Mathematical Background for Wavelet Estimators of Cross-Covariance and Cross-Correlation

Brandon Whitcher Peter Guttorp Donald B. Percival

NRCSE

Technical Report Series

NRCSE-TRS No. 038

The **NRCSE** was established in 1996 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.

Mathematical Background for Wavelet Estimators of Cross-Covariance and Cross-Correlation

Brandon Whitcher
EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Peter Guttorp
University of Washington, Department of Statistics, Box 354322, Seattle, WA 98195-2933

Donald B. Percival
University of Washington, Applied Physics Laboratory, Box 355640, Seattle, WA 98195-5640
MathSoft, Inc., 1700 Westlake Avenue North, Seattle, Washington 98109-9891

June 8, 1999

Abstract

In this technical report we provide mathematical support to the claims in Whitcher, Guttorp, and Percival (1999). First, the decomposition of covariance by the wavelet covariance is firmly established. Central limit theorems for MODWT estimators of the wavelet covariance and correlation are then provided along with the definition of the variance for estimators of wavelet covariance under the assumption of Gaussianity.

1 Wavelet-Based Estimators of Covariance and Correlation

Here we define the basic quantities of interest for estimating association between two time series using the MODWT. The decomposition of covariance on a scale by scale basis of the wavelet covariance is shown, and central limit theorems are provided for the wavelet covariance and correlation.

1.1 Definition and Properties of the Wavelet Cross-Covariance

Let $\{U_t\} \equiv \{\ldots, U_{-1}, U_0, U_1, \ldots\}$ be a stochastic process whose dth order backward difference $(1-B)^d U_t = Z_t$ is a stationary Gaussian process with zero mean and spectral density function $S_Z(\cdot)$, where d is a non-negative integer. Let

$$\overline{W}_{j,t}^{(U)} = \tilde{h}_{j,l} * U_t \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} U_{t-l}, \quad t = \ldots, -1, 0, 1, \ldots,$$

be the stochastic process obtained by filtering $\{U_t\}$ with the MODWT wavelet filter $\{\tilde{h}_{j,l}\}$. Percival and Walden (2000, Sec. 8.2) showed that if $L \geq 2d$, then $\{\overline{W}_{j,t}^{(U)}\}$ is a stationary process with zero mean and spectrum given by $S_{j,U}(\cdot)$.

Let $\{X_t\} \equiv \{\ldots, X_{-1}, X_0, X_1, \ldots\}$ and $\{Y_t\} \equiv \{\ldots, Y_{-1}, Y_0, Y_1, \ldots\}$ be stochastic processes whose d_Xth and d_Yth order backward differences are stationary Gaussian processes as defined above, and define $d \equiv \max\{d_X, d_Y\}$. Let $S_{XY}(\cdot)$ denote their cross spectrum and, $S_X(\cdot)$ and $S_Y(\cdot)$ denote their autospectra, respectively. The wavelet cross-covariance of $\{X_t, Y_t\}$ for scale $\lambda_j = 2^{j-1}$ and lag τ is defined to be

$$\gamma_{\tau,XY}(\lambda_j) \equiv \operatorname{Cov}\left\{\overline{W}_{j,t}^{(X)}, \overline{W}_{j,t+\tau}^{(Y)}\right\}, \tag{1}$$

where $\{\overline{W}_{j,t}^{(X)}\}$ and $\{\overline{W}_{j,t}^{(Y)}\}$ are the scale λ_j MODWT coefficients for $\{X_t\}$ and $\{Y_t\}$, respectively. The MODWT coefficients have mean zero, when $L \geq 2d$, and therefore $\gamma_{\tau,XY}(\lambda_j) = E\{\overline{W}_{j,t}^{(X)} \overline{W}_{j,t+\tau}^{(Y)}\}$. When $\tau = 0$ we obtain the wavelet covariance between $\{X_t, Y_t\}$, which we denote as $\gamma_{XY}(\lambda_j) = \gamma_{0,XY}(\lambda_j)$ to simplify notation.

Setting $\tau = 0$ and Y_t to X_t or X_t to Y_t, Equation (1) reduces to the wavelet variance for X_t or Y_t denoted as, respectively, $\nu_X^2(\lambda_j)$ or $\nu_Y^2(\lambda_j)$ (Percival 1995). The wavelet variance decomposes the process variance on a scale by scale basis, and the wavelet cross-covariance give a similar decomposition for the process cross-covariance.

