SOME CONVERSE LIMIT THEOREMS FOR EXCHANGEABLE BOOTSTRAPS

JON A. WELLNER[1]

University of Washington

The bootstrap Glivenko-Cantelli and bootstrap Donsker theorems of Giné and Zinn (1990) contain both necessary and sufficient conditions for the asymptotic validity of Efron's nonparametric bootstrap. In the more general case of exchangeably weighted bootstraps, Praestgaard and Wellner (1993) and Van der Vaart and Wellner (1996) give analogues of the sufficiency half of the Theorems of Giné and Zinn (1990), but did not address the corresponding necessity parts of the theorems. Here we establish a new lower bound for exchangeably weighted processes and show that the necessity half of the a.s. bootstrap Glivenko-Cantelli theorem holds for exchangeably weighted bootstraps. We also make some progress toward the conjectured necessity parts of the bootstrap Donsker theorems with exchangeable weights.

AMS subject classifications: Primary: 60F05, 60F17; secondary 60J65, 60J70.
Keywords and phrases: asymptotic distributions, bootstrap, consistency, converse, empirical measures, empirical processes, exchangeable, Donsker theorem, Glivenko-Cantelli theorem, multiplier inequalities.

1 Introduction

Giné and Zinn (1990) established several beautiful limit theorems for Efron's nonparametric bootstrap of the general empirical process. One of the key tools used by Giné and Zinn (1990) was the *multiplier inequality* used earlier in Giné and Zinn (1983) with Gaussian multipliers to "Gaussianize" the empirical process and to relate the Gaussianized process to the symmetrized empirical process, and hence to the empirical process itself via symmetrization and de-symmetrization inequalities.

Efron's nonparametric bootstrap, which involves resampling from the empirical measure $\mathbb{P}_n$, can be viewed as one instance of an exchangeably weighted bootstrap with the weights being the components of a random vector which has a Multinomial distribution with n cells, n trials, and vector of "success" probabilities $(1/n, \ldots, 1/n)$. The first limit theory for exchangeably weighted bootstraps was established by Mason and Newton (1992); they treated exchangeable bootstrapping of the mean and of the classical empirical process. Praestgaard and Wellner (1993) extended the direct half

[1]Research supported in part by National Science Foundation grant DMS-9532039 and NIAID grant 2R01 AI291968-04.

of several of the theorems of Giné and Zinn (1990) by use of a new type of multiplier inequality involving symmetrization with ranks rather than with Rademacher random variables. Similarly, the direct half of the bootstrap Glivenko - Cantelli theorems of Giné and Zinn (1990) was established in Van der Vaart and Wellner (1996); see Lemma 3.6.16, page 357. In both Praestgaard and Wellner (1993) and Van der Vaart and Wellner (1996), the analogue of the converse half of the theorems of Giné and Zinn (1990) could not be proved for the general exchangeable bootstrap because of the lack of an appropriate analogue of the lower bound in the i.i.d. version of the multiplier inequality.

In this paper we make some progress toward filling these gaps. We first establish an appropriate lower bound in the case of exchangeably weighted sums of i.i.d. random elements (Section 2). We then show how this lower bound yields the converse half of bootstrap Glivenko - Cantelli theorems (Section 3). In the case of bootstrap Donsker theorems (Section 4), we are able to establish some of the integrability needed for the converse half, but still lack an appropriate analogue of the Hoffmann-Jørgensen theorem needed to obtain the crucial uniform integrability needed to apply the L_1 inequalities established in Section 2.

For other approaches to proving bootstrap limit theorems in other cases of interest, see Csörgő and Mason (1989), Einmahl and Mason (1992), and Shorack (1996), (1997).

2 Multiplier inequalities: a new lower bound

First we give a statement of the multiplier inequalities for the case of i.i.d. multipliers.

Let $Z_1, \ldots, Z_n$ be independent and identically distributed (i.i.d.) processes indexed by a set $\mathcal{F}$ with mean 0 (i.e. $EZ_i(f) = 0$ for all $f \in \mathcal{F}$), and let $\xi_1, \ldots, \xi_n, \ldots$ be i.i.d. real-valued random variables which are independent of $Z_1, \ldots, Z_n, \ldots$; in the case of empirical processes, $Z_i = \delta_{X_i} - P$ for i.i.d. random elements $X_i \in \mathcal{X}$ where the basic probability space is $(\mathcal{X}, \mathcal{A}, P)$. For processes Z_i we write $||Z_i||_{\mathcal{F}}$ for $\sup_{f \in \mathcal{F}} |Z_i(f)|$. As in Van der Vaart and Wellner (1996), we will assume throughout that $X_1, X_2, \ldots$ are defined as the coordinate projections on the "first" component of the product probability space $(\mathcal{X}^\infty \times \mathcal{Z}, \mathcal{A}^\infty \times \mathcal{C}, P^\infty \times Q)$, and let $\xi_1, \xi_2, \ldots$ depend on the last coordinate only.

