*Time Series Analysis: Methods and Applications, Vol. 30*
ISSN: 0169-7161


DOI: 10.1016/B978-0-444-53858-1.00001-6

# Bootstrap Methods for Time Series

*Jens-Peter Kreiss*[1] *and Soumendra Nath Lahiri*[2]

[1]*Technische Universität Braunschweig, Institut für Mathematische Stochastik, Pockelsstrasse 14, D-38106 Braunschweig, Germany*
[2]*Department of Statistics, TAMU-3143 Texas A & M University, College Station, TX 77843, USA*

**Abstract**

The chapter gives a review of the literature on bootstrap methods for time series data. It describes various possibilities on how the bootstrap method, initially introduced for independent random variables, can be extended to a wide range of dependent variables in discrete time, including parametric or nonparametric time series models, autoregressive and Markov processes, long range dependent time series and nonlinear time series, among others. Relevant bootstrap approaches, namely the intuitive residual bootstrap and Markovian bootstrap methods, the prominent block bootstrap methods as well as frequency domain resampling procedures, are described.

Further, conditions for consistent approximations of distributions of parameters of interest by these methods are presented. The presentation is deliberately kept non-technical in order to allow for an easy understanding of the topic, indicating which bootstrap scheme is advantageous under a specific dependence situation and for a given class of parameters of interest. Moreover, the chapter contains an extensive list of relevant references for bootstrap methods for time series.

**Keywords:** bootstrap methods, discrete Fourier transform, linear and nonlinear time series, long range dependence, Markov chains, resampling, second order correctness, stochastic processes.

## 1. Introduction

The bootstrap method, initially introduced by Efron (1979) for independent variables and later extended to deal with more complex dependent variables by several authors, is a class of nonparametric methods that allow the statistician to carry out statistical

inference on a wide range of problems without imposing much structural assumptions on the underlying data-generating random process. By now, there exist several books and monographs, e.g., Hall (1992), Efron and Tibshirani (1993), Shao and Tu (1995), Davison and Hinkley (1997), and Lahiri (2003a), among others, which describe different aspects of the bootstrap methodology at varying levels of sophistication and generality. Moreover, several papers in the literature give overviews of various aspects of bootstrapping time series. Among them are Berkowitz and Kilian (2000), Bose and Politis (1995), Bühlmann (2002), Carey (2005), Härdle et al. (2003), Li and Maddala (1996), and Politis (2003). These papers consider bootstrap and resampling methods for general stochastic processes and time series models. The review papers by Paparoditis and Politis (2009) and by Ruiz and Pascual (2002) especially focus on financial time series, while McMurry and Politis considers resampling methodology for functional data. In this article, we aim to provide an easy-to-read description of some of the key ideas and issues and present latest results on a set of selected topics in the context of time series data showing temporal dependence.

The basic idea behind the bootstrap methods is very simple, and it can be described in general terms as follows. Let $X_1, \ldots, X_n$ be a stretch of a time series with joint distribution $P_n$. For estimating a population parameter $\theta$, suppose that we have constructed an estimator $\hat{\theta}_n$ (e.g., using the generalized method of moments) based on $X_1, \ldots, X_n$. A common problem that the statistician must deal with is to assess the accuracy of $\hat{\theta}_n$, for example, by using an estimate of its mean squared error (MSE) or an interval estimate of a given confidence level. However, any such measure of accuracy depends on the sampling distribution of $\hat{\theta}_n - \theta$, which is typically unknown in practice and often very complicated. Bootstrap methods provide a general recipe for estimating the distribution of $\hat{\theta}_n$ and its functionals without restrictive model assumptions on the time series.

We now give a general description of the basic principle underlying the bootstrap methods. As before, suppose that the data are generated by a part of a time series $\{X_1, \ldots, X_n\} \equiv \mathbf{X}_n$ with joint distribution $P_n$. Given $\mathbf{X}_n$, first construct an estimate $\hat{P}_n$ of $P_n$. Next, generate random variables $\{X_1^*, \ldots, X_n^*\} \equiv \mathbf{X}_n^*$ from $\hat{P}_n$. If $\hat{P}_n$ is a reasonably "good" estimator of $P_n$, then the relation between $\{X_1, \ldots, X_n\}$ and $P_n$ is closely reproduced (in the bootstrap world) by $\{X_1^*, \ldots, X_n^*\}$ and $\hat{P}_n$. Define the bootstrap version $\hat{\theta}_n^*$ of $\hat{\theta}_n$ by replacing $X_1, \ldots, X_n$ with $X_1^*, \ldots, X_n^*$, and similarly, define $\theta^*$ by replacing $P_n$ in $\theta = \theta(P_n)$ by $\hat{P}_n$. Then, the conditional distribution (function) $\hat{G}_n$ or $G_n^*$ (say) of $\hat{\theta}_n^* - \theta^*$ (given $\mathbf{X}_n$) gives the bootstrap estimator of the distribution (function) $G_n$ (say) of $\hat{\theta}_n - \theta$. Here, $\theta^*$ is some properly chosen parameter, which in many applications can be computed from $\hat{P}_n$ along the same lines as $\theta$ is computed from $P_n$. In almost all applications, the bootstrap is used to approximate distributions of the type $c_n(\hat{\theta}_n - \theta)$, where the to infinity increasing sequence $(c_n)$ of non-negative real numbers is chosen such that the sequence of distributions converges to a nondegenerate limit.

To define the bootstrap estimators of a functional of the distribution of $\hat{\theta}_n - \theta$, such as the variance or the quantiles of $\hat{\theta}_n - \theta$, we may simply use the "plug-in" principle and employ the corresponding functional to the conditional distribution of $\hat{\theta}_n^* - \theta^*$. Thus, the bootstrap estimator of the variance $\sigma_n^2$ of $\hat{\theta}_n - \theta$ is given by the conditional

variance $\hat{\sigma}_n^2$ of $\hat{\theta}_n^* - \theta^*$, i.e., by

$$\begin{aligned}\hat{\sigma}_n^2 &= \text{the bootstrap estimator of } \sigma_n^2 \\ &= \text{Var}(\hat{\theta}_n^* - \theta^* | \mathbf{X}_n) \\ &= \int x^2 d\hat{G}_n(x) - \left[\int x d\hat{G}_n(x)\right]^2.\end{aligned}$$

Similarly, if $q_{\alpha,n}$ denotes the $\alpha \in (0,1)$ quantile of (the distribution of ) $\hat{\theta}_n - \theta$, then its bootstrap estimator is given by

$$\hat{q}_{\alpha,n} = \hat{G}_n^{-1}(\alpha), \text{ the } \alpha \text{ quantile of the conditional distribution of } \hat{\theta}_n^* - \theta^*.$$

In general, having chosen a particular bootstrap method for a specific application, it is very difficult (and often, impractical) to derive closed-form analytical expressions for the bootstrap estimators of various population quantities. This is where the computer plays an indispensable role. Bootstrap estimators of the distribution of $\hat{\theta}_n - \theta$ can be computed numerically using Monte-Carlo simulation. First, a large number (usually in hundreds) of independent copies $\{\hat{\theta}_n^{*k} : k = 1, \ldots, K\}$ of $\hat{\theta}_n^*$ are constructed by *repeated* resampling. The empirical distribution of these bootstrap replicates gives the desired Monte-Carlo approximation to the true bootstrap distribution of $\hat{\theta}_n^* - \theta^*$ and to its functionals. Specifically, for the variance parameter $\sigma_n^2 = \text{Var}(\hat{\theta}_n - \theta)$, the Monte-Carlo approximation to the bootstrap estimator $\hat{\sigma}_n^2$ is given by

$$[\hat{\sigma}_n^{\text{MC}}]^2 \equiv (K-1)^{-1} \sum_{k=1}^{K} \left[\hat{\theta}_n^{*k} - K^{-1} \sum_{j=1}^{K} \hat{\theta}_n^{*j}\right]^2,$$

the sample variance of the replicates $\{\hat{\theta}_n^{*k} - \theta^* : k = 1, \ldots, K\}$. Similarly, the Monte-Carlo approximation to the bootstrap estimator $\hat{q}_{\alpha,n}$ is given by

$$\hat{q}_{n,\alpha}^{\text{MC}} \equiv \hat{\theta}_n^{*(\lfloor K\alpha \rfloor)} - \theta^*,$$

the $\lfloor K\alpha \rfloor$ order statistic of the replicates $\{\hat{\theta}_n^{*k} - \theta^* : k = 1, \ldots, K\}$, where for any real number $x$, $\lfloor x \rfloor$ denotes the largest integer not exceeding $x$. From this point of view, the introduction of the bootstrap has been very timely; almost none of the interesting applications of the bootstrap would have been possible without the computing power of present day computers.

The rest of the paper is organized as follows. Section 2 presents and discusses residual bootstrap methods for parametric and nonparametric models. The proposals mainly apply the classical bootstrap approach of *drawing with replacement* to residuals of a fitted model to the data. As a special case, Section 3 considers in detail an approach by fitting autoregressions of increasing order to the observed data. A rather relevant model class of dependent observations to which bootstrap procedures successfully can be applied are Markov chains (cf. Section 4).

Section 5 discusses in detail the prominent block bootstrap methods for time series. So far, all discussed bootstrap methods are in time domain. Of course, frequency domain bootstrap methods exist and are presented in Section 6. Mixtures of both frequency and time domain bootstrap methods are described in Section 7. A final Section 8 concentrates on bootstrap methods for time series with long-range dependence.

## 2. Residual bootstrap for parametric and nonparametric models

Since the original bootstrap idea of Efron (1979) for i.i.d. random variables of *drawing with replacement* cannot be applied directly to dependent observations, because by obvious reasons, it suggests itself to apply the classical bootstrap principle to residuals of an (optimal) predictor of the $X_t's$.

