The Local Partial Autocorrelation Function and Some Applications[*]

Rebecca Killick

Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK
e-mail: r.killick@lancs.ac.uk

Marina I. Knight

Department of Mathematics, University of York, York, YO10 5DD, UK
e-mail: marina.knight@york.ac.uk

Guy P. Nason

School of Mathematics, University Walk, University of Bristol, BS8 1TW, UK
e-mail: g.p.nason@bris.ac.uk

and

Idris A. Eckley

Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK
e-mail: i.eckley@lancs.ac.uk

Abstract: The classical regular and partial autocorrelation functions are powerful tools for stationary time series modelling and analysis. However, it is increasingly recognized that many time series are not stationary and the use of classical global autocorrelations can give misleading answers. This article introduces two estimators of the local partial autocorrelation function and establishes their asymptotic properties. The article then illustrates the use of these new estimators on both simulated and real time series. The examples clearly demonstrate the strong practical benefits of local estimators for time series that exhibit nonstationarities.

Keywords and phrases: locally stationary time series, integrated local wavelet periodogram, wavelets, practical estimation, Haar cross-correlation wavelet.

Received April 2019.

1. Introduction

Much work has been undertaken to develop both theory and methods for the use of the autocorrelation and partial autocorrelation for mean zero second-order stationary time series. See, for example, Priestley (1983), Brockwell and

Davis (1991) or Chatfield (2003). For stationary time series, both autocorrelations are fundamental for eliciting second-order structure and are particularly useful for subsequent modelling and prediction. Unfortunately, in many applied situations, for example neurophysiology (Fiecas and Ombao, 2016) or biology (Hargreaves et al., 2018), the stationarity assumption is not tenable and, hence, use of the classical stationary-based autocorrelations is highly questionable. Indeed, it is not possible for a, potentially complicated, time-varying parameter to be adequately summarised by a single coefficient. Before practical analysis, one should therefore attempt to assess whether the series is stationary or not. Many techniques and software packages exist that enable such assessment, see reviews in Dahlhaus (2012) or Cardinali and Nason (2018) or newer techniques that measure, rather than test, degree of nonstationarity, e.g. Das and Nason (2016).

A large literature on nonstationary time series modelling has developed since the 1950s. See, for example, Page (1952), Silverman (1957), Whittle (1963), Priestley (1965), Tong (1974) and Dahlhaus (1997). Alternative model forms including the piecewise stationary time series of Adak (1998); the wavelet models of Nason, von Sachs and Kroisandt (2000); and the SLEX models of Ombao et al. (2002) have been proposed. A comprehensive review of locally stationary series can be found in Dahlhaus (2012). As part of these developments, the local autocovariance, for non- or locally stationary processes, has been studied in the literature and details on specific estimators can be found in Hyndman and Wand (1997), Nason (2013b), Cardinali (2014) and Zhao (2015), for example. However, to date, little attention seems to have been paid to *local* partial autocorrelation and the benefits it could bring. An exception is Degerine and Lambert (1996) and Degerine and Lambert-Lacroix (2003), who extended the classical partial autocorrelation to encompass nonstationary processes. Their seminal work mentioned estimation, including the windowing idea that we use in Section 3, but provided no theory for their estimator nor evaluation via simulation or on real time series. More recently, Yang, Holan and Wikle (2016) use a hierarchical Bayesian modelling approach to estimate process time-frequency structure, linking the time-dependent partial autocorrelations to the coefficients of a time-varying autoregressive process.

Autocorrelation and partial autocorrelation are intimately related, presenting complementary views on the underlying structure within a time series. For example, arguably, partial autocorrelation provides direct information on the order and underlying structure of autoregressive-type processes. Hence, it is likely that, local partial autocorrelation will be just as important for non-stationary processes as classical partial autocorrelation has been for stationary ones. Our work provides the first estimator theory on two estimators and also a freeware R software package, `lpacf`, for local partial autocorrelation that complements existing local autocorrelations, such as `lacf` in the `locits` package.

As in the stationary case, for real-life statistical analysis one needs both local autocorrelation *and* partial autocorrelation and this article fills the gap for the latter. We introduce two new estimators of the local partial autocorrelation function, supplying new results on their theoretical properties. We further exhibit

our estimators on a simulated series and three real time series that demonstrate the importance of using a local approach.

2. The Local Partial Autocorrelation Function

2.1. The (process) local partial autocorrelation function, q_T, for a locally stationary process

Let $\{X_{t,T}\}$ be a zero-mean locally stationary process such as the locally stationary Fourier process, Dahlhaus (1997, Definition 2.1), or the locally stationary wavelet process, Nason, von Sachs and Kroisandt (2000, Definition 1) (for ease of reference, these definitions can also be found in Appendix B). Locally stationary process theory often has quantities of interest such as the time-varying spectrum, $f(z,\omega)$ at (rescaled) time $z \in (0,1)$ and frequency ω, or local autocovariance $c(z,\tau)$ at location z and lag τ, which are estimated via a process quantity (f_T or c_T), which depends on the sample size T and asymptotically approaches to the quantity of interest as $T \to \infty$. Consider, for example, $f_T(z,\omega)$ from Neumann and von Sachs (1997, formula (3.3)) or $c_T(z,\tau)$ from Nason, von Sachs and Kroisandt (2000, Section 2.3). We follow this paradigm by first introducing the process local partial autocorrelation, q_T.

The (process) partial autocorrelation function, $q_T(z,\tau)$, of a zero-mean *locally* stationary process can be understood informally as

$$q_T(z,\tau) = \text{corr}\left(X_{[zT],T}, X_{[zT]+\tau,T} | \text{"in-between" data}\right),$$

where $[x]$ denotes the integer part of the real number x. A formal definition follows.

Definition 2.1. The local process partial autocorrelation of a zero-mean locally stationary process $\{X_{t,T}\}_{t=0}^{T-1}$, at rescaled time $z \in (0,1)$ and lag τ, is defined by

$$q_T(z,\tau) = \text{corr}\left\{X_{[zT]+\tau,T} - P_{[zT],\tau}(X_{[zT]+\tau,T}), X_{[zT],T} - P_{[zT],\tau}(X_{[zT],T})\right\}, \tag{1}$$

where $P_{[zT],\tau}(\cdot)$ is the projection operator onto $\overline{\text{sp}}(X_{[zT]+1,T}, \ldots, X_{[zT]+\tau-1,T})$, where $\overline{\text{sp}}$ is the closed span defined by Brockwell and Davis (1991).

The next proposition shows an alternative useful representation of q_T.

Proposition 2.2. *Let $\{X_{t,T}\}$ be a zero-mean locally stationary process. Then the process local partial autocorrelation, q_T, can be expressed as*

$$q_T(z,\tau) = \varphi_{[zT],\tau,\tau;T}\left[\frac{\text{Var}\{X_{[zT],T} - P_{[zT],\tau}(X_{[zT],T})\}}{\text{Var}\{X_{[zT]+\tau,T} - P_{[zT],\tau}(X_{[zT]+\tau,T})\}}\right]^{1/2}, \tag{2}$$

where $\varphi_{[zT],\tau,\tau;T}$ is from projecting $X_{[zT]+\tau,T}$ onto $\overline{sp}(X_{[zT],T}, \ldots, X_{[zT]+\tau-1,T})$, the projection being

$$\hat{X}_{[zT]+\tau,T} = \varphi_{[zT],\tau,1;T}X_{[zT]+\tau-1,T} + \ldots + \varphi_{[zT],\tau,\tau;T}X_{[zT],T}. \tag{3}$$

Proof. See Appendix D.1.

Formulae (1) and (2) are natural generalisations of their stationary equivalents, compare for example with Definitions 3.4.1 and 3.4.2 from Brockwell and Davis (1991).

2.2. Equivalent expressions for the process local partial autocorrelation function, q_T

As a step to estimation, we will express q_T by exploiting a well-known connection between partial autocorrelation and linear prediction. We introduce the following notation $P_{[zT],\tau}(X_{[zT],T}) = \hat{X}^{(b)}_{[zT],T}$ and $P_{[zT],\tau}(X_{[zT]+\tau,T}) = \hat{X}^{(f)}_{[zT]+\tau,T}$. These are simply the respective linear predictors of $X_{[zT],T}$ (*b*ack-casted), and $X_{[zT]+\tau,T}$ (*f*orecasted), using the predictor set $X_{[zT]+1,T}, \ldots, X_{[zT]+\tau-1,T}$. The numerator and denominator in (2) can be re-expressed as a Mean Squared Prediction Error (MSPE). Consequently, we can rewrite $q_T(z,\tau)$ as

$$q_T(z,\tau) = \varphi_{[zT],\tau,\tau;T} \left\{ \frac{\text{MSPE}(\hat{X}^{(b)}_{[zT],T}, X_{[zT],T})}{\text{MSPE}(\hat{X}^{(f)}_{[zT]+\tau,T}, X_{[zT]+\tau,T})} \right\}^{1/2}. \tag{4}$$

For details see Appendix D.2. For stationary processes the square root term in (4) equals one and $q_T(z,\tau)$ coincides with the classical $q(\tau)$.

In general, given t observations of a zero-mean locally stationary process, $X_{0,T}, \ldots, X_{t-1,T}$, the mean squared prediction error of a linear predictor of $X_{t,T}$, $\hat{X}_{t,T} = \sum_{s=0}^{t-1} b_{t-1-s,T} X_{s,T}$, can be written as

$$\text{MSPE}\left(\hat{X}_{t,T}, X_{t,T}\right) = \boldsymbol{b}_t^T \Sigma_{t;T} \boldsymbol{b}_t,$$

where $\boldsymbol{b}_t = (b_{t-1,T}, \ldots, b_{0,T}, -1)^T$ and $\Sigma_{t,T}$ is the covariance of $X_{0,T}, \ldots, X_{t,T}$, see, e.g., Fryzlewicz, Van Bellegem and von Sachs (2003, Section 3.3). In our case, the back-casted and forecasted values of $X_{[zT],T}$ and $X_{[zT]+\tau,T}$ are also linear predictors using the window of observations $X_{[zT]+1,T}, \ldots, X_{[zT]+\tau-1,T}$, and can be expressed as

$$\hat{X}^{(b)}_{[zT],T} = \sum_{p=1}^{\tau-1} b^{(b)}_{p,T} X_{[zT]+p,T} \quad \text{and} \quad \hat{X}^{(f)}_{[zT]+\tau,T} = \sum_{p=1}^{\tau-1} b^{(f)}_{\tau-1-p,T} X_{[zT]+p,T},$$

respectively. Here, the $\boldsymbol{b}^{(b)}$, $\boldsymbol{b}^{(f)}$ coefficient vectors are obtained through minimisation of the corresponding mean squared prediction error using the same principle as in the stationary case.

We next give a proposition that paves the way towards a natural definition of the local partial autocorrelation function q in Section 2.3.

Proposition 2.3. *Let* $\{X_{t,T}\}$ *be a zero-mean locally stationary process. Then* q_T *can also be expressed as*

$$
q_T\left(z,\tau\right) = \varphi_{[zT],\tau,\tau;T} \left\{ \frac{(\boldsymbol{b}_{[zT]}^{(b)})^T \Sigma_{[zT];T}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}}{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \Sigma_{[zT]+\tau;T}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}} \right\}^{1/2}, \tag{5}
$$

where $\varphi_{[zT],\tau,\tau;T}$ *is as in* (3)*, and* $\boldsymbol{b}_{[zT]}^{(b)} = (-1, \tilde{b}_{1,T}^{(b)}, \ldots, \tilde{b}_{\tau-1,T}^{(b)})^T$ *and* $\boldsymbol{b}_{[zT]+\tau}^{(f)} = (b_{\tau-2,T}^{(f)}, \ldots, b_{0,T}^{(f)}, -1)^T$ *are* $\tau \times 1$ *coefficient vectors. To simplify notation we have suppressed the dependency of the* $\boldsymbol{b}$*-vector components on* $[zT]$ *and also dependency of* $\boldsymbol{b}_{[zT]}^{(b)}$*,* $\Sigma_{[zT];T}^{(b)}$*,* $\boldsymbol{b}_{[zT]+\tau}^{(f)}$*,* $\Sigma_{[zT]+\tau;T}^{(f)}$ *on* τ*, even though it is still present. The* $\tau \times \tau$ *covariance matrices* $\Sigma_{[zT];T}^{(b)}$ *and* $\Sigma_{[zT]+\tau;T}^{(f)}$ *are given in Appendix* C*.*

Proof. See Appendix D.2.

We will use expression (5) as the basis of an estimator in Section 2.4. The last element of the vector $\boldsymbol{\varphi}_{[zT],\tau;T}$, denoted $\varphi_{[zT],\tau,\tau;T}$, can be obtained as the solution to the (local) Yule-Walker equations $\Sigma_{[zT];T} \boldsymbol{\varphi}_{[zT],\tau;T} = \boldsymbol{r}_{[zT];T}$, where $\Sigma_{[zT];T}$ is the $\tau \times \tau$ covariance matrix given in Appendix C and $\boldsymbol{r}_{[zT];T}$ is the $\tau \times 1$ covariance vector of $X_{[zT]+\tau,T}$ with $\left(X_{[zT]+\tau-1,T}, \ldots, X_{[zT],T}\right)$. This is equivalent to obtaining a solution $\hat{X}_{[zT]+\tau,T}$ that achieves minimum mean squared prediction error over the class of linear predictors. For stationary processes the covariance matrix Γ_τ is Toeplitz. However, for locally stationary processes the covariance matrix $\Sigma_{[zT];T}$ only has an approximate Toeplitz structure. Once again, for ease of notation, we have suppressed the dependency on the lag τ from the vector $\boldsymbol{r}_{[zT];T}$ and covariance matrix $\Sigma_{[zT];T}$, the latter given in Appendix C.

2.3. The wavelet local partial autocorrelation function, q

The local (process) partial autocorrelation introduced in Section 2.1 can be applied to any zero-mean locally stationary process. However, for the theory we develop below, we need to establish the underlying asymptotic quantity, which is intimately related to the data generating model. Hence, from now on, we assume that the process $\{X_{t,T}\}$ is a zero-mean locally stationary wavelet process and define the local partial autocorrelation function, q, which we show later to be the asymptotic limit of q_T from (2).

Definition 2.4. Let $\{X_{t,T}\}$ be a zero-mean locally stationary wavelet process as defined by Fryzlewicz, Van Bellegem and von Sachs (2003) with local auto-covariance $c(z,\tau)$ and spectrum $\{S_j(z)\}_j$ that satisfy

$$
\sum_{\tau=0}^{\infty} \sup_z |c(z,\tau)| < \infty, C_1 := \mathrm{ess} \inf_{z,\omega} \sum_{j>0} S_j(z)|\hat{\psi}_j(\omega)|^2 > 0,
$$

where $\hat{\psi}_j(\omega) = \sum_s \psi_{j,0}(s) \exp(i\omega s)$. Then, the local partial autocorrelation function is

$$q(z, \tau) = \varphi_{[zT],\tau,\tau} \left\{ \frac{(\boldsymbol{b}^{(b)}_{[zT]})^T B^{(b)}_{[zT]} \boldsymbol{b}^{(b)}_{[zT]}}{(\boldsymbol{b}^{(f)}_{[zT]+\tau})^T B^{(f)}_{[zT]+\tau} \boldsymbol{b}^{(f)}_{[zT]+\tau}} \right\}^{1/2}, \qquad (6)$$

where

1. the quantity $\varphi_{[zT],\tau,\tau}$ is the last element in the vector $\boldsymbol{\varphi}_{[zT],\tau}$ (of length τ) obtained as the solution to the local Yule-Walker equations i.e. $B_{[zT]}\boldsymbol{\varphi}_{[zT],\tau} = \boldsymbol{r}_{[zT]}$,

2. the matrices $B^{(f)}_{[zT]+\tau}$ and $B^{(b)}_{[zT]}$ are the local approximations of $\Sigma^{(f)}_{[zT]+\tau;T}$ and $\Sigma^{(b)}_{[zT];T}$, as in the proof of Lemma A.1 from Fryzlewicz, Van Bellegem and von Sachs (2003). The $\boldsymbol{r}_{[zT]}$ are also local approximations to $\boldsymbol{r}_{[zT];T}$ from Section 2.2 but using $c(z, \tau)$.

3. the coefficient vectors $\boldsymbol{b}^{(f)}_{[zT]+\tau}$ and $\boldsymbol{b}^{(b)}_{[zT]}$ are obtained as the solution to the forecasting and back-casting prediction equations, or equivalently through minimisation of the MSPE. See Section 3.1 and Proposition 3.1 from Fryzlewicz, Van Bellegem and von Sachs (2003) for details.

Next, Proposition 2.5 shows that the (process) local partial autocorrelation, q_T, converges to the local partial autocorrelation, q, defined by (6).

Proposition 2.5. *Let $\{X_{t,T}\}$ be a zero-mean locally stationary wavelet process as defined by Definition 2.4, with spectrum $\{S_j(z)\}_{j=1}^{\infty}$ constructed with nondecimated discrete wavelet system $\{\psi_{j,k}(t)\}$. Let local partial autocorrelations q_T and q be defined as in (4) and (6) respectively. Then, as $T \to \infty$, uniformly in $\tau \in \mathbb{Z}$ and $z \in (0, 1)$, we have $|q_T(z, \tau) - q(z, \tau)| = \mathcal{O}(T^{-1})$.*

Proof. See Appendix D.3.

This result parallels the local autocovariance result of Nason, von Sachs and Kroisandt (2000), where it is shown that $|c_T(z, \tau) - c(z, \tau)| = \mathcal{O}(T^{-1})$ as $T \to \infty$ uniformly in $\tau \in \mathbb{Z}$ and $z \in (0, 1)$.

2.4. Wavelet local partial autocorrelation estimation

We now consider the important problem of local partial autocorrelation estimation. We begin by first noting that all the quantities on the right-hand side of (6) for $q(z, \tau)$ are based on the local autocovariance $c(z, \tau)$. A natural estimator of q can thus be obtained by replacing all occurrences of $c(z, \tau)$ by the wavelet-based estimator $\hat{c}(z, \tau)$ from Nason (2013b, Section 3.3) as follows.

Definition 2.6. The wavelet-based local partial autocorrelation estimator is defined as

$$\tilde{q}(z, \tau) = \tilde{\varphi}_{[zT],\tau,\tau} \left\{ \frac{(\tilde{\boldsymbol{b}}^{(b)}_{[zT]})^T \tilde{B}^{(b)}_{[zT]} \tilde{\boldsymbol{b}}^{(b)}_{[zT]}}{(\tilde{\boldsymbol{b}}^{(f)}_{[zT]+\tau})^T \tilde{B}^{(f)}_{[zT]+\tau} \tilde{\boldsymbol{b}}^{(f)}_{[zT]+\tau}} \right\}^{1/2}, \qquad (7)$$

where the matrix $\tilde{\mathrm{B}}^{(b)}_{[zT]}$, $\tilde{\mathrm{B}}^{(f)}_{[zT]+\tau}$ and vector $\tilde{\boldsymbol{b}}^{(b)}_{[zT]}$, $\tilde{\boldsymbol{b}}^{(f)}_{[zT]+\tau}$ estimates are obtained from their population quantities in Sections 2.2 and 2.3 by plugging in the wavelet-based local autocovariance estimator $\hat{c}$ from Nason (2013b). Similarly, the vector $\tilde{\boldsymbol{\varphi}}_{[zT],\tau}$ is obtained as the solution to the local Yule-Walker equations in Definition 2.4 again replacing c by $\hat{c}$.

We next establish the consistency of $\tilde{q}$ for q.

Proposition 2.7. *Let* $\{X_{t,T}\}$ *be a zero-mean locally stationary wavelet process under the assumptions given in Definition 2.4. The local partial autocorrelation estimator* $\tilde{q}(z,\tau)$ *from* (7) *is consistent for the true local partial autocorrelation* $q(z,\tau)$, *in that* $\tilde{q}(z,\tau) - q(z,\tau) = o_p(1)$ *as* $T \to \infty$.

Proof. See Appendix D.4.

Our wavelet-based estimator, $\tilde{q}(z,\tau)$, develops earlier work on forecasting by Fryzlewicz (2003) in a new direction. However, the estimator is not simple to implement and, as we will see later, does not perform as well as the following alternative approach, which applies a window to the classical partial autocorrelation.

3. Windowed Estimation of Local Partial Autocorrelation

3.1. The integrated local wavelet periodogram

We introduce an alternative estimator, $\tilde{q}_W(z,\tau)$, that is simpler to implement than $\tilde{q}(z,\tau)$, and turns out to perform better. This new estimator is constructed by windowing the classical partial autocorrelation (designed for stationary processes) over an interval centred at time $[zT]$ with length $L(T)$, where $L(T) \to \infty$ and $L(T)/T \to 0$, as $T \to \infty$. Proposition 3.4, in Section 3.2, establishes the asymptotic behaviour of $\tilde{q}_W(z,\tau)$ by approximating the integrated local wavelet periodogram of a (zero-mean) locally stationary wavelet process by its equivalent stationary version at a fixed rescaled time (see Theorem 1). The proof of the theorem introduces new bounds for quantities involving cross-correlation wavelets, as well as a new exact formula for cross-correlation Haar wavelets.

All key definitions and results are presented below, while full proofs are in Appendix E.

Definition 3.1. Let $\{X_{t,T}\}$ be a locally stationary wavelet process as in Definition 1 from Nason, von Sachs and Kroisandt (2000) with evolutionary wavelet spectrum $\{S_j(z)\}_{j=1}^{\infty}$ for $z \in (0,1)$, Lipschitz constants $\{L_j\}_{j=1}^{\infty}$, process constants $\{C_j\}_{j=1}^{\infty}$ and underlying discrete nondecimated wavelets $\{\psi_{j,k}\}$. The integrated local periodogram on the interval $[[zT] - L(T)/2 + 1, [zT] + L(T)/2]$ is

$$J_{L(T)}(z,\phi) = \sum_{j=1}^{\infty} \phi_j I^*_{L(T)}(z,j).$$

Here $\{\phi_j\}_{j=1}^\infty \in \Phi$ and Φ is a set of complex-valued bounded sequences equipped with uniform norm $||\phi||_\infty := \sup_j |\phi_j|$, $z \in (0, 1)$ and, for $j \in \mathbb{N}$, $I_{L(T)}^*(z, j)$ is the uncorrected, tapered local wavelet periodogram given by

$$I_{L(T)}^*(z, j) = H_{L(T)}^{-1} \left| \sum_{t=0}^{L(T)-1} h\{t/L(T)\} X_{[zT]+t-L(T)/2+1,T} \psi_{j,[zT]}(t) \right|^2, \text{ for } j \in \mathbb{N},$$

$$(8)$$

with $h : [0, 1] \to \mathbb{R}_+$ and normalizing factor $H_{L(T)} := \sum_{j=0}^{L(T)-1} h^2\{j/L(T)\} \sim L(T) \int_0^1 h^2(x)\, dx$.

We next approximate the integrated local (wavelet) periodogram $J_{L(T)}(z, \phi)$ by the corresponding statistics of a stationary process $\{Y_t\}$ with the same local corresponding statistics at $t = zT$, for fixed z. Conceptually, this is a common approach useful in establishing asymptotic properties for functions of locally stationary processes (Dahlhaus and Giraitis, 1998), which in this work we advance to include *wavelet*-based expansions. Specifically, define

$$J_{L(T)}^Y(\phi) = \sum_{j=1}^\infty \phi_j I_{L(T)}^{*,Y}(j),$$

where

$$I_{L(T)}^{*,Y}(j) := H_{L(T)}^{-1} \left| \sum_{s=0}^{L(T)-1} h\{s/L(T)\} Y_{[zT]-L(T)/2+1+s,T} \psi_{j,[zT]}(s) \right|^2$$

is the wavelet periodogram on the segment $[zT]-L(T)/2+1, \ldots, [zT]+L(T)/2$ of the *stationary* process

$$Y_s = \sum_{j=1}^\infty W_j(z) \sum_{k=-\infty}^\infty \psi_{j,k}(s)\xi_{j,k},$$

$$(9)$$

where $\psi_{j,k}$ is the same wavelet sequence as previously, $\{\xi_{j,k}\}$ a set of independent identically distributed random variables with mean zero and unit variance and $W_j(z)$ is such that $W_j^2(z) = S_j(z)$ for all $z \in (0, 1)$ and $j \in \mathbb{N}$. The next theorem is the key result establishing the asymptotic properties of the integrated local wavelet periodogram.