Lemma 2.1 *(Multiplier inequality for i.i.d. multipliers). Let $Z_1, \ldots, Z_n$ be i.i.d stochastic processes with $E^*||Z_i||_{\mathcal{F}} < \infty$ independent of the Rademacher variables $\epsilon_1, \ldots, \epsilon_n$. Then for every i.i.d. sample $\xi_1, \ldots, \xi_n$ of mean-zero* *random variables independent of* $Z_1, \ldots, Z_n$*, and any* $1 \le n_0 \le n$*,*

$$\begin{aligned} \frac{1}{2}\|\xi\|_1 E^*\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \epsilon_i Z_i\Big\|_{\mathcal{F}} &\le E^*\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i Z_i\Big\|_{\mathcal{F}} \\ &\le 2(n_0-1)E^*\|Z_1\|_{\mathcal{F}} E \max_{1\le i\le n}\frac{|\xi_i|}{\sqrt{n}} \\ &\quad + 2\sqrt{2}\|\xi\|_{2,1}\max_{n_0\le k\le n} E^*\Big\|\frac{1}{\sqrt{k}}\sum_{i=n_0}^k \epsilon_i Z_i\Big\|_{\mathcal{F}}. \end{aligned} \tag{1}$$

Here $\|\xi_1\|_{2,1} \equiv \int_0^\infty \sqrt{P(|\xi_1| > t)}dt$ is assumed to be finite. For a proof of Lemma 2.1, see Giné and Zinn (1983) or Van der Vaart and Wellner (1996). The important thing to be observed here is that both the upper and lower bounds for the expected norm of the multiplier process $\sum_{i=1}^n \xi_i Z_i$ are provided in terms of the symmetrized process $\sum_{i=1}^n \epsilon_i Z_i$.

Now we turn to the case of exchangeable multipliers. We will assume that the vector $\underline{\xi}_n \equiv (\xi_{n1}, \ldots, \xi_{nn})$ is exchangeable and $\|\xi_{n1}\|_{2,1} < \infty$ for each n. The upper bound half of the following inequality is a consequence of the multiplier inequality proved by Praestgaard and Wellner (1993).

Lemma 2.2 *(Multiplier inequality for exchangeable multipliers). Suppose that* $Z_1, \ldots, Z_n$ *are i.i.d. stochastic processes with* $E^*\|Z_i\|_{\mathcal{F}} < \infty$ *independent of the Rademacher variables* $\epsilon_1, \ldots, \epsilon_n$ *and of the random permutation* $\underline{R} = (R_1, \ldots, R_n)$ *of the first* n *integers. Then for every exchangeable random vector* $\underline{\xi}_n = (\xi_{n1}, \ldots, \xi_{nn})$ *independent of* $Z_1, \ldots, Z_n$*, and any* $1 \le n_0 \le n$*,*

$$\begin{aligned} \frac{1}{2}\|\xi_{n1}\|_1 E^*\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \epsilon_i Z_i\Big\|_{\mathcal{F}} &\le E^*\Big\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_{ni} Z_i\Big\|_{\mathcal{F}} \\ &\le 2(n_0-1)E^*\|Z_1\|_{\mathcal{F}} E \max_{1\le i\le n}\frac{|\xi_{ni}|}{\sqrt{n}} \\ &\quad + 2\|\xi_{n1}\|_{2,1}\max_{n_0\le k\le n} E^*\Big\|\frac{1}{\sqrt{k}}\sum_{i=n_0}^k Z_{R_i}\Big\|_{\mathcal{F}}. \end{aligned} \tag{2}$$

Proof The inequality on the right follows from Praestgaard and Wellner (1993), Lemma 4.1, page 2063. It remains only to prove the inequality on the left. To do this, let $Z_1', \ldots, Z_n'$ be an independent copy of $Z_1, \ldots, Z_n$ (canonically formed on an appropriate product probability space). Then by the triangle inequality

$$2E^*\Big\|\sum_{i=1}^n \xi_{ni} Z_i\Big\|_{\mathcal{F}} \ge E^*\Big\|\sum_{i=1}^n \xi_{ni}(Z_i - Z_i')\Big\|_{\mathcal{F}}$$

$$
\begin{aligned}
&= E^*\Big\|\sum_{i=1}^n \xi_{ni}\epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \\
&= E^*\Big\|\sum_{i=1}^n |\xi_{ni}|\mathrm{sign}(\xi_{ni})\epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \\
&= E^*\Big\|\sum_{i=1}^n |\xi_{ni}|\epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \qquad (3)
\end{aligned}
$$

where the last equality holds because the vector $(\mathrm{sign}(\xi_{n1})\epsilon_1, \ldots, \mathrm{sign}(\xi_{nn})\epsilon_n)$ has the same distribution as $(\epsilon_1, \ldots, \epsilon_n)$ and, moreover, is independent of $(\xi_{n1}, \ldots, \xi_{nn})$. Hence by convexity of the norm $\|\cdot\|_{\mathcal{F}}$ and Jensen's inequality the right side of (3) is bounded below by

$$
E|\xi_{n1}|E^*\Big\|\sum_{i=1}^n \epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \geq E|\xi_{n1}|E^*\Big\|\sum_{i=1}^n \epsilon_i Z_i\Big\|_{\mathcal{F}}
$$

by convexity again and since the Z_i''s have mean 0; this is quite similar to the proofs of Lemma 2.3.1, page 108, and Lemma 2.9.1, page 177, Van der Vaart and Wellner (1996) where the measurability details of the proof are given in detail. ∎

3 Bootstrap Glivenko-Cantelli theorems

Now suppose that $X_1, X_2, \ldots$ are i.i.d. P on $(\mathcal{X}, \mathcal{A})$, and let $\mathbb{P}_n$ be the empirical measure of the first n of the X_i's;

$$
\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}\,.
$$

The classical Efron nonparametric bootstrap empirical measure $\widehat{\mathbb{P}}_n$ is just

$$
\widehat{\mathbb{P}}_n = \frac{1}{n}\sum_{i=1}^n \delta_{\widehat{X}_i}
$$

where $\widehat{X}_1, \ldots, \widehat{X}_n$ is a sample drawn (with replacement) from $\mathbb{P}_n$. Giné and Zinn (1990) established the following bootstrap Glivenko-Cantelli theorem. Their notation NLDM(P) stands for "nearly linearly deviation measureable for P"; we refer to Giné and Zinn (1984), page 935 for the detailed definition.

