Alternative Estimators of Wavelet Variance

Eric M. Aldrich

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science

University of Washington

2005

Program Authorized to Offer Degree:
Statistics

University of Washington

Graduate School

This is to certify that I have examined this copy of a master's thesis by

Eric M. Aldrich

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

Donald Bame Percival

Peter Guttorp

Christopher Stephen Bretherton

Date: ______________________________

University of Washington

Abstract

Alternative Estimators of Wavelet Variance

by Eric M. Aldrich

Chair of Supervisory Committee:
Affiliate Associate Professor Donald B. Percival
Statistics

The wavelet variance is a scale-based decomposition of the process variance that is particularly well-suited for analyzing intrinsically stationary processes. This decomposition has proven to be useful for studying various geophysical time series, including some related to subtidal sea level variations, vertical shear in the ocean and variations in soil composition along a transect. Previous work has established the large sample properties of an unbiased estimator of the wavelet variance formed using the nonboundary wavelet coefficients from the maximal overlap discrete wavelet transform (MODWT). The present work considers two alternative estimators, one of which is unbiased, and the other, biased. The new unbiased estimator is appropriate for asymmetric wavelet filters such as the Daubechies filters of width four and higher and is obtained from the nonboundary coefficients that result from running a wavelet filter through a time series in both a forward and a backward direction. The biased estimator is constructed in a similar fashion, but utilizes all wavelet coefficients that result from filtering a time series in forward and backward directions. While the two alternative estimators have the same asymptotic distribution as the original unbiased estimator (with some restrictions in the case of the biased estimator), they can have substantially better statistical properties in small sample sizes. Formulae for evaluating the mean squared errors of the usual unbiased estimator and the two alternative estimators are developed and verified for several fractionally differenced processes via Monte Carlo experiments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

Wavelet analysis is a useful method of decomposing and analyzing a time series across different scales and in some ways is analogous to spectral methods, which provide a decomposition across frequencies. More specifically, wavelet analysis consists of transforming the series into wavelet coefficients of different levels $j = 1, 2, \ldots$, with each level corresponding to a particular scale $\tau_j = 2^{j-1}$. The coefficients are formed by filtering the time series using a sequence of filters, one for each level j. To be precise, a level 1 Maximal Overlap Discrete Wavelet Transform (MODWT) wavelet filter of width L_1 is a sequence of values, $\overrightarrow{h}_{1,0}, \overrightarrow{h}_{1,1}, \ldots, \overrightarrow{h}_{1,L_1-1}$, such that

$$\sum_{l=0}^{L_1-1} \overrightarrow{h}_{1,l} = 0, \quad \sum_{l=0}^{L_1-1} \overrightarrow{h}_{1,l}^2 = \frac{1}{2} \quad \text{and} \quad \sum_{l=0}^{L_1-1} \overrightarrow{h}_{1,l} \overrightarrow{h}_{1,l+2n} = 0$$

for all nonzero integers n, where $\overrightarrow{h}_{1,0} \neq 0 \neq \overrightarrow{h}_{1,L_1-1}$ (see, e.g., Percival and Walden 2000 [hereafter PW], eq. 163a). Similarly, the level 1 MODWT scaling filter is defined via the quadrature mirror relationship

$$\overrightarrow{g}_{1,l} \equiv (-1)^{l+1} \overrightarrow{h}_{1,L_1-1-l}$$

(PW, eq. 163d). If we let

$$\overrightarrow{H}_1(f) \equiv \sum_{l=0}^{L_1-1} \overrightarrow{h}_{1,l} e^{-i2\pi fl} \quad \text{and} \quad \overrightarrow{G}_1(f) \equiv \sum_{l=0}^{L_1-1} \overrightarrow{g}_{1,l} e^{-i2\pi fl}$$

denote the transfer functions of $\overrightarrow{h}_{1,l}$ and $\overrightarrow{g}_{1,l}$, respectively, then the transfer functions for the level j MODWT wavelet and scaling filters can be obtained by

$$\overrightarrow{H}_j(f) \equiv \overrightarrow{H}_1(2^{j-1}f) \prod_{l=0}^{j-2} \overrightarrow{G}_1(2^l f) \quad \text{and} \quad \overrightarrow{G}_j(f) \equiv \prod_{l=0}^{j-1} \overrightarrow{G}_1(2^l f)$$

(PW, eq. 169b). The level j wavelet and scaling filters result by taking the inverse DFT of the appropriate transfer function, and will be denoted $\overrightarrow{h}_{j,l}$ and $\overrightarrow{g}_{j,l}$. It can be shown that $\overrightarrow{h}_{j,l}$ satisfies

$$\sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} = 0, \quad \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l}^2 = \frac{1}{2^j} \quad \text{and} \quad \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} \overrightarrow{h}_{1,l+2^j n} = 0$$

for all nonzero integers n, where $L_j = (2^j - 1)(L_1 - 1) + 1$ and $\overrightarrow{h}_{j,0} \neq 0 \neq \overrightarrow{h}_{j,L_j-1}$ (PW, pp. 169 and 202). It will often be convenient to regard $\overrightarrow{h}_{j,l}$ and $\overrightarrow{g}_{j,l}$ as infinite sequences, where $\overrightarrow{h}_{j,l} = \overrightarrow{g}_{j,l} = 0$ for $l < 0$ and $l \geq L_j$.

Among the many possibilities of wavelet filters that satisfy the conditions above, we will restrict our attention to those of the Daubechies class, for which the squared gain function of a level 1 filter is defined as

$$\overrightarrow{\mathcal{H}}_1^{(D)}(f) = \left|\overrightarrow{H}_1(f)\right|^2 \equiv 2\sin^L(\pi f) \sum_{l=0}^{\frac{L_1}{2}-1} \begin{pmatrix} \frac{L_1}{2} - 1 + l \\ l \end{pmatrix} \cos^{2l}(\pi f) \tag{1.1}$$

(Daubechies 1992, sec. 6.1). The advantage of confining our analysis to Daubechies wavelet filters is that they can be interpreted as containing a number of embedded differencing operations, and thus allow us to handle intrinsically stationary processes rather than just stationary ones. In particular, it can be shown that the squared gain function for the difference filter, $\{d_0 = 1, d_1 = -1\}$, is $\mathcal{D}(f) = 4\sin^2(\pi f)$, and hence that Equation (1.1) can be expressed as

$$\overrightarrow{\mathcal{H}}_1^{(D)}(f) = \mathcal{D}^{\frac{L_1}{2}}(f)\mathcal{A}_{L_1}(f)$$

where

$$\mathcal{A}_{L_1}(f) \equiv \frac{1}{2^{L_1-1}} \sum_{l=0}^{\frac{L_1}{2}-1} \begin{pmatrix} \frac{L_1}{2} - 1 + l \\ l \end{pmatrix} \cos^{2l}(\pi f)$$

(PW, pp. 105–106). In other words, the squared gain function for a level 1 Daubechies wavelet filter of length L_1 can be interpreted as the cascade of $L_1/2$ difference filters with a low-pass filter, $\mathcal{A}_{L_1}(f)$, that acts as a weighted average (note that the value $L_1/2$ is always an integer, since L_1 must always be an even value in order to satisfy the conditions of a wavelet filter above). This interpretation can likewise be extended to arbitrary levels, where a level j Daubechies wavelet filter is interpreted as a filter cascade of $L_1/2$ difference filters and a modified low-pass filter (PW, pp. 304 and 535–536).

For a particular wavelet filter of any class, $\overrightarrow{h}_{j,l}$, we define the level j MODWT wavelet coefficients for an intrinsically stationary stochastic process, X_t, as

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} X_{t-l}, \quad \text{for} \quad t \in \mathbb{Z} \tag{1.2}$$

(PW, eq. 296a). In the case that $\overrightarrow{h}_{j,l}$ is of the Daubechies class, each of the wavelet coefficients in Equation (1.2) can be interpreted as differences of weighted averages over time scale $\tau_j = 2^{j-1}$, localized at a particular point in time (as determined by the index t). Hence, we see the particular advantage of restricting our attention to Daubechies wavelet filters.

As in spectral analysis, once a time series has been transformed into wavelet coefficients, and hence decomposed into differences of weighted averages over particular time scales, it is often of interest to the practitioner to determine which time scales best describe or provide the greatest contribution to the original series. One method of achieving this objective is to determine the variance of the wavelet coefficients for specified scales, or the wavelet variance, defined as

$$\nu^2(\tau_j) \equiv \text{var}\left\{\overline{W}_{j,t}\right\},$$

which is independent of t for intrinsically stationary processes (PW, p. 296). In analogy to the spectral density function, the wavelet variance provides an additive decomposition of the process variance

$$\sum_{j=1}^{\infty} \nu^2(\tau_j) = \text{var}\left\{X_t\right\}$$

(PW, eq. 296d), and hence $\nu^2(\tau_j)$ provides a good indication of the relative importance of scale τ_j to the original process, X_t.

In practice, given a time series that is regarded as a realization of a portion $X_0, X_1, \ldots, X_{N-1}$ of $\{X_t\}$, the MODWT coefficients are computed via

$$\overrightarrow{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} X_{t-l \bmod N}, \quad t = 0, 1, \ldots N-1, \tag{1.3}$$

(PW, eq. 169a). This is only a minor alteration of Equation (1.2), the difference being in the modulo operation which extends the series X_t in a circular fashion beyond both endpoints (i.e. $X_{-1} = X_{N-1}$, $X_{-2} = X_{N-2}$, $X_N = X_0$, $X_{N+1} = X_1$, etc.), often referred to as a circular boundary condition. From Equation (1.3) we see that the first $L_j - 1$ wavelet coefficients depend on the circular boundary condition, and are in some sense 'biased' since they rely on the assumption that $X_{t-l \bmod N}$ is a good surrogate for X_{t-l}. We will refer to these wavelet coefficients as 'boundary' coefficients, while those which do not rely on the circular boundary condition will be referred to as 'nonboundary' coefficients.

Historically, the wavelet variance has been estimated using only those values of $\overrightarrow{W}_{j,t}$ that are equivalent to $\overline{W}_{j,t}$, or the nonboundary wavelet coefficients:

$$\overrightarrow{\nu}_u^2(\tau_j) \equiv \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \overrightarrow{W}_{j,t}^2 = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \overline{W}_{j,t}^2, \tag{1.4}$$

where $M_j \equiv N - L_j + 1$ (PW, eq. 306b). It is trivial to see that $\overrightarrow{\nu}_u^2(\tau_j)$ is an unbiased estimator of $\nu^2(\tau_j)$, and hence it will be referred to as the traditional unbiased estimator of the wavelet variance.

In terms of the wavelet variance for stationary processes, we can assume, without loss of generality, that $\mathrm{E}\{X_t\} = 0$. This can be justified by first recognizing that

$$\begin{aligned}
\sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l}(X_{t-l \bmod N} - \mu_X) &= \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} X_{t-l \bmod N} - \mu_X \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} && (1.5)\\
&= \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} X_{t-l \bmod N} && (1.6)\\
&= \overrightarrow{W}_{j,t},
\end{aligned}$$

where Equation (1.6) follows from Equation (1.5) since $\sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} = 0$. Thus, the wavelet coefficients are invariant to the value of μ_X, and hence the wavelet variance, since it is defined in terms of the wavelet coefficients (Equation (1.4)), is also invariant to μ_X. The same will be true of the alternative estimators proposed in the ensuing chapters (by making appropriate substitutions for $\overrightarrow{h}_{j,l}$ and X_t), and hence we will assume that $\mathrm{E}\{X_t\} = 0$ in what follows.

This thesis will examine two alternative estimators for the wavelet variance and their statistical properties relative to those of the traditional unbiased estimator. Specifically, Chapter 2 will propose a new unbiased estimator and develop theory for a previously proposed biased estimator (Greenhall, Howe and Percival 1999). Chapter 3 will examine the asymptotic properties of these estimators, first under the assumption that the underlying series, X_t, is stationary, and then will extend the asymptotic results to certain nonstationary cases. Chapter 4 will follow by discussing simulation results and comparing the statistical properties of the estimators. Chapter 5 will state our conclusions and give some directions for future research.

Chapter 2

ALTERNATIVE ESTIMATORS OF WAVELET VARIANCE

2.1 Forward-Backward Unbiased Estimator

In this section we introduce an alternative unbiased estimator of the wavelet variance and derive an expression for its variance. To begin, we now refer to $\{\overrightarrow{h}_{j,l}\}$ as a forward wavelet filter and use it to define an associated backward wavelet filter,

$$\overleftarrow{h}_{j,l} \equiv \overrightarrow{h}_{j,L_j-1-l}, \quad l = 0, 1, \ldots, L_j - 1.$$

The backward filter is used to form a set of backward wavelet coefficients,

$$\overleftarrow{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \overleftarrow{h}_{j,l} X_{t-l \bmod N}, \quad t = 0, 1, \ldots, N-1, \tag{2.1}$$

which are combined with the forward wavelet coefficients, $\overrightarrow{W}_{j,t}$, to obtain

$$\overleftrightarrow{\nu}^2_u(\tau_j) \equiv \frac{1}{2M_j} \sum_{t=L_j-1}^{N-1} \left(\overrightarrow{W}^2_{j,t} + \overleftarrow{W}^2_{j,t} \right).$$

We will refer to this as the forward-backward unbiased estimator of the wavelet variance and will compare its variance, $\mathrm{var}\,(\overleftrightarrow{\nu}^2_u(\tau_j))$, with that of the traditional unbiased estimator, $\mathrm{var}\,(\overrightarrow{\nu}^2_u(\tau_j))$. In particular,

$$\mathrm{var}\,(\overrightarrow{\nu}^2_u(\tau_j)) = \frac{1}{M_j^2} \sum_{t'=L_j-1}^{N-1} \sum_{t=L_j-1}^{N-1} \mathrm{cov}\,(\overrightarrow{W}^2_{j,t'}, \overrightarrow{W}^2_{j,t}),$$

while

$$\begin{aligned} \mathrm{var}\,(\overleftrightarrow{\nu}^2_u(\tau_j)) &= \frac{1}{4M_j^2} \sum_{t'=L_j-1}^{N-1} \sum_{t=L_j-1}^{N-1} \Big(\mathrm{cov}\,(\overrightarrow{W}^2_{j,t'}, \overrightarrow{W}^2_{j,t}) + \mathrm{cov}\,(\overleftarrow{W}^2_{j,t'}, \overleftarrow{W}^2_{j,t}) \\ &\qquad + 2\mathrm{cov}\,(\overrightarrow{W}^2_{j,t'}, \overleftarrow{W}^2_{j,t}) \Big). \end{aligned}$$

Given two sequences $\{a_l\}$ and $\{b_l\}$ of width L_j, let us define

$$A_t = \sum_{l=0}^{L_j-1} a_l X_{t-l} \text{ and } B_t = \sum_{l=0}^{L_j-1} b_l X_{t-l}$$

and develop an expression for $\operatorname{cov}(A_{t'}^2, B_t^2)$. Now,

$$\begin{aligned}
\operatorname{cov}(A_{t'}^2, B_t^2) &= \operatorname{cov}\left(\left(\sum_k a_k X_{t'-k}\right)^2, \left(\sum_l b_l X_{t-l}\right)^2\right) \\
&= \operatorname{cov}\left(\sum_k \sum_{k'} a_k a_{k'} X_{t'-k} X_{t'-k'}, \sum_l \sum_{l'} b_l b_{l'} X_{t-l} X_{t-l'}\right) \\
&= \sum_k \sum_{k'} \sum_l \sum_{l'} a_k a_{k'} b_l b_{l'} \operatorname{cov}(X_{t'-k} X_{t'-k'}, X_{t-l} X_{t-l'}).
\end{aligned}$$

Under the assumption that $\{X_t\}$ is a zero mean Gaussian stationary process with autocovariance sequence $s_\tau = \operatorname{cov}(X_{t+\tau}, X_t)$, the Isserlis theorem (Isserlis 1918) says that

$$\operatorname{cov}(X_{t'-k} X_{t'-k'}, X_{t-l} X_{t-l'}) = s_{k-l+t-t'} s_{k'-l'+t-t'} + s_{k-l'+t-t'} s_{k'-l+t-t'}, \tag{2.2}$$

so we have

$$\operatorname{cov}(A_{t'}^2, B_t^2) = \sum_k \sum_{k'} \sum_l \sum_{l'} a_k a_{k'} b_l b_{l'} \left(s_{k-l+t-t'} s_{k'-l'+t-t'} + s_{k-l'+t-t'} s_{k'-l+t-t'}\right). \tag{2.3}$$

