The Annals of Probability
1990, Vol. 18, No. 2, 851–869

BOOTSTRAPPING GENERAL EMPIRICAL MEASURES

By Evarist Giné[1] and Joel Zinn[2]

Texas A & M University

It is proved that the bootstrapped central limit theorem for empirical processes indexed by a class of functions $\mathscr{F}$ and based on a probability measure P holds a.s. if and only if $\mathscr{F} \in \mathrm{CLT}(P)$ and $\int F^2\,dP < \infty$, where $F = \sup_{f \in \mathscr{F}} |f|$, and it holds in probability if and only if $\mathscr{F} \in \mathrm{CLT}(P)$. Thus, for a large class of statistics, no local uniformity of the CLT (about P) is needed for the bootstrap to work. Consistency of the bootstrap (the bootstrapped law of large numbers) is also characterized. (These results are proved under certain weak measurability assumptions on $\mathscr{F}$.)

1. Introduction. Efron (1979) introduced the "bootstrap," a resampling method for approximating the distribution functions of statistics $H_n(X_1, \dots, X_n; P)$, where the random variables X_i are independent, identically distributed with common law P [i.i.d.(P)]. Since the empirical measure

$$P_n(\omega) = n^{-1} \sum_{i=1}^{n} \delta_{X_i(\omega)} \tag{1.1}$$

is (a.s.) close to P, one may hope that, if $\hat{X}_{n1}, \dots, \hat{X}_{nn}$ are i.i.d.($P_n(\omega)$) (i.e., the $\hat{X}_{ni}$ are obtained by sampling from the data, with replacement), then the distribution of $\hat{H}_n(\omega) = H_n(\hat{X}_{n1}, \dots, \hat{X}_{nn}; P_n(\omega))$ is ω-a.s. asymptotically close to that of $H_n(X_1, \dots, X_n; P)$. In turn, the distribution of the bootstrapped statistic, $\hat{H}_n(\omega)$ can be approximated by Monte Carlo simulation. This suggestive method has been validated with limit theorems for many particular $\hat{H}_n(\omega)$ by Efron (1979), Bickel and Freedman (1981), Singh (1981), Beran (1982, 1984), Bretagnolle (1983), Gaenssler (1987) and others. In this article we offer a justification of the bootstrap for functions H_n of a special type, namely for continuous functions of the empirical measure viewed as an element of $\ell^\infty(\mathscr{F})$, for classes of functions $\mathscr{F}$. Such H include the Kolmogorov–Smirnov and the Cramér–von Mises statistics (in any number of dimensions) as well as the statistics considered in Beran and Millar (1986).

Let $(S, \mathscr{S}, P)$ be a probability space, let X_i: $(S^{\mathbb{N}}, \mathscr{S}^{\mathbb{N}}, P^{\mathbb{N}}) \to (S, \mathscr{S}, P)$ be the coordinate functions [i.i.d.(P)], let $P_n(\omega)$ be as in (1.1) for $\omega \in S^{\mathbb{N}}$, let $\hat{X}^{\omega}_{nj}$, $j = 1, \dots, n$, be i.i.d.($P_n(\omega)$), let $\hat{P}_n(\omega)$ be the empirical measure based on $\{\hat{X}^{\omega}_{nj}\}_{j=1}^{n}$, i.e.,

$$\hat{P}_n(\omega) = n^{-1} \sum_{j=1}^{n} \delta_{\hat{X}^{\omega}_{nj}}, \tag{1.2}$$

Received March 1988; revised January 1989.

[1]Research partially supported by National Science Foundation Grant No. DMS-86-19411.

[2]Research partially supported by National Science Foundation Grant No. DMS-86-01250 and Texas Advanced Research Program Grant No. 3285.

AMS 1980 *subject classifications*. Primary 60F17, 62E20; secondary 60B12.

Key words and phrases. Bootstrapping, empirical processes, central limit theorem, law of large numbers.

and let $\mathscr{F}$ be a class of measurable functions on $(S, \mathscr{S})$ such that

$$F = \sup_{f \in \mathscr{F}} |f| \tag{1.3}$$

is finite for all $s \in S$. We then prove that, under some measurability on $\mathscr{F}$, the conditions

$$\int F^2 \, dP < \infty \tag{1.4}$$

and

$$n^{1/2}(P_n - P) \to G_P \quad \text{weakly in } l^\infty(\mathscr{F}) \tag{1.5}$$

are necessary and sufficient for

$$n^{1/2}\big(\hat{P}_n(\omega) - P_n(\omega)\big) \to G \quad \text{weakly in } l^\infty(\mathscr{F}),\ \omega\text{-a.s.} \tag{1.6}$$

for a centered Gaussian process G independent of ω. Then G coincides with G_P, the Gaussian limit in (1.5).

Thus, this result completely settles, modulo measurability, the question of the validity of the bootstrap for the CLT for empirical processes indexed by classes of functions (or sets).

The main feature of this theorem, aside from its generality, is that no assumptions are made on local uniformity (about P) of the CLT (1.5) for the bootstrap CLT (1.6) to hold [this was unexpected in view of, e.g., the comments in Bickel and Freedman (1981), page 1209]. Another new feature is necessity of the integrability condition (1.4) and the usual CLT (1.5) for the bootstrap.

The proof relies on several results and techniques from probability in Banach spaces. Among other such results and techniques, we use symmetrization by randomization in an essential way [an idea in Pisier (1985) has been useful in connection with this], results of Le Cam (1970) on Poissonization and on the CLT in Banach spaces, integrability of Gaussian processes [e.g., Fernique (1984)], Hoffmann-Jørgensen's (1974) inequality and convergence of moments in the CLT in Banach spaces [de Acosta and Giné (1979)], results on empirical processes from Giné and Zinn (1984, 1986) and, particularly, a result of Ledoux, Talagrand and Zinn [cf. Ledoux and Talagrand (1988b)] on the almost sure weak convergence of $\sum_{i=1}^n g_i X_i(\omega)/n^{1/2}$, g_i i.i.d. with $\int_0^\infty (P\{|g_1| > t\})^{1/2}\, dt < \infty$ (i.e., $g_1 \in L_{2,1}$). Actually, it is this last result that is at the base of our proof. The Ledoux–Talagrand–Zinn result uses for its proof a recent extension of Yurinski's decomposition as applied to $E_g\|\sum g_i x_i\| - E\|\sum g_i X_i\|$. This was observed by Ledoux and Talagrand (1988a) in the proof of one of the main results about the law of the iterated logarithm in Banach spaces.

The above techniques (except for the result of Ledoux, Talagrand and Zinn) are also used to obtain a similar result for the bootstrap in probability. The a.s. results are given in Section 2 and Section 3 contains the "in probability" result.

The bootstrapped law of the large numbers, much easier to prove than the CLT, is also characterized.

2. The a.s. bootstrapped limit theorems. Given P, a probability measure on a measure space $(S, \mathscr{S})$, we let

$$(2.1)\quad \rho_P^2(f,g) = \int (f-g)^2\, dP - \left(\int (f-g)\, dP\right)^2, \qquad f,g \in \mathscr{L}_2(P),$$

$$(2.2)\quad e_P^2(f,g) = \int (f-g)^2\, dP, \qquad f,g \in \mathscr{L}_2(P)$$

and, given a collection $\mathscr{F}$ of P-square integrable functions on $(S, \mathscr{S})$, we let

$$(2.3)\quad \begin{aligned} \mathscr{F}_\delta' &= \{f-g\colon f,g \in \mathscr{F}, e_P(f,g) \le \delta\}, \qquad \delta > 0,\\ \mathscr{F}' &= \{f-g\colon f,g \in \mathscr{F}\} \end{aligned}$$

and

$$(2.4)\quad (\mathscr{F}')^2 = \{(f-g)^2\colon f,g \in \mathscr{F}\}.$$

$G_P := \{G_P(f)\colon f \in \mathscr{F}\}$ denotes a centered Gaussian process indexed by $\mathscr{F}$, with covariance

$$(2.5)\quad EG_P(f)G_P(g) = \int fg\,dP - \int f\,dP \int g\,dP, \qquad f,g \in \mathscr{F}$$

and $Z_P := \{Z_P(f)\colon f \in \mathscr{F}\}$ denotes the centered Gaussian process with

$$(2.6)\quad EZ_P(f)Z_P(g) = \int fg\,dP, \qquad f,g \in \mathscr{F}.$$

We recall Hoffmann-Jørgensen's (1984) definition of weak convergence in $\ell^\infty(\mathscr{F})$, the space of bounded functions $\mathscr{F} \to \mathbb{R}$ with the sup norm topology: a sequence $\{Y_n\}_{n=1}^\infty$ of random elements of $\ell^\infty(\mathscr{F})$ *converges weakly* in $\ell^\infty(\mathscr{F})$ if there exists a Radon probability measure γ on $\ell^\infty(\mathscr{F})$ such that for all $H\colon \ell^\infty(\mathscr{F}) \to \mathbb{R}$ bounded and continuous,

$$\lim_{n\to\infty} E^*H(Y_n) = \int H\, d\gamma.$$

Then we say that $\mathscr{F} \in \mathrm{CLT}(P)$ if the sequence $\{n^{1/2}(P_n - P)(f)\colon f \in \mathscr{F}\}$ converges weakly in $\ell^\infty(\mathscr{F})$ to a Radon centered Gaussian probability measure γ_P on $\ell^\infty(\mathscr{F})$. γ_P is the law of G_P which, by virtue of the Radonicity of γ_P, admits a version with bounded uniformly continuous paths on $(\mathscr{F}, \rho_P)$, and $(\mathscr{F}, \rho_P)$ is totally bounded [see, e.g., Giné and Zinn (1986)]. We continue denoting this version by G_P.