Theorem 3.1 *(Glivenko-Cantelli theorem for Efron's Bootstrap). Suppose that $\mathcal{F}$ is NLDM(P). Then the following are equivalent:*
(a) $P(F) < \infty$ and $\|\mathbb{P}_n - P\|_{\mathcal{F}} \to 0$ in probability.
(b) $P^\infty-$ a.s. $\|\widehat{\mathbb{P}}_n - \mathbb{P}_n\|_{\mathcal{F}} \to 0$ in probability.

Proof See Giné and Zinn (1990), page 860. ■

Here is a slight reformulation of the bootstrap Glivenko-Cantelli theorem of Giné and Zinn (1990) which avoids the hypothesis that $\mathcal{F}$ is $NLDM(P)$. Let $BL_1(R)$ be the collection of all functions $h : R \to [0,1]$ such that $|h(x_1) - h(x_2)| \leq |x_1 - x_2|$ for all $x_1, x_2 \in R$.

Theorem 3.2 *(Modified Glivenko-Cantelli theorem for Efron's Bootstrap). The following are equivalent:*
(a) $P^*\|f - Pf\|_{\mathcal{F}} < \infty$ *and* $\|\mathbb{P}_n - P\|^*_{\mathcal{F}} \to 0$ *in probability.*
(b) $(P^\infty)^*-$ *a.s.* $\|\widehat{\mathbb{P}}_n - \mathbb{P}_n\|_{\mathcal{F}} \to 0$ *in probability and* $\widehat{E}h(\|\widehat{\mathbb{P}}_n - \mathbb{P}_n\|_{\mathcal{F}})^* - \widehat{E}h(\|\widehat{\mathbb{P}}_n - \mathbb{P}_n\|_{\mathcal{F}})_* \to_{a.s.} 0$ *for every* $h \in BL_1(R)$.

Proof This follows as a corollary of Theorem 3.3 below. ■

Now we turn to exchangeably weighted bootstraps. Suppose that $\underline{W}_n = (W_{n1}, \ldots, W_{nn})$ satisfies the following conditions:

A1. $\underline{W}_n = (W_{n1}, \ldots, W_{nn})$ is exchangeable for each n.

A2. $W_{ni} \geq 0$ for $i = 1, \ldots, n$ and $\sum_{i=1}^n W_{ni} = n$.

A3. $\max_{1 \leq i \leq n}(W_{ni}/n) \to_p 0$.

A4. $\lim_{n\to\infty} E|W_{n1} - 1| = b > 0$, $\limsup_{n\to\infty} E|W_{n1} - 1|^2 < \infty$.

For such a vector of exchangeable weights $\underline{W}_n$, consider the *exchangeably weighted bootstrap empirical measure* $\widehat{\mathbb{P}}_n^W$ given by

$$\widehat{\mathbb{P}}_n^W = \frac{1}{n}\sum_{i=1}^n W_{ni}\delta_{X_i}\,.$$

It is easily seen that the classical nonparametric bootstrap empirical measure is the special case of $\widehat{\mathbb{P}}_n^W$ obtained by taking $\underline{W}_n = \underline{M}_n$ where $\underline{M}_n = (M_{n1}, \ldots, M_{nn}) \sim Mult_n(n, (1/n, \ldots, 1/n))$.

Theorem 3.3 *(Glivenko-Cantelli theorem for the Exchangeable Bootstrap). Suppose that* $\{\underline{W}_n\}$ *satisfies A1 - A4. Then the following are equivalent:*
(a) $P^*\|f - Pf\|_{\mathcal{F}} < \infty$ *and* $\|\mathbb{P}_n - P\|^*_{\mathcal{F}} \to 0$ *in probability.*
(b) $P^\infty-$ *a.s.* $\|\widehat{\mathbb{P}}_n^W - \mathbb{P}_n\|^*_{\mathcal{F}} \to 0$ *in probability and* $\widehat{E}h(\|\widehat{\mathbb{P}}_n^W - \mathbb{P}_n\|_{\mathcal{F}})^* - \widehat{E}h(\|\widehat{\mathbb{P}}_n^W - \mathbb{P}_n\|_{\mathcal{F}})_* \to_{a.s.} 0$ *for every* $h \in BL_1(R)$.
Moreover, if either (a) or (b) holds it follows that

$$E_W\|\widehat{\mathbb{P}}_n^W - \mathbb{P}_n\|_{\mathcal{F}} \to_{a.s.*} 0\,. \tag{4}$$

Proof That (a) implies (b) was proved in Van der Vaart and Wellner (1996), Lemma 3.6.16, page 357. It remains only to prove (b) implies (a).