Now

$$\begin{aligned}
\sum_k \sum_{k'} \sum_l \sum_{l'} a_k a_{k'} b_l b_{l'} s_{k-l+t-t'} s_{k'-l'+t-t'} &= \sum_k \sum_l a_k b_l s_{k-l+t-t'} \sum_{k'} \sum_{l'} a_{k'} b_{l'} s_{k'-l'+t-t'} \\
&= \left(\sum_k \sum_l a_k b_l s_{k-l+t-t'}\right)^2 \\
&= \left(\sum_{v=-(L_j-1)}^{L_j-1} s_{v+t-t'} \sum_{u=0}^{L_j-1-v} a_{u+v} b_u\right)^2,
\end{aligned}$$

where, in the last expression, $a_l = b_l = 0$ for $l < 0$ and $l \geq L_j$ (note that the inner summation above is a cross-correlation and hence can be computed efficiently by two discrete Fourier transforms (DFTs) and one inverse DFT). The part of Equation (2.3) involving $s_{k-l'+t-t'} s_{k'-l+t-t'}$ reduces in a similar manner, yielding

$$\operatorname{cov}(A_{t'}^2, B_t^2) = 2\left(\sum_{v=-(L_j-1)}^{L_j-1} s_{v+t-t'} \sum_{u=0}^{L_j-1-v} a_{u+v} b_u\right)^2.$$

Using this result, we obtain

$$\operatorname{var}\left(\overrightarrow{\nu}^2_u(\tau_j)\right) = \frac{2}{M_j^2} \sum_{t'=L_j-1}^{N-1} \sum_{t=L_j-1}^{N-1} \left(\sum_{v=-(L_j-1)}^{L_j-1} s_{v+t-t'} \sum_{u=0}^{L_j-1-v} \overrightarrow{h}_{j,u+v} \overrightarrow{h}_{j,u} \right)^2$$

and

$$\begin{aligned} \operatorname{var}\left(\overleftrightarrow{\nu}^2_u(\tau_j)\right) \quad = \quad & \frac{1}{2M_j^2} \sum_{t'=L_j-1}^{N-1} \sum_{t=L_j-1}^{N-1} \left[\left(\sum_{v=-(L_j-1)}^{L_j-1} s_{v+t-t'} \sum_{u=0}^{L_j-1-v} \overrightarrow{h}_{j,u+v} \overrightarrow{h}_{j,u} \right)^2 \right. \\ & + \left(\sum_{v=-(L_j-1)}^{L_j-1} s_{v+t-t'} \sum_{u=0}^{L_j-1-v} \overleftarrow{h}_{j,u+v} \overleftarrow{h}_{j,u} \right)^2 \\ & \left. +2 \left(\sum_{v=-(L_j-1)}^{L_j-1} s_{v+t-t'} \sum_{u=0}^{L_j-1-v} \overrightarrow{h}_{j,u+v} \overleftarrow{h}_{j,u} \right)^2 \right]. \end{aligned}$$

Comparison of the variances (and hence the mean squared errors) of these estimators is deferred until Chapter 4, where they are plotted for a range of intrinsically stationary processes in conjunction with estimates that are computed via simulation. For the curious reader, it might suffice to mention that the MSE (i.e., the variance) of the forward-backward estimator is no larger than that of the traditional unbiased estimator for intrinsically stationary processes, with the greatest improvement achieved for sample sizes that are small relative to the length of the wavelet filter used in the variance calculations (i.e., sample sizes that are between one and three times the filter length).

2.2 Forward-Backward Biased Estimator

The alternative biased estimator that we consider here is formed by applying the MODWT to a time series using reflection boundary conditions; i.e., given a time series X_0, X_1, ..., X_{N-1} of length N, we form the series of length $2N$ consisting of

$$X_0, X_1, \ldots, X_{N-2}, X_{N-1}, X_{N-1}, X_{N-2}, \ldots, X_1, X_0 \tag{2.4}$$

and circularly filter it using a forward MODWT wavelet filter $\{\overrightarrow{h}_{j,l}\}$ of width L_j. Let X'_0, X'_1, ..., X'_{2N-1} represent the reflected series, and define

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} X'_{t-l \bmod 2N}, \quad t = 0, 1, \ldots, 2N-1. \tag{2.5}$$

The alternative biased estimator is given by

$$\overleftrightarrow{\nu}^2_b(\tau_j) \equiv \frac{1}{2N} \sum_{t=0}^{2N-1} \widetilde{W}^2_{j,t},$$

which we will refer to as the forward-backward biased estimator of the wavelet variance. This estimator was originally proposed by Greenhall, Howe and Percival (1999) but its statistical properties were explored only in a minimal fashion by limited computer experiments. To obtain expressions for the bias and variance of this estimator, we make use of the following results, which are stated without proof (see Ravishanker and Dey 2002, p. 174):

$$\begin{aligned} \mathrm{E}\left(\boldsymbol{x}^T A \boldsymbol{x}\right) &= \mathrm{tr}\,(A\Sigma) && (2.6)\\ \mathrm{var}\left(\boldsymbol{x}^T A \boldsymbol{x}\right) &= 2\mathrm{tr}\,(A\Sigma)^2 && (2.7)\\ \mathrm{E}\left(\boldsymbol{x}^T A \boldsymbol{x} \cdot \boldsymbol{x}^T B \boldsymbol{x}\right) &= \mathrm{tr}\,(A\Sigma)\,\mathrm{tr}\,(B\Sigma) + 2\mathrm{tr}\,(A\Sigma B\Sigma) && (2.8) \end{aligned}$$

where $\boldsymbol{x} \sim \boldsymbol{N}_N(\boldsymbol{0}, \Sigma)$ and where $\mathrm{tr}\,(M)^2 = \mathrm{tr}\,(MM)$.

Suppose that $h_{j,l}$ is a wavelet filter of width L_j used to compute level j MODWT wavelet coefficients, $\widetilde{\boldsymbol{W}}_j$, for the reflected series $\boldsymbol{X}'$. It is useful to express these coefficients as

$$\widetilde{\boldsymbol{W}}_j \equiv \widetilde{\mathcal{W}}_j \boldsymbol{X}',$$

where

$$\widetilde{\mathcal{W}}_j = \begin{bmatrix} h^{\circ}_{j,0} & h^{\circ}_{j,2N-1} & h^{\circ}_{j,2N-2} & h^{\circ}_{j,2N-3} & \cdots & h^{\circ}_{j,3} & h^{\circ}_{j,2} & h^{\circ}_{j,1} \\ h^{\circ}_{j,1} & h^{\circ}_{j,0} & h^{\circ}_{j,2N-1} & h^{\circ}_{j,2N-2} & \cdots & h^{\circ}_{j,4} & h^{\circ}_{j,3} & h^{\circ}_{j,2} \\ h^{\circ}_{j,2} & h^{\circ}_{j,1} & h^{\circ}_{j,0} & h^{\circ}_{j,2N-1} & \cdots & h^{\circ}_{j,5} & h^{\circ}_{j,4} & h^{\circ}_{j,3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ h^{\circ}_{j,2N-3} & h^{\circ}_{j,2N-4} & h^{\circ}_{j,2N-5} & h^{\circ}_{j,2N-6} & \cdots & h^{\circ}_{j,0} & h^{\circ}_{j,2N-1} & h^{\circ}_{j,2N-2} \\ h^{\circ}_{j,2N-2} & h^{\circ}_{j,2N-3} & h^{\circ}_{j,2N-4} & h^{\circ}_{j,2N-5} & \cdots & h^{\circ}_{j,1} & h^{\circ}_{j,0} & h^{\circ}_{j,2N-1} \\ h^{\circ}_{j,2N-1} & h^{\circ}_{j,2N-2} & h^{\circ}_{j,2N-3} & h^{\circ}_{j,2N-4} & \cdots & h^{\circ}_{j,2} & h^{\circ}_{j,1} & h^{\circ}_{j,0} \end{bmatrix},$$

and where $h^{\circ}_{j,l}$ is the filter $h_{j,l}$ periodized to width $2N$ (see PW, sec. 2.6, for a discussion of periodization). Using this result we can obtain an expression for the expectation of $\overleftrightarrow{\nu}^2_b(\tau_j)$:

$$\begin{aligned} \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) &\equiv \mathrm{E}\left(\frac{1}{2N}\widetilde{\boldsymbol{W}}_j^T\widetilde{\boldsymbol{W}}_j\right) \\ &= \frac{1}{2N}\mathrm{E}\left(\boldsymbol{X}'^T\widetilde{\mathcal{W}}_j^T\widetilde{\mathcal{W}}_j\boldsymbol{X}'\right) \\ &= \frac{1}{2N}\mathrm{tr}\left(\widetilde{\mathcal{W}}_j^T\widetilde{\mathcal{W}}_j\Sigma_{\boldsymbol{X}'}\right), \end{aligned} \tag{2.9}$$

where Equation (2.9) follows by using Equation (2.6) and where $\Sigma_{\boldsymbol{X}'}$ is the covariance matrix of $\boldsymbol{X}'$.

Let us now define $\overrightarrow{\mathcal{W}}_{NB,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ to be the matrices such that

$$\overrightarrow{\boldsymbol{W}}_{NB,j} = \overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X} \quad \text{and} \quad \overleftarrow{\boldsymbol{W}}_{NB,j} = \overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X},$$

where $\overrightarrow{\boldsymbol{W}}_{NB,j}$ and $\overleftarrow{\boldsymbol{W}}_{NB,j}$ are, respectively, vectors of the forward and backward nonboundary wavelet coefficients of $\boldsymbol{X}$, and where the forms of $\overrightarrow{\mathcal{W}}_{NB,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ are equivalent to that of $\widetilde{\mathcal{W}}_j$, with the exception that they are, respectively, comprised of the values of $\overrightarrow{h}_{j,l}$ and $\overleftarrow{h}_{j,l}$, periodized to length N. Appealing to Equation (2.9) we now see that

$$\begin{aligned} \mathrm{bias}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) &= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \nu^2(\tau_j) & (2.10) \\ &= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \mathrm{E}\left(\overleftrightarrow{\nu}^2_u(\tau_j)\right) & (2.11) \\ &= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \frac{1}{2M_j}\mathrm{E}\left(\sum_{t=L_j-1}^{N-1}\overrightarrow{W}^2_{j,t}\right) - \frac{1}{2M_j}\mathrm{E}\left(\sum_{t=L_j-1}^{N-1}\overleftarrow{W}^2_{j,t}\right) \\ &= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \frac{1}{2M_j}\mathrm{E}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right) \end{aligned}$$

$$
\begin{aligned}
& \qquad\qquad\qquad -\frac{1}{2M_j}\mathrm{E}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right) \\
= \;& \mathrm{E}\left(\overleftrightarrow{\nu}_b^2(\tau_j)\right) - \frac{1}{2M_j}\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right) - \frac{1}{2M_j}\mathrm{tr}\left(\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right) \\
= \;& \frac{1}{2N}\mathrm{tr}\left(\widetilde{\mathcal{W}}_j^T\widetilde{\mathcal{W}}_j\Sigma_{\boldsymbol{X}'}\right) - \frac{1}{2M_j}\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right) \\
& \qquad\qquad\qquad -\frac{1}{2M_j}\mathrm{tr}\left(\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right), \qquad (2.12)
\end{aligned}
$$

where Equation (2.11) follows from Equation (2.10) since $\overleftrightarrow{\nu}_u^2(\tau_j)$ is unbiased and where $\Sigma_{\boldsymbol{X}}$ is the covariance matrix of $\boldsymbol{X}$.

In a similar fashion, the variance of $\overleftrightarrow{\nu}_b^2(\tau_j)$ is

$$
\begin{aligned}
\mathrm{var}\left(\overleftrightarrow{\nu}_b^2(\tau_j)\right) &\equiv \mathrm{var}\left(\frac{1}{2N}\widetilde{\boldsymbol{W}}^T\widetilde{\boldsymbol{W}}\right) \\
&= \frac{1}{4N^2}\mathrm{var}\left(\boldsymbol{X}'^T\widetilde{\mathcal{W}}_j^T\widetilde{\mathcal{W}}_j\boldsymbol{X}'\right) \\
&= \frac{1}{2N^2}\mathrm{tr}\left(\widetilde{\mathcal{W}}_j^T\widetilde{\mathcal{W}}_j\Sigma_{\boldsymbol{X}'}\right)^2. \qquad (2.13)
\end{aligned}
$$

While Equations (2.12) and (2.13) can be used to evaluate the statistical properties of the forward-backward biased estimator of the wavelet variance, more efficient methods exist and are developed in Appendix B. As mentioned in the previous section, such evaluations will be considered in Chapter 4, where we will see that the forward-backward biased estimator generally has smaller MSE than the forward-backward unbiased and traditional unbiased estimators for stationary and first order difference stationary processes, with the greatest improvement achieved for sample sizes that are small relative to the length of the wavelet filter.

Chapter 3

ASYMPTOTIC THEORY

In this chapter we will first state some lemmata that will be useful in subsequent proofs. We will then examine the asymptotic properties of the forward-backward wavelet variance estimators under the assumption that the underlying process is stationary and will conclude by extending these results to certain nonstationary cases.

3.1 Mathematical Background

In what follows, we will assume that $\{a_l\}$ is a filter of width $L \geq 2$ such that $\sum_{l=0}^{L-1} a_l = 0$ and $\sum_{l=0}^{L-1} a_l^2 \leq 1$, that U_0, U_1, …, U_{N-1} is a portion of a stationary stochastic process, $\{U_t\}$, of length N such that $U_t \sim \mathrm{N}(\mu_U, \sigma_U^2)\ \forall t$, and that $A_t = \sum_{l=0}^{L-1} a_l U_{t-l \bmod N}$. We note that since

$$\begin{aligned}
\sum_{l=0}^{L-1} a_l(U_{t-l \bmod N} - \mu_U) &= \sum_{l=0}^{L-1} a_l U_{t-l \bmod N} - \mu_U \sum_{l=0}^{L-1} a_l & (3.1)\\
&= \sum_{l=0}^{L_j-1} a_l U_{t-l \bmod N} & (3.2)\\
&= A_t,
\end{aligned}$$

that A_t doesn't depend on μ_U. Thus, we can assume, without loss of generality, that $\mathrm{E}\,(U_t) = 0$.

Lemma 1. $A_t^2 \leq \sum_{l=0}^{L-1} U_{t-l \bmod N}^2$.

Proof. By Cauchy's inequality

$$\begin{aligned} A_t^2 &= \left(\sum_{l=0}^{L-1} a_l U_{t-l \bmod N}\right)^2 \\ &\leq \sum_{l=0}^{L-1} a_l^2 \sum_{l=0}^{L-1} U_{t-l \bmod N}^2 \\ &\leq \sum_{l=0}^{L-1} U_{t-l \bmod N}^2. \end{aligned}$$

$\square$

Lemma 2. $A_t \sim \mathrm{N}(0, \sigma_A^2)$, where $\sigma_A^2 \leq L\sigma_U^2$.

Proof. Since A_t is a linear combination of zero mean Gaussian random variables, its distribution is also Gaussian with mean zero. Further, by Lemma 1

$$\begin{aligned} \sigma_A^2 &= \mathrm{E}\left(A_t^2\right) \\ &\leq \mathrm{E}\left(\sum_{l=0}^{L-1} U_{t-l \bmod N}^2\right) \\ &= \sum_{l=0}^{L-1} \mathrm{E}\left(U_{t-l \bmod N}^2\right) \\ &= L\sigma_U^2. \end{aligned}$$

where $\mathrm{E}\left(A_t^2\right) = \sigma_A^2$ and $\mathrm{E}\left(U_t^2\right) = \sigma_U^2$ because $E(A_t) = \mathrm{E}\left(U_t\right) = 0$. $\square$

Lemma 3. $\mathrm{E}\left(\sum_{t=0}^{L-2} A_t^2\right) \leq (L-1)L\sigma_U^2$.

Proof. By Lemma 2,

$$\begin{aligned} \mathrm{E}\left(\sum_{t=0}^{L-2} A_t^2\right) &= \sum_{t=0}^{L-2} \mathrm{E}\left(A_t^2\right) \\ &= \sum_{l=0}^{L_j-2} \sigma_A^2 \\ &\leq (L-1)L\sigma_U^2 \end{aligned}$$

$\square$

Lemma 4. $\mathrm{var}\left(\sum_{t=0}^{L-2} A_t^2\right) \leq 2(L-1)^2 L^2 \sigma_U^4$.