Theorem 1 *Let $\{X_t\}$ and $\{Y_t\}$ be two weakly stationary processes with autospectra given by $S_X(f)$ and $S_Y(f)$, respectively. If we require $L > 2d$, then for any integer $J \geq 1$ we have*

$$C_{\tau,XY} \equiv \operatorname{Cov}\{X_t, Y_{t+\tau}\} = \operatorname{Cov}\left\{\overline{V}_{J,t}^{(X)}, \overline{V}_{J,t+\tau}^{(Y)}\right\} + \sum_{j=1}^{J} \gamma_{\tau,XY}(\lambda_j),$$

where $\overline{V}^{(X)}_{J,t} \equiv \tilde{g}_{J,l} * X_t$ *and* $\overline{V}^{(Y)}_{J,t} \equiv \tilde{g}_{J,l} * Y_t$ *are obtained by filtering* $\{X_t\}$ *and* $\{Y_t\}$ *using the MODWT scaling filter* $\{\tilde{g}_{J,l}\}$*, respectively. As* $J \to \infty$*, we have*

$$C_{\tau,XY} = \sum_{j=1}^{\infty} \gamma_{\tau,XY}(\lambda_j),$$

which gives the required decomposition.

Proof of Theorem 1 Before proving Theorem 1, we require the following lemma.

Lemma 1 *For all* $\epsilon > 0$*, there exists a* J_ϵ *such that* $\left|\mathrm{Cov}\left\{\overline{V}^{(X)}_{J,t}, \overline{V}^{(Y)}_{J,t}\right\}\right| < \epsilon$ *for* $J > J_\epsilon$*.*

Proof of Lemma 1 For the orthonormal DWT $\sum_l g^2_{J,l} = 1$ and by definition $\tilde{g}_{J,l} = g_{J,l}/2^{J/2}$. Therefore we have $\sum_l \tilde{g}^2_{J,l} = 1/2^J$. Parseval's relation tells us that

$$\int_{-1/2}^{1/2} \widetilde{\mathcal{G}}_J(f)\, df = \int_{-1/2}^{1/2} \left|\widetilde{G}_J(f)\right|^2 \, df = \sum_{l=0}^{L_J-1} \tilde{g}^2_{J,l} = \frac{1}{2^J}.$$

We know the amplitude spectrum $A_{XY}(f) \equiv |S_{XY}(f)|$ is a non-negative real valued function. Hence, if $A_{XY}(\cdot)$ is bounded by some finite number C, then for $J > J_\epsilon$,

$$\left|\mathrm{Cov}\left\{\overline{V}^{(X)}_{J,t}, \overline{V}^{(Y)}_{J,t}\right\}\right| \leq \int_{-1/2}^{1/2} \widetilde{\mathcal{G}}_J(f)\, |S_{XY}(f)| \, df = C \int_{-1/2}^{1/2} \widetilde{\mathcal{G}}_J(f)\, df = \frac{C}{2^J} < \epsilon.$$

If $A_{XY}(\cdot)$ cannot be bounded by any finite number C, there at least exists a constant C_ϵ such that $\int_{A_{XY}(f) \geq C_\epsilon} A_{XY}(f)\, df < \epsilon/2$, using a Lebesgue integral. A rough bound on the squared gain function of the scaling filter for Daubechies wavelets is $\widetilde{\mathcal{G}}_J(f) \leq 1$, so for all $J > J_\epsilon$,

$$\begin{aligned}
\left|\int_{-1/2}^{1/2} \widetilde{\mathcal{G}}_J(f) S_{XY}(f)\, df\right| &\leq \int_{A_{XY}(f) \geq C_\epsilon} \widetilde{\mathcal{G}}_J(f)\, |S_{XY}(f)| \, df \\
&\quad + \int_{A_{XY}(f) < C_\epsilon} \widetilde{\mathcal{G}}_J(f)\, |S_{XY}(f)| \, df \\
&\leq \int_{A_{XY}(f) \geq C_\epsilon} A_{XY}(f)\, df + C_\epsilon \int_{A_{XY}(f) < C_\epsilon} \widetilde{\mathcal{G}}_J(f)\, df \\
&\leq \frac{\epsilon}{2} + C_\epsilon \int_{-1/2}^{1/2} \widetilde{\mathcal{G}}_J(f)\, df \leq \frac{\epsilon}{2} + \frac{C_\epsilon}{2^J} < \epsilon.
\end{aligned}$$