Suppose that we show that (b) implies $P^*\|f - Pf\|_{\mathcal{F}} < \infty$ and that (4) holds. It follows that

$$E^*\|\widehat{\mathbb{P}}_n^W - \mathbb{P}_n\|_{\mathcal{F}} \to 0 . \tag{5}$$

Now note that by A2 we can write

$$\widehat{\mathbb{P}}_n^W - \mathbb{P}_n = \frac{1}{n}\sum_{i=1}^n (W_{ni} - 1)(\delta_{X_i} - P) \equiv \frac{1}{n}\sum_{i=1}^n \xi_{ni} Z_i$$

with $\xi_{ni} \equiv W_{ni} - 1$ and $Z_i \equiv \delta_{X_i} - P$. Hence it follows from the Multiplier Lemma 2.2 and a standard symmetrization inequality (see e.g. Van der Vaart and Wellner (1996), Lemma 2.3.1, page 108) that

$$\begin{aligned} E^*\|\widehat{\mathbb{P}}_n^W - \mathbb{P}_n\|_{\mathcal{F}} &\geq \frac{1}{2} E|W_{n1} - 1| E^*\|\frac{1}{n}\sum_{i=1}^n \epsilon_i Z_i\|_{\mathcal{F}} \\ &\geq \frac{1}{4} E|W_{n1} - 1| E^*\|\frac{1}{n}\sum_{i=1}^n Z_i\|_{\mathcal{F}} \\ &= \frac{1}{4} E|W_{n1} - 1| E^*\|\mathbb{P}_n - P\|_{\mathcal{F}} . \end{aligned}$$

In view of the non-degeneracy condition A4, this together with (5) yields

$$E^*\|\mathbb{P}_n - P\|_{\mathcal{F}} \to 0 ,$$

which implies the convergence in probability part of (a) by Markov's inequality.

Now we show that (b) implies that $P^*\|f - Pf\|_{\mathcal{F}} < \infty$ and that (4) holds. Suppose that (b) holds. Let $X_1', X_2', \ldots$ be an independent copy of $X_1, X_2, \ldots$, and let $\epsilon_1, \ldots, \epsilon_n$ be independent Rademacher random variables which are independent of the $\{X_i\}$, $\{X_i'\}$, and $\{W_{ni}, i = 1, \ldots, n, n \geq 1\}$. Set $Z_i \equiv \delta_{X_i} - P$, $Z_i' \equiv \delta_{X_i'} - P$, and $\xi_{ni} \equiv W_{ni} - 1$. Then we have

$$\begin{aligned} \Big\|\frac{1}{n}\sum_{i=1}^n \xi_{ni} Z_i\Big\|_{\mathcal{F}} + \Big\|\frac{1}{n}\sum_{i=1}^n \xi_{ni} Z_i'\Big\|_{\mathcal{F}} &\geq \Big\|\frac{1}{n}\sum_{i=1}^n \xi_{ni}(Z_i - Z_i')\Big\|_{\mathcal{F}} \\ &=_d \Big\|\frac{1}{n}\sum_{i=1}^n \xi_{ni}\epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \\ &= \Big\|\frac{1}{n}\sum_{i=1}^n |\xi_{ni}|\mathrm{sign}(\xi_{ni})\epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \\ &=_d \Big\|\frac{1}{n}\sum_{i=1}^n |\xi_{ni}|\epsilon_i(Z_i - Z_i')\Big\|_{\mathcal{F}} \end{aligned}$$

Since the left side of this display converges to 0 in probability a.s. with respect to $P^{\infty}_{Z,Z'}$, so does the right side. That is, for every $\delta > 0$,

$$P_{\xi,\epsilon}\left(\left\|\frac{1}{n}\sum_{i=1}^{n}|\xi_{ni}|\epsilon_i(Z_i - Z_i')\right\|_{\mathcal{F}} > \delta\right) \to 0 \qquad \text{almost surely}.$$

By independence of $\underline{\xi}$, $\underline{\epsilon}$, and $\underline{Z}, \underline{Z}'$, this yields

$$E_\xi P_\epsilon\left(\left\|\frac{1}{n}\sum_{i=1}^{n}|\xi_{ni}|\epsilon_i(Z_i - Z_i')\right\|_{\mathcal{F}} > \delta\right) \to 0 \qquad \text{almost surely},$$

and this implies that

$$P_\epsilon\left(\left\|\frac{1}{n}\sum_{i=1}^{n}|\xi_{ni}|\epsilon_i(Z_i - Z_i')\right\|_{\mathcal{F}} > \delta\right) \to 0$$

in probability with respect to $\underline{\xi}$ and almost surely with respect to $\underline{Z}, \underline{Z}'$. But now the summands $Y_i \equiv |\xi_{ni}|\epsilon_i(Z_i - Z_i')$, $i = 1, \ldots, n$, are, for fixed $|\xi_{ni}|$ and $Z_i - Z_i'$, independent and symmetric. Hence by Lévy's inequality (e.g. Proposition A.1.2, Van der Vaart and Wellner (1996), page 431) it follows that

$$P_\epsilon\left(\max_{1\le i\le n}\left\|\frac{1}{n}|\xi_{ni}|\epsilon_i(Z_i - Z_i')\right\|_{\mathcal{F}} > \delta\right) \to 0$$

in probability with respect to $\underline{\xi}$ and almost surely with respect to $\underline{Z}, \underline{Z}'$. But the norm is equal to

$$\frac{1}{n}|\xi_{ni}|\left\|(Z_i - Z_i')\right\|_{\mathcal{F}}$$

and since this does not depend on ϵ_i, it follows that

$$\frac{1}{n}\max_{1\le i\le n}|\xi_{ni}|\left\|(Z_i - Z_i')\right\|_{\mathcal{F}} \to_p 0 \tag{6}$$

in probability with respect to $\underline{\xi}$ and almost surely with respect to $\underline{Z}, \underline{Z}'$. That is,

$$P_\xi(\frac{1}{n}\max_{1\le i\le n}|\xi_{ni}|\left\|(Z_i - Z_i')\right\|_{\mathcal{F}} > \epsilon) \to 0 \tag{7}$$

almost surely with respect to $\underline{Z}, \underline{Z}'$.