Suppose for the following that we are given observations $X_1, \ldots, X_n$. For some fixed $p \in \mathbb{N}$ denote by $\widehat{m}_n(X_{t-1}, \ldots, X_{t-p})$, a parametric or nonparametric estimator of the conditional expectation $E[X_t | X_{t-1}, \ldots, X_{t-p}]$. This estimator leads to residuals

$$\widehat{e}_t := X_t - \widehat{m}_n(X_{t-1}, \ldots, X_{t-p}), t = p+1, \ldots, n, \tag{1}$$

and in a next step to a bootstrap time series

$$X_t^* = \widehat{m}_n(X_{t-1}^*, \ldots, X_{t-p}^*) + e_t^*, t = 1, \ldots, n. \tag{2}$$

The bootstrap innovations $e_1^*, \ldots, e_n^*$ follow a Laplace distribution over the set $\{\widehat{e}_{p+1}^c, \ldots, \widehat{e}_n^c\}$ of centered estimated residuals $\widehat{e}_{p+1}, \ldots, \widehat{e}_n$.

Here, we presumed that all residuals more or less share the same variance. In a heteroscedastic situation, one might think of some kind of a localized selection of bootstrap residuals or a *wild* bootstrap approach. The latter means that bootstrap innovations are generated according to

$$e_t^* := \widehat{e}_t \cdot \eta_t^*, t = p+1, \ldots, n, \tag{3}$$

where the (bootstrap) random variables $(\eta_t^*)$ possess zero mean and unit variance, only. Typically, it is not necessary to specify some distribution for the $\eta_t^*$'s. If a distributional assumption is made, this ranges from rather simple discrete (even two-point) distributions to standard normal distribution. For reasons of better higher order performance for properly studentized statistics, one additionally should ensure $E^*(\eta_t^*)^3 = 1$. The simple discrete distribution taking values $z_1 = (1+\sqrt{5})/2$ and $z_2 = (1-\sqrt{5})/2$ with probabilities $p_1 = (\sqrt{5}-1)/(2\sqrt{5})$ and $p_2 = (\sqrt{5}+1)/(2\sqrt{5})$, respectively, satisfies the assumption of zero mean and unit second and third moments.

If we decide to use a fully nonparametric estimator in (1), the probabilistic properties of the bootstrap time series (2) could be rather delicate to investigate, because we, in principle, could not control the behavior of nonparametric estimators in regions far away from the origin, because we do not have many underlying observations in such regions. This typically leads to not very reliable estimators in that regions, and therefore, the stability of the bootstrap process cannot easily be guaranteed (recall that

models of type (2) typically need some quite restrictive growth conditions on the behavior of the function $\widehat{m}_n(x_{t-1}, \ldots, x_{t-p})$). But in order to establish asymptotic consistency of this bootstrap proposal, we need at least some stability and typically moreover some mixing or weak dependence properties for the triangular array of dependent observations in the bootstrap world. Such conditions would be rather helpful in order to prove asymptotic results for the bootstrap process.

One way out of this problem is to define instead of (2), a regression model in the bootstrap world, i.e., to generate bootstrap observations according to

$$X_t^* = \widehat{m}_n(X_{t-1}, \ldots, X_{t-p}) + e_t^*, t = 1, \ldots, n. \tag{4}$$

Along this proposal, we do not obtain a time series in the bootstrap world any longer, but an advantage of this proposal over (2) is that the design variables (which are lagged original observations themselves) now indeed mimic the $p$-dimensional marginal distribution of the underlying data by construction.

The investigation of a residual bootstrap procedure is much simpler; hence, we decide to use a fully parametric estimator in (1). For example, an optimal linear approximation of the conditional expectation, i.e., an autoregressive fit of order $p$ to the underlying data. The estimator $\widehat{m}_n$ in this case simplifies to $\widehat{m}_n(x_1, \ldots, x_p) = \sum_{k=1}^{p} \widehat{a}_k x_{t-k}$. Using Yule-Walker parameter estimates $\widehat{a}_k$ in such a simple situation always leads to a stable and causal process in the bootstrap world (cf. Kreiss and Neuhaus (2006), Satz 8.7 and Bemerkung 8.8). But, of course, one can apply the idea of a parametric fit to the conditional expectation to other models including moving-average and ARMA models.

The question of main interest is in which situations and to what extent the described bootstrap proposals asymptotically work.

In order to ensure that a fitted *parametric* model generates according to (2) bootstrap data that are able to mimic all dependence properties of the underlying observations, one has to assume that the data-generating process itself belongs to the parametric class, i.e., possess a representation of the form

$$X_t = m_\theta(X_{t-1}, \ldots, X_{t-p}) + e_t, \, t \in \mathbb{Z}, \tag{5}$$

with i.i.d. innovations and parametric conditional mean function $m_\theta$, which of course is quite restrictive. However, it can be stated that the parametric residual bootstrap consistently mimics the process (5). An obvious extension of the residual bootstrap (including an estimator of the conditional deviation (volatility)) leads to a residual bootstrap which consistently mimics the following slight deviation of model (5)

$$X_t = m_\theta(X_{t-1}, \ldots, X_{t-p}) + s_\theta(X_{t-1}, \ldots X_{t-q}) \cdot e_t, \, t \in \mathbb{Z}. \tag{6}$$

In case, the data-generating process does not belong to class (5) or (6), a residual bootstrap making use of such a model fit asymptotically can only work if the asymptotic distribution of the parameter of interest does not vary if switching from the true underlying process to a process of type (5) or (6), respectively.

The simplest situation in this context one might think of is a causal (linear) autoregressive model of fixed and known order $p$ and with i.i.d. innovations $(e_t)$ (having zero

mean and at least finite second-order moments) for the data-generating process, i.e.,

$$X_t = \sum_{k=1}^{p} a_k X_{t-k} + e_{t-k}, t \in \mathbb{Z}. \tag{7}$$

Of course in such a situation, it suffices to consider an autoregressive process of the same order $p$ with consistently estimated parameters $\widehat{a}_k$ (e.g., Yule-Walker estimates) and consistently estimated distribution of the innovations in the bootstrap world. If the statistic of interest is the centered autocovariance or centered autocorrelation function evaluated at some lags, then it is known that the asymptotic distribution for these quantities is not the same for linear AR(p) processes of type (7) and, for example, general mixing processes. This means that the residual bootstrap based on an autoregressive fit in general does not lead to consistent results.

As long as one is interested in the distribution of the coefficients of the (linear) autoregressive fit itself and as long as the underlying model follows (7), even the wild bootstrap proposal (4) leads to valid approximation results. The bootstrap estimators in such a situation just are the coefficients of a linear regression of $X_t^*$ on $X_{t-1}, \ldots, X_{t-p}$. The reason is that the asymptotic distributions of Yule-Walker and least-squares estimators for the coefficients in linear autoregression and linear regression with i.i.d. errors coincide. For more general statistics, it is of course not true that the wild bootstrap proposal (4) leads to asymptotically valid results, because in the bootstrap world, we even do not generate a stochastic process.

The application of a residual resampling scheme (2) in principle is of course not limited to causal (linear) autoregressive processes but easily can be extended to a broad class of further parametric models (including ARMA, threshold, ARCH, and GARCH models). Relevant references for ARMA models are Bose (1988), Bose (1990), and Franke and Kreiss (1992). The multivariate ARMA situation is considered in Paparoditis and Streitberg (1991). Basawa et al. (1991), Datta (1996), and Heimann and Kreiss (1996) dealt with the situation of general AR(1) models in which the parameter value is not restricted to the stationary case. For first-order autoregressions with positive innovations, Datta and McCormick (1995a) considered a bootstrap proposal for an estimator specific to the considered situation. Finally, Franke et al. (2006) considered the application of the bootstrap to order selection in autoregression, and Paparoditis and Politis (2005) considered bootstrap methods for unit root testing in autoregressions. It is worth mentioning that the assumption of i.i.d. innovations is rather essential for the asymptotic validity of the described bootstrap proposals for most statistics of interest. For a bootstrap test for a unit root in autoregressions with weakly dependent errors, see Psaradakis (2001).

Finally, let us come back to the fully nonparametric situation. If the data-generating process follows a nonparametric model equation of the form

$$X_t = m(X_{t-1}, \ldots, X_{t-p}) + s(X_{t-1}, \ldots, X_{t-q}) \cdot e_t, t \in \mathbb{Z}, \tag{8}$$

again with i.i.d. innovations $(e_t)$ (having zero mean and unit variance) and known orders $p, q$, in order to define a bootstrap process according to (2) or (4), we have to apply nonparametric estimators of the underlying functional parameters $m : \mathbb{R}^p \to \mathbb{R}$ and $s : \mathbb{R}^q \to [0, \infty]$, which are conditional mean and conditional volatility function of

the process. For smooth mean functions $m$ and smooth volatility functions $v$, kernel-based estimators successfully could be applied, while for more general situations, wavelet-based estimators may be used. It can be expected that for almost all statistical quantities, a residual bootstrap based on a nonparametric model fit for (8) will lead to a consistent resampling procedure.

As far as nonparametric estimators are of interest, one can take advantage of the so-called *whitening by windowing* effect, which in many situations of interest implies that the dependence structure of the underlying process does not show up in asymptotic distributions of nonparametric estimates. Because of this, one might also take regression-type standard as well as wild residual bootstrap procedures like (4) into consideration, which are often much easier to implement because they completely ignore the underlying dependence structure. We refer to Franke et al. (2002a) and Franke et al. (2002b) for nonparametric kernel-based-bootstrap methods. Neumann and Kreiss (1998) and Kreiss (2000) considered to what extent the nonparametric regression type bootstrap procedures successfully can be applied to situations (8) as long as nonparametric estimators and tests for conditional mean and/or volatility functions in nonparametric autoregressions are considered. A local bootstrap approach to kernel estimation for dependent observations is suggested and investigated in Paparoditis and Politis (2000).

Nonparametric bootstrap applications to goodness-of-fit testing problems for mean and volatilty functions in models of the form (8) are derived and discussed in Kreiss and Neumann (1999) and Kreiss et al. (2008). Paparoditis and Politis (2003) applied the concept of block bootstrap (cf. Section 5) to residuals in order to deal with rather relevant unit root testing problems.

## 3. Autoregressive-sieve bootstrap

The main idea of autoregressive (AR)-sieve bootstrap follows the lines of residual bootstrap described in Section 2. Instead of applying the *drawing with replacement* idea to residuals of an in some sense optimal predictor, we restrict for the AR-sieve bootstrap to (optimal) linear predictors, given an increasing number of past values of the underlying process itself.

