Skip-sampling: subsampling in the frequency domain

By Tucker McElroy
Research and Methodology Directorate, U.S. Census Bureau,
4600 Silver Hill Road, Washington, D.C. 20233-9100, USA
tucker.s.mcelroy@census.gov

and Dimitris N. Politis
Department of Mathematics and Halicioglu Data Science Institute,
University of California–San Diego, La Jolla, CA 92093-0112, USA
dpolitis@ucsd.edu

Summary

Over the last 35 years, several bootstrap methods for time series have been proposed. Popular time domain methods include the block-bootstrap, the stationary bootstrap, the linear process bootstrap, etc.; subsampling for time series is also available, and is closely related to the block-bootstrap. the frequency domain bootstrap (FDB) has been performed either by resampling the periodogram ordinates or by resampling the ordinates of the Discrete Fourier Transform (DFT). The paper at hand proposes a novel construction of subsampling the DFT ordinates, and investigates its theoretical properties and realm of applicability. Numerical studies show that the new method performs comparably to the FDB for linear spectral means and ratio statistics, while at the same time yielding significant computational savings as well as numerical stability.

Key words: Discrete Fourier Transform, Spectral Density, Time Series.

1 Introduction

Efron (1979) developed the bootstrap for independent and identically distributed (i.i.d.) data, and paved the way for practical nonparametric statistics in the modern era. Soon after, practitioners were able to apply resampling ideas in a variety of non-i.i.d. situations including the interesting case of dependent data.

Time series analysis has a time domain and a frequency domain aspect to it. Consequently, proposals for resampling time series can be in either of these two flavors. There have been several proposals with regards to time domain resampling plans; these include the block-bootstrap in its many variations, the stationary bootstrap, the linear process bootstrap, etc.; see Chapter 12 of McElroy and Politis (2020) for a concise description. Subsampling for time series is closely related to the block-bootstrap—see Politis et al. (1999), Lahiri (2003a), or Kreiss and Paparoditis (2024).

Franke and Härdle (1992) proposed resampling the periodogram ordinates, i.e., a frequency-domain bootstrap. The motivation behind this approach is that periodogram ordinates at different Fourier frequencies are approximately independent. Research on frequency-domain bootstrap was pursued by several researchers including Janas and Dahlhaus (1994), Dahlhaus and Janas (1996), Kreiss and Paparoditis (2003), Meyer et al. (2020), and Yu et al. (2023).

Interestingly, there is an early unpublished work by Hurvich and Zeger (1987) who proposed resampling the ordinates of the Discrete Fourier Transform (DFT), as they are also approximately independent. Actually, the aforementioned (approximate) independence of periodogram ordinates is a consequence of the (approximate) independence of DFT ordinates, since the periodogram is a function of the DFT. Hence, resampling the DFT can be thought of as a more fundamental construction; a rigorous development of DFT-based bootstrap can be found in Kirch and Politis (2011).

Since resampling the DFT is a fundamental construction, the question of possibly subsampling the DFT presents itself; this is the subject of the paper at hand. The basic idea is to divide the DFT (based on a sample of size T) into q vectors of length b, each consisting of the DFT ordinates at frequencies separated by q/T. Each such vector is asymptotically independent of one another, and distributed as a DFT vector based on a sample of size b. If the statistic at hand is computable based on the DFT alone, it could be re-computed on the smaller DFTs, and an empirical distribution of such subsample statistics, appropriately centered and normalized, would estimate the original statistic's sampling distribution.

The above construction will be termed skip-sampling of the DFT because of the process of skipping over some frequencies in putting together the subsample DFT vectors. The following section gives the precise construction as well as some background on the properties of the DFT. Some theoretical results on skip-sampling of the DFT are given in Section 3, while applications to spectral means and ratio statistics are given in Section 4; all proofs are deferred to Appendix A. Simulation results on the finite-sample performance of skip-sampling are reported in Section 5; some additional simulations are given in Appendix B. Finally, some useful results on DFT symmetries are given in Appendix C.

2 Problem setup

Let $X_1, \ldots, X_T$ be an observed sample from a strictly stationary time series $\{X_t\}$ defined on a probability space with probability measure P, and set $\mathbf{X} = [X_1, \ldots, X_T]'$. Time series analysis in the frequency domain hinges on the Discrete Fourier Transform (DFT), which maps the data vector $\mathbf{X}$ to a vector with (approximately) independent entries. To define the DFT, let $[\cdot]$ denote integer part, and consider the set of Fourier frequencies

$$\Lambda_T = \{\lambda_\ell = 2\pi\ell/T, \text{ where } \ell = [T/2] - T + 1, \ldots, [T/2]\};$$

this index range corresponds to $-[T/2], \ldots, [T/2]$ when T is odd, or $-[T/2] + 1, \ldots, [T/2]$ when T is even. Define a $T \times T$ unitary matrix Q (i.e., $Q^{-1} = Q^*$, the conjugate transpose) with complex-valued entries $Q_{jk} = T^{-1/2} \, e^{ij\lambda_{[T/2]-T+k}}$ for $j, k = 1, \ldots, T$. The DFT vector (see Proposition of 7.2.7 of McElroy and Politis (2020) for more details) is defined as $\widetilde{\mathbf{X}} = Q^* \, \mathbf{X}$, which means that the

jth component of the DFT vector is

$$\widetilde{X}_j = T^{-1/2} \sum_{k=1}^{T} e^{-ik\lambda_{[T/2]-T+j}} X_k. \tag{1}$$

The DFT map is invertible, because clearly $\mathbf{X} = Q\,\widetilde{\mathbf{X}}$.

We now provide details on the novel construction that is at the heart of this paper's methodology. For simplicity, consider positive integers q and b such that $b\,q = T$; if such a choice is not feasible, then one could let $q = [T/b]$, and work as if the data were just $X_1, \ldots, X_{bq}$, i.e., discard the data points $X_{bq+1}, \ldots, X_T$, and re-define T to equal $b\,q$. Define sub-components of the DFT vector by

$$\widetilde{\mathbf{X}}^{(j)} = [\widetilde{X}_j, \widetilde{X}_{q+j}, \ldots, \widetilde{X}_{(b-1)q+j}]' \tag{2}$$

for $j = 1, \ldots, q$. In terms of the entire DFT, this operation can be expressed as transforming $\widetilde{\mathbf{X}}$ into a $b \times q$ matrix such that the jth column is $\widetilde{\mathbf{X}}^{(j)}$, and the rows correspond to the Fourier frequencies $\Lambda_{T,\ell} = \{\lambda_{[T/2]-T+(\ell-1)q+1}, \ldots, \lambda_{[T/2]-T+\ell q}\}$ for $\ell = 1, \ldots, b$. Because of the construction of keeping every qth Fourier frequency and skipping over the intervening ones, this operation can be called skip-sampling on the DFT, and $\widetilde{\mathbf{X}}^{(j)}$ is called the jth skip-sample DFT; it is a complex vector of length b, obtained by evaluating the DFT only at Fourier frequencies $\lambda_{[T/2]-T+(\ell-1)q+j}$, where $\ell = 1, \ldots, b$.

Recall that the DFT $\widetilde{\mathbf{X}}$ contains all the information carried in the sample $\mathbf{X}$, since we can re-create $\mathbf{X}$ as $Q\,\widetilde{\mathbf{X}}$. However, the jth skip-sample DFT $\widetilde{\mathbf{X}}^{(j)}$ contains only a part of the information carried by the sample $\mathbf{X}$; putting all the skip-sample DFTs $\widetilde{\mathbf{X}}^{(j)}$ together for $j = 1, \ldots, q$, we can capture the whole information again. In this sense, working with the skip-sample DFTs $\widetilde{\mathbf{X}}^{(j)}$ for $j = 1, \ldots, q$ can be considered a form of subsampling in the frequency domain. This should be contrasted to the usual subsampling of a time series in the time domain, which is done by carving the sample $X_1, \ldots, X_T$ into smaller blocks, each consisting of b consecutive data points; see Politis and Romano (1994).

