Submitted to the Annals of Statistics

WAVELET VARIANCE ANALYSIS FOR GAPPY TIME SERIES*

BY DEBASHIS MONDAL AND DONALD B. PERCIVAL

University of Washington

The wavelet variance is a scale-based decomposition of the process variance for a time series and has been used to analyze, for example, time deviations in atomic clocks, variations in soil properties in agricultural plots, accumulation of snow fields in the polar regions and marine atmospheric boundary layer turbulence. We propose two new unbiased estimators of the wavelet variance when the observed time series is 'gappy,' i.e., is sampled at regular intervals, but certain observations are missing. We deduce the large sample properties of these estimators and discuss methods for determining an approximate confidence interval for the wavelet variance. We apply our proposed methodology to series of gappy observations related to atmospheric pressure data and Nile River minima.

1. Introduction. In recent years, there has been great interest in using wavelets to analyze data arising from various scientific fields. The pioneering work of Donoho, Johnstone and co-workers on wavelet shrinkage sparked this interest, and wavelet methods have been used to study a large number of problems in signal and image processing including density estimation,

*Supported by NSF Grant DMS–02–22115.

AMS 2000 subject classifications: Primary 62F12, 62M10, 62M15; secondary 60G15, 60G35, 62P12

Keywords and phrases: Cumulant, Fractionally differenced process, Local stationarity, Nile River minima, Semi-variogram, TAO data

deconvolution, edge detection, nonparametric regression and smooth estimation of evolutionary spectra. See, for example, [6], [11], [10], [15], [20], [23], [27] and references therein. Wavelets also give rise to the concept of the wavelet variance (also called the wavelet power spectrum) which decomposes the sample variance of a time series on a scale by scale basis and provides a time- and scale-based analysis of variance. Here 'scale' refers to a fixed interval or span of time [30]. The wavelet variance is particularly useful as an exploratory tool to identify important scales, to assess properties of long memory processes, to detect inhomogeneity of variance in time series and to estimate time-varying power spectra (thus complementing classical Fourier analysis). Applications include the analysis of time series related to electroencephalographic sleep state patterns of infants [7], the El Niño–Southern Oscillation [37], soil variations [25], solar coronal activity [33], the relationship between rainfall and runoff [24], ocean surface waves [26], surface albedo and temperature in desert grassland [29], heart rate variability [32] and the stability of the time kept by atomic clocks [19].

1.1. *Variance decomposition.* If X_t $(t \in \mathbb{Z})$ is a second-order stationary process, a fundamental property of the wavelet variance is that it breaks up the process variance into pieces, each of which represents the contribution to the overall variance due to variablity on a particular scale. In mathematical notation,

$$\operatorname{var}(X_t) = \sum_{j=1}^{\infty} \nu_X^2(\tau_j),$$

where $\nu_X^2(\tau_j)$ is the wavelet variance associated with dyadic scale $\tau_j = 2^{j-1}$; see equation (2.5) for the precise definition. Roughly speaking, $\nu_X^2(\tau_j)$ is a measure of how much a weighted average of X_t over an interval of τ_j differs from a similar average in an adjacent interval. A plot of $\nu_X^2(\tau_j)$ against τ_j thus reveals which scales are important contributors to the process variance. The wavelet variance is also well-defined if X_t is intrinsically stationary, which means that X_t is nonstationary but its backward differences of a certain order d are stationary. For such a process the wavelet variance at individual scales τ_j exists and serves as a meaningful description of variablity of the process.

1.2. *Scalogram.* If X_t is intrinsically stationary and has an associated spectral density function (SDF) S_X, the wavelet variance provides a simple regularization of S_X in the sense that

$$\nu_X^2(\tau_j) \approx 2 \int_{2^{-j-1}}^{2^{-j}} S_X(f)\, df.$$

The wavelet variance thus summarizes the information in the SDF using just one value per octave frequency band and is particularly useful when the SDF is relatively featureless within each octave band. Suppose for example that X_t is a pure power law process, which means that its SDF is proportional to $|f|^\alpha$. Then, with a suitable choice of the wavelet filter, $\nu_X^2(\tau_j)$ is approximately proportional to $\tau_j^{-\alpha-1}$. The scalogram is a plot of $\log\{\nu_X^2(\tau_j)\}$ versus $\log(\tau_j)$. If it is approximately linear, a power law process is indicated, and the exponent α of the power law can be determined from the slope of the line. Thus for this and other simple models there is no loss in using the summary given by the wavelet variance.

1.3. *Local Stationarity.* Wavelet analysis is particularly useful to handle data that exhibit inhomogeneities. For example if the assumption of stationarity is in question, an alternative assumption is that the time series is locally stationary and can be divided into homogenous blocks. The wavelet variance can be used to check the need for this more complicated approach. Moreover, when stationarity is questionable, as an alternative to dividing the time series into disjoint blocks, we can compute wavelet power spectra within a data window and compare its values as the window slides through the time series. Note that we are not advocating infill mechanism here since a large number of time series observed in geophysics tend to have future observations and obtaining further observations at finer grid (rather than in future) is prohibitive. See Section 7.2 for an example of a time series for which the homogeneity assumption is questionable.

1.4. *Gappy series.* In practice, time series collected in various fields often deviate from regular sampling by having missing values ('gaps') amongst otherwise regularly sampled observations. As is also the case with the classical Fourier transform, the usual discrete wavelet transform is designed for regularly sampled observations and cannot be applied directly to time series with gaps. In geophysics, gaps are often handled by interpolating the data, see e.g., [40], but such schemes are faced with the problem of bias and of deducing what effect interpolation has had on any resulting statistical inference. There are various definitions for nonstandard wavelet transforms that could be applied to gappy data, with the 'lifting' scheme being a prominent example [36]. The general problem with this approach is that the wavelet coefficients are not truly associated with particular scales of interest, thus making it hard to draw meaningful scale-dependent inferences. The methodologies developed here overcome these problems. Wavelet analysis has also been discussed in the context of irregular time series [12], and in the context of signals with continuous gaps [14]. Related works address the problem of the spectral analysis of gappy data [35]. The statistical properties of some of these methodologies are unknown and not easy to derive. We return to this in Section 8 and indicate how we can use our wavelet variance estimator to estimate the SDF for gappy data.

This paper is laid out as follows. In Section 2 we discuss estimation of the wavelet variance for gap-free time series. In Section 3 and 4 we describe estimation and construction of confidence intervals for the wavelet variance based upon gappy time series. In Section 5 we compare various estimates and perform some simulation studies on autoregressive and fractionally differenced processes, while Section 6 describes schemes for estimating wavelet variance for time series with stationary dth backward differences. We consider two examples involving gappy time series related to atmospheric pressure and Nile River minima in Section 7. Finally we end with some discussion in Section 8.

2. Wavelet variance estimation for non-gappy time series. Let $h_{1,l}$ denote a unit level Daubechies wavelet filter of width L normalized such that $\sum_l h_{1,l}^2 = \frac{1}{2}$ [9]. The transfer function for this filter, i.e., its discrete Fourier transform (DFT)

$$H_1(f) = \sum_{l=0}^{L-1} h_{1,l} e^{-\mathrm{i}2\pi fl},$$

has a corresponding squared gain function by definition satisfying

$$\mathcal{H}_1(f) \equiv |H_1(f)|^2 = \sin^L(\pi f) \sum_{l=0}^{\frac{L}{2}-1} \binom{\frac{L}{2} - 1 + l}{l} \cos^{2l}(\pi f). \tag{2.1}$$

We note that $h_{1,l}$ can be expressed as the convolution of $\frac{L}{2}$ first difference filters and a single averaging filter that can be obtained by performing $\frac{L}{2}$

cumulative summations on $h_{1,l}$. The jth level wavelet filter $h_{j,l}$ is defined as the inverse DFT of

$$H_j(f) = H(2^{j-1}f)\prod_{l=0}^{j-2} e^{-i2\pi 2^l f(L-1)} H(\tfrac{1}{2} - 2^l f). \tag{2.2}$$

The width of this filter is given by $L_j \equiv (2^j - 1)(L-1) + 1$. We denote the corresponding squared gain function by $\mathcal{H}_j$. Since $H_j(0) = 0$, it follows that

$$\sum_{l=0}^{L_j-1} h_{j,l} = 0. \tag{2.3}$$

For a nonnegative integer d, let X_t $(t \in \mathbb{Z})$ be a process with dth order stationary increments, which implies that

$$Y_t \equiv \sum_{k=0}^{d} \binom{d}{k} (-1)^k X_{t-k} \tag{2.4}$$

is a stationary process. Let S_X and S_Y represent the SDFs for X_t and Y_t. These SDFs are defined over the Fourier frequencies $f \in [-\frac{1}{2}, \frac{1}{2}]$ and are related by $S_Y(f) = [2\sin(\pi f)]^{2d} S_X(f)$. We can take the wavelet variance at scale $\tau_j = 2^{j-1}$ to be defined as

$$\nu_X^2(\tau_j) \equiv \int_{-1/2}^{1/2} \mathcal{H}_j(f) S_X(f)\, df. \tag{2.5}$$

By virtue of (2.1) and (2.2), the wavelet variance is well defined for $L \geq 2d$. When $d = 0$ so that X_t is a stationary process with autocovariance sequence (ACVS) $s_{X,k} \equiv \operatorname{cov}\{X_t, X_{t+k}\}$, then we can rewrite the above as

$$\nu_X^2(\tau_j) = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} s_{X,l-l'}. \tag{2.6}$$

When $d = 1$, the increment process $Y_t = X_t - X_{t-1}$ rather than X_t itself is stationary, in which case the above equation can be replaced by one involving the ACVS for Y_t and the cumulative sum of $h_{j,l}$ [8]. Alternatively, let $\gamma_{X,k} =$

$\frac{1}{2}\text{var}\,(X_0 - X_k)$ denote the semi-variogram of X_t. Then the wavelet variance can be expressed as

$$\nu_X^2(\tau_j) = -\sum_{l=0}^{L_j-1}\sum_{l'=0}^{L_j-1} h_{j,l}h_{j,l'}\gamma_{X,l-l'}. \tag{2.7}$$

The above equation also holds when X_t is stationary.

Given an observed time series that can be regarded as a realization of $X_0, \ldots, X_{N-1}$ and assuming the sufficient condition $L > 2d$, an unbiased estimator of $\nu_X^2(\tau_j)$ is given by

$$\hat{\nu}_X^2(\tau_j) \equiv \frac{1}{M_j}\sum_{t=L_j-1}^{N-1} W_{j,t}^2,$$

where $M_j \equiv N - L_j + 1$, and

$$W_{j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l}X_{t-l}.$$

The wavelet coefficient process $W_{j,t}$ is stationary with mean zero, an SDF given by $\mathcal{H}_j(f)S_X(f)$ and an ACVS to be denoted by $s_{j,k}$. The following theorem holds [30].

THEOREM 2.1. *Let $W_{j,t}$ be a mean zero Gaussian stationary process satisfying the square integrable condition*

$$A_j \equiv \int_{-1/2}^{1/2} \mathcal{H}_j^2(f)S_X^2(f)\,df = \sum_{k=-\infty}^{\infty} s_{j,k}^2 < \infty.$$

Then $\hat{\nu}_X^2(\tau_j)$ is asymptotically normal with mean $\nu_X^2(\tau_j)$ and large sample variance $2A_j/M_j$.

In practical applications, A_j is estimated by

$$\hat{A}_j = \tfrac{1}{2}\hat{s}_{j,0}^2 + \sum_{k=1}^{M_j-1} \hat{s}_{j,k}^2,$$

where

$$\hat{s}_{j,k} = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1-|k|} W_{j,t} W_{j,t+|k|}$$

is the usual biased estimator of the ACVS for a process whose mean is known to be zero. Theorem 2.1 provides a simple basis for constructing confidence intervals for the wavelet variance $\nu_X^2(\tau_j)$.

