H on $(E \times \mathbb{R}, d_H)$ implies condition (ii) in Theorem 2 is satisfied. Since $I_{X_t\le y}$ is bounded, condition (i) in Theorem 2 is also satisfied. Therefore, Theorem 3 follows once we verify condition (iii) of Theorem 2.

To verify (iii) we use Lemma 4. As before, we identify the function $f = I_{X_s\le x} \in \mathcal{F}$ with the point $(s,x) \in E \times \mathbb{R}$. Hence, for the centered Gaussian process

$$\{G_f : f \in \mathcal{F}\}$$

in (iii) of Theorem 2, for $(s,x) \in E \times \mathbb{R}$, we take the process

$$\tilde{G}_{(s,x)} = G_{(s,x)} + \tilde{H}_s.$$

In our definition of $\tilde{G}$ we are assuming:

(a) $\{\tilde{H}_s : s \in E\}$ is a Gaussian process whose law is that of the process $\{H_s : s \in E\}$ given in the theorem, and independent of everything in our empirical model, and

(b) $\{G_{(s,x)} : (s,x) \in E \times \mathbb{R}\}$ is a uniformly continuous and sample bounded version of the Gaussian process, also denoted by $G_{(s,x)}$, but defined above on $E_0 \times Q$. The extension to all of $E \times \mathbb{R}$ again follows by the fact that $E_0 \times Q$ is dense in $(E \times \mathbb{R}, \tau)$.

Therefore, $\tilde{G}$ is sample bounded and uniformly continuous on $E \times \mathbb{R}$ with respect to its L_2 distance

$$d_{\tilde{G}}((s,x),(t,y)) = \{\tau^2((s,x),(t,y)) + \rho^2(s,t)\}^{1/2}.$$

Condition (iii) of Theorem 2 now follows easily from Lemma 4 since for (t,y) fixed,

$$\{(s,x) : \lambda((s,x),(t,y)) \le \varepsilon\} \supseteq \{(s,x) : d_{\tilde{G}}((s,x),(t,y)) \le \varepsilon\},$$

and for a random variable Z bounded by one, we have

$$\sup_{t>0} t^2 \Pr(|Z| > t) \le \Pr(|Z| > 0). \qquad \square$$

7. Another theorem and some examples. Let $\{X_t : t \in E\}$ be a sample continuous process such that:

(I) $\sup_{t \in E} |F_t(x) - F_t(y)| \le k|x-y|^\beta$ for all $x, y \in \mathbb{R}$ and some $k < \infty$ and some $\beta \in (0,1]$. Note that this condition implies that for every t, F_t is continuous and hence that $\tilde{F}_t = F_t$.

(III) $|X_t - X_s| \le \Gamma\phi(s,t)$ for all $s, t \in E$, and for some $\eta > 0$ and all $x \ge x_0$

$$\Pr(\Gamma \ge x) \le x^{-\eta}.$$

(III) For β as in (I), and η as in (II), there exists $\alpha \in (0, \beta/2)$ such that

$$\eta\left(\frac{1}{\alpha} - \frac{2}{\beta}\right) \ge 2$$

and $(\phi(s,t))^\alpha \le \rho(s,t)$, $s,t \in E$, where $\rho(s,t)$ is the L_2 distance of a sample bounded, uniformly continuous, centered Gaussian process on (E,ρ), which we denote by $\{H(t) : t \in E\}$.

THEOREM 5. *Let $\{X_t : t \in E\}$ be a sample continuous process satisfying* (I)–(III) *above. Then*

$$\{I_{X_s \le x} - \Pr(X_s \le x) : s \in E, x \in \mathbb{R}\}$$

satisfies the central limit theorem in $\ell_\infty(E \times \mathbb{R})$. This CLT also holds under (I)–(II), *provided* (II) *is strengthened to hold for all $\eta > 0$ and $x \ge x_\eta$, and for some $\alpha \in (0, \beta/2)$, we have $(\phi(s,t))^\alpha \le \rho(s,t)$, $s,t \in E$, where $\rho(s,t)$ is as in* (III).

REMARK 8. If the process $\{X_t : t \in E\}$ in Theorem 5 is a Gaussian process, then the CLT of Theorem 5 holds provided (I) is satisfied, (II) is such that $|X_t - X_s| \le \Gamma\phi(s,t)$ for all $s, t \in E$ and $\Gamma < \infty$ and and for some $\alpha \in (0, \beta/2)$, we have $(\phi(s,t))^\alpha \le \rho(s,t)$, $s, t \in E$, where $\rho(s,t)$ is as in (III).

PROOF OF THEOREM 5. The theorem follows by verifying the L condition in Theorem 3 with respect to ρ and $\{H(t) : t \in E\}$ as given in (III). Since (I) implies the distribution functions F_t are all continuous, the distributional transforms in (10) are simply the distributions themselves. Therefore, applying (I), with α and ρ as given in (III), we have for all $t \in E$ that

$$\begin{aligned} &\Pr{}^*\Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} |F_t(X_s) - F_t(X_t)| \ge \varepsilon^2\Big) \\ &\qquad \le \Pr{}^*\bigg(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} |X_s - X_t| \ge \bigg(\frac{\varepsilon^2}{k}\bigg)^{1/\beta}\bigg). \end{aligned} \tag{22}$$

Hence (II) implies

$$\begin{aligned} &\Pr{}^*\Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} |F_t(X_s) - F_t(X_t)| \ge \varepsilon^2\Big) \\ &\qquad \le \Pr\bigg(\Gamma \ge \bigg(\frac{\varepsilon^2}{k}\bigg)^{1/\beta} \Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} \phi(s,t)\Big)^{-1}\bigg) \end{aligned} \tag{23}$$

and since $\alpha > 0$ is such that $\eta(\frac{1}{\alpha} - \frac{2}{\beta}) \ge 2$ and

$$\Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} \phi(s,t)\Big)^{-1} \ge \Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} \rho(s,t)\Big)^{-1/\alpha} \ge \varepsilon^{-1/\alpha},$$