Theorem 1. Let $\{X_{t,T}\}$ be a zero-mean Gaussian locally stationary wavelet process as defined by Definition 3.1. Suppose $\sum_j C_j^2 2^{2j} < \infty$, $\{W_j\}_j$ is Lipschitz continuous with Lipschitz constants L_j such that $\sum_j L_j^2 2^{2j} < \infty$ and $\sum_j W_j^2(z) 2^{2j} < \infty$ at any rescaled time z, $L(T)/T \to 0$ as $T \to \infty$, and $\phi \in \Phi$ is a sequence of bounded variation. Further, assume h is a rectangular kernel. Then, using the family of discrete Haar wavelets, we have

$$\mathbb{E}\left\{J_{L(T)}(z,\phi)\right\} = \mathbb{E}\left\{J_{L(T)}^{Y}(\phi)\right\} + \mathcal{O}\left\{L(T)^{-1}\right\}, \tag{10}$$

$$J_{L(T)}(z,\phi) - \mathbb{E}\left\{J_{L(T)}(z,\phi)\right\} = o_p\left(L(T)^{-1/2}\right), \tag{11}$$

$$J_{L(T)}^{Y}(\phi) - \mathbb{E}\left\{J_{L(T)}^{Y}(\phi)\right\} = o_p\left(L(T)^{-1/2}\right). \tag{12}$$

Proof. Appendix E.5 contains the full proof.

An important difference between earlier literature in *this* area and our work is the introduction of windowing. We provide new results on windowed versions of the cross-correlation wavelets, which we denote $i_{N,z}$, where, to simplify notation, we replace $L(T)$ by N and sometimes omit z. To prove Theorem 1 we need bounds on quantities involving $i_{N,z}$ which we can obtain via their link with cross-correlation wavelets and, in particular, our new *closed form* expression for the cross-correlation Haar wavelet. For completeness, we define the truncated cross-correlation wavelet here and some of the key bounds.

Definition 3.2. For $N \in \mathbb{N}$, scales $j, \ell \in \mathbb{N}$ and rescaled time $z \in (0,1)$, the windowed cross-scale autocorrelation wavelets $i_{N,z}(j, \ell, \cdot)$ over the interval $[[zT] - N/2 + 1, [zT] + N/2]$ are

$$i_{N,z}(j, \ell, k) = \sum_{t=0}^{N-1} \psi_{j,[zT]-t} \psi_{\ell,k-[zT]-t-1+N/2}, \tag{13}$$

where $\{\psi_{j,m}\}_{j,m}$ is a family of discrete wavelets and $k \in \mathbb{Z}$.

The similarity between the cross-scale autocorrelation wavelets $\Psi_{j,\ell}(\cdot)$, defined in Fryzlewicz (2003, Definition 5.4.2) as $\Psi_{j,\ell}(\tau) = \sum_{k \in \mathbb{Z}} \psi_{j,k} \psi_{\ell,k+\tau}$ for $j, \ell \in \mathbb{N}$ and $\tau \in \mathbb{Z}$, and their windowed version, $i_{N,z}(j, \ell, \cdot)$ defined above, is key to how we subsequently bound quantities involving $i_{N,z}$. The exact new formulas for Haar cross-scale autocorrelation wavelets are established in Appendix E.1, along with a pictorial description in Figure 6 in Appendix E.7.

As bounds for $i_N(j, \ell, \cdot)$ are a key component of the proof of Theorem 1, these are provided by the next three results. The first bound for i_N is valid for all discrete wavelets based on Daubechies (1992) compactly supported wavelets, although we later only use it for Haar wavelets.

Lemma 1. Using previous notation and assumptions, let $b_1 = [zT] + N/2 + 1$ and $b_2 = [zT] + N/2 + N_\ell - 1$. Then

$$|i_{N,z}(j, \ell, k)| \leq |\Psi_{j,\ell}(k - 2[zT] + N/2 - 1)|, \tag{14}$$

holds when

$$[zT] > N_j - 2 \text{ and } k < b_1 \tag{15}$$

or when

$$k > b_2, \tag{16}$$

for integers k, $z \in (0,1)$, $j, \ell \in \mathbb{N}$ and N_j is the length of the discrete wavelet $\psi_{j,.}$ for all Daubechies compactly supported wavelets.

When $b_1 \leq k \leq b_2$ we have (i) for Daubechies' wavelets with two or more vanishing moments:

$$
\begin{aligned}
|i_{N,z}(j,\ell,k)| \leq{} & |\Psi_{j,\ell}(k - 2[zT] + N/2 - 1)| \\
& + \frac{2^{-(j+\ell)/2}}{[zT] - N} \{\gamma + \log([zT] - N) - \log([zT]) + \log N \\
& + \mathcal{O}([zT]^{-1}) + \mathcal{O}\{([zT] - N)^{-1}\} + \mathcal{O}(N^{-1})\},
\end{aligned}
\tag{17}
$$

where γ is the Euler-Mascheroni constant and (ii) for Haar wavelets we have:

$$
|i_{N,z}(j,\ell,k)| \leq 2^{-(j+\ell)/2} \{\min(N_\ell, N_j) + N_\ell\}.
\tag{18}
$$

Proof. See Appendix E.3.

We use Lemma 1 to prove the next two useful results about i_N.

Lemma 2. Using previous notation and assumptions, and assuming $\{\psi_{j,k}\}$ are discrete Haar wavelets

$$
\sum_{k=-\infty}^{\infty} \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)|^2 = \mathcal{O}(2^{2\ell}),
\tag{19}
$$

$$
\sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)i_{N,z}(j,m,n)| \right\}^2 = \mathcal{O}\{2^{(\ell+m)}\}.
\tag{20}
$$

Proof. See Appendix E.4.

These properties of the integrated local wavelet periodogram allow us to establish the asymptotic behaviour of $\tilde{q}_W(z,\tau)$ in the following section.

3.2. Windowed local partial autocorrelation estimation

We now define a local partial autocorrelation estimator by using the classical (stationary) partial autocorrelation computed on a window of length $L(T)$ centred at time $[zT]$. The theoretical properties of this windowed estimator are derived and we investigate its empirical behaviour.

Definition 3.3. Let $\hat{q}$ be the usual partial autocorrelation estimator as defined by Brockwell and Davis (1991, Definition 3.4.3) for example. Define the window $\mathcal{I}(z,L) := [z - L(T)/2T, z + L(T)/2T]$ for some interval length function $L(T)$ and location $z \in (0,1)$. Define the windowed estimator, $\tilde{q}_W(z,\tau)$, of the local partial autocorrelation function at rescaled time z and lag τ, to be the classical partial autocorrelation function evaluated on observations contained in $\mathcal{I}(z,L)$ and denoted by

$$
\tilde{q}_W(z,\tau) = \hat{q}_{\mathcal{I}(z,L)}(\tau).
$$

Our definition uses a rectangular window, but some of our applications later use an Epanechnikov window, others could be substituted.

The integrated wavelet periodogram approximation derived in Theorem 1 ensures that our windowed estimator can benefit from the established asymptotic distributional properties of the partial autocovariance in the stationary setting.

Proposition 3.4. *Let $\{X_{t,T}\}$ be a zero-mean Gaussian locally stationary wavelet process under the conditions set out by Theorem 1. Then, for the windowed local partial autocorrelation estimator $\tilde{q}_W(z, \tau)$ from Definition 3.3, assuming $L(T) \to \infty$ and $L(T)/T \to 0$, as $T \to \infty$, we have that $\tilde{q}_W(z, \tau)$ converges in distribution to $\hat{q}^Y(\tau)$, where Y is a stationary process with the same characteristics at rescaled time z as the process $\{X_{t,T}\}$ (constructed as in equation (9)) and $\hat{q}^Y(\tau) = \hat{\varphi}^Y_{\tau,\tau}$ is the classical Yule-Walker partial autocorrelation function estimator.*

Proof. Appendix E.6 contains the proof, which relies on the integrated wavelet periodogram approximation from Theorem 1.

When dealing with processes that can be locally well modelled by an autoregressive structure, the result above amounts to establishing the asymptotic normality of our windowed local partial autocovariance estimator for large lags (see next corollary).

Corollary 1. Under the assumptions from Proposition 3.4 and assuming that $\{X_{t,T}\}$ can be locally well modelled by an autoregressive structure of order say p, then for lags τ larger than p we have that $L(T)^{1/2}\tilde{q}_W(z, \tau)$ converges in distribution to a standard normal random variable.

Proof. The proof follows directly from Proposition 3.4 and classical theory on the asymptotic behaviour of Yule-Walker estimates for stationary autoregressive processes (see for instance Theorem 8.1.2 from Brockwell and Davis (1991)).

3.3. Choice of Control Parameters

As with many nonparametric estimation methods in the literature, we have to make various choices in an attempt to obtain good estimators $\tilde{q}_W(z, \tau)$. Unfortunately there is no universal automatic best choice, at least in the real world. For the wavelet estimator, $\tilde{q}$, we have to specify an underlying wavelet, a method for handling boundaries and also a smoothing parameter, e.g. s in Section 3.3 in Nason (2013b). However, a further advantage of the windowed estimator is that we really only have to choose the window width $L(T)$ and the window kernel. Dahlhaus and Giraitis (1998) show that the Epanechnikov window is a good choice, which we advocate here.

Unfortunately, rates of convergence of the estimator, although providing theoretical insight, do not really help with the practical selection of the window width. A promising direction for practical bandwidth selection be via methods

such as the locally stationary process bootstrap for pre-periodogram-like quantities, as proposed by Kreiss and Paparoditis (2014), but development of this is beyond the scope of the current paper.

Below, we use a manually-selected window width, by observing choices that achieve a good balance between estimates that are too rough, and those that appear too smooth (and change little on further smoothing). Section 4.3 and Appendix G provide some empirical evidence that the window width choice is not too hard, and the results are not particularly sensitive to it. Such manually-selected procedures are well-acknowledged in the literature, e.g. page 1 of Chaudhuri and Marron (1999). They could, with some work, be considerably expanded into something like that article's SiZer methodology. We are also guided by the 'automatic' bandwidth selection method from Nason (2013b) as implemented in the R locits package.

4. Local partial autocorrelation estimates in practice

4.1. Simulated time-varying autoregressive process example

We illustrate our local partial autocorrelation function estimators on a simple, well-understood, simulated time-varying autoregressive process, TVAR(1). Figure 1 shows the classical partial autocorrelation function and our two estimators applied to a single $T = 512$ realization from a time-varying autoregressive process with lag one coefficient linearly changing from 0.9 to -0.9 over the series. The 95% confidence bands are constructed under the null hypothesis of white noise and are the standard ones as displayed by, e.g., established R software.

Unsurprisingly, the classical partial autocorrelation is misleading. If anything, it indicates an incorrect strong lag two structure, which shows the importance of not using techniques designed for stationary series on nonstationary ones. Our two developed local partial autocorrelation estimators track the true time-varying autoregressive parameters. The wavelet-based estimate seems a bit worse, particularly for the lag two partial autocorrelation after about time 350. This was confirmed by a small simulation study, based on 100 realizations drawn from the TVAR process. The average root-mean-square error for the wavelet estimator at lags one and two was 0.95 and 0.19, respectively, whereas for the windowed estimator it was 0.88 and 0.07 respectively. Both estimators are less accurate near to the ends of the series, which is a common problem with such estimators, see Cheng and Hall (2003), for example. However, the windowed estimator usually appears less affected.

The TVAR process used in Figure 1 exhibits a large range of time-varying parameter values from -0.9 to $+0.9$. However, we repeated the example for less extreme parameter changes. Unless the parameter change is very small, and the process is close to stationary, the classical partial autocorrelation still misleads. For smaller parameter changes, the classical partial autocorrelation often gets the process order correct, but gives a partial autocorrelation value that is often close to the average of the local partial autocorrelations.
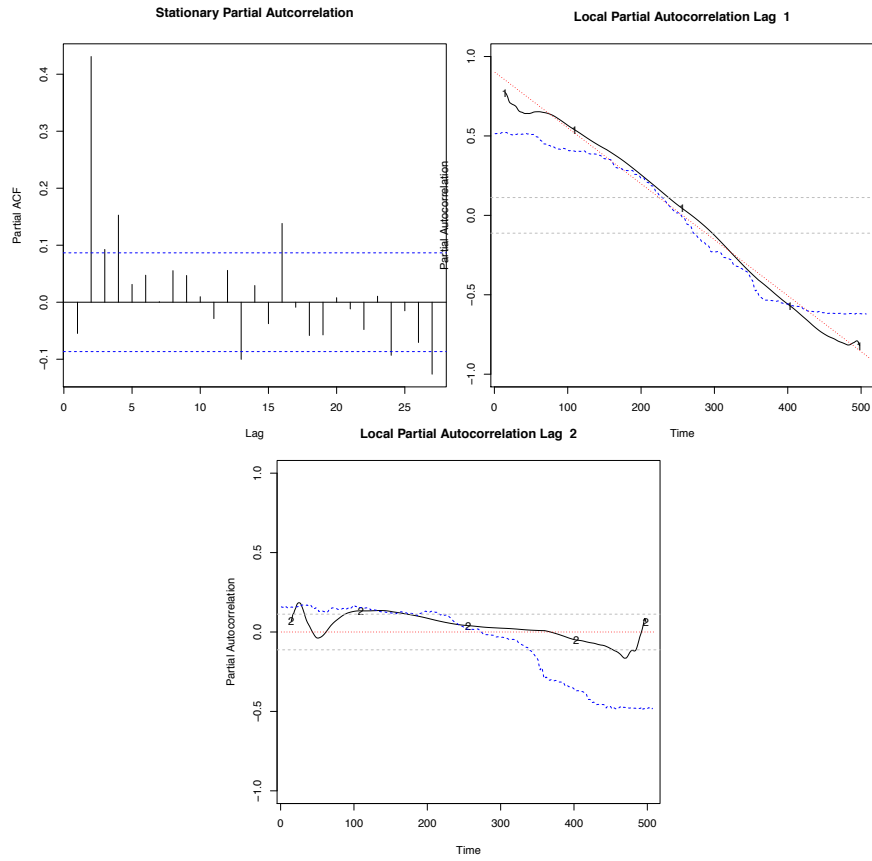
FIG 1. *Partial autocorrelation function (pacf) estimators applied to a time-varying autoregressive process realization. Top left: classical (stationary) pacf. Top right: Lag one local pacf. Bottom: Lag two local pacf. Theoretical: red dotted line, Daubechies D5 wavelet-based estimate ($\tilde{q}$, Section 2.4) is dashed blue line, Epanechnikov windowed estimate ($\tilde{q}_W$, Section 3) is solid black line. Bandwidth selected using* AutoBestBW *from* locits *package.*

4.2. U.K. National Accounts data

The ABML time series obtained from the Office for National Statistics contains values of the U.K. gross-valued added (GVA), which is a component of gross domestic product (GDP). Our ABML series is recorded quarterly from quarter one 1955 to quarter three 2010 and consists of $T = 223$ observations. As with
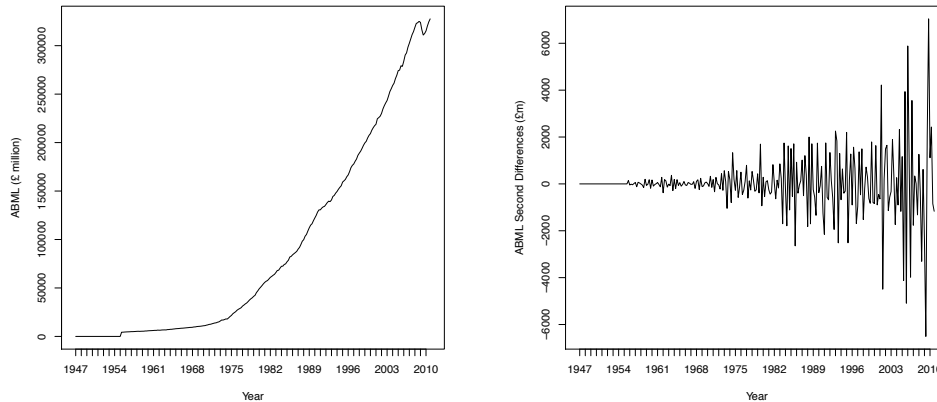

FIG 2. *Left: ABML time series. Right: second-order differences of ABML.*

many economic time series, ABML exhibits a clear trend, which we removed using second-order differences; these are shown in Figure 2. Naturally, other methods for removing the trend could be tried. The second-order differences strongly suggest that the series is not second-order stationary, with the series variance increasing markedly over time. Use of methods from Nason (2013b) show that the autocorrelation also changes over time. In particular, the lag one autocorrelation undergoes a major and rapid shift around 1991.

Much of the increase in variance observed in Figure 2 is probably due to inflation. However, we have also analysed two different inflation-corrected versions of ABML, one provided by the U.K. Office of National Statistics, and both of these are also not second-order stationary, as determined by tests of stationarity in Priestley and Subba Rao (1969) and Nason (2013b).

Our new estimation methodology enables us to obtain the windowed local partial autocorrelation estimator, $\tilde{q}_W(z, \tau)$, shown in Figure 3 and computed on the ABML second-order differences. The partial autocorrelations are time-varying and suggest that a time-varying AR(4) process might be suitable. There are times, such as the 1970s, when the higher-order partial autocorrelations are not outside of the approximate significant bands, indicating that a simpler time-varying AR(2) might be appropriate. These results are (i) economically interesting as the local variance, autocorrelation and partial autocorrelation all change over time, (ii) highlight the concerns with having no access to second-order conditional information (as was the case until now) and (iii) further pose
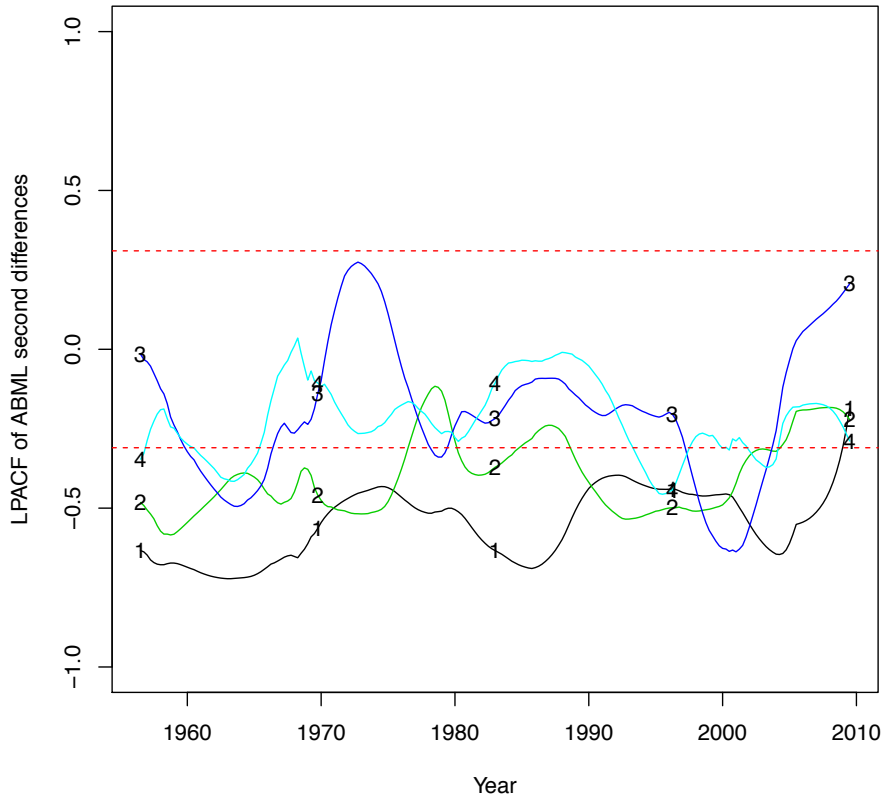
FIG 3. *Windowed partial autocorrelation, $\tilde{q}_W(z,\tau)$, of ABML second difference series for lags one to four indicated on each curve (Epanechnikov window with $L = 40$). Horizontal red dotted lines are approximate 95% confidence intervals.*

the challenge of accurately forecasting such data. Although the topic of time series forecasting is outside the scope of this paper, many authors acknowledge the superiority of wavelet-based forecasting (Aminghafari and Poggi, 2007; Schlüter and Deuschle, 2010) and we envisage the proposed local partial autocorrelation estimator could further improve results.

4.3. Precipitation in Eastport, U.S.

Understanding precipitation patterns is important for detecting climate change indications and for policy decisions. The left panel in Figure 4 shows monthly precipitation in millimetres from January 1887 until December 1950 (768 observations) at a location in Eastport. The data can be found in Hipel and McLeod (1994) and have been analysed in many publications including Rao, Hamed and Chen (2012); Dhakal et al. (2015). Our windowed local partial autocorrelation
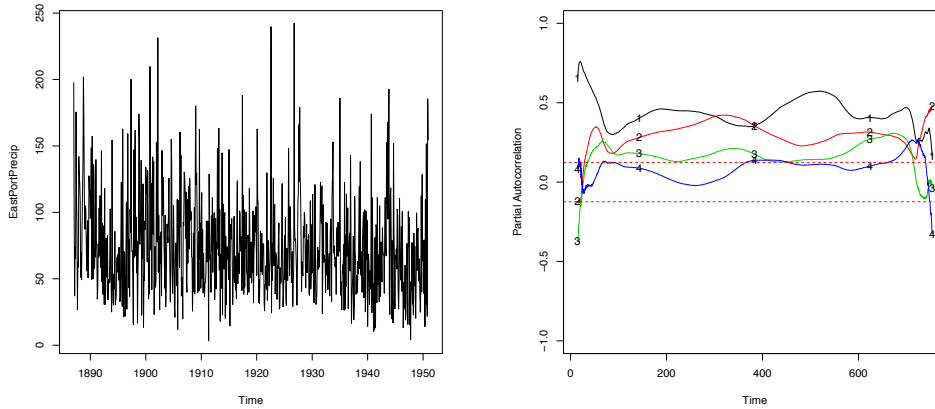
Fig 4. *Left: Precipitation (mm) in Eastport, U.S. Right: Windowed partial autocorrelation,*
$\tilde{q}_W(z, \tau)$, *of left for lags one to four indicated on each curve (Epanechnikov window with $L =$*
250, *chosen by* `AutoBestBW` *from* `locits` *package.). Horizontal red dotted lines are approximate*
95% confidence intervals.

estimate of the Eastport data is given in the right panel of Figure 4 and shows
clear nonstationarity at lags one through three. Some authors analyse this se-
ries as if it were stationary and our analysis suggests that this is inappropriate.
Indeed, if one applies a formal hypothesis test of nonstationarity on appropri-
ate lengths of the series, such as that proposed by Cardinali and Nason (2018),
there is strong evidence for nonstationarity. From a modelling point of view,
the estimated local partial autocorrelation behaviour might support fitting a
time-varying AR(3) model.

To provide some empirical support to the notion that window width is not
visually critical to the interpretation of the local partial autocorrelation, Ap-
pendix G shows the smoothed local partial autocorrelation plots similar to that
in the right-hand plot of Figure 4, but at three smaller window widths of 160, 80
and 40. The plot at window width of 160 is not that different to the one above at
$L = 250$ and, indeed, the $L = 80$ plot is not that dissimilar. However, the $L = 40$
plot almost certainly contains too much 'noise' and should be disregarded.

4.4. Euro-Dollar exchange rate

Following the introduction of the Euro currency in 1999 several authors includ-
ing Ahamada and Boutahar (2002) and Garcin (2017) have considered different
properties of this series, which have an influence on setting monetary policy in
various jurisdictions.

We analyze log returns of the monthly Euro-Dollar exchange rate as provided
by EuroStat at
http://ec.europa.eu/eurostat/web/products-datasets/-/ei_mfrt_m
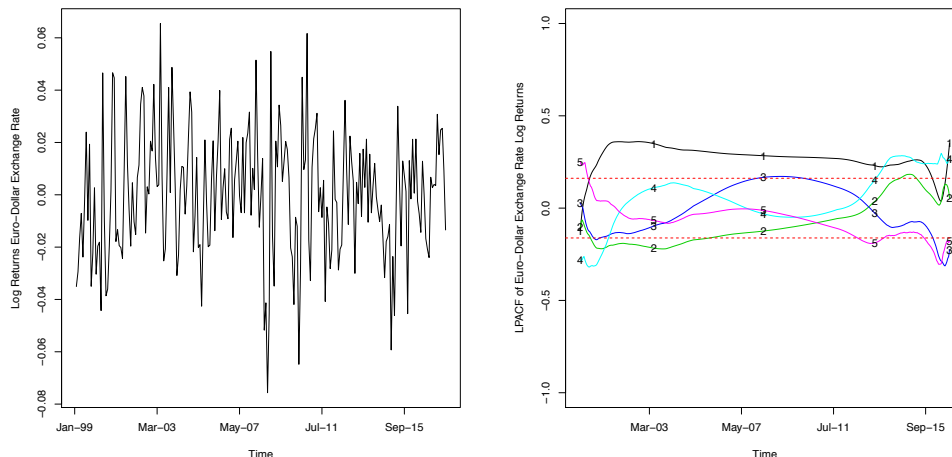from January 1999 until October 2017. The log returns and corresponding local

FIG 5. *Left: Monthly Euro-Dollar log returns. Right: Windowed partial autocorrelation,*
$\tilde{q}_W(z, \tau)$, *of left for lags one to five indicated on each curve. Horizontal red dotted lines*
are approximate 95% confidence intervals. (Epanechnikov window with $L = 147$)

partial autocorrelation function estimates are given in Figure 5. This demonstrates that the log returns do not appear to be time varying (outside of the boundary locations) and exhibit only lag one partial autocorrelation. This apparent stationarity is confirmed with formal tests using the `locits` (Nason (2013a), Nason (2013b)) and `fractal` (Constantine and Percival (2016), Priestley and Subba Rao (1969)) packages in R. Interestingly this relationship holds throughout the financial crisis, from 2008 to 2011.