Proof. Note that $\mathrm{var}\,(A_t^2) = 2\sigma_A^4$ because of the assumption that $A_t \sim \mathrm{N}(0, \sigma_A^2)$. Appealing to Lemma 2,

$$\mathrm{var}\,\left(A_t^2\right) \leq 2L^2\sigma_U^4 \equiv k.$$

It follows that

$$\begin{aligned}
\mathrm{var}\left(\sum_{t=0}^{L-2} A_t^2\right) &= \sum_{t=0}^{L-2} \mathrm{var}\,\left(A_t^2\right) + 2\sum_{t=0}^{L-3}\sum_{t'=t+1}^{L-2} \mathrm{cov}\,\left(A_t^2, A_{t'}^2\right) \\
&\leq \sum_{t=0}^{L-2} \mathrm{var}\,\left(A_t^2\right) + 2\sum_{t=0}^{L-3}\sum_{t'=t+1}^{L-2} \sqrt{\mathrm{var}\,\left(A_t^2\right)\mathrm{var}\,\left(A_{t'}^2\right)} \\
&< \sum_{t=0}^{L-2} k + 2\sum_{t=0}^{L-3}\sum_{t'=t+1}^{L-2} \sqrt{kk} \\
&= (L-1)k + (L-2)(L-1)k \\
&= (L-1)^2 k \\
&= 2(L-1)^2 L^2 \sigma_U^4.
\end{aligned}$$

□

3.2 Stationary Case

We will now show that, if $\{X_t\}$ is a stationary Gaussian stochastic process of length N, the three estimators of wavelet variance have the same asymptotic distribution.

Lemma 5.

$$\sum_{t=0}^{N-1} \overrightarrow{W}_{j,t}^2 = \sum_{t=0}^{N-1} \overleftarrow{W}_{j,t}^2. \tag{3.3}$$

Proof. Let

$$\overrightarrow{H}_j(f) \equiv \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} e^{-i2\pi fl} \text{ and } \overleftarrow{H}_j(f) \equiv \sum_{l=0}^{L_j-1} \overleftarrow{h}_{j,l} e^{-i2\pi fl}$$

be the transfer functions for the forward and backward filters $\{\overrightarrow{h}_{j,l}\}$ and $\{\overleftarrow{h}_{j,l}\}$. Note that

$$\begin{aligned}
\overleftarrow{H}_j(f) = \sum_{l=0}^{L_j-1} \overleftarrow{h}_{j,l} e^{-i2\pi fl} &= \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,L_j-1-l} e^{-i2\pi fl} \\
&= \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} e^{-i2\pi f(L_j-1-l)} \\
&= e^{-i2\pi f(L_j-1)} \left(\sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} e^{-i2\pi fl} \right)^* = e^{-i2\pi f(L_j-1)} \overrightarrow{H}_j^*(f),
\end{aligned}$$

from which it follows that $\overleftarrow{\mathcal{H}}_j(f) = |\overleftarrow{H}_j(f)|^2 = |\overrightarrow{H}_j(f)|^2 = \overrightarrow{\mathcal{H}}_j(f)$. Let $\{\mathcal{X}_k : k = 0, 1, \ldots, N-1\}$ be the DFT of $X_0, X_1, \ldots, X_{N-1}$. Parseval's theorem says that Equation (3.3) is equivalent to the statement

$$\sum_{k=0}^{N-1} |\mathcal{X}_k|^2 \overrightarrow{\mathcal{H}}_j\left(\tfrac{k}{N}\right) = \sum_{k=0}^{N-1} |\mathcal{X}_k|^2 \overleftarrow{\mathcal{H}}_j\left(\tfrac{k}{N}\right),$$

from which the desired result now follows. □

Theorem 6. The asymptotic distributions of $\overrightarrow{\nu}_u^2(\tau_j)$ and $\overleftrightarrow{\nu}_u^2(\tau_j)$ are the same if $\{X_t\}$ is stationary.

Proof. Using Lemma 5, we have

$$\begin{aligned}
\overleftrightarrow{\nu}^2_u(\tau_j) &= \frac{1}{2M_j}\sum_{t=L_j-1}^{N-1}\left(\overrightarrow{W}^2_{j,t}+\overleftarrow{W}^2_{j,t}\right)\\
&= \frac{1}{2M_j}\left(\sum_{t=0}^{N-1}\left(\overrightarrow{W}^2_{j,t}+\overleftarrow{W}^2_{j,t}\right)-\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}+\overleftarrow{W}^2_{j,t}\right)\right)\\
&= \frac{1}{M_j}\sum_{t=0}^{N-1}\overrightarrow{W}^2_{j,t}-\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}+\overleftarrow{W}^2_{j,t}\right)\\
&= \frac{1}{M_j}\sum_{t=L_j-1}^{N-1}\overrightarrow{W}^2_{j,t}+\frac{1}{M_j}\sum_{t=0}^{L_j-2}\overrightarrow{W}^2_{j,t}-\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}+\overleftarrow{W}^2_{j,t}\right)\\
&= \overrightarrow{\nu}^2_u(\tau_j)+\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}-\overleftarrow{W}^2_{j,t}\right).
\end{aligned}$$

It follows from Slutsky's theorem (Ferguson, 1996, p. 39) that $\overrightarrow{\nu}^2_u(\tau_j)$ and $\overleftrightarrow{\nu}^2_u(\tau_j)$ have the same asymptotic distribution if we can show that

$$\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}-\overleftarrow{W}^2_{j,t}\right)\xrightarrow{p}0.$$

Appealing to Equations (1.3) and (2.1) and Lemmata 3 and 4,

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}-\overleftarrow{W}^2_{j,t}\right)\right) &\le \frac{1}{2M_j}\left\{\mathrm{E}\left(\sum_{t=0}^{L_j-2}\overrightarrow{W}^2_{j,t}\right)+\mathrm{E}\left(\sum_{t=0}^{L_j-2}\overleftarrow{W}^2_{j,t}\right)\right\}\\
&< \frac{C_j}{M_j}=\frac{N}{M_j}\frac{C_j}{N},
\end{aligned}\tag{3.4}$$

and

$$\begin{aligned}
\mathrm{var}\left(\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t}-\overleftarrow{W}^2_{j,t}\right)\right) &= \frac{1}{4M_j^2}\left\{\mathrm{var}\left(\sum_{t=0}^{L_j-2}\overrightarrow{W}^2_{j,t}\right)+\mathrm{var}\left(\sum_{t=0}^{L_j-2}\overleftarrow{W}^2_{j,t}\right)\right\}\\
&\quad+\frac{2}{4M_j^2}\mathrm{cov}\left(\sum_{t=0}^{L_j-2}\overrightarrow{W}^2_{j,t},\sum_{t=0}^{L_j-2}\overleftarrow{W}^2_{j,t}\right)\\
&\le \frac{1}{4M_j^2}\left\{2K_j+2\sqrt{\mathrm{var}\left(\sum_{t=0}^{L_j-2}\overrightarrow{W}^2_{j,t}\right)\mathrm{var}\left(\sum_{t=0}^{L_j-2}\overleftarrow{W}^2_{j,t}\right)}\right\}\\
&= \frac{K_j}{M_j^2}=\frac{N^2}{M_j^2}\frac{K_j}{N^2}.
\end{aligned}$$

where $C_j = (L_j - 1)L_j\sigma_X^2$ and $K_j = 2(L_j - 1)^2 L_j^2 \sigma_X^4$. If we define

$$A_N = \frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t} - \overleftarrow{W}^2_{j,t}\right)$$

and $\mu_{A_N} = \mathrm{E}\left[A_N\right]$, by Chebyshev's inequality

$$P\left(|A_N - \mu_{A_N}| \geq \epsilon\right) \leq \frac{\mathrm{var}\left(A_N\right)}{\epsilon^2} \leq \frac{N^2}{M_j^2}\frac{K_j}{N^2\epsilon^2}.$$

Taking the limit of this probability, as $N \to \infty$, we see that $A_N - \mu_{A_N} \xrightarrow{p} 0$ (since $\frac{N}{M_j} = \frac{N}{N-L_j+1} \to 1$ as $N \to \infty$). Further, it follows from Equation (3.4) that $\lim_{N\to\infty} \mu_{A_N} = 0$. Thus, another application of Slutsky's theorem yields

$$\frac{1}{2M_j}\sum_{t=0}^{L_j-2}\left(\overrightarrow{W}^2_{j,t} - \overleftarrow{W}^2_{j,t}\right) = A_N = (A_N - \mu_{A_N}) + \mu_{A_N} \xrightarrow{p} 0,$$

satisfying the necessary condition for $\overleftrightarrow{\nu}^2_u(\tau_j) - \overrightarrow{\nu}^2_u(\tau_j) \xrightarrow{d} 0$. □

Theorem 7. The asymptotic distributions of $\overrightarrow{\nu}^2_u(\tau_j)$ and $\overleftrightarrow{\nu}^2_b(\tau_j)$ are the same if $\{X_t\}$ is stationary.

Proof.

$$\begin{aligned}
\overleftrightarrow{\nu}^2_b(\tau_j) &= \frac{1}{2N}\left(\sum_{t=0}^{2N-1}\widetilde{W}^2_{j,t}\right) \\
&= \frac{1}{2N}\left(\sum_{t=0}^{L_j-2}\widetilde{W}^2_{j,t} + \sum_{t=L_j-1}^{N-1}\widetilde{W}^2_{j,t} + \sum_{t=N}^{N+L_j-2}\widetilde{W}^2_{j,t} + \sum_{t=N+L_j-1}^{2N-1}\widetilde{W}^2_{j,t}\right) \\
&= \frac{1}{2N}\left(\sum_{t=0}^{L_j-2}\widetilde{W}^2_{j,t} + \sum_{t=L_j-1}^{N-1}\overrightarrow{W}^2_{j,t} + \sum_{t=N}^{N+L_j-2}\widetilde{W}^2_{j,t} + \sum_{t=L_j-1}^{N-1}\overleftarrow{W}^2_{j,t}\right) \\
&= \frac{1}{2N}\left(\sum_{t=0}^{L_j-2}\widetilde{W}^2_{j,t} + \sum_{t=N}^{N+L_j-2}\widetilde{W}^2_{j,t} + 2M_j\overleftrightarrow{\nu}^2_u(\tau_j)\right) \\
&= \frac{1}{2N}\sum_{t=0}^{L_j-2}\widetilde{W}^2_{j,t} + \frac{1}{2N}\sum_{t=N}^{N+L_j-2}\widetilde{W}^2_{j,t} + \frac{M_j}{N}\overleftrightarrow{\nu}^2_u(\tau_j) \qquad (3.5)
\end{aligned}$$

Since $\frac{M_j}{N} = \frac{N-L_j+1}{N} \to 1$ as $N \to \infty$, it follows from Slutsky's theorem that $\overleftrightarrow{\nu}^2_b(\tau_j)$ and $\overleftrightarrow{\nu}^2_u(\tau_j)$ have the same asymptotic distribution if

$$\frac{1}{2N}\sum_{t=0}^{L_j-2}\widetilde{W}^2_{j,t} \xrightarrow{p} 0 \quad \text{and} \quad \frac{1}{2N}\sum_{t=N}^{N+L_j-2}\widetilde{W}^2_{j,t} \xrightarrow{p} 0.$$

If we define $B_N = \frac{1}{2N}\sum_{t=0}^{L_j-2}\widetilde{W}_{j,t}^2$ and $\mu_{B_N} = \mathrm{E}[B_N]$, by Chebyshev's inequality, Equation (2.5) and Lemma 4

$$P\left(|B_N - \mu_{B_N}| \geq \epsilon\right) \leq \frac{\mathrm{var}\left(B_N\right)}{\epsilon^2} \leq \frac{K_j}{N^2\epsilon^2},$$

where it is important to note that this result holds since $\sigma^2_{\boldsymbol{X}'} = \sigma^2_{\boldsymbol{X}}$. Taking the limit of this probability, as $N \to \infty$, we see that $B_N - \mu_{B_N} \xrightarrow{p} 0$. Further, from Equation (2.5) and Lemma 3 we see that

$$\mathrm{E}\left(\frac{1}{2N}\sum_{t=0}^{L_j-2}\widetilde{W}_{j,t}^2\right) \leq \frac{C_j}{2N},$$

from which it follows that $\lim_{N\to\infty}\mu_{B_N} = 0$. Thus, Slutsky's theorem yields

$$\frac{1}{2N}\sum_{t=0}^{L_j-2}\widetilde{W}_{j,t}^2 = B_N = \left(B_N - \mu_{B_N}\right) + \mu_{B_N} \xrightarrow{p} 0. \tag{3.6}$$

By a mere alteration of the indices above, it is simple to show

$$\frac{1}{2N}\sum_{t=N}^{N+L_j-2}\widetilde{W}_{j,t}^2 \xrightarrow{p} 0. \tag{3.7}$$

From Equations (3.6) and (3.7), we see that the requisite conditions have been satisfied for $\overleftrightarrow{\nu}_b^2(\tau_j) - \overleftrightarrow{\nu}_u^2(\tau_j) \xrightarrow{d} 0$. Combining this result with the fact that $\overleftrightarrow{\nu}_u^2(\tau_j)$ and $\overrightarrow{\nu}_u^2(\tau_j)$ have the same asymptotic distribution (Theorem 6), we see that $\overleftrightarrow{\nu}_b^2(\tau_j)$ and $\overrightarrow{\nu}_u^2(\tau_j)$ also have the same asymptotic distribution. $\square$

3.3 Extensions to Nonstationary Cases

It is now appropriate to extend the results of the previous section to certain nonstationary cases. Accordingly, we will define the first order and second order backward differences of $\{X_t\}$ to be

$$Y_t = X_t - X_{t-1}$$

and

$$Z_t = Y_t - Y_{t-1} = X_t - 2X_{t-1} + X_{t-2},$$

respectively. Similarly, let Y_t' and Z_t' be the corresponding differences for X_t', as defined in Equation (2.4). Given a portion, $X_0, X_1, \ldots, X_{N-1}$, of $\{X_t\}$, assuming circular boundary conditions yields

$$\begin{aligned} Y_0, Y_1, \ldots, Y_{N-1} &= X_0 - X_{N-1}, X_1 - X_0, \ldots, X_{N-1} - X_{N-2}, \\ Z_0, Z_1, \ldots, Z_{N-1} &= X_0 - 2X_{N-1} + X_{N-2}, X_1 - 2X_0 + X_{N-1}, \ldots, \\ &\quad X_{N-1} - 2X_{N-2} + X_{N-3}, \end{aligned}$$

$$\begin{aligned} &Y_0', Y_1', \ldots, Y_{N-1}', Y_N', Y_{N+1}', \ldots, Y_{2N-1}' \\ &\quad = X_0' - X_{2N-1}', X_1' - X_0', \ldots, X_{N-1}' - X_{N-2}', X_N' - X_{N-1}', \\ &\qquad X_{N+1}' - X_N', \ldots, X_{2N-1}' - X_{2N-2}' \\ &\quad = 0, X_1 - X_0, \ldots, X_{N-1} - X_{N-2}, 0, X_{N-2} - X_{N-1}, \ldots, X_0 - X_1 \\ &\quad = 0, Y_1, \ldots, Y_{N-1}, 0, -Y_{N-1}, \ldots, -Y_1 \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} &Z_0', Z_1', Z_2', \ldots, Z_{N-1}', Z_N', Z_{N+1}', Z_{N+2}', \ldots, Z_{2N-1}' \\ &\quad = X_0' - 2X_{2N-1}' + X_{2N-2}', X_1' - 2X_0' + X_{2N-1}', X_2' - 2X_1' + X_0', \ldots, \\ &\qquad X_{N-1}' - 2X_{N-2}' + X_{N-3}', X_N' - 2X_{N-1}' + X_{N-2}', X_{N+1}' - 2X_N' + X_{N-1}', \\ &\qquad X_{N+2}' - 2X_{N+1}' + X_N', \ldots, X_{2N-1}' - 2X_{2N-2}' + X_{2N-3}' \\ &\quad = Y_1, Y_1, Z_2, \ldots, Z_{N-1}, -Y_{N-1}, -Y_{N-1}, Z_{N-1}, \ldots, Z_2. \end{aligned} \tag{3.9}$$

Lemma 8. Define $\overrightarrow{c}_{j,l} = \sum_{k=0}^{l} \overrightarrow{h}_{j,k}$ and $\overleftarrow{c}_{j,l} = \sum_{k=0}^{l} \overleftarrow{h}_{j,k}$ for $l \geq 0$ and $\overrightarrow{c}_{j,l} = \overleftarrow{c}_{j,l} = 0$ for $l < 0$. Then

(a) $\overrightarrow{c}_{j,l}$ and $\overleftarrow{c}_{j,l}$ are filters of width $L_j - 1$

(b) $\overrightarrow{W}_{j,t} = \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} Y_{t-l \bmod N}$

(c) $\overleftarrow{W}_{j,t} = \sum_{l=0}^{L_j-2} \overleftarrow{c}_{j,l} Y_{t-l \bmod N}$

(d) $\widetilde{W}_{j,t} = \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} Y'_{t-l \bmod 2N}$

Proof.