(III) therefore implies that

$$\sup_{t\in E} \Pr{}^*\Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}} |F_t(X_s) - F_t(X_t)| \ge \varepsilon^2\Big) \le k^{\eta/\beta}\varepsilon^{-2\eta/\beta}\varepsilon^{\eta/\alpha} \le k^{\eta/\beta}\varepsilon^2, \tag{24}$$

provided $0 < \varepsilon < \varepsilon_0$ is sufficiently small to imply $k^{-1/\beta}\varepsilon^{2/\beta - 1/\alpha} > x_0$. To obtain the final conclusion of the theorem assume $\alpha \in (0, \beta/2)$ and η is sufficiently large that $\eta(1/\alpha - 2/\beta) > 2$. Then, for $0 < \varepsilon < \varepsilon_\eta$ sufficiently small that $k^{-1/\beta}\varepsilon^{2/\beta-1/\alpha} > x_\eta$ we again have (24). Since these estimates are uniform in $\varepsilon \in (0, \varepsilon_0 \wedge \varepsilon_\eta)$, (24) then implies the L condition, and the proof is complete. $\square$

COROLLARY 2. *Let $\{Y_t : t \in [0, T]\}$ be a sample continuous γ-fractional Brownian motion for $0 < \gamma < 1$ such that $Y_0 = 0$ with probability one, and set $X_t = Y_t + Z$, where Z is independent of $\{Y_t : t \in [0, T]\}$ and has a bounded density function. Then,*

$$\{I_{X_s \le x} - \Pr(X_s \le x) : s \in [0, T], x \in \mathbb{R}\}$$

satisfies the central limit theorem in $\ell_\infty([0, T] \times \mathbb{R})$.

REMARK 9. The addition of the random variable Z in the previous corollary implies the densities of $Y_t + Z, t \in E$ are all bounded by the same bound as that of the density of Z, and hence condition (I) holds with $\beta = 1$. In particular, Z is not used in any other way. Furthermore, below we will see that something of this sort is necessary, since we will show that the CLT of the previous corollary fails for the fractional Brownian motion process Y itself, that is, when $Z = 0$.

PROOF OF COROLLARY 2. The L_2 distance for $\{X_t : t \in [0, T]\}$ is given by

$$\mathbb{E}\big((X_s - X_t)^2\big)^{1/2} = c_\gamma |s - t|^\gamma, \qquad s, t \in [0, T], \tag{25}$$

and without loss of generality we may assume the process to be normalized so that $c_\gamma = 1$. Furthermore, it is well known that these processes are Hölder continuous on $[0, T]$; that is, for every $\theta < \gamma$ we have

$$|X_t - X_s| \le \Gamma |t - s|^\theta, \qquad s, t \in [0, T], \tag{26}$$

where

$$\mathbb{E}(\exp\{c\Gamma^2\}) < \infty \tag{27}$$

for some $c > 0$. That Γ has exponential moments is due to the Fernique–Landau–Shepp theorem, and hence the corollary follows as in Remark 8, provided we take the Gaussian process H to be an $\alpha\theta$ fractional Brownian motion for any fixed $\theta < \gamma$ and any fixed $\alpha \in (0, \frac{1}{2})$ as $\beta = 1$. Hence the corollary is proven. $\square$

COROLLARY 3. *Let $I = [0, T]$ and $\{Y_{(s,t)} : (s, t) \in I \times I\}$ be a sample continuous Brownian sheet, that is, the centered Gaussian process on $I \times I$ with covariance $E(Y_{(s,t)} Y_{(u,v)}) = (s \wedge u)(t \wedge v)$ such that with probability one $Y_{(0,t)} = Y_{(s,0)} = 0$ for $s, t \in I$. Also, set $X_{(s,t)} = Y_{(s,t)} + Z$, where Z is independent of $\{Y_{(s,t)} : (s, t) \in I \times I\}$ and has a bounded density function. Then,*

$$\{I_{X_{(s,t)} \le x} - \Pr(X_{(s,t)} \le x) : (s, t) \in I \times I, x \in \mathbb{R}\}$$

satisfies the central limit theorem in $\ell_\infty((I \times I) \times \mathbb{R})$.

PROOF. First of all observe that since Z has a bounded density, and is independent of the Brownian sheet Y, we have (I) holding with $\beta = 1$. Furthermore, from Theorem 1 in the paper Yeh (1960), these processes are Hölder continuous on $I \times I$; that is, for $(s, t), (u, v) \in I \times I$ and $0 < \gamma < 1/2$, we have

$$\big|X_{(s,t)} - X_{(u,v)}\big| \le \Gamma \phi((s, t), (u, v)), \tag{28}$$

where

$$\phi((s, t), (u, v)) = \left(\left(\frac{u - s}{T}\right)^2 + \left(\frac{v - t}{T}\right)^2\right)^{\gamma/2}$$

and

$$\mathbb{E}(\exp\{c\Gamma^2\}) < \infty \tag{29}$$

for some $c > 0$. That Γ has exponential moments is due to the Fernique–Landau–Shepp theorem, and hence the corollary will follow as in Remark 8, provided we take the Gaussian process $H_{(s,t)}$ to be

$$H_{(s,t)} = Y_s + Z_t, \qquad (s,t) \in I \times I, \tag{30}$$

where the processes $\{Y_s : s \in I\}$ and $\{Z_t : t \in I\}$ are independent θ-fractional Brownian motions. To determine θ we fix $\gamma = 1/4$, and normalizing the Y_s and Z_t processes suitably, we have

$$\rho^2((s,t),(u,v)) = \mathbb{E}\big(\big(H_{(u,v)} - H_{(s,t)}\big)^2\big) = \left(\left|\frac{u-s}{T}\right|^{2\theta} + \left|\frac{v-t}{T}\right|^{2\theta}\right).$$

Hence for any $\alpha \in (0, 1/2)$ and $\gamma = 1/4$, we take $\theta \in (0, \alpha/4)$, which implies that

$$\phi^\alpha((s,t),(u,v)) \le \rho((s,t),(u,v)).$$

Since each such θ yields suitable fractional Brownian motion choices for Y and Z, the corollary is proven. $\square$

8. Examples where our CLT fails.

8.1. *Fractional Brownian motions.* Since the class of sets in our CLT arises using the Vapnik–Cervonenkis class of half lines, one might think that perhaps if i.i.d. copies of the process $\{X_t : t \in E\}$ satisfied the CLT in $C(E)$, then the class of sets $\mathcal{C}$ of Theorem 3 would satisfy the CLT(P). Our first example shows this fails, even if the process X_t is Brownian motion on $[0,1]$ tied down at $t = 0$. In this example the process fails condition (I) in Theorem 5 since $\Pr(X_0 = 0) = 1$. To prove this we show the necessary condition for $\mathcal{C}$ to satisfy the CLT(P) appearing in (ii)(a) of Theorem 1 fails. More precisely, since measurability is an issue here, the next lemma shows that there is a countable subclass $\mathcal{C}_Q$ of sets in $\mathcal{C}$ such that by Theorem 3 of Talagrand (1988), $\mathcal{C}_Q \notin$ CLT(P). Thus $\mathcal{C}$ fails the CLT(P), as otherwise all subclasses also are in CLT(P).