3. Wavelet variance estimation for gappy time series. We consider first the case $d = 0$, so that X_t itself is stationary with ACVS $s_{X,k}$ and variogram $\gamma_{X,k}$. Consider a portion $X_0, \dots, X_{N-1}$ of this process. Let δ_t be the corresponding gap pattern, assumed to be a portion of a binary stationary process independent of X_t. The random variable δ_t assumes the values of 0 or 1 with nonzero probabilities, with zero indicating that the corresponding realization for X_t is missing. Define

$$\beta_k^{-1} = \Pr(\delta_t = 1 \text{ and } \delta_{t+k} = 1),$$

which is necessarily greater than zero. For $0 \le l, l' \le L_j - 1$, let

$$\hat{\beta}_{l,l'}^{-1} \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \delta_{t-l} \delta_{t-l'}.$$

We assume that $\hat{\beta}_{l,l'}^{-1} > 0$ for all l and l'. For a fixed j, this condition will hold asymptotically almost surely, but it can fail for finite N for a time series with too many gaps, a point that we return to in Section 8. By the weak law of large numbers, $\hat{\beta}_{l,l'}^{-1}$ is a consistent estimator of $\beta_{l-l'}^{-1}$ as $N \to \infty$.

Consider the following two statistics:

$$\hat{u}_X(\tau_j) \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{l,l'} X_{t-l} X_{t-l'} \delta_{t-l} \delta_{t-l'} \tag{3.1}$$

and

$$\hat{v}_X(\tau_j) \equiv -\frac{1}{2M_j} \sum_{t=L_j-1}^{N-1} \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{l,l'} \left(X_{t-l} - X_{t-l'}\right)^2 \delta_{t-l} \delta_{t-l'}. \tag{3.2}$$

When $\delta_t = 1$ for all t (the gap-free case), both statistics collapse to $\hat{\nu}_X^2(\tau_j)$. Conditioning on the observed gap pattern $\delta = (\delta_0, \ldots, \delta_{N-1})$, it follows that

$$E\{\hat{u}_X(\tau_j) \mid \delta\} = E\{\hat{v}_X(\tau_j) \mid \delta\} = \nu_X^2(\tau_j)$$

and hence that both statistics are unconditionally unbiased estimators of $\nu_X^2(\tau_j)$; however, whereas $\hat{\nu}_X^2(\tau_j) \geq 0$ necessarily in the gap-free case, these two estimators can be negative.

REMARK: In the gappy case, the covariance type estimator $\hat{u}_X(\tau_j)$ does not remain invariant if we add a constant to the original process X_t, whereas the variogram type estimator $\hat{v}_X(\tau_j)$ does. In practical applications, this fact becomes important if the sample mean of the time series is large compared to its sample standard deviation, in which case it is important to use $\hat{u}_X(\tau_j)$ only after centering the series by subtracting off the sample mean.

4. Large sample properties of $\hat{u}_X(\tau_j)$ and $\hat{v}_X(\tau_j)$. For a fixed j, define the following stochastic processes:

$$Z_{u,j,t} \equiv \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \beta_{l-l'} X_{t-l} X_{t-l'} \delta_{t-l} \delta_{t-l'}, \tag{4.1}$$

and

$$Z_{v,j,t} \equiv -\tfrac{1}{2} \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \beta_{l-l'} (X_{t-l} - X_{t-l'})^2 \delta_{t-l} \delta_{t-l'}. \tag{4.2}$$

The processes $Z_{u,j,t}$ and $Z_{v,j,t}$ are both stationary with mean $\nu_X^2(\tau_j)$, and both collapse to $W_{j,t}^2$ in the gap-free case. Our estimators $\hat{u}_X(\tau_j)$ and $\hat{v}_X(\tau_j)$

are essentially sample means of $Z_{u,j,t}$ or $Z_{v,j,t}$, with $\beta_{l-l'}$ replaced by $\hat{\beta}_{l,l'}$. At this point we assume the following technical condition about our gap process.

ASSUMPTION 4.1. *For fixed j, let $V_{p,t} = \delta_{t-l}\delta_{t-l'}$ for $p = (l, l')$ and $l, l' = 0, \ldots, L_j - 1$. We assume that the covariances of $V_{p_1,t}$ and $V_{p_2,t}$ satisfy $\sum_k |\operatorname{cov}(V_{p_1,t}, V_{p_2,t+k})| < \infty$ and the higher order cumulants satisfy*

$$\sum_{t_1=0}^{N-1} \cdots \sum_{t_n=0}^{N-1} |\operatorname{cum}(V_{p_1,t_1}, \ldots, V_{p_n,t_n})| = o(N^{n/2}) \tag{4.3}$$

for $n = 3, 4, \ldots$ and for fixed $p_1, \ldots, p_n$.

REMARK: Assumption 4.1 holds for a wide range of binary processes. For example, if δ_t is derived by thresholding a stationary Gaussian process whose covariances are absolutely summable, then the higher order cumulants of $V_{p,t}$ are absolutely summable. Note that Assumption 4.1 is weaker than the assumption that the cumulants are absolutely summable. This latter assumption has been used to prove central limit theorems in other contexts; see, e.g., assumption 2.6.1 of [4].

The following central limit theorems (Theorem 4.2 and 4.3) provide the basis for inference about the wavelet variance using the estimators $\hat{u}_X(\tau_j)$ and $\hat{v}_X(\tau_j)$. We defer proofs to the Appendix, but we note that they are based on calculating mixed cumulants and require a technique sometimes known as a diagram method. This method has been used widely to prove various central and non-central limit theorems involving functionals of Gaussian random variables. See e.g., [3], [16], [17], [13], [21] and the references therein. While building upon previous works, the proofs involve some unique and significantly different arguments that can be used to strengthen asymptotic results in other contexts, e.g., wavelet covariance estimation.

THEOREM 4.2. *Suppose X_t is a stationary Gaussian process whose SDF is square integrable, and suppose δ_t is a strictly stationary binary process (independent of X_t) such that Assumption 4.1 holds. Then $\hat{u}_X(\tau_j)$ is asymptotically normal with mean $\nu_X^2(\tau_j)$ and large sample variance $S_{u,j}(0)/M_j$, where $S_{u,j}$ is the SDF for $Z_{u,j,t}$, with a formula stated in the Appendix.*

REMARK: The Gaussian assumption on X_t can be dropped if we add appropriate mixing conditions, an approach that has been taken in the gap-free case [34]. Since our estimators are essentially averages of stationary processes (4.1) and (4.2), asymptotic normality for the estimators (3.1) and (3.2) will follow if both X_t and the gap process δ_t possess appropriate mixing conditions. Moreover, construction of confidence intervals for the wavelet variance when X_t is non-Gaussian and the asymptotic normality of the estimators holds is same as what is described below. This incorporates robustness into the methods developed in this paper.

Given a consistent estimator of $S_{u,j}(0)$, the above theorem can be used to construct an asymptotically correct confidence interval for $\nu_X^2(\tau_j)$. We use a multitaper spectral approach [34]. Let

$$\widetilde{Z}_{u,j,t} \equiv \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{l,l'} X_{t-l} X_{t-l'} \delta_{t-l} \delta_{t-l'}, \quad t = L_j - 1, \ldots, N-1.$$

Let $\lambda_{k,t}$, $t = 0, \ldots, M_j - 1$, for $k = 0, \ldots, K-1$ be the first K orthonormal

Slepian tapers, where K is an odd integer. Define

$$J_{u,j,k} = \sum_{t=0}^{M_j-1} \lambda_{k,t} \widetilde{Z}_{u,j,t+L_j-1}, \quad \lambda_{k,+} = \sum_{t=0}^{M_j-1} \lambda_{k,t}, \quad \widetilde{u}_j = \frac{\sum_{k=0,2,\ldots}^{K-1} J_{u,j,k}\lambda_{k,+}}{\sum_{k=0,2,\ldots}^{K-1} \lambda_{k,+}^2}.$$

We estimate $S_{u,j}(0)$ by

$$\hat{S}_{u,j}(0) = \frac{1}{K} \sum_{k=0}^{K-1} (J_{u,j,k} - \widetilde{u}_j \lambda_{k,+})^2.$$

Following the recommendation of [34], we choose $K = 5$ and set the bandwidth parameter so that the Slepian tapers are band-limited to the interval $[-\frac{7}{2M_j}, \frac{7}{2M_j}]$. Previous Monte Carlo studies (see [34]) show that $\hat{S}_{u,j}(0)$ performs well.

We now turn to the large sample properties of the second estimator $\hat{v}_X(\tau_j)$, which closely resemble those for $\hat{u}_X(\tau_j)$.

THEOREM 4.3. *Suppose X_t or its increments is a stationary Gaussian process whose SDF is such that $\sin^2(\pi f) S_X(f)$ is square integrable. Assume the same conditions on δ_t as in Theorem 4.2. Then $\hat{v}_X(\tau_j)$ is asymptotically normal with mean $v_X^2(\tau_j)$ and large sample variance $S_{v,j}(0)/M_j$, where $S_{v,j}$ is the SDF for $Z_{v,j,t}$, with a formula stated in the Appendix.*

Based upon

$$\widetilde{Z}_{v,j,t} \equiv -\tfrac{1}{2} \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{l,l'} (X_{t-l} - X_{t-l'})^2 \delta_{t-l} \delta_{t-l'},$$

we can estimate $S_{v,j}(0)$ using the same multitaper approach as before.

4.1. *Efficiency study.* The estimators $\hat{u}_X(\tau_j)$ and $\hat{v}_X(\tau_j)$ both work for stationary processes, whereas the latter can also be used for nonstationary processes with stationary increments. If $\hat{v}_X(\tau_j)$ performed better than $\hat{u}_X(\tau_j)$ in the stationary case, then the latter would be an unattractive estimator because it is restricted to just stationary processes. To address this issue, consider the asymptotic relative efficiency of the two estimators, which is given by the ratio of $S_{v,j}(0)$ to $S_{u,j}(0)$. For selected cases, this ratio can be computed to sufficient accuracy using the relationships

$$S_{u,j}(0) = \sum_{k=-\infty}^{\infty} s_{u,j,k} \text{ and } S_{v,j}(0) = \sum_{k=-\infty}^{\infty} s_{v,j,k},$$

where $s_{u,j,k}$ and $s_{v,j,k}$ are the ACVSs corresponding to SDFs $S_{u,j}$ and $S_{v,j}$. We consider two cases, in both of which we use a level $j = 3$ Haar wavelet filter and assume that δ_t is a sequence of independent and identically distributed Bernoulli random variables with $\Pr(\delta_t = 1) = 0.9$. In the first case, we let X_t to be a first order autoregressive (AR(1)) process with $s_{X,k} = \phi^{|k|}$. The left-hand plot of Figure 1 shows the asymptotic relative efficiency as a function of ϕ. Except for ϕ close to unity, $\hat{u}_X(\tau_j)$ outperforms $\hat{v}_X(\tau_j)$. When ϕ is close to unity, the differencing inherent in $\hat{v}_X(\tau_j)$ makes it a more stable estimator than $\hat{u}_X(\tau_j)$, which is inituitively reasonable because the AR(1) process starts to resemble a random walk. For the second case, let X_t to be a stationary fractionally differenced (FD) process with $s_{X,k}$ satisfying

$$s_{X,0} = \frac{\Gamma(1-2\alpha)}{\Gamma(1-\alpha)\Gamma(1-\alpha)} \text{ and } s_{X,k} = s_{X,k-1}\frac{k+\alpha-1}{k-\alpha}$$

for $k = 1, 2, \ldots$; see, e.g., [18], [22]. Here $\alpha < \frac{1}{2}$ is the long memory parameter, with $\alpha = 0$ corresponding to white noise and α close to $\frac{1}{2}$ corresponding to a highly correlated process whose ACVS damps down to zero very slowly. The right-hand plot of Figure 1 shows the asymptotic relative efficiency as a function of α. As α approaches $\frac{1}{2}$, the variogram-based estimator $\hat{v}_X(\tau_j)$

outperforms $\hat{u}_X(\tau_j)$. These two cases tell us that $\hat{u}_X(\tau_j)$ is not uniformily better than $\hat{v}_X(\tau_j)$ for stationary processes and that, even for these processes, differencing can help stabilize the variance. Experimentation with other Daubechies filters leads us to the same conclusions.

5. Monte Carlo study. The purpose of this Monte Carlo study is to access the adequacy of the normal approximation in Theorem 4.2 and 4.3 for simple situations. We also look at the performance of the estimates of $S_{u,j}(0)$ and $S_{v,j}(0)$.

