

ADAPTING TO UNKNOWN DISTURBANCE AUTOCORRELATION IN REGRESSION WITH LONG MEMORY*

by

Javier Hidalgo and Peter M Robinson
London School of Economics and Political Science

Contents:

Abstract

1. Introduction

2. Adaptive Estimation of β

3. Monte Carlo Study of the Finite-Sample
Behaviour

4. Proof of Theorem 2.1

Appendix A

Appendix B

References

Tables 1 – 4

Discussion Paper
No. EM/01/427
September 2001

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020 – 7955 6698

* The first author's research was supported by ESRC Grant R000238212. The second author's research was supported by ESRC Grant R000235892 and R000238212 and a Leverhulme Trust Personal Research Professorship. Stepana Lazarova carried out the Monte Carlo work, modifying a computer program written by Josu Arteché, who carried out the Monte Carlo in the original version of this paper. The paper has been improved in response to comments of three referees and a co-editor.

Abstract

We show that it is possible to adapt to nonparametric disturbance auto-correlation in time series regression in the presence of long memory in both regressors and disturbances by using a smoothed nonparametric spectrum estimate in frequency-domain generalized least squares. When the collective memory in regressors and disturbances is sufficiently strong, ordinary least squares is not only asymptotically inefficient but asymptotically non-normal and has a slow rate of convergence, whereas generalized least squares is asymptotically normal and Gauss-Markov efficient with standard convergence rate. Despite the anomalous behaviour of nonparametric spectrum estimates near a spectral pole, we are able to justify a standard construction of frequency-domain generalized least squares, earlier considered in case of short memory disturbances. A small Monte Carlo study of finite sample performance is included.

Keywords: Time series regression; long memory; adaptive estimation.

JEL No.: C22

© by the authors. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without special permission provided that full credit, including © notice, is given to the source.

Contact address: Dr Javier Hidalgo, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE. Email: [f.j.hidalgo@lse.ac.uk/](mailto:f.j.hidalgo@lse.ac.uk)

1. INTRODUCTION

Adaptive estimation in semiparametric models entails an estimate of the parametric component achieving the same first-order efficiency when the nonparametric nuisance function is consistently estimated as when this function is known. It seems that such adaptive estimation was first established by Hannan (1963), in the context of linear time series regression with disturbance having nonparametric autocorrelation, the Gauss-Markov bound being attained. Hannan considered an approximate frequency-domain generalized least squares (*GLS*) estimate, using a smoothed nonparametric estimate of the disturbance spectral density. Even though the spectral estimate converges slowly, Hannan showed that nevertheless his *GLS* estimate can have the same first order limit distribution as if the spectral density were completely known, or a known parametric function. The same idea was extended to more general semiparametric time series models, especially ones useful in econometrics, by Hannan (1965), Hannan and Terrell (1972, 1973), Robinson (1976) and more recently by Robinson (1991), who also allowed for a data dependent smoothing in the spectral estimation and considered optimal choice of the smoothing number, and Phillips (1991) who considered this type of estimate in the presence of unit root regressors.

The regularity conditions required by Hannan (1963) and the subsequent authors are in many respects quite mild, in particular the disturbance spectral density need only be bounded and bounded away from zero and satisfy mild smoothness conditions, so that stationary and invertible autoregressive moving averages are easily covered. However, there is now increasing awareness of the possibility of long memory in disturbances, which contradicts such specifications. The concept of long memory can be loosely defined in terms of the $I(d)$ processes. We say that a covariance stationary, invertible process is $I(d)$ if it has spectral density that behaves like $C\lambda^{-2d}$ as frequency $\lambda \rightarrow 0+$, for $0 < C < \infty$ and $-1/2 < d < 1/2$. Then there is said to be negative memory if $-1/2 < d < 0$, short memory if $d = 0$, and long memory if $0 < d < 1/2$. For $d > 1/2$, a process can be said to be (non-stationary) $I(d)$ if its k th integer difference is $I(d - k)$, for $-1/2 < d - k < 1/2$, to cover unit root ($I(1)$) processes, for example. We can also define long memory relative to a spectral pole at some non-zero frequency, and will return to this possibility subsequently. In the context of the linear regression model

$$y_t = \alpha + \beta' x_t + u_t, \quad t = 1, 2, \dots, \quad (1.1)$$

where the scalar α and the $p \times 1$ vector β are unknown and the prime denotes transposition, long memory in the disturbance u_t can occur if the $p \times 1$ regressor vector x_t does not wholly account for long memory in the dependent variable y_t . Generally, long memory processes can be thought of as a broad and flexible class

which can bridge the gap between short memory and unit root processes (which have recently played a considerable role in econometric modelling), especially as the latter occupy only one point in d -space, whereas stationary long memory processes take up an interval.

Initial study of the implications of long memory u_t in (1.1) focussed on ordinary least squares (*OLS*) estimates of β . In case of deterministic, such as polynomial-in- t regressors, Yajima (1988, 1991) found that *OLS* estimates can still be asymptotically normal, even if u_t is non-Gaussian but a linear process. However their asymptotic variance, indeed their rate of convergence, is adversely affected by the long memory in u_t . Moreover, while the *OLS* estimates may not necessarily have poor efficiency, they cannot achieve the Gauss-Markov bound, even for polynomial-in- t regressors, unlike in the case of short memory u_t considered by Grenander (1954). In econometric applications, it is often reasonable to regard x_t , or at least some elements of it, as stochastic. Here, even if x_t and u_t are independent stationary processes, indeed possibly Gaussian ones, sufficiently strong collective memory leads to *OLS* having a slow rate of convergence and non-standard limiting distribution, specifically, as found by Robinson (1994a), if u_t is $I(d)$ and x_t is $I(d_x)$ and $d + d_x > 1/2$. This outcome is consistent with the familiar one in econometrics in which $d = 0$ and $d_x = 1$, but the implications are rather more serious because not only, as there, is the limit distribution less convenient for use in inference than the normal, but it depends on d and d_x , which are typically regarded as unknown when fractional models are entertained. Note also that not only is the conventional econometric approach to inference, based on *OLS* and an autocorrelation-consistent variance estimate (employing a smoothed nonparametric estimate of the spectral density of $x_t u_t$, see e.g. Andrews, 1991) inefficient (as usual) due to its dependence on an inefficient point estimate, but it is not even available due to the asymptotic non-normality.

To resolve this difficulty, Robinson and Hidalgo (1997) considered a class of frequency-domain weighted least squares (*WLS*) estimates. Based on observations (x'_t, y_t) , $t = 1, \dots, n$, define the discrete Fourier transforms of x_t and y_t

$$w_x(\lambda) = \frac{1}{(2\pi n)^{1/2}} \sum_{t=1}^n x_t e^{it\lambda}, \quad w_y(\lambda) = \frac{1}{(2\pi n)^{1/2}} \sum_{t=1}^n y_t e^{it\lambda},$$

and the periodogram matrix and vector

$$I_{xx}(\lambda) = w_x(\lambda) w_x^*(\lambda), \quad I_{xy}(\lambda) = w_x(\lambda) w_y^*(\lambda),$$

where "*" indicates transposition combined with complex conjugation. Now write

$$\hat{\beta}_\phi = \left(\sum_{j=1}^{n-1} I_{xx}(\lambda_j) \phi(\lambda_j) \right)^{-1} \left(\sum_{j=1}^{n-1} I_{xy}(\lambda_j) \phi(\lambda_j) \right), \quad (1.2)$$

where $\phi(\lambda)$ is real-valued, even and periodic of period 2π and $\lambda_j = 2\pi j/n$, for integer j . This is one of the *WLS* estimates considered by Robinson and Hidalgo (1997) (the other replacing sums by integrals and having the same first-order asymptotic properties, but being less convenient computationally). Omission of the frequency $j = 0$ (and $j = n$) entails sample-mean correction, and when $\phi(\lambda) \equiv 1$, $\widehat{\beta}_\phi$ reduces to the *OLS* estimate of β for the model (1.1). Assume that u_t has absolutely continuous spectral distribution function, so it has a spectral density $f(\lambda)$ satisfying

$$\gamma_j = \text{Cov}(u_1, u_{1+j}) = \int_{-\pi}^{\pi} f(\lambda) \cos(j\lambda) d\lambda, \quad j = 0, \pm 1, \dots$$

If, for all λ , $f(\lambda) > 0$ and $f(\lambda)$ is known then $\widehat{\beta}_{f^{-1}}$ is an approximate *GLS* estimate, achieving the Gauss-Markov bound under suitable conditions, in particular

$$n^{1/2} \left(\widehat{\beta}_{f^{-1}} - \beta \right) \xrightarrow{d} N(0, \Omega^{-1}), \quad (1.3)$$

where $\Omega = (2\pi)^{-1} \int_{-\pi}^{\pi} f^{-1}(\lambda) dF(\lambda)$ and the matrix $F(\lambda)$ satisfies

$$\Gamma(j) = E((x_1 - Ex_1)(x_{1+j} - Ex_1)') = \int_{-\pi}^{\pi} e^{ij\lambda} dF(\lambda)$$

such that $F(\lambda)$ has Hermitian nonnegative definite increments and is uniquely defined by the requirement that it is continuous from the right.

The result (1.3) was established by Robinson and Hidalgo (1997) under regularity conditions which permit arbitrarily strong stationary long memory in both x_t and u_t . Intuitively, whereas *OLS* has an asymptotic variance involving the integral $\int_{-\pi}^{\pi} f(\lambda) dF(\lambda)$, which will not converge if the spectral mass of x_t and u_t near zero frequency is collectively sufficiently great, this is no obstacle to convergence of Ω because long memory in u_t entails a zero, not a pole, in $f^{-1}(\lambda)$. Thus, not only does *GLS* have its usual desirable asymptotic Gauss-Markov efficiency properties, but it is also motivated in the present circumstances by being asymptotically normal, even when *OLS* is not and has slow rate of convergence. Moreover, even when *OLS* is $n^{1/2}$ -consistent, it is easy to construct examples in which *OLS* has very poor efficiency relative to *GLS*. Dahlhaus (1995) has also considered *GLS* under long memory u_t , but for deterministic x_t where, as noted previously, *OLS* is typically asymptotically normal however strong the memory of u_t , so that *GLS* appears to fill a lesser gap than in the stochastic case.

As presented, (1.3) assumes $f(\lambda)$ is known, which is unrealistic in applications. Robinson and Hidalgo (1997) showed that (1.3) still goes through if $f(\lambda)$ is a known function of λ and finitely many unknown parameters, and the latter are

replaced by $n^{1/2}$ -consistent estimates. However, correct parametric specification of $f(\lambda)$ is here essential. For example if u_t is specified as a fractional autoregressive integrated moving average (*FARIMA*) process but either the autoregressive or the moving average order is under-stated, or both are over-stated, then the estimate of β will not be asymptotically efficient and (1.3) will not hold. Indeed, the estimate may not then even be $n^{1/2}$ -consistent and asymptotically normal, because the misspecification of autoregressive and/or moving average orders leads to an inconsistent spectral estimate, which in particular may be biased downwards such that the reciprocal of the estimated $f(\lambda)$ does not sufficiently compensate for the long memory in the manner discussed above. In view of the earlier discussion we would therefore like to adapt to nonparametric $f(\lambda)$. This is accomplished in the present paper; we establish (1.3) after replacing f by a smoothed nonparametric estimate. We have found this to be a very difficult task, far more than when u_t has short memory, especially in the context of the standard specification of semiparametric *GLS* (one originally proposed for short memory u_t) which we use, that does not resort to special devices such as trimming near frequency zero in order to deal with the unpleasant properties of spectral estimates near a spectral pole. Indeed, even for known f , the result (1.3) was found by Robinson and Hidalgo (1997) to be considerably more difficult to establish under long memory than short memory, and some of our regularity conditions are no stronger than ones in that paper.

As in Robinson and Hidalgo (1997), we assume x_t is stationary. Nonstationary x_t can be considered, indeed Phillips (1991) analyzed a semiparametric *GLS* estimate with $I(1)$ x_t (but short memory u_t) but his findings were that *GLS* loses its classical properties (c.f. (1.3)) then, indeed that a narrow-band estimate (about zero frequency) does as well, so that the unequal weighting entailed in *GLS* is redundant. This is due to the dominance of low frequencies in $I(1)$ processes, and so a similar outcome might be expected in case of other nonstationary processes, especially in view of results of Robinson and Marinucci (1998). This type of outcome is not very interesting from the point of view of adaptive estimation, especially if (as in the $x_t \sim I(1) / u_t \sim I(0)$ case) the limit distribution of *GLS* is not even normal. Now from our discussion of the convergence of Ω , it appears that even for some nonstationary x_t , specifically $I(d_x)$ ones for $1/2 \leq d_x < 1$, the zero in $f^{-1}(\lambda)$ might 'compensate', so as to lead to an outcome like (1.3). However this would clearly require sufficiently strong long memory in u_t relative to that in x_t ($d > d_x - 1/2$ if $u_t \sim I(d)$), even in case of known f no limit results are yet available, and a modified construction would be needed such as one involving tapering, in view of other literature on nonstationary fractional processes. In the circumstances it seems desirable to deal with the stationary case first, stressing again that our inclusion of $I(d_x)$ processes for $d_x < 1/2$ is considerably broader

than the class of $I(0)$ processes sometimes erroneously referred to in the literature as synonymous with stationarity, while our conditions are in other respects mostly undeniably mild. The stationarity assumption may itself be reasonable in the context of some financial data, while in any case it can sometimes be hard to distinguish a record of a nonstationary process from that of a stationary fractional one, for example compare realizations of a unit root process with ones of a $FARIMA(1, 0.45, 0)$ with autoregressive parameter 0.95.

Our assumptions require f to have a pole, if at all, at frequency zero only, but because we use a standard construction of a GLS estimate which gives no special treatment to this or any other frequency, it is clear that the results will still go through if f has a pole at some other frequency (see e.g. Hosoya, 1996) or indeed at several other frequencies, as when a cyclic or seasonal phenomenon prevails, even in the presence of arbitrarily strong stationary long memory in x_t at these and other frequencies. We have focused on the zero-frequency case in part because it seems relatively important in econometric applications (referring to the 'typical spectral shape' of an economic variable), and in part for simplicity of exposition, our proofs even here being extremely technical. Notice also that it seems straightforward to extend our proofs to justify analogous estimates in more general models of econometric interest, such as lagged regression, constrained regression, nonlinear regression, band spectrum regression, distributed lag, simultaneous equations and continuous time models, in view of work of Hannan (1963, 1965), Hannan and Terrell (1972, 1973), Robinson (1976, 1991), Xiao and Phillips (1999), but the linear regression (1.1) affords a relatively compact treatment.

It should be stressed that our results can differ significantly from those that pertain in case of deterministic x_t . In the first place it is necessary to observe that central limit theory for both OLS and GLS when x_t is deterministic need not obtain when u_t is not a linear process (indeed it is not clear to what extent it obtains for our GLS estimates with stochastic x_t). If u_t is, for example, an instantaneous nonlinear function of a Gaussian long memory process v_t , then by extension of results for the sample mean of Rosenblatt (1961) or Taqqu (1975), one expects the limit distributions of OLS and GLS to be non-normal in case u_t has Hermite rank greater than 1 (that is, its expansion in Hermite polynomials of v_t contains no linear component). Even when OLS and GLS are asymptotically normal, they have rate of convergence which is not only affected by trending behaviour in deterministic (such as polynomial-in- t) x_t but, unlike our GLS estimates with stochastic x_t , is also adversely affected by long memory u_t when the limiting spectral distribution function of the normalized x_t has a jump at frequency zero (as in the polynomial-in- t case); see Yajima (1988, 1991). Dahlhaus (1995) studied a class of weighted estimates deriving from Adenstedt's (1974) treatment of the simple location model, studying to what extent they achieve the

asymptotic Gauss-Markov bound in the presence of long (and negative) memory u_t . These estimates depend on the spectrum of u_t only through d , even in case of a semiparametric model similar to our (2.1) below, and Dahlhaus showed that insertion of a suitable estimate of d does not change the normal limit distribution, when u_t is Gaussian. Dahlhaus indicated that his approach applies to polynomial-in- t regressors, but calculations of Yajima (1988) suggest that the efficiency gains over *OLS* may be relatively slight in this case. For more general deterministic regressors, whose limiting spectral distribution function is not known *a priori* to be a step function with step only at frequency zero, an adaptive method of estimation like ours seems to be necessary, and it remains to extend Hannan's (1963) approach, which assumed "Grenander's conditions" on x_t ; however, at least when x_t has a spectral jump at frequency zero the technical problems appear to differ from those in our case.

The following section introduces a nonparametric estimate \hat{f} of f and gives conditions under which (1.3) holds in a theorem whose proof in Section 4 makes use of a series of propositions appearing in Appendix A, which themselves use a series of technical lemmas described in Appendix B. Section 3 contains a Monte Carlo study of finite-sample performance.