If $\mathscr{F}$ satisfies certain measurability conditions, then P_n can be randomized (i.e., we can replace $\delta_{X_i} - P$ by $\xi_i \delta_{X_i}$ with ξ_i symmetric, independent of X_i and satisfying certain integrability conditions) and Fubini's theorem can be used freely. These conditions, spelled out in Giné and Zinn (1984), are that $\mathscr{F}$ be nearly linearly deviation measurable for P, NLDM(P) for short, and that both $\mathscr{F}^2$ and $\mathscr{F}'^2$ are nearly linearly supremum measurable for P, NLSM(P). In this paper if $\mathscr{F}$ satisfies all of the above conditions with respect to P we write $\mathscr{F} \in M(P)$. To see why $\mathscr{F} \in M(P)$ suffices we note, as in Giné and Zinn

(1984), Remark 2.4 (2), page 935, that the measurability of the map

$$(a_1,\dots,a_n,x_1,\dots,x_n) \to \sup_{f\in G}\left\{\sum_{j=1}^{n} a_j f(X_j)\right\}$$

implies, for example, the measurability for any $M < \infty$ of the map

$$(x_1,\dots,x_n) \to \sup_{f\in G}\left\{\sum_{j=1}^{n} f(x_j) I_{F(X_j)\le M}\right\}$$

by considering the composition of the map

$$(x_1,\dots,x_n) \to (I(F(x_1)\le M),\dots,I(F(x_n)\le M),x_1,\dots,x_n)$$

with the measurable map given by hypothesis. Actually close consideration of the proofs shows that even weaker hypotheses suffice, but the best measurability is not our concern here. We further note that if $\mathscr{F}$ is countable, or if $\{P_n\}_{n=1}^{\infty}$ are stochastically separable in $\mathscr{F}$, or more generally, if $\mathscr{F}$ is image admissible Suslin [Dudley (1984), page 101], then $\mathscr{F}\in M(P)$.

The following proposition is the first step in the proof of the bootstrap CLT. It is a version of Le Cam's Poissonization lemma [Le Cam (1970); reproduced in Araujo and Giné (1980), Theorem 3.4.8] for expectations.

2.1. LEMMA. *Let B be a separable Banach space and let $\|\cdot\|$ be a measurable pseudonorm on B. For some $n\in\mathbb{N}$, let $\{X_i\}_{i=1}^n$ be independent symmetric B-valued random variables and let $\{\mathscr{L}(X_i)\}_{i=1}^n$ be their laws. Then*

$$E\left\|\sum_{i=1}^{n} X_i\right\| \le 2\int \|x\|\, d\,\mathrm{Pois}\left(\sum_{i=1}^{n}\mathscr{L}(X_i)\right)(x). \tag{2.7}$$

[We recall that for a finite measure ν, $\mathrm{Pois}\,\nu = e^{-\nu(B)}\sum_{n=0}^{\infty}\nu^n/n!$ where $\nu^n = \nu * \cdots * \nu$, that $\mathrm{Pois}\sum\nu_i = (\mathrm{Pois}\,\nu_1)*\cdots*(\mathrm{Pois}\,\nu_n)$, and that if $\nu = \frac{1}{2}(\delta_x + \delta_{-x})$ for some $x\in B$, then $\mathrm{Pois}\,\nu = \mathscr{L}(\tilde{N}x)$ where $\tilde{N} = N - N'$ with N and N' independent Poisson real random variables with expectation $1/2$; we will call $\tilde{N}$ a symmetrized Poisson random variable.] Here is a proof of inequality (2.7): If X_{ij} are independent, $X_{i0} = 0$, $\mathscr{L}(X_{ij}) = \mathscr{L}(X_i)$ for $j > 0$ and N_i are Poisson with parameter 1, independent and independent of $\{X_{ij}\}$, then Fubini's theorem and convexity ($E\|X+Y\|\ge E\|X\|$ if X and Y are independent and $EY = 0$) give

$$\begin{aligned}(1-e^{-1})E\left\|\sum X_i\right\| &\le E\left\|\sum (N_i\wedge 1)X_{i1}\right\| \\ &= E_N\left(E_X\left\|\sum (N_i\wedge 1)X_{i1}\right\|\right) \le E_N\left(E_X\left\|\sum_i\sum_{j=0}^{N_i} X_{ij}\right\|\right) \\ &= E\left\|\sum_i\sum_{j=0}^{N_i} X_{ij}\right\| = \int\|x\|\, d\,\mathrm{Pois}\left(\sum\mathscr{L}(X_i)\right)(x).\end{aligned}$$

2.2. PROPOSITION. *Let B be a Banach space, let $\|\cdot\|$ be a measurable pseudonorm, let $n \in \mathbb{N}$, let $\{x_i\}_{i=1}^n \subset B$, let $\hat{X}_{nj}$, $j = 1, \dots, n$, be i.i.d. B-valued random variables with $\mathscr{L}(\hat{X}_{nj}) = n^{-1}\sum_{i=1}^n \delta_{x_i}$ and let $\{\varepsilon_j\}_{j=1}^n$, $\{\tilde{N}_j\}_{j=1}^n$ be, respectively, a Rademacher sequence and a sequence of independent symmetrized Poisson real random variables with parameter $1/2$, both independent of $\{\hat{X}_{nj}\}$. Then*

$$\frac{1}{\sqrt{2}}(1 - e^{-1})E\left\|\sum_{i=1}^n \varepsilon_i x_i\right\| \le E\left\|\sum_{j=1}^n \varepsilon_j \hat{X}_{nj}\right\| \le 2E\left\|\sum_{i=1}^n \tilde{N}_i x_i\right\|. \tag{2.8}$$

PROOF. We can write

$$\hat{X}_{nj} = \sum_{i=1}^n x_i I_{A_{ij}},$$

where, for each j, the sets $A_{1j}, A_{2j}, \dots, A_{nj}$ are disjoint, the sequences $\{A_{ij}\}_{i=1}^n$, $j = 1, \dots, n$, are independent, and $PA_{ij} = 1/n$, $i, j = 1, \dots, n$. Let $\{\varepsilon_{ij}\}$ be a Rademacher array independent of $\{A_{ij}\}$. Then, by disjointness, the vectors

$$\varepsilon_j\big(x_1 I_{A_{1j}}, \dots, x_n I_{A_{nj}}\big) \quad \text{and} \quad \big(\varepsilon_{1j} x_1 I_{A_{1j}}, \dots, \varepsilon_{nj} x_n I_{A_{nj}}\big), \qquad j = 1, \dots, n,$$

all have the same distribution and, of course, they are independent for different j's. Moreover, by independence of $\{\varepsilon_{ij}\}$ and independence between $\{\varepsilon_{ij}\}$ and $\{A_{ij}\}$, the vector $(\sum_{j=1}^n \varepsilon_{1j} I_{A_{1j}}, \dots, \sum_{j=1}^n \varepsilon_{nj} I_{A_{nj}})$ is symmetric. Let $\{\varepsilon_j'\}$ be a Rademacher sequence independent of $\{\varepsilon_{ij}\}$ and $\{A_{ij}\}$. Then these two observations give

$$\begin{aligned} E\left\|\sum_{j=1}^n \varepsilon_j \hat{X}_{nj}\right\| &= E\left\|\sum_{j=1}^n \varepsilon_j \sum_{i=1}^n x_i I_{A_{ij}}\right\| = E\left\|\sum_{j=1}^n \sum_{i=1}^n \varepsilon_{ij} x_i I_{A_{ij}}\right\| \\ &= E\left\|\sum_{i=1}^n \left(\sum_{j=1}^n \varepsilon_{ij} I_{A_{ij}}\right) x_i\right\| = E\left\|\sum_{i=1}^n \varepsilon_i' \left(\sum_{j=1}^n \varepsilon_{ij} I_{A_{ij}}\right) x_i\right\| \\ &= E\left\|\sum_{i=1}^n \varepsilon_i' \left|\sum_{j=1}^n \varepsilon_{ij} I_{A_{ij}}\right| x_i\right\|. \end{aligned} \tag{2.9}$$

We now notice that by Khintchine's inequality [see Szarek (1976) or Haagerup (1981) for the best constant]

$$\begin{aligned} E\left|\sum_{j=1}^n \varepsilon_{ij} I_{A_{ij}}\right| &\ge \frac{1}{\sqrt{2}} E\left(\sum_{j=1}^n I_{A_{ij}}\right)^{1/2} \ge \frac{1}{\sqrt{2}} P\left\{\sum_{j=1}^n I_{A_{ij}} \ne 0\right\} \\ &= \frac{1}{\sqrt{2}}\left[1 - \left(1 - \frac{1}{n}\right)^n\right] \ge \frac{1}{\sqrt{2}}(1 - e^{-1}). \end{aligned}$$

Hence, by Jensen's inequality and (2.9), and since $E|\sum_{j=1}^n \varepsilon_{ij} I_{A_{ij}}|$ does not depend on i,

$$E\left\|\sum_{j=1}^{n}\varepsilon_j\hat{X}_{nj}\right\| \geq \frac{1}{\sqrt{2}}(1-e^{-1})E\left\|\sum_{i=1}^{n}\varepsilon_i' x_i\right\|,$$

which is the first inequality in (2.8). This proof is essentially taken from Pisier [(1975), proof of Proposition 5.1].