5.1. *Autoregressive process of order 1.* In the first example, we simulate 1000 time series of length 1024 from an AR(1) process with $\phi = 0.9$. For each time series, we simulate δ_t independent and identically from a Bernoulli distribution with $\Pr(\delta_t = 1) = p = 0.9$. For each simulated gappy series, we estimate wavelet variances at scales indexed by $j = 1, \ldots, 6$ using $\hat{u}_X(\tau_j)$ and $\hat{v}_X(\tau_j)$ with the Haar wavelet filter. We also estimate the variance of the wavelet variances by using the multitaper method described in Section 4 and also from the sample variance of the Monte Carlo estimates. We then compare estimated values with the corresponding large sample approximations. Table 1 summarizes this experiment. Let $\hat{u}_{X,r}(\tau_j)$ and $\hat{v}_{X,r}(\tau_j)$ be the wavelet variance estimates for the rth realization, and let $\hat{S}_{u,j,r}(0)$ and $\hat{S}_{v,j,r}(0)$ be the corresponding multitaper estimates of $S_{u,j}(0)$ and $S_{v,j}(0)$. We note from Table 1 that the sample means of $\hat{u}_{X,r}(\tau_j)$ and $\hat{v}_{X,r}(\tau_j)$ are in excellent agreement with the true wavelet variance $\nu_X^2(\tau_j)$. The sample standard deviations of $\hat{u}_{X,r}(\tau_j)$ and $\hat{v}_{X,r}(\tau_j)$ are also in good agreement with $M_j^{-\frac{1}{2}} S_{u,j}^{\frac{1}{2}}(0)$ and $M_j^{-\frac{1}{2}} S_{v,j}^{\frac{1}{2}}(0)$. In particular, the ratios of the standard deviation of the

$\hat{u}_{X,r}(\tau_j)$'s to their large sample approximations are quite close to unity, ranging between 0.884 and 1.005. The corresponding ratios for $\hat{v}_{X,r}(\tau_j)$ range between 0.926 and 1.002. We also consider the performance of the multitaper estimates. In particular, we find the sample means of $M_j^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}_{u,j,r}(0)$ and $M_j^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}_{v,j,r}(0)$ are close to their respective theoretical values, but with a slight downward bias. Figure 2 plots the realization of the time series for which the sum of squares of errors $\sum_j\{\hat{u}_{X,r}(\tau_j)-\nu_X^2(\tau_j)\}^2$ is closest to the average sum of squares of errors, namely, $1000^{-1}\sum_r\sum_j\{\hat{u}_{X,r}(\tau_j)-\nu_X^2(\tau_j)\}^2$. For this typical realization, we also plot the estimated and theoretical wavelet variances with corresponding 95% confidence intervals. The black (gray) solid line in Figure 2 gives the estimated (theoretical) confidence intervals based on $\hat{u}_X(\tau_j)$, with the dotted lines indicating corresponding intervals based upon $\hat{v}_X(\tau_j)$. We see reasonable agreement between the theoretical and estimated values.

5.2. *Kolmogrov turbulence.* In the second example, we generate 1000 time series of length 1024 from an FD($\frac{5}{6}$) process, which is a nonstationary process that has properties very similar to Kolmogorov turbulence and hence is of interest in atmospheric science and oceanography. For each time series, we simulate the gaps δ_t as before. In this example increments of X_t rather the X_t itself are stationary. Therefore we employ only $\hat{v}_X(\tau_j)$ and consider how well its variance is approximated by the large sample result stated in Theorem 4.3. Table 2 summarizes the results of this experiment using the Haar wavelet filter. Again we find that, for each level j, the average $\hat{v}_{X,r}(\tau_j)$ is in excellent agreement with the true $\nu_X^2(\tau_j)$; the sample standard deviation of $\hat{v}_{X,r}(\tau_j)$ is in good argeement with its large sample approximation; and the sample mean of $M_j^{-\frac{1}{2}}\hat{S}_{v,j}^{\frac{1}{2}}(0)$ is close to $M_j^{-\frac{1}{2}}S_{v,j}^{\frac{1}{2}}(0)$, with a slight downward bias. Figure 3 has the same format as Figure 2 and again indicates reasonable agreement between theoretical and estimated values.

6. Generalization of basic theory.

6.1. *Gappy dth order stationary increment process.* In this subsection, we extend the basic theory developed in Section 3 and Section 4 to handle estimation of the wavelet variance for dth order stationary increment processes. First we note that Theorems 4.2 and 4.3 hold for a wider class of wavelet filters than just the Daubchies filters. In particular, both theorems continue to hold for any filter $h_{j,l}$ that has finite width and sums to zero (if the original process X_t has mean zero, Theorem 4.2 only requires $h_{j,l}$ to be of finite width). This provides us with an estimation theory for wavelet variances other than those defined by a Daubechies wavelet filter. For example, at the unit scale, we can entertain the filter $\{-\frac{1}{4}, \frac{1}{2}, -\frac{1}{4}\}$, which can be considered to be a discrete approximation of the Mexican hat wavelet. Moreover, as useful byproducts, we obtain the following schemes that deal with estimation of the Daubechies wavelet variance for a general dth order backward stationary increment process.

Assume as in (2.4) that X_t for $t \in \mathbb{Z}$ is a process with dth order stationary increments Y_t. Let μ_Y be the mean, $s_{Y,k}$ the ACVS and $\gamma_{Y,k}$ the semivariogram of Y_t. For $L \geq 2d$, an expression for the Daubechies wavelet variance that is analogous to (2.6) is

$$\nu_X^2(\tau_j) = \sum_{l=0}^{L_j-d-1} \sum_{l'=0}^{L_j-d-1} b_{j,l,d} b_{j,l',d} s_{Y,l-l'}, \tag{6.1}$$

where $b_{j,l,r}$ is the rth order cumulative summation of Daubechies wavelet filter $h_{j,l}$, i.e.,

$$b_{j,l,0} = h_{j,l}, \quad b_{j,l,k} = \sum_{r=0}^{l} b_{j,r,k-1},$$

for $l = 0, \ldots, L_j - k - 1$ (see [8]). Moreover, if $L > 2d$, we obtain the alternative expression

$$\nu_X^2(\tau_j) = -\sum_{l=0}^{L_j-d-1} \sum_{l'=0}^{L_j-d-1} b_{j,l,d} b_{j,l',d} \gamma_{Y,l-l'}. \tag{6.2}$$

We can now proceed to estimate $\nu_X^2(\tau_j)$ as follows. First we carry out dth successive differencing of the observed X_t to obtain an observed Y_t. This will generate a new gap pattern that has more gaps than the old gap structure, but the new gap pattern will still be stationary and independent of Y_t. We then mimic the stationary ($d = 0$) case described as in Section 3 with $b_{j,l,d}$ replacing $h_{j,l}$, the new gap pattern replacing δ_t, and Y_t replacing X_t in the estimators (3.1) and (3.2). As a simple illustration of this scheme, consider the case $d = 2$. For $t = 2, 3, \ldots$, compute $Y_t = X_t - 2X_{t-1} + X_{t-2}$ whenever $\delta_t = \delta_{t-1} = \delta_{t-2} = 1$. Let $\eta_t = 1$ if $\delta_t = \delta_{t-1} = \delta_{t-2} = 1$ and $= 0$ otherwise. Note that the new gap pattern η_t is stationary and independent of Y_t. Let

$$\hat{\rho}_{l,l'}^{-1} = \frac{1}{M_j} \sum_{t=L_j-3}^{N-1} \eta_{t-l} \eta_{t-l'},$$

where now M_j is redefined to be $N - L_j + 3$. Again $\hat{\rho}_{l,l}^{-1}$ is a consistent estimator of $\rho_{l-l'}^{-1} = \Pr(\eta_{t-l} = 1,\, \eta_{t-l'} = 1)$. As before, assume $\hat{\rho}_{l,l'}^{-1} > 0$ for $l, l' = 0, \ldots, L_j - 3$. The new versions of the estimators of $\nu_X^2(\tau_j)$ are then given by

$$\hat{\hat{u}}_X(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-3}^{N-1} \sum_{l=0}^{L_j-3} \sum_{l'=0}^{L_j-3} b_{j,l,2} b_{j,l',2} \hat{\rho}_{l,l'} Y_{t-l} Y_{t-l'} \eta_{t-l} \eta_{t-l'},$$

and

$$\hat{\hat{v}}_X(\tau_j) = -\frac{1}{2M_j}\sum_{t=L_j-3}^{N-1}\sum_{l'=0}^{L_j-3}\sum_{l'=0}^{L_j-3} b_{j,l,2}b_{j,l',2}\hat{\rho}_{l,l'}(Y_{t-l}-Y_{t-l'})^2\eta_{t-l}\eta_{t-l'}.$$

The large sample properties of these estimators are given by obvious analogs to Theorems 4.2 and 4.3.

THEOREM 6.1. *Suppose X_t is a process whose d order increments Y_t are a stationary Gaussian process with square integrable SDF, and suppose δ_t is a strictly stationary binary process (independent of X_t) such that the derived binary process η_t satisfies Assumption 4.1. Then, if $L \geq 2d$, $\hat{\hat{u}}_X(\tau_j)$ is asymptotically normal with mean $\nu_X^2(\tau_j)$ and large sample variance $S_{d,u,j}(0)/M_j$. where $S_{d,u,j}$ is the SDF for $\sum_l\sum_{l'} b_{j,l,d}b_{j,l',d}\rho_{l,l'}Y_{t-l}Y_{t-l'}\eta_{t-l}\eta_{t-l'}$.*

THEOREM 6.2. *Suppose X_t is a process whose $d+1$ order increments are a stationary Gaussian process with square integrable SDF, and suppose δ_t is a strictly stationary binary process (independent of X_t) such that the derived binary process η_t satisfies Assumption 4.1. Then, if $L > 2d$, $\hat{\hat{v}}_X(\tau_j)$ is asymptotically normal with mean $\nu_X^2(\tau_j)$ and large sample variance $S_{d,v,j}(0)/M_j$. where $S_{d,v,j}$ is the SDF for $-\frac{1}{2}\sum_l\sum_{l'} b_{j,l,d}b_{j,l',d}\rho_{l,l'}(Y_{t-l}-Y_{t-l'})^2\eta_{t-l}\eta_{t-l'}$.*

The proofs of Theorems 6.1 and 6.2 are similar to that of, respectively, Theorems 4.2 and 4.3 and thus are omitted.

REMARK: Since each extra differencing produces more gaps, an estimate that requires less differencing will be more efficient. This is where the semi-variogram estimator $\hat{v}_X(\tau_j)$ comes in handy. Let C_t denote the $(d-1)$th backward differences of X_t. Then C_t is not stationary but its increments are. Let the semi-variogram of C_t be denoted by $\gamma_{C,k}$. Then by virtue of

(6.2), we can write for $L \geq 2d$

$$\nu_X^2(\tau_j) = -\sum_{l=0}^{L_j-d}\sum_{l'=0}^{L_j-d} b_{j,l,d-1} b_{j,l',d-1} \gamma_{C,l-l'}. \tag{6.3}$$

Thus alternatively we can proceed as follows. We carry out $d-1$ successive differences of X_t to obtain C_t and then use the semi-variogram estimator with the new gap structure and with the Daubechies filter replaced by $b_{j,l,d-1}$. Unlike the stationary case, this estimator often outperforms the covariance-type estimator that requires one more order of differencing.

6.2. *Systematic gaps.* We have focused on geophysical applications which tend to have gaps that are stochastic in nature. When systematic gaps occur, e.g., in financial time series when no trading takes place on weekends, we note that our estimates (3.1) and (3.2) produce valid unbiased estimate of the true wavelet variance as long as $\hat{\beta}_{l,l'}^{-1} > 0$ for $l, l' = 0, \ldots, L_j - 1$ (for the financial example, this condition on $\hat{\beta}$ holds when the length of the time series N is sufficiently large); moreover, our large sample theory can be readily adjusted to handle those gaps. First, we redefine the theoretical β by taking the deterministic limit of $\hat{\beta}$ as N tends to infinity. Next we observe that the processes $Z_{u,j,t}$ and $Z_{v,j,t}$ defined via (4.1) and (4.2) are no longer stationary under this systematic gap pattern. To see this consider $j = 2$ and the Haar wavelet filter for which $L_2 = 4$. Then $Z_{u,2,t}$ for a Friday depends on the observations obtained from Tuesday to Friday while $Z_{u,2,t}$ for a Monday depends only on values of the time series observed on Monday and the previous Friday. As a consequence we can not invoke Theorem 4.2 or 4.3 directly. However, because the gaps have a period of a week, we can retrieve stationarity by summing $Z_{u,j,t}$ and $Z_{v,j,t}$ over 7 days; i.e., $\sum_{m=0}^{6} Z_{u,j,t+m}$

and $\sum_{m=0}^{6} Z_{v,j,t+m}$ are stationary processes. For large M_j the summations of t in estimators (3.1) and (3.2) are essentially sums over these stationary processes, plus terms that are asymptotically negligible. Thus we can prove asymptotic normality of (3.1) and (3.2) from the respective asymptotic normality of the averages of $\sum_{m=0}^{6} Z_{u,j,t+m}$ and $\sum_{m=0}^{6} Z_{v,j,t+m}$. The proofs are similar to those for Theorems 4.2 and 4.3, with some simplification because the gaps are deterministic. Alternatively, we can invoke Theorem 1 of [21] to deduce proofs. Large sample confidence intervals can be constructed using the multitaper procedure described in Section 4.