These examples highlight the versatility of our method and its potential use to identify stationary behaviour as well as departures from it.

5. Discussion

This article develops two new estimators of the local partial autocorrelation function and studied their theoretical properties when applied to a locally stationary wavelet process. We established consistency for the wavelet-based estimator and asymptotic distribution for the windowed estimator. The latter result relied on new results on the integrated local wavelet periodogram, the (windowed) Haar cross-correlation wavelets and related quantities. We demonstrated the utility of the estimators for eliciting local structure on simulated data, the U.S. Eastport precipitation time series and the U.K. ABML time series. We also demonstrated the versatility of our method in the (desirable) presence of stationarity for the Euro-Dollar exchange rates. Most of the theoretical results relating to the generic local partial autocorrelation function estimator presented here are based on Haar wavelets, but many results and definitions also apply to other Daubechies' compactly supported wavelets. The associated software package `lpacf` contains functionality to compute the estimators for all

such wavelets, up to ten vanishing moments as contained within the `wavethresh` package (Nason, 2013c). The package will be released on to the Comprehensive R Archive Network (CRAN) in due course.

References

ABADIR, K. M. and MAGNUS, J. R. (2005). *Matrix Algebra (Econometric Exercises)*. Cambridge University Press, Cambridge.

ADAK, S. (1998). Time-dependent spectral analysis of nonstationary time series. *J. Am. Statist. Ass.* **93** 1488–1501.

AHAMADA, I. and BOUTAHAR, M. (2002). Tests for covariance stationarity and white noise, with an application to Euro/US dollar exchange rate: An approach based on the evolutionary spectral density. *Economics Letters* **77** 177 - 186.

AMINGHAFARI, M. and POGGI, J. M. (2007). Forecasting time series using wavelet. *Int. J. Wavelets, Multiresolution, Inf. Process.* **5** 709–724.

BILLINGSLEY, P. (1999). *Convergence of Probability Measures*. Wiley, New York.

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*. Springer, New York.

CARDINALI, A. (2014). Local Covariance Estimation using Costationarity. In *Topics in Nonparametric Statistics. Springer Proceedings in Mathematics and Statistics* **74** 53–60. Springer, New York.

CARDINALI, A. and NASON, G. P. (2018). Practical powerful wavelet packet tests for second-order stationarity. *App. Comp. Harm. Anal.* **44** 558–583.

CHATFIELD, C. (2003). *The Analysis of Time Series: An Introduction*. Chapman and Hall/CRC, London.

CHAUDHURI, P. and MARRON, J. S. (1999). SiZer for exploration of structures in curves. *J. Am. Statist. Ass.* **94** 807–823.

CHENG, M. Y. and HALL, P. (2003). Reducing variance in nonparametric surface estimation. *J. Mult. Anal.* **86** 375–397.

CONSTANTINE, W. and PERCIVAL, D. (2016). fractal: Fractal Time Series Modeling and Analysis R package version 2.0-1.

DAHLHAUS, R. (1997). Fitting Time Series Models to Nonstationary Processes. *Ann. Statist.* **25** 1–37.

DAHLHAUS, R. (2012). Locally Stationary Processes. In *Handbook of Statistics*, (T. Subba Rao, S. Subba Rao and C. R. Rao, eds.) **30** 351–413. Elsevier.

DAHLHAUS, R. and GIRAITIS, L. (1998). On the optimal segment length for paramter estimates for locally stationary time series. *J. Time Ser. Anal.* **19** 629–655.

DAS, S. and NASON, G. P. (2016). Measuring the degree of nonstationarity of a time series. *Stat* **5** 295–305.

DAUBECHIES, I. (1992). *Ten Lectures on Wavelets*. SIAM, Philadelphia.

DEGERINE, S. and LAMBERT, S. (1996). Evolutive instantaneous spectrum associated with the partial autocorrelation function for nonstationary time se-

ries. In *IEEE Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis* 457–460.

DEGERINE, S. and LAMBERT-LACROIX, S. (2003). Characterization of the partial autocorrelation function of nonstationary time series. *J. Mult. Anal.* **87** 46–59.

DHAKAL, N., JAIN, S., GRAY, A., DANDY, M. and STANCIOFF, E. (2015). Nonstationarity in seasonality of extreme precipitation: A nonparametric circular statistical approach and its application. *Water Resources Research* **51** 4499–4515.

DUNFORD, N. and SCHWARTZ, J. (1958). *Linear operators. Part I: General Theory.* Wiley, New York.

FAN, J. and YAO, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer, New York.

FIECAS, M. and OMBAO, H. (2016). Modeling the evolution of dynamic brain processes during an associative learning experiment. *J. Am. Statist. Ass.* **111** 1440–1453.

FRYZLEWICZ, P. Z. (2003). Wavelet Techniques for Time Series and Poisson Data PhD thesis, University of Bristol, U.K.

FRYZLEWICZ, P., VAN BELLEGEM, S. and VON SACHS, R. (2003). Forecasting non-stationary time series by wavelet process modelling. *Ann. Inst. Statist. Math.* **55** 737–764.

GARCIN, M. (2017). Estimation of time-dependent Hurst exponents with variational smoothing and application to forecasting foreign exchange rates. *Physica A: Statistical Mechanics and its Applications* **483** 462–479.

HARGREAVES, J. K., KNIGHT, M. I., PITCHFORD, J. W., OAKENFULL, R. and DAVIS, S. J. (2018). Clustering nonstationary circadian plant rhythms using locally stationary wavelet representations. *Multiscale Model. and Simul.* **16** 184–214.

HIPEL, K. W. and MCLEOD, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems. Developments in Water Science.* Elsevier Science.

HYNDMAN, R. J. and WAND, M. P. (1997). Nonparametric Autocovariance Function Estimation. *Aust. NZ J. Stat.* **39** 313–324.

KREISS, J. P. and PAPARODITIS, E. (2014). Bootstrapping locally stationary processes. *J. R. Statist. Soc.* B **77** 267–290.

NASON, G. P. (2013a). locits: Tests of stationarity and localized autocovariance R package version 1.4.

NASON, G. P. (2013b). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *J. R. Statist. Soc.* B **75** 879–904.

NASON, G. (2013c). wavethresh: Wavelets statistics and transforms. R package version 4.6.6.

NASON, G. P., VON SACHS, R. and KROISANDT, G. (2000). Wavelet Processes and Adaptive Estimation of the Evolutionary Wavelet Spectrum. *J. R. Statist. Soc.* B **62** 271-292.

NEUMANN, M. and VON SACHS, R. (1997). Wavelet thresholding in anisotropic

function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.* **25** 38–76.

OMBAO, H. C., RAZ, J., VON SACHS, R. and GUO, W. (2002). The SLEX Model of Non-Stationary Random Processes. *Ann. Inst. Statist. Math.* **54** 171–200.

PAGE, C. H. (1952). Instantaneous power spectra. *J. Appl. Phys.* **23** 103–106.

PRIESTLEY, M. B. (1965). Evolutionary Spectra and Non-Stationary Processes. *J. R. Statist. Soc.* B **27** 204–237.

PRIESTLEY, M. B. (1983). *Spectral Analysis and Time Series.* Academic Press, London.

PRIESTLEY, M. B. and SUBBA RAO, T. (1969). A Test for Stationarity of Time Series. *J. R. Statist. Soc.* B **31** 140–149.

RAO, A. R., HAMED, K. H. and CHEN, H. L. (2012). *Nonstationarities in Hydrologic and Environmental Time Series. Water Science and Technology Library.* Springer Netherlands.

SCHLÜTER, S. and DEUSCHLE, C. (2010). Using wavelets for time series forecasting: Does it pay off? *IWQW Discussion Papers* **04**.

SILVERMAN, R. A. (1957). Locally stationary random processes. *IRE Trans. Information Theory* **IT-3** 182–187.

TONG, H. (1974). On time dependent linear transformations of non-stationary stochastic processes. *J. Appl. Prob.* **11** 53–62.

WALTER, G. G. and SHEN, X. (2000). *Wavelets and other orthogonal systems.* CRC Press, Boca Raton, FL.

WHITTLE, P. (1963). Recursive relations for predictors of non-stationary processes. *J. R. Statist. Soc.* B **27** 523–532.

YANG, W. H., HOLAN, S. H. and WIKLE, C. K. (2016). Bayesian lattice filters for time-varying autoregression and time–frequency analysis. *Bayesian Anal.* **11** 977–1003.

ZHAO, Z. (2015). Inference for local autocorrelations in locally stationary models. *J. Bus. Econ. Stat.* **33** 296–306.

Appendix A: Résumé: partial autocorrelation for stationary series

Let $\{X_t\}_{t \in \mathbb{Z}}$ be a zero-mean second-order stationary process with autocovariance function $\gamma(\tau)$. Loosely speaking, the partial autocorrelation function at lag τ is the correlation between X_1 and $X_{\tau+1}$ whilst adjusting for the "in-between" observations, $X_2, \ldots, X_\tau$. Brockwell and Davis (1991, p. 54) define the closed span $\overline{\mathrm{sp}}(X_t, t \in H)$ of any subset $\{X_t, t \in H\}$ of a Hilbert space $\mathcal{H}$ to be the smallest closed subspace of $\mathcal{H}$ which contains each $X_t, t \in H$. Then, following Brockwell and Davis (1991, p. 98), the lag τ partial autocorrelation function $q(\tau)$ of $\{X_t\}$ is defined by $q(\tau) = \mathrm{corr}\{X_{\tau+1} - P_{1,\tau}(X_{\tau+1}), X_1 - P_{1,\tau}(X_1)\}$, where $P_{1,\tau}(\cdot)$ denotes the projection operator onto $\overline{\mathrm{sp}}(X_2, \ldots, X_\tau)$. See also Fan and Yao (2003, p. 43)

Alternatively, if $\gamma(0) > 0$ and $\gamma(h) \to 0$ as $h \to \infty$, then the partial autocorrelation function, $q(\tau)$, can be obtained as the final entry of the vector φ_τ

which is the solution to the well-known Yule-Walker equations $\Gamma_\tau \boldsymbol{\varphi}_\tau = \boldsymbol{\gamma}_\tau$. Here $\Gamma_\tau = \{\gamma(i-j)\}_{i,j=1}^\tau$ is a $\tau \times \tau$ covariance matrix and $\boldsymbol{\gamma}_\tau = \{\gamma(i)\}_{i=1}^\tau$ is a vector of covariances. Equivalently, $q(\tau) = \varphi_{\tau,\tau}$ where $\varphi_{\tau,\tau}$ is the coefficient of X_1 when projecting $X_{\tau+1}$ on the space spanned by $X_1, \ldots, X_\tau$, i.e. the projection $\hat{X}_{\tau+1} = \varphi_{\tau,1} X_\tau + \ldots + \varphi_{\tau,\tau} X_1$.

For a sampled series $\{X_t\}_{t=1}^T$, the sample partial autocorrelation at lag τ is often estimated by solving $\hat{\Gamma}_\tau \hat{\boldsymbol{\varphi}}_\tau = \hat{\boldsymbol{\gamma}}_\tau$, where $\hat{\gamma}$ are the usual sample autocovariances, and taking $\hat{q}(\tau) := \hat{q}_{[1,T]}(\tau) = \hat{\varphi}_{\tau,\tau}$. Here we use the index notation $[1, T]$ in order to indicate the range of observations on which the estimation of $\hat{\Gamma}_\tau$ and $\hat{\boldsymbol{\gamma}}_\tau$ is based. The properties of $\hat{q}(\cdot)$ are well-known, see Brockwell and Davis (1991, Section 8.10). In particular, $T^{1/2}\{\hat{q}(\tau) - q(\tau)\}$ has a limiting Gaussian distribution, as $T \to \infty$, with mean zero and variance proportional to the last term on the diagonal of Γ_τ^{-1}.

Appendix B: Definitions of locally stationary processes

Definition 2.1 of Dahlhaus (1997) is as follows.

"A sequence of stochastic processes $X_{t,T}$ $(t = 0, \ldots, T - 1)$ is called locally stationary with transfer function A^0 and trend μ if there exists a representation

$$X_{t,T} = \mu(t/T) + \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^0(\lambda) d\xi(\lambda), \qquad (21)$$

where the following holds.

(i) $\xi(\lambda)$ is a stochastic process on $[-\pi, \pi]$ with $\overline{\xi(\lambda)} = \xi(-\lambda)$ and

$$\text{cum}\{d\xi(\lambda_1), \ldots, d\xi(\lambda_k)\} = \eta\left(\sum_{j=1}^{k} \lambda_j\right) g_k(\lambda_1, \ldots, \lambda_{k-1}) d\lambda_1 \ldots d\lambda_k,$$

where $\text{cum}\{\cdots\}$ denotes the cumulant of kth order, $g_1 = 0$, $g_2(\lambda) = 1$, $|g_k(\lambda_1, \ldots, \lambda_{k-1})| \leq \text{const}_k$ for all k and $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ is the periodic 2π extension of the Dirac delta function.

(ii) There exists a constant K and a 2π-periodic function $A : [0, 1] \times \mathbb{R} \to \mathbb{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$ and

$$\sup_{t,\lambda} \left| A_{t,T}^0(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1} \qquad (22)$$

for all T; $A(u, \lambda)$ and $\mu(u)$ are assumed to be continuous in u."

Definition 1 of Nason, von Sachs and Kroisandt (2000), including improvements from Fryzlewicz (2003), is as follows.

"The locally stationary wavelet processes are a sequence of doubly indexed stochastic processes $\{X_{t,T}\}_{t=0,\ldots,T-1}$, $T = 2^J \geq 1$, having the representation in the mean-square sense

$$X_{t,T} = \sum_{j=1}^{J} \sum_{k} w_{j,k;T}^0 \psi_{j,k}(t) \xi_{j,k}, \qquad (23)$$

where $\xi_{j,k}$ is a random orthonormal increment sequence and where $\{\psi_{j,k}(t)\}_{j,k}$ is a discrete non-decimated family of wavelets based on a mother wavelet $\psi(t)$ of compact support. The quantities in (23) have the following properties:

(a) $\mathbb{E}(\xi_{j,k}) = 0$ for all j, k. Hence $E(X_{t,T}) = 0$ for all t, T,

(b) $\mathrm{cov}\left(\xi_{j,k}, \xi_{\ell,m}\right) = \delta_{j,\ell}\delta_{k,m}$ where $\delta_{i,j}$ is the Kronecker delta.

(c) There exists, for each $j \geq 1$ a Lipschitz continuous function $W_j(z)$ for $z \in (0,1)$ which fulfils the following properties:

$$\sum_{j=1}^{\infty} |W_j(z)|^2 < \infty \quad \text{uniformly in } z \in (0,1); \tag{24}$$

the Lipschitz constants L_j are uniformly bounded in j and

$$\sum_{j=1}^{\infty} 2^j L_j < \infty; \tag{25}$$

there exists a sequence of constant C_j such that for each T

$$\sup_k \left| w_{j,k;T}^0 - W_j(k/T) \right| \leq C_j/T, \tag{26}$$

where $\{C_j\}$ fulfils $\sum_{j=1}^{\infty} C_j < \infty$."

Appendix C: Miscellaneous covariance matrices

The $\tau \times \tau$ covariance matrices are given by

$$\Sigma_{[zT];T}^{(b)} = \begin{pmatrix} \mathrm{Cov}(X_{[zT],T}, X_{[zT],T}) & \cdots & \mathrm{Cov}(X_{[zT],T}, X_{[zT]+\tau-1,T}) \\ \vdots & \cdots & \vdots \\ \mathrm{Cov}(X_{[zT]+\tau-1,T}, X_{[zT],T}) & \cdots & \mathrm{Cov}(X_{[zT]+\tau-1,T}, X_{[zT]+\tau-1,T}) \end{pmatrix}$$

and

$$\Sigma_{[zT]+\tau;T}^{(f)} = \begin{pmatrix} \mathrm{Cov}(X_{[zT]+1,T}, X_{[zT]+1,T}) & \cdots & \mathrm{Cov}(X_{[zT]+1,T}, X_{[zT]+\tau,T}) \\ \vdots & \cdots & \vdots \\ \mathrm{Cov}(X_{[zT]+\tau,T}, X_{[zT]+1,T}) & \cdots & \mathrm{Cov}(X_{[zT]+\tau,T}, X_{[zT]+\tau,T}) \end{pmatrix}.$$

$$\Sigma_{[zT];T} = \begin{pmatrix} \mathrm{Cov}(X_{[zT]+\tau-1,T}, X_{[zT]+\tau-1,T}) & \cdots & \mathrm{Cov}(X_{[zT],T}, X_{[zT]+\tau-1,T}) \\ \vdots & \cdots & \vdots \\ \mathrm{Cov}(X_{[zT]+\tau-1,T}, X_{[zT],T}) & \cdots & \mathrm{Cov}(X_{[zT],T}, X_{[zT],T}) \end{pmatrix}$$

Appendix D: Proofs from Section 2.

D.1. Proof of Proposition 2.2

Proof. We shall use the notation $P_{[zT],\tau}(X_{[zT],T}) = \hat{X}_{[zT],T}^{(b)}$ and $P_{[zT],\tau}(X_{[zT]+\tau,T}) = \hat{X}_{[zT]+\tau,T}^{(f)}$, since these are the linear predictors of $X_{[zT],T}$ (*b*ack-casted), respectively $X_{[zT]+\tau,T}$ (*f*orecasted), using the set of predictors $X_{[zT]+1,T}, \ldots, X_{[zT]+\tau-1,T}$.

Decomposing the projection space $\overline{\mathrm{sp}}(X_{[zT],T}, \ldots, X_{[zT]+\tau-1,T})$ into $\overline{\mathrm{sp}}(X_{[zT]+1,T},$
$\ldots,$
$X_{[zT]+\tau-1,T})$ and its orthogonal complement, we can also write $\hat{X}_{[zT]+\tau,T}$ as

$$\hat{X}_{[zT]+\tau,T} = \hat{X}_{[zT]+\tau,T}^{(f)} + P_{[zT],\tau}^{\perp}(X_{[zT]+\tau,T}), \tag{27}$$

where $P_{[zT],\tau}^{\perp}(\cdot)$ denotes the projection onto the orthogonal complement space above. Since this space is $\overline{\mathrm{sp}}(X_{[zT],T}-\hat{X}_{[zT],T}^{(b)})$, it then follows that $P_{[zT],\tau}^{\perp}(X_{[zT]+\tau,T}) = a(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)})$, and using equation (27) we obtain

$$\mathrm{cov}\left(\hat{X}_{[zT]+\tau,T}, X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right) = a\,\mathrm{var}\left(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right).$$

Due to orthogonality of projection spaces, we have $\left(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right) \perp \overline{\mathrm{sp}}(X_{[zT]+1,T}, \ldots, X_{[zT]+\tau-1,T})$ and from the equation above and equation (3), it follows that $a = \varphi_{[zT],\tau,\tau;T}$.

Hence $P_{[zT],\tau}^{\perp}(X_{[zT]+\tau,T})) = \varphi_{[zT],\tau,\tau;T}(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)})$ and we obtain

$$
\begin{aligned}
\varphi_{[zT],\tau,\tau;T} &= \frac{\mathrm{cov}\left(P_{[zT],\tau}^{\perp}(X_{[zT]+\tau,T})), X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)}{\mathrm{var}\left(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)} \\[2mm]
&= \frac{\mathrm{cov}\left(X_{[zT]+\tau,T}, X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)}{\mathrm{var}\left(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)},
\end{aligned}
$$

as $\mathrm{cov}(X_{[zT]+\tau,T} - P_{[zT],\tau}^{\perp}(X_{[zT]+\tau,T}), Y) = 0, \forall Y \in \overline{\mathrm{sp}}(X_{[zT],T}-\hat{X}_{[zT],T}^{(b)})$. Hence,

$$\varphi_{[zT],\tau,\tau;T} = \frac{\mathrm{cov}\left(X_{[zT]+\tau,T} - \hat{X}_{[zT]+\tau,T}^{(f)}, X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)}{\mathrm{var}\left(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)} \tag{28}$$

because $\hat{X}_{[zT]+\tau,T}^{(f)} \perp \overline{\mathrm{sp}}(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)})$.

Recall from equation (1) that

$$q_T(z,\tau) = \mathrm{corr}\left(X_{[zT]+\tau,T} - P_{[zT],\tau}(X_{[zT]+\tau,T}), X_{[zT],T} - P_{[zT],\tau}(X_{[zT],T})\right),$$

or equivalently $q_T(z,\tau) = \mathrm{corr}\left(X_{[zT]+\tau,T} - \hat{X}_{[zT]+\tau,T}^{(f)}, X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)$ which, combined with equation (28), yields

$$q_T(z,\tau) = \varphi_{[zT],\tau,\tau;T}\left\{\frac{\mathrm{var}\left(X_{[zT],T} - \hat{X}_{[zT],T}^{(b)}\right)}{\mathrm{var}\left(X_{[zT]+\tau,T} - \hat{X}_{[zT]+\tau,T}^{(f)}\right)}\right\}^{1/2}$$

as desired.

D.2. Proof of Proposition 2.3

Proof. As $\hat{X}_{[zT],T}^{(b)}$ and $\hat{X}_{[zT]+\tau,T}^{(f)}$ are projections of $X_{[zT],T}$ and $X_{[zT]+\tau,T}$, respectively, on the space $\overline{\mathrm{sp}}(X_{[zT]+1,T}, \ldots, X_{[zT]+\tau-1,T})$ it follows that $\mathbb{E}(X_{[zT],T}-$

$\hat{X}^{(b)}_{[zT],T}) = 0$ and $\mathbb{E}(X_{[zT]+\tau,T} - \hat{X}^{(f)}_{[zT]+\tau,T}) = 0$. Hence, the numerator and denominator in (2) can be re-expressed as a Mean Squared Prediction Error (MSPE), since

$$
\begin{aligned}
\text{var}(X_{[zT],T} - \hat{X}^{(b)}_{[zT],T}) &= \mathbb{E}(\hat{X}^{(b)}_{[zT],T} - X_{[zT],T})^2 = \text{MSPE}(\hat{X}^{(b)}_{[zT],T}, X_{[zT],T}), \\
\text{var}(X_{[zT]+\tau,T} - \hat{X}^{(f)}_{[zT]+\tau,T}) &= \mathbb{E}(\hat{X}^{(f)}_{[zT]+\tau,T} - X_{[zT]+\tau,T})^2 \\
&= \text{MSPE}(\hat{X}^{(f)}_{[zT]+\tau,T}, X_{[zT]+\tau,T}).
\end{aligned}
$$

Using these expressions we can rewrite $q_T(z,\tau)$ from (2) as

$$
q_T(z,\tau) = \varphi_{[zT],\tau,\tau;T} \left\{ \frac{\text{MSPE}(\hat{X}^{(b)}_{[zT],T}, X_{[zT],T})}{\text{MSPE}(\hat{X}^{(f)}_{[zT]+\tau,T}, X_{[zT]+\tau,T})} \right\}^{1/2}.
$$

Now use the fact that the MSPE of a linear predictor of $X_{t,T}$ can be written as

$$
\text{MSPE}\left(\hat{X}_{t,T}, X_{t,T}\right) = \mathbb{E}\left(\hat{X}_{t,T} - X_{t,T}\right)^2 = \boldsymbol{b}_t^T \Sigma_{t,T} \boldsymbol{b}_t,
$$

where $\boldsymbol{b}_t = (b_{t-1,T}, \dots, b_{0,T}, -1)^T$ and $\Sigma_{t,T}$ is the covariance of $X_{0,T}, \dots, X_{t,T}$ (Fryzlewicz, Van Bellegem and von Sachs, 2003, Section 3.3). In our case, the back-casted and forecasted values of $X_{[zT],T}$ and $X_{[zT]+\tau,T}$ are also linear predictors using the window of observations $X_{[zT]+1,T}, \dots, X_{[zT]+\tau-1,T}$, and thus their corresponding MSPE can be expressed as

$$
\text{MSPE}(\hat{X}^{(b)}_{[zT],T}, X_{[zT],T}) = (\boldsymbol{b}^{(b)}_{[zT]})^T \Sigma^{(b)}_{[zT];T} \boldsymbol{b}^{(b)}_{[zT]}, \tag{29}
$$

$$
\text{MSPE}(\hat{X}^{(f)}_{[zT]+\tau,T}, X_{[zT]+\tau,T}) = (\boldsymbol{b}^{(f)}_{[zT]+\tau})^T \Sigma^{(f)}_{[zT]+\tau;T} \boldsymbol{b}^{(f)}_{[zT]+\tau}, \tag{30}
$$

where, as above, the $\tau \times 1$ coefficient vectors are $\boldsymbol{b}^{(b)}_{[zT]} = (-1, \tilde{b}^{(b)}_{1,T}, \dots, \tilde{b}^{(b)}_{\tau-1,T})^T$ and $\boldsymbol{b}^{(f)}_{[zT]+\tau} = (b^{(f)}_{\tau-2,T}, \dots, b^{(f)}_{0,T}, -1)^T$ and the $\tau \times \tau$ covariance matrices $\Sigma^{(b)}_{[zT];T}$ and $\Sigma^{(f)}_{[zT]+\tau;T}$ appear in Appendix C.

