The Annals of Statistics
1999, Vol. 27, No. 1, 386–404

THEORETICAL COMPARISONS OF BLOCK BOOTSTRAP METHODS[1]

BY S. N. LAHIRI

Iowa State University

In this paper, we compare the asymptotic behavior of some common block bootstrap methods based on nonrandom as well as random block lengths. It is shown that, asymptotically, bootstrap estimators derived using any of the methods considered in the paper have the *same* amount of bias to the first order. However, the variances of these bootstrap estimators may be different even in the first order. Expansions for the bias, the variance and the mean-squared error of different block bootstrap variance estimators are obtained. It follows from these expansions that using overlapping blocks is to be preferred over nonoverlapping blocks and that using random block lengths typically leads to mean-squared errors larger than those for nonrandom block lengths.

1. Introduction. In recent years, different block bootstrap methods have been proposed in the literature in the context of bootstrapping dependent data, in attempts to reproduce different aspects of the dependence structure of the observed data in the "resampled data." In this paper, we consider some of the most commonly used block bootstrap methods and compare their large sample properties in estimating the bias and the variance of a large class of estimators that can be represented in terms of certain multidimensional smooth functions of sample means. The methods that we consider in this paper are:

1. The moving block bootstrap (MBB), proposed by Künsch (1989) and Liu and Singh (1992).
2. The nonoverlapping block bootstrap (NBB), based on the work of Carlstein (1986).
3. The circular block bootstrap (CBB), proposed by Politis and Romano (1992).
4. The stationary bootstrap (SB), proposed by Politis and Romano (1994).

The first three of these methods resample blocks of observations with a nonrandom block length. The last one, namely, the SB, differs from the rest in that it uses a random block length and hence, has a slightly more complicated structure. (A description of all the methods are presented in Section 2.) It is well known [cf. Hall, Horowitz and Jing (1995)] that the bias and the variance of a block bootstrap estimator crucially depend on the block length (which serves as the smoothing parameter in this case) and that either may be the dominating term in the expansion for the mean-squared error (MSE) of a block bootstrap estimator, depending on the order of magnitude of the block length.

Received October 1997; revised July 1998.
[1]Supported in part by NSF Grant DMS-95-05124.
AMS 1991 *subject classifications*. Primary 62G05; secondary 62E05.
Key words and phrases. Block bootstrap, mean squared error, stationary bootstrap.

The main results of the paper show that for a given block length (expected block length, for the last method), all the four methods have the *same* amount of bias asymptotically. However, the leading terms in the variance part are not the same for all methods. The first-order terms in the expansions for the variances of the MBB and the CBB estimators are identical, making them asymptotically equivalent also in the MSE sense at all levels of the block length considered in the paper. In terms of the variance, the NBB estimators have an asymptotic efficiency of 2/3 compared to the corresponding MBB or CBB estimators. Consequently, when the block size l is such that the bias and the variance are of the same order, the MSE of an NBB estimator continue to be larger than that of the corresponding MBB and CBB estimators, making the NBB less efficient.

Before comparing the performance of the SB method, we pause to note that the expansions for the biases and the variances of the block bootstrap estimators also yield asymptotic expressions for the MSE-optimal block sizes. As a corollary to the expansions obtained in the paper, we determine the MSE-optimal block sizes for (the bias and variance estimation based on) each of the block bootstrap methods mentioned above. For the MBB and the NBB, some results in this direction have been obtained by Bühlmann and Künsch (1994), Hall, Horowitz and Jing (1995) and Lahiri (1996b). The results for the CBB and the SB seem to be new. The major technical difficulties encountered in obtaining the expansions for the MSE's of block bootstrap estimators arise in the SB case. Because of the randomness of the block length, the number of blocks resampled under the SB method is also random. In Section 4, we develop some conditioning arguments to obtain bounds on conditional moments of a random sum of block sums, where each block itself has a random number of terms. These bounds yield suitable stochastic expansions for the SB estimators and expansions for the MSEs of the SB estimators. For the latter purpose, we also need to use some elementary Hilbert space techniques to obtain closed form expressions for the variance part of the SB estimators. Politis and Romano (1994) identified the right order of the optimal expected block length for estimating the variance of the sample mean by the SB method. Our results refine theirs by providing an exact expression for the optimality constant.

From these expansions, it follows that the variances of the SB estimators are always *at least* twice as large as the variances of the corresponding NBB estimators and *at least* three times as large as those of the MBB and CBB estimators. Indeed, the results of Section 3 show that for all three block bootstrap methods based on nonrandom block lengths, the variances depend only on the value of the spectral density (say) $f(\cdot)$ at the origin, but for the SB method, the variance gets contributions from the spectral density at *all* frequencies in the interval $[-\pi, \pi]$. As a result, relative magnitudes of $f(0)$ and $f(x)$, $x \neq 0$ determine the degree of inferiority of SB estimators with respect to the other block bootstrap methods. It is observed that for (expected) block lengths of order larger than $n^{1/3}$, the asymptotic relative efficiency (ARE) (defined as the limit of the ratios of MSEs) of a SB estimator compared to a NBB estimator lie in the range $(0, 1/2)$ and the ARE of a SB estimator compared to a MBB or CBB lie in the interval $(0, 1/3)$. In fact, an ARE arbitrarily close to zero is possible, for example, if the ratio $f(w)/f(0)$ takes arbitrarily large values over a set of the form $\varepsilon < |w| < \pi/2 - \varepsilon$, $\varepsilon > 0$, making the variance of the SB estimator relatively very large.

We also carry out comparisons of the bootstrap methods in terms of their "best" possible performances, that is, by comparing the (asymptotically) minimum values of the MSE for each of the block bootstrap methods. This would be the case if we used each block bootstrap method with the corresponding optimal block length. In this case also, both the MBB and the CBB outperform the NBB and the NBB outperforms the SB. The ARE of the NBB compared to the MBB and the CBB is $(2/3)^{2/3}$, and the ARE of the SB compared to the NBB is less than $2^{-2/3}$ at the respective optimal block lengths.

We conclude this section with a brief literature review. Different versions of block bootstrap methods have been formulated by Hall (1985), Carlstein (1986), Künsch (1989), Liu and Singh (1992), Politis and Romano (1992, 1994) and Carlstein, Do, Hall, Hesterberg and Künsch (1995), among others. Properties of block bootstrap methods have been investigated by Davison and Hall (1993), Shao and Yu (1993), Bühlman (1994), Naik-Nimbalkar and Rajarshi (1994), Lahiri (1991, 1995, 1996a), Götze and Künsch (1996), among others. The important problem of choosing optimal block lengths has been addressed by Bühlmann and Künsch (1994), Hall, Horowitz and Jing (1995) and Lahiri (1996b).

The rest of the paper is organized as follows. In Section 2, we describe the block bootstrap methods considered in the paper. In Section 3, we state the main results of the paper. Results of a small simulation study are reported in Section 4 while proofs of all results are given in Section 5.

2. Description of block bootstrap methods.

In this section, we briefly describe the block bootstrap methods and introduce the "smooth function model" that will serve as the theoretical framework for our investigation. Let $\{X_i\}_{i=-\infty}^{\infty}$ be a $\mathbb{R}^d$-valued stationary process with $EX_1 = \mu$ and let $\mathbf{X}_n = \{X_1, \dots, X_n\}$ denote the available observations. Suppose that $\hat{\theta}_n$ is an estimator of the parameter of interest θ. The "smooth function model" [cf. Bhattacharya and Ghosh (1978), Hall (1992)] assumes that $\theta = H(\mu)$ and $\hat{\theta}_n = H(\bar{X}_n)$ where $\bar{X}_n \equiv n^{-1}\sum_{i=1}^{n} X_i$ and $H\colon \mathbb{R}^d \to \mathbb{R}$ is a smooth function. Considering suitable transformations of the original observations, this formulation allows us to consider a wide class of estimators, such as, the sample lag cross-covariance estimators, the sample autocorrelation estimators, the generalized M-estimators of Bustos (1982) and the Yule–Walker estimators for autoregressive processes.

Next we describe the block bootstrap methods mentioned in the introduction. Let $l \equiv l_n$ be an integer, satisfying $1 < l < n$. Here, l denotes the (expected) block length for the block bootstrap methods. Given the observations $\mathbf{X}_\mathbf{n}$, we form the time series $\{X_{0i}\}_{i\ge 1}$ by periodic extension where for $i \ge 1$, $X_{0i} = X_j$ if $i = mn + j$ for some integers $m \ge 0$ and $1 \le j \le n$. Also, define blocks of length $k \ge 1$ based on the time series $X_{01}, X_{02}, \dots$ by

$\mathscr{B}(i, k) = (X_{0i}, \ldots, X_{0,i+k-1})$, $i \geq 1$, $k \geq 1$. Different versions of block bootstrap methods are obtained by resampling from suitable subcollections of all "observable" blocks $\{\mathscr{B}(i, k)\colon i \geq 1, \ k \geq 1\}$. We describe the details of each method below.

