Wavelet Variance Analysis for Gappy Time Series
Wavelet Variance Analysis for Gappy Data

Debashis Mondal · Donald B. Percival

Received: date / Revised: date

Abstract 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.

Keywords Cumulant · Fractionally differenced process · Local stationarity · Nile River minima · Semi-variogram · TAO data

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

Supported by NSF Grant DMS–02–22115.

D. Mondal
Department of Statistics, University of Chicago, 5734 South University Avenue, Chicago, IL 60637, U.S.A.
E-mail: mondal@galton.uchicago.edu

D. B. Percival
Applied Physics Laboratory, University of Washington, Box 355640, WA 98195, U.S.A.
E-mail: dbp@apl.washington.edu.
processing including density estimation, deconvolution, edge detection, non-parametric regression and smooth estimation of evolutionary spectra. See, for example, Candès and Donoho (2002), Donoho et al. (1995), Donoho and Johnstone (1998), Genovese and Wasserman (2005), Hall and Penev (2004), Kalifa and Mallat (2003), Neumann and von Sachs (1997) 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 (Percival, 1995). 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 (Chiann and Morettin, 1998), the El Niño–Southern Oscillation (Torrence and Compo, 1998), soil variations (Lark and Webster, 2001), solar coronal activity (Rybáčk and Dorotovič, 2002), the relationship between rainfall and runoff (Labat et al., 2001), ocean surface waves (Massel, 2001), surface albedo and temperature in desert grassland (Pelgrum et al., 2000), heart rate variability (Pichot et al., 1999) and the stability of the time kept by atomic clocks (Greenhall et al, 1999).

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 variability on a particular scale. In mathematical notation,

\[
\text{var}(X_t) = \sum_{j=1}^{\infty} \nu^2_X(\tau_j),
\]

where \( \nu^2_X(\tau_j) \) is the wavelet variance associated with dyadic scale \( \tau_j = 2^{j-1} \); see equation (2.5) for the precise definition. Roughly speaking, \( \nu^2_X(\tau_j) \) is a measure of how much a weighted average of \( X_t \) over an interval of \( \tau_j \) differs from a similar average in an adjacent interval. A plot of \( \nu^2_X(\tau_j) \) against \( \tau_j \) 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 \( \tau_j \) exists and serves as a meaningful description of variability 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

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

The wavelet variance thus summarizes the information in the SDF using just one value per octave band $f \in [2^{-j-1}, 2^{-j}]$ 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, $\nu^2_X(\tau_j)$ is approximately proportional to $\tau_j^{-\alpha-1}$. The scalogram is a plot of $\log\{\nu^2_X(\tau_j)\}$ versus $\log(\tau_j)$. If it is approximately linear, a power law process is indicated, and the exponent $\alpha$ 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 (see Section 7.2 for an example of a time series for which the homogeneity assumption is questionable). 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. The typical situation is geophysics is that more observations are collected with the passage of time rather than by, e.g., sampling more finely over a fixed finite interval, so we do not consider procedures where more data are entertained via an in-fill mechanism.

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., Vio et al. (2000), 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 (Sweldens, 1997). 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 (Foster, 1996), and in the context of signals with continuous gaps (Frick and Tchamitchian, 1998). Related works address the problem of the spectral analysis of gappy data (Stoica et al., 2000). 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 order 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 $\sum_l h_{1,l}^2 = \frac{1}{2}$ (Daubechies, 1992). The transfer function for this filter, i.e., its discrete Fourier transform (DFT)

$$H_1(f) = \sum_{l=0}^{L-1} h_{1,l} e^{-i2\pi f l},$$

has a corresponding squared gain function by definition satisfying

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

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 $j$th level wavelet filter $h_{j,l}$ is defined as the inverse DFT of

$$H_j(f) = H(2^{j-1} f) \prod_{l=0}^{j-2} e^{-i2\pi 2^l f (L-1)} H\left(\frac{L}{2} - 2^l f\right). \quad (2.2)$$
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 $H_j$. Since $H_j(0) = 0$, it follows that

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

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

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

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

$$\nu^2_X(\tau_j) \equiv \int_{-1/2}^{1/2} H_j(f) S_X(f) \, df. \quad (2.5)$$

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) $s_{X,k} \equiv \text{cov}(X_t, X_{t+k})$, then we can rewrite the above equation as

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

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 $h_{j,l}$ (Craigmile and Percival, 2005). Alternatively, let $\gamma_{X,k} = \frac{1}{2} \text{var}(X_0 - X_k)$ denote the semi-varioam of $X_t$. Then the wavelet variance can be expressed as

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

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 $\nu^2_X(\tau_j)$ is given by

$$\hat{\nu}^2_X(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} W_{j,t}^2,$$
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 $H_j(f)S_X(f)$ and an ACVS to be denoted by $s_{j,k}$. The following theorem holds (Percival, 1995).

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

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

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

In practical applications, $A_j$ is estimated by

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

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 $\nu^2_X(\tau_j)$.

### 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 $X_0, \ldots, X_{N-1}$ of this process. Let $\delta_t$ be the corresponding gap pattern, assumed to be a portion of a binary stationary process independent of $X_t$. The random variable $\delta_t$ assumes the values of 0 or 1 with nonzero probabilities, with zero indicating that the corresponding realization for $X_t$ is missing. Define

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

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

$$\hat{\beta}^{-1}_{l,l'} = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \delta_{t-l}\delta_{t-l'}. $$
We assume that $\hat{\beta}_{j,l}^{-1} > 0$ 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, $\hat{\beta}_{j,l}^{-1}$ is a consistent estimator of $\beta_{j,l}^{-1}$, as $N \to \infty$.

