The Annals of Statistics
1996, Vol. 24, No. 5, 1914–1933

SECOND-ORDER CORRECTNESS OF THE BLOCKWISE BOOTSTRAP FOR STATIONARY OBSERVATIONS

BY F. GÖTZE AND H. R. KÜNSCH

University of Bielefeld and ETH Zürich

We show that the blockwise bootstrap approximation for the distribution of a studentized statistic computed from dependent data is second-order correct provided we choose an appropriate variance estimator. We also show how to adapt the BC_a confidence interval of Efron to the dependent case. For the proofs we extend the results of Götze and Hipp on the validity of the formal Edgeworth expansion for a sum to the studentized mean.

1. Introduction. Efron's (1979) bootstrap is distribution-free, but the assumption of independence is crucial since it ignores the time order of the observations. For serially dependent observations, Künsch (1989) has proposed a blockwise bootstrap which samples blocks of length l of consecutive observations with replacement. He has shown that this procedure estimates the asymptotic variance and the asymptotic normal distribution consistently if the statistic considered is a smooth function of vector means. The only assumptions are suitable mixing and moment conditions for the observations and an increase of $l = l(n)$ to infinity, but at a slower speed than n. However, for such simple statistics there are other methods to estimate the asymptotic variance which require fewer computations, for example, the blockwise jackknife or an estimate of the spectrum at zero for the estimated influence function [see also Künsch (1989)]. Hence the additional effort required for the bootstrap is only justified if the resulting approximation to the distribution of the statistic is better than the one relying on asymptotic normality. With i.i.d. data, Efron's bootstrap has indeed this property; see, for example, Singh (1981) and Hall (1988). Here we show that this is also true for the blockwise bootstrap as announced some time ago in an abstract [Götze and Künsch (1990)]. In the meantime, Lahiri (1996) also has given a proof of a similar result.

The essential reason for second-order correctness is the asymptotically correct skewness of the blockwise bootstrap distribution [cf. Künsch (1989), formula (3.19)]. This gain in accuracy is, however, covered by the error in the bootstrap variance which is in our case of the order $\mathscr{O}(l^{-1}) + \mathscr{O}_P(l^{1/2}n^{-1/2})$ and thus always larger than the order $\mathscr{O}(n^{-1/2})$ of the skewness term. So for

Received October 1993; revised September 1995.
AMS 1991 *subject classifications*. Primary 62M10; secondary 62G15, 62E20.
Key words and phrases. Resampling, Edgeworth expansion, studentization, BC_a confidence interval, time series, dependent data, strong mixing.

a better than normal approximation we have to make the true and the bootstrap variance equal, for example, by standardizing the statistic if its variance is known [Lahiri (1991)] or otherwise by studentizing. In the latter case, it is crucial how the variances are estimated; see Davison and Hall (1993). The essential requirement is that the variance estimators for the original and the bootstrap sample should have the same bias. This should become clearer in Section 2 where we discuss formal Edgeworth expansions.

Our result is somewhat surprising because dependence makes the estimation of the distribution of a statistic inherently more difficult. Note that we do not assume a specific type of dependence like an AR model where second-order correctness can be obtained by applying Efron's bootstrap for the estimated innovations; see Bose (1988). Nevertheless our bootstrap approximation is correct in the $n^{-1/2}$ term like for i.i.d. situations. The effect of dependence is felt only in the following term of the expansion, which is at best of $\mathscr{O}(l/n)$ instead of the usual $\mathscr{O}(n^{-1})$. The optimal block size l depends on the lag weights w_k of the variance estimator in formula (3). Choosing $w_0 = 1$ and $w_k = 2$ $(1 \le k < l)$ gives as optimal order $l = \mathscr{O}(n^{1/4})$, but for these weights the variance estimate could be negative. Other weights avoiding this have the optimal order $l = \mathscr{O}(n^{1/3})$; see the discussion at the end of Section 2.

As a direct consequence of our result, the coverage probability of one-sided bootstrap-t intervals is correct up to $o(n^{-1/2})$. However, these intervals are not equivariant under monotone transformations of the parameter. In Section 3 we propose an extension of the BC_a intervals of Efron (1987) to the dependent case. These intervals combine transformation equivariance with second-order coverage probabilities.

The proof will be given in Section 4. The most difficult part is to show the validity of the Edgeworth expansion for the studentized statistic for dependent observations. Götze and Hipp (1983) have shown the validity of the Edgeworth expansion for the arithmetic mean under a set of conditions covering a broad range of situations. We are going to use the same conditions. The difficulty in passing from the arithmetic to the studentized mean comes from the fact that the estimated variance is the sum of l different arithmetic means, namely, the sample covariances for lags $0, \dots, l-1$, and l increases with n.

2. Notations and heuristics. Suppose that we have observed a stationary process $(Y_j, j \in \mathbb{Z})$ with values in $\mathbb{R}^k$ for $j = 1, \dots, n$. Let us denote $\mathbb{E}[Y_j]$ by μ, $n^{-1} \sum_{j=1}^n Y_j$ by $\overline{Y}_n$ and $n^{1/2}(\overline{Y}_n - \mu)$ by S_n. We consider a statistic

$$U_n = H(\overline{Y}_n) = H(\mu + n^{-1/2} S_n)$$

as an estimator of $H(\mu)$, where $H: \mathbb{R}^k \to \mathbb{R}$ is a given function. This framework is not as narrow as it may seem at first glance. We can replace the original observations Y_j by

$$Y_j' \triangleq \varphi(Y_j, Y_{j+1}, \dots, Y_{j+m-1}) \tag{1}$$

with a fixed m and φ and apply everything to the transformed observations Y_j'. So, for instance, the following version of the sample autocovariances,

$$\hat{\gamma}_m = n^{-1}\sum_{j=1}^{n} Y_j Y_{j+m} - n^{-2}\sum_{j=1}^{n} Y_j \sum_{j=1}^{n} Y_{j+m},$$

fits into our framework. By taking nonlinear functions of $\hat{\gamma}_m$, we arrive at autocorrelations, partial autocorrelations and Yule–Walker estimators in autoregressive processes.

If H is differentiable and the gradient DH does not vanish at μ, then $n^{1/2}(U_n - H(\mu))$ is asymptotically equivalent to

$$DH(\mu)^T S_n = n^{-1/2}\sum_{i=1}^{n} Z_i \triangleq T_n \quad \text{where } Z_i \triangleq DH(\mu)^T(Y_i - \mu).$$

Moreover, under standard assumptions about moments of (Z_i) and the decay of mixing coefficients, T_n is asymptotically $\mathscr{N}(0, \sigma_\infty^2)$-distributed with

$$\sigma_\infty^2 \triangleq \sum_{j=-\infty}^{\infty} \mathbb{E}[Z_0 Z_j] = \lim_n \sigma_n^2, \tag{2}$$

where

$$\sigma_n^2 \triangleq \operatorname{Var}[T_n] = \sum_{j=-n}^{n} (1 - |j|/n)\mathbb{E}[Z_0 Z_j].$$

For studentizing U_n we need an estimator of σ_∞^2. We estimate the covariances $\mathbb{E}[Z_0 Z_j]$ and then truncate the infinite sum corresponding to (2) with a lag window

$$\hat{\sigma}_n^2 = DH(\overline{Y}_n)^T \sum_{k=0}^{l-1} w_k n^{-1} \sum_{j=1}^{n-l} (Y_j - \overline{Y}_n)(Y_{j+k} - \overline{Y}_n)^T DH(\overline{Y}_n). \tag{3}$$

The lag weights $w_k = w_k(l)$ are supposed to be of the form $w_0 = 1$ and $w_k = 2\omega(k/l)$ for $0 < k < l$ with $\omega: [0, 1) \to [0, 1]$ continuous and $\omega(0) = 1$. If $l \to \infty$ and $l/n \to 0$, then $\hat{\sigma}_n^2$ is consistent under mild conditions on the process (Y_j). The optimal rate for l will be given at the end of this section. The studentized statistic is

$$U_{n,\text{stud}} = n^{1/2}(U_n - H(\mu))\hat{\sigma}_n^{-1},$$

which is asymptotically $\mathscr{N}(0, 1)$-distributed.

Our aim is to produce a better than normal approximation of the distribution of $U_{n,\text{stud}}$ with the help of the blockwise bootstrap of Künsch (1989). It depends on a block length which we choose equal to the width l of the lag window in (3). The reason for this will be explained shortly. We then resample with replacement from all l-tuples of consecutive observations in the original sample. Assuming for simplicity that $n = bl$ with $b \in \mathbb{N}$, we let the starting blocks $N_1, \dots, N_b$ be i.i.d. uniform on $\{0, 1, \dots, n - l\}$ and take as the bootstrap sample $(Y_1^*, \dots, Y_n^*)$,

$$Y^*_{(j-1)l+i} = Y_{N_j+i} \qquad (1 \le j \le b,\, 1 \le i \le l).$$

The bootstrap statistic is then

$$U_n^* = H(\overline{Y}_n^*).$$

Denoting the conditional expectation of any function of the Y_j^*s for given $Y_1, \dots, Y_n$ by $\mathbb{E}^*$, a simple calculation similar to Lemma 3.1 of Künsch (1989) shows that

$$\mu_n^* \triangleq \mathbb{E}^*[\overline{Y}_n^*] = l^{-1}\mathbb{E}^*[Y_1^* + \cdots + Y_l^*] = (n - l + 1)^{-1} \sum_{j=0}^{n-l} l^{-1} \sum_{i=1}^{l} Y_{j+i}$$

$$= (n - l + 1)^{-1} \sum_{j=1}^{n} \min(j/l, 1, (n + 1 - j)/l) Y_j = \overline{Y}_n + \mathscr{O}_P(l^{1/2} n^{-1}).$$

Hence $n^{1/2}(\overline{Y}_n^* - \overline{Y}_n)$ has a (conditional) bias of $\mathscr{O}(l^{1/2}n^{-1/2})$, which is larger than the first term in the Edgeworth expansion of $n^{1/2}(\overline{Y}_n - \mu)$. So for second-order correctness we must center U_n^* at $H(\mu_n^*)$ instead of $H(\overline{Y}_n)$.