MBB. The MBB method of Künsch (1989) and Liu and Singh (1992) resamples blocks randomly, with replacement from the subcollection $\{\mathscr{B}(i, k)\colon i = 1, \ldots, n-l+1, k = l\}$ of overlapping blocks. Let $I_{11}, \ldots, I_{1b}$ be conditionally iid random variables with the discrete uniform distribution on $\{1, \ldots, n-l+1\}$, that is, $P_*(I_{11} = i) = (n-l+1)^{-1}$, $1 \leq i \leq n-l+1$, where $b = [\![n/l]\!]$. Here, and in the following, we use P_* and E_* to denote, respectively, the conditional probability, and conditional expectation, given $\mathbf{X_n}$, and for any real number x, we write $[\![x]\!]$ to denote the largest integer not exceeding x. Thus, the resampled blocks for the MBB are given by $\mathscr{B}(I_{11}, l), \ldots, \mathscr{B}(I_{1b}, l)$ and arranging the elements in all b blocks in a sequence, we get the bootstrap sample $X^*_{1,1}, \ldots, X^*_{1,l}, X^*_{1,l+1}, \ldots, X^*_{1,2l}, \ldots, X^*_{1,bl}$. Here, we use the subscript 1 in $X^*_{1,i}$ to specify the bootstrap samples selected using the MBB. Similarly, we will use the subscripts 2, 3 and 4 in the following to specify the bootstrap samples obtained by using the NBB, the CBB and the SB, respectively. The bootstrap version of the centered estimator $T_n \equiv \hat{\theta}_n - \theta$ under the MBB is given by $T^{*(1)}_{n,l} = H(\bar{X}^{*(1)}_{n,l}) - H(E_*\bar{X}^{*(1)}_{n,l})$ where $\bar{X}^{*(1)}_{n,l} = n_1^{-1}\sum_{i=1}^{n_1} X^*_{1,i}$ is the bootstrap sample mean and $n_1 \equiv bl$ denotes the bootstrap sample size. The MBB estimator of the sampling distribution of T_n is given by the conditional distribution of $T^{*(1)}_{n,l}$, given $\mathbf{X_n}$.

NBB. In the NBB, one restricts attention to the collection of disjoint blocks $\{\mathscr{B}((i-1)l+1, l)\colon 1 \leq i \leq b\}$. The main idea behind the NBB is due to Carlstein (1986), who developed a variance estimation technique based on nonoverlapping blocks. To describe it, let $I_{21}, \ldots, I_{2b}$ be conditionally iid random variables with $P_*(I_{21} = i) = b^{-1}$, $1 \leq i \leq b$. Then the bootstrap sample $X^*_{2,1}, \ldots, X^*_{2,bl}$ for the NBB is obtained by arranging the bl elements in the b resampled blocks $\mathscr{B}((I_{21}-1)l+1, l), \ldots, \mathscr{B}((I_{2b}-1)l+1, l)$. The NBB version of T_n is given by $T^{*(2)}_{n,l} = H(\bar{X}^{*(2)}_{n,l}) - H(E_*\bar{X}^{*(2)}_{n,l})$, where $\bar{X}^{*(2)}_{n,l} = n_1^{-1}\sum_{i=1}^{n_1} X^*_{2,i}$.

CBB. The CBB method, proposed by Politis and Romano (1992), resamples from the collection of blocks $\{\mathscr{B}(i, l)\colon 1 \leq i \leq n\}$. Thus, in contrast to the MBB and the NBB, the CBB uses elements from the periodically extended series $\{X_{0i}\}_{i \geq 1}$ beyond X_{0n}. Let $I_{31}, \ldots, I_{3b}$ be conditionally iid r.v.'s with $P_*(I_{31} = i) = n^{-1}$, $1 \leq i \leq n$. Then the bootstrap sample $X^*_{3,1}, \ldots, X^*_{3,bl}$ under the CBB method is given by the elements of the resampled blocks $\mathscr{B}(I_{31}, l), \ldots, \mathscr{B}(I_{3b}, l)$. Politis and Romano (1992) showed that for the CBB sample mean $\bar{X}^{*(3)}_{n,l}$, $E_*\bar{X}^{*(3)}_{n,l} = \bar{X}_n$. Hence, the CBB version of T_n is given by $T^{*(3)}_{n,l} = H(\bar{X}^{*(3)}_{n,l}) - H(\bar{X}_n)$.

SB. Unlike the MBB, NBB and CBB, the SB method of Politis and Romano (1994) uses a *random* block length to generate the bootstrap sample. Let $L_{ni} \equiv L_i$, $i \geq 1$ be conditionally iid r.v.'s having the geometric distribution with parameter $p = l^{-1} \in (0, 1)$, that is, $P_*(L_1 = k) = p(1-p)^{k-1}$, $k = 1, 2, \ldots$. Also, let $I_{41}, \ldots, I_{4n}$ be conditionally iid r.v.'s with the discrete uniform distribution on $\{1, \ldots, n\}$. Then, the SB resamples $K \equiv \inf\{k \geq 1: L_1 + \cdots + L_k \geq n\}$ blocks, given by $\mathscr{B}(I_{41}, L_1), \ldots, \mathscr{B}(I_{4K}, L_K)$. Since $E_* L_1 = 1/p = l$ under the geometric distribution of L_1, on the average, the lengths of the resampled blocks tend to infinity with n like the other block bootstrap methods. The first n elements in the array $\mathscr{B}(I_{41}, L_1), \ldots, \mathscr{B}(I_{4K}, L_K)$ yield the SB sample $X^*_{4,1}, \ldots, X^*_{4,n}$. Politis and Romano (1994) show that the bootstrap sample generated by this resampling scheme is stationary (hence the name "stationary bootstrap") and that $E_*(\bar{X}^{*(4)}_{n,l}) = \bar{X}_n$. Hence, the bootstrap version of T_n under the SB is given by $T^{*(4)}_{n,l} = H(\bar{X}^{*(4)}_{n,l}) - H(\bar{X}_n)$.

Each of the block bootstrap methods described above provides estimators of different population characteristics of the estimator $\hat{\theta}_n$. In this paper, we compare the performance of these methods for estimating the parameters $\phi_{1n} \equiv \text{Bias}(\hat{\theta}_n) = E\hat{\theta}_n - \theta$ and $\phi_{2n} \equiv \text{Var}(\hat{\theta}_n) = E(\hat{\theta}_n - E\hat{\theta}_n)^2$. The bootstrap estimators of $\text{Bias}(\hat{\theta}_n)$ and $\text{Var}(\hat{\theta}_n)$ based on the block bootstrap methods described above are respectively given by

$$\hat{\phi}_{1n}(j; l) \equiv \widehat{\text{BIAS}}_j(l) = E_* T^{*(j)}_{n,l}, \qquad j = 1, 2, 3, 4;$$

$$\hat{\phi}_{2n}(j; l) \equiv \widehat{\text{VAR}}_j(l) = \text{Var}_*(T^{*(j)}_{n,l}), \qquad j = 1, 2, 3, 4.$$

3. Main results. In this section, we compare the block bootstrap estimators $\widehat{\text{BIAS}}_j(l)$ and $\widehat{\text{VAR}}_j(l)$, $1 \leq j \leq 4$ in terms of their asymptotic MSEs. Define the strong mixing coefficient of $\{X_n\}$ by $\alpha(n) \equiv sup\{|P(A \cap B) - P(A)P(B)|: A \in \mathscr{F}_1^k, B \in \mathscr{F}_{k+n}^\infty, k \geq 1\}$, where for any $1 \leq a \leq b \leq \infty$, $\mathscr{F}_a^b$ denotes the σ-field generated by $\{X_i: a \leq i \leq b\}$. Also, for $r \geq 1$ and $\delta > 0$, let $\Delta(r, \delta) = \sum_{n=1}^\infty n^{2r-1}\alpha(n)^{\delta/(2r+\delta)}$. As a convention, we also assume that unless otherwise stated, limits in the order symbols are taken as n tends to infinity.

CONDITIONS. (D_r) $H: \mathbb{R}^d \to \mathbb{R}$ and $\max\{|D^\nu H(x)|: |\nu| = r\} \leq C(1 + \|x\|^{a_0})$, $x \in \mathbb{R}^d$ for some integer $a_0 \geq 1$.
(M_r) $E\|X_1\|^{2r+\delta} < \infty$ and $\Delta(r; \delta) < \infty$ for some $\delta > 0$.