LEMMA 7. *Let $\mathcal{C} = \{C_{t,x} : 0 \le t \le 1, -\infty < x < \infty\}$, where $C_{t,x} = \{z \in C[0,1] : z(t) \le x\}$, and assume $\{X_t : t \in [0,1]\}$ is a sample continuous Brownian motion tied down at zero. Also, let $\mathcal{C}_Q$ denote the countable subclass of $\mathcal{C}$ given by $\mathcal{C}_Q = \{C_{t,y} \in \mathcal{C} : t, y \in Q\}$, where Q denotes the rational numbers. Then, for each integer $n \ge 1$, with probability one*

$$\Delta^{\mathcal{C}_Q}(\{B_1, \dots, B_n\}) = 2^n,$$

where $\Delta^{\mathcal{C}_Q}(\{B_1, \dots, B_n\}) = \mathrm{card}\{C \cap \{B_1, \dots, B_n\} : C \in \mathcal{C}_Q\}$, and $B_1, \dots, B_n$ are independent copies of $\{X_t : t \in [0,1]\}$.

PROOF. Fix k, $1 \le k \le n$, and integers $1 \le j_1 < \cdots < j_k \le n$. The first thing we want to show is that with probability one there are suitable $C_{t,x} \in \mathcal{C}_Q$ such that the k functions $\{B_{j_1}, \ldots, B_{j_k}\} = C_{t,x} \cap \{B_1, \ldots, B_n\}$. Of course, since the functions $\{B_{j_1}, \ldots, B_{j_k}\}$ are random, the choice of $C_{t,x}$ also may need be random, and for this we use the law of the iterated logarithm (LIL). This will show that with probability one all nonempty subsets of $\{B_1, \ldots, B_n\}$ are in $\Delta^{\mathcal{C}_Q}(\{B_1, \ldots, B_n\})$, and to get the empty set with probability one is trivial, that is, the sample functions are continuous on [0, 1], but the choice of x in $C_{t,x}$ can be made arbitrarily negative. Now for the details.

Let

$$\mathbf{u} = (u_1, \ldots, u_n),$$

where $u_{j_1} = u_{j_2} = \cdots = u_{j_k} = 1$ and all other $u_j = 2$. Then $\|\mathbf{u}\| = (\sum_{j=1}^n u_j^2)^{1/2} = (4n - 3k)^{1/2}$. Now set $\mathbf{v} = (v_1, \ldots, v_n)$, where

$$v_{j_1} = v_{j_2} = \cdots = v_{j_k} = \frac{1}{2(4n - 3k)^{1/2}}$$

and all other $v_j = \frac{1}{(4n-3k)^{1/2}}$. Then $\mathbf{v} = \mathbf{u}/(2\|\mathbf{u}\|)$ and $\|\mathbf{v}\| = 1/2$.

For $x > 0$, let $Lx = \log_e x$ and set

$$W(s) = \frac{(B_1(s), \ldots, B_n(s))}{(2sLL(1/s))^{1/2}}$$

for $0 < s \le 1$. Then the multi-dimensional compact LIL implies with probability one that

$$\liminf_{s \downarrow 0} \|\mathbf{v} - W(s)\| = 0,$$

and hence with probability one there are infinitely many rational numbers $t \downarrow 0$ such that

$$C_{t,x(t)} \cap \{B_1, \ldots, B_n\} = \{B_{j_1}, \ldots, B_{j_k}\},$$

where $x(t) \in Q$ for $t \in Q$ and

$$\left| x(t) - \frac{3(2tLL(1/t))^{1/2}}{4(4n - 3k)^{1/2}} \right| < \frac{(2tLL(1/t))^{1/2}}{16(4n - 3k)^{1/2}}.$$

Since k and the set $\{j_1, \ldots, j_k\}$ were arbitrary, and with probability one we can pick out the subset $\{B_{j_1}, \ldots, B_{j_k}\}$; the lemma follows as the intersection of 2^n subsets of probability one has probability one. □

The failure of the CLT also holds for all sample continuous fractional Brownian motions $\{X_H(t) : t \in [0, 1]\}$ which are tied down at zero. The proof of this again depends on the law of the iterated logarithm for n independent copies of this process at $t = 0$, which then allows us to prove an analog of the previous lemma.

The LIL result at $t = 0$ for a single copy follows, for example, by Theorem 4.1 of Goodman and Kuelbs (1991), and then one can extend that result to n independent copies by classical proofs as in Kuelbs (1976). The details of this last step are lengthy, but at this stage are more or less routine in the subject, and hence are omitted. Of course, the CLT for i.i.d. copies of these processes is obvious as they are Gaussian.

8.2. *A uniform CLT example*. In the previous examples, when the distribution function F_t of X_t jumped, the oscillation of the processes at that point caused a failure of our CLT. Hence one other possible idea is that if the process $\{X_t : t \in [0, 1]\}$ is Lip-1 on $[0, 1]$, then our CLT might hold. For example, this is true for the Lip-1 process $X_t = tU, t \in [0, 1]$, where U is uniform on $[0, 1]$. Moreover, in this example the densities of F_t still are unbounded near $t = 0$.

