Asymptotic Decorrelation of Discrete Wavelet Packet Transform of Generalized Long-memory Stochastic Volatility

Alex C. Gonzaga
Department of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Manila

We derive the asymptotic properties of discrete wavelet packet transform (DWPT) of generalized long-memory stochastic volatility (GLMSV) model, a relatively general model of stochastic volatility that accounts for persistent (or long-memory) and seasonal (or cyclic) behavior at several frequencies. We derive the rates of convergence to zero of between-scale and within-scale wavelet packet coefficients at different subbands. Wavelet packet coefficients in the same subband can be shown to be approximately uncorrelated by appropriate choice of basis vectors using a white noise test. These results may be used to simplify the variance-covariance matrix into a diagonalized matrix, whose diagonal elements have the least distinct variances to compute.

Keywords: discrete wavelet packet transform, generalized long-memory stochastic volatility, asymptotic decorrelation

1. Introduction

Long-memory or long-range dependence has been recognized in the time-varying volatility of many time series. This characteristic may not be modeled parsimoniously using the autoregressive conditional heteroscedasticity (ARCH) or stochastic volatility (SV) models (see e.g. (Hurvich and Ray, 2003), (Deo and Hurvich, 2001), (Jensen, 2004), (Bollerslev and Mikkelsen, 1996), (Gonzaga and Hauser, 2011), and (Gonzaga, 2013)). Stochastic volatility (SV) models are both time-varying and stochastic, and serve as the closest equivalent to the autoregressive (AR) process or autoregressive moving average (ARMA) process in the second moment (Wang, 2009)). In this context, the long-memory stochastic volatility (LMSV) model was proposed by (Breidt, et al., 1998) and (Harvey, 1998) by replacing the autoregressive moving average (ARMA) in the SV model by the fractionally integrated (FI) process. However, it does not account for periodic or seasonal components. This cyclical behavior is often seen as prominent peaks in the periodogram of the series, which is commonly observed in the volatility of financial time series such as the intraday data of stock prices (see e.g. (Andersen and Bollerslev, 1997) and (Bisaglia, et al., 2003)). Modelling and forecasting volatility dynamics may be aversely affected when these persistence and periodicities are not accounted for. Hence, we consider the generalized long-memory stochastic volatility (GLMSV), which is known to account for k persistent periodicities in the volatility series (Gonzaga and Hauser, 2011).

Asymptotic decorrelation of wavelet coefficients of long-memory processes has been considered in literature (see e.g. (Gonzaga and Kawanaka, 2006), (Gonzaga and Hauser, 2011), (Craigmile and Percival, 2005), (Fan, 2003) and (Yu, 2013)). The wavelet coefficients of a fractionally integrated (FI) process have been conveniently assumed independent within and across scales since the spectral density of an FI process is approximately flat outside the zero frequency. However, owing to varying locations of the poles in models with seasonality at certain points, within-scale correlations of wavelet coefficients may vary significantly depending on the locations of these poles. For this purpose, we exploit the ability of the discrete wavelet packet transform (DWPT) to approximately decorrelate the generalized long-memory stochastic volatility (GLMSV) model, a relatively general model of stochastic volatility that accounts for persistent (or long-memory) and seasonal (or cyclic) behavior at several frequencies. (Gonzaga and Hauser, 2011) previously determined the asymptotic properties of the discrete wavelet transform (DWT) of GLMSV model. On the other hand, (Yu, 2013) derived the order of convergence of DWPT coefficients of seasonal long-memory process.

In this paper, we derive the asymptotic properties of discrete wavelet packet transform (DWPT) of generalized long-memory stochastic volatility (GLMSV) model. We derive the rates of convergence to zero of between-scale and within-scale wavelet packet coefficients from different subbands. Wavelet packet coefficients in the same subband can be shown to be approximately uncorrelated by appropriate choice of basis vectors using a white noise test. These results may be used to simplify the variance-covariance matrix into a diagonalized matrix, whose diagonal elements have the least distinct variances to compute.

This paper is organized as follows. In Section 1, we define and discuss the properties of the generalized long-memory stochastic volatility (GLMSV) process. In Section 2, we present the covariance structure of the discrete wavelet packet transform (DWPT) of GLMSV. In Section 3, we derive the asymptotic properties of the DWPT of GLMSV. Finally, some concluding remarks are given in Section 4.

