The Annals of Probability
1991, Vol. 19, No. 2, 758–782

GAUSSIAN CHARACTERIZATION OF UNIFORM DONSKER CLASSES OF FUNCTIONS

By Evarist Giné[1] and Joel Zinn[2]

College of Staten Island of CUNY and University of Connecticut, and Texas A & M University

It is proved that, for classes of functions $\mathscr{F}$ satisfying some measurability, the empirical processes indexed by $\mathscr{F}$ and based on $P \in \mathscr{P}(S)$ satisfy the central limit theorem uniformly in $P \in \mathscr{P}(S)$ if and only if the P-Brownian bridges G_P indexed by $\mathscr{F}$ are sample bounded and ρ_P uniformly continuous uniformly in $P \in \mathscr{P}(S)$. Uniform exponential bounds for empirical processes indexed by universal bounded Donsker and uniform Donsker classes of functions are also obtained.

1. Introduction, notation, definitions. Let $(S, \mathscr{S})$ be a measurable space, let $\mathscr{P}(S)$ be the set of all probability measures on $(S, \mathscr{S})$ and let $\mathscr{F}$ denote a collection of real-valued measurable functions on S such that $\sup_{f \in \mathscr{F}} |f(s) - c_f| < \infty$ for all $s \in S$ and some $c_f < \infty$, $f \in \mathscr{F}$. $\mathscr{F} \in \mathrm{CLT}(P)$ or $\mathscr{F}$ is P-Donsker for $P \in \mathscr{P}(S)$ if the empirical processes based on P and indexed by $\mathscr{F}$ satisfy the central limit theorem as random elements in $l^\infty(\tilde{\mathscr{F}})$, the Banach space of all the bounded real-valued functions on $\tilde{\mathscr{F}} = \{f - c_f: f \in \mathscr{F}\}$, with the sup norm ([7], [9], [10]). $\mathscr{F}$ is universal Donsker if $\mathscr{F} \in \mathrm{CLT}(P)$ for all $P \in \mathscr{P}(S)$ ([8]). Many universal Donsker classes of functions satisfy a stronger property, namely, that the CLT holds not only for all P but also *uniformly* in P, as we prove in Section 3 below (definitions follow shortly). In this paper we characterize this property in terms of Gaussian processes (Section 2). The equivalent Gaussian property is much easier to check than the original definition (see Section 3 for examples). Gaussian characterizations of CLT-related properties in finite dimensions can be traced back to the CLT in type 2 spaces of Hoffmann-Jørgensen and Pisier [13] (type 2 is a Gaussian property) and to the Jain–Marcus CLT (which can be viewed as a consequence of a certain map being type 2; see [14], [22]); Pisier's type 2 characterization of Vapnik–Červonenkis classes [17] and Zinn's Gaussian characterization of universal bounded Donsker classes [23] are examples of results of this type for empirical processes.

The classes of functions we study in this article might in fact provide an adequate framework for the parametric bootstrap as well as for more sophisti-

Received May 1989; revised February 1990.
[1]Partially supported by NSF Grant DMS-86-19411.
[2]Partially supported by NSF Grant DMS-86-01250 and Texas Advanced Research Program No. 3285.
AMS 1980 *subject classifications*. Primary 60F17, 60B12; secondary 62E20.
Key words and phrases. Empirical processes, uniformity in P in the central limit theorem, uniform Donsker classes of functions, uniformly pregaussian classes of functions, exponential inequalities.

cated stochastic procedures as in Beran and Millar [3] and [4]. The two crucial properties of Vapnik–Červonenkis classes of sets that these authors use, namely an empirical triangular array central limit theorem and an exponential bound for the empirical process that holds uniformly in $P \in \mathscr{P}(S)$ also hold for the classes considered in this paper (and the latter only holds for these classes); see Corollaries 2.7 and 2.11.

We proved some of the results in this paper [the equivalence (a) $\Leftrightarrow$ (b) in Theorem 2.6] in 1987, in connection with our work [11] on the bootstrap, but this turned out to be irrelevant for our research at the time. A year later we became aware of the work of Sheehy and Wellner [21], where a similar (but more general) concept is introduced, and noticed that our previous work essentially contained the solution to one of their problems (in particular disproving their conjecture in Remark 8, Section 1 of [21]). The scope of their work, together with the other possible applications mentioned in the previous paragraph, convinced us of the interest of these results, that we then developed in the present form (which has been influenced by [21]).

Given $\mathscr{F}$ as above and $P \in \mathscr{P}(S)$, we let, as in [9] and [10],

$$(1.1)\qquad \begin{aligned} e_P^2(f,g) &= \int_S (f-g)^2\,dP,\\ \rho_P^2(f,g) &= \int_S (f-g)^2\,dP - \left(\int_S (f-g)\,dP\right)^2, \qquad f,g,\in \mathscr{F}, \end{aligned}$$

$$(1.2)\qquad \mathscr{F}' = \{f-g\colon f,g \in \mathscr{F}\}, \qquad (\mathscr{F}')^2 = \left\{(f-g)^2\colon f,g \in \mathscr{F}\right\}.$$

If d is a pseudo-distance on $\mathscr{F}$ (usually $d = e_P$ or $d = \rho_P$) and $\delta > 0$, we let

$$(1.3)\qquad \mathscr{F}'(\delta,d) = \{f-g\colon f,g \in \mathscr{F}, d(f,g) \le \delta\}$$

and if Φ is a real-valued function on $\mathscr{F}$,

$$(1.4)\qquad \|\Phi\|_{\mathscr{F}} = \sup_{f\in\mathscr{F}} |\Phi(f)|, \quad \|\Phi\|_{\mathscr{F}'(\delta,d)} = \sup_{f,g\in\mathscr{F},\, f-g\in\mathscr{F}'(\delta,d)} |\Phi(f) - \Phi(g)|.$$

For probability measures ν on $(S,\mathscr{S})$ and measurable functions f we often write $E_\nu f$ or $\nu(f)$ for $\int_S f\,d\nu$, but if $\nu = P^{\mathbb{N}}$ we will use E_P instead of $E_{P^{\mathbb{N}}}$. To every P in $\mathscr{P}(S)$ such that $\mathscr{F} \subset \mathscr{L}_2(P)$, we associate two centered Gaussian processes G_P and Z_P indexed by $\mathscr{F}$ [or more generally by $\mathscr{L}_2(P)$, but we only consider their restrictions to $\mathscr{F}$, or at most to $\mathscr{F}'$]: those given by the covariances

$$(1.5)\qquad EG_P(f)G_P(g) = E_P(fg) - (E_P f)(E_P g), \qquad f,g \in \mathscr{F},$$

$$(1.6)\qquad EZ_P(f)Z_P(g) = E_P(fg), \qquad f,g \in \mathscr{F}.$$

Note that if g is $N(0,1)$ independent of G_P, then a version of Z_P is

$$(1.7)\qquad Z_P(f) = G_P(f) + gE_P f, \qquad f \in \mathscr{F}.$$

If G_P (or Z_P) has a version with bounded ρ_P uniformly continuous (e_P uniformly continuous) sample paths, then G_P (or Z_P) will always denote such a version; and if $P = \sum \alpha_i \delta_{s_i}$ has finite or countable support, then the versions we always take are

$$G_P = \sum \alpha_i^{1/2} g_i(\delta_{s_i} - P), \qquad Z_P = \sum \alpha_i^{1/2} g_i \delta_{s_i}, \tag{1.8}$$

where $\{g_i\}$ are i.i.d. $N(0,1)$ and δ_{s_i} is point mass at $s_i \in S$.

Given a function $\mathbb{X}$: $S^{\mathbb{N}} \to l^\infty(\mathscr{F})$, its outer law under P is the set function

$$\mathscr{L}^*_{P,\mathscr{F}}(\mathbb{X}) = (P^{\mathbb{N}})^* \circ \mathbb{X}^{-1}. \tag{1.9}$$

If no confusion may arise, we write $\mathscr{L}^*_P$ or even $\mathscr{L}^*$ for $\mathscr{L}^*_{P,\mathscr{F}}$. Following Hoffmann-Jørgensen ([12]), a sequence $\mathbb{X}_n$ of $l^\infty(\mathscr{F})$-valued functions defined on $S^{\mathbb{N}}$ converges in law or weakly to a Radon probability measure γ on $l^\infty(\mathscr{F})$, and we write

$$\mathscr{L}^*_{P,\mathscr{F}}(\mathbb{X}_n) \to_w \gamma$$

if

$$\int_S^* H(\mathbb{X}_n)\, dP^{\mathbb{N}} \to \int_S H\, d\gamma \tag{1.10}$$

for all functions H: $l^\infty(\mathscr{F}) \to \mathbb{R}$ bounded and continuous, where $\int^*$ denotes upper integral.

Let X_i: $S^{\mathbb{N}} \to S$ be the coordinate functions, $i \in \mathbb{N}$. The variables $\{X_i\}_{i=1}^\infty$ are independent with respect to $P^{\mathbb{N}}$ for every $P \in \mathscr{P}(S)$. The normalized empirical measure based on $P \in \mathscr{P}(S)$, ν_n^P, is

$$\nu_n^P = n^{-1/2} \sum_{i=1}^n (\delta_{X_i} - P), \tag{1.11}$$

where $\delta_{X_i(\omega)}$ is point mass at $X_i(\omega) \in l^\infty(\mathscr{F})$, $\omega \in S^{\mathbb{N}}$. We consider ν_n^P as a $l^\infty(\mathscr{F})$-valued random element defined on the probability space $(S^{\mathbb{N}}, \mathscr{S}^{\mathbb{N}}, P^{\mathbb{N}})$ [or, if needed, on the product of this space with $([0,1], \mathscr{B}, \lambda)$, λ Lebesgue measure]. $\mathscr{F}$ is P-Donsker or $\mathscr{F} \in \mathrm{CLT}(P)$ if both the law $\mathscr{L}_{\mathscr{F}}(G_P)$ of G_P is Radon in $l^\infty(\mathscr{F})$ (i.e., $\mathscr{F}$ is P pregaussian) and

$$\mathscr{L}^*_{P,\mathscr{F}}(\nu_n^P) \to_w \mathscr{L}_{\mathscr{F}}(G_P). \tag{1.12}$$

We recall that $\mathscr{L}_{\mathscr{F}}(G_P)$ is Radon in $l^\infty(\mathscr{F})$ if and only if G_P admits a version with bounded uniformly continuous trajectories in $(\mathscr{F}, \rho_P)$ ([1]; see also [10]). $\mathscr{F}$ is universal Donsker if $\mathscr{F} \in \mathrm{CLT}(P)$ for all P ([8]). When no confusion is possible, we write $\mathscr{L}(G_P)$ for $\mathscr{L}_{\mathscr{F}}(G_P)$, and $\mathscr{L}^*(\nu_n^P)$ for $\mathscr{L}^*_{P,\mathscr{F}}(\nu_n^P)$.

Let

$$BL_1^{\mathscr{F}} = BL_1(l^\infty(\mathscr{F})) = \Big\{H: l^\infty(\mathscr{F}) \to \mathbb{R}, \|H\|_\infty \le 1, \sup_{x,y \in l^\infty(\mathscr{F})} |H(x) - H(y)|/\|x-y\|_{\mathscr{F}} \le 1\Big\}. \tag{1.13}$$

For measures μ, ν defined on subsigma algebras of the Borel sets of $l^\infty(\mathscr{F})$, in analogy with the corresponding definition for Polish spaces (e.g., [2], Section

1.2), we let

$$(1.14)\qquad d_{BL_1^*}(\mu,\nu) = \sup_{H\in BL_1^{\mathscr{F}}}\left|\int^* H\,d\mu - \int^* H\,d\nu\right|.$$

The proof of Theorem 1.3, Chapter 1, in [10] (together with the well-known fact that $d_{BL_1^*}$ metrizes weak convergence in $\mathbb{R}^d$) with minor changes gives the following. (For sufficiency the changes are indicated in the proof of Claim 5, Theorem 2.3; for necessity one proceeds as in [10], using the Kirszbraun–McShane theorem—Proposition 1.3, page 2 in [2].)