Therefore, on combining equation (4) with (29) and (30) we obtain as desired

$$
q_T(z,\tau) = \varphi_{[zT],\tau,\tau;T} \left\{ \frac{(\boldsymbol{b}^{(b)}_{[zT]})^T \Sigma^{(b)}_{[zT];T} \boldsymbol{b}^{(b)}_{[zT]}}{(\boldsymbol{b}^{(f)}_{[zT]+\tau})^T \Sigma^{(f)}_{[zT]+\tau;T} \boldsymbol{b}^{(f)}_{[zT]+\tau}} \right\}^{1/2}.
$$

D.3. Proof of Proposition 2.5

Proof. The proof treats the convergence of $\varphi_{[zT],\tau,\tau;T}$ and the quotient that forms the square-root in equation (4) separately. Firstly, we address the quotient convergence.

A: Quotient Convergence: By Proposition 3.1 from Fryzlewicz, Van Bellegem and von Sachs (2003) we have

$$\text{MSPE}\left(\hat{X}^{(b)}_{[zT],T}, X_{[zT],T}\right) = (\boldsymbol{b}^{(b)}_{[zT]})^T \Sigma^{(b)}_{[zT];T} \boldsymbol{b}^{(b)}_{[zT]} = \left\{(\boldsymbol{b}^{(b)}_{[zT]})^T \text{B}^{(b)}_{[zT]} \boldsymbol{b}^{(b)}_{[zT]}\right\} \{1 + o_T(1)\}$$

and

$$\begin{aligned}\text{MSPE}\left(\hat{X}^{(f)}_{[zT]+\tau,T}, X_{[zT]+\tau,T}\right) &= (\boldsymbol{b}^{(f)}_{[zT]+\tau})^T \Sigma^{(f)}_{[zT]+\tau;T} \boldsymbol{b}^{(f)}_{[zT]+\tau} \\ &= \left\{(\boldsymbol{b}^{(f)}_{[zT]+\tau})^T \text{B}^{(f)}_{[zT]+\tau} \boldsymbol{b}^{(f)}_{[zT]+\tau}\right\} \{1 + o_T(1)\}.\end{aligned}$$

Hence

$$\frac{\text{MSPE}\left(\hat{X}^{(b)}_{[zT],T}, X_{[zT],T}\right)}{\text{MSPE}\left(\hat{X}^{(f)}_{[zT]+\tau,T}, X_{[zT]+\tau,T}\right)} = \frac{\left\{(\boldsymbol{b}^{(b)}_{[zT]})^T \text{B}^{(b)}_{[zT]} \boldsymbol{b}^{(b)}_{[zT]}\right\}}{\left\{(\boldsymbol{b}^{(f)}_{[zT]+\tau})^T \text{B}^{(f)}_{[zT]+\tau} \boldsymbol{b}^{(f)}_{[zT]+\tau}\right\}} \{1 + o_T(1)\}$$

B: Convergence of $\varphi_{[zT],\tau,\tau;T}$. We defined $\varphi_{[zT],\tau,\tau;T}$ as the last element in $\varphi_{[zT],\tau;T}$ that is the solution to the Yule-Walker equations $\Sigma_{[zT];T}\varphi_{[zT],\tau;T} = \boldsymbol{r}_{[zT];T}$.

Let $[\mathcal{O}(T^{-1})]$ be an appropriately-sized matrix whose elements are all $\mathcal{O}(T^{-1})$ and $\boldsymbol{\mathcal{O}(T^{-1})}$ be a similarly defined vector.

Now consider

$$\begin{aligned}\Sigma_{[zT];T}\left(\varphi_{[zT],\tau;T} - \varphi_{[zT],\tau}\right) &= \Sigma_{[zT];T}\varphi_{[zT],\tau;T} - \Sigma_{[zT];T}\varphi_{[zT],\tau} \\ &= \boldsymbol{r}_{[zT];T} - \left(\text{B}_{[zT]} + [\mathcal{O}(T^{-1})]\right)\varphi_{[zT],\tau} \\ &= \boldsymbol{r}_{[zT];T} - \boldsymbol{r}_{[zT]} - [\mathcal{O}(T^{-1})]\varphi_{[zT],\tau} \\ &\qquad (\text{as } |c_T(z,\tau) - c(z,\tau)| = \mathcal{O}(T^{-1})) \\ &= \boldsymbol{\mathcal{O}(T^{-1})} - [\mathcal{O}(T^{-1})]\varphi_{[zT],\tau}.\end{aligned}$$

Observe that $[\mathcal{O}(T^{-1})]\varphi_{[zT],\tau} = \left(\sum_j \sum_i K_{j,i}\varphi_{[zT],\tau,j}/T\right)\boldsymbol{1}$, where $K_{j,i}$ is the (j,i)th constant, and in what follows we shall seek to bound this quantity.

From the Cauchy-Schwarz inequality $\|\varphi_{[zT],\tau}\|_1 \leq |\tau|^{1/2}\|\varphi_{[zT],\tau}\|_2 = C_\tau$ as τ is fixed, and by standard properties of the spectral norm

$$\begin{aligned}\|\varphi_{[zT],\tau}\|_2^2 &= \varphi_{[zT],\tau}^T \varphi_{[zT],\tau} \\ &\leq \varphi_{[zT],\tau}^T \Sigma_{[zT];T}\varphi_{[zT],\tau}\|\Sigma^{-1}_{[zT];T}\| \\ &= \text{MSPE}\left(\hat{X}_{[zT]+\tau-1,T}, X_{[zT]+\tau-1,T}\right)\|\Sigma^{-1}_{[zT];T}\| < \infty,\end{aligned}$$

as the spectral norm $\|\Sigma^{-1}_{[zT];T}\|$ is bounded using Lemma A.3 from Fryzlewicz, Van Bellegem and von Sachs (2003).

Thus $[\mathcal{O}(T^{-1})]\varphi_{[zT],\tau} = \boldsymbol{\mathcal{O}(T^{-1})}$ and it follows that

$$\Sigma_{[zT];T}\left(\varphi_{[zT],\tau;T} - \varphi_{[zT],\tau}\right) = \boldsymbol{\mathcal{O}(T^{-1})},$$

which is equivalent to $\boldsymbol{\varphi}_{[zT],\tau;T} - \boldsymbol{\varphi}_{[zT],\tau} = \Sigma_{[zT];T}^{-1}\boldsymbol{\mathcal{O}(T^{-1})}$. By bounds of Rayleigh quotients (Abadir and Magnus, 2005, pg.181), $\|\boldsymbol{\varphi}_{[zT],\tau;T} - \boldsymbol{\varphi}_{[zT],\tau}\|_2$ $= \|\Sigma_{[zT];T}^{-1}\boldsymbol{\mathcal{O}(T^{-1})}\|_2 \leq \mu^{1/2}\|\boldsymbol{\mathcal{O}(T^{-1})}\|_2 = (\mu|\tau|)^{1/2}/T = \mathcal{O}(T^{-1})$ as τ is fixed. Here μ is the largest eigenvalue of $\left(\Sigma_{[zT];T}^{-1}\right)^T \Sigma_{[zT];T}^{-1}$, i.e. $\mu = \|\Sigma_{[zT];T}^{-1}\|^2$, and so $\mu < \infty$. It follows that $\boldsymbol{\varphi}_{[zT],\tau;T} - \boldsymbol{\varphi}_{[zT],\tau} = \boldsymbol{\mathcal{O}(T^{-1})}$.

Putting parts A and B together:

$$
\begin{aligned}
|q_T(z,\tau) - q(z,\tau)| = \Bigg| & \varphi_{[zT],\tau;T} \left\{ \frac{\mathrm{MSPE}\left(\hat{X}_{[zT],T}^{(b)}, X_{[zT],T}\right)}{\mathrm{MSPE}\left(\hat{X}_{[zT]+\tau,T}^{(f)}, X_{[zT]+\tau,T}\right)} \right\}^{1/2} \\
& -\varphi_{[zT],\tau,\tau} \left\{ \frac{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}}{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}} \right\}^{1/2} \Bigg| \\
= \Bigg| & \left\{\varphi_{[zT],\tau,\tau} + \mathcal{O}(T^{-1})\right\} \\
& \left[\frac{\left\{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}\right\}\{1 + o_T(1)\}}{\left\{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}\right\}\{1 + o_T(1)\}} \right]^{1/2} \\
& -\varphi_{[zT],\tau,\tau} \left\{ \frac{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}}{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}} \right\}^{1/2} \Bigg| \\
= \Bigg| & \varphi_{[zT],\tau,\tau} \left\{ \frac{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}}{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}} \right\}^{1/2} \\
& \left(\left[\frac{\{1 + o_T(1)\}}{\{1 + o_T(1)\}} \right]^{1/2} - 1 \right) \\
& +\mathcal{O}(T^{-1}) \left\{ \frac{\mathrm{MSPE}\left(\hat{X}_{[zT],T}^{(b)}, X_{[zT],T}\right)}{\mathrm{MSPE}\left(\hat{X}_{[zT]+\tau,T}^{(f)}, X_{[zT]+\tau,T}\right)} \right\}^{1/2} \Bigg| \\
= & \ \mathcal{O}(T^{-1}).
\end{aligned}
$$

For the last equality the first term is asymptotically zero since $\frac{\{1+o_T(1)\}}{\{1+o_T(1)\}} \to 1$ as $T \to \infty$, $\varphi_{[zT],\tau,\tau} < \infty$ and

$$
\left\{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}\right\}^{1/2} \left\{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}\right\}^{-1/2} < \infty,
$$

or, more concisely, $q(z,\tau) < \infty$. The second term is $\mathcal{O}(T^{-1})$ as each expectation is finite. This concludes the proof.

D.4. Proof of Proposition 2.7

Proof. First recall that we defined the local partial autocorrelation as

$$q\left(z,\tau\right) = \varphi_{[zT],\tau,\tau;T} \left\{ \frac{\mathrm{Var}\{X_{[zT],T} - P_{[zT],\tau}(X_{[zT],T})\}}{\mathrm{Var}\{X_{[zT]+\tau,T} - P_{[zT],\tau}(X_{[zT]+\tau,T})\}} \right\}^{1/2},$$

where the coefficient $\varphi_{[zT],\tau,\tau;T}$ is obtained in a manner akin to the (stationary) partial autocorrelation coefficient by expressing $X_{[zT]+\tau,T}$ as an AR(τ) process and solving the associated Yule-Walker equations. The fraction under the square root quantifies the ratio between the backward and forward variances associated to the AR(τ) process. The Yule-Walker equations here are localized at the rescaled time z, in the sense that they involve observations over the interval $[[zT],[zT]+\tau]$.

Recall that, in estimating the local partial autocorrelation, we use the $\hat{c}(z,\tau)$ estimator of Nason, von Sachs and Kroisandt (2000), which was shown there to be consistent for the (true) local autocovariance $c(z,\tau)$. By the classical stationary theory, it follows that the estimated Yule-Walker coefficients of the AR(τ) process (solution vector to the local Yule-Walker equations) are consistent estimators of the true coefficients, hence $\tilde{\varphi}_{[zT],\tau,\tau;T} \xrightarrow{P} \varphi_{[zT],\tau,\tau}$, and the forward and backward variances are also estimated consistently.

Using the continuous mapping theorem (Billingsley, 1999) and assuming that the variance is non-zero, it follows that the square-root of the ratio of estimated backward and forward variances

$$\left\{ \frac{(\tilde{\boldsymbol{b}}_{[zT]}^{(b)})^T \tilde{\mathrm{B}}_{[zT]}^{(b)} \tilde{\boldsymbol{b}}_{[zT]}^{(b)}}{(\tilde{\boldsymbol{b}}_{[zT]+\tau}^{(f)})^T \tilde{\mathrm{B}}_{[zT]+\tau}^{(f)} \tilde{\boldsymbol{b}}_{[zT]+\tau}^{(f)}} \right\}^{1/2}$$

is a consistent estimator of the true ratio of variances

$$\left\{ \frac{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}}{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}} \right\}^{1/2}.$$

This together with the consistency of $\tilde{\varphi}_{[zT],\tau,\tau;T}$, yields

$$
\begin{aligned}
&\tilde{\varphi}_{[zT],\tau,\tau;T} \left\{ \frac{(\tilde{\boldsymbol{b}}_{[zT]}^{(b)})^T \tilde{\mathrm{B}}_{[zT]}^{(b)} \tilde{\boldsymbol{b}}_{[zT]}^{(b)}}{(\tilde{\boldsymbol{b}}_{[zT]+\tau}^{(f)})^T \tilde{\mathrm{B}}_{[zT]+\tau}^{(f)} \tilde{\boldsymbol{b}}_{[zT]+\tau}^{(f)}} \right\}^{1/2} \xrightarrow{P} \\
&\varphi_{[zT],\tau,\tau} \left\{ \frac{(\boldsymbol{b}_{[zT]}^{(b)})^T \mathrm{B}_{[zT]}^{(b)} \boldsymbol{b}_{[zT]}^{(b)}}{(\boldsymbol{b}_{[zT]+\tau}^{(f)})^T \mathrm{B}_{[zT]+\tau}^{(f)} \boldsymbol{b}_{[zT]+\tau}^{(f)}} \right\}^{1/2}.
\end{aligned}
\tag{31}
$$

Appendix E: Proofs from Section 3.

First note that by substituting $s = [zT] - t$ in (13) and by denoting the rectangular kernel by h, we obtain:

$$i_{N,z}(j, \ell, k) = \sum_{s=[zT]-N+1}^{[zT]} \psi_{j,s} \psi_{\ell, s+k-2[zT]+N/2-1} h\left(\frac{[zT]-s}{N}\right). \quad (32)$$

The above expression is by no means restricted to h being a rectangular kernel, and other kernels may be used, as explained in the article main text.

This new formulation of $i_{N,z}$ is very similar to that of the cross-correlation wavelet $\Psi_{j,\ell}$, except that the summation limits are $[zT - N + 1]$ and $[zT]$ instead of $-\infty$ and ∞. This similarity is true for all Daubechies' compactly supported wavelets. We shall use this similarity to bound $i_{N,z}$ using Lemma 1.

E.1. Cross-scale autocorrelation Haar wavelets

Proposition E.1. *The cross-scale autocorrelation Haar wavelets $\Psi_{j,\ell}(\cdot)$ at scales $\ell < j$ are given by*

$$\Psi_{j,\ell}(\tau) = 2^{-(j-\ell)/2} \begin{cases} 0 & \text{for } \tau < -2^\ell, \\ -(2^{-\ell}\tau + 1) & \text{for } -2^\ell \leq \tau < -2^{\ell-1}, \\ 2^{-\ell}\tau & \text{for } -2^{\ell-1} \leq \tau < 0, \\ 0 & \text{for } 0 \leq \tau < 2^{j-1} - 2^\ell, \\ 2^{-\ell}(2\tau - 2^j + 2^{\ell+1}) & \text{for } 2^{j-1} - 2^\ell \leq \tau < 2^{j-1} - 2^{\ell-1}, \\ 2^{-\ell}(2^j - 2\tau) & \text{for } 2^{j-1} - 2^{\ell-1} \leq \tau < 2^{j-1}, \\ 0 & \text{for } 2^{j-1} \leq \tau < 2^j - 2^\ell, \\ 2^{-\ell}(2^j - \tau - 2^\ell) & \text{for } 2^j - 2^\ell \leq \tau < 2^j - 2^{\ell-1}, \\ 2^{-\ell}(\tau - 2^j) & \text{for } 2^j - 2^{\ell-1} \leq \tau < 2^j, \\ 0 & \text{for } 2^j \leq \tau. \end{cases}$$

$$(33)$$

For $\ell > j$ we have $\Psi_{j,\ell}(\tau) = \Psi_{\ell,j}(-\tau)$ and a precise formula appears in equation (35) (Appendix E.1). Note $\Psi_{j,j}(\tau) = \Psi_j(\tau)$, the regular autocorrelation wavelet.

Proof. For completeness, the definition of (continuous-time) Haar wavelets is

$$\psi_H(x) = \begin{cases} -1 & \text{if } 0 \leq x < 1/2, \\ 1 & \text{if } 1/2 \leq x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Nason, von Sachs and Kroisandt (2000) show that $\Psi_j(\tau) = \Psi_H(2^{-j}|\tau|)$ where $\Psi_j(\tau)$ is the regular discrete autocorrelation wavelet and $\Psi_H(u)$ is the continu-

ous Haar autocorrelation wavelet given by

$$\Psi_H(u) = \int_{-\infty}^{\infty} \psi_H(x)\psi_H(x-u)\,dx = \begin{cases} 1 - 3|u| & \text{for } |u| \in [0, \frac{1}{2}], \\ |u| - 1 & \text{for } |u| \in (\frac{1}{2}, 1]. \end{cases}$$

Hence, we can derive the following integral equation for $\Psi_j(\tau)$ for $\tau \geq 0$:

$$\Psi_j(\tau) = \Psi_H(2^{-j}\tau) = \int_{-\infty}^{\infty} \psi_H(x)\psi_H(x - 2^{-j}\tau)\,dx$$

from the definition of $\Psi_H(u)$. Then make the substitution $x = 2^{-j}y$ to obtain:

$$\Psi_j(\tau) = \int_{-\infty}^{\infty} \psi_H(2^{-j}y)\psi_H(2^{-j}y - 2^{-j}\tau)\,2^{-j}\,dy$$

$$= \int_{-\infty}^{\infty} 2^{-j/2}\psi_H(2^{-j}y)\,2^{-j/2}\psi_H\{2^{-j}(y - \tau)\}\,dy$$

$$= \int_{-\infty}^{\infty} \psi_{j,0}(y)\psi_{j,0}(y - \tau)\,dy,$$

where $\psi_{j,0}(y) = 2^{-j/2}\psi_H(2^{-j}y)$.

Hence, by a similar argument it is the case that

$$\Psi_{j,\ell}(\tau) = \int_{-\infty}^{\infty} \psi_{j,0}(y)\psi_{\ell,0}(y - \tau)\,dy.$$

For Haar wavelets, since we know the precise form of ψ_H we should be able to obtain an analytical formula for $\Psi_{j,\ell}$. To do this we consider $\ell < j$ and see that

$$\Psi_{j,\ell}(\tau) = \int_{-\infty}^{\infty} 2^{-j/2}\psi_H(2^{-j}y)2^{-\ell/2}\psi_H\{2^{-\ell}(y - \tau)\}\,dy.$$

Now let $x = 2^{-\ell}y$ and we obtain

$$\Psi_{j,\ell}(\tau) = \int_{-\infty}^{\infty} 2^{-j/2}\psi_H\{2^{-(j-\ell)}x\}2^{\ell/2}\psi_H(x - 2^{-\ell}\tau)\,dx$$

$$= \int_{-\infty}^{\infty} \psi_{j-\ell}(x)\psi(x - 2^{-\ell}\tau)\,dx.$$

Hence, it makes sense to introduce the following *core* function:

$$\Omega_i(u) = \int_{-\infty}^{\infty} \psi_i(x)\psi(x - u)\,dx,$$

for integers $i = 0, 1, 2, \ldots$. Clearly,

$$\Psi_{j,\ell}(\tau) = \Omega_{j-\ell}(2^{-\ell}\tau), \tag{34}$$

for $\ell < j$. Also $\Omega_0(u) = \Psi_H(u)$.

Using Lemma 4 and (34), we can now specify an exact formula for $\Psi_{j,\ell}(\tau)$. For $\ell < j$ the result is shown in (33). Corollary 2 shows the formula for $\ell > j$.

Corollary 2. The regions on the right-hand side of (33) correspond exactly to the regions III_a to III_j in (75). For completeness (and usefulness in working out derived quantities) we can write down $\Psi_{j,\ell}(\tau)$ for $\ell > j$ explicitly as

$$\Psi_{j,\ell}(\tau) = 2^{-(\ell-j)/2} \begin{cases} 0 & \text{for } \tau \leq -2^\ell, \\ -2^{-j}(\tau + 2^\ell) & \text{for } -2^\ell < \tau \leq -2^\ell + 2^{j-1}, \\ 2^{-j}(2^\ell + \tau - 2^j) & \text{for } -2^\ell + 2^{j-1} < \tau \leq -2^\ell + 2^j, \\ 0 & \text{for } -2^\ell + 2^j < \tau \leq -2^{\ell-1}, \\ 2^{-j}(2^\ell + 2\tau) & \text{for } -2^{l-1} < \tau \leq -2^{\ell-1} + 2^{j-1}, \\ 2^{-j}(2^{j+1} - 2^\ell - 2\tau) & \text{for } -2^{\ell-1} + 2^{j-1} < \tau \leq -2^{\ell-1} + 2^j, \\ 0 & \text{for } -2^{\ell-1} + 2^j < \tau \leq 0, \\ -2^{-j}\tau & \text{for } 0 < \tau \leq 2^{j-1}, \\ -(1 - 2^{-j}\tau) & \text{for } 2^{j-1} < \tau \leq 2^j. \end{cases}$$
(35)

Proposition E.1 shows that the support of the cross-correlation wavelet is $\{k \in \mathbb{Z} : -2^\ell <= k < 2^j\}$ for $\ell < j$.

E.2. Subsidiary result used in the proof of Lemma 1

Lemma 3. For a, b such that $2a, 2b \in \mathbb{N}$ and $a, b > 0$ it is the case that

$$\int_{-\pi}^{\pi} \frac{\{1 - \cos(2a\omega)\}\{1 - \cos(2b\omega)\}}{1 - \cos(\omega)} \, d\omega = 4\pi \min(a, b).$$
(36)

Proof. The result is obtained by combining known results on the Fejér and Dirichlet kernels as follows. The Fejér kernel can be defined by:

$$F_n(\omega) = \frac{\sin^2(n\omega/2)}{2\pi n \sin^2(\omega/2)} = \frac{1 - \cos(n\omega)}{2\pi n \{1 - \cos(\omega)\}},$$

for $\omega \in [-\pi, \pi]$, see Walter and Shen (2000) Section 4.2, for example. The Fejér kernel can also be written in the following alternative form

$$F_n(\omega) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(\omega),$$
(37)

where $D_k(\omega) = \pi^{-1} \left(\frac{1}{2} + \sum_{p=1}^{k} \cos p\omega \right)$ is the Dirichlet kernel, see Section 1.2.1 of Walter and Shen (2000).

Let the integral on the left-hand side of (36) be $I1$. Then:

$$I1 = 2\pi \int_{-\pi}^{\pi} \{1 - \cos(2a\omega)\} 2b F_{2b}(\omega) \, d\omega$$

$$= 4\pi b \left\{ \int_{-\pi}^{\pi} F_{2b}(\omega) \, d\omega - \int_{-\pi}^{\pi} \cos(2a\omega) F_{2b}(\omega) \, d\omega \right\}$$

$$= 4\pi b \left\{ 1 - \int_{-\pi}^{\pi} \cos(2a\omega)(2b)^{-1} \sum_{k=0}^{2b-1} D_k(\omega) \, d\omega \right\},$$

by substituting (37) and since $\int_{-\pi}^{\pi} F_n(\omega) \, d\omega = 1$. Then

$$I1 = 4\pi b \left\{ 1 - (2b)^{-1} \sum_{k=0}^{2b-1} \int_{-\pi}^{\pi} \cos(2a\omega) D_k(\omega) \, d\omega \right\}$$

$$= 4\pi b \{ 1 - (2b)^{-1} \sum_{k=0}^{2b-1} I2_k \}, \tag{38}$$

where $I2_k = \int_{-\pi}^{\pi} D_k(\omega) \cos(2a\omega) \, d\omega$. Now,

$$I2_k = \pi^{-1} \int_{-\pi}^{\pi} \left(\frac{1}{2} + \sum_{p=1}^{k} \cos p\omega \right) \cos(2a\omega) \, d\omega$$

$$= (2\pi)^{-1} \int_{-\pi}^{\pi} \cos(2a\omega) \, d\omega + \pi^{-1} \sum_{p=1}^{k} \int_{-\pi}^{\pi} \cos(p\omega) \cos(2a\omega) \, d\omega \tag{39}$$

$$= (2\pi)^{-1} \left[\frac{\sin(2a\omega)}{2a} \right]_{-\pi}^{\pi} + \pi^{-1} \sum_{p=1}^{k} \left[\frac{2a\cos(p\omega)\sin(2a\omega) - p\cos(2a\omega)\sin(p\omega)}{4a^2 - p^2} \right]_{-\pi}^{\pi}$$

$$= 0, \tag{40}$$

for $p^2 \neq 4a^2$ for $p \neq 2a$ (and recall $a > 0$). For $p = 2a$ the integral in (39) is

$$\int_{-\pi}^{\pi} \cos^2(2a\omega) \, d\omega = \pi + \sin(4a\pi)/4a = \pi, \tag{41}$$

since $2a \in \mathbb{N}$. Hence,

$$I2_k = \begin{cases} 0 & \text{for } k < 2a \\ 1 & \text{for } k \geq 2a. \end{cases} \tag{42}$$

Hence, substituting (42) into (38) gives, for $b > a$

$$I1 = 4\pi b \{ 1 - (2b)^{-1}(2b - 2a) \} = 4\pi b - 2\pi(2b - 2a) = 4\pi a. \tag{43}$$

Since the integral (36) is symmetric in a and b we also have $I1 = 4\pi b$ for $b \leq a$. Hence, the result in equation (36) follows.