Then we have the following result on the bias part of the bootstrap estimators $\hat{\phi}_{1n}(j; l)$ and $\hat{\phi}_{2n}(j; l)$, $j = 1, 2, 3, 4$.

THEOREM 3.1. *Assume that l is such that $l^{-1} + n^{-1/2} l = o(1)$ as $n \to \infty$.*

(a) *Suppose that Condition D_r holds with $r = 3$ and that M_r holds with $r = 3 + a_0$, where a_0 is as specified by Condition D_r. Then there is a constant A_1 [cf. (5.1) in Section 5] such that*

$$\text{Bias}(\widehat{\text{BIAS}}_j(l)) = n^{-1} l^{-1} A_1 + o(n^{-1} l^{-1}) \quad \textit{for } j = 1, 2, 3, 4.$$

(b) *Suppose that Condition D_r holds with $r = 2$ and that M_r holds with $r = 4 + 2a_0$, where a_0 is as specified by Condition D_r. Then there is a constant A_2 [cf.* (5.2) *in Section* 5] *such that*

$$\mathrm{Bias}(\widehat{\mathrm{VAR}}_j(l)) = n^{-1}l^{-1}A_2 + o(n^{-1}l^{-1}) \quad \textit{for } j = 1, 2, 3, 4.$$

Thus, it follows from Theorem 3.1 that the biases of the bootstrap estimators of ϕ_{1n} and ϕ_{2n} are identical up to the first-order terms *for all* of the block bootstrap methods considered here. In particular, contrary to the common belief, the stationarity of the SB observations $X^*_{4,1}, X^*_{4,2}, \ldots$ does not contribute significantly toward reducing the bias of the resulting bootstrap estimators. Also, using either overlapping or nonoverlapping blocks incurs the same amount of bias asymptotically. Since the bias of the block bootstrap estimators essentially results from replacing the original data sequence $X_1, \ldots, X_n$ by independent copies of *smaller* subsequences, all the methods perform similarly as long as the (expected) length l of these subsequences are asymptotically equivalent.

Next we compare the variances of the block bootstrap estimators of ϕ_{1n} and ϕ_{2n}.

THEOREM 3.2. *Assume that the conditions of Theorem* 3.1 *on the block length parameter l and on the index r in Conditions D_r and M_r for the respective parts hold. Then there exist symmetric real valued functions g_1, g_2 [cf.* (5.3) *in Section* 5] *such that*:

(a)
$$\mathrm{Var}(\widehat{\mathrm{BIAS}}_j(l)) = \{4\pi^2 g_1(0)/3\}n^{-3}l + o(n^{-3}l), \qquad j = 1, 3;$$
$$\mathrm{Var}(\widehat{\mathrm{BIAS}}_j(l)) = \{2\pi^2 g_1(0)\}n^{-3}l + o(n^{-3}l), \qquad j = 2;$$
$$\mathrm{Var}(\widehat{\mathrm{BIAS}}_j(l)) = (2\pi)\left[2\pi g_1(0) + \int_{-\pi}^{\pi}(1 + e^{iw})g_1(w)\,dw\right](n^{-3}l) + o(n^{-3}l), \qquad j = 4;$$

(b)
$$\mathrm{Var}(\widehat{\mathrm{VAR}}_j(l)) = \{(2\pi)^2 g_2(0)/3\}n^{-3}l + o(n^{-3}l), \qquad j = 1, 3;$$
$$\mathrm{Var}(\widehat{\mathrm{VAR}}_j(l)) = \{(2\pi)^2 g_2(0)/2\}n^{-3}l + o(n^{-3}l), \qquad j = 2;$$
$$\mathrm{Var}(\widehat{\mathrm{VAR}}_j(l)) = (2\pi)\left[2\pi g_2(0) + \int_{-\pi}^{\pi}(1 + e^{iw})g_2(w)\,dw\right](n^{-3}l) + o(n^{-3}l), \qquad j = 4.$$

Theorem 3.2 shows that the MBB and the CBB (i.e., $j = 1, 3$) estimators of $\phi_{1n} = \mathrm{Bias}(\hat{\theta}_n)$ and $\phi_{2n} = \mathrm{Var}(\hat{\theta}_n)$ have 2/3-times smaller variances than the corresponding NBB estimators. Since the blocks in the MBB and the CBB are allowed to overlap, the amount of variability in the resampled blocks is less, leading to a smaller variance for these estimators. This advantage of the MBB over the NBB was first noted by Künsch (1989) (cf. Remark 3.3) for the sample mean and, more generally, by Hall, Horowitz and Jing (1995). It is interesting to note that in spite of all the differences in their resampling mechanisms, all four-block bootstrap methods have the same *order of magnitude* for the variances of the resulting estimators. This is particularly surprising in the case of the SB method, since it introduces additional randomness in the resampled blocks. The effect of this additional randomness shows up in the constant of the leading term in the expansion for the variances of the SB estimators. Since $\Delta_k \equiv \int_{-\pi}^{\pi}(1+e^{iw})g_k(w)\,dw \geq 0$ for $k = 1, 2$, it follows that the SB estimators have asymptotically the largest variances among all four-block bootstrap estimators for a given sequence of values of l. Furthermore, as pointed out in the introduction, there exist random processes for which the constants Δ_1, Δ_2 can be very large, leading to arbitrarily large MSEs for the SB estimators.

From Theorems 3.1 and 3.2, we see that for *each* of the block bootstrap methods considered here, as the (expected) block length l increases, the bias of a block bootstrap estimator decreases while its variance part increases. As a result, for each block bootstrap estimator, there is a critical value of the block length parameter l that minimizes the MSE. We call the value of l minimizing the leading terms in the expansion of the MSE as the (first-order) MSE-optimal block length. Let $l_{1j}^o = \arg\min\{\mathrm{MSE}(\widehat{\mathrm{BIAS}}_j(l))\colon n^{\varepsilon} < l < n^{(1-\varepsilon)/2}\}$ and $l_{2j}^o = \arg\min\{\mathrm{MSE}(\widehat{\mathrm{VAR}}_j(l))\colon n^{\varepsilon} < l < n^{(1-\varepsilon)/2}\}$, $1 \leq j \leq 4$, where $0 < \varepsilon < 1/2$ is a given number. The following result gives the optimal block lengths l_{kj}^o, $k = 1, 2$, $j = 1, 2, 3, 4$ for estimating ϕ_{1n}, ϕ_{2n} by the four-block bootstrap methods.

COROLLARY 3.1. *Assume that the conditions of Theorems* 3.1 *and* 3.2 *hold and that* $A_k \neq 0$, $g_k(0) \neq 0$, $k = 1, 2$. *Then, for* $k = 1, 2$,

$$l_{kj}^o = \left(3A_k^2/[2\pi^2 g_k(0)]\right)^{1/3} n^{1/3}(1+o(1)), \qquad j = 1, 3;$$

$$l_{kj}^o = \left(A_k^2/[\pi^2 g_k(0)]\right)^{1/3} n^{1/3}(1+o(1)), \qquad j = 2;$$

$$l_{kj}^o = \left(A_k^2 \Big/ \left[2\pi^2 g_k(0) + \pi\int_{-\pi}^{\pi}(1+e^{iw})g_k(w)\,dw\right]\right)^{1/3} n^{1/3}(1+o(1)), \qquad j = 4.$$

The formulas in Corollary 3.2 for the MBB have been noted before by Bühlmann and Künsch (1994), Hall, Horowitz and Jing (1995) and Lahiri (1996b) under similar moment and mixing conditions. For the SB variance estimator of the sample mean, Politis and Romano (1994) show that the order of the MSE-optimal expected block length $l = p^{-1}$ must be $n^{1/3}$. Corollary 3.2 specifies the optimality constant that multiplies $n^{1/3}$ for the optimal value of l.