(a) $\overrightarrow{c}_{j,L_j-1} = \sum_{k=0}^{L_j-1} \overrightarrow{h}_{j,k} = 0 = \sum_{k=0}^{L_j-1} \overleftarrow{h}_{j,k} = \overleftarrow{c}_{j,L_j-1}$, $\overrightarrow{c}_{j,L_j-2} = \sum_{k=0}^{L_j-2} \overrightarrow{h}_{j,k} = -\overrightarrow{h}_{j,L_j-1} \neq 0 \neq -\overleftarrow{h}_{j,L_j-1} = \sum_{k=0}^{L_j-2} \overleftarrow{h}_{j,k} = \overleftarrow{c}_{j,L_j-2}$, and $\overrightarrow{c}_{j,0} = \overrightarrow{h}_{j,0} \neq 0 \neq \overleftarrow{h}_{j,0} = \overleftarrow{c}_{j,0}$. Thus, $\overrightarrow{c}_{j,l}$ and $\overleftarrow{c}_{j,l}$ are infinite sequences where $\overrightarrow{c}_{j,l} = \overleftarrow{c}_{j,l} = 0$ for $l < 0$ and $l \geq L_j - 1$ and $\overrightarrow{c}_{j,l} \neq 0 \neq \overleftarrow{c}_{j,l}$ for $l = 0$ and $l = L_j - 2$, which shows that $\overrightarrow{c}_{j,l}$ and $\overleftarrow{c}_{j,l}$ are filters of width $L_j - 1$.

(b)

$$
\begin{aligned}
\overrightarrow{W}_{j,t} &= \sum_{l=0}^{L_j-1} \overrightarrow{h}_{j,l} X_{t-l \bmod N} = \sum_{l=0}^{L_j-1} \left(\overrightarrow{c}_{j,l} - \overrightarrow{c}_{j,l-1}\right) X_{t-l \bmod N} \\
&= \sum_{l=0}^{L_j-1} \overrightarrow{c}_{j,l} X_{t-l \bmod N} - \sum_{l=0}^{L_j-1} \overrightarrow{c}_{j,l-1} X_{t-l \bmod N} \\
&= \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} X_{t-l \bmod N} - \sum_{l=-1}^{L_j-2} \overrightarrow{c}_{j,l} X_{t-l-1 \bmod N} \\
&= \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} \left(X_{t-l \bmod N} - X_{t-l-1 \bmod N}\right) \\
&= \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} Y_{t-l \bmod N},
\end{aligned}
$$

where use is made of the fact that $\overrightarrow{c}_{j,l} = 0$ for $l < 0$ and $l \geq L_j - 1$.

Parts (c) and (d) are proven in a manner similar to part (b), with appropriate substitutions of $\overleftarrow{c}_{j,l}$ and Y_t'. $\square$

Lemma 9. Given $\overrightarrow{c}_{j,l} = \sum_{k=0}^{l} \overrightarrow{h}_{j,k}$ and $\overleftarrow{c}_{j,l} = \sum_{k=0}^{l} \overleftarrow{h}_{j,k}$, if the width of $\overrightarrow{h}_{1,l}$ and $\overleftarrow{h}_{1,l}$, L_1, is greater than 2 (i.e. at least 4),

(a) $\sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} = \sum_{l=0}^{L_j-2} \overleftarrow{c}_{j,l} = 0$

(b) $\sum_{l=0}^{L_j-2} \overrightarrow{c}^{\,2}_{j,l} = \sum_{l=0}^{L_j-2} \overleftarrow{c}^{\,2}_{j,l} = c_{j,L_1}$

where c_{j,L_1} is a finite constant, dependent on choice of Daubechies filter and level.

Proof.

(a) Let d_l denote the backward difference filter, $\{d_0 = 1, d_1 = -1\}$. We will consider d_l and $\overrightarrow{h}_{j,l}$ infinite sequences (i.e. $d_l = 0$ for $l < 0$ and $l \geq 2$, and $\overrightarrow{h}_{j,l} = 0$ for $l < 0$ and $l \geq L_j$). Then

$$\overrightarrow{h}_{j,l} = \overrightarrow{c}_j * d_l = \sum_{k=0}^{L_j-2} d_k \overrightarrow{c}_{j,l-k} \tag{3.10}$$

where the $*$ operator denotes convolution. If we let

$$\overrightarrow{\mathcal{C}}_j(f) = \left| \sum_{l=0}^{Lj-2} \overrightarrow{c}_{j,l} e^{-i2\pi fl} \right|^2 \qquad \text{and} \qquad \mathcal{D}(f) = \left| \sum_{l=0}^{1} d_l e^{-i2\pi fl} \right|^2$$

be the squared gain functions for $\overrightarrow{c}_{j,l}$ and d_l (recalling a similar definition for $\overrightarrow{\mathcal{H}}_j(f)$ in the previous section), Equation (3.10) implies that

$$\overrightarrow{\mathcal{H}}_j(f) = \overrightarrow{\mathcal{C}}_j(f)\mathcal{D}(f). \tag{3.11}$$

Thus,

$$\begin{aligned}
\left| \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} e^{-i2\pi fl} \right|^2 &= \overrightarrow{\mathcal{C}}_j(f) = \overrightarrow{\mathcal{H}}_j(f)/\mathcal{D}(f) \\
&= \mathcal{D}^{\frac{L_1}{2}}(f)\overrightarrow{\mathcal{M}}j(f)/\mathcal{D}(f) \\
&= \mathcal{D}^{\frac{L_1}{2}-1}(f)\overrightarrow{\mathcal{M}}j(f) \\
&= 4\sin^{L_1-2}(\pi f)\overrightarrow{\mathcal{M}}j(f)
\end{aligned}$$

where

$$
\begin{aligned}
\overrightarrow{\mathcal{M}}j(f) &= \left[\prod_{k=0}^{j-2} 4\cos^2(\pi 2^k f)\right]^{\frac{L_1}{2}} \widetilde{\mathcal{A}}_{L_1}(2^{j-1}f)\widetilde{\mathcal{G}}_{j-1}^{(D)}(f) \\
\widetilde{\mathcal{A}}_{L_1}(f) &= \frac{1}{2_1^L}\sum_{l=0}^{\frac{L_1}{2}-1}\binom{\frac{L_1}{2}-1+l}{l}\cos^{2l}(\pi f) \\
\widetilde{\mathcal{G}}_j(f) &= \prod_{l=0}^{j-1}\cos_1^L(\pi 2^l f)\sum_{l=0}^{\frac{L_1}{2}-1}\binom{\frac{L_1}{2}-1+l}{l}\cos^{2l}(\pi f) \\
\mathcal{D}(f) &= 4\sin^2(\pi f)
\end{aligned}
$$

(see PW, sec. 4.8 and pp. 163, 535–6 for details). Thus,

$$
\left|\sum_{l=0}^{L_j-2}\overrightarrow{c}_{j,l}\right|^2 = \overrightarrow{\mathcal{C}}_j(0) = 0
$$

$$
\Rightarrow \sum_{l=0}^{L_j-2}\overrightarrow{c}_{j,l} = 0
$$

(b) Since $\overrightarrow{c}_{j,l}$ is comprised of a finite number of nonzero, finite values, $\sum_{l-0}^{L_j-2}\overrightarrow{c}_{j,l}^{\,2}$ is a finite constant, dependent on Daubechies filter and level, and can be denoted c_{j,L_1}.

The proof for $\overleftarrow{c}_{j,l}$ is similar, where the proof of part (a) is simplified by the fact that $\overrightarrow{\mathcal{C}}_j(f) = \mathcal{D}(f)\overrightarrow{\mathcal{H}}_j(f) = \mathcal{D}(f)\overleftarrow{\mathcal{H}}_j(f) = \overleftarrow{\mathcal{C}}_j(f)$, since $\overrightarrow{\mathcal{H}}_j(f) = \overleftarrow{\mathcal{H}}_j(f)$ (as shown in the proof of Lemma 3.3). □

Theorem 10. The asymptotic distributions of $\overrightarrow{\nu}_u^2(\tau_j)$ and $\overleftrightarrow{\nu}_u^2(\tau_j)$ are the same if $\{X_t\}$ is a stochastic process whose first backward differences, $\{Y_t\}$, are stationary and Gaussian.

Proof. We first note that by Lemmata 8, 3 and 4,

$$
\mathrm{E}\left(\sum_{l=0}^{L_j-3}\overrightarrow{W}_{j,t}^2\right) = \mathrm{E}\left(\sum_{l=0}^{L_j-3}\overleftarrow{W}_{j,t}^2\right) \le (L_j-2)(L_j-1)\sigma_Y^2
$$

and

$$
\mathrm{var}\left(\sum_{l=0}^{L_j-3}\overrightarrow{W}_{j,t}^2\right) = \mathrm{var}\left(\sum_{l=0}^{L_j-3}\overleftarrow{W}_{j,t}^2\right) \le 2(L_j-2)^2(L_j-1)^2\sigma_Y^2,
$$

where $\sigma_Y^2 = \operatorname{var}(Y_t)$. Thus, the remainder of the proof is equivalent to that of Theorem 6 if every instance of L_j is decremented by 1. $\square$

Theorem 11. The asymptotic distributions of $\overrightarrow{\nu}_u^2(\tau_j)$ and $\overleftrightarrow{\nu}_b^2(\tau_j)$ are the same if $\{X_t\}$ is a stochastic process whose first backward differences, $\{Y_t\}$, are stationary and Gaussian.

Proof. Again, by appealing to Lemmata 8, 3 and 4,

$$\mathrm{E}\left(\sum_{l=0}^{L_j-3} \widetilde{W}_{j,t}^2\right) = \mathrm{E}\left(\sum_{l=N}^{N+L_j-3} \widetilde{W}_{j,t}^2\right) \leq (L_j-2)(L_j-1)\sigma_Y^2$$

and

$$\operatorname{var}\left(\sum_{l=0}^{L_j-3} \widetilde{W}_{j,t}^2\right) = \operatorname{var}\left(\sum_{l=N}^{N+L_j-3} \widetilde{W}_{j,t}^2\right) \leq 2(L_j-2)^2(L_j-1)^2\sigma_Y^4.$$

Thus, the remainder of the proof is equivalent to that of Theorem 7 if every instance of L_j is decremented by 1. $\square$

Lemma 12. Define $\overrightarrow{cc}_{j,l} = \sum_{k=0}^{l} \overrightarrow{c}_{j,k} = \sum_{k=0}^{l}\sum_{m=0}^{k} \overrightarrow{h}_{j,m}$ and $\overleftarrow{cc}_{j,l} = \sum_{k=0}^{l} \overleftarrow{c}_{j,k} = \sum_{k=0}^{l}\sum_{m=0}^{k} \overleftarrow{h}_{j,m}$ for $l \geq 0$ and $\overrightarrow{cc}_{j,l} = \overleftarrow{cc}_{j,l} = 0$ for $l < 0$. Then

(a) $\overrightarrow{cc}_{j,l}$ and $\overleftarrow{cc}_{j,l}$ are filters of width $L_j - 2$

(b) $\overrightarrow{W}_{j,t} = \sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} Z_{t-l \bmod N}$

(c) $\overleftarrow{W}_{j,t} = \sum_{l=0}^{L_j-3} \overleftarrow{cc}_{j,l} Z_{t-l \bmod N}$

(d) $\widetilde{W}_{j,t} = \sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} Z'_{t-l \bmod 2N}$

Proof.

(a) By Lemma 9a, $\overrightarrow{cc}_{j,L_j-2} = \sum_{k=0}^{L_j-2} \overrightarrow{c}_{j,k} = 0 = \sum_{k=0}^{L_j-2} \overleftarrow{c}_{j,k} = \overleftarrow{cc}_{j,L_j-2}$, $\overrightarrow{cc}_{j,L_j-3} = \sum_{k=0}^{L_j-3} \overrightarrow{c}_{j,k} = -\overrightarrow{c}_{j,L_j-2} \neq 0 \neq -\overleftarrow{c}_{j,L_j-2} = \sum_{k=0}^{L_j-3} \overleftarrow{c}_{j,k} = \overleftarrow{cc}_{j,L_j-3}$, and $\overrightarrow{cc}_{j,0} = \overrightarrow{c}_{j,0} \neq 0 \neq \overleftarrow{c}_{j,0} = \overleftarrow{cc}_{j,0}$. Thus, $\overrightarrow{cc}_{j,l}$ and $\overleftarrow{cc}_{j,l}$ are infinite sequences where $\overrightarrow{cc}_{j,l} = \overleftarrow{cc}_{j,l} = 0$ for $l < 0$ and $l \geq L_j - 2$ and $\overrightarrow{cc}_{j,l} \neq 0 \neq \overleftarrow{cc}_{j,l}$ for $l = 0$ and $l = L_j - 3$, which shows that they are filters of width $L_j - 2$.

(b)

$$\begin{aligned}
\overrightarrow{W}_{j,t} &= \sum_{l=0}^{L_j-2} \overrightarrow{c}_{j,l} Y_{t-l \bmod N} = \sum_{l=0}^{L_j-2} \left(\overrightarrow{cc}_{j,l} - \overrightarrow{cc}_{j,l-1}\right) Y_{t-l \bmod N} \\
&= \sum_{l=0}^{L_j-2} \overrightarrow{cc}_{j,l} Y_{t-l \bmod N} - \sum_{l=0}^{L_j-2} \overrightarrow{cc}_{j,l-1} Y_{t-l \bmod N} \\
&= \sum_{l=0}^{L_j-2} \overrightarrow{cc}_{j,l} Y_{t-l \bmod N} - \sum_{l=-1}^{L_j-3} \overrightarrow{cc}_{j,l} Y_{t-l-1 \bmod N} \\
&= \sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} \left(Y_{t-l \bmod N} - Y_{t-l-1 \bmod N}\right) \\
&= \sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} Z_{t-l \bmod N},
\end{aligned}$$

where use is made of the fact that $\overrightarrow{cc}_{j,l} = 0$ for $l < 0$ and $l \geq L_j - 2$.

Parts (c) and (d) are proven in a manner similar to part (b), with appropriate substitutions of $\overleftarrow{cc}_{j,l}$ and Z'_t. □

Lemma 13. Given $\overrightarrow{cc}_{j,l} = \sum_{k=0}^{l} \sum_{m=0}^{k} \overrightarrow{h}_{j,m}$ and $\overleftarrow{cc}_{j,l} = \sum_{k=0}^{l} \sum_{m=0}^{k} \overleftarrow{h}_{j,k}$, if $L_1 \geq 6$, then

(a) $\sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} = \sum_{l=0}^{L_j-3} \overleftarrow{cc}_{j,l} = 0$

(b) $\sum_{l=0}^{L_j-3} \overrightarrow{cc}^2_{j,l} = \sum_{l=0}^{L_j-3} \overleftarrow{cc}^2_{j,l} = cc_{j,L_1}$

where cc_{j,L_1} is a finite constant, dependent on choice of Daubechies filter and level.

Proof.

(a) As before, if we let d_l denote the backward difference filter, $\{d_0 = 1, d_1 = -1\}$, then

$$\overrightarrow{c}_{j,l} = \overrightarrow{cc} * d_l = \sum_{k=0}^{L_j-3} d_k \overrightarrow{cc}_{l-k} \tag{3.12}$$

where the $*$ operator denotes convolution. If we let

$$\overrightarrow{\mathcal{CC}}_j(f) = \left| \sum_{l=0}^{Lj-3} \overrightarrow{cc}_{j,l} e^{-i2\pi fl} \right|^2 \qquad \text{and} \qquad \mathcal{D}(f) = \left| \sum_{l=0}^{1} d_l e^{-i2\pi fl} \right|^2$$

be the squared functions for $\overrightarrow{cc}_{j,l}$ and d_l (recalling a similar definition for $\overrightarrow{\mathcal{C}}_j(f)$ above), Equation (3.12) can be expressed as

$$\overrightarrow{\mathcal{C}}_j(f) = \overrightarrow{\mathcal{CC}}_j(f)\mathcal{D}(f)$$

Referring to Equation (3.11), we see

$$\begin{aligned}\overrightarrow{\mathcal{H}}_j(f) &= \overrightarrow{\mathcal{C}}_j(f)\mathcal{D}(f) \\ &= \overrightarrow{\mathcal{CC}}_j(f)\mathcal{D}(f)\mathcal{D}(f)\end{aligned}$$

Thus,

$$\begin{aligned}\left|\sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} e^{-i2\pi fl}\right|^2 &= \overrightarrow{\mathcal{CC}}_j(f) = \overrightarrow{\mathcal{H}}_j(f)/\mathcal{D}^2(f) \\ &= \mathcal{D}^{\frac{L_1}{2}}(f)\overrightarrow{\mathcal{M}}j(f)/\mathcal{D}^2(f) \\ &= \mathcal{D}^{\frac{L_1}{2}-2}(f)\overrightarrow{\mathcal{M}}j(f) \\ &= 4\sin^{L_1-4}(\pi f)\overrightarrow{\mathcal{M}}j(f)\end{aligned}$$

where $\overrightarrow{\mathcal{M}}j(f)$ is defined in the proof of Lemma 9a. Thus,

$$\left|\sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l}\right|^2 = \overrightarrow{\mathcal{CC}}_j(0) = 0$$

$$\Rightarrow \sum_{l=0}^{L_j-3} \overrightarrow{cc}_{j,l} = 0$$

(b) Since $\overrightarrow{cc}_{j,l}$ is comprised of a finite number of nonzero, finite values, $\sum_{l-0}^{L_j-3} \overrightarrow{cc}^2_{j,l}$ is a finite constant, dependent on Daubechies filter and level, and can be denoted cc_{j,L_1}.