7. Examples.

7.1. *Analysis of TAO data.* We apply our techniques to daily atmospheric pressure data (Figure 4) collected by NOAA's Tropical Atmospheric Ocean (TAO) buoy array over a period of 578 days. There were 527 days of observed values and 51 days during which no observations were made. Shorts gaps in this time series are mainly due to satellite transmission problems. Equipment malfunctions that require buoy repairs result in longer gaps. It is reasonable to assume that the gaps are independent of the pressure values and are a realization of a stationary process. Of particular interest are contributions to the overall variability due to different dynamical phenomena, including an annual cycle, interseasonal oscillations and a menagerie of tropical waves and disturbances associated with small time scales. We employ wavelet variance estimators (3.1) and (3.2) using the Haar wavelet filter.

Estimated wavelet variances for levels $j = 1, \ldots 8$ are plotted in Figure 4 along with the 95% confidence intervals (solid and dotted lines for, respectively, $\hat{u}_X(\tau_j)$ and $\hat{v}_X(\tau_j)$). We see close agreement between these two estimation procedures. Note that the wavelet variance is largest at scales τ_7 and τ_8, which correspond to periods of, respectively, 128–256 days and 256–512 days. Large variability at these scales is due to a strong yearly cycle in the data. Apart from this, we also see a much weaker peak at scale τ_5, which corresponds to a period of 32–64 days and captures the interseasonal variability. Note also that there is hardly any variability at scale τ_1, although there is some at scales τ_2, τ_3 and τ_4, indicating relatively important contributions to the variance due to disturbances at all but the very smallest scale. Finally, we note that we obtained similar results using the Daubechies $L = 4$ extremal phase and $L = 8$ least asymmetric wavelet filters.

7.2. *Nile river minima.* This time series (Figure 5) consists of measurements of minimum yearly water level of the Nile River over the years 622–1921, with 622–1284 representing the longest segment without gaps [38]. The rate of gaps is about 43% after year 1285. Several authors have previously analyzed the initial gap free segment (see, e.g., [1] and [31]). The entire series, including the gappy part, have been analyzed based on a parametric state space model [28], in contrast to our nonparametric approach. Historical records indicate a change around 715 in the way the series was measured. For the gap free segment, there is more variability at scales τ_1 and τ_2 before 715 than after [41]. Therefore we restrict ourselves to the period 716–1921. Figure 5 plots wavelet variance estimates up to scale τ_8 along with 95% confidence intervals using $\hat{v}_X(\tau_j)$ with the Haar wavelet filter. Here solid lines stand for the gap free segment 716–1284, and dotted lines for the gappy segment 1286–1921. Except at scales τ_1, τ_6 and τ_8, we see reasonably good agreement between the estimates from the two segments. Substantial uncertainties due to the large number of gaps are reflected in the larger confidence intervals of the wavelet variances for the gappy segment. Under the assumption that the statistical properties of the Nile River were the same throughout 716–1921, we could combine the analyses of the two segments to produce overall estimates and confidence intervals for the wavelet variances; however, this assumption is questionable at certain scales. Over the years 1286–1470, there are only six gaps. Separate analysis of this segment suggests more variability at scales τ_1 and τ_2 than what was observed in 716–1284. In addition, construction of the first Aswan Dam starting in 1899 changed the nature of the Nile River in the subsequent years. However, a wavelet variance analysis over 1286–1898 (omitting the years after the dam was built) does not differ much from that of 1286–1921. This indicates that the apparent increase in variability at the largest scales of this time series from segment 716–1284 to 1286–1921 cannot be attributed just to the influence of the dam.

8. Discussion.

8.1. In Section 3, we made the crucial assumption that $\hat{\beta}^{-1}_{l,l'} > 0$ for $l, l' = 0, \ldots, L_j - 1$ for a given level j. For small sample sizes, this condition might fail to hold. This situation arises mainly when half or more of the observations are missing and is often accompanied by systematic periodic patterns in the gaps. For example $\hat{\beta}^{-1}_{0,1}$ is zero if δ_t alternates between zero and one, which indicates that the observed time series does not contain relevant information about $\nu^2_X(\tau_1)$. Such gap patterns with at least 50% missing observations require different treatment. Generalized prolate spheroidal sequences have been used to handle spectral estimation of irregularly sampled processes, which in essence corresponds to the construction of special filters [5]. One solution to the $\hat{\beta}_{l,l'}^{-1} = 0$ problem would be to use this approach to construct approximations to the Daubechies filters.

8.2. Given two time series $X_{1,t}$ and $X_{2,t}$, we can define wavelet cross covariances that yield a scale-based analysis of the cross covariance between the two series in a manner similar to a wavelet variance analysis. For estimation of the wavelet cross covariance, see [42] and the references therein. The methodology described in this paper can be extended to estimate the wavelet cross covariance for multivariate gappy time series and can be used to investigate the scale-based relationships among such series.

8.3. Estimation of the SDF for gappy time series is a long-standing difficult problem. In Section 1 we noted that the wavelet variances provide a simple and useful estimator of the integral of the SDF over octave bands. In particular, the Blackman–Tukey ([2], Sec. 18) pilot spectra coincide with the Haar wavelet variances. Recently Tsakiroglou and Walden [39] extended the pilot spectra of Blackman and Tukey by utilising the (maximum overlap) discrete wavelet packet transform. The result is a spectrum estimator that is competitive with existing estimators of the SDF. In the same vein, our wavelet variance estimators can be used to derive estimators of the SDF for gappy time series.

APPENDIX A: PROOFS

We first need the followings propositions and lemmas. To avoid a triviality, we assume throughout that $\mathrm{var}\{X_t\} > 0$.

PROPOSITION A.1. *Let X_t be a real-valued zero mean Gaussian process with ACVS $s_{X,k}$ and with SDF S_X that is square integrable over $[-\frac{1}{2}, \frac{1}{2}]$. Then the bivariate process $\mathbf{U}_t \equiv [X_{t-k}X_{t-k'}, X_{t-l}X_{t-l'}]^T$, for any choice of k, k', l and l', has a spectral matrix $\mathcal{S}_{\mathbf{U}}$ that is continuous.*

PROOF. Using the Isserlis theorem, we have

$$\mathrm{cov}\left(X_{t-k}X_{t-k'}, X_{t-l+\tau}X_{t-l'+\tau}\right) = s_{X,k-l+\tau}s_{X,k'-l'+\tau} + s_{X,k-l'+\tau}s_{X,k'-l+\tau}.$$

By the Fourier transform we obtain

$$\begin{aligned} S_{k,k',l,l'}(f) &= e^{\mathrm{i}2\pi f(k'-l')} \int_{-1/2}^{1/2} e^{\mathrm{i}2\pi f'(k-l-k'+l')} S_X(f')S_X(f-f')\,\mathrm{d}f' \\ &\quad + e^{\mathrm{i}2\pi f(k'-l)} \int_{-1/2}^{1/2} e^{\mathrm{i}2\pi f'(k+l-k'-l')} S_X(f')S_X(f-f')\,\mathrm{d}f'. \end{aligned}$$

Because $\exp\{\mathrm{i}2\pi f(k'-l')\}$ is a continuous function of f, we can establish the continuity of the first term above if we can show that

$$A_{k,k',l,l'}(f) \equiv \int_{-1/2}^{1/2} e^{\mathrm{i}2\pi f'(k-l-k'+l')} S_X(f')S_X(f-f')\,\mathrm{d}f'$$

is a continuous function, from which the continuity of the second term – and hence of $S_{k,k',l,l'}$ itself – follows immediately. The Chauchy–Schwarz inequality says that

$$\begin{aligned} &\left|A_{k,k',l,l'}(f+\rho) - A_{k,k',l,l'}(f)\right| \\ &= \left|\int_{-1/2}^{1/2} e^{\mathrm{i}2\pi f'(k-l-k'+l')} S_X(f')\left[S_X(f+\rho-f') - S_X(f-f')\right]\,\mathrm{d}f'\right| \\ &\leq \left(\int_{-1/2}^{1/2} S_X^2(f')\,\mathrm{d}f' \int_{-1/2}^{1/2} \left|S_X(f+\rho-f') - S_X(f-f')\right|^2\,\mathrm{d}f'\right)^{1/2}. \end{aligned}$$

By hypothesis $\int_{-1/2}^{1/2} S_X^2(f')\,\mathrm{d}f'$ is finite, while

$$\int_{-1/2}^{1/2} \left|S_X(f+\rho-f') - S_X(f-f')\right|^2 \,\mathrm{d}f' \to 0 \text{ as } \rho \to 0$$

by Lemma 1.11, p. 37 of [44]. Hence $A_{k,k',l,l'}$ and $S_{k,k',l,l'}$ are continuous. $\square$

PROPOSITION A.2. *Let $h_{j,l}$ be any filter of finite width L_j with squared gain function $\mathcal{H}_j$. Define $S_{k,k',l,l'}$ as in Proposition A.1 in terms of a squared integrable S_X. Then we must have*

$$\sum_{k,k'} \sum_{l,l'} h_{j,k} h_{j,k'} h_{j,l} h_{j,l'} S_{k,k',l,l'}(0) > 0.$$

PROOF. Using the definition of $S_{k,k',l,l'}$, it follows that

$$\sum_{k,k'} \sum_{l,l'} h_{j,k} h_{j,k'} h_{j,l} h_{j,l'} S_{k,k',l,l'}(0) = 2 \int_{-1/2}^{1/2} \mathcal{H}_j^2(f') S_X^2(f')\,\mathrm{d}f',$$

which is strictly positive because $\mathcal{H}_j$ is zero only on a set of Lebesgue measure zero and $\operatorname{var} X_t > 0$. $\square$

PROPOSITION A.3. *Let X_t be a real-valued zero mean Gaussian process with ACVS $s_{X,k}$ and SDF S_X satisfying*

$$\int_{-1/2}^{1/2} \sin^4(2\pi f) S_X^2(f)\,\mathrm{d}f < \infty.$$

Then the bivariate process $\mathbf{U}_t = \left[\frac{1}{2}(X_{t-k} - X_{t-k'})^2, \frac{1}{2}(X_{t-l} - X_{t-l'})^2\right]^T$, for any choice of k, k', l and l', has a spectral matrix $\mathcal{S}_{\mathbf{U}}$ that is continuous.