Consider the following two statistics:

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

(3.1)

and

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

(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 3.1** 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:

$$
Z_{u,j,t} \equiv \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{t,l'} X_{t-l} X_{t-l'} \delta_{t-l} \delta_{t-l'},
$$

(4.1)

and

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

(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 $\hat{\beta}_{t,l'}$ replaced by $\hat{\beta}_{t,l'}$. 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 $V_{p_2,t}$ are absolutely summable and the higher order cumulants satisfy

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

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

Remark 4.1 Assumption 4.1 holds for a wide range of binary processes. For example, if $\delta_t$ 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 Brillinger (1981).

The following central limit theorems (Theorem 4.2 and 4.4) 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 called 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., Breuer and Major (1983), Giraitis and Surgailis (1985), Giraitis and Taqqu (1998), Fox and Taqqu (1987), Ho and Sun (1987) 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 $\delta_t$ 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 $\nu^2_X(\tau_j)$ 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 4.3 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 (Serroukh et al., 2000). 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 $\delta_t$ 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 $\nu^2_X(\tau_j)$. We use a
multitaper spectral approach (Serroukh et al., 2000). Let
\[ \tilde{Z}_{u,j,t} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{j,l'} X_{t-l} X_{t-l'} \delta_{t-l} \delta_{t-l'}, \quad t = L_j - 1, \ldots, N - 1. \]

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
\[ J_{u,j,k} = \sum_{t=0}^{M_j-1} \lambda_{k,t} \tilde{Z}_{u,j,t} + \sum_{t=0}^{M_j-1} \lambda_{k,t} \]
and
\[ \bar{u}_j = \frac{\sum_{k=0,2,\ldots}^{K-1} J_{u,j,k} \lambda_{k,+}}{\sum_{k=0,2,\ldots}^{K-1} \lambda_{k,+}^2}. \]

We estimate \( S_{u,j}(0) \) by
\[ \hat{S}_{u,j}(0) = \frac{1}{K} \sum_{k=0}^{K-1} (J_{u,j,k} - \bar{u}_j \lambda_{k,+})^2. \]

Following the recommendation of Serroukh et al. (2000), we choose \( K = 5 \) and set the bandwidth parameter so that the Slepian tapers are band-limited to the interval \( \left[-\frac{7}{2M_j}, \frac{7}{2M_j}\right] \). Previous Monte Carlo studies show that \( \hat{S}_{u,j}(0) \) performs well (Serroukh et al., 2000).

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

**Theorem 4.4** 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 \( \delta_t \) 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
\[ \tilde{Z}_{v,j,t} = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_j-1} h_{j,l} h_{j,l'} \hat{\beta}_{j,l'} (X_{t-l} - X_{t-l'})^2 \delta_{t-l} \delta_{t-l'}, \]
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

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

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 $\delta_t$ 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 $\phi$. Except for $\phi$ close to unity, $\hat{u}_X(\tau_j)$ outperforms $\hat{v}_X(\tau_j)$. When $\phi$ 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 intuitively 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

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

for $k = 1, 2, \ldots$; see, e.g., Granger and Joyeux (1980) and Hosking (1981). Here $\alpha < \frac{1}{2}$ is the long memory parameter, with $\alpha = 0$ corresponding to white noise and $\alpha$ 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 $\alpha$. As $\alpha$ 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 uniformly 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 4.4 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 \( \delta_t \) 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 A.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 \( r \)-th 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 A.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 \( M^{-1/2}_j S_{1/2}^u(\tau_j) \) and \( M^{-1/2}_j S_{1/2}^v(\tau_j) \). In particular, the ratios of the standard deviation of \( \hat{u}_{X,r}(\tau_j) \)'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 \( M^{-1/2}_j \hat{S}_{1/2}^u_{u,j,r}(0) \) and \( M^{-1/2}_j \hat{S}_{1/2}^v_{v,j,r}(0) \) to be 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 FDI(\( \text{FD}(\frac{5}{6}) \)) 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 \( \delta_t \) as before. In this example increments of \( X_t \) rather \( X_t \) 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 4.4. Table A.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 agreement with
its large sample approximation; and the sample mean of \( M_j^{-\frac{1}{2}} \hat{S}_{v,j}^2(0) \) is close to \( M_j^{-\frac{1}{2}} S_{v,j}^2(0) \), 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 \( d \)th 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 \( d \)th order stationary increment processes. First we note that Theorems 4.2 and 4.4 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 Daubchies 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 Daubchies wavelet variance for a general \( d \)th order backward stationary increment process.

Assume as in (2.4) that \( X_t \) for \( t \in \mathbb{Z} \) is a process with \( d \)th order stationary increments \( Y_t \). Let \( \mu_Y \) be the mean, \( s_{Y,k} \) the ACVS and \( \gamma_{Y,k} \) the semi-variogram of \( Y_t \). For \( L \geq 2d \), an expression for the Daubchies 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 \( r \)th order cumulative summation of the Daubchies 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 Craigmile and Percival, 2005). 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 \( d \)th order 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 in Section 3 with $b_{j,l,d}$ replacing $h_{j,t}$, the new gap pattern replacing $\delta_t$, 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. 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 $\nu^2_X(\tau_j)$ are then given by

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

and

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

The large sample properties of these estimators are given by obvious analogs to Theorems 4.2 and 4.4.

**Theorem 6.1** Suppose $X_t$ is a process whose $d$th order increments $Y_t$ are a stationary Gaussian process with square integrable SDF, and suppose $\delta_t$ is a strictly stationary binary process (independent of $X_t$) such that the derived binary process $\eta_t$ satisfies Assumption 4.1. Then, if $L \geq 2d$, $\hat{\nu}^2_X(\tau_j)$ is asymptotically normal with mean $\nu^2_X(\tau_j)$ and large sample variance $S_{d,u,j}(0)/M_j$, where $S_{d,u,j}$ is the SDF for $\sum_{l'=0}^{L_j-3} b_{j,l,l'} \rho_{l,l'} Y_{t-l} Y_{t-l'} \eta_{t-l} \eta_{t-l'}$.

