HIGHER ORDER APPROXIMATIONS FOR WALD STATISTICS IN TIME SERIES REGRESSIONS WITH INTEGRATED PROCESSES

BY

ZHIJIE XIAO and PETER C. B. PHILLIPS

COWLES FOUNDATION PAPER NO. 1094

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

2004

http://cowles.econ.yale.edu/
Higher order approximations for Wald statistics in time series regressions with integrated processes

Zhijie Xiaoa, Peter C.B. Phillipsb,

\textsuperscript{a}Department of Economics, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, IL 61820, USA
\textsuperscript{b}Cowles Foundation for Research in Economics, Yale University, P.O. Box 208281, New Haven, CT 06520, USA; University of Auckland & University of York

Received 16 June 1998; received in revised form 6 September 2001; accepted 13 November 2001

Abstract

Asymptotic expansions are developed for Wald test statistics in time series regressions with integrated processes. These expansions provide an opportunity to reduce size distortion in testing by suitable bandwidth selection, and automated rules for doing so are calculated. A band spectral regression and the associated Wald test are also considered. Both the first order and second order properties of the estimator are studied.

JEL classification: C14; C22

Keywords: Bandwidth selection; Higher order approximation; Moment expansion; Spectral regression

1. Introduction

This paper studies the Wald statistic in time series regression models of the form

\[ y_t = \beta' x_t + u_t, \quad t = 1, \ldots, n, \]

where \( u_t \) is an \( I(0) \) process with zero mean and continuous spectral density \( f_{uu}(\lambda) \). The regressor \( x_t \) is a vector of integrated time series of order one (\( I(1) \)) or zero (\( I(0) \)). Both \( I(0) \) and \( I(1) \) regressors are studied here, but the main focus of the paper is on the case with \( I(1) \) regressors. In this event, if \( x_t \) is \( I(1) \), then both \( x_t \) and \( y_t \) are nonstationary and the linear combination \( y_t - \beta' x_t \) is stationary, so that (1.1) is a cointegrating regression in the sense of Engle and Granger (1987). More will be
said later about the alternate specifications of the model as we proceed to develop the higher order theory.

We consider the frequency domain version of (1.1), viz.

\[ w_y(\lambda_s) = \beta' w_x(\lambda_s) + w_u(\lambda_s), \quad \lambda_s = 2\pi s/n, \quad s = -n/2 + 1, \ldots, n/2, \quad (1.2) \]

where \( w_y(\lambda_s) \), \( w_x(\lambda_s) \), and \( w_u(\lambda_s) \) are discrete Fourier transforms of \( y_t \), \( x_t \), and \( u_t \), at the fundamental frequencies \( \lambda_s \) defined by \( w_\lambda(\lambda_s) = \left(1/\sqrt{2\pi n}\right) \sum_{t=1}^{n} a_t e^{i\lambda_s t} \). For convenience we simply assume that \( n \) is even. If the residual process in (1.1), \( u_t \), is stationary with a continuous spectrum that is bounded away from the origin, the residuals of regression (1.2), \( w_u(\lambda_s) \), are asymptotically uncorrelated but generally heteroskedastic. Regression in the frequency domain permits a nonparametric treatment of the regression errors and, utilizing a consistent estimate of \( f_{uu}(\lambda) \), delivers a semi-parametric estimator of \( \beta \) that is asymptotically equivalent to GLS in the case of stationary \( u_t \) and \( x_t \). Hence, it is unnecessary to parameterize the autocorrelation structure of \( u_t \) to achieve efficient estimation. In addition, these methods allow attention to be focused on the most relevant frequency in the regression, thereby providing a selective approach that accommodates more general formulations in which the parameter \( \beta \) may not be constant across frequency bands (Of course, in this case, the time domain counterpart of regression (1.2) will be more complicated—see Corbae et al., 2002.) When \( x_t \) is \( I(1) \) and independent of \( u_t \), OLS regression is asymptotically efficient in (1.1), as shown by Phillips and Park (1988), but efficiency gains still accrue from GLS in finite samples (see Section 2 for a discussion).

Regression analysis in the frequency domain was introduced by Hannan (1963a), following ideas of Whittle (1951), and extended to nonlinear models by Hannan (1971) and Robinson (1976). Robinson (1991) studied automatic frequency domain inference on semiparametric and nonparametric models where bandwidth selection for the spectral density estimate is determined from the data. Phillips (1991) showed how to apply this method to cointegrating regressions, developed a limit theory for frequency domain estimators in this case and established some optimality properties for such regressions. In addition, a recent paper by Corbae et al. (2002) provides asymptotic results on the behavior of discrete Fourier transforms of integrated processes and deterministic trends and explores spectral regression methods in the presence of deterministic trends.

The present paper studies the use of Wald statistics in linear time series regressions like (1.1). Finite sample problems of over-rejection in the use of Wald tests in such regressions have long been recognized by econometricians, an example being the well studied phenomenon of over-rejection of homogeneity conditions in empirical consumer demand analysis (see Barten, 1969; Byron, 1970; Lluch, 1971; Deaton, 1974). Monte Carlo results in such cases have shown that the Wald test is biased toward rejecting the null hypothesis (e.g. Laitinen, 1978; Meisner, 1979; Bera et al., 1981). One of the mechanisms for improving asymptotic \( \chi^2 \) approximations of Wald tests is the use of higher order expansions. The statistical theory of asymptotic expansions for Wald tests has been extensively studied in econometrics (see Sargan and Mikhail, 1971; Sargan, 1976; Phillips, 1977; Phillips, 1987; Phillips and Park, 1988; Rothenberg, 1984a,b; and Linton, 1995a, b). However, higher order expansions have not so far been developed and used in nonstationary time series environments, with an exception
in Phillips (1987), largely because of the difficulty in developing valid higher order extensions of the underlying functional central limit theory on which the nonstationary regression asymptotics typically depend.

This paper seeks to implement a simple and usable approach to the development of a higher order theory for Wald tests in time series regressions with $I(1)$ regressors, avoiding the need for higher order extensions of the functional limit theory. Our approach is formal, rather than rigorous. In particular, we develop formal moment approximations, in terms of moments of approximating expansions, rather than rigorous asymptotic expansions of the moments of the Wald statistics. These formal expansions lead quite simply to formulae that can be used in practical applications. In efficient semiparametric time series regression a critical element in the construction of the estimator and associated tests is the choice of the bandwidth in the estimation of the spectrum. The idea behind the present development is to construct a bandwidth selection criterion by minimizing the second order effect on the expected value of the Wald statistic. A second order adjusted Wald statistic can then be constructed to correct the size distortion of the Wald test in finite samples. Results from this approach are given also for stationary time series regressions.