If we again assume that the underlying process is stationary and, moreover, has positive variance $\gamma(0) > 0$ and asymptotically (as $h \to \infty$) vanishing autocovariances $\gamma(h)$, then we obtain from Brockwell and Davis (1991), Prop. 5.1.1, that the matrix $\mathbf{\Gamma}_p = (\gamma(i-j))_{i,j=1,2,\ldots,p}$ is positive definite, and therefore, immediately the best (in mean square sense) linear predictor of $X_{j+1}$ given $p$ past values $\mathbf{X}_{j,p} = (X_j, \ldots, X_{j-p+1})$ exists, which is unique and is given by $\widehat{X}_{j+1} = \sum_{j=1}^{p} a_j(p) X_{t-j}$. The coefficients $(a_j(p) j = 1, 2, \ldots, p)$ efficiently can be calculated from

$$(a_1(p), a_2(p), \ldots, a_p(p))^T = \mathbf{\Gamma}_p^{-1}(\gamma(1), \gamma(2), \ldots, \gamma(p))^T.$$

Now, one way to generate bootstrap pseudo-time series is to select a set of $p$ starting values $X_1^*, X_2^*, \ldots, X_p^*$ and, given the past $X_1^*, X_2^*, \ldots, X_j^*$, $j \geq p$, to generate the next observation $X_{j+1}^*$ using an estimated version of the best linear predictor $\widehat{X}_{j+1} = \sum_{s=1}^{p} a_s(p) X_{j+1-s}^*$ plus an error term which is selected randomly from the set of centered estimated prediction errors $X_{t+1} - \widehat{X}_{t+1} = X_{t+1} - \sum_{s=1}^{p} a_s(p) X_{t+1-s}$. This idea together with the order $p$ converging to infinity as sample size $n$ increases

lead to the so-called AR-sieve bootstrap procedure, which can be summarized in the following steps.

**Step 1:** Select an order $p = p(n) \in \mathbb{N}$, $p \ll n$, and fit a $p$th order autoregressive model to $X_1, X_2, \ldots, X_n$. Denote by $\widehat{a}(p) = (\widehat{a}_j(p), j = 1, 2, \ldots, p)$, the Yule-Walker autoregressive parameter estimators, that is $\widehat{a}(p) = \widehat{\Gamma}(p)^{-1}\widehat{\gamma}_p$, where for $0 \le h \le p$,

$$\widehat{\gamma}_X(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(X_t - \overline{X}_n)(X_{t+|h|} - \overline{X}_n),$$

$\overline{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t$, $\widehat{\Gamma}(p) = (\widehat{\gamma}_X(r-s))_{r,s=1,2,\ldots,p}$ and $\widehat{\gamma}_p = (\widehat{\gamma}_X(1), \ldots, \widehat{\gamma}_X(p))'$.

**Step 2:** Let $\widetilde{\varepsilon}_t(p) = X_t - \sum_{j=1}^{p}\widehat{a}_j(p)X_{t-j}$ $t = p+1, p+2, \ldots, n$, be the residuals of the autoregressive fit and denote by $\widehat{F}_n$ the empirical distribution function of the centered residuals $\widehat{\varepsilon}_t(p) = \widetilde{\varepsilon}_t(p) - \overline{\varepsilon}$, where $\overline{\varepsilon} = (n-p)^{-1}\sum_{t=p+1}^{n}\widetilde{\varepsilon}_t(p)$

Let $(X_1^*, X_2^*, \ldots, X_n^*)$ be a set of observations from the time series $\mathbf{X}^* = \{X_t^* : t \in \mathbb{Z}\}$, where $X_t^* = \sum_{t=1}^{p}\widehat{a}_j(p)X_{t-j}^* + e_t^*$ and the $e_t^*$'s are independent random variables having identical distribution $\widehat{F}_n$.

**Step 3:** Let $T_n^* = T_n(X_1^*, X_2^*, \ldots, X_n^*)$ be the same estimator as the estimator $T_n$ of interest based on the pseudo-time series $X_1^*, X_2^*, \ldots, X_n^*$, and $\vartheta^*$ the analogue of $\vartheta$ associated with the bootstrap process $\mathbf{X}^*$. The AR-sieve bootstrap approximation of $\mathcal{L}_n = \mathcal{L}(c_n(\hat{\theta}_n - \theta))$ is then given by $\mathcal{L}_n^* = \mathcal{L}^*(c_n(T_n^* - \vartheta^*))$.

Using Yule-Walker estimators in Step 1 of the AR-sieve bootstrap is rather convenient. Besides simple, stable, and fast computation (using the Durbin–Levinson algorithm), it ensures that the complex polynomial $\widehat{A}_p(z) = 1 - \sum_{j=1}^{p}\widehat{a}_j(p)z^j$ has no roots on or within the unit disc $\{z \in \mathbb{C} : |z| \le 1\}$, i.e., the bootstrap process $\mathbf{X}^*$ is always a stationary and causal autoregressive process (cf. Kreiss and Neuhaus (2006), Satz 8.7 and Bemerkung 8.8).

The described AR-sieve bootstrap has been introduced by Kreiss (1988) and has been investigated from several points of view in Paparoditis and Streitberg (1991), Kreiss (1992), Paparoditis (1996), Bühlmann (1997), Kreiss (1997), Bühlmann (1998), Choi and Hall (2000), Gonçalves and Kilian (2007), Poskitt (2008), and recently in Kreiss et al. (2011). Park (2002) gives an invariance principle for the sieve bootstrap and Bose (1988) worked out the edgeworth correction of bootstrap in autoregressions. Kapetanios (2010) applied the idea of sieve bootstrap to long-memory processes.

The question of course is under what assumptions on the underlying stochastic process $(X_t : t \in \mathbb{Z})$ and for what kind of statistics $T_n(X_1, \ldots, X_n)$ can we successfully approximate the distribution $\mathcal{L}_n$ by that of $\mathcal{L}_n^*$? In almost all papers concerning AR-sieve bootstrap, it is assumed that $(X_t)$ is a *linear* autoregression of possibly infinite order, i.e.,

$$X_t = \sum_{j=1}^{\infty} a_j X_{t-j} + e_t, \tag{9}$$

with $(e_t)$ an i.i.d. sequence and absolutely summable coefficients $a_j$, which moreover typically are assumed to decrease polynomially or even exponentially fast. An exception is the sample mean $\overline{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t$, where Bühlmann (1997) showed that for this

specific statistic, the assumption of i.i.d. innovations $(e_t)$ can be relaxed to martingale differences.

Kreiss et al. (2011) used the fact that every purely nondeterministic, zero mean stationary process possessing a strictly positive and continuous spectral density has a unique Wold-type autoregressive representation of the form

$$X_t = \sum_{j=1}^{\infty} a_j X_{t-j} + \varepsilon_t, \tag{10}$$

with absolutely summable coefficients $a_k$ and a white noise process $(\varepsilon_t)$ consisting of zero mean, uncorrelated random variables. The representation (10) does by far not mean that the underlying process is a linear, causal AR($\infty$) process driven by i.i.d. innovations!

Kreiss et al. (2011) have shown that under rather mild regularity assumptions, the AR-sieve bootstrap asymptotically correctly mimics the behavior of the following so-called companion autoregressive process $(\widetilde{X}_t : t \in \mathbb{Z})$ defined according to

$$\widetilde{X}_t = \sum_{j=1}^{\infty} a_j \widetilde{X}_{t-j} + \widetilde{\varepsilon}_t, \tag{11}$$

where the innovation process $(\widetilde{\varepsilon}_t)$ consists of i.i.d. random variables whose marginal distribution coincides with that of $(\varepsilon_t)$, i.e., $\mathcal{L}(\varepsilon_t) = \mathcal{L}(\widetilde{\varepsilon}_t)$ and the coefficients are those of the Wold-type autoregressive representation (10). Note that the first- and second-order properties of the two stochastic processes $(\widetilde{X}_t)$ and $(X_t)$ are the same, i.e., autocovariances and the spectral density coincide. However, all probability characteristics beyond second-order quantities are not necessarily the same and, in general, will substantially differ. Kreiss et al. (2011) showed for a rather general class of statistics that the AR-sieve bootstrap asymptotically works if the asymptotic distribution of the statistics of interest is the same for the underlying process $(X_t)$ and the companion autoregressive process $(\widetilde{X}_t)$. This rather plausible check criterion for the AR-sieve bootstrap to work leads, for example, for the arithmetic mean under very mild assumptions (much weaker than martingale differences for the innovations) to consistency of the AR-sieve proposal. For autocorrelations, this check criterion shows that AR-sieve bootstrap works if the underlying process possesses any linear representation with i.i.d. errors not depending on whether this representation can be inverted to an AR($\infty$)–representation with i.i.d. errors or not. For further details, we refer to Kreiss et al. (2011).

## 4. Bootstrap for Markov chains

Extension of the Bootstrap methods from i.i.d. random variables to Markov chains was initiated by Kulperger and Prakasa Rao (1989) for the finite state space case. Suppose that $\{X_n\}_{n\geq 0}$ be a stationary Markov chain with a finite state space $S = \{s_1, \ldots, s_\ell\}$, where $\ell \in \mathbb{N}$ and where $\mathbb{N} \equiv \{1, 2, \ldots\}$ denotes the set of all natural integers. Let the $\ell \times \ell$ transition probability matrix of the chain be given by $\mathbb{P} = ((p_{ij}))$ and the stationary distribution by $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_\ell)$. Thus, for any $1 \leq i, j \leq \ell$, $p_{ij} = P(X_1 = s_j | X) = s_i)$

and $\pi_i = P(X_0 = s_i)$. The joint distribution of the chain is completely determined by the finitely many unknown parameters, given by the components of $\boldsymbol{\pi}$ and $\mathbb{P}$. Given a sample $X_0, \ldots, X_{(n-1)}$ of size $n$ from the Markov chain, we can estimate the population parameters $\pi_i$'s and $p_{ij}$'s as

$$\hat{\pi}_i = n^{-1} \sum_{k=0}^{n-1} \mathbb{1}(X_k = s_i) \quad \hat{p}_{ij} = n^{-1} \sum_{k=0}^{n-2} \mathbb{1}(X_k = s_i, X_{k+1} = s_j)/\hat{\pi}_i, \quad (12)$$