Let $e_1,\dots,e_n$ be the canonical basis of $\mathbb{R}^n$ and for $a=\sum a_i e_i$, let $|||a||| := \|\sum a_i x_i\|$, which is a pseudonorm on $\mathbb{R}^n$. Consider now the random vectors

$$Y_j=\sum_{i=1}^{n}\varepsilon_{ij}I_{A_{ij}}e_i,\qquad j=1,\dots,n,$$

which are independent, symmetric and

$$\mathscr{L}(Y_j)=\frac{1}{2n}\sum_{i=1}^{n}\left(\delta_{e_i}+\delta_{-e_i}\right) \tag{2.10}$$

(i.e., Y_j takes the values $\pm e_i$, $i=1,\dots,n$, each with probability $1/2n$). Then, $\|\sum_{i=1}^{n}(\sum_{j=1}^{n}\varepsilon_{ij}I_{A_{ij}})x_i\| = |||\sum_{j=1}^{n}Y_j|||$. This, (2.9), (2.10) and Lemma 2.1 give

$$E\left\|\sum_{j=1}^{n}\varepsilon_j\hat{X}_{nj}\right\| = E\left|\left|\left|\sum_{j=1}^{n}Y_j\right|\right|\right| \leq 2\int |||x|||\, d\,\mathrm{Pois}\left(\tfrac{1}{2}\sum_{i=1}^{n}\left(\delta_{e_i}+\delta_{-e_i}\right)\right)(x)$$

$$= 2E\left|\left|\left|\sum_{i=1}^{n}\tilde{N}_i e_i\right|\right|\right| = 2E\left\|\sum_{i=1}^{n}\tilde{N}_i x_i\right\|,$$

which is the right-hand side inequality in (2.8). □

What is needed from the result of Ledoux, Talagrand and Zinn is the main part of their proof, namely Lemma 5 in Ledoux and Talagrand (1988b). In the empirical case one needs to complete the proof of tightness in a way different from the original; we incorporate this in the proof of our theorem. First, the lemma:

2.3. LEMMA. *Let $(S,\mathscr{S},P)$ be a probability space, $\mathscr{F}$ an* NLDM(P) *class of functions on S with $E_P F^2<\infty$, $|||\cdot|||$ any of the pseudonorms $\|\cdot\|_{\mathscr{F}}$, $\|\cdot\|_{\mathscr{F}'_\delta}$, $\delta>0$, $X_i\colon S^{\mathbb{N}}\to S$ the coordinate functionals and $\{\xi_i\}$ a sequence of i.i.d. symmetric real random variables with $E\xi_1^2<\infty$, independent of $\{X_i\}$ (actually defined on another probability space). Let E_ξ denote integration with respect to only the variables $\{\xi_i\}$. Then,*

$$a.s.\quad \limsup_n n^{-1/2}E_\xi\left|\left|\left|\sum_{i=1}^{n}\xi_i X_i(\omega)\right|\right|\right| \leq 4\limsup_n n^{-1/2}E\left|\left|\left|\sum_{i=1}^{n}\xi_i X_i\right|\right|\right|. \tag{2.11}$$

The bootstrap CLT is as follows.

2.4. THEOREM. *Let $\mathscr{F} \in M(P)$ and let P be a probability measure on $(S, \mathscr{S})$. Let P_n, $\hat{P}_n(\omega)$, $\omega \in S^{\mathbb{N}}$, and G_P be as defined in* (1.1), (1.2) *and* (2.5). *Then the following are equivalent*:

(a) $\int F^2\, dP < \infty$ *and* $\mathscr{F} \in \mathrm{CLT}(P)$.

(b) *There exists a centered Gaussian process G on $\mathscr{F}$ whose law is Radon in $l^\infty(\mathscr{F})$ such that, $P^{\mathbb{N}}$-a.s., $n^{1/2}(\hat{P}_n(\omega) - P_n(\omega)) \to G$ weakly in $l^\infty(\mathscr{F})$.*

If either (a) *or* (b) *holds, then $G = G_P$.*

PROOF. (a) $\Rightarrow$ (b). Obviously, if N is a Poisson real random variable, then $\int_0^\infty (P\{N > t\})^{1/2}\, dt < \infty$. So, Lemma 1.2.4 in Giné and Zinn (1986) holds for $g_k = \tilde{N}_k$, a sequence of i.i.d. symmetrized Poisson real random variables with parameter $1/2$; hence, their Theorem 1.2.8 [(a) $\Rightarrow$ (e)] gives

$$(2.12)\qquad (\mathscr{F}, e_P) \text{ is totally bounded}$$

and

$$(2.13)\qquad \lim_{\delta \to 0} \limsup_n E \left\| \sum_{i=1}^n \tilde{N}_i \delta_{X_i} / n^{1/2} \right\|_{\mathscr{F}_\delta'} = 0.$$

[Here $\{X_i\}$ is independent of $\{\tilde{N}_i\}$, and is as defined in the introduction, i.e., for $i \in \mathbb{N}$, X_i is the ith coordinate of $(S^{\mathbb{N}}, \mathscr{S}^{\mathbb{N}}, P^{\mathbb{N}})$.] Let E_N denote integration only with respect to $\{\tilde{N}_i\}$. Then, (2.13) and Lemma 2.3 give

$$(2.14)\qquad P^{\mathbb{N}}\text{-a.s.}\qquad \lim_{\delta \to 0} \limsup_n E_N \left\| \sum_{i=1}^n \tilde{N}_i \delta_{X_i(\omega)} / n^{1/2} \right\|_{\mathscr{F}_\delta'} = 0.$$

(2.14) and Proposition 2.2 then give (letting $E_{\varepsilon, A}$ denote integration only with respect to $\{\varepsilon_j\}$ and $\{A_{ij}\}$)

$$(2.15)\qquad P^{\mathbb{N}}\text{-a.s.}\qquad \lim_{\delta \to 0} \limsup_n E_{\varepsilon, A} \left\| \sum_{j=1}^n \varepsilon_j \delta_{\hat{X}_{nj}(\omega)} / n^{1/2} \right\|_{\mathscr{F}_\delta'} = 0$$

and by symmetrization [we will use without further mention that for $\{U_i\}$ independent, independent of $\{\varepsilon_i\}$, $E\|\sum (U_i - EU_i)\| \le 2E\|\sum \varepsilon_i U_i\|$ and $E\|\sum \varepsilon_i (U_i - EU_i)\| \le 2E\|\sum (U_i - EU_i)\|$],

$$(2.16)\qquad P^{\mathbb{N}}\text{-a.s.}\qquad \lim_{\delta \to 0} \limsup_n E_A \|n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))\|_{\mathscr{F}_\delta'} = 0.$$

If $\mathscr{F} \in \mathrm{CLT}(P)$, so does $\mathscr{F}' \in \mathrm{CLT}(P)$. Then, Theorem 1.4.6 in Giné and Zinn (1986) gives $\sup_{f \in \mathscr{F}} |(P_n(\omega) - P)(f^2)| \to 0$ and $\sup_{f, g \in \mathscr{F}} |(P_n(\omega) - P)(f - g)| \to 0$ in probability. Since $\int F^2\, dP < \infty$ these limits hold a.s. [e.g., by a reverse submartingale argument as in Pollard (1981)]. Therefore

$$(2.17)\qquad \sup_{f, g \in \mathscr{F}} |(P_n(\omega) - P)(fg)| \to 0 \quad \text{a.s.}$$

and of course

$$\|P_n(\omega) - P\|_{\mathscr{F}} \to 0 \quad \text{a.s.} \tag{2.18}$$

[We should note here that the proof of Theorem 1.4.6, loc. cit. contains a typographical error (which in the end, is of no consequence for its validity): The relation between entropies should read $N_{n,2}(\varepsilon, \overline{\mathscr{F}}(\lambda)^2) \le N_{n,2}(\varepsilon/2\lambda, \overline{\mathscr{F}}(\lambda))$.] Call the subsets of $\mathscr{S}^{\mathbb{N}}$ where (2.17) and (2.18) hold, respectively, Ω_1 and Ω_2, and let Ω_3 be the intersection for all $\alpha > 0$ rational of the subsets of $\mathscr{S}^{\mathbb{N}}$ for which eventually $\max_{i \le n} F(X_i(\omega)) \le \alpha n^{1/2}$. It follows from the Lindeberg–Feller theorem [as, e.g., in Singh (1981)] that for $\omega \in \Omega_1 \cap \Omega_2 \cap \Omega_3, n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))(\Sigma_{\text{finite}} a_i f_i) \to \Sigma a_i G_P(f_i)$ weakly, for all $\{a_i\} \subset \mathbb{R}$, $\{f_i\} \subset \mathscr{F}$. Thus, (2.16) and (2.12) imply the bootstrap CLT (b) with $G = G_P$ by, e.g., Theorem 1.1.3 in Giné and Zinn (1986) (which, although given for the i.i.d. sequence case, holds, with the same proof, for triangular arrays as well).