2. ADAPTIVE ESTIMATION OF β

We first discuss the estimation of f , which (see Condition C1 below) satisfies the property

$$f(\lambda) \sim C\lambda^{-2d}, \quad \text{as } \lambda \rightarrow 0+, \quad (2.1)$$

for $0 < C < \infty$ and $0 \leq d < 1/2$. Though we cannot thus expect to estimate $f(\lambda)$ well near $\lambda = 0$ when $d > 0$, we nevertheless use a conventional spectrum estimate. Because the u_t are unobservable we form the residuals

$$\tilde{u}_t = y_t - \bar{y} - \tilde{\beta}'(x_t - \bar{x}), \quad (2.2)$$

where $\bar{x} = n^{-1} \sum_{t=1}^n x_t$, $\bar{y} = n^{-1} \sum_{t=1}^n y_t$ and $\tilde{\beta} = \hat{\beta}_\phi$ is a preliminary estimate (see (1.2)), with $\phi(\lambda)$ chosen such that $\tilde{\beta}$ is $n^{1/2}$ -consistent. For this, it suffices to choose $\phi(\lambda)$ to satisfy the conditions of Robinson and Hidalgo (1997), where essentially a zero of order greater than or equal to $1/2$ at $\lambda = 0$ is needed in order to guard against the possibility of long memory in both x_t and u_t ; the *OLS* choice $\phi(\lambda) = 1$ is inadequate, as mentioned in Section 1. It suffices to choose $\phi(\lambda) \equiv \varphi(\lambda)$, where

$$\varphi(\lambda) = |1 - e^{i\lambda}|. \quad (2.3)$$

Our estimate of $f(\lambda)$ is

$$\widehat{f}(\lambda) = \bar{c}^{-1} \frac{1}{2m+1} \sum_{j=-m}^m c_j \widetilde{I}(\lambda + \lambda_j), \quad (2.4)$$

where

$$\widetilde{I}(\lambda) = |w_{\tilde{u}}(\lambda)|^2, \quad w_{\tilde{u}}(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^n \tilde{u}_t e^{it\lambda},$$

$c_j = c(j/m)$ for a function $c(u)$ restricted by Condition C7 below, with $\bar{c} = (2m+1)^{-1} \sum_{j=-m}^m c_j$, and $m = m(n)$ is a bandwidth sequence which increases slowly with n in a manner prescribed in Condition C8 below.

The estimate $\widehat{\beta}_{\widehat{f}^{-1}}$, with its averaging over Fourier frequencies λ_j , is essentially the specification considered by Robinson (1991) in case of short-memory u_t , differing from that of Hannan (1963) which entailed smoothed nonparametric estimation for x_t , and then averaging over frequencies $2\pi j/M$, $1 \leq j \leq M$, where M increases more slowly than n in asymptotic theory. Undoubtedly Hannan's estimate can be justified under conditions very similar to those we employ for $\widehat{\beta}_{\widehat{f}^{-1}}$.

Our regularity conditions are as follows.

Condition C1 $f(\lambda) > 0$ for all $\lambda \in [0, \pi]$ and for some $\eta > 0$, as $\lambda \rightarrow 0+$

$$f(\lambda) = C\lambda^{-2d} + O(\lambda^{\eta-2d}), \quad 0 < C < \infty, \quad 0 \leq d < 1/2.$$

Condition C2

$$u_t = \sum_{j=0}^{\infty} \tau_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} \tau_j^2 < \infty,$$

where $E(\varepsilon_t | \mathbb{F}_{t-1}) = 0$, $E(\varepsilon_t^2 | \mathbb{F}_{t-1}) = E(\varepsilon_t^2) = \sigma^2$, $E(|\varepsilon_t|^\ell | \mathbb{F}_{t-1}) = E(|\varepsilon_t|^\ell) = \mu_\ell < \infty$ a.s., $\ell = 3, \dots, 12$, \mathbb{F}_t being the σ -field of events generated by ε_s , $s \leq t$, and for $4 \leq \ell \leq 12$, the joint cumulant of ε_{t_i} , $i = 1, \dots, \ell$, satisfies

$$\text{cum}(\varepsilon_{t_i}, i = 1, \dots, \ell) = \begin{cases} \kappa_\ell, & t_1 = t_2 = \dots = t_\ell, \\ 0, & \text{otherwise.} \end{cases}$$

Condition C3 For $\tau(\lambda) = \sum_{j=0}^{\infty} \tau_j e^{ij\lambda}$,

$$\frac{d}{d\lambda} \log(|\tau(\lambda)|) = O(\lambda^{-1}) \quad \text{as } \lambda \rightarrow 0+$$

and $\tau(\lambda)$ is continuously differentiable in $\lambda \in (\eta, \pi)$, for any $\eta > 0$. Also, if $d = 0$, $\tau(\lambda)$ is continuously differentiable in $[0, \pi]$.

Condition C4 $\{x_t\}$ is fourth-order stationary with $\Gamma(j) \rightarrow 0$ as $|j| \rightarrow \infty$, and

$$\lim_{|u| \rightarrow \infty} \max_{|v|, |w| < \infty} |\kappa_{abcd}(0, u, v, w)| = 0, \quad 1 \leq a, b, c, d \leq p,$$

where $\kappa_{abcd}(0, u, v, w)$ is the fourth cumulant of $x_{a0}, x_{bu}, x_{cv}, x_{dw}$, and x_{it} is the i th element of x_t .

Condition C5 The processes $\{x_t\}$ and $\{u_t\}$ are mutually independent.

Condition C6 Ω is positive definite.

Condition C7 $c(u)$ is an even, nonnegative function on $[-1, 1]$, twice continuously differentiable in $(0, 1)$, bounded away from zero on $[0, 1 - \varepsilon)$ for some $\varepsilon \in (0, 1)$, and such that $c(1) = 0$.

Condition C8 As $n \rightarrow \infty$, $n^{-1}m^2 + nm^{-3} \rightarrow 0$.

Condition C9 $\widehat{\beta}_\phi = \beta + O_p(n^{-1/2})$.

Conditions C1 and C3 are for the most part substantially stronger than those in Robinson and Hidalgo (1997). There, no power law behaviour near zero frequency was required, nor was f required to be differentiable (or even continuous) away from frequency zero (note that $f(\lambda) = (\sigma^2/2\pi) |\tau(\lambda)|^2$.) The stronger conditions C1 and C3 are introduced in part to cope with the difficulty of using the slowly converging nonparametric estimate $\widehat{f}(\lambda)$, and in part for notational convenience. As mentioned in Section 1, it seems readily possible to extend our results to cover poles in f at finitely many other frequencies. Note that we allow $d = 0$ in C1, so the short memory case, treated by Hannan (1963) and others, is also covered. Condition C1 only differs from conditions used in the semiparametric literature that focusses on low frequencies only (see Robinson, 1995a,b) by its requirement that f be everywhere positive, which is a natural one in the present context. The first part of Condition C3 was used in the semiparametric context by Robinson (1995b), for frequencies λ in a neighbourhood of the origin, while the second part is standard in smoothed spectral density estimation.

Conditions C2 and C4-C6 are the same as in Robinson and Hidalgo (1997) except for the considerably stronger moment condition on u_t that C2 entails. This is introduced in order to mitigate the bad behaviour of $\widehat{f}(\lambda)$ near $\lambda = 0$ when $d > 0$ (we will return to its implication subsequently in Section 4), and could be relaxed if trimming out of low frequencies or some kind of pre-whitening/re-colouring spectral estimation is used. However, such modifications are practically unattractive, especially if they involve choice of trimming numbers or weight functions, with the consequent ambiguity, and at least when actually $d = 0$, trimming is

likely to sacrifice some finite-sample efficiency due to the unwarranted discarding of information. We have thus preferred to focus on the 'standard construction' $\widehat{\beta}_{\widehat{f}^{-1}}$, and note that the 12th moment condition of C2 is satisfied if u_t is Gaussian, which is necessary for $\widehat{\beta}_{\widehat{f}^{-1}}$ to also attain the asymptotic Cramer-Rao efficiency bound. Nevertheless when $d > 0$ inclusion of $\widehat{f}^{-1}(\lambda)$ for small λ is likely to impair performance, and estimates with better finite-sample properties may be worth seeking. One possible candidate entails obtaining a semiparametric estimate \widehat{d} of d (for example, by log-periodogram regression) from the \widetilde{u}_t sequence, consequent filtering of \widetilde{u}_t to approximately eliminate the long memory, (that is, forming $\widetilde{v}_t = (1 - L)^{\widehat{d}} \{\widetilde{u}_t \mathcal{I}(t > 0)\}$), spectrum estimation based on the \widetilde{v}_t , and finally multiplying this spectrum estimate by $\lambda^{-2\widehat{d}}$, to obtain an estimate of f which can be inserted in (1.2). However, our view is that the immediate priority is to justify the computationally relatively simple, standard construction $\widehat{\beta}_{\widehat{f}^{-1}}$. Condition C2 essentially requires that u_t be a linear process with innovations that are independent and identically distributed up to 12th moments. Stationarity of x_t in C4 was discussed in Section 1, while given the stationarity the aesthetically simple condition $\Gamma(j) \rightarrow 0$ is very mild: it implies that $F(\lambda)$ is continuous (see Doob, 1953, p. 494) and on the other hand is itself implied if $F(\lambda)$ is absolutely continuous (from the Riemann-Lebesgue lemma). The cumulant condition on x_t in C4 which holds automatically under Gaussianity, also seems mild, while C6 is an unavoidable identifiability condition. Condition C5 warrants discussion. It can be relaxed to a milder type of orthogonality between x_t and u_t , but at the cost of strengthening other conditions and of greater complexity. In a cointegration context, of course, even uncorrelatedness of x_t and u_t would likely be unacceptable, but in this situation our estimates $\widehat{\beta}_\phi$, including *OLS*, will be inconsistent in our stationary x_t context; here, Robinson (1994b) showed that β can be estimated consistently by a narrow-band frequency domain estimate when $d_x > d$.

Examples of weights satisfying Condition C7 are

$$c(u) = 1 - |u|, \quad c(u) = \frac{1}{2} (1 + \cos(\pi u)),$$

but the uniform $c(u) = 1/2, u \in [-1, 1]$, is excluded. Condition C8 is unsurprisingly stronger than the corresponding bandwidth condition employed by Robinson (1991) for short memory u_t . As always, there will be sensitivity to choice of m . The question of optimal choice of m seems not to be entirely resolved in the current circumstances, though it seems that the usual solutions, valid for a smooth spectrum (see e.g. Robinson (1991, Theorem 5.1) will be justified away from frequency zero, while the optimality theory of Robinson (1994c) should apply near zero frequency. Some modification of the data-dependent cross-validation proce-

ture justified by Robinson (1991) should therefore be valid here. Condition C9 was discussed earlier, and is satisfied by $\widehat{\beta}_\varphi$, see (2.3).

Even the broad details of the proof of the following theorem are somewhat technical, so they are postponed until Section 4, with the remainder of the proof in Appendices A and B.

Theorem 2.1. *Let (1.1) and Conditions C1-C9 hold and let \widehat{f} be computed from (2.4) with \widetilde{u}_t given by (2.2) and $\widetilde{\beta}$ be given by (1.2) with φ as in (2.3). Then as $n \rightarrow \infty$*

$$n^{1/2} \left(\widehat{\beta}_{\widehat{f}_{-1}} - \beta \right) \xrightarrow{d} N(0, \Omega^{-1}),$$

and Ω is consistently estimated by

$$\frac{1}{n} \sum_{j=1}^{n-1} \widehat{f}^{-1}(\lambda_j) I_{xx}(\lambda_j). \quad (2.5)$$

3. MONTE CARLO STUDY OF FINITE-SAMPLE BEHAVIOUR

A small Monte Carlo study was carried out in order to investigate how well the efficiency of $\widehat{\beta}_{\widehat{f}_{-1}}$ approximates its optimal asymptotic level in finite samples, as well as to look at the effect of iterating the procedure and the efficiency of least squares. In (1.1) we took $p = 1$, $\alpha = 0$, $\beta = 1$ and generated u_t and x_t (with mean zero) as Gaussian *FARIMA* processes with spectral densities $f(\lambda) = (2\pi)^{-1} |1 - e^{i\lambda}|^{-2d}$ and $dF(\lambda)/d\lambda = (2\pi)^{-1} |1 - e^{i\lambda}|^{-2d_x}$ (see Adenstedt (1974)). Our results are invariant to α , β , Ex_t and the variances of u_t and x_t . We employed

$$d = 0, 0.05 (0.1) 0.45, \quad d_x = 0.05 (0.1) 0.45, \quad (3.1)$$

in order to assess performance across different levels of memory in both u_t and x_t . For each combination, 1000 replications of series of lengths $n = 64, 128$ and 256 were generated, by the method of Davies and Harte (1987). The estimate $\widehat{\beta}_{\widehat{f}_{-1}}$ was computed for $m = n/32, n/16$ and $n/8$, with $c(u) = .5(1 + \cos(\pi u))$ for each sample size n , and with the preliminary estimate $\widetilde{\beta} = \widehat{\beta}_\phi$ given both by $\phi(\lambda) = |1 - e^{i\lambda}|$ as in (2.3) (Method 1) and $\phi(\lambda) = 1$ (Method 2). In Method 1 $\widetilde{\beta}$ is $n^{1/2}$ -consistent for all d, d_x , as indicated in Section 2, whereas in Method 2, the *OLS* estimate $\widetilde{\beta}$ is $n^{1/2}$ -consistent only when $d + d_x < 1/2$; of course (3.1) includes also cases $d + d_x > 1/2$, under which our condition C9 fails, but it is nevertheless of interest to see how the simple and familiar choice of *OLS* for the preliminary estimate performs.

In Tables 1 and 2 we report the ratios $R = \text{Monte Carlo } MSE(\widehat{\beta}_{\widehat{f}_{-1}}) / (n^{-1}\Omega^{-1})$ based on Methods 1 and 2 respectively, across the various combinations of n, m, d and d_x . We calculated Ω analytically using the formula

$$\Omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{2(d-d_x)} d\lambda = \frac{2^{4(d-d_x)}}{\pi} \mathcal{B}\left(d - d_x + \frac{1}{2}, d - d_x + \frac{1}{2}\right),$$

where $\mathcal{B}(\cdot, \cdot)$ is the Beta function. Almost uniformly, $R > 1$, with R tending to approach 1 with n increasing as expected. For the most part, for given n R tends to fall with increasing m , the most noticeable exceptions occurring when both d is small and d_x is large, with the variation across m not always being monotonic, such that in some cases an “optimal” m (in the sense that the MSE most closely approximates its limiting value) lies within the range of m employed. It is not surprising that the implications for choice of m should vary across d , but perhaps surprising that they should vary across d_x . We conjecture that the fact that the results deteriorate as d_x increases when $d = 0$ is a second order effect due to the slow convergence of $n^{-1} \sum_{t=1}^n x_t^2$ to σ_x^2 when $d_x \geq 1/4$ (see Rosenblatt, (1961)), GLS then being asymptotically equivalent to OLS under our simple model for u_t ; as d increases it tends to “compensate” for d_x and the effect diminishes. So far as the comparison between Methods 1 and 2 is concerned, as expected Method 2 performs better for the smaller values of $d + d_x$, where OLS is $n^{1/2}$ -consistent (indeed with $d = 0$ in our model it is efficient), and surprisingly does so also in some cases when $d + d_x \geq 1/2$. However, the discrepancies in the former situation tend to be small, whereas on the occasions where Method 1 beats Method 2 they are sometimes quite substantial.

One can iterate the GLS procedures, on each step replacing $\widetilde{\beta}$ in (2.2) by the current estimate from the previous GLS step in order to re-estimate f prior to the next GLS step. Such iterations cannot improve asymptotic first-order efficiency, but might be expected to approach a form of frequency-domain Gaussian pseudo maximum likelihood estimate. Because the iterations tended to settle down after two iterations following the initial GLS estimation (suggesting fairly quick convergence), Table 3 contains results for R based on twice iterating either of Methods 1 and 2 (with $n = 256$). For Method 1, the iterations worsen the MSE ratio in all cases, although in 62 cases out of 90 the discrepancy is smaller than 2%. The relative performance of iteration tended to be even worse for $n = 64, 128$. This behaviour may be due to the fact that the estimate to which the iterations numerically converge is implicitly-defined, and the finite-sample properties of estimates of this type can sometimes be inferior to those of simpler, closed-form,

estimates. Moreover, in his original paper, Hannan (1963) referred to the possibility of iterating his, related, *GLS* procedure, but noted "experience suggests that this would not be worthwhile". For Method 2, iteration worsens matters in 42 cases out of 90, though by more than 2% in only 6 cases. Note the improvements achieved here when $d + d_x \geq 1/2$, which is the situation in which the *GLS* procedure has not been theoretically justified owing to C9 not holding; the larger $d + d_x$, the better. This appears to illustrate a general result of Robinson (1988), which indicates that certain iterations can improve a n^γ -consistent estimate to an $n^{1/2}$ -consistent one, in finitely many steps.

The finite sample efficiency of *GLS* relative to the preliminary estimates is also of interest, in view of the extra computation and the bandwidth choice that it requires. For the same model employed in the current simulations, Robinson and Hidalgo (1997) analytically calculated asymptotic efficiency relative to $\widehat{\beta}_\varphi$ under (2.3) (the preliminary estimate for Method 1) over the same (d, d_x) grid as in the present paper. To complement this, and recognize the popularity of *OLS*, Table 4 reports Monte Carlo efficiency $R^* = MSE(GLS)/MSE(OLS)$ for the same n , m and parameter values as before, with *GLS* computed according to Method 2. Notice that when $n = 64$ *GLS* is noticeably worse than *OLS* for smallish d , due to the imprecision introduced by the nonparametric spectral estimation. This effect diminishes with increasing n , such that with $n = 256$, R^* barely exceeds 1 for $d = 0$, in which case *OLS* is of course asymptotically efficient. For larger d , even when $n = 64$, *GLS* noticeably improves upon *OLS*, with monotonic reduction in R^* as m decreases, while there is global improvement in *GLS* as n increases. *GLS* tends to improve with increasing d_x , though some evidence of a trough around $d_x = 0.35$ is detected.