2. The GLMSV Model

The generalized long-memory stochastic volatility (GLMSV) model provides a general framework in modeling volatility dynamics incorporating persistence (or long-memory) and multiple seasonalities (or periodicities). We consider the stochastic volatility model given by

$$r_t = \sigma \exp\{X_t/2\} e_t, \tag{1}$$

with $\sigma > 0$, $\{e_t\}$ *are iid shocks with zero mean and unit variance, and* $\{X_t\}$ *is a* k-GARMA process, which is independent of $\{e_t\}$. From (Woodward, et al., 1998), a k-GARMA(p,d,u,q) process {Xt} is given by

$$\Phi(B)\prod_{i=1}^{k}(1-2u_iB+B^2)^{d_i}X_t = \Theta(B)\varepsilon_t, \tag{2}$$

where B denotes the backshift operator, $\{\varepsilon_t\}\sim iid\ N(0,\sigma_\varepsilon^2)$, $(B) = 1-\varphi_1B^1 - \ldots - \varphi_pB^p$ and $\Theta(B) = 1 + \theta_1B^1 + \ldots + \theta_qB^q$ are polynomials of order p and q, respectively, with all roots inside the unit circle; d and u are vectors of length k, with $d_i \neq 0$ and distinct u_i, with $|u_i| \leq 1$, $i=1,\ldots,k$. The components of vector d are called the long-memory parameters, and the frequencies corresponding to the seasonal or periodic components, $v_i = \frac{\cos^{-1}(u_i)}{2\pi} \in [0,0.5]$, i =1, 2, …, k, are called the Gegenbauer frequencies. From (Woodward, et al., 1998) and (Giraitis and Leipus, 1995), a k-GARMA(p,d,u,q) process is causal and invertible if for i =1,…, k.

$$|d_i| < \begin{cases} \frac{1}{2}, & 0 < v_i < \frac{1}{2} \\ \frac{1}{4}, & v_i = 0 \text{ or } \frac{1}{2} \end{cases}, \tag{3}$$

and the spectral density is given by

$$S_X(f) = \sigma_\varepsilon^2 \left|\frac{\Theta(e^{-i2\pi f})}{\Phi(e^{-i2\pi f})}\right|^2 \prod_{i=1}^{k}\left|2\left(\cos(2\pi f)-u_i\right)\right|^{-2d_i},\ \ f \in (-0.5,0.5]. \tag{4}$$

Let $Y_t = \log r_t^2$ be the log squared process. Hence, we have

$$Y_t = \mu + X_t + \eta_t, \tag{5}$$

where $\mu = \log \sigma^2 + E(\log e_t^2)$ and $\eta_t = \log e_t^2 - E(\log e_t^2)$ is an iid mean-zero process with finite variance σ_η^2 independent of $\{X_t\}$.

The autocovariance function of $\{Y_t\}$ is simply the sum of the covariances of the long-memory process and the noise given by

$$\gamma_Y(s) = \gamma_X(s) + \sigma_\eta^2 I_{\{s=0\}}\,. \tag{6}$$

Hence, the spectral density is

$$S_Y(f) = S_X(f) + \sigma_\eta^2, \tag{7}$$

where the constant σ_η^2 is the spectral density of the iid process $\{\eta_t\}$, that is $S_\eta(f) = \sigma_\eta^2$.

3. Covariances of DWPT Coefficients

We present some properties of the discrete wavelet packet transform (DWPT) coefficients of the log-squared GLMSV process. The ensuing nomenclature is based on (Percival and Walden, 2000).

Let $\{h_{1,l}\}_{l=0}^{L-1}$ denote a Daubechies' compactly supported wavelet filter of even length L. The scaling filter $\{g_{1,l}\}_{l=0}^{L-1}$ is defined by $g_{1,l} = (-1)^{l+1} h_{1,L-l-l}$.. From (Daubechies, 1992), the filter $\{h_{1,l}\}$ has squared gain function defined by

$$|H_{1,L}(f)|^2 = 2\sin^L(\pi f)\sum^{-}\begin{pmatrix} - - \;\; + \\ \\ \end{pmatrix}\cos^l(\pi f) \tag{8}$$

such that

$$|H_{1,L}(f)|^2 + |G_{1,L}(f)|^2 = 2 \text{ and } |G_{1,L}(f)|^2 = H_{1,L}(0.5 - f)|^2 \tag{9}$$

for $f\,| \leq 1/2$.

From (Percival and Walden, 2000), we write the DWPT coefficients $\{D_{j,n,t} \mid j = 0, ..., J, n = 0, ... 2^j - 1, t = 0, ... N2^{-j} - 1\}$ of the signal $\{Y_t\}_{t=0}^{N-1}$ in the form

$$D_{j,n,t} = \sum_{l=0}^{L_j-1} u_{j,n,l} Y_{2^j[t+1]-1-l \bmod M}, \tag{10}$$

where $L_j = (2^j - 1)(L-1) + 1$, and $\{u_{j,n,l}\}$ is the filter corresponding to the node (j,n) in the wavelet packet table. The filter $\{u_{j,n,l}\}$ can be computed from $\{h_{1,\ell}\}$ and

$\{ g_{1,\ell} \}$ by letting $u_{1,0,l} = g_{1,l}$ and $u_{1,1,l} = h_{1,l}$. Then for $j > 1$ and for each node (j,n), we recursively obtain $u_{j,n,l}$ using the equation

$$u_{j,n,l} = \sum_{l=0}^{L-1} u_{n,k} u_{j-1,\lfloor n/2 \rfloor, l-2^{j-1}k}, \quad 1 = 0, ..., L_j - 1, \tag{11}$$

where

$$u_{n,l} = \begin{cases} g_{1,l}, & \text{if n mod(4) = 0 or 3} \\ h_{1,l}, & \text{if n mod(4) = 1 or 2} \end{cases} \tag{12}$$