Now let R be a random permutation of the first n integers. Let $b_{ni} \equiv (1/n)\|(Z_i - Z_i')\|_{\mathcal{F}}$, and suppose that $I \in \{1, \ldots, n\}$ satisfies $\max_{1\le i\le n} b_{ni} = b_{nI}$. Then by exchangeability of the W_{ni}'s we have

$$\max_{1\le i\le n}|\xi_{ni}|b_{ni} =_d \max_{1\le i\le n}|\xi_{n,R(i)}|b_{ni}.$$

Hence, conditioning on the W_{ni}'s, it follows that the probability in (7) equals

$$
\begin{aligned}
E_W P(\max_{1\le i\le n} |\xi_{nR(i)}| b_{n,i} > \epsilon | \underline{W}) &\ge E_W P(|\xi_{nR(I)}| b_{nI} > \epsilon | \underline{W}) \\
&= E_W \left(\frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| b_{nI} > \epsilon]} \right) \\
&\ge E_W \left(\frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| > \sqrt{\epsilon}]} 1_{[b_{nI} > \sqrt{\epsilon}]} \right) \\
&= E_W \left(\frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| > \sqrt{\epsilon}]} 1_{[b_{nI} > \sqrt{\epsilon}]} \right) \\
&= E_W \left(\frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| > \sqrt{\epsilon}]} \right) 1_{[b_{nI} > \sqrt{\epsilon}]} \\
&= P(|\xi_{n1}| > \sqrt{\epsilon}) 1_{[b_{nI} > \sqrt{\epsilon}]} \,. \qquad (8)
\end{aligned}
$$

Now for any non-negative random variable Y and $a > 0$ we have

$$E(Y) = E(Y 1_{[Y \le a]}) + E(Y 1_{[Y > a]}) \le a + \sqrt{E(Y^2)} \sqrt{P(Y > a)}$$

and, choosing $a = cE(Y)$ with $c < 1$, this yields

$$P(Y > cE(Y)) \ge \frac{(1-c)^2 (EY)^2}{E(Y^2)} \;; \qquad (9)$$

this is the Paley-Zygmund argument. Using this with $Y = |\xi_{n1}|$ and $cEY = \sqrt{\epsilon}$ in combination with (7) and (8) yields

$$
\begin{aligned}
& P_\xi\left(\frac{1}{n} \max_{1\le i\le n} |\xi_{ni}| \|(Z_i - Z_i')\|_{\mathcal{F}} > \epsilon\right) \\
(10) \quad \ge \;& \frac{(1 - \sqrt{\epsilon}/E|\xi_{n1}|)^2 \{E|\xi_{n1}|\}^2}{E|\xi_{n1}|^2} 1\{n^{-1} \max_{1\le i\le n} \|(Z_i - Z_i')\|_{\mathcal{F}} > \sqrt{\epsilon}\} \,.
\end{aligned}
$$

But by hypothesis A4 the first term on the right side of (10) converges to a positive constant. Since the left side converges to 0 almost surely by (7), it follows that

$$n^{-1} \max_{1\le i\le n} \left\| (Z_i - Z_i') \right\|_{\mathcal{F}} \to_{a.s.} 0 \,. \qquad (11)$$

But it is well known that (11) holds if and only if

$$E^* \left(\|(Z_1 - Z_1')\|_{\mathcal{F}} \right) < \infty \,;$$

by convexity of the norm together with $EZ_1' = 0$, this implies that

$$E^*\left(\|Z_1\|_{\mathcal{F}}\right) = P^*\|f(X_1) - Pf\|_{\mathcal{F}} < \infty .$$

This together with A4 implies that (4) holds via uniform integrability. ∎

4 Bootstrap Donsker theorems: conjectures and partial proofs

Suppose that $\mathbb{P}_n$, $\widehat{\mathbb{P}}_n$, and $\widehat{\mathbb{P}}_n^W$ are defined as in Section 3, and define the empirical process $\mathbb{G}_n$, the nonparametric bootstrap empirical process $\widehat{\mathbb{G}}_n$, and the exchangeably weighted empirical process $\widehat{\mathbb{G}}_n^W$, by

$$(12)\quad \mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - P), \qquad \widehat{\mathbb{G}}_n = \sqrt{n}(\widehat{\mathbb{P}}_n - \mathbb{P}_n), \qquad \widehat{\mathbb{G}}_n^W = \sqrt{n}(\widehat{\mathbb{P}}_n^W - \mathbb{P}_n).$$

In this section we will consider the processes $\mathbb{G}_n$, $\widehat{\mathbb{G}}_n$, and $\widehat{\mathbb{G}}_n^W$ as processes indexed by functions $f \in \mathcal{F}$, where $\mathcal{F} \subset \mathcal{L}_2(P) = \mathcal{L}_2(\mathcal{X}, \mathcal{A}, P)$. The following bootstrap Donsker theorem results from Giné and Zinn (1990) together with the measurability improvements of Van der Vaart and Wellner (1996).

Theorem 4.1 *(Almost sure Donsker theorem for Efron's Bootstrap). The following are equivalent:*
(a) $P^*\|f - Pf\|_{\mathcal{F}}^2 < \infty$ *and* $\mathcal{F}$ *is* $P-$*Donsker; i.e.* $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ *in* $l^\infty(\mathcal{F})$.
(b) $\sup_{h\in BL_1} |\widehat{E}h(\widehat{\mathbb{G}}_n) - Eh(\mathbb{G}_P)| \to_{a.s.^*} 0$ *and* $\widehat{E}h(\widehat{\mathbb{G}}_n)^* - \widehat{E}h(\widehat{\mathbb{G}}_n)_* \to_{a.s.} 0$ *for every* $h \in BL_1$.