Let us introduce the quantities

$$A_j \triangleq l^{-1/2} \sum_{i=1}^{l} (Y_{j+i} - \mu_n^*), \qquad S_n^* \triangleq n^{1/2}(\overline{Y}_n^* - \mu_n^*) = b^{-1/2} \sum_{j=1}^{b} A_{N_j}.$$

By a Taylor expansion of H at μ_n^* we see that the linear part of $n^{1/2}(U_n^* - H(\mu_n^*))$ is

$$T_n^* \triangleq DH(\mu_n^*)^T S_n^* = b^{-1/2} \sum_{j=1}^{b} B_{N_j},$$

where $B_j \triangleq DH(\mu_n^*)^T A_j$. By the independence of $N_1, \dots, N_b$ it follows that

$$\sigma_n^{*2} \triangleq \mathrm{Var}^*[T_n^*] = \mathrm{Var}^*[B_{N_1}] = (n - l + 1)^{-1} \sum_{j=0}^{n-l} B_j^2.$$

A straightforward computation shows that up to boundary effects, σ_n^{*2} is equal to $\hat{\sigma}_n^2$ of (3) with $w_k = 2(1 - k/l)$. This is the reason why we choose the window width in (3) and the block length of the bootstrap to be equal. However, our results could be obtained also without this special choice. Asymptotic normality of the linear part T_n^* can be proved using Lindeberg's condition. So essentially first-order correctness of the bootstrap follows from the convergence of σ_n^{*2} to σ_∞^2; see Künsch (1989). Because $\sigma_n^{*2} - \sigma_n^2$ is always at least of the order $\mathscr{O}_P(n^{-1/3})$ (except when the Z_js are uncorrelated), removing the error due to the wrong variance is of interest.

The obvious estimate of σ_n^{*2} is

$$\hat{\sigma}_n^{*2} \triangleq b^{-1} \sum_{j=1}^{b} \hat{B}_{N_j}^2, \quad \text{where } \hat{B}_j \triangleq l^{-1/2} \sum_{i=1}^{l} DH(\bar{Y}_n^*)^T (Y_{j+i} - \bar{Y}_n^*). \tag{4}$$

Then the studentized bootstrap statistic is

$$U_{n,\text{stud}}^* = n^{1/2}(U_n^* - H(\mu_n^*))/\hat{\sigma}_n^*.$$

Note that in using $\hat{B}_j$ instead of B_j we have mimicked the uncertainty of estimating μ by $\bar{Y}_n$ in (3). However, in other aspects $\hat{\sigma}_n^{*2}$ does not correspond exactly to $\hat{\sigma}_n^2$. Namely, we can write

$$\hat{\sigma}_n^{*2} = n^{-1} DH(\bar{Y}_n^*)^T \sum_{[i/l]=[j/l]} (Y_i^* - \bar{Y}_n^*)(Y_j^* - \bar{Y}_n^*) DH(\bar{Y}_n^*), \tag{5}$$

where $[x]$ denotes the integer part of x. So for $\hat{\sigma}_n^{*2}$ we take only pairs in the same blocks, whereas for $\hat{\sigma}_n^2$ we take all pairs at lag distance $< l$ with a lag weight w_k. The reason for this difference is that for H linear, $\mathbb{E}^*[\hat{\sigma}_n^{*2}] = \sigma_n^{*2}(1 - b^{-1})$, whereas the analogue of (5) with the original observations would have a bias of $\mathcal{O}(l^{-1})$. Such a large bias would destroy second-order correctness; see the discussion at the end of this section or Davison and Hall (1993).

We now derive formal Edgeworth expansions for $U_{n,\text{stud}}$ and $U_{n,\text{stud}}^*$. First we approximate $U_{n,\text{stud}}$ and $U_{n,\text{stud}}^*$ by quadratic statistics. A Taylor expansion of the numerator and denominator of $U_{n,\text{stud}}$ gives

$$n^{1/2}(U_n - H(\mu)) = T_n + n^{-1/2} \tfrac{1}{2} S_n^T D^2 H(\mu) S_n + n^{-1} \xi_{n,1}$$

and

$$\hat{\sigma}_n^2 = \sigma_n^2 + n^{-1/2}(V_n + S_n^T d) + (\tau_n^2 - \sigma_n^2) + n^{-1} \xi_{n,2},$$

where

$$\tau_n^2 \triangleq \sum_{k=0}^{l} w_k \mathbb{E}[Z_0 Z_k],$$

$$V_n \triangleq \sum_{k=0}^{l} w_k n^{-1/2} \sum_{j=1}^{n-l} (Z_j Z_{j+k} - \mathbb{E}[Z_0 Z_k]),$$

$$d \triangleq 2D^2 H(\mu) \sum_{k=-\infty}^{\infty} \mathbb{E}\left[(Y_0 - \mu)(Y_k - \mu)^T\right] DH(\mu).$$

Hence

$$\begin{aligned} U_{n,\text{stud}} &= T_n \sigma_n^{-1} + n^{-1/2}(P(S_n) - \tfrac{1}{2} T_n V_n \sigma_n^{-3}) \\ &\quad - \tfrac{1}{2} T_n (\tau_n^2 - \sigma_n^2) \sigma_n^{-3} + n^{-1} \xi_{n,3} \\ &\triangleq E_n - \tfrac{1}{2} T_n (\tau_n^2 - \sigma_n^2) \sigma_n^{-3} + n^{-1} \xi_{n,3} \quad \text{say,} \end{aligned} \tag{6}$$

where

$$P(S_n) \triangleq \tfrac{1}{2} S_n^T D^2 H(\mu) S_n \sigma_n^{-1} - \tfrac{1}{2} T_n S_n^T d \sigma_n^{-3}.$$

Because $S_n = \mathcal{O}_P(1)$ and $V_n = \mathcal{O}_P(l^{1/2})$, the remainder terms are $\xi_{n,1} = \mathcal{O}_P(1)$, $\xi_{n,2} = \mathcal{O}_P(l)$ and $\xi_{n,3} = \mathcal{O}_P(l)$.

The difference $\tau_n^2 - \sigma_n^2$ reflects the bias of $\hat{\sigma}_n^2$ as an estimator of σ_n^2 when H is linear. Its order depends on the smoothness at zero of the function ω generating the weights w_k and on the decay of the covariances $\mathbb{E}[Z_0 Z_k]$. Under the assumptions of Section 3 the covariances decay exponentially. Thus if $\omega(x) - 1 \sim -cx^j$ as $x \to 0$ with $c > 0$, then $\tau_n^2 - \sigma_n^2$ is $O(l^{-j})$. If $\omega \equiv 1$ and $\log(n) = o(l)$, then $\tau_n^2 - \sigma_n^2$ is $\mathscr{O}(n^{-1})$. The implications of this for the error in the bootstrap approximation will be considered at the end of this section.

Similarly we have

$$\begin{aligned} U_{n,\mathrm{stud}}^* &= T_n^* \sigma_n^{*-1} + n^{-1/2}\left(P^*(S_n^*) - \tfrac{1}{2} l^{1/2} T_n^* V_n^* \sigma_n^{*-3}\right) + b^{-1}\xi_{n,3}^* \\ &\triangleq E_n^* + b^{-1}\xi_{n,3}^* \quad \text{say,} \end{aligned}$$

where

$$\begin{aligned} P^*(S_n^*) &\triangleq \tfrac{1}{2} S_n^{*T} D^2 H(\mu_n^*) S_n^* \sigma_n^{*-1} - \tfrac{1}{2} T_n^* S_n^{*T} d^* \sigma_n^{*-3}, \\ d^* &\triangleq 2D^2 H(\mu_n^*) \mathbb{E}^*\left[A_{N_1} A_{N_1}^T\right] DH(\mu_n^*), \\ V_n^* &\triangleq b^{-1/2} \sum_{j=1}^{b} \left(B_{N_j}^2 - \sigma_n^{*2}\right). \end{aligned}$$

The remainder term $\xi_{n,3}^*$ is $\mathscr{O}_P^*(1)$.