□

Without loss of generality, we set $\tau = 0$ and simply shift $\{\overline{W}^{(Y)}_{j,t}\}$ with respect to $\{\overline{W}^{(X)}_{j,t}\}$ to get $\tau \neq 0$. Because $\{\overline{W}^{(X)}_{j,t}\}$ and $\{\overline{W}^{(Y)}_{j,t}\}$ are obtained by filtering the processes $\{X_t\}$ and $\{Y_t\}$ with a Daubechies compactly supported wavelet filter of even length $L > 2d$, respectively, we know that $\{\overline{W}^{(X)}_{j,t}\}$ and $\{\overline{W}^{(Y)}_{j,t}\}$ are stationary processes with autospectra defined by $S_{j,X}(f) \equiv \widetilde{\mathcal{H}}_j(f) S_X(f)$ and $S_{j,Y}(f) \equiv \widetilde{\mathcal{H}}_j(f) S_Y(f)$ where $\widetilde{\mathcal{H}}_j(f) \equiv \widetilde{\mathcal{H}}(2^{j-1}f) \prod_{l=0}^{j-2} \widetilde{\mathcal{G}}(2^l f)$ is the squared gain function for

$\{\tilde{h}_j\}$. Note, the squared gain functions associated with unit scale for the wavelet and scaling filters are given by $\widetilde{\mathcal{H}}(f) \equiv |\widetilde{H}(f)|^2$ and $\widetilde{\mathcal{G}}(f) \equiv |\widetilde{G}(f)|^2$.

The covariance between $\{\overline{W}_{j,t}^{(X)}\}$ and $\{\overline{W}_{j,t}^{(Y)}\}$ is given by

$$\gamma_{XY}(\lambda_j) = \int_{-1/2}^{1/2} \widetilde{\mathcal{H}}_j(f) S_{XY}(f)\, df.$$

This is a straightforward generalization of the univariate case; see Whitcher (1998) for more details. The covariance between $\{\overline{V}_{j,t}^{(X)}\}$ and $\{\overline{V}_{j,t}^{(Y)}\}$ is given by

$$\text{Cov}\left\{\overline{V}_{J,t}^{(X)}, \overline{V}_{J,t}^{(Y)}\right\} = \int_{-1/2}^{1/2} \widetilde{\mathcal{G}}_J(f) S_{XY}(f)\, df,$$

where $\widetilde{\mathcal{G}}_J(f) \equiv \prod_{l=0}^{J-1} \widetilde{\mathcal{G}}(2^l f)$ is the squared gain function for $\{\tilde{g}_J\}$. Because of the following identity for squared gain functions $\widetilde{\mathcal{H}}(f) + \widetilde{\mathcal{G}}(f) = 1$ for all f (Percival and Walden 2000, Sec. 4.3), we have

$$\text{Cov}\{X_t, Y_t\} = \int_{-1/2}^{1/2} \left[\widetilde{\mathcal{G}}(f) + \widetilde{\mathcal{H}}(f)\right] S_{XY}(f)\, df = \text{Cov}\left\{\overline{V}_{1,t}^{(X)}, \overline{V}_{1,t}^{(Y)}\right\} + \gamma_{XY}(\lambda_1),$$

and the case when $J = 1$ holds. We now proceed to prove the main assertion by induction. Assume the property holds for $J - 1$; i.e.,

$$\text{Cov}\{X_t, Y_t\} = \text{Cov}\left\{\overline{V}_{J-1,t}^{(X)}, \overline{V}_{J-1,t}^{(Y)}\right\} + \sum_{j=1}^{J-1} \gamma_{XY}(\lambda_j).$$

So we have

$$\begin{aligned}
\text{Cov}\left\{V_{J-1,t}^{(X)}, V_{J-1,t}^{(Y)}\right\} &= \int_{-1/2}^{1/2} \left[\prod_{l=0}^{J-2} \widetilde{\mathcal{G}}(2^l f)\right] S_{XY}(f)\, df \\
&= \int_{-1/2}^{1/2} \left[\widetilde{\mathcal{G}}(2^{J-1} f) + \widetilde{\mathcal{H}}(2^{J-1} f)\right] \left[\prod_{l=0}^{J-2} \widetilde{\mathcal{G}}(2^l f)\right] S_{XY}(f)\, df \\
&= \int_{-1/2}^{1/2} \left[\widetilde{\mathcal{G}}_J(f) + \widetilde{\mathcal{H}}_J(f)\right] S_{XY}(f)\, df \\
&= \text{Cov}\left\{\overline{V}_{J,t}^{(X)}, \overline{V}_{J,t}^{(Y)}\right\} + \gamma_{XY}(\lambda_J).
\end{aligned}$$

The decomposition of covariance between $\{X_t, Y_t\}$ has now been established for a finite number of scales. From Lemma 1, as $J \to \infty$ the remaining covariance between the scaling coefficients goes to zero. Hence, the theorem is established.