It is clear from the definitions of l_{kj}^o that each block bootstrap method provides the most accurate estimator of the parameter ϕ_{kn} when it is used with the corresponding optimal block length. In the next result, we compare the block bootstrap methods at their *best* possible performances, that is, when *each* method is used to estimate a given parameter with its *MSE-optimal block length.*

THEOREM 3.3. *Assume that the conditions of Corollary* 3.1 *hold. Then*:

(a)
$$\mathrm{MSE}(\widehat{\mathrm{BIAS}}_j(l_{1j}^o)) = 3^{1/3}\big[2\pi^2 g_1(0)A_1\big]^{2/3} n^{-8/3} + o(n^{-8/3}), \qquad j = 1, 3;$$
$$\mathrm{MSE}(\widehat{\mathrm{BIAS}}_j(l_{1j}^o)) = 3\big[\pi^2 g_1(0)A_1\big]^{2/3} n^{-8/3} + o(n^{-8/3}), \qquad j = 2;$$
$$\mathrm{MSE}(\widehat{\mathrm{BIAS}}_j(l_{1j}^o)) = 3\left[\left\{2\pi^2 g_1(0) + \pi\int_{-\pi}^{\pi}(1+e^{iw})g_1(w)\,dw\right\}A_1\right]^{2/3} \times n^{-8/3} + o(n^{-8/3}), \qquad j = 4.$$

(b) *Expansions for* $\mathrm{MSE}(\widehat{\mathrm{VAR}}_j(l_{2j}^o))$*'s are obtained by replacing* $g_1(\cdot)$ *and* A_1 *in part* (a) *by* $g_2(\cdot)$ *and* A_2, *respectively.*

Theorem 3.3 shows that when each method is used with the corresponding MSE-optimal value of l, the MBB and the CBB has a MSE that is $(2/3)^{2/3}$ times smaller than the MSE for the NBB, and the MSE of a NBB estimator is, in turn, at least $2^{-2/3}$-times smaller than that of a SB estimator. Noting that $\int_{-\pi}^{\pi}(1+e^{iw})g_k(w)\,dw$ can take arbitrarily large positive values for some spectral densities $f(\cdot)$, it follows that even at the respective optimal blocking parameter values, the asymptotic relative efficiency of a SB estimator compared to a NBB estimator can still be arbitrarily close to zero.

REMARK 3.1. Recently Carlstein, Do, Hall, Hesterberg and Künsch (1995) have proposed a block bootstrap method, called the matched block bootstrap. Under some structural assumptions on the underlying process [e.g., AR(p) or Markov], they show that the resulting estimator of the variance of the sample mean has a variance of comparable order as the NBB and a bias that is of smaller order than $n^{-1}l^{-1}$. Thus, the minimum MSE of the matched block bootstrap is of a smaller order than the minimum MSEs for the four methods considered in this paper. Consequently, the matched block bootstrap outperforms the other methods at the respective optimal block sizes.

REMARK 3.2. It is possible to carry out a comparison of the block bootstrap methods for more complicated functionals of the sampling distribution of suitably normalized or studentized $\hat{\theta}_n$ than the bias and the variance functionals. In particular, for estimating the distribution function of a Studentized version of $\hat{\theta}_n$, an expansion for the MSE of the MBB estimator is known [cf. Lahiri (1996b)]. Similar arguments also yield expansions for the NBB and the CBB estimators as well. Furthermore, the Edgeworth expansion results of Lahiri (1997) for the SB can be used to derive an expansion for the MSE of the SB distribution function estimator. However, since exactly the same conclusions continue to hold in that case, we present the results in this simpler and representative set up and do not pursue such extensions here.

4. Simulation results. In this section, we present the results of a small simulation study to get an idea about the implications of the theoretical findings of Section 3 in finite sample situations. In the simulation study, we considered three different models generating the observations, namely:

1. ARMA (1, 1) model: $X_t - 0.3X_{t-1} = \varepsilon_t + 0.4\varepsilon_{t-1}$;
2. AR (1) model: $X_t = 0.3X_{t-1} + \varepsilon_t$;
3. MA (1) model: $X_t = \varepsilon_t + 0.4\varepsilon_{t-1}$;

where the innovations $\{\varepsilon_t\}$ were independent $N(0, 1)$ random variables. Performance of the block bootstrap methods were compared at two sample sizes, namely, $n = 100$ and $n = 400$ for estimating the parameter $\phi_{2n} \equiv \mathrm{Var}(\bar{X}_n)$. MSE's of different block bootstrap estimators were computed using 2000 simulation runs.

Figures 1 and 2 show the MSEs of the MBB, NBB and the SB estimators of ϕ_{2n} as a function of the (expected) block length l for the three models based on 500 bootstrap relicates at each block length. We chose to present the result for smaller values of l for highlighting their differences, though a similar pattern continues to hold over a larger range of l. Also, we left out the CBB estmators because of their almost identical performance to the MBB estimators over

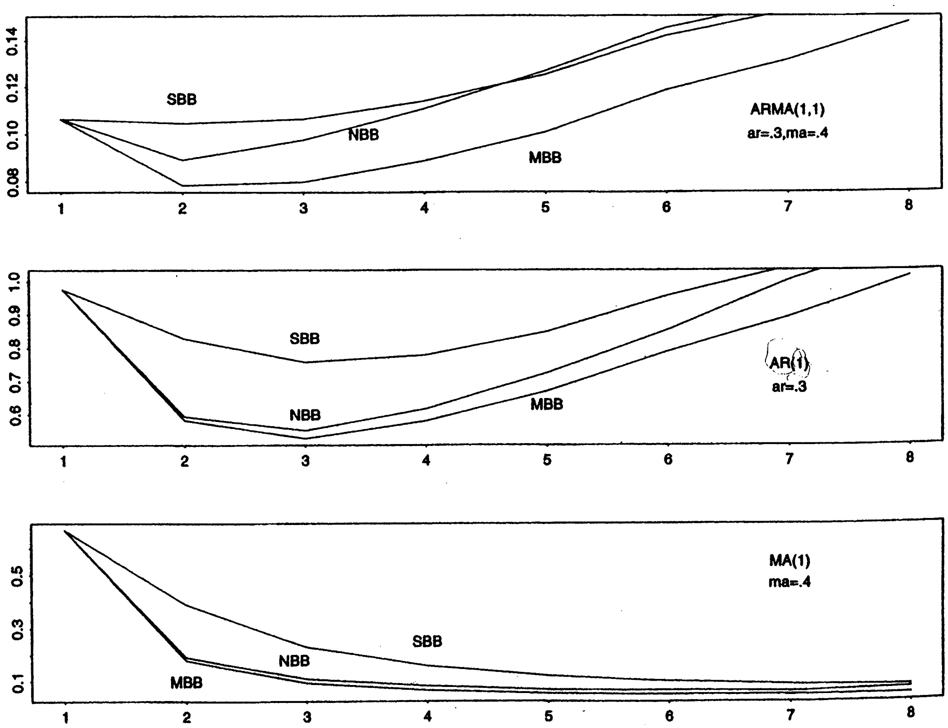

FIG. 1. *MSE's of block bootstrap estimators of $Var(\bar{X}_n)$ as a function of block length parameter l at sample size $n = 100$.*

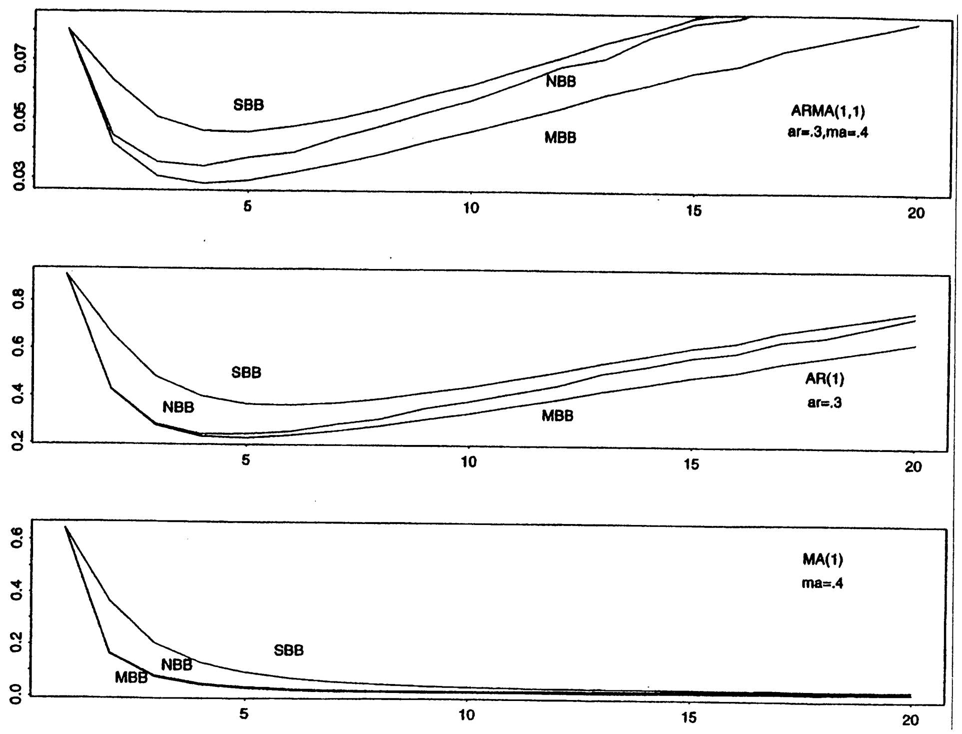

FIG. 2. *MSE's of block bootstrap estimators of* $Var(\bar{X}_n)$ *as a function of block length parameter* l *at sample size* $n = 400$.

the range of values of l considered. From the figures, it is apparent that the SBB estimators tend to have larger MSEs than the MSEs of the MBB and NBB estimators under all three models at both sample sizes. Furthermore, the relative magnitude of the MSEs of different methods shows the *ordering* predicted by the theory, at all levels of the block length parameter l, starting from $l = 2$. Thus, it appears that one is likely to gain some advantages by using the MBB or the CBB over the other two-block bootstrap methods even for moderate sample sizes.