Theorem 4.2 *(In probability Donsker theorem for Efron's Bootstrap). The following are equivalent:*
(a) $\mathcal{F}$ *is* $P-$*Donsker; i.e.* $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ *in* $l^\infty(\mathcal{F})$.
(b) $\sup_{h\in BL_1} |\widehat{E}h(\widehat{\mathbb{G}}_n) - Eh(\mathbb{G}_P)| \to_{p^*} 0$ *and* $\widehat{\mathbb{G}}_n$ *is asymptotically measurable:* $Eh(\widehat{\mathbb{G}}_n)^* - Eh(\widehat{\mathbb{G}}_n)_* \to 0$ *for every* $h \in BL_1$.

Proof See Giné and Zinn (1990), page 857; for a version of Theorem 4.1 under additional measurability hypotheses; and see Giné and Zinn (1990), page 862 for a version of Theorem 4.2 under additional measurability hypotheses. The above statements are from Van der Vaart and Wellner (1996), page 347 (where complete proofs are also given). ∎

The strong point of these theorems is they provide necessary and sufficient conditions in order for the process $\widehat{\mathbb{G}}_n$ to converge weakly either almost surely or in probability respectively.

Now we turn to exchangeably weighted bootstraps. Our goal is to establish the analogues of the converse halves of of Theorems 4.1 and 4.2 in the case of the exchangeable bootstrap process $\widehat{\mathbb{G}}_n^W$. The resulting Theorems 4.3 and 4.4 below strengthen Theorems 2.1 and 2.2 of Praestgaard and

Wellner (1993) (or see Theorem 3.6.13 of Van der Vaart and Wellner (1996)) to the level of equivalence established by Giné and Zinn (1990) for Efron's bootstrap.

Suppose that $\underline{W}_n = (W_{n1}, \ldots, W_{nn})$ satisfies the following conditions:

B1. $\underline{W}_n = (W_{n1}, \ldots, W_{nn})$ is exchangeable for each n.

B2. $W_{ni} \geq 0$ for $i = 1, \ldots, n$ and $\sum_{i=1}^n W_{ni} = n$.

B3. $\sup_{n\geq 1} \|W_{n,1}\|_{2,1} < \infty$, and $\lim_{\lambda\to\infty} \limsup_{n\to\infty} \sup_{t\geq\lambda} t^2 P(W_{n1} \geq t) = 0$.

B4. $\lim_{n\to\infty} E|W_{n1} - 1| = b > 0$, and $n^{-1}\sum_{i=1}^n (W_{ni} - 1)^2 \to_p c^2 > 0$.

In view of Lemma 4.7 of Praestgaard and Wellner (1993), B3 and B4 together imply that $\{W_{n1}\}$ is uniformly square-integrable and hence so is $n^{-1}\sum_{i=1}^n (W_{ni} - 1)^2$. Therefore B3 and B4 together imply that $E(W_{n1} - 1)^2 \to c^2$.

Theorem 4.3 *(Conjectured Almost sure Donsker theorem for the Exchangeable Bootstrap). Suppose that $\{\underline{W}_n\}$ satisfies B1 - B4. Then the following are equivalent:*
(a) $P^*\|f - Pf\|_{\mathcal{F}}^2 < \infty$ and $\mathcal{F}$ is $P-$Donsker; i.e. $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ in $l^\infty(\mathcal{F})$.
(b) $\sup_{h\in BL_1} |\hat{E}h(\hat{\mathbb{G}}_n^W) - Eh(c\mathbb{G}_P)| \to_{a.s.^*} 0$ and $\hat{E}h(\hat{\mathbb{G}}_n^W)^* - \hat{E}h(\hat{\mathbb{G}}_n^W)_* \to_{a.s.} 0$ for every $h \in BL_1$.

Theorem 4.4 *(Conjectured In probability Donsker theorem for the Exchangeable Bootstrap). Suppose that $\{\underline{W}_n\}$ satisfies B1 - B4. Then the following are equivalent:*
(a) $\mathcal{F}$ is $P-$Donsker; i.e. $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ in $l^\infty(\mathcal{F})$.
(b) $\sup_{h\in BL_1} |\hat{E}h(\hat{\mathbb{G}}_n^W) - Eh(c\mathbb{G}_P)| \to_{P^*} 0$ and $\hat{\mathbb{G}}_n^W$ is asymptotically measurable.

Proof (Partial.) That (a) implies (b) in both Theorems 4.3 and 4.4 was proved in Van der Vaart and Wellner (1996), Theorem 3.6.13, page 355, under additional measurability hypotheses. It remains only to prove that (b) implies (a) in both cases.