$1 \le i, j \le \ell$. The bootstrap observations $X_0^*, \ldots, X_{n-1}^*$ can now be generated using the estimated transition matrix and the marginal distribution. Specifically, first generate a random variable $X_0^*$ from the discrete distribution on $\{1, \ldots, \ell\}$ that assigns mass $\hat{\pi}_i$ to $s_i$, $1 \le i \le \ell$. Next, having generated $X_0^*, \ldots, X_{k-1}^*$ for some $1 \le k < n-1$, generate $X_k^*$ from the discrete distribution on $\{1, \ldots, \ell\}$ that assigns mass $\hat{p}_{ij}$ to $j$, $1 \le j \le \ell$, where $s_i$ is the value of $X_{k-1}^*$. The bootstrap version of a given random variable $T_n = t_n(\mathbf{X}_n; \theta)$ based on $(X_0, \ldots, X_{n-1})$ and a parameter $\theta$ of interest is now defined as

$$T_n^* = t_n(X_0^*, \ldots, X_{n-1}^*; \hat{\theta}_n)$$

where $\hat{\theta}_n$ is an estimator of $\theta$ based on $X_0, \ldots, X_{n-1}$. For example, for $T_n = n^{1/2}(\bar{X}_n - \mu)$, where $\bar{X}_n = n^{-1} \sum_{k=0}^{n-1} X_i$ and $\mu = EX_0$, we set $T_n^* = n^{1/2}(\bar{X}_n^* - \hat{\mu}_n)$, where $\bar{X}_n^*$ is the average of the $n$ bootstrap variables $X_k^*$'s and where $\hat{\mu}_n = \sum_{i=1}^{\ell} \hat{\pi}_i X_i$, the (conditional) expectation of $X_0^*$ given $\mathbf{X}_n$. This approach has been extended to the countable case by Athreya and Fuh (1992).

More recently, different versions of the Bootstrap method for Markov processes based on estimated transition probability functions have been extended to the case, where the state space is Euclidean. In this case, one can use the nonparametric function estimation methodology to estimate the marginal distribution and the transition probability function. For consistency of the method, see Rajarshi (1990), and for the second-order properties of the method, see Horowitz (2003). A "local" version of the method (called the Local Markov Bootstrap or MLB, in short) has been put forward by Paparoditis and Politis (2001b). The idea here is to construct the bootstrap chain by sequential drawing – having selected a set of bootstrap observations, the next observation is randomly selected from a "neighborhood of close values" of the observation(s) in the immediate past. Paparoditis and Politis (2001b) showed that the resulting bootstrap chain was stationary and Markov and also that it enjoyed some robustness with regard to the Markovian assumption. For more on the properties of the MLB, see Paparoditis and Politis (2001b).

A completely different approach to bootstrapping Markov chains was introduced by Athreya and Fuh (1992). Instead of using estimated transition probabilities, they formulate a resampling scheme based on the idea of regeneration. A well-known result (Athreya and Ney, 1978) on Markov chains literature says that for a large class of Markov chains satisfying the so-called *Harris recurrence condition*, successive returns to a recurrent state gives a decomposition of the chain into i.i.d. cycles (of random lengths). The regeneration-based bootstrap resamples these i.i.d. cycles to generate the bootstrap observations. Here, we describe it for a Markov Chain $\{X_n\}_{n \ge 0}$ with values in a general state space $S$, equipped with a countably generated $\sigma$-field $\mathcal{S}$. Let $\mathbb{P}(x, dy)$

denote the transition probability function, and let $\pi(\cdot)$ denote the stationary distribution of the Markov chain. Suppose that $\{X_n\}_{n\geq 0}$ is positive recurrent with a known "accessible atom" $A \in \mathcal{S}$; Here, a set $A \in \mathcal{S}$ is called an "accessible atom" if it satisfies

$$\pi(A) > 0 \quad \text{and} \quad \mathbb{P}(x,\cdot) = \mathbb{P}(y,\cdot) \quad \text{for all} \quad x, y \in A.$$

For a Harris recurrent Markov chain with a countable state space, this condition holds trivially. Define the successive return times to $A$ by

$$\tau_1 = \inf\{m \geq 1 : X_m \in A\} \quad \text{and}$$
$$\tau_{k+1} = \inf\{m \geq \tau_k : X_m \in A\},\ k \geq 1.$$

Then, by strong Markov property, the blocks $\mathbb{B}_k = \{X_i : \tau_k + 1 \leq i \leq \tau_{k+1}\}, k \geq 1$ are i.i.d. variables with values in the taurus $\cup_{k\geq 1} S^k$. The *regeneration-based bootstrap* resamples the collection of blocks

$$\left\{\mathbb{B}_k : \mathbb{B}_k \subset \{X_0, \ldots, X_{n-1}\}\right\}$$

with replacement to generate the bootstrap observations. Validity of the method for the sample mean in the countable state space case is established by Athreya and Fuh (1992). For second-order properties of the regeneration-based bootstrap, see Datta and McCormick (1995b), and its refinements in Bertail and Clemencon (2006). Bertail and Clemencon (2006) show that the regeneration-based bootstrap, with a proper definition of the bootstrap version, achieves almost the same level of accuracy as in the case of i.i.d. random variables for linear statistics. As a result, for Markov chains satisfying the requisite regularity conditions, one should use the regeneration-based bootstrap (with blocks of random lengths) instead of the block bootstrap methods described below which are applicable to more general processes but are not as accurate.

## 5. Block bootstrap methods

For time series that are *not* assumed to have a specific structural form, Künsch (1989) formulated a general bootstrap method, currently known as the *moving block bootstrap* or MBB, in short. Quite early in the bootstrap literature, Singh (1981) showed that resampling single observations, as considered by Efron (1979) for independent data, failed to produce valid approximations in presence of dependence. As a remedy for the limitation of the single-data-value resampling scheme for dependent time series data, Künsch (1989) advocated the idea of resampling blocks of observations at a time (see also Bühlmann and Künsch (1995)). By retaining the neighboring observations together within the blocks, the dependence structure of the random variables at short lag distances is preserved. As a result, resampling blocks allows one to carry this information over to the bootstrap variables. The same resampling plan was also independently suggested by Liu and Singh (1992), who coined the term "moving block bootstrap."

We now briefly describe the MBB. Suppose that $\{X_t\}_{t\in\mathbb{N}}$ is a stationary weakly dependent time series and that $\{X_1, \ldots, X_n\} \equiv \mathbf{X}_n$ are observed. Let $\ell$ be an integer

satisfying $1 \le \ell < n$. Define the overlapping blocks $\mathbb{B}_1, \ldots, \mathbb{B}_N$ of length $\ell$ contained in $\mathbf{X}_n$ as

$$\begin{aligned}
\mathbb{B}_1 &= (X_1, X_2, \ldots, X_\ell), \\
\mathbb{B}_2 &= \quad (X_2, \ldots, X_\ell, X_{\ell+1}), \\
&\ldots \qquad\qquad\qquad\qquad\qquad \ldots \\
\mathbb{B}_N &= \qquad\qquad\qquad\qquad\qquad (X_{n-\ell+1}, \ldots, X_n),
\end{aligned}$$

where $N = n - \ell + 1$. For simplicity, suppose that $\ell$ divides $n$. Let $b = n/\ell$. To generate the MBB samples, we select $b$ blocks at random with replacement from the collection $\{\mathbb{B}_1, \ldots, \mathbb{B}_N\}$. Since each resampled block has $\ell$ elements, concatenating the elements of the $b$ resampled blocks serially yields $b \cdot \ell$ bootstrap observations $X_1^*, \ldots, X_n^*$. Note that if we set $\ell = 1$, then the MBB reduces to the ordinary bootstrap method of Efron (1979) for i.i.d. data. However, for a valid approximation in the dependent case, it is typically required that

$$\ell^{-1} + n^{-1}\ell = o(1) \quad \text{as} \quad n \to \infty. \tag{13}$$

Some typical choices of $\ell$ are $\ell = Cn^{1/k}$ for $k = 3, 4$, where $C \in \mathbb{R}$ is a constant. Next, suppose that the random variable of interest is of the form $T_n = t_n(\mathbf{X}_n; \theta(P_n))$, where $P_n = \mathcal{L}(\mathbf{X}_n)$ denotes the joint probability distribution of $\mathbf{X}_n$. The MBB version of $T_n$ based on blocks of size $\ell$ is defined as

$$T_n^* = t_n(X_1^*, \ldots, X_n^*; \theta(\hat{P}_n)),$$

where $\hat{P}_n = \mathcal{L}(X_1^*, \ldots, X_n^* | \mathbf{X}_n)$, the conditional joint probability distribution of $X_1^*, \ldots, X_n^*$, given $\mathbf{X}_n$, and where we suppress the dependence on $\ell$ to ease the notation. In the general case, where $n$ is not a multiple of $\ell$, one may resample $b = b_0$ blocks, where $b_0 = \min\{k \ge 1 : k\ell \ge n\}$ and retain the first $n$ resampled data-values to define the bootstrap replicate of $T_n$.

To illustrate the construction of $T_n^*$ in a specific example, suppose that $T_n$ is the centered and scaled sample mean $T_n^{1/2}(\bar{X}_n - \mu)$. Then, the MBB version of $T_n$ is given by $T_n^* = n^{1/2}(\bar{X}_n^* - \tilde{\mu}_n)$, where $\bar{X}_n^*$ is the sample mean of the bootstrap observations and where $\tilde{\mu}_n = E_*(\bar{X}_n^*)$. It is easy to check that

$$\begin{aligned}
\tilde{\mu}_n &= N^{-1} \sum_{i=1}^{N} \big(X_i + \cdots + X_{i+\ell-1}\big)/\ell \\
&= N^{-1} \left[ \sum_{i=\ell}^{N} X_i + \sum_{i=1}^{\ell-1} \frac{i}{\ell} \big(X_i + X_{n-i+1}\big) \right],
\end{aligned} \tag{14}$$

which is different from $\bar{X}_n$ for $\ell > 1$. Lahiri (1991) established second-order correctness of the MBB approximation for the normalized sample mean, where the bootstrap sample mean is centered at $\tilde{\mu}_n$. The "naive" centering of $\bar{X}_n^*$ at $\bar{X}_n$ is not appropriate as it leads to a loss of accuracy of the MBB approximation (Lahiri, 1992). Second-order

correctness of the MBB approximation for studentized statistics has been established independently by Götze and Künsch (1996) for stationary processes and by Lahiri (1996) in multiple linear regression models with dependent errors.