The paper is organized as follows. The model and test statistics are studied in the next section. Some preliminary results for spectral density estimation and spectral regression with integrated regressors are given in Section 3. The expansion for the Wald statistic and a modified Wald test in the univariate case is given in Section 4. The case of a narrow band spectral regression is studied in Section 5. Section 6 gives the results in the multivariate case. The results of a small Monte Carlo experiment are reported in Section 7. Section 8 concludes and some proofs are given in the Appendix.

2. Background and assumptions

Our subject is the Wald statistic in regression (1.1), where the regressor $x_t$ is either an $I(1)$ or an $I(0)$ process. An intercept term can also be included in (1.1) simply by dropping the zero frequency in (1.2) in estimation. Our main assumptions are as follows, although, for some results, further specific assumptions may be needed and these are stated as they are required. When $x_t$ is $I(1)$, we assume that $\Delta x_t = v_t$, and that $x_t$ is initialized at $t=0$ by any $O_p(1)$ random variables. The quantities $v_t$ and $u_t$ are independent stationary and ergodic $k$-vector and scalar time series with zero mean, finite fourth moments and spectral densities $f_{vv}(\lambda)$ and $f_{uu}(\lambda)$, which are bounded away from zero. We assume that $\sum_{h} |h|^q |\gamma_q(h)| < \infty$, for some $q > 2$, where $\gamma_q(h) = \operatorname{cat}(u_t, u_{t+h})$. The vector $x_t$ and partial sums of $u_t$ both satisfy invariance principles with independent limit processes, so that, as $n \to \infty$, $n^{-1/2} x_{t[n]} \Rightarrow B_x(r) \equiv BM(2\pi f_{vv}(0))$, a vector Brownian motion of dimension $k$ with covariance matrix $2\pi f_{vv}(0)$, and $n^{-1/2} \sum_{t=1}^{[nr]} u_t \Rightarrow B_u(r) \equiv BM(2\pi f_{uu}(0))$. In the case that $x_t$ is an $I(0)$ regressor, we assume that $x_t$ satisfies the same conditions as those given above for $u_t$.

---

1 This summability condition not only implies that the spectral density is continuous and bounded, but also imply the uniform boundedness of $f_q(\lambda)$ which is defined below (see, e.g. Brillinger, 1980).
Table 1
Finite sample comparison between OLS and GLS

<table>
<thead>
<tr>
<th>n</th>
<th>OLS</th>
<th>GLS1</th>
<th>GLS2</th>
<th>GLS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>/VT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.001469</td>
<td>-0.000921</td>
<td>-0.001120</td>
<td>-0.001292</td>
</tr>
<tr>
<td>128</td>
<td>0.045074</td>
<td>0.046907</td>
<td>0.047975</td>
<td></td>
</tr>
<tr>
<td>256</td>
<td>0.013086</td>
<td>0.012652</td>
<td>0.012956</td>
<td>0.012995</td>
</tr>
<tr>
<td>/VT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.00087</td>
<td>-0.000116</td>
<td>-0.000018</td>
<td>-0.000012</td>
</tr>
<tr>
<td>128</td>
<td>0.024189</td>
<td>0.024747</td>
<td>0.024798</td>
<td></td>
</tr>
<tr>
<td>256</td>
<td>0.005566</td>
<td>0.00522</td>
<td>0.00563</td>
<td></td>
</tr>
<tr>
<td>/VT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.000018</td>
<td>-0.000012</td>
<td>-0.00001</td>
<td>-0.000012</td>
</tr>
<tr>
<td>128</td>
<td>0.037645</td>
<td>0.037445</td>
<td>0.038616</td>
<td></td>
</tr>
<tr>
<td>256</td>
<td>0.00994</td>
<td>0.00998</td>
<td>0.01068</td>
<td></td>
</tr>
</tbody>
</table>

Under these assumptions, a frequency domain efficient estimator of \( \beta \) is obtained from a regression with weighted averages of periodogram estimates at the fundamental frequencies, viz

\[
\hat{\beta} = \left( \sum_s I_{xx}(\lambda_s) \hat{f}_{uu}^{-1}(\lambda_s) \right)^{-1} \left( \sum_s I_{xy}(\lambda_s) \hat{f}_{uu}^{-1}(\lambda_s) \right),
\]

where \( I_{xx}(\lambda_s) = w_x(\lambda_s)w_x(\lambda_s)^* \) and \( I_{xy}(\lambda_s) = w_x(\lambda_s)w_y(\lambda_s)^* \) (the suffix * signifies the complex conjugate transpose), \( \hat{f}_{uu}(\lambda_s) \) is a nonparametric estimator of \( f_{uu}(\lambda_s) \) and the summation \( \sum_s \) is taken over \( [-\pi, \pi] \). When \( x_t \) is an integrated process, it is well known that the frequency domain GLS estimator (2.1) is more efficient than the OLS estimator and achieves asymptotically the Gauss–Markov efficiency bound under the given smoothness condition on the residual spectral density. In the case that \( x_t \) is an integrated process, OLS estimation of \( \beta \) is (first order) asymptotically equivalent to the GLS estimator (Phillips and Park, 1988). However, such equivalence holds only asymptotically and, in general, the GLS regression estimator (2.1) still has better small sample properties. Table 1 reports some Monte Carlo results on the comparison between OLS and (2.1) in a simple version of (1.1) with autocorrelated errors. The results indicate that, although OLS is asymptotically efficient when the regressor is \( I(1) \), (2.1) generally has superior finite sample performance.

We are interested in testing the linear hypothesis

\[
H_0: R\beta = r,
\]
where \( r \) is an \( p \times 1 \) vector of constant and \( R \) is an \( p \times k \) matrix. The regression Wald statistic corresponding to \( \hat{\beta} \) in (2.1) is given by

\[
W = (R\hat{\beta} - r)'[R\hat{\Sigma}^{-1}R']^{-1}(R\hat{\beta} - r),
\]

(2.3)

where

\[
\hat{\Sigma} = \sum_s I_{sx}(\lambda_s)f_{uu}^{-1}(\lambda_s).
\]

All of the theory we develop carries over with minor changes to the case of analytic nonlinear restrictions on \( \beta \) in place of (2.2) and associated changes in the Wald statistic (2.3).

The following lemma is closely related to earlier work in Phillips (1991) and gives the limit theory of the estimator \( \hat{\beta} \) and the corresponding Wald statistic (2.3) when \( x_t \) is an \( I(1) \) process.