1.1. THEOREM. *$\mathscr{F}\in$ CLT(P) if and only if both $\mathscr{F}$ is P pregaussian and*

$$(1.15)\qquad \lim_{n\to\infty} d_{BL_1^*}\big(\mathscr{L}^*_{P,\mathscr{F}}(\nu_n^P), \mathscr{L}_{\mathscr{F}}(G_P)\big) = 0.$$

In applications (e.g., Corollary 1.7 in [21]) uniformity in $P\in\mathscr{P}(S)$ of the limit (1.15) is not useful unless it can be combined with some kind of uniformity of G_P such as $d_{BL_1^*}(\mathscr{L}(G_P), \mathscr{L}(G_Q))$ being small if Q is close to P in some weak sense (e.g., in the sense of the Hellinger distance in [21]). If $\mathscr{F}$ satisfies the following definition, then this type of behavior for G_P is assured (see Corollary 2.7).

1.2. DEFINITION. $\mathscr{F}$ is *uniformly pregaussian*, $\mathscr{F}\in$ UPG for short, if for all $P\in\mathscr{P}(S)$, G_P has a version with bounded ρ_P uniformly continuous paths, and for these versions, both

$$(1.16)\qquad \sup_{P\in\mathscr{P}(S)} E\|G_P\|_{\mathscr{F}} < \infty$$

and

$$(1.17)\qquad \lim_{\delta\to 0}\sup_{P\in\mathscr{P}(S)} E\|G_P\|_{\mathscr{F}'(\delta,\rho_P)} = 0.$$

$\mathscr{F}$ is *finitely uniformly pregaussian*, $\mathscr{F}\in$ UPG$_f$ for short, if both

$$(1.16)'\qquad \sup_{P\in\mathscr{P}_f(S)} E\|G_P\|_{\mathscr{F}} < \infty$$

and

$$(1.17)'\qquad \lim_{\delta\to 0}\sup_{P\in\mathscr{P}_f(S)} E\|G_P\|_{\mathscr{F}'(\delta,\rho_P)} = 0,$$

where $\mathscr{P}_f(S) = \{P\in\mathscr{P}(S)\colon P \text{ has finite support}\}$, and G_P are as in (1.8).

In the course of the proof of the main theorem we show that $\mathscr{F}\in$ UPG if and only if $\mathscr{F}\in$ UPG$_f$.

It is convenient to remark that the statements (1.16) and (1.17) are equivalent to $\sup_{P\in\mathscr{P}(S)} E\|G_P\|^r_{\mathscr{F}} < \infty$ and $\lim_{\delta\to 0}\sup_{P\in\mathscr{P}(S)} E\|G_P\|^r_{\mathscr{F}'(\delta,\rho_P)} = 0$ for any $r > 0$, as well as to $\lim_{\lambda\to\infty}\sup_{P\in\mathscr{P}(S)}\Pr\{\|G_P\|_{\mathscr{F}} > \lambda\} = 0$ and $\lim_{\delta\to 0}\sup_{P\in\mathscr{P}(S)}\Pr\{\|G_P\|_{\mathscr{F}'(\delta,\rho_P)} > \varepsilon\} = 0$ for all $\varepsilon > 0$. (Hence, Definition 1.2

coincides with Definition 1.4 in [21].) This is a (well-known) consequence of Borell's inequality ([5]; see [18], Theorem 2.1, for its version for expectations): If G is a sample bounded centered Gaussian process, $\|\cdot\|$ denotes sup norm and $M :=$ median of $\|G\|$, then for all $t > 0$, $\Pr\{|\,\|G\| - M| > \varepsilon\} \le \exp(-t^2/2\sigma^2)$, where $\sigma^2 = \sup EG^2(t)$. Then, since $\sigma \le c_1 M$ and $E|\,\|G\| - M|^p = \int_0^\infty P\{|\,\|G\| - M| > t^{1/p}\}\,dt \le \int_0^\infty \exp\{-t^{2/p}/2\sigma^2\}\,dt \le c_2\sigma^p$ for suitable constants $c_1, c_2 < \infty$ independent of G, it follows that $E\|G\|^p \le K_p M^p$ for some K_p independent of G. This observation takes care of the nonobvious parts of the above equivalences. A similar remark applies to (1.16)′ and (1.17)′.

Theorem 1.1 and the observation following it suggest:

1.3. DEFINITION. $\mathscr{F} \in \mathrm{CLT}(P)$ *uniformly in* P, or $\mathscr{F} \in \mathrm{CLT}_u$, or $\mathscr{F}$ is a *uniform Donsker class* if both

$$\mathscr{F} \in \mathrm{UPG}$$

and

$$\lim_{n\to\infty} \sup_{P\in\mathscr{P}(S)} d_{BL_1^*}\big(\mathscr{L}^*_{P,\mathscr{F}}(\nu_n^P), \mathscr{L}_{\mathscr{F}}(G_P)\big) = 0. \tag{1.18}$$

In [21], (1.18) is replaced by a uniform invariance principle in probability for adequate versions of ν_n^P and G_P. The above definition seems more natural and it also seems to be all that is needed in most applications. It would be surprising if Definition 1.3 were not equivalent to Definition 1.5 in [21] for $\mathscr{P} = \mathscr{P}(S)$. We are not interested in this question here.

Sometimes the processes Z_P and the distances e_P are easier to work with than G_P and ρ_P. In some sense, replacing them in Definition 1.2 if $\mathscr{F}$ is *uniformly bounded* gives a more adequate definition of UPG and UPG_f (see the last part of Section 2).

In Section 2 we prove our main result, which is that, under measurability, $\mathscr{F} \in \mathrm{UPG}_f \Rightarrow \mathscr{F} \in \mathrm{CLT}_u$. (The converse implication is trivial.) We also obtain a slight improvement of a theorem of Sheehy and Wellner [21] on uniformity of the CLT over subsets of probability measures. We also prove exponential bounds that hold uniformly in $P \in \mathscr{P}(S)$, both for universal bounded Donsker classes ([22]) and for uniform Donsker classes.

In Section 3 we show that many interesting classes of functions are in CLT_u, but that there are universal Donsker classes which are not CLT_u.

For the results in Section 2, we need $\mathscr{F}$ to satisfy enough measurability so that, for each P, $\|\sum_{i=1}^n(\delta_{X_i} - P)/n^{1/2}\|_{\mathscr{F}(\delta,\rho_P)}$ is completion measurable and Fubini's theorem can be applied to $\|\sum_{i=1}^n \xi_i\delta_{X_i}/n^{1/2}\|_{\mathscr{F}(\delta,\rho_P)}$, where $\{\xi_i\}$ are i.i.d. real-valued symmetric (usually normal or Bernoulli) independent of $\{X_i\}$, actually defined on $([0,1], \mathscr{B}, \lambda)$. In other words, we need $\mathscr{F}$ to be nearly linearly deviation measurable (NLDM(P)) for each P, in the notation of [9] and [10]. When this holds, we say that $\mathscr{F}$ is *measurable*. For example, $\mathscr{F}$ is measurable if it is countable, or if the empirical processes ν_n^P are stochastically separable or if $\mathscr{F}$ is image admissible Suslin ([7]).

2. Results and proofs. In some instances in the proof of the main result we will use finite dimensional approximation. We will then require two lemmas for $\mathbb{R}^d$-valued random variables. We only sketch their proofs since they are standard (and the lemmas themselves are well known).

2.1. LEMMA. *Let $\mathscr{P}_M^d = \{P \text{ on } \mathbb{R}^d\colon \operatorname{supp} P \subset \{\|x\| \le M\}\}$ for $M < \infty$. For $P \in \mathscr{P}_M^d$, let $\{\xi_i^P\}_{i=1}^\infty$ be i.i.d. random variables with law P, and let $\Phi_P = \operatorname{Cov}(P)$. Then*

$$\lim_{n\to\infty} \sup_{P\in\mathscr{P}_M^d} d_{BL_1^*}\left[\mathscr{L}\left(\sum_{i=1}^n (\xi_i^P - E\xi_i^P)/n^{1/2}\right), N(0,\Phi_P)\right] = 0, \tag{2.1}$$

where $N(0,\Phi_P)$ is the centered Gaussian law of $\mathbb{R}^d$ with covariance Φ_P.

PROOF (Sketch). This follows from standard results on speed of convergence in the multidimensional CLT ([20]). An elementary proof can be obtained combining the following two observations: (i) The usual Lindeberg proof of the CLT (e.g., [2], pages 37 and 67) readily gives

$$d_3\left[\mathscr{L}\left(\sum_{i=1}^n (\xi_i^P - E\xi_i^P)/n^{1/2}\right), N(0,\Phi_P)\right] \le KM(\operatorname{trace}\Phi_P)/n^{1/2},$$

where K is a universal constant and

$$d_3(\mu,\nu) := \sup\left\{\left|\int f\,d(\mu-\nu)\right|\colon \sum_{|\alpha|\le 3}\|D^\alpha f\|_\infty \le 1\right\},$$

$\alpha = (\alpha_1,\dots,\alpha_d) \in (\mathbb{N}\cup\{0\})^d$, $|\alpha| = \sum_{i=1}^d \alpha_i$ and $D^\alpha f = \partial^{|\alpha|} f/\partial^{\alpha_1}x_1\dots\partial^{\alpha_d}x_d$. And (ii), $f \in BL_1(\mathbb{R}^d)$ can be uniformly approximated by C^∞ functions whose partial derivatives have not too large $\|\cdot\|_\infty$-norms [specifically, if $\|f\|_{BL} \le 1$, where $\|f\|_{BL} = \|f\|_\infty + \sup_{x\ne y}|f(x)-f(y)|/|x-y|$, and if $f_\varepsilon(x) = \int f(x-\varepsilon y)e^{-\|y\|^2/2}\,dy/(2\pi)^{d/2} := (f*\varphi_\varepsilon)(x)$, $x\in\mathbb{R}^d$, then

$$\|f - f_\varepsilon\|_\infty \le (2\pi)^{-d/2}\int(2\wedge\varepsilon\|y\|)e^{-\|y\|^2/2}\,dy \le c(d)\varepsilon := h_d(\varepsilon) \to 0$$

as $\varepsilon\to 0$ and $\|D^\alpha f_\varepsilon\|_\infty \le \varepsilon^{-|\alpha|}\int|D^\alpha\varphi|\,dy$, φ being the density of $N(0,I)$]. □

The same principles apply to give the following inequality (for which we do not claim optimality).

2.2. LEMMA. *Let Φ and $\overline{\Phi}$ be two covariances on $\mathbb{R}^d\times\mathbb{R}^d$ and let $N(0,\Phi)$, $N(0,\overline{\Phi})$ be the corresponding centered Gaussian laws in $\mathbb{R}^d$. Let $\|\Phi-\overline{\Phi}\|_\infty = \max_{i,j\le d}\|\Phi(i,j)-\overline{\Phi}(i,j)\|$, where $\Phi(i,j) := \Phi(e_i,e_j)$ and $\{e_i\}$ is the canonical basis of $\mathbb{R}^d$. Then*

$$d_{BL_1^*}(N(0,\Phi),N(0,\overline{\Phi})) \le c(d)\|\Phi-\overline{\Phi}\|_\infty^{1/4}, \tag{2.2}$$

where $c(d)$ is an absolute constant that depends on d.

PROOF (Sketch). By the previous proof it is enough to show

$$d_3(N(0,\Phi), N(0,\bar{\Phi})) \le c(d)\|\Phi - \bar{\Phi}\|_\infty.$$

To prove this inequality one proceeds as in Lindeberg's proof of the CLT with the random variables $\sum_{i=1}^n X_i/n^{1/2}$ and $\sum_{i=1}^n Y_i/n^{1/2}$, X_i, Y_i all independent, $\mathscr{L}(X_i) = N(0,\Phi)$, $\mathscr{L}(Y_i) = N(0,\bar{\Phi})$, and let $n \to \infty$. Details are omitted since they are routine. □

The following is our main result.

2.3. THEOREM. *Let $\mathscr{F}$ be a measurable class of functions. Then*

$$\mathscr{F} \in UPG_f \Rightarrow \mathscr{F} \in CLT_u.$$

PROOF. Assume $\mathscr{F} \in \mathrm{UPG}_f$. We divide the proof into several steps.