For each node (j,n) of the wavelet packet table, we consider cj,n a vector whose components are 0's and 1's as defined in (Percival and Walden, 2000), p. 215. We then write the transfer function of $\{u_{j,n,l}\}$ as

$$U_{j,n}(f) = \prod_{m=0}^{j-1} M_{c_{j,n,m}}(2^m f), \tag{13}$$

where $\mathbf{c}_{j,n,m}$ is the m^{th} element of $\mathbf{c}_{j,n}$, and $M_0(f) = G_{1,L}(f)$ and $M_1(f) = H_{1,L}(f)$.

From (10), the covariance of the DWPT coefficients of $\{Y_t\}_{t=0}^{N-1}$ is given by

$$\text{cov}(D_{j,n,t}{}^{Y}, D_{j',n',t'}{}^{Y}) = \sum_{\ell=0}^{L_j-1} \sum_{\ell'=0}^{L_{j'}-1} u_{j,n,\ell} u_{j',n',\ell'} \gamma_Y (2^j(t+1) - 2^{j'}(t'+1) + \ell' - \ell), \tag{14}$$

where $\gamma_Y(s)$ is the autocovariance function of Y_t at lag s. As in (Percival and Walden, 2000, 348) (14) may be written as

$$\text{cov}(D_{j,n,t}{}^{Y}, D_{j',n',t'}{}^{Y}) = \int_{-1/2}^{1/2} e^{i2\pi f(2^{j'}(t'+1) - 2^j(t+1))} U_{j,L}(f) \overline{U}_{j',L}(f) S_Y(f) df, \tag{15}$$

where $\overline{U}$ denotes the complex conjugate of U. From (15), for $j=j'$, the within-scale autocovariance function of nonboundary DWPT coefficients is given by

$$\text{cov}(D^Y_{j,n,t}, D^Y_{j,n,(t+s)}) = \int_{-1/2}^{1/2} e^{i2\pi f 2^j s} \, |U_{jn}(f)|^2 \, S_Y(f) df \tag{16}$$

where $U_{j,n}(f)$ is the transfer function of the filter $\{u_{j,n,1}\}$. $U_{j,n}(f)$ depends only on $G(\cdot)$ and $H(\cdot)$ such that the squared gain function $|U_{j,n}(f)|^2$ is nominally bandpass over the frequency interval $I_{j,n} = \left(\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}} \right]$. Since $S_Y(f)$ have poles at some frequencies between 0 and 0.5 and $|U_{j,n}(f)|^2$ is bandpass on the interval $I_{j,n}$, $n=0,\ldots,2_j-1$. Appropriate choice of the nodes (j,n) will ensure that (16) is close to zero, thereby decorrelating the wavelet packet coefficients.

We decorrelate the wavelet packet coefficients in the same subband, by choosing the wavelet filters $\{u_{j,n,1}\}$ appropriately. For this purpose, we consider the algorithm proposed in (Percival, et al., 2000), which successively selects the DWPT coefficients at each level by performing a white noise test on wavelet packet subbands at each level. For our purpose, we use the Portmanteau test with test statistic given in (Percival, et al., 2000) in the form

$$Q = N_j(N_j + 2)\sum_{s=1}^{K} \frac{\hat{\rho}^2(s)}{r-s}, \tag{17}$$

where N_j is the length of the vector $D_{j,n}$, $K=\max\{2,\min\{20,N_j/10\}\}$, the summation is the weighted sum of the first K sample autocorrelations given by

$$\hat{\rho}(s) = \frac{\mathrm{cov}(D_{j,n,t}, D_{j,n,t+s})}{\mathrm{var}(D_{j,n,t})}. \tag{18}$$

We reject white noise if $P(\chi_m^2 > Q) < \alpha$, for some value of the level of significance α.

We apply this white noise test to obtain a decorrelated wavelet packet coefficients as outlined in (Percival, et al., 2000). Suppose that $\boldsymbol{D}_{j,n}$, $n = 0,\ldots,2^j-1$, are the vectors of wavelet coefficients in the jth row of the wavelet packet table. We define $\mathbf{D}_{0,0}$ to be the given input signal. For $j < J$, starting with $j=1$, we test the vector $\boldsymbol{D}_{j,n}$ for white noise. If the test fails to reject, we retain $\boldsymbol{D}_{j,n}$. If the test rejects, we split $\boldsymbol{D}_{j,n}$ into $\boldsymbol{D}_{j+1,2n}$ and $D_{j+1,2n+1}$, and test both the resulting subbands for white noise. We repeat this process until $j = J$ in which we retain $\boldsymbol{D}_{J,n}$. We denote the resulting vector of DWPT coefficients by

$$\mathbf{D} = (\boldsymbol{D}_{j,n}, (j,n)$$

$$\mathbf{D} = (\boldsymbol{D}_{j,n}, (j,n) \in \mathrm{B}), \tag{19}$$

which is approximately uncorrelated.