**Lemma 1.** As \( n \to \infty \):

1. \( n(\hat{\beta} - \beta) \to [\int B_sB_s']^{-1}[\int B_sdB_s] = MN(0, 2\pi f_{uu}(0)[\int B_sB_s']^{-1}). \)
2. Under \( H_0 \), \( W \to \chi_p^2 \).

As the lemma shows, the asymptotic distribution of the regression estimator is mixed normal with matrix mixing variate \( [\int B_sB_s']^{-1} \) that depends on the Brownian motion \( B_s(r) \). This mixed normal distribution facilitates the construction of a regression Wald test about \( \beta \) by using an estimate of the conditional covariance matrix \( \hat{\Sigma} = \sum_s I_{sx}(\lambda_s)f_{uu}^{-1}(\lambda_s) \). As a result, the Wald statistic has an asymptotic \( \chi^2 \) distribution under \( H_0 \), and the null hypothesis can be tested using conventional \( \chi^2 \) limit theory distribution. However, the asymptotic \( \chi^2 \) distribution by no means always provides a good approximation to the distribution of \( W \), and Monte Carlo evidence indicates that Wald tests often over-reject the null hypothesis in finite samples. This paper proceeds to derive a higher order expansion for the expected Wald statistic. The expansion helps to explain the finite sample performance of the Wald test and to make compensating adjustments using second order effects.

### 3. Some preliminary expansions

We consider the ‘leave-one-out’ type of nonparametric estimator for the spectral density of the residual process \( u_t \), viz

\[
\hat{f}_{uu}(\lambda) = \frac{1}{m} \sum_{j \in B(\lambda_s), j \neq \lambda_s} K(\lambda_j - \lambda_s)w(\lambda_i)w(\lambda_j)^* = \frac{1}{m} \sum_{j \in B(\lambda_s), j \neq \lambda_s} K(\lambda_j - \lambda_s)I_{i\lambda}(\lambda_j)
\]

\[
= \sum_{j \neq s} c_{\lambda_j}I_{\lambda \lambda}(\lambda_j).
\]
where
\[
\omega_{ij} = \begin{cases} 
K(\lambda_j - \lambda_s)/m, & \lambda_j \in B(\lambda_s), \\
0 & \text{otherwise},
\end{cases}
\]
and \(B(\lambda_s) = \{\omega; \lambda_s - \pi/2M < \omega \leq \lambda_s + \pi/2M\}\) is a frequency band of width \(\pi/M\) centered on \(\lambda_s = 2\pi s/n\). Let \(m = \lfloor n/2M \rfloor\), where \(\lfloor \cdot \rfloor\) signifies integer part. Then, each band \(B(\lambda_s)\) contains \(m\) fundamental frequencies \(\lambda_s\). \(K(\cdot)\) is a spectral window and satisfies conventional properties of being a real, even function with \(1 = m \sum_{\lambda_j \in B(\omega)} K(\lambda_j - \omega) = 1\). Many candidate kernel functions are available and are discussed in standard texts of spectral analysis (e.g., Hannan, 1970; Brillinger, 1980; Priestley, 1981). The periodogram ordinates \(\hat{I}_{uu}(\lambda_s)\) are calculated using consistent estimates of the residuals. In our analysis, we use the residuals from an OLS regression on (1.2).

It is convenient to decompose the error term in the nonparametric spectral density estimator \(\hat{f}_{uu}(\lambda_s)\) into a bias effect due to smoothing, \(B_s\), a variance term arising from the periodogram, \(V_s\), and an error coming from the preliminary estimation of \(w_u(\lambda_s), P_s\). Thus, we have
\[
\hat{f}_{uu}(\lambda_s) = f_{uu}(\lambda_s) + B_s + V_s + P_s, \tag{3.1}
\]
where \(B_s = \sum_{j \neq s} \omega_{ij} [f_{uu}(\lambda_j) - f_{uu}(\lambda_s)]\), \(V_s = \sum_{j \neq s} \omega_{ij} [I_{uu}(\lambda_j) - f_{uu}(\lambda_j)]\), \(P_s = \sum_{j \neq s} \omega_{ij} [\hat{I}_{uu}(\lambda_j) - I_{uu}(\lambda_j)]\). Asymptotic results for the bias and variance in spectral density estimation are well known (e.g. Hannan, 1970) and are stated in the following lemma for convenience.

**Lemma 2.**
\[
B_s \sim -M^{-q} k_q f_q(\lambda_s),
\]
\[
\sqrt{m} V_s \overset{d}{\sim} N \left(0, \frac{1}{2} \int_{-\infty}^{\infty} k(x)^2 dx f_{uu}^2(\lambda_s) \right),
\]
where \(q\) is the characteristic exponent of the kernel function defined as \(\lim_{s \to 0} \{1 - k(x)|x|^q = k_q < \infty, f_q(\lambda) = 1/2\pi \sum_{s=-\infty}^{\infty} |h|^q f_k(h)e^{-i\lambda h}\}, \sim\) denotes asymptotic equivalence and \(\overset{d}{\sim}\) denotes asymptotically distributed.

Most commonly used kernel functions have the property that \(q = 2\) in which case the bias term is of order \(O(M^{-2})\). For example, this applies for the smoothed periodogram estimate corresponding to the Daniell kernel with \(k(x) = \sin(\pi x/2)/\pi x/2\), and spectral window
\[
K(\lambda) = \begin{cases} 
\pi^{-1} & \text{for } |\lambda| \leq \pi/2, \\
0 & \text{otherwise}.
\end{cases}
\]
Since our interest is primarily in the analysis of second order effects, in what follows we will simply give results based on the smoothed periodogram estimate for which \(a = 1/2 \int_{-\infty}^{\infty} k(x)^2 dx = 1\), and \(\sqrt{m} V_s \overset{d}{\sim} N(0, f_{uu}^2(\lambda_s))\). The use of other kernels then involves only minor scale adjustments in the formulae.
In spectral regression for stationary time series, the error term coming from preliminary estimation of $w_u(\lambda_s)$ is of smaller order of magnitude and can be dropped. However, different results arise in regression models with $I(1)$ regressors. In particular, the periodogram average of an $I(1)$ process around the zero frequency has a larger order of magnitude than at high frequencies. As a result, in the spectral density estimator $\hat{f}_{uu}(\lambda_s)$, the errors from preliminary estimation are amplified around the zero frequency and thus enter the second order effect. The order of magnitude for $P_t$ is summarized in the following lemma.