To see this let $X_{t,j}, j = 1, \dots, n, t \in [0, 1]$, be i.i.d. copies of $X_t = tU, t \in [0, 1]$, and define

$$W_n(C_{t,y}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} [I(X_{(\cdot),j} \in C_{t,y}) - \Pr(X_{(\cdot),j} \in C_{t,y})],$$

where $\mathcal{C} = \{C_{t,y} : t \in [0, 1], y \in \mathbb{R}\}$ and $C_{t,y} = \{z \in C[0, 1] : z(t) \le y\}$. Therefore, $W_n(C_{t,y}) = 0$ for all $y \in \mathbb{R}$ when $t = 0$, and also when $y/t \ge 1$. Moreover, if we define $\mathcal{G} = \{(-\infty, r] : 0 \le r \le 1\}$, and

$$\phi(I_{C_{t,y}}) = I_{(-\infty,1]} \qquad \text{if } y/t > 1, 0 \le t \le 1, \text{ or } y = 0 \text{ and } t = 0,$$

$$\phi(I_{C_{t,y}}) = I_{(\infty,0]} \qquad \text{if } y/t \le 0 \text{ but not } y = 0 \text{ and } t = 0$$

and

$$\phi(I_{C_{t,y}}) = I_{(-\infty,y/t]} \qquad \text{if } 0 < y/t < 1, 0 < y < 1, 0 < t < 1.$$

Then $\phi(\mathcal{C}) = \{I_{U \le r} : 0 \le r \le 1\} \equiv \mathcal{G}$ and ϕ maps L_2 equivalence classes of $\mathcal{C}$ onto $\mathcal{G}$ with respect to the law of $\{X_t : t \in [0, 1]\}$ for sets in $\mathcal{C}$, and the law of U for sets in $\mathcal{G}$. Now $\mathcal{G}$ satisfies the CLT$(\mathcal{L}(U))$ by the classical empirical CLT [e.g., see Theorem 16.4 of Billingsley (1968)], and since ϕ preserves covariances we thus have $W_n(C_{t,y})$ converges weakly to the Gaussian centered process $W(C_{t,y}) = Y(\phi(C_{t,y}))$ on $\mathcal{C}$, where $Y((-\infty, s]) = B(s) - sB(1)$ is the tied down Wiener process on $[0, 1]$; that is, $B(\cdot)$ is a Brownian motion.

8.3. *A Lip-*1 *example without the CLT*. In this example we see that the Lip-1 property for $\{X_t : t \in [0, 1]\}$ is not always sufficient for our CLT. Here $X_0 = 0$, and for $0 < t \le 1$, we define

$$X_t = t \sum_{j=1}^{\infty} (\alpha_j(t) + 2) I_{E_j}(t) U,$$

where:

(i) $E_j = (2^{-j}, 2^{-(j-1)}]$ for $j \geq 1$.

(ii) $\{\alpha_j(t) : j \geq 1\}$ are independent random processes with $\alpha_j(\cdot)$ defined on E_j such that for $j \geq 1$,

$$\Pr\bigl(\alpha_j(t) = \sin(2\pi 2^j t), t \in E_j\bigr) = 1/2$$

and

$$\Pr\bigl(\alpha_j(t) = \sin(2\pi 2^{j+1} t), t \in E_j\bigr) = 1/2.$$

(iii) U is a uniform random variable on $[3/2, 2]$, independent of the $\{\alpha_j\}$.

Since the α_j's are zero at endpoints of the E_j and we have set $X(0) = 0$, it is easy to see $X(t)$ has continuous paths on $[0, 1]$. Moreover, $X(t)$ is Lip-1 on $[0, 1]$, and $X(t)$ has a density for each $t \in (0, 1]$, but our CLT fails.

The failure of the empirical CLT can be shown by verifying a lemma of the sort we have above for Brownian motion, and again we see the lack of uniformly bounded densities is a determining factor.

For each integer $n \geq 1$, let $X_1, \dots, X_n$ be independent copies of X, and again take $\mathcal{C} = \{C_{t,x} : 0 \leq t \leq 1, \infty < x < \infty\}$, where $C_{t,x} = \{z \in C[0, 1] : z(t) \leq x\}$. Also, define $\mathcal{C}_Q$ as in Lemma 7. Then, we have the following lemma, and combined with the argument in Section 8.1, we see the empirical CLT fails when this $X(\cdot)$ is used.

LEMMA 8. *For each integer $n \geq 1$, with probability one*

$$\Delta^{\mathcal{C}_Q}(\{X_1, \dots, X_n\}) = 2^n,$$

where $\Delta^{\mathcal{C}_Q}(\{X_1, \dots, X_n\}) = \operatorname{card}\{C \cap \{X_1, \dots, X_n\} : C \in \mathcal{C}_Q\}$.

PROOF. As in the proof of Lemma 7, assume one wants the k functions $\{X_{i_1}, \dots, X_{i_k}\}$ with probability one, where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. If we write

$$X_i(t) = t \sum_{j=1}^{\infty} \bigl(\alpha_{i,j}(t) + 2\bigr) I_{E_j}(t) U_i,$$

where the $\{\alpha_{i,j} : j \geq 1\}$ are independent copies of $\{\alpha_j(t) : j \geq 1\}$ and $\{U_i : i \geq 1\}$ are independent copies of U, independent of all the $\alpha_{i,j}$'s, then this can be arranged by taking

$$\alpha_{i,j}(t) = \sin(2\pi 2^{j+1} t), \qquad i \in \{i_1, \dots, i_k\},$$

and

$$\alpha_{i,j}(t) = \sin(2\pi 2^j t), \qquad i \in \{1, \dots, n\} \cap \{i_1, \dots, i_k\}^c,$$

provided we set $t = t_j = 2^{-j} + \frac{1}{4}(2^{-(j-1)} - 2^{-j})$. The probability of this configuration on the interval E_j is $1/2^n$, and hence with probability one the Borel–Cantelli lemma implies there are infinitely many (random) $\{t_j \downarrow 0\}$ such that

$$\alpha_{i_1,j}(t_j) = \cdots = \alpha_{i_k,j}(t_j) = 0$$

and

$$\alpha_{i,j}(t_j) = 1$$

for all other $i \in \{1, \ldots, n\}$. Thus with probability one there are infinitely many rational numbers $t \downarrow 0$ such that

$$C_{t,x(t)} \cap \{X_1, \ldots, X_n\} = \{X_{i_1}, \ldots, X_{i_k}\},$$

provided

$$x = x(t) = \frac{17t}{4}.$$

Of course, $x(t)$ is then also in Q, and since k and the set $\{i_1, \ldots, i_k\}$ were arbitrary, and we can pick out $\{X_{i_1}, \ldots, X_{i_k}\}$ with probability one using sets in $\mathcal{C}_Q$, the lemma follows as the intersection of 2^n subsets of probability one has probability one.