(b) $\Rightarrow$ (a). We show first that if (b) holds then $\int F^2 dP < \infty$. Note that the convergence in (b) is actually weak convergence of Radon measures (for each ω for which there is convergence) and therefore the CLT theory for separable Banach spaces applies. The system $\{Y_{nj}(\omega) = n^{-1/2}\delta_{\hat{X}^{\omega}_{nj}}\}$ is infinitesimal ω-a.s.: $P^{\mathbb{N}}$-a.s., for all $\varepsilon > 0$, $P_n^n\{\|f(\hat{X}^{\omega}_{n1})\|_{\mathscr{F}} > \varepsilon n^{1/2}\} = \Sigma_{i=1}^n I(F(X_i(\omega)) > \varepsilon n^{1/2})/n \to 0$ by the law of large numbers (by monotonicity, it is enough to consider rational $\varepsilon > 0$). Hence, since ω-a.s. the sequence $\{\Sigma_{j=1}^n Y_{nj}(\omega)\}$ is shift convergent in law to a Gaussian limit, it follows from, e.g., Araujo and Giné (1980), Theorem 3.5.4 that

$$nP_n\{\|f(\hat{X}^{\varepsilon}_{n1})\|_{\mathscr{F}} > n^{1/2}\} \to 0 \quad \text{a.s.,}$$

that is,

$$\sum_{i=1}^{n} I\big(F(X_i(\omega)) > n^{1/2}\big) \to 0 \quad \text{a.s.} \tag{2.19}$$

Since if $\Sigma_{i=1}^m I(F(X_i(\omega)) > n^{1/2}) < 1$ then $\Sigma_{i=1}^m I(F(X_i(\omega)) > n^{1/2}) = 0$, (2.19) implies that ω-a.s. there is $n(\omega) < \infty$ such that for $n > n(\omega)$,

$$F(X_n(\omega))/n^{1/2} \le \max_{i \le n} F(X_i(\omega))/n^{1/2} \le 1.$$

This and the Borel–Cantelli lemma give $\Sigma P\{F(X_n) > n^{1/2}\} < \infty$, that is,

$$EF^2(X_1) < \infty. \tag{2.20}$$

Let $f \in \mathscr{F}' \cup \mathscr{F}$. Then by hypothesis $\mathscr{L}(n^{1/2}(\hat{P}_n f - P_n f)) \to_w \mathscr{L}(G(f))$ and by the converse CLT in $\mathbb{R}$ for triangular arrays, together with (2.19), we have

$$\lim_{n \to \infty} \left(\sum_{i=1}^{n} f(X_i)^2/n - \left(\sum_{i=1}^{n} f(X_i)/n \right)^2 \right) = E(G(f))^2 \quad \text{a.s.}$$

But, by (2.20) and the law of large numbers, this limit is $E(f(X_1))^2 - (Ef(X_1))^2$. We have thus shown

$$G = G_P. \tag{2.21}$$

Moreover, since G, hence G_P, has a Radon law and since (2.20) holds, we also have that $(\mathscr{F}, e_P)$ is totally bounded.

Next we prove $P^{\mathbb{N}}$-a.s. uniform integrability of $\{\|n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))\|_{\mathscr{F}}\}_{n=1}^{\infty}$. By Theorem 3.2 in de Acosta and Giné (1979) it is enough to show

$$\sup_n E_A \max_{j \le n} \|\delta_{\hat{X}^{\omega}_{nj}} - P_n(\omega)\|^2_{\mathscr{F}}/n < \infty \quad \text{a.s.,} \tag{2.22}$$

where E_A denotes integration with respect to $\{I_{A_{ij}}\}$. But the random variable in (2.22) is bounded by

$$\begin{aligned} \sup_n E_A \|\delta_{\hat{X}^{\omega}_{n1}} - P_n(\omega)\|^2_{\mathscr{F}} &= \sup_n \frac{1}{n} \sum_{i=1}^n \|\delta_{X_i(\omega)} - P_n(\omega)\|^2_{\mathscr{F}} \\ &\le 4 \sup_n \frac{1}{n} \sum_{i=1}^n F^2(X_i(\omega)) < \infty \quad \text{a.s.} \end{aligned}$$

(by the law of large numbers, since $\int F^2\, dP < \infty$). We thus have, by uniform integrability,

$$P^{\mathbb{N}}\text{-a.s.,} \quad \begin{cases} E_A\left\| n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))\right\|_{\mathscr{F}} \to E\|G_P\|_{\mathscr{F}} \\ E_A\left\| n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))\right\|_{\mathscr{F}'_\delta} \to E\|G_P\|_{\mathscr{F}'_\delta} \end{cases} \quad \text{for all } \delta > 0. \tag{2.23}$$

Denote by $|||\cdot|||$ any of the pseudonorms $\|\cdot\|_{\mathscr{F}'_\delta}$, $\delta > 0$, or $\|\cdot\|_{\mathscr{F}}$. By Proposition 2.2 we have, with $c = (1 + e^{-1})/\sqrt{2}$,

$$\begin{aligned} P^{\mathbb{N}}\text{-a.s.,} \quad cE_\varepsilon\left|\left|\left| \sum_{i=1}^n \varepsilon_i \delta_{X_i(\omega)}/n^{1/2} \right|\right|\right| &\le E_{\varepsilon, A}\left|\left|\left| \sum_{j=1}^n \varepsilon_j \delta_{\hat{X}^{\omega}_{nj}}/n^{1/2} \right|\right|\right| \\ &\le E_{\varepsilon, A}\left|\left|\left| \sum_{j=1}^n \varepsilon_j \left(\delta_{\hat{X}^{\omega}_{nj}} - P_n(\omega)\right)/n^{1/2} \right|\right|\right| \\ &\quad + \left(E\left| \sum_{i=1}^n \varepsilon_j/n^{1/2}\right|\right) |||P_n(\omega)||| \\ &\le 2E_A|||n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))||| + |||P_n(\omega)|||. \end{aligned} \tag{2.24}$$

(2.23) and (2.24) give

$$\begin{aligned} \limsup_{n\to\infty} \Pr\left\{\left|\left|\left| \sum_{i=1}^n \frac{\varepsilon_i \delta_{X_i}}{n^{1/2}} \right|\right|\right| \ge M\right\} &\le \frac{1}{M} \limsup_{n\to\infty} E\left(E_\varepsilon \left|\left|\left| \sum_{i=1}^n \frac{\varepsilon_i \delta_{X_i}}{n^{1/2}} \right|\right|\right| \wedge M\right) \\ &\le \frac{1}{M} E \limsup_{n\to\infty} E_\varepsilon \left|\left|\left| \sum_{i=1}^n \frac{\varepsilon_i \delta_{X_i}}{n^{1/2}} \right|\right|\right| \wedge M \\ &\le \frac{E\,|||\delta_{X_1}||| + 2E\,|||G_P|||}{cM} \to 0 \quad \text{as } M \to \infty. \end{aligned}$$

The above inequality, by Hoffmann-Jørgensen's inequality and $EF^2(X_1) < \infty$, implies

$$\sup_n E\left\|\left\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n^{1/2}\right\|\right\| < \infty. \tag{2.25}$$

In particular $E\|\Sigma_{i=1}^n \varepsilon_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$, hence $E\|P_n - P\| \to 0$, or

$$|||P_n - P||| \to 0 \quad \text{a.s.}$$

[cf. Pollard (1981)]. Hence,

$$\lim_{\delta\to 0}\lim_{n\to\infty}\|P_n(\omega)\|_{\mathscr{F}'_\delta} = \lim_{\delta\to 0}\|Ef(X_1)\|_{\mathscr{F}'_\delta} \le \lim_{\delta\to 0}\|Ef^2(X_1)\|_{\mathscr{F}'_\delta}^{1/2} = 0. \tag{2.26}$$

Using (2.26) in (2.24) we obtain that, $P^{\mathbb{N}}$-a.s.,

$$\lim_{\delta\to 0}\limsup_n E_\varepsilon\left\|\sum_{i=1}^n \varepsilon_i \delta_{X_i(\omega)}/n^{1/2}\right\|_{\mathscr{F}'_\delta} \le (2+c')c^{-1}\lim_{\delta\to 0} E\|G_P\|_{\mathscr{F}'_\delta} = 0. \tag{2.27}$$

Bounded convergence and Fatou's lemma then give

$$\lim_{\delta\to 0}\limsup_n E\left(\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n^{1/2}\|_{\mathscr{F}'_\delta} \wedge M\right) = 0$$

for all $M > 0$, which, by Theorem 1.2.8 in Giné and Zinn (1986), implies that $\mathscr{F} \in \mathrm{CLT}(P)$. □

2.5. Remark. A corollary of Theorem 2.4 is that if X_i are i.i.d. B-valued random variables, B a separable Banach space, then

$$E\|X_1\|^2 < \infty \quad \text{and} \quad X_1 \in \mathrm{CLT} \Leftrightarrow \sum_{j=1}^n \left(\hat{X}_{nj} - \bar{X}_n\right)/n^{1/2} \to G_X \quad \text{weakly a.s.}$$

Actually the proof of this result is somewhat simpler than that of Theorem 2.4 since in this case $E\|X_1\| < \infty$ already implies $|||P_n - P||| \to 0$ a.s. [see the material following (2.25)].