CLAIM 1. It suffices to prove the theorem for classes $\mathscr{F}$ which are uniformly bounded by 1, i.e., such that $F(s) \le 1$ for all $s \in S$.

PROOF OF CLAIM 1. First note that $\mathscr{F} \in \mathrm{UPG}_f$ if and only if $\tilde{\mathscr{F}} \in \mathrm{UPG}_f$, where $\tilde{\mathscr{F}} = \{\tilde{f} = c(f - c_f): f \in \mathscr{F}\}$ for some $c \neq 0$ and arbitrary finite constants c_f, and the same is true of CLT_u. Now,

$$\mathscr{F} \in \mathrm{CLT}_u \Rightarrow \mathscr{F} \text{ is universal Donsker} \Rightarrow \sup_{f\in\mathscr{F}} \operatorname{diam}(f) < \infty,$$

where $\operatorname{diam}(f) = \sup_{s\in S} f(s) - \inf_{s\in S} f(s)$ ([8]); moreover,

$$\mathscr{F} \in \mathrm{UPG}_f \Rightarrow \sup_{f\in\mathscr{F}} \operatorname{diam}(f) < \infty,$$

as observed in [22] [if $P(s_1, s_2) = \frac{1}{2}(\delta_{s_1} + \delta_{s_2})$, then $\sup_{s_1,s_2\in S} E\|G_{P(s_1,s_2)}\|_{\mathscr{F}} < \infty$ implies $\sup_{f\in\mathscr{F}} (\operatorname{diam}(f))^2/4 = \sup_{f\in\mathscr{F}} \sup_{s_1,s_2\in S} EG^2_{P(s_1,s_2)}(f) < \infty$]. Therefore, if in the definition of $\tilde{\mathscr{F}}$, we take $c_f = \inf(f)$ and $c = [\sup_{f\in\mathscr{F}}(\operatorname{diam}(f))]^{-1}$, $\tilde{\mathscr{F}}$ is a class of functions uniformly bounded by 1. By the first remark it suffices to prove the theorem for $\tilde{\mathscr{F}}$. We assume $F \le 1$ in the rest of this proof.

CLAIM 2. Let $\mathscr{G} = \mathscr{F} \cup \mathscr{F}^2 \cup \mathscr{F}' \cup (\mathscr{F}')^2$. Then

$$\sup_{P\in\mathscr{P}(S)} E_P\|P_n - P\|_{\mathscr{G}} = O(n^{-1/2}),$$

where $P_n = n^{-1}\sum_{i=1}^n \delta_{X_i}$.

PROOF OF CLAIM 2. We prove it only for $\mathscr{G} = (\mathscr{F}')^2$, since subsets of this proof give the rest. By Claim 1, for $f, g, \bar{f}, \bar{g} \in \mathscr{F}$ we have

$$E_{P_n}|(f-g)^2 - (\bar{f}-\bar{g})^2|^2 \le 16 E_{P_n}|(f-g) - (\bar{f}-\bar{g})|^2. \tag{2.3}$$

Let $\{g_i\}$ be an i.i.d. $N(0,1)$ sequence independent of $\{X_i\}$ [actually defined on $([0,1], \mathscr{B}, \lambda)$: for each P we take $(\Omega, \Sigma, \Pr_P) = (S^{\mathbb{N}}, \mathscr{S}^{\mathbb{N}}, P^{\mathbb{N}}) \times ([0,1], \mathscr{B}, \lambda)$ to be our general probability space]. For each $\omega \in S^{\mathbb{N}}$ fixed and $n \in \mathbb{N}$, the process $(h(X_1(\omega)), \dots, h(X_n(\omega))) \to \sum_{i=1}^n g_i h(X_i(\omega))/n^{1/2}$, $h \in (\mathscr{F}')^2$, attains the value 0 for one h and has a separable index set, so (2.3) and the Slepian–Fernique lemma (as stated in [10], Theorem 4.4, Chapter 1) give

$$E_g \left\| \sum_{i=1}^n g_i \delta_{X_i}/n^{1/2} \right\|_{(\mathscr{F}')^2} \le 8 E_g \left\| \sum_{i=1}^n g_i \delta_{X_i}/n^{1/2} \right\|_{\mathscr{F}'}, \tag{2.4}$$

where E_g denotes integration only with respect to the variables g_i (or λ). Therefore, if $\{\varepsilon_i\}$ is a Rademacher sequence also defined on $([0,1], \mathscr{B}, \lambda)$ and independent of $\{g_i\}$, we obtain

$$\begin{aligned} E_P \|P_n - P\|_{(\mathscr{F}')^2} &\le 2 E_{\Pr_P} \left\| \sum_{i=1}^n \frac{\varepsilon_i \delta_{X_i}}{n} \right\|_{(\mathscr{F}')^2} \quad \text{(by symmetrization)} \\ &\le \frac{2}{E|g_1|} E_{\Pr_P} \left\| \sum_{i=1}^n \frac{g_i \delta_{X_i}}{n} \right\|_{(\mathscr{F}')^2} \\ &\qquad \text{(by Jensen's inequality after replacing } g_i \text{ by } \varepsilon_i |g_i|) \\ &\le \frac{16}{n^{1/2} E|g_1|} E_P E_g \left\| \sum_{i=1}^n \frac{g_i \delta_{X_i}}{n^{1/2}} \right\|_{\mathscr{F}'} \quad [\text{by (2.4)}] \\ &\le \frac{32}{n^{1/2} E|g_1|} E_P E_g \left\| \sum_{i=1}^n \frac{g_i \delta_{X_i}}{n^{1/2}} \right\|_{\mathscr{F}} \quad \text{(by the triangle inequality)} \\ &\le \frac{32}{n^{1/2} E|g_1|} \sup_{Q \in \mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}} \quad \left(\text{since } Z_{P_n} = \sum_{i=1}^n \frac{g_i \delta_{X_i}}{n^{1/2}} \right) \\ &\le \frac{32\sqrt{\pi/2}}{n^{1/2}} \sup_{Q \in \mathscr{P}_f(S)} \left(E\|G_Q\|_{\mathscr{F}} + \left(\frac{2}{\pi}\right)^{1/2} \right) \\ &\qquad [\text{by (1.7) and Claim 1}] \\ &= O(n^{-1/2}) \text{ uniformly in } P \in \mathscr{P}(S) \quad (\text{since } \mathscr{F} \in \mathrm{UPG}_f). \end{aligned} \tag{2.5}$$

CLAIM 3. $(\mathscr{F}, e_P)$ is totally bounded for all $P \in \mathscr{P}(S)$ and

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \sup_{P \in \mathscr{P}(S)} P^{\mathbb{N}}\left\{ \|\nu_n^P\|_{\mathscr{F}'(\delta, e_P)} > \varepsilon \right\} = 0 \quad \text{for all } \varepsilon > 0, \tag{2.6}$$

in particular, $\mathscr{F}$ is universal Donsker.

PROOF OF CLAIM 3. Since $\mathscr{F} \in \mathrm{UPG}_f$, we have

$$\sup\left\{ E\|G_{P_n}(\omega)\|_{\mathscr{F}} : P \in \mathscr{P}(S), \omega \in S^{\mathbb{N}}, n \in \mathbb{N} \right\} < \infty.$$

Hence the covering numbers $N(\varepsilon, \mathscr{F}, e_{P_n}(\omega))$ [:= smallest number of $e_{P_n}(\omega)$ balls of radius less than or equal to ε and centers in $\mathscr{F}$ needed to cover $\mathscr{F}$] are uniformly bounded by Sudakov's minorization theorem (e.g., [10], Theorem

4.3, Chapter 1), concretely there is $c < \infty$ such that for all $\omega \in S^{\mathbb{N}}$, $n \in \mathbb{N}$, $P \in \mathscr{P}(S)$ and $\varepsilon > 0$,

$$\log N\big(\varepsilon, \mathscr{F}, e_{P_n}(\omega)\big) < c/\varepsilon^2. \tag{2.7}$$

A well-known consequence of Claim 2 ([19]) is that $\|P_n - P\|_{(\mathscr{F}')^2} \to 0$ a.s. Since

$$\left|e^2_{P_n(\omega)}(f, g) - e^2_P(f, g)\right| = \left|P_n(f-g)^2 - P(f-g)^2\right|,$$

we have

$$\sup_{f, g \in \mathscr{F}} \left|e^2_{P_n(\omega)}(f, g) - e^2_P(f, g)\right| \to 0 \qquad P^{\mathbb{N}} \text{ a.s. for all } P \in \mathscr{P}(S). \tag{2.8}$$

This and (2.7) give

$$\sup_{P \in \mathscr{P}(S)} \log N(\varepsilon, \mathscr{F}, e_P) \le c/\varepsilon^2, \qquad \varepsilon > 0. \tag{2.9}$$

Hence, $(\mathscr{F}, e_P)$ is totally bounded (uniformly in P). In order to prove (2.6), we first symmetrize: Using Lemma 2.5 and the proof of Lemma 2.7(b) in [9] we have, for $\{\xi_i\}$, a Rademacher sequence independent of $\{X_i\}$, [i.e., defined on $((0,1), \mathscr{B}, \lambda)$]

$$P^{\mathbb{N}}\left\{\|\nu_n^P\|_{\mathscr{F}'(\delta, e_P)} > 4\varepsilon\right\} \le 2(1 - \delta^2/4\varepsilon^2)^{-1} \Pr_P\left\{\left\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n^{1/2}\right\|_{\mathscr{F}'(\delta, e_P)} > \varepsilon\right\}.$$

Then

$$\begin{aligned}
&\Pr_P\left\{\left\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n^{1/2}\right\|_{\mathscr{F}'(\delta, e_P)} > \varepsilon\right\} \\
&\quad \le \Pr_P\left\{\left\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n^{1/2}\right\|_{\mathscr{F}'(2^{1/2}\delta, e_{P_n})} > \varepsilon\right\} \\
&\qquad + P^{\mathbb{N}}\left\{\sup_{f, g \in \mathscr{F}} \left|e^2_{P_n(\omega)}(f, g) - e^2_P(f, g)\right| > \delta^2\right\} \\
&\quad := (\mathrm{I})_P + (\mathrm{II})_P.
\end{aligned}$$

Now Claim 2, concretely (2.5), implies

$$\lim_{n \to \infty} \sup_{P \in \mathscr{P}(S)} (\mathrm{II})_P = 0 \quad \text{for all } \delta > 0.$$

Next, noting that by (1.7),

$$E\|Z_Q\|_{\mathscr{F}'(\delta, e_Q)} \le E\|G_Q\|_{\mathscr{F}'(\delta, e_Q)} + (2/\pi)^{1/2}\delta \le E\|G_Q\|_{\mathscr{F}'(\delta, \rho_Q)} + (2/\pi)^{1/2}\delta,$$

we can proceed as in (2.5) and obtain

$$\begin{aligned} E_{\Pr_P}\left\| \sum_{i=1}^{n} \varepsilon_i \delta_{X_i}/n^{1/2} \right\|_{\mathscr{F}'(\delta, e_{P_n})} &\leq (1/E\|g_1|) E_P E_g \left\| \sum_{i=1}^{n} g_i \delta_{X_i}/n^{1/2} \right\|_{\mathscr{F}'(\delta, e_{P_n})} \\ &\leq (1/E|g_1|) \sup_{Q \in \mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}'(\delta, e_Q)} \\ &\leq \sqrt{\pi/2} \sup_{Q \in \mathscr{P}_f(S)} E\|G_Q\|_{\mathscr{F}'(\delta, \rho_Q)} + \sqrt{\pi/2}\, \delta^{1/2} \end{aligned}$$

for all $P \in \mathscr{P}(S)$. Hence, since $\mathscr{P} \in \mathrm{UPG}_f$,

$$\lim_{\delta \to 0} \sup_{P \in \mathscr{P}(S)} (\mathrm{I})_P = 0 \quad \text{for all } \varepsilon > 0,$$

thus proving (2.6). Finally, $\mathscr{F}$ is a universal Donsker class by, e.g., Theorem 1.3, Chapter 1 in [10].

CLAIM 4. $\mathscr{F} \in \mathrm{UPG}$.