5. Proofs. For proving the results, we need to introduce some notation. Let $\mathbb{Z}$ denote the set of all integers and let $\mathbb{Z}_+ = \{0, 1, \dots, \}$. For $\alpha = (\alpha_1, \dots, \alpha_d)' \in (\mathbb{Z}_+)^d$, $x = (x_1, \dots, x_d)' \in \mathbb{R}^d$, let $|\alpha| = \alpha_1 + \dots + \alpha_d$, $\alpha! = \prod_{i=1}^d \alpha_i!$, $x^\alpha = \prod_{i=1}^d (x_i)^{\alpha_i}$ and let D^α denote the αth order partial differential operator $\partial^{\alpha_1+\dots+\alpha_d}/\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}$. For any real numbers x, y, let $x \wedge y = \min\{x, y\}$, $x \vee y = \max\{x, y\}$, and $x_+ = \max\{x, 0\}$. Let $\zeta_s = (E\|X_1\|^s)^{1/s}$, $s > 0$. Write $c_\alpha = D^\alpha H(\mu)/\alpha!$, $\hat{c}_{\alpha, j} = D^\alpha H(\hat{\mu}(j; l))/\alpha!$, and $\hat{\mu}(j; l) = E_* \bar{X}_{n,l}^{*(j)}$, where $1 \le j \le 4$ and $\alpha \in \mathbb{Z}_+^d$. Write Σ_∞ for the limiting covariance matrix of the scaled sample mean $n^{1/2}\bar{X}_n$. Let Z_∞ denote a random vector having the $N(0, \Sigma_\infty)$

distribution on $\mathbb{R}^d$. Let $f(\cdot)$ denote the spectral density matrix of $\{X_n\}$. We index the components of the $d \times d$ matrix-valued function $f(x)$ by unit vectors $\alpha, \beta \in \mathbb{Z}_+^d$ as $f(x; \alpha, \beta)$, $|\alpha| = 1 = |\beta|$. Let $\tilde{Z}_i = \sum_{|\alpha|=1} c_\alpha (X_i - \mu)^\alpha$, $i \geq 1$. For integers $i \geq 1$, $k \geq 1$, define the partial sum $S(i, k) = X_{0i} + \cdots + X_{0(i+k-1)}$ of the extended time series $\{X_{0j}\}_{j \geq 1}$. Set $N = n - l + 1$. For a set A, let $1(A)$ denote the indicator function of the set A. We will use $C, C(\cdot)$ to denote generic positive constants depending on their arguments (if any) but not on n. For notational simplicity, we shall always set $\mu = 0$ in expressions involving the r.v.'s X_i's.

Then the constants A_i appearing in Theorem 3.1 are defined as

$$A_1 = -\sum_{|\alpha|=1} \sum_{|\beta|=1} c_{\alpha+\beta} \left(\sum_{j=-\infty}^{\infty} |j| E X_1^\alpha X_{1+j}^\beta \right); \tag{1}$$

$$A_2 = -\sum_{j=-\infty}^{\infty} |j| E \tilde{Z}_1 \tilde{Z}_{1+j}. \tag{2}$$

Also, the function g_1 of Theorem 3.2 is defined as

$$\begin{aligned} g_1(w) = \sum_{|\alpha|=1} \sum_{|\beta|=1} \sum_{|\gamma|=1} \sum_{|\iota|=1} & c_{\alpha+\beta} c_{\gamma+\iota} \\ & \times \{ f(w; \alpha, \gamma) \bar{f}(w; \beta, \iota) + f(w; \alpha, \iota) \bar{f}(w; \beta, \gamma) \\ & + f(w; \beta, \gamma) \bar{f}(w; \alpha, \iota) + f(w; \beta, \iota) \bar{f}(w; \alpha, \gamma) \}, \end{aligned} \tag{3}$$

$-\pi \leq w \leq \pi$, where $\bar{z}$ denotes the conjugate of a complex number z. The function $g_2(w)$ is obtained by replacing $c_{\alpha+\beta}$, $c_{\gamma+\iota}$ in the definition of $g_1(w)$ by $c_\alpha c_\beta$, $c_\gamma c_\iota$, respectively.

In the following, to save space we shall only outline the main steps in the proofs of the results.

LEMMA 5.1. *Let $\{t_n\}$ be a sequence of real numbers such that $t_n \to \infty$ as $n \to \infty$ and let $r \geq 1$ be an integer. Then*:

(i) $P(L_1 > t_n l) \leq C \exp(-t_n)$;
(ii) $P(|K - np| > \sqrt{np}(\log n)) \leq \exp(-C(\log n)^2)$;
(iii) $E(L_K)^r \leq C(r) l^{r+1}$;
(iv) $E(K)^r \leq C(r)(np)^r$;
(v) $E(K^{r-1} \sum_{i=1}^K L_i^r) \leq C(r) n^r$.

PROOF. Since L_1 has a geometric distribution, it can be shown [cf. Lemma 4.4, Lahiri (1997)] that $P(L_1 > t_n p^{-1}) \leq \exp([p^{-1} t_n - 1] \log(1 - p)) \leq C \exp(-t_n)$, proving part (i). Part (ii) follows from Lemma 4.5 of Lahiri (1997), which in turn, implies part (iv). For part (iii), we have

$$\begin{aligned} P(L_K = m) &= \sum_{k=1}^{n} P(L_K = m, K = k) \\ &= P(L_1 = m, L_1 \geq n) \end{aligned}$$

$$+ P(L_1 = m) \sum_{k=2}^{n} P(n - m \le L_1 + \cdots + L_{k-1} < n)$$

$$\le P(L_1 = m)\left[1 + \sum_{k=2}^{n} \sum_{j=k}^{k+m-1} P(K = j)\right] \le (m+1)P(L_1 = m).$$

Hence, $E(L_K)^r \le \sum_{m=1}^{\infty} m^r (m+1) P(L_1 = m) \le C(r) l^{r+1}$.

For part (v), note that K is a stopping time with respect to $\{\sigma\langle L_1, \dots, L_k\rangle\}_{k=1}^{n}$. Hence, by (iv), Wald's lemma [cf. Theorem 1.3, Woodroofe (1982)] and Hölder's inequality,

$$E\left(K^{r-1} \sum_{i=1}^{K} L_i^r\right) \le \{E(K^{2r-2})\}^{1/2}\left[E\left\{\sum_{i=1}^{K}(L_i^r - EL_1^r)\right\}^2 + (EK^2)(EL_1^r)^2\right]^{1/2}$$

$$\le C(r)(np)^{r-1}[E(K)EL_1^{2r} + (EK^2)(EL_1^{2r})]^{1/2}$$

$$\le C(r)(np)^r l^r. \qquad \square$$

LEMMA 5.2. *Assume that $l = O(n^{1-\varepsilon})$ for some $0 < \varepsilon < 1$ and that Condition M_r holds for some positive integer r. Then:*

(i) $E\{E_* \|S(I_{j1}; l)\|^{2r}\} \le C(r,d)\zeta_{2r+\delta}^{2r} \Delta(r;\delta) l^r$ *for* $j = 1, 2, 3$;

(ii) $E\{E_* \|S(I_{41}; L_1)\|^{2r}\} \le C(r,d)\zeta_{2r+\delta}^{2r} \Delta(r;\delta) l^r$;

(iii) $E\|\hat{\mu}(j,l)\|^{2r} \le C(r,d)\zeta_{2r+\delta}^{2r}\Delta(r;\delta) n^{-r}$ *for* $j = 1, 2, 3, 4$.

(iv) $E\{E_* \|\bar{X}_{n,l}^{*(j)}\|^{2r}\} \le C(r,d)\zeta_{2r+\delta}^{2r}\Delta(r;\delta)\varepsilon_n(r;j)$ *for* $j = 1, 2, 3, 4$ *where* $\varepsilon_n(r;j) = n^{-r}$, $1 \le j \le 3$ *and* $\varepsilon_n(r;4) = n^{-r}(1 + l(np)^{-2r})$.