The law of large numbers has a proof similar to that of Theorem 2.4 but simpler since, in this case, the lemma of Ledoux, Talagrand and Zinn is not needed and some further simplifications are also possible.

2.6. Theorem. *Let $\mathscr{F}$ be* NLDM(P). *Then the following are equivalent*:

(a) *$\int F\,dP < \infty$ and $\|P_n - P\|_{\mathscr{F}} \to 0$ in probability.*
(b) *$P^{\mathbb{N}}$-a.s., $\|\hat{P}_n(\omega) - P_n(\omega)\|_{\mathscr{F}} \to 0$ in probability.*

Proof. (Sketch). $\int F\,dP < \infty$, $\|P_n - P\|_{\mathscr{F}} \to 0$ pr. $\Rightarrow \|P_n - P\|_{\mathscr{F}} \to 0$ a.s. [e.g., Pollard (1981)] $\Rightarrow \|\Sigma_{i=1}^n \varepsilon_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$ a.s. [Giné and Zinn (1984), page 980] $\Rightarrow E\|\Sigma_{i=1}^n \tilde{N}_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$ [as noted in Giné, Marcus and Zinn (1990), by a proof similar to that in Lemma 2.9 of Giné and Zinn (1984), since $E|\tilde{N}| < \infty$] $\Rightarrow \|\Sigma_{i=1}^n \tilde{N}_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$ a.s. [by, e.g., a reverse martingale argument as in Pollard (1981)] $\Rightarrow P^{\mathbb{N}}$-a.s. $\|\Sigma_{i=1}^n \tilde{N}_i \delta_{X_i(\omega)}/n\|_{\mathscr{F}} \to 0$ a.s. (Fubini) $\Rightarrow P^{\mathbb{N}}$-a.s. $E_N\|\Sigma_{i=1}^n \tilde{N}_i \delta_{X_i(\omega)}/n\|_{\mathscr{F}} \to 0$. (To see this we use Hoffman-

Jørgensen's inequality [Hoffman-Jørgensen (1974)] to reduce to showing $E_N \max_{i \le n} |\tilde{N}_i| \, \|\delta_{X_i(\omega)}\|_{\mathscr{F}}/n \to 0$, $P^{\mathbb{N}}$-a.s. But, for any $c > 0$

$$E_N \max_{i \le n} |\tilde{N}_i| \frac{\|\delta_{X_i(\omega)}/n\|_{\mathscr{F}}}{n} \le c \max_{i \le n} \frac{\|\delta_{X_i(\omega)}\|_{\mathscr{F}}}{n} + E_N \sum_{i=1}^{n} |\tilde{N}_i| I_{|\tilde{N}_i| > c} \frac{\|\delta_{X_i(\omega)}\|_{\mathscr{F}}}{n}.$$

The first term goes to zero since $F \in L_1$ and the second equals $(\sum_{i=1}^{n} \|\delta_{X_i(\omega)}\|_{\mathscr{F}}/n) E|\tilde{N}| I_{|\tilde{N}| > \varepsilon}$. But the first term in this last quantity is P^N-a.s. bounded by the strong law of large numbers and the fact that $F \in L_1$. The second can be made arbitrarily small by taking c large.) $\Rightarrow P^{\mathbb{N}}$-a.s., $E_{\varepsilon, A} \|\sum_{j=1}^{n} \varepsilon_j \delta_{\hat{X}^{\omega}_{nj}}/n\|_{\mathscr{F}} \to 0$ (Proposition 2.2) $\Rightarrow P^{\mathbb{N}}$-a.s., $E_A \|\hat{P}_n(\omega) - P_n(\omega)\|_{\mathscr{F}} \to 0$ (desymmetrization).

For the converse, observe first that, as in Theorem 2.4,

$$\|\hat{P}_n(\omega) - P_n(\omega)\|_{\mathscr{F}} \to 0 \ \omega\text{-a.s.} \Rightarrow \int F \, dP < \infty \quad \text{and}$$

$$E_A \|\hat{P}_n(\omega) - P_n(\omega)\|_{\mathscr{F}} \to 0 \ \omega\text{-a.s.}$$

But, by symmetrization, as in (2.24),

$$E_{\varepsilon, A} \Big\| \sum_{j=1}^{n} \varepsilon_j \delta_{\hat{X}_{nj}}/n \Big\|_{\mathscr{F}} \le 2 E_A \|\hat{P}_n(\omega) - P_n(\omega)\|_{\mathscr{F}} + \left(E \left| \sum_{j=1}^{n} \varepsilon_j / n \right| \right) \|P_n(\omega)\|_{\mathscr{F}}$$

and these two variables tend to zero a.s. (note that, since $\int F \, dP < \infty$, $\|P_n(\omega)\|_{\mathscr{F}}$ is a.s. bounded). Hence Proposition 2.2 implies $E_\varepsilon \|\sum_{i=1}^{n} \varepsilon_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$ a.s. So for all $M > 0$, $E(\|\sum_{i=1}^{n} \varepsilon_i \delta_{X_i}/n\|_{\mathscr{F}} \wedge M) \to 0$, i.e., $\|\sum_{i=1}^{n} \varepsilon_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$ in pr., which, since $\int F \, dP < \infty$, implies $\|P_n - P\|_{\mathscr{F}} \to 0$ a.s. [Giné and Zinn (1984), page 980]. □

3. The bootstrapped (in probability) limit theorems. We first give the appropriate notion of bootstrap in probability in the context of empirical processes and show how it can be used.

In Giné and Zinn [(1986), theorem 1.1.3], we give a natural and short proof of: $\mathscr{F} \in \mathrm{CLT}(P)$ iff $(\mathscr{F}, \rho_P)$ is totally bounded and the usual eventual equicontinuity condition holds. This proof actually shows that $\mathscr{F} \in \mathrm{CLT}(P)$ iff $\mathscr{F}$ is P-pre-Gaussian and

$$\sup_{H \in BL_1(\ell^\infty(\mathscr{F}))} |E^* H(n^{1/2}(P_n - P)) - EH(G_P)| \to 0, \tag{3.1}$$

where $BL_1(\ell^\infty(\mathscr{F})) = \{H: \ell^\infty(\mathscr{F}) \to \mathbb{R}, |H(x) - H(y)| \le \|x - y\|_{\mathscr{F}}, \|H\|_\infty \le 1\}$. With some abuse of notation, we may call the quantity in (3.1),

$$d_{BL^*}\big(\mathscr{L}(n^{1/2}(P_n - P)), \mathscr{L}(G_P)\big)$$

as in the case when these are true probability laws [$n^{1/2}(P_n - P)$ may not be measurable as an $\ell^\infty(\mathscr{F})$-valued random element]. The above observation extends also to more general limit theorems (e.g., non-i.i.d., different normings). In particular $n^{1/2}(\hat{P}_n(\omega) - P_n(\omega)) \to G_P$ weakly in $\ell^\infty(\mathscr{F})$, ω-a.s. iff

$$d_{BL^*}\big(\mathscr{L}(n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))), \mathscr{L}(G_P)\big) \to 0 \quad \text{a.s.} \tag{3.2}$$

So, it is justifiable to say that the bootstrapped CLT(P) holds *in probability* iff the limit (3.2) takes place in outer probability.

To see the usefulness of this notion, suppose that $\|P_n - P\|_{\mathscr{F}}$ is measurable, that $\|G_P\|_{\mathscr{F}}$ has a continuous distribution and that $\mathscr{F}$ satisfies both the CLT(P) and the bootstrapped CLT(P) in probability. Since $\overline{H} = H \circ \|\cdot\|_{\mathscr{F}} \in BL_1(\ell^\infty(\mathscr{F}))$ if $H \in BL_1(\mathbb{R})$, we have

$$(3.3)\qquad d_{BL^*}\big(\mathscr{L}(n^{1/2}\|\hat{P}_n(\omega) - P_n(\omega)\|_{\mathscr{F}}), \mathscr{L}(\|G_P\|_{\mathscr{F}})\big) \to 0 \quad \text{in pr.}$$

By passing back and forth to a.s. convergent subsequences, since d_{BL^*} metrizes weak convergence in $\mathbb{R}$, we get from (3.3) that

$$(3.4)\qquad \sup_{x\in\mathbb{R}} |F_{n^{1/2}\|\hat{P}_n(\omega)-P_n(\omega)\|_{\mathscr{F}}}(x) - F_{\|G_P\|_{\mathscr{F}}}(x)| \to 0 \quad \text{in pr}$$

(where $\mathscr{F}_\xi$ denotes the distribution function of the real random variable ξ). By the assumptions, we also have

$$(3.5)\qquad \sup_{x\in\mathbb{R}} |F_{n^{1/2}\|P_n - P\|_{\mathscr{F}}}(x) - F_{\|G_P\|_{\mathscr{F}}}(x)| \to 0.$$

So, if $c_n(\alpha) = c_n(\alpha, \omega)$ is defined by

$$c_n(\alpha) = \inf\{t: F_{n^{1/2}\|\hat{P}_n(\omega)-P_n(\omega)\|_{\mathscr{F}}}(t) \ge 1 - \alpha\},$$

then (3.4) and (3.5) give

$$(3.6)\qquad F_{n^{1/2}\|P_n - P\|_{\mathscr{F}}}(c_n(\alpha)) \to 1 - \alpha \quad \text{in pr}$$

and also

$$(3.7)\qquad \Pr\{n^{1/2}\|P_n - P\|_{\mathscr{F}} \le c_n(\alpha)\} \to 1 - \alpha.$$

In conclusion the bootstrap in probability as described above allows the construction of asymptotic confidence regions for P.