In the converse direction, we are currently able only to show that (b) of Theorem 4.3 implies the integrability condition of (a) of Theorem 4.3. Suppose that (b) holds. Then, with $(Z_1', Z_2', \ldots)$ an independent copy of $(Z_1, Z_2, \ldots)$, and $\epsilon_1, \epsilon_2, \ldots$ a sequence of i.i.d Rademacher random variables independent of the $\underline{W}$'s, Z_i's, and Z_i''s,

$$\hat{\mathbb{G}}_n^W - \hat{\mathbb{G}}_n^{W\prime} = \frac{1}{\sqrt{n}}\sum_{i=1}^n (W_{ni} - 1)(Z_i - Z_i')$$

$$=_d \frac{1}{\sqrt{n}} \sum_{i=1}^n |W_{ni}-1| \mathrm{sign}(W_{ni}) \epsilon_i (Z_i - Z_i')$$
$$=_d \frac{1}{\sqrt{n}} \sum_{i=1}^n |W_{ni}-1| \epsilon_i (Z_i - Z_i') .$$

In view of (b) it follows that this difference converges weakly for almost all sequences $Z_1, Z_2, \ldots, Z_1', Z_2', \ldots,$ to the tight Gaussian process $c(\mathbb{G}_P - \mathbb{G}_P')$ with values in $C_u(\mathcal{F})$, the space of uniformly continuous functions from $\mathcal{F}$ to R. Thus the supremum $c\|\mathbb{G}_P - \mathbb{G}_P'\|_{\mathcal{F}}$ has moments of all orders, and for every $\epsilon > 0$ there exists an x sufficiently large so that

$$P(c\|\mathbb{G}_P - \mathbb{G}_P'\|_{\mathcal{F}} \geq x) \leq \frac{\epsilon}{x^2} .$$

Hence, by the Portmanteau theorem, for $n \geq N_{\omega,\omega'}$,

$$P^*_{W,\epsilon}(\| \sum_{i=1}^n |W_{ni}-1| \epsilon_i (Z_i(\omega) - Z_i'(\omega'))\|_{\mathcal{F}} > x\sqrt{n})$$
$$\leq 2P(c\|\mathbb{G}_P - \mathbb{G}_P'\|_{\mathcal{F}} \geq x) \leq \frac{2\epsilon}{x^2}.$$

But by Lévy's inequality used conditionally on the $\underline{W}$'s, the left side is

$$\begin{aligned} & E_W P_\epsilon \left(\| \sum_{i=1}^n |W_{ni}-1| \epsilon_i (Z_i - Z_i') \|_{\mathcal{F}} > x\sqrt{n} \right) \\ \geq\ & \frac{1}{2} E_W P_\epsilon \left(\max_{1\leq i\leq n} |W_{ni}-1| |\epsilon_i| \|Z_i - Z_i'\|_{\mathcal{F}} > x\sqrt{n} \right) \\ =\ & \frac{1}{2} E_W P_\epsilon \left(\max_{1\leq i\leq n} |W_{ni}-1| \|Z_i - Z_i'\|_{\mathcal{F}} > x\sqrt{n} \right) . \end{aligned}$$

Thus, letting $b_{ni} \equiv n^{-1/2}\|(Z_i - Z_i')\|_{\mathcal{F}}$ and $\xi_{ni} \equiv W_{ni} - 1$, we have

$$P_W \left(\max_{1\leq i\leq n} |\xi_{ni}| b_{ni} > x\right) \leq \frac{4\epsilon}{x^2} . \tag{13}$$

Now let $I \in \{1, \ldots, n\}$ satisfy $\max_{1\leq i\leq n} b_{ni} = b_{nI}$. Then by exchangeability of the W_{ni}'s it follows that

$$\max_{1\leq i\leq n} |\xi_{ni}| b_{ni} =_d \max_{1\leq i\leq n} |\xi_{n,R(i)}| b_{ni}$$

where $\underline{R}$ is a random permutation which is independent of the $\underline{W}$'s and the Z_i's, and Z_i''s. Thus the left side of (13) is equal to

$$E_W P_R \left(\max_{1\leq i\leq n} |\xi_{n,R(i)}| b_{ni} > x \right) \geq E_W P_R \left(|\xi_{n,R(I)}| b_{nI} > x \right)$$

$$= E_W \left\{ \frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| b_{nI} > x]} \right\}$$

$$\geq E_W \left(\frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| > \frac{1}{2} E|\xi_{n1}|]} 1_{[b_{nI} > 2x/E|\xi_{n1}|]} \right)$$

$$= E_W \left(\frac{1}{n} \sum_{j=1}^{n} 1_{[|\xi_{nj}| > \frac{1}{2} E|\xi_{n1}|]} \right) 1_{[b_{nI} > 2x/E|\xi_{n1}|]}$$

$$= P(|\xi_{n1}| > (1/2)E|\xi_{n1}|) \cdot 1\{b_{nI} > 2x/E|\xi_{n1}|\}$$

$$\geq \frac{(1/4)[E|\xi_{n1}|]^2}{[E|\xi_{n1}|^2]} 1\{b_{nI} > 2x/E|\xi_{n1}|\}. \tag{14}$$

where we have used the inequality (9) with $Y = |\xi_{n1}|$ and $c = 1/2$ in the last step. Now the first term on the right side of (14) has a positive limit inferior as $n \to \infty$ by B3-B4. Because the right side is smaller than $4\epsilon/x^2$ by (13), it follows that

$$\limsup_{n \to \infty} \frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} \|Z_i - Z_i'\|_{\mathcal{F}} \leq \frac{4x}{b} < \infty \tag{15}$$

almost surely where $b \equiv \lim_{n\to\infty} E|W_{n1} - 1| = \lim_{n\to\infty} E|\xi_{n1}|$. By a standard argument, this implies that

$$E^* \left(\|(Z_1 - Z_1')\|_{\mathcal{F}}^2 \right) < \infty .$$

By convexity of the norm together with $EZ_1' = 0$, this implies that

$$E^* \left(\|Z_1\|_{\mathcal{F}}^2 \right) = P^* \|f(X_1) - Pf\|_{\mathcal{F}}^2 < \infty .$$