PROOF. The proof is similar to that of Proposition A.1. $\square$

PROPOSITION A.4. *Let $h_{j,l}$ be as in Proposition A.2. Assume the conditions of Proposition A.3, and let $S_{k,k',l,l'}$ be the (k,k',l,l') component of $\mathcal{S}_{\mathbf{U}}$*

in that proposition. Then

$$\sum_{k,k'} \sum_{l,l'} h_{j,k} h_{j,k'} h_{j,l} h_{j,l'} S_{k,k',l,l'}(0) > 0.$$

PROOF. The proof is similar to that of Proposition A.2. □

LEMMA A.5. *Let $U_{l,l',t}$ and $V_{l,l',t}$ be stationary processes that are independent of each other for any choice of k, k', l and l'. Let*

$$\begin{aligned} U_{l,l',t} &= \psi_{l,l'} + \int_{-1/2}^{1/2} e^{\mathrm{i}2\pi ft} d\mathcal{U}_{l,l'}(f) \\ V_{l,l',t} &= \omega_{l,l'} + \int_{-1/2}^{1/2} e^{\mathrm{i}2\pi ft} d\mathcal{V}_{l,l'}(f) \end{aligned}$$

be their respective spectral representations. For any k, k', l and l', let $S_{k,k',l,l'}$ and $G_{k,k',l,l'}$ denote the respective cross spectrum between $U_{k,k',t}$ and $U_{l,l',t}$ and between $V_{k,k',t}$ and $V_{l,l',t}$. Let $a_{l,l'}$ be fixed real numbers. Define

$$Q_t = \sum_{l,l'} a_{l,l'} (U_{l,l',t} V_{l,l',t} - \psi_{l,l'} \omega_{l,l'}).$$

Then Q_t is a second order stationary process whose spectral density function is given by

$$\begin{aligned} \text{(A.1)} \quad S_Q(f) &\equiv \sum_{k,k'} \sum_{l,l'} a_{k,k'} a_{l,l'} [\psi_{k,k'} \psi_{l,l'} G_{k,k',l,l'}(f) + \omega_{k,k'} \omega_{l,l'} S_{k,k',l,l'}(f) \\ &+ S * G_{k,k',l,l'}(f)], \end{aligned}$$

where

$$S * G_{k,k',l,l'}(f) \equiv \int_{-1/2}^{1/2} G_{k,k',l,l'}(f - f') S_{k,k',l,l'}(f')\, df'.$$

PROOF. A full proof is straightforward, but tedious. The key steps are to note that

$$\text{(A.2)} \quad \operatorname{cov}\{Q_t, Q_{t+m}\} = \sum_{k,k',l,l'} a_{k,k'} a_{l,l'} \operatorname{cov}\{U_{k,k',t} V_{k,k',t}, U_{l,l',t+m} V_{l,l',t+m}\},$$

to use the spectral representations of $U_{l,l',t}$ and $V_{l,l',t}$ and the independence assumption to obtain

$$\begin{aligned}&\operatorname{cov}\{U_{k,k',t}V_{k,k',t}, U_{l,l',t+m}V_{l,l',t+m}\}\\&\qquad=\int_{-1/2}^{1/2} e^{\mathrm{i}2\pi fm}\Big[\psi_{k,k'}\psi_{l,l'}G_{k,k',l,l'}(f)+\omega_{k,k'}\omega_{l,l'}S_{k,k',l,l'}(f)\\&\qquad+\int_{-1/2}^{1/2} G_{k,k',l,l'}(f-f')S_{k,k',l,l'}(f')\,\mathrm{d}f'\Big]\,\mathrm{d}f,\end{aligned}$$

and to plug the above formula into equation (A.2). □

PROPOSITION A.6. *Let X_t be a real-valued zero mean Gaussian stationary process with ACVS $s_{X,m}$ and SDF S_X that is square integrable over $[-\frac{1}{2},\frac{1}{2}]$. Let δ_t be a binary-valued strictly stationary process that is independent of X_t and satisfies Assumption 4.1. Let $Z_{u,j,t}$ be as in equation (4.1). Then $Z_{u,j,t}$ is a second order stationary process whose SDF at zero is strictly positive.*

PROOF. Let $U_{l,l',t}=X_{t-l}X_{t-l'}$, $\mathbf{U}_t=[U_{k,k',t},U_{l,l',t}]^T$, $V_{l,l',t}=\delta_{t-l}\delta_{t-l'}$ and $a_{l,l'}=h_{j,l}h_{j,l'}\beta_{l-l'}$. By Proposition A.1, $\mathbf{U}_t$ has a continuous cross spectrum $S_{k,k',l,l'}$. By Lemma A.5, the SDF $S_{u,j}(f)$ of $Z_{u,j,t}$ is given by the right-hand side of equation (A.2), where $\psi_{l,l'}=\mathrm{E}\,X_{t-l}X_{t-l'}$, $\omega_{l,l'}=\mathrm{E}\,\delta_{t-l}\delta_{t-l'}=\beta_{l-l'}^{-1}$ and $G_{k,k',l,l'}$ is the cross spectrum between $\delta_{t-k}\delta_{t-k'}$ and $\delta_{t-l}\delta_{t-l'}$. Since $a_{l,l'}\omega_{l,l'}=h_{j,l}h_{j,l'}$, by Proposition A.2

$$\sum_{k,k'}\sum_{l,l'} a_{k,k'}a_{l,l'}\omega_{k,k'}\omega_{l,l'}S_{k,k',l,l'}(0)=\sum_{k,k'}\sum_{l,l'} h_{j,k}h_{j,k'}h_{j,l}h_{j,l'}S_{k,k',l,l'}(0)>0.$$

Now $\sum_{k,k'}\sum_{l,l'} a_{k,k'}a_{l,l'}\psi_{k,k'}\psi_{l,l'}G_{k,k',l,l'}(f)$ and $\sum_{k,k'}\sum_{l,l'} a_{k,k'}a_{l,l'}S{*}G_{k,k',l,l'}(f)$ are nonnegative because $G_{k,k',l,l'}$ and $S_{k,k',l,l'}$ are entries of spectral density matrices. Hence $S_{u,j}(0)>0$. □

PROPOSITION A.7. *Let X_t be a real-valued Gaussian stationary process with zero mean, and SDF S_X that satisfies*

$$\int_{-1/2}^{1/2} \sin^4(\pi f) S_X^2(f)\,\mathrm{d}f < \infty.$$

Let δ_t be a binary-valued strictly stationary process that is independent of X_t and satisfies Assumption 4.1. Let $Z_{v,j,t}$ be as in equation (4.2). Then $Z_{v,j,t}$ is a second order stationary process whose SDF at zero is strictly positive.

PROOF. The proof closely parallels that of Proposition A.6. Here we take $U_{l,l',t} = -\frac{1}{2}(X_{t-l} - X_{t-l'})^2$ and use Propositions A.3 and A.4 instead of Propositions A.1 and A.2. $\square$

Next we state the following theorem from [4], p. 21.

THEOREM A.8. *Consider a two way array of random variables (RVs) $\Theta_{i,j}$, $j = 1, \ldots, J_i$ and $i = 1, \ldots, n$. Consider the n RVs $\Upsilon_i = \prod_{j=1}^{J_i} \Theta_{i,j}$ for $i = 1, \ldots, n$. Then the joint cumulant of $\Upsilon_1, \ldots, \Upsilon_n$ is given by the formula*

$$\operatorname{cum}(\Upsilon_1, \ldots, \Upsilon_n) = \sum_{\chi} \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_r)$$

where the summation is over all indecomposable partitions $\chi = \chi_1 \cup \cdots \cup \chi_r$ of the (not necessarily rectangular) two way table

$$\begin{array}{ccc} (1,1) & \cdots & (1,J_1) \\ \vdots & & \vdots \\ (n,1) & \cdots & (n,J_n). \end{array} \tag{A.3}$$

Next we need the following lemmas.

LEMMA A.9. *Assume that X_t satisfies the conditions stated in Theorem 4.2. Let $U_{p,t} = X_{t-l}X_{t-l'}$ and $\mathrm{E}\, U_{p,t} = \psi_p$, where $p = (l, l')$. Then for $n \geq 3$ and fixed $p_1, \ldots, p_n$,*

$$\sum_{t_1,\ldots t_n} |\operatorname{cum}(U_{p_1,t_1} - \psi_{p_1}, \ldots, U_{p_n,t_n} - \psi_{p_n})| = o(M^{n/2}), \tag{A.4}$$

where each t_i ranges from 0 to $M-1$ (here and below M is shorthand for M_j in the main text).

PROOF. Since a cumulant is invariant under the addition of constants,

$$\operatorname{cum}(U_{p_1,t_1} - \psi_{p_1}, \ldots, U_{p_n,t_n} - \psi_{p_n}) = \operatorname{cum}(U_{p_1,t_1}, \ldots, U_{p_n,t_n}).$$

Consider the $n \times 2$ table of RVs given by

$$\begin{array}{cc} \Theta_{1,1} = X_{t_1-l_1} & \Theta_{1,2} = X_{t_1-l_1'} \\ \vdots & \vdots \\ \Theta_{n,1} = X_{t_n-l_n} & \Theta_{n,2} = X_{t_n-l_n'}. \end{array}$$

As $U_{p,t}$ is the product of the two Gaussian RVs in row p of the table, we invoke Theorem A.8 to break up $\operatorname{cum}(U_{p_1,t_1}, \ldots, U_{p_n,t_n})$. Moreover, because all cumulants of order $r \geq 3$ are zero due to Gaussianity, we can restrict ourselves to indecomposable partitions $\chi = \chi_1 \cup \cdots \cup \chi_n$ of the two way table (A.3) with $J_1 = \cdots = J_m = 2$ so that $|\chi_k| = 2$ for all k. Let $\sum_{t_1,\ldots,t_n} \operatorname{cum}(U_{p_1,t_1}, \ldots, U_{p_n,t_n}) \equiv \sum_\chi I_{U,M}(\chi)$ with

$$I_{U,M}(\chi) = \sum_{t_1,\ldots,t_n} \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_n).$$

Since n is fixed and the number of indecomposable partitions depends only on n, it then suffices to show that $I_{U,M}(\chi) = o(M^{n/2})$ for any fixed χ. As χ is an indecomposable partition, without loss of generality (WLOG), we can properly order the index of table (A.3) so that $\chi_k = \{(k, \eta_k), (k+1, \xi_{k+1})\}$ for $k = 1, \ldots, n-1$ and $\chi_n = \{(n, \eta_n), (1, \xi_1)\}$, where η_k takes values of 1 or 2 for $k = 1, \ldots, n$ and $\xi_k = 3 - \eta_k$. We set, for $k = 1, \ldots, n$,

$$e_k = \begin{cases} l_{(k+1) \bmod n} - l_k & \text{if } \xi_{(k+1) \bmod n} = \eta_k = 1, \\ l'_{(k+1) \bmod n} - l_k & \text{if } \xi_{(k+1) \bmod n} = 2, \eta_k = 1, \\ l_{(k+1) \bmod n} - l'_k & \text{if } \xi_{(k+1) \bmod n} = 1, \eta_k = 2 \quad \text{and} \\ l'_{(k+1) \bmod n} - l'_k & \text{if } \xi_{(k+1) \bmod n} = \eta_k = 2. \end{cases}$$

Then we can write

$$\operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_k) = s_{X, t_{k+1} - t_k - e_k}$$

for $k = 1, \ldots, n-1$ and $\operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_n) = s_{X, t_1 - t_n - e_n}$. Hence

$$I_{U,M}(\chi) \equiv \sum_{t_1, \ldots, t_n} s_{X, t_1 - t_n - e_n} \prod_{i=1}^{n-1} s_{X, t_{i+1} - t_i - e_i}. \tag{A.5}$$

For a fixed K, write $I_{U,M}(\chi) = I'_{U,M}(\chi) + I''_{U,M}(\chi)$, where $I'_{U,M}(\chi)$ is the sum of (A.5) taken over t_i, $i = 1, \ldots, n$, such that $|t_{i+1} - t_i| \leq K$ for $i = 1, \ldots, n-1$ and $|t_1 - t_n| \leq K$. Set $q_i = t_{i+1} - t_i$ for $i = 1, \ldots, n-1$. Since $s_{X,\tau}$ is bounded in magnitute by $s_{X,0}$, we obtain

$$|I'_{U,M}(\chi)| \leq s_{X,0} \sum_{|q_i| \leq K, \, i=1,\ldots,n-1} \sum_{t_n} 1 \leq s_{X,0} K^{n-1} M.$$

The rest of the proof runs parallel to that of Lemma 6 of Giraitis and Surgailis (1985). Thus we show that $I''_{U,M}(\chi) \leq \epsilon(K) M^{n/2}$ where $\epsilon(K) \to 0$

as $K \to \infty$. We repeatedely use the Cauchy–Schwartz inequality to obain

$$\begin{aligned}
& I''_{U,M}(\chi) \\
= \quad & \sum_{t_1,\ldots,t_n} s_{X,t_1-t_n-e_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-e_i} \\
= \quad & \sum_{t_1,\ldots,t_{n-1}} \prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i-e_i} \sum_{t_n} s_{X,t_1-t_n-e_n} s_{X,t_n-t_{n-1}-e_{n-1}} \\
\leq \quad & \sum_{t_1,\ldots,t_{n-1}} \prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}} \\
= \quad & \sum_{t_1,\ldots,t_{n-2}} \prod_{i=1}^{n-3} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \sum_{t_{n-1}} s_{X,t_{n-1}-t_{n-2}-e_{n-2}} \\
& \quad \Big(\sum_{t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}} \\
\leq \quad & \sum_{t_1,\ldots,t_{n-2}} \prod_{i=1}^{n-3} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_{n-1}} s^2_{X,t_{n-1}-t_{n-2}-e_{n-2}}\Big)^{\frac{1}{2}} \\
& \quad \Big(\sum_{t_{n-1},t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}} \\
\vdots \\
\leq \quad & \Big(\sum_{t_1,t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \prod_{i=2}^{n} \Big(\sum_{t_{i-1},t_i} s^2_{X,t_i-t_{i-1}-e_{i-1}}\Big)^{\frac{1}{2}}
\end{aligned}$$