The proof for $\overleftarrow{cc}_{j,l}$ is similar, where the proof of part (a) is simplified by the fact that $\overrightarrow{\mathcal{CC}}_j(f) = \mathcal{D}^2(f)\overrightarrow{\mathcal{H}}_j(f) = \mathcal{D}^2(f)\overleftarrow{\mathcal{H}}_j(f) = \overleftarrow{\mathcal{CC}}_j(f)$, since $\overrightarrow{\mathcal{H}}_j(f) = \overleftarrow{\mathcal{H}}_j(f)$ (as shown in the proof of Lemma 5). □

Theorem 14. The asymptotic distributions of $\overrightarrow{\nu}^2_u(\tau_j)$ and $\overleftrightarrow{\nu}^2_u(\tau_j)$ are the same if $\{X_t\}$ is a stochastic process whose second backward differences, $\{Z_t\}$, are stationary and Gaussian.

Proof. We first note that by Lemmata 12, 3 and 4,

$$\mathrm{E}\left(\sum_{l=0}^{L_j-4}\overrightarrow{W}_{j,t}^2\right)=\mathrm{E}\left(\sum_{l=0}^{L_j-4}\overleftarrow{W}_{j,t}^2\right)\le (L_j-3)(L_j-2)\sigma_Z^2$$

and

$$\mathrm{var}\left(\sum_{l=0}^{L_j-4}\overrightarrow{W}_{j,t}^2\right)=\mathrm{var}\left(\sum_{l=0}^{L_j-4}\overleftarrow{W}_{j,t}^2\right)\le 2(L_j-3)^2(L_j-2)^2\sigma_Z^2.$$

where $\sigma_Z^2=\mathrm{var}\,(Z_t)$. Thus, the remainder of the proof is equivalent to that of Theorem 6 if every instance of L_j is decremented by 2. $\square$

It would be desireable to obtain a similar result for the asymptotic distribution of $\overleftrightarrow{\nu}_b^2(\tau_j)$, that is, to show that $\overleftrightarrow{\nu}_b^2(\tau_j)-\overrightarrow{\nu}_u^2(\tau_j)\xrightarrow{d}0$. Unfortunately, in the case of stationary second differences, the values Z_0', Z_1',Z_N' and Z_{N+1}' are not members of a stationary process (see Equation (3.9)) and thus there is no guarantee that $\mathrm{E}\,\left(Z'^2_{t-l \bmod N}\right)$ or σ_Z^2 are finite. Hence we are unable to obtain upper bounds for

$$\mathrm{E}\left(\sum_{l=0}^{L_j-4}\overrightarrow{W}_{j,t}^2\right),\quad \mathrm{E}\left(\sum_{l=N}^{N+L_j-4}\overrightarrow{W}_{j,t}^2\right),\quad \mathrm{var}\left(\sum_{l=0}^{L_j-4}\overrightarrow{W}_{j,t}^2\right)\quad\text{and}\quad\mathrm{var}\left(\sum_{l=N}^{N+L_j-4}\overrightarrow{W}_{j,t}^2\right)$$

which prevents us from concluding that

$$\frac{1}{2N}\sum_{t=0}^{L_j-4}\widetilde{W}_{j,t}^2\xrightarrow{p}0\qquad\text{and}\qquad\frac{1}{2N}\sum_{t=N}^{N+L_j-4}\widetilde{W}_{j,t}^2\xrightarrow{p}0.$$

Thus, there is no guarantee that $\overleftrightarrow{\nu}_b^2(\tau_j)-\overrightarrow{\nu}_u^2(\tau_j)\xrightarrow{d}0$.

As an example, let $\{X_t\}$ be a random run process defined by:

$$X_t\equiv\begin{cases}\sum_{u=0}^{t}\sum_{u'=0}^{u}\epsilon_{u'}, & \text{for } t\ge 0,\\ 0, & \text{for } t=-1,-2,\\ \sum_{u=0}^{|t|-3}\sum_{u'=0}^{u}\epsilon_{-u'-1}, & \text{for } t\le 3,\end{cases}$$

where ϵ_t is a white noise process with mean zero and variance σ_ϵ^2. It is straightforward to show that the first order backward difference process shifted by one time unit, $W_t=X_{t+1}-X_t$, is a random walk process:

$$W_t\equiv\begin{cases}\sum_{u=1}^{t}\epsilon_u', & \text{for } t\ge 1,\\ 0, & \text{for } t=0,\\ -\sum_{u=0}^{|t|-1}\epsilon_{-u-1}', & \text{for } t\le -1,\end{cases}$$

where $\epsilon'_t = \epsilon_{t-1}$ is a white noise process with mean zero and variance σ^2_ϵ. Since $Y_t = X_t - X_{t-1} = W_{t-1}$, Y_t is also a random walk process. It follows that $\mathrm{E}\,(Y_t) = \sum_{u=1}^{t} \mathrm{E}\,(\epsilon'_u) = 0$ for $t > 0$. Further, for $t > 0$

$$\begin{aligned} \mathrm{E}\,(Y_t Y_{t+m}) &= \mathrm{E}\left(\left(\sum_{u=1}^{t} \epsilon'_u\right)\left(\sum_{k=0}^{t+m} \epsilon'_k\right)\right) \\ &= \sum_{u=0}^{t}\sum_{k=0}^{t+m} \mathrm{E}\,(\epsilon'_u \epsilon'_k) \end{aligned}$$

from which we see that $\mathrm{var}\,(Y_t) = \mathrm{E}\,(Y_t^2) = t\sigma^2_\epsilon$ and $\mathrm{cov}\,(Y_t, Y_{t+m}) = \mathrm{E}\,(Y_t Y_{t+m}) = t\sigma^2_\epsilon$, since $\mathrm{E}\,(\epsilon'_u \epsilon'_k)$ is nonzero only when $u = k$, which occurs t times. If $\overleftrightarrow{\nu}^2_b(\tau_j)$ is computed with the unit level D6 wavelet filter, then Equation (3.5) reduces to

$$\overleftrightarrow{\nu}_b(\tau_1) = \frac{1}{2N}\sum_{t=0}^{4} \widetilde{W}^2_{1,t} + \frac{1}{2N}\sum_{t=N}^{N+4} \widetilde{W}^2_{1,t} + \frac{N-5}{N}\overleftrightarrow{\nu}_u(\tau_1),$$

since $L_1 = 6$ and $M_1 = N - L_1 + 1 = N - 5$. By Lemma 8,

$$\sum_{t=0}^{2} \widetilde{W}^2_{1,t} = \sum_{t=0}^{4}\left(\sum_{l=0}^{5} \overrightarrow{h}^{(D6)}_{1,l} X'_{t-l \bmod N}\right)^2 = \sum_{t=0}^{4}\left(\sum_{l=0}^{4} \overrightarrow{c}^{(D6)}_{1,l} Y'_{t-l \bmod N}\right)^2$$

and

$$\sum_{t=N}^{N+4} \widetilde{W}^2_{1,t} = \sum_{t=N}^{N+4}\left(\sum_{l=0}^{5} \overrightarrow{h}^{(D6)}_{1,l} X'_{t-l \bmod N}\right)^2 = \sum_{t=N}^{N+4}\left(\sum_{l=0}^{4} \overrightarrow{c}^{(D6)}_{1,l} Y'_{t-l \bmod N}\right)^2$$

where $\overrightarrow{c}^{(D6)}_{1,l} = \sum_{k=0}^{l} \overrightarrow{h}^{(D6)}_{1,k}$. Thus, the expected difference between $\overleftrightarrow{\nu}_b(\tau_1)$ and $\overleftrightarrow{\nu}_u(\tau_1)$ is

$$\begin{aligned} &\frac{1}{2N}\sum_{t=0}^{4} \mathrm{E}\left(\widetilde{W}^2_{1,t}\right) + \frac{1}{2N}\sum_{t=N}^{N+4} \mathrm{E}\left(\widetilde{W}^2_{1,t}\right) - \frac{5}{N}\nu^2(\tau_1) \\ &\quad \geq \frac{1}{2N}\mathrm{E}\left(\widetilde{W}^2_{1,N+2}\right) - \frac{5}{N}\nu(\tau_1) \\ &\quad = \frac{\mathrm{E}\left(\left(\overrightarrow{c}^{(D6)}_{1,0} Y'_{N+2} + \overrightarrow{c}^{(D6)}_{1,1} Y'_{N+1} + \overrightarrow{c}^{(D6)}_{1,2} Y'_{N} + \overrightarrow{c}^{(D6)}_{1,3} Y'_{N-1} + \overrightarrow{c}^{(D6)}_{1,4} Y'_{N-2}\right)^2\right)}{2N} - \frac{5}{N}\nu(\tau_1) \\ &\quad = \frac{\mathrm{E}\left(\left((\overrightarrow{c}^{(D6)}_{1,4} - \overrightarrow{c}^{(D6)}_{1,0})Y_{N-2} - (\overrightarrow{c}^{(D6)}_{1,1} - \overrightarrow{c}^{(D6)}_{1,3})Y_{N-1}\right)^2\right)}{2N} - \frac{5}{N}\nu(\tau_1) \\ &\quad = \frac{c_1^2 \mathrm{E}\,(Y^2_{N-1})}{2N} + \frac{c_2^2 \mathrm{E}\,(Y^2_{N-2})}{2N} - 2\frac{c_1 c_2 \mathrm{E}\,(Y_{N-1}Y_{N-2})}{2N} - \frac{5}{N}\nu(\tau_1) \\ &\quad = \frac{c_1^2 \sigma^2_\epsilon (N-1)}{2N} + \frac{c_2^2 \sigma^2_\epsilon (N-2)}{2N} - 2\frac{c_1 c_2 \sigma^2_\epsilon (N-2)}{2N} - \frac{5}{N}\nu(\tau_1) \end{aligned}$$

$$\begin{aligned}
&\to \frac{1}{2}(c_1^2 - 2c_1c_2 + c_2^2)\sigma_\epsilon^2 \qquad\qquad (\text{as } N \to \infty)\\
&= \frac{1}{2}(c_1 - c_2)^2\sigma_\epsilon^2\\
&= 0.022\sigma_\epsilon^2 \neq 0,
\end{aligned}$$

where $c_1 = \overrightarrow{c}_{1,1}^{(D6)} - \overrightarrow{c}_{1,3}^{(D6)} = 0.421$, $c_2 = \overrightarrow{c}_{1,4}^{(D6)} - \overrightarrow{c}_{1,0}^{(D6)} = 0.210$ and $\sigma_\epsilon^2 > 0$, by assumption. Thus we see that the difference between $\overleftrightarrow{\nu}_b^2(\tau_j)$ and $\overleftrightarrow{\nu}_u^2(\tau_j)$ need not converge to zero as $N \to \infty$ when $\{X_t\}$ is a nonstationary process with second order stationary backward differences.

Chapter 4

MONTE CARLO EXPERIMENTS

As shown in the previous chapter, the forward-backward unbiased wavelet variance estimator has the same asymptotic distribution as the traditional unbiased estimator, with the same holding true for the forward-backward biased estimator when the time series of interest is either stationary or first order difference stationary. We now determine the statistical properties of these new estimators for small sample sizes. To achieve this purpose, we simulated stochastic processes, X_t, such that

$$(1-B)^{\delta} X_t = \epsilon_t,$$

where

$$(1-B)^{\delta} = \sum_{k=0}^{\infty} \binom{\delta}{k} (-1)^k B^k \text{ and } \epsilon_t \sim \mathrm{N}(0, \sigma^2).$$

Processes, X_t, that satisfy the conditions above are known as fractionally differenced processes with parameter δ (simply referred to as FD(δ) processes). The rationale for using such realizations in our analysis is that they are representative of a wide variety of processes that arise in practice (see Table 4.1).

In our simulation study, we generated FD processes (via the Davies-Harte method – see, for example, PW 2000, sec. 7.8) of varying lengths for a range of δ values and evaluated each of the three wavelet variance estimators using six different filters: levels 1, 2 and 3 of the D4 and LA8 filters (Daubechies 1992, p.195). In particular, we let δ take the values $-\frac{1}{2}$, $-\frac{1}{4}$, 0, $\frac{1}{4}$, $\frac{9}{20}$, $\frac{1}{2}$, $\frac{5}{6}$ 1, 1.5, and 2. For each δ we simulated 1000 realizations of length 350, thus allowing the length of each realization to be seven times greater than width of the longest filter used in our analysis, the level 3 LA8 wavelet filter. We then specified a filter (one of the six mentioned above) and a series length $N \leq 350$ (repeating the analysis for each N from 1 to 7 times the length of the LA8 wavelet filter with level corresponding to the level of the chosen filter), and computed the three wavelet variance estimates using the first

Table 4.1: FD processes for various values of parameter δ.

Parameter	Process
$\delta < \frac{1}{2}$	X_t is stationary
$\frac{i}{2} \leq \delta < \frac{i+2}{2}$, X_t for $i = 1, 3, \ldots$	X_t is $\frac{i+1}{2}$ order difference stationary
$\delta > 0$, X_t	X_t is a long memory process
$\delta = 0$	X_t is a white noise process
$\delta = \frac{1}{2}$	X_t is a random walk process
$\delta = \frac{3}{2}$	X_t is a random run process

N values of each realization, thus obtaining 1000 estimates for each of the estimators. We finished by computing the sample mean and MSE of the 1000 estimates as well as the ratios of the MSEs of each of the alternative estimators to the MSE of the traditional unbiased estimator. By repeating this procedure 40 times for each choice of δ, filter, and series length, we were able to obtain point and interval estimates for the MSEs of each estimator and the ratios of MSEs of the new estimators to the traditional unbiased estimator. An example of such interval estimates (spanning one standard error on each side of the point estimates) are depicted in Figures 4.1 and 4.2, where the blue and red symbols correspond to the forward-backward unbiased and biased estimators, respectively. Note that Figure 4.2 is equivalent to Figure 4.1, except that it uses a log scale for the vertical axis so as to include the MSE values of the alternative biased estimator for $\delta = 2$, which are too large to depict on a linear scale. From these figures, we see that the alternative estimators provide improvements in MSE, with the forward-backward biased dominating the forward-backward unbiased in all cases except for values of $\delta \geq 1.5$. In particular, in light of the results from all the filters used in our study (see Appendix A), it is clear that the former estimator is not generally useful for values of $\delta \geq 1.5$, whereas the latter provides some improvement for all values of δ under consideration (and should provide similar results for any value of δ). Further, the efficiency of both estimators improves as sample size decreases, with the MSEs converging (faster in the case of the forward-backward unbiased estimator) to the traditional

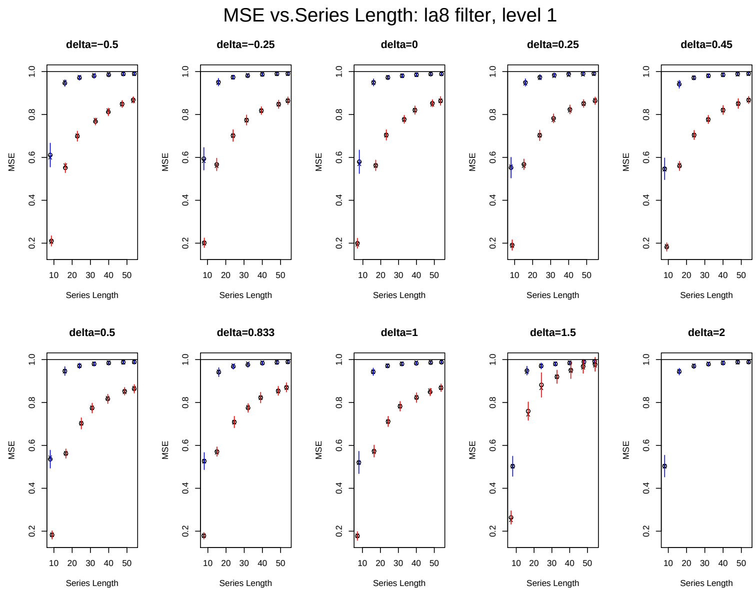

Figure 4.1: MSE vs. N for the level 1 LA8 wavelet filter and 8 different values of δ. Note that in the bottom right panel, the vertical axis is chosen so that values corresponding to the alternative biased estimator are not pictured (see Figure 4.2).