4. Asymptotic Decorrelation of DWPT coefficients

In the following theorem, we show that the autocovariance function, $\mathrm{cov}(D_{j,n,t}^{\ Y}, D_{j',n',t'}^{\ Y})$, is approximately zero when $j \neq j'$ for sufficiently large L. This is borne out of the fact that the squared gain function $|U_{j,n}(f)|^2$ is nominally bandpass over the frequency interval $I_{j,n} = \left(\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}} \right]$. These passbands are disjoint when $j \neq j'$. Hence, (15) goes to zero as L goes to infinity. This approximation depends only on the length of the wavelet filter L, and not on the length of the series.

Theorem 1. Let $\{D_{j,n,t}{}^{Y}\}$ be the DWPT of the log squared process $\{Y_t\}_{t=0}^{N-1}$, then for $j \neq j'$ the autocovariance of between-scale wavelet packet coefficients satisfies

$$\left|\mathrm{cov}(D_{j,n,t}{}^{Y}, D_{j',n',t'}{}^{Y})\right|^2 = O\left(\frac{1}{L^{1/2}}\right). \tag{20}$$

Proof:

Let $S_Y(f)$ be the spectral density function of the log squared process. Without loss of generality, assume that $j > j' \geq 1$. From (15), we have

$$\left|\mathrm{cov}(D_{j,n,t}{}^{Y}, D_{j',n',t'}{}^{Y})\right|^2 \leq C_0\left(\int_0^{0.5} |U_{j,n}(f)||U_{j',n'}(f)|S_Y(f)df\right)^2, \tag{21}$$

for some constant C_0 independent of L. (From now on, we denote by C_i (i =1,2,…) a constant independent of L.) Since $j > j' \geq 1$, then the parent node of $\boldsymbol{c}_{j',n'}$ differs in at least one component with the vector $\boldsymbol{c}_{j,n}$. The product representations in (13) for $U_{j,n}$ and $U_{j',n'}$ differ in at least one factor, such that one factor is H(·) and another factor is G(·). Hence, for some m^*, $|H(2^{m^*}f)||G(2^{m^*}f)|$ is a factor of $|U_{j,n'}(f)||U_{j',n'}(f)|$. Now, let

$$W = \frac{|U_{j,n}(f)|^2|U_{j',n'}(f)|^2}{|H(2^{m^*}f)|^2|G(2^{m^*}f)|^2}. \tag{22}$$

Since $m^* \leq j' < j$, then $|M_{c_{j,n}}(2^{m^*+1})|^2 \ldots M_{c_{j,n}}(2^j f)|^2$ is a factor of W, where $c_{j,n}$ is either 0 or 1.

By Cauchy-Schwarz inequality, we have

$$\left|\mathrm{cov}(D_{j,n,t}{}^{Y}, D_{j',n',t'}{}^{Y})\right|^2 \leq$$

$$C_0\left(\int_0^{0.5} |\mathrm{H}(2^{m^*}f)|^2\,|\mathrm{G}(2^{m^*}f)|^2\,|M_{c_{j,n}}(2^{m^*+1}f)|^2 \cdots |M_{c_{j,n}}\left(2^j f\right)|^2\,S_Y(f)df\right)$$

$$x\left(\int_0^{0.5} W^* S_Y(f)df\right), \tag{23}$$

where W^* is a factor of W excluding $|M_{c_{j,n}}(2^{m^*+1})|^2 \ldots M_{c_{j,n}}(2^j f)|^2$. So it is a product of $(j'-m^*-1)$ factors of the form $|M_{c_{j,n}}(\cdot)|^2$. From (9), $0 \leq W^* \leq 2^{j'-m^*-1}$, and so the second factor in (23) is finite, that is

$$\int_0^{0.5} W^* S_Y(f)df \le 2^{j'-m^*-1} Var(Y_t) = 2^{j'-m^*-1}\left(Var\left(X_t\right)+\sigma_\eta^2\right) < \infty . \qquad (24)$$

We now show that the first factor in (23) converges to 0 as L → ∞, and obtain its order of convergence. By partitioning the interval of integration, we get

$$\begin{aligned}
&\int_0^{0.5} |\mathrm{H}(2^{m^*} f)|^2\, |\mathrm{G}(2^{m^*} f)|^2\, |M_{c_{j,n}}(2^{m^*+1} f)|^2 \cdots |M_{c_{j,n}}\left(2^j f\right)|^2\, S_Y(f)df \\
&\le \mathrm{C1} \int_0^{2^{-m^*-2}} |H(2^{m^*} f)|^2\, S_Y(f)df \qquad (25)\\
&+ \mathrm{C2} \sum_{r=m^*}^{j} \int_{2^{-r-2}}^{2^{-r-1}} |G_{1,L}\left(2^{m^*} f\right)|^2 |M_{c_{j,n}}(2^{m^*+1} f)|^2 \cdots |M_{c_{j,n}}\left(2^j f\right)|^2\, S_Y(f)df .
\end{aligned}$$