Now use the fact that t_i ranges from 0 to $M-1$ and $|t_{i+1}-t_i| > K$ for $i = 1, \ldots, n-1$ and $|t_1 - t_n| > K$. Thus for example

$$\sum_{t_{i-1},t_i} s^2_{X,t_i-t_{i-1}-e_{i-1}} \leq \text{ constant } \sum_{|\tau|>K} \sum_{t_i} s^2_{X,\tau} = \text{ constant } M \sum_{|\tau|>K} s^2_{X,\tau},$$

where $\sum_{|\tau|>K} s^2_{X,\tau}$ goes to zero as $K \to \infty$ because of the square integrability assumption. Hence we have

$$I''_{U,M}(\chi) \leq \text{ constant } M^{\frac{1}{2}n} \Big(\sum_{|\tau|>K} s^2_{X,\tau}\Big)^{\frac{1}{2}n} = \epsilon(K) M^{\frac{1}{2}n},$$

and the required result follows by choosing $K = \lfloor \log(M) \rfloor$. □

LEMMA A.10. *Assume that X_t satisfies the conditions stated in Theorem 4.3. Let $U_{p,t} = -\frac{1}{2}(X_{t-l} - X_{t-l'})^2$ and $\mathrm{E}\, U_{p,t} = \psi_p$, in which $p = (l, l')$. Then for $n \geq 3$ and fixed $p_1, \ldots, p_n$,*

$$\sum_{t_1,\ldots t_n} |\operatorname{cum}(U_{p_1,t_1} - \psi_{p_1}, \ldots, U_{p_n,t_n} - \psi_{p_n})| = o(M^{n/2}),$$

where each t_i ranges from 0 to $M-1$.

PROOF. The proof goes as that of Lemma A.9 with the modification that $U_{p,t}$ can be written as the product of $X_{t-l} - X_{t-l'}$ and $-\frac{1}{2}(X_{t-l} - X_{t-l'})$, where the Gaussian process $X_{t-l} - X_{t-l'}$ has a squared integrable SDF. □

LEMMA A.11. *Let $U_{p,t}$ be either as in Lemma A.9 or as in Lemma A.10. Assume*

$$\kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_n) = \operatorname{cum}(U_{p_1,t_1} - \psi_{p_1}, \ldots, U_{p_n,t_n} - \psi_{p_n}).$$

Define for $i = 1, 2, \ldots, n-1$

$$\kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_i) = \sum_{t_{i+1},\ldots,t_n} M^{-\frac{1}{2}(n-i-1)} \kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_n),$$

where the summation in t_j ranges from 0 to $M-1$. Then, for $i = 1, 2, \ldots, n$, $\kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_i)$ is bounded and satisfies

$$\sum_{t_1,\ldots,t_i} \kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_i) = o\left(M^{\frac{1}{2}(i+1)}\right). \tag{A.6}$$

PROOF. We retain all the notation of Lemma A.9. Thus

$$\kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_n) = \sum_{\chi} s_{X,t_1-t_n-e_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-e_i}$$

Since equation (A.6) follows from (A.4), it suffices to show that for any fixed χ

$$\sum_{t_{\lambda_1},\ldots,t_{\lambda_i}} M^{-\frac{1}{2}(i-1)} s_{X,t_1-t_n-e_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-e_i} \tag{A.7}$$

is bounded for any distinct choice of $\lambda_1, \dots, \lambda_i$ that belong to $\{1, \dots, n\}$ and $i < n$.

Consider $i = 1$. WLOG assume $\lambda_1 = n$. Then

$$\begin{aligned}
&\sum_{t_n} s_{X,t_1-t_n-e_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-e_i} \\
\leq\; &\prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}},
\end{aligned}$$

which is bounded because of the square integrability assumption. Thus (A.7) is bounded.

Now consider $i = 2$. WLOG assume $\lambda_1 = n$. Now we have two cases. In the first case $\lambda_2 = 1$ or $n-1$ so that the pair $t_{\lambda_1}, t_{\lambda_2}$ appears together in a single term involving s_X in (A.7). If we assume WLOG $\lambda_2 = n-1$ we obtain

$$\begin{aligned}
&\sum_{t_{n-1},t_n} M^{-\frac{1}{2}} s_{X,t_1-t_n-e_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-e_i} \\
\leq\; &\sum_{t_{n-1}} M^{-\frac{1}{2}} \prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}} \\
\leq\; &\prod_{i=1}^{n-3} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_{n-1}} s^2_{X,t_{n-1}-t_{n-2}-e_{n-2}}\Big)^{\frac{1}{2}} \\
&\quad \Big(M^{-1} \sum_{t_{n-1},t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}}
\end{aligned}$$

Clearly the above expression is bounded because $\sum_t s^2_{X,t}$ is so and therefore $\lim_{M\to\infty} M^{-1} \sum_{t_{n-1},t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}} = \sum_\tau s^2_{X,\tau-e_{n-1}} < \infty$. Thus (A.7) is bounded. In the second case assume $\lambda_2 = n-2$. Thus $t_{\lambda_1}, t_{\lambda_2}$ appear in two distinct terms involving s_X in (A.7). Hence

$$\begin{aligned}
&\sum_{t_{n-2},t_n} s_{X,t_1-t_n-e_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-e_i} \\
\leq\ & \sum_{t_{n-2}} \prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}} \\
\leq\ & \prod_{i=1}^{n-4} s_{X,t_{i+1}-t_i-e_i} \Big(\sum_{t_n} s^2_{X,t_1-t_n-e_n}\Big)^{\frac{1}{2}} \Big(\sum_{t_n} s^2_{X,t_n-t_{n-1}-e_{n-1}}\Big)^{\frac{1}{2}} \\
&\qquad \Big(\sum_{t_{n-1}} s^2_{X,t_{n-1}-t_{n-2}-e_{n-2}}\Big)^{\frac{1}{2}} \Big(\sum_{t_{n-2}} s^2_{X,t_{n-2}-t_{n-3}-e_{n-3}}\Big)^{\frac{1}{2}}
\end{aligned}$$

Clearly this is bounded. Note that we do not need to use the $M^{-\frac{1}{2}}$ factor. Thus boundedness of (A.7) holds.

The pattern for the general proof is now clear. Note that, because χ is an indecomposable partition, there can be at most $i-1$ pairs of λ_i, namely $(\lambda_j, \lambda_{j+1})$ for $j = 1, \ldots, i-1$ such that each of $(i-1)$ pairs $(t_{\lambda_j}, t_{\lambda_{j+1}})$ appears in distinct $i-1$ terms involving s_X in the equation (A.7). Thus summing over t_{λ_j} for $j = 1, \ldots, i$ in the left hand side of (A.7) and repeated use of Cauchy–Schwartz inequality will give rise to the $(i-1)$ terms $M^{-1} \sum_{t_{\lambda_j}, t_{\lambda_{j+1}}} s^2_{X, t_{\lambda_{j+1}} - t_{\lambda_j} - e_{\lambda_j}}$, $j = 1, \ldots, i-1$. Note that all these terms are bounded and hence boundedness of (A.7) follows. Of course, if there are less than $(i-1)$ such pairs $(\lambda_j, \lambda_{j+1})$, we no longer need to use the factor $M^{-\frac{1}{2}(i-1)}$ (in fact in the exponent we just need half the number of such pairs). This completes the proof. □

Proof of Theorem 4.2. Take $U_{p,t} = X_{t-l}X_{t-l'}$, $V_{p,t} = \delta_{t-l}\delta_{t-l'}$ and $a_p = h_{j,l}h_{j,l'}\beta_{l-l'}$, where $p = (l, l')$. Take $Q_t = \sum_p a_p(U_{p,t}V_{p,t} - \psi_p\omega_p)$ as in Lemma A.5. Note that $\hat{u}_X(\tau_j) - \nu_X^2(\tau_j)$ is the average of Q_t over $L_j -$

$1 \le t \le N-1$ with $\beta_{l-l'}$ replaced by its consistent estimate $\hat{\beta}_{l,l'}$. Since Q_t is stationary, we first prove a CLT for $R = M^{-\frac{1}{2}} \sum_{t=0}^{M-1} Q_t$ and then invoke Slutsky's theorem to complete the proof that $\hat{u}_X(\tau_j)$ is asymptotically normal. We use Žurbenko ([43], p. 2) to write the log of the characteristic function of R as

$$\log F(\lambda) = \sum_{n=1}^{\infty} \frac{\mathrm{i}^n \lambda^n}{n!} \sum_{t_1,\ldots,t_n} \frac{B_n(t_1,\ldots,t_n)}{M^{n/2}},$$

where B_n is the nth order cumulant of Q_t, and each t_i ranges from 0 to $M-1$. Since Q_t is centered, $B_1(t_1) = 0$. By Proposition A.6, the autocovariances $s_{Q,\tau}$ of Q_t are absolutely summable and $M^{-1} \sum_{t_1} \sum_{t_2} B_2(t_1,t_2) \to \sum_\tau s_{Q,\tau} = S_Q(0) > 0$. In order to prove the CLT for R, it suffices to show that $\sum_{t_1,\ldots,t_n} M^{-n/2} B_n(t_1,\ldots,t_n) \to 0$ for $n = 3, 4, \ldots$.

First using p. 19 of [4], we break up the nth order cumulant as follows:

$$\begin{aligned} B_n(t_1,\ldots,t_n) &= \sum_{p_1} \cdots \sum_{p_n} a_{p_1} \cdots a_{p_n} \\ &\qquad \operatorname{cum}(U_{p_1,t_1}V_{p_1,t_1} - \psi_{p_1}\omega_{p_1}, \ldots, U_{p_n,t_n}V_{p_n,t_n} - \psi_{p_n}\omega_{p_n}). \end{aligned}$$

Let $D_{1,p,t} = (U_{p,t} - \psi_p)(V_{p,t} - \omega_p)$, $D_{2,p,t} = \omega_p(U_{p,t} - \psi_p)$ and $D_{3,p,t} = \psi_p(V_{p,t} - \omega_p)$. Then $U_{p,t}V_{p,t} - \psi_p\omega_p = D_{1,p,t} + D_{2,p,t} + D_{3,p,t}$. Using p. 19 of [4] again, we have

$$\begin{aligned} &\operatorname{cum}(U_{p_1,t_1}V_{p_1,t_1} - \psi_{p_1}\omega_{p_1}, \ldots, U_{p_n,t_n}V_{p_n,t_n} - \psi_{p_n}\omega_{p_n}) \\ &\qquad\qquad = \sum_{c_1,\ldots,c_n} \operatorname{cum}(D_{c_1,p_1,t_1}, \ldots, D_{c_n,p_n,t_n}), \end{aligned}$$

where each c_i ranges from 1 to 3. Therefore, it suffices to show that, for fixed $p_1,\ldots,p_n$ and $c_1,\ldots,c_n$, $\operatorname{cum}(D_{c_1,p_1,t_1},\ldots,D_{c_n,p_n,t_n}) = o(M^{n/2})$. Since the cumulant of n RVs is invariant under a reordering of the RVs, assume $c_1 = c_2 = \cdots = c_m = 1, c_{m+1} = c_{m+2} = \cdots = c_{m'} = 2, c_{m'+1} = c_{m'+2} = \cdots =$

$c_n = 3$, and consider a two way table $\Theta_{i,j}$ with n rows. Rows $i = 1, \ldots, m$ each contain exactly two RVs, namely, $U_{p_i,t_i} - \psi_{p_i}$ and $V_{p_i,t_i} - \omega_{p_i}$ (note that the product of the RVs in row i is D_{1,p_i,t_i}). The remaining $n - m$ rows contain one RV each, namely, $U_{p_i,t_i} - \psi_{p_i}$ (which is proportional to D_{2,p_i,t_i}) for $i = (m+1), \ldots, m'$, and $V_{p_i,t_i} - \omega_{p_i}$ (proportional to D_{3,p_i,t_i}) for $i = m'+1, \ldots, n$. Theorem 4 says $\mathrm{cum}(D_{c_1,p_1,t_1}, \ldots, D_{c_n,p_n,t_n})$ is proportional to $\sum_\chi \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_r)$. We complete the proof by showing that for any fixed χ

$$\sum_{t_1,\ldots,t_n} \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_r) = o(M^{n/2}). \tag{A.8}$$