4. PROOF OF THEOREM 2.1

Denote $I_{xx}(\lambda_j)$ and $\widehat{f}(\lambda_j)$ by $I_{xx,j}$ and \widehat{f}_j respectively. It suffices to prove that

$$\frac{1}{n^{1/2}} \sum_{j=1}^{n-1} \widehat{f}_j^{-1} q_j \xrightarrow{d} N(0, \Omega) \quad (4.1)$$

and

$$\frac{1}{n} \sum_{j=1}^{n-1} \widehat{f}_j^{-1} I_{xx,j} \xrightarrow{p} \Omega, \quad (4.2)$$

where $q_j = w_x(\lambda_j) w_u^*(\lambda_j)$, with $w_u(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^n u_t e^{it\lambda}$. We establish only (4.1), omitting for the sake of brevity the much easier proof of (4.2).

Introduce the infeasible estimate of $f(\lambda)$,

$$\check{f}(\lambda) = \frac{1}{\bar{c}(2m+1)} \sum_{j=-m}^m c_j I_{uu}(\lambda + \lambda_j),$$

where $I_{uu}(\lambda) = |w_u(\lambda)|^2$ and abbreviate $f(\lambda_j)$ and $\check{f}(\lambda_j)$ as f_j and \check{f}_j respectively. Now (4.1) follows if

$$\frac{1}{n^{1/2}} \sum_{j=1}^{n-1} \check{f}_j^{-1} q_j \xrightarrow{d} N(0, \Omega) \quad (4.3)$$

and

$$\frac{1}{n^{1/2}} \sum_{j=1}^{n-1} \left(\widehat{f}_j^{-1} - \check{f}_j^{-1} \right) q_j \xrightarrow{p} 0. \quad (4.4)$$

A key difficulty in establishing (4.3) and (4.4) is that \check{f}_j , and its feasible version \widehat{f}_j , do not well approximate f_j for λ_j suitably near to zero frequency, where $f(\lambda)$ can have a pole, while dealing with f_j itself in this region poses technical difficulty. The latter problem is dealt with by suitable truncation with respect to frequency, somewhat after the manner of Robinson (1995a, b) in his study of estimating (2.1). For the first, Propositions A.2, A.3 and A.6 deal with behaviour of the centred quantities $\check{f}_j - E\check{f}_j$, the latter proposition proving the modification of (4.4) in which the bracketed factor is replaced by $\check{f}_j^{-1} - (E\check{f}_j)^{-1}$. However, $E\check{f}_j$ does not approximate f_j for small enough j - Proposition A.1 finds that, as $m \rightarrow \infty$ with n , according to C8, it varies like f_j for $j = 2m+1, \dots, [n/2]$, where $[.]$ denotes integer part, but like f_m for $j = 1, \dots, 2m$. When $d > 0$ we cannot therefore expect to have an asymptotically unbiased estimator for f_j at $j = 1, \dots, 2m$. So the modification of (4.4) with the bracketed factor replaced by $(E\check{f}_j)^{-1} - f_j^{-1}$ needs special care for these frequencies. Nevertheless, because the latter form an asymptotically negligible fraction of the Fourier frequencies used in the estimation, Proposition A.4 is able to show that the error thereby included is asymptotically negligible.

However, the latter requires $m = o(n^{1/2})$, under the mild Condition C4, as can be observed from Lemma B.7. On the other hand, since $\widehat{\beta}_{\widehat{f}_{-1}}$ involves the \widehat{f}_j as random denominators, then as in other semiparametric estimation problems, uniform convergence of the \widehat{f}_j is required. Following, for instance, Brillinger (1981, p.445), whose approach is employed in the proof of Proposition A.3, we can expect the rate of such uniform convergence to be $O_p(nm^{-k})$ when $E u_t^{4k} < \infty$. So, the latter order of magnitude, together with our condition $m = o(n^{1/2})$, implies that for uniform convergence we need $k \geq 3$, that is finite 12th moments. Although

Lemma B.7 indicates that the 12th moment condition on u_t could be relaxed, this is possible only at the expense of strengthening other conditions, notably Conditions C4 and C8. More specifically, allowing $F(\lambda) - F(0) = O(\lambda^{1-2d_x})$, or alternatively $\Gamma(j) = O(j^{2d_x-1})$, with $d_x < 1/2$, implies that we can choose $m = O(n^{(1+\delta)/2})$ for $0 < \delta < 1/3$. However, this would require strengthening the weak Condition C4. So, we have preferred to keep Condition C2 as it stands since from the aforementioned result of Brillinger, we would in any case need u_t to have finite 8th moments.

The remainder of the proof is mainly concerned with showing that \hat{f} can replace \check{f} , see (4.4). Though condition C9 plays a major role here, the proof is considerably more complex than the corresponding parts of the proofs of Hannan (1963), Robinson (1991), due to the allowance for long memory in u_t and x_t . In particular, care is needed in dealing with fact that discrete Fourier transforms behave differently close to zero frequency than they do far from it. The remainder of the current section deals principally with this matter of proving (4.4), but we first establish (4.3), whose left side is

$$\frac{1}{n^{1/2}} \sum_{j=1}^{n-1} f_j^{-1} q_j + \frac{1}{n^{1/2}} \sum_{j=1}^{n-1} \left(\check{f}_j^{-1} - (E\check{f}_j)^{-1} \right) q_j + \frac{1}{n^{1/2}} \sum_{j=1}^{n-1} \left((E\check{f}_j)^{-1} - f_j^{-1} \right) q_j. \quad (4.5)$$

The second and third terms of (4.5) are $o_p(1)$ by Propositions A.6 and A.4 respectively of Appendix A, while the first term converges in distribution to $N(0, \Omega)$ by the proof of Theorem 1 of Robinson and Hidalgo (1997). Thus (4.3) is proved.

We now prove (4.4). By the Schwarz inequality the left side of (4.4) has squared modulus bounded by

$$\left(\sum_{j=1}^{n-1} \left| (E\check{f}_j)^{-1} (\check{f}_j - \hat{f}_j) \right|^2 \left| \hat{f}_j^{-1} (E\check{f}_j)^2 \check{f}_j^{-1} \right|^2 \right) \left(\frac{1}{n} \sum_{j=1}^{n-1} \left\| (E\check{f}_j)^{-1} q_j \right\|^2 \right), \quad (4.6)$$

where $\|\cdot\|$ denotes the Euclidean norm, and the presence of the factors $E\check{f}_j$ and its reciprocal is due to remarks above. Let n be so large that $2m < [n/2]$. Writing $a_{j,n} = (E\check{f}_j)^{-1} f_j$, the expectation of the second bracketed factor of (4.6) is

$$\frac{2}{n} \sum_{j=1}^{2m} E \left\| a_{j,n} \frac{q_j}{f_j} \right\|^2 + \frac{2}{n} \sum_{j=2m+1}^{[n/2]} E \left\| a_{j,n} \frac{q_j}{f_j} \right\|^2.$$

The first term is $O(1)$ by Lemma B.7 since by Proposition A.1 part (a) $K^{-1} < |\lambda_m^{2d} (E\check{f}_j)| < K$, where K henceforth denotes an arbitrarily large but finite positive constant, and by C1 $f_j \lambda_m^{2d} \leq K(m/j)^{2d}$ for $1 \leq j \leq 2m$, so that

$|a_{j,n}| \leq K(m/j)^{2d}$. The second term of the last displayed expression is also $O(1)$ as we now show. By Proposition A.1 parts (b) and (c), $|a_{j,n} - 1| \leq Km/j$, so that term is bounded by

$$\begin{aligned} \frac{K}{n} \sum_{j=2m+1}^{[n/2]} (1 + mj^{-1})^2 E \left\| \frac{q_j}{f_j} \right\|^2 &\leq \frac{K}{n} \sum_{j=2m+1}^{[n/2]} E (f_j^{-1} I_{uu,j}) \operatorname{tr} (E (I_{xx,j})) \\ &\leq \frac{K}{n} \sum_{j=2m+1}^{[n/2]} \operatorname{tr} (E (I_{xx,j})) = O(1) \end{aligned}$$

where the first inequality follows by C5 and the fact that $f_j^{-1} \leq K$ (by C1 and C3), and the second inequality from Robinson's (1995a) Theorem 2, $E(f_j^{-1} I_{uu,j}) \leq K$ and $\sum_{j=2m+1}^{[n/2]} E I_{xx,j} \leq (2\pi)^{-1} \sum_{j=1}^n E x_t x_t' = O(n)$.

Thus, by Markov's inequality, the second factor of (4.6) is $O_p(1)$, and the proof of the Theorem is completed if the first factor of (4.6) is $o_p(1)$. We prove below that

$$\sum_{j=1}^{n-1} \left| (E \check{f}_j)^{-1} (\check{f}_j - \hat{f}_j) \right|^2 = o_p(1), \quad (4.7)$$

$$\sup_{j=1, \dots, [n/2]} \left| (E \check{f}_j)^{-1} \hat{f}_j - 1 \right| = o_p(1). \quad (4.8)$$

In addition, by Propositions A.1 and A.3,

$$\sup_{j=1, \dots, [n/2]} \left| (E \check{f}_j)^{-1} \check{f}_j - 1 \right| = o_p(1). \quad (4.9)$$

The proof then follows in the same spirit as in Hannan (1963), see also Hannan (1970, p.489). By (4.8) and (4.9),

$$\left| (E \check{f}_j)^{-1} \hat{f}_j \right| \geq K^{-1} > 0 \text{ and } \left| (E \check{f}_j)^{-1} \check{f}_j \right| \geq K^{-1} > 0 \quad (4.10)$$

with probability approaching one as $n \rightarrow \infty$. The first factor of (4.6) is thus bounded by

$$\begin{aligned} \sup_{j=1, \dots, [n/2]} \left| \hat{f}_j (E \check{f}_j)^{-2} \check{f}_j \right|^{-2} \sum_{j=1}^{n-1} \left| (E \check{f}_j)^{-1} (\check{f}_j - \hat{f}_j) \right|^2 &\leq K \sum_{j=1}^{n-1} \left| (E \check{f}_j)^{-1} (\check{f}_j - \hat{f}_j) \right|^2 \\ &= o_p(1), \end{aligned}$$

using (4.7), so the first factor of (4.6), and thus the left side of (4.4), is $o_p(1)$.

It remains to prove (4.7) and (4.8). We begin with (4.7). By (1.1)

$$\widehat{f}_j - \check{f}_j = 2 \left(\widetilde{\beta} - \beta \right)' b_j + \left(\widetilde{\beta} - \beta \right)' B_j \left(\widetilde{\beta} - \beta \right), \quad (4.11)$$

where, denoting $\operatorname{Re}(z)$ as the real part of the complex number z ,

$$b_j = \frac{1}{\bar{c}(2m+1)} \sum'_{\ell(j)} c_\ell \operatorname{Re}(q_{j+\ell}), \quad B_j = \frac{1}{\bar{c}(2m+1)} \sum'_{\ell(j)} c_\ell I_{xx,j+\ell},$$

in which $\sum'_{\ell(j)} = \sum_{\ell=-m, \ell+j \neq 0}^m$. Notice that the exclusion of the non-mean-corrected quantities $q_{j+\ell}$, $I_{xx,j+\ell}$ for $j+\ell \neq 0$ corresponds to the sample-mean-correction in (2.2).

Noting that, by boundedness of the c_ℓ and periodicity of $I_{xx}(\lambda)$,

$$\|B_j\| \leq \frac{K}{m} \operatorname{tr} \left(\sum_{k=1}^n I_{xx,k} \right) \leq \frac{K}{m} \operatorname{tr} \left(\sum_{t=1}^n x_t x_t' \right) = O_p \left(\frac{n}{m} \right) \quad (4.12)$$

uniformly in j , whereas

$$\sum_{j=1}^n \|B_j\| \leq K \operatorname{tr} \left(\sum_{k=1}^n I_{xx,k} \right) = O_p(n),$$

the contribution to (4.7) due to the final term of (4.11) is bounded by

$$\left\| \widetilde{\beta} - \beta \right\|^4 \sum_{j=1}^{n-1} (E\check{f}_j)^{-2} \|B_j\|^2 \leq K \left\| \widetilde{\beta} - \beta \right\|^4 \sup_j \|B_j\| \sum_{j=1}^n \|B_j\| = O_p \left(\frac{1}{m} \right),$$

since $(E\check{f}_j)^{-1} \leq K$ by Proposition A.1 in Appendix A and C9 holds.

Next, the contribution to (4.7) due to b_j in (4.11) is bounded by

$$4n \left\| \widetilde{\beta} - \beta \right\|^2 \left(\sum_{j=1}^{\lfloor n^{2/3} \rfloor} \left\| (E\check{f}_j)^{-1} n^{-1/2} b_j \right\|^2 + \sum_{j=1+\lfloor n^{2/3} \rfloor}^{\lfloor n/2 \rfloor} \left\| (E\check{f}_j)^{-1} n^{-1/2} b_j \right\|^2 \right). \quad (4.13)$$

By elementary inequalities the first sum on (4.13) is bounded by

$$\frac{K}{m^2} \sum_{j=1}^{2m} \left\| \frac{1}{n^{1/2}} \sum'_{\ell(j)} c_\ell \lambda_m^{2d} \operatorname{Re}(q_{j+\ell}) \right\|^2 + \frac{K}{m^2} \sum_{j=2m+1}^{\lfloor n^{2/3} \rfloor} \left\| \frac{1}{n^{1/2}} \sum'_{\ell(j)} c_\ell f_{j+\ell}^{-1} \operatorname{Re}(q_{j+\ell}) \right\|^2, \quad (4.14)$$

because $(E\check{f}_j)^{-1} \leq K \lambda_m^{2d}$ for $1 \leq j \leq 2m$ by Proposition A.1 part (a) and $(E\check{f}_j)^{-1} \leq K f_{j+\ell}^{-1}$, $|\ell| \leq m$, for $2m < j < \lfloor n/2 \rfloor$ by Proposition A.1 part (b). But

for $1 \leq j \leq 2m$ and $j + \ell \neq 0$, we have $f_{j+\ell} \lambda_m^{2d} < K (m/|j + \ell|)^{2d}$, so by Lemma B.7 the expectation of the first term of (4.14) is bounded by

$$\frac{K}{m^2} \sum_{j=1}^{2m} 1 = O\left(\frac{1}{m}\right) = o(1),$$

using C8, whereas the second term of (4.14) has expectation bounded by

$$\frac{K}{m^2} \sum_{j=2m+1}^{\lceil n^{2/3} \rceil} 1 = O\left(\frac{n^{2/3}}{m^2}\right) = o(1),$$

using C8 and proceeding as in the proof of the first term on the right of (B.11) but taking the weights $a_{j,n}$ there to be bounded by K .

Now consider the second sum on the right of (4.13). Let $w_{\varepsilon,j} = (2\pi n)^{-1/2} \sum_{t=1}^n \varepsilon_t e^{it\lambda_j}$. By elementary inequalities and $\operatorname{Re}(q_{j+\ell}) = (q_{j+\ell} + q_{j+\ell}^*)/2$, that sum is bounded by

$$\begin{aligned} & K n^{-1} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \left\| \frac{1}{(2m+1)} \sum'_{\ell(j)} c_\ell \frac{f_{j+\ell}^{1/2}}{(E\check{f}_j)} \left(\frac{w_{u,j+\ell}}{f_{j+\ell}^{1/2}} - \frac{(2\pi)^{1/2} w_{\varepsilon,j+\ell}}{\sigma} \right) w_{x,j+\ell}^* \right\|^2 \\ & + K n^{-1} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \left\| \frac{1}{(2m+1)} \sum'_{\ell(j)} c_\ell \frac{f_{j+\ell}^{1/2}}{(E\check{f}_j)} w_{\varepsilon,j+\ell} w_{x,j+\ell}^* \right\|^2. \end{aligned} \quad (4.15)$$

Because by C1 and C3 and Proposition A.1 parts (b) and (c) $(E\check{f}_j)^{-1} \leq K f_j^{-1} \leq K f_{j+\ell}^{-1/2}$, and because $E w_{\varepsilon,j+\ell} w_{\varepsilon,k+\ell}^* = (\sigma^2/2\pi) \mathcal{I}(j=k)$ by C2, see Brillinger (1981, Exercise 4.8.12), the expectation of the second term of (4.15) is

$$\begin{aligned} \frac{K}{n} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \frac{1}{(2m+1)^2} \sum'_{\ell(j)} c_\ell^2 \operatorname{tr} \{E I_{xx,j+\ell}\} & \leq \frac{K}{nm^2} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \sum_{\ell=-m}^m \operatorname{tr} \{E I_{xx,j+\ell}\} \\ & \leq \frac{K}{nm} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \operatorname{tr} \{E I_{xx,j}\} = O(m^{-1}), \end{aligned}$$

since the sum is bounded by $\sum_{j=1}^{n-1} E I_{xx,j} = (2\pi)^{-1} \sum_{t=1}^n E(x_t x_t') = O(n)$. On the other hand, by the Schwarz inequality the expectation of the first term of (4.15)

is bounded by

$$\begin{aligned} & \frac{K}{nm^2} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \left(\sum_{\ell=-m}^m E \left| \frac{w_{u,j+\ell}}{f_{j+\ell}^{1/2}} - \frac{(2\pi)^{1/2} w_{\varepsilon,j+\ell}}{\sigma} \right|^2 \right) \sum_{\ell=-m}^m \text{tr} \{EI_{xx,j+\ell}\} \\ & \leq \frac{K}{nm^2} \sum_{j=1+\lceil n^{2/3} \rceil}^{\lfloor n/2 \rfloor} \sum_{\ell=-m}^m \frac{\log(j+\ell)}{j+\ell} \sum_{\ell=-m}^m \text{tr} \{EI_{xx,j+\ell}\} = O\left(\frac{\log^2 n}{m}\right) \end{aligned}$$

by Robinson's (1995a) Theorem 2, $\sum_{\ell=1}^n \ell^{-1} \log \ell = O(\log^2 n)$ and proceeding as in the proof of the second term of (4.15). Thus from C9 (4.13) is $o_p(1)$, to complete the proof that (4.7) is $o_p(1)$.