For the formal Edgeworth expansion of $U_{n,\mathrm{stud}}$ we introduce

$$\kappa_n \triangleq n^{1/2}\mathbb{E}\left[T_n^3\right], \qquad \alpha_n \triangleq \mathbb{E}[T_n V_n] = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n-l} \sum_{k=0}^{l} w_k \mathbb{E}\left[Z_i Z_j Z_{j+k}\right].$$

Furthermore let Σ_n denote the covariance matrix of the $(k+1)$-dimensional vector $(T_n/\sigma_n, S_n^T)^T$ and let c_α denote the coefficients of the polynomial $P(S_n)$ with respect to the variables S_n and T_n. The formal Edgeworth expansion $\Psi_n(a)$ of $U_{n,\mathrm{stud}}$ is the same as the one for E_n defined in (6). It is defined by its characteristic function

$$\begin{aligned} \tilde{\Psi}_n(t) &= \left(1 + n^{-1/2}\sigma_n^{-3}\left(\left(\tfrac{1}{6}\kappa_n - \tfrac{1}{2}\alpha_n\right)(it)^3 - \tfrac{1}{2}\alpha_n it\right)\right)\exp\left[-t^2/2\right] \\ &\quad + n^{-1/2} it \sum_{\alpha} c_\alpha (-1)^{|\alpha|} D_w^\alpha \exp\left[-w^T \Sigma_n w/2\right]\big|_{w=(t,0,\dots,0)}. \end{aligned} \tag{7}$$

Similarly the formal Edgeworth expansion of $U_{n,\mathrm{stud}}^*$ (conditional on the observations) is defined by its characteristic function

$$\begin{aligned} \tilde{\Psi}_n^*(t) &= \left(1 + n^{-1/2}\kappa_n^* \sigma_n^{*-3}\left(-\tfrac{1}{3}(it)^3 - \tfrac{1}{2}it\right)\right)\exp\left[-t^2/2\right] \\ &\quad + n^{-1/2} it \sum_{\alpha} c_\alpha^* (-1)^{|\alpha|} D_w^\alpha \exp\left[-w^T \Sigma_n^* w/2\right]\big|_{w=(t,0,\dots,0)}. \end{aligned}$$

Here

$$\kappa_n^* \triangleq n^{1/2}\mathbb{E}^*\left[T_n^{*3}\right] = l^{1/2}\mathbb{E}^*[T_n^* V_n^*] = l^{1/2}\mathbb{E}^*\left[B_{N_1}^3\right]$$

and Σ_n^* and c_α^* are defined in analogy to Σ_n and c_α. Hence the two formal Edgeworth expansions are close if all coefficients are close, that is, the moments up to order 3 should be close. From Künsch (1989) and the foregoing discussion it follows that

$$\begin{aligned} \mu_n^* &= \mu + \mathscr{O}_P(n^{-1/2}), \\ \Sigma_n &= \Sigma_\infty + \mathscr{O}(n^{-1}), \\ \Sigma_n^* &= \Sigma_\infty + \mathscr{O}(l^{-1}) + \mathscr{O}_P(b^{-1/2}). \end{aligned}$$

Moreover it is easily seen that

$$\alpha_n = \kappa_n + \mathscr{O}(n^{-1}) = \kappa_\infty + \mathscr{O}(n^{-1}), \qquad \kappa_\infty \triangleq \sum_{i,j=-\infty}^{\infty} \mathbb{E}[Z_0 Z_i Z_j].$$

Finally

$$\kappa_n^* = l^{1/2}(n-l+1)^{-1} \sum_{j=0}^{n-l} B_j^3.$$

Hence because B_j is a standardized sum of length l we have

$$\mathbb{E}[\kappa_n^*] = \kappa_\infty + \mathscr{O}(l^{-1}).$$

In Section 4 we will show that

$$\kappa_n^* - \mathbb{E}[\kappa_n^*] = \mathscr{O}_P(l^{1/2} b^{-1/2}) = \mathscr{O}_P(l n^{-1/2}).$$

Taking all this together we obtain the order of the difference between the two Edgeworth expansions:

$$\sup_a |\Psi_n(a) - \Psi_n^*(a)| = \mathscr{O}(l^{-1} n^{-1/2}) + \mathscr{O}_P(l n^{-1}). \tag{8}$$

The error between the Edgeworth expansions and the distribution functions will be shown to be

$$\sup_a |\mathbb{P}\{U_{n,\mathrm{stud}} \le a\} - \Psi_n(a)| = \mathscr{O}(l n^{-1+\varepsilon}) + \mathscr{O}(\tau_n^2 - \sigma_n^2) \tag{9}$$

and

$$\sup_a |\mathbb{P}^*\{U_{n,\mathrm{stud}}^* \le a\} - \Psi_n^*(a)| = \mathscr{O}_P(l n^{-1+\varepsilon}) \tag{10}$$

for any $\varepsilon > 0$ provided all moments exist. Together (8)–(10) imply the desired result that the studentized bootstrap is second-order correct.

We close this section with a brief discussion of the choice of l and ω. First we note that when $\tau_n^2 - \sigma_n^2 = \mathscr{O}(l^{-1})$, some of the error terms in (8)–(10) are of larger order than $\mathscr{O}(n^{-1/2})$ for any choice of l, that is, second-order correctness does not hold. In particular we must not use the triangular weights $\omega(x) = 1 - x$ even though they are suggested by the form of the bootstrap variance. As we have seen, the order of $\tau_n^2 - \sigma_n^2$ is minimal for the rectangular weights $\omega \equiv 1$. In this case the optimal order of l is seen to be $\mathscr{O}(n^{1/4})$, which leads to an error in the bootstrap approximation of

$\mathcal{O}_P(n^{-3/4+\varepsilon})$. The only disadvantage of the rectangular weights is that $\hat{\sigma}_n^2$ is not guaranteed to be positive. Positivity can be achieved with weights $\omega(x) - 1 \sim cx^2$ with $c > 0$. Then the optimal order of l becomes $\mathcal{O}(n^{1/3})$ with an error of $\mathcal{O}_P(n^{-2/3+\varepsilon})$ in the bootstrap approximation.

3. Confidence intervals. We consider the problem of constructing a one-sided confidence interval $(-\infty, \overline{U}_n)$ for $H(\mu)$ with level α. One possibility is to take

$$\overline{U}_n = U_n + n^{-1/2}\hat{\sigma}_n G_n^{*-1}(\alpha), \tag{11}$$

where $G_n^*(a) = \mathbb{P}^*\{U_{n,\text{stud}}^* \le a\}$. Our results imply that this interval has actual coverage probability $\alpha + o(n^{-1/2})$, but it is not equivariant under monotone transformations of the parameter. The following generalization of Efron's (1987) BC_a interval to the dependent case has both properties. We use the bootstrap distribution $G_{n,0}^*(a) = \mathbb{P}^*\{U_n^* \le a\}$ without any standardization or studentization, but we adjust the level to $\beta = \beta(\alpha)$:

$$\overline{U}_n \triangleq G_{n,0}^{*-1}(\beta), \qquad \beta(\alpha) \triangleq \Phi\big(z_{1-\alpha} + a_n z_{1-\alpha}^2 + b_n z_{1-\alpha} + c_n + d_n\big). \tag{12}$$

Here $z_\alpha \triangleq \Phi^{-1}(\alpha)$ and, with the quantities defined in Section 2,

$$\begin{aligned} a_n &\triangleq n^{-1/2}\kappa_n^*/\big(6\sigma_n^{*3}\big),\\ b_n &\triangleq \hat{\sigma}_n/\sigma_n^* - 1,\\ c_n &\triangleq 2\Phi^{-1}\big(G_{n,0}^*(H(\mu_n^*))\big),\\ d_n &\triangleq n^{1/2}(U_n - H(\mu_n^*))/\sigma_n^*. \end{aligned}$$

The expressions a_n and c_n are invariant under monotone transformations, but b_n and d_n are not. However, since terms of $o_P(n^{-1/2})$ do not matter, we can replace $U_n - H(\mu_n^*)$ in d_n by $DH(\mu_n^*)^T(\overline{Y}_n - \mu_n^*)$. Similarly, we can modify b_n by replacing $DH(\overline{Y}_n)$ in definition (3) of $\hat{\sigma}_n^2$ by $DH(\mu_n^*)$. With these modifications transformation equivariance holds, and the coverage probability follows from Theorem 3.1.

THEOREM 3.1. *Under the conditions of Theorem* 4.1, $\mathbb{P}\{\overline{U}_n \le H(\mu)\} = \alpha + o(n^{-1/2})$.

PROOF. Set $\overline{U}_n' \triangleq U_n - n^{-1/2}\hat{\sigma}_n\Psi_n^{-1}(\alpha)$. Then by Theorem 4.1, $\mathbb{P}\{\overline{U}_n' \le H(\mu)\} = \alpha + o(n^{-1/2})$. Thus it is sufficient to show that

$$\begin{aligned} G_{n,0}^*\big(\overline{U}_n'\big) &= \mathbb{P}^*\big\{n^{1/2}\sigma_n^{*-1}(U_n^* - H(\mu_n^*)) \le d_n - \Psi_n^{-1}(\alpha) - b_n\Psi_n^{-1}(\alpha)\big\}\\ &= \beta + o_P(n^{-1/2}). \end{aligned}$$

By Lahiri's (1991) result on second-order correctness for the standardized statistic we may replace the bootstrap probability by the Edgeworth expansion Ξ_n of $n^{1/2}\sigma_n^{-1}(U_n - H(\mu))$. Thus

$$G_{n,0}^*\big(\overline{U}_n'\big) = \Xi_n\big(d_n - \Psi_n^{-1}(\alpha) - b_n\Psi_n^{-1}(\alpha)\big) + o_P(n^{-1/2}).$$

Now we have to exploit the specific form of Ψ_n and Ξ_n. From (7) it follows that there are coefficients γ_1, γ_2 such that

$$\Psi_n(z) = \Phi(z) + n^{-1/2}\varphi(z)(\gamma_1 + \gamma_2 z^2),$$

whence

$$\Psi_n^{-1}(\alpha) = z_\alpha - n^{-1/2}(\gamma_1 + \gamma_2 z_\alpha^2) + o(n^{-1/2}).$$

Similarly there are coefficients γ_3, γ_4 such that

$$\begin{aligned}\Xi_n(z) &= \Phi(z) + n^{-1/2}\varphi(z)(\gamma_3 + \gamma_4 z^2)\\ &= \Phi(z + n^{-1/2}(\gamma_3 + \gamma_4 z^2)) + o(n^{-1/2}).\end{aligned}$$