To see $X(t)$ is Lip-1 on [0, 1], observe that the intervals $\{E_j : j \geq 1\}$ are disjoint, $X(t)$ is differentiable on their interiors and an easy computation implies

$$\sup_{j \geq 1} \sup_{2^{-j} < t < 2^{-(j-1)}} |X'(t)| \leq [4\pi + 3]U.$$

Furthermore, $X(t)$ is continuous on [0, 1], so the mean value theorem and an elementary argument shows $X(t)$ is Lip-1 on [0, 1], with Lipschitz constant bounded by $(4\pi + 3)U$ with probability one. Furthermore, since U is uniform on $[3/2, 2]$ and independent of the $\{\alpha_j\}$, then $X(t)$ has a density for each $t \in (0, 1]$. $\square$

8.4. *Variations of the L condition and the CLT.* Here we produce examples where the sets $\mathcal{C}$, or more precisely the class of indicator functions given by $\mathcal{C}$, are P-pre-Gaussian, and yet $\mathcal{C} \notin \mathrm{CLT}(P)$. More importantly, they also satisfy the modified L condition, that is, we say $\{X_s : s \in E\}$ satisfies the modified L condition if for all $\varepsilon > 0$, there exists $L < \infty$ such that

$$\sup_{t \in E} \Pr{}^*\Big(\sup_{\rho(s,t) \leq \varepsilon} |F_t(X_s) - F_t(X_t)| > \varepsilon^2\Big) \leq L\varepsilon^2. \tag{31}$$

Of course, if the distribution functions $\{F_t : t \in E\}$ are all continuous, this is the L condition. Hence these examples also provide motivation for the use of the distributional transforms in the L condition of (10) used in Theorem 3.

Notation for the examples in this subsection is as follows. Let $E = \{1, 2, 3, \ldots\}$, and assume $D(E) = \{z : z(t) = 0 \text{ or } 1, t \in E\}$ with $\mathcal{C} = \{C_{t,y} : t \in E, y \in \mathbb{R}\}$, where $C_{t,y} = \{z \in D(E) : z(t) \leq y\}$. Then, since the functions in $D(E)$ take only the values zero and one, we have

$$\mathcal{C} = \mathcal{C}_0 \cup \{\{D(E)\}\},$$

where $\mathcal{C}_0 = \{\tilde{C}_{t,0} : t \in E\}$ and for $t \in E$, $\tilde{C}_{t,0} = \{z \in D(E) : z(t) = 0\}$. Also, let Σ denote the minimal sigma-field of subsets of $D(E)$ containing $\mathcal{C}$, and let P denote the probability on $(D(E), \Sigma)$ such that $\Pr(\tilde{C}_{t,0}) = p_t$ and the events $\{\tilde{C}_{t,0} : t \in E\}$ are independent events, that is, P is a product measure on the coordinate spaces of $D(E) = \{0, 1\}^{\mathbb{N}}$ with the tth coordinate of $D(E)$ having the two point probability that puts mass p_t on zero, and $1 - p_t$ on one.

PROPOSITION 2. *Let $\mathcal{C}$ be defined as above. Then*:

(i) *$\mathcal{C}$ is P-pre-Gaussian whenever $p_t = o((\log(t+2))^{-1})$ as $t \to \infty$.*
(ii) *$\mathcal{C} \in \mathrm{CLT}(P)$ if and only if for some $r > 0$,*

$$\sum_{t=1}^{\infty} \bigl(p_t(1 - p_t)\bigr)^r < \infty.$$

(iii) *If $p_t = (\log(t+2))^{-2}$, and $\{H(t) : t \in E\}$ consists of centered independent Gaussian random variables with $\mathbb{E}(H(t)^2) = (\log(t+2))^{-3/2}$, then $\{X_s : s \in E\}$ satisfies the modified L condition, $\mathcal{C}$ is P-pre-Gaussian and $\mathcal{C} \notin \mathrm{CLT}$. In particular, in view of Theorem 3, it does not satisfy the L condition.*

PROOF. Since $\mathcal{C}$ differs from $\mathcal{C}_0$ by the single set $D(E)$ and $P(D(E)) = 1$, it is easy to see that $\mathcal{C} \in \mathrm{CLT}(P)$ if and only if $\mathcal{C}_0 \in \mathrm{CLT}(P)$. Therefore, since the events of $\mathcal{C}_0$ are independent, Theorem 3.9.1 in Dudley [(1999), page 122] implies that $\mathcal{C}_0 \in \mathrm{CLT}(P)$ if and only if for some $r > 0$,

$$\sum_{t=1}^{\infty} \bigl(p_t(1 - p_t)\bigr)^r < \infty.$$

Hence (ii) holds.

Now the centered Gaussian process $\{G_P(C) : C \in \mathcal{C}_0\} = \{G_P(C_{t,0}) : t \in E\}$, and since the random variables $\{G_P(C_{t,0}) : t \in E\}$ are mean zero and $\mathbb{E}(G_P(C_{t,0})^2) = p_t(1 - p_t)$, we have $\mathcal{C}_0$ is P-pre-Gaussian provided

$$p_t = o\bigl((\log(t+2))^{-1}\bigr) \qquad \text{as } t \to \infty.$$

Hence $\mathcal{C}$ is P-pre-Gaussian whenever $p_t = o((\log(t+2))^{-1})$ as $t \to \infty$, and (i) holds. To verify (iii) we take $p_t = (\log(t+2))^{-2}$, and $\{H(t) : t \in E\}$ to be centered independent Gaussian random variables with $\mathbb{E}(H(t)^2) = (\log(t+2))^{-3/2}$. If $\rho^2(s,t) = \mathbb{E}((H(s) - H(t))^2)$, then, for $s \neq t$,

$$\begin{aligned}
\rho^2(s,t) &= \bigl(\log(s+2)\bigr)^{-3/2} + \bigl(\log(t+2)\bigr)^{-3/2} \\
&\geq \max\bigl\{\bigl(\log(s+2)\bigr)^{-3/2}, \bigl(\log(t+2)\bigr)^{-3/2}\bigr\}.
\end{aligned}$$

In addition, we have $|F_t(X_s) - F_t(X_t)| = 0$ if $X_s = X_t = 0$ or $X_s = X_t = 1$, and $|F_t(X_s) - F_t(X_t)| = p_t$ if $X_t \neq X_s$.