PROOF OF CLAIM 4. If $\mathscr{G}$ is as in Claim 2, it follows from uniform boundedness of $\mathscr{F}$ (hence of $\mathscr{G}$) that

$$\|P_n - P\|_{\mathscr{G}} \to 0 \qquad P^{\mathbb{N}} \text{ a.s. for all } P \in \mathscr{P}(S), \tag{2.10}$$

(see, e.g., [19]). Given P, fix $\omega \in S^{\mathbb{N}}$ so that convergence in (2.10) takes place. Then, because of (2.10), for any finite number $f_1, \dots, f_r$ of functions in $\mathscr{F}$,

$$\mathscr{L}\big(G_{P_{n(\omega)}}(f_1), \dots, G_{P_{n(\omega)}}(f_r)\big) \to_w \mathscr{L}\big(G_P(f_1), \dots, G_P(f_r)\big).$$

So, if we show that $\{\mathscr{L}_{\mathscr{F}}(G_{P_{n(\omega)}})\}$ is Cauchy in $d_{BL_1^*}$, we will have proved both that the law of G_P is Radon and that

$$\mathscr{L}\big(G_{P_{n(\omega)}}\big) \to_w \mathscr{L}(G_P) \quad \text{in } l^\infty(\mathscr{F}). \tag{2.11}$$

Let $H \in BL_1^{\mathscr{F}}$. Since $\rho_P \leq e_P$, by (2.8) and (2.9) in the proof of Claim 3, given $\tau > 0$ there is $n > 0$, there are $f_1, \dots, f_N \in \mathscr{F}$ with $N < \infty$, and a partition of $\mathscr{F}$, $A_1, \dots, A_N$ with $f_i \in A_i$ such that A_i is contained in the $\rho_{P_{m(\omega)}}$ ball about f_i of radius τ for all $m \geq n$. Let $\pi_\tau f = f_i$ if $f \in A_i$, $i = 1, \dots, N$. Let $G_{P_{n(\omega)},\tau}(f) = G_{P_{n(\omega)}}(\pi_\tau f)$ and write

$$\begin{aligned} &|EH(G_{P_{n(\omega)}}) - EH(G_{P_{m(\omega)}})| \\ &\quad \leq |EH(G_{P_{n(\omega)}}) - EH(G_{P_{n(\omega)},\tau})| + |EH(G_{P_{h(\omega)},\tau}) - EH(G_{P_{m(\omega)}\tau})| \\ &\qquad + E|H(G_{P_{m(\omega)},\tau}) - EH(G_{P_{m(\omega)}})| \\ &\quad := (\mathrm{I})_{\tau,n} + (\mathrm{II})_{\tau,n,m} + (\mathrm{I})_{\tau,m}. \end{aligned} \tag{2.12}$$

We have

$$(\mathrm{I})_{\tau,n} \leq E\|G_{P_{n(\omega)}}\|_{\mathscr{F}'(\tau, \rho_{P_{n(\omega)}})}, \qquad (\mathrm{I})_{\tau,m} \leq E\|G_{P_{m(\omega)}}\|_{\mathscr{F}'(\tau, \rho_{P_{m(\omega)}})},$$

hence

$$(2.13) \qquad (\mathrm{I})_{\tau,m} \le \sup_{Q\in\mathscr{P}_f(S)} E\|G_Q\|_{\mathscr{F}'(\tau,\rho_Q)}, \qquad m \ge n.$$

As for $(\mathrm{II})_{\tau,m,n}$, we will apply Lemma 2.2. Note that, by polarity,

$$\|\Phi_{P_n} - \overline{\Phi}_{P_m}\|_\infty \le \tfrac{1}{2}\max_{i,j\le N}|P_n(f_i-f_j)^2 - P_m(f_i-f_j)^2| + \tfrac{1}{2}\max_{i,j\le N}|(P_n(f_i-f_j))^2 - (P_m(f_i-f_j))^2| + \max_{i\le N}|P_n f_i^2 - P_m f_i^2| + \max_{i\le N}|(P_n f_i)^2 - (P_m f_i)^2|,$$

so that, by (2.10) (recall $\mathscr{F}$ is uniformly bounded), we have

$$\lim_{n\to\infty}\sup_{m\ge n}\|\Phi_{P_n(\omega)} - \overline{\Phi}_{P_m(\omega)}\|_\infty = 0.$$

Lemma 2.2 then gives

$$(2.14) \qquad \lim_{n\to\infty}\sup_{m\ge n}\sup_{H\in BL_1^{\mathscr{F}}}(\mathrm{II})_{\tau,n,m} = 0.$$

Hence, by (2.13), $\mathscr{F}\in \mathrm{UPG}_f$, and (2.14), we obtain from (2.12) that

$$\lim_{n\to\infty}\sup_{m\ge n} d_{BL_1^*}\big(\mathscr{L}(G_{P_n(\omega)}), \mathscr{L}(G_{P_m(\omega)})\big) = 0.$$

Therefore, the sequence $\{\mathscr{L}_{\mathscr{F}}(G_{P_n(\omega)})\}$ is Cauchy in $d_{BL_1^*}$ and (2.11) is proved. A consequence of (2.11) is that $E\|G_{P_n(\omega)}\|_{\mathscr{F}} \to E\|G_P\|_{\mathscr{F}}$ (uniform integrability follows since $\mathscr{F}\in\mathrm{UPG}_f$ implies $\sup_{Q\in\mathscr{P}_f(S)}E\|G_Q\|_{\mathscr{F}}^2 < \infty$, as remarked after Definition 1.2). Similarly, $E\|G_{P_n(\omega)}\|_{\mathscr{F}'(\delta,\rho_P)} \to E\|G_P\|_{\mathscr{F}'(\delta,\rho_P)}$; but for n large enough, again by (2.10), $\|G_{P_n(\omega)}\|_{\mathscr{F}'(\delta,\rho_P)} \le \|G_{P_n(\omega)}\|_{\mathscr{F}'(2^{1/2}\delta,\rho_{P_n(\omega)})}$ and therefore, since $\mathscr{F}\in\mathrm{UPG}_f$, it follows that $\lim_{\delta\to 0}\sup_{P\in\mathscr{P}(S)}E\|G_P\|_{\mathscr{F}'(\delta,\rho_P)} = 0$. So, $\mathscr{F}\in$ UPG.

CLAIM 5.

$$\lim_{n\to\infty}\sup_{P\in\mathscr{P}(S)} d_{BL_1^*}\big(\mathscr{L}^*(\nu_n^P), \mathscr{L}(G_P)\big) = 0.$$

PROOF OF CLAIM 5. By Claim 4 we have

$$(2.15) \qquad \sup_{P\in\mathscr{P}(S)} E\|G_P\|_{\mathscr{F}<\infty}, \qquad \lim_{\delta\to\infty}\sup_{P\in\mathscr{P}(S)} E\|G_P\|_{\mathscr{F}'(\delta,e_P)} = 0.$$

We show that (2.15) and (2.6) prove our claim by following the steps in the proof of (ii) $\Rightarrow$ (i) (Theorem 1.3, Chapter 1, [10]), with some simplifications. Given $\tau > 0$, let $f_1,\dots,f_{N_P(\tau)}$ be the centers of a minimal covering of $\mathscr{F}$ by e_P balls of radius τ and centers in $\mathscr{F}$, $N_P(\tau) := N(\tau,\mathscr{F},e_P) < \infty$ [by (2.9)]. Let π_τ^P: $\mathscr{F}\to\mathscr{F}$ be a mapping satisfying $\pi_\tau^P f = f_j$ for some j, and $e_P(\pi_\tau^P f, f) \le \tau$. Let $Y_j^P(f) = f(X_j) - Pf$, $Y_{j,\tau}^P(f) = Y_j^P(\pi_\tau^P f)$, $f\in\mathscr{F}$, $j = 1,\dots$. Let G_P be its version with bounded ρ_P (hence e_P) uniformly continuous paths, and let

$G_{P,\tau}(f) = G_P(\pi_\tau^P f)$ as before, $f \in \mathscr{F}$. Let $H \in BL_1^{\mathscr{F}}$. Then, as in (1.13), in the place cited, we have

$$
\begin{aligned}
|E^*H(\nu_n^P) - EH(G_P)| &\le \left| E^*H\left(\sum_{j=1}^n Y_j^P/n^{1/2}\right) - EH\left(\sum_{j=1}^n Y_{j,\tau}^P/n^{1/2}\right)\right| \\
&\quad + \left| EH\left(\sum_{j=1}^n Y_{j,\tau}^P/n^{1/2}\right) - EH(G_{P,\tau})\right| \qquad (2.16)\\
&\quad + |EH(G_{P,\tau}) - EH(G_P)| \\
&:= (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}).
\end{aligned}
$$

By Lemma 2.1, since $\sup_P N_P(\tau) < \infty$ [(2.9)],

$$\lim_{n\to\infty} \sup_{P \in \mathscr{P}(S)} \sup_{H \in BL_1^{\mathscr{F}}} (\mathrm{II}) = 0 \quad \text{for all } \tau > 0. \qquad (2.17)$$

For every $\varepsilon > 0$ and $H \in BL_1^{\mathscr{F}}$, since $|H(x) - H(y)| \le 2 \wedge \|x - y\|_\infty$, we have

$$(\mathrm{I}) \le \varepsilon + 2P^{\mathbb{N}}\left\{\left\|\nu_n^P\right\|_{\mathscr{F}'(\tau, e_P)} > \varepsilon\right\}.$$

Hence (2.6) gives

$$\lim_{\tau\to 0} \limsup_{n\to\infty} \sup_{P \in \mathscr{P}(S)} \sup_{H \in BL_1^{\mathscr{F}}} (\mathrm{I}) = 0. \qquad (2.18)$$

Similarly,

$$(\mathrm{III}) \le \varepsilon + 2\Pr\left\{\|G_P\|_{\mathscr{F}'(\tau, e_P)} > \varepsilon\right\}$$

so that by (2.15),

$$\lim_{\tau\to 0} \limsup_{n\to\infty} \sup_{P \in \mathscr{P}(S)} \sup_{H \in BL_1^{\mathscr{F}}} (\mathrm{III}) = 0. \qquad (2.19)$$

(2.16)–(2.19) prove the claim. □

Sometimes it is easier to deal with Z_P and e_P than with G_P and ρ_P. This suggests the following modification of Definition 1.2.

2.4. Definition. $\mathscr{F} \in \mathrm{UPG}_f'$ if

$$\sup_{P \in \mathscr{P}_f(S)} E\|Z_P\|_{\mathscr{F}} < \infty \qquad (2.20)$$

and

$$\lim_{\delta\to 0} \sup_{P \in \mathscr{P}_f(S)} E\|Z_P\|_{\mathscr{F}'(\delta, e_P)} = 0. \qquad (2.21)$$

If these properties hold with $\mathscr{P}_f(S)$ replaced by $\mathscr{P}(S)$, then we write $\mathscr{F} \in \mathrm{UPG}'$.

To compare UPG'_f with UPG_f, let us note first that if $\mathscr{F} \in \mathrm{UPG}'_f$, then $\mathscr{F}$ is uniformly bounded: Since $P(s) = \delta_s \in \mathscr{P}_f(S)$, we have

$$\sin_{P \in \mathscr{P}_f(S)} E\|Z_P\|_{\mathscr{F}} \geq E|Z_{P(s)}(f)| = \sqrt{2/\pi}\,|f(s)|, \qquad s \in S.$$

But if $\mathscr{F}$ is uniformly bounded and $\mathscr{F} \in \mathrm{UPG}_f$, then $\mathscr{F} \in \mathrm{UPG}'_f$. By (1.7), $E\|Z_P\|_{\mathscr{F}} \leq E\|G_P\|_{\mathscr{F}} + \|Pf\|_{\mathscr{F}}$ and $E\|G_P\|_{\mathscr{F}'(\delta, \rho_P)} \geq E\|G_P\|_{\mathscr{F}'(\delta, e_P)} \geq E\|Z_P\|_{\mathscr{F}'(\delta, e_P)} - (2/\pi)^{1/2}\delta$. It also follows from (1.7) that, for uniformly bounded classes, Z_P can be replaced by G_P in Definition 2.4.