3.1. THEOREM. *Assuming $\mathscr{F} \in M(P)$, the following are equivalent*:

(a) $\mathscr{F} \in$ CLT(P).

(b) *There exists a centered Gaussian process G on $\mathscr{F}$ whose law is Radon in $\ell^\infty(\mathscr{F})$ such that*

$$(3.8)\qquad d_{BL_1^*}\big(\mathscr{L}(n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))), \mathscr{L}(G)\big) \to 0 \quad \textit{in pr}^*.$$

If either (a) *or* (b) *holds, then $G = G_P$, i.e., $\mathscr{F}$ satisfies the bootstrapped* CLT(P) *in probability.*

PROOF. (a) ⇒ (b). Using the decomposition (1.13) in Theorem 1.1.3 from Giné and Zinn (1986) of

$$E^*H(n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))) - EH(G_P), \qquad H \in BL_1(\ell^\infty(\mathscr{F})),$$

and the bootstrapped CLT in finite dimensions, it follows that, in order to establish (3.8) it suffices to prove that

$$\lim_{\delta\to 0}\limsup_n \Pr{}^*\{E_A\|n^{1/2}(\hat{P}_n(\omega)-P_n(\omega))\|_{\mathscr{F}'_\delta}>\varepsilon\}=0 \quad \text{for all } \varepsilon>0. \tag{3.9}$$

Symmetrization and Proposition 2.2 give

$$E_A\left\|n^{1/2}(\hat{P}_n(\omega)-P_n(\omega))\right\|_{\mathscr{F}'_\delta}\le 2E_{\varepsilon,A}\left\|\sum_{j=1}^n \varepsilon_j\delta_{\hat{X}^\omega_{nj}}/n^{1/2}\right\|_{\mathscr{F}'_\delta}$$

$$\le 4E_N\left\|\sum_{i=1}^n \tilde{N}_i\delta_{X_i(\omega)}/n^{1/2}\right\|_{\mathscr{F}'_\delta}.$$

Now, by the multiplier Lemma 1.2.4 and Theorem 1.1.8 in Giné and Zinn (1986), the above inequality yields

$$\lim_{\delta\to 0}\limsup_n EE_A\|n^{1/2}(\hat{P}_n(\omega)-P_n(\omega))\|_{\mathscr{F}'_\delta}$$

$$\le 4\|\tilde{N}\|_{2,1}\lim_{\delta\to 0}\limsup_n E\left\|\sum_{i=1}^n \varepsilon_i\delta_{X_i}/n^{1/2}\right\|_{\mathscr{F}'_\delta}=0.$$

This gives (3.9), hence (3.8) with $G=G_P$.

(b) $\Rightarrow$ (a). If (b) holds, for every subsequence of $\mathbb{N}$ there is a further subsequence, say $\{n_k\}$ such that

$$d_{BL^*}\Big(\mathscr{L}\big(n_k^{1/2}(\hat{P}_{n_k}(\omega)-P_n(\omega))\big),\mathscr{L}(G)\Big)\to 0 \quad \omega\text{-a.s.} \tag{3.10}$$

Then, by infinitesimality and Gaussian limits, we have, as in the proof of Theorem 2.4, for all $\delta>0$,

$$\sum_{i=1}^{n_k} I\big(F(X_i(\omega))>\delta n_k^{1/2}\big)\to 0 \quad \text{a.s.} \tag{3.11}$$

($=0$ eventually a.s.). This implies

$$\sum_{i=1}^{n} I\big(F(X_i(\omega))>\delta n^{1/2}\big)\to 0 \quad \text{in pr.}$$

Now, previous arguments show that this limit holds in expectation, i.e.,

$$n\Pr\{F(X)>\delta n^{1/2}\}\to 0. \tag{3.12}$$

For every subsequence $\{n_k\}$ for which (3.10) holds, we can use (3.11) and the converse CLT in $\mathbb{R}$ to obtain, as in the proof of Theorem 2.4,

$$\lim_{n_k\to 0}\left(\frac{\sum_{i=1}^{n_k}f(X_i)^2}{n_k}-\left(\frac{\sum_{i=1}^{n_k}f(X_i)}{n_k}\right)^2\right)=E(G(f))^2 \quad \text{a.s.}$$

for all $f\in\mathscr{F}'\cup\mathscr{F}$. Hence this limit holds for the whole sequence $\mathbb{N}$ in probability. If $Ef^2(X)<\infty$ the limit is actually $E(G_P(f))^2$ by the law of large numbers. If $Ef^2(X)=\infty$ then, by Lemma 2 in Giné and Zinn (1989), the empirical second moment dominates the square of the empirical first (absolute) moment, and we get

$$\lim_{n\to\infty} \sum_{i=1}^{n} f^2(X_i)/n = E(G(f))^2 \quad \text{in pr.}$$

Then, by the converse CLT (centering part), the truncated centers must converge, i.e., $Ef^2(X_1)I(|f(X_1)| \le \sqrt{n})$ converges, implying $Ef^2(X) < \infty$, a contradiction. We have just proved $Ef^2(X) < \infty$, $f \in \mathscr{F}$, and

$$G = G_P. \tag{3.13}$$

Consider now a subsequence $\{n_k\}$ for which (3.10) holds. Then, for any $p > 0$ and $a > 0$,

$$\begin{aligned} E_A \max_{j\le n_k}\Big(\|\delta_{\hat{X}^\omega_{nj}} - P_{n_k}(\omega)\|_{\mathscr{F}}/n_k^{1/2}\Big)^p &\le 2^p \max_{j\le n_k}\big(F(X_i(\omega))/n_k^{1/2}\big)^p \\ &\le 2^p\left[a + \sum_{i=1}^{n_k}\big[F(X_i(\omega))/n_k^{1/2}\big] I\big(F(X_i(\omega)) > a n_k^{1/2}\big)\right]^p \end{aligned}$$

and by (3.11) this last quantity is eventually $(2a)^p$ a.s. Hence

$$\sup_k E_A \max_{j\le n_k}\|\delta_{\hat{X}^\omega_{n_kj}} - P_{n_k}(\omega)\|^p_{\mathscr{F}}/n_k^{p/2} < \infty \quad \text{a.s.}$$

This allows us to follow for $\{n_k\}$ exactly the same steps as in the proof of (b) ⇒ (a) in Theorem 2.4, from inequality (2.22) on, to conclude that

$$d_{BL_1^*}\left(\mathscr{L}\left(\sum_{i=1}^{n_k}\varepsilon_i\delta_{X_i}/n_k^{1/2}\right),\ \mathscr{L}(Z_P)\right) \to 0.$$

Hence, since every subsequence has a further subsequence $\{n_k\}$ for which this limit holds, we obtain

$$d_{BL_1^*}\left(\mathscr{L}\left(\sum_{i=1}^{n}\varepsilon_i\delta_{X_i}/n^{1/2}\right),\ \mathscr{L}(Z_P)\right) \to 0,$$

i.e., $\mathscr{F} \in \mathrm{CLT}(P)$. □

3.2. Remark. A similar result holds in the case of normings $a_n \ne n^{1/2}$ and Gaussian limits: $\mathscr{F} \in \mathrm{CLT}(P; a_n)$ with limit G iff $\mathscr{L}\{\sum_{j=1}^n (f(\hat{X}^\omega_{nj}) - P_n(\omega)(f))/a_n\colon f \in \mathscr{F}\} \to_w \mathscr{L}(G)$ in probability. The proof is analogous to that of Theorem 3.1 and is omitted. However, such a result cannot hold in the case of a stable non-Gaussian limit [Giné and Zinn (1989)].

3.3. Remark. Note that the proof of Theorem 3.1 is more elementary than the proof of Theorem 2.4: The deeper Lemma 2.3 is not needed for the bootstrap in probability.