Finally consider (4.8). By the triangle inequality the left side of (4.8) is bounded by

$$\sup_j \left| (E\check{f}_j)^{-1} (\check{f}_j - \widehat{f}_j) \right| + \sup_j \left| (E\check{f}_j)^{-1} (\check{f}_j - E(\check{f}_j)) \right|.$$

The second term is $o_p(1)$ by (4.9), whereas, by the triangle inequality and (4.11), the first term is bounded by

$$2 \left\| \widetilde{\beta} - \beta \right\| \sup_j \left\| (E\check{f}_j)^{-1} b_j \right\| + \left\| \widetilde{\beta} - \beta \right\|^2 (E\check{f}_j)^{-1} \sup_j \|B_j\|. \quad (4.16)$$

The second term of (4.16) is $o_p(1)$ because of C9, (4.12) and Proposition A.1, whereas the first term is bounded by

$$\left\{ \left\| \widetilde{\beta} - \beta \right\|^2 \sum_{j=1}^{\lfloor n/2 \rfloor} \left\| (E\check{f}_j)^{-1} b_j \right\|^2 \right\}^{1/2} = o_p(1)$$

since we have already shown that (4.13) = $o_p(1)$. This concludes the proof of (4.8) and the theorem. \square

APPENDIX A

Our first proposition is concerned with the asymptotic behaviour of $E\check{f}_j$ across three suitably chosen, exhaustive subsets of the Fourier frequencies. We find it convenient to approximate $E\check{f}_j$, over two of these subsets, by

$$\widetilde{f}_j = \frac{1}{\bar{c}(2m+1)} \sum'_{\ell(j)} c_\ell f_{j+\ell}, \quad 1 \leq j \leq \lfloor n/2 \rfloor; \quad \widetilde{f}_j = \widetilde{f}_{n-j}, \quad 1 + \lfloor n/2 \rfloor \leq j < n,$$

whose varying orders of magnitude are indicated in the proposition. We shall then employ \check{f}_j subsequently in the proposition, largely as abbreviating notation for these orders of magnitude.

Proposition A.1 *Let C1-C3 and C7-C8 hold. Then as $n \rightarrow \infty$,*

(a)

$$\sup_{1 \leq j \leq 2m} \left| \lambda_m^{2d} \left(E(\check{f}_j) - \tilde{f}_j \right) \right| = O\left(\frac{\log m}{m^{1-2d}} \right),$$

where $K^{-1} \leq \left| \lambda_m^{2d} \tilde{f}_j \right| \leq K$ for $j = 1, \dots, 2m$, independently of j and n .

(b) For arbitrarily small $\delta > 0$,

$$\sup_{2m < j \leq \lfloor \delta n / 2 \rfloor} \left| f_j^{-1} \left(E(\check{f}_j) - \tilde{f}_j \right) \right| = O\left(\frac{\log m}{m} \right),$$

where $K^{-1} \leq \left| f_j^{-1} \tilde{f}_j \right| \leq K$ for $j = 2m + 1, \dots, \lfloor n/2 \rfloor$ and

$\sup_{j=2m+1, \dots, \lfloor n/2 \rfloor} (j/m) \left| f_j^{-1} \tilde{f}_j - 1 \right| = O(1)$ as $n \rightarrow \infty$.

(c) For arbitrarily small $\delta > 0$, as $n \rightarrow \infty$,

$$\sup_{\lfloor \delta n / 2 \rfloor < j \leq \lfloor n/2 \rfloor} \left| E(\check{f}_j) - f_j \right| = O\left(\frac{m}{n} \right).$$

Proof. We first prove (a). By definition

$$E(\check{f}_j) - \tilde{f}_j = \frac{1}{\bar{c}(2m+1)} \sum'_{\ell(j)} c_\ell \{ E I_{j+\ell} - f_{j+\ell} \},$$

abbreviating $I_{uu,j}$ as I_j . By C7 and Theorems 1 and 2 of Robinson (1995a), see in particular his (4.2), the right side of the last displayed equation is bounded in absolute value by

$$\begin{aligned} & \frac{K}{m} \sum_{\substack{\ell=-m \\ 0 \neq |j+\ell| \leq r}}^m f_{j+\ell} + \frac{K}{m} \sum'_{\substack{\ell=-m \\ r < |j+\ell|}}^m f_{j+\ell} \left(\frac{\log |j+\ell|}{|j+\ell|} \right) \\ & \leq \frac{Kr}{m} \sum_{\ell=1}^r f_\ell + \frac{K \log m}{m} \sum_{\ell=r+1}^{2m} \ell^{-1} f_\ell, \end{aligned}$$

for some $r > 0$. (Observe that the component $|(j+\ell)/n|^\eta$ in Robinson's (1995a) Theorem 2 does not appear since we employ $f_{j+\ell}$ instead of its approximation

$C\lambda_{j+\ell}^{-2d}$.) By C1, and with $r = \log^{1/2} m$, the right side of the last displayed inequality is bounded by

$$\frac{K \log m}{m} \sum_{\ell=1}^{2m} \ell^{-1} \lambda_{\ell}^{-2d} = O((\log m) \lambda_m^{-2d} / m^{1-2d}).$$

To complete the proof of part (a) we need to show that $K^{-1} \leq \left| \lambda_m^{2d} \tilde{f}_j \right| \leq K$.

First, $\lambda_m^{2d} \tilde{f}_j$ has absolute value upper-bounded by

$$K m^{2d-1} \sum_{\ell=1}^{2m} \ell^{-2d} \leq K,$$

whereas, by C7 and for n sufficiently large, $\min_{[m/4] \leq \ell \leq [m/2]} c_{\ell} > K^{-1}$, so

$$\left| \lambda_m^{2d} \tilde{f}_j \right| \geq K^{-1} m^{2d-1} \sum_{\ell=[m/4]}^{[m/2]} (j+\ell)^{-2d} \geq K^{-1}.$$

Now consider part (b). By definition of $f_j^{-1} \left(E(\check{f}_j) - \tilde{f}_j \right)$ and triangle inequality we have

$$\left| f_j^{-1} \left(E(\check{f}_j) - \tilde{f}_j \right) \right| \leq f_j^{-1} \frac{1}{\bar{c}(2m+1)} \sum_{\ell=-m}^m c_{\ell} |E(I_{j+\ell}) - f_{j+\ell}|,$$

which, by Theorem 2 of Robinson (1995a) and its routine extension for $2m < j < [n\delta/2]$ with arbitrarily small $\delta > 0$, is bounded by

$$f_j^{-1} \frac{K}{(2m+1)} \sum_{\ell=-m}^m f_{j+\ell} \left| \frac{\log(j+\ell)}{j+\ell} \right| = O(m^{-1} \log m),$$

because $|\log(j+\ell)/(j+\ell)| = O(m^{-1} \log m)$ and $f_j^{-1} f_{j+\ell} \leq K$ by C1 and C3 uniformly in ℓ for all $2m < j \leq [n\delta/2]$.

We prove finally that $K^{-1} \leq \left| f_j^{-1} \tilde{f}_j \right| \leq K$. By C1,

$$\left| f_j^{-1} \tilde{f}_j \right| \leq \frac{K}{m} \sum_{\ell=-m}^m \left| \frac{f_{j+\ell}}{f_j} \right| \leq \frac{K}{m} \sum_{\ell=-m}^m \left| \frac{j}{j+\ell} \right|^{2d} \leq K,$$

because for $j > 2m$ and $|\ell| \leq m$, we have $2^{-1} < |1 + \ell/j|$. On the other hand, proceeding as in part (a)

$$\left| f_j^{-1} \tilde{f}_j \right| \geq K^{-1} m^{-1} \sum_{\ell=[m/4]}^{[m/2]} \left| \frac{j}{j+\ell} \right|^{2d} \geq K^{-1},$$

because for $j > 2m$ and $|\ell| \leq m$, $|1 + \ell/j| < 3/2$. Next we show that $\sup_{2m < j \leq [n/2]} \binom{j}{m} \left| f_j^{-1} \tilde{f}_j - 1 \right| = O(1)$ as $n \rightarrow \infty$. By definition

$$\begin{aligned} \binom{j}{m} \left| f_j^{-1} \tilde{f}_j - 1 \right| &= \binom{j}{m} \frac{1}{\bar{c}(2m+1)} \left| \sum_{\ell=-m}^m c_\ell \left(\frac{f_{j+\ell}}{f_j} - 1 \right) \right| \\ &\leq \binom{j}{m} \frac{K}{(2m+1)} \sum_{\ell=-m}^m \frac{|\ell|}{j} \leq K \end{aligned} \quad (\text{A.1})$$

by the mean value theorem and triangle inequality, so that the $\sup_{2m < j \leq [n/2]} \binom{j}{m} \left| f_j^{-1} \tilde{f}_j - 1 \right| = O(1)$. This concludes the proof of part (b).

Finally we prove (c). First, by definition

$$\begin{aligned} E(\check{f}_j) - f_j &= \frac{1}{\bar{c}(2m+1)} \sum_{\ell=-m}^m c_\ell (E(I_{j+\ell}) - f_{j+\ell}) \\ &\quad + \frac{1}{\bar{c}(2m+1)} \sum_{\ell=-m}^m c_\ell (f_{j+\ell} - f_j). \end{aligned} \quad (\text{A.2})$$

Because $[\delta n/2] < j \leq [n/2]$ and $f(\lambda)$ is continuously differentiable for $\delta < |\lambda| \leq \pi$ by C3, the second term on the right of (4.2) is bounded by

$$\frac{K}{(2m+1)} \sum_{\ell=-m}^m c_\ell \frac{|\ell|}{n} (1 + o(1)) = O\left(\frac{m}{n}\right).$$

Finally, the first term on the right of (A.2) is $O(n^{-1} \log n)$ by an obvious extension of the proof of Robinson's (1995a) Theorem 2 to $0 < \delta \leq |\lambda| \leq \pi$. This concludes the proof of part (c) and of the Proposition. \square

Proposition A.2 *Assuming C1-C3 and C7-C8, as $n \rightarrow \infty$*

$$\tilde{f}_j^{-2} \text{Cov}(\tilde{f}_j, \tilde{f}_j) = \begin{cases} O(m^{-1}) & \text{if } 2m < |j| \leq [n/2] \\ O(g(m)) & \text{if } 1 \leq |j| \leq 2m, \end{cases}$$

where $g(m) = m^{4d-2} \mathcal{I}(d > 1/4) + m^{-1} \log m \mathcal{I}(d = 1/4) + m^{-1} \mathcal{I}(d < 1/4)$.

Proof. We begin by estimating $\text{Cov}(\tilde{f}_j, \tilde{f}_j)$, which is

$$\frac{1}{4\pi^2 n^2 \bar{c}^2 (2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_\ell c_k \sum_{t_1} \sum_{t_2} \sum_{t_3}^n \sum_{t_4=1} \{ \gamma_{t_1-t_3} \gamma_{t_2-t_4} + \gamma_{t_1-t_4} \gamma_{t_2-t_3} \\ + \text{cum}(u_{t_1}, u_{t_2}, u_{t_3}, u_{t_4}) \} e^{i(t_1-t_2)\lambda_{\ell+j} - i(t_3-t_4)\lambda_{k+j}}. \quad (\text{A.3})$$

We deal only with the contributions from the first and third terms in braces, that from the second being handled similarly to the first. Applying Brillinger (1981, (2.6.3) p. 26, and (2.10.3) p. 39), the contribution from the third term of (A.3) is bounded in absolute value by

$$\frac{K}{n^2 (2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_\ell c_k \left| \int_{[-\pi, \pi]^3} \tau(-\lambda) \tau(-\mu) \tau(-\zeta) \tau(\lambda + \mu + \zeta) \right. \\ \times D(\lambda + \lambda_{\ell+j}) D(\mu - \lambda_{\ell+j}) D(\zeta - \lambda_{k+j}) D(\lambda_{k+j} - (\lambda + \mu + \zeta)) d\lambda d\mu d\zeta \\ = \frac{K}{n^2 (2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_\ell c_k |\tau(\lambda_{\ell+j}) \tau(-\lambda_{\ell+j}) \tau(\lambda_{k+j}) \tau(-\lambda_{k+j})| \quad (\text{A.4}) \\ \times \int_{[-\pi, \pi]^3} \frac{\tau(-\lambda) \tau(-\mu) \tau(-\zeta) \tau(\lambda + \mu + \zeta)}{\tau(\lambda_{\ell+j}) \tau(-\lambda_{\ell+j}) \tau(\lambda_{k+j}) \tau(-\lambda_{k+j})} \\ \times D(\lambda + \lambda_{\ell+j}) D(\mu - \lambda_{\ell+j}) D(\zeta - \lambda_{k+j}) D(\lambda_{k+j} - (\lambda + \mu + \zeta)) d\lambda d\mu d\zeta|,$$

where $D(w) = \sum_{\ell=0}^{n-1} e^{-i\ell w}$ is the Dirichlet kernel and we note that $f(\lambda) > 0$ implies that $|\tau(\lambda)| > 0$. Because

$$\int_{[-\pi, \pi]^3} D(\lambda + \lambda_{\ell+j}) D(\mu - \lambda_{\ell+j}) D(\zeta - \lambda_{k+j}) D(\lambda_{k+j} - (\lambda + \mu + \zeta)) d\lambda d\mu d\zeta = (2\pi)^3 n,$$

using the identity

$$q_1 q_2 q_3 q_4 = (q_1 q_2 - 1)(q_3 q_4 - 1) + \sum_{\ell=1}^2 (q_{2\ell-1} - 1)(q_{2\ell} - 1) + \sum_{i=1}^4 (q_i - 1) + 1,$$

and proceeding as in the proof of (4.28) in Robinson (1995b), the right side of (A.4) is bounded by

$$\frac{K}{(2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_\ell c_k f_{\ell+j} f_{k+j} (n^{-1} + |\ell+j|^{-1} |k+j|^{-1}) \quad (\text{A.5}) \\ + (|\ell+j|^{-1/2} |k+j|^{-1} + |\ell+j|^{-1} |k+j|^{-1/2} + n^{-1/2} |\ell+j|^{-1/2} |k+j|^{-1/2}),$$

since $\left|f_{\ell+j}^{-1/2} \tau(\lambda_{\ell+j})\right| < K$.

Next, we examine the first term of (A.3), which can be written

$$\begin{aligned} & \frac{1}{4\pi^2 n^2 \bar{c}^2 (2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_{\ell} c_k \left\{ \int D(\theta) D(\lambda_{k-\ell} - \theta) f(\theta + \lambda_{\ell+j}) d\theta \right. \\ & \quad \left. \times \int D(-\theta) D(\theta + \lambda_{\ell-k}) f(\theta - \lambda_{k+j}) d\theta \right\}. \end{aligned} \quad (\text{A.6})$$

We study the first term inside the braces of (A.6), the second being identical. Adding and subtracting $f_{\ell+j} \int D(\theta) D(\lambda_{k-\ell} - \theta) d\theta$, this term is

$$\int D(\theta) D(\lambda_{k-\ell} - \theta) (f(\theta + \lambda_{\ell+j}) - f_{\ell+j}) d\theta + f_{\ell+j} \int D(\theta) D(\lambda_{k-\ell} - \theta) d\theta.$$

By Theorem 2 of Robinson (1995a), the first term of this expression is $O(n(\log|k+j|)f_{\ell+j}/|\ell+j|)$, whereas the second term is zero unless $k = \ell$, in which case it is $2\pi n f_{\ell+j}$. Thus, as $n \rightarrow \infty$, the modulus of (A.6) is bounded by

$$\begin{aligned} & \frac{K}{(2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_{\ell} c_k f_{\ell+j} f_{k+j} \frac{\log^2|k+j|}{|\ell+j||k+j|} \\ & + \frac{K}{(2m+1)^2} \sum'_{\ell(j)} c_{\ell}^2 f_{\ell+j}^2 \left(\frac{\log|\ell+j|}{|\ell+j|} + 1 \right). \end{aligned} \quad (\text{A.7})$$

Similarly, the second term of (A.3) is, as $n \rightarrow \infty$, bounded by

$$\begin{aligned} & \frac{K}{(2m+1)^2} \sum'_{\ell(j)} \sum'_{k(j)} c_{\ell} c_k f_{\ell+j} f_{k+j} \frac{\log^2|k+j|}{|\ell+j||k+j|} \\ & + \frac{K}{(2m+1)^2} \sum'_{\ell(j), k(j)} c_{\ell} c_k f_{\ell+j} f_{k+j} \left(\frac{\log|\ell+j|}{|\ell+j|} + 1 \right), \end{aligned} \quad (\text{A.8})$$

where $\sum'_{\ell(j), k(j)}$ denotes the terms in the double sum $\sum'_{\ell(j)} \sum'_{k(j)}$ such that $k + \ell + 2j = 0$. Thus,

$$|Cov(\check{f}_j, \check{f}_j)| \leq (\text{A.5}) + (\text{A.7}) + (\text{A.8}). \quad (\text{A.9})$$

We now turn to the proof for $2m < |j| \leq [n/2]$. Applying Lemma B.1 with $\psi = 0, 1/2, 1$, the first term on the right of (A.9) is $O(m^{-3/2} f_j^2)$ by C8. Next, consider the second term on the right of (A.9), namely (A.7). Its first term is bounded by

$$K \frac{\log^2|j|}{j^2} \left(\frac{1}{(2m+1)} \sum'_{\ell(j)} c_{\ell} f_{\ell+j} \right)^2 = O\left(\frac{\log^2|j|}{j^2} f_j^2 \right),$$

from Lemma B.1 with $\psi = 0$. The second term of (A.7) is dominated by

$$K \frac{1}{m^2} \sum'_{\ell(j)} c_\ell^2 f_{\ell+j}^2 = O(m^{-1} f_j^2),$$

because by the mean value theorem $f_j^{-2} |f_{\ell+j}^2 - f_j^2| \leq K(\ell/j) < K$ for $2m < j \leq [n/2]$, so $f_{\ell+j}^2 = O(f_j^2)$ and thus (A.7) is $O(m^{-1} f_j^2)$.