**Lemma 3.**

$$P_t \sim -\frac{1}{2m\pi} \int dB_sB_s' \left[ \int B_sB_s' \right]^{-1} \int B_s dB_u, \quad \text{for } |\lambda_s| \leq \pi/(2M)$$

and

$$P_t = o_p(m^{-1}) \quad \text{for } |\lambda_s| > \pi/(2M).$$

In view of the $\sqrt{m}$ rate of convergence of the spectral estimates and the $M^{-q}$ order of magnitude of the bias, the expansions we deal with will typically be in terms of powers of $m^{-1/2}$ and $M^{-q}$. In these expansions, terms that are of order $O_p(m^{-1})$ or $O_p(M^{-2q})$ turn out to be of principal interest and will be referred to as “higher order” terms.

An expansion for the inverse of the spectral density estimator, $\hat{f}_{uu}(\lambda_s)^{-1}$, can be obtained based on (3.1) and Lemmas 2 and 3. The results are given in Lemma 4 and readers are referred to Xiao and Phillips (1998) for details of the derivation.

**Lemma 4.**

$$\hat{f}_{uu}(\lambda_s)^{-1} = f_{uu}(\lambda_s)^{-1} - \hat{f}_{uu}(\lambda_s)^{-3} V_t - f_{uu}(\lambda_s)^{-2} B_t$$

$$+ f_{uu}(\lambda_s)^{-3} B_t^2 - \hat{f}_{uu}(\lambda_s)^{-2} P_t + R_1(\lambda_s), \quad (3.2)$$

where

$$\hat{f}_{uu}(\lambda_s) = m^{-1} \sum_{j \neq s} K(\lambda_j - \lambda_s) f_{uu}(\lambda_j), \quad \hat{f}_{uu}(\lambda_s) = m^{-1} \sum_{j \neq s} K(\lambda_j - \lambda_s) I_{uu}(\lambda_j),$$

and

$$R_1(\lambda_s) = -\hat{f}_{uu}(\lambda_s)^{-1} f_{uu}(\lambda_s)^{-3} B_t^3 - \hat{f}_{uu}(\lambda_s)^{-1} \hat{f}_{uu}(\lambda_s)^{-3} V_t^3$$

$$+ \hat{f}_{uu}(\lambda_s)^{-1} \hat{f}_{uu}(\lambda_s)^{-2} P_t. \quad (3.3)$$

By assumption, $f_{uu}(\lambda)$ is bounded away from zero and so, from Lemmas 2–4, we deduce that $R_1(\lambda_s) = o_p(m^{-1} + M^{-2q})$.

The following lemma gives a useful limiting result for periodogram averages of $I(1)$ processes. Its derivation is based on a related result of Phillips (1991) and it will be used in the analysis of higher order asymptotics throughout later sections.
Lemma 5. If the function $g(\lambda)$ has an absolutely and uniformly convergent Fourier representation $g(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \phi_k e^{ik\lambda}$, then

$$\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) g(\lambda_s) \xrightarrow{d} \frac{1}{2\pi} g(0) \int_0^1 B_s B'_s.$$ 

In particular, if $g(\lambda)$ is a power function of $f_{uu}(\lambda)$ and its derivatives to order $q$, the above result holds.

Remark 1. As in Phillips (1991), contributions from the zero-frequency ordinates dominate the limit behavior of these periodogram averages of $I(1)$ series and the limits are characterized in terms of quadratic functionals of the Brownian motion weak limit of $n^{-1/2} X_{[n]}$.

4. The Wald expansion in the univariate case

We start the development with the scalar case ($k = 1$) to simplify the derivations and illustrate the main results. The multivariate case will be treated later in Section 6. As mentioned in the introduction, here and elsewhere in the paper we proceed with formal moment approximations, rather than rigorous moment expansions.

In the univariate case, the linear hypothesis $R\beta = r$ reduces to $H_0: \beta = \beta_0$, and the corresponding regression Wald statistic is simply $W = (\hat{\beta} - \beta_0)^* \hat{\Sigma}(\hat{\beta} - \beta_0)$. Under the null hypothesis that $\beta = \beta_0$,

$$\hat{\beta} - \beta_0 = \left[ \sum_s I_{xx}(\lambda_s) \hat{f}_{uu}(\lambda_s) \right]^{-1} \left[ \sum_s I_{uu}(\lambda_s) \hat{f}_{uu}(\lambda_s) \right],$$

and

$$W = Z^* \hat{H}^{-1} Z,$$  \hspace{1cm} (4.1)

where

$$\hat{H} = \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1},$$  \hspace{1cm} (4.2)

$$H = \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1},$$  \hspace{1cm} (4.3)

$$Z = \frac{1}{n} \sum_s I_{uu}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1}.$$  \hspace{1cm} (4.4)

By an application of Lemma 4, we decompose $D = \hat{H} - H$ and $Z$ as follows:

$$D = -\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) \tilde{f}(\lambda_s)^{-2} V_s + \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) \tilde{f}(\lambda_s)^{-3} V_s^2$$

$$- \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s + \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2$$

$$- \frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} P_s + R_3$$

$$= -D_{Y1} + D_{Y2} - D_{B1} + D_{B2} - D_{P1} + R_3$$  \hspace{1cm} (4.5)
and
\[
Z = \frac{1}{n} \sum_s I_{su}(\hat{\lambda}_s)f_{su}(\hat{\lambda}_s)^{-1} - \frac{1}{n} \sum_s I_{su}(\hat{\lambda}_s)f_{su}(\hat{\lambda}_s)^{-2}V_s \\
+ \frac{1}{n} \sum_s I_{su}(\hat{\lambda}_s)f_{su}(\hat{\lambda}_s)^{-3}V_s^2 - \frac{1}{n} \sum_s I_{su}(\hat{\lambda}_s)f_{su}(\hat{\lambda}_s)^{-2}B_s \\
+ \frac{1}{n} \sum_s I_{su}(\hat{\lambda}_s)f_{su}(\hat{\lambda}_s)^{-3}B_s^2 - \frac{1}{n} \sum_s I_{su}(\hat{\lambda}_s)f_{su}(\hat{\lambda}_s)^{-2}P_s + R_4 \\
= Z_0 - ZV_1 + ZV_2 - ZB_1 + ZB_2 - ZP_1 + R_4,
\]
(4.6)
where \(R_3 = 1/n^2 \sum_s I_{sx}(\hat{\lambda}_s)R_3(\hat{\lambda}_s), \ R_4 = 1/n \sum_s I_{su}(\hat{\lambda}_s)R_1(\hat{\lambda}_s), \) and \(R_1(\hat{\lambda}_s) = o_p(m^{-1} + M^{-q})\) is defined in Lemma 4.