3.4. REMARK. Beran, Le Cam and Millar (1987) show that whenever a bootstrapped limit theorem holds in probability, then the empirical distributions of the bootstrapped laws also converge weakly in probability. This justifies using Monte Carlo simulation to approximate the bootstrapped distributions. Concretely, our Theorem 3.1 and the corollary in Section 4 of their paper give:

Let $\hat{\nu}_n^\omega = n^{1/2}(\hat{P}_n(\omega) - P_n(\omega))$, which is an $\ell^\infty(\mathscr{F})$-valued random variable, and for $j_n \to \infty$ consider i.i.d. copies of $\hat{\nu}_n^\omega$, say $\{\hat{\nu}_{n,j}^\omega\}_{j=1}^{j_n}$. Then, if $\mathscr{F} \in \mathrm{CLT}(P)$, we have

$$d_{BL^*}\left(\frac{1}{j_n}\sum_{j=1}^{j_n}\delta_{\hat{\nu}_{n,j}^\omega}, \mathscr{L}(G_P)\right) \to 0 \quad \text{in probability}$$

[in probability refers to $(\mathscr{L}(\hat{\nu}_n^\omega))^{j_n} \otimes \Pr$, for each n].

Finally we show that the weak law of large numbers for empirical processes, can also be bootstrapped in probability. It may be worth mentioning that an example of $\mathscr{F}$ and P for which the WLLN holds but the strong law does not hold is: $P =$ uniform distribution on $[0,1]$ $\mathscr{F} = \{w(t)I_{(0,1]}: t \in (0, 1/2]\}$ with w decreasing $tw(t) \to 0$ but $\int_0^{1/2} w(t)\,dt = \infty$, i.e., the weighted empirical process [Theorem 7.3 in Andersen, Giné and Zinn (1988)]. Some additional notation for Theorem 3.5: Given random variables ξ, η, d_{pr} denotes their Ky Fan distance, which metrizes convergence in probability, $d_{\mathrm{pr}}(\xi, \eta) = \inf[\varepsilon: \Pr\{|\xi - \eta| > \varepsilon\} < \varepsilon]$. If the random variables involve $\hat{X}_{nj}^\omega, \varepsilon_j, N_j$, then d_{pr_A}, $d_{\mathrm{pr}_{\varepsilon,A}}$ and d_{pr_N} indicate that the distance d_{pr} is taken with respect to the conditional probability given $X_1(\omega), \dots, X_n(\omega)$.

3.5. THEOREM. *Let* $\mathscr{F}$ *be* NLSM(P). *The following are equivalent*:

(i) $\|\sum_{i=1}^n (f(X_i) - PfI(F \le n))/n\|_{\mathscr{F}} \to 0$ *in pr.*
(ii) $d_{\mathrm{pr}_A}(\|\sum_{j=1}^n (\delta_{\hat{X}_{nj}^\omega} - P_n(\omega))/n\|_{\mathscr{F}}, 0) \to 0$ *in pr.*

and if (i) *or* (ii) *holds then also*

$$E_A\left\|\sum_{j=1}^n \left(\delta_{\hat{X}_{nj}^\omega} - P_n(\omega)\right)/n\right\|_{\mathscr{F}} \to 0 \quad \textit{in pr.}$$

PROOF. (i) $\Rightarrow$ (ii). (a) We first show (i) $\Rightarrow$ $\|\sum_{i=1}^n \varepsilon_i \delta_{X_i}/n\|_{\mathscr{F}} \to 0$ in probability. To this end we note that

$$\left\|\sum_{i=1}^n (f(X_i) - PfI(F \le n))/n\right\|_{\mathscr{F}} \to 0 \quad \text{in pr.}$$

implies

$$\left\|\sum_{i=1}^n (f(X_i) - f(X_i'))/n\right\|_{\mathscr{F}} \to 0 \quad \text{in pr.}$$

by the triangle inequality for $\|\cdot\|_{\mathscr{F}}$, where $\{X_i, X_j'\}_{i,j=1}^\infty$ are i.i.d., and this implies [see the proof of Corollary 2.13 in Giné and Zinn (1984)] that

$$n \Pr\big(\|\delta_{X_1} - \delta_{X_1'}\|_{\mathscr{F}} > n\big) \to 0 \quad \text{as } n \to \infty.$$

But then

$$n \Pr\big(\|\delta_{X_1}\|_{\mathscr{F}} > 2n\big)\Pr\big(\|\delta_{X_1'}\|_{\mathscr{F}} \le n\big) \to 0 \quad \text{as } n \to \infty$$

and hence

$$t \Pr\big(\|\delta_{X_1}\|_{\mathscr{F}} > t\big) \to 0 \quad \text{as } t \to \infty.$$

Also, from symmetrization procedures [Lemma 2.7, Giné and Zinn (1984)] we know

$$\Pr\left\{\left\|\sum_{i=1}^{n} \varepsilon_i\big(f(X_i) - PfI(F \le n)\big)\right\|_{\mathscr{F}} > \varepsilon n\right\}$$

$$\le 2\max_{k \le r}\Pr\left\{\left\|\sum_{i=1}^{k} \big(f(X_i) - PfI(F \le n)\big)\right\|_{\mathscr{F}} > \frac{\varepsilon n}{2}\right\}$$

$$+ 2 \max_{r<k\le n} \Pr\left\{\left\|\sum_{i=1}^{k} \big(f(X_i) - PfI(F \le n)\big)\right\|_{\mathscr{F}} > \frac{\varepsilon n}{2}\right\}.$$

The first term on the right goes to zero since $n \to \infty$. The second term can be made less than any $\varepsilon > 0$ if r (and therefore k) is large enough, since the WLLN [i.e., (i)] is assumed to hold. Further, since $t\Pr(\|\delta_{X_1}\|_{\mathscr{F}} > t) \to 0$ as $t \to \infty$,

$$|PfI(F \le n)| \le \int_0^\infty \Pr\big(|f(x)|I(F(X) \le n) > t\big)\,dt \le 1 + \int_1^n \frac{K}{t}\,dt \le K' \ln n,$$

where K and K' are fixed constants. But then

$$\left|\sum_{i=1}^{n} \frac{\varepsilon_i PfI(F \le n)}{n}\right| \le K' \frac{|\sum_{i=1}^{n}\varepsilon_i|}{(n/\ln n)},$$

which converges to zero a.s. by, e.g., the Marcinkiewicz–Zygmund SLLN. Hence, for all $\varepsilon > 0$, $\Pr(\|\sum_{i=1}^{n}\varepsilon_i f(X_i)\|_{\mathscr{F}} > \varepsilon n) \to 0$ as $n \to \infty$.

(b) Le Cam's Poissonization lemma [Le Cam (1970); see also Araujo and Giné (1980) Lemma 3.4.8] in probability gives

$$d_{\mathrm{pr}_{\varepsilon,A}}\left(\left\|\sum_{j=1}^{n} \varepsilon_j \delta_{\hat{X}_{nj}^{\omega}}/n\right\|_{\mathscr{F}}, 0\right) \le 2d_{\mathrm{pr}_N}\left(\left\|\sum_{i=1}^{n} \tilde{N}_i \delta_{X_i(\omega)}/n\right\|_{\mathscr{F}}, 0\right).$$

(c) If $\|\sum_{i=1}^{n}\varepsilon_1\delta_{X_i}/n\|_{\mathscr{F}} \to 0$ pr. then, as pointed out in Giné, Marcus and Zinn (1990), Remark 4.2, $\|\sum_{i=1}^{n}\tilde{N}_i\delta_{X_i}/n\|_{\mathscr{F}} \to 0$ in pr. because $E\tilde{N}^{1+\delta} < \infty$. Hence, by (a), for all $\varepsilon > 0$

$$E_X \Pr_N\left\{\left\|\sum_{i=1}^{n} \tilde{N}_i \delta_{X_i}(\omega)/n\right\|_{\mathscr{F}} > \varepsilon\right\} \to 0. \tag{3.14}$$

But, $d_{\mathrm{pr}_N}(\|\Sigma_{i=1}^n \tilde{N}_i \delta_{X_i(\omega)}/n\|_{\mathscr{F}}, 0) \le \varepsilon \vee \Pr_N\{\|\Sigma_{i=1}^n \tilde{N}_i \delta_{X_i(\omega)}/n\|_{\mathscr{F}} > \varepsilon\}$, for all $\varepsilon > 0$ by definition of the Ky Fan distance. Therefore (3.14) implies

$$E_X d_{\mathrm{pr}_N}\left(\left\|\sum_{i=1}^n \tilde{N}_i \delta_{X_i(\omega)}/n\right\|_{\mathscr{F}}, 0\right) \to 0.$$

Now, (b) and (c) give

$$E_X d_{\mathrm{pr}_{\varepsilon,A}}\left(\left\|\sum_{i=1}^n \varepsilon_j \delta_{\hat{X}^\omega_{nj}}/n\right\|_{\mathscr{F}}, 0\right) \to 0. \tag{3.15}$$

(d) Now we must desymmetrize in (3.15). For every subsequence of $\mathbb{N}$, there exists a further subsequence $\{n_k\}$ such that

$$\left\|\sum_{i=1}^{n_k} \varepsilon_j \delta_{\hat{X}^\omega_{n_kj}}/n_k\right\|_{\mathscr{F}} \to 0 \quad \text{in } \mathrm{pr}_{\varepsilon, A}\ \omega\text{-a.s.}$$

Hence $\Sigma_{i=1}^{n_k} I(\|X_i\| > an_k) = 0$ eventually a.s., for all $a > 0$. Therefore,

$$E_A \max_{j\le n_k} \frac{\|\delta_{\hat{X}^\omega_{nj}}\|_{\mathscr{F}}}{n_k} \le a + \int_a^\infty \sum_1^{n_k} I(\|X_i\| > n_k t)\,dt \le a \quad \text{eventually a.s.}$$

$$\Rightarrow E_A \max_{j\le n_k} \|\delta_{\hat{X}^\omega_{n_kj}}\|_{\mathscr{F}}/n_k \to 0 \quad \omega\text{-a.s.}$$

$$\Rightarrow E_{\varepsilon,A}\left\|\sum_{j=1}^{n_k} \varepsilon_j \delta_{X^\omega_{n_kj}}/n_k\right\|_{\mathscr{F}} \to 0 \quad \omega\text{-a.s.}$$

[by (3.15) and Hoffman-Jørgensen's inequality]

$$\Rightarrow E_A\left\|\sum_{j=1}^{n_k} \left(\delta_{\hat{X}^\omega_{n_kj}} - P_{n_k}(\omega)\right)/n_k\right\|_{\mathscr{F}} \to 0 \quad \omega\text{-a.s.}$$

[see the inequalities following (2.15)]

$$\Rightarrow E_A\left\|\sum_{j=1}^{n} \left(\delta_{\hat{X}^\omega_{nj}} - P_n(\omega)\right)/n\right\|_{\mathscr{F}} \to 0 \quad \text{in probability}$$

which is even more than the actual statement (ii).