We can think of the jth skip-sample DFT $\widetilde{\mathbf{X}}^{(j)}$ as a smaller proxy for the entire DFT $\widetilde{\mathbf{X}}$. These proxies are asymptotically independent from one another but this independence property is not contingent on the fact that we skip along by exactly q frequencies; we could imagine a different organization of the components of $\widetilde{\mathbf{X}}$. Choosing the jth element of the set $\Lambda_{T,\ell}$, namely $\lambda_{[T/2]-T+(\ell-1)q+j}$, to represent the whole subset $\Lambda_{T,\ell}$ will be called *regular-draw* skip-sampling. However, we could instead have used another of the q elements of $\Lambda_{T,\ell}$ as its representative. Having q choices of representative for each set $\Lambda_{T,\ell}$ would lead to $S = q^b$ possible skip-sample statistics; we will call this scheme, whereby we use all S skip-sample statistics, *total-draw* skip-sampling.

Total-draw skip-sampling is, in general, not computationally feasible, as it involves the computation of q^b DFT ordinates with b and q diverging. However, we could construct an alternative scheme, wherein one member of each row is selected to form the first skip-sample DFT, followed by the second skip-sample DFT, thereby drawing one member from each row, ensuring these draws are distinct from the first skip-sample DFT. Such a mechanism will also have the asymptotic independence property. The key is that each skip-sample has one element from each row, which ensures that the frequencies are dispersed throughout $[-\pi, \pi]$, thereby avoiding estimation bias as discussed more below.

One such variant of regular-draw skip-sampling selects a random frequency in each row; we will call this *random-draw* skip-sampling, and denote it as

$$\widetilde{\mathbf{X}}^{(*j)} = [\widetilde{X}_{J_1}, \widetilde{X}_{q+J_2}, \ldots, \widetilde{X}_{(b-1)q+J_b}]'$$

where $J_1, \ldots, J_b$ are i.i.d. Uniform taking values in $\{1, \ldots, q\}$. We can now generate a large number (say B) such random-draw skip-sample DFTs so that the set $\{\widetilde{\mathbf{X}}^{(*j)}$ for $j = 1, \ldots, B\}$ is sufficient to approximate the set of total-draw skip-sample DFTs. This Monte Carlo construction is analogous to the stochastic approximation to the standard time domain subsampling; see Ch. 2.4 of Politis et al. (1999). The randomized strategy can produce many more skip-sample statistics, i.e., B instead of the q regular-draw skip-sample statistics; the result is more granularity in our estimates of the target distribution, which is practically useful as our simulation experiments confirm. Section 3 studies in detail the new skip-sampling methodology in its three variants: regular-draw, total-draw, and random-draw skip-sampling.

Although neither $\widetilde{\mathbf{X}}^{(j)}$ nor $\widetilde{\mathbf{X}}^{(*j)}$ will necessarily possess the Symmetry Property (see Definition C.1 in Appendix C), by employing the techniques of Remark C.1 we can ensure that applying the b-dimensional version of matrix Q to the symmetrized skip-sample DFT (so as to invert the DFT and bring us back to the time domain) will yield a real-valued vector of length b; this can be useful for statistics that are easier to formulate in the time domain. However, there is an interesting class of statistics that are defined in the frequency domain; two prime examples are discussed in Section 4.

3 Skip-sampling: methodology and key results

3.1 Framework

Let $X_1, X_2, \ldots, X_T$ be an observed sample from a strictly stationary time series $\{X_t\}$ with mean μ and absolutely summable autocovariance $\gamma_k = \mathrm{Cov}(X_0, X_k)$; the spectral density $f(\lambda) = \sum_{-\infty}^{\infty} \gamma_k e^{-ik\lambda}$ is well-defined and continuous on $[-\pi, \pi]$, i.e., belongs to $C[-\pi, \pi]$. A crude estimate of $f(\lambda)$ is the periodogram $I_T(\lambda) = \sum_{k=-T+1}^{T-1} \hat{\gamma}_k e^{-ik\lambda}$, where the sample autocovariance is defined as $\hat{\gamma}_k = T^{-1} \sum_{t=1}^{T-|k|} (X_t - \bar{X})(X_{t+|k|} - \bar{X})$, and $\bar{X} = T^{-1} \sum_{t=1}^{T} X_t$ is the sample mean.

When evaluated at a (nonzero) Fourier frequency $\lambda_\ell = 2\pi\ell/T$, the periodogram equals the squared magnitude of the DFT. To see that, note the identity $I_T(\lambda) = T^{-1} |\sum_{t=1}^{T} (X_t - \bar{X}) e^{-it\lambda}|^2$. One of the columns of the matrix Q consists of constant elements, and the other columns are orthogonal to it. Hence, $I_T(0) = 0$, and when $\ell \neq 0$ we have $I_T(\lambda_\ell) = T^{-1} |\sum_{t=1}^{T} X_t e^{-it\lambda_\ell}|^2$.

Let θ be a real-valued parameter of interest; extensions to multivariate parameters are straightforward. Let $\widehat{\theta}_T$ be an estimator of θ based on the data $X_1, \ldots, X_T$. The statistic $\widehat{\theta}_T$ is a function of the data vector $\mathbf{X}$; since the latter is in one-to-one correspondence with the DFT vector $\widetilde{\mathbf{X}}$ we may write

$$\widehat{\theta}_T = \mathcal{H}_T \left(\widetilde{\mathbf{X}} \right), \tag{3}$$

for some function $\mathcal{H}_T$ mapping $\mathbf{C}^T$ to $\mathbf{R}$. We will assume the following condition:

Assumption (A): For some nondegenerate limit distribution J, we have $a_T(\widehat{\theta}_T - \theta)$ converges in law to J as $T \to \infty$, where $a_T = T^\delta L(T)$ for some $\delta > 0$ and some slowly varying function L.

Letting $J_T(x) = P[a_T(\widehat{\theta}_T - \theta) \le x]$, Assumption (A) implies that $J_T(x) \to J(x)$ for all points x at which J is continuous.

3.2 Regular-draw skip-sampling

We can now define the jth regular-draw skip-sample statistic

$$\widehat{\theta}_b^{(j)} = \mathcal{H}_b\left(\widetilde{\mathbf{X}}^{(j)}\right) \tag{4}$$

for $j = 1, \ldots, q$. The idea is that $\widehat{\theta}_b^{(j)}$ will have the same asymptotic distribution as $\widehat{\theta}_b^{(1)}$ when $b \to \infty$. Furthermore, under standard conditions (see, e.g. Lahiri (2003b)), the DFT ordinates evaluated at different Fourier frequencies will be asymptotically independent; this would render the skip-sample statistics $\widehat{\theta}_b^{(1)}, \ldots, \widehat{\theta}_b^{(q)}$ (for fixed q) approximately independent as well.

We formulate these stylized facts in the following assumption, which operates under the condition

$$\frac{b}{T} + \frac{1}{b} \to 0 \text{ as } T \to \infty. \tag{5}$$

Recall that $q = T/b$; hence, eq. (5) implies that $q \to \infty$.

Assumption (A*): Under condition (5) the following are true: (a) For any j, $P[a_b(\widehat{\theta}_b^{(j)} - \theta) \le x] - J_b(x) = o(1)$ for all points x at which J is continuous; and (b) for any $j \ne k$, and any bounded functions g_1, g_2, we have $\mathrm{Cov}(g_1(\widehat{\theta}_b^{(j)}), g_2(\widehat{\theta}_b^{(k)})) \to 0$.