Therefore, for all $t \in E$ fixed and $\varepsilon > 0$, we have

$$\Pr{}^*\Bigl(\sup_{\{s : \rho(s,t) \leq \varepsilon\}} |F_t(X_s) - F_t(X_t)| > \varepsilon^2\Bigr) = 0 \qquad \text{if } p_t \leq \varepsilon^2$$

and

$$\begin{aligned}\Pr^*\Big(&\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}}|F_t(X_s)-F_t(X_t)|>\varepsilon^2\Big)\\&\le\Pr^*\Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}}I_{X_t\ne X_s}>0\Big)\qquad\text{for } p_t>\varepsilon^2.\end{aligned}$$

Of course, E countable makes the outer probabilities in the above, ordinary probabilities, but for simplicity we retained the outer probability notation used in (9) and (10). Now $\rho(s,t)\le\varepsilon$ implies

$$\max\{(\log(s+2))^{-3/4},(\log(t+2))^{-3/4}\}\le\varepsilon.$$

Thus if $p_t=(\log(t+2))^{-2}>\varepsilon^2$, we have $\{\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}}I_{X_t\ne X_s}>0\}=\varnothing$. Combining the above we have for each fixed $t\in E$ and $\varepsilon>0$ that

$$\Pr^*\Big(\sup_{\{s\,:\,\rho(s,t)\le\varepsilon\}}|F_t(X_s)-F_t(X_t)|>\varepsilon^2\Big)=0,$$

and hence the modified L condition for $\{X_s:s\in E\}$ holds. Thus (iii) follows. □

APPENDIX: TALAGRAND'S CONTINUITY RESULT FOR GAUSSIAN PROCESSES

The proof of Theorem 4 in Section 5 is as follows.

PROOF OF THEOREM 4. First we will show (i) and (ii) are equivalent. If (ii) holds, then by Fatou's lemma we have

$$0=\lim_{n\to\infty}\mathbb{E}\Big(\sup_{d(s,t)\le 1/n}|X_s-X_t|\Big)\ge\mathbb{E}\Big(\liminf_{n\to\infty}\sup_{d(s,t)\le 1/n}|X_s-X_t|\Big).$$

Thus, with probability one

$$\liminf_{n\to\infty}\sup_{d(s,t)\le 1/n}|X_s-X_t|=0,$$

and since the random variables $\sup_{d(s,t)\le 1/n}|X_s-X_t|$ decrease as n increases, this implies with probability one

$$\lim_{n\to\infty}\sup_{d(s,t)\le 1/n}|X_s-X_t|=0,$$

which implies (i).

If we assume (i), then since (T,d_X) is assumed totally bounded, we have

$$Z=\sup_{t\in T}|X_t|<\infty$$

with probability one, and the Fernique–Landau–Shepp theorem implies Z is integrable. Since

$$\sup_{d(s,t)\le\varepsilon} |X_s - X_t| \le 2Z,$$

and (i) implies

$$\lim_{\varepsilon\to 0}\sup_{d(s,t)\le\varepsilon} |X_s - X_t| = 0$$

with probability one, the dominated convergence theorem implies (ii). Thus (i) and (ii) are equivalent.

Now we assume (i) and (ii), and choose $\varepsilon_k \downarrow 0$ such that

$$\sup_s \mathbb{E} \sup_{\{t\,:\,d(t,s)\le\varepsilon_k\}} X_t \le \sup_s \mathbb{E} \sup_{\{t\,:\,d(t,s)\le\varepsilon_k\}} (X_t - X_s) \le \mathbb{E} \sup_{d(s,t)\le\varepsilon_k} (X_t - X_s) \le 2^{-k}.$$

Since we are assuming (i) and that (T, d) is totally bounded, the sample paths of $\{X_t : t \in E\}$ are uniformly continuous and bounded on (T, d). Hence by Sudakov's inequality, if $N(T, d, \varepsilon)$ equals the minimal number of open balls of radius ε that cover (T, d), then

$$\lim_{\varepsilon\downarrow 0} \varepsilon(\log N(T, d, \varepsilon))^{1/2} = 0.$$

Therefore, we also are free to assume the ε_k are such that for all $k \ge 1$ we have $\varepsilon_k(\log N(\varepsilon_k))^{1/2} \equiv \varepsilon_k(\log(N(T, d, \varepsilon_k))^{1/2} < \frac{1}{2}$. Moreover, since (T, d) is totally bounded, and (i) holds, by Theorem 2.1.1 of Talagrand (2005) there exists an admissible sequence of partitions $\{\tilde{\mathcal{A}}_n : n \ge 0\}$ of (T, d) such that for a universal constant L we have

$$\frac{1}{2L}\sup_{t\in T}\sum_{n\ge k} 2^{n/2}\Delta(A_n(t)) \le \mathbb{E}\Big(\sup_{t\in T} X_t\Big).$$

Now choose $\{n_k : k \ge 1\}$ to be a strictly increasing sequence of integers such that $n_1 > 4$ and

$$2^{\sum_{2\le j\le k} 1/\varepsilon_j^2} \le n_{k-1}. \tag{32}$$

Based on the n_k's we define an increasing sequence of partitions, $\mathcal{B}_n$. For $0 \le n \le n_1$ we let $\mathcal{B}_n = \tilde{\mathcal{A}}_n$. For $n_1 < n \le n_2$ we proceed as follows.