**Theorem 6.2** Suppose $X_t$ is a process whose increments of order $d+1$ are a stationary Gaussian process with square integrable SDF, and suppose $\delta_t$ is as in the previous theorem. Then, if $L > 2d$, $\hat{\nu}^2_X(\tau_j)$ is asymptotically normal with mean $\nu^2_X(\tau_j)$ and large sample variance $S_{d,u,j}(0)/M_j$, where $S_{d,u,j}$ is the SDF for $-\frac{1}{2} \sum_{l'=0}^{L_j-3} b_{j,l,l'} \rho_{l,l'} (Y_{t-l} - Y_{t-l'})^2 \eta_{t-l} \eta_{t-l'}$.

The proofs of Theorems 6.1 and 6.2 are similar to those of, respectively, Theorems 4.2 and 4.4 and thus are omitted.

**Remark 6.3** Since each extra differencing produces more gaps, an estimate that requires less differencing will be more efficient. This is where the semivariogram estimator $\hat{\nu}^2_X(\tau_j)$ comes in handy. Let $C_t$ denote the backward differences of order $d-1$ $X_t$. 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 $L \geq 2d$

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

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 estimates of the true wavelet variance as long as $\beta_{l,l'} > 0$ for $l,l' = 0, \ldots, L_j - 1$ (for the financial example, this condition on $\beta$ 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 $\beta$ by taking the deterministic limit of $\beta$ 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 4.4 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 4.4, with some simplification because the gaps are deterministic (an alternative approach is to use Theorem 1 of Ho and Sun, 1987). 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 over a period of 578 days by the Tropical Atmospheric Ocean (TAO) buoy array operated by the National Oceanic and Atmospheric Administration. 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 $\tau_7$ and $\tau_8$, 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 $\tau_5$, which corresponds to a period of 32–64 days and captures the interseasonal variability. Note also that there is hardly any variability at scale $\tau_1$, although there is some at scales $\tau_3$, $\tau_4$, and $\tau_8$, indicating relatively important contributions to the variance due to disturbances at all but the very smallest scale. (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 (Toussoun, 1925). The rate of gaps is about 43% after year 1285. Several authors have previously analyzed the initial gap free segment (see, e.g., Beran, 1994, and Percival and Walden, 2000). The entire series, including the gappy part, has been analyzed based on a parametric state space model (Palma and Del Pino, 1999), 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 $\tau_1$ and $\tau_2$ before 715 than after (Whitcher et al., 2002). Therefore we restrict ourselves to the period 716–1921. Figure 5 plots wavelet variance estimates up to scale $\tau_8$ 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 $\tau_1$, $\tau_6$ and $\tau_8$, we see reasonably good agreement between estimates from the two segments. Substantial uncertainties due to the large number of gaps are reflected in the larger confidence intervals 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 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 $\tau_1$ and $\tau_2$ 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. Thus the apparent increase in variability at the largest scales from segment 716–1284 to 1286–1921 cannot be attributed just to the influence of the dam.

8 Discussion

In Section 3 we made the crucial assumption that, for a fixed $j$, $\hat{\beta}_{l,l'}^{-1} > 0$ when $l, l' = 0, \ldots, L_j - 1$. For small sample sizes, this condition might fail to hold. This situation arises mainly when half or more of the observations are missing and can be due to systematic periodic patterns in the gaps. For example, if $\delta_t$ alternates between zero and one, then $\hat{\beta}_{0,1}^{-1}$ is zero, reflecting the fact that the observed time series does not contain relevant information about $\nu_2^2(\tau_1)$. A methodology different from what we have discussed might be able to handle some gap patterns for which $\hat{\beta}_{l,l'}^{-1} = 0$. In particular, generalized prolate spheroidal sequences have been used to handle spectral estimation of irregularly sampled processes (Bronez, 1988). This approach in essence corresponds to the construction of special filters and could be used to construct approximations to the Daubechies filters when $\hat{\beta}_{l,l'}^{-1} = 0$.

Estimation of the SDF for gappy time series is a long-standing difficult problem. In Section 1 we noted that the wavelet variance provides a simple and useful estimator of the integral of the SDF over a certain octave band. In particular, the Blackman–Tukey pilot spectrum (Blackman and Tukey, 1958, Sec. 18) coincides with the Haar wavelet variance. Recently Tsakiroglou and Walden (2002) generalized this pilot spectrum by utilising the (maximal overlap) discrete wavelet packet transform. The result is an SDF estimator that is competitive with existing estimators. With a similar generalization, our wavelet variance estimator for gappy time series can be adapted to serve as an SDF estimator. Moreover, Nason et al. (2000) used shrinkage of squared wavelet coefficients to estimate spectra for locally stationary processes. In the same vein, we can apply wavelet shrinkage to the $Z_{u,j,t}$ or $Z_{v,j,t}$ processes to estimate time-varying spectra when the original time series is observed with gaps.
Finally we note a generalization of interest in the analysis of multivariate gappy time series. Given two time series \( X_{1,t} \) and \( X_{2,t} \), the wavelet cross covariance yields a scale-based analysis of the cross covariance between the two series in a manner similar to wavelet variance analysis (for estimation of the wavelet cross covariance, see Whitcher et al., 2000, and the references therein). The methodology described in this paper can be readily adapted to estimate the wavelet cross covariance for multivariate time series with gaps.

### A Proofs

We first need the following propositions and lemmas. To avoid a triviality, we assume throughout that \( \{X_t\} > 0 \).

**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 \( U_t \equiv [X_{t-k}X_{t-k'}, X_{t-l}X_{t-l'}]^T \), for any choice of \( k, k', l \) and \( l' \), has a spectral matrix \( S_U \) that is continuous.