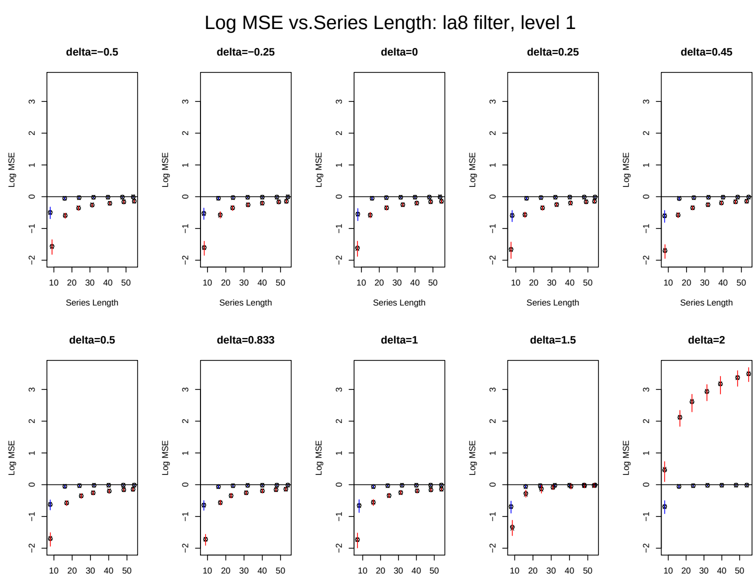

Figure 4.2: Log MSE vs. N for the level 1 LA8 wavelet filter and 8 different values of δ.

unbiased estimator as sample size increases. Not surprisingly, these results corroborate our findings in the preceding chapter: the asymptotic distributions (and hence the MSEs) of the alternative and traditional unbiased estimators are equivalent for intrinsically stationary processes, with the same being true for the alternative biased estimator only for processes that are stationary or first order difference stationary ($\delta < 1.5$).

Also depicted in Figures 4.1 and 4.2 are the true MSE ratios, obtained from the expressions for bias and variance that were derived in Appendix B. Given the accuracy of the average MSEs obtained via simulation, it is generally very difficult to differentiate the true MSEs (x's) from the means of the simulated MSEs (o's).

Figures corresponding to the other filters are located in Appendix A and depict results very similar to those above. From these figures we note that convergence of the MSE of the forward-backward unbiased estimator is faster as wavelet filter level increases, while convergence of the MSE of the biased estimator remains roughly the same. In addition, it appears that the LA8 filter provides greater improvement in MSE (for both of the alternative estimators) than the D4 filter; however this conclusion must take into consideration that the sample size is always twice as long (relative to the filter length) for the D4 filter. Since the results indicate that MSE improves for the estimators as sample size decreases with respect to filter length, we would thus expect to see lower MSEs when using the LA8 filter.

Chapter 5

CONCLUSION

This thesis has proposed two alternative estimators of the wavelet variance that are formed by utilizing both forward and backward wavelet coefficients: that is, wavelet coefficients that have been obtained by filtering a time series in both forward (typical) and backward directions. The first of the two alternative estimators, the forward-backward unbiased estimator, is obtained by evaluating the mean of the squared nonboundary forward and backward wavelet coefficients, while the second, the forward-backward biased estimator, is obtained by evaluating the mean of all of the squared forward and backward wavelet coefficients obtained by assuming a reflection boundary condition (i.e. assuming that unrealized portions of the series of interest are reflections rather than circular repetitions of the given values).

Expressions for the theoretical MSEs for the alternative estimators as well as for the traditional unbiased estimator were derived and asymptotic theory showed that the forward-backward unbiased estimator converges in distribution to the traditional unbiased estimator for processes that are intrinsically stationary, while the same is true of the forward-backward biased estimator only for stationary and first order difference stationary processes.

A simulation study confirmed the theoretical results and gave strong evidence for improvements in MSE for the alternative estimators, especially for small sample sizes. In particular, for series that were stationary or first order difference stationary, the forward-backward biased estimator had substantially lower MSE than its competitors, and, for second order difference stationary processes, the forward-backward unbiased estimator improved upon the traditional biased estimator in most cases.

Future work might entail a more detailed comparison of the alternative estimators for FD processes with $\delta \in (1, 1.5)$. In the preceding analysis we have seen that the MSE of the forward-backward biased estimator is less than that of the forward-backward unbiased estimator for all cases where values of δ were chosen to be no greater than 1 and that the reverse is true for all cases where δ was chose to be no less than 1.5. However it might be of interest to determine the relative behavior of these estimators in the region (1,1.5). Additional possibilities include developing formulas for computing the true bias and variance of $\overleftrightarrow{\nu}_b^2(\tau_j)$ and computing the variance of $\overleftrightarrow{\nu}_u^2(\tau_j)$ for second order difference stationary processes (one method would be to use the less efficient development in Section 2.2 and formulate the covariance matrix of Z' in addition to extending the formulas in Section 2.1), extending the results in this thesis to the non-Gaussian case by making assumptions similar to Serroukh, Walden and Percival (2000), and determining if the Equivalent Degrees of Freedom (EDOF), which are defined as

$$\mathrm{edof}(X) = \frac{2\mathrm{E}\left[X\right]^2}{\mathrm{var}\left(X\right)},$$

can be used in conjunction with a χ^2 distribution (as the degrees of freedom) to approximate the actual distribution of the wavelet variance estimators. For this latter extension, it is important to note that the exact distribution can be obtained for any quadratic form (see Johnson, Kotz and Balikrishnan, 1995), and hence the exact distributions of the wavelet variance estimators can be compared with χ^2 approximations to explore the adequacy of EDOF as a summary statistic.

The work herein is of interest because it allows practitioners to better estimate the wavelet variance for a wide range of processes, especially in situations where sample size is small relative to filter length (i.e. when the wavelet variance is poorly estimated in the first place). Among the alternative estimators, the forward-backward unbiased is a more conservative choice because it can be used reliably in conjunction with any intrinsically stationary process, although its statistical properties are not as good as those of the forward-backward biased estimator for processes that are stationary and first order difference stationary. Ultimately, it is possible that the limitations of the forward-backward biased estimator can be overcome due to the fact that one can often visually recognize processes that require more than one order of differencing to achieve stationarity and apply enough differencing operations so as to render them suitable for the biased estimator.

BIBLIOGRAPHY

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

Ferguson, T. S. (1996), *A Course in Large Sample Theory*, Boca Raton, FL: Chapman & Hall/CRC.

Greenhall, C. A., Howe, D. A., and Percival, D. B. (1999), "Total Variance, an Estimator of Long-Term Frequency Stability," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 46, 1183–1191.

Isserlis, L. (1918), "On a Formula for the Product-Moment Coefficient of Any Order of a Normal Frequency Distribution in Any Number of Variables," *Biometrika*, 12, 134–139.

Johnson, N., Kotz, S., and Balikrishnan, N. (1995), *Continuous Univariate Distributions*, New York, NY: John Wiley & Sons.

Percival, D. B. (1995), "On Estimation of the Wavelet Variance," *Biometrika*, 82, 619–631.

Percival, D. B. and Walden, A. T. (2000), *Wavelet Methods for Time Series Analysis*, Cambridge, UK: Cambridge University Press.

Ravishanker, N. and Dey, D. K. (2002), *A First Course in Linear Model Theory*, Boca Raton, FL: Chapman & Hall/CRC.

Serroukh, A., Walden, A. T., and Percival, D. B. (2000), "Statistical properties of the wavelet variance estimator for non-Gaussian/non-linear time series," *Journal of the American Statistical Association*, 95, 184–196.

Appendix A

FIGURES

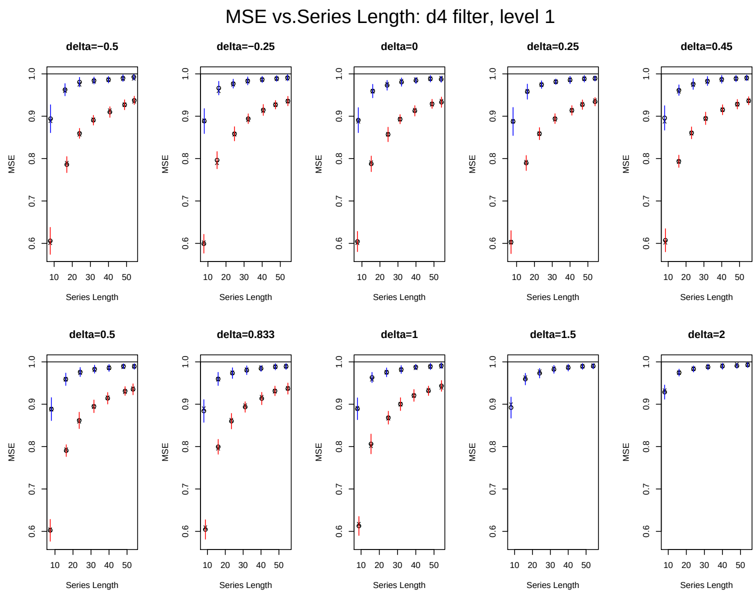

Figure A.1: MSE vs. N for the level 1 D4 wavelet filter and 10 different values of δ.

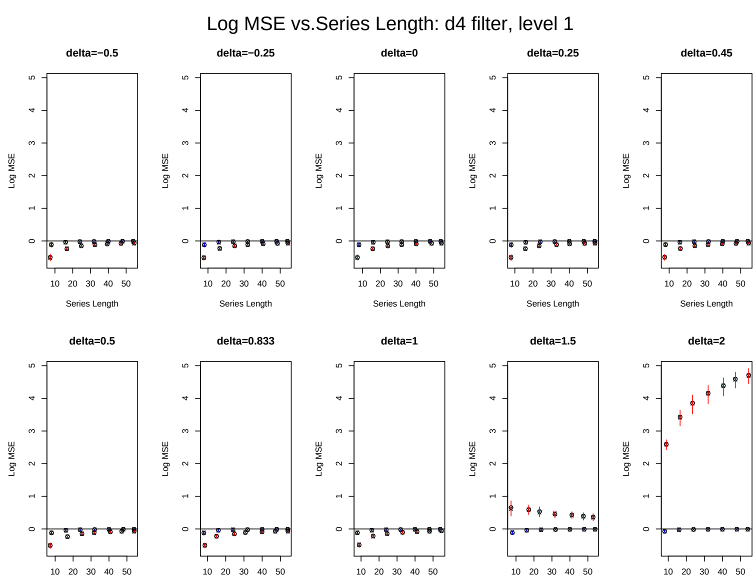

Figure A.2: Log MSE vs. N for the level 1 D4 wavelet filter and 10 different values of δ.

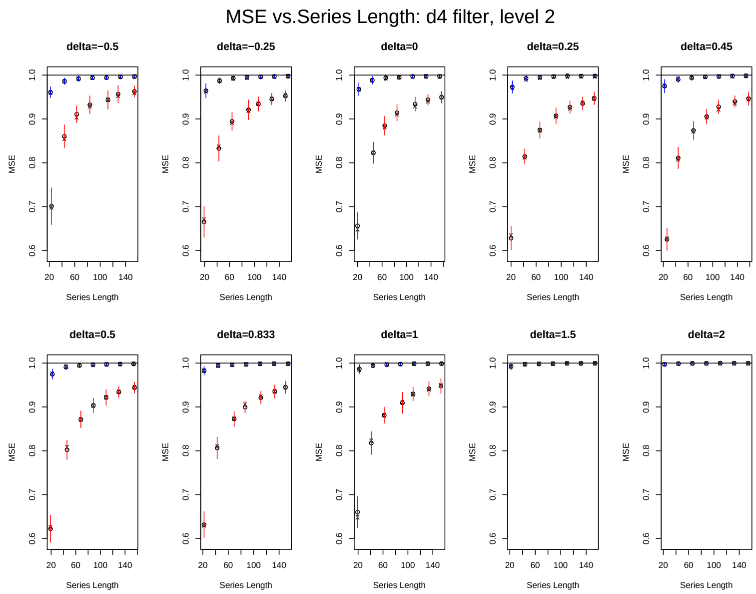

Figure A.3: MSE vs. N for the level 2 D4 wavelet filter and 10 different values of δ.

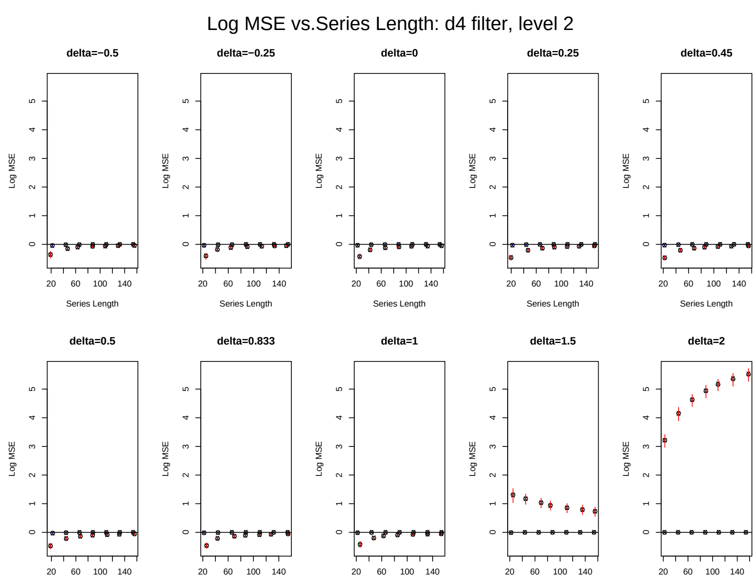

Figure A.4: Log MSE vs. N for the level 2 D4 wavelet filter and 10 different values of δ.

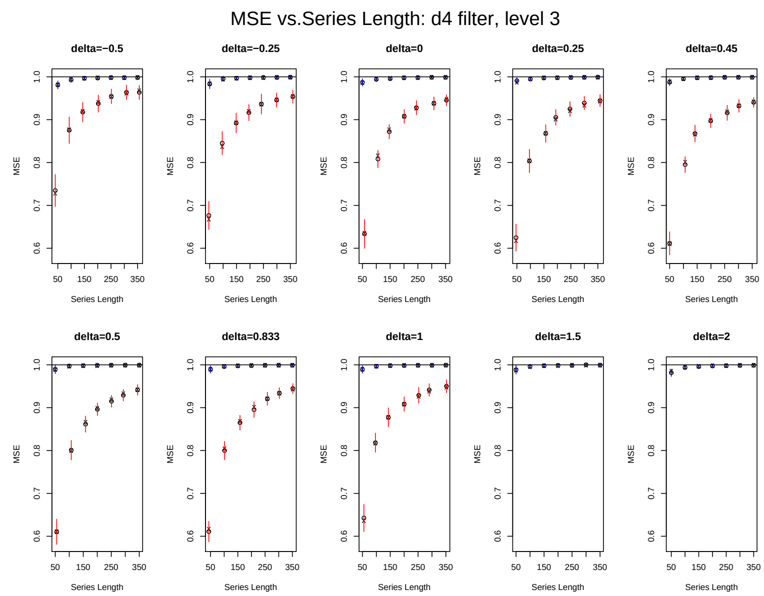

Figure A.5: MSE vs. N for the level 3 D4 wavelet filter and 10 different values of δ.