Several variants of the block bootstrap method exist in the literature. One of the early versions of the block bootstrap, implicit in the work of Carlstein (1986), restricts attention to the collection of nonoverlapping blocks in the data, and resamples from this smaller collection to generate the bootstrap observations. This is known as the *nonoverlapping block bootstrap* (NBB). To describe it briefly, suppose that $\ell$ is an integer in $(1, n)$ satisfying (13). Also, for simplicity, suppose that $\ell$ divides $n$ and set $b = n/\ell$. The NBB samples are generated by selecting $b$ blocks at random with replacement from the collection $\{\tilde{\mathbb{B}}_1, \ldots, \tilde{\mathbb{B}}_b\}$, where

$$\begin{aligned}
\tilde{\mathbb{B}}_1 &= (X_1, \ldots, X_\ell), \\
\tilde{\mathbb{B}}_2 &= \qquad\qquad (X_{\ell+1}, \ldots, X_{2\ell}), \\
&\cdots \qquad\qquad\qquad\qquad\qquad\qquad \cdots \\
\tilde{\mathbb{B}}_b &= \qquad\qquad\qquad\qquad\qquad\qquad (X_{(b-1)\ell+1}, \ldots, X_n).
\end{aligned}$$

Because the blocks in the NBB construction do not overlap, it is easier to analyze theoretical properties of NBB estimators than those of MBB estimators of a population parameter. However, the NBB estimators typically have higher MSEs at any block size $\ell$ compared to their MBB counterparts (cf. Lahiri (1999)).

Other variants of the block bootstrap include the *circular block bootstrap* (CBB) and the *stationary bootstrap* (SB) of Politis and Romano (1992, 1994), the *matched block bootstrap* (MaBB) of Carlstein et al. (1998), the *tapered block bootstrap* (TBB) of Paparoditis and Politis (2001a), among others. The CBB and the SB are primarily motivated by the need to remove the uneven weighting of the observations at the beginning and at the end in the MBB (cf. (14)) and are based on the idea of periodic extension of the observed segment of the time series. Further, while most block bootstrap methods are based on blocks of a deterministic length $\ell$, the SB is based on blocks of random lengths that have a Geometric distribution with expected length $\ell$ satisfying (13). The biases of the variance estimators generated by the MBB, NBB, CBB, and SB are of the order $O(\ell^{-1})$, while the variances are of the order $O(n^{-1}\ell)$, where $\ell$ denotes the block size and $n$ the sample size. It turns out that the MBB and the CBB have asymptotically equivalent performance and are also the most accurate of these four methods. For relative merits of these four methods, see Lahiri (1999), Politis and White (2004), and Nordman (2009). The MaBB uses a stochastic mechanism to reduce the edge effects from joining independent blocks in the MBB, while the TBB shrinks the boundary values in a block towards a common value, like the sample mean, to achieve the same. Although somewhat more complex than the MBB or the CBB, both the MaBB and the TBB yield more accurate variance estimators, with biases of the order $O(\ell^{-2})$ and variances of the order $O(n^{-1}\ell)$. In this sense, both MaBB and TBB are considered second-generation block bootstrap methods.

Performance of the block bootstrap methods crucially depends on the choice of the block size and on the dependent structure of the process. Explicit formulas for MSE-optimal block sizes for estimating the variances of smooth functions of sample means are known for the MBB, CBB, NBB, and SB (Hall et al., 1995; Lahiri, 1999). Thus,

one can use these expressions to formulate plug-in estimators of the optimal block sizes (Patton et al., 2009; Politis and White, 2004). For the variance estimation problem, Bühlmann and Künsch (1999) formulated a method based on linearization of an estimator using its influence function, which is somewhat more general than the direct plug-in approach. But perhaps the most widely used method in this context is given by Hall et al. (1995) who develop a general empirical method for estimating the optimal block sizes for estimating *both* the variance and the distribution function. The Hall et al. (1995) method uses the subsampling method to construct an estimator of the MSE as a function of the block size, and then minimize it to produce the estimator of the optimal block size. An alternative method based on the Jackknife-after-bootstrap method (Efron, 1992; Lahiri, 2002) has been recently proposed by Lahiri et al. (2007). They call it a *nonparametric plug-in* (NPPI) method, as it works like a plug-in method, but at the same time, it does not require the user to find an exact expression for the optimal block size analytically. The key construction of the NPPI method combines more than one resampling method suitably and, thereby, implicitly estimates the population parameters that appear in the formulas for the optimal block sizes. Further, the NPPI method is applicable to block bootstrap estimation problems involving the variance, the distribution function, and the quantiles. However, it is a computationally intensive method as it uses a combination of bootstrap and Jackknife methods.

For further discussion of the block length selection rules for block bootstrap methods, see Lahiri (2003a, Chapter 7) and the references therein.

## 6. Frequency domain bootstrap methods

An alternative bootstrap method that completely avoids the difficult problem of block length selection is given by the *Frequency Domain Bootstrap* (FDB).

One can apply the FDB for inference on population parameters of a second-order stationary process that can be expressed as a functional of its spectral density. Here, we give a short description of the FDB (see Paparoditis (2002) for an overview on frequency domain bootstrap methods). Given the data $\mathbf{X}_n$, define its Fourier transform

$$Y_n(w) = n^{-1/2} \sum_{t=1}^{n} X_t \exp(-\iota w t), \quad w \in (-\pi, \pi]. \tag{15}$$

The formulation of the FDB is based on the following well-known results:

(i) the Fourier transforms $Y_n(\lambda_1), \ldots, Y_n(\lambda_k)$ are *asymptotically independent* for any set of distinct ordinates $-\pi < \lambda_1 < \cdots < \lambda_k \le \pi$ (cf. Brockwell and Davis (1991), Lahiri (2003b));

(ii) The original observations $\mathbf{X}_n$ admit a representation in terms of the transformed values $\mathbf{Y}_n = \{Y_n(w_j) : j \in \mathcal{I}_n\}$ as (cf. Brockwell and Davis (1991)),

$$X_t = n^{-1/2} \sum_{j \in \mathcal{I}_n} Y_n(w_j) \exp(\iota t w_j), \quad t = 1, \ldots, n \tag{16}$$

where $\iota = \sqrt{-1}$, $w_j = 2\pi j/n$, and $\mathcal{I}_n = \{-\lfloor (n-1) \rfloor/2, \ldots, \lfloor (n-1) \rfloor/2\}$.

Thus, one can express a given variable $R_n = r_n(\mathbf{X}_n; \theta)$ also in terms of the transformed values $\mathbf{Y}_n$ and resample from the $Y$-values to define the FDB version of $R_n$. Variants of the FDB method have been proposed and studied by Hurvich and Zeger (1987) and Franke and Härdle (1992). Under some regularity conditions, Dahlhaus and Janas (1996) established second-order correctness of the FDB for a class of estimators called the "ratio statistics." Ratio statistics are defined as the ratio of two "spectral mean" estimators of the form $\int_0^\pi g(w) I_n(w) dw$, where $g : [0, \pi) \to \mathbb{R}$ is an integrable function and where $I_n(w) = |Y(w)|^2$ is the periodogram of $\mathbf{X}_n$. A common example of a ratio estimator is the lag-$k$ sample autocorrelation coefficient, $k \geq 1$, given by

$$\hat{\rho}_n(k) = r_n(k)/r_n(0),$$

where, for any $m \geq 0$, $r_n(m) = n^{-1} \sum_{i=1}^{n-m} X_i X_{i+m}$ is a (mean-uncorrected) version of the sample autocovariance function at lag $m$. It is easy to check that $r_n(m) = 2 \int_0^\pi \cos(mw) I_n(w) dw$, and therefore, $\hat{\rho}_n(k)$ is a ratio-statistic estimating the population $k$th order lag autocorrelation coefficient $\rho(k) = EX_1 X_{1+k}/EX_1^2$, when $\{X_n\}$ is a zero-mean second-order stationary process.

Although the FDB avoids the problem of block length selection, second-order accuracy of the FDB distributional approximations is available only under restrictive regularity conditions (cf. Dahlhaus and Janas (1996)). Further, it is known (cf. Lahiri (2003a, Section 9.2)) that accuracy of the FDB for spectral means and ratio estimators is rather sensitive to deviations from the model assumptions. Frequency domain bootstrap methods can also be applied to testing problems, cf. Dette and Paparoditis (2009).

Paparoditis and Politis (1999) applied the idea of a localized bootstrap approach to periodogram statistics, while a more general version of the FDB is proposed by Kreiss and Paparoditis (2003), which adds an intermediate autoregressive model fitting step in an attempt to capture higher order cross-cumulants of the DFTs. Kreiss and Paparoditis (2003) show that the modified version of the FDB provides a valid approximation for a wider class of spectral mean estimators that includes the class of ratio estimators covered by the FDB. We elaborate on this in the next section.

## 7. Mixture of two bootstrap methods

So far, we discussed several bootstrap proposals which are either defined in time domain (like block-, residual, AR-sieve and Markovian bootstrap) or defined in frequency domain (like periodogram-bootstrap). In this section, we briefly discuss mixtures of two bootstrap proposals (so-called hybrid bootstrap procedures). The rational behind such proposals is to bring together advantages of resampling approaches from both fields.

The hybrid bootstrap procedure proposed in Kreiss and Paparoditis (2003) can be understood as an extension of AR-sieve bootstrap as well as an extension of frequency domain bootstrap. As described in Section 3, AR-sieve bootstrap uses an autoregressive fit in order to obtain residuals of this fit. It can be argued that these residuals under reasonably assumptions on the data-generating process can be regarded to behave

approximately like i.i.d. random variables. Since such an i.i.d. property for the residuals does (if at all) at most holds approximately, it might be advisable to add a further nonparametric step to the AR-sieve bootstrap which is able to correct for data features which cannot or are not represented by the autoregressive fit.