We could ask what kind of uniform CLT does $\mathscr{F}$ satisfy if $\tilde{\mathscr{F}} \in \mathrm{UPG}'_f$, where $\tilde{\mathscr{F}}$ is as in the proof of Claim 1, Theorem 2.3. For the purpose of Theorem 2.6, let us make a change in the definition of $\tilde{\mathscr{F}}$.

2.5. Definition. Let $c_f := \inf_{s \in S} f(s)$. Assume $|c_f| < \infty$ for all $f \in \mathscr{F}$. Then we define $\tilde{\mathscr{F}} = \mathscr{F}$ if $\mathscr{F}$ is uniformly bounded and $\tilde{\mathscr{F}} = \{f - c_f\}$ otherwise. Analogously $\tilde{f} = f$ in the first case and $\tilde{f} = f - c_f$ in the second. For every $P \in \mathscr{P}(S)$, we let $\tilde{e}_P(f, g) = e_P(\tilde{f}, \tilde{g})$.

2.6. Theorem. *Let $\mathscr{F}$ be a measurable class of functions with $|c_f| < \infty$ for all $f \in \mathscr{F}$. The following are equivalent:*

(a) $\tilde{\mathscr{F}} \in \mathit{UPG}'_f$.

(b) $(\mathscr{F}, \tilde{e}_P)$ *is totally bounded for all P and*

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \sup_{P \in \mathscr{P}(S)} P^{\mathbb{N}}\left\{\left\|\nu_n^P\right\|_{\mathscr{F}'(\delta, \tilde{e}_P)} > \varepsilon\right\} = 0$$

for all $\varepsilon > 0$.

(c) $\tilde{\mathscr{F}} \in \mathit{UPG}'$ *and* $\lim_{n \to \infty} \sup_{P \in \mathscr{P}(S)} d_{BL_1^*}[\mathscr{L}^*_{P, \mathscr{F}}(\nu_n^P), \mathscr{L}_{\mathscr{F}}(G_P)] = 0$.

Proof. The proof is analogous to (and essentially contained in) the proof of Theorem 2.3 except for (b) $\Rightarrow \tilde{\mathscr{F}} \in \mathrm{UPG}'$. So, this is the only part we prove. By Theorem 1.3, Chapter 1 in [10], (b) implies that $\mathscr{F}$ is universal Donsker. Therefore, by [22], Theorem 2.3,

$$\sup_{P \in \mathscr{P}(S)} E\|G_P\|_{\mathscr{F}} < \infty$$

and $\tilde{\mathscr{F}}$ is uniformly bounded. Since $E\|Z_P\|_{\tilde{\mathscr{F}}} \leq E\|G_P\|_{\tilde{\mathscr{F}}} + \|P\tilde{f}\|_{\tilde{\mathscr{F}}} = E\|G_P\|_{\mathscr{F}} + \|P\tilde{f}\|_{\tilde{\mathscr{F}}}$, it follows that

$$\sup_{P \in \mathscr{P}(S)} E\|Z_P\|_{\tilde{\mathscr{F}}} < \infty. \tag{2.22}$$

By Theorem 2.8, Chapter 1 in [10],

$$\mathscr{L}^*_{\Pr_P, \mathscr{F}}\left\{\sum_{i=1}^{n} \varepsilon_i \delta_{X_i}/n^{1/2}\right\} \to_w \mathscr{L}_{\mathscr{F}}(Z_P) \text{ in } l^\infty(\tilde{\mathscr{F}}).$$

Hence, by the Portmanteau theorem, which is still valid for this type of convergence, as is easy to check, we have

$$\liminf_{n\to\infty} (\Pr_P)_* \left\{ \sum_{i=1}^{n} \varepsilon_i \delta X_i / n^{1/2} \in G \right\} \geq \Pr\{ Z_P \in G \}$$

for open subsets G of $l^\infty(\tilde{\mathscr{F}})$. This, the symmetrization Lemma 2.7(a) of [9] and (b) give

$$\begin{aligned} 0 &= \lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{P\in\mathscr{P}(S)} \Pr_P \left\{ \left\| \sum_{i=1}^{n} \varepsilon_i \delta_{X_i} / n^{1/2} \right\|_{\tilde{\mathscr{F}}'(\delta, e_P)} > \varepsilon \right\} \\ &\geq \lim_{\delta\to 0} \sup_{P\in\mathscr{P}(S)} \limsup_{n\to\infty} \Pr_P \left\{ \left\| \sum_{i=1}^{n} \varepsilon_i \delta_{X_i} / n^{1/2} \right\|_{\tilde{\mathscr{F}}'(\delta, e_P)} > \varepsilon \right\} \\ &\geq \lim_{\delta\to 0} \sup_{P\in\mathscr{P}(S)} \Pr\{ \|Z_P\|_{\tilde{\mathscr{F}}'(\delta, e_P)} > \varepsilon \}. \end{aligned} \tag{2.23}$$

Borell's inequality or its Maurey–Pisier formulation ([18], Theorem 2.1), combined with (2.23) shows, as mentioned after Definition 1.2,

$$\lim_{\delta\to 0} \sup_{P\in\mathscr{P}(S)} E\|Z_P\|_{\tilde{\mathscr{F}}'(\delta, e_P)} = 0. \tag{2.24}$$

(2.22) and (2.24) mean $\tilde{\mathscr{F}} \in$ UPG′. □

Compared with Theorem 2.3 and its proof, Theorem 2.6 is more complete in the sense that (b) ⇔ (a) means that a uniformity condition in the CLT not containing a priori a Gaussian uniformity condition is indeed equivalent to a Gaussian uniformity condition. The problem in Theorem 2.3 is that we do not know how to obtain $\mathscr{F} \in$ UPG from only (2.6) plus $(\mathscr{F}, \rho_P)$ or $(\mathscr{F}, e_P)$ totally bounded for all $P \in \mathscr{P}(S)$: that only seems to give $\tilde{\mathscr{F}} \in$ UPG′.

Condition (c) in Theorem 2.6 is slightly weaker than $\mathscr{F} \in \mathrm{CLT}_u$: the difference is that in (c), Gaussian uniformity is with respect to $\mathscr{F}'(\delta, \tilde{e}_P) \subseteq \mathscr{F}'(\delta, \rho_P)$. However, this weaker uniformity suffices in many instances.

Corollary 2.7 provides a framework for the application of the above theorems in statistics. It is similar to Corollaries 1.4 and 1.7 in [21].

2.7. Corollary. *Let $\mathscr{F}$ be a measurable UPG$_f$ class and let $\mathscr{G} = \mathscr{F} \cup \mathscr{F}^2 \cup \mathscr{F}' \cup (\mathscr{F}')^2$. Let $\{R_n\}_{n=0}^\infty$ be probability measures on $(S, \mathscr{S})$ such that $\|R_n - R_0\|_{\mathscr{G}} \to 0$. Then $\mathscr{L}^*(\nu_n^{R_n}) \to_w \mathscr{L}(G_{R_0})$ in $l^\infty(\mathscr{F})$.*

Proof. The hypothesis and Lemma 2.2 imply, as in the proofs of (2.13) and (2.14) above, that

$$d_{BL_1^*}\big(\mathscr{L}(G_{R_n}), \mathscr{L}(G_{R_0})\big) \to 0.$$

Now, this and Theorem 2.3 give, by the triangle inequality,

$$d_{BL_1^*}\big(\mathscr{L}^*(\nu_n^{R_n}), \mathscr{L}(G_{R_0})\big) \le d_{BL_1^*}\big(\mathscr{L}(\nu_n^{R_n}), \mathscr{L}(G_{R_n})\big)$$
$$+ d_{BL_1^*}\big(\mathscr{L}(G_{R_n}), \mathscr{L}(G_{R_0})\big) \to 0. \qquad \square$$

For instance if R_n is taken to be $P_n(\omega)$, then Corollary 2.7 gives an easy proof of the bootstrap CLT in [11] for these classes $\mathscr{F}$. This is not too interesting in view of the general results in [11], but $P_n(\omega)$ is not the only possible choice of R_n [e.g., one could take $R_0 = P_\theta$ and $R_n = P_{\hat{\theta}_n(\omega)}$ for a suitable estimator $\hat{\theta}_n(\omega)$ of θ]. Another application of Corollary 2.7 is to show that its conclusion holds if $H(R_n, R_0) \to 0$, where H is Hellinger distance (see the proof of Corollary 1.7 in [21]).

In [21], Sheehy and Wellner consider uniformty of the CLT over subsets Π of $\mathscr{P}(S)$. This is a more versatile concept than that of CLT_u considered here (see [21] for applications to the regularity of P_n as an estimator of P). It seems however too general to allow for a description as complete as that just obtained for CLT_u. These authors prove (Theorem 1.2, [21]) that if $\mathscr{F}$ satisfies Pollard's metric entropy condition, then $\mathscr{F}$ verifies the CLT uniformly in $P \in \Pi$ for any class Π such that

$$\lim_{\lambda\to\infty} \sup_{P\in\Pi} E_P F^2 I(F > \lambda) = 0. \tag{2.25}$$

Not surprisingly, the method of proof of Theorem 2.3 above allows replacement of the metric entropy condition by a weaker intrinsically Gaussian condition:

2.8. THEOREM. *Let $\Pi \subset \mathscr{P}(S)$ be a set of probability measures on $(S, \mathscr{S})$ and let $\mathscr{F}$ be a class of measurable real functions on S, NLDM(P) for all $P \in \Pi$. Let $\bar{F} = F \vee 1$, where F is the envelope of $\mathscr{F}$. Assume*

$$\sup_{Q\in\mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}} \big/ \big(E_Q\bar{F}^2\big)^{1/2} < \infty, \tag{2.26}$$

$$\lim_{\delta\to 0} \sup_{Q\in\mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}'(\delta(E_Q\bar{F}^2)^{1/2}, e_Q)} \big/ \big(E_Q\bar{F}^2\big)^{1/2} = 0 \tag{2.27}$$

and

$$\lim_{\lambda\to\infty} \sup_{P\in\Pi} E_P F^2 I(F \ge \lambda) = 0. \tag{2.28}$$

Then $\mathscr{F} \in$ CLT(P) for all $P \in \Pi$ and

$$\sup_{P\in\Pi} E\|Z_P\|_{\mathscr{F}} < \infty \quad \Big(\textit{hence } \sup_{P\in\Pi} E\|G_P\|_{\mathscr{F}} < \infty\Big), \tag{2.29}$$

$$\lim_{\delta\to 0} \sup_{P\in\Pi} E\|Z_P\|_{\mathscr{F}'(\delta, e_P)} = 0 \quad \Big(\textit{hence } \lim_{\delta\to 0} \sup_{P\in\Pi} E\|G_P\|_{\mathscr{F}'(\delta, e_P)} = 0\Big) \tag{2.30}$$

and

$$\lim_{n\to\infty}\sup_{P\in\Pi} d_{BL_1^*}\big(\mathscr{L}^*_{P,\mathscr{F}}(\nu_n^P), \mathscr{L}_{\mathscr{F}}(G_P)\big) = 0. \tag{2.31}$$

PROOF (Sketch). (2.28) implies both

$$1 \le c := \sup_{P\in\Pi} E_P \overline{F}^2 < \infty \tag{2.32}$$

and

$$\lim_{\lambda\to\infty}\sup_{P\in\Pi} E_P \overline{F}^2 I(\overline{F} \ge \lambda) = 0. \tag{2.33}$$

This and the usual truncation technique give

$$\lim_{n\to\infty}\sup_{P\in\Pi} P\left\{\left|\frac{E_{P_n}\overline{F}^2}{E_P\overline{F}^2} - 1\right| > \varepsilon\right\} = 0 \tag{2.34}$$

for all $\varepsilon > 0$. Truncation and symmetrization reduce the proof of

$$\lim_{n\to\infty}\sup_{P\in\Pi} (P^{\mathbb{N}})^*\{\|P_n - P\|_{(\mathscr{F}')^2} > \varepsilon\} = 0 \quad \text{for all } \varepsilon > 0 \tag{2.35}$$

to proving

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{P\in\Pi} (\Pr_P)^*\left\{\sup_{f\in(\mathscr{F}')^2}\left|\sum_{i=1}^n \varepsilon i f(X_i) I\big(\overline{F}(X_i) \le (\delta n)^{1/2}\big)/n\right| > \varepsilon\right\} = 0. \tag{2.36}$$

Proceeding as in (2.5), this probability can be bounded from above by

$$\frac{8(\delta\pi c)^{1/2}}{\varepsilon}\sup_{Q\in\mathscr{P}_f(S)}\frac{E\|Z_Q\|_{\mathscr{F}}}{\big(E_Q\overline{F}^2\big)^{1/2}} + P^N\left\{\frac{E_{P_n}\overline{F}^2}{E_P\overline{F}^2} > 2\right\}$$

and this gives (2.36), hence (2.35) by the hypothesis (2.26) and by (2.34).