**Proof** Using the Isserlis theorem, we have

\[
\text{cov} \left( X_{t-k}X_{t-k'}, X_{t-l}X_{t-l'} \right) = s_{X,k-l} s_{X,k'-l'} + s_{X,k-l'} s_{X,k'-l}.
\]

By the Fourier transform we obtain

\[
S_{k,k',l,l'}(f) = e^{i2\pi f(k'-l')} \int_{-1/2}^{1/2} e^{i2\pi f(k-l-k'+l')} S_X(f)S_X(f-f') df' + e^{i2\pi f(k-l)} \int_{-1/2}^{1/2} e^{i2\pi f(k+l-k'-l')} S_X(f)S_X(f-f') df'.
\]

Because \( \exp(i2\pi f(k'-l')) \) is a continuous function of \( f \), we can establish the continuity of the first term above if we can show that

\[
A_{k,k',l,l'}(f) \equiv \int_{-1/2}^{1/2} e^{i2\pi f(k-l-k'+l')} S_X(f)S_X(f-f') df'
\]

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

\[
\left| A_{k,k',l,l'}(f + \rho) - A_{k,k',l,l'}(f) \right| = \left| \int_{-1/2}^{1/2} e^{i2\pi f(k-l-k'+l')} S_X(f') \left[ S_X(f + \rho - f') - S_X(f - f') \right] df' \right| \\
\leq \left( \int_{-1/2}^{1/2} S_X^2(f') df' \right)^{1/2} \left( \int_{-1/2}^{1/2} \left| S_X(f + \rho - f') - S_X(f - f') \right|^2 df' \right)^{1/2}.
\]

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

\[
\int_{-1/2}^{1/2} \left| S_X(f + \rho - f') - S_X(f - f') \right|^2 df' \to 0 \quad \text{as} \quad \rho \to 0
\]

by Lemma 1.11, p. 37 of Zygmund (1978). Hence \( A_{k,k',l,l'} \) and \( S_{k,k',l,l'} \) are continuous. \( \square \)
Proposition A.2  Let $h_{j,t}$ 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

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

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

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

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

Proposition A.3  Let $X_t$ be a real-valued zero mean Gaussian process with ACVS $s_X$ and SDF $S_X$ satisfying

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

Then the bivariate process $U_t = [\{\frac{1}{2}(X_{t-k} - X_{t-k'})^2, \frac{1}{2}(X_{t-l} - X_{t-l'})^2\}]^T_T$, for any choice of $k,k',l$ and $l'$, has a spectral matrix $S_U$ that is continuous.

Proof  The proof is similar to that of Proposition A.1. \hfill \Box

Proposition A.4  Let $h_{j,t}$ 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 $S_U$ in that proposition. Then

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

Proof  The proof is similar to that of Proposition A.2. \hfill \Box

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

$$U_{l',l,t} = \psi_{l',l} + \int_{-1/2}^{1/2} e^{i2\pi ft} dU_{l',l}(f)$$

$$V_{l',l,t} = \omega_{l',l} + \int_{-1/2}^{1/2} e^{i2\pi ft} dV_{l',l}(f)$$

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 $U_{k,k',l}$ and $U_{l,l',t}$ and between $V_{k,k',l}$ and $V_{l,l',t}$. Let $\alpha_{l',l}$ be fixed real numbers. Define

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

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

$$S_Q(f) = \sum_{k,k'} \sum_{l,l'} \alpha_{k,k'} \alpha_{l,l'} \left[ \psi_{k,k'} \psi_{l,l'} G_{k,k',l,l'}(f) + \omega_{k,k'} \omega_{l,l'} S_{k,k',l,l'}(f) \right] + S * G_{k,k',l,l'}(f),$$

where

$$S * G_{k,k',l,l'}(f) = \int_{-1/2}^{1/2} \mathcal{G}_{k,k',l,l'}(f - f') S_{k,k',l,l'}(f') df'.$$
Proof A full proof is straightforward, but tedious. The key steps are to note that
\[
\text{cov} \{ Q_t, Q_{t+m} \} = \sum_{k,k',l,l'} a_{k,k'} a_{l,l'} \text{cov} \{ U_{k,k',t} V_{k,k',t}, U_{l,l',t+m} V_{l,l',t+m} \}, \quad (A.2)
\]
to use the spectral representations of \( U_{l,l',t} \) and \( V_{l,l',t} \) and the independence assumption to obtain
\[
\text{cov} \{ U_{k,k',t} V_{k,k',t}, U_{l,l',t+m} V_{l,l',t+m} \}
= \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} e^{2\pi f \delta t} \left[ \psi_{k,k'} \psi_{l,l'} G_{k,k',l,l'}(f) + \omega_{k,k'} \omega_{l,l'} S_{k,k',l,l'}(f) \right] + \int_{-1/2}^{1/2} G_{k,k',l,l'}(f-f') S_{k,k',l,l'}(f') df' df,
\]
and to plug the above formula into equation (A.2). \( \square \)

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 \( \delta_t \) 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 \( U_{l,l',t} = X_{t-l} X_{t-l'} \). Then \( U_t = [U_{k,k',t}, U_{l,l',t}]^T \), \( V_{l,l',t} = \delta_{t-l} \delta_{t-l'} \) and \( a_{l,l'} = h_{j,l} h_{j,l'} \delta_{l-l'} \). By Proposition A.1, \( U \) has a continuous cross spectrum \( S_{k,k',l,l'} \). Then by Lemma A.5, the SDF \( S_{u,j,t} \) of \( Z_{u,j,t} \) is given by the right-hand side of equation (A.2), where \( \psi_{l,l'} = E X_{t-l} X_{t-l'} \), \( \omega_{l,l'} = E \delta_{t-l} \delta_{t-l'} = \delta_{l-l'} \) 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
\[
\sum_{k,k'} \sum_{l,l'} a_{k,k'} a_{l,l'} \omega_{k,k'} \omega_{l,l'} S_{k,k',l,l'}(0) = \sum_{k,k'} \sum_{l,l'} h_{j,k} h_{j,l} h_{j,l'} h_{j,l'} S_{k,k',l,l'}(0) > 0.
\]
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 \( \sum_{k,k'} \sum_{l,l'} a_{k,k'} a_{l,l'} S_{k,k',l,l'}(f) \) are nonnegative because \( G_{k,k',l,l'} \) and \( S_{k,k',l,l'} \) are entries of spectral density matrices. Hence \( S_{u,j}(0) > 0 \). \( \square \)