On the other hand, frequency domain bootstrap as described above mainly uses the fact that periodogram ordinates asymptotically behave like i.i.d. random variables. But neglecting the existing and only asymptotically vanishing dependence structure between contiguous periodogram ordinates leads to drawbacks of frequency domain bootstrap. Therefore, an additional step of fitting a parametric model (e.g., an autoregressive model) to the data and applying – in the spirit of Tukey's pre-whitening – a frequency domain bootstrap approach to the residuals of the fit partly is able to remove this remedy. If, for example, the true underlying spectral density has some dominant peaks, then pre-whitening leads to a considerable improvement of nonparametric spectral density estimators. An autoregressive fit really is able to catch the peaks of the spectral density rather well and the curve $I_n(\lambda)/\hat{f}_{\mathrm{AR}}(\lambda)$, cf. Step 5 below, is much smoother than $I_n(\lambda)$, thus much easier to estimate nonparametrically.

Based on this motivation, an autoregressive-aided frequency domain hybrid bootstrap can be described along the following five steps. It is worth mentioning that fitting an autoregression should be understood as a (convenient) example. Of course, fitting other parametric models may be regarded as a pre-stage of frequency domain bootstrap.

**Step 1:** Given the observations $X_1, \ldots, X_n$, we fit an autoregressive process of order $p$, where $p$ may depend on the particular sample at hand.

This leads to estimated parameters $\hat{a}_1(p), \ldots, \hat{a}_p(p)$ and $\hat{\sigma}(p)$, which are obtained from the common Yule-Walker equations. Consider the estimated residuals

$$\hat{\varepsilon}_t = X_t - \sum_{\nu=1}^{p} \hat{a}_\nu(p) X_{t-\nu},\ t = p+1, \ldots, n,$$

and denote by $\hat{F}_n$ the empirical distribution of the standardized quantities $\widehat{\varepsilon}_{p+1}, \ldots, \widehat{\varepsilon}_n$, i.e., $\widehat{F}_n$ has mean zero and unit variance.

**Step 2:** Generate bootstrap observations $X_1^+, X_2^+, \ldots, X_n^+$, according to the following autoregressive model of order $p$

$$X_t^+ = \sum_{\nu=1}^{p} \widehat{a}_\nu(p) X_{t-\nu}^+ + \widehat{\sigma}(p) \cdot \varepsilon_t^+ ,$$

where $(\varepsilon_t^+)$ constitutes a sequence of i.i.d. random variables with cumulative distribution function $\widehat{F}_n$ (conditionally on the given observations $X_1, \ldots, X_n$).

The bootstrap process $\mathbf{X}^+ = (X_t^+ : t \in \mathbb{Z})$ possesses the following spectral density:

$$\hat{f}_{\mathrm{AR}}(\lambda) = \frac{\hat{\sigma}^2(p)}{2\pi} \left| 1 - \sum_{\nu=1}^{p} \hat{a}_\nu(p) e^{-i\nu\lambda} \right|^{-2} ,\ \lambda \in [0, \pi].$$

Note that because we make use of the Yule-Walker parameter estimators in Step 1, it is always ensured that $\hat{f}_{\mathrm{AR}}$ is well-defined, i.e., the polynomial

$1 - \sum_{v=1}^{p} \hat{a}_v(p) z^v$ has no complex roots with magnitude less than or equal to one. Moreover, the bootstrap autocovariances $\gamma^+(h) = E^+ X_1^+ X_{1+h}^+, h = 0, 1, \ldots, p$ coincide with the empirical autocovariances $\hat{\gamma}_n(h)$ of the underlying observations. It should be noted that it is convenient, but not necessary to work with Yule-Walker parameter estimates. Any $\sqrt{n}$-consistent parameter estimates would suffice.

**Step 3:** Compute the periodogram of the bootstrap observations, i.e.,

$$I_n^+(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_t^+ e^{-i\lambda t} \right|^2 , \lambda \in [0, \pi].$$

**Step 4:** Define the following nonparametric estimator $\widehat{q}$

$$\widehat{q}(\lambda) = \frac{1}{n} \sum_{j=-N}^{N} K_h\Big(\lambda - \lambda_j\Big) \frac{I_n(\lambda_j)}{\widehat{f}_{\mathrm{AR}}(\lambda_j)} , \quad \text{for} \quad \lambda \in [0, \pi),$$

while for $\lambda = \pi$, $\hat{q}(\pi)$ is defined as twice the quantity on the right-hand side of the above equation taking into account that no Fourier frequencies greater than $\pi$ exist. Here and above, the $\lambda_j$'s denote the Fourier frequencies, $K : [-\pi, \pi] \to [0, \infty)$ denotes a probability density (kernel), $K_h(\cdot) = h^{-1} K(\cdot / h)$, and $h > 0$ is the so-called bandwidth.

**Step 5:** Finally, the bootstrap periodogram $I_n^*$ is defined as follows:

$$I_n^*(\lambda) = \widehat{q}(\lambda) I_n^+(\lambda), \ \lambda \in [0, \pi].$$

Under some standard assumptions, the validity of this hybrid bootstrap was shown in Kreiss and Paparoditis (2003) for spectral means (e.g., sample autocovariance and spectral distribution function)

$$\int_0^{\pi} \varphi(\omega) I_n(\omega) d\omega, \tag{17}$$

where it is necessary to fit (at least asymptotically) the correct model and for ratio statistics (e.g., sample autocorrelation)

$$\int_0^{\pi} \varphi(\omega) I_n(\omega) d\omega \Big/ \int_0^{\pi} I_n(\omega) d\omega \tag{18}$$

and kernel spectral estimators, where it is not necessary to fit the correct model.

As can be seen from Kreiss and Paparoditis (2003), the described hybrid bootstrap procedure works well, and indeed the effect that on one hand the nonparametric correction step in frequency domain corrects for features which cannot be represented by the autoregressive model and that on the other hand the superior properties of the

autoregressive bootstrap procedure show up can be observed. Especially, it is observed that the frequency domain part of the described hybrid bootstrap leads to a much less dependence of the hybrid bootstrap on the selected autoregressive order $p$ than for the parametric autoregressive bootstrap itself.

The so far described hybrid bootstrap procedure is applicable to statistics, which can be written as functions of the periodogram only. But of course, relevant statistics in time series analysis do not share this property as, for example, the simple sample mean of the observations. Therefore, one is interested in a resampling procedure which still uses some computational parts in frequency domain but which are able to produce bootstrap observations $X_1^*, \ldots, X_n^*$ in time domain. When we switch to the frequency domain, as is, for example, suggested in Step 3 above, then we have to take into account the fact that the periodogram $I_n^+$ does not contain all information about the bootstrap process $X^+$ that is contained in the bootstrap observations $X_1^+, \ldots, X_n^+$. But, we can write $I_n^+(\omega) = |J_n^+(\omega)|^2$, where

$$J_n^+(\omega) = \frac{1}{\sqrt{2\pi n}} \sum_{s=1}^{n} X_s^+ \exp^{-is\omega} \tag{19}$$

denotes the discrete Fourier-transform (DFT). And of course, there is a one-to-one correspondence between the $n$ observations of a time series and the DFT evaluated at the Fourier frequencies $\omega_j = 2\pi \frac{j}{n}$ (cf. (16)). The solution now is to apply a nonparametric correction in the frequency domain to the DFT instead of the periodogram and then use the one-to-one correspondence to get back to the time domain. The modified hybrid bootstrap procedure reads as follows:

**Step 1:** Fit an $AR(p)$ model to the data, compute the estimated residuals $\widehat{\epsilon}_t = X_t - \sum_{v=1}^{p} \widehat{a}_v(p) X_{t-v}, t = p+1, \ldots, n$.

**Step 2:** Generate bootstrap observations $X_1^+, \ldots, X_n^+$ according to $X_t^+ = \sum_{v=1}^{p} \widehat{a}_v(p) X_{t-v}^+ + \widehat{\sigma}(p)\epsilon_t^+$, $\epsilon_t^+$ i.i.d. with empirical distribution of standardized residuals.

**Step 3:** Compute the DFT $J_n^+(\omega)$ and the nonparametric correction term $\widetilde{q}(\omega) = \widehat{q}^{1/2}(\omega)$ at the fourier frequencies $\omega_j = 2\pi \frac{j}{n}$, $j = 1, \ldots, n$.

**Step 4:** Compute the inverse DFT of the corrected DFT $\widetilde{q}(\omega_1) J_n^+(\omega_1), \ldots, \widetilde{q}(\omega_n) J_n^+(\omega_n)$ to obtain bootstrap observations $X_1^*, \ldots, X_n^*$ according to

$$X_t^* = \sqrt{\frac{2\pi}{n}} \sum_{j=1}^{n} \widetilde{q}(\omega_j) J_n^+(\omega_j) e^{it\omega_j}, \quad t = 1, \ldots, n. \tag{20}$$

This modified hybrid bootstrap proposal works for spectral means and ratio statistics as the not modified hybrid bootstrap procedure of Kreiss and Paparoditis (2003) does. Instead of using representations of statistics in frequency domain, we now simply can compute statistics in the time domain. The paper Jentsch and Kreiss (2010), to which we refer for details, discusses the modified hybrid bootstrap procedure for the multivariate case which in many respects is different.

So far, we only have considered autoregressions as parametric models to which we apply nonparametric corrections in frequency domain. It is of course not necessary

that the underlying model follows an autoregressive scheme of finite or infinite order, because of the additional nonparametric correction step. Moreover, it is not necessary to stay with autoregressive models; this has been done for simplicity only. So concerning a hybrid bootstrap procedure, one may think of any parametric model fit in a first step and a nonparametric correction as has been described in a second step. In the univariate situation, the resulting hybrid bootstrap procedure will result in asymptotically correct approximation results for statistics of observations from linear processes, which can be written as functions of autocorrelations or of the standardized (having integral one) spectral density as well as typically for the sample mean. The main reason for that is that asymptotic distributions of such statistics only depend on second-order terms of the underlying stochastic process, and these quantities are correctly mimicked by a hybrid bootstrap proposals. In the multivariate case, the mentioned result concerning the dependence of asymptotic distribution on second-order terms of linear time series does not hold any more, and therefore, the multivariate situation is much more involved (cf. Jentsch and Kreiss (2010)). A related method that allows resampling in frequency domain to obtain bootstrap replicates in time domain is considered in Kirch and Politis (2011). The papers Sergides and Paparoditis (2008) and Kreiss and Paparoditis (2011) considered an autoregressive-aided frequency domain hybrid bootstrap procedure and the modified hybrid bootstrap procedure along the lines described in this section for locally stationary time series.