As in the proof of Claim 3 in Theorem 2.3,

$$\begin{aligned}
&P^{\mathbb{N}}\{\|\nu_n^P\|_{\mathscr{F}'(\delta, e_P)} > 4\varepsilon\} \\
&\quad\le 2\left(1 - \frac{\delta^2}{4\varepsilon^2}\right)\Pr_P\left\{\left\|\sum_{i=1}^n \frac{\varepsilon_i\delta_{X_i}}{n^{1/2}}\right\|_{\mathscr{F}'(\delta, e_P)} > \varepsilon\right\} \\
&\quad\le 2\left(1 - \frac{\delta^2}{4\varepsilon^2}\right)\Bigg[(\Pr_P)^*\left\{\left\|\sum_{i=1}^n \frac{\varepsilon_i\delta_{X_i}}{n^{1/2}}\right\|_{\mathscr{F}'(2^{1/2}\delta(E_{P_n}\overline{F}^2)^{1/2}, e_{P_n})} > \left(\frac{2}{3c}\right)^{1/2}\varepsilon\big(E_{P_n}\overline{F}^2\big)^{1/2}\right\} \\
&\qquad + (P^{\mathbb{N}})^*\{\|P_n - P\|_{(\mathscr{F}')^2} > \delta^2\} + P^{\mathbb{N}}\left\{\left|\frac{E_{P_n}\overline{F}^2}{E_P\overline{F}^2} - 1\right| > \frac{1}{2}\right\}\Bigg]
\end{aligned}$$

and by Jensen's inequality,

$$E_\varepsilon\left\|\sum_{i=1}^n \varepsilon_i\delta_{X_i}/n^{1/2}\left(E_{P_n}\bar{F}^2\right)^{1/2}\right\|_{\mathscr{F}'(2^{1/2}\delta(E_{P_n}\bar{F}^2)^{1/2},\, e_{P_n})}$$

$$\le \sqrt{\pi/2}\sup_{Q\in\mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}'(2^{1/2}\delta(E_Q\bar{F}^2)^{1/2},\, e_{Q_n})}/\left(E_Q\bar{F}^2\right)^{1/2}.$$

Therefore, (2.34), (2.35) and hypothesis (2.27) give

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{P\in\Pi} P^{\mathbb{N}}\left\{\left\|\nu_n^P\right\|_{\mathscr{F}'(\delta, e_P)} > \varepsilon\right\} = 0 \quad \text{for all } \varepsilon > 0. \tag{2.37}$$

This is the key step in the proof of (2.31).

Sudakov's theorem and (2.26) imply $\log N(\varepsilon(E_{P_{n(\omega)}}\bar{F}^2)^{1/2}, \mathscr{F}, e_{P_{n(\omega)}}) < C/\varepsilon^2$ for all $\varepsilon > 0$ and some $C < \infty$; on the other hand (2.35) and (2.32) imply that for $P \in \Pi$, $\sup_{f,g\in\mathscr{F}}|e_{P_n}^2(f,g)/E_{P_n}\bar{F}^2 - e_P^2(f,g)/E_P\bar{F}^2| \to 0$ $P^{\mathbb{N}}$ a.s. These two observations then yield

$$\sup_{P\in\Pi}\log N(\varepsilon, \mathscr{F}, e_P) \le K/\varepsilon^2 \tag{2.38}$$

for all $\varepsilon > 0$ and some $K < \infty$. Now (2.38) gives, as in the proof of Claim 4, that both (2.29) and (2.30) hold. Finally, the limit (2.31) follows from (2.29), (2.30) and (2.37) as in Claim 5, by a finite dimensional result analogous to Lemma 2.1, easy to prove using that lemma together with a classical truncation argument [based on (2.28)], which we omit. □

2.9. EXAMPLE. In Section 3 we show that many interesting universal Donsker classes are UPG_f and/or UPG'_f. Suppose that $\mathscr{H} \in \mathrm{UPG}'_f$ and that its envelope H is a bounded function. Let F be any measurable function and assume for simplicity that $F(s) \ge 1$, $s \in S$. Then the class $\mathscr{F} = \{Fh\colon h \in \mathscr{H}\}$ satisfies hypotheses (2.26) and (2.27) of Theorem 2.8, and therefore it satisfies the conclusion of that theorem for any class Π for which $\sup_{P\in\Pi} E_P F^2 I(F \ge \lambda) \to 0$ as $\lambda \to \infty$. To verify (2.26) and (2.27), let $Q = \sum\alpha_i\delta_{s_i}$ with $\sum\alpha_i = 1$, $\alpha_i > 0$, and let $R = \sum\beta_i\delta_{s_i}$, with $\beta_i = \alpha_i F^2(s_i)/\sum\alpha_i F^2(s_i)$. Then

$$Z_Q(Fh)/\left(E_Q F^2\right)^{1/2} = \sum\alpha_i^{1/2} g_i F(s_i)h(s_i)/\left(\sum\alpha_i F^2(s_i)\right)^{1/2} = Z_R(h)$$

and

$$e_Q^2(Fh_1, Fh_2) = \sum\alpha_i F^2(s_i)\left(h_1(s_i) - h_2(s_i)\right)^2 = \left(E_Q F^2\right)e_R^2(h_1, h_2).$$

Hence (2.26) and (2.27) follow from (2.20) and (2.21) in the definition of UPG'_f.

Corollary 1.7 of [21] and Corollary 2.7 extend to the classes of functions $\mathscr{F}$ that satisfy the hypotheses of Theorem 2.8 (with a proof similar to that of Corollary 2.7).

Finally, we consider exponential bounds that work uniformly in P for empirical processes indexed by universal bounded Donsker classes ([23]) and by UPG'_f classes. We recall that $\mathscr{F}$ is a universal bounded Donsker class if the sequence $\{\|\nu_n^P\|_{\mathscr{F}}\}_{n=1}^\infty$ is stochastically bonded for all $P \in \mathscr{P}(S)$. Several equivalent definitions are given in [23] for these classes. In particular, a measurable class $\mathscr{F}$ is universal bounded Donsker if and only if $|c_f| < \infty$ for all $f \in \mathscr{F}$ and $\sup_{P \in \mathscr{P}(S)} E\|Z_P\|_{\tilde{\mathscr{F}}} < \infty$, where c_f and $\tilde{\mathscr{F}}$ are as in Definition 2.5.

For real random variables ξ, we define

$$\psi(\xi) = \inf\{c: E\exp(|\xi|^2/c^2) \le 2\}.$$

ψ is the Orlicz pseudonorm corresponding to the Young function $e^{x^2} - 1$.

2.10. THEOREM. *Let $\mathscr{F}$ be a measurable universal bounded Donsker class of functions, $\tilde{\mathscr{F}}$ be as in Definition* 2.5, $M = \sup\{$*median of* $\|Z_P\|_{\tilde{\mathscr{F}}}: P \in \mathscr{P}(S)\}$ *and let* $\tilde{F} = \sup_{f \in \tilde{\mathscr{F}}}|f|$. *Then*

$$\sup_{P \in \mathscr{P}(S)} \sup_{n \in \mathbb{N}} \psi\left(\|\nu_n^P\|_{\mathscr{F}}\right) \le (2\pi)^{1/2}\left(2\|\tilde{F}\|_\infty + (\log 2)^{-1/2}M\right). \tag{2.39}$$

PROOF. As mentioned above, the results in [23] show that both $\|\tilde{F}\|_\infty$ and M are finite. Let, for $P \in \mathscr{P}_f(S)$, $M_P =$ median of $\|Z_P\|_{\tilde{\mathscr{F}}}$ and $\sigma_P^2 = \sup_{f \in \tilde{\mathscr{F}}} E(Z_P(f))^2$. Borell's inequality ([5]), namely

$$\Pr\{|\,\|Z_P\|_{\tilde{\mathscr{F}}} - M_P| > t\} \le \exp\{-t^2/2\sigma_P^2\}, \qquad t \in \mathbb{R}_+,$$

implies that for $\alpha \le 1/(4\sigma_P^2)$,

$$E\exp\left\{\alpha\left(\|Z_P\|_{\tilde{\mathscr{F}}} - M_P\right)^2\right\} = 1 + \int_1^\infty \Pr\left\{\|Z_P\|_{\tilde{\mathscr{F}}} - M_P| > \left(\alpha^{-1}\log u\right)^{1/2}\right\} du$$

$$\le 1 + \int_1^\infty u^{-1/2\alpha\sigma_P^2}\,du = 1 + \left[\left(2\alpha\sigma_P^2\right)^{-1} - 1\right]^{-1} \le 2.$$

Therefore,

$$\psi(\|Z_P\|_{\tilde{\mathscr{F}}}) \le \psi(\|Z_P|_{\tilde{\mathscr{F}}} - M_P) + \psi(M_P) \le 2\sigma_P + (\log 2)^{-1/2}M_P.$$

Then, since $\sup_{P \in \mathscr{P}_f(S)} \sigma_P = |\tilde{F}\|_\infty$, we obtain

$$\sup_{P \in \mathscr{P}_f(S)} \psi(|Z_P|_{\tilde{\mathscr{F}}}) \le 2^{1/2}|\tilde{F}|_\infty + (\log 2)^{-1/2}M. \tag{2.40}$$

If P_n' is an independent copy of P_n and $\{\varepsilon_i\}, \{g_i\}$ are as in previous proofs, the uniformity of the bound (2.40), the norm properties of ψ and the convexity of the function $e^{\|x\|^2_{\tilde{\mathscr{F}}}}$ readily give (via the usual tools, namely, Jensen and Fubini)

$$\begin{aligned}\psi\left(\|\nu_n^P\|_{\mathscr{F}}\right) &\le \psi\left(n^{1/2}\|P_n - P_n'\|_{\mathscr{F}}\right) \le 2\psi\left(\left\|\sum_{i=1}^n \varepsilon_i\delta_{X_i}/n^{1/2}\right\|_{\mathscr{F}}\right)\\ &\le (2\pi)^{1/2}\psi\left(\left\|\sum_{i=1}^n g_i\delta_{X_i}/n^{1/2}\right\|_{\tilde{\mathscr{F}}}\right)\\ &\le (2\pi)^{1/2}\left(2\|\tilde{F}\|_\infty + (\log 2)^{-1/2}M\right)\end{aligned} \tag{2.41}$$

for all n and P. [Note the crucial role of uniformity for the estimate (2.40) in the last inequality of (2.41).] □

REMARK. A similar result (and proof) also holds for an arbitrary type 2 operator between two Banach spaces.