The orders of magnitude of the terms in expansions (4.5) and (4.6) are analyzed in Lemma A1 in the appendix. Expanding \(\hat{H}^{-1}\) around \(H^{-1}\) to the third term in a geometric expansion, substituting the truncated forms of \(D; Z\), and \(\hat{H}^{-1}\) into (4.1), and collecting terms up to \(O_p(m^{-1} + M^{-q})\), we obtain a formal moment expansion for \(E[W]\), which we summarize for convenience in the following formal result.

**Theorem 1.**
\[
E[W] = Q_0 + m^{-1}Q_1 + M^{-q}Q_2 + o(m^{-1} + M^{-q}) \\
= \tilde{W} + o(m^{-1} + M^{-q}),
\]
where \(Q_0, Q_1, \) and \(Q_2\) are \(O(1)\) quantities defined as
\[
Q_0 = E[H^{-1}Z_0^*Z_0], \\
Q_1 = Q_{11} + Q_{12}, \\
Q_{11} = m\{E[H^{-1}Z_0^*Z_{V_1}] + 2E[H^{-1}Z_0^*Z_{V_2}] \\
+ E[H^{-3}D_{V_1}^2Z_0^*Z_0] - E[H^{-2}D_{V_2}Z_0^*Z_0]\}, \\
Q_{12} = m\{E[H^{-1}D_{P_1}^2Z_0^*Z_0] - 2E[H^{-1}Z_0^*Z_{P_1}]\}, \\
Q_2 = M^q\{E[H^{-2}D_{B_1}Z_0^*Z_0] - 2E[H^{-1}Z_0^*Z_{B_1}]\},
\]
and \(\tilde{W} = Q_0 + m^{-1}Q_1 + M^{-q}Q_2\) is the truncated expectation of the Wald statistic.

**Remark 2.** In this expansion, the leading term, \(Q_0\), is just the expected value of the infeasible version of \(W\) in which the true spectral densities appear.

**Remark 3.** The second order effects (up to \(O(m^{-1} + M^{-q})\)) include: \(Q_{11}\), a variance effect; \(Q_{12}\), second order effect coming from the preliminary estimation; and \(Q_2\), a bias effect. Here, the bias term \(Q_2\) dominates the squared bias terms in order of magnitude and thus enters the second order effect.
Remark 4. Among the second order effects, the first term, $m^{-1}Q_1$, is a variance term and is always positive. The effect coming from preliminary estimation, $m^{-1}Q_{12}$, is asymptotically positive. The third term, $M^{-q}Q_2$, is a bias term that depends on the curvature of the spectral density function $f_{uu}(\omega)$. In fact, if we perform the expansion further to the order $M^{-2q}$, it can be shown that the $O(M^{-2q})$ term is also positive. Thus, it is apparent from the expansion that, in most cases, the Wald test is likely to over-reject the null hypothesis.

When $Q_2$ is negative, we get two terms in the second order effect with different signs: $m^{-1}Q_1$ and $M^{-q}Q_2$. Whether or not the Wald test will over-reject $H_0$ depends on which of these terms dominates. In this case, a (second order) optimal bandwidth can be selected by minimizing the absolute value of the second order effects, that is, $M$ can be chosen by equating $m^{-1}Q_1$ to $M^{-q} |Q_2|$, giving the optimal bandwidth

$$M = \left( \frac{|Q_2|}{2Q_1} \right)^{1/(q+1)} n^{-1/(q+1)}.$$  

Choosing the bandwidth by this formula enhances second order efficiency.

When $Q_2$ is positive, both terms in the second order effect of $E[W]$ are positive. As a result, at least to the second order, the Wald statistic (2.2) tends to over-reject the null hypothesis. In this case, a bandwidth selection criterion can be defined by minimizing the expected value of the second order effect in the Wald statistic, i.e. $M$ can be chosen to minimize $m^{-1}Q_1 + M^{-q}Q_2$, yielding

$$M = \left( \frac{Q_2}{2Q_1} \right)^{1/(q+1)} n^{1/(q+1)}.$$  

Choosing the bandwidth in such a way, the truncated expected Wald statistic is

$$Q_0 + n^{-q/(q+1)} [q^{1/(q+1)} + q^{-q/(q+1)}] (2Q_1)^{q/(q+1)} Q_2^{1/(q+1)}.$$  

A second order adjusted Wald statistic can also be constructed based on the above expansion. If $\hat{Q}_1$ and $\hat{Q}_2$ are consistent estimates of $Q_1$ and $Q_2$, we define the following second order modified Wald statistic:

$$W_M = W - n^{-q/(q+1)} \left[ q^{1/(q+1)} + q^{-q/(q+1)} \right] (2\hat{Q}_1)^{q/(q+1)} \hat{Q}_2^{1/(q+1)}.$$  

This modified Wald statistic is asymptotically $\chi^2$ and has an $O_p(n^{-q/(q+1)})$ second order effect removed from its expected value.

The asymptotics of the second order terms can be analyzed using the results of Lemma 5 (see Lemma A.2 in the appendix for this analysis), leading us to the following approximation for the truncated expected Wald statistic. Again, we state these formal approximation results in summary form as a theorem.

**Theorem 2.**

$$\tilde{W} \sim 1 + \frac{4}{m} + \frac{k_q}{M^{q/2}} f_{uu}(0)^{-1} f_q(0).$$
Since the effect at the zero frequency dominates, the second order effect in $\tilde{W}$ is asymptotically determined by the spectral density of the residual process at the origin. As a result, the size of the test based on $W$ will be largely affected by the curvature of the spectral density function $f_{uu}(\cdot)$ at the origin. When $f_q(0)$ is positive, a simple formula for bandwidth selection can be defined that is based on the unknown spectrum and its smoothness component, viz.

$$M = \left( \frac{q k_q f_{uu}(0)^{-1} f_q(0)}{8} \right)^{1/(q+1)} n^{1/(q+1)}. \quad (4.7)$$

If we plug consistent estimates $\hat{f}_{uu}(0)$ and $\hat{f}_q(0)$ into the formula, then a practical formula for bandwidth selection is obtained. A simple second order modified Wald test can also be constructed using this formula as follows:

$$W_{SO} = W - n^{-q/(q+1)}[q^{1/(q+1)} + q^{-q/(q+1)}](8)^{q/(q+1)}[k_q \hat{f}_q(0)/\hat{f}_{uu}(0)]^{1/(q+1)}. \quad (4.8)$$

In the extreme case that $f_q(0) = 0$, the leading bias term $M^{-q}k_q f_{uu}(0)^{-1} f_q(0)$ is zero, indicating that the optimal $m$ will be of larger order of magnitude. In this case, in principle, further expansion to higher order is needed. In the special case where $u_t$ are i.i.d., we choose $m = n$ and thus the OLS estimator is optimal.