This concludes the proof. ∎

To complete the proof of the conjectured Theorem 4.3, it would suffice to show that

$$\{\|\widehat{\mathbb{G}}_n^W\|_{\mathcal{F}}^* : \ n \geq 1\} \ \text{ is uniformly integrable}. \tag{16}$$

and

$$\{\|\widehat{\mathbb{G}}_n^W\|_{\mathcal{F}}^* = \|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_{ni}(\delta_{X_i^\omega} - \mathbb{P}_n^\omega)\|_{\mathcal{F}} : \ n \geq 1\} \quad \text{is } P^\infty - \text{a.s. uniformly integrable}. \tag{17}$$

If (16) and (17) could be proved, then the proof of the conjectured Theorems 4.3 and 4.4 are easily finished by use of the left inequality in Lemma 2.2. By general weak convergence theory, if (b) of Theorem 4.3 holds, then, for any $\delta_n \to 0$ and every $\epsilon > 0$,

$$P_W\left(\|\widehat{\mathbb{G}}_n^W\|_{\mathcal{F}_{\delta_n}} > \epsilon\right) \to_{a.s.*} 0 \qquad \text{as } n \to \infty\,;$$

see e.g. Van der Vaart and Wellner (1996), Theorem 1.5.7, page 37. This together with (17) implies that

$$E_W^*\|\widehat{\mathbb{G}}_n^W\|_{\mathcal{F}_{\delta_n}} \to_{a.s.*} 0 \qquad \text{as } n \to \infty\,,$$

and then, in turn,

$$E^*\|\widehat{\mathbb{G}}_n^W\|_{\mathcal{F}_{\delta_n}} \to 0 \qquad \text{as } n \to \infty\,. \tag{18}$$

Combining (18) with the left inequality in Lemma 2.2 (with $\|\cdot\|_{\mathcal{F}}$ replaced by $\|\cdot\|_{\mathcal{F}_{\delta_n}}$), $\xi_{ni} = W_{ni} - 1$, and $Z_i \equiv \delta_{X_i} - P$, we would be able to conclude that

$$\frac{1}{2}E|\xi_{n1}|E^*\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \epsilon_i Z_i\|_{\mathcal{F}_{\delta_n}} \le E^*\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_{ni} Z_i\|_{\mathcal{F}_{\delta_n}} = E^*\|\widehat{\mathbb{G}}_n^W\|_{\mathcal{F}_{\delta_n}} \to 0$$

as $n \to \infty$, and since $\lim_{n\to\infty} E|\xi_{n1}| = b > 0$ by B4, this yields (by a standard symmetrization inequality)

$$\frac{1}{2}E^*\|\mathbb{G}_n\|_{\mathcal{F}_{\delta_n}} \le E^*\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \epsilon_i Z_i\|_{\mathcal{F}_{\delta_n}} \to 0 \qquad \text{as } n \to \infty\,.$$

It would follow (e.g. by Van der Vaart and Wellner (1996), Corollary 2.3.12, page 115) that $\mathcal{F}$ is $P-$Donsker, so that (a) of Theorem 4.3 would hold.

Unfortunately, we have not been able to establish (16) and (17). The missing link seems to be a suitable replacement in the exchangeable case for either Hoffmann-Jørgensen's inequality (see Van der Vaart and Wellner (1996), Proposition A.1.5, page 433), or the uniform in n weak-L_2 condition for $\|\mathbb{G}_n\|_{\mathcal{F}}$ of Van der Vaart and Wellner (1996), Lemma 2.3.9, page 113.

Acknowledgements. We owe thanks to Evarist Giné, Jens Praestgaard, and Aad van der Vaart for conversations about these problems and (conjectured) theorems.

REFERENCES

Bickel, P. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics* **9**, 1196-1217.

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 - 321.

Csörgő, S. and Mason, D. M. (1989). Bootstrapping empirical functions. *Annals of Statistics* **17**, 1447-1471.

Einmahl, U. and Mason, D. M. (1992). Approximations to permutation and exchangeable processes. *Journal of Theoretical Probability* **5**, 101-126.

Giné, E. (1997). Lectures on Some Aspects of the Bootstrap. In *Lectures on Probability Theory and Statistics*, Ecole d'Eté de Probabilités de Saint-Flour, XXVI-1996. *Lecture Notes in Mathematics* **1665**, 37 - 151.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Annals of Probability* **12**, 929-989.

Giné, E. and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes. *Lecture Notes in Mathematics* **1211**, 50-113.

Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Annals of Probability* **18**, 851-869.

Mason, D. M. and Newton, M. (1992). A rank statistics approach to the consistency of a generalized bootstrap. *Annals of Statistics* **20**, 1611-1624.

Praestgaard, J. and Wellner, J.A. (1993). Exchangeably weighted bootstraps of the general empirical processes. *Annals of Probabability* **21**, 2053-2086.

Shorack, G. R. (1996). Linear rank statistics, finite sampling, permutation tests and winsorizing. *Annals of Statistics* **24**, 1371 - 1385.

Shorack, G. R. (1997). Uniform CLT, WLLN, LIL and bootstrapping in a data analytic approach to trimmed L-statistics. *Journal of Statistical Planning and Inference* **30**, 1 - 44.

Van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.

UNIVERSITY OF WASHINGTON
STATISTICS
BOX 354322
SEATTLE, WASHINGTON 98195-4322
U.S.A
jaw@stat.washington.edu