2.11. COROLLARY. (a) *If $\mathscr{F}$ is a measurable, universal bounded Donsker class then for all $t > 0$,*

$$\sup_{P \in \mathscr{P}(S)} \sup_{n \in \mathbb{N}} P^{\mathbb{N}}\left\{\|\nu_n^P\|_{\mathscr{F}} > t\right\} \le 2\exp\left\{-t^2/2\pi\left(2\|\tilde{F}\|_\infty + (\log 2)^{-1/2}M\right)^2\right\}. \tag{2.42}$$

(b) *Let $\mathscr{F}$ be such that $|c_f| < \infty$ for all $f \in \mathscr{F}$ and $\tilde{\mathscr{F}}$ is a measurable* UPG'_f *class. Let for each $\tau > 0$, $M(\tau) = \sup\{$median of $\|Z_P\|_{\tilde{\mathscr{F}}'(\tau, e_P)}$: $P \in P_f(S)\}$ (which by definition tends to 0 at $\tau \to 0$) and let $\overline{M} = \sup\{$median of $\|Z_P\|_{(\tilde{\mathscr{F}}')^2}$: $P \in \mathscr{P}_f(S)\}$ [which is finite by* (2.3) *and the Slepian–Fernique lemma]. Then inequality* (2.42) *holds and moreover, for all $\delta > 0$, $t > \delta/2$ and $n \in \mathbb{N}$,*

$$\begin{aligned}\sup_{P \in \mathscr{P}(S)} & P^{\mathbb{N}}\left\{\|\nu_n^P\|_{\mathscr{F}'(\delta, e_P)} > 4t\right\} \\ &\le 4\left(1 - \delta^2/4t^2\right)^{-1}\left[\exp\left\{-t^2/2\pi\left(4\delta^2 + (\log 2)^{-1/2}M(2^{1/2}\delta)\right)\right\}\right. \\ &\qquad \left. + \exp\left\{-\delta^4 n/2\pi\left(8\|\tilde{F}\|_\infty^2 + (\log 2)^{-1/2}\overline{M}\right)^2\right\}\right].\end{aligned} \tag{2.43}$$

REMARK. $\overline{M}$ can be bounded in terms of M and $\|\tilde{F}\|_\infty$. A possible way to proceed is as follows: By Borell's inequality,

$$\overline{M} \le 2^{-1}\pi^{1/2}\sigma + \sup_{P \in \mathscr{P}_f(S)} E\|Z_P\|_{(\tilde{\mathscr{F}}')^2},$$

where $\bar{\sigma}^2 = \sup_{P \in \mathscr{P}_f(S)} \sup_{h \in (\tilde{\mathscr{F}}')^2} Ph^2 \le 4\|\tilde{F}\|_\infty^2$, so that by (2.4) and the Slepian–Fernique lemma, $\overline{M} \le 2\|\tilde{F}\|_\infty^2 + 16\|\tilde{F}\|_\infty \sup_{P \in \mathscr{P}_f(S)} E\|Z_P\|_{\tilde{\mathscr{F}}}$ and $E\|Z_P\|_{\tilde{\mathscr{F}}} \le cM_p$ for a universal constant c (again by Borell's inequality).

PROOF OF COROLLARY 2.11. Inequality (2.42) follows directly from (2.39) and Markov's inequality applied to the random variable $\exp\{\|\nu_n^P\|_{\mathscr{F}}^2/\psi^2(\|\nu_n^P\|_{\mathscr{F}})\}$. To prove (2.43), we proceed in analogy with the proof of (2.6), which gives

$$\begin{aligned}P^{\mathbb{N}}&\left\{\|\nu_n^P\|_{\mathscr{F}'(\delta, e_P)} > 4t\right\} \\ &\le 2\left(1 - \delta^2/4t^2\right)^{-1}\left[\Pr\left\{\left\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n^{1/2}\right\|_{\mathscr{F}'(2^{1/2}\delta, e_{P_n})} > t\right\}\right. \\ &\qquad \left. + P^{\mathbb{N}}\left\{\|\nu_n^P\|_{(\tilde{\mathscr{F}}')^2} > \delta^2 n^{1/2}\right\}\right].\end{aligned} \tag{2.44}$$

Then we note that $\sup_{f \in \tilde{\mathscr{F}}'(2^{1/2}\delta, e_{P_n})} E(Z_{P_n}(f))^2 \le 2\delta^2$ and that the median of $\|Z_{P_n}\|_{\tilde{\mathscr{F}}'(2^{1/2}\delta, e_{P_n})}$ is bounded above by $M(2^{1/2}\delta)$, so that we obtain, as in the proof of Theorem 2.10 and part (a) of this corollary,

$$\sup_{P\in\mathscr{P}(S)}\sup_{n\in\mathbb{N}}\Pr\left\{\left\|\sum_{i=1}^{n}\varepsilon_i\delta_{X_i}/n^{1/2}\right\|_{\tilde{\mathscr{F}}'(2^{1/2}\delta,\,e_{P_n})}>t\right\} \tag{2.45}$$

$$\le 2\exp\left\{-t^2/2\pi\left(2\cdot 2\delta^2+(\log 2)^{-1/2}M(2^{1/2}\delta)\right)^2\right\}.$$

Since $(\tilde{\mathscr{F}}')^2$ is universal bounded Donsker by (2.4), direct application of inequality (2.42) gives

$$P^{\mathbb{N}}\left\{\|\nu_n^P\|_{(\tilde{\mathscr{F}}')^2}>\delta^2 n^{1/2}\right\}\le 2\exp\left\{-\delta^4 n/2\pi\left(8\|\tilde{F}\|_\infty^2+(\log 2)^{-1/2}\overline{M}\right)^2\right\}. \tag{2.46}$$

Now, (2.43) follows from (2.44)–(2.46). □

The constants in the above inequality are not the best possible: the extraneous coefficient 2π is due to the randomization procedure, which conceivably could be made more efficient; the factor 2 of $\|\tilde{F}\|_\infty$ can be decreased at the expense of the coefficient 2 in front of the exponential. An advantage of these inequalities is that they apply in situations which are not covered by entropy conditions. Also (given, of course, the right ingredient from Gaussian theory—in this case, the deep Borell's inequality) it is difficult to think of a simpler proof of a Kiefer type inequality for general empirical processes.

3. Some uniform Donsker classes of functions. Let $N(\varepsilon, \mathscr{F}, e_P)$, $\varepsilon > 0$, denote the covering numbers of $(\mathscr{F}, e_P)$ (as defined in Claim 3 above). We then have:

3.1. PROPOSITION. *Let $\mathscr{F}$ be measurable and such that $\sup_{f\in\mathscr{F}}(\operatorname{diam} f)<\infty$. Then the conditions*

$$\sup_{Q\in\mathscr{P}_f(S)}\int_0^\infty\left(\log N\left(\varepsilon,\tilde{\mathscr{F}},e_Q\right)\right)^{1/2}d\varepsilon<\infty \tag{3.1}$$

and

$$\lim_{\delta\to 0}\sup_{Q\in\mathscr{P}_f(S)}\int_0^\delta\left(\log N\left(\varepsilon,\tilde{\mathscr{F}},e_Q\right)\right)^{1/2}d\varepsilon=0 \tag{3.2}$$

imply $\mathscr{F}\in$ UPG$_f$ (hence also $\tilde{\mathscr{F}}\in$ UPG$'_f$). Therefore, $\mathscr{F}\in$ CLT$_u$.

PROOF. Since $N(\varepsilon, \tilde{\mathscr{F}}, e_Q)\ge N(\varepsilon, \mathscr{F}, \rho_Q)$ for all Q, Proposition 3.1 follows just from Dudley's theorem on sample continuity of Gaussian processes ([6]; see also the version in [15] in terms of expected values) and from Theorem 2.3. □

As a consequence of Proposition 3.1, if $\tilde{\mathscr{F}}$ satisfies Pollard's entropy condition

$$(3.3)\qquad \int_0^\infty \sup_{Q\in\mathscr{P}_f(S)} \left(\log N\left(\varepsilon, \tilde{\mathscr{F}}, e_Q\right)\right)^{1/2} d\varepsilon < \infty,$$

then $\mathscr{F}\in \mathrm{UPG}_f$. In particular, this is true if $\mathscr{F}$ is the class of indicators of a Vapnik–Červonenkis family of sets ([7]).

3.2. PROPOSITION. *Let* $\mathscr{F}=\{f_k\}_{k=2}^\infty$ *with* $\|f_k\|_\infty = o(1/(\log k)^{1/2})$. *Then* $\mathscr{F}\in$ UPG (*hence* $\mathscr{F}\in$ UPG$'$). *Therefore* $\mathscr{F}\in \mathrm{CLT}_u$.

PROOF. Let $\alpha_k = (\log k)^{1/2}\|f_k\|_\infty \to 0$, $\bar\alpha = \sup_{k\geq 2}\alpha_k$, $\bar\alpha_N = \sup_{k\geq N}\alpha_k \to 0$. A classical computation shows that if g_k are $N(0,1)$ (not necessarily independent), then

$$(3.4)\qquad E\sup_k|\alpha_k g_k|/(\log k)^{1/2} \leq c\bar\alpha$$

for some $c<\infty$. Since $G_P(f_k) = (\mathrm{Var}_P(f_k))^{1/2}g_k$ and $(E_P f_k^2)^{1/2} \leq \alpha_k/(\log k)^{1/2}$, (3.4) gives

$$(3.5)\qquad \sup_{P\in\mathscr{P}(S)} E\|G_P\|_{\mathscr{F}} < \infty.$$

Moreover, these observations also imply

$$E\|G_P\|_{\mathscr{F}'(\delta,\rho_P)} \leq 2E\sup_{k\geq N}|G_P(f_k)| + E_P \sup_{k,l\leq N,\, f_k-f_l\in\mathscr{F}'(\delta,\rho_P)} |G_P(f_k)-G_P(f_l)|$$
$$\leq 2c\bar\alpha_N + \delta N^2.$$

So, for all N,

$$(3.6)\qquad \limsup_{\delta\to 0}\sup_{P\in\mathscr{P}(S)} E\|G_P\|_{\mathscr{F}'(\delta,\rho_P)} \leq 2c\bar\alpha_N \to 0.$$

(3.5) and (3.6) imply $\mathscr{F}\in$ UPG. □

Propositions 3.1 and 3.2 show that all the classes of functions considered in Figure 1 of [8] are uniform Donsker except perhaps $(1.4)^{co}$ (although it is obvious that if $\mathscr{F}$ is universal Donsker, then so is its convex hull, we do not know if this property holds for uniform Donsker classes). So, there is a wide variety of classes $\mathscr{F}$ for which CLT(P) holds uniformly in P. We may ask whether there are any uniformly bounded, universal Donsker classes of functions which are not UPG'_f (hence, not UPG_f). The answer is positive, as the following example shows.

3.3. EXAMPLE. Let H be a separable infinite dimensional Hilbert space and let H_1 be its unit ball with center $0\in H$. Take $S=H_1$, $\mathscr{S}$ equals the Borel sets of H_1 and $\mathscr{F}=H_1$ acting on S by inner product. Then $\mathscr{F}$ is universal Donsker since bounded random variables with values in H satisfy the central limit theorem (e.g., [2], Section 3.7 and [10], Lemma 5.4 and

Chapter 4). We show that $\mathscr{F} \notin \mathrm{UPG}'_f$. Let $\{e_i\}$ be an orthonormal basis for H and for each $N \in \mathbb{N}$,

$$Q_N = N^{-1} \sum_{i=1}^{N} \delta_{e_i}.$$

Then, for $\delta^2 N \geq 2$ and for some $c > 0$ independent of δ and N, if $x_i = \langle x, e_i \rangle$, $x \in H$, we have

$$E\|Z_Q\|_{\mathscr{F}'(\delta, e_{Q_N})} = E \sup_{\substack{\|x\| \leq 2 \\ \sum_{i=1}^N x_i^2 \leq \delta^2 N}} \left| \sum_{i=1}^{N} g_i x_i \right| / N^{1/2}$$

$$= \left(2 \wedge \sqrt{\delta^2 N}\right) E \left(\sum_{i=1}^{N} g_i^2 \right)^{1/2} / N^{1/2} \geq c\left(2 \wedge \sqrt{\delta^2 N}\right).$$

Hence,

$$\liminf_{\delta \to 0} \sup_{Q \in \mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}'(\delta, e_Q)} \geq 2c,$$

that is, $\mathscr{F} \notin \mathrm{UPG}'_f$. (And, since $\mathscr{F}$ is uniformly bounded, $\mathscr{F}$ is not in UPG_f either.)