First we choose a maximal set $\{s_1, \dots, s_{N(\varepsilon_2)}\}$ of (T, d) for which $d(s_i, s_j) \ge \varepsilon_2$. Furthermore, by our choice of $\{\varepsilon_k : k \ge 1\}$ via Sudakov's inequality, we have that $N(\varepsilon_2) \le 2^{1/\varepsilon_2^2}$. To define the partitions for $n_1 < n \le n_2$ we next consider the partition of T formed by the sets

$$C_j = B(s_j, \varepsilon_2) \cap \left(\bigcup_{k=1}^{j-1} B(s_k, \varepsilon_2)\right)^c, \qquad 1 \le j \le N(\varepsilon_2), \tag{33}$$

and the sets $B(s,\varepsilon)$ are ε balls centered at s. Then by Theorem 2.1.1 of Talagrand (2005) for every integer $1 \le j \le N(\varepsilon_2)$ there exists an admissible sequence of partitions for (C_j, d), which we denote by $\mathcal{B}^{s_j}_{n_1,n}$, such that

$$2^{-2} \ge \mathbb{E} \sup_{\{t \in C_j\}} X_t \ge \frac{1}{2L} \sup_{\{t\in C_j\}} \sum_{n \ge 0} 2^{n/2} \Delta(B^{s_j}_{n_1,n}(t))$$

$$\ge \frac{1}{2L} \sup_{\{t\in C_j\}} \sum_{n_1 < n \le n_2} 2^{n/2} \Delta\big(A_{n_1}(t) \cap B^{s_j}_{n_1,n}(t)\big).$$

Since the sets C_j form a partition of T, if $B_{n_1,n}$ is one of the sets, $B^{s_j}_{n_1,n}(t)$, then

$$2^{-2} \ge \frac{1}{2L} \sup_{t \in T} \sum_{n_1 < n \le n_2} 2^{n/2} \Delta\big(A_{n_1}(t) \cap B_{n_1,n}(t)\big),$$

and we define the increasing sequence of partitions $\mathcal{B}_{n_1,n}$ to be all sets of the form $A_{n_1}(t) \cap B_{n_1,n}(t)$, where $t \in T$ and $B_{n_1,n} \in \mathcal{B}^{s_j}_{n_1,n}(t)$ for some $j \in [1, N(\varepsilon_2)]$. Furthermore, since the C_j's are disjoint, for $n_1 < n \le n_2$ we have

$$\text{Card}(\mathcal{B}_{n_1,n}) \le 2^{2^{n_1}} 2^{2^n} N(\varepsilon_2) \le 2^{2^{n+1}} 2^{1/\varepsilon_2^2} \le 2^{2^{n+1}} n_1 \le 2^{2^{n+1}} 2^{2^n} = 2^{2^{n+2}}, \tag{34}$$

and for $n_1 < n \le n_2$ we define $\mathcal{B}_n = \mathcal{B}_{n_1,n}$.

Iterating what we have done for $n_1 < n \le n_2$, we have increasing partitions $\mathcal{B}_{n_{k-1},n}$, $n_{k-1} < n \le n_k$, for which

$$(2L)2^{-k} \ge \sup_t \sum_{n \ge 0} 2^{n/2} \Delta(B_{n_{k-1},n})$$

$$\ge \sup_t \sum_{n_{k-1} < n \le n_k} 2^{n/2} \Delta\big(B_{n_{k-1},n}(t) \cap B_{n_{k-2},n_{k-1}}(t) \cap \cdots \cap B_{n_1,n_2}(t) \cap A_{n_1}(t)\big),$$

and for $n_{k-1} < n \le n_k$ we define $\mathcal{B}_n = \mathcal{B}_{n_{k-1},n}$. Therefore, we now have an increasing sequence of partitions $\{\mathcal{B}_n : n \ge 0\}$ such that

$$(2L)\sum_{k \ge r} 2^{-k} \ge \sum_{k \ge r} \sup_t \sum_{n_{k-1} < n \le n_k} 2^{n/2} \Delta\left(B_{n_{k-1},n}(t) \cap \left(\bigcap_{j=2}^{k-1} B_{n_{j-1},n_j}(t)\right) \cap A_{n_1}(t)\right)$$

$$\ge \sup_t \sum_{k \ge r} \sum_{n_{k-1} < n \le n_k} 2^{n/2} \Delta\left(B_{n_{k-1},n}(t) \cap \left(\bigcap_{j=2}^{k-1} B_{n_{j-1},n_j}(t)\right) \cap A_{n_1}(t)\right)$$

and, letting $B_n(t)$ denote the generic set of $\mathcal{B}_n$ containing t, we have

$$(2L)\sum_{k \ge r} 2^{-k} \ge \sup_t \sum_{k \ge r} \sum_{n_{k-1} < n \le n_k} 2^{n/2} \Delta(B_n(t)). \tag{35}$$

Now we count the number of elements in each partition. Since (1) holds, the partitions $\mathcal{B}_{n_{k-1},n}^{s_j}$ are assumed admissible, and the C_j's used at the subsequent iterations are always disjoint, we have for $n_{k-1} < n \le n_k$ that

$$\mathrm{Card}(\mathcal{B}_{n_{k-1},n}) \le 2^{2^{n_1}} \left[\prod_{j=2}^{k-1} 2^{2^{n_j}} N(\varepsilon_j)\right] 2^{2^n} N(\varepsilon_k) \le 2^{\sum_{1\le j\le k-1} 2^{n_j} + 2^n} n_{k-1} \le 2^{2^{n+2}}.$$

Given the increasing sequence of partitions $\{\mathcal{B}_n : n \ge 0\}$, we now define the partitions $\mathcal{A}_n$ to be the single set T for $n = 0, 1$ and $\mathcal{A}_n = \mathcal{B}_{n-2}$ for $n \ge 2$. Since we have $\mathrm{Card}(\mathcal{B}_n) \le 2^{2^{n+2}}$ for $n \ge 0$ we thus have that the $\mathcal{A}_n$'s are admissible and using (35) above they satisfy (iii).