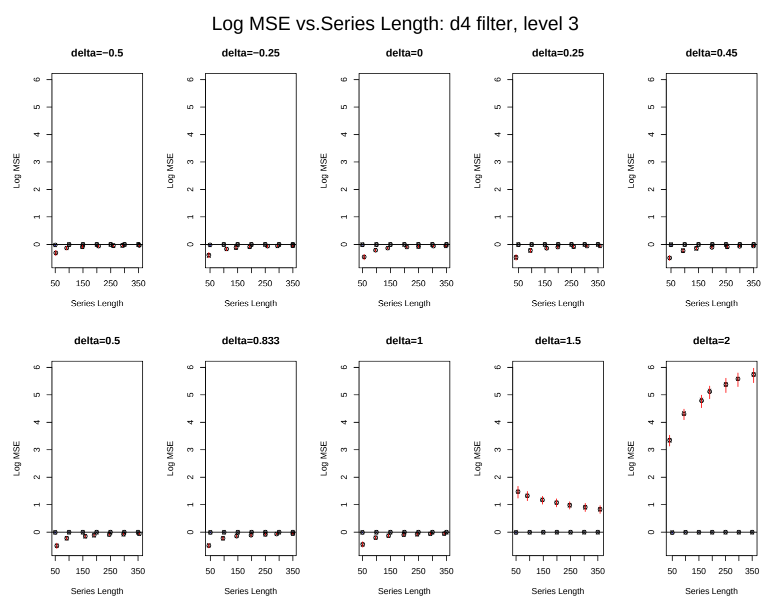

Figure A.6: Log MSE vs. N for the level 3 D4 wavelet filter and 10 different values of δ.

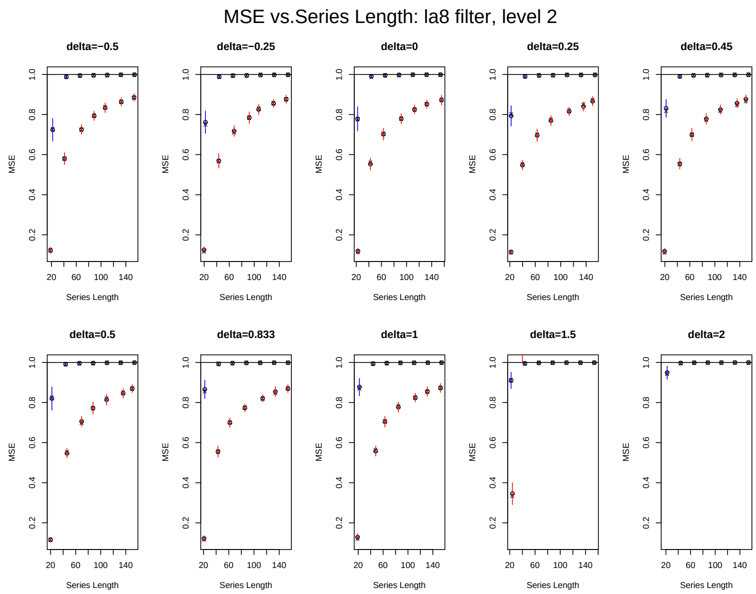

Figure A.7: MSE vs. N for the level 2 LA8 wavelet filter and 10 different values of δ.

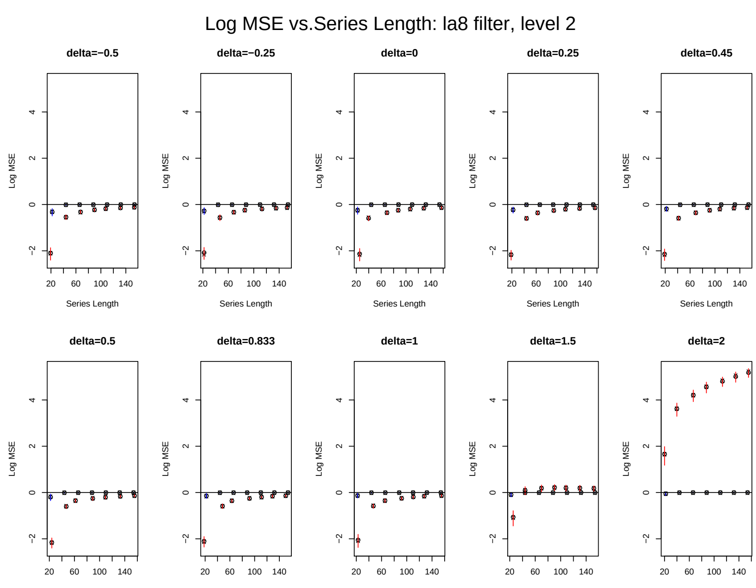

Figure A.8: Log MSE vs. N for the level 2 LA8 wavelet filter and 10 different values of δ.

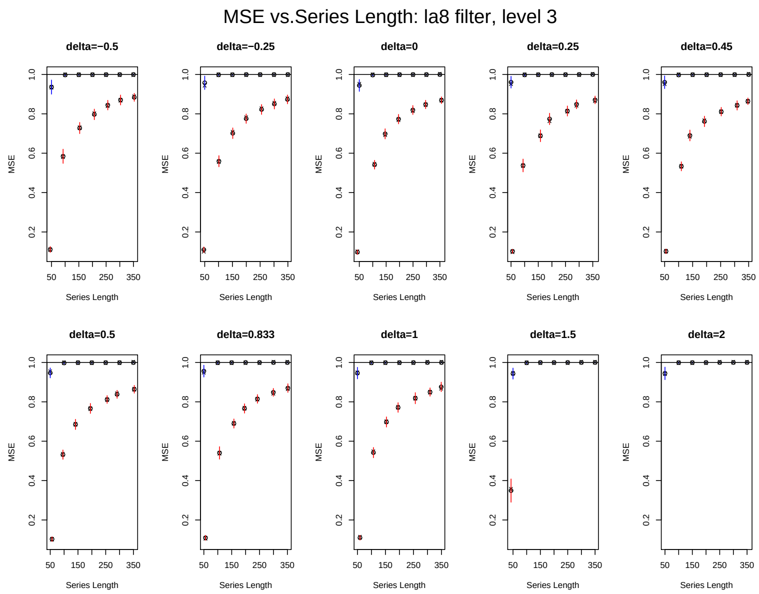

Figure A.9: MSE vs. N for the level 3 LA8 wavelet filter and 10 different values of δ.

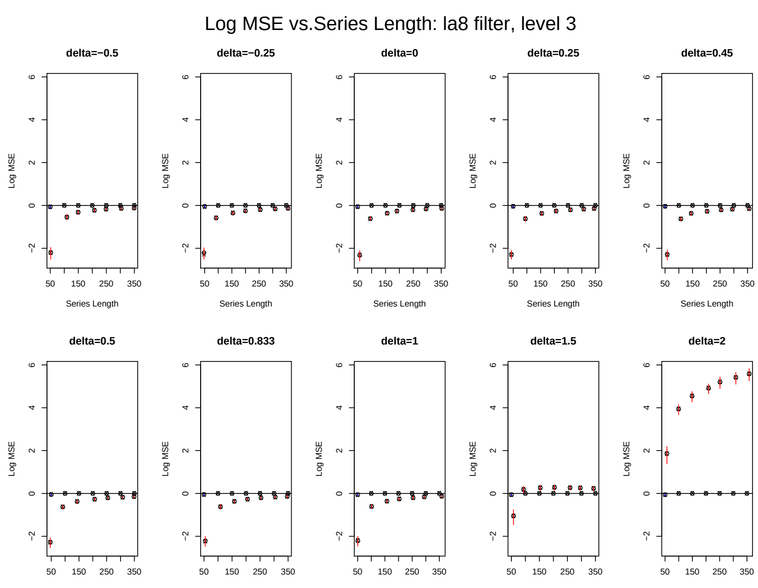

Figure A.10: Log MSE vs. N for the level 3 LA8 wavelet filter and 10 different values of δ.

Appendix B

EFFICIENT METHODS FOR COMPUTING THE WAVELET VARIANCE

As mentioned in Section 2.2 and Chapter 4, there exist more efficient methods for computing the expectation, bias and variance of $\overleftrightarrow{\nu}^2_b(\tau_j)$. To begin, Equations (2.6) and (2.8) show that

$$\begin{aligned}\text{cov}\,\left(\boldsymbol{x}^T A\boldsymbol{x}, \boldsymbol{x}^T B\boldsymbol{x}\right) &= \text{E}\,\left(\boldsymbol{x}^T A\boldsymbol{x}\cdot\boldsymbol{x}^T B\boldsymbol{x}\right) - \text{E}\left(\boldsymbol{x}^T A\boldsymbol{x}\right)\text{E}\,\left(\boldsymbol{x}^T B\boldsymbol{x}\right)\\ &= \text{tr}\,(A\Sigma)\,\text{tr}\,(B\Sigma) + 2\text{tr}\,(A\Sigma B\Sigma) - \text{tr}\,(A\Sigma)\,\text{tr}\,(B\Sigma)\\ &= 2\text{tr}\,(A\Sigma B\Sigma)\,, \end{aligned} \tag{B.1}$$

a result which will be useful in the ensuing discussion.

Assuming $\{X_t\}$ is stationary, we now express the level j forward MODWT wavelet coefficients of the reflected series $\boldsymbol{X}'$ as

$$\widetilde{\boldsymbol{W}}_j \equiv \left(\overrightarrow{\boldsymbol{W}}_{B,j}, \overrightarrow{\boldsymbol{W}}_{NB,j}, \overleftarrow{\boldsymbol{W}}_{B,j}, \overleftarrow{\boldsymbol{W}}_{NB,j}\right)^T,$$

where $\overrightarrow{\boldsymbol{W}}_{B,j}$ are the coefficients for indices $t = 0, \ldots, L_j - 2$, $\overrightarrow{\boldsymbol{W}}_{NB,j}$ for $t = L_j - 1, \ldots, N-1$, $\overleftarrow{\boldsymbol{W}}_{B,j}$ for $t = N, \ldots, N + L_j - 2$, and $\overleftarrow{\boldsymbol{W}}_{NB,j}$ for $t = N + L_j - 1, \ldots, 2N$. It is important to recognize that $\overrightarrow{\boldsymbol{W}}_{NB,j}$ and $\overleftarrow{\boldsymbol{W}}_{NB,j}$ are the same as, respectively, the level j forward and backward MODWT wavelet coefficients of $\boldsymbol{X}$ for indices $t = L_j - 1, \ldots, N-1$ (as defined in Section 2.2).

With appropriate definitions of matrices $\overrightarrow{\mathcal{W}}_{B,j}$, $\overrightarrow{\mathcal{W}}_{NB,j}$, $\overleftarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ we can express the forward MODWT wavelet coefficients of $\boldsymbol{X}'$ as

$$\widetilde{\boldsymbol{W}}_j \equiv \left(\overrightarrow{\boldsymbol{W}}_{B,j}, \overrightarrow{\boldsymbol{W}}_{NB,j}, \overleftarrow{\boldsymbol{W}}_{B,j}, \overleftarrow{\boldsymbol{W}}_{NB,j}\right)^T \equiv \left(\overrightarrow{\mathcal{W}}_{B,j}, \overrightarrow{\mathcal{W}}_{NB,j}, \overleftarrow{\mathcal{W}}_{B,j}, \overleftarrow{\mathcal{W}}_{NB,j}\right)^T \boldsymbol{X}.$$

The construction of these matrices is deferred until later. Using the development above, we can obtain an expression for the expectation of $\overleftrightarrow{\nu}^2_b(\tau_j)$:

$$
\begin{aligned}
\mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) &\equiv \mathrm{E}\left(\frac{1}{2N}\widetilde{\boldsymbol{W}}^T\widetilde{\boldsymbol{W}}\right) \\
&= \frac{1}{2N}\mathrm{E}\Bigg(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}^T_{B,j}\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X} + \boldsymbol{X}^T\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X} \\
&\qquad + \boldsymbol{X}^T\overleftarrow{\mathcal{W}}^T_{B,j}\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X} + \boldsymbol{X}^T\overleftarrow{\mathcal{W}}^T_{NB,j}\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\Bigg) \\
&= \frac{1}{2N}\Bigg\{\mathrm{E}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}^T_{B,j}\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X}\right) + \mathrm{E}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right) \\
&\qquad + \mathrm{E}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}^T_{B,j}\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X}\right) + \mathrm{E}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}^T_{NB,j}\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\Bigg\} \\
&= \frac{1}{2N}\Bigg\{\mathrm{tr}\left(\overrightarrow{\mathcal{W}}^T_{B,j}\overrightarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right) + \mathrm{tr}\left(\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right) \\
&\qquad + \mathrm{tr}\left(\overleftarrow{\mathcal{W}}^T_{B,j}\overleftarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right) + \mathrm{tr}\left(\overleftarrow{\mathcal{W}}^T_{NB,j}\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)\Bigg\}, \qquad \text{(B.2)}
\end{aligned}
$$

where Equation (B.2) follows from Equation (2.6). Appealing to Equation (B.2) we see that

$$
\begin{aligned}
\mathrm{bias}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) &= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \nu^2(\tau_j) \qquad \text{(B.3)} \\
&= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \mathrm{E}\left(\overleftrightarrow{\nu}^2_u(\tau_j)\right) \qquad \text{(B.4)} \\
&= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \frac{1}{2M_j}\mathrm{E}\left(\sum_{t=L_j-1}^{N-1}\overrightarrow{W}^2_{j,t}\right) - \frac{1}{2M_j}\mathrm{E}\left(\sum_{t=L_j-1}^{N-1}\overleftarrow{W}^2_{j,t}\right) \\
&= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \frac{1}{2M_j}\mathrm{E}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right) \qquad \text{(B.5)} \\
&\qquad - \frac{1}{2M_j}\mathrm{E}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}^T_{NB,j}\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right) \\
&= \mathrm{E}\left(\overleftrightarrow{\nu}^2_b(\tau_j)\right) - \frac{1}{2M_j}\mathrm{tr}\left(\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right) \\
&\qquad - \frac{1}{2M_j}\mathrm{tr}\left(\overleftarrow{\mathcal{W}}^T_{NB,j}\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right) \\
&= \frac{1}{2N}\left\{\mathrm{tr}\left(\overrightarrow{\mathcal{W}}^T_{B,j}\overrightarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right) + \mathrm{tr}\left(\overleftarrow{\mathcal{W}}^T_{B,j}\overleftarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right)\right\} \\
&\qquad + \frac{1}{2}\left(\frac{1}{N} - \frac{1}{M_j}\right)\Bigg\{\mathrm{tr}\left(\overrightarrow{\mathcal{W}}^T_{NB,j}\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)
\end{aligned}
$$

$$\mathrm{tr}\left(\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)\Bigg\},\tag{B.6}$$

where Equation (B.4) follows from Equation (B.3) since $\overleftrightarrow{\nu}_u^2(\tau_j)$ is unbiased.