From (Fan, 2003), for $f \in [0,2^{-m^*-2})$

$$|H_{1,L}(2^{m^*} f)|^2 = \cos(2^{m^*+1}\pi f)\sum_{k=L}^{\infty}\frac{(4k-1)!!}{(4k)!!}\sin^{4k}(2^{m^*+1}\pi f) . \qquad (26)$$

Hence, the first integral in (25) becomes

$$\begin{aligned}
&\int_0^{2^{-m^*-2}} |H(2^{m^*} f)|^2\, S_Y(f)df \\
&= \lim_{\delta\to 0} C_3 \sum_{k=L}^{\infty}\frac{(4k-1)!!}{(4k)!!}\int_0^{2^{-m^*-2}-\delta} \cos\left(2^{m^*+1}\pi f\right)\sin^{4k}\left(2^{m^*+1}\pi f\right) S_Y\left(f\right) df , \qquad (27)
\end{aligned}$$

where by Lemma 4.2 in (Fan, 2003), $\frac{(4k-1)!!}{(4k)!!} \approx k^{-1/2}$ for large k. Hence, if each Gegenbauer frequency $v_k \notin \left[0, 2^{-m^*-2}\right)$, then $S_Y(f)$ is finite on $[0,2^{-m^*-2}]$ such that

$$\begin{aligned}
&\int_0^{2^{-m^*-2}} |H(2^{m^*} f)|^2\, S_Y(f)df \\
&= \lim_{\delta\to 0} C_4 \sum_{k=L}^{\infty}\frac{1}{k^{1/2}}\int_0^{2^{-m^*-2}-\delta} \cos\left(2^{m^*+1}\pi f\right)\sin^{4k}\left(2^{m^*+1}\pi f\right) df \qquad (28)\\
&= C_5 \sum_{k=L}^{\infty}\frac{\sin^{4k+1}\left(2^{-1}\pi\right)}{k^{1/2}(4k+1)} \le C_5 \sum_{k=L}^{\infty}\frac{1}{k^{1/2}(4k+1)} = O\left(\frac{1}{L^{3/2}}\right).
\end{aligned}$$

On the other hand, if there is some Gegenbauer frequency $v_i \in [0,2^{-m^*-2}]$, since $S_Y(f)$ is integrable on $[0,2^{-m^*-2}]$ and $sin^{4k}(2^{m^*+1}\pi f)$ is continuous and increasing on $[0,2^{-m^*-2}]$, the mean value theorem gives us

$$\int_0^{2^{-m^*-2}} |H(2^{m^*} f)|^2 S_Y(f)df$$

$$= \lim_{\delta \to 0} C_7 \sum_{k=L}^{\infty} \frac{1}{k^{1/2}} \left\{ \sin^{4k}\left(2^{-1}\pi\right) \right\} = C_8 \sum_{k=L}^{\infty} \frac{1}{k^{1/2}} = O\left(\frac{1}{L^{1/2}}\right). \quad (29)$$

Now, from (9) and (26), we have for $f \in (2^{-r-2}, 2^{-r-1})$

$$|M_{c_{j,n}}(2^j f)|^2 \leq |G_{1,L}(2^r f)| = -\cos(2^{r+1}\pi f) \sum_{k=L}^{\infty} \frac{(4k-1)!!}{(4k)!!} \sin^{4k}(2^{r+1}\pi f), \quad (30)$$

where $r = m^*, \ldots, j$. Hence, the second term in (25) satisfies

$$\sum_{r=m^*}^{j} \int_{2^{-r-2}}^{2^{-r-1}} |G_{1,L}\left(2^{m^*} f\right)|^2 |M_{c_{j,n}}(2^{m^*+1} f)|^2 \cdots |M_{c_{j,n}}\left(2^j f\right)|^2 S_Y(f)df$$

$$\leq \lim_{\delta \to 0^+} C_9 \sum_{k=L}^{\infty} \frac{1}{k^{1/2}} \int_{2^{-r-2}+\delta}^{2^{-r-1}} \cos\left(2^{r+1}\pi f\right) \sin^{4k}\left(2^{r+1}\pi f\right) S_Y(f)df. \quad (31)$$

If $v_i \notin \left(2^{-r-2}, 2^{-r-1}\right]$ for all $i = 1, \ldots, k$, then $S_Y(f)$ is finite on $(2^{-r-2}, 2^{-r-1}]$. On the other hand, if there is $v_i \in (2^{-r-2}, 2^{-r-1}]$ for some $i = 1, \ldots, k$, then $S_Y(f)$ is integrable on $(2^{-r-2}, 2^{-r-1}]$. In both cases, the mean value theorem applies. Using Eq. (44) and Eq. (47) in (Gonzaga and Hauser, 2011), (31) satisfies

$$\lim_{\delta \to 0^+} C_9 \sum_{k=L}^{\infty} \frac{1}{k^{1/2}} \int_{2^{-r-2}+\delta}^{2^{-r-1}} \cos\left(2^{r+1}\pi f\right) \sin^{4k}\left(2^{r+1}\pi f\right) S_Y(f)df = \left(\frac{1}{L}\right). \quad (32)$$

The result follows from (28), (29) and (32). QED.