PROOF. Proof of the lemma for the nonrandom block lengths (i.e., for $j = 1, 2, 3$) follows using cumulant expansions and some standard arguments [cf. Lemmas 5.2 and 5.3, Lahiri (1996b)]. Hence, we sketch a proof for the case $j = 4$ only. By Lemmas 5.1, 5.2(i) and the equality of I_{31} and I_{41} in distribution,

$$E[E_* \|S(I_{41}; L_1)\|^{2r}]$$

$$\le \sum_{m=1}^{2rl \log n} E\{E_* \|S(I_{31}; m)\|^{2r} p(1-p)^{m-1}\}$$

$$+ C(r) \max_{1 \le k \le n} \{E\|S(1;k)\|^{2r}\} \exp(-2r \log n)$$

$$\le C(r,d)\Delta(r;d)\zeta_{2r+\delta}^{2r} l^r,$$

proving (ii). Since $\hat{\mu}(4;l) = \bar{X}n$, (iii) holds trivially for $j = 4$. As for (iv), define $N_1 = L_1 + \cdots + L_K$, and $R_{4n}^* = \sum_{i=1}^{L_K} w_i^* X_{n,J(i)}$ where $J(i) = I_{4,K} + i - 1$ and $w_i^* = 0$ if $L_1 + \cdots + L_{K-1} + i \le n$ and $= 1$, otherwise. Without loss of generality, assume that the bootstrap variables $I_{41}, \dots, I_{4n}$ and $L_1, \dots, L_n$ are also defined on the underlying probability space $(\Omega, \mathscr{F}, P)$, supporting the sequence $\{X_n\}$. Let $\mathscr{L}_n = \sigma\langle L_1, \dots, L_n\rangle$, $\mathscr{X}_n = \sigma\langle X_1, \dots, X_n\rangle$

and $\mathscr{Y}_n = \mathscr{L}_n \vee \mathscr{X}_n$, $n \geq 1$. Then, by the independence of the variables I_{4j}'s, L_j's and $\{X_n\}$, we get $E\{E_*\|R_{4n}^*\|^{2r}\} \equiv E\{E(\|R_{4n}^*\|^{2r}|\mathscr{X}_n)\} = E(\|R_{4n}^*\|^{2r}) = E[E\{E(\|R_{4n}^*\|^{2r}|\mathscr{Y}_n)|\mathscr{L}_n\}] \leq C(r)E[\max\{E\|\sum_{i=1}^m a_i X_i\|^{2r}\colon\ 1 \leq m \leq L_K,\ a_i \in \{0,1\}\}] \leq C(r)\zeta_{2r+\delta}^{2r}\Delta(r;\delta)EL_K^r$. Hence, by Lemma 5.1 and Lemma III.3.1 of Ibragimov and Hasminskii (1981), we have, for $j = 4$,

$$
\begin{aligned}
&E(E_*\|\bar{X}_{n,l}^{*(j)}\|^{2r})\\
&\quad = n^{-2r}E\left\{E_*\left\|\sum_{i=1}^K S(I_{4i}, L_i) + R_{4n}^*\right\|^{2r}\right\}\\
&\quad \leq C(r)n^{-2r}\left[E(E_*\|R_{4n}^*\|^{2r}) + E\left\{\left(E_*\left|\sum_{i=1}^K L_i\right|^{2r}\right)\|\bar{X}_n\|^{2r}\right\}\right.\\
&\qquad \left. + C(r)n^{-2r}E\left(E\left\{E\left(\left\|\sum_{i=1}^K (S(I_{4i}, L_i) - L_i\bar{X}_n)\right\|^{2r}\middle|\mathscr{Y}_n\right)\middle|\mathscr{X}_n\right\}\right)\right]\\
&\quad \leq C(r)\zeta_{2r+\delta}^{2r}\Delta(r;\delta)n^{-r}[1 + n^{-2r}E(L_K)^{2r}]\\
&\qquad + C(r)n^{-2r}E\left\{K^{r-1}\sum_{i=1}^K E(\|S(I_{4i}, L_i)\|^{2r}|\mathscr{L}_n)\right\} \qquad (4)\\
&\quad \leq C(r)\zeta_{2r+\delta}^{2r}\Delta(r;\delta)n^{-r}[1 + n^{-2r}l^{2r+1}]\\
&\qquad + C(r)n^{-2r}E\left[K^{r-1}\sum_{i=1}^K \{(E\|S(1, L_i)\|^{2r}|\mathscr{L}_n)1(L_i \leq 4rl\log n)\right.\\
&\qquad\qquad + \max\{E\|S(1,m)\|^{2r}\colon 1 \leq m \leq n\}\\
&\qquad\qquad \left. \times 1(L_i \geq 4rl\log n)\}\right]\\
&\quad \leq C(r)\zeta_{2r+\delta}^{2r}\Delta(r;\delta)n^{-r}[1 + l(np)^{-2r}].
\end{aligned}
$$

This completes the proof of Lemma 5.2. □

LEMMA 5.3. *Let $g\colon (-\pi, \pi] \to [0, \infty)$ be a continuous function that is symmetric about zero. Then, with $p = l^{-1}$, $q = 1 - p$, and $Q_{jn}(w) = e^{ijw}(1 - q^j e^{ijw})^{-1}$, $j = 1, 2$:*

(i) $$\lim_{n\to\infty}\int_{-\pi}^{\pi}[Q_{1n}(w)]g(w)\,dw = \pi g(0) + \int_{-\pi}^{\pi}[2^{-1}\cos w + (\cos(w/2))^2]g(w)\,dw;$$

(ii) $$\lim_{n\to\infty} p\int_{-\pi}^{\pi}[Q_{1n}(w)Q_{2n}(w)]g(w)\,dw = g(0)\left[2\int_{-\infty}^{\infty}(1+4y^2)^{-2}\,dy - \int_{-\infty}^{\infty}(1+4y^2)^{-1}\,dy\right];$$

(iii) $$\int_{-\pi}^{\pi}[Q_{2n}(w)]g(w)\,dw = O(1).$$

PROOF. Since g is real and symmetric, for any $M > 1$, we have

$$\begin{aligned}
&\int_{-\pi}^{\pi}[Q_{1n}(w)]g(w)\,dw\\
&\quad=\left(\int_{|w|\le Mp}+\int_{Mp<|w|<M^{-1}}+\int_{M^{-1}<|w|<\pi}\right)\\
&\qquad\left[(p^2+4q\sin^2(w/2))^{-1}\{(p+2q\sin^2(w/2))\cos w\right.\\
&\qquad\qquad\left.+q(\sin w)^2\}g(w)\right]dw\\
&\quad\equiv I_1(M)+I_2(M)+I_3(M),\quad\text{say.}
\end{aligned}\tag{5}$$

Using the change of variable $y = w/2p$ and the bounded convergence theorem, one can show that for any $M > 1$, as $n \to \infty$,

$$\begin{aligned}
I_1(M)&=2\int_{|y|<M/2}\frac{(1+2qp^{-1}\sin^2 yp)\cos 2py+qp^{-1}(\sin^2 2yp)}{1+4qy^2[(py)^{-2}\sin^2 py]}g(2py)\,dy\\
&=2g(0)\int_{|y|<M/2}(1+4y^2)^{-1}dy+o(1).
\end{aligned}\tag{6}$$

Next, noting that $x/3 < \sin x$ for all $x \in (0, \pi/2]$, for any $M > 1$, we have

$$\begin{aligned}
I_2(M)&\le C\left[\int_{M<y<(Mp)^{-1}}y^{-2}g(py)\,dy+\int_{0<w<M^{-1}}g(w)\,dw\right]\\
&\le C\max\{g(w)\colon 0<w<M^{-1}\}M^{-1}.
\end{aligned}\tag{7}$$

Finally, by the bounded convergence theorem, for any $M > 1$,

$$\begin{aligned}
\lim_{n\to\infty} I_3(M)&=\int_{M^{-1}<|w|<\pi}(4\sin^2 w/2)^{-1}\\
&\qquad\times[(2\sin^2 w/2)\cos w+(\sin w)^2]g(w)\,dw.
\end{aligned}\tag{8}$$

Part (i) follows from (5)–(8), by letting $M \to \infty$. The proof of part (ii) is similar to part (i), except that we need to split the integral into three parts, now ranging over the sets $[-Mp, Mp]$, $\{w\colon Mp < |w| < \pi/2\}$ and $\{w\colon \pi/2 \le |w| < \pi\}$, $M > 1$ and establish negligibility of the last two terms. We omit the details. The proof of (iii) is similar and is also omitted to save space. □

PROOF OF THEOREM 3.1. We provide an outline of the proof only for the bias estimators $\hat{\phi}_{1n}(j;l)$, $j = 1, 2, 3, 4$. Using Taylor's expansion, Hölder's inequality, and Lemma 5.2, for $j \in \{1, 2, 3\}$, one can show that

$$\hat{\phi}_{1n}(j;l)=b^{-1}l^{-2}\sum_{|\alpha|=2}c_\alpha E_*(S(I_{j1};l))^\alpha+\hat{R}_{1n}(j;l),\tag{9}$$

where the remainder term $\hat{R}_{1n}(j;l)$ satisfies the inequality $E(\hat{R}_{1n}(j;l))^2 \le C[n^{-3} + n^{-4}l^2]$. Now one can establish Theorem 3.1 for $j \in \{1, 2, 3\}$ using (9) and the fact that $\mathrm{BIAS}(\hat{\theta}_n) = n^{-1}\sum_{|\alpha|=2} c_\alpha EZ_\infty^\alpha + O(n^{-2})$. We omit the details.