In a similar fashion, the variance of $\overleftrightarrow{\nu}_b^2(\tau_j)$ is

$$\begin{aligned}
\mathrm{var}\left(\overleftrightarrow{\nu}_b^2(\tau_j)\right) &\equiv \mathrm{var}\left(\frac{1}{2N}\widetilde{\boldsymbol{W}}^T\widetilde{\boldsymbol{W}}\right)\\
&= \frac{1}{4N^2}\mathrm{var}\Bigg(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X}+\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\\
&\qquad +\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X}+\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\Bigg)\\
&= \frac{1}{4N^2}\Bigg\{\mathrm{var}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X}\right)+\mathrm{var}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\\
&\quad +\mathrm{var}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X}\right)+\mathrm{var}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\\
&\quad +2\mathrm{cov}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X},\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\\
&\quad +2\mathrm{cov}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X},\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X}\right)\\
&\quad +2\mathrm{cov}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\boldsymbol{X},\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\\
&\quad +2\mathrm{cov}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X},\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X}\right)\\
&\quad +2\mathrm{cov}\left(\boldsymbol{X}^T\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\boldsymbol{X},\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\\
&\quad +2\mathrm{cov}\left(\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\boldsymbol{X},\boldsymbol{X}^T\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\boldsymbol{X}\right)\Bigg\}\\
&= \frac{1}{4N^2}\Bigg\{2\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right)^2+2\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)^2\\
&\quad +2\mathrm{tr}\left(\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right)^2+2\mathrm{tr}\left(\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)^2\\
&\quad +4\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)\\
&\quad +4\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right)\\
&\quad +4\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{B,j}^T\overrightarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)\\
&\quad +4\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\overleftarrow{\mathcal{W}}_{B,j}^T\overleftarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\right)\\
&\quad +4\mathrm{tr}\left(\overrightarrow{\mathcal{W}}_{NB,j}^T\overrightarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\overleftarrow{\mathcal{W}}_{NB,j}^T\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)
\end{aligned}$$

$$+4\mathrm{tr}\left(\overleftarrow{\mathcal{W}}_{B,j}^{T}\overleftarrow{\mathcal{W}}_{B,j}\Sigma_{\boldsymbol{X}}\overleftarrow{\mathcal{W}}_{NB,j}^{T}\overleftarrow{\mathcal{W}}_{NB,j}\Sigma_{\boldsymbol{X}}\right)\Bigg\}. \tag{B.7}$$

To complete the expressions for bias and variance above, we must specify the forms of the matrices $\overrightarrow{\mathcal{W}}_{B,j}$, $\overrightarrow{\mathcal{W}}_{NB,j}$, $\overleftarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$. Assume that these matrices are constructed from the elements of a filter $h_{j,l}$ of width $L_j = 8$. Given that $L_j < N$, it is straightforward to show (using Equation (1.2)) that

$$\overrightarrow{\mathcal{W}}_{B,j} = \begin{bmatrix} h_{j,0}+h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6} & h_{j,7} & 0 & \ldots & 0 \\ h_{j,1}+h_{j,2} & h_{j,0}+h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6} & h_{j,7} & 0 & 0 & \ldots & 0 \\ h_{j,2}+h_{j,3} & h_{j,1}+h_{j,4} & h_{j,0}+h_{j,5} & h_{j,6} & h_{j,7} & 0 & 0 & 0 & \ldots & 0 \\ h_{j,3}+h_{j,4} & h_{j,2}+h_{j,5} & h_{j,1}+h_{j,6} & h_{j,0}+h_{j,7} & 0 & 0 & 0 & 0 & \ldots & 0 \\ h_{j,4}+h_{j,5} & h_{j,3}+h_{j,6} & h_{j,2}+h_{j,7} & h_{j,1} & h_{j,0} & 0 & 0 & 0 & \ldots & 0 \\ h_{j,5}+h_{j,6} & h_{j,4}+h_{j,7} & h_{j,3} & h_{j,2} & h_{j,1} & h_{j,0} & 0 & 0 & \ldots & 0 \\ h_{j,6}+h_{j,7} & h_{j,5} & h_{j,4} & h_{j,3} & h_{j,2} & h_{j,1} & h_{j,0} & 0 & \ldots & 0 \end{bmatrix}$$

Thus, for the general case, we see that the m, nth element of $\overrightarrow{\mathcal{W}}_{B,j}$ is $h_{j,m-n} + h_{j,m+n-1}$, where $1 \leq m \leq L_j - 1$ and $1 \leq n \leq N$. Similarly, if $L_j = 8$,

$$\overrightarrow{\mathcal{W}}_{NB,j} = \begin{bmatrix} h_{j,7} & h_{j,6} & h_{j,5} & h_{j,4} & h_{j,3} & h_{j,2} & h_{j,1} & h_{j,0} & 0 & 0 & \ldots & 0 \\ 0 & h_{j,7} & h_{j,6} & h_{j,5} & h_{j,4} & h_{j,3} & h_{j,2} & h_{j,1} & h_{j,0} & 0 & \ldots & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & \ldots & 0 & h_{j,7} & h_{j,6} & h_{j,5} & h_{j,4} & h_{j,3} & h_{j,2} & h_{j,1} & h_{j,0} & 0 \\ 0 & \ldots & 0 & 0 & h_{j,7} & h_{j,6} & h_{j,5} & h_{j,4} & h_{j,3} & h_{j,2} & h_{j,1} & h_{j,0} \end{bmatrix},$$

$$\overleftarrow{\mathcal{W}}_{B,j} = \begin{bmatrix} 0 & \ldots & 0 & h_{j,7} & h_{j,6} & h_{j,5} & h_{j,4} & h_{j,3} & h_{j,2} & h_{j,0}+h_{j,1} \\ 0 & \ldots & 0 & 0 & h_{j,7} & h_{j,6} & h_{j,5} & h_{j,4} & h_{j,0}+h_{j,3} & h_{j,1}+h_{j,2} \\ 0 & \ldots & 0 & 0 & 0 & h_{j,7} & h_{j,6} & h_{j,0}+h_{j,5} & h_{j,1}+h_{j,4} & h_{j,2}+h_{j,3} \\ 0 & \ldots & 0 & 0 & 0 & 0 & h_{j,0}+h_{j,7} & h_{j,1}+h_{j,6} & h_{j,2}+h_{j,5} & h_{j,3}+h_{j,4} \\ 0 & \ldots & 0 & 0 & 0 & h_{j,0} & h_{j,1} & h_{j,2}+h_{j,7} & h_{j,3}+h_{j,6} & h_{j,4}+h_{j,5} \\ 0 & \ldots & 0 & 0 & h_{j,0} & h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4}+h_{j,7} & h_{j,5}+h_{j,6} \\ 0 & \ldots & 0 & h_{j,0} & h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6}+h_{j,7} \end{bmatrix}$$

and

$$\overleftarrow{\mathcal{W}}_{NB,j} = \begin{bmatrix} 0 & \dots & 0 & 0 & h_{j,0} & h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6} & h_{j,7} \\ 0 & \dots & 0 & h_{j,0} & h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6} & h_{j,7} & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & h_{j,0} & h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6} & h_{j,7} & 0 & \dots & 0 \\ h_{j,0} & h_{j,1} & h_{j,2} & h_{j,3} & h_{j,4} & h_{j,5} & h_{j,6} & h_{j,7} & 0 & 0 & \dots & 0 \end{bmatrix}.$$

As before, we see that the m,nth elements of $\overrightarrow{\mathcal{W}}_{NB,j}$, $\overleftarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ are, respectively, $h_{j,m-n+L_j-1}$, $h_{j,m+n-N-1} + h_{m-n+N}$ and $h_{j,m+n-N+L_j-2}$, where $1 \le m \le N - L_j + 1$ for $\overrightarrow{\mathcal{W}}_{NB,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$, $1 \le m \le L_j - 1$ for $\overleftarrow{\mathcal{W}}_{B,j}$ and $1 \le n \le N$ for all of the matrices. These expressions hold for filters of any width and can be used in conjunction with Equations (B.6) and (B.7) to evaluate the statistical properties of the forward-backward biased estimator of the wavelet variance.

Similar results hold if $\{X_t\}$ is first order difference stationary. In this case $\boldsymbol{Y}$ (as defined in the beginning of Section 3.3) and $\Sigma_{\boldsymbol{Y}}$ must be substituted for $\boldsymbol{X}$ and $\Sigma_{\boldsymbol{X}}$ in the development above, and it can be shown that (again assuming that $h_{j,l}$ is a filter of width $L_j = 8$)

$$\overrightarrow{\mathcal{W}}_{B,j} = \begin{bmatrix} 0 & -c_{j,1} & -c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} & -c_{j,6} & 0 & \dots & 0 \\ 0 & c_{j,0} - c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} & -c_{j,6} & 0 & 0 & \dots & 0 \\ 0 & c_{j,1} - c_{j,3} & c_{j,0} - c_{j,4} & -c_{j,5} & -c_{j,6} & 0 & 0 & 0 & \dots & 0 \\ 0 & c_{j,2} - c_{j,4} & c_{j,1} - c_{j,5} & c_{j,0} - c_{j,6} & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & c_{j,3} - c_{j,5} & c_{j,2} - c_{j,6} & c_{j,1} & c_{j,0} & 0 & 0 & 0 & \dots & 0 \\ 0 & c_{j,4} - c_{j,6} & c_{j,3} & c_{j,2} & c_{j,1} & c_{j,0} & 0 & 0 & \dots & 0 \\ 0 & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} & c_{j,1} & c_{j,0} & 0 & \dots & 0 \end{bmatrix},$$

$$\overrightarrow{\mathcal{W}}_{NB,j} = \begin{bmatrix} 0 & c_{j,6} & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} & c_{j,1} & c_{j,0} & 0 & 0 & \dots & 0 \\ 0 & 0 & c_{j,6} & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} & c_{j,1} & c_{j,0} & 0 & \dots & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & 0 & \dots & 0 & c_{j,6} & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} & c_{j,1} & c_{j,0} & 0 \\ 0 & 0 & \dots & 0 & 0 & c_{j,6} & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} & c_{j,1} & c_{j,0} \end{bmatrix},$$

$$\overleftarrow{\mathcal{W}}_{B,j} = \begin{bmatrix} 0 & \dots & 0 & c_{j,6} & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} & c_{j,1} \\ 0 & \dots & 0 & 0 & c_{j,6} & c_{j,5} & c_{j,4} & c_{j,3} & c_{j,2} - c_{j,0} \\ 0 & \dots & 0 & 0 & 0 & c_{j,6} & c_{j,5} & c_{j,4} - c_{j,0} & c_{j,3} - c_{j,1} \\ 0 & \dots & 0 & 0 & 0 & 0 & c_{j,6} - c_{j,0} & c_{j,5} - c_{j,1} & c_{j,4} - c_{j,2} \\ 0 & \dots & 0 & 0 & 0 & -c_{j,0} & -c_{j,1} & c_{j,6} - c_{j,2} & c_{j,5} - c_{j,3} \\ 0 & \dots & 0 & 0 & -c_{j,0} & -c_{j,1} & -c_{j,2} & -c_{j,3} & c_{j,6} - c_{j,4} \\ 0 & \dots & 0 & -c_{j,0} & -c_{j,1} & -c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} \end{bmatrix}$$

and

$$\overleftarrow{\mathcal{W}}_{NB,j} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & -c_{j,0} & -c_{j,1} & -c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} & -c_{j,6} \\ 0 & 0 & \dots & 0 & -c_{j,0} & -c_{j,1} & -c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} & -c_{j,6} & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & 0 & -c_{j,0} & -c_{j,1} & -c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} & -c_{j,6} & 0 & \dots & 0 \\ 0 & -c_{j,0} & -c_{j,1} & -c_{j,2} & -c_{j,3} & -c_{j,4} & -c_{j,5} & -c_{j,6} & 0 & 0 & \dots & 0 \end{bmatrix},$$

where $c_{j,l} = \sum_{k=0}^{l} h_{j,k}$ and $c_{j,L_j-1} = \sum_{k=0}^{L_j-1} h_{j,k} = 0$. It is straightforward to verify these results by choosing $N = 10$ and doing the (somewhat tedious) manipulations by hand. In general,we see that the m, nth elements of $\overrightarrow{\mathcal{W}}_{B,j}$, $\overrightarrow{\mathcal{W}}_{NB,j}$, $\overleftarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ are, respectively, $c_{j,m-n} - c_{j,m+n-2}$, c_{m-n+L_j-1}, $c_{j,m-n+N} - c_{j,m+n-N-2}$ and $-c_{m+n-N+L_j-3}$, where $1 \leq m \leq L_j - 1$ for $\overrightarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{B,j}$, $1 \leq m \leq N - L_j + 1$ for $\overrightarrow{\mathcal{W}}_{NB,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ and $1 \leq n \leq N$ for all of the matrices. As in the stationary case, these expressions hold for filters of any width.

Finally, we extend these results to processes that are second order difference stationary. Substituting $\boldsymbol{Z}$ (as defined as the beginning of Section 3.3) and $\Sigma_{\boldsymbol{Z}}$ for $\boldsymbol{X}$ and $\Sigma_{\boldsymbol{X}}$, it is straightforward to show (assuming $L_j = 8$ and noting that $Y_t = \sum_{l=0}^{t} Z_l$)

$$\overrightarrow{\mathcal{W}}_{B,j} = \begin{bmatrix} g_0 & g_0 & g_1 & g_2 & g_3 & g_4 & g_5 & 0 & \dots & 0 \\ g_0 + g_1 & g_0 + g_1 & g_2 & g_3 & g_4 & g_5 & 0 & 0 & \dots & 0 \\ g_1 + g_2 & g_1 + g_2 & g_0 + g_3 & g_4 & g_5 & 0 & 0 & 0 & \dots & 0 \\ g_2 + g_3 & g_2 + g_3 & g_1 + g_4 & g_0 + g_5 & 0 & 0 & 0 & 0 & \dots & 0 \\ g_3 + g_4 & g_3 + g_4 & g_2 + g_5 & g_1 & g_0 & 0 & 0 & 0 & \dots & 0 \\ g_4 + g_5 & g_4 + g_5 & g_3 & g_2 & g_1 & g_0 & 0 & 0 & \dots & 0 \\ g_5 & g_5 & g_4 & g_3 & g_2 & g_1 & g_0 & 0 & \dots & 0 \end{bmatrix},$$

$$\overrightarrow{\mathcal{W}}_{NB,j} = \begin{bmatrix} 0 & 0 & g_5 & g_4 & g_3 & g_2 & g_1 & g_0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & g_5 & g_4 & g_3 & g_2 & g_1 & g_0 & 0 & \dots & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & 0 & 0 & \dots & 0 & g_5 & g_4 & g_3 & g_2 & g_1 & g_0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & g_5 & g_4 & g_3 & g_2 & g_1 & g_0 \end{bmatrix},$$

$$\overleftarrow{\mathcal{W}}_{B,j} = \begin{bmatrix} -g_0 & \dots & -g_0 & g_5 - g_0 & g_4 - g_0 & g_3 - g_0 & g_2 - g_0 & g_1 - g_0 \\ -d_0 & \dots & -d_0 & -d_0 & g_5 - d_0 & g_4 - d_0 & g_3 - d_0 & g_2 - d_0 \\ -g_1 - g_2 & \dots & -d_1 & -d_1 & -d_1 & g_5 - d_1 & g_4 - d_1 & g_0 + g_3 - d_1 \\ -d_2 & \dots & -d_2 & -d_2 & -d_2 & -d_2 & g_0 + g_5 - d_2 & g_1 + g_4 - d_2 \\ -d_3 & \dots & -d_3 & -d_3 & -d_3 & g_0 - d_3 & g_1 - d_3 & g_2 + g_5 - d_3 \\ -d_4 & \dots & -d_4 & -d_4 & g_0 - d_4 & g_1 - d_4 & g_2 - d_4 & g_3 - d_4 \\ -g_5 & \dots & -g_5 & g_0 - g_5 & g_1 - g_5 & g_2 - g_5 & g_3 - g_5 & g_4 - g_5 \end{bmatrix}$$

and

$$\overleftarrow{\mathcal{W}}_{NB,j} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 & -g_0 & -g_1 & -g_2 & -g_3 & -g_4 & -g_5 \\ 0 & 0 & 0 & \dots & 0 & -g_0 & -g_1 & -g_2 & -g_3 & -g_4 & -g_5 & 0 \\ \vdots & & & & & & & & & & & \vdots \\ 0 & 0 & 0 & -g_0 & -g_1 & -g_2 & -g_3 & -g_4 & -g_5 & 0 & \dots & 0 \\ 0 & 0 & -g_0 & -g_1 & -g_2 & -g_3 & -g_4 & -g_5 & 0 & 0 & \dots & 0 \end{bmatrix},$$

where $g_l = cc_{j,l}$, $d_l = cc_{j,l} + cc_{j,l+1}$, $cc_{j,l} = \sum_{k=0}^{l} c_{j,k}$ and $cc_{j,L_j-2} = \sum_{k=0}^{L_j-2} c_{j,k} = 0$. These results demonstrate that in general the m,nth elements of $\overrightarrow{\mathcal{W}}_{B,j}$, $\overrightarrow{\mathcal{W}}_{NB,j}$, $\overleftarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ are, respectively, $cc_{j,m-n} - cc_{j,m+n-3}$, cc_{m-n+L_j-1}, $cc_{j,m-n+N} - cc_{j,m+n-N-3} - cc_{j,m-1} - cc_{j,m-2}$ and $-cc_{m+n-N+L_j-4}$, where $1 \le m \le L_j - 1$ for $\overrightarrow{\mathcal{W}}_{B,j}$ and $\overleftarrow{\mathcal{W}}_{B,j}$, $1 \le m \le N - L_j + 1$ for $\overrightarrow{\mathcal{W}}_{NB,j}$ and $\overleftarrow{\mathcal{W}}_{NB,j}$ and $1 \le n \le N$ for all of the matrices. As before, these expressions hold for filters of any width.

Appendix C

CODE AND REVISIONS

The monte carlo experiments in Chapter 4 were performed with the `R` statistical computing environment. All code used in the analysis can be obtained from the author's web-page,

http://www.ealdrich.com/Grad/Stat/Thesis/,

and should be sufficient for complete reproduction of the results herein. In addition, version 0.2-1 of the `wavelets` package for `R` and Peter Craigmile's `R` code for simulating FD(δ) processes via the Davies-Harte method are also posted at the preceding URL, since both are required for the analysis. Although both of these sets of code should be available from other sources, potential changes might render them incompatible with the code written for this thesis, and hence versions concurrent with the writing of this document will be preserved. It should be noted that the concurrent `wavelets` package for `R` is in a very rudimentary state and that subsequent versions of this package should be available from the Comprehensive `R` Archive Network (CRAN): http://cran.r-project.org.

This thesis (as with most) contains a number of errors and will be revised in the future. As such, both original and revised versions will be maintained at the URL above and are free for public distribution.