In the following theorem, we show that the within-scale ($j = j'$) autocovariance function, $\operatorname{cov}(D_{j,n,t}{}^{Y}, D_{j,n',t'}{}^{Y})$, is approximately zero when $n \neq n'$ for sufficiently large L. This is because the squared gain function $|U_{j,n}(f)|^2$ is nominally bandpass over the frequency interval $I_{j,n} = \left(\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$. These passbands are disjoint when $n \neq n'$. Hence, (15) goes to zero as L goes to infinity. This approximation depends only on the length of the wavelet filter L, and not on the length of the series.

Theorem 2. Let $\left\{D_{j,n,t}{}^{Y}\right\}$ be the DWPT of the log squared process $\left\{Y_t\right\}_{t=0}^{N-1}$, then for $n \neq n'$ the autocovariance of within-scale wavelet packet coefficients satisfies

$$\left|\operatorname{cov}(D_{j,n,t}{}^{Y}, D_{j,n',t'}{}^{Y})\right|^2 = O\left(\frac{1}{L^{1/2}}\right). \tag{33}$$

Proof:

Let $S_Y(f)$ be the spectral density function of the log squared process. For $n \neq n'$, (15) becomes

$$\left|\operatorname{cov}(D_{j,n,t}{}^{Y}, D_{j,n',t'}{}^{Y})\right|^2 \leq C_0\left(\int_0^{0.5} |U_{j,n}(f)||U_{j,n'}(f)|S_Y(f)df\right)^2, \tag{34}$$

for some constant C_0 independent of L. For each j, $\mathbf{c}_{j,n}$ is one of the 2^j distinct j-sequences of 0's and 1's. Hence, the vector $\mathbf{c}_{j,n}$ differs in at least one component with the vector $\mathbf{c}_{j,n}$. The product representations in (13) for $U_{j,n}$ and $U_{j,n'}$ differ in at least one factor, such that one factor is $H(\cdot)$ and another factor is $G(\cdot)$. Hence, for some m*, $1 \leq m^* \leq j$, $|H(2^{m^*}f)||G(2^{m^*}f)$ is a factor of $|U_{j,n}(f)||U_{j,n'}(f)|$. By analogous argument to the proof of Theorem 1,

$$\left|\operatorname{cov}(D_{j,n,t}{}^{Y}, D_{j,n',t'}{}^{Y})\right|^2 = O\left(\frac{1}{L^{1/2}}\right). \text{ QED.} \tag{35}$$

In the following theorem, we approximate the spectral density of $D_{j,n,t}{}^{Y}$ for sufficiently large L. We also show that the within-scale covariances converge to zero for appropriate choice of the wavelet packet subbands, which are deemed uncorrelated by a Portmanteau test.

Theorem 3. Let $\left\{D_{j,n,t}{}^{Y}\right\}$ be the DWPT coefficients of the log-squared process selected in (19), then for $j = j'$ and $n \neq n'$, the autocovariance of within-scale of wavelet packet coefficients satisfies

$$\left|\sigma_j^2(s)\right| = \lim_{L\to\infty}\left|\operatorname{cov}(D_{j,n,t}{}^{Y}, D_{j,n,t+s}{}^{Y})\right| = 2\left|\int_0^{0.5} \cos(2\pi fs)\left[S_Y\left(2^{-j}f + n2^{-j-1}\right)\right]df\right|, \tag{36}$$

which is the absolute autocovariance at lag s of a stationary process with spectral density function

$$S_Y(f) = S_y(2^{-j}f + n2^{-j-1}). \tag{37}$$

Moreover, the absolute autocorrelation at lag s satisfies $\left|\rho_j(s)\right| = O\left(\frac{1}{\sqrt{N_j}}\right)$, where

$$\rho_j(s) = \frac{\sigma_j^2(s)}{\sigma_j^2(0)}. \tag{38}$$

Proof of Theorem 3

From (16), the within-scale autocovariance of $\left\{D_{j,n,t}{}^{Y}\right\}$ is given by

$$\mathrm{cov}(D_{j,n,t}{}^{Y}, D_{j,n,(t+s)}{}^{Y}) = 2\int_0^{0.5} \cos\left(2\pi f 2^j s\right) | U_{j,n}(f)|^2 \; S_Y(f) df. \tag{39}$$