E.3. Proof of Lemma 1

Proof. It is obvious that inequality (14) holds when $i_{N,z}(j, \ell, k) = 0$. This occurs when the lower limit in the sum (13), plus the extra $k - 2[zT] + N/2 - 1$ exceeds the support of $\psi_{\ell,\cdot}$. In other words, $i_{N,z}(j, \ell, k) = 0$ when:

$$
\begin{aligned}
&[zT] - N + 1 + k - 2[zT] + N/2 - 1 > N_\ell - 1 \\
&\implies k - [zT] - N/2 > N_\ell - 1 \\
&\implies k > [zT] + N/2 + N_\ell - 1 = b_2.
\end{aligned}
\tag{44}
$$

It can also be shown that $i_{N,z}(j, \ell, k) = 0$ when $k < [zT] - N/2 + 1$ but this inequality is not of interest in this proof .

For the inequalities in (15) we decompose Ψ into three terms:

$$
\Psi_{j,\ell}(k - 2[zT] + N/2 - 1) = L + i_{N,z}(j, \ell, k) + U,
\tag{45}
$$

where

$$
L = \sum_{s=-\infty}^{[zT]-N} \psi_{j,s} \psi_{\ell, s+k-2[zT]+N/2-1}
\tag{46}
$$

and

$$
U = \sum_{s=[zT]+1}^{\infty} \psi_{j,s} \psi_{\ell, s+k-2[zT]+N/2-1}.
\tag{47}
$$

Clearly, the inequality (14) is satisfied when $i_{N,z}(j, \ell, k) = \Psi_{j,\ell}(k - 2[zT] + N/2 - 1)$ which occurs when $L = U = 0$. We now investigate the conditions when $L = U = 0$.

(A) When is $U = 0$? When the lower limit of the sum defining U in (47) exceeds the support of $\psi_{j,\cdot}$, i.e.

$$
[zT] + 1 > N_j - 1 \implies [zT] > N_j - 2,
\tag{48}
$$

or when the lower limit exceeds the support of $\psi_{\ell,\cdot}$, i.e.

$$
[zT] + 1 + k - 2[zT] + N/2 - 1 > N_\ell - 1 \implies k > [zT] - N/2 + N_\ell - 1.
\tag{49}
$$

(B) When is $L = 0$? When the upper limit of the sum defining L in (46) is less than the lower support bound of $\psi_{j,\cdot}$, which is zero, i.e.

$$
[zT] - N < 0 \implies [zT] < N,
\tag{50}
$$

or when the upper limit is less than the support of $\psi_{\ell,\cdot}$, i.e.

$$
[zT] - N + k - 2[zT] + N/2 - 1 < 0 \implies k < [zT] + N/2 + 1 = b_1.
\tag{51}
$$

Hence, $U = L = 0$ when inequalities (48) and (51) are satisfied. Note: we are not particularly interested in inequalities (49) and (50). For the former, the inequality (49) would have to be allied with (50) (as (51) would be contradictory

to (49)) and, asymptotically (50) will not hold (as we expect the rate of increase of T to be much bigger than N).

So far we have demonstrated the Lemma up to inequalities (15) and (16) and now we look to establish the second part of the Lemma.

To establish (17) it can be shown that, for Daubechies' wavelets with two or more vanishing moments,

$$|i_{N,z}(j,\ell,k)| \leq 2^{-(j+\ell)/2}K^2 \sum_{s=[zT]-N+1}^{[zT]} s^{-1}(s-[zT]+N)^{-1} \qquad (52)$$
$$= 2^{-(j+\ell)/2}\left(\mathcal{H}_N - \mathcal{H}_{[zT]} + \mathcal{H}_{[zT]-N}\right)/([zT]-N),$$

where $\mathcal{H}_n$ is the nth Harmonic number, and K is a constant (maximum absolute value of the wavelet). Now using the following approximation for $\mathcal{H}_n$

$$\mathcal{H}_n = \log n + \gamma + \mathcal{O}(n^{-1}),$$

where γ is the Euler-Mascheroni constant, we can obtain the result in (17).

Now we consider Haar wavelets. First, let us recall what the discrete Haar wavelet is. We have

$$\psi_{j,k} = \begin{cases} 2^{-j/2} & \text{for } 0 \leq k < N_j/2, \\ -2^{-j/2} & \text{for } N_j/2 \leq k < N_j, \\ 0 & \text{otherwise.} \end{cases}$$

For Haar wavelets $N_j = 2^j$ for $j \in \mathbb{N}$. Next we will require the discrete Fourier transform of the Haar wavelet given by:

$$\hat{\psi}_j(\omega) = \sum_{s=-\infty}^{\infty} \psi_{j,s}e^{-i\omega s},$$

for $\omega \in (-\pi, \pi)$. The inverse of this transform is:

$$\psi_{j,s} = (2\pi)^{-1}\int_{-\pi}^{\pi} \hat{\psi}_j(\omega)e^{i\omega s}d\omega \qquad (53)$$

for $s \in \mathbb{Z}$.

Now let us work out the precise form of the Fourier transform of the discrete

Haar wavelet:

$$
\hat{\psi}_j(\omega) = 2^{-j/2}\left(\sum_{s=0}^{N_j/2-1} e^{-i\omega s} - \sum_{s=N_j/2}^{N_j-1} e^{-i\omega s}\right)
$$

$$
= 2^{-j/2}\left\{\sum_{s=0}^{N_j/2-1} e^{-i\omega s} - \sum_{s=0}^{N_j/2-1} e^{-i\omega(s+N_j/2)}\right\}
$$

$$
= 2^{-j/2}\sum_{s=0}^{N_j/2-1} e^{-i\omega s}\left(1 - e^{-i\omega N_j/2}\right)
$$

$$
= 2^{-j/2}\left(1 - e^{-i\omega N_j/2}\right)\sum_{s=0}^{N_j/2-1} e^{-i\omega s}
$$

$$
= 2^{-j/2}\left(1 - e^{-i\omega N_j/2}\right)\frac{1 - \exp(-i\omega N_j/2)}{1 - \exp(-i\omega)}
$$

$$
= 2^{-j/2}\frac{\left(1 - e^{-i\omega N_j/2}\right)^2}{1 - \exp(-i\omega)}
$$

We now directly examine formula (32) with a rectangular kernel, as discussed in the main body of the paper. To simplify notation, we let $B = [zT] - N$ and $r = k - 2[zT] + N/2 - 1$. In (32) replace the discrete wavelets $\psi_{j,s}$ and $\psi_{\ell,s+r}$ by their Fourier inverse representations given by (53) to obtain:

$$
i_{N,z}(j,\ell,k) = \sum_{s=B+1}^{B+N}(2\pi)^{-2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\hat{\psi}_j(\omega)e^{i\omega s}\hat{\psi}_\ell(\nu)e^{i\nu(s+r)}d\omega d\nu \tag{54}
$$

$$
= (2\pi)^{-1}\int_{-\pi}^{\pi}\hat{\psi}_\ell(\nu)e^{i\nu r}\left\{(2\pi)^{-1}\times\right. \tag{55}
$$

$$
\left.\int_{-\pi}^{\pi}\hat{\psi}_j(\omega)\sum_{s=B+1}^{B+N} e^{i(\omega+\nu)s}d\omega\right\}d\nu \tag{56}
$$

$$
= (2\pi)^{-1}\int_{-\pi}^{\pi}\hat{\psi}_\ell(\nu)e^{i\nu r}\left\{(2\pi)^{-1}\times\right. \tag{57}
$$

$$
\left.\int_{-\pi}^{\pi}\hat{\psi}_j(\omega)e^{i(\omega+\nu)(B+1)}\frac{1 - e^{i(\omega+\nu)N}}{1 - e^{i(\omega+\nu)}}d\omega\right\}d\nu
$$

$$
= (2\pi)^{-1}\int_{-\pi}^{\pi}\hat{\psi}_\ell(\nu)e^{i\nu r}G_{j,B,N}(\nu)d\nu, \tag{58}
$$

where

$$G_{j,B,N}(\nu) = (2\pi)^{-1} \int_{-\pi}^{\pi} \hat{\psi}_j(\omega) e^{i(\omega+\nu)(B+1)} \frac{1 - e^{i(\omega+\nu)N}}{1 - e^{i(\omega+\nu)}} d\omega$$

$$= \sum_{s=B+1}^{B+N} e^{i\nu s} \psi_{j,s}$$

$$= \hat{\psi}_j(\nu) - \sum_{s=-\infty}^{B} \psi_{j,s} e^{i\nu s} - \sum_{s=B+N+1}^{\infty} \psi_{j,s} e^{i\nu s}.$$

We now examine what happens to $G_{j,B,N}(\nu)$ under four different cases depending on how the support of the wavelet $\psi_{j,s}$ overlaps the interval $[B+1, B+N]$ or not. Note: the support of the wavelet is the interval $[0, N_j - 1]$. Note: we are mostly interested in the situation when T, N are large and hence not so interested in a potential fifth case when $[B+1, B+N] \subseteq [0, N_j - 1]$.

Case-I: Suppose $[0, N_j - 1] \subseteq [B+1, B+N]$. That is, the support of the wavelet lies entirely within the interval $[B+1, B+N]$. Then

$$G_{j,B,N}(\nu) = \sum_{s=B+1}^{B+N} e^{i\nu s} \psi_{j,s} = \sum_{s=0}^{N_j-1} \psi_{j,s} e^{i\nu s} = \hat{\psi}_j(\nu). \qquad (59)$$

Case-II: Suppose $0 < B+1$ but $B+1 < N_j - 1$, that is right-hand end of the wavelet support overlaps $[B+1, B+N]$ but the left-hand end does not. Then:

$$G_{j,B,N}(\nu) = \sum_{s=B+1}^{N_j-1} \psi_{j,s} e^{i\nu s} = \hat{\psi}_j(\nu) - \sum_{s=0}^{B} \psi_{j,s} e^{i\nu s} \qquad (60)$$

$$= \hat{\psi}_j(\nu) - G_{j,B}(\nu), \qquad (61)$$

where

$$G_{j,B}(\nu) = \sum_{s=0}^{B} \psi_{j,s} e^{i\nu s}. \qquad (62)$$

Case-III: Suppose $[0, N_j - 1]$ and $[B, B+N]$ do not overlap. Then $G_{j,B,N}(\nu) = 0$.

Case-IV: Suppose that $0 < B+N$ and $B+N < N_j - 1$, that is the left-hand end of the wavelet support overlaps $[B+1, B+N]$ but the right-hand end does not. And $B+1 < 0$ which is not of interest as it means that $[zT] - N + 1 < 0$ which should not happen, for large T, as T increases faster than N.

Now let us derive $G_{j,B}(\nu)$ from (62) for $B > 0$:

$$G_{j,B}(\nu) = \begin{cases} 2^{-j/2} \sum_{s=0}^{B} e^{i\nu s} & \text{if } 0 \le B \le N_j/2 - 1, \\ 2^{-j/2} \left(\sum_{s=0}^{N_j/2-1} e^{i\nu s} - \sum_{s=N_j/2}^{B} e^{i\nu s} \right) & \text{if } N_j/2 - 1 < B \le N_j - 1, \\ 2^{-j/2} \sum_{s=0}^{N_j/2-1} \left\{ e^{i\nu s} - e^{i\nu(s+N_j/2)} \right\} & \text{if } N_j - 1 < B. \end{cases}$$

$$(63)$$

Computing the sums in (63) gives:

$$
G_{j,B}(\nu) = \frac{2^{-j/2}}{\{\exp(i\nu)-1\}}
\begin{cases}
\exp\{i\nu(B+1)\}-1, & \text{for } B \in S_1, \\
2\exp(i\nu N_j/2)-\exp\{i\nu(B+1)\}-1, & \text{for } B \in S_2, \\
\{1-\exp(i\nu N_j/2)\}\{\exp(i\nu N_j/2)-1\}, & \text{for } B \in S_3,
\end{cases}
\tag{64}
$$

where $S_1 = \{B : 0 \le B \le N_j/2 - 1\}, S_2 = \{B : N_j/2 - 1 < B \le N_j - 1\}, S_3 = \{B : N_j - 1 < B\}$.

Now returning to the main formula (54).

Case-I: suppose $[0, N_j - 1] \subseteq [B+1, B+N]$ then $G_{j,B,N}(\nu) = \hat{\psi}_j(\nu)$ as given by (59). Hence, substituting into (54) gives:

$$
i_{N,z}(j, \ell, k) = (2\pi)^{-1} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} G_{j,B,N}(\nu) \, d\nu
\tag{65}
$$

$$
= (2\pi)^{-1} \int_{-\pi}^{\pi} 2^{-\ell/2} \frac{(1 - e^{-i\nu N_\ell/2})^2}{1 - e^{-i\nu}} e^{i\nu r} 2^{-j/2} \frac{(1 - e^{-i\nu N_j/2})^2}{1 - e^{-i\nu}} \, d\nu
$$

$$
= \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} e^{i\nu r} \frac{(1 - e^{-i\nu N_\ell/2})^2 (1 - e^{-i\nu N_j/2})^2}{(1 - e^{-i\nu})^2} \, d\nu
\tag{66}
$$

We now bound $i_{N,z}(j, \ell, k)$ by the integral of the absolute value of its integrand, i.e.

$$
|i_{N,z}(j, \ell, k)| \le \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{2\{1 - \cos(N_\ell \nu/2)\} 2\{1 - \cos(N_j \nu/2)\}}{2\{1 - \cos(\nu)\}} \, d\nu
$$

$$
= \frac{2^{1-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{\{1 - \cos(N_\ell \nu/2)\}\{1 - \cos(N_j \nu/2)\}}{1 - \cos(\nu)} \, d\nu,
$$

because $|e^{i\nu r}| = 1$, $|(1 - e^{-i\nu})|^2 = 2\{1 - \cos(\nu)\}$, and so on. Using Lemma 3 with $a = N_\ell/4, b = N_j/4$ gives

$$
|i_{N,z}(j, \ell, k)| \le 2^{-(j+\ell)/2} \min(N_\ell, N_j)
\tag{67}
$$

$$
=
\begin{cases}
2^{-(j-\ell)/2} & \text{for } \ell \le j, \\
2^{-(\ell-j)/2} & \text{for } \ell > j,
\end{cases}
\tag{68}
$$

as $N_j = 2^j$ for Haar wavelets.

Case-IIa. Consider the case when $0 \le B \le N_j/2 - 1$. From (61) we have $G_{j,B,N}(\nu) = \hat{\psi}_j(\nu) - G_{j,B}(\nu)$. Hence:

$$
\begin{aligned}
i_{N,z}(j,\ell,k) &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} G_{j,B,N}(\nu)\, d\nu \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} \left[\hat{\psi}_j(\nu) - G_{j,B}(\nu) \right] d\nu \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} \hat{\psi}_j(\nu)\, d\nu \\
&\quad - \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} \frac{2^{j/2}}{e^{i\nu} - 1} \left[e^{i\nu(B+1)} - 1 \right] d\nu \\
&= (66) - \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{\left(1 - e^{-i\nu N_\ell/2}\right)^2 \left(1 - e^{i\nu(B+1)}\right) e^{i\nu r}}{\left(1 - e^{i\nu}\right)\left(1 - e^{-i\nu}\right)}\, d\nu \\
&\leq (66) - \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{\left(1 - e^{-i\nu N_\ell/2}\right)^2 \left(1 - e^{i\nu N_j/2}\right) e^{i\nu r}}{\left(1 - e^{i\nu}\right)\left(1 - e^{-i\nu}\right)}\, d\nu
\end{aligned}
$$

We now bound $i_{N,z}(j,\ell,k)$ by the integral of the absolute value of its integrand, i.e.

$$
\begin{aligned}
|i_{N,z}(j,\ell,k)| &\leq (68) + \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{\left|\left(1 - e^{-i\nu N_\ell/2}\right)^2\right| \left|\left(1 - e^{i\nu N_j/2}\right)\right| |e^{i\nu r}|}{\left|\left(1 - e^{i\nu}\right)\left(1 - e^{-i\nu}\right)\right|}\, d\nu \\
&\leq (68) + \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{2\left(1 - \cos(\nu N_\ell/2)\right)2}{2\left(1 - \cos(\nu)\right)}\, d\nu \\
&= (68) + \frac{2^{-(j+\ell)/2}}{\pi} \int_{-\pi}^{\pi} 2\pi(N_\ell/2) F_{N_\ell/2}(\nu)\, d\nu \\
&= (68) + 2^{-(j+\ell)/2} N_\ell \\
&= 2^{-(j+\ell)/2} \left(\min(N_\ell, N_j) + N_\ell\right),
\end{aligned}
$$

Case-IIb. Consider the case when $N_j/2 - 1 < B \leq N_j - 1$. Again using (61) we have $G_{j,B,N}(\nu) = \hat{\psi}_j(\nu) - G_{j,B}(\nu)$ and from the corresponding value of (64), we obtain (based on the same logic as above in Cases I and IIa):

$$i_{N,z}(j,\ell,k) \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} G_{j,B,N}(\nu) \, d\nu$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} \left[\hat{\psi}_j(\nu) - G_{j,B}(\nu) \right] d\nu$$

$$= (66) - \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} \frac{2^{j/2}}{e^{i\nu} - 1} \left[2e^{i\nu N_j/2} - e^{i\nu(B+1)} - 1 \right] d\nu$$

$$\leq (66) - \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\psi}_\ell(\nu) e^{i\nu r} \frac{2^{j/2}}{e^{i\nu} - 1} \left[2e^{i\nu N_j/2} - e^{i\nu N_j} - 1 \right] d\nu$$

$$= (66)$$

$$+ \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{\left(1 - e^{-i\nu N_\ell/2}\right)^2 \left(2e^{i\nu N_j/2} - 1 - e^{i\nu N_j}\right) e^{i\nu r}}{(1 - e^{i\nu})(1 - e^{-i\nu})} \, d\nu$$

We now bound $i_{N,z}(j,\ell,k)$ by the integral of the absolute value of its integrand, i.e.

$$|i_{N,z}(j,\ell,k)| \leq (68)$$

$$+ \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{\left| \left(1 - e^{-i\nu N_\ell/2}\right)^2 \right| \left| \left(1 - 2e^{i\nu N_j/2} + e^{i\nu N_j}\right) \right| \left| e^{i\nu r} \right|}{\left| (1 - e^{i\nu})(1 - e^{-i\nu}) \right|} \, d\nu$$

$$\leq (68) + \frac{2^{-(j+\ell)/2}}{2\pi} \int_{-\pi}^{\pi} \frac{2\left(1 - \cos(\nu N_\ell/2)\right) 2}{2\left(1 - \cos(\nu)\right)} \, d\nu$$

$$= (68) + \frac{2^{1-(j+\ell)/2}}{\pi} \int_{-\pi}^{\pi} \pi(N_\ell/2) F_{N_\ell/2}(\nu) \, d\nu$$

$$= (68) + 2^{-(j+\ell)/2} N_\ell$$

$$= 2^{-(j+\ell)/2} \left(\min(N_\ell, N_j) + N_\ell \right),$$

Case-III. Clearly, $|i_{N,z}(j,\ell,k)| \leq 0 \leq (68)$ and Case-IV does not apply.

E.4. Proof of Lemma 2

This lemma has two parts, hence we next prove the first part.

Proof. In what follows we use the two bounds (14) for $k < b_1$ and $k > b_2$ and (18) for $b_1 \leq k \leq b_2$ from Lemma 1. Denote $\mathcal{B} = [b_1, b_2]$ and express

$$\sum_{k=-\infty}^{\infty} \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)|^2 = \sum_{k \in \mathcal{B}} \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)|^2 + \sum_{k \in \mathcal{\cancel{B}}} \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)|^2.$$

For the first sum:

$$\sum_{k \in \mathcal{B}} \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k)|^2 \leq \sum_{k \in \mathcal{B}} \sum_{j=1}^{\infty} 2^{-(j+\ell)} \left(\min(N_\ell, N_j) + N_\ell \right)^2$$

$$= \sum_{k \in \mathcal{B}} \left\{ \sum_{j=1}^{\ell-1} 2^{-(j+\ell)} (N_j + N_\ell)^2 + \sum_{j=\ell}^{\infty} 2^{-(j+\ell)} (2N_\ell)^2 \right\}$$

$$= 2^{-\ell} \sum_{k \in \mathcal{B}} \left\{ \sum_{j=1}^{\ell-1} 2^{-j} \left(2^{2j} + 2^{2\ell} + 2^{1+j+\ell} \right) + 4 \sum_{j=\ell}^{\infty} 2^{-j} 2^{2\ell} \right\}$$

$$= 2^{-\ell} \sum_{k \in \mathcal{B}} \left\{ \sum_{j=1}^{\ell-1} 2^j + 2^{2\ell} \sum_{j=1}^{\ell-1} 2^{-j} \right.$$

$$\left. + 2^{1+\ell} \sum_{j=1}^{\ell-1} 1 + 2^{2+2\ell} \sum_{j=\ell}^{\infty} 2^{-j} \right\}$$

$$= 2^{-\ell} \sum_{k \in \mathcal{B}} \left\{ 2^\ell - 2 + 2^{2\ell} \left(1 - 2^{-\ell+1} \right) \right.$$

$$\left. + (\ell - 1) 2^{1+\ell} + 2^{2+2\ell} 2^{-\ell+1} \right\}$$

$$= (N_\ell - 1) 2^{-\ell} \left\{ (2\ell + 5) 2^\ell - 2 + 2^{2\ell} \right\}$$

$$= (N_\ell - 1) \left(2\ell + 5 - 2^{-\ell+1} + 2^\ell \right)$$

$$= \mathcal{O} \left(2^{2\ell} \right).$$

For the second sum:

$$\sum_{k \in \mathcal{B}} \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k)|^2 \leq \sum_{k \in \mathcal{B}} \sum_{j=1}^{\infty} |\Psi_{j,\ell} (k - 2[zT] + N/2 - 1)|^2$$

$$\leq \sum_j B_\ell(j, j),$$

where $B_\ell(j, p)$ is the fourth-order cross-correlation wavelet absolute value product of order $r = 0$, defined as $B_\ell^{(r)}(j, i) = \sum_{p=-\infty}^{\infty} |p|^r |\Psi_{j,\ell}(p) \Psi_{i,\ell}(p)|$ for $r = 0, 1$ and scales $\ell, j, i \in \mathbb{N}$.

Splitting the sum of j and using Proposition F.1 leads to the following:

$$\sum_j B_\ell(j,j) = \sum_{j=0}^{\ell-1} B_\ell(j,j) + \sum_{j=\ell+1}^{\infty} B_\ell(j,j) + B_\ell(\ell,\ell)$$

$$= \sum_{j=0}^{\ell-1} 2^{-\ell}\left(2^{2j-1}+1\right) + \sum_{j=\ell+1}^{\infty} 2^{-j}\left(2^{2\ell-1}+1\right) + \frac{1}{3}2^{-\ell}\left(2^{2\ell}+5\right)$$

$$= 2^{-\ell}\left[\frac{1}{2}\frac{1}{3}\left(2^{2\ell}-1\right)+\ell\right] + \left(2^{2\ell-1}+1\right)2^{-\ell} + \frac{1}{3}2^{-\ell}\left(2^{2\ell}+5\right)$$

$$= \mathcal{O}\left(2^{\ell}\right).$$

Hence $\sum_{k=-\infty}^{\infty}\sum_{j=1}^{\infty}|i_{N,z}(j,\ell,k)|^2 = \mathcal{O}(2^{2\ell})$.

Now we prove the second part of Lemma 2.