The verification of Assumption (A*) generally requires some further knowledge about the statistic and the time series dynamics. In Section 4 we establish the joint asymptotic normality and independence of distinct skip-sample statistics in the case of linear spectral mean statistics, from which Assumption (A*) follows.

The quantity $a_T(\widehat{\theta}_T - \theta)$ is sometimes called a root. Our core result is a consistency theorem for skip-sampling which gives conditions under which the empirical distribution of the regular-draw skip-sample roots $a_b(\widehat{\theta}_b^{(j)} - \widehat{\theta}_T)$ for $j = 1, \ldots, q$ can be used to approximate the distribution of the original root. To develop it, define the two regular-draw skip-sampling distributions:

$$U_{b,T}(x) = q^{-1} \sum_{j=1}^{q} \mathbf{1}\{a_b(\widehat{\theta}_b^{(j)} - \theta) \le x\} \text{ and } L_{b,T}(x) = q^{-1} \sum_{j=1}^{q} \mathbf{1}\{a_b(\widehat{\theta}_b^{(j)} - \widehat{\theta}_T) \le x\},$$

where $\mathbf{1}$ denotes the indicator function. Of the two skip-sampling distributions, $U_{b,T}$ is termed an oracle, as it requires knowledge of θ for its construction. By contrast, $L_{b,T}$ is a bona fide statistic that can be used for estimation purposes.

Theorem 3.1 *Assume condition (5) and Assumptions (A) and (A^*). Then, $L_{b,T}(x)$ converges in probability to $J(x)$ for all points x at which J is continuous.*

Remark 3.1 In many situations, the limit law J will be $N(0, v)$. In this case, it may be of interest to use skip-sampling to estimate the asymptotic variance v, and use the normal tables (instead of the quantiles of the skip-sampling distribution $L_{b,T}$) in order to construct confidence intervals and tests. The regular-draw skip-sampling estimator of v is

$$\widehat{v}_b = \frac{a_b^2}{q} \sum_{j=1}^{q} (\widehat{\theta}_b^{(j)} - \bar{\widehat{\theta}}_b)^2, \tag{6}$$

where $\bar{\widehat{\theta}}_b = q^{-1} \sum_{j=1}^{q} \widehat{\theta}_b^{(j)}$. The consistency of $\widehat{v}_b$ requires some different conditions that are outlined in Corollary 3.1 below. Such conditions can be verified for the two prominent types of periodogram-based statistics, namely spectral means and ratio statistics; see Sections 4.2 and 4.3. Additional examples of potential applicabilty of frequency domain resampling (including skip-sampling) are given in Corollary 3.1 of Bertail and Dudek (2021) and its related discussion.

For the next result, we strengthen (5) to the following condition:

$$\frac{b}{T} + \frac{1}{b} + \frac{a_b^2}{T} \to 0 \text{ as } T \to \infty. \tag{7}$$

Corollary 3.1 *Assume Assumption (A) with $\sup_T E\widehat{\theta}_T^4 < \infty$, and $a_T^2\, Var[\widehat{\theta}_T] \to v > 0$ as $T \to \infty$. Let b be a sequence satisfying (7), and assume that, for any $i, j = 1, \ldots, q$ and $i \neq j$, the following set of assumptions holds:*

$$\begin{cases} E[\widehat{\theta}_b^{(j)}] = \theta + o(a_b^{-1}) \\ a_b^2\, Var[\widehat{\theta}_b^{(j)}] = v + o(1) \\ Cov[\widehat{\theta}_b^{(i)}, \widehat{\theta}_b^{(j)}] = O(T^{-1}). \end{cases} \tag{8}$$

Further assume that when $i \neq j$,

$$Cov\left\{a_b^2(\widehat{\theta}_b^{(i)} - \theta)^2,\ a_b^2(\widehat{\theta}_b^{(j)} - \theta)^2\right\} = o(1). \tag{9}$$

Then, $\widehat{v}_b$ converges in probability to v as $T \to \infty$.

Remark 3.2 The validity of assumption (9) can be motivated by the asymptotic independence of $\widehat{\theta}_b^{(i)}$ and $\widehat{\theta}_b^{(j)}$. Moreover, the set of assumptions (8) implies $a_b^2 E(\widehat{\theta}_b^{(i)} - \theta)^2 = v + o(1)$. Hence, by Markov's inequality, $a_b^2(\widehat{\theta}_b^{(i)} - \theta)^2 = v + o_P(1)$, implying that $\left[a_b^2(\widehat{\theta}_b^{(i)} - \theta)^2 - v\right]\left[a_b^2(\widehat{\theta}_b^{(j)} - \theta)^2 - v\right] = o_P(1)$. Eq. (9) can be be viewed as a stronger version of this result. In Section 4 we provide a verification of (9) in the case of spectral mean statistics.

3.3 Total-draw and random-draw skip-sampling

There are $S = q^b$ possible total skip-samples. Order them in an arbitrary way and denote the jth skip-sample by $\widetilde{\mathbf{X}}^{(\sharp j)}$ where $j = 1, \ldots, S$. Consider a pair of total-draw skip-samples, say $\widetilde{\mathbf{X}}^{(\sharp j)}$ and $\widetilde{\mathbf{X}}^{(\sharp k)}$, and suppose they have r DFT components in common where $0 \leq r \leq b$. The number of such pairs is $S(q-1)^{b-r} b!/(r!(b-r)!)$; there are S choices for the first DFT vector, there is a choice of r out of b components of the second DFT vector that are in common with the first vector, and for each of the remaining $b - r$ components there are $q - 1$ possible choices. Summing over $r = 0, \ldots, b$ yields the total number of pairs, S^2. As in eq. (4), we define the jth total-draw skip-sample statistic as

$$\widehat{\theta}_b^{(\sharp j)} = \mathcal{H}_b\left(\widetilde{\mathbf{X}}^{(\sharp j)}\right) \tag{10}$$

for $j = 1, \ldots, S$. We also generalize Assumption (A*) to the following:

Assumption (A$^\sharp$): Under condition (5) the following are true: (a) For any j, $P[a_b(\widehat{\theta}_b^{(\sharp j)} - \theta) \leq x] - J_b(x) = o(1)$ for all points x at which J is continuous; and (b) for any bounded functions g_1, g_2, if $\widehat{\theta}_b^{(\sharp j)}$ and $\widehat{\theta}_b^{(\sharp k)}$ have r DFT components in common (for $r = 0, \ldots, b$), then $\mathrm{Cov}(g_1(\widehat{\theta}_b^{(\sharp j)}), g_2(\widehat{\theta}_b^{(\sharp k)})) = O(r/b)$.

Assumption (A$^\sharp$) implies Assumption (A*), since regular-draw skip-sample statistics are a particular sub-class of total-draw skip-sample statistics, corresponding to $r = 0$ shared DFT components. Define the two total-draw skip-sampling distributions, the oracle and the bona fide statistic, respectively as

$$U_{b,T}^\sharp(x) = S^{-1} \sum_{j=1}^{S} \mathbf{1}\{a_b(\widehat{\theta}_b^{(\sharp j)} - \theta) \leq x\} \text{ and } L_{b,T}^\sharp(x) = S^{-1} \sum_{j=1}^{S} \mathbf{1}\{a_b(\widehat{\theta}_b^{(\sharp j)} - \widehat{\theta}_T) \leq x\}.$$

As with Theorem 3.1, the consistency of $L_{b,T}^\sharp(x)$ follows from that of $U_{b,T}^\sharp(x)$.