Proposition A.7 Let \( X_t \) be a real-valued Gaussian stationary process with zero mean, and SDF \( S_X \) that satisfies
\[
\int_{-1/2}^{1/2} \sin^4(\pi f) S_X^2(f) df < \infty.
\]
Let \( \delta_t \) 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 (4.2). Then \( Z_{u,j,t} \) 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 Brillinger (1981), p. 21.

Theorem A.8 Consider a two way array of random variables (RVs) \( \Theta_{i,j}, j = 1, \ldots, J \) and \( i = 1, \ldots, n \). Consider the n RVs \( T_i = \prod_{j=1}^{J} \Theta_{i,j} \) for \( i = 1, \ldots, n \). Then the joint cumulant of \( T_1, \ldots, T_n \) is given by the formula
\[
\text{cum}(T_1, \ldots, T_n) = \sum_{\chi} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r)
\]
where the summation is over all indecomposable partitions \( \chi = \chi_1 \cup \cdots \cup \chi_r \) of the (not necessarily rectangular) two way table

\[
(1, 1) \cdots (1, J_1) \\
\vdots \\
(n, 1) \cdots (n, J_n).
\]

(A.3)

Next we need the following lemmas.

**Lemma A.9** Assume that \( X_i \) satisfies the conditions stated in Theorem 4.2. Let \( U_{p,t} = X_{t-1}X_{t-i'} \) and \( E_{p,t} = \psi_p \) where \( p = (l, l') \). Then for \( n \geq 3 \) and fixed \( p_1, \ldots, p_n \),

\[
\sum_{t_1, \ldots, t_n} |\text{cum}(U_{p_1, t_1} - \psi_{p_1}, \ldots, U_{p_n, t_n} - \psi_{p_n})| = o(N^{n/2}),
\]

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,

\[
\text{cum}(U_{p_1, t_1} - \psi_{p_1}, \ldots, U_{p_n, t_n} - \psi_{p_n}) = \text{cum}(U_{p_1, t_1}, \ldots, U_{p_n, t_n}).
\]

Consider the \( n \times 2 \) table of RVs given by

\[
\Theta_{1,1} = X_{t_1 - t_1} \quad \Theta_{1,2} = X_{t_1 - t_1'} \\
\vdots \\
\Theta_{n,1} = X_{t_n - t_n} \quad \Theta_{n,2} = X_{t_n - t_n}'.
\]

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 \( \text{cum}(U_{p_1, t_1}, \ldots, U_{p_n, t_n}) \). 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 \( \sum_{t_1, \ldots, t_n} \text{cum}(U_{p_1, t_1}, \ldots, U_{p_n, t_n}) \equiv \sum \chi I_{U,M}(\chi) \) with

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

Since \( n \) is fixed and the number of indecomposable partitions depends only on \( n \), it then suffices to show that \( I_{U,M}(\chi) = o(N^{n/2}) \) for any fixed \( \chi \). As \( \chi \) 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 \( \eta_k \) takes values of 1 or 2 for \( k = 1, \ldots, n \) and \( \xi_k = 3 - \eta_k \). We set, for \( k = 1, \ldots, n \),

\[
e_k = \begin{cases} 
(l_{k+1} \mod n) - l_k, & \text{if } \xi_{k+1} \mod n = \eta_k = 1 \\
(l_{k+1} \mod n) - l_k, & \text{if } \xi_{k+1} \mod n = 2, \eta_k = 1 \\
(l_{k+1} \mod n) - l_k', & \text{if } \xi_{k+1} \mod n = 1, \eta_k = 2 \\
(l_{k+1} \mod n) - l_k', & \text{if } \xi_{k+1} \mod n = \eta_k = 2 
\end{cases}
\]

Then we can write

\[
\text{cum}(\Theta_{i,j} : (i, j) \in \chi_k) = s_{X, t_k+e_k - t_k - e_k}
\]

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

\[
I_{U,M}(\chi) \equiv \sum_{t_1, \ldots, t_n} s_{X, t_1 - t_n - e_n} \prod_{l=1}^{n-1} s_{X, t_{l+1} - t_l - e_l}.
\]

(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 magnitude 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 repeatedly use the Cauchy-Schwarz inequality to obtain

$$I''_{U,M}(\chi)$$

$$\leq \sum_{t_1,\ldots,t_n} s_{X,t_1-t_n} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i}$$

$$= \sum_{t_1,\ldots,t_n} \prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i} \prod_{i=1}^{n} s_{X,t_1-t_n} s_{X,t_n-t_{n-1}}$$

$$\leq \sum_{t_1,\ldots,t_n} \prod_{i=1}^{n-2} s_{X,t_{i+1}-t_i} \left( \sum_{t_n} s_{X,t_1-t_n} \right)^2 \left( \sum_{t_n} s_{X,t_n-t_{n-1}} \right)^2$$

$$= \sum_{t_1,\ldots,t_n} \prod_{i=1}^{n-3} s_{X,t_{i+1}-t_i} \left( \sum_{t_n} s_{X,t_1-t_n} \right)^2 \left( \sum_{t_n} s_{X,t_n-t_{n-1}} \right)^2$$

$$\leq \left( \sum_{t_1,\ldots,t_n} \prod_{i=1}^{n} s_{X,t_{i+1}-t_i} \right)^{\frac{1}{2}} \prod_{i=1}^{n-1} \left( \sum_{t_n} s_{X,t_1-t_n} \right)^{\frac{1}{2}}$$