## 8. Bootstrap under long-range dependence

Let $\{X_t\}_{t\in\mathbb{N}}$ be a stationary process with $EX_1^2 \in (0,\infty)$, autocovariance function $r(\cdot)$, and spectral density function $f(\cdot)$. We say that the process $\{X_t\}_{t\in\mathbb{N}}$ is long-range dependent (LRD) if $\sum_{k=1}^{\infty} |r(k)| = \infty$ or if $f(\lambda) \to \infty$ as $\lambda \to 0$. Otherwise, $\{X_t\}_{t\in\mathbb{N}}$ is said to be short-range dependent (SRD). We also use the acronym LRD (SRD) for long- (respectively, short) range dependence. Limit behaviors of many common statistics and tests under LRD are different from their behaviors under SRD. For example, the sample mean of $n$ observations from a LRD process may converge to the population mean at a rate *slower* than $O_p(n^{-1/2})$, and similarly, with proper centering and scaling, the sample mean may have a *non-normal* limit distribution even when the population variance is finite. More specifically, we consider the following result on the sample mean under LRD. Let $\{Z_t\}_{t\in\mathbb{N}}$ be a zero mean unit variance Gaussian process with an autocovariance function $r_1(\cdot)$ satisfying

$$r_1(k) \sim Ck^{-\alpha} \quad \text{as} \quad k \to \infty, \tag{21}$$

for some $\alpha \in (0,1)$, where for any two sequences $\{s_n\}_{n\geq 1}$ in $\mathbb{R}$ and $\{t_n\}_{n\geq 1}$ in $(0,\infty)$, we write $s_n \sim t_n$ if $s_n/t_n \to 1$ as $n \to \infty$. Note that here $\sum_{k=1}^{\infty} |r_1(k)| = \infty$, and hence, the process $\{Z_t\}$ is LRD. Next suppose that the $X_t$ process derives from the $Z_t$ process through the transformation

$$X_t = H_q(Z_t), \quad t \in \mathbb{N}, \tag{22}$$

for some integer $q \geq 1$, where $H_q(x)$ is the $q$th Hermite polynomial, i.e., for $x \in \mathbb{R}$, $H_q(x) = (-1)^q \big(\exp(x^2/2)\big) \frac{d^q}{dx^q} \big(\exp(-x^2/2)\big)$. Results in Taqqu (1975, 1979) and Dobrushin and Major (1979) imply the following result on the sample mean:

THEOREM 1. *Suppose that $\{X_t\}_{t\in\mathbb{N}}$ admits the representation (22) for some $q \geq 1$. If $\alpha \in (0, q^{-1})$, then*

$$n^{q\alpha/2}(\bar{X}_n - \mu) \to^d W_q \tag{23}$$

*where $\mu = EX_1$ and where $W_q$ is defined in terms of a multiple Wiener-Ito integral with respect to the random spectral measure $W$ of the Gaussian white noise process as*

$$W_q = A^{-q/2} \int \frac{\exp(\iota(x_1 + \cdots + x_q)) - 1}{\iota(x_1 + \cdots + x_q)} \prod_{k=1}^{q} |x_k|^{(\alpha-1)/2} dW(x_1) \ldots dW_q(x_q) \tag{24}$$

*with $A = 2\Gamma(\alpha)\cos(\alpha\pi/2)$.*

For $q = 1$, $W_q$ has a normal distribution with mean zero and variance $2/[(1-\alpha)(2-\alpha)]$. However, for $q \geq 2$, $W_q$ has a non-normal distribution. Although the bootstrap methods described in the earlier sections are successful in a variety of problems under SRD, they need not provide a valid answer under LRD. The following result gives the behavior of the MBB approximation under LRD:

THEOREM 2. *Let $\bar{X}_n^*$ denote the MBB sample mean based on blocks of size $\ell$ and resample size $n$. Suppose that the conditions of Theorem 1 hold and that $n^\delta \ell^{-1} + \ell n^{1-\delta} = o(1)$ as $n \to \infty$ for some $\delta \in (0,1)$. Then,*

$$\sup_{x\in\mathbb{R}} \Big| P_* \Big( c_n(\bar{X}_n^* - \hat{\mu}) \leq x \Big) - P\Big( n^{q\alpha/2}(\bar{X}_n - \mu) \leq x \Big) \Big| = o(1) \quad as \quad n \to \infty \tag{25}$$

*for some sequence $\{c_n\}_{n\geq 1} \in (0,\infty)$ if and only if $q = 1$.*

Theorem 2 is a consequence of the results in Lahiri (1993). It shows that for any choice of the scaling sequence, the MBB method fails to capture the distribution of the sample mean whenever the limit distribution of $\bar{X}_n$ is non-normal. With minor modifications of the arguments in Lahiri (1993), it can be shown that the same conclusion also holds for the NBB and the CBB. Intuitively, this may not be very surprising. The heuristic arguments behind the construction of these block bootstrap methods show (cf. Section 5) that all three methods attempt to estimate the initial approximation $P_\ell^\infty$ to the joint distribution $P$ of $\{X_t\}_{t\in\mathbb{N}}$, but $P_\ell^\infty$ itself gives an inadequate approximation to $P$ under LRD. Indeed, for the same reason, the MBB approximation fails even for $q = 1$ with the natural choice of the scaling sequence $c_n = n^{q\alpha/2}$. In this case, the (limit) distribution can be captured by using the MBB only with specially constructed scaling sequences $\{c_n\}_{n\geq 1}$, where $c_n \sim [n/\ell^{1+q\alpha}]^{1/2}$ as $n \to \infty$. For the sample mean of an LRD *linear* process with a normal limit, Kim and Nordman (2011) recently established the validity of MBB. Formulation of a suitable bootstrap method that works for both

normal and non-normal cases is still an open problem. For related results on subsampling and empirical likelihood methods under LRD, see Hall et al. (1998), Nordman et al. (2007), and the references therein.

## Acknowledgment

The research is partially supported by US NSF grant number DMS 1007703.

## References

Athreya, K.B., Fuh, C.D., 1992. Bootstrapping markov chains: countable case. J. Stat. Plan. Inference 33, 311–331.

Athreya, K.B., Ney, P., 1978. A new approach of the limit theory of recurrent Markov chains. Trans. Am. Math. Soc. 245, 493–501.

Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H., Taylor, R.L., 1991. Bootstrapping unstable first-order autoregressive processes. Ann. Stat. 19, 1098–1101.

Berkowitz, J., Kilian, L., 2000. Recent developments in bootstrapping time series. Econom. Rev. 19, 1–48.

Bertail, P., Clemencon, S., 2006. Regenerative block bootstrap for Markov chains. Bernoulli 12, 689–712.

Bose, A., 1988. Edgeworth correction by bootstrap in autoregressions. Ann. Stat. 16, 1709–1722.

Bose, A., 1990. Bootstrap in moving average models. Ann. Inst. Stat. Math. 42, 753–768.

Bose, A., Politis, D.N., 1995. A review of the bootstrap for dependent samples. In: Bhat, B.R., Prakasa Rao, B.L.S. (Eds.), Stochastic Processes and Statistical Inference. New Age International Publishers, New Delhi, pp. 39–51.

Brockwell, P., Davis, R.A. (Eds.), 1991. Time Series: Theory and Methods, Springer, New York.

Bühlmann, P., 1997. Sieve bootstrap for time series. Bernoulli 3, 123–148.

Bühlmann, P., 1998. Sieve bootstrap for smoothing in nonstationary time series. Ann. Stat. 26, 48–83.

Bühlmann, P., 2002. Bootstraps for time series. Stat. Sci. 17, 52–72.

Bühlmann, P., Künsch, H.R., 1995. The blockwise bootstrap for general parameters of a stationary time series. Scand. J. Statist. 22, 35–54.

Bühlmann, P., Künsch, H.R., 1999. Block length selection in the bootstrap for time series. Comput. Stat. Data Anal. 31, 295–310.

Carey, V.J., 2005. Resampling methods for dependent data. J. Amer. Statist. Assoc. 100, 712–713.

Carlstein, E., 1986. The use of subseries values for estimating the variance of a general statistics from a stationary time series. Ann. Statist. 14, 1171–1179.

Carlstein, E., Do, K.-A., Hall, P., Hesterberg, T., Künsch, H.R., 1998. Matched-block bootstrap for dependent data. Bernoulli 4, 305–328.

Choi, E., Hall, P., 2000. Bootstrap confidence regions computed from autoregressions of arbitrary order. J. R. Stat. Soc. Ser. B 62, 461–477.

Dahlhaus, R., Janas, D., 1996. A frequency domain bootstrap for ratio statistics in time series analysis. Ann. Statist. 24, 1934–1963.

Datta, S., 1996. On asymptotic properties of bootstrap for AR(1) processes. J. Stat. Plan. Infer 53, 361–374.

Datta, S., McCormick, W.P., 1995a. Bootstrap inference for a first-order autoregression with positive innovations. J. Am. Statist. Assoc. 90, 1289–1300.

Datta, S., McCormick, W.P., 1995b. Some continuous edgeworth expansions for markov chains with application to bootstrap. J. Multivar. Anal. 52, 83–106.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and Their Application, Cambridge University Press, Cambridge, UK.

Dette, H., Paparoditis, E., 2009. Bootstrapping frequency domain tests in multivariate time series with an application to testing equality of spectral densities. J. R. Stat. Soc. Ser. B 71, 831–857.

Dobrushin, R.L., Major, P., 1979. Non-central limit theorems for non-linear functionals of Gaussian fields. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 50, 27–52.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Stat. 7, 1–26.