□

If we think of $\{\tilde{g}_J\}$ as a low-pass filter covering the nominal frequency band $[-2^{-(J+1)}, 2^{-(J+1)}]$, this statement is intuitively plausible since the scaling filter $\{\tilde{g}_J\}$ is capturing smaller and smaller portions of the cross spectrum as $J \to \infty$.

1.2 Estimating the Wavelet Cross-Covariance

Suppose $X_0, \ldots, X_{N-1}$ and $Y_0, \ldots, Y_{N-1}$ can be regarded as realizations of portions of the processes $\{X_t\}$ and $\{Y_t\}$, whose d_Xth and d_Yth order backward differences form stationary Gaussian processes. As before, let $d = \max\{d_X, d_Y\}$.

Let $\widetilde{W}_{j,t} = \overline{W}_{j,t}$ for those indices t where $\widetilde{W}_{j,t}$ is unaffected by the boundary – this is true as long as $t \geq L_j - 1$. Thus, if $N \geq L_j$, we can define a biased estimator $\tilde{\gamma}_{\tau,XY}(\lambda_j)$ of the wavelet cross-covariance based upon the MODWT via

$$\tilde{\gamma}_{\tau,XY}(\lambda_j) \equiv \begin{cases} \widetilde{N}_j^{-1} \sum_{l=L_j-1}^{N-\tau-1} \widetilde{W}_{j,l}^{(X)} \widetilde{W}_{j,l+\tau}^{(Y)}, & \tau = 0, \ldots, \widetilde{N}_j - 1; \\ \widetilde{N}_j^{-1} \sum_{l=L_j-1-\tau}^{N-1} \widetilde{W}_{j,l}^{(X)} \widetilde{W}_{j,l+\tau}^{(Y)}, & \tau = -1, \ldots, -(\widetilde{N}_j - 1); \\ 0, & \text{otherwise}, \end{cases}$$

where $\widetilde{N}_j \equiv N - L_j + 1$. The bias is due to the denominator $1/\widetilde{N}_j$ remaining constant for all lags, though it disappears at lag $\tau = 0$. Let us now state the large sample properties of the wavelet covariance estimator; i.e., $\tilde{\gamma}_{XY}(\lambda_j) \equiv \tilde{\gamma}_{0,XY}(\lambda_j)$. We may generalize to the wavelet cross-covariance by simply replacing shifting $\{\overline{W}_{j,t}^{(Y)}\}$ with respect to $\{\overline{W}_{j,t}^{(X)}\}$ and appealing to the same result.

Theorem 2 *If $L > 2d$, and suppose $\{\overline{W}_{j,t}^{(X)}, \overline{W}_{j,t}^{(Y)}\}$ is a bivariate Gaussian stationary process with autospectra satisfying $\int_{-1/2}^{1/2} S_{j,X}^2(f) < \infty$ and $\int_{-1/2}^{1/2} S_{j,Y}^2(f) < \infty$, then*

$$\tilde{\gamma}_{XY}(\lambda_j) \sim N\left(\gamma_{XY}(\lambda_j), \widetilde{N}_j^{-1} S_{j,(XY)}(0)\right);$$

i.e., the MODWT-based estimator of wavelet covariance is normally distributed with mean $\gamma_{XY}(\lambda_j)$ and large sample variance $\widetilde{N}_j^{-1} S_{j,(XY)}(0)$. The quantity $S_{j,(XY)}(0)$ is the spectral density function for $\{\overline{W}_{j,t}^{(X)} \overline{W}_{j,t}^{(Y)}\}$ (the product of the scale λ_j MODWT coefficients) at zero frequency.

Proof of Theorem 2 With $L > 2d$, both series of MODWT coefficients $\{\overline{W}_{j,t}^{(X)}\}$ and $\{\overline{W}_{j,t}^{(Y)}\}$ have zero mean. Square integrability of the autospectra implies that $\{s_{j,\tau,X}\} \longleftrightarrow S_{j,X}(\cdot)$ and $\{s_{j,\tau,Y}\} \longleftrightarrow S_{j,Y}(\cdot)$; i.e., the autocovariance sequences and autospectra are Fourier transform pairs. Because $L > 2d$, the squared gain function for Daubechies wavelet filters guarantees we have $S_{j,X}(0) = 0 = \sum_{\tau=-\infty}^{\infty} s_{j,\tau,X}$. A similar statement holds for $\{\widetilde{W}_{j,t}^{(Y)}\}$ and, therefore, $\{s_{j,\tau,X}\}$ and $\{s_{j,\tau,Y}\}$ are absolutely summable.