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

$$\sum_{t_1,\ldots,t_n} s_{X,t_1-t_n-e_{i-1}} \leq \text{constant} \sum_{|\tau| > K} \sum_{t_n} s_{X,\tau} = \text{constant} M \sum_{|\tau| > K} s_{X,\tau},$$

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

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

and the required result follows by choosing $K = |\log(M)|$. \hfill \Box

**Lemma A.10** Assume that $X_t$ satisfies the conditions stated in Theorem 4.4. Let $U_{p,t} = -\frac{1}{2} (X_{t-l} - X_{t-l})^2$ and $E_{U_{p,t}} = \psi_{p,t}$, in which $p = (l, l')$. Then for $n \geq 3$ and fixed $p_1, \ldots, p_n$,

$$\sum_{l_1,\ldots,l_n} |\text{cum}(U_{p_1,l_1} - \psi_{p_1}, \ldots, U_{p_n,l_n} - \psi_{p_n})| = o(M^{n/2}),$$
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 \( X_{t-i} - X_{t-i} \) and \(-\frac{1}{2}(X_{t-i} - X_{t-i})^2\), where the Gaussian process \( X_{t-i} - X_{t-i} \) 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

\[
\kappa_0(p_1, \ldots, p_n, t_1, \ldots, t_n) = \text{cum}(U_{p_1, t_1} - \psi_{p_1}, \ldots, U_{p_n, t_n} - \psi_{p_n}).
\]

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 \( \kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_i) \) is bounded and satisfies

\[
\sum_{t_1, \ldots, t_i} \kappa_n(p_1, \ldots, p_n, t_1, \ldots, t_i) = o\left(M^{\frac{1}{2}(i+1)}\right), \quad i = 1, 2, \ldots, n.
\]  

(A.6)

**Proof** We retain all the notation of Lemma A.9. Thus

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

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

\[
\sum_{t_{\lambda_1}, \ldots, t_{\lambda_i}} M^{-\frac{1}{2}(i-1)}\pi_{X,t_1-t_n} - e_n \prod_{i=1}^{n-1} \pi_{X,t_{i+1}-t_{i} - e_i}
\]

(A.7)

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

Consider \( i = 1 \). WLOG assume \( \lambda_1 = n \). Then

\[
\sum_{t_n} \pi_{X,t_1-t_n} - e_n \prod_{i=1}^{n-1} \pi_{X,t_{i+1}-t_{i} - e_i}
\]

\[
\leq \prod_{i=1}^{n-2} \pi_{X,t_{i+1}-t_i - e_i} \left( \sum_{t_n} \pi_{X,t_1-t_n} - e_n \right)^{\frac{1}{2}} \left( \sum_{t_n} \pi_{X,t_{n-1}-t_{n-1} - e_{n-1}} \right)^{\frac{1}{2}},
\]

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

\[
\sum_{t_{n-1}, t_n} M^{-\frac{1}{2}}\pi_{X,t_1-t_n} - e_n \prod_{i=1}^{n-1} \pi_{X,t_{i+1}-t_{i} - e_i}
\]

\[
\leq \sum_{t_{n-1}} M^{-\frac{1}{2}} \prod_{i=1}^{n-2} \pi_{X,t_{i+1}-t_i - e_i} \left( \sum_{t_n} \pi_{X,t_1-t_n} - e_n \right)^{\frac{1}{2}} \left( \sum_{t_n} \pi_{X,t_{n-1}-t_{n-1} - e_{n-1}} \right)^{\frac{1}{2}}
\]

\[
\leq \prod_{i=1}^{n-3} \pi_{X,t_{i+1}-t_i - e_i} \left( \sum_{t_n} \pi_{X,t_1-t_n} - e_n \right)^{\frac{1}{2}} \left( \sum_{t_n} \pi_{X,t_{n-1}-t_{n-1} - e_{n-1}} \right)^{\frac{1}{2}}
\]