Hence

$$\begin{aligned}G_{n,0}^*(\overline{U}_n') &= \Phi(z_{1-\alpha} + n^{-1/2}(\gamma_1 + \gamma_3 + (\gamma_2 + \gamma_4)z_{1-\alpha}^2) + b_n z_{1-\alpha} + d_n)\\ &\quad + o_P(n^{-1/2}).\end{aligned}$$

Because $c_n = 2\Phi^{-1}(\Xi_n(0)) + o_P(n^{-1/2}) = 2n^{-1/2}\gamma_3 + o_P(n^{-1/2})$, the claim of the theorem is now equivalent to $\gamma_1 = \gamma_3$ and $\gamma_2 + \gamma_4 = a_n$. This follows from a long but straightforward computation, which we omit. □

Let us briefly comment on the differences between our formulae and those in Efron (1987). Our c_n is twice the bias correction z_0 of Efron. For the statistic $U_n = H(\overline{Y}_n)$ the empirical influence function is simply $DH(\overline{Y}_n)(Y_j - \overline{Y}_n) \sim B_j$, so our a_n is the analogue of Efron's acceleration constant a. Of course, since our formula takes dependence into account, the two expressions are different. The two terms $b_n z_{1-\alpha}$ and d_n do not occur in the independent case. They are needed to correct for differences between the real and the bootstrap world. Finally the computation of the adjusted level $\beta(\alpha)$ is not exactly the same, but Efron's formula (3.9) is asymptotically equivalent to $z[\alpha] = 2z_0 + z_\alpha + az_\alpha^2$, which agrees with (12).

4. Rigorous results and proofs.

4.1. *Edgeworth expansion for the studentized statistic.* We shall assume that the sequence of random vectors $R_j \triangleq (Y_j, Z_j) \in \mathbb{R}^k \times \mathbb{R}$ satisfies the following conditions used in Götze and Hipp (1983), which subsequently will be denoted by GH.

(A1) $\mathbb{E}Y_j = 0$, $j = 1, 2, \dots$.

(A2) $\beta_s \triangleq \mathbb{E}\|Y_j\|^{s+\delta} < \infty$ for some integer $s \geq 8$, and $\delta > 0$ arbitrary small.

(A3) There exists a sequence $\mathscr{D}_k$, $k \in \mathbb{Z}$, of sub-σ fields of $\mathscr{A}$ and a constant $d > 0$ such that for $j, m = 1, 2, \dots$, with $m > d^{-1}$, the r.v. Y_j can be approximated by a $\mathscr{D}_{j-m,j+m} \triangleq \sigma(\mathscr{D}_p : |p - j| \leq m)$-measurable random vector $\overline{Y}_{j,m}$ with

$$\mathbb{E}\|Y_j - \overline{Y}_{j,m}\| \leq d^{-1}\exp[-dm].$$

(A4) There exists a $d > 0$ such that for all $m, j = 1, 2, \dots, A \in \mathscr{D}_{-\infty, j}$, $B \in \mathscr{D}_{j+m, \infty}$,

$$|\mathbb{P}\{A \cap B\} - \mathbb{P}\{A\}\mathbb{P}\{B\}| \le d^{-1} \exp[-dm]$$

(Rosenblatt mixing).

(A5) There exists a $d > 0$ such that for all $m, j = 1, 2, \dots, d^{-1} < m < j$ and $|t| \ge d$,

$$\mathbb{E}|\mathbb{E}\big(\exp\big[it(Z_{j-m} + \cdots + Z_{j+m})\big] | \mathscr{D}_l : l \ne j\big)| \le \exp[-d]$$

and

$$\liminf_n \operatorname{Var}(Z_1 + \cdots + Z_n)/n > 0.$$

(A6) There exists a $d > 0$ such that for all $m, j, p = 1, 2, \dots$ and $A \in \mathscr{D}_{j-p, j+p}$,

$$\mathbb{E}|\mathbb{P}\{A|\mathscr{D}_l : l \ne j\} - \mathbb{P}\{A|\mathscr{D}_l : 0 < |l - j| \le m + p\}| \le d^{-1} \exp[-dm].$$

Moreover for the function H which defines our statistic U_n we make the following assumption.

(A7) $H: \mathbb{R}^k \to \mathbb{R}$ is three times differentiable, $DH(\mu) \ne 0$ and there are constants $C, A > 0$ such that

$$\|D^3 H(x)\| \le C(1 + \|x\|^A) \quad \text{for every } x \in \mathbb{R}^k.$$

Condition (A1) means no restriction since we can always put $\mu = 0$. Moreover, since Z_j is a linear function of Y_j, (A1)–(A3) hold also for Z_j (after adjusting the constants if necessary). Our main result is Theorem 4.1:

THEOREM 4.1. *Under the conditions* (*A*1)–(*A*7) *the Edgeworth approximation for* $U_{n,\text{stud}}$ *defined in* (7) *holds, that is, for* $s \ge 8$, $l \le n^{1/3}$ *and* $\log(n) = o(l)$ *we have*

$$\sup_a |\mathbb{P}\{U_{n,\text{stud}} \le a\} - \Psi_n(a)| = \mathscr{O}(ln^{-1+2/s})$$

provided $\omega \equiv 1$. *If* $\omega(x) - 1 \sim -cx^2$ *with* $c > 0$, *the error contains an additional* $\mathscr{O}(l^{-2})$ *term.*

REMARK 4.1. A direct application of the multivariate result in GH for a stochastic expansion in terms of the vector (S_n, V_n) is not very desirable, although it does yield expansions up to arbitrary degree. However, we would have to check the conditional Cramér condition (A5) for (S_n, V_n), which can be very difficult even for simple time series models, in particular when the statistic U_n is constructed from transformed observations as in (1). For examples where (A5) is verified, see Bose (1988) and Götze and Hipp (1994).

REMARK 4.2. This result also can be generalized for the case of random fields Z_{ij}, $ij \in \mathbb{Z}^2$. In order not to exclude interesting examples, the mixing condition (A4) has to be relaxed as follows: For finite subsets I_1, I_2 of $\mathbb{Z}^2$ whose Euclidean distance is greater than or equal to m and for $A \in \sigma(I_1)$, $B \in \sigma(I_2)$,

$$|\mathbb{P}\{A \cap B\} - \mathbb{P}\{A\}\mathbb{P}\{B\}| \le c|I_1|^{\delta_1}|I_2|^{\delta_2}e^{-dm}.$$

The proofs still carry over with this weaker condition; see Jensen (1993). Thus we have an analogous result for statistics $U_n \triangleq H(n^{-2}\sum_{i,j=1}^n Z_{ij})$.

PROOF OF THEOREM 4.1. In order to limit the moment conditions, we have to use truncations as in the proofs of GH. Define $T: \mathbb{R}^{k+1} \to \mathbb{R}^{k+1}$ by

$$T(x) \triangleq xn^{\beta}\chi(\|x\|n^{-\beta})/\|x\|,$$

where $\chi \in C^\infty(0,\infty)$ satisfies $\chi(r) \triangleq r$ for $r \le 1$, χ is increasing and $\chi(r) \equiv 2$ for $r \ge 2$. The value of $\frac{1}{2} > \beta > 0$ will be determined later. Define $R_j^\dagger \triangleq (Z_j^\dagger, Y_j^\dagger) \triangleq T(R_j)$. Let $U_{n,\mathrm{stud}}^\dagger$ denote the statistic $U_{n,\mathrm{stud}}$ applied to $R_j^\dagger$, $j = 1,\dots,n$. Note that

$$0 = \mathbb{E}R_j = \mathbb{E}R_j 1_{\{\|R_j\| < n^\beta\}} + \mathbb{E}R_j 1_{\{\|R_j\| \ge n^\beta\}} = J_1 + J_2,$$

say, where $\|J_2\| = \mathscr{O}(n^{-(s-1)\beta})$ by Chebyshev's inequality. Hence we obtain

$$\begin{aligned} &\|\mathbb{E}R_j^\dagger\| = \mathscr{O}(n^{-(s-1)\beta}),\\ (13)\qquad &\|\mathbb{E}S_n^\dagger\| = \mathscr{O}(n^{1/2-(s-1)\beta}) = \mathscr{O}(n^{-1-\varepsilon}),\\ &\|\mathbb{E}V_n^\dagger\| = \mathscr{O}(ln^{1/2-(s-2)\beta}) = \mathscr{O}(n^{-1-\varepsilon}), \end{aligned}$$

choosing β such that

$$(14)\qquad (s-2)\beta > 11/6 + \varepsilon, \qquad s\beta > 2 + \varepsilon.$$

Then we have $\mathbb{P}\{\|R_j\| > n^\beta \text{ for some } 1 \le j \le n\} = \mathscr{O}(n^{1-\beta s}) = o(n^{-1})$. These estimates together with the arguments in Lemma 3.30 of GH finally lead to

$$|\kappa_p(a^T S_n) - \kappa_p(a^T S_n^\dagger)| = \mathscr{O}(n^{1-p/2-(s-p)\beta})$$

for the cumulants of order $p = 2, 3, \dots, s$. By expansion of moments in terms of cumulants we get [similar to page 224 and to Lemma 3.18 of GH or Bulinskii and Zhurbenko (1976)]

$$(15)\qquad \sup_n \mathbb{E}\|S_n\|^{s+\delta} < \infty, \qquad \sup_n \mathbb{E}\|S_n^\dagger\|^p \le c(p)$$

for every $p \ge 2$. Furthermore,

$$(16)\qquad \mathbb{E}|V_n|^{s/2} = \mathscr{O}(l^{s/4}) \quad \text{and} \quad \mathbb{E}|V_n^\dagger|^p = \mathscr{O}(l^{p/2})$$

for any p. The bounds in (16) can be shown as follows. Decompose $V_n = \sum_{j=1}^{n/l} W_j(l/n)^{1/2}$, where $W_j \triangleq \sum^* (Z_\nu Z_{\nu+k} - \mathbb{E}Z_\nu Z_{\nu+k})w_k l^{-1/2}$ and the summation extends over all ν, k such that $jl \le \nu < (j+1)l$, $0 \le k \le l$. The sum over j may be decomposed into three sums (with indices $j = 0, 1, 2 \bmod 3$) which consist of almost independent summands. This implies by standard estimates $\mathbb{E}|V_n|^q = \mathscr{O}(\mathbb{E}|W_0|^q)$. Finally we may replace $|x|^q$ by a function ϕ of class C^∞ with $|\phi^{(p)}(x)| \le c(1 + |x|^{q-p})$, $p \le q$, and obtain by induction on q, $|\mathbb{E}\,\phi(W_0 l^{-1/2})| = \mathscr{O}(1)$ uniformly in this class. Here we use Taylor expansion, $\mathbb{E}W_0^2 = \mathscr{O}(l)$ and Tikhomirov's (1980) gap expansion arguments.