Proof. Consider

$$\sum_{k=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\left\{\sum_{j=1}^{\infty}|i_{N,z}(j,\ell,k)i_{N,z}(j,m,n)|\right\}^2$$

$$= \sum_k \sum_n \sum_{j=1}^{\infty}|i_{N,z}(j,\ell,k)i_{N,z}(j,m,n)|\times$$

$$\sum_{p=1}^{\infty}|i_{N,z}(p,\ell,k)i_{N,z}(p,m,n)|$$

$$= \sum_{j,p=1}^{\infty}\sum_k |i_{N,z}(j,\ell,k)i_{N,z}(p,\ell,k)| \sum_n |i_{N,z}(j,m,n)i_{N,z}(p,m,n)|$$

$$= \sum_{j,p=1}^{\infty}\left\{\sum_{k\in\mathcal{B}}|i_{N,z}(j,\ell,k)i_{N,z}(p,\ell,k)|\right.$$

$$\left.+\sum_{k\notin\mathcal{B}}|i_{N,z}(j,\ell,k)i_{N,z}(p,\ell,k)|\right\}$$

$$\times\left\{\sum_{n\in\mathcal{B}}|i_{N,z}(j,m,n)i_{N,z}(p,m,n)|\right.$$

$$\left.+\sum_{n\notin\mathcal{B}}|i_{N,z}(j,m,n)i_{N,z}(p,m,n)|\right\}$$

$$= T_{\mathcal{BB}}(j,\ell,p,m) + T_{\mathcal{B\!\!\!/B}}(j,\ell,p,m) + T_{\mathcal{B\!\!\!/B}}(j,\ell,p,m) + T_{\mathcal{B\!\!\!/B\!\!\!/}}(j,\ell,p,m).$$

The term $T_{\mathcal{B\!\!\!/B\!\!\!/}} = \sum_{k\notin\mathcal{B}}\sum_{n\notin\mathcal{B}}$ is the case where $i_{N,z}$ can be bounded by Ψ and is addressed in detail in Lemma 7. The term $T_{\mathcal{BB}} = \sum_{k\in\mathcal{B}}\sum_{n\in\mathcal{B}}$ is dealt with

in Lemma 5 using the bound (18) for $i_{N,z}$ from Lemma 1 and the cross term is dealt with in Lemma 6. Each of these lemmas (below) show that each of the product terms is of order no worse than $\mathcal{O}\{2^{2(\ell+m)}\}$.

E.5. Proof of Theorem 1

Proof. First recall that we are under the zero-mean locally stationary wavelet process framework as described in Appendix B, with $\{X_{t,T}\}_{t=0}^{T-1}$ a doubly-index stochastic process with representation given by

$$X_{t,T} = \sum_{j=1}^{\infty} \sum_{k=-\infty}^{\infty} w_{j,k;T}^0 \psi_{j,k}(t)\xi_{j,k}.$$

The integrated local periodogram was defined as

$$J_N(z,\phi) = \sum_{j=1}^{\infty} \phi_j I_N^*(z,j)$$

where $\{\phi_j\}_{j=1}^{\infty} \in \Phi$, with Φ a set of complex-valued bounded sequences equipped with uniform norm $||\phi||_\infty := \sup_j |\phi_j|$ and in order to avoid notational clutter N replaces the interval length notation $L(T)$ present in the main body of the paper. The quantity $I_N^*(z,j)$ denotes the uncorrected tapered local wavelet periodogram

$$I_N^*(z,j) = H_N^{-1} \left| \sum_{t=0}^{N-1} h\left(\frac{t}{N}\right) X_{t-N/2+1,T}\psi_{j,[zT]}(t) \right|^2,$$

with $h : [0,1] \to \mathbb{R}_+$ a data taper, $H_N := \sum_{j=0}^{N-1} h^2(j/N) \sim N \int_0^1 h^2(x)\,dx$ the normalizing factor and $h(\cdot)$ is assumed symmetric and with a bounded second derivative.

As in Dahlhaus and Giraitis (1998) we approximate $J_N(z,\phi)$ by the corresponding statistics of a stationary process with the same local corresponding statistics at $t = zT$, z fixed. Let

$$J_N^Y(\phi) = \sum_{j=1}^{\infty} \phi_j I_N^{*,Y}(j)$$

where

$$I_N^{*,Y}(j) := H_N^{-1} \left| \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) Y_{[zT]-N/2+1+s,T}\psi_{j,[zT]}(s) \right|^2$$

is the wavelet periodogram on the segment $[zT] - N/2 + 1, \ldots, [zT] + N/2$ of the *stationary* process

$$Y_s = \sum_{j=1}^{\infty} W_j(z) \sum_{k=-\infty}^{\infty} \psi_{j,k}(s)\xi_{j,k}.$$

Note: The next section uses sequences of bounded variation The total variation of a sequence $\{\phi_j\}_{j=1}^{\infty}$ is defined by $\mathrm{TV}(\{\phi_j\}) = \sum_{j=1}^{\infty} |\phi_{j+1} - \phi_j|$ and the space of all sequences of finite total variation is denoted by bv, see Dunford and Schwartz (1958) for example.

From equations (11) and (12), we obtain $J_N(z, \phi) - \mathbb{E}(J_N(z, \phi)) = J_N^Y(\phi) - \mathbb{E}(J_N^Y(\phi)) + o_p(N^{-1/2})$ and using equation (10) it follows that

$$J_N(z, \phi) = J_N^Y(\phi) + \mathcal{O}\left(N^{-1}\right) + o_p(N^{-1/2})$$

which reveals the approximation we make and should be compared to equation (4.4) in Dahlhaus and Giraitis (1998), where a term $\mathcal{O}\left(\frac{N}{T}\right)$ appears instead of $\mathcal{O}\left(N^{-1}\right)$.

Using the uncorrected tapered local periodogram expression in equation (8) and the LSW definition in equation (23), by rearranging formulae we can write the integrated wavelet periodogram:

$$J_N(z, \phi) = \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \hat{d}_{N,z}(\ell, k, m, n) \xi_{\ell,k} \xi_{m,n},$$

where

$$\hat{d}_{N,z}(\ell, k, m, n) = H_N^{-1} w_{\ell,k}^0 w_{m,n}^0 \sum_{j=1}^{\infty} \phi_j i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)$$

and from (32)

$$i_{N,z}(j, \ell, k) = \sum_{t=0}^{N-1} h\left(\frac{t}{N}\right) \psi_{j,[zT]-t} \psi_{\ell,k-[zT]-t-1+N/2}.$$

Using the properties of the $\{\xi_{\ell,k}\}_{\ell,k}$ field, we obtain

$$\mathbb{E}\{J_N(z, \phi)\} = \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \hat{d}_{N,z}(\ell, k, m, n) \delta_{l,m} \delta_{k,n}$$

$$= \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \hat{d}_{N,z}(\ell, k, \ell, k)$$

$$= \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} H_N^{-1} \left[\{w_{\ell,k}^0 - W_\ell(k/T)\} + \{W_\ell(k/T) - W_\ell(z)\} \right.$$

$$\left. + W_\ell(z) \right]^2 \sum_{j=1}^{\infty} \phi_j |i_{N,z}(j, \ell, k)|^2$$

$$= \underbrace{\sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} H_N^{-1} W_\ell^2(z) \sum_{j=1}^{\infty} \phi_j |i_{N,z}(j, \ell, k)|^2}_{\mathbb{E}(J_N^Y(\phi))} + \text{cross terms},$$

where the assumption $\sum_j W_j(z)^2 2^{2j} < \infty$ (at any rescaled time z) ensures that $\mathbb{E}(J_N^Y(\phi))$ is finite.

Therefore, using the LSW property that $\sup_k \left| w_{\ell,k;T}^0 - W_\ell(k/T) \right| \leq C_\ell/T$, leads to

$$|\mathbb{E}\{J_N(z,\phi)\} - \mathbb{E}\{J_N^Y(\phi)\}| \leq \text{ sum of the modulus of the cross terms,}$$

all upper bounded by terms of the form

$$A = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty \frac{C_\ell^2}{T^2} \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2,$$

$$B = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty \left[\frac{C_\ell}{T} \{W_\ell(k/T) - W_\ell(z)\} \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 \right],$$

$$C = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty \left[\{W_\ell(k/T) - W_\ell(z)\}^2 \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 \right],$$

$$D = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty \left[\{W_\ell(k/T) - W_\ell(z)\} W_\ell(z) \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 \right].$$

In order to further bound these quantities, we use Lemma 2 that proves

$$\sum_{k=-\infty}^\infty \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 = \mathcal{O}(2^{2\ell}),$$

hence we obtain

$$\begin{aligned}
A = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty \frac{C_\ell^2}{T^2} \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 &= \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \frac{C_\ell^2}{T^2} \sum_{k=-\infty}^\infty \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 \\
&\leq \frac{\|\phi\|}{H_N} \frac{1}{T^2} \sum_{\ell=1}^\infty C_\ell^2 2^{2\ell} = \mathcal{O}(N^{-1}T^{-2}),
\end{aligned}$$

where we have used $\sum_{\ell=1}^\infty C_\ell^2 2^{2\ell} < \infty$.

$$\begin{aligned}
B = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty &\left[\frac{C_\ell}{T} \{W_\ell(k/T) - W_\ell(z)\} \sum_{j=1}^\infty |i_{N,z}(j,\ell,k)|^2 \right] \\
&\leq \frac{\|\phi\|}{H_N} \sum_{\ell=1}^\infty \sum_{k=-\infty}^\infty \sum_{j=1}^\infty \frac{C_\ell}{T} L_\ell \left| \frac{k - [zT]}{T} \right| |i_{N,z}(j,\ell,k)|^2 \\
&\leq \frac{\|\phi\|}{H_N} \frac{1}{T} \sum_{\ell=1}^\infty C_\ell L_\ell 2^{2\ell} = \mathcal{O}(N^{-1}T^{-1}),
\end{aligned}$$

using the Lipschitz continuity of $\{W_j\}_j$, $T^{-1}|k - [zT]| \in (0,1)$ and the Hölder inequality

$$\sum_{\ell=1}^{\infty} C_\ell L_\ell 2^{2\ell} < \left(\sum_{\ell=1}^{\infty} C_\ell^2 2^{2\ell}\right)^{1/2} \left(\sum_{\ell=1}^{\infty} L_\ell^2 2^{2\ell}\right)^{1/2} < \infty$$

coupled with the assumptions in the theorem.

We now need to bound

$$C = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \left[\{W_\ell(k/T) - W_\ell(z)\}^2 \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)|^2\right]$$

$$\leq \frac{\|\phi\|}{H_N} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{j=1}^{\infty} L_\ell^2 \frac{|(k - [zT])^2|}{T^2} |i_{N,z}(j,\ell,k)|^2.$$

As $\sum_\ell L_\ell^2 2^{2\ell} < \infty$ and as $|\frac{k-[zT]}{T}|^2 \in (0,1)$, we obtain that $C = \mathcal{O}(N^{-1})$.

The term

$$D = \frac{\|\phi\|}{H_N} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \left[\{W_\ell(k/T) - W_\ell(z)\} W_\ell(z) \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k)|^2\right]$$

$$\leq \frac{\|\phi\|}{H_N} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{j=1}^{\infty} L_\ell W_\ell(z) \frac{|(k - [zT])|}{T} |i_{N,z}(j,\ell,k)|^2$$

can be bounded using Hölder's inequality

$$\sum_{\ell=1}^{\infty} W_\ell(z) L_\ell 2^{2\ell} < \left(\sum_{\ell=1}^{\infty} W_\ell^2(z) 2^{2\ell}\right)^{1/2} \left(\sum_{\ell=1}^{\infty} L_\ell^2 2^{2\ell}\right)^{1/2} < \infty$$

based on the assumptions in the theorem and recalling that we assumed $\{Y_t\}$ to be stationary. Hence $D = \mathcal{O}(N^{-1})$.

This completes the proof of $\mathbb{E}\{J_N(z,\phi)\} = \mathbb{E}\{J_N^Y(\phi)\} + \mathcal{O}(N^{-1})$.

Now let us establish consistency and its rate. Start by considering

$$\text{var}\{J_N(z,\phi)\} = \mathbb{E}\{J_N(z,\phi) - \mathbb{E}(J_N(z,\phi))\}^2$$

$$= \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{\ell'=1}^{\infty} \sum_{k'=-\infty}^{\infty} \sum_{m'=1}^{\infty} \sum_{n'=-\infty}^{\infty}$$

$$\hat{d}_{N,z}(\ell,k,m,n)\hat{d}_{N,z}(\ell',k',m',n') \, \text{cov}(\xi_{\ell,k}\xi_{m,n}, \xi_{\ell',k'}\xi_{m',n'}).$$

Using Isserlis, we can decompose

$$
\begin{aligned}
\operatorname{cov}(\xi_{\ell,k}\xi_{m,n}, \xi_{\ell',k'}\xi_{m',n'}) &= \mathbb{E}(\xi_{\ell,k}\xi_{m,n}\xi_{\ell',k'}\xi_{m',n'}) \\
&\quad - \mathbb{E}(\xi_{\ell,k}\xi_{m,n})\mathbb{E}(\xi_{\ell',k'}\xi_{m',n'}) \\
&= \mathbb{E}(\xi_{\ell,k}\xi_{m,n})\mathbb{E}(\xi_{\ell',k'}\xi_{m',n'}) \\
&\quad + \mathbb{E}(\xi_{\ell,k}\xi_{\ell',k'})\mathbb{E}(\xi_{m,n}\xi_{m',n'}) \\
&\quad + \mathbb{E}(\xi_{\ell,k}\xi_{m',n'})\mathbb{E}(\xi_{\ell',k'}\xi_{m,n}) - \mathbb{E}(\xi_{\ell,k}\xi_{m,n})\mathbb{E}(\xi_{\ell',k'}\xi_{m',n'}) \\
&= \delta_{\ell,\ell'}\delta_{k,k'}\delta_{m,m'}\delta_{n,n'} + \delta_{\ell,m'}\delta_{k,n'}\delta_{m,\ell}\delta_{n,k'},
\end{aligned}
$$

hence

$$
\begin{aligned}
\operatorname{var}\{J_N(z,\phi)\} &= \sum_{\ell=1}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty} \hat{d}_{N,z}^2(\ell,k,m,n) \\
&\quad + \sum_{\ell=1}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty} \hat{d}_{N,z}(\ell,k,m,n)\hat{d}_{N,z}(m,n,\ell,k) \\
&= 2\sum_{\ell=1}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty} \hat{d}_{N,z}^2(\ell,k,m,n),
\end{aligned}
$$

as $\hat{d}_{N,z}^2(m,n,\ell,k) = \hat{d}_{N,z}^2(\ell,k,m,n)$. Hence, we seek to bound

$$
\sum_{\ell=1}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty} \hat{d}_{N,z}^2(\ell,k,m,n). \tag{69}
$$

Let us now expand the above

$$
\begin{aligned}
\frac{1}{2}\operatorname{var}\{J_N(z,\phi)\} &= \sum_{\ell=1}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty} \Big[\{w_{\ell,k}^0 - W_\ell(k/T)\} + \{W_\ell(k/T) - W_\ell(z)\} \\
&\quad + W_\ell(z)\Big]^2 \\
&\quad \times \Big[\{w_{m,n}^0 - W_m(n/T)\} + \{W_m(n/T) - W_m(z)\} + W_m(z)\Big]^2 \\
&\quad \times \left\{\sum_{j=1}^{\infty}\phi_j|i_{N,z}(j,\ell,k)i_{N,z}(j,m,n)|\right\}^2.
\end{aligned}
$$

Using an inequality of the type $(a+b+c)^2 \le 3(a^2+b^2+c^2)$, the above quantity is upper bounded by a linear combination of a finite number of terms, of the following types

$$AA = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{C_\ell^2}{T^2} \frac{C_m^2}{T^2} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$BB = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{C_\ell^2}{T^2} \{W_m(n/T) - W_m(z)\}^2$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$CC = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \{W_\ell(k/T) - W_\ell(z)\}^2 \{W_m(n/T) - W_m(z)\}^2$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$DD = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \{W_\ell(k/T) - W_\ell(z)\}^2 W_m^2(z)$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$EE = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} W_\ell^2(z) W_m^2(z)$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2.$$

In order to bound the above quantities, we use Lemma 2 which proves that

$$\sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2 = \mathcal{O}(2^{(\ell+m)}). \qquad (70)$$

Using this result we can bound each term in turn, as follows.

The first term

$$AA = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{C_\ell^2}{T^2} \frac{C_m^2}{T^2} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$= \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \frac{C_\ell^2}{T^2} \frac{C_m^2}{T^2} \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \frac{1}{T^4} \sum_{\ell=1}^{\infty} C_\ell^2 2^{2\ell} \sum_{m=1}^{\infty} C_m^2 2^{2m} = \mathcal{O}(N^{-2} T^{-4})$$

where we used $\sum_{\ell=1}^{\infty} C_\ell^2 2^{2\ell} < \infty$.

The term

$$BB = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{C_\ell^2}{T^2} \{W_m(n/T) - W_m(z)\}^2 \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$= \frac{\|\phi^2\|}{H_N^2} \frac{1}{T^2} \sum_{\ell=1}^{\infty} C_\ell^2 \sum_{m=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \{W_m(n/T) - W_m(z)\}^2 \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \frac{1}{T^2} \sum_{\ell=1}^{\infty} C_\ell^2 \sum_{m=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} L_m^2 \left| \frac{n}{T} - z \right|^2 \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \frac{1}{T^2} \sum_{\ell=1}^{\infty} C_\ell^2 2^{2\ell} \sum_{m=1}^{\infty} L_m^2 2^{2m} = \mathcal{O}(N^{-2} T^{-2}),$$

where we have used the Lipschitz continuity of W_m, $|\frac{n}{T} - z|^2 \in (0,1)$, $\sum_\ell C_\ell^2 2^{2\ell} < \infty$ and $\sum_m L_m^2 2^{2m} < \infty$.

Using the same set of arguments, we bound

$$CC = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \{W_\ell(k/T) - W_\ell(z)\}^2 \{W_m(n/T) - W_m(z)\}^2$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} L_\ell^2 \left| \frac{k}{T} - z \right|^2 L_m^2 \left| \frac{n}{T} - z \right|^2 \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} L_\ell^2 2^{2\ell} \sum_{m=1}^{\infty} L_m^2 2^{2m} = \mathcal{O}(N^{-2}).$$

The term

$$DD = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \{W_\ell(k/T) - W_\ell(z)\}^2 W_m^2(z)$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2,$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{m=1}^{\infty} W_m^2(z) \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} L_\ell^2 \left| \frac{k}{T} - z \right|^2$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} L_\ell^2 2^{2\ell} \sum_{m=1}^{\infty} W_m^2(z) 2^{2m} = \mathcal{O}(N^{-2}),$$

as $|\frac{k}{T} - z|^2 \in (0,1)$, $\sum_\ell L_\ell^2 2^{2\ell} < \infty$ and $\sum_m W_m^2(z) 2^{2m} < \infty$ at a set z (recall the process $\{Y_t\}$ was assumed stationary).

Similarly,

$$EE = \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} W_\ell^2(z) W_m^2(z)$$

$$\times \left\{ \sum_{j=1}^{\infty} \phi_j |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} W_\ell^2(z) \sum_{m=1}^{\infty} W_m^2(z) \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty}$$

$$\times \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j,\ell,k) i_{N,z}(j,m,n)| \right\}^2$$

$$\leq \frac{\|\phi^2\|}{H_N^2} \sum_{\ell=1}^{\infty} W_\ell^2(z) 2^{2\ell} \sum_{m=1}^{\infty} W_m^2(z) 2^{2m} = \mathcal{O}(N^{-2}).$$

We therefore obtain that

$$\text{var}\{N^{1/2} J_N(z,\phi)\} \leq K/N, \quad \text{for some constant } K, \tag{71}$$

hence $J_N(z,\phi) - \mathbb{E}\{J_N(z,\phi)\} = o_P(N^{-1/2})$. The result for the process $\{Y_t\}$ follows similarly, which concludes the proof of Theorem 1.

E.6. Proof of Proposition 3.4

Proof. Theorem 1 established the limit properties of the approximation we make for $J_{L(T)}(z,\phi)$. From (11) and (12), we obtain $J_{L(T)}(z,\phi) - \mathbb{E}\left\{ J_{L(T)}(z,\phi) \right\} =$

$J_{L(T)}^Y(\phi) - \mathbb{E}\left\{J_{L(T)}^Y(\phi)\right\} + o_p\left\{L(T)^{-1/2}\right\}$ and using equation (10) it follows that

$$J_{L(T)}(z,\phi) = J_{L(T)}^Y(\phi) + \mathcal{O}\left\{L(T)^{-1}\right\} + o_p\left\{L(T)^{-1/2}\right\}. \tag{72}$$

Now recall we defined our windowed local partial autocorrelation estimator $\tilde{q}_W(z,\tau) = \hat{q}_{[z-L(T)/2T, z+L(T)/2T]}(\tau)$, hence $\tilde{q}_W(z,\tau) = \left(\hat{\Gamma}_{z,\tau}^{-1}\hat{\underline{\gamma}}_{z,\tau}\right)_\tau$ where both $\hat{\Gamma}_{z,\tau}$ and $\hat{\underline{\gamma}}_{z,\tau}$ are a matrix, respectively vector of local tapered covariances $\hat{c}(z,\tau)$.

The elements of the covariance matrix $(\hat{\Gamma})$ and vector $(\hat{\gamma})$ are $\hat{c}(z,\tau) = \sum_j \hat{S}_j(z)\Psi_j(\tau)$ and thus can be written as integrated periodograms $J_{L(T)}$, since $\hat{\mathbf{S}}_j(z) = A_J^{-1}\mathbf{I}_{L(T)}^*(z)$ and $\mathbf{I}_{L(T)}^*(z) = \left(I_{L(T)}^*(z,1),\ldots,I_{L(T)}^*(z,J)\right)^T$. Using the result in equation (72), in the manner of Dahlhaus and Giraitis (1998), it follows that $\tilde{q}_W(z,\tau)$ has the same asymptotic distribution as in the stationary case.

E.7. *Additional Results Required For the Proofs from Section 3*

Lemma 4. The core function, $\Omega_i(u)$ for Haar wavelets is given by

$$\Omega_i(u) = 2^{-i/2}\begin{cases} 0 & \text{for } u < -1, \\ -(u+1) & \text{for } -1 \leq u < -\frac{1}{2}, \\ u & \text{for } -\frac{1}{2} \leq u < 0, \\ 0 & \text{for } 0 \leq u < 2^{i-1} - 1, \\ 2u - 2^i + 2 & \text{for } 2^{i-1} - 1 \leq u < 2^{i-1} - \frac{1}{2}, \\ 2^i - 2u & \text{for } 2^{i-1} - \frac{1}{2} \leq u < 2^{i-1}, \\ 0 & \text{for } 2^{i-1} \leq u < 2^i - 1, \\ 2^i - u - 1 & \text{for } 2^i - 1 \leq u < 2^i - \frac{1}{2}, \\ u - 2^i & \text{for } 2^i - \frac{1}{2} \leq u < 2^i, \\ 0 & \text{for } 2^i \leq u, \end{cases} \tag{73}$$

for $u \in \mathbb{R}$ and $i \in \mathbb{N} \cup \{0\}$.

Figure 6 shows a depiction of $\Omega_i(u)$.

Proof. Using the substitution $y = x - u$ we can break the integral for Ω_i into
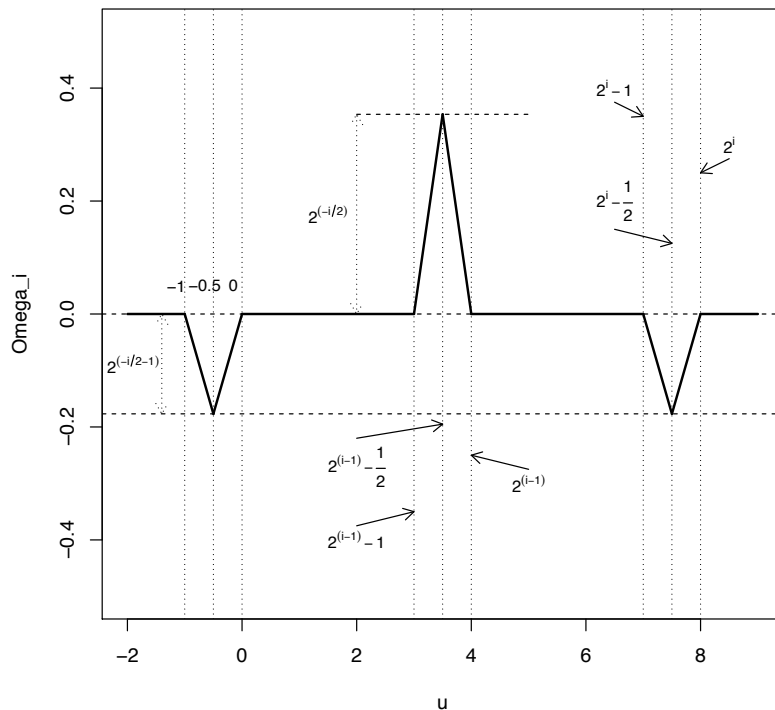
FIG 6. *Depiction of $\Omega_i(u)$. The function is symmetric about $2^{(i-1)} - \frac{1}{2}$ and the extent of the function is from -1 on the left to 2^i on the right. The width of all of the triangles is always 1 for all i. As i increases the function gets stretched to the right (but also anchored on the left at $u = -1$), the peaks decrease in size like $2^{-i/2}$.*

two pieces as follows:

$$
\begin{aligned}
\Omega_i(u) &= \int_{-\infty}^{\infty} 2^{-i/2} \psi(2^{-i}x)\psi(x - u)\,dx \\
&= 2^{-i/2} \int_{-\infty}^{\infty} \psi\{2^{-i}(y + u)\}\psi(y)\,dy \\
&= 2^{-i/2} \int_{0}^{1} \psi\{2^{-i}(y + u)\}\psi(y)\,dy \\
&= 2^{-i/2} \left[-\int_{0}^{\frac{1}{2}} \psi\{2^{-i}(y + u)\}\,dy + \int_{\frac{1}{2}}^{1} \psi\{2^{-i}(y + u)\}\,dy \right] \\
&= 2^{-i/2}\{-I + II\},
\end{aligned}
\tag{74}
$$

where I, II are the integrals in the final line of (74).