Theorem 3.2 *Assume condition (5) and Assumptions (A) and (A$^\sharp$). Then, $L_{b,T}^\sharp(x)$ converges in probability to $J(x)$ for all points x at which J is continuous.*

Remark 3.3 Analogously to Remark 3.1, we may define the total-draw skip-sampling estimator of the large-sample variance v as

$$\widehat{v}_b^\sharp = \frac{a_b^2}{S} \sum_{j=1}^{S} (\widehat{\theta}_b^{(\sharp j)} - \bar{\widehat{\theta}}_b^\sharp)^2, \tag{11}$$

where $\bar{\widehat{\theta}}_b^\sharp = S^{-1} \sum_{j=1}^{S} \widehat{\theta}_b^{(\sharp j)}$. An analogue of Corollary 3.1 can now be formulated and proven. For brevity we omit further discussion here but show in Section 4 the consistency of the total-draw estimator $\widehat{v}_b^\sharp$ for linear spectral means.

Finally, recall that the computation of $L_{b,T}^\sharp(x)$ is in general infeasible as $S = q^b$ can be prohibitively large. As discussed earlier, we may resort to a stochastic approximation. To elaborate, let B be a large positive integer, and select $\widetilde{\mathbf{X}}^{(*1)}, \ldots, \widetilde{\mathbf{X}}^{(*B)}$ at random (with replacement) from the set $\{\widetilde{\mathbf{X}}^{(\sharp j)},\ j = 1, \ldots, S\}$. Then, let $\widehat{\theta}_b^{(*j)} = \mathcal{H}_b(\widetilde{\mathbf{X}}^{(*j)})$ for $j = 1, \ldots, B$, and define the random-draw skip-sampling distribution $L^*_{b,T}(x) = B^{-1}\sum_{j=1}^{B} \mathbf{1}\{a_b(\widehat{\theta}_b^{(*j)} - \widehat{\theta}_T) \le x\}$.

Corollary 3.2 *Assume the conditions of Theorem 3.2, and $B \to \infty$ as $T \to \infty$. Then, $L^*_{b,T}(x)$ converges in probability to $J(x)$ for all points x at which J is continuous.*

Corollary 3.2 remains true even if the sampling of the $\widetilde{\mathbf{X}}^{(*j)}$ is performed without replacement; see the proof of Corollary 2.4.1 of Politis et al. (1999).

Remark 3.4 Define the random-draw skip-sampling variance estimator

$$\widehat{v}_b^* = \frac{a_b^2}{B}\sum_{j=1}^{B}(\widehat{\theta}_b^{(*j)} - \bar{\widehat{\theta}}_b^*)^2 \tag{12}$$

where $\bar{\widehat{\theta}}_b^* = B^{-1}\sum_{j=1}^{B}\widehat{\theta}_b^{(*j)}$. It is easy to see that if $\widehat{v}_b^\sharp$ is consistent for the large-sample variance v, then so is $\widehat{v}_b^*$ as long as $B \to \infty$.

4 Application to statistics defined in the frequency domain

4.1 Framework

Throughout this section, suppose that $\{X_t\}$ is a stationary non-Gaussian processes. In many cases of interest, the parameter θ is a functional of the spectral density f, i.e., $\theta = \mathcal{G}(f)$ for some $\mathcal{G} : C[-\pi, \pi] \to \mathbf{R}$. Recall that the periodogram is asymptotically unbiased but inconsistent for $f(\lambda)$, as its variance does not tend to zero; see Chapter 9 of McElroy and Politis (2020). However, there are several situations where θ can be consistently estimated using the periodogram as a basis. In this case, our statistic $\widehat{\theta}_T$ may be defined as a functional of I_T, i.e.,

$$\widehat{\theta}_T = \mathcal{G}_T(I_T) \tag{13}$$

where, for each T, we have a functional $\mathcal{G}_T : C[-\pi, \pi] \to \mathbf{R}$. In simpler cases, the functional $\mathcal{G}_T$ might not depend on T, as in the case of spectral means and ratio statistics discussed in detail in Sections 4.2 and 4.3.

Since the periodogram evaluated at (nonzero) Fourier frequencies equals the squared magnitude of the DFT, it is apparent that the statistic (13) is a special case of the general statistic (3). In terms of feasible statistical computing, we will further assume, as it is invariably the case, that the statistic $\widehat{\theta}_T$ is computable based on the periodogram evaluated just on the Fourier frequencies.

4.2 Spectral means

Consider a bounded function $g(\lambda)$ of domain $[-\pi, \pi]$ that has bounded variation, and denote $\langle g \rangle = (2\pi)^{-1} \int_{-\pi}^{\pi} g(\lambda)\, d\lambda$. A linear spectral mean (Dahlhaus, 1985) is a parameter of the form

$$\theta = \langle g\, f \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\lambda)\, f(\lambda)\, d\lambda. \tag{14}$$

The prime example of a linear spectral mean is the autocovariance at lag k, where $g(\lambda) = e^{ik\lambda}$.

As already mentioned, the periodogram is asymptotically unbiased but inconsistent for $f(\lambda)$, as its variance does not tend to zero. However, plugging in I_T instead of f in eq. (14) yields a consistent estimator of the spectral mean, since integration works like summation in terms of reducing the variance. As a matter of fact, the integral in eq. (14) is typically approximated by a Riemann sum over the Fourier frequencies. Consequently, a linear spectral mean θ satisfying eq. (14) is practically estimated by

$$\widehat{\theta}_T = T^{-1} \sum_{n \in R_T} g(2\pi n/T)\, I_T(2\pi n/T), \tag{15}$$

where we define the index range $R_T = \{[T/2] - T + 1, \ldots, [T/2]\}$.

The recent paper by Yu et al. (2023) focuses on classical time domain subsampling, which is consistent for arbitrary statistics evaluated on stationary time series data under α-mixing; see Politis and Romano (1994). While employing the classical block-subsampling method, Yu et al. (2023) manage to relax the α-mixing and strict stationarity conditions when the statistic is a spectral mean; they also work out the moments of subsampling statistics, and make comparisons to the hybrid method of Meyer et al (2020). Our skip-sampling method, in contrast, is established upon summability conditions on autocumulants.

If the kth moment (for $k \geq 2$) exists, then the order k autocumulant function is defined as $\gamma_{h_1,\ldots,h_{k-1}} = \mathrm{cum}\{X_{t+h_1}, X_{t+h_2}, \ldots, X_{t+h_{k-1}}, X_t\}$, and we can formulate an autocumulant condition as in Taniguchi and Kakizawa (2000):

Assumption (Bk): for each $j = 1, \ldots, k-1$ we have

$$\sum_{h_1 \in \mathbb{Z}} \cdots \sum_{h_{k-1} \in \mathbb{Z}} (1 + |h_j|)\, |\gamma_{h_1,\ldots,h_{k-1}}| < \infty.$$

For a process satisfying Assumption (Bk) for some $k \geq 4$ the tri-spectral density

$$F(\omega_1, \omega_2, \omega_3) = \sum_{h_1 \in \mathbb{Z}} \sum_{h_2 \in \mathbb{Z}} \sum_{h_3 \in \mathbb{Z}} \gamma_{h_1,h_2,h_3} \exp\{-i(h_1\omega_1 + h_2\omega_2 + h_3\omega_3)\}$$

is well-defined. Moreover, if Assumption (Bk) holds for all $k \geq 2$, then Theorem A.1 of McElroy and Roy (2022) shows that, as $T \to \infty$,

$$T^{1/2} \left(\widehat{\theta}_T - \theta \right) \text{ converges weakly to } \mathcal{N}\left(0, \langle g\, g^{\star}\, f^2 \rangle + \langle\langle g\, g\, F \rangle\rangle\right). \tag{16}$$

In the above, we have used the short-hand $g^\star(\lambda) = g(\lambda) + g^\sharp(\lambda)$, where $g^\sharp$ is the reflection of g about the y-axis, and we have denoted

$$\langle\langle g\, g\, F \rangle\rangle = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g(\lambda)\, g(\omega)\, F(\lambda, -\lambda, \omega)\, d\lambda d\omega.$$

Dahlhaus (1985) proved eq. (16) for linear processes. Working under Assumption (Bk) allows us to go beyond the setting of linearity.