Let $S_{j,XY}(f) \equiv \widetilde{\mathcal{H}}_j(f) S_{XY}(f)$ denote the MODWT filtered cross spectrum. From the magnitude squared coherence being bounded by unity, and using the Cauchy–Schwarz inequality, we know that

$$\begin{aligned} \int_{-1/2}^{1/2} |S_{j,XY}(f)|^2 \, df &\leq \int_{-1/2}^{1/2} S_{j,X}(f) S_{j,Y}(f) \, df \\ &\leq \left(\int_{-1/2}^{1/2} S_{j,X}^2(f) \, df \int_{-1/2}^{1/2} S_{j,Y}^2(f) \, df \right)^{1/2} < \infty. \end{aligned}$$

So the cross-covariance sequence and cross spectrum associated with scale λ_j are also a Fourier pair and, again, by using a Daubechies wavelet filters with $L > 2d$, we have $S_{j,XY}(0) = 0$. Therefore, the cross-covariance sequence for $\{\overline{W}_{j,t}^{(X)}, \overline{W}_{j,t}^{(Y)}\}$ is absolutely summable.

We first note that the MODWT estimate of the wavelet covariance $\tilde{\gamma}_{XY}(\lambda_j)$ is essentially a sample mean for the time series $\overline{W}_{j,t}^{(XY)} \equiv \overline{W}_{j,t}^{(X)}\overline{W}_{j,t}^{(Y)}$ (cf. Equation (1)). This process also has an absolutely summable cumulant sequence by Theorem 2.9.1 of Brillinger (1981, p. 38). Then Theorem 4.4.1 of Brillinger (1981, p. 94) tells us that $\tilde{\gamma}_{XY}(\lambda_j)$ is asymptotically normal with mean $\gamma_{XY}(\lambda_j)$ and large sample variance given by $\widetilde{N}_j^{-1} S_{j,(XY)}(0)$, where $S_{j,(XY)}(0)$ is the spectral density for $\overline{W}_{j,t}^{(XY)}$ evaluated at $f = 0$.

□

Since we are exclusively interested in Gaussian processes, $S_{j,(XY)}(0)$ may be re-expressed as a function of the auto and cross spectra of the wavelet coefficients $\{W_{j,l}^{(X)}\}$ and $\{W_{j,l}^{(Y)}\}$. The variance of the estimated MODWT wavelet covariance at scale λ_j can be computed directly via

$$
\begin{aligned}
\mathrm{Var}\{\tilde{\gamma}_{XY}(\lambda_j)\} &= \frac{1}{\widetilde{N}_j^2} \sum_{l=L_j-1}^{N-1} \sum_{m=L_j-1}^{N-1} \mathrm{Cov}\left\{\widetilde{W}_{j,l}^{(X)}\widetilde{W}_{j,l}^{(Y)}, \widetilde{W}_{j,m}^{(X)}\widetilde{W}_{j,m}^{(Y)}\right\} \\
&= \frac{1}{\widetilde{N}_j} \sum_{\tau=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \left(1 - \frac{|\tau|}{\widetilde{N}_j}\right) \mathrm{Cov}\left\{\widetilde{W}_{j,l}^{(X)}\widetilde{W}_{j,l}^{(Y)}, \widetilde{W}_{j,l+\tau}^{(X)}\widetilde{W}_{j,l+\tau}^{(Y)}\right\} \\
&\equiv \frac{1}{\widetilde{N}_j} \sum_{\tau=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \left(1 - \frac{|\tau|}{\widetilde{N}_j}\right) s_{j,\tau,XY}, \qquad (2)
\end{aligned}
$$

where $s_{j,\tau,XY}$ is the autocovariance sequence for the product of the scale λ_j MODWT coefficients with respect to $\{X_t\}$ and $\{Y_t\}$.

Using the Isserlis theorem and properties of the Fourier transform, the spectrum of $\{Z_t\}$ at $f = 0$ is $S_Z(0) = \int_{-1/2}^{1/2} S_U(f)S_V(f)\,df + \int_{-1/2}^{1/2} S_{UV}^2(f)\,df$ (Whitcher 1998). Since we have the Fourier relationship $\{s_{\tau,Z}\} \longleftrightarrow S_Z(\cdot)$, we necessarily have $S_Z(0) = \sum_{\tau=-\infty}^{\infty} s_{\tau,Z}$, when $f = 0$. Re-examining Equation (2) and utilizing Cesàro summability (Titchmarsh 1939, p. 411), we can say

$$
\begin{aligned}
\lim_{\widetilde{N}_j \to \infty} \widetilde{N}_j \,\mathrm{Var}\{\tilde{\gamma}_{XY}(\lambda_j)\} &= \lim_{\widetilde{N}_j \to \infty} \sum_{\tau=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \left(1 - \frac{|\tau|}{\widetilde{N}_j}\right) s_{j,\tau,XY} \\
&= \sum_{\tau=-\infty}^{\infty} s_{j,\tau,XY} = S_{j,(XY)}(0),
\end{aligned}
$$

where

$$
S_{j,(XY)}(0) = \int_{-1/2}^{1/2} S_{j,X}(f)S_{j,Y}(f)\,df + \int_{-1/2}^{1/2} S_{j,XY}^2(f)\,df
$$