Now assume (iii) holds and (T, d) is totally bounded. We give a sketch of the case (iii) implies (ii). Corollary 1.6 in Fernique (1985) reduces our task to showing

$$\lim_{\eta\to 0} \sup_t \mathbb{E} \sup_{s\in B_d(t,\eta)} (X_s - X_t) = 0 \tag{36}$$

and

$$\lim_{\delta\to 0} \delta^2 \log_2 N(T, \delta) = 0. \tag{37}$$

In the computation below the existence of K follows from Theorem 2.1.1 of Talagrand (2005). To show (36) we estimate

$$\begin{aligned}
\sup_t \mathbb{E} \sup_{s\in B_d(t,\eta)} (X_s - X_t) &= \sup_t \mathbb{E} \sup_{s\in B_d(t,\eta)} X_s \\
&\le K \sup_t \sum_{n\ge 0} 2^{n/2} \Delta\big(A_n(t) \cap B_d(t, \eta)\big) \\
&\le K \Bigg(\sup_t \sum_{0\le n\le k} 2^{n/2} \Delta\big(A_n(t) \cap B_d(t, \eta)\big) \\
&\qquad + \sup_t \sum_{k<n} 2^{n/2} \Delta\big(A_n(t) \cap B_d(t, \eta)\big)\Bigg) \\
&\le K \left((2\eta) C 2^{k/2} + \sup_t \sum_{k<n} 2^{n/2} \Delta\big(A_n(t) \cap B_d(t, \eta)\big)\right) \\
&\le K \left((2\eta) C 2^{k/2} + \sup_t \sum_{k<n} 2^{n/2} \Delta(A_n(t))\right),
\end{aligned}$$

where in the third inequality C is such that $\sum_{0\le n\le k} 2^{n/2} \le C 2^{k/2}$. Hence,

$$\lim_{\eta\to 0} \sup_t \mathbb{E} \sup_{s\in B_d(t,\eta)} (X_s - X_t) \le \sup_t K \sum_{k<n} 2^{n/2} \Delta(A_n(t)) \qquad \text{for every } k.$$

By the hypothesis, this last quantity converges to 0 as $k \to \infty$.

To handle (37), by the hypothesis we can choose k such that

$$\sup_{t\in T}\sum_{n\geq k}2^{n/2}\Delta(A_n(t))\leq\varepsilon. \tag{38}$$

Hence,

$$2^{k/2}\sup_{t\in T}\Delta(A_k(t))=2^{k/2}\sup_{B\in\mathcal{A}_k}\Delta(B)\leq\varepsilon.$$

For each $B\in\mathcal{A}_k$ pick a point, $t_{k,B}$. Let $\delta_k=2\sup_{B\in\mathcal{A}_k}\Delta(B)$. Then, if $B_d(t,\delta)=\{s\in T: d(s,t)\leq\delta\}$ and $\frac{2\varepsilon}{2^{k/2}}=\delta'_k$, we have

$$B(t_{k,B})\subseteq B_d(t_{k,B},\delta_k)\subseteq B_d(t_{k,B},\delta'_k)\qquad\text{for every } B\in\mathcal{A}_k.$$

Since $\mathcal{A}_k$ is a partition, $T=\bigcup_{B\in\mathcal{A}_k}B_d(t_{k,B},\delta'_k)$. Hence,

$$\log_2\mathcal{N}(T,\delta'_k)\leq\log_2(2^{2^k})=\left(\frac{2\varepsilon}{\delta'_k}\right)^2.$$

By interpolating we get (37). □

Acknowledgments. It is a pleasure to thank the referee for his careful reading of the manuscript. His comments and suggestions led to a number of refinements in the presentation.

REFERENCES

ANDERSEN, N. T., GINÉ, E., OSSIANDER, M. and ZINN, J. (1988). The central limit theorem and the law of iterated logarithm for empirical processes under local conditions. *Probab. Theory Related Fields* **77** 271–305. MR0927241

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. MR0233396

DUDLEY, R. M. (1999). *Uniform Central Limit Theorems. Cambridge Studies in Advanced Mathematics* **63**. Cambridge Univ. Press, Cambridge. MR1720712

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach. Probability and Mathematical Statistics* **1**. Academic Press, New York. MR0215390

FERNIQUE, X. (1975). Regularité des trajectoires des fonctions aléatoires gaussiennes. In *École D'Été de Probabilités de Saint-Flour, IV*-1974. *Lecture Notes in Math*. **480** 1–96. Springer, Berlin. MR0413238

FERNIQUE, X. (1985). Sur la convergence étroite des mesures gaussiennes. *Z. Wahrsch. Verw. Gebiete* **68** 331–336. MR0771470

GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes (with discussion). *Ann. Probab.* **12** 929–998. MR0757767

GOODMAN, V. and KUELBS, J. (1991). Rates of clustering for some Gaussian self-similar processes. *Probab. Theory Related Fields* **88** 47–75. MR1094077

KUELBS, J. (1976). A strong convergence theorem for Banach space valued random variables. *Ann. Probab.* **4** 744–771. MR0420771

LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes. Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Springer, Berlin. MR1102015

RÜSCHENDORF, L. (2009). On the distributional transform, Sklar's theorem, and the empirical copula process. *J. Statist. Plann. Inference* **139** 3921–3927. MR2553778

SKLAR, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika* (*Prague*) **9** 449–460. MR0345164

SWANSON, J. (2007). Weak convergence of the scaled median of independent Brownian motions. *Probab. Theory Related Fields* **138** 269–304. MR2288071

SWANSON, J. (2011). Fluctuations of the empirical quantiles of independent Brownian motions. *Stochastic Process. Appl.* **121** 479–514. MR2763093

TALAGRAND, M. (1988). Donsker classes of sets. *Probab. Theory Related Fields* **78** 169–191. MR0945108

TALAGRAND, M. (2005). *The Generic Chaining*: *Upper and Lower Bounds of Stochastic Processes*. Springer, Berlin. MR2133757

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

VERVAAT, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* **23** 245–253. MR0321164

YEH, J. (1960). Wiener measure in a space of functions of two variables. *Trans. Amer. Math. Soc.* **95** 433–450. MR0125433

J. KUELBS
T. KURTZ
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706
USA
E-MAIL: kuelbs@math.wisc.edu
kurtz@math.wisc.edu

J. ZINN
DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77840
USA
E-MAIL: jzinn@math.tamu.edu