Let us consider integral I first, making the substitution $x = 2^{-i}(y + u)$

$$
I = \int_{0}^{\frac{1}{2}} \psi\{2^{-i}(y + u)\}\,dy = 2^i \int_{2^{-i}u}^{2^{-i}(u + \frac{1}{2})} \psi(x)\,dx.
$$

So, integral I is the result of integrating the product of $\psi(x)$ with the moving window $[2^{-i}u, 2^{-i}(u + \frac{1}{2})]$. To derive integral I we first note that $I = 0$ if $2^{-i}u \geq 1$ or $2^{-i}(u + \frac{1}{2}) \leq 0$ which translates into $I = 0$ if $u \geq 2^i$ or $u \leq -\frac{1}{2}$. We break down the remainder of the case $-\frac{1}{2} < u < 2^i$ into five subregions.

(Ia) If $2^i - \frac{1}{2} \leq u < 2^i$ then

$$I = 2^i \int_{2^{-i}u}^{1} \psi(x)\,dx = 2^i \int_{2^{-i}u}^{1} dx = 2^i(1 - 2^{-i}u) = 2^i - u.$$

(Ib) If $2^{i-1} \leq u < 2^i - \frac{1}{2}$ then

$$I = 2^i \int_{2^{-i}u}^{2^{-i}(u+\frac{1}{2})} dx = 2^i(2^{-i}u + 2^{-i-1} - 2^{-i}u) = 2^{-1}.$$

(Ic) If $2^{i-1} - \frac{1}{2} \leq u \leq 2^{i-1}$ then

$$I = 2^i \left\{ \int_{2^{-i}u}^{\frac{1}{2}} (-1)\,dx + \int_{\frac{1}{2}}^{2^{-i}(u+\frac{1}{2})} dx \right\} = 2u - 2^i + \frac{1}{2}.$$

(Id) If $0 \leq u < 2^{i-1} - \frac{1}{2}$ then

$$I = 2^i \int_{2^{-i}u}^{2^{-i}(u+\frac{1}{2})} (-1)\,dx = -\frac{1}{2}.$$

(Ie) If $-\frac{1}{2} \leq u < 0$ then

$$I = 2^i \int_{0}^{2^{-i}(u+\frac{1}{2})} (-1)\,dx = -(u + \frac{1}{2}).$$

For the second integral we note that:

$$II(u) = \int_{\frac{1}{2}}^{1} \psi\{2^{-i}(y + u)\}\,dy = 2^i \int_{2^{-i}(u+\frac{1}{2})}^{2^{-i}(u+1)} \psi(x) = I(u + \frac{1}{2}).$$

In other words, we have already done the work to evaluate $II(u)$. Hence,

$$\begin{cases} 0 & \text{for } u < -1, \\ -(u+1) & \text{for } -1 \leq u < -\frac{1}{2} \ (IIa), \\ -\frac{1}{2} & \text{for } -\frac{1}{2} \leq u < 2^{i-1} - 1 \ (IIb), \\ 2u - 2^i + \frac{3}{2} & \text{for } 2^{i-1} - 1 \leq u < 2^{i-1} - \frac{1}{2} \ (IIc), \\ \frac{1}{2} & \text{for } 2^{i-1} - \frac{1}{2} \leq u < 2^i - 1 \ (IId), \\ 2^i - u - \frac{1}{2} & \text{for } 2^i - 1 \leq u < 2^i - \frac{1}{2} \ (IIe), \\ 0 & \text{for } 2^i - \frac{1}{2} \leq u. \end{cases}$$

Now we need to put the two results together into (74) and work out the regions of overlap in the two sets of intervals. Hence, define the following regions:

$$
\begin{cases}
III_a & = \{u : u < -1\}, \\
III_b & = II_a = \{u : -1 \le u < -\frac{1}{2}\}, \\
III_c & = I_a \cap II_b = \{u : -\frac{1}{2} \le u < 0\}, \\
III_d & = I_b \cap II_b = \{u : 0 \le u < 2^{i-1} - 1\}, \\
III_e & = I_b \cap II_c = \{u : 2^{i-1} - 1 \le u < 2^{i-1} - \frac{1}{2}\}, \\
III_f & = I_c \cap II_d = \{u : 2^{i-1} - \frac{1}{2} \le u < 2^{i-1}\}, \\
III_g & = I_d \cap II_d = \{u : 2^{i-1} \le u < 2^i - 1\}, \\
III_h & = I_d \cap II_e = \{u : 2^i - 1 \le u < 2^i - \frac{1}{2}\}, \\
III_i & = I_e = \{u : 2^i - \frac{1}{2} \le u < 2^i\}, \\
III_j & = \{u : 2^i \le u\}.
\end{cases}
\tag{75}
$$

Putting together the two integrals with these new domains gives the result in (73).

Lemma 5. Under the conditions and notations set out so far, for the nondecimated family of discrete Haar wavelets we have

$$
T_{\mathcal{BB}}(\ell, m) = \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2 = \mathcal{O}\{2^{2(\ell+m)}\}.
$$

Proof. We bound $T_{\mathcal{BB}}(\ell, m)$ and include the indices ℓ, m explicitly. For $\ell < m$ we have

$$
T_{\mathcal{BB}}(\ell, m) = \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2
$$

$$
\le \sum_{k \in \mathcal{B}, n \in \mathcal{B}} \left[\sum_{j=1}^{\ell-1} 2^{-(\ell+j-m-j)/2} \{\min(N_j, N_\ell) + N_\ell\} \{\min(N_j, N_m) + N_m\} \right.
$$

$$
+ \sum_{j=\ell}^{m-1} 2^{-(j+\ell)/2} \{\min(N_j, N_\ell) + N_\ell\} 2^{-(m+j)/2} \{\min(N_j, N_m) + N_m\}
$$

$$
\left. + \sum_{j=m}^{\infty} 2^{-(j+\ell)/2} \{\min(N_j, N_\ell) + N_\ell\} 2^{-(j+m)/2} \{\min(N_j, N_m) + N_m\} \right]^2,
$$

using bound (18). Hence

$$
T_{\mathcal{BB}}(\ell, m) \leq \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \Bigg\{ 2^{-(\ell+m)/2} \sum_{j=1}^{\ell-1} 2^{-j} \left(2^{2j} + 2^{j+m} + 2^{j+\ell} + 2^{\ell+m} \right)
$$

$$
+ 2^{1-(\ell+m)/2} \sum_{j=\ell}^{m-1} 2^{-j} \left(2^{\ell+j} + 2^{\ell+m} \right) + 2^{2-(\ell+m)/2} \sum_{j=m}^{\infty} 2^{-j} 2^{\ell+m} \Bigg\}^2
$$

$$
= \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \Bigg\{ 2^{-(\ell+m)/2} \sum_{j=1}^{\ell-1} 2^{j} + 2^{-(\ell-m)/2} \sum_{j=1}^{\ell-1} 2^{0} + 2^{-(m-\ell)/2} \sum_{j=1}^{\ell-1} 2^{0}
$$

$$
+ 2^{(\ell+m)/2} \sum_{j=1}^{\ell-1} 2^{-j} + 2^{1-(m-\ell)/2} \sum_{j=\ell}^{m-1} 2^{0} + 2^{1+(\ell+m)/2} \sum_{j=\ell}^{m-1} 2^{-j}
$$

$$
+ 2^{2+(\ell+m)/2} \sum_{j=m}^{\infty} 2^{-j} \Bigg\}^2
$$

$$
= \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \Bigg\{ 2^{-(\ell+m)/2} \left(2^{\ell} - 2 \right) + 2^{-(\ell-m)/2} \left(\ell - 1 \right)
$$

$$
+ 2^{-(m-\ell)/2} \left(\ell - 1 \right) + 2^{(\ell+m)/2} \left(1 - 2^{-(\ell-1)} \right)
$$

$$
+ 2^{1-(m-\ell)/2} \left(m - \ell \right) + 2^{2+(\ell+m)/2} \left(2^{-\ell} - 2^{-m} \right) + 2^{2+(\ell+m)/2} 2^{1-m} \Bigg\}^2
$$

$$
= \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \Bigg\{ \left(\ell - 1 + 1 + 8 + 2m - 2\ell - 4 \right) 2^{-(m-\ell)/2} - 2^{1-(\ell+m)/2}
$$

$$
+ \left(l - 1 - 2 + 4 \right) 2^{-(\ell-m)/2} + 2^{(\ell+m)/2} \Bigg\}^2
$$

$$
= \sum_{k \in \mathcal{B}} \sum_{n \in \mathcal{B}} \Bigg\{ \left(4 - \ell + 2m \right) 2^{-(m-\ell)/2} - 2^{1-(\ell+m)/2}
$$

$$
+ \left(l + 1 \right) 2^{-(\ell-m)/2} + 2^{(\ell+m)/2} \Bigg\}^2
$$

$$
= \left(N_{\ell} - 1 \right) \left(N_m - 1 \right) \Bigg\{ \left(4 - \ell + 2m \right) 2^{-(m-\ell)/2} - 2^{1-(\ell+m)/2} \tag{76}
$$

$$
+ \left(l + 1 \right) 2^{-(\ell-m)/2} + 2^{(\ell+m)/2} \Bigg\}^2
$$

$$
= \left(N_{\ell} - 1 \right) \left(N_m - 1 \right) \Bigg\{ \frac{2^{\ell+m} - 2}{2^{(\ell+m)/2}} + \frac{\left(\ell + 1 \right) 2^{-(\ell-m)} + \left(4 - \ell + 2m \right)}{2^{-(\ell-m)/2}} \Bigg\}^2
$$

$$
= \left(N_{\ell} - 1 \right) \left(N_m - 1 \right) \Bigg\{ \frac{2^{2(\ell+m)} - 2^{1+\ell+m} + 4}{2^{\ell+m}}
$$

$$
+ \frac{\left(\ell + 1 \right)^2 2^{-2(\ell-m)} + \left(4 - \ell + 2m \right) \left(l + 1 \right) 2^{1-(\ell-m)} + \left(4 - \ell + 2m \right)^2}{2^{-(\ell-m)}}
$$

$$
+ \frac{\left(2^{\ell+m} - 2 \right) \left(\left(\ell + 1 \right) 2^{-(\ell-m)} + \left(4 - \ell + 2m \right) \right)}{2^m} \Bigg\} \tag{77}
$$

$$
= \left(2^{\ell} - 1 \right) \left(2^m - 1 \right) \Bigg\{ 2^{\ell+m} - 2 + 2^{2-(\ell+m)} + \left(\ell + 1 \right)^2 2^{-(\ell-m)}
$$

$$
+ 2 \left(4 - \ell + 2m \right) \left(l + 1 \right) + \left(4 - \ell + 2m \right)^2 2^{\ell-m} + \left(\ell + 1 \right) 2^m
$$

$$
+ \left(4 - \ell + 2m \right) 2^{\ell} - 2 \left(\ell + 1 \right) 2^{-\ell} - 2 \left(4 - \ell + 2m \right) 2^{-m} \Bigg\}
$$

$$
= \mathcal{O}\{ 2^{2(\ell+m)} \}
$$

Due to symmetry, if $\ell \geq m$ then we have (77) with the indices m and ℓ interchanged.

Lemma 6. Under the conditions and notations set out so far, for the nondecimated family of discrete Haar wavelets we have that the order of the cross terms is:

$$T_{\mathcal{B}\mathcal{B}}(\ell, m) = \sum_{k \notin \mathcal{B}} \sum_{n \in \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2 = \mathcal{O}(2^{2(\ell+m)})$$

$$T_{\mathcal{B}\mathcal{B}}(\ell, m) = \sum_{k \in \mathcal{B}} \sum_{n \notin \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2 = \mathcal{O}(2^{2(\ell+m)}).$$

Proof. We bound $T_{\mathcal{B}\mathcal{B}}(\ell, m)$ and include the indices ℓ, m explicitly. For $\ell < m$ we have

$$T_{\mathcal{B}\mathcal{B}}(\ell, m) = \sum_{k \notin \mathcal{B}} \sum_{n \in \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2$$

$$= \sum_{p,j=1}^{\infty} U(j, p, m) V(j, p, \ell), \tag{78}$$

where

$$U(j, p, m) = \sum_{n \in \mathcal{B}} |i_{N,z}(j, m, n) i_{N,z}(p, m, n)|,$$

and

$$V(j, p, \ell) = \sum_{k \notin \mathcal{B}} |i_{N,z}(j, \ell, k) i_{N,z}(p, \ell, k)|. \tag{79}$$

Now, for $n \in \mathcal{B}$ we can apply inequality (18) to obtain

$$U(j, p, m) \leq \sum_{n \in \mathcal{B}} 2^{-(j+m)/2} \left(\min(N_j, N_m) + N_m \right) 2^{-(p+m)/2} \left(\min(N_p, N_m) + N_m \right)$$

$$= (N_m - 1) \, 2^{-(j+m)/2} \left(2^{\min(j,m)} + 2^m \right) 2^{-(p+m)/2} \left(2^{\min(p,m)} + 2^m \right)$$

$$= (N_m - 1) \, 2^{-m} 2^{-(j+p)/2} \left(2^{\min(j,m)} + 2^m \right) \left(2^{\min(p,m)} + 2^m \right)$$

$$= (N_m - 1) \, 2^{-m} 2^{-(j+p)/2}$$

$$\times \left\{ 2^{\min(j,m)+\min(p,m)} + 2^{m+\min(j,m)} + 2^{m+\min(p,m)} + 2^{2m} \right\}.$$

Now let us examine the sum over k in (78) using the Ψ bound for $i_{N,z}$ from (14).

$$
\begin{aligned}
V(j, p, \ell) &\leq \sum_{k \notin \mathcal{B}} |\Psi_{j,\ell}(k - 2[zT] + N/2 - 1)\Psi_{p,\ell}(k - 2[zT] + N/2 - 1)| \\
&\leq \sum_{k} |\Psi_{j,\ell}(k - 2[zT] + N/2 - 1)\Psi_{p,\ell}(k - 2[zT] + N/2 - 1)| \\
&= \sum_{q} |\Psi_{j,\ell}(q)||\Psi_{p,\ell}(q)| = B_\ell(j, p),
\end{aligned}
$$

where $B_\ell(j, p)$ is the fourth-order cross-correlation wavelet absolute value product of order $r = 0$, defined as $B_\ell^{(r)}(j, i) = \sum_{p=-\infty}^{\infty} |p|^r |\Psi_{j,\ell}(p)\Psi_{i,\ell}(p)|$ for $r = 0, 1$ and scales $\ell, j, i \in \mathbb{N}$.

Hence

$$
\begin{aligned}
T_{\mathcal{B}\mathcal{B}}(\ell, m) \leq (N_m - 1)2^{-m} &\left[\sum_{j=1}^{m} \sum_{p=1}^{m} 2^{-(j+p)/2} \left(2^{j+p} + 2^{m+j} + 2^{m+p} + 2^{2m} \right) B_\ell(j, p) \right. \\
&+ \sum_{j=1}^{m} \sum_{p=m+1}^{\infty} 2^{-(j+p)/2} \left(2^{j+m} + 2^{m+j} + 2^{2m} + 2^{2m} \right) B_\ell(j, p) \\
&+ \sum_{j=m+1}^{\infty} \sum_{p=1}^{m} 2^{-(j+p)/2} \left(2^{m+p} + 2^{2m} + 2^{m+p} + 2^{2m} \right) B_\ell(j, p) \\
&+ \left. \sum_{j=m+1}^{\infty} \sum_{p=m+1}^{\infty} 2^{-(j+p)/2} \left(2^{2m} + 2^{2m} + 2^{2m} + 2^{2m} \right) B_\ell(j, p) \right]
\end{aligned}
$$

Using $B_\ell(j, p) \leq K 2^{-(j+p)/2} 2^{2\ell}$ (see Proposition F.1), we obtain

$$T_{\mathcal{BB}}(\ell, m) \leq K 2^{2\ell}(N_m - 1)2^{-m} \left[\sum_{j=1}^{m} \sum_{p=1}^{m} 2^{-(j+p)/2} \left(2^{(j+p)/2} + 2^{m+(j-p)/2} \right. \right.$$

$$\left. + 2^{m+(p-j)/2} + 2^{2m-(j+p)/2} \right)$$

$$+ \sum_{j=1}^{m} \sum_{p=m+1}^{\infty} 2^{-(j+p)/2} \left(2^{m+(j-p)/2} + 2^{m+(j-p)/2} + 2^{1+2m-(j+p)/2} \right)$$

$$+ \sum_{j=m+1}^{\infty} \sum_{p=1}^{m} 2^{-(j+p)/2} \left(2^{m+(p-j)/2} + 2^{1+2m-(j+p)/2} + 2^{m+(p-j)/2} \right)$$

$$\left. + \sum_{j=m+1}^{\infty} \sum_{p=m+1}^{\infty} 2^{-(j+p)/2} 2^{2+2m-(j+p)/2} \right]$$

$$= K 2^{2\ell}(N_m - 1)2^{-m} \left[\sum_{j=1}^{m} \sum_{p=1}^{m} \left(2^0 + 2^{m-p} + 2^{m-j} + 2^{2m-j-p} \right) \right.$$

$$+ \sum_{j=1}^{m} \sum_{p=m+1}^{\infty} \left(2^{1+m-p} + 2^{1+2m-j-p} \right) + \sum_{j=m+1}^{\infty} \sum_{p=1}^{m} \left(2^{1+m-j} + 2^{1+2m-j-p} \right)$$

$$\left. + \sum_{j=m+1}^{\infty} \sum_{p=m+1}^{\infty} 2^{2+2m-j-p} \right]$$

$$= K 2^{2\ell} (N_m - 1) 2^{-m} \left[m^2 + 2m \left(2^m - 1 \right) + \left(2^{2m} - 2^{m+1} + 1 \right) + 2m \right.$$

$$\left. + 2 \left(2^m - 1 \right) + 2m + 2 \left(2^m - 1 \right) + 2^2 \right]$$

$$= K 2^{2\ell} (N_m - 1) 2^{-m} \left(m + 1 + 2^m \right)^2$$

$$= \mathcal{O} \left(2^{2(\ell+m)} \right).$$

By similar calculations the other cross term $T_{\mathcal{BB}}(\ell, m) = \mathcal{O}(2^{2(\ell+m)})$.

Lemma 7. Under the conditions and notations set out so far, for the nondecimated family of discrete Haar wavelets we have that

$$T_{\mathcal{BB}}(\ell, m) = \sum_{k \notin \mathcal{B}} \sum_{n \notin \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2 = \mathcal{O}\{2^{2(\ell+m)}\}.$$

Proof. Using the definition of V from (79),

$$
\begin{aligned}
T_{\mathcal{B}\mathcal{B}}(\ell, m) &= \sum_{k \notin \mathcal{B}} \sum_{n \notin \mathcal{B}} \left\{ \sum_{j=1}^{\infty} |i_{N,z}(j, \ell, k) i_{N,z}(j, m, n)| \right\}^2 \\
&= \sum_{p,j=1}^{\infty} V(j, p, m) V(j, p, \ell) \\
&\leq \sum_{p,j=1}^{\infty} B_m(j, p) B_\ell(j, p) \\
&\leq K^2 \sum_{p,j=1}^{\infty} 2^{2(\ell+m)} 2^{-(j+p)} \\
&= K' 2^{2(\ell+m)} \text{ for some constant } K' \\
&= \mathcal{O}\left\{ 2^{2(\ell+m)} \right\}.
\end{aligned}
$$

Appendix F: Some Exact Formulae for Haar wavelets

Fourth-order absolute value wavelet cross-correlations for Haar wavelets. In what follows we demonstrate new results on the the fourth-order absolute value wavelet cross-correlations, for Haar wavelets which were used in showing the previous results in Appendix E.5.

Recall these were defined as $B_\ell^{(r)}(j, i) = \sum_{p=-\infty}^{\infty} |p|^r |\Psi_{j,\ell}(p) \Psi_{i,\ell}(p)|$ for $r = 0, 1$ and scales $\ell, j, i \in \mathbb{N}$.

The B products are symmetric in their arguments, $B_\ell^{(r)}(j, i) = B_\ell^{(r)}(i, j)$. Note that for $r = 0$, for ease of notation, these $B^{(0)}$ quantities appeared as B in the previous proofs.

Proposition F.1. *For Haar wavelets. (Part A) For $i, j > \ell$:*

$$
B_\ell^{(0)}(j, i) = \begin{cases} 2^{-j}(2^{2\ell-1} + 1) & \text{for } j = i, \\ 2^{-j}(2^{2\ell-1} + 1)2^{-3/2} & \text{for } i = j + 1, \\ 2^{-j/2}2^{-i/2}(2^{2\ell-1} + 1)/6 & \text{for } |j - i| > 1. \end{cases}
$$

Also, for all i, j, ℓ such that $i, j > \ell$, $B_\ell^{(0)}(j, i)$ is bounded by

$$
B_\ell^{(0)}(j, i) \leq 2^{-j/2} 2^{-i/2} 2^{2\ell}.
$$

(Part B) For $i, j < \ell$:

$$
B_\ell^{(0)}(j, i) = \begin{cases} 2^{-\ell}(2^{2j-1} + 1) & \text{for } i = j < \ell, \\ \frac{3}{2} 2^{-\ell} 2^{-j/2} 2^{5i/2 - 1} & \text{for } i < j < \ell. \end{cases}
$$

(Part C) For $i < \ell < j$:

$$B_\ell^{(0)}(j,i) = \begin{cases} \frac{1}{8}2^{-(\ell+1)/2}2^{3i/2}(2^{i-\ell}+2) & \text{for } j = \ell+1, \\ \frac{1}{8}2^{-j/2}2^{3i/2}(2-2^{i-\ell}) & \text{for } j > \ell+1. \end{cases} \tag{80}$$

(Part D) For $\ell = j$ and $i > \ell$:

$$B_\ell^{(0)}(\ell,i) = 2^{-(i-\ell)/2}\begin{cases} \frac{17}{9}2^{\ell-3} & \text{for } i = \ell+1, \\ \frac{17}{27}2^{\ell-3} & \text{for } i > \ell+1. \end{cases} \tag{81}$$

For $i < \ell$ we have the following bound:

$$B_\ell^{(0)}(\ell,i) \leq 2^{3i/2}2^{-\ell/2}.$$

(Part E) Finally, when all indices are equal we can use (34) from Nason, von Sachs and Kroisandt (2000) to show

$$B_\ell^{(0)}(\ell,\ell) = \sum_p \Psi_\ell^2(p) = A_{\ell,\ell} = \frac{1}{3}2^{-\ell}(2^{2\ell}+5),$$

for $\ell > 0$ and A is the matrix from Nason, von Sachs and Kroisandt (2000).

The symmetry of B permits evaluation of $B_\ell^{(0)}(j,i)$ for other orderings of (i,j).

An overall bound for all i,j,ℓ is $B_\ell^{(0)}(j,i) \leq K2^{-(j+i)/2}2^{2\ell}$ for some positive constant K.