(ii) $\Rightarrow$ (i). If (ii) holds, we obtain as in the CLT that

$$n \Pr(F > n) \to 0. \tag{3.16}$$

Recall that for any $\{n_k\}$ for which $d_{\mathrm{pr}_A}(\|\Sigma_{j=1}^{n_k}(\delta_{\hat{X}^\omega_{n_kj}} - P_{n_k}(\omega))/n_k\|_{\mathscr{F}}, 0) \to 0$ a.s.

$$\sum_{i=1}^{n_k} I(F(X_i) > an_k) = 0 \quad \text{eventually, a.s.} \tag{3.17}$$

So, as above $E_A \max_{j\le n_k}\|\delta_{\hat{X}^\omega_{nj}}/n\|_{\mathscr{F}} \to 0$ a.s. Also, $\|P_{n_k}(\omega)/n_k\|_{\mathscr{F}} = \|\Sigma_{i=1}^{n_k} f(X_i)/n_k^2\|_{\mathscr{F}} \to 0$ (since, eventually, this norm is less than or equal to a). So, $E_A \max_{j\le n_k}\|(\delta_{\hat{X}^\omega_{nj}} - P_{n_k}(\omega))/n_k\|_{\mathscr{F}} \to 0$ a.s. Hence, by Hoffmann-

Jørgensen's inequality,

$$E_A\left\|\frac{\sum_{j=1}^{n_k}\left(\delta_{\hat{X}_{n_kj}^{\omega}} - P_{n_k}(\omega)\right)}{n_k}\right\| \to 0 \quad \text{a.s.} \tag{3.18}$$

Now, as in (2.24),

$$cE_\varepsilon\left\|\sum_{i=1}^{n_k}\varepsilon_i\delta_{X_i}/n_k\right\|_{\mathscr{F}} \le 2E_A\|(\hat{P}_{n_k}(\omega) - P_{n_k}(\omega))\|_{\mathscr{F}} + \|P_{n_k}(\omega)\|_{\mathscr{F}}/n_k^{1/2},$$

so,

$$\begin{aligned}\limsup \Pr\left\{\left\|\sum_{i=1}^{n_k}\frac{\varepsilon_i\delta_{X_i}}{n_k}\right\|_{\mathscr{F}} \ge \varepsilon\right\} &\le \frac{1}{\varepsilon}\limsup E\left[\left(2c^{-1}E_A\|\hat{P}_{n_k}(\omega) - P_{n_k}(\omega)\|_{\mathscr{F}}\right)\wedge\varepsilon\right]\\ &\quad + \frac{1}{\varepsilon}E\left(\frac{c^{-1}\|P_{n_k}(\omega)\|_{\mathscr{F}}}{n_k^{1/2}}\wedge\varepsilon\right)\\ &= (\mathrm{I}) + (\mathrm{II}).\end{aligned}$$

(I) $\to 0$ by (3.18) and the dominated convergence theorem, and (II) $\to 0$ because, by (3.16), $\sum_{i=1}^{n}F(X_i)/n^{3/2} \to 0$ in probability. Hence,

$$\left\|\sum_{i=1}^{n}\varepsilon_i\delta_{X_i}/n\right\|_{\mathscr{F}} \to 0 \quad \text{in probability.} \tag{3.19}$$

Finally, (i) follows by a standard desymmetrization:

$$\left\|\sum_{i=1}^{n} I(F(X_i) > n)\delta_{X_i}/n\right\|_{\mathscr{F}} \to 0$$

in probability by (3.16), hence we can truncate in (3.19) and then take expectations and use the symmetrization inequalities given immediately after (2.15) to obtain $E\|\sum_{i=1}^{n}(f(X_i)I(F(X_i) \le n) - PfI(F \le n))/n\|_{\mathscr{F}} \to 0$. Again, using (3.16) we obtain (i). □

3.6 Remark. The weak law of large numbers with normings other than n (i.e., $n^{1/p}$ or even more general a_n's) can also be bootstrapped in probability, in complete analogy with Theorem 3.5. [See, e.g., Andersen, Giné and Zinn (1988) for examples of Marcinkiewicz-type laws of large numbers for empirical processes.]

Acknowledgment. The authors would like to thank Professor Richard M. Dudley for a careful reading of this paper and for motivating us to clarify the exposition.

REFERENCES

Alexander, K. S. (1987). The central limit theorem for empirical processes on Vapnik–Červonenkis classes. *Ann. Probab.* **15** 178–203.

Andersen, N. T., Giné, E. and Zinn, J. (1988). The central limit theorem for empirical processes under local conditions: The case of Radon limits without Gaussian component. *Trans. Amer. Math. Soc.* **308** 603–635.

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

BERAN, R. (1982). Estimated sampling distributions: The bootstrap and competitors. *Ann. Statist.* **10** 212–225.

BERAN, R. (1984). Bootstrap methods in statistics. *Jahresber. Deutsch. Math.-Verein.* **86** 14–30.

BERAN, R., LE CAM, L. and MILLAR, P. W. (1987). Convergence of stochastic empirical measures. *J. Multivariate Anal.* **23** 159–168.

BERAN, R. and MILLAR, P. W. (1986). Confidence sets for a multivariate distribution. *Ann. Statist.* **14** 431–443.

BICKEL, P. J. and FREEDMAN, D. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1216.

BRETAGNOLLE, J. (1983). Lois limites du bootstrap de certaines fonctionelles. *Ann. Inst. H. Poincaré. Probab. Statist.* **19** 281–296.

DE ACOSTA, A. and GINÉ, E. (1979). Convergence of moments and related functionals in the general central limit theorem in Banach spaces. *Z. Wahrsch. Verw. Gebiete* **48** 213–231.

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

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

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

GAENSSLER, P. (1987). Bootstrapping empirical measures indexed by Vapnik–Červonenkis classes of sets. In *Probability Theory and Mathematical Statistics* (Yu. V. Prohorov, V. A. Statulevicius, V. V. Sazanov and B. Grigelronis, eds.) **1** 467–481. VNU Science Press, Utrecht.

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

GINÉ, E. and ZINN, J. (1986). Lectures on the central limit theorem for empirical processes. *Probability and Banach Spaces. Lecture Notes in Math.* **1221** 50–113. Springer, Berlin.

GINÉ, E. and ZINN, J. (1989). Necessary conditions for the bootstrap of the mean. *Ann. Statist.* **17** 684–691.

GINÉ, E., MARCUS, M. B. and ZINN, J. (1990). On random multipliers in the central limit theorem with p-stable limit, $0 < p < 2$. *Proc. Sixth International Conference on Probability in Banach Spaces* 120–149. Birkhäuser, Boston.

HAAGERUP, U. (1981). The best constants in the Khinchine inequality. *Studia Math.* **70** 231–283.

HOFFMANN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–186.

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

LE CAM, L. (1970). Remarques sur le théorème limite central dans les espaces localement convexes. In *Les Probabilités sur les Structures Algébriques* 233–249. CNRS, Paris.

LEDOUX, M. and TALAGRAND, M. (1988a). Characterization of the law of the iterated logarithm in Banach spaces. *Ann. Probab.* **16** 1242–1264.

LEDOUX, M. and TALAGRAND, M. (1988b). Un critère sur les petites boules dans le théorème limite central. *Probab. Theory Related Fields* **77** 29–47.

PISIER, G. (1975). Le théorème limite central et la loi du logarithme iteré dans les espaces de Banach. *Séminaire Maurey–Schwartz 1975–1976, exposés 3 et 4.* Ecole Polytechnique, Paris.

POLLARD, D. (1981). Limit theorems for empirical processes. *Z. Wahrsch. Verw. Gebiete* **57** 181–185.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

SZAREK, S. J. (1976). On the best constant in the Khinchine inequality. *Studia Math.* **58** 197–208.

DEPARTMENT OF MATHEMATICS
THE COLLEGE OF STATEN ISLAND OF CUNY
130 STUYVESANT PLACE
STATEN ISLAND, NEW YORK 10301

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