We partition the interval of integration into $I_{j,n}$ and $[0,0.5] - I_{j,n}$, so that

$$\begin{aligned}\mathrm{cov}(D_{j,n,t}{}^{Y}, D_{j,n,(t+s)}{}^{Y}) &= 2\int_{I_{j,n}} \cos\left(2\pi f 2^j s\right) | U_{j,n}(f)|^2 S_Y(f) df \\ &+ 2\int_{[0,0.5]-I_{j,n}} \cos\left(2\pi f 2^j s\right) | U_{j,n}(f)|^2 S_Y(f) df.\end{aligned} \tag{40}$$

From (9) and (13), $0 \le |U_{j,n}(f)|^2 \le 2^j$ for any f. Hence,

$$\begin{aligned}&\int_0^{0.5} \cos\left(2\pi f 2^j s\right) \left|| U_{j,n}(f)|^2 \; S_Y(f)\right| df \\ &\le \int_0^{0.5} 2^j S_Y(f) df = 2^j Var(Y_t) = 2^j \left(Var\left(X_t\right) + \sigma_\eta^2\right) < \infty.\end{aligned} \tag{41}$$

Moreover, from Eq. (6) in (Yu, 2013), we have

$$\lim_{L\to\infty} \left|U_{j,n,L}(f)\right| = \chi_{j,n}(f), \tag{42}$$

where

$$\chi_{j,n}(f) = \begin{cases} 2^j & |f| \in I_{j,n} \\ 0 & |f| \, [0,0.5] - I_{j,n} \end{cases}. \tag{43}$$

Thus, by the Lebesgue's Dominated Convergence Theorem, we have

$$\lim_{L\to\infty} \mathrm{cov}(D_{j,n,t}{}^{Y}, D_{j,n,(t+s)}{}^{Y}) = 2^{j+1} \int_{I_{j,n}} \cos\left(2\pi f 2^j s\right) S_Y(f) df \tag{44}$$

Using the transformation $f' = 2^j(f - n2^{-j-1})$, we get

$$\sigma_j^2(s) = \lim_{L\to\infty} \operatorname{cov}(D_{j,n,t}{}^Y, D_{j,n,(t+s)}{}^Y)$$
$$= 2(-1)^{ns}\int_0^{0.5} \cos(2\pi f's) S_Y\left(2^{-j}f' + n2^{-j-1}\right)df'. \quad (45)$$

Hence,

$$\left|\sigma_j^2(s)\right| = \lim_{L\to\infty}\left|\operatorname{cov}(D_{j,n,t}{}^Y, D_{j,n,(t+s)}{}^Y)\right| = 2\left|\int_0^{0.5}\cos(2\pi fs) S_Y\left(2^{-j}f + n2^{-j-1}\right)df\right|, \quad (46)$$

which is the absolute autocovariance at lag s of a stationary process with spectral density function

$$S(f) = S_Y(2^{-j}f + n2^{-j-1}). \quad (47)$$

The absolute autocorrelation at lag s is

$$\left|\rho_j(s)\right| = \left|\frac{\sigma_j^2(s)}{\sigma_j^2(0)}\right| = \frac{2\left|\int_0^{0.5}\cos(2\pi fs) S_Y\left(2^{-j}f + n2^{-j-1}\right)df\right|}{Var(Y_t)}. \quad (48)$$

Now, the wavelet packet subbands in (19) are selected if

$$Q = N_j(N_j + 2)\sum_{s=1}^{K}\frac{\rho_j^2(s)}{N_j - s} < \chi_K^2(\alpha), \quad (49)$$

where N_j is the length of the vector $\mathbf{D}_{j,n}$, K= max{2,min{20,N_j/10}}, and $\chi_K^2(\alpha)$ is the value of the chi-square random variable with K degrees of freedom leaving an area of α to the right. From (49), we have

$$\sum_{s=1}^{K}\rho_j^2(s) < \frac{N_j - K}{N_j(N_j + 2)}\chi_K^2(\alpha). \quad (50)$$

But K is independent of j for large N_j, and from (Tadeusz, 2010),

$$K + 2\log(1/\alpha) - 5/2 \le \chi_K^2(\alpha) \le K + 2\log(1/\alpha) + 2\sqrt{K\log(1/\alpha)}. \quad (51)$$

Therefore, $\left|\rho_j(s)\right| = O\left(\frac{1}{\sqrt{N_j}}\right)$. QED

The preceding theorem shows that the within-scale correlations of wavelet packet coefficients of the GLMSV model may not be approximately zero for large L, since unlike a fractionally integrated process with a spectral density that is almost flat on the passband, the spectral density of a k-factor GARMA process varies with the locations of the Gegenbauer frequencies. However, by appropriate choice of the wavelet basis vectors, the autocorrelations may be approximately zero by appropriately selecting wavelet packet subbands that best decorrelate the series using only the minimum number of levels j in the wavelet packet table, thereby minimizing the number of variances to compute.