We prove the above in the following steps.

step 1: Since $\Theta_{i,j}$ is centered, its first order cumulant is zero, so we can restrict ourselves to cases where $|\chi_k| \geq 2$ for all k. If any group of RVs in $\Theta_{i,j} : (i,j) \in \chi_k$ is independent of the remaining RVs in that set, then $\mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_k) = 0$. Since the $U_{p_i,t_i} - \psi_{p_i}$'s and $V_{p_i,t_i} - \omega_{p_i}$'s are independent, we need only consider χ_k containing either just $U_{p_i,t_i} - \psi_{p_i}$'s or just $V_{p_i,t_i} - \omega_{p_i}$'s.

step 2: Consider $m = 0$. In this case each row in $\Theta_{i,j}$ has only one RV, and thus all of $\Theta_{i,j}$ together form the only indecomposable partition $\chi = \chi_1$. Now if $m' = 0$, then by Assumption 4.1

$$\begin{aligned}\sum_{t_1,\ldots,t_n} & \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi) \\ &= \sum_{t_1,\ldots,t_n} \mathrm{cum}(V_{p_1,t_1} - \omega_{p_1}, \ldots, V_{p_n,t_n} - \omega_{p_n}) = o(M^{n/2}).\end{aligned}$$

On the other hand if $m' = n$, then by Lemma A.9

$$\sum_{t_1,\ldots,t_n} \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi) \\ = \sum_{t_1,\ldots,t_n} \operatorname{cum}(U_{p_1,t_1} - \omega_{p_1}, \ldots, U_{p_n,t_n} - \omega_{p_n}) = o(M^{n/2}).$$

Finally we rule out the case $1 \le m' < n$ because then χ contains both $U_{p_i,t_i} - \psi_{p_i}$'s and $V_{p_i,t_i} - \omega_{p_i}$'s and hence $\operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi) = 0$.

step 3: Finally consider $m \ge 1$. Assume that $\chi_1, \ldots, \chi_q$ partition $\{\Theta_{1,1}, \ldots, \Theta_{m',1}\}$ (these are all $U_{p_i,t_i} - \psi_{p_i}$) and that $\chi_{q+1}, \ldots, \chi_r$ partition $\{\Theta_{1,2}, \ldots, \Theta_{m,2}, \Theta_{m'+1,1}, \ldots, \Theta_{n,1}\}$ (these are all $V_{p_i,t_i} - \omega_{p_i}$). To check that (A.8) holds, we need to consider five cases.

case 1: When $m' > m$ we sum over $t_{m+1}, \ldots, t_{m'}$ in the left hand side of (A.8) and use (A.6). In order to keep track of all the individual t_i for which $(i, 1)$ belongs to χ_k for $k = 1, \ldots, q$, we set $0 = \rho_0 \le \rho_1 \le \cdots \le \rho_q = m$, $m = \sigma_0 \le \sigma_1 \le \cdots \le \sigma_q = m'$ and assume, for $k = 1, \ldots, q$, $\chi_k = \{(\rho_{k-1} + 1, 1), \ldots, (\rho_k, 1), (\sigma_{k-1} + 1, 1), \ldots, (\sigma_k, 1)\}$. Then, for $k = 1, \ldots, q$, we obtain by Lemma A.11

$$\begin{aligned} &\sum_{t_{\sigma_{k-1}+1},\ldots,t_{\sigma_k}} \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_k) \\ &= \sum_{t_{\sigma_{k-1}+1},\ldots,t_{\sigma_k}} \kappa_{\rho_k+\sigma_k-\rho_{k-1}-\sigma_{k-1}}(p_i, t_i : (i,1) \in \chi_k) \\ &= M^{\frac{1}{2}(\sigma_k-\sigma_{k-1}-1)^+} \kappa_{\rho_k+\sigma_k-\rho_{k-1}-\sigma_{k-1}}(p_i, : (i,1) \in \chi_k, t_{\rho_{k-1}+1}, \ldots, t_{\rho_k}). \end{aligned}$$

Now boundedness of $\kappa_{\rho_k+\sigma_k-\rho_{k-1}-\sigma_{k-1}}(p_i, (i,1) \in \chi_k, t_{\rho_{k-1}+1}, \ldots, t_{\rho_k})$ yields

$$\begin{aligned} &M^{-\frac{1}{2}(m'-m-1)} \sum_{t_1,\ldots,t_n} \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\ &\propto \sum_{t_1,\ldots,t_m} \sum_{t_{m'+1},\ldots,t_n} \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_{q+1}) \cdots \operatorname{cum}(\Theta_{i,j} : (i,j) \in \chi_r) = o(M^{\frac{1}{2}n}). \end{aligned}$$

The last equality follows from Assumption 4.1.

case 2: If $m' = m$ and $|\chi_k| > 2$ for some k in $q+1, \ldots, r$, then using the boundedness of $\mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_{k'})$ for $k' = 1, \ldots, q$, we obtain

$$\begin{aligned}
&\sum_{t_1,\ldots,t_n} \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
\leq\ & C_0 \sum_{t_1,\ldots,t_n} \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_{q+1}) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_r) = o(M^{n/2}).
\end{aligned}$$

In the above C_0 is a constant and the last equality follows from Assumption 4.1.

case 3: Consider $|\chi_k| = 2$ for $k = q+1, \ldots, r$ and assume $m' = m$. Clearly $2m > n > m$ and $r - q = n - m$. Let $(m+i, 1)$ be contained in χ_{q+i} for $i = 1, \ldots, n$. We sum over $t_{m+1}, \ldots, t_n$ to obtain

$$\begin{aligned}
&\sum_{t_1,\ldots,t_n} \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
&\leq \text{constant} \sum_{t_1,\ldots,t_m} \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_q) \\
&\quad \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_{q+n-m+1}) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
&\leq \text{constant} \sum_{t_1,\ldots,t_m} \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \mathrm{cum}(\Theta_{i,j} : (i,j) \in \chi_q) = o(M^{\frac{1}{2}n}).
\end{aligned}$$

In the above derivation we need the fact that $\sum_{t_i} \mathrm{cum}(V_{p_i,t_i}, V_{p_\tau,t_\tau})$ is bounded and note that the constant is changing from line to line.

case 4: Consider the case $n = m$. Again if any $|\chi_k| > 2$ for $k = 1, \ldots, q$, we are done by using (A.4) along with the fact that cumulants of $V_{p_i,t_i} - \omega_{p_i}$ are bounded.

case 5: The last case is $n = m$ and $|\chi_k| = 2$ for all k. The proof requires Theorem 4 to write down the left hand side of (A.8) in terms of covariances of U_{p_i,t_i} and V_{p_i,t_i} and hinges on the fact that ACVS of U_{p_i,t_i} and V_{p_i,t_i} are absolutely summable.

□

PROOF. OF THEOREM 4.3. Take $U_{p,t} = -\frac{1}{2}(X_{t-l} - X_{t-l'})^2$, $V_{p,t} = \delta_{t-l}\delta_{t-l'}$, $a_p = h_{j,l}h_{j,l'}\beta_{l-l'}$, where $p = (l, l')$. Use Lemma 3 in place of Lemma 2 and complete the proof as in Theorem 4.2 by checking all the steps.

□

ACKNOWLEDGEMENTS

We thank Peter Guttorp and Chris Bretherton for discussions.

REFERENCES

[1] BERAN, J. (1994). *Statistics for Long-Memory Processes.* Monographs on Statistics and Applied Probability, Vol. **61**. Chapman and Hall, New York.

[2] BLACKMAN, R. B. AND TUKEY, J. W. (1958). *The Measurement of Power Spectra.* Dover Publications, New York.

[3] BREUER, P. AND MAJOR, P. (1983). Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13**, 3, 425–441.

[4] BRILLINGER, D. R. (1981). *Time Series.* Holden-Day Inc., Oakland, Calif.

[5] BRONEZ, T. P. (1988). Spectral estimation of irregularly sampled multidimensional processes by genelarized prolate spheroidal sequences. *IEEE Transactions on Acoustics, Speech, and Signal Processing 36*, 1862–1873.

[6] CANDÈS, E. J. AND DONOHO, D. L. (2002). Recovering edges in ill-posed inverse problems: optimality of curvelet frames. *Ann. Statist.* **30**, 3, 784–842. Dedicated to the memory of Lucien Le Cam.

[7] CHIANN, C. AND MORETTIN, P. A. (1998). A wavelet analysis for time series. *Nonparametric Statistics 10*, 1–46.

[8] CRAIGMILE, P. F. AND PERCIVAL, D. B. (2005). Asymptotic decorrelation of between-scale wavelet coefficients. *IEEE Trans. Inform. Theory* **51**, 8, part 2, 1039–1048.

[9] Daubechies, I. (1992). *Ten Lectures on Wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. **61**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

[10] Donoho, D. L. and Johnstone, I. M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26**, 3, 879–921.

[11] Donoho, D. L., Johnstone, I. M., Kerkyacharian, G., and Picard, D. (1995). Wavelet shrinkage: asymptopia? *J. Roy. Statist. Soc. Ser. B* **57**, 2, 301–369. With discussion and a reply by the authors.

[12] Foster, G. (1996). Wavelets for period analysis of unevenly sampled time series. *The Astronomical Journal 112*, 1709–1729.

[13] Fox, R. and Taqqu, M. S. (1987). Central limit theorems for quadratic forms in random variables having long-range dependence. *Probab. Theory Related Fields* **74**, 2, 213–240.

[14] Frick, P. Grossmann, A. and Tchamitchian, P. (1998). Wavelet analysis of signals with gaps. *J. Math. Phys. 39*, 4091–4107.

[15] Genovese, C. R. and Wasserman, L. (2005). Confidence sets for nonparametric wavelet regression. *Ann. Statist.* **33**, 2, 698–729.

[16] Giraitis, L. and Surgailis, D. (1985). CLT and other limit theorems for functionals of Gaussian processes. *Z. Wahrsch. Verw. Gebiete* **70**, 2, 191–212.

[17] Giraitis, L. and Taqqu, M. S. (1998). Central limit theorems for quadratic forms with time-domain conditions. *Ann. Probab.* **26**, 1, 377–398.

[18] Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1**, 1, 15–29.

[19] Greenhall, C. A., Howe, D. A., and Percival, D. B. (1999). Total variance, an estimator of long-term frequency stability. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 46*, 1183–1191.

[20] Hall, P. and Penev, S. (2004). Wavelet-based estimation with multiple sampling rates. *Ann. Statist.* **32**, 5, 1933–1956.

[21] Ho, H. C. and Sun, T. C. (1987). A central limit theorem for noninstantaneous filters of a stationary Gaussian process. *J. Multivariate Anal.* **22**, 1, 144–155.

[22] Hosking, J. R. M. (1981). Fractional differencing. *Biometrika* **68**, 1, 165–176.

[23] Kalifa, J. and Mallat, S. (2003). Thresholding estimators for linear inverse problems and deconvolutions. *Ann. Statist.* **31**, 1, 58–109.

[24] Labat, D., Ababou, R., and Mangin, A. (2001). Introduction of wavelet analyses to rainfall/runoffs relationship for a karstic basin:the case of licq–atherey karstic system. *Ground Water 39*, 605–615.

[25] Lark, R. M. and Webster, R. (2001). Changes in variance and correlation of soil properties with scale and location: analysis using an adapted maximal overlap discrete wavelet transform. *European Journal of Soil Science 52*, 547–562.

[26] Massel, S. R. (2001). Wavelet analysis for processing of ocean surface wave records. *Ocean Engineering 28*, 957–987.

[27] Neumann, M. H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.* **25**, 1, 38–76.

[28] Palma, W. and Del Pino, G. (1999). Statistical analysis of incomplete long-range dependent data. *Biometrika* **86**, 4, 965–972.

[29] Pelgrum, H., Schmugge, T., Rango, A., Ritchie, J., and Kustas, B. (2000). Length-scale analysis of surface albedo, temperature, and normalized difference vegetation index in desert grassland. *Water Resources Research 36*, 1757–1766.