\[
(M^{-1} \sum_{t_{n-1}, t_n} \pi_{X,t_1-t_n} - e_n \prod_{i=1}^{n-1} \pi_{X,t_{i+1}-t_{i} - e_i})^{\frac{1}{2}}
\]
Clearly the above expression is bounded because \( \sum_t s_{X,t}^2 \) is so and therefore
\[
\lim_{M \to \infty} M^{-1} \sum_{t_{n-1},t_n} s_{X,t_{n-1}-t_n-\epsilon_{n-1}} = \sum_{\tau} s_{X,\tau-\epsilon_{n-1}} < \infty.
\]
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
\[
\sum_{t_{n-2},t_n} s_{X,t_{n-1}-t_n-\epsilon_{n-1}} \prod_{i=1}^{n-1} s_{X,t_{i+1}-t_i-\epsilon_i} \\
\leq \prod_{i=1}^{n-4} s_{X,t_{i+1}-t_i-\epsilon_i} \left( \sum_{t_n} s_{X,t_{n-1}-t_n-\epsilon_n} \right)^2 \left( \sum_{t_n} s_{X,t_{n-1}-t_n-\epsilon_n} \right)^{\frac{1}{4}} \left( \sum_{t_{n-2}} s_{X,t_{n-2}-t_{n-3}-\epsilon_{n-3}} \right)^{\frac{1}{2}}
\]
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 \( \chi \) is an indecomposable partition, there can be at most \( t - 1 \) pairs of \( \lambda_i \), 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 \( t_{\lambda_j} \) for \( j = 1, \ldots, i \) in the left hand side of (A.7) and repeated use of Cauchy–Schwarz inequality will give rise to the \( (i - 1) \) terms
\[
M^{-1} \sum_{t_{\lambda_j}, t_{\lambda_{j+1}}} s_X^{\lambda_j \lambda_{j+1}} f_{t_{\lambda_j}, t_{\lambda_{j+1}} - \tau} - \epsilon_{\lambda_j} \leq 0
\]
where \( \epsilon_{\lambda_j} \) is so and therefore bounded. Therefore, 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 \( U_{p,t} = X_{t,-1}X_{t,-1}, V_{p,t} = \delta_{t,-1}\delta_{t,-1} \) and \( \alpha_p = h_{t_i} h_{t_i+1} \beta_{t_i+1} \), where \( p = (l, l') \). Take \( Q_t = \sum_{p} a_p \langle U_{p,t}V_{p,t} - \psi_p \omega_p \rangle \) as in Lemma A.5. Note that \( \tilde{\nu}_X(\tau_j) = \nu_X(\tau_j) \) is the average of \( Q_t \) over \( L_j-1 \leq t \leq N - 1 \) with \( \beta_{t_i-l'} \) replaced by its consistent estimate \( \beta_{t_i-l'} \). 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 \( \tilde{\nu}_X(\tau_j) \) is asymptotically normal. We use Zuberenko (1986), p. 2, to write the log of the characteristic function of \( R \) as
\[
\log F(\lambda) = \sum_{n=1}^{\infty} \frac{i^n \lambda^n}{n!} \sum_{t_1, \ldots, t_n} B_n(t_1, \ldots, t_n) M^{-n/2},
\]
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, \( B_1(t) = 0 \). By Proposition A.6, the autocovariances \( s_{Q,\tau} \) of \( Q_t \) are absolutely summable and \( M^{-1} \sum_{t_1} s_{Q,\tau} B_2(t_1, t_2) \to \sum_{\tau} s_{Q,\tau} = S_Q(0) > 0 \). 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 Brillinger (1981), we break up the nth order cumulant as follows:
\[
B_n(t_1, \ldots, t_n) = \sum_{p_1} \cdots \sum_{p_n} \alpha_{p_1} \cdots \alpha_{p_n} \cum(U_{p_1, t_1} V_{p_1, t_1} - \psi_{p_1} \omega_{p_1}, \ldots, U_{p_n, t_n} V_{p_n, t_n} - \psi_{p_n} \omega_{p_n}).
\]
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 Brillinger (1981) again, we have

$$
\text{cum}(U_{p_1,t_1}V_{p_1,t_1} - \psi_{p_1}\omega_{p_1}, \ldots, U_{p_n,t_n}V_{p_n,t_n} - \psi_{p_n}\omega_{p_n}) = \sum_{c_1,\ldots,c_n} \text{cum}(D_{c_1,p_1,t_1}, \ldots, D_{c_n,p_n,t_n}),
$$

where each $c_i$ ranges from 1 to 3. Therefore, it suffices to show that, for fixed $p_1, \ldots, p_n$ and $c_1, \ldots, c_n$, $\text{cum}(D_{c_1,p_1,t_1}, \ldots, D_{c_n,p_n,t_n}) = o(M^{n/2})$. Since the cumulant of n RVs is invariant under a reordering of the RVs, assume $c_1 = c_2 = \cdots = c_m = 1, c_{m+1} = c_{m+2} = \cdots = c_{m' + 1} = 2, c_{m' + 2} = \cdots = c_n = 3$, 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 $\text{cum}(D_{1,p_1,t_1}, \ldots, D_{3,p_n,t_n})$ is proportional to $\sum_{\chi} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) = O(M^{n/2})$.

(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 $\text{cum}(\Theta_{i,j} : (i,j) \in \chi_k) = 0$. 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 $\chi_k$ 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

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

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

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

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

**step 3:** Finally consider $m \geq 1$. Assume that $\chi_1, \ldots, \chi_q$ partition the random variables $\{\Theta_{1,1}, \ldots, \Theta_{m',1}\}$ (these are all $U_{p_i,t_i} - \psi_{p_i}$) and that $\chi_{q + 1}, \ldots, \chi_r$ partition the random variables $\{\Theta_{m,2}, \Theta_{m',1,2}, \ldots, \Theta_{n,1}\}$ (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 $\chi_k$ for $k = 1, \ldots, q$, we set $0 = \rho_0 \leq \rho_1 \leq \cdots \leq \rho_q = m$, $m = \sigma_0 \leq \sigma_1 \leq \cdots \leq \sigma_q = m'$ 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
\[
\sum_{t_{\sigma_{k-1}+1} \cdots t_{\sigma_k}} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_k) \\
= \sum_{t_{\sigma_{k-1}+1} \cdots t_{\sigma_k}} \kappa_{\rho_k+\sigma_k-\rho_{k-1}-\sigma_{k-1}}(p_i, t_i : (i,1) \in \chi_k) \\
= M^{\frac{1}{2}}(\sigma_k-\sigma_{k-1})^+ \kappa_{\rho_k+\sigma_k-\rho_{k-1}-\sigma_{k-1}}(p_i, (i,1) \in \chi_k, t_{\rho_{k-1}+1} \cdots t_{\rho_k}).
\]
Now boundedness of $\kappa_{\rho_k+\sigma_k-\rho_{k-1}-\sigma_{k-1}}(p_i, (i,1) \in \chi_k, t_{\rho_{k-1}+1} \cdots t_{\rho_k})$ yields
\[
M^{-\frac{1}{2}}(m'-m-1) \sum_{t_1 \cdots t_m} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
\propto \sum_{t_1 \cdots t_m} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_q+1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) = o(M^{n/2}).
\]
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 $\text{cum}(\Theta_{i,j} : (i,j) \in \chi_k)$ for $k' = 1, \ldots, q$, we obtain
\[
\sum_{t_1 \cdots t_n} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
\leq C_0 \sum_{t_1 \cdots t_n} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_q+1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) = o(M^{n/2}).
\]
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 $\chi_{q+i}$ for $i = 1, \ldots, n$. We sum over $t_{m+1} \cdots t_n$ to obtain
\[
\sum_{t_1 \cdots t_m} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
\leq \text{constant} \sum_{t_1 \cdots t_m} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_q) \\
\sum_{t_1 \cdots t_m} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_{q+n-m+1}) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_r) \\
\leq \text{constant} \sum_{t_1 \cdots t_m} \text{cum}(\Theta_{i,j} : (i,j) \in \chi_1) \cdots \text{cum}(\Theta_{i,j} : (i,j) \in \chi_q) = o(M^{n/2}).
\]
In the above derivation we need the fact that $\sum_{t_i} \text{cum}(V_{p_i,t_i}, V_{p_i,t_j})$ 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 4.4. Take $U_{p,t} = -\frac{1}{2}(X_{t-l} - X_{t-l'})^2$, $V_{p,t} = \delta_{t-l}\delta_{l-l'}$ and $a_p = h_{ij}h_{jl}h_{ij}h_{jl}$, where $p = (l, l')$. Use Lemma 3 in place of Lemma 2 and complete the proof as in Theorem 2.1 by checking all the steps. □