Recall from the discussion following (2) that the frequencies involved in $\widetilde{\mathbf{X}}^{(j)}$ are $\lambda_{[T/2]-T+(\ell-1)q+j}$ for $\ell = 1, \ldots, b$. Hence, evaluating the linear spectral mean on the jth regular-draw skip-sample DFT yields

$$b^{-1} \sum_{\ell=1}^{b} g(\lambda_{[b/2]-b+\ell})\, I_T(\lambda_{[T/2]-T+(\ell-1)q+j}). \tag{17}$$

We can rewrite (17) to more closely resemble (15): it is easy to show that $[T/2] = [b/2]q + [q/2]\mathbf{1}\{b \text{ or } T \text{ is odd}\}$, and therefore $[T/2] - T + (\ell-1)q + j = ([b/2] - b + \ell)q + \tilde{j}$ for $\ell = 1, \ldots, b$, where $\tilde{j} = j - q + [q/2]\mathbf{1}\{b \text{ or } T \text{ is odd}\}$. Using $nq/T = n/b$, (17) equals $b^{-1} \sum_{n \in R_b} g(2\pi n/b)\, I_T(2\pi n/b + 2\pi\tilde{j}/T)$ with $R_b = \{[b/2] - b + 1, \ldots, [b/2]\}$. The displacement by $2\pi\tilde{j}/T$ means that the periodogram is no longer an even function of n (unless $\tilde{j} = q$), so the asymptotic variance has no contribution from $g^\sharp$. In order to correct this, we symmetrize the periodogram by replacing $\tilde{j}$ by $-\tilde{j}$ when $n < 0$; making this change (and dropping the $n = 0$ term) suggests the definition of the jth regular-draw skip-sample statistic

$$\widehat{\theta}_b^{(j)} = b^{-1} \sum_{\ell=1}^{[b/2]} g^\star(2\pi\ell/b)\, I_T(2\pi\ell/b + 2\pi\tilde{j}/T). \tag{18}$$

Next, we summarize some moment properties of these skip-sample statistics.

Theorem 4.1 *Assume that $\{X_t\}$ is strictly stationary and satisfies Assumption (Bk) for $k = 2, \ldots, 8$. Consider a linear spectral mean θ satisfying eq. (14), and some fixed bounded function $g(\lambda)$ having bounded variation. Let b be a sequence satisfying (5). Then, for any $j = 1, \ldots, q$, we have*

$$E[\widehat{\theta}_b^{(j)}] = \theta + O(b^{-1}) + O(T^{-1})$$
$$Var[\widehat{\theta}_b^{(j)}] = b^{-1}\, \langle g\, g^\star\, f^2 \rangle + T^{-1}\, \langle\langle g\, g\, F \rangle\rangle + O(T^{-2}) + O(b^{-2}) + O(b^{-1}T^{-1}).$$

Also, for $i, j = 1, \ldots, q$ and $i \neq j$, we have

$$Cov\left(\widehat{\theta}_b^{(i)}, \widehat{\theta}_b^{(j)}\right) = O(T^{-1});\ Cov\left({\widehat{\theta}_b^{(i)}}^2, \widehat{\theta}_b^{(j)}\right) = O(T^{-1});\ Cov\left({\widehat{\theta}_b^{(i)}}^2, {\widehat{\theta}_b^{(j)}}^2\right) = O(T^{-1}).$$

Denote the left-hand-side of (16) by the root $S_T(\theta) = T^{1/2}\, (\widehat{\theta}_T - \theta)$. As a consequence of Theorem 4.1, the jth oracle skip-sample root $S_T^{(j)}(\theta) = b^{1/2}\, (\widehat{\theta}_b^{(j)} - \theta)$ has asymptotic variance

$$\langle g\, g^\star\, f^2 \rangle + (b/T)\, \langle\langle g\, g\, F \rangle\rangle. \tag{19}$$

Theorem 4.1 allows us to take advantage of the avenue suggested by Remark 3.1, i.e., use the skip-sampling estimator of the asymptotic variance of eq. (16), and then use the normal tables for inference. Here the rate $a_T = T^{1/2}$. Letting $b = o(T^{1/2})$ to satisfy (7), Corollary 3.1 is applicable as long as assumption (9) is verified; but this follows from the three asymptotic covariances of Theorem 4.1.

However, there is an additional issue: since $b/T \to 0$, the second term of (19) asymptotically drops out, which is undesirable in terms of capturing the variance given in (16). To elaborate, for roots such that $\langle\langle g\, g\, F\rangle\rangle = 0$ the asymptotic variance of $S_T^{(j)}(\theta)$ is correct, but otherwise must be adjusted to account for the non-trivial contribution from the tri-spectrum. Whenever $\langle\langle g\, g\, F\rangle\rangle = 0$, we will say the asymptotic distribution (16) is tri-spectrum free.

Corollary 4.1 *Assume the assumptions of Theorem 4.1 and $b = o(T^{1/2})$. Then the skip-sampling estimator $\hat{v}_b$ from eq. (6) is consistent for the asymptotic variance appearing in eq. (16) when the latter is tri-spectrum free.*

Remark 4.1 Requiring that the asymptotic distribution be tri-spectrum free is common with several resampling methods in the frequency domain. For example, the original frequency domain bootstrap of Franke and Härdle (1992) fails to capture the second term of (19) even for linear processes; see Paparoditis (2002) for a review. In general, by the Wold decomposition we have $X_t = EX_0 + \sum_{j\geq 0} \psi_j \epsilon_{t-j}$, where the sequence ϵ_t is mean zero, uncorrelated with variance σ^2, i.e., a white noise, but not necessarily i.i.d. If the process $\{X_t\}$ is linear (and causal), then $\{\epsilon_t\}$ is i.i.d. as well, and the expression for the variance greatly simplifies. In this case, $\langle\langle g\, g\, F\rangle\rangle = (\eta - 3)\, \langle g\, f\rangle^2$, where $\eta = E[\epsilon_t^4]/\sigma^4$. This yields a classical result: for linear time series, if the innovation kurtosis is that of a Gaussian (i.e., $\eta = 3$), or in the special case when the linear spectral mean is zero (i.e., $\langle g\, f\rangle = 0$), then the asymptotic distribution (16) is tri-spectrum free and Corollary 4.1 is applicable.

If $\{X_t\}$ is linear but $\eta \neq 3$, it may still be possible to conduct inference on spectral means via a hybrid procedure employing skip-sampling as a component. For example, let $\widehat{\eta}$ be the estimator of η based on the technique of Fragkeskou and Paparoditis (2016), and let $\widehat{f}$ be a consistent estimator of the spectral density f. Then, we can estimate the asymptotic variance appearing in eq. (16) by $\widehat{v}_b + (\widehat{\eta} - 3)\, \langle g\, \widehat{f}\rangle^2$. Alternative hybrid methods are also available, see Kreiss and Paparoditis (2003), or Meyer et al. (2020). A further result for skip-sampling can be formulated if we extend the cumulant conditions to all orders.

Theorem 4.2 *Assume that $\{X_t\}$ is strictly stationary and satisfies Assumption (Bk) for all $k \geq 2$. Consider a linear spectral mean θ satisfying eq. (14), and some fixed bounded function $g(\lambda)$ having bounded variation. Let b be a sequence satisfying (5). Then, for any $j \neq k = 1, \ldots, q$, the skip-sample roots $S_T^{(j)}(\theta)$ and $S_T^{(k)}(\theta)$ are jointly asymptotically normal with limiting covariance matrix equal to $\langle g g^\star f^2\rangle$ times the bivariate identity matrix.*

Statement (a) of Assumption (A*) follows from the weak convergence in Theorem 4.2, and statement (b) follows from the asymptotic independence. Hence, the conditions of Theorem 4.2 imply the validity of Theorem 3.1.