The third term on the right of (A.9) (that is, (A.8)) is, proceeding as with (A.7), $O(m^{-3/2} f_j^2)$, by C8 and the fact that, for $j > 2m$, the last term of (A.8) is zero. Thus, we conclude that (A.3) is $O(m^{-1} f_j^2)$ and that $\tilde{f}_j^{-2} \text{Cov}(\check{f}_j, \check{f}_j) = O(m^{-1})$, since by Proposition A.1 parts (b) and (c), $\tilde{f}_j^{-2} = O(f_j^{-2})$.

Next, for $1 \leq |j| \leq 2m$. Using Lemma B.2 instead of Lemma B.1, it is straightforward to observe that (A.9) is

$$O(n^{4d} m^{-2} (m^{1-4d} \mathcal{I}(d < 1/4) + \log m \mathcal{I}(d = 1/4) + \mathcal{I}(d > 1/4))).$$

But by Proposition A.1 part (a), $\tilde{f}_j^{-2} = O(\lambda_m^{4d})$, which concludes the proof. \square

Proposition A.3 *Let C1-C3, C7 and C8 hold. Then, as $n \rightarrow \infty$,*

$$(a) \sup_{j=1, \dots, 2m} \left| \tilde{f}_j^{-1} (\check{f}_j - E\check{f}_j) \right| = o_p(1)$$

$$(b) \sup_{j=2m+1, \dots, k} \left| \tilde{f}_j^{-1} (\check{f}_j - E\check{f}_j) \right| = O_p(k^{1/6} / m^{1/2}).$$

Proof. We first show (b). Observing that by triangle inequality,

$$\sup_{j=2m+1, \dots, k} \left| \tilde{f}_j (\check{f}_j - E\check{f}_j) \right| \leq \sup_{j=2m+1, \dots, k} \left| \tilde{f}_j h_{1j} \right| + \sup_{j=2m+1, \dots, k} \left| \tilde{f}_j h_{2j} \right|$$

where

$$h_{1j} = \frac{1}{2m+1} \sum'_{\ell(j)} c_\ell f_{\ell+j} \left\{ \left(\frac{I_{\ell+j}}{f_{\ell+j}} - 2\pi \frac{I_{\varepsilon\varepsilon, \ell+j}}{\sigma^2} \right) - \left(\frac{E(I_{\ell+j})}{f_{\ell+j}} - 1 \right) \right\} \quad (\text{A.10})$$

and

$$h_{2j} = \frac{1}{2m+1} \sum'_{\ell(j)} c_\ell f_{\ell+j} \left(2\pi \frac{I_{\varepsilon\varepsilon, \ell+j}}{\sigma^2} - 1 \right), \quad (\text{A.11})$$

the proof of (b) is in two parts. Since $(\sup_j |a_j|)^v \leq \sum_j |a_j|^v$, we first show that

$$\left(\sup_{j=2m+1, \dots, k} \left| \tilde{f}_j h_{1j} \right| \right)^2 \leq \sum_{j=2m+1}^k \left| \tilde{f}_j h_{1j} \right|^2 = O_p\left(\frac{\log^2 k}{m}\right) \quad (\text{A.12})$$

and then that

$$\left(\sup_{j=2m+1, \dots, k} \left| \tilde{f}_j h_{2j} \right| \right)^6 \leq \sum_{j=2m+1}^k \left| \tilde{f}_j h_{2j} \right|^6 = O_p \left(\frac{k}{m^3} \right). \quad (\text{A.13})$$

Using the inequality in (A.12), and that $(a+b)^2 \leq 2a^2+2b^2$, $\left(\sup_{j=2m+1, \dots, k} \left| \tilde{f}_j h_{1j} \right| \right)^2$ is bounded by

$$\begin{aligned} & 2 \sum_{j=2m+1}^k \tilde{f}_j^{-2} \left(\frac{1}{2m+1} \sum'_{\ell(j)} c_\ell f_{\ell+j} \left(\frac{I_{\ell+j}}{f_{\ell+j}} - 2\pi \frac{I_{\varepsilon\varepsilon, \ell+j}}{\sigma^2} \right) \right)^2 \\ & + 2 \sum_{j=2m+1}^k \tilde{f}_j^{-2} \left(\frac{1}{2m+1} \sum'_{\ell(j)} c_\ell f_{\ell+j} \left(\frac{E(I_{\ell+j})}{f_{\ell+j}} - 1 \right) \right)^2. \end{aligned} \quad (\text{A.14})$$

But proceeding step-by-step as in the proof of (4.8) in Robinson (1995b), the expectation of the first term of (A.14) is bounded by

$$\frac{K}{m^2} \sum_{j=2m+1}^k \tilde{f}_j^{-2} f_j^2 \sum_{\ell=-m}^m \left(\frac{f_{\ell+j}}{f_j} \right)^2 \frac{\log(\ell+j)}{\ell+j} \leq \frac{K}{m^2} \sum_{\ell=-m}^m \log^2 k = O(m^{-1} \log^2 k),$$

because $|f_j^{-1} f_{\ell+j}| \leq K$ by C1 and C3 and $|\tilde{f}_j^{-1} f_j| \leq K$ by Proposition A.1 parts (b) and (c). Next, using Robinson's (1995a) Theorem 2 and its obvious extension for all $j = 2m+1, \dots, [n/2]$, the second term of (A.14) is bounded by

$$K \sum_{j=2m+1}^k \left| \tilde{f}_j^{-1} \frac{1}{2m+1} \sum_{\ell=-m}^m c_\ell f_{\ell+j} \frac{\log(\ell+j)}{\ell+j} \right|^2 \leq K \sum_{j=2m+1}^k \left| \tilde{f}_j^{-1} f_j \frac{\log k}{(j-m)} \right|^2 = O\left(\frac{\log^2 k}{m} \right),$$

applying Lemma B.1 with $\psi = 0$, and using the fact that $\sup_{j=2m+1, \dots, k} |\tilde{f}_j^{-1} f_j| \leq K$ by Proposition A.1 parts (b) and (c). Next, using the inequality in (A.13), the expectation of $\left(\sup_{j=2m+1, \dots, k} \left| \tilde{f}_j^{-1} h_{2j} \right| \right)^6$ is bounded by

$$\sum_{j=2m+1}^k \tilde{f}_j^{-6} f_j^6 E \left| \frac{1}{2m+1} \sum_{\ell=-m}^m c_\ell \left(\frac{f_{\ell+j}}{f_j} \right) \left(2\pi \frac{I_{\varepsilon\varepsilon, \ell+j}}{\sigma^2} - 1 \right) \right|^6 = O(km^{-3})$$

from the proof of Brillinger's (1981) Theorem 7.4.4., and $\sup_{j=2m+1, \dots, k} |\tilde{f}_j^{-1} f_j| \leq K$ and $|f_{\ell+j}/f_j| \leq K$, as argued above. This concludes the proof of part (b).

To prove part (a), we first consider the case $d \geq 1/4$. Write $\chi_j = \tilde{f}_j^{-1} (\check{f}_j - E\check{f}_j)$ and $\tilde{\gamma}_r = \hat{\gamma}_r - E\hat{\gamma}_r$, where $\hat{\gamma}_r = n^{-1} \sum_{t=1}^{n-|r|} u_t u_{t+|r|}$. By direct calculation,

$$\begin{aligned} |\chi_j|^2 &= \left| \frac{1}{\tilde{f}_j \bar{c} (2m+1)} \sum'_{\ell(j)} c_\ell (I_{\ell+j} - E(I_{\ell+j})) \right|^2 \\ &= \left| \frac{1}{2\pi \tilde{f}_j \bar{c}} \sum_{r=1-n}^{n-1} C_j(r) e^{ir\lambda_j} \tilde{\gamma}_r \right|^2, \end{aligned} \quad (\text{A.15})$$

where

$$C_j(r) = \frac{1}{(2m+1)} \sum'_{\ell(j)} c_\ell e^{ir\lambda_\ell}. \quad (\text{A.16})$$

Note first that

$$\sup_{j=1, \dots, 2m} |\chi_j|^2 \leq K \lambda_m^{4d} \sup_{j=1, \dots, 2m} \left| \sum_{r=1-n}^{n-1} C_j(r) e^{ir\lambda_j} \tilde{\gamma}_r \right|^2$$

by definition of \tilde{f}_j and χ_j in (A.15) and by Proposition A.1 part (a). By the Schwarz inequality, the right side of the last displayed inequality is bounded by

$$(K \lambda_m^{4d}) \sup_{j=1, \dots, 2m} \left| \sum_{r=1-n}^{n-1} |C_j(r)| \right| \sup_{j=1, \dots, 2m} \left| \sum_{r=1-n}^{n-1} |C_j(r)| \tilde{\gamma}_r^2 \right|.$$

The second factor is $O(m^{-1}n)$ by Lemma B.4, whereas the third factor is $O_p(m^{-1}n(n^{4d-2}\mathcal{I}(d \geq 1/4) + n^{-1} \log n \mathcal{I}(d = 1/4)))$ by the properties of $\tilde{\gamma}_r^2$ established in Lemma B.5. The conclusion is now immediate since $d < 1/2$.

To complete the proof we are left with the case $d < 1/4$. The proof proceeds step-by-step as in part (b), noting that k there is now $2m$, $\tilde{f}_j^{-1} f_{\ell+j} = O(m^{2d}/|\ell+j|^{2d})$ and $m^{4d-1} = o(1)$ as $d < 1/4$. \square

For the remaining propositions, let q_j and $w_{x,j}$ denote typical components of the $p \times 1$ vectors $I_{x_u,j}$ and $w_x(\lambda_j)$ respectively, while $F(\lambda)$ and $\Gamma(s)$ will here denote the spectral distribution and autocovariance function of the corresponding element of x_t .

Proposition A.4 *Assuming C1-C5 and C7-C8, as $n \rightarrow \infty$,*

$$\frac{1}{n^{1/2}} \sum_{j=1}^{[n/2]} \left[\frac{1}{\tilde{f}_j} - \frac{1}{E(\check{f}_j)} \right] q_j \xrightarrow{p} 0.$$

Proof. The left side of the above expression is

$$\frac{1}{n^{1/2}} \sum_{j=2m+1}^{\lfloor n/2 \rfloor} \left[\frac{1}{f_j} - \frac{1}{E(\check{f}_j)} \right] q_j - \frac{1}{n^{1/2}} \sum_{j=1}^{2m} \frac{1}{E(\check{f}_j)} q_j + \frac{1}{n^{1/2}} \sum_{j=1}^{2m} \frac{1}{f_j} q_j. \quad (\text{A.17})$$

The third term of (A.17) is $o_p(1)$ on taking $a_{j,n} = 1$ for all j in Lemma B.7 since C8 entails $m/n \rightarrow 0$ and thus $F(\lambda_m^{1/2}) - F(0) \rightarrow 0$, $F(\cdot)$ being continuous from the right. Now consider the second term of (A.17). Because $a_{j,n} = f_j (E(\check{f}_j))^{-1} \leq K(m/j)^{2d}$ by Proposition A.1 part (a), it follows from Lemma B.7 that the second term of (A.17) is also $o_p(1)$. Finally the first term of (A.17) is

$$\frac{1}{n^{1/2}} \sum_{j=2m+1}^{\lfloor n^{2/3}/\log m \rfloor} a_{j,n} \frac{1}{f_j} q_j + \frac{1}{n^{1/2}} \sum_{j=1+\lfloor n^{2/3}/\log m \rfloor}^{\lfloor n/2 \rfloor} a_{j,n} \frac{1}{f_j} q_j, \quad (\text{A.18})$$

where

$$a_{j,n} = \frac{E(\check{f}_j) - f_j}{f_j} \frac{1}{f_j^{-1} E(\check{f}_j)}.$$

The first term of (A.18) is $o_p(1)$ as we now show. Since by Proposition A.1 part (b),

$$\sup_{2m < j \leq \lfloor \delta n/2 \rfloor} \left| f_j^{-1} (E(\check{f}_j) - \check{f}_j) \right| = O(m^{-1} \log m), \quad K^{-1} < \sup_{2m < j \leq \lfloor n/2 \rfloor} \left| f_j^{-1} \check{f}_j \right| < K,$$

it follows that $K^{-1} < f_j^{-1} E(\check{f}_j) < K$ as $n \rightarrow \infty$, and, using (A.1) for $2m < j \leq \lfloor n^{2/3}/\log m \rfloor$, that

$$|a_{j,n}| \leq K \left(\frac{|E(\check{f}_j) - \check{f}_j|}{f_j} + \frac{|\check{f}_j - f_j|}{f_j} \right) \leq K \left(\frac{\log m}{m} + \frac{m}{j} \right) \leq K \left(\frac{m}{j} \right)^{2d}$$

since $d < 1/2$ and C8 holds. Now proceed as with the second term of (A.17) to conclude that the first term of (A.18) is $o_p(1)$ by Lemma B.7.

Finally consider the second term of (A.18). By Proposition A.1 parts (b) and (c) and (A.1)

$$\sup_{\lfloor n^{2/3}/\log m \rfloor < j \leq \lfloor n/2 \rfloor} \left| f_j^{-1} \check{f}_j - 1 \right| = O \left(\log m \left(\frac{m}{n^{2/3}} + \frac{1}{m} \right) \right),$$

so that $\sup_{\lfloor n^{2/3}/\log m \rfloor < j \leq \lfloor n/2 \rfloor} |a_{j,n}^2| \leq K((m^2/n^{4/3} + 1/m^2) \log^2 m) \leq K((m^2 \log^2 m)/n^{4/3})$. Thus by Lemma B.6 and C8, the second term of (A.18) is $o_p(1)$ since $d < 1/2$. \square

The next Proposition is highly technical and only used in some steps of Proposition A.6.

Proposition A.5 Assume C1-C8, and let $a_{j,n}$ be an array of positive bounded constants. Then, as $n \rightarrow \infty$, for $k = 1, 2$,

$$\frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]} a_{j,n} \tilde{f}_j^{-k} h_{2j}^k \frac{w_{\varepsilon,j}}{f_j^{1/2}} w_{x,j}^* = o_p(1).$$

Proof. We deal only with the case $k = 2$, the case $k = 1$ being similarly handled. By definition of h_{2j} given in (A.11), the left side of the last displayed equation is

$$\begin{aligned} & \frac{4\pi^2}{\sigma^4 (2m+1)^2 n^{1/2}} \sum_{\ell_1, \ell_2 = -m}^m \sum_{j=2m+1}^{[n/2]} c_{\ell_1} c_{\ell_2} a_{j,n} \frac{f_j^2}{\tilde{f}_j^2} \left(\frac{f_{\ell_1+j}}{f_j} \frac{f_{\ell_2+j}}{f_j} \right) \\ & (I_{\varepsilon\varepsilon, \ell_1+j} - EI_{\varepsilon\varepsilon, \ell_1+j}) (I_{\varepsilon\varepsilon, \ell_2+j} - EI_{\varepsilon\varepsilon, \ell_2+j}) \frac{w_{\varepsilon,j}}{f_j^{1/2}} w_{x,j}^*. \end{aligned}$$

Denote $d_j = (2\pi n)^{1/2} w_{\varepsilon,j}$. Since $f_j^{-1} (f_{\ell_1+j} + f_{\ell_2+j}) < K$ for $j > 2m$, $f_j^{-1} < K$ by C1 and C3, $K^{-1} < f_j^{-1} \tilde{f}_j < K$ by Proposition A.1 parts (b) and (c) and $|a_{j,n}^2| < K$, the variance of the last displayed expression is bounded by

$$\begin{aligned} & \frac{K}{m^4 n^6} \sum_{\ell_i = -m, i=1, \dots, 4}^m \sum_{j_1, j_2 = 2m+1}^{[n/2]} |E(w_{x,j_1}^* w_{x,j_2})| \tag{A.19} \\ & \quad \times |cum(d_{\ell_1+j_1} d_{\ell_1+j_1}^* d_{\ell_2+j_1} d_{\ell_2+j_1}^* d_{j_1}, d_{\ell_3+j_2} d_{\ell_3+j_2}^* d_{\ell_4+j_2} d_{\ell_4+j_2}^* d_{j_2}^*)| \\ & = \frac{K}{m^4 n^6} \sum_v \sum_{\ell_i = -m, i=1, \dots, 4}^m \sum_{j_1, j_2 = 2m+1}^{[n/2]} |E(w_{x,j_1}^* w_{x,j_2})| \prod_{l=1}^q |cum(d_{s_1+s_2}; (s_1, s_2) \in v_l)| \end{aligned}$$

with $s_1 = 0, \ell_1, \ell_2, \ell_3, \ell_4$ and $s_2 = j_1, j_2$, and where the summation in v is over all indecomposable partitions $v = v_1 \cup \dots \cup v_q$, $q = 1, \dots, 5$, of the table

$$\begin{array}{ccccc} d_{\ell_1+j_1} & d_{\ell_1+j_1}^* & d_{\ell_2+j_1} & d_{\ell_2+j_1}^* & d_{j_1} \\ d_{\ell_3+j_2} & d_{\ell_3+j_2}^* & d_{\ell_4+j_2} & d_{\ell_4+j_2}^* & d_{j_2}^* \end{array} \tag{A.20}$$

see Brillinger (1981, p.20 and Theorem 2.6.1).