4. Conclusions

We derived the order of convergence to zero of between-scale and within-scale wavelet packet coefficients at different subbands. Wavelet packet coefficients in the same subband were shown to be approximately uncorrelated by appropriately selecting wavelet packet subbands that best decorrelate the series using only the minimum number of levels in the wavelet packet table thereby minimizing the number of distinct variances. These results effectively simplify the variance-covariance matrix into a diagonalized matrix, whose diagonal elements have the least distinct variances to compute. Moreover, since it is not possible to obtain an exact expression for the autocovariance of a GLMSV process with two or more poles, estimation in the wavelet domain circumvents the problem of convergence in time domain, which is greatly amplified in the presence of high-frequency data.

ACKNOWLEDGEMENT

This paper is part of a research project funded by the National Research Council of the Philippiness.

REFERENCES

ANDERSEN, T. and BOLLERSLEV, T. ,1997, Intraday periodicity and volatility persistence in financial markets, Journal of Empirical Finance, 4: 115-158.

ANDERSEN T, BOLLERSLEV T, DIEBOLD FX, 2003, Modeling and Forecasting Realized Volatility, Econometrica , 71: 529-626.

ARTECHE J, 2004, Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models, Journal of Econometrics, 119: 131-154.

ARTECHE J, ROBINSON P, 2000, Semiparametric Inference in Seasonal and cyclical long memory processes, Journal of Time Series Analysis, 21(1):1-25.

BISAGLIA L, BORDIGNON S, LISI F, 2003, k-factor garma models for intraday volatility forecasting, Applied Economics Letters, 10: 251-254.

BOLLERSLEV T, MIKKELSEN HO, 1996, Modeling and pricing long memory in stock market volatility, Journal of Econometrics, 73:151-184.

BREIDT F.J., CRATO N, DE LIMA P, 1998, On the detection and estimation of long memory in stochastic volatility, Journal of Econometrics, 83: 325-348.

CRAIGMILE P.F., PERCIVAL DB, 2005, Asymptotic decorrelation of between-scale wavelet coefficients, IEEE Trans. Info. Theory, 51: 1039-1048.

DAUBECHIES I, 1992, Ten Lectures on Wavelets, SIAM, Philadelphia.

DEO R, HURVICH C, 2001, On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models, Econometric Theory, 17:686-710.

DEO R, HURVICH C, LU Y, 2006, Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment, Journal of Econometrics, 131: 29-58.

FAN Y, 2003, On the Approximate Decorrelation Property of Discrete Wavelet Transform for Fractionally Differenced Process, IEEE Transactions on Information Theory, 49:516-521.

GIRAITIS L, LEIPUS R, 1995, A generalized fractionally differencing approach in long-memory modeling, Lithuanian Mathematical Journal, 35(1):53-65.

GONZAGA A, KAWANAKA A, 2006, Asymptotic Decorrelation of Between-Scale Wavelet Coefficients of Generalized Fractional Process, Digital Signal Processing, 16: 320-329.

GONZAGA, A. AND HAUSER, M., 2011, A wavelet Whittle estimator of generalized long-memory stochastic volatility, Statistical Methods and Applications, 20(1) 23-48.

GONZAGA, A., 2013, Forecasting long-memory stochastic volatility, International Journal of Applied Mathematics and Statistics, 51(21): 195-2004.

HARVEY A., 1998, Long memory in stochastic volatility. In: Knight J, Satchell S (Eds.) Forecasting Volatility in Financial Markets. Butterworth-Heineman, Oxford, pp. 307-320.

HURVICH C., RAY B, 2003, The Local Whittle estimator of long-memory stochastic volatility, Journal of Financial Econometrics, 1(3):445-470.

JENSEN M., 2004, Semiparametric Bayesian Inference of Long-memory stochastic volatility models, Journal of Time Series Analysis, 25(6): 895-122.

PERCIVAL D.B., SARDY S. AND DAVISON A.C., 2001, Wavestrapping Time Series: Adaptive Wavelet-Based Bootstrapping, in W. J. Fitzgerald, R. L. Smith, A. T. Walden and P. C. Young (Eds.), Nonlinear and Nonstationary Signal Processing, Cambridge, England: Cambridge University Press 442-470.

PERCIVAL D., WALDEN A., 2000, Wavelet Methods for Time Series Analysis, Cambridge University Press.

TADEUSZ, I., 2010, Inequalities for quantiles of chi-square distribution, Probability and Mathematical Statistics, 30(2): 339-351.

WANG, P., 2009, Financial Econometrics (2nd Edition), Routledge, New York.

WHITCHER B., 2004, Wavelet-based estimation for seasonal long-memory, Technometrics, 46(2): 225-238.

WOODWARD W., CHENG Q, GRAY H., 1998, A k-factor GARMA Long-memory model. Journal of Time Series Analysis, 19(4): 485-504.

YU, X., 2013, Asymptotic decorrelation of wavelet packet transform for certain long-memory processes, IEEE Transactions on Information Theory, 59(8): 5051-5062.