Remark 4.2 (Total-draw skip-sampling) Analogues of Theorem 4.1 and Corollary 4.1 can be established for total-draw skip-sampling; under the same assumptions, and with $\widehat{\theta}_b^{(\sharp j)}$ defined using symmetrization of the periodogram as in (18), $E[\widehat{\theta}_b^{(\sharp j)}] = \theta + O(b^{-1}) + O(T^{-1})$ and $\mathrm{Cov}\left(\widehat{\theta}_b^{(\sharp i)\,u}, \widehat{\theta}_b^{(\sharp j)\,v}\right)$ for $u, v = 1, 2$ is $O(r/b^2)$ for $r = 1, \ldots, b$ and $O(1/T)$ for $r = 0$, where r is the number DFT components in common between $\widehat{\theta}_b^{(\sharp i)}$ and $\widehat{\theta}_b^{(\sharp j)}$. Furthermore, the total-draw skip-sampling estimator $\widehat{v}_b^{\sharp}$ is consistent for the large-sample variance v when the latter is tri-spectrum free.

4.3 Ratio statistics

Consider a parameter θ that is obtained as the finite ratio of two linear spectral means, i.e., $\theta = \langle p\, f\rangle / \langle m\, f\rangle$ for some fixed bounded functions $p(\lambda)$ and $m(\lambda)$ having bounded variation on $[-\pi, \pi]$. We can estimate θ by the ratio statistic

$$\widehat{\theta}_T = \frac{\sum_{n \in R_T} p(2\pi n/T)\, I_T(2\pi n/T)}{\sum_{n \in R_T} m(2\pi n/T)\, I_T(2\pi n/T)}. \tag{20}$$

The prime example of a ratio statistic is the sample autocorrelation at lag k, where $p(\lambda) = e^{ik\lambda}$ and $m(\lambda) = 1$.

Ratio statistics have an asymptotic distribution that can be tri-spectrum free under some conditions, such as linearity of the time series, and are thus amenable to frequency domain resampling. In fact, Dahlhaus and Janas (1996) showed that the original frequency domain bootstrap of Franke and Härdle (1992) is not only consistent, but higher-order accurate for ratio statistics from linear time series. To elaborate, $\widehat{\theta}_T - \theta$ equals the ratio of two linear spectral mean statistics (15) with respective weighting functions $g = p - m\theta$ (for the numerator) and m (for the denominator), as shown in the proof of Corollary 9.6.9 of McElroy and Politis (2020). Using Slutsky's theorem for the denominator, Remark 4.1 implies that, provided the process is linear, ratio statistics satisfy a simplified version of (16), viz.

$$T^{1/2}\left(\widehat{\theta}_T - \theta\right) \text{ converges weakly to } \mathcal{N}\left(0, \langle g\, g^{\star}\, f^2\rangle / \langle m\, f\rangle^2\right) \text{ as } T \to \infty; \tag{21}$$

the notation $g^{\star}$ was defined right after eq. (16). In analogy with the previous subsection, we define the jth regular-draw skip-sample ratio statistic via

$$\widehat{\theta}_b^{(j)} = \frac{b^{-1} \sum_{\ell=1}^{[b/2]} p^{\star}(2\pi\ell/b)\, I_T(2\pi\ell/b + 2\pi\tilde{j}/T)}{b^{-1} \sum_{\ell=1}^{[b/2]} m^{\star}(2\pi\ell/b)\, I_T(2\pi\ell/b + 2\pi\tilde{j}/T)}.$$

Theorem 4.3 *Assume that $\{X_t\}$ is a strictly stationary linear process that satisfies Assumption (Bk) for $k = 2, \ldots, 8$. Consider a finite ratio of linear spectral means $\theta = \langle p\, f\rangle / \langle m\, f\rangle$, for some fixed bounded functions $p(\lambda)$ and $m(\lambda)$ having bounded variation, and let $\widehat{\theta}_T$ be the ratio statistic (20). Let b be a sequence* *satisfying (5), and set $g = p - m\theta$. Then, for any $j = 1, \ldots, q$, we have*

$$E[\widehat{\theta}_b^{(j)}] = \theta + O(b^{-1}) + O(T^{-1})$$
$$Var[\widehat{\theta}_b^{(j)}] = b^{-1} \langle g\, g^\star f^2 \rangle / \langle m\, f \rangle^2 + O(T^{-2}) + O(b^{-2}) + O(b^{-1}T^{-1}).$$

Also, for $i, j = 1, \ldots, q$ and $i \neq j$, we have

$$Cov\left(\widehat{\theta}_b^{(i)}, \widehat{\theta}_b^{(j)}\right) = O(T^{-1});\ Cov\left({\widehat{\theta}_b^{(i)}}^{2}, \widehat{\theta}_b^{(j)}\right) = O(T^{-1});\ Cov\left({\widehat{\theta}_b^{(i)}}^{2}, {\widehat{\theta}_b^{(j)}}^{2}\right) = O(T^{-1}).$$

Theorem 4.3 confirms the validity of assumption set (8) in the context of ratio statistics; the following corollary then ensues.

Corollary 4.2 *Assume the assumptions of Theorem 4.3 and $b = o(T^{1/2})$. Then the skip-sampling estimator $\hat{v}_b$ from eq. (6) is consistent for the asymptotic variance appearing in eq. (21).*

Akin to Remark 4.2, total-draw skip-sampling can be applied for ratio statistics, yielding analogues of Theorem 4.3 and Corollary 4.2; details are omitted.

5 Numerical Studies

In this section, simulations are carried out to compare skip-sampling to the frequency domain bootstrap (FDB), focusing on the lag one autocovariance and autocorrelation; these statistics are a spectral mean and a ratio of spectral means, respectively. For conciseness, we only present results here on random-draw skip-sampling. However, Appendix B contains numerical results based upon regular-draw skip-sampling as well. The random-draw skip-sampling estimator is constructed as outlined in Sections 3.3, 4.2, and 4.3. Recall that for large B the random-draw method behaves like the total-draw method, whose use is justified by Corollary 3.2 and Remark 4.2; below, we use $B = 1{,}000$ random skip-samples.

Following Kirch and Politis (2011), the FDB is constructed from the DFT: the DFT is transformed by the reciprocal square root of a spectral density estimate (we use a Parzen window, as well as a trapezoidal window with cutoff of .25, employing a positive definite modification); this transformed DFT is de-meaned and an empirical distribution function is formed from its real and imaginary parts. Bootstrap draws from the real and imaginary empirical distribution functions are then spliced into complex vectors with the Symmetry Property (see Definition C.1 in Appendix C) enforced by employing Remark C.1. Finally, these draws are back-transformed using the same square root spectral density estimate, and the appropriate root is constructed.

The FDB employs a loop over the number of bootstrap draws, which can be expensive for larger sample sizes; in contrast, the skip-sampling estimator can be computed for a suite of choices of b in substantially less time. Also, because the FDB involves division by the square root of the estimated spectral density, huge values of the ratio can occur when the latter is close to zero, thereby greatly distorting the variability across simulations; this was egregious for FDB with the trapezoidal window, but was less of a concern for FDB with the Parzen window. Skip-sampling, in contrast, does not suffer from this issue of occasional bad estimates, being more stable overall. We report the root mean squared error (RMSE) based on $1,000$ Monte Carlo simulations.