We shall now show that the right side of (A.19) is $o(1)$. When $q = 1, 2$ or 3 , using Brillinger's (1981) Theorem 4.3.2, in particular his expression (4.3.15), the right side of (A.19) is bounded by

$$\begin{aligned} \frac{K}{m^4 n^3} \sum_{\ell_i = -m, i=1, \dots, 4}^m \sum_{j_1, j_2 = 2m+1}^{[n/2]} |E(w_{x,j_1}^* w_{x,j_2})| & \leq \frac{K}{n^3} \left(\sum_{j=2m+1}^{[n/2]} E(I_{xx,j}) \right)^2 \\ & = o(1) \end{aligned}$$

by the Schwarz inequality and $\sum_{j=1}^{\lfloor n/2 \rfloor} EI_{xx,j} = O(n)$.

Next, when $q = 4$, the right side of (A.19) is bounded by

$$\frac{K}{m^4 n^2} \sum_v \sum_{\ell_i = -m, i=1, \dots, 4}^m \sum_{j_1, j_2 = 2m+1}^{\lfloor n/2 \rfloor} |E(w_{x,j_1}^* w_{x,j_2})| \prod_{l=1}^4 \mathcal{I} \left(\sum_{(s_1, s_2) \in v_l} s_1 + s_2 = 0 \right), \quad (\text{A.21})$$

since by Brillinger (1981, (4.3.15) and Exercise 4.8.12)

$$\text{cum}(d_{j_1}, \dots, d_{j_\ell}) = n \mathcal{I} \left(\sum_{s=1}^{\ell} j_s = 0, \text{mod}(n) \right).$$

Consider a typical indecomposable partition in v , for instance

$$\{((1, 1), (1, 2), (2, 1)) \cup ((1, 3), (2, 1), (2, 2)) \cup ((1, 4), (2, 5)) \cup ((1, 5), (2, 4))\},$$

where (i, j) is the (i, j) -th element in the table (A.20). Then (A.21) differs from zero only if $j_2 + \ell_3 = 0, j_1 + \ell_2 = 0, j_2 + j_1 = \ell_2$ and $j_1 - j_4 = \ell_4$, which implies, for example, that $\ell_3 - \ell_2 = \ell_4$. Thus (A.21) is bounded by

$$\frac{K}{mn^2} \sum_{j_1, j_2 = 2m+1}^{\lfloor n/2 \rfloor} |E(w_{x,j_1}^* w_{x,j_2})| \leq \frac{K}{m} \left(\frac{1}{n} \sum_{j=2m+1}^{\lfloor n/2 \rfloor} E(I_{xx,j}) \right)^2 = O(m^{-1})$$

by the Schwarz inequality and $\sum_{j=1}^{\lfloor n/2 \rfloor} EI_{xx,j} = O(n)$.

Finally, when $q = 5$, the right side of (A.19) is bounded by

$$\frac{K}{m^4 n} \sum_v \sum_{\ell_i = -m, i=1, \dots, 4}^m \sum_{j_1, j_2 = 2m+1}^{\lfloor n/2 \rfloor} |E(w_{x,j_1}^* w_{x,j_2})| \prod_{l=1}^5 \mathcal{I} \left(\sum_{(s_1, s_2) \in v_l} s_1 + s_2 = 0 \right) \quad (\text{A.22})$$

by Brillinger (1981, (4.3.15) and Exercise 4.8.12). Two typical indecomposable partitions in v are

$$v_1 = \{((1, 1), (2, 1)) \cup ((1, 2), (2, 2)) \cup ((1, 3), (2, 3)) \cup ((1, 4), (2, 4)) \cup ((1, 5), (2, 5))\}$$

or

$$v_2 = \{((1, 1), (2, 1)) \cup ((1, 2), (2, 2)) \cup ((1, 3), (2, 3)) \cup ((1, 4), (2, 5)) \cup ((1, 5), (2, 4))\}.$$

For the partition v_1 , (A.22) differs from zero only if $j_1 = j_2$ and thus $\ell_1 = \ell_3$ and $\ell_2 = \ell_4$, whereas for the partition v_2 , (A.22) differs from zero only if $\ell_1 = \ell_2 =$

$\ell_3 = \ell_4$. Thus, the expectation of (A.22) is bounded by

$$\begin{aligned} & \frac{K}{m^2 n} \sum_{j=2m+1}^{[n/2]} E(I_{xx,j}) + \frac{K}{m^3 n} \sum_{j_1, j_2=2m+1}^{[n/2]} |E(w_{x,j_1}^* w_{x,j_2})| \\ &= \frac{K}{m^3 n} \left(\sum_{j=2m+1}^{[n/2]} E(I_{xx,j}) \right)^2 + O(m^{-2}) \\ &= O(nm^{-3} + m^{-2}) = o(1) \end{aligned}$$

using the Schwarz inequality, $\sum_{j=1}^{[n/2]} E(I_{xx,j}) = O(n)$ and C8. \square

Proposition A.6 *Assuming C1-C5 and C7-C8, as $n \rightarrow \infty$,*

$$(a) \quad \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \left(\frac{1}{E(\check{f}_j)} - \frac{1}{\check{f}_j} \right) q_j \xrightarrow{p} 0 \quad (\text{A.23})$$

$$(b) \quad \frac{1}{n^{1/2}} \sum_{j=1}^{2m} \left(\frac{1}{E(\check{f}_j)} - \frac{1}{\check{f}_j} \right) q_j \xrightarrow{p} 0. \quad (\text{A.24})$$

Proof. We begin with (a). Denote $\hat{e}_j = \check{f}_j - E(\check{f}_j)$. The left side of (A.23) is

$$\begin{aligned} \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \frac{\hat{e}_j q_j}{\check{f}_j E(\check{f}_j)} &= \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \frac{\hat{e}_j}{\check{f}_j} \frac{(f_j^{-1/2} q_j - \tilde{q}_j)}{f_j^{-1/2} E(\check{f}_j)} \\ &+ \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \frac{(\hat{e}_j - h_{2j})}{\check{f}_j} \frac{\tilde{q}_j}{f_j^{-1/2} E(\check{f}_j)} \quad (\text{A.25}) \\ &+ \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \frac{h_{2j}}{\check{f}_j} \frac{\tilde{q}_j}{f_j^{-1/2} E(\check{f}_j)}, \end{aligned}$$

where h_{2j} is as defined in (A.11) and $\tilde{q}_j = (2\pi)^{1/2} \sigma^{-1} w_{\varepsilon,j} w_{x,j}^*$.

By the Schwarz inequality the squared modulus of the first term on the right of (A.25) is bounded by

$$K \left(\frac{1}{n} \sum_{j=2m+1}^{[n/2]-1} j^{-1} \left| \frac{\tilde{f}_j^{-1} \hat{e}_j}{\tilde{f}_j^{-1} \check{f}_j} \right|^2 \right) \sum_{j=2m+1}^{[n/2]-1} j \left| \frac{f_j^{-1/2} q_j - \tilde{q}_j}{f_j^{-1} E^2(\check{f}_j)} \right|^2.$$

The factor in parentheses is $O_p(m^{-1}n^{-1} \log n)$ by Proposition A.2 and because by (4.10) and Proposition A.1 parts (b, c) $g_{m, [n/2]} = \inf_{j=2m+1, \dots, [n/2]-1} \left| \tilde{f}_j^{-1} \check{f}_j \right| \geq K^{-1}$. Next, $\left| f_j^{-1/2} q_j - \tilde{q}_j \right|^2 = \left| f_j^{-1/2} w_{u,j} - (2\pi)^{1/2} \sigma^{-1} w_{\varepsilon,j} \right|^2 I_{xx,j}$, which by Robinson's (1995a) Theorem 2 and C5, has first moment bounded by $Kj^{-1} E(I_{xx,j}) \log j$. Moreover, by Proposition A.1 parts (b) and (c), $K^{-1} \leq f_j^{-2} E^2(\check{f}_j) \leq K$, so the last sum of the last displayed expression has expectation bounded by $K \log n \sum_{j=2m+1}^{[n/2]} E(I_{xx,j}) = O(n \log n)$. Thus, the first term on the right of (A.25) is $O_p(m^{-1/2} \log n) = o_p(1)$ by C8.

Next, by the Schwarz inequality and $K^{-1} < f_j^{-1} E(\check{f}_j) < K$, the squared modulus of the second term on the right of (A.25) is bounded by

$$\left(K g_{m, [n/2]}^{-2} \right) \left(\sum_{j=2m+1}^{[n/2]-1} \left(\tilde{f}_j^{-1} (\hat{e}_j - h_{2j}) \right)^2 \right) \left(\frac{1}{n} \sum_{j=2m+1}^{[n/2]-1} \left| \frac{\tilde{q}_j}{f_j^{1/2}} \right|^2 \right) = o_p(1)$$

because the second factor on the left is $O_p(m^{-1} \log^2 n)$ proceeding as in the proof of Proposition A.3 part (b), c.f. (A.12), and observing that $\hat{e}_j - h_{2j} = h_{1j}$, with h_{1j} defined in (A.10). On the other hand, the last factor is $O_p(1)$, since its expectation is bounded by $Kn^{-1} \sum_j E I_{xx,j} = O(1)$ since $f_j^{-1} \leq K$. So the second term on the right of (A.25) is also $o_p(1)$.

To complete the proof of part (a), it remains to examine the third term on the right of (A.25). Using the identity

$$\frac{1}{a} = \frac{1}{b} - \frac{a-b}{b^2} + \frac{(b-a)^2}{b^2 a},$$

that term is

$$\begin{aligned} & \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \left(\tilde{f}_j^{-1} E(\check{f}_j) \right)^{-1} \frac{\tilde{f}_j^{-1} h_{2j} \tilde{q}_j}{f_j^{-1/2} E(\check{f}_j)} \\ & - \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \left(\tilde{f}_j^{-1} E(\check{f}_j) \right)^{-2} \frac{\hat{e}_j}{\tilde{f}_j} \frac{\tilde{f}_j^{-1} h_{2j} \tilde{q}_j}{f_j^{-1/2} E(\check{f}_j)} \\ & + \frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \left(\tilde{f}_j^{-1} E(\check{f}_j) \right)^{-2} \frac{\tilde{f}_j^{-1} h_{2j}}{\tilde{f}_j^{-1} \check{f}_j} \left(\frac{\hat{e}_j}{\tilde{f}_j} \right)^2 \frac{\tilde{q}_j}{f_j^{-1/2} E(\check{f}_j)}. \end{aligned} \tag{A.26}$$

Because $K^{-1} < \tilde{f}_j^{-1} E(\check{f}_j) < K$ and $K^{-1} < f_j^{-1} E(\check{f}_j) < K$ by Proposition A.1 parts (b) and (c), the first term of (A.26) is $o_p(1)$ by Proposition A.5 with $k = 1$

there. By the Schwarz inequality the squared modulus of the third term of (A.26) is bounded by

$$\left(K g_{m, [n/2]}^{-2}\right) \left(\frac{1}{n} \sum_{j=2m+1}^{[n/2]-1} \left(\tilde{f}_j^{-1} \hat{e}_j\right)^4\right) \sum_{j=2m+1}^{[n/2]-1} \frac{\left|\tilde{f}_j^{-1} h_{2j} \tilde{q}_j\right|^2}{f_j^{-1} E^2(\check{f}_j)} = O_p(nm^{-3}) \quad (\text{A.27})$$

as we now show. The last sum on the left of (A.27) has expectation bounded by

$$K \sum_{j=2m+1}^{[n/2]-1} E \left| \tilde{f}_j^{-1} h_{2j} w_{\varepsilon, j} \right|^2 E(I_{xx, j})$$

by C5 and Proposition A.1 parts (b) and (c) $K^{-1} < f_j^{-1} E^2(\check{f}_j)$. But the last displayed expression is $O(nm^{-1})$ since by the Schwarz inequality and Brillinger's (1981) Theorem 7.7.4,

$$E \left| \tilde{f}_j^{-1} h_{2j} w_{\varepsilon, j} \right|^2 \leq \left(E \left| \tilde{f}_j^{-1} h_{2j} \right|^4 \right)^{1/2} \left(E |w_{\varepsilon, j}|^4 \right)^{1/2} = O(m^{-1})$$

and $\sum_{j=2m+1}^{[n/2]-1} E(I_{xx, j}) = O(n)$. The second bracketed factor in (A.27) is $O_p(m^{-2})$, as we now show. We have

$$\left\{ \tilde{f}_j^{-1} \hat{e}_j \right\}^4 \leq K \left[\left(\tilde{f}_j^{-1} h_{1j} \right)^4 + \left(\tilde{f}_j^{-1} h_{2j} \right)^4 \right]. \quad (\text{A.28})$$

Proceeding as with the proof of (A.13), and that $\sum_{j=1}^n a_j^4 \leq n^{1/3} \left(\sum_{j=1}^n a_j^6 \right)^{2/3}$ by Hölder's inequality, the contribution from the second term on the right of (A.28) to the second bracketed factor in (A.27) is $O_p(m^{-2})$, whereas the contribution from the first term on the right of (A.28) is $O_p(m^{-2} n^{-1} \log^4 n)$ by (A.12) and that $\sum_j a_j^4 \leq \sup_j a_j^2 \sum_j a_j^2 \leq \left(\sum_j a_j^2 \right)^2$.

Because $\sum_j E \left(\tilde{f}_j^{-1} h_{1j} \right)^2 = O(m^{-1} \log^2 n)$ by (A.12), and using $\hat{e}_j = h_{1j} + h_{2j}$, the second term of (A.26) is

$$\frac{1}{n^{1/2}} \sum_{j=2m+1}^{[n/2]-1} \left(\tilde{f}_j^{-1} E(\check{f}_j) \right)^{-1} \frac{\tilde{f}_j^{-2} h_{2j}^2 \tilde{q}_j}{f_j^{-1/2} E(\check{f}_j)} + O_p(m^{-1} \log^2 n).$$

But because $K^{-1} < \tilde{f}_j^{-1} E(\check{f}_j) < K$ and $K^{-1} < f_j^{-1} E(\check{f}_j) < K$ by Proposition A.1 parts (b) and (c), the first term of the last displayed expression is $o_p(1)$ by Proposition A.5 with $k = 2$ there, which concludes the proof of part (a).

Regarding part (b), the left side of (A.24) is

$$\frac{1}{n^{1/2}} \sum_{j=1}^{2m} \frac{\widehat{e}_j}{\check{f}_j} \frac{\left(f_j^{-1/2} q_j - \widetilde{q}_j\right)}{f_j^{-1/2} E(\check{f}_j)} + \frac{1}{n^{1/2}} \sum_{j=1}^{2m} \frac{\widehat{e}_j}{\check{f}_j} \frac{\widetilde{q}_j}{f_j^{-1/2} E(\check{f}_j)}. \quad (\text{A.29})$$

By the Schwarz inequality the expectation of the squared modulus of the second term of (A.29) is bounded by

$$K g_{1,2m}^{-2} \sum_{j=1}^{2m} E \left| \widetilde{f}_j^{-1} \widehat{e}_j \right|^2 \frac{1}{n} \sum_{j=1}^{2m} E \left| \frac{\widetilde{q}_j}{f_j^{-1/2} E(\check{f}_j)} \right|^2,$$

where $g_{1,2m} = \inf_{j=1, \dots, 2m} \left| \widetilde{f}_j^{-1} \check{f}_j \right|$. By (4.10) and Proposition A.1 part (a), $g_{1,2m} \geq K^{-1}$, whereas by Proposition A.2 part (b), the second factor of the last displayed expression is $O\left((m^{4d-1} \mathcal{I}(d > 1/4) + \log m \mathcal{I}(d = 1/4) + \mathcal{I}(d < 1/4))\right)$. Finally, by C1 and Proposition A.1 part (a) $\left| f_j^{1/2} E^{-1}(\check{f}_j) \right| \leq K \lambda_m^d (m/j)^d \leq K \lambda_m^d (m/j)^{2d}$, the third factor of the last displayed expression is

$$O\left(\lambda_m^{2d} \left(\frac{m^2}{n}\right)^{2d} \left(F(\lambda_m^{1/2}) - F(0)\right)\right),$$

proceeding as in the proof of the first term on the right of (B.10) with $f_j^{-1} w_{u,j}$ replaced by $w_{\varepsilon,j}$. Thus, the second term of (A.29) is $o_p(1)$ since C8 entails $m/n \rightarrow 0$ and thus $F(\lambda_m^{1/2}) - F(0) \rightarrow 0$, since $F(\cdot)$ is continuous from the right.