Acknowledgements We thank Peter Guttorp and Chris Bretherton for discussions.
References


### Table A.1 Summary of Monte Carlo results for AR(1) process

<table>
<thead>
<tr>
<th>j</th>
<th>$\nu_X^2(\tau_j)$</th>
<th>mean of $\hat{u}_{X,r}(\tau_j)$</th>
<th>mean of $\hat{v}_{X,r}(\tau_j)$</th>
<th>$M_j^{-1}\hat{S}_{u,j}^+(0)$</th>
<th>s.d. of $\hat{u}_{X,r}(\tau_j)$</th>
<th>mean of $M_j^{-1}\hat{S}_{u,j,r}^+(0)$</th>
<th>s.d. of $\hat{v}_{X,r}(\tau_j)$</th>
<th>mean of $M_j^{-1}\hat{S}_{v,j,r}^+(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0500</td>
<td>0.0690</td>
<td>0.1085</td>
<td>0.0087</td>
<td>0.0055</td>
<td>0.0027</td>
<td>0.0141</td>
<td>0.0119</td>
</tr>
<tr>
<td>2</td>
<td>0.0689</td>
<td>0.1084</td>
<td>0.1593</td>
<td>0.0104</td>
<td>0.0101</td>
<td>0.0044</td>
<td>0.0047</td>
<td>0.0168</td>
</tr>
<tr>
<td>3</td>
<td>0.1079</td>
<td>0.1592</td>
<td>0.1910</td>
<td>0.0230</td>
<td>0.0204</td>
<td>0.0039</td>
<td>0.0039</td>
<td>0.0330</td>
</tr>
<tr>
<td>4</td>
<td>0.1585</td>
<td>0.1916</td>
<td>0.1715</td>
<td>0.0347</td>
<td>0.0338</td>
<td>0.0175</td>
<td>0.0205</td>
<td>0.0704</td>
</tr>
<tr>
<td>5</td>
<td>0.1907</td>
<td>0.1911</td>
<td>0.1710</td>
<td>0.1716</td>
<td>0.1715</td>
<td>0.1710</td>
<td>0.1715</td>
<td>0.1715</td>
</tr>
<tr>
<td>6</td>
<td>0.1710</td>
<td>0.1716</td>
<td>0.1715</td>
<td>0.1715</td>
<td>0.1715</td>
<td>0.1715</td>
<td>0.1715</td>
<td>0.1715</td>
</tr>
</tbody>
</table>

### Table A.2 Summary of Monte Carlo results for FD($\frac{5}{6}$) process

<table>
<thead>
<tr>
<th>j</th>
<th>$\nu_X^2(\tau_j)$</th>
<th>mean of $\hat{v}_{X,r}(\tau_j)$</th>
<th>$M_j^{-1}\hat{S}_{v,j}^+(0)$</th>
<th>s.d. of $\hat{v}_{X,r}(\tau_j)$</th>
<th>mean of $M_j^{-1}\hat{S}_{v,j,r}^+(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2594</td>
<td>0.2023</td>
<td>0.0141</td>
<td>0.0203</td>
<td>0.0129</td>
</tr>
<tr>
<td>2</td>
<td>0.3078</td>
<td>0.3099</td>
<td>0.0141</td>
<td>0.0399</td>
<td>0.0186</td>
</tr>
<tr>
<td>3</td>
<td>0.4427</td>
<td>0.0857</td>
<td>0.0141</td>
<td>0.0857</td>
<td>0.0386</td>
</tr>
<tr>
<td>4</td>
<td>0.6831</td>
<td>0.1899</td>
<td>0.0141</td>
<td>0.1899</td>
<td>0.0847</td>
</tr>
<tr>
<td>5</td>
<td>1.0762</td>
<td>0.4281</td>
<td>0.0141</td>
<td>0.4281</td>
<td>0.1877</td>
</tr>
<tr>
<td>6</td>
<td>1.7050</td>
<td>0.4281</td>
<td>0.0141</td>
<td>0.4281</td>
<td>0.4275</td>
</tr>
</tbody>
</table>
Fig. A.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. A.2 Plot of a typical simulated gappy AR(1) time series and wavelet variances at various scales.
Fig. A.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. A.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. A.5 Nile River minima (left) and Haar wavelet variance estimates (right) for scales indexed by $j = 1, \ldots, 8$. 