Thus we obtain for the Edgeworth expansion $\Psi_n^\dagger(a)$ for $U_{n,\mathrm{stud}}^\dagger$ defined similarly as $\Psi_n(a)$ by standard technical arguments [see Bhattacharya and Rao (1986)] $\sup_a|\Psi_n(a) - \Psi_n^\dagger(a)| = \mathscr{O}(n^{-1})$ and, therefore,

$$\begin{aligned} \Delta_n &\triangleq \sup_a |\mathbb{P}\{U_{n,\mathrm{stud}} \le a\} - \Psi_n(a)| \\ &\le \sup_a |\mathbb{P}\{U_{n,\mathrm{stud}}^\dagger \le a\} - \Psi_n^\dagger(a)| + \mathscr{O}(n^{-1}). \end{aligned} \tag{17}$$

Next we argue that the Edgeworth expansions for $U_{n,\mathrm{stud}}^\dagger$ and for $E_{n,\tau}^\dagger \triangleq E_n^\dagger + T_n^\dagger\sigma_n^{-1}(\tau_n^2 - \sigma_n^2)\sigma_n^{-2}$ [see (6)] agree up to $\mathscr{O}(ln^{-1+\varepsilon})$, that is, we can neglect the remainder term $n^{-1}\xi_{n,3}$ in the Taylor expansion (6). By standard arguments [see, e.g., Chibisov (1972)] this is true if $\mathbb{P}\{|\xi_{n,3}^\dagger| > ln^\varepsilon\} = \mathscr{O}(ln^{-1+\varepsilon})$. This follows from the structure of the remainder terms in the Taylor expansions and Chebychev's inequality, since $S_n^\dagger$ and $V_n^\dagger$ have moments of all orders with bounds given by (15) and (16). Finally, since the Edgeworth expansion of $E_n^\dagger$ and of $E_{n,\tau}^\dagger$, say $\Psi_{n,\tau}^\dagger$, differ by $\mathscr{O}(\tau_n^2 - \sigma_n^2)$, we obtain

$$\Delta_n \le \sup_a |\mathbb{P}\{E_{n,\tau}^\dagger \le a\} - \Psi_{n,\tau}^\dagger(a)| + \mathscr{O}(\tau_n^2 - \sigma_n^2) + \mathscr{O}(ln^{-1+\varepsilon}). \tag{18}$$

In order to simplify the notation we shall replace, in the following, $E_{n,\tau}^\dagger$ by $E_n^\dagger$ and $\Psi_{n,\tau}^\dagger$ by $\Psi_n^\dagger$, since both statistics behave exactly similarly in terms of estimating their c.f. By the Berry–Esseen lemma we can estimate the error in the Edgeworth expansion by characteristic functions. Let $\varphi_n^\dagger(t) \triangleq \mathbb{E}\exp[itE_n^\dagger]$. Then (16) together with (18) leads to

$$\begin{aligned} \Delta_n &\le c\left(\int_{|t|\le n^\varepsilon} + \int_{n^\varepsilon<|t|<n^{1-2/s}}\right)|\varphi_n^\dagger(t) - \tilde{\Psi}_n^\dagger(t)||t|^{-1}\,dt + \mathscr{O}(ln^{-1+\varepsilon}) \\ &= I_1 + I_2 + \mathscr{O}(ln^{-1+\varepsilon}), \quad \text{say}. \end{aligned} \tag{19}$$

Estimation of I_1. By Lemma 3.33 of GH we have

$$|\mathbb{E}\exp[itT_n^\dagger/\sigma_n] - \tilde{\Psi}_{n,\mathrm{lin}}^\dagger(t)| \le cn^{-1}(1 + \beta_s)(1 + |t|^{12})\exp[-ct^2] \tag{20}$$

for some absolute constant $c > 0$, where

$$\tilde{\Psi}_{n,\mathrm{lin}}^\dagger \triangleq \exp\left[-\frac{1}{2}t^2\right]\left(1 + n^{-1/2}(it)^3\frac{\kappa_3^\dagger}{6\sigma_n^3}\right).$$

Now expanding $\mathbb{E}\exp[itE_n^\dagger]$ in terms of $tn^{-1/2}(P(S_n^\dagger) - T_n^\dagger V_n^\dagger/2)$ we obtain

$$\begin{aligned}\mathbb{E}&\exp\big[itE_n^\dagger\big]\\ &= \mathbb{E}\exp\big[itT_n^\dagger/\sigma_n\big] + itn^{-1/2}\mathbb{E}\exp\big[itT_n^\dagger/\sigma_n\big]\big(P(S_n^\dagger) - \tfrac{1}{2}T_n^\dagger V_n^\dagger\big)\sigma_n^{-3}\\ &\quad + \mathscr{O}(t^2n^{-1})\Big(\mathbb{E}P\big(S_n^\dagger\big)^2 + \mathbb{E}\big(T_n^\dagger V_n^\dagger\big)^2\Big)\\ &\triangleq \mathbb{E}\exp\big[itT_n^\dagger/\sigma_n\big] + J_{1,t} + J_{2,t} + R_t, \quad \text{say.}\end{aligned} \tag{21}$$

By (15) we obtain $\sup_n \mathbb{E}P(S_n^\dagger)^2 \le c < \infty$, and expanding $\mathbb{E}T_n^{\dagger 2}V_n^{\dagger 2}$ into individual summands we obtain, by conditions (A3) and (A4) similarly to Bulinskii and Zhurbenko (1976), by counting multiplicities,

$$\mathbb{E}T_n^{\dagger 2}V_n^{\dagger 2} = \mathscr{O}(l). \tag{22}$$

Furthermore, let $T_I^\dagger \triangleq \sum_{j\in I} Z_j^\dagger n^{-1/2}/\sigma_n$, $T_{n,I}^\dagger = T_n^\dagger/\sigma_n - T_I^\dagger$ and $H_{p,\nu,j}^\dagger \triangleq \frac{1}{2}\sigma_n^{-3}w_\nu(Z_p^\dagger Z_{p+\nu}^\dagger - \mathbb{E}Z_pZ_{p+\nu})Z_j^\dagger$. We have

$$J_{2,t} = -itn^{-3/2}\sigma_n^{-3}\sum_{p,\nu,j}{}^{*}\,\mathbb{E}H_{p,\nu,j}^\dagger\exp\big[itT_I^\dagger\big]\exp\big[itT_{n,I}^\dagger\big],$$

where $\sum^*$ denotes the summation over $1 \le j \le n$, $1 \le p \le n - l$ and $1 \le \nu \le l$ and where I denotes the set of indices in $\langle 1, n\rangle$ which have distance less than m from $\{j, p, p+\nu\}$ with $m \triangleq K\log n$, for some K to be chosen later. Using an expansion technique introduced by Tikhomirov (1980) we obtain, by expanding $\exp[itT_I^\dagger]$ up to second order,

$$\begin{aligned}J_{2,t} &= -itn^{-3/2}\sum_{p,\nu,j}{}^{*}\,\mathbb{E}H_{p,\nu,j}^\dagger\left(1 + itT_I^\dagger + \frac{(it)^2}{2}T_I^{\dagger 2}\right)\exp\big[itT_{n,I}^\dagger\big]\\ &\quad + \mathscr{O}\left(t^4n^{-3/2}\sum_{p,\nu,j}{}^{*}\,\mathbb{E}\big|H_{p,\nu,j}^\dagger T_I^{\dagger 3}\big|\right)\\ &\triangleq J_{3,t} + J_{4,t} + J_{5,t} + R_{6,t}.\end{aligned}$$

In the term $R_{6,t}$ we have to estimate ln^2m^3 summands separately. Thus $R_{6,t} = \mathscr{O}(lm^3t^4n^{-1})$. For the other terms we shall employ the approximate independence of, for example, $H_{p,\nu,j}^\dagger$ and $\exp[itT_{n,I}^\dagger]$ using conditions (A3) and (A4) in order to factor $J_{3,t}$ as a product of expectations plus an error of $\mathscr{O}(n^C\exp[-dm])$.