1.3 Wavelet Cross-Correlation

We can define the wavelet cross-correlation for scale λ_j and lag τ as

$$\rho_{\tau,XY}(\lambda_j) \equiv \frac{\mathrm{Cov}\left\{\overline{W}_{j,t}^{(X)}, \overline{W}_{j,t+\tau}^{(Y)}\right\}}{\left(\mathrm{Var}\left\{\overline{W}_{j,t}^{(X)}\right\} \mathrm{Var}\left\{\overline{W}_{j,t+\tau}^{(Y)}\right\}\right)^{1/2}} = \frac{\gamma_{\tau,XY}(\lambda_j)}{\nu_X(\lambda_j)\nu_Y(\lambda_j)}.$$

Since this is just a correlation coefficient between two random variables, $-1 \le \rho_{\tau,XY}(\lambda_j) \le 1$ for all τ, j. The wavelet cross-correlation is roughly analogous to its Fourier counterpart – the magnitude squared coherence – but it is related to bands of frequencies (scales). Just as the cross-correlation is used to determine lead/lag relationships between two processes, the wavelet cross-correlation will provide a lead/lag relationship on a scale by scale basis.

1.4 Estimating the Wavelet Cross-Correlation

Since the wavelet cross-correlation is simply made up of the wavelet cross-covariance for $\{X_t, Y_t\}$ and wavelet variances for $\{X_t\}$ and $\{Y_t\}$, the MODWT estimator of the wavelet cross-correlation is simply

$$\tilde{\rho}_{\tau,XY}(\lambda_j) \equiv \frac{\tilde{\gamma}_{\tau,XY}(\lambda_j)}{\tilde{\nu}_X(\lambda_j)\tilde{\nu}_Y(\lambda_j)}, \tag{3}$$

where $\tilde{\gamma}_{\tau,XY}(\lambda_j)$ is the wavelet covariance, and $\tilde{\nu}_X^2(\lambda_j)$ and $\tilde{\nu}_Y^2(\lambda_j)$ are the wavelet variances. When $\tau = 0$ we obtain the MODWT estimator of the wavelet correlation between $\{X_t, Y_t\}$, denoted as $\tilde{\rho}_{XY}(\lambda_j)$ for simplicity.

Large sample theory for the cross-correlation is more difficult to come by than for the cross-covariance. Brillinger (1979) constructed approximate confidence intervals for the auto and cross-correlation sequences of bivariate stationary time series. We use his technique to establish a central limit theorem for the MODWT estimated wavelet cross-correlation. To simplify notation the following gives a central limit theorem for the wavelet correlation ($\tau = 0$) but easily generalizes to arbitrary lags.

Theorem 3 *Let $L > 2d$, and suppose $\{\overline{W}_{j,t}^{(X)}, \overline{W}_{j,t}^{(Y)}\}$ is a bivariate Gaussian stationary process with square integrable autospectra, then*

$$\tilde{\rho}_{XY}(\lambda_j) \sim N\left(\rho_{XY}(\lambda_j), \tilde{N}_j^{-1} R_j\right);$$

i.e., the MODWT-based estimator $\tilde{\rho}_{XY}(\lambda_j)$ of the wavelet correlation is asymptotically normally *distributed with mean $\rho_{XY}(\lambda_j)$ and large sample variance given by*

$$\begin{aligned} R_j \equiv \mathrm{Var}\{\tilde{\rho}_{XY}(\lambda_j)\} \;\approx\; & \frac{1}{\widetilde{N}_j} \sum_{\tau=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \{\, \rho_{\tau,X}(\lambda_j)\rho_{\tau,Y}(\lambda_j) + \rho_{\tau,XY}(\lambda_j)\rho_{\tau,YX}(\lambda_j) \\ & - 2\rho_{0,XY}(\lambda_j)[\rho_{\tau,X}(\lambda_j)\rho_{\tau,YX}(\lambda_j) + \rho_{\tau,Y}(\lambda_j)\rho_{\tau,YX}(\lambda_j)] \\ & + \rho^2_{0,XY}(\lambda_j)[\tfrac{1}{2}\rho^2_{\tau,X}(\lambda_j) + \rho^2_{\tau,XY}(\lambda_j) + \tfrac{1}{2}\rho^2_{\tau,Y}(\lambda_j)]\,\}, \end{aligned}$$

where $\rho_{\tau,X}(\lambda_j) \equiv E\{\overline{W}^{(X)}_{j,t}\overline{W}^{(X)}_{j,t+|\tau|}\}/[2\lambda_j\nu^2_X(\lambda_j)]$ is the scale λ_j wavelet autocorrelation for the process $\{X_t\}$.