For data generating processes (DGP) we consider: (i) a Gaussian MA(1), (ii) an MA(1) with Student t innovations, (iii) an ARCH(1) with Gaussian inputs, and (iv) a non-invertible Gaussian MA(1). We study the lag one autocovariance and the lag one autocorrelation, being examples of spectral means and ratio statistics respectively. These statistics are asymptotically normal, and we use FDB and skip-sampling to estimate the asymptotic variance v (cf. Remark 3.1). For an MA(1) process with moving average parameter ϑ_1 and innovation variance σ^2, its spectral density is $f(\lambda) = |1+\vartheta_1 e^{-i\lambda}|^2 \sigma^2 = \left((1+\vartheta_1^2) + 2\cos(\lambda)\,\vartheta_1\right)\sigma^2$. Hence, with $g(\lambda) = \cos(\lambda)$ it follows that

$$\langle g\, f\rangle = \vartheta_1\,\sigma^2 = \gamma_1 \quad \text{and} \quad \langle g^2\, f^2\rangle = \frac{\sigma^4}{2}\left(3\vartheta_1^2 + (1+\vartheta_1^2)^2\right).$$

Using eq. (16) and Remark 4.1, the asymptotic variance of the lag one autocovariance estimator is $v = \left(1 + 5\vartheta_1^2 + \vartheta_1^4 + (\eta-3)\vartheta_1^2\right)\sigma^4$. For a Gaussian random variable $\eta = 3$, whereas for a Student t random variable with $\nu > 4$ degrees of freeedom, $\eta = 3 + 6/(\nu-4)$. Similarly, from eq. (21) we find that the asymptotic variance of the lag one autocorrelation estimator is

$$v = \frac{(1+4\theta^2+\theta^4)(1+2\theta^2/(1+\theta^2)^2) - 7\theta^2}{(1+\theta^2)^2}.$$

Hence, for the first two DGPs (and the fourth DGP) we can compute the true asymptotic variance v, and make RMSE comparisons with FDB and skip-sampling estimates of this variance. The third DGP is non-linear, so we make a direct calculation of $\langle\langle ggF\rangle\rangle$. To our knowledge, calculation of the GARCH trispectrum is still an open problem, although the spectral density of the squared process is known (see He and Teräsvirta (1999)) and this will be sufficient for our purposes. Letting $g(\lambda) = e^{i\lambda}$,

$$\langle\langle ggF\rangle\rangle = \sum_{k=-\infty}^{\infty} \gamma_{k+1,k,1} = \sum_{k=-\infty}^{\infty} \mathrm{Cov}[X_t X_{t-1}, X_{t-k}X_{t-k-1}] - \gamma_k^2 - \gamma_{k+1}\gamma_{k-1}.$$

It can be shown that the covariance is zero unless $k=0$; letting τ_h be the lag h autocovariance of $\{X_t^2\}$, we seek to compute τ_1. For an ARCH(1) process of parameter β and input errors $\{Z_t\}$, i.e., $X_t = Z_t(1+\beta X_{t-1}^2)^{1/2}$, He and Teräsvirta (1999) show that $\tau_1 = \beta\tau_0$ and $\tau_0 = (\eta-1)(1-\beta)^{-2}(1-\eta\beta^2)^{-1}$, where $\eta = E[Z_t^4]$. There is an implicit stationarity condition that $\eta\beta^2 < 1$; for standard normal inputs we have $\eta = 3$, and choosing $\beta = .5$ the stationarity condition is satisfied. Since $\gamma_h = \mathbf{1}\{h=0\}(1-\beta)^{-1}$, the asymptotic variance of the lag one autocovariance estimator is $v = \langle gg^\star f^2\rangle + \langle\langle ggF\rangle\rangle = \gamma_0 + (\tau_1 - \gamma_0) = 2\beta(1-\beta)^{-2}(1-3\beta^2)^{-1}$. Because the lag one autocorrelation is zero, the variance of its estimator turns out to be $v = 1$.

We consider two sample sizes: $T = 240$ and $T = 1200$. For the smaller sample, the bandwidth fractions used in the FDB are $h = .1$ and $h = .3$, whereas the values of q are chosen from $\{40, 30, 24, 20, 16, 12, 10, 8, 6, 5, 4, 3\}$ (which yields $b \in \{6, 8, 10, 12, 15, 20, 24, 30, 40, 48, 60, 80\}$). For the larger sample, the bandwidth fractions are $h = .05$ and $h = .1$ and the same values of b (so the q values are 5 times as large). For DGP (i) we set $\vartheta = .8$, and for DGP (ii) we set $\vartheta = -.4$ and $\nu = 6$. For DGP (iii) we set $\beta = 1/3$, and for DGP (iv) we set $\vartheta = -1$, which means the spectral density at frequency zero equals zero. Tables 1 and 2 report the RMSE results for both sample sizes and both statistics.

	First DGP		Second DGP		Third DGP		Fourth DGP	
$T = 240$	acvf	acf	acvf	acf	acvf	acf	acvf	acf
FDB Trapezoid ($h = .1$)	8.29	0.19	1.00	0.15	1.09	0.22	138.44	0.88
FDB Trapezoid ($h = .3$)	210.58	0.76	126.41	0.42	143.98	0.71	3932.47	0.35
FDB Parzen ($h = .1$)	1.62	0.10	0.99	0.14	1.14	0.21	2.55	0.10
FDB Parzen ($h = .3$)	1.81	0.13	1.13	0.17	1.04	0.26	2.78	0.12
Skip ($b = 6$)	2.64	0.07	1.23	0.31	1.18	0.44	5.26	0.23
Skip ($b = 8$)	2.34	0.12	1.13	0.25	1.11	0.34	4.70	0.20
Skip ($b = 10$)	2.21	0.13	1.09	0.22	1.13	0.30	4.31	0.18
Skip ($b = 12$)	2.15	0.13	1.09	0.20	1.09	0.27	4.09	0.17
Skip ($b = 15$)	2.27	0.11	1.08	0.15	1.06	0.25	3.36	0.10
Skip ($b = 20$)	2.06	0.13	1.06	0.18	1.04	0.25	3.69	0.14
Skip ($b = 24$)	2.07	0.13	1.05	0.18	1.07	0.25	3.60	0.14
Skip ($b = 30$)	2.05	0.14	1.05	0.19	1.03	0.26	3.33	0.14
Skip ($b = 40$)	2.07	0.14	1.09	0.20	1.02	0.29	3.27	0.15
Skip ($b = 48$)	2.11	0.15	1.11	0.21	1.02	0.31	3.21	0.16
Skip ($b = 60$)	2.14	0.17	1.14	0.23	1.02	0.34	3.19	0.17
Skip ($b = 80$)	2.27	0.20	1.23	0.27	1.07	0.40	3.33	0.20

Table 1: RMSE for Frequency Domain Bootstrap (FDB) and Random-draw skip-sampling, for estimating the variance v of both the lag one autocovariance (acvf) and autocorrelation (acf), for four DGPs (Gaussian MA(1), Student t MA(1), Gaussian ARCH(1), and non-invertible MA(1)) of sample size $T = 240$. Bandwidth fraction for FDB is h and the size of skip-samples is b.