The sum $J_{4,t}$ is split according to the case where $|j-p| \le 2m$ or $|j-(p+\nu)| \le 2m$ and the case where both distances are larger than $2m$. In the first case, we have $\mathscr{O}(m^2nl)$ summands which yield a contribution of $\mathscr{O}(lm^2nt^2n^{-2})$. In the second case, we may use the approximate independence of $W_{p,\nu}^\dagger \triangleq Z_p^\dagger Z_{p+\nu}^\dagger - \mathbb{E}Z_pZ_{p+\nu}$ and $Z_j^\dagger Z_{j+\mu}^\dagger\exp[itT_{n,I}^\dagger]$ as well as of $W_{p,\nu}^\dagger Z_{p+\mu}^\dagger$ (resp. $W_{p,\nu}^\dagger Z_{p+\nu+\mu}^\dagger$) for $|\mu| \le m$ and the other factors. Thus $J_{4,t} = \mathscr{O}(lmt^2n^{-1})$. In $J_{5,t}$ the same splitting yields a bound of $\mathscr{O}(lm^3n|t|^3n^{-3/2})$ for the first part and in the second part negligible contributions from terms of type $W_{p,\nu}^\dagger Z_j^\dagger Z_{j+\mu_1}^\dagger Z_{j+\mu_2}^\dagger$. In this part, sums (over j, p) of factorizing expectations of the type $\mathbb{E}W^\dagger_{p,\nu}Z^\dagger_{p+\mu_1}\mathbb{E}Z^\dagger_j Z^\dagger_{j+\mu_2}$ contribute to the expansion. Collecting these terms and the error bounds we get

$$\begin{aligned} J_{2,t} = & -\tfrac{1}{2}n^{-1/2}\mathbb{E}T_n^\dagger V_n^\dagger \sigma_n^{-3}\big(it + (it)^3\big)\mathbb{E}\exp\big[itT_n^\dagger\big] \\ & + \mathscr{O}\Bigg(tn^{-3/2}\sum_{p,\nu,j,\mu_1,\mu_2} G^\dagger_{p,\nu,j,\mu_1,\mu_2}\mathbb{E}\Big|\Big(\mathbb{E}\exp\big[itT_n^\dagger\big] - \mathbb{E}\exp\big[itT^\dagger_{n,I}\big]\Big)\Big|\Bigg) \\ & + \mathscr{O}\bigg(\frac{t^4}{n}m^3 l\bigg), \end{aligned} \tag{23}$$

where $G^\dagger_{p,\nu,j,\mu_1,\mu_2} \triangleq H^\dagger_{p,\nu,j}(1 + t^2n^{-1}Z^\dagger_{j+\nu_1}Z^\dagger_{p+\nu_2})$ and the summation extends over $|\nu| \le l$, $|\mu_1|$, $|\mu_2| \le m$. Finally the second line in (23), say $J_{7,t}$, can be handled easily by estimating $\exp[itT_I^\dagger] - 1$ using (15) and (16). This leads to $J_{7,t} = \mathscr{O}(t^2n^{-1}l^{1/2} + lm^3t^4n^{-1})$. Furthermore, in order to estimate $J_{1,t}$ we obtain by Lemma 3.33 of GH for the multivariate expansion of $\mathbb{E}\exp[ia^TS_n^\dagger/\sigma_n]$ in terms of the multivariate c.f. expansion $\tilde{\Psi}_{n,\mathrm{mult}}(a)$ of length 2 with first component of a being equal to t,

$$\begin{aligned} & \Big|\mathbb{E}P(S_n^\dagger)\exp\big[itT_n^\dagger/\sigma_n\big] - \sum_\alpha c_\alpha D^\alpha\tilde{\Psi}_{n,\mathrm{mult}}(a)\Big| \\ & \quad \le \sum_\alpha |c_\alpha|\Big|D^\alpha\Big(\mathbb{E}\exp\big[ia^TS_n^\dagger/\sigma_n\big] - \tilde{\Psi}_{n,\mathrm{mult}}(a)\Big)\Big| \\ & \quad \le c\Big(1 + \sup_j \mathbb{E}\|R_j^\dagger\|^4\Big)\big(1 + |t|^{6+|\alpha|}\big)e^{-ct^2}n^{-1/2-\varepsilon}, \end{aligned} \tag{24}$$

where $c > 0$ denotes an absolute constant and the partial derivatives are taken at the point $(t, 0, 0, \dots, 0)$. The coefficients c_α correspond to the monomial coefficients of the polynomial $P(S_n^\dagger)$. Collecting the expansion results (20)–(24) and letting $\alpha_3^\dagger \triangleq \sum_\mu \sum_\nu w_\nu \mathbb{E}Z_1^\dagger Z^\dagger_{1+\nu}Z^\dagger_{1+\nu+\mu}$, we get by definition of the expansion $\tilde{\Psi}_n^\dagger(t)$,

$$\begin{aligned} & \big|\varphi_n^\dagger(t) - \tilde{\Psi}_n^\dagger(t)\big| \\ & \quad \le \Big|\mathbb{E}\exp\big[itT_n^\dagger/\sigma_n\big] - \tilde{\Psi}^\dagger_{n,\mathrm{lin}}(t)\Big| \\ & \quad\quad + n^{-1/2}|t|\Big|\mathbb{E}P(S_n^\dagger)\exp\big[itT_n^\dagger/\sigma_n\big] - \sum_\alpha c_\alpha D^\alpha\tilde{\Psi}_{n,\mathrm{mult}}(a)\Big| \\ & \quad\quad + \tfrac{1}{2}(|t| + |t|^3)n^{-1/2}|\alpha_3^\dagger|\Big|\mathbb{E}\exp\big[itT_n^\dagger/\sigma_n\big] - \exp\big[-\tfrac{1}{2}t^2\big]\Big| \\ & \quad\quad + \mathscr{O}(t^2mn^{-1}) + \mathscr{O}(t^2ln^{-1}). \end{aligned} \tag{25}$$

Hence we obtain by integration over $|t| \le n^\varepsilon$, $I_1 = \mathscr{O}(ln^{-1+2\varepsilon})$.

Estimation of I_2. Let N denote a number such that $n \ge N \ge m$, where $m \triangleq K\log n$ and (K is chosen sufficiently large below) such that for $n^\varepsilon < |t| < n^{1-\varepsilon}$,

$$N \triangleq \bigg[\bigg(\frac{n}{t^2} + 1\bigg)m^2\bigg]. \tag{26}$$

The reasons for this choice will become apparent later in the proof. Now we replace the random vector $R_j \triangleq (Z_j, Y_j)$ by some $\mathscr{D}_{j-m,j+m}$-measurable random vectors $R_j^{\dagger\dagger} = (Z_j^{\dagger\dagger}, Y_j^{\dagger\dagger})$ with

$$\mathbb{E}\|R_j - R_j^{\dagger\dagger}\| < d^{-1}\exp[-dm]$$

using condition (A3). Hence also

$$\mathbb{E}\|T(R_j) - T(R_j^{\dagger\dagger})\| < d^{-1}\exp[-dm].$$

Denote by $E_n^{\dagger\dagger}$ the statistic E_n for truncated and $\mathscr{D}_{j-m,j+m}$-measurable vectors $T(R_j^{\dagger\dagger})$. We obtain by the truncation of random variables

$$\begin{aligned}&|\mathbb{E}\exp\left[itE_n^{\dagger}\right] - \mathbb{E}\exp\left[itE_n^{\dagger\dagger}\right]|\\ &\quad\le n^A \sup_j \mathbb{E}\|T(R_j) - T(R_j^{\dagger\dagger})\| = \mathscr{O}(n^{-2})\end{aligned} \tag{27}$$

for some $A > 0$ provided K in the definition of m is chosen sufficiently large. Thus we may assume w.l.o.g. that all r.v. are truncated by n^{β} and are $\mathscr{D}_{j-m,j+m}$-measurable. In order to simplify the notation, we shall from now on drop the † notation and write $R_j \triangleq (Z_j, Y_j)$, S_n, T_n, V_n again. We decompose

$$(S_n, V_n) = \left(\overline{S}_N, \overline{V}_N\right) + \left(\overline{S}_{n-N}, \overline{V}_{n-N}\right),$$

where

$$\left(\overline{S}_N, \overline{V}_N\right) \triangleq \sum_{j=1}^{N} n^{-1/2}\left(Z_j, Y_j, \sum_{\nu=0}^{l} w_\nu(Z_j Z_{j+\nu} - \mathbb{E}Z_j Z_{j+\nu})\right)$$

and $(\overline{S}_{n-N}, \overline{V}_{n-N})$ denotes the complementary part of the sums.

Let $Q(S_n, V_n) \triangleq P(S_n) - \frac{1}{2}T_n V_n$. In the next step we modify an argument used in Lemma 3.43 of GH. We expand

$$Q(S_n, V_n) = Q\left(\overline{S}_{n-N}, \overline{V}_{n-N}\right) - \tfrac{1}{2}T_n\overline{V}_N + \sum_{j=1}^{r} \overline{S}_N^j Q_j\left(\overline{S}_{n-N}, \overline{V}_{n-N}\right),$$

where Q_j denotes an appropriate vector of polynomials and the dot indicates a multilinear form. Thus we have, by expansion of exponential terms depending on $\overline{V}_N$ and $\overline{S}_N$ in powers of $Z_j Z_{j+\nu}$ (resp. Z_j and Y_j),

$$\begin{aligned}|\mathbb{E}\exp[itE_n]| \le &\Bigg|E\exp[itT_n]\exp\left[itn^{-1/2}Q\left(\overline{S}_{n-N}, \overline{V}_{n-N}\right)\right]\\ &\times \sum_{\alpha,\beta}{}^{*} c_{\alpha,\beta} Z^\alpha Y^\beta Q_{\alpha\beta}\left(\overline{S}_{n-N}, \overline{V}_{n-N}\right)\Bigg|\\ &+\frac{1}{p!}\mathbb{E}|Q(S_n, V_n) - Q\left(\overline{S}_{n-N}, \overline{V}_{n-N}\right)|^p\left(|t|n^{-1/2}\right)^p,\end{aligned} \tag{28}$$

where the sum Σ^* extends over all tuples

$$\alpha \triangleq (\alpha_1, \dots, \alpha_{N+l}, 0, \dots, 0), \qquad \beta \triangleq (\beta_1, \beta_2, \dots, \beta_N, 0, \dots, 0)$$

with $|\alpha| + |\beta| \le r(p-1)$, r denotes the degree of P and $Q_{\alpha,\beta}$ denote polynomials of the vector $(\overline{S}_{n-N}, \overline{V}_{n-N})$.