Proof of Theorem 2 Since $L > 2d$, we have that both sets of wavelet coefficients $\{\widetilde{W}^{(X)}_{j,t}\}$ and $\{\widetilde{W}^{(Y)}_{j,t}\}$ have mean zero. Let us define $A_{j,t} \equiv [\widetilde{W}^{(X)}_{j,t}]^2$, $B_{j,t} \equiv [\widetilde{W}^{(Y)}_{j,t}]^2$, and $C_{j,t} \equiv \widetilde{W}^{(X)}_{j,t}\widetilde{W}^{(Y)}_{j,t}$, and subsequently define their sample means

$$\begin{aligned} \overline{A}_j &\equiv \frac{1}{\widetilde{N}_j}\sum_{t=L_j-1}^{N-1} A_{j,t} = \tilde{\nu}^2_X(\lambda_j), \\ \overline{B}_j &\equiv \frac{1}{\widetilde{N}_j}\sum_{t=L_j-1}^{N-1} B_{j,t} = \tilde{\nu}^2_Y(\lambda_j), \quad \text{and} \\ \overline{C}_j &\equiv \frac{1}{\widetilde{N}_j}\sum_{t=L_j-1}^{N-1} C_{j,t} = \tilde{\gamma}_{XY}(\lambda_j). \end{aligned}$$

The vector-valued process $\{A_{j,t}, B_{j,t}, C_{j,t}\}$ has an absolutely summable joint cumulant sequence by Theorem 2.9.1 of Brillinger (1981, p. 38). Hence, from Theorem 4.4.1 of Brillinger (1981, p. 94) the vector of sample means $\{\overline{A}_j, \overline{B}_j, \overline{C}_j\}$ are asymptotically normally distributed with mean vector $\{\nu^2_X(\lambda_j), \nu^2_Y(\lambda_j), \gamma_{XY}(\lambda_j)\}$, and large sample variance given by $\widetilde{N}_j^{-1}\mathbf{S}_{j,ABC}(0)$, where $\mathbf{S}_{j,ABC}(\cdot)$ is the 3×3 spectral matrix for $\{A_{j,t}, B_{j,t}, C_{j,t}\}$.

The MODWT estimator of the wavelet correlation $\tilde{\rho}_{XY}(\lambda_j)$ is essentially a function of these sample means $g(\overline{A}_j, \overline{B}_j, \overline{C}_j)$, where $g(x,y,z) \equiv z/\sqrt{xy}$. Appealing to Mann and Wald (1943), we have that $\tilde{\rho}_{XY}(\lambda_j)$ is asymptotically normally distributed with mean $\rho_{XY}(\lambda_j)$ and large sample variance

$$\widetilde{N}_j^{-1}\,\dot{g}\left(\nu^2_X(\lambda_j), \nu^2_Y(\lambda_j), \gamma_{XY}(\lambda_j)\right)^T\,\mathbf{S}_{j,ABC}(0)\,\dot{g}\left(\nu^2_X(\lambda_j), \nu^2_Y(\lambda_j), \gamma_{XY}(\lambda_j)\right) \tag{4}$$

where $\dot{g}(\cdot,\cdot,\cdot)$ is the gradient of $g(\cdot,\cdot,\cdot)$. Now let us re-express Equation (4) into the desired result using the fact that we are only interested in Gaussian processes. Because we are evaluating $\mathbf{S}_{j,ABC}(\cdot)$ at $f = 0$, it is in fact a symmetric matrix of the form

$$\mathbf{S}_{j,ABC}(0) = \begin{bmatrix} S_{j,AA}(0) & S_{j,AB}(0) & S_{j,AC}(0) \\ S_{j,AB}(0) & S_{j,BB}(0) & S_{j,BC}(0) \\ S_{j,AC}(0) & S_{j,BC}(0) & S_{j,CC}(0) \end{bmatrix},$$