Efron, B., 1992. Jackknife-after-bootstrap standard errors and influence functions (disc. pp. 111–127). J. R. Stat. Soc. Ser. B 54, 83–111.
Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap, Chapman and Hall, New York.
Franke, J., Härdle, W., 1992. On bootstrapping kernel spectral estimates. Ann. Stat. 20, 121–145.
Franke, J., Kreiss, J.-P., 1992. Bootstrapping stationary autoregressive moving-average models. J. Time Ser. Anal. 13, 297–317.
Franke, J., Kreiss, J.-P., Mammen, E., 2002a. Bootstrap of kernel smoothing in nonlinear time series. Bernoulli 8, 1–37.
Franke, J., Kreiss, J.-P., Mammen, E., Neumann, M.H., 2002b. Properties of the nonparametric autoregressive bootstrap. J. Time Ser. Anal. 23, 555–585.
Franke, J., Kreiss, J.-P., Moser, M., 2006. Bootstrap order selection for autoregressive processes. Stat. Decis. 24, 305–325.
Gonçalves, S., Kilian, L., 2007. Asymptotic and bootstrap inference for AR($\infty$) processes with conditional heteroskedasticity. Econom. Rev. 26, 609–641.
Götze, F., Künsch, H., 1996. Second-order correctness of the blockwise bootstrap for stationary observations. Ann. Statist. 24, 1914–1933.
Hall, P., 1992. The Bootstrap and Edgeworth Expansion, Springer, New York.
Hall, P., Horowitz, J.L., Jing, B.-Y., 1995. On blocking rules for the bootstrap with dependent data. Biometrika 82, 561–574.
Hall, P., Jing, B.-Y., Lahiri, S.N., 1998. On the sampling window method under long range dependence. Stat. Sin. 8, 1189–1204.
Härdle, W., Horowitz, J., Kreiss, J.-P., 2003. Bootstrap for time series. Int. Stat. Rev. 71, 435–459.
Heimann, G., Kreiss, J.-P., 1996. Bootstrapping general first order autoregression. Stat. Prob. Lett. 30, 87–98.
Horowitz, J.L., 2003. Bootstrap methods for markov processes. Econometrica 71, 1049–1082.
Hurvich, C.M., Zeger, S.L., 1987. Frequency Domain Bootstrap Methods for Time Series. Preprint, Department of Statistics and Operations Research, New York University.
Jentsch, C., Kreiss, J.-P., 2010. The multiple hybrid bootstrap: Resampling multivariate linear processes. J. Mult. Anal. 101, 2320–2345.
Kapetanios, G., 2010. A generalization of a sieve bootstrap invariance principle to long memory processes. Quant. Qual. Anal. Soc. Sci. 4, 19–40.
Kim, Y.-M., Nordman, D.J., 2011. Properties of a block bootstrap method under long range dependence. Sankhya Ser. A. 73, 79–109.
Kirch, C., Politis, D.N., 2011. TFT-Bootstrap: Resampling time series in the frequency domain to obtain replicates in the time domain. Ann. Stat. 39, 1427–1470.
Kreiss, J.-P., 1988. Asymptotic Statistical Inference for a Class of Stochastic Processes, Habilitationsschrift, Universität Hamburg.
Kreiss, J.-P., 1992. Bootstrap procedures for AR($\infty$)-processes. In: Lecture Notes in Economics and Mathematical Systems, vol. 376 (Proc. Bootstrapping and Related Techniques, Trier), pp. 107–113.
Kreiss, J.-P., 1997. Asymptotical Properties of Residual Bootstrap for Autoregression. Preprint, TU Braunschweig.
Kreiss, J.-P., 2000. Nonparametric estimation and bootstrap for financial time series. In: Chan, W.S., Li, W.K., Tong, H., (Eds.), Statistics and Finance: An Interface. Imperial College Press, London.
Kreiss, J.-P., Neuhaus, G., 2006. Einführung in die Zeitreihenanalyse, Springer, Heidelberg.
Kreiss, J.-P., Neumann, M.H., 1999. Bootstrap tests for parametric volatility structure in nonparametric autoregression. In: Grigelionis, B. et al., (Eds.), Prob. Theory Math. Stat. pp. 393–404.
Kreiss, J.-P., Neumann, M.H., Yao, Q., 2008. Bootstrap tests for simple structures in nonparametric time series regression. Stat. Inter. 1, 367–380.
Kreiss, J.-P., Paparoditis, E., 2003. Autoregressive aided periodogram bootstrap for time series. Ann. Stat. 31, 1923–1955.
Kreiss, J.-P., Paparoditis, E., 2011. Bootstrapping Locally Stationary Time Series. Technical Report.
Kreiss, J.-P., Paparoditis, E., Politis, D.N., 2011. On the range of validity of the autoregressive sieve bootstrap. Ann. Stat. 39, 2103–2130.
Kulperger, R.J., Prakasa Rao, B.L.S., 1989. Bootstrapping a finite state Markov chain. Sankhya Ser. A 51, 178–191.
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. Stat. Prob. Lett. 11, 335–341.
Lahiri, S.N., 1992. Edgeworth correction by moving block bootstrap for stationary and nonstationary data. In: LePage, R., Billard, L. (Eds.), Exploring the Limits of Bootstrap. Wiley, New York, pp. 183–214.
Lahiri, S.N., 1993. On the moving block bootstrap under long range dependence. Stat. Prob. Lett. 18, 405–413.
Lahiri, S.N., 1996. On edgeworth expansions and the moving block bootstrap for studentized M–estimators in multiple linear regression models. J. Multivar. Anal. 56, 42–59.
Lahiri, S.N., 1999. Theoretical comparisons of block bootstrap methods. Ann. Statist. 27, 386–404.
Lahiri, S.N., 2002. On the jackknife after bootstrap method for dependent data and its consistency properties. Econom. Theory 18, 79–98.
Lahiri, S.N., 2003a. Resampling Methods for Dependent Data, Springer, New York.
Lahiri, S.N., 2003b. A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short- and long-range dependence. Ann. Stat. 31, 613–641.
Lahiri, S.N., Furukawa, K., Lee, Y.-D., 2007. A nonparametric plug-in method for selecting the optimal block length. Stat. Method 4, 292–321.
Li, H., Maddala, G.S., 1996. Bootstrapping time series models. Econom. Rev. 15, 115–158.
Liu, R.Y., Singh, K., 1992. Moving blocks jackknife and bootstrap capture weak dependence. In: LePage, R., Billard, L. (Eds.), Exploring the Limits of Bootstrap. Wiley, New York.
Neumann, M. H., Kreiss, J.-P., 1998. Regression-type inference in nonparametric autoregression. Ann. Stat. 26, 1570–1613.
Nordman, D.J., 2009. A note on the stationary bootstrap. Ann. Stat. 37, 359–370.
Nordman, D., Sibbersten, P., Lahiri, S.N., 2007. Empirical likelihood confidence intervals for the mean of a long range dependent process. J. Time Ser. Anal. 28, 576–599.
Paparoditis, E., 1996. Bootstrapping autoregressive and moving average parameter estimates of infinite order vector autoregressive processes. J. Multivar. Anal. 57, 277–296.
Paparoditis, E., 2002. Frequency domain bootstrap for time series. In: Dehling, H., Mikosch, T., Sørensen, M. (Eds.), Empirical Process Techniques for Dependent Data. Birkhäuser, Boston, pp. 365–381.
Paparoditis, E., Politis, D.N., 1999. The local bootstrap for periodogram statistics. J. Time Ser. Anal. 20, 193–222.
Paparoditis, E., Politis, D.N., 2000. The local bootstrap for kernel estimators under general dependence conditions. Ann. Inst. Statist. Math. 52, 139–159.
Paparoditis, E., Politis, D.N., 2001a. The tapered block bootstrap. Biometrika 88, 1105–1119.
Paparoditis, E., Politis, D.N., 2001b. A markovian local resampling scheme for nonparametric estimators in time series analysis. Econom. Theory 17, 540–566.
Paparoditis, E., Politis, D.N., 2003. Residual-based block bootstrap for unit root testing. Econometrica 71, 813–856.
Paparoditis, E., Politis, D.N., 2005. Bootstrapping unit root tests for autoregressive time series. J. Am. Statist. Assoc. 100, 545–553.
Paparoditis, E., Politis, D.N., 2009. Resampling and subsampling for financial time series. In: Andersen, T., Davis, R., Kreiss, J.-P., Mikosch, T. (Eds.), Handbook of Financial Time Series. Springer, New York, pp. 983–999.
Paparoditis, E., Streitberg, B., 1991. Order identification statistics in stationary autoregressive moving average models: vector autoccorrelations and the bootstrap. J. Time Ser. Anal. 13, 415–435.
Park, J.Y., 2002. An invariance principle for siebe bootstrap in time series. Econom. Theory 18, 469–490.
Patton, A., Politis, D.N., White, H., 2009. Correction to "Automatic block-length selection for the dependent bootstrap by D.N. Politis and H. White". Econom. Rev. 28, 372–375.
Politis, D.N., 2003. The impact of bootstrap methods on time series analysis. Stat. Sci. 18, 219–230.
Politis, D.N., Romano, J.P., 1992. A circular block resampling procedure for stationary data. In: Lepage, R., Billard, L., (Eds.), Exploring the Limits of Bootstrap. Wiley, New York, pp. 263–270.
Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. J. Am. Statist. Assoc. 89, 1303–1313.
Politis, D.N., White, H., 2004. Automatic block-length selection for the dependent bootstrap. Econom. Rev. 23, 53–70.
Poskitt, D.S., 2008. Properties of the sieve bootstrap for fractionally integrated and non-invertible processes. J. Time Ser. Anal. 29, 224–250.
Psaradakis, Z., 2001. Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors. J. Time Ser. Anal. 22, 577–594.

Rajarshi, M.B., 1990. Bootstrap in Markov sequences based on estimates of transition density. Ann. Inst. Statist. Math. 42, 253–268.

Ruiz, E., Pascual, L., 2002. Bootstrapping financial time series. J. Econ. Surveys 16, 271–300. Reprinted in: Contributions to Financial Econometrics: Theoretical and Practical Issues (Eds.: McAleer, M. and Oxley, L.), Blackwell.

Sergides, M., Paparoditis, E., 2008. Bootstrapping the local periodogram of locally stationary processes. J. Time Ser. Anal. 29, 264–299. Corrigendum: Journal of Time Series Analysis 30, 260–261.

Shao, J., Tu, D., 1995. The Jackknife and Bootstrap, Springer, New York.

Singh, K., 1981. On the asymptotic accuracy of Efron's bootstrap. Ann. Stat. 9, 1187–1195.

Taqqu, M.S., 1975. Weak convergence to fractional Brownian motion and to the Rosenblatt process. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 31, 287–302.

Taqqu, M.S., 1979. Convergence of integrated processes of arbitrary hermite rank. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 50, 53–83.