**Remark 5.** If $x_t$ is an $I(0)$ process, the higher order asymptotics are different. It can be shown in this case that

$$E[W] \sim 1 + \frac{2}{m} + \frac{k_u}{M^q} \Omega^{-1} \Gamma,$$

where

$$\Omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{uu}(\omega) f_{uu}(\omega)^{-1} d\omega,$$

$$\Gamma = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{uu}(\omega) f_{uu}(\omega)^{-2} f_q(\omega) d\omega,$$

and the corresponding bandwidth formula based on minimization of the second order term in the expected Wald statistic is

$$M = \left( \frac{q k_q \Omega^{-1} \Gamma^{1/(q+1)}}{4} \right) n^{1/(q+1)}.$$

Such a formula depends on integrated forms of the spectral density functions, rather than their values at a single point. As a result, we cannot simply substitute consistent nonparametric density estimates for a single point into the formula to obtain a bandwidth choice rule. Instead, we have to estimate the spectral density at points over the interval $[-\pi, \pi]$ and approximate $\Omega$, $\Gamma$, by Riemann sum discrete approximations. Another way to find a bandwidth in this case is to prespecify a parametric model for the error process $u_t$, estimate the parameters and utilize these estimates in a plug-in procedure to obtain an estimator of the optimal bandwidth. The bandwidth rule is then
optimal for this parametric model, and has the correct order of magnitude even when the parametric model is misspecified.

5. Band spectral regression

One of the advantages of frequency domain regression is that it provides a selective procedure which allows us to focus on the most relevant frequencies. In economic applications, it is possible that a model applies more accurately at some frequencies than others, a famous example being the permanent income hypothesis model where permanent consumption is modeled as a linear function of permanent income. In this model, the marginal propensity to consume out of permanent income can well be very different from that out of transitory income. Since it is reasonable to associate permanent income with zero- and low-frequency components in the data, regressions based on high-frequency components can be expected to deliver different estimates of the marginal propensity to consume than those from low frequency components.

The idea of band spectral regression has a long history. Hannan (1963b) proposed averaging over a nondegenerate subset of frequency in regression and Engle (1974) applied it to test the permanent income hypothesis, naming the procedure band spectral regression. Later, Corbae et al. (1994) extended it to cases with stochastic trends, and Corbae et al. (2002) studied the effects of deterministic detrending in the time and frequency domains.

In what follows, we consider for model (1.2) a simple band spectral regression based on averaging over a frequency band \( B \) which contains a subset of fundamental frequencies and is symmetric about the origin. As we have seen in previous sections, due to the dominating effects of the zero frequency components in the nonstationary regressor case, estimators based on a frequency band that includes the zero frequency will have similar asymptotic properties to those of \( \hat{\beta} \) in the previous sections. Thus, the following discussion considers a band \( B \) that concentrates on some non-zero frequencies and we focus on the nonstationary case. In particular, for some frequency \( \omega > 0 \), we set \( B = B_\omega \cup B_{-\omega} \) with

\[
B_\omega = \left[ \omega - \frac{\pi}{2L}, \omega + \frac{\pi}{2L} \right],
\]

where \( L/n + 1/L \to 0 \) as \( n \to \infty \), and define

\[
\ell = \#\{\lambda_\ell \in B_\omega\}.
\]

The band \( B \) then has \( 2\ell \) fundamental frequencies with \( \ell \to \infty \) as \( n \to \infty \), while the band itself shrinks to the symmetric frequencies \( \omega \) and \(-\omega\). We will assume that \( \ell/n^{1/4} \to 0 \). The band \( B \) spectral regression estimator is

\[
\hat{\beta}_B = \left[ \sum_{\lambda_\ell \in B} I_{xx}(\lambda_\ell)\hat{f}_{uu}(\lambda_\ell)^{-1} \right]^{-1} \left[ \sum_{\lambda_\ell \in B} I_{xy}(\lambda_\ell)\hat{f}_{yu}(\lambda_\ell)^{-1} \right],
\]
so that $\sqrt{2\ell} (\hat{\beta}_B - \beta) = \hat{H}_B^{-1} \hat{\Psi}_B$, where
\[
\hat{H}_B = \frac{1}{2\ell} \sum_{\lambda \in B} I_{\text{ct}}(\lambda) \hat{f}_{\text{mu}}(\lambda) - 1 \quad \text{and} \quad \hat{\Psi}_B = \frac{1}{\sqrt{2\ell}} \sum_{\lambda \in B} I_{\text{ct}}(\lambda) \hat{f}_{\text{mu}}(\lambda) - 1.
\]

Under some regularity conditions, notably a linear process condition on $\mu_t$ (see Assumption A in the appendix), we have the following first order result for $\hat{\beta}_B$. 

**Lemma 6.**

\[
\sqrt{2\ell} (\hat{\beta}_B - \beta) \overset{d}{\rightarrow} \text{MN}(0, h_1(\omega)^{-1} f_{\text{mu}}(\omega)),
\]  

where
\[
h_1(\omega) = f_{\text{mu}}(\omega) + \frac{1}{2\pi} \frac{B_1(1)^2}{|1 - e^{i\omega}|^2} = \frac{f_{\text{xx}}(\omega)}{|1 - e^{i\omega}|^2} + \frac{1}{2\pi} \frac{B_1(1)^2}{|1 - e^{i\omega}|^2}.
\]

Here the mixing variate in (5.1) involves $B_1(1)^2$ and appears via $h_1(\omega)$. The Wald statistic for $H_0 (\hat{\beta} = \beta_0)$ in this band spectral regression can be written as $W_B = \hat{\Psi}_B^* \hat{H}_B^{-1} \hat{\Psi}_B$. The expansion for $W_B$ follows steps similar to those for $W$ in Section 4. However, since the band $B$ concentrates on the non-zero frequencies $\omega$ and $-\omega$, the periodogram averages have different orders of magnitude and need to be standardized differently. Lemma A.3 in the appendix gives orders of magnitude for some of the higher order components and is useful in the formal moment expansion for the band spectral regression Wald statistic that follows.