Next, consider the first term of (A.29), which, by the Schwarz inequality, has squared modulus bounded by

$$\left(K g_{1,2m}^{-2}\right) \left(\sum_{j=1}^{2m} \frac{1}{j} \left|\widetilde{f}_j^{-1} \widehat{e}_j\right|^2\right) \left(\frac{1}{n} \sum_{j=1}^{2m} \frac{j \left|f_j^{-1/2} q_j - \widetilde{q}_j\right|^2}{f_j^{-1} E^2(\check{f}_j)}\right) \quad (\text{A.30})$$

which is $o_p(1)$ as we now show. The first factor of (A.30) is bounded by (4.10) and Proposition A.1 parts (a), whereas the second factor is

$$O_p\left(\log m \left(m^{4d-2} \mathcal{I}(d > 1/4) + m^{-1} \log m \mathcal{I}(d = 1/4) + m^{-1} \mathcal{I}(d < 1/4)\right)\right)$$

by Proposition A.2 part (b) and $\sum_{j=1}^m j^{-1} \leq K \log m$. Finally the last factor of (A.30) is $O_p(\lambda_m^{4d} n^{2d} \log n)$ since $E(\check{f}_j) \geq K^{-1} \lambda_m^{-2d}$ by Proposition A.1 part (a) and

$$\frac{1}{n} \sum_{j=1}^{2m} \frac{j \left|f_j^{-1/2} q_j - \widetilde{q}_j\right|^2}{f_j^{-1} E(\check{f}_j)} = O_p(n^{2d} \log n),$$

proceeding as in the proof of the first term on the right of (A.25). Thus (A.30) = $o_p(1)$ by C8 and $d < 1/2$. This completes the proof of part (b) and the proposition. \square

APPENDIX B

Lemma B.1 *Assume C1. For $2m < j \leq [n/2]$ and $\psi \geq 0$, as $n \rightarrow \infty$*

$$\frac{1}{2m+1} \sum'_{p(j)} (p+j)^{-\psi} c_p f_{p+j} = O(f_j m^{-\psi}). \quad (\text{B.1})$$

Proof. The left side of (B.1) times $f_j^{-1} m^\psi$ is bounded by

$$K \frac{m^\psi}{2m+1} \sum'_{p(j)} (p+j)^{-\psi} \frac{f_{p+j}}{f_j} = K \frac{m^\psi}{2m+1} \sum'_{p(j)} (p+j)^{-\psi} \left(1 + O\left(\left|\frac{p}{j}\right|\right) \right),$$

since by C1, $f_j^{-1} f_{p+j} = 1 + O(|p/j|)$. The result follows since $m^\psi (p+j)^{-\psi} \leq K$. \square

Lemma B.2 *Assume C1, and let $v = 2d + \psi$. If $1 \leq j \leq 2m$, as $n \rightarrow \infty$*

$$\frac{1}{2m+1} \sum'_{p(j)} |p+j|^{-\psi} c_p f_{|p+j|} = O\left(m^{-1} n^{2d} (\mathcal{I}(v > 1) + \log m \mathcal{I}(v = 1))\right). \quad (\text{B.2})$$

Proof. Since by C1, $f_{|p+j|} = O(\lambda_{|p+j|}^{-2d})$, the left side of (B.2) is bounded by

$$\frac{K}{2m+1} \sum'_{p(j)} |p+j|^{-\psi} \lambda_{|p+j|}^{-2d} \leq K \frac{n^{2d}}{(2m+1)} \sum'_{p(j)} |p+j|^{-v}.$$

(B.2) now follows as the sum on the right is $O(1) \mathcal{I}(v > 1) + O(\log m) \mathcal{I}(v = 1)$. \square

Lemma B.3 *Let $1 \leq |r| \leq n-1$. Then, as $n \rightarrow \infty$*

$$\left| \frac{1}{2m+1} \sum_{j=-m}^m c_j e^{ir\lambda_j} \right| = O\left(\left(\frac{n}{rm}\right)^2\right). \quad (\text{B.3})$$

Proof. Since $c_m = 0$ by C7, the left side of (B.3) is, by Abel summation by parts,

$$\begin{aligned} & \frac{1}{2m+1} \left| c_0 + 2 \sum_{j=1}^{m-1} (c_j - c_{j+1}) \sum_{\ell=1}^j \cos(r\lambda_\ell) \right| \\ &= \frac{1}{2m+1} \left| c_0 + 2 \sum_{j=1}^{m-1} (c_j - c_{j+1}) \left(\frac{\sin((j+1/2)\lambda_r)}{2\sin(\lambda_r/2)} - \frac{1}{2} \right) \right| \\ &= \frac{1}{2m+1} \left| (c_0 - c_1) + 2 \sum_{j=1}^{m-1} (c_j - c_{j+1}) \frac{\sin((j+1/2)\lambda_r)}{2\sin(\lambda_r/2)} \right|, \end{aligned} \quad (\text{B.4})$$

see Zygmund (1990, p.2). Because $c(u)$ is twice continuously differentiable on $(0, 1)$ by C7, $c_j - c_{j+1} = m^{-1}d_j + O(m^{-2})$, where $d_j = d(j/m)$ and $d(u) = (\partial/\partial u)c(u)$. Thus, by Brillinger (1981, p.15), the right side of (B.4) is bounded by

$$\frac{K}{m|\sin(\lambda_r/2)|} \left| \frac{1}{m} + \int_0^1 d(u) \sin\left(\frac{2\pi r m}{n}u\right) du \right| = \frac{K}{m|\sin(\lambda_r/2)|} \left(\frac{1}{m} + \frac{n}{rm} \right)$$

using integration by parts. The result now follows since $(\frac{\pi n}{r} \sin(\lambda_r/2))^{-1} = O(1)$ for $|r| \leq [n/2]$. \square

Lemma B.4 *Let $C_j(r)$ be as defined in (A.16). For $1 < j \leq [n/2]$,*

$$\sum_{r=1-n}^{n-1} |C_j(r)| = O(m^{-1}n). \quad (\text{B.5})$$

Proof. By definition of $C_j(r)$ and triangle inequality, the left side of (B.5) is

$$\begin{aligned} & \left\{ \sum_{|r| \leq [n/m]} + \sum_{[n/m] < |r| \leq n-1} \right\} \left(\left| \frac{1}{(2m+1)} \sum_{\ell=-m}^m c_\ell e^{ir\lambda_\ell} \right| + \frac{c_j}{(2m+1)} \mathcal{I}(|j| \leq m) \right) \\ &= O(m^{-1}n) + K \sum_{[n/m] < |r| \leq n-1} \left(\frac{n}{rm}\right)^2 \end{aligned}$$

by Lemma B.3. The conclusion now follows because $\sum_{[n/m] < |r| \leq [n/2]} |r|^{-2} = O(n^{-1}m)$. \square

Lemma B.5 Assume C1-C3 and let $d \geq 1/4$, then

$$E\tilde{\gamma}_r^2 = O(n^{4d-2}\mathcal{I}(d > 1/4) + n^{-1} \log n \mathcal{I}(d = 1/4)).$$

Proof. Using formulae in Anderson (1970, p. 452), $E\tilde{\gamma}_r^2$ is

$$\frac{1}{n} \sum_{p=-(n-r-1)}^{n-r-1} \left(1 - \frac{|p|}{n-r}\right) [\gamma_p^2 + \gamma_{p+r}\gamma_{p-r} + \kappa_u(r, -p, r-p)],$$

where $\kappa_u(v_1, v_2, v_3)$ is the cumulant of $u_t, u_{t+v_1}, u_{t+v_2}, u_{t+v_3}$. First, by C1 and C3 $\gamma_p = O(p^{2d-1})$ as $p \rightarrow \infty$ by Lemma 4 of Fox and Taqqu (1986). Thus, the first two terms of the last displayed expression are $O(n^{4d-2}\mathcal{I}(d > 1/4) + n^{-1} \log n \mathcal{I}(d = 1/4))$. Using Brillinger's (1981) formulae as in Proposition A.2, the last term is bounded by

$$\begin{aligned} & \frac{K}{n} \int_{[-\pi, \pi]^3} D(\lambda + \mu) \tau(\lambda) \tau(\mu) \tau(\zeta) \tau(-\lambda - \mu - \zeta) d\lambda d\mu d\zeta \\ & \leq \frac{K}{n} \int_{[-\pi, \pi]^2} D(\lambda + \mu) \tau(\lambda) \tau(\mu) d\lambda d\mu \end{aligned} \quad (\text{B.6})$$

using $\int_{[-\pi, \pi]} |\tau(-\lambda - \mu - \zeta) \tau(\zeta)| d\zeta < K$ and integrability of $|\tau(\zeta)|^2$. But the modulus of the right side of (B.6) is bounded by

$$\frac{K}{n} \int_0^\pi \int_{\lambda < \mu} |D(\lambda + \mu)| (|\tau(\lambda)|^2 + |\tau(\mu)|^2) d\lambda d\mu \leq \frac{K \log n}{n}$$

since $\int_0^\pi |D(\vartheta)| d\vartheta = O(\log n)$ by Zygmund (1990, p.67), $|\tau(\mu)| \leq K |\tau(\lambda)|$ by C1 and C3 and integrability of $|\tau(\lambda)|^2$. \square

For the remaining lemmas, let q_j and $w_{x,j}$ denote typical components of the $p \times 1$ vectors $I_{xu,j}$ and $w_x(\lambda_j)$ respectively, while $F(\lambda)$ and $\Gamma(s)$ will here denote the spectral distribution and autocovariance function of the corresponding element of x_t .

Lemma B.6 Assume C1-C5, and let $a_{j,n}$ be a triangular array of constants. For n sufficiently large

$$\max_{\ell \leq r < s \leq \lfloor n/2 \rfloor} E \left| \sum_{j=r}^s a_{j,n} \frac{q_j}{f_j} \right|^2 \leq K \left(\frac{n}{\ell}\right)^{2d} n \max_{\ell \leq j \leq \lfloor n/2 \rfloor} a_{j,n}^2, \quad (\text{B.7})$$

where $\ell^{-1} + \ell/n \rightarrow 0$.

Proof. Suppressing reference to n in $a_{j,n}$,

$$E \left| \sum_{j=r}^s a_j \frac{I_{xu,j}}{f_j} \right|^2 = \frac{1}{n} \int_{-\pi}^{\pi} f(\lambda) g_{r,s}(\lambda) d\lambda, \quad (\text{B.8})$$

where $g_{q,h}(\lambda) = E \left| \sum_{j=q}^h a_j f_j^{-1} w_{x,j} D(\lambda_j - \lambda) \right|^2$.

Writing $\vartheta_n = \pi\ell/n$, the contribution of

$$\frac{1}{n} \int_{\vartheta_n < |\lambda| \leq \pi} f(\lambda) g_{r,s}(\lambda) d\lambda$$

to the right side of (B.8) is bounded by

$$\begin{aligned} \frac{K}{n} f(\vartheta_n) \int_{-\pi}^{\pi} g_{r,s}(\lambda) d\lambda &\leq K f(\vartheta_n) \sum_{j=r}^s a_j^2 f_j^{-2} E(I_{xx,j}) \\ &\leq K f_{\ell} \left(\max_{\ell \leq j \leq [n/2]} a_j^2 \right) \sum_{j=1}^n E(I_{xx,j}) \leq K \left(\frac{n}{\ell} \right)^{2d} n \max_{\ell \leq j \leq [n/2]} a_j^2, \end{aligned}$$

because $\sum_{j=1}^n E(I_{xx,j}) = (2\pi)^{-1} \sum_{j=1}^n E(x_t x_t') = O(n)$ and $f_j^{-1} \leq K f_s^{-1}$ by C1.

On the other hand, by elementary inequalities, the contribution of $n^{-1} \int_{-\vartheta_n}^{\vartheta_n} f(\lambda) g_{r,s}(\lambda) d\lambda$ to the right of (B.8) is bounded by

$$\begin{aligned} &\int_{-\vartheta_n}^{\vartheta_n} f(\lambda) \left(\frac{1}{n} \sum_{j=r}^s a_j^2 f_j^{-1} E I_{xx,j} \right) \left(\sum_{j=r}^s f_j^{-1} |D(\lambda_j - \lambda)|^2 \right) d\lambda \quad (\text{B.9}) \\ &\leq \max_{\ell \leq j \leq [n/2]} a_j^2 \left(\frac{1}{n} \sum_{j=1}^{[n/2]} f_j^{-1} E I_{xx,j} \right) \left(\sup_{|\lambda| < \vartheta_n} \sum_{j=\ell}^{[n/2]} f_j^{-1} |D(\lambda_j - \lambda)|^2 \right) \left(\int_{-\vartheta_n}^{\vartheta_n} f(\lambda) d\lambda \right). \end{aligned}$$

The second factor on the right of (B.9) is bounded by K , as follows from the proofs of Propositions 3 and 7 of Robinson and Hidalgo (1997). Noting that

$$|D(\lambda)| \leq K |\lambda|^{-1}, \quad 0 < |\lambda| < \pi$$

and that $\lambda_{\ell} - \pi\ell/n = \lambda_{\ell}/2$, the third factor on the right of (B.9) is bounded by

$$K n^{2-2d} \sum_{j \geq [\ell/2]+1} j^{2d-2} \leq K n (n/\ell)^{1-2d}.$$

Finally the last factor on the right of (B.9) is ϑ_n^{1-2d} . Thus, it follows that (B.9) is bounded by $K n \max_{\ell \leq j \leq [n/2]} a_j^2$ so that the left side of (B.7) is bounded by $K (n/\ell)^{2d} n \max_{\ell \leq j \leq [n/2]} a_j^2$. \square

Lemma B.7 Assume C1-C5, and let $a_{j,n}$ be an array of positive constants bounded by $K(m/j)^{2d}$ and ℓ be as in Lemma B.6. Then, for n sufficiently large

$$\max_{1 \leq r < s \leq \ell} E \left| \sum_{j=r}^s a_{j,n} \frac{q_j}{f_j} \right|^2 \leq Kn \left(\frac{m^2}{n} \right)^{2d} \left(\ell^{-1} \log^2 \ell + \left(F(\lambda_\ell^{1/2}) - F(0) \right) \right).$$

Proof. Suppressing reference to n in $a_{j,n}$,

$$\begin{aligned} E \left| \sum_{j=r}^s a_j \frac{q_j}{f_j} \right|^2 &= \frac{1}{n} \int_{-\pi}^{\pi} E \left| \sum_{j=r}^s a_j \frac{w_{u,j}}{f_j} D(\lambda_j - \lambda) \right|^2 dF(\lambda) \\ &= \frac{1}{n} \sum_{j=r}^s \int_{-\pi}^{\pi} E \left| a_j \frac{w_{u,j}}{f_j} D(\lambda_j - \lambda) \right|^2 dF(\lambda) \\ &\quad + \frac{2}{n} \int_{-\pi}^{\pi} E \sum_{\substack{j=r, k=r \\ j < k}}^s a_j a_k \frac{w_{u,j}}{f_j} \frac{w_{u,-k}}{f_k} D(\lambda_j - \lambda) D(\lambda - \lambda_k) dF(\lambda). \end{aligned} \quad (\text{B.10})$$

Since $E \left| f_j^{-1/2} w_{u,j} \right|^2 \leq K$ by Robinson's (1995a) Theorem 1 and Theorem 2 part (a), the first term on the right of (B.10) is bounded by

$$\begin{aligned} &K \sum_{j=r}^s \left(\frac{m}{j} \right)^{4d} \left(\frac{j}{n} \right)^{2d} \int G(\lambda_j - \lambda) dF(\lambda) \\ &= K \left(\frac{m^2}{n} \right)^{2d} \sum_{j=r}^s \frac{1}{j^{2d}} \left\{ \int_{|\lambda| \leq 2\lambda_j^{1/2}} + \int_{|\lambda| > 2\lambda_j^{1/2}} \right\} G(\lambda_j - \lambda) dF(\lambda) \end{aligned} \quad (\text{B.11})$$

where we use the fact that $K^{-1} < f_j \lambda_j^{2d} < K$ by C1 and define

$G(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^n e^{it\lambda} \right|^2$, Fejér's kernel. Because, for $0 < |\lambda| \leq \pi$, $G(\lambda) \leq K(n\lambda^2)^{-1}$, see Zygmund (1990, p. 88), and $|\lambda| > 2\lambda_j^{1/2}$ implies $|\lambda - \lambda_j| > \lambda_j^{1/2}/2$, the contribution from the second term in braces on the right of (B.11) is bounded by

$$K \left(\frac{m^2}{n} \right)^{2d} \sum_{j=r}^s \frac{1}{j^{1+2d}} \int_{|\lambda| > 2\lambda_j^{1/2}} dF(\lambda) \leq K(\Gamma(0)) \left(\left(\frac{m^2}{n} \right)^{2d} \mathcal{I}(d > 0) + \log s \mathcal{I}(d = 0) \right),$$

whereas since $\sum_{j=r}^s G(\lambda_j - \lambda) \leq \sum_{j=1}^n G(\lambda_j - \lambda) = Kn$ the contribution from the first term in braces on the right of (B.11) is bounded by

$$K \left(\frac{m^2}{n} \right)^{2d} n \left(F(\lambda_\ell^{1/2}) - F(0) \right).$$