Let us estimate first the last remainder term, say I in (28). By (15), (16) and (26) we have in extension of (16) by properties of truncated moments $\mathbb{E}|\overline{V}_N|^p = \mathscr{O}(l^{p/2}(N/n)^{s/2})$ and $\mathbb{E}|\overline{S}_N|^p = \mathscr{O}((N/n)^{s/2})$ for $p \geq s$. Thus

$$
\begin{aligned}
I &= \mathscr{O}\Big(l^{p/2}\big(|t|N^{1/2}/n\big)^p\Big) \\
&= \mathscr{O}\Big(\big(mn^{-1/2}l^{1/2}\big)^p\Big) \quad \text{for } |t| \leq \sqrt{n} \qquad (29)\\
&= \mathscr{O}\Big(\big(m|t|n^{-1}l^{1/2}\big)^p\Big) \quad \text{for } \sqrt{n} < |t| < n^{1-\varepsilon}l^{-1/2}.
\end{aligned}
$$

Thus $I = o(n^{-1})$ for $|t| \leq n^{1-\varepsilon}l^{-1/2}$ provided that

$$
p \geq s \quad \text{and} \quad s\varepsilon > 2. \qquad (30)
$$

In order to evaluate the expansion terms in (28) we may proceed as in the proof of Lemma 3.43 of GH (pages 233–235). Surrounding the at most $r(p-1)$ indices j, where $\alpha_j > 0$ or $\beta_j > 0$, and the block $\langle N+1,\dots,n\rangle$ by an "independence neighborhood" I of size $3m$, define $\{j_1^0,\dots,j_{r(p-1)}^0\} \triangleq \{j\colon \alpha_j > 0 \text{ or } \beta_j > 0\}$ and $I \triangleq \{j \in \{1,\dots,N-m\}\colon |j - j_\nu^0| \geq 3m,\ \nu = 1,\dots,r(p-1)\}$. Divide I into blocks $A_1, B_1,\dots,A_q, B_q$ as follows. Define $j_1,\dots,j_q$, $j_1 \triangleq \inf I$ and $j_{\nu+1} = \inf\{j \geq j_\nu + 7m\colon j \in I\}$. Let q denote the smallest integer for which the inf is undefined. Write

$$
\begin{aligned}
A_\nu &\triangleq \prod\Big\{\exp\big[itn^{-1/2}Z_j\big]\colon j \in I, |j - j_\nu| \leq m\Big\}, \\
B_\nu &\triangleq \prod\Big\{\exp\big[itn^{-1/2}Z_j\big]\colon j \in I, j_\nu + m + 1 \leq j \leq j_{\nu+1} - m - 1\Big\}, \qquad (31)\\
R &\triangleq Z^\alpha Y^\alpha \prod_{j \notin I}\exp\big[itZ_j n^{-1/2}\big]\exp\Big[itQ\big(\overline{S}_{n-N}, \overline{V}_{n-N}\big)\Big],
\end{aligned}
$$

where $\nu = 1,\dots,q-1$. Then

$$
\exp\Big[it\Big(T_n + n^{-1/2}Q\big(\overline{S}_{n-N}, \overline{V}_{n-N}\big)\Big)\Big] = \prod_1^q A_\nu B_\nu R,
$$

where $|A_\nu| \leq 1$, $|B_\nu| \leq 1$, $|R| \leq n^{\beta(p-1)r}$ and A_ν is $\mathscr{D}_{j_\nu - 2m, j_\nu + 2m}$-measurable, B_ν is $\mathscr{D}_{j_\nu+1, j_{\nu+1}-1}$-measurable and R is measurable with respect to $\{\mathscr{D}_l\colon \exists j \notin I$ with $|l - j| \leq m\}$. We have

$$
\begin{aligned}
&\left|\mathbb{E}R\prod_1^q A_\nu B_\nu - \mathbb{E}R\prod_1^q \mathbb{E}\big(A_\nu|\mathscr{D}_j\colon |j - j_\nu| \leq 3m\big)B_\nu\right| \\
&\quad \leq \sum_{s=1}^q \left|\mathbb{E}R\prod_{\nu=1}^{s-1} A_\nu B_\nu\Big(A_s - \mathbb{E}\big(A_s|\mathscr{D}_j\colon |j - j_s| \leq 3m\big)\Big)\right. \qquad (32)\\
&\qquad \left. \times \prod_{\mu=s+1}^q \mathbb{E}\big(A_\mu|\mathscr{D}_j\colon |j - j_\mu| \leq 3m\big)B_\mu\right|.
\end{aligned}
$$

Here we may replace A_s by $\mathbb{E}(A_s|\mathscr{D}_j\colon j \neq j_s)$ since the product of all remaining factors in (32) is measurable with respect to $\mathscr{D}_j\colon j \neq j_s$ (i.e., is constant with respect to this conditioning by our construction). Thus invoking condition (A6) we obtain that (32) can be bounded from above by $n^B d^{-1}\exp[-dm] = o(n^{-A})$

for arbitrary large $A > 0$ provided K in the definition of m is chosen large enough. Recall that the functions $\mathbb{E}(A_\nu|\mathscr{D}_j: 0 < |j - j_\nu| \le 3m)$, $\nu = 1, \dots, q$, are weakly dependent since $j_{\nu+1} - j_\nu \ge 7m$, $\nu = 1, \dots, q-1$. Using condition (A4) and the bounds for R, A_ν and B_ν we obtain

$$\begin{aligned} &\left|\mathbb{E}R\left[\prod_1^q B_\nu \mathbb{E}\big(A_\nu|\mathscr{D}_j: 0 < |j - j_\nu| \le 3m\big)\right]\right| \\ (33)\qquad &\le n^{\beta(p-1)r}\mathbb{E}\prod_1^q \left|\mathbb{E}\big(A_\nu|\mathscr{D}_j: 0 < |j - j_\nu| \le 3m\big)\right| \\ &\le n^B \prod_1^q \mathbb{E}\left|\mathbb{E}\big(A_\nu|\mathscr{D}_j: 0 < |j - j_\nu| \le 3m\big)\right| + 4n^B q d^{-1}\exp[-dm]. \end{aligned}$$

Condition (A5) implies that for $|t| \ge d$, $\mathbb{E}|\mathbb{E}(A_\nu|\mathscr{D}_j: j \ne j_\nu)| \le \exp[-d]$. So by Lemma 3.2 of GH and assumption (A6),

$$\begin{aligned} &\mathbb{E}|\mathbb{E}\big(A_\nu|\mathscr{D}_j: |j - j_\nu| \le 3m\big)| \\ &\quad\le \mathbb{E}\left|\mathbb{E}\big(A_\nu|\mathscr{D}_j: |j - j_\nu| \ne 0\big)\right| + \mathscr{O}\big(n^B d^{-1}\exp[-dm]\big) \\ &\quad\le \max\big(\exp[-dt^2 n^{-1}], \exp[-d]\big) + \mathscr{O}\big(n^B d^{-1}\exp[-dm]\big). \end{aligned}$$

We finally obtain with $q \ge cN/m$ for some $c > 0$ using (26):

$$\begin{aligned} (34)\qquad |\mathbb{E}\exp[itE_n]| &\le n^A\big(\max(\exp[-dt^2 n^{-1}], \exp[-d])\big)^{N/m} + \mathscr{O}(n^{-2}) \\ &= n^A \exp[-d'm] = \mathscr{O}(n^{-2}) \end{aligned}$$

for some K sufficiently large and $d' > 0$, thus completing the proof. □

4.2. *Edgeworth expansions for the studentized bootstrap.* Assume that $(Z_j, Y_j) \in \mathbb{R}^{k+1}$, $j = 1, \dots, n$, denotes a sequence of dependent r.v.s satisfying condition (A1)–(A4). Let A_j, B_j, $N_1 \cdots N_b$ and $U^*_{n,\text{stud}}$ be as defined in Section 2. Then we have the following theorem.

THEOREM 4.2. *Let $\Psi_n(a)$ denote the Edgeworth expansion defined in* (7) *and assume that* $(\log n)^K \le l \le n^{1/3}$ *with K large enough and that conditions* (A1)–(A4) *hold with s replaced by qs, $q \ge 3$ and $s \ge 8$. Then*

$$\sup_a |\mathbb{P}\{U^*_{n,\text{stud}} \le a\} - \Psi_n(a)| = \mathscr{O}_P(n^{-1+\varepsilon} l) + \mathscr{O}(l^{-1} n^{-1/2})$$

for $\varepsilon \triangleq 2/s$ and $l \le n^{1/3}$.