**Theorem 3.**

\[
E[W_B] = \Phi_0 + m^{-1} \Phi_1 + M^{-q} \Phi_2 + o(m^{-1} + M^{-q})
\]
\[
= \hat{W}_B + o(m^{-1} + M^{-q}),
\]

where $\Phi_0, \Phi_1$, and $\Phi_2$ are $O(1)$ quantities defined as
\[
\Phi_0 = E[\hat{H}_B^{-1} \Psi_0^* \Psi_0],
\]
\[
\Phi_1 = m \{ E[\hat{H}_B^{-1} \Psi_{v1}^* \Psi_{v1}] + 2E[\hat{H}_B^{-1} \Psi_0^* \Psi_{v2}] 
+ E[\hat{H}_B^{-3} A_{v1}^2 \Psi_0^* \Psi_0] - E[\hat{H}_B^{-2} A_{v2} \Psi_0^* \Psi_0] \},
\]
\[
\Phi_2 = M^q \{ E[\hat{H}_B^{-2} A_{v1} \Psi_0^* \Psi_0] - 2E[\hat{H}_B^{-1} \Psi_0^* \Psi_{b1}] \}.
\]

$H_B = (1/2\ell) \sum_{\lambda \in B} I_{\text{ct}}(\lambda) \hat{f}_{\text{mu}}(\lambda)^{-1}$, $\Psi_0 = (1/\sqrt{2\ell}) \sum_{\lambda \in B} I_{\text{ct}}(\lambda) \hat{f}_{\text{mu}}(\lambda)^{-1}$, and $\Psi_{v1}$, $\Psi_{v2}$, $\Psi_{b1}$, $A_{v1}$, $A_{v2}$, and $A_{v3}$ are quantities defined in Lemma A.3 in the appendix, and $\hat{W}_B = \Phi_0 + m^{-1} \Phi_1 + M^{-q} \Phi_2$ is the truncated expected Wald statistic.

**Remark 6.** This expansion is similar to that in Section 4. However, given the narrow band spectral regression estimator that we consider in this section, the quantities in the formula have different definitions and different asymptotics. As discussed earlier, for testing (2.2) using the regression Wald test, a bandwidth selection criterion can
be defined by minimizing the expected value of the second order effect in the Wald Statistic, giving the following bandwidth choice:

$$M = \left( \frac{q \Phi_2}{2 \Phi_1} \right)^{1/(q+1)} n^{1/(q+1)}.$$  

Some asymptotic results for the quantities in Theorem 3 are given in Lemma A.4 in the appendix. These results deliver the following explicit approximation for the truncated expected Wald statistic $\tilde{W}_B$,

$$\tilde{W}_B \sim 1 + \frac{3}{m} + \frac{k_q}{M_q} f_{uw}(\omega)^{-1} f_q(\omega).$$

**Remark 7.** Since we are considering a narrow band spectral regression, the higher order terms converge to quantities evaluated at the corresponding frequency $\omega$. Consequently, the second order effect in $\tilde{W}_B$ is asymptotically determined by the spectral density of the residual process at frequency $\omega$. In particular, similar to the results in Section 4, the size of the test based on $\tilde{W}_B$ will be largely affected by the curvature of the spectral density function $f_{uw}(\cdot)$ at $\omega$. Thus, a formula for bandwidth selection and a simple second order modified Wald test can be obtained in a similar way as (4.7) and (4.8) based on consistent estimates of $f_{uw}(\omega)$ and $f_q(\omega)$.

6. The multivariate case

This section considers the multivariate case and the Wald test for a general linear hypothesis with $p$ restrictions. Thus, the null hypothesis has the general format $H_0: R\beta = r$, where $r$ is a $p \times 1$ vector and $R$ is a $p \times k$ matrix. The quantities $x_t$ and $\beta$ are now $k \times 1$ vectors, $I_{xx}(\lambda)$ and $f_{xx}(\lambda)$ are $k \times k$ matrices, and $I_{xy}, I_{xu}$ are $k \times 1$ vectors. The corresponding regression Wald statistic is given by (2.3). Under the null hypothesis that $R\beta = r$, we have

$$W = (\hat{\beta} - \beta)' R'[R\hat{\Sigma}^{-1} R']^{-1} R(\hat{\beta} - \beta).$$

For convenience, we continue to use $\hat{H}, H,$ and $Z$, which are defined by formulae (4.2) – (4.4) in the scalar case, for their matrix counterparts. Then, the Wald statistic can be written as

$$W = Z^* \hat{H}^{-1} R'[R\hat{H}^{-1} R']^{-1} R\hat{H}^{-1} Z.$$

The expansion for $W$ follows lines similar to those of Section 4, and the higher order asymptotics are similar to those in Theorems 1 and 2. However, the expansion is now complicated by the multivariate nature of the model. Detailed expansions are given in the appendix and we only state the major steps of the expansions here. Denoting $R\hat{H}^{-1} R'$ by $J_0$, and $R\hat{H}^{-1} R'$ by $J_0$, we expand $J_0^{-1}$ around $J_0$ to the third order and get the following expansion for $W$,

$$Z^* \hat{H}^{-1} R'[J_0^{-1} - J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1} + J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1}(\hat{J}_0 - J_0)J_0^{-1} - R_3^{-1}]^{-1} R\hat{H}^{-1} Z,$$  

(6.1)
and collecting terms up to Op(\(\varepsilon\)), notice that \(\hat{J}_0 - J_0 = R(\hat{H}^{-1} - H^{-1})R', \hat{H}^{-1}\) can be expanded around \(H\) as \(\hat{H}^{-1} = H^{-1} - H^{-1}DH^{-1} + H^{-1}DH^{-1}DH^{-1} - R_0\), where \(R_0 = H^{-1}DH^{-1}DH^{-1}DH^{-1}\), and \(D\) and \(Z\) follow similar expansions as \((4.6), (4.7)\), substituting all these preliminary expansions into \((6.1)\) and collecting terms up to Op\((m^{-1} + M^{-q})\), we get the following formal moment approximation for the Wald statistic.

**Theorem 4.**

\[
E[W] = S_0 + m^{-1}S_1 + M^{-q}S_2 + o_p(m^{-1} + M^{-q})
= \bar{W}_G + o_p(m^{-1} + M^{-q}),
\]

where \(S_0, S_1\) and \(S_2\) are \(O(1)\) quantities defined as

\[
S_0 = E[Z_0^*H^{-1}R(RH^{-1}R')^{-1}RH^{-1}Z_0],
\]

\[
S_1 = m\{E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Z_{11}] + E[Z_0^*H^{-1}Dv_1H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}Z_0]
- E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}R'J_0^{-1}RH^{-1}Z_0]
+ E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}R'J_0^{-1}RH^{-1}Z_0]
- 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}Z_0]
+ 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}Dv_1H^{-1}Z_0]
+ E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_2H^{-1}R'J_0^{-1}RH^{-1}Z_0]
+ 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_2H^{-1}R'J_0^{-1}RH^{-1}Z_0] - 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Z_{12}]
+ 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}Z_0] - 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_2H^{-1}Z_0]
- E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}R'J_0^{-1}RH^{-1}Z_0]\},
\]

\[
S_2 = m^2\{2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Z_{12}] - 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Z_0]
- 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_1H^{-1}Z_0]
- 2E[Z_0^*H^{-1}R'J_0^{-1}RH^{-1}Dv_2H^{-1}Z_0]\},
\]

and \(\bar{W}_G = S_0 + m^{-1}S_1 + M^{-q}S_2\) is the truncated expected Wald statistic.