Proof. <u>Part A:</u> for $i,j > \ell$. We can work out the exact formula for $B_\ell(j,j)$ by direct application of the formula (33) for $\Psi_{j,\ell}(p)$ for $j > \ell$.

$$\begin{aligned}
B_\ell^{(0)}(j,j) &= \sum_p \Psi_{j,\ell}^2(p) \\
&= 2^{-j+\ell}\left\{ \sum_{p=-2^\ell}^{-2^{\ell-1}-1}(2^{-\ell}p+1)^2 + \sum_{p=-2^{\ell-1}}^{-1}(2^{-\ell}p)^2 \right. \\
&\quad + 2^{-2\ell}\sum_{p=2^{j-1}-2^\ell}^{2^{j-1}-2^{\ell-1}-1}(2p-2^j+2^{\ell+1})^2 \\
&\quad + 2^{-2\ell}\sum_{p=2^{j-1}-2^{\ell-1}}^{2^{j-1}-1}(2^j-2p)^2 \\
&\quad + 2^{-2\ell}\sum_{p=2^j-2^\ell}^{2^j-2^{\ell-1}-1}(2^j-p-2^\ell)^2 \\
&\quad \left. + 2^{-2\ell}\sum_{p=2^j-2^{\ell-1}}^{2^j-1}(p-2^j)^2 \right\}.
\end{aligned} \tag{82}$$

Range	j	i
I	$-2^\ell \leq p < 0$	$-2^\ell \leq p < 0$
II	$2^{j-1} - 2^\ell \leq p < 2^{j-1}$	$2^{i-1} - 2^\ell \leq p < 2^{i-1}$
III	$2^j - 2^\ell \leq p < 2^j$	$2^i - 2^\ell \leq p < 2^i.$

Hence, after some algebra,

$$B_\ell^{(0)}(j,j) = 2^{-j+\ell}(2^\ell + 2^{\ell-1} + 2^{1-\ell} + 2^{-\ell})/3 = 2^{-j}(2^{2\ell-1} + 1). \qquad (83)$$

Next, we examine:

$$B_\ell^{(0)}(j,i) = B_\ell^{(0)}(j,j+1) = \sum_p |\Psi_{j,\ell}(p)\Psi_{j+1,\ell}(p)|,$$

for $i = j + 1$. Examining Table 1 shows that Range I for both the j and $j + 1$ cross-correlation wavelets always overlap and Range II for $i = j + 1$ overlaps with Range III for j. Hence:

$$B_\ell^{(0)}(j, i = j+1) = \sum_{p \in \text{Range I}} |\Psi_{j,\ell}(p)\Psi_{j+1,\ell}(p)| + \sum_{p \in \text{Range 3}} |\Psi_{j,\ell}(p)\Psi_{j+1,\ell}(p)|$$

$$= 2^{-(j-\ell)/2} 2^{-(i-\ell)/2} \left\{ \sum_{p=-2^\ell}^{-2^{\ell-1}-1} (2^{-\ell}p + 1)^2 + \sum_{p=-2^{\ell-1}}^{-1} (2^{-\ell}p)^2 \right.$$

$$+ 2^{-2\ell} \sum_{p=2^j-2^\ell}^{2^j-2^{\ell-1}-1} |(2p - 2^i + 2^{\ell+1})(2^j - p - 2^\ell)|$$

$$+ 2^{-2\ell} \left. \sum_{p=2^j-2^{\ell-1}}^{2^j-1} |(2^i - 2p)(p - 2^j)| \right\}$$

$$= 2^\ell 2^{-j/2} 2^{-(j+1)/2} \times 2^{-2-\ell}(2^{2\ell} + 2)$$

$$= 2^{-j} 2^{-3/2}(2^{2\ell-1} + 1). \qquad (84)$$

Finally for $|i - j| > 1$ we examine:

$$B_\ell^{(0)}(j,i) = \sum_p |\Psi_{j,\ell}(p)\Psi_{i,\ell}(p)|$$

$$= 2^{-(j-\ell)/2} 2^{-(i-\ell)/2} \left\{ \sum_{p=-2^\ell}^{-2^{\ell-1}-1} (2^{-\ell}p + 1)^2 + \sum_{p=-2^{\ell-1}}^{-1} (2^{-\ell}p)^2 \right\}$$

$$= 2^{-j/2} 2^{-i/2}(2^{2\ell-1} + 1)/6. \qquad (85)$$

Part B: for the case $i, j < \ell$. First, for $j = i$ we have

$$B_\ell^{(0)}(j,j) = \sum_p \Psi_{j,\ell}^2(p) = \sum_p \Psi_{\ell,j}^2(-p) = \sum_p \Psi_{\ell,j}^2(p) = B_L^{(0)}(J,J),$$

where $L = j$ and $J = \ell$ and we used $\Psi_{j,\ell}(p) = \Psi_{\ell,j}(-p)$ since we will use the formula for $\Psi_{\ell,j}$ where $\ell > j$ in (33). Since $L = j < \ell = J$ this puts us into the situation of (83) which gives:

$$B_\ell^{(0)}(j, j) = 2^{-J}(2^{2L-1} + 1) = 2^{-\ell}(2^{2j-1} + 1),$$

as required. For $i < j < \ell$ we use the formula for Ψ for $j < \ell$ given by (35). Due to the form of $\Psi_{j,\ell}(p)$ for $i < j < \ell$ we can split the sum into three parts corresponding to the the three non-zero parts of the autocorrelation wavelet given in (35). The condition $i < j < \ell$ is helpful as the interval associated with i nests within that of j, and the j interval nests within that associated with ℓ. First, we will deal with the last two terms of (35) which do not depend on ℓ ("front part").

$$
\begin{aligned}
\text{front} &= \sum_{p=1}^{2^i} |\Psi_{j,\ell}(p)||\Psi_{i,\ell}(p)| \\
&= \sum_{p=1}^{2^{i-1}} |-2^j p||-2^i p| + \sum_{p=2^{i-1}+1}^{2^i} |-2^{-j} p||1 - 2^{-i} p|
\end{aligned}
\tag{86}
$$

Note how the first term in each of these sums is $|-2^{-j}p|$ because this is the formula for $\Psi_{j,\ell}(p)$ over the interval $[1, 2^i]$ because this interval is always contained within $[1, 2^{j-1}]$ since $i < j$. We also ensure that the last absolute value term in the sum is positive (if it was negative then we'd switch signs of the contents of the absolute value as $|x| = -x$ if x is negative). Continuing

$$\text{front} = 2^{-i-j} \sum_{p=1}^{2^{i-1}} p^2 + 2^{-j} \sum_{p=2^{i-1}+1}^{2^i} p(1 - 2^{-i}p) = 2^{-j+2i-3}.$$

The middle term is constructed in a similar way except that ℓ enters into the equation. However, the concept that the j interval can only ever overlap the i

interval on its first half still remains. Hence,

$$
\begin{aligned}
\text{middle} &= \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^i} |\Psi_{j,\ell}(p)||\Psi_{i,\ell}(p)| \\
&= \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^{i-1}} 2^{-j}(2^\ell + 2p)2^{-i}(2\ell + 2p) \\
&\quad + \sum_{p=-2^{\ell-1}+2^{i-1}+1}^{-2^{\ell-1}+2^i} 2^{-j}(2^\ell + 2p)2^{-i}(2^{i+1} - 2^\ell - 2p) \\
&= 2^{-j-i} \left\{ \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^{i-1}} (2^\ell + 2p)^2 \right. \\
&\quad \left. + \sum_{p=-2^{\ell-1}+2^{i-1}+1}^{-2^{\ell-1}+2^i} (2^\ell + 2p)(2^{i+1} - 2^\ell - 2p) \right\} \\
&= 2^{-j}2^{2i-1}. \tag{87}
\end{aligned}
$$

The back part uses precisely the same part as the middle

$$
\begin{aligned}
\text{back} &= \sum_{p=-2^\ell+1}^{-2^\ell+2^{i-1}} |-2^{-j}(p+2^\ell)||-2^{-i}(p+2^\ell)| \\
&\quad + \sum_{p=-2^\ell+2^{i-1}+1}^{-2^\ell+2^i} |-2^{-j}(p+2^\ell)||2^{-i}(2^\ell + p - 2^i)| \\
&= 2^{-j-i} \left\{ \sum_{p=-2^\ell+1}^{-2^\ell+2^{i-1}} (p+2^\ell)^2 + \sum_{p=-2^\ell+2^{i-1}+1}^{-2^\ell+2^i} (p+2^\ell)(2^i - p - 2^\ell) \right\} \\
&= 2^{-2}2^{-j}2^{2i-1}. \tag{88}
\end{aligned}
$$

Adding the front, middle and back components together and multiplying through by the constant that appears at the front of (35) gives:

$$
\begin{aligned}
B_\ell^{(0)}(j,i) &= 2^{-(\ell-j)/2}2^{-(\ell-i)/2}(2^{-j}2^{-2}2^{2i-1} + 2^{-j}2^{2i-1} + 2^{-2}2^{-j}2^{2i-1}) \\
&= 2^{-\ell}2^{i/2}2^{j/2} \times 2^{-j}2^{2i-1}(2^{-2} + 1 + 2^{-2}) \\
&= \frac{3}{2}2^{-\ell}2^{-j/2}2^{5i/2-1}, \tag{89}
\end{aligned}
$$

as required.

<u>Part C:</u> for the case $i < \ell < j$. When considering the formula for $B_\ell^{(0)}(j,i)$ we use formula (33) for $\Psi_{j,\ell}(p)$ because $\ell < j$ and formula (35) for $\Psi_{i,\ell}(p)$ because $i < \ell$.

Let us first consider positive p first. In this case, the only values of $\Psi_{i,\ell}(p)$ which are nonzero are for $0 < p \leq 2^i$ and the only values of $\Psi_{j,\ell}(p)$ are for $2^{j-1} - 2^\ell \leq p$. So, if $2^i < 2^{j-1} - 2^\ell$ then there is no overlap between these two cross-correlation wavelets, under what conditions can this occur (subject to $i < \ell < j$)? Suppose $j = \ell + 1$ then $2^{j-1} - 2^\ell = 0$ and 2^i is never less than or equal to zero. Suppose $j = \ell + 2$, then $2^{j-1} - 2^\ell = 2^\ell$ which is always greater than 2^i because we know that $\ell > i$. Hence, we have to consider two cases (i) $j = \ell + 1$ where the positive parts of the cross-correlation functions overlap and $j > \ell + 1$ where they do not.

We will now work out this 'overlap' contribution where $j = \ell + 1$:

$$\text{`overlap'} = \sum_{p=1}^{2^i} |\Psi_{\ell+1,\ell}(p)||\Psi_{i,\ell}(p)|.$$

The limits of the sum correspond to where the term $\Psi_{i,\ell}(p)$ is non-zero for $p \geq 0$. Now we consider which parts of $\Psi_{\ell+1,\ell}(p)$ are relevant to the sum. With $j = \ell + 1$ the fifth and sixth ranges in (33) turn into

$$0 \leq \tau < 2^{\ell-1}, \tag{90}$$

$$2^{\ell-1} \leq \tau < 2^\ell, \tag{91}$$

respectively. Since $i < \ell$ we have $2^i < 2^\ell$ and so the two ranges in (90) and (91) are the only ones we need to consider for $\Psi_{\ell+1,\ell}(p)$ for positive p. Now the highest that i can be is $i = \ell - 1$ so that maximum p is $2^i = 2^{\ell-1}$ so, in actuality, it is only range (90) that is active for any $p > 0$.

Hence, without the constant $2^{-(j-\ell)/2}2^{-(\ell-i)/2} = 2^{-j/2}2^{i/2}$ at the front of (33) and (35), we can write

$$\text{`overlap'} = \sum_{p=1}^{2^i} |2^{-\ell}(2p - 2^{\ell+1} + 2^{\ell+1})||\Psi_{i,\ell}(p)|$$

$$= 2^{-\ell+1} \sum_{p=1}^{2^i} p|\Psi_{i,\ell}(p)|$$

$$= 2^{-\ell+1} \left\{ \sum_{p=1}^{2^{i-1}} p| - 2^{-i}p| + \sum_{p=2^{i-1}+1}^{2^i} p|1 - 2^{-i}p| \right\}. \tag{92}$$

Note that $1 - 2^{-i}p$ is positive on the range of p in the second sum. Hence,

$$\text{`overlap'} = 2^{-\ell+1} \left\{ 2^{-i} \sum_{p=1}^{2^{i-1}} p^2 + \sum_{p=2^{i-1}+1}^{2^i} p(1 - 2^{-i}p) \right\}$$

$$= 2^{-\ell}2^{2i}2^{-2}, \tag{93}$$

for $i < \ell$ and $j = \ell + 1$.

For $p < 0$, $i < \ell < j$ the cross-correlation wavelet $\Psi_{j,\ell}(p)$ splits into two parts random each of those two parts the cross-correlation wavelet $\Psi_{i,\ell}(p)$ splits into two further parts. Hence, the sum on the $p < 0$ terms (the 'back' bit) is given by:

$$\text{'Back'} = \sum_{p=-2^\ell+1}^{-2^\ell+2^i} |\Psi_{j,\ell}(p)||\Psi_{i,\ell}(p)| + \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^i} |\Psi_{j,\ell}(p)||\Psi_{i,\ell}(p)|$$

$$= \sum_{p=-2^\ell+1}^{-2^\ell+2^i} |-(2^{-\ell}p+1)||\Psi_{i,\ell}(p)| + \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^i} |2^{-\ell}p||\Psi_{i,\ell}(p)| \qquad (94)$$

We can now split each of these two sums into the two sets of two ranges dictated by the domain of $\Psi_{i,\ell}(p)|$ as follows:

$$\text{'Back'} = \sum_{p=-2^\ell+1}^{-2^\ell+2^{i-1}} |1 + 2^{-\ell}p||2^{-i}(p+2^\ell)|$$

$$+ \sum_{p=-2^\ell+2^{i-1}+1}^{-2^\ell+2^i} |1 + 2^{-\ell}p||2^{-i}(2^\ell + p - 2^i)|$$

$$+ \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^{i-1}} |2^{-\ell}p||2^{-i}(2^\ell + 2p)|$$

$$+ \sum_{p=-2^{\ell-1}+2^{i-1}+1}^{-2^{\ell-1}+2^i} |2^{-\ell}p||2^{-i}(2^{i+1} - 2^\ell - 2p)| \qquad (95)$$

For all of the eight components in the four sums in (95) all except the fourth, fifth and seventh are always positive over their respective sum's range of p values. Hence, we replace those terms in the absolute values by their negative (e.g. $|x| = -x$ for $x < 0$) and obtain:

$$\text{'Back'} = 2^{-i} \sum_{p=-2^\ell+1}^{-2^\ell+2^{i-1}} (1 + 2^{-\ell}p)(p+2^\ell)$$

$$+ 2^{-i} \sum_{p=-2^\ell+2^{i-1}+1}^{-2^\ell+2^i} (1 + 2^{-\ell}p)(2^i - p - 2^\ell)$$

$$+ 2^{-\ell-i} \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^{i-1}} (-p)(2^\ell + 2p)$$

$$+ 2^{-\ell-i} \sum_{p=-2^{\ell-1}+2^{i-1}+1}^{-2^{\ell-1}+2^i} (-p)(2^{i+1} - 2^\ell - 2p)$$

$$= 2^{i-3}(2 - 2^{i-\ell}). \qquad (96)$$

Now multiplying by the $2^{-j/2}2^{i/2}$ which we omitted earlier gives:

$$B_\ell^{(0)}(j,i) = 2^{-j/2}2^{i/2}2^{i-3}(2 - 2^{i-\ell})$$
$$= 2^{-j/2}2^{3i/2-3}(2 - 2^{i-\ell}), \tag{97}$$

as required for the second equation in (80) for $i < \ell$ and $\ell + 1 < j$.

For the first equation in (80) we have to add $2^{-j/2}2^{i/2}$ times equation (93) to obtain:

$$B_\ell^{(0)}(j,i) = 2^{-j/2}2^{3i/2-3}(2 - 2^{i-\ell}) + 2^{-j/2}2^{i/2}2^{-\ell}2^{2i}2^{-2}$$
$$= \frac{1}{8}2^{-(\ell+1)/2}2^{3i/2}(2^{i-\ell} + 2), \tag{98}$$

as required.

<u>Part D:</u> First consider $j = \ell < i$. Then:

$$B_\ell^{(0)}(j,i) = \sum_p |\Psi_{j,\ell}(p)||\Psi_{i,\ell}(p)| = \sum_p |\Psi_\ell(p)||\Psi_{i,\ell}(p)|,$$

where $\Psi_\ell(p)$ is the ordinary autocorrelation wavelet from Nason, von Sachs and Kroisandt (2000). The domain of $\Psi_\ell(p)$ is from $-2^\ell < p2^\ell$. The range of $\Psi_\ell(p)$ and $\Psi_{i,\ell}(p)$ agrees for $p < 0$ but the two wavelets only overlap for $p > 0$ if the lower end of the nonzero range of $\Psi_{i,\ell}(p)$, namely $2^{i-1} - 2^\ell$ is smaller than the upper end of the autocorrelation wavelet 2^ℓ, recalling that $i > \ell$. This only occurs when $i = \ell + 1$ and for $i \geq \ell + 2$ there is no overlap for $p > 0$.

For this latter case let use look at the negative range of p values:

$$
\begin{aligned}
B_\ell^{(0)}(\ell, i) &= \sum_{p=-2^\ell+1}^{-2^{\ell-1}} |\Psi_\ell(p)||\Psi_{i,\ell}(p)| + \sum_{p=-2^{\ell-1}+1}^{0} |\Psi_\ell(p)||\Psi_{i,\ell}(p)| \\
&= 2^{-(i-\ell)/2} \left\{ \sum_{p=-2^\ell+1}^{-2^{\ell-1}} |2^{-\ell}|p| - 1|| - (2^{-\ell}p + 1)| \right. \\
&\quad + \left. \sum_{p=-2^{\ell-1}+1}^{0} |1 - 3|p|2^{-\ell}||2^{-\ell}p| \right\} \\
&= 2^{-(i-\ell)/2} \left\{ \sum_{p=-2^\ell+1}^{-2^{\ell-1}} |-2^{-\ell}p - 1||2^{-\ell}p + 1| \right. \\
&\quad + \left. \sum_{p=-2^{\ell-1}+1}^{0} |1 + 3p2^{-\ell}|(-2^{-\ell}p) \right\} \\
&= 2^{-(i-\ell)/2} \left\{ \sum_{p=-2^\ell+1}^{-2^{\ell-1}} (2^{-\ell}p + 1)^2 \right. \\
&\quad + \left. \sum_{p=-2^{\ell-1}+1}^{0} |1 + 3p2^{-\ell}|(-2^{-\ell}p) \right\} \\
&= 2^{-(i-\ell)/2} \{C2 + C1\}. \tag{99}
\end{aligned}
$$

The sum $C2 = \sum_{p=-2^\ell+1}^{-2^{\ell-1}} (2^{-\ell}p + 1)^2 = \frac{1}{3}2^{-3-\ell}(2 + 3 \cdot 2^\ell + 2^{2\ell})$. The sum $C1$ can be shown to be approximated by $(2^{3+\ell} - 9 \cdot 2^{1-\ell} - 27)/216$ in Mathematica with error bounded by

$$
2|1 + 2 \cdot 2^{-\ell}(-2^\ell/3 - 1)|| - 2^\ell(-2^\ell/3 - 1)| \le 5 \cdot 2^{-\ell}.
$$

Hence, for $i \ge \ell + 2$ we have

$$
B_\ell^{(0)}(\ell, i) = \tfrac{17}{27}2^{\ell-3}. \tag{100}
$$

Now, for $i = \ell + 1$ let us examine the contribution for $p > 0$:

$$
\begin{aligned}
\text{`Front'} &= \sum_{p=1}^{2\ell} |\Psi_\ell(p)||\Psi_{\ell+1,\ell}(p)| \\
&= \sum_{p=1}^{2^{\ell-1}} |\Psi_\ell(p)||\Psi_{\ell+1,\ell}(p)| + \sum_{p=2^{\ell-1}+1}^{2^\ell} |\Psi_\ell(p)||\Psi_{\ell+1,\ell}(p)| \\
&= 2^{-1/2} \left\{ \sum_{p=1}^{2^{\ell-1}} |2 - 3|p|2^{-\ell}||2^{-\ell}(2p - 2^{\ell+1} + 2^{\ell+1})| \right. \\
&\quad \left. + \sum_{p=2^{\ell-1}+1}^{2^\ell} |2^{-\ell}|p| - 1||2^{-\ell}(2^{\ell+1} - 2p)| \right\} \\
&= 2^{-1/2} \left\{ 2^{-\ell+1} \sum_{p=1}^{2^{\ell-1}} |1 - 3p2^{-\ell}|p + 2^{-\ell} \sum_{p=2^{\ell-1}+1}^{2^\ell} |2^{-\ell}p - 1||2^{\ell+1} - 2p| \right\} \\
&= (27 - 9 \cdot 2^{1-\ell} + 2^{\ell+3})/108 + 2^{-2-\ell}(2^{2\ell} - 3 \cdot 2^\ell + 2)/3 \\
&= \frac{17}{27}2^{\ell-2}.
\end{aligned}
\tag{101}
$$

Hence, adding together (100) and (101) gives:

$$
B_\ell^{(0)}(\ell, i) = \frac{17}{9}2^{\ell-3}, \tag{102}
$$

for $i = \ell + 1$ as required.

For the case of $i < \ell$ we use the following bound:

$$
B_\ell^{(0)}(\ell, i) = \sum_p |\Psi_\ell(p)||\Psi_{i,\ell}(p)| = \sum_{p=-2^\ell+1}^{2^i} |\Psi_\ell(p)||\Psi_{i,\ell}(p)| \le \sum_{p=-2^\ell+1}^{2^i} |\Psi_{i,\ell}(p)|,
$$

as $|\Psi_\ell(p)|$ takes its maximum value of 1 at $p = 0$. We can work out this sum directly taking care to discover when the argument inside the absolute value sign is negative or positive giving:

$$
\begin{aligned}
2^{-(i-\ell)/2}B_\ell^{(0)}(\ell, i) &\le 2^{-i} \sum_{p=-2^\ell+1}^{-2^\ell+2^{i-1}} (p + 2^\ell) + 2^{-i} \sum_{p=-2^\ell+2^{i-1}+1}^{-2^\ell+2^i} (2^i - 2^\ell - p) \\
&\quad + 2^{-i} \sum_{p=-2^{\ell-1}+1}^{-2^{\ell-1}+2^{i-1}} (2^\ell + 2p) + 2^{-i} \sum_{p=-2^{\ell-1}+2^{i-1}+1}^{-2^{\ell-1}+2^i} (2^{i+1} - 2^\ell - 2p) \\
&\quad + 2^{-i} \sum_{p=1}^{2^{i-1}} p + \sum_{p=2^{i-1}+1}^{2^i} (1 - 2^{-i}p),
\end{aligned}
\tag{103}
$$

Hence

$$B_\ell^{(0)}(\ell, i) \leq 2^{3i/2} 2^{-\ell/2}. \tag{104}$$

as required.

Appendix G: LPACF of Eastport Precipitation Data at Different Window Widths

The plots below were produced by the following functions executed using the `lpacf` package with binwidths of 160, 80 and 40.

```
function(binwidth=250){
#
# Compute the Epanechnikov kernel smoothed local PACF using
# a specified binwidth, using parallel processing function
# mclapply on all points.
#
# Then plot the answer: only the first four lags
#  using colours 1 thru 4.
#
plot(lpacf.Epan(EastPortPrecip, allpoints=TRUE, binwidth=binwidth,
            lapplyfn=mclapply), lags=1:4, lcol=1:4)
#
# Construct and plot "standard" confidence intervals
#
ci <- 1.96/sqrt(binwidth)
abline(h = c(-ci, ci), lty = 2, col = 2)
}
```
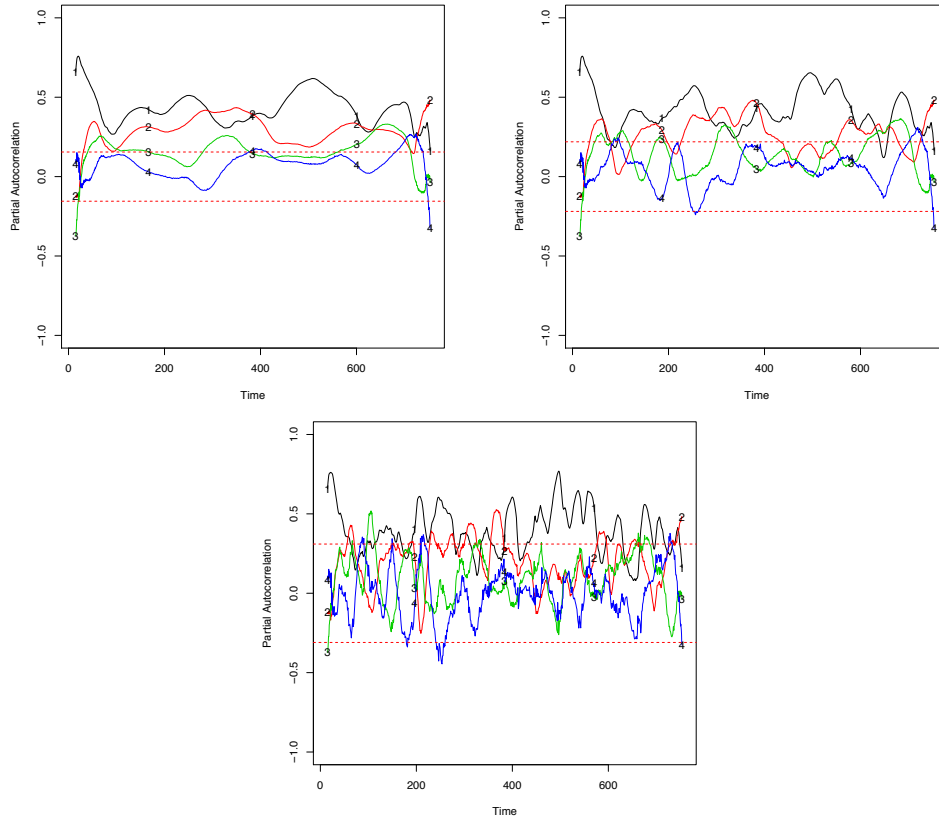
FIG 7. *Windowed partial autocorrelation of Eastport Precipitation Data, $\tilde{q}_W(z, \tau)$, of left for lags one to four indicated on each curve. Horizontal red dotted lines are approximate 95% confidence intervals. All plots were created with the Epanechnikov window with sizes a) $L = 160$, b) $L = 80$, c) $L = 40$.*