Next the second term on the right of (B.10) is bounded in absolute value by

$$\begin{aligned} & \left(\frac{m^2}{n}\right)^{2d} \frac{K}{n} \sum_{\substack{j=r, k=r \\ j < k}}^s \frac{1}{j^d k^d} \left| E \frac{w_{u,j} w_{u,-k}}{f_j^{1/2} f_k^{1/2}} \right| \times \left\{ \int_{|\lambda| \leq \frac{1}{2} \lambda_j} + \int_{\frac{1}{2} \lambda_j < |\lambda| \leq \frac{1}{2} \lambda_{j+k}} \right. \\ & \left. + \int_{\frac{1}{2} \lambda_{j+k} < |\lambda| \leq \lambda_{k+\ell}} + \int_{\lambda_{k+\ell} < |\lambda|} \right\} |D(\lambda_j - \lambda) D(\lambda - \lambda_k)| dF(\lambda), \quad (\text{B.12}) \end{aligned}$$

since $K^{-1} < f_j \lambda_j^{2d} < K$ by C1. Take $d > 0$. By Robinson's (1995b) Theorem 2 part (c), the contribution from the first and fourth integrals is bounded by

$$\begin{aligned} & K \left(\frac{m^2}{n}\right)^{2d} n \sum_{\substack{j=r, k=r \\ j < k}}^s \frac{1}{j^{1+d} k^d} \left(\frac{\log k}{j(2k-j)} \int_{|\lambda| \leq \frac{1}{2} \lambda_j} dF(\lambda) + \frac{\log k}{\ell(\ell+k-j)} \right) \\ & \leq K \left(\frac{m^2}{n}\right)^{2d} n \left(\left(F(\lambda_\ell^{1/2}) - F(0) \right) + \ell^{-1} \log^2 \ell \right), \end{aligned}$$

whereas the contribution from the second and third integrals is bounded by

$$\begin{aligned} & K \left(\frac{m^2}{n}\right)^{2d} n (F(\lambda_\ell) - F(0)) \sum_{\substack{j=r, k=r \\ j < k}}^s \frac{1}{j^{1+d} k^d} (k-j)^{-1} \log(k) \\ & \leq K \left(\frac{m^2}{n}\right)^{2d} n \left(F(\lambda_\ell^{1/2}) - F(0) \right). \end{aligned}$$

Now take $d = 0$. Since $E(w_{u,j} w_{u,-k}) = O(n^{-1} \log n)$ by C3 and Brillinger's (1981) Theorem 4.3.2, it is straightforward to observe that (B.12) is

$$K s \log^2 n (F(\lambda_\ell) - F(0)) \leq K n \left(F(\lambda_\ell^{1/2}) - F(0) \right). \quad \square$$

References

- [1] ADENSTEDT, R.K. (1974): "On Large-Sample Estimation of the Mean of a Stationary Random Sequence," *Annals of Statistics*, 2, 1095-1107.
- [2] ANDERSON, T.W. (1970): *The Statistical Analysis of Time Series*. New York: John Wiley.
- [3] ANDREWS, D.W.K. (1991): "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.
- [4] BRILLINGER, D.R. (1981): *Time Series, Data Analysis and Theory*. San Francisco: Holden Day.
- [5] DAHLHAUS, R. (1995): "Efficient Location and Regression Estimation for Long Range Dependent Regression Models," *Annals of Statistics*, 23, 1029-1047.
- [6] DAVIES, R.B. AND D.S. HARTE (1987): "Tests for Hurst Effect," *Biometrika*, 74, 95-101.
- [7] DOOB, J.L. (1953): *Stochastic Processes*. New York: John Wiley.
- [8] FOX, R. AND M.S. TAQQU (1986): "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series," *Annals of Statistics*, 14, 517-532.
- [9] GRENANDER, U. (1954): "On the Estimation of Regression Coefficients in the case of an Autocorrelated Disturbance," *Annals of Mathematical Statistics*, 25, 252-272.
- [10] HANNAN, E.J. (1963): "Regression for Time Series". In M. Rosenblatt (ed.), *Time Series Analysis*, 17-37. New York: John Wiley.
- [11] HANNAN, E.J. (1965): "The Estimation of Relations Involving Distributed Lags," *Econometrica*, 33, 206-224.
- [12] HANNAN, E.J. (1970): *Multiple Time Series*. New York: John Wiley.
- [13] HANNAN, E.J. AND R.D. TERRELL (1972): "Time Series Regression with Linear Constraints," *International Economic Review*, 13, 189-200.

- [14] HANNAN, E.J. AND R.D. TERRELL (1973): "Multiple Equation Systems with Stationary Errors," *Econometrica*, 41, 299-320.
- [15] HOSOYA, Y. (1996): "The Quasi-Likelihood Approach to Statistical Inference in Multiple Time Series with Long-range Dependence," *Journal of Econometrics*, 73, 217-236.
- [16] PHILLIPS, P.C.B. (1991): "Spectral Regression for Cointegrated Time Series," *Nonparametric and Semiparametric Methods in Econometrics and Statistics* (W.A. Barnett, J. Powell and G. Tauchen, eds.), Cambridge: Cambridge University Press, 413-435.
- [17] ROBINSON, P.M. (1976): "Instrumental Variables Estimation of Differential Equations," *Econometrica*, 44, 765-776.
- [18] ROBINSON, P.M. (1988): "The Stochastic Difference between Econometric Statistics," *Econometrica*, 56, 531-548.
- [19] ROBINSON, P.M. (1991): "Automatic Frequency Domain Inference on Semiparametric and Nonparametric Models," *Econometrica*, 59, 1329-1363.
- [20] ROBINSON, P.M. (1994a): "Time Series with Strong Dependence," In C.A. Sims, ed., *Advances in Econometrics: Sixth World Congress*, Vol.1, 47-95. Cambridge: Cambridge University Press.
- [21] ROBINSON, P.M. (1994b): "Semiparametric Analysis of Long-Memory Time Series," *Annals of Statistics*, 22, 515-539.
- [22] ROBINSON, P.M. (1994c): "Rates of Convergence and Optimal Spectral Bandwidth for Long Memory Dependence," *Probability Theory and Related Fields*, 99, 443-473.
- [23] ROBINSON, P.M. (1995a): "Log-Periodogram Regression for Time Series with Long Range Dependence," *Annals of Statistics*, 23, 1048-1072.
- [24] ROBINSON, P.M. (1995b): "Gaussian Semiparametric Estimation of Long-Range Dependence," *Annals of Statistics*, 23, 1630-1661.
- [25] ROBINSON, P.M. AND F.J. HIDALGO (1997): "Time Series Regression with Long Range Dependence," *Annals of Statistics*, 25, 77-104.
- [26] ROBINSON, P.M. AND D. MARINUCCI (1998): "Semiparametric Frequency-Domain Analysis of Fractional Cointegration," Preprint.

- [27] ROSENBLATT, M. (1961): "Independence and Dependence," *Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability*. Berkeley: University of California Press, 411-443.
- [28] TAQQU, M.S. (1975): "Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31, 287-302.
- [29] XIAO, Z. AND P.C.B. PHILLIPS (1999): "Higher Order Approximations for Wald Statistics in Time Series Regressions with Integrated Processes," Preprint.
- [30] YAJIMA, Y. (1988): "On Estimation of a Regression Model with Long-Memory Stationary Errors," *Annals of Statistics*, 16, 791-807.
- [31] YAJIMA, Y. (1991): "Asymptotic Properties of the LSE in a Regression Model with Long Memory Stationary Errors," *Annals of Statistics*, 19, 158-177.
- [32] ZYGMUND, A. (1990): *Trigonometric Series*. Cambridge: Cambridge University Press.

TABLE 1: *MSE* ratio R for METHOD 1

$m = n/32$

$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.325	1.380	1.505	1.859	3.878	1.221	1.258	1.349	1.625	3.269	1.141	1.169	1.231	1.428	2.698
.05	1.306	1.342	1.420	1.613	2.304	1.209	1.232	1.288	1.436	1.996	1.132	1.153	1.192	1.295	1.712
.15	1.285	1.302	1.334	1.404	1.580	1.196	1.206	1.228	1.280	1.422	1.121	1.135	1.156	1.195	1.295
.25	1.276	1.287	1.303	1.331	1.395	1.191	1.196	1.206	1.229	1.280	1.113	1.124	1.140	1.165	1.209
.35	1.276	1.285	1.298	1.315	1.346	1.191	1.195	1.202	1.216	1.243	1.107	1.117	1.131	1.153	1.189
.45	1.285	1.295	1.309	1.330	1.359	1.197	1.201	1.209	1.224	1.250	1.105	1.114	1.127	1.150	1.189

$m = n/16$

$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.231	1.276	1.384	1.699	3.522	1.159	1.191	1.273	1.523	3.032	1.103	1.129	1.185	1.369	2.572
.05	1.220	1.249	1.315	1.487	2.114	1.149	1.171	1.220	1.355	1.867	1.097	1.116	1.152	1.247	1.641
.15	1.209	1.224	1.251	1.313	1.475	1.140	1.150	1.171	1.221	1.352	1.088	1.103	1.125	1.164	1.260
.25	1.208	1.218	1.234	1.261	1.321	1.137	1.143	1.155	1.180	1.232	1.081	1.094	1.115	1.145	1.195
.35	1.211	1.222	1.237	1.256	1.289	1.138	1.143	1.152	1.170	1.203	1.074	1.087	1.107	1.138	1.187
.45	1.222	1.235	1.253	1.277	1.311	1.143	1.148	1.157	1.176	1.208	1.070	1.082	1.101	1.133	1.190

$m = n/8$

$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.149	1.192	1.293	1.586	3.283	1.111	1.141	1.217	1.449	2.865	1.080	1.105	1.163	1.344	2.522
.05	1.141	1.169	1.232	1.395	1.981	1.104	1.124	1.170	1.295	1.774	1.075	1.095	1.133	1.228	1.615
.15	1.136	1.152	1.181	1.244	1.402	1.098	1.109	1.132	1.182	1.308	1.070	1.089	1.118	1.164	1.267
.25	1.139	1.153	1.173	1.207	1.274	1.098	1.107	1.125	1.157	1.218	1.067	1.088	1.121	1.169	1.242
.35	1.146	1.162	1.183	1.213	1.258	1.100	1.110	1.127	1.158	1.210	1.062	1.085	1.122	1.182	1.274
.45	1.158	1.177	1.203	1.239	1.290	1.106	1.116	1.134	1.168	1.226	1.059	1.080	1.117	1.184	1.300

TABLE 2: *MSE* ratio R for METHOD 2

$m = n/32$															
$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.275	1.302	1.376	1.621	3.170	1.205	1.233	1.306	1.537	2.985	1.141	1.167	1.225	1.413	2.646
.05	1.271	1.285	1.323	1.443	1.946	1.199	1.217	1.260	1.381	1.862	1.135	1.155	1.193	1.293	1.702
.15	1.282	1.288	1.298	1.327	1.431	1.199	1.209	1.228	1.270	1.386	1.128	1.144	1.168	1.213	1.322
.25	1.317	1.330	1.341	1.352	1.375	1.212	1.224	1.240	1.264	1.309	1.126	1.142	1.166	1.203	1.267
.35	1.379	1.414	1.447	1.473	1.488	1.239	1.259	1.285	1.315	1.349	1.130	1.148	1.177	1.220	1.287
.45	1.485	1.561	1.640	1.710	1.759	1.285	1.318	1.361	1.411	1.464	1.142	1.167	1.204	1.261	1.347

$m = n/16$															
$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.211	1.239	1.317	1.570	3.134	1.154	1.183	1.256	1.484	2.893	1.104	1.129	1.185	1.366	2.556
.05	1.206	1.222	1.264	1.393	1.913	1.147	1.166	1.211	1.332	1.805	1.099	1.118	1.155	1.251	1.646
.15	1.210	1.217	1.229	1.265	1.384	1.143	1.154	1.175	1.222	1.342	1.092	1.109	1.135	1.181	1.289
.25	1.226	1.235	1.243	1.256	1.291	1.146	1.156	1.173	1.203	1.257	1.088	1.105	1.132	1.175	1.247
.35	1.253	1.271	1.289	1.307	1.328	1.156	1.168	1.187	1.218	1.264	1.085	1.103	1.133	1.184	1.267
.45	1.297	1.329	1.366	1.406	1.446	1.172	1.189	1.215	1.256	1.315	1.086	1.106	1.139	1.200	1.306

$m = n/8$															
$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.140	1.175	1.262	1.526	3.101	1.109	1.138	1.210	1.432	2.802	1.080	1.106	1.162	1.343	2.519
.05	1.135	1.158	1.210	1.353	1.892	1.103	1.123	1.168	1.288	1.751	1.076	1.097	1.135	1.232	1.623
.15	1.138	1.152	1.176	1.228	1.368	1.100	1.114	1.140	1.191	1.315	1.073	1.094	1.126	1.178	1.292
.25	1.149	1.165	1.185	1.216	1.274	1.102	1.117	1.142	1.184	1.251	1.071	1.096	1.135	1.197	1.290
.35	1.167	1.189	1.217	1.253	1.298	1.108	1.125	1.154	1.204	1.276	1.068	1.095	1.143	1.226	1.357
.45	1.191	1.223	1.265	1.321	1.385	1.118	1.138	1.174	1.237	1.335	1.066	1.094	1.147	1.249	1.431

TABLE 3: *MSE* ratio R after two iterations for $n = 256$

$m = n/32$

METHOD 1						METHOD 2				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.165	1.192	1.253	1.446	2.703	1.165	1.192	1.253	1.447	2.706
.05	1.157	1.178	1.218	1.321	1.736	1.157	1.178	1.218	1.322	1.740
.15	1.146	1.161	1.184	1.227	1.333	1.146	1.161	1.185	1.228	1.336
.25	1.137	1.149	1.168	1.197	1.248	1.137	1.150	1.168	1.198	1.251
.35	1.131	1.141	1.157	1.183	1.225	1.131	1.142	1.158	1.184	1.227
.45	1.127	1.137	1.151	1.176	1.219	1.128	1.137	1.152	1.177	1.222

$m = n/16$

METHOD 1						METHOD 2				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.110	1.135	1.192	1.373	2.567	1.110	1.135	1.192	1.374	2.568
.05	1.104	1.124	1.161	1.257	1.652	1.104	1.124	1.161	1.257	1.653
.15	1.096	1.112	1.136	1.179	1.283	1.096	1.112	1.136	1.180	1.284
.25	1.088	1.103	1.126	1.161	1.221	1.088	1.103	1.126	1.161	1.222
.35	1.081	1.095	1.116	1.152	1.210	1.081	1.095	1.117	1.152	1.212
.45	1.076	1.088	1.109	1.144	1.207	1.076	1.089	1.109	1.144	1.208

$m = n/8$

METHOD 1						METHOD 2				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.082	1.107	1.164	1.345	2.521	1.082	1.107	1.164	1.345	2.521
.05	1.077	1.098	1.136	1.233	1.624	1.077	1.098	1.136	1.233	1.624
.15	1.073	1.094	1.124	1.175	1.286	1.073	1.094	1.124	1.175	1.287
.25	1.069	1.092	1.128	1.183	1.269	1.069	1.092	1.128	1.184	1.269
.35	1.065	1.088	1.128	1.196	1.304	1.065	1.088	1.128	1.196	1.305
.45	1.061	1.083	1.123	1.195	1.326	1.061	1.083	1.123	1.195	1.327

TABLE 4: *MSE* ratio R^* for METHOD 2

$m = n/32$															
$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.228	1.236	1.258	1.305	1.389	1.119	1.123	1.135	1.160	1.200	1.049	1.042	1.038	1.036	1.036
.05	1.223	1.226	1.238	1.271	1.338	1.108	1.109	1.117	1.139	1.176	1.032	1.022	1.014	1.011	1.013
.15	1.164	1.146	1.127	1.123	1.148	1.038	1.017	1.002	1.004	1.026	0.948	0.915	0.884	0.863	0.859
.25	1.040	0.988	0.935	0.896	0.888	0.907	0.853	0.806	0.781	0.787	0.807	0.743	0.681	0.637	0.623
.35	0.862	0.779	0.703	0.650	0.629	0.732	0.650	0.582	1.541	0.537	0.630	0.541	0.463	0.411	0.394
.45	0.658	0.562	0.486	0.440	0.422	0.541	0.450	0.381	0.342	0.336	0.444	0.352	0.282	0.239	0.226

$m = n/16$															
$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.170	1.184	1.212	1.268	1.369	1.089	1.096	1.111	1.141	1.190	1.017	1.012	1.008	1.007	1.007
.05	1.166	1.175	1.193	1.235	1.317	1.078	1.081	1.093	1.119	1.164	0.999	0.991	0.985	0.984	0.986
.15	1.107	1.092	1.079	1.082	1.119	1.007	0.989	0.979	0.984	1.011	0.916	0.887	0.860	0.843	0.843
.25	0.980	0.930	0.878	0.844	0.845	0.876	0.826	0.782	0.761	0.770	0.778	0.717	0.661	0.623	0.616
.35	0.797	0.713	0.637	0.586	0.571	0.701	0.622	0.556	0.518	0.518	0.604	0.519	0.446	0.400	0.389
.45	0.586	0.489	0.413	0.367	0.352	0.509	0.421	0.354	0.317	0.313	0.423	0.334	0.267	0.228	0.219

$m = n/8$															
$n = 64$						$n = 128$					$n = 256$				
$d \backslash d_x$.05	.15	.25	.35	.45	.05	.15	.25	.35	.45	.05	.15	.25	.35	.45
.00	1.107	1.125	1.159	1.223	1.338	1.050	1.059	1.078	1.113	1.173	1.003	1.002	1.001	1.002	1.004
.05	1.102	1.115	1.139	1.190	1.285	1.040	1.045	1.060	1.091	1.144	0.986	0.982	0.979	0.981	0.985
.15	1.046	1.036	1.029	1.040	1.090	0.972	0.959	0.953	0.963	0.998	0.906	0.882	0.862	0.852	0.858
.25	0.926	0.880	0.835	0.809	0.820	0.846	0.802	0.766	0.752	0.769	0.770	0.716	0.669	0.642	0.647
.35	0.750	0.672	0.601	0.556	0.548	0.676	0.604	0.546	0.517	0.524	0.598	0.519	0.454	0.419	0.423
.45	0.546	0.455	0.384	0.341	0.331	0.490	0.407	0.346	0.316	0.321	0.419	0.335	0.273	0.241	0.244