For $T = 240$ we see in Table 1 that skip-sampling is competitive with FDB in many situations, although we may wish to discount the $h = .3$ results for the FDB with Trapezoidal window, as a few poor estimates blow up the RMSE. (This is blow up is especially prevalent for DGP (iv), where the non-invertibility of the spectral density causes a challenge for the Trapezoidal FDB, since one must divide by an estimate of the spectral density.) In practice one does not know what bandwidth will work the best; skip-sampling yields a performance that is fairly uniform across b. For autocovariance estimation, performance is similar for the Parzen FDB and skip-sampling in the cases of DGPs (ii) and (iii), and slightly worse for skip-sampling in DGPs (i) and (iv). While the Parzen FDB is superior in the autocorrelation estimation for DGPs (ii) and (iii), skip-sampling is on par for DGPs (i) and (iv). The results of Table 1 for $T = 1200$ indicate an improved performance of skip-sampling relative to Parzen FDB; for instance, the acvf case of DGP (ii) has better performance for skip-sampling. Results for regular-draw skip-sampling are found in Tables B.1 and B.2 of Appendix B; it is apparent that the estimates from random-draw skip-sampling (which is a proxy for total-draw) are more accurate as compared to regular-draw. The situation can be compared to time domain subsampling

$T = 1200$	First DGP acvf	First DGP acf	Second DGP acvf	Second DGP acf	Third DGP acvf	Third DGP acf	Fourth DGP acvf	Fourth DGP acf
FDB Trapezoid ($h = .05$)	0.83	0.06	0.70	0.08	0.41	0.12	37.22	0.75
FDB Trapezoid ($h = .1$)	0.94	0.07	0.79	0.10	0.46	0.14	3.85	0.20
FDB Parzen ($h = .05$)	0.82	0.05	0.68	0.07	0.39	0.11	1.26	0.05
FDB Parzen ($h = .1$)	0.88	0.06	0.75	0.08	0.44	0.13	1.36	0.06
Skip ($b = 6$)	2.51	0.05	0.64	0.30	0.44	0.42	3.62	0.23
Skip ($b = 8$)	1.95	0.10	0.57	0.23	0.43	0.31	3.07	0.19
Skip ($b = 10$)	1.66	0.11	0.54	0.19	0.43	0.25	2.61	0.17
Skip ($b = 12$)	1.46	0.11	0.54	0.17	0.42	0.21	2.41	0.15
Skip ($b = 15$)	1.64	0.06	0.67	0.09	0.51	0.17	1.54	0.05
Skip ($b = 20$)	1.17	0.08	0.55	0.12	0.43	0.15	1.94	0.11
Skip ($b = 24$)	1.12	0.08	0.58	0.11	0.42	0.14	1.88	0.10
Skip ($b = 30$)	1.07	0.07	0.60	0.11	0.42	0.14	1.75	0.09
Skip ($b = 40$)	1.04	0.07	0.61	0.10	0.42	0.13	1.65	0.08
Skip ($b = 48$)	1.02	0.06	0.62	0.10	0.43	0.13	1.62	0.08
Skip ($b = 60$)	1.04	0.06	0.64	0.10	0.42	0.13	1.54	0.08
Skip ($b = 80$)	1.03	0.06	0.68	0.10	0.44	0.14	1.53	0.08

Table 2: RMSE for Frequency Domain Bootstrap (FDB) and Random-draw skip-sampling, for estimating the variance v of both the lag one autocovariance (acvf) and autocorrelation (acf), for four DGPs (Gaussian MA(1), Student t MA(1), Gaussian ARCH(1), and non-invertible MA(1)) of sample size $T = 1200$. Bandwidth fraction for FDB is h and the size of skip-samples is b.

where subsampling non-overlapping blocks is consistent but less efficient compared to using all available blocks; see eq. (3.46) of Politis et al. (1999).

Remark 5.1 (Numerical stability) As already mentioned, the numerical instability of the FDB can be attributed to the required normalization of periodogram ordinated by division by a local estimate of the spectral density (which may turn out to be close to zero). Skip-sampling does not require this normalization and is thus simpler to implement, as well as being numerically stable.

Remark 5.2 (Computational cost) An advantage of skip-sampling over the FDB is its lower computational cost. With sample size $T = 1200$, the FDB (with Trapezoidal window and $h = .05$) takes 16.73 seconds for a typical run using DGP (i) on a 2.40 GHz processor. In contrast, regular-draw skip-sampling requires .23 seconds (with $q = 200$), and .39 seconds with the random-draw skip-sampling method. For the smaller sample size of $T = 240$ (and $q = 40$), the FDB (with Trapezoidal window and $h = .1$) takes .77 seconds while random-draw skip-sampling takes .11 seconds and regular-draw less than .01 seconds.

Acknowledgement. This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau. Many thanks are due to the Editor, Associate Editor and two anonymous reviewers for several helpful comments. The research of the second author was partially supported by NSF grant DMS 19-14556.

References

[1] Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory*, Holden-Day, New York.

[2] Bertail, P. and Dudek, A.E. (2021). Consistency of the frequency domain bootstrap for differentiable functionals, *Electronic Journal of Statistics* **15**, 1–36.

[3] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods, (2nd ed.)*, Springer, New York.

[4] Dahlhaus, R. (1985). Asymptotic normality of spectral estimates, *Journal of Multivariate Analysis* **16**, 412–431.

[5] Dahlhaus, R. and Janas, D. (1996). A frequency domain bootstrap for ratio statistics in time series analysis, *Ann. Statist.* **24**(5), 1934–1963.

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

[7] Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates, *Ann. Statist.* **20**, 121–145.

[8] Fragkeskou, M. and Paparoditis, E. (2016). Inference for the fourth-order innovation cumulant in linear time series, *Journal of Time Series Analysis*, **2**(37), 240–266.

[9] He, C. and Teräsvirta, T. (1999). Fourth moment structure of the GARCH (p, q) process, *Econometric Theory*, 15(6), 824–846.

[10] Hurvich, C. M. and Zeger, S. L. (1987). Frequency domain bootstrap methods for time series, Working Paper, New York University.

[11] Janas, D. and R. Dahlhaus (1994). A frequency domain bootstrap for time series. In: *Computationally Intensive Statistical Methods. Proceedings of the 26th Symposium on the Interface*, (J. Sall and A. Lehman, eds.). Interface Foundation of North America, Faifax Station, VA, pp. 423–425

[12] Kirch, C. and D.N. Politis (2011). TFT-bootstrap: resampling time series in the frequency domain to obtain replicates in the time domain, *Ann. Statist.* **39**(3), 1427–1470.

[13] Kreiss, J.-P. and Paparoditis, E. (2003). Autoregressive aided periodogram bootstrap for time series, *Ann. Statist.* **31**(6), 1923–1955.

[14] Kreiss, J.-P. and Paparoditis, E. (2024). *Bootstrap for Time Series: Theory and Applications*, Springer, Heidelberg (to appear).

[15] Lahiri, S.N. (2003a). *Resampling Methods for Dependent Data*, Springer, New York.

[16] Lahiri, S.N (2003b). A necessary and sufficient condition for asymptotic independence of discrete Fourier transforms under short-and long-range dependence, *Ann. Statist.* **31**(2), 613–641.

[17] McElroy, T.S. and Politis, D.N. (2020). *Time Series: A First Course with Bootstrap Starter*, Chapman and Hall/CRC Press, Boca Raton.

[18] McElroy, T.S. and Roy, A. (2022). Model identification via total Frobenius norm of multivariate spectra, *Journal of the Royal Statistical Society, Ser. B*, **84**, 473–495.

[19] Meyer, M. Paparoditis, E. and Kreiss, J.-P. (2020). Extending the validity of frequency domain bootstrap methods to general stationary processes, *Ann. Statist.* **48**(4), 2404–2427.

[20] Paparoditis, E. (2002). Frequency domain bootstrap for time series. In *Empirical Process Techniques for Dependent Data* (H. Dehling et al., eds.) 365–381. Birkhäuser, Boston.

[21] Paparoditis, E. and Politis, D.N. (1999). The local bootstrap for periodogram statistics, *Journal of Time Series Analysis* **20**, 193–222.

[22] Politis, D.N., and Romano, J.P. (1994), Large sample confidence regions based on subsamples under minimal assumptions, *Ann. Statist.*, 22: 2031-2050.

[23] Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling*, Springer, New York.

[24] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*, Springer, New York.

[25] Yu, H., Kaiser, M.S. and Nordman, D.J. (2023). A subsampling perspective for extending the validity of state-of-the-art bootstraps in the frequency domain, *Biometrika* (to appear).