where the elements of the matrix are

$$
\begin{aligned}
S_{j,AA}(0) &= 2\int_{-1/2}^{1/2} S_{j,X}^2(f)\,df, \quad S_{j,BB}(0) = 2\int_{-1/2}^{1/2} S_{j,Y}^2(f)\,df,\\
S_{j,CC}(0) &= \int_{-1/2}^{1/2} S_{j,X}(f)S_{j,Y}(f)\,df + \int_{-1/2}^{1/2} S_{j,XY}^2(f)\,df,\\
S_{j,AB}(0) &= 2\int_{-1/2}^{1/2} S_{j,XY}(f)S_{j,YX}(f)\,df,\\
S_{j,AC}(0) &= 2\int_{-1/2}^{1/2} S_{j,X}(f)S_{j,YX}(f)\,df, \quad \text{and}\\
S_{j,BC}(0) &= 2\int_{-1/2}^{1/2} S_{j,Y}(f)S_{j,YX}(f)\,df.
\end{aligned}
$$

The gradient is explicitly given by

$$
\begin{aligned}
&\dot{g}\left(\nu_X^2(\lambda_j), \nu_Y^2(\lambda_j), \gamma_{XY}(\lambda_j)\right) =\\
&\qquad \left[-\frac{\gamma_{XY}(\lambda_j)}{2\nu_X^2(\lambda_j)\sqrt{\nu_X^2(\lambda_j)\nu_Y^2(\lambda_j)}} \;\; -\frac{\gamma_{XY}(\lambda_j)}{2\nu_Y^2(\lambda_j)\sqrt{\nu_X^2(\lambda_j)\nu_Y^2(\lambda_j)}} \;\; \frac{1}{\sqrt{\nu_X^2(\lambda_j)\nu_Y^2(\lambda_j)}} \right]^T,
\end{aligned}
$$

and, through matrix multiplication and application of Parseval's relation to each auto and cross spectra in $\mathbf{S}_{j,ABC}(0)$, we may express Equation (4) as

$$
\begin{aligned}
\frac{1}{\widetilde{N}_j} \sum_{\tau=-(\widetilde{N}_j-1)}^{\widetilde{N}_j-1} \Bigg\{ & \frac{\gamma_{XY}^2(\lambda_j)}{4\nu_X^6(\lambda_j)\nu_Y^2(\lambda_j)}\, 2s_{j,\tau,X}^2 + \frac{\gamma_{XY}^2(\lambda_j)}{2\nu_X^4(\lambda_j)\nu_Y^4(\lambda_j)}\, 2C_{j,\tau,XY}C_{j,\tau,YX}\\
&+ \frac{\gamma_{XY}^2(\lambda_j)}{4\nu_X^2(\lambda_j)\nu_Y^6(\lambda_j)}\, 2s_{j,\tau,Y}^2 + \frac{1}{\nu_X^2(\lambda_j)\nu_Y^2(\lambda_j)} \left(s_{j,\tau,X}s_{j,\tau,Y} + C_{j,\tau,XY}^2\right)\\
&- \frac{\gamma_{XY}(\lambda_j)}{\nu_X^4(\lambda_j)\nu_Y^2(\lambda_j)}\, 2s_{j,\tau,X}C_{j,\tau,YX} - \frac{\gamma_{XY}(\lambda_j)}{\nu_X^2(\lambda_j)\nu_Y^4(\lambda_j)}\, 2s_{j,\tau,Y}C_{j,\tau,YX} \Bigg\}.
\end{aligned}
$$

Each of the autocovariance terms are equivalent to the wavelet autocovariance for scale λ_j (defined by letting $X_t = Y_t$ in Equation (1)) and each cross-covariance term is equivalent to the wavelet cross-covariance for scale λ_j. This yields the desired result.

□

References

Brillinger, D. R. (1979). Confidence intervals for the crosscovariance function. In *Mathematical Statistics*, Volume 5 of *Selecta Statistica Canadiana*, pp. 1–16. Hamilton, Ontario: McMaster University Printing Services.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory.* Holden-Day Series in Time Series Analysis. San Francisco: Holden-Day. Expanded edition.

Mann, H. B. and A. Wald (1943). On stochastic limit and order relationships. *The Annals of Mathematical Statistics 14*, 217–226.

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

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

Titchmarsh, E. C. (1939). *The Theory of Functions* (2 ed.). Oxford: Oxford University Press.

Whitcher, B. (1998). *Assessing Nonstationary Time Series Using Wavelets*. Ph. D. thesis, University of Washington.

Whitcher, B., P. Guttorp, and D. B. Percival (1999). Wavelet analysis of covariance with application to atmospheric time series. Submitted to the *Journal of Geophysical Research – Atmospheres*.