So we have uniform Donsker $\substack{\Rightarrow \\ \not\Leftarrow}$ universal Donsker. The situation is different for the bounded Donsker property. In fact we have from [23] that

$$\sup_n \sup_{P \in \mathscr{P}(S)} E_P \|\nu_n^P\|_{\mathscr{F}} < \infty \Leftrightarrow \sup_{P \in \mathscr{P}(S)} E\|G_P\|_{\mathscr{F}} < \infty$$

$$\Leftrightarrow \sup_n E_P \|\nu_n^P\|_{\mathscr{F}} < \infty \quad \text{for all } P,$$

that is, uniform bounded Donsker $\Leftrightarrow$ universal bounded Donsker. Note also that universal bounded Donsker $\not\Rightarrow$ universal Donsker: take $\|f_k\| = 1/(\log k)$, $k \geq 3$, in Proposition 3.2 to obtain a class that is universal bounded Donsker but not universal Donsker.

Obviously in Example 3.3 it is enough to take $S = \{e_i\}_{i=1}^{\infty}$. In this case it becomes the example of Proposition 6.3 of Dudley [8] [replacing $I_{A(j)}$ by e_j in Dudley's example constitutes only a reformulation of the same example].

Proposition 3.2 provides the natural candidates for classes $\mathscr{F}$ which are CLT_u and yet do not satisfy the entropy conditions (3.1) and (3.2) (after all, c_0 is the counterexample space!). It could be argued that this is an extreme type of classes $\mathscr{F}$ and therefore that (3.1) and (3.2) (or even the slightly stronger condition in [21]) do give all the CLT_u classes that will ever be needed. In fact we can even produce such classes in Hilbert space, which, from many points of view, is as far from c_0 as it can be. For this we use a result of Mityagin [16] on metric entropy of ellipsoids together with the following simple lemma.

3.4. Lemma. *Let H and H_1 be as in Example* 3.3, *let $S = \{x_k : k \in \mathbb{N}\} \subset H$ with $\|x_k\| \to 0$ and let $\mathscr{F} = H_1$, acting on S by inner product. Then $\mathscr{F} \in \mathrm{UPG}'_f$.*

PROOF. To see this, let $Q = \sum \alpha_k \delta_{x_k}$ with $\sum \alpha_k = 1$, $\alpha_k \geq 0$. Then $Z_Q = \sum \alpha_k^{1/2} g_k \delta_{x_k}$, $\{g_k\}$ i.i.d. $N(0,1)$. We have

$$E\|Z_Q\|_{\mathscr{F}} = E\|\sum \alpha_k^{1/2} g_k x_k\| \leq \left(\sum \alpha_k \|x_k\|^2\right)^{1/2} \leq \sup_{k \in \mathbb{N}} \|x_k\| \tag{3.7}$$

and

$$\begin{aligned} E\|Z_Q\|_{\mathscr{F}'(\delta, e_Q)} &= E \sup_{\sum \alpha_k \langle x_k, z\rangle^2 \leq \delta^2, \|z\| \leq 2} \left|\sum \alpha_k^{1/2} \langle x_k, z\rangle g_k\right| \\ &\leq E \sup_{\sum_{k=1}^n \alpha_k \langle x_k, z\rangle^2 \leq \delta^2} \left|\sum_{k=1}^n \alpha_k^{1/2} \langle x_k, z\rangle g_k\right| \\ &\quad + 2E \sup_{\|z\| \leq 1} \left|\sum_{k=n+1}^\infty \alpha_k^{1/2} \langle x_k, z\rangle g_k\right| \\ &\leq \delta E\left(\sum_{k=1}^n g_k^2\right)^{1/2} + 2E\left\|\sum_{k=n+1}^\infty \alpha_k^{1/2} g_k x_k\right\| \\ &\leq \delta n^{1/2} + 2\left(\sum_{k=n+1}^\infty \alpha_k \|x_k\|^2\right)^{1/2} \\ &\leq \delta n^{1/2} + 2 \sup_{k>n} \|x_k\|. \end{aligned} \tag{3.8}$$

Now, (3.7) gives $\sup_{Q \in \mathscr{P}_f(S)} E\|Z_Q\|_{\mathscr{F}} < \infty$ and (3.8) gives

$$\lim_{\delta \to 0} \sup_Q E\|Z_Q\|_{\mathscr{F}'(\delta, e_Q)} \leq \lim_{n \to \infty} \lim_{\delta \to 0} \left(\delta n^{1/2} + 2 \sup_{k>n} \|x_k\|\right) = 0.$$

Hence $\mathscr{F} \in \mathrm{UPG}'_f$. □

It would be interesting to know exactly what compact subsets S of H verify the property that the CLT holds uniformly on all P with supports in S. Lemma 3.4 gives a class of compact sets for which this is true and shows that this property may be unrelated to S being a GC set since the sets in the lemma if we take $x_k = b_k e_k$ with $b_k \to 0$ are GC if and only if $b_k = o(1/(\log k)^{1/2})$ (Dudley [6], Proposition 6.7).

3.5. EXAMPLE. Let S and $\mathscr{F}$ be as in Lemma 3.4 with $x_k = b_k e_k$ and $b_k = (\log k)^{-\delta/2}$ for some $\delta \in (0,1)$ and $k \geq 3$. Although $\mathscr{F} \in \mathrm{UPG}'_f$ by Lemma 3.4, we will show that $\mathscr{F}$ verifies

$$\sup_{Q \in \mathscr{P}_f(S)} \int_0^\infty \left(\log N(\varepsilon, \mathscr{F}, e_Q)\right)^{1/2} d\varepsilon = \infty, \tag{3.9}$$

that is, $\mathscr{F}$ does not verify (3.1). To see this, we let $P = \sum_{k=1}^\infty \alpha_k \delta_{b_k e_k}$ with $\alpha_k = c/k(\log k)^{1+\delta}$, and (3.9) will follow for the subset $\{P_n(\omega)\}$ of $\mathscr{P}_f(S)$ for some $\omega \in \Omega$. The distance e_P is defined by the norm $|||z|||^2 = c\sum_{k=1}^\infty z_k^2/k(\log k)^{1+2\delta}$. The metric entropy of the unit ball $\mathscr{F} = H_1$ with respect to e_P is the same as the metric entropy of the ellipsoid

$$K = \left\{ u \in H: \sum_{k=1}^{\infty} k(\log k)^{1+2\delta} u_k^2 \leq c \right\}$$

with respect to the Hilbert space norm [as is easily seen by the change of variables $u_k = (ck(\log k)^{1+2\delta})^{-1/2} z_k$]. By Mityagin [16], Section 2, Theorem 3, this entropy is bounded by $\int_0^{1/2\varepsilon} t^{-1} m(t)\, dt$ with $m(t) = \sup\{k: k^{1/2}(\log k)^{\delta+1/2} \leq \varepsilon\}$, that is,

$$\log N(\varepsilon, \mathscr{F}, e_P) = \log N(\varepsilon, K, \|\cdot\|) \gtrsim \varepsilon^{-2}(\log \varepsilon^{-1})^{-1-2\delta}.$$

But the square root of this function is not integrable at zero, that is,

$$(3.10) \qquad \int_0^{\infty} (\log N(\varepsilon, \mathscr{F}, e_P))^{1/2}\, d\varepsilon = \infty.$$

Now we notice that if $P_n(\omega) = n^{-1}\sum_{i=1}^n \delta_{X_i(\omega)}$, $n \in \mathbb{N}$, is the empirical measure corresponding to P,

$$\sup\left\{ \left| e_{P_n(\omega)}^2(f, g) - e_P^2(f, g) \right| : f, g \in \mathscr{F} \right\}$$

$$\leq 2 \sup_{\|z\| \leq 1} \left| \left\langle \frac{\sum_{i=1}^n X_i}{n}, z \right\rangle^2 - \left\langle \int x\, dP, z \right\rangle^2 \right|$$

$$\leq 4 \left\| \frac{\sum_{i=1}^n (X_i - EX_i)}{n} \right\| \to 0 \quad \text{a.s.}$$

by the law of large numbers in H. Therefore, for all ω in a set of probability 1,

$$\liminf_{n \to \infty} N(\varepsilon/2, \mathscr{F}, e_{P_n(\omega)}) \geq N(\varepsilon, \mathscr{F}, e_P).$$

Then, Fatou's lemma and (3.10) yield

$$\infty = \int_0^{\infty} (\log N(\varepsilon, \mathscr{F}, e_P))^{1/2}\, d\varepsilon \leq \liminf_{n \to \infty} \int_0^{\infty} \left(\log N(\varepsilon/2, \mathscr{F}, e_{P_n(\omega)})\right)^{1/2} d\varepsilon$$

$$\leq 2 \sup_{Q \in \mathscr{P}_f(S)} \int_0^{\infty} (\log N(\varepsilon, \mathscr{F}, e_Q))^{1/2}\, d\varepsilon,$$

that is, $\mathscr{F}$ does not satisfy (3.1).

Acknowledgments. We thank Richard M. Dudley and David Pollard for interesting conversations on the subject of this article. We also thank Ann Sheehy and Jon Wellner for making their unpublished work available to us. This research was started when the first named author was affiliated with Texas A & M University, and he acknowledges the A & M Mathematics Department for providing a stimulating research environment.

REFERENCES

[1] ANDERSON, N. T. and DOBRIĆ, V. (1987). The central limit theorem for stochastic processes. *Ann. Probab.* **15** 164–177.

[2] ARAUJO, A. and GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York.

[3] BERAN, R. and MILLAR, P. W. (1987). Stochastic estimation and testing. *Ann. Statist.* **15** 1131–1154.

[4] BERAN, R. and MILLAR, P. W. (1989). A stochastic minimum distance test for multivariate parametric models. *Ann. Statist.* **17** 125–140.

[5] BORELL, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30** 205–216.

[6] DUDLEY, R. M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1** 290–330.

[7] DUDLEY, R. M. (1984). A course on empirical processes. *École d'Éte de Probabilités de Saint-Flour XII–1982. Lecture Notes in Math.* **1097** 2–142. Springer, Berlin.

[8] DUDLEY, R. M. (1987). Universal Donsker classes and metric entropy. *Ann. Probab.* **15** 1306–1326.

[9] GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–989.

[10] GINÉ, E. and ZINN, J. (1986). Lectures on the central limit theorem for empirical processes. *Singularities and Constructive Methods for Their Treatment. Lecture Notes in Math.* **1221** 50–113. Springer, Berlin.

[11] GINÉ, E. and ZINN, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18** 851–869.

[12] HOFFMANN-JØRGENSEN, J. (1984). Stochastic processes on Polish spaces. Unpublished manuscript.

[13] HOFFMANN-JØRGENSEN, J. and PISIER, G. (1976). The law of large numbers and the central theorem in Banach spaces. *Ann. Probab.* **4** 587–599.

[14] JAIN, N. and MARCUS, M. B. (1975). Central limit theorems for $C(S)$-valued random variables. *J. Funct. Anal.* **19** 216–231.

[15] MARCUS, M. B. and PISIER, G. (1981). *Random Fourier Series with Applications to Harmonic Analysis*. Princeton Univ. Press.

[16] MITYAGIN, B. S. (1961). Approximation dimension and bases in nuclear spaces. *Russian Math. Surveys* **16** 59–127.

[17] PISIER, G. (1984). Remarques sur les classe de Vapnik–Červonenkis. *Ann. Inst. H. Poincaré Probab. Statist.* **20** 287–298.

[18] PISIER, G. (1986). Probabilistic methods in the geometry of Banach spaces. *Probability and Analysis, Varenna (Como) 1985. Seminar. Lecture Notes in Math.* **1206** 167–241. Springer, Berlin.

[19] POLLARD, D. (1981). Limit theorems for empirical processes. *Z. Wahrsch. Verw. Gebiete* **57** 181–196.

[20] SAZONOV, V. V. (1981). *Normal Approximation: Some Recent Advances. Lecture Notes in Math.* **879**. Springer, Berlin.

[21] SHEEHY, A. and WELLNER, J. (1988). Uniformity in P of some limit theorems for empirical measures and processes. Technical Report 134, Dept. Statistics, Univ. Washington.

[22] ZINN, J. (1977). A note on the central limit theorem in Banach spaces. *Ann. Probab.* **5** 283–286.

[23] ZINN, J. (1990). Universal Donsker classes and type 2. In *Proc. Sixth Internat. Conf. Probab. Banach Spaces* 283–288. Birkhäuser, Boston.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CONNECTICUT
STORRS, CONNECTICUT 06268

DEPARTMENT OF MATHEMATICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843