Next we consider the case $j = 4$. As in (9), one can show that for $j = 4$,

$$\hat{\phi}_{1n}(j;l) = \sum_{|\alpha|=2} c_\alpha E_*(\bar{X}_{n,l}^{*(j)} - \bar{X}_n)^\alpha + \hat{R}_{1n}(j;l), \tag{10}$$

where the remainder term $\hat{R}_{1n}(j;l)$ admits the bound $E(\hat{R}_{1n}(j;l))^2 \le Cn^{-3}$. Next for $0 \le j \le n-1$ and $\alpha,\ \beta \in \mathbb{Z}_+^d$ with $|\alpha| = 1 = |\beta|$, write $\hat{\sigma}(j;\alpha,\beta) = n^{-1}\sum_{i=1}^{n-j} X_i^\alpha X_{i+j}^\beta$. Note that $\{X_{4,i}^*\}_{i\ge 1}$ is *stationary* and that for all $j < L_1$, $X_{4,1}^*$ and $X_{4,(j+1)}^*$ are in the same resampled block. Hence, arguing as in Politis and Romano (1994), for any $|\alpha| = |\beta| = 1$, $1 \le j \le n-1$, one gets

$$\begin{aligned}
&E_*(X_{4,1}^*)^\alpha (X_{4,(j+1)}^*)^\beta \\
&\quad = E\Big[E\{((X_{4,1}^*)^\alpha (X_{4,(j+1)}^*)^\beta 1(L_1 \le j) \\
&\qquad\qquad + (X_{4,1}^*)^\alpha (X_{4,(j+1)}^*)^\beta 1(L_1 > j))\,|\,\mathscr{Y}_n\}\Big|\mathscr{X}_n\Big] \\
&\quad = (\bar{X}_n)^{\alpha+\beta} P(L_1 \le j) + \left(n^{-1}\sum_{i=1}^n X_{0i}^\alpha X_{0(i+j)}^\beta\right) P(L_1 > j) \\
&\quad = (\bar{X}_n)^{\alpha+\beta}(1-q^j) + q^j\{\hat{\sigma}(j;\alpha,\beta) + \hat{\sigma}(n-j;\beta,\alpha)\}.
\end{aligned}$$

Hence, using the stationarity of $X_{4,i}^*$'s, one can show that

$$\begin{aligned}
&\sum_{|\alpha|=1}\sum_{|\beta|=1} c_{\alpha+\beta} E_*(\bar{X}_{n,l}^{*(4)} - \bar{X}_n)^{\alpha+\beta} \\
&\quad = n^{-1}\sum_{|\alpha|=1}\sum_{|\beta|=1} c_{\alpha+\beta}\Bigg[\sum_{j=0}^{n-1} q_{nj}\big(\hat{\sigma}(j;\alpha,\beta) + \hat{\sigma}(j;\beta,\alpha)\big) \\
&\qquad\qquad - \left\{1 + 2\sum_{j=1}^{n-1}(1-n^{-1}j)q^j\right\}(\bar{X}_n)^{\alpha+\beta}\Bigg],
\end{aligned} \tag{11}$$

where $q_{nj} = (1-n^{-1}j)q^j + (n^{-1}j)q^{(n-j)}$, $1 \le j \le n-1$ and $q_{n0} = 1/2$. Next, note that by Taylor's expansion, $|1 - q^j - jp| \le j^2p^2/2$ for all $j \ge 1$, and that $p^{-1}(1-q_{nj})(1-n^{-1}j) \to j$ as $n \to \infty$, for all $j \ge 1$. Now, using (10), (11) and the fact that $\sum_{j=1}^\infty j^2|EX_1^\alpha X_{1+j}^\beta| < \infty$, one can complete the proof for the case $j = 4$. □

PROOF OF THEOREM 3.2. For $j = 1, 2, 3$, a proof of the expansion for Var $(\hat{\phi}_{1n}(j;l))$ under the conditions of Theorem 3.2 can be constructed using (9) and the arguments in the proof of Theorem 2.1 of Lahiri (1996b). [See also Hall, Horowitz and Jing (1995).] Here, we obtain the desired expansion for $\mathrm{Var}(\hat{\phi}_{1n}(j;l))$, $j = 4$ only. Since $E\{n^{-1}(1 + 2\sum_{j=1}^{n-1}(1-n^{-1}j)q^j)^2\|\bar{X}_n\|^4\} = O(n^{-4}l^2)$, in view of (10) and (11), it is enough to show that $\mathrm{Var}(n^{-1}\sum_{|\alpha|=1}\sum_{|\beta|=1} c_{\alpha+\beta}[\sum_{j=0}^{n-1} q_{nj}(\hat{\sigma}(j;\alpha,\beta) + \hat{\sigma}(j;\beta,\alpha))]) = (2\pi)[2\pi g_1(0) + \int_{-\pi}^{\pi}(1+e^{iw})\cdot g_1(w)\,dw](n^{-3}l) + o(n^{-3}l)$.

Using Bartlett's (1946) formula for the covariance of sample autocovariance estimators [cf. (5.3.21), (5.3.22), Priestley (1981)], one can show that for unit vectors $\alpha, \beta, \gamma, \iota \in \mathbb{Z}_+^d$,

$$\begin{aligned}
&\operatorname{cov}\big(\hat{\sigma}(j;\alpha,\beta), \hat{\sigma}(k;\gamma,\iota)\big)\\
&\quad = n^{-1} \sum_{m=-(n-j)+1}^{(n-j-v-1)} \{1 - n^{-1}(\eta_{jv}(m) + k)\}\\
&\qquad \times \{(EX_1^\alpha X_{1+m}^\gamma)(EX_1^\beta X_{1+m+v}^\iota) + (EX_1^\alpha X_{m+k+1}^\iota)(EX_1^\beta X_{1+m-j}^\gamma)\}\\
&\qquad + R_{21n}(j,k;\alpha,\beta,\gamma,\iota),
\end{aligned} \tag{12}$$

where $v = k - j$, $\eta_{jv}(m) = m1(m > 0) - (m+v)1(-(n-j)+1 \le m < -v)$, and the remainder term $R_{21n}(j,k;\alpha,\beta,\gamma,\iota)$ satisfies the inequality $n^{-2}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} q_{nj}q_{nk}|R_{21n}(j,k;\alpha,\beta,\gamma,\iota)| = O(n^{-3})$.

Next, let $a \equiv a_n = [\![l(\log n)^2 + n^{1/3}]\!]$ and $d_{jvm} \equiv d_{jvmn} = n^{-1}\{1 - n^{-1}(\eta_{jv}(m) + j + v)\}$. Note that

$$\begin{aligned}
&\max\{q_{nj}: a \le j \le n-a\}\\
&\quad \le q^a = O\Big(\exp(-(\log n)^2)\Big);\\
&|q_{nj}q_{n(j+v)} - q^j q^{j+v}|\\
&\quad \le 2n^{-1}(jq^j + (j+v)q^{j+v}), \qquad 1 \le j, v \le 2a;\\
&\quad \times n^{-2}\sum_{j=n-a}^{n-2}\sum_{v=1}^{n-1-j} q_{nj}q_{n(j+v)} \sum_{m=-(n-j)+1}^{(n-j)-v-1} |d_{jvm}(EX_1^\alpha X_{1+m}^\gamma)(EX_1^\beta X_{1+v+m}^\iota)|\\
&\quad \le n^{-3}\sum_{j=n-a}^{n-2}\sum_{v=1}^{a}\sum_{m=-a+1}^{a-v-1}\{n^{-1}(|n-j| + |m| + v)\}|EX_1^\alpha X_{1+m}^\gamma||EX_1^\beta X_{1+v+m}^\iota|\\
&\quad \le 3n^{-4}a^2\bigg(\sum_{m=-\infty}^{\infty}|EX_1^\alpha X_{1+m}^\gamma|\bigg)\bigg(\sum_{u=-\infty}^{\infty}|EX_1^\beta X_{1+u}^\iota|\bigg);
\end{aligned}$$

and by similar arguments, $n^{-2}\sum_{j=0}^{a}\sum_{v=n-2a}^{(n-j)-1} q_{nj}q_{n(j+v)} \sum_{m=-(n-j)+1}^{n-j-v-1}|d_{jvm}|\,|EX_1^\alpha X_{1+m}^\gamma|\,|EX_1^\beta X_{1+v+m}^\iota| = O(n^{-4}a^2)$. Hence, using these bounds, Condition M_r and Perseval's identity [cf. Chapter 4, Rudin (1985)] for Hilbert spaces, one can show that for any unit vectors $\alpha, \beta, \gamma, \iota \in \mathbb{Z}_+^d$,