PROOF. We show (10) by using the same arguments as in Theorem 1 with R_j, $j = 1, \dots, n$, replaced by $C_{N_j} \triangleq (B_{N_j}, A^T_{N_j})^T$, $j = 1, \dots, b$. So we have to check that the conditions (A1)–(A6) hold conditionally on $Y_1, \dots, Y_n$, uniformly for all $Y_1, \dots, Y_n$ in a set whose probability tends to 1 for n going to infinity. Condition (A1) holds by the definition of A_j and B_j. In order to check the moment condition (A2), consider

$$f_s \triangleq \mathbb{E}^*\big(\|C_{N_1}\|^s\big) = (n - l + 1)^{-1}\big(\|C_0\|^s + \cdots + \|C_{n-l}\|^s\big).$$

Assume that $E\|(Z_j, Y_j)\|^{qs} < \infty$ and let $f_s = l^{-1}\sum_{\nu=1}^{l} f_{s,\nu} + \mathbb{E}\|C_1\|^s$, where

$$f_{s,\nu} \triangleq b^{-1}\sum_{\mu=0}^{b-1}\left(\|C_{\mu l+\nu}\|^s - \mathbb{E}\|C_1\|^s\right).$$

By Theorem 2.11 in GH we have, for any $x > 0$,

$$\mathbb{P}\{|f_{s,\nu} - \mathbb{E}f_{s,\nu}| > x\} = \mathscr{O}\left(\left(\frac{l}{n}\right)^{(q-2)/2}\right)$$

and therefore

$$\mathbb{P}\{|f_s - \mathbb{E}f_s| > x\} = \mathscr{O}\left(l\left(\frac{l}{n}\right)^{(q-2)/2}\right) = o(1)$$

for $q \geq 3$. The conditions (A3), (A4) and (A6) obviously hold true by independence when we choose $\mathscr{D}_j \triangleq \sigma(N_j)$, $j = 1, 2, \dots, b$. Hence it remains to check condition (A5) in this case. So we have to show that for some $0 < \zeta < 1/2$,

$$\mathbb{P}\left\{\sup_{d<|t|<b^{1/2}}\left|\mathbb{E}^*\exp\left[itB_{N_1}\right]\right| \leq 1 - \zeta\right\} = 1 - o(n^{-1}). \tag{35}$$

By definition we have with $N \triangleq n - l + 1$,

$$\begin{aligned}
&\left|\mathbb{E}^*\exp\left[itB_{N_1}\right]\right| \\
&\leq \left|N^{-1}\sum_{j=1}^{N}\left(\exp\left[itB_j\right] - \mathbb{E}\exp\left[itB_j\right]\right)\right| + \sup_j\left|\mathbb{E}\exp\left[itB_j\right]\right| \\
&\triangleq I_1 + I_2, \quad \text{say.}
\end{aligned} \tag{36}$$

By assumption (A4) and Lemma 3.33 of GH with $\alpha = 0$ and n replaced by l we have for $|t| \leq l^{\varepsilon}$, $I_2 \leq c_1\exp[-c_2 t^2]$ for some absolute constants $c_1, c_2 > 0$. For large t we get from the proof of Lemma 3.43 of GH with n replaced by l and $\alpha = 0$, $I_2 \leq \exp[-c_2 l^{\delta}]$ for some absolute constants $c_2 > 0$ and $\delta > 0$ and every t with $l^{\varepsilon} < |t| < b^{1/2}$, provided that $l \gg m \triangleq (\log n)^K$, K sufficiently large but independent of n. Hence we obtain

$$I_2 \leq 1 - 2\zeta \tag{37}$$

for every $d < |t| < b^{1/2}$ and some $0 < \zeta < 1/2$.

In order to estimate $I_1 = I_1(t)$ let $\xi_j \triangleq \exp[itB_j] - \mathbb{E}\exp[itB_j]$ denote a complex bounded random variable with mean zero. By condition (A3) we can approximate ξ_j by a mean zero random variable $\xi_j^{\dagger}$ which is $\mathscr{D}_{j-m,j+l+m}$ measurable such that

$$\mathbb{E}|\xi_j - \xi_j^{\dagger}| \leq |t| l^{1/2} d^{-1}\exp[-dm]. \tag{38}$$

We then write

$$\begin{aligned} (39)\qquad & \mathbb{P}\left\{\left|N^{-1}\sum_{j=1}^{N}\xi_j\right| > x\right\} \\ & \le \mathbb{P}\left\{\left|N^{-1}\sum_{j=1}^{N}\xi_j^{\dagger}\right| > \frac{x}{2}\right\} + \mathbb{P}\left\{\left|N^{-1}\sum_{j=1}^{N}\left(\xi_j^{\dagger} - \xi_j\right)\right| > \frac{x}{2}\right\} \\ & \triangleq I_3 + I_4, \quad \text{say.} \end{aligned}$$

By (38) we have

$$\sup_{|t|\le b^{1/2}} I_4 \le N^{-1}\sum_{j=1}^{N}\mathbb{E}|\xi_j - \xi_j^{\dagger}|/x = \mathscr{O}(n^{1/2}\exp[-dm]) = o(n^{-1})$$

for $m = (\log n)^K$ and K sufficiently large. For I_3 we repeat the argument used to check (A2). If

$$(2l)^{-1}\sum_{\nu=1}^{2l} 2lN^{-1}\sum_{p=0}^{N/(2l)}\xi_{2pl+\nu}^{\dagger} > x/2,$$

then there exists a ν with

$$m_\nu \triangleq 2lN^{-1}\sum_{p=0}^{N/(2l)}\xi_{2pl+\nu}^{\dagger} > x/2.$$

By assumption (A4) we may assume that $\xi_{2pl+\nu}^{\dagger}$, $p = 0, 1, \dots$, are independent up to an error of $\mathscr{O}(N\exp[-dm])$. Thus (39) implies

$$\mathbb{P}\{I_1(t) > x\} \le (2l)\mathbb{P}\{m_\nu > x/2\} + \mathscr{O}(N\exp[-dm]).$$

By standard exponential inequalities for bounded independent r.v. (in m_ν) we get

$$(40)\qquad \mathbb{P}\{I_1(t) > x\} \le \mathscr{O}(l)\exp\left[-cx^2N/l\right] + \mathscr{O}(N\exp[-dm])$$

for some absolute constant $c > 0$. We now have to prove (35) using (40). To this end, split the intervals $\{t: d < |t| < b^{1/2}\}$ into n^C intervals of equal length. Let A_n denote the set of midpoints of these intervals. Then we have by standard arguments

$$\begin{aligned} & \mathbb{P}\left\{\sup_{d<|t|<b^{1/2}} I_1(t) > \zeta\right\} \\ & \le \mathbb{P}\left\{\sup_{t\in A_n} I_1(t) > \zeta/2\right\} + \mathbb{P}\left\{\sup_{|t-s|\le n^{-C}b^{1/2}} |I_1(t) - I_1(s)| > \zeta/2\right\} \\ & = I_4 + I_5, \quad \text{say.} \end{aligned}$$

By (40) we have $I_4 = \mathscr{O}(ln^C\exp[-c'N/l]) + o(n^{-1})$ for some absolute constant $c' > 0$ choosing $x \triangleq \zeta/2$. In order to estimate I_5, note that we have for $\xi_j^{\dagger}(t)$,

$$\mathbb{E}\sup_{|t-s|\le\varepsilon}|\xi_j^{\dagger}(t) - \xi_j^{\dagger}(s)| \le c\varepsilon$$

by (15) and, therefore,

$$I_5 \leq \mathcal{O}(n^{-C} b^{1/2}),$$

which finally yields the desired condition (35) by choosing $C > 0$ sufficiently large. In order to complete the proof, we have to show (8). However, this is straightforward because the behavior of the bootstrap moments follows from the argument to show (A2) used at the beginning of this proof. □

REFERENCES

BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*, 2nd ed. Wiley, New York.

BOSE, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.* **16** 1709–1722.

BULINSKII, A. V. and ZHURBENKO, I. G. (1976). The central limit theorem for random fields. *Sov. Math. Dokl.* **17** 14–17.

CHIBISOV, D. M. (1972). An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion. *Theory Probab. Appl.* **17** 620–630.

DAVISON, A. C. and HALL, P. (1993). On studentizing and blocking methods for implementing the bootstrap with dependent data. *Austral. J. Statist.* **35** 215–224.

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

EFRON, B. (1987). Better bootstrap confidence intervals (with discussion). *J. Amer. Statist. Assoc.* **82** 171–200.

GÖTZE, F. and HIPP, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **64** 211–239.

GÖTZE, F. and HIPP, C. (1994). Asymptotic distribution of statistics in time series. *Ann. Statist.* **22** 2062–2088.

GÖTZE, F. and KÜNSCH, H. R. (1990). Blockwise bootstrap for dependent observations: higher order approximations for studentized statistics (abstract CP-401). *IMS Bulletin* **19** 443.

HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.* **16** 927–985.

JENSEN, J. L. (1993). A note on asymptotic expansions for sums over a weakly dependent random field with application to the Poisson and Strauss process. *Ann. Inst. Statist. Math.* **45** 353–360.

KÜNSCH, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241.

LAHIRI, S. N. (1991). Second order optimality of stationary bootstrap. *Statist. Probab. Lett.* **11** 335–341.

LAHIRI, S. N. (1996). On Edgeworth expansion and moving block bootstrap for studentized *M*-estimators in multiple linear regression models. *J. Multivariate Anal.* **56** 42–59.

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

TIKHOMIROV, A. N. (1980). On the convergence rate in the central limit theorem for weakly dependent random variables. *Theory Probab. Appl.* **25** 790–809.

FAKULTÄT FÜR MATHEMATIK
UNIVERSITY OF BIELEFELD
D-33501 BIELEFELD
GERMANY
E-MAIL: goetze@mathematik.uni-bielefeld.de

SEMINAR FÜR STATISTIK
ETH ZÜRICH
CH-8092 ZÜRICH
SWITZERLAND