[30] Percival, D. B. (1995). On estimation of the wavelet variance. *Biometrika* **82**, 3, 619–631.

[31] Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge Series in Statistical and Probabilistic Mathematics, Vol. **4**. Cambridge University Press, Cambridge.

[32] Pichot, V., Gaspoz, J.-M., Molliex, S., Antoniadis, A., Busso, T., Roche, F., Costes, F., Quintin, L., Lacour, J.-R., and Barthélémy, J.-C. (1999). Wavelet transform to quantify heart rate variability and to assess its instantaneous changes. *Journal of Applied Physiology 86*, 1081–1091.

[33] Rybák, J. and Dorotovič, I. (2002). Temporal variability of the coronal green-line index(1947-1998). *Solar Physics 205*, 177–187.

[34] Serroukh, A., Walden, A. T., and Percival, D. B. (2000). Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series. *J. Amer.*

Statist. Assoc. **95**, 449, 184–196.

[35] Stoica, P., Larsson, E. G., and Li, J. (2000). Adaptive filter–bank approack to restoration and spectral analysis of gapped data. *The Astronomical Journal 163*, 2163–2173.

[36] Sweldens, W. (1997). The lifting scheme: a construction of second generation wavelets. *SIAM J. Math. Anal.* **29**, 2, 511–546.

[37] Torrence, C. and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* **79**, 1, 61–78.

[38] Toussoun, O. (1925). *Mémoire sur l'Histoire du Nil*Vol. **9.** .

[39] Tsakiroglou, E. and Walden, A. T. (2002). From Blackman–Tukey pilot estimators to wavelet packet estimators: a modern perspective on an old spectrum estimation idea. *Signal Processing 82*, 1425–1441.

[40] Vio, R., Strohmer, T., and Wamsteker, W. (2000). On the reconstruction of irregularly sampled time series. *Publications of the Astronomical Society of the Pacific 112*, 74–90.

[41] Whitcher, B. J., Byers, S. D., Guttorp, P., and Percival, D. B. (2002). Testing for homogeneity of variance in time series: long memory, wavelets and the Nile river. *Water Resources Research 38*, 1054–1070.

[42] Whitcher, B. J., Guttorp, P., and Percival, D. B. (2000). Wavelet analysis of covariance with application to atmospheric time series. *Journal of Geophysical Research 105*, 14941–14962.

[43] Žurbenko, I. G. (1986). *The Spectral Analysis of Time Series.* North-Holland Series in Statistics and Probability, Vol. **2**. North-Holland Publishing Co., Amsterdam. With a foreword by A. N. Kolmogorov.

[44] Zygmund, A. (1978). *Trigonometric Series.* Cambridge University Press, Cambridge. Reprinting of the 1968 version of the second edition with Volumes I and II bound together.

TABLE 1
Summary of Monte Carlo results for AR(1) process

	j					
	1	2	3	4	5	6
$\nu_X^2(\tau_j)$	0.0500	0.0689	0.1079	0.1585	0.1907	0.1710
mean of $\hat{u}_{X,r}(\tau_j)$	0.0502	0.0690	0.1084	0.1593	0.1911	0.1716
mean of $\hat{v}_{X,r}(\tau_j)$	0.0503	0.0692	0.1085	0.1592	0.1910	0.1715
$M_j^{-\frac{1}{2}} S_{u,j}^{\frac{1}{2}}(0)$	0.0087	0.0057	0.0104	0.0230	0.0347	0.0429
s.d. of $\hat{u}_{X,r}(\tau_j)$	0.0076	0.0055	0.0101	0.0204	0.0338	0.0431
mean of $M_j^{-\frac{1}{2}} \hat{S}_{u,j,r}^{\frac{1}{2}}(0)$	0.0071	0.0047	0.0086	0.0175	0.0288	0.0340
$M_j^{-\frac{1}{2}} S_{v,j}^{\frac{1}{2}}(0)$	0.0027	0.0047	0.0102	0.0207	0.0345	0.0428
s.d. of $\hat{v}_{X,r}(\tau_j)$	0.0025	0.0044	0.0099	0.0205	0.0337	0.0428
mean of $M_j^{-\frac{1}{2}} \hat{S}_{v,j,r}^{\frac{1}{2}}(0)$	0.0022	0.0039	0.0085	0.0173	0.0285	0.0339

TABLE 2
Summary of Monte Carlo results for FD($\frac{5}{6}$) process

	j					
	1	2	3	4	5	6
$\nu_X^2(\tau_j)$	0.2594	0.3078	0.4427	0.6831	1.0762	1.7050
mean of $\hat{v}_{X,r}(\tau_j)$	0.2599	0.3081	0.4421	0.6832	1.0771	1.7179
$M_j^{-\frac{1}{2}} S_{v,j}^{\frac{1}{2}}(0)$	0.0141	0.0203	0.0399	0.0857	0.1899	0.4281
s.d. of $\hat{v}_{X,r}(\tau_j)$	0.0129	0.0186	0.0386	0.0847	0.1877	0.4275
mean of $M_j^{-\frac{1}{2}} \hat{S}_{v,j,r}^{\frac{1}{2}}(0)$	0.0119	0.0168	0.0330	0.0704	0.1567	0.3489

DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
BOX 354322, SEATTLE, WA 98195, U.S.A.
E-MAIL: debashis@stat.washington.edu

APPLIED PHYSICS LABORATORY
UNIVERSITY OF WASHINGTON
BOX 355640, WA 98195, U.S.A.
E-MAIL: dbp@apl.washington.edu
URL: http://faculty.washington.edu/dbp/

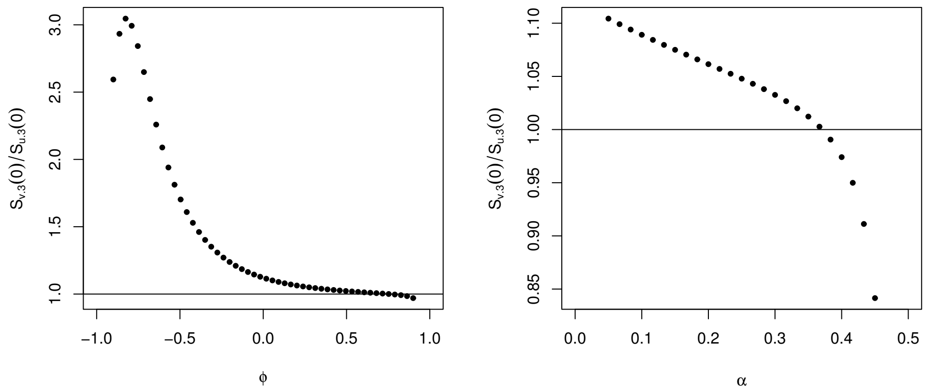

FIG 1. *Plot of asymptotic efficiency of $\hat{u}_X(\tau_3)$ with respect to $\hat{v}_X(\tau_3)$ under autoregressive (left) and fractionally differenced (right) models.*

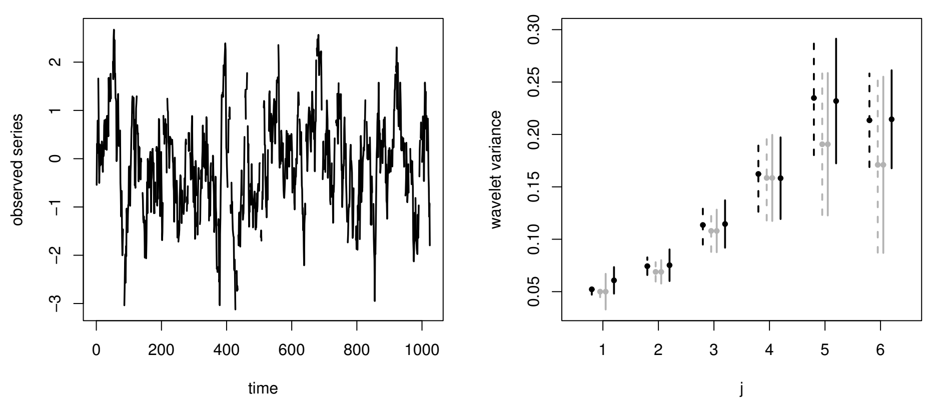

FIG 2. *Plot of a typical simulated gappy AR(1) time series and wavelet variances at various scales.*

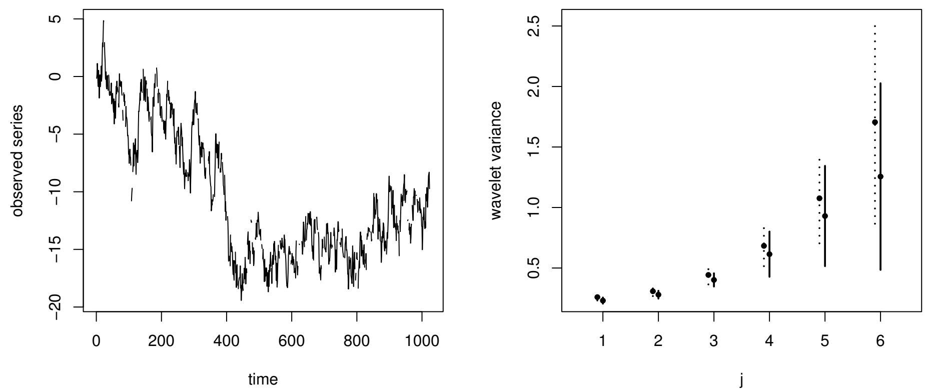

FIG 3. *Plot of a typical simulated gappy FD($\frac{5}{6}$) time series and wavelet variances at various scales. Solid lines indicate the estimated intervals while dotted lines indicate the true intervals.*

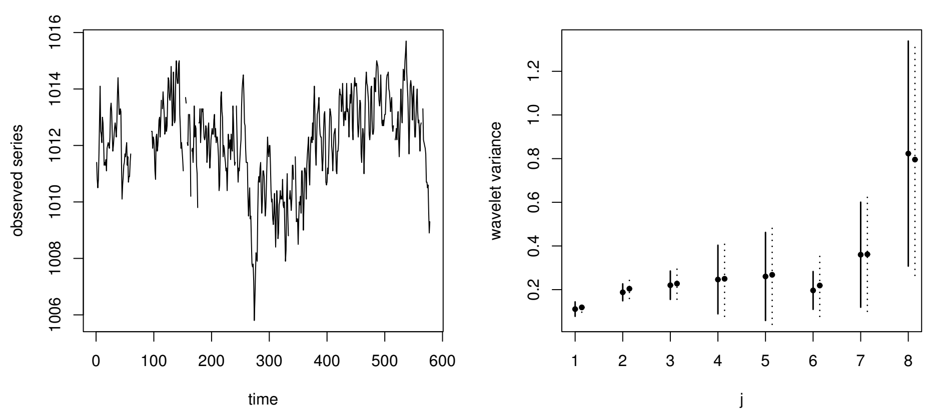

FIG 4. *Atmospheric pressure data (left) from NOAA's TAO buoy array and Haar wavelet variance estimates (right) for scales indexed by $j = 1, \ldots, 8$.*

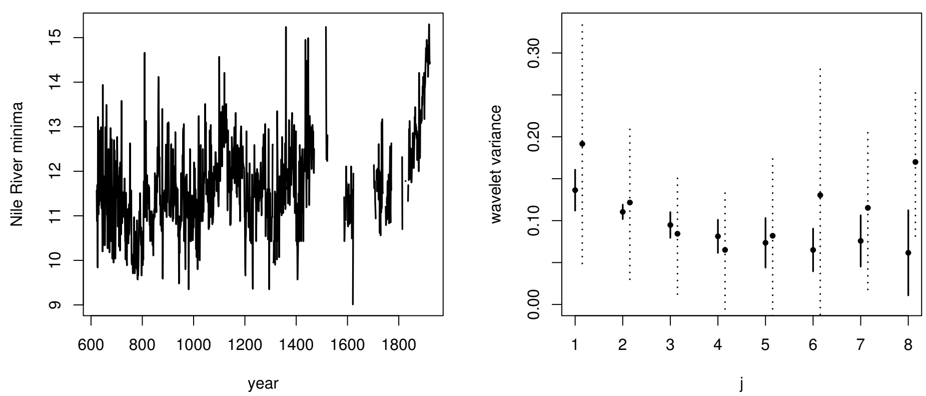

FIG 5. *Nile river minima (left) and Haar wavelet variance estimates (right) for scales indexed by $j = 1, \ldots, 8$.*