$$\begin{aligned}
&n^{-2}\sum_{j=0}^{n-2}\sum_{v=1}^{n-1-j} q_{nj}q_{n(j+v)}\Bigg\{\sum_{m=-(n-j)+1}^{(n-j)-v-1} d_{jvm}(EX_1^\alpha X_{1+m}^\gamma)(EX_1^\beta X_{1+v+m}^\iota)\Bigg\}\\
&\quad = n^{-3}\bigg(\sum_{v=1}^{a} 2^{-1}q^v + \sum_{j=1}^{a}\sum_{v=1}^{a} q^{2j+v}\bigg)
\end{aligned}$$

$$\times\left\{\sum_{m=-(n-j)+1}^{(n-j)-v-1}(EX_1^{\alpha}X_{1+m}^{\gamma})(EX_1^{\beta}X_{1+v+m}^{\iota})\right\}+O(n^{-4}a^2)$$

$$\begin{aligned}
&= n^{-3}\left(2^{-1}\sum_{v=1}^{\infty}q^v+\sum_{j=1}^{\infty}\sum_{v=1}^{\infty}q^{2j+v}\right)\\
&\quad\times\left[\sum_{m=-\infty}^{\infty}\left(\int_{-\pi}^{\pi}e^{imw}f(w;\alpha,\gamma)\,dw\right)\left(\int_{-\pi}^{\pi}\exp(-i(v+m)w)\bar{f}(w;\beta,\iota)\,dw\right)\right]\\
&\quad+O(n^{-4}a^2)\\
&= n^{-3}\left(\pi+\frac{2\pi q^2}{(1-q^2)}\right)\sum_{v=1}^{\infty}q^v\left(\int_{-\pi}^{\pi}f(w;\alpha,\gamma)\exp(-ivw)\bar{f}(w;\beta,\iota)\,dw\right)\\
&\quad+O(n^{-4}a^2)\\
&= n^{-3}\left(\pi+\frac{2\pi q^2}{(1-q^2)}\right)\int_{-\pi}^{\pi}qe^{iw}(1-qe^{iw})^{-1}f(-w;\alpha,\gamma)\bar{f}(-w;\beta,\iota)\,dw\\
&\quad+O(n^{-4}a^2).
\end{aligned}\tag{13}$$

Using similar arguments, one can show that for any unit vectors $\alpha,\beta,\gamma,\iota\in\mathbb{Z}_+^d$,

$$\begin{aligned}
&n^{-2}\sum_{j=0}^{n-2}\sum_{v=1}^{n-1-j}q_{nj}q_{n(j+v)}\left\{\sum_{m=-(n-j)+1}^{(n-j)-v-1}d_{jvm}(EX_1^{\alpha}X_{1+m+j+v}^{\iota})(EX_1^{\beta}X_{1+m-j}^{\gamma})\right\}\\
&\qquad= n^{-3}\int_{-\pi}^{\pi}\left\{\pi+\frac{2\pi(qe^{iw})^2}{1-(qe^{iw})^2}\right\}\left\{\frac{qe^{iw}f(w;\beta,\gamma)\bar{f}(w;\alpha,\iota)}{1-qe^{iw}}\right\}dw\\
&\qquad\quad+O(n^{-4}a^2),
\end{aligned}\tag{14}$$

$$\begin{aligned}
&n^{-3}\sum_{j=0}^{n-1}q_{nj}^2\sum_{m=-(n-j)+1}^{(n-j)-1}\left(1-\frac{|m|+j}{n}\right)\\
&\qquad\times\{(EX_1^{\alpha}X_{1+m}^{\gamma})(EX_1^{\beta}X_{1+m}^{\iota})+(EX_1^{\alpha}X_{1+m+j}^{\iota})(EX_1^{\beta}X_{1+m-j}^{\gamma})\}\\
&\qquad= n^{-3}\left\{\left(\frac{\pi}{2}+\frac{2\pi q^2}{1-q^2}\right)\int_{-\pi}^{\pi}f(w;\alpha,\gamma)\bar{f}(w;\beta,\iota)\,dw\right.\\
&\qquad\quad\left.+\int_{-\pi}^{\pi}\left(\frac{\pi}{2}+\frac{2\pi(qe^{iw})^2}{1-(qe^{iw})^2}\right)f(-w;\beta,\gamma)\bar{f}(-w;\alpha,\iota)\,dw\right\}\\
&\qquad\quad+O(n^{-4}a^2).
\end{aligned}\tag{15}$$

The proof of Theorem 3.2 can now be completed using (13)–(15) and Lemma 5.3. We omit the details. □

PROOF OF COROLLARY 3.1. Corollary 3.1 follows from Theorems 3.1 and 3.2 and the fact that the function $h(x) = c_1 x + c_2 x^{-2}$, $x > 0$, $c_1 \geq 0$, $c_2 \geq 0$ is minimized at $x^* = (2c_2/c_1)^{1/3}$ and $h(x^*) = (3/\sqrt[3]{4})c_1^{2/3}c_2^{1/3}$. □

PROOF OF THEOREM 3.3. Theorem 3.3 follows from Theorems 3.1 and 3.2 and Corollary 3.1. □

REFERENCES

BARTLETT, M. S. (1946). On the theoretical specification of sampling properties of autocorrelated time series. *J. Roy. Statist. Soc. Suppl.* **8** 27–41.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BÜHLMAN, P. (1994). Blockwise bootstrapped empirical process for stationary sequences. *Ann. Statist.* **22** 995–1012.

BÜHLMAN, P. and KÜNSCH, H. R. (1994). Block length selection in the bootstrap for time series. Research Report 72, Seminar für Statistik, ETH, Zürich.

BUSTOS, O. (1982). General M-estimates for contaminated pth order autoregressive processes: consistency and asymptotic normality. *Z. Wahrsch. Verw. Gebiete* **59** 491–504.

CARLSTEIN, E. (1986). The use of subseries methods for estimating the variance of a general statistic from a stationary time series. *Ann. Statist.* **14** 1171–1179.

CARLSTEIN, E., DO, K-A., HALL, P., HESTERBERG, T. and KÜNSCH, H. R. (1995). Matched-block bootstrap for dependent data. Research Report 74, Seminar für Statistik, ETH, Zürich.

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.

GÖTZE, F. and KÜNSCH, H. R. (1996). Blockwise bootstrap for dependent observations: higher order approximations for Studentized statistics. *Ann. Statist.* **24** 1914–1933.

HALL, P. (1985). Resampling a coverage pattern. *Stochastic Process. Appl.* **20** 231–246.

HALL, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer, New York.

HALL, P., HOROWITZ, J. L. and JING, B. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82** 561–574.

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1981). *Statistical Estimation: Asymptotic Theory.* Springer, New York.

KÜNSCH, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1261.

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

LAHIRI, S. N. (1995). On the asymptotic behaviour of the moving block bootstrap for normalized sums of heavy-tail random variables. *Ann. Statist.* **23** 1331–1349.

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

LAHIRI, S. N. (1996b). On empirical choice of the optimal block size for block bootstrap methods. Preprint. Dept. Statistics, Iowa State Univ., Ames.

LAHIRI, S. N. (1997). On second-order properties of the stationary bootstrap method for Studentized statistics. Preprint. Dept. Statistics, Iowa State Univ., Ames.

LIU, R. Y. and SINGH, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap* (R. Lepage and L. Billard, eds.) 225–248. Wiley, New York.

NAIK-NIMBALKAR, U. V. and RAJARSHI, M. B. (1994). Validity of blockwise bootstrap for empirical processes with stationary observations. *Ann. Statist.* **22** 980–994.

POLITIS, D. and ROMANO, J. P. (1992). A circular block resampling procedure for stationary data. In *Exploring the Limits of Bootstrap* (R. Lepage and L. Billard, eds.) 263–270. Wiley, New York.

POLITIS, D. and ROMANO, J. P. (1994). The stationary bootstrap. *J. Amer. Statist. Assoc.* **89** 1303–1313.

PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series* **1**. Academic Press, New York.

RUDIN, W. (1985). *Real and Complex Analysis.* McGraw-Hill, New York.

SHAO, Q. M. and YU, H. (1993). Bootstrapping the sample means for stationary mixing sequences. *Stochastic Process. Appl.* **48** 175–190.

WOODROOFE, M. B. (1982). *Nonlinear Renewal Theory in Sequential Analysis*. SIAM, Philadelphia.

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011-1210
E-MAIL: snlahiri@iastate.edu