**Remark 8.** The orders of magnitude of the second order terms are the same as those in the univariate case. However, these second order effects do depend on the linear hypothesis. The number of restrictions, \(p\), in the null hypothesis appears in both the first and the second order asymptotics. Lemma A.5 in the appendix gives asymptotic results for the higher order terms in the expansion. As a result, the truncated expected Wald statistic \(\bar{W}_G\) has the following approximation, similar to that given in Theorem 2:

\[
\bar{W}_G \sim p + \frac{2p}{m} + \frac{pkq}{M^q} f_{m^{-1}}(0)f_q(0).
\]
Remark 9. Just as in previous sections, an optimal bandwidth rule can be determined by minimizing the second order term, giving formula (4.8), and a second order modified Wald test can then be constructed as in (4.9).

Remark 10. When $x_t$ is a stationary process, the zero frequency no longer has a dominating effect and the value of spectral density elsewhere appears in the asymptotics. As a result, certain terms cannot be cancelled out and the higher order asymptotics of the Wald test have a more complicated representation than in the univariate case. We have been able to show that the Wald statistic in this case has the following moment approximation:

$$E[W] \sim p + \frac{2p}{m} + \frac{k_q}{M^q} \text{tr}\{ (R\Omega^{-1}R')^{-1} R\Omega^{-1} \Gamma \Omega^{-1} R' \},$$

where $\Omega$ and $\Gamma$ are $k \times k$ matrices of integrals of spectral densities given by the same formulae as those in Remark 5.

7. Monte Carlo results

We conducted a small Monte Carlo experiment to evaluate the second order theory for Wald tests in cointegrating regressions of the type considered here. In particular, we examined the size properties of Wald statistics using different bandwidth choices and the second order adjusted Wald test, and looked at the effect of the second order bandwidth selection criterion on test size.

The data were generated according to the model

$$y_t = \beta x_t + u_t,$$
$$x_t = x_{t-1} + v_t,$$
$$u_t = \sigma^2 v_{t-1} + \epsilon_t,$$

where $v_t$ and $\epsilon_t$ are iid $N(0, \sigma_v^2)$ and $N(0, \sigma^2)$ variates and are independent of each other. Three values of autoregressive coefficient, $\alpha = 0.3, 0.6, \text{and } 0.8$ were considered. The null hypothesis $H_0: \beta = 1$ was tested. The corresponding Wald test is asymptotically $\chi^2_1$ and the nominal 5% level critical value for the test based on the asymptotic distribution is 3.84.

Three sample sizes ($n = 64, 128, 256$) were used in the experiment, the number of replications was 10,000, and the Daniell window was used in the density estimation. The size properties of four Wald statistics are examined in the Monte Carlo analysis. They are as follows:

- $W_0$: Wald statistic (2.3) using the optimal bandwidth formula (4.8).
- $W_1$: Wald statistic (2.3) using a fixed bandwidth choice $M = 8$.
- $W_2$: Wald statistic (2.3) using a fixed bandwidth choice $M = 10$.
- $W_{SO}$: second order adjusted Wald statistic (4.9) using the optimal bandwidth.
Different combinations of $\sigma_i^2$ and $\sigma_c^2$ were tried. Although in cases where $x_i$ is endogenous the size distortion is known to be larger for small ratios of $\sigma_i^2/\sigma_c^2$, with exogenous regressors, the effect of $\sigma_i^2/\sigma_c^2$ on the empirical size of the Wald statistic in our model is very small and thus we report results only for $\sigma_i^2 = \sigma_c^2 = 1$ here. The empirical size of the Wald statistic, $W_0$, is smaller than that of the Wald statistics ($W_1$ and $W_2$). The empirical critical values of $W_0$ are also closer to the nominal asymptotic critical values than those of $W_1$ and $W_2$. Size distortion is further reduced in the second order adjusted Wald statistic. As the sample size increases from 64 to 256, the empirical size decreases and the empirical critical values move closer to the
nominal values. However, there is noticeably more size distortion in all cases as the 
autocorrelation in the residual process $u_t$ increases.

8. Conclusion

This paper develops higher order expansions for Wald test statistics in efficient, fre-
quency domain semi-parametric regression models with integrated regressors. These 
expansions address some of the problems presented by first order asymptotic theory. 
In particular, they provide an opportunity to reduce size distortion in statistical testing 
by suitable bandwidth choices and second order adjustments. Since the effect at the 
zero frequency dominates the asymptotics for periodogram averages of $I(1)$ processes, 
the second order effect in the expected value of the Wald statistic is asymptotically 
determined by the spectral density of the residual process at the origin. As a result, 
the size properties of Wald tests are largely affected by the curvature of the spectral 
density function at the origin.

A bandwidth selection criterion that is based on minimizing the second order effect 
on the expected value of the Wald statistic is proposed to improve statistical testing and 
second order modified Wald tests are proposed that use consistent estimates of the sec-
don order terms. Expansions for the Wald test are also developed for a narrow band 
spectral regression estimator. When the frequency band under consideration concen-
trates to some non-zero frequency, the estimator of the cointegrating vector converges 
to its true value at a different rate and, the second order effects in the expected value of 
the Wald statistics are functions of the spectral density at the corresponding frequency. 
Interestingly, some of the effects of the nonstationarity that dominates low frequency 
limit behavior also carry over to high frequency asymptotics in band spectral regression 
models.

Acknowledgements

We thank Peter Robinson, Roger Koenker and two referees for helpful comments 
on earlier versions of this paper. Phillips thanks the NSF for research support under 
Grant Nos. SBR94-22922 and SBR97-30295 and Xiao thanks the Hong Kong Research 
Grants Council for research support under Grant No. CUHK4078/98 and the Cowles 
Foundation for hospitality during a visit in May–June 1998.

Appendix A

This appendix provides derivations of results given in the body of the paper. Second 
order terms are often only sketched and rigorous derivations are not always given. 
Further specific assumptions may also be needed in some cases for rigorous results for 
second order terms to apply. For notational convenience, we denote $A \sim B$ if $A = B(1 + 
o_p(1))$. 
A.1. Proof of Lemma 1

Notice that, by a geometric expansion on \( \hat{f}_{uu}(\lambda_s)^{-1} \),

\[
n^{-1} \sum_s I_{ss}(\lambda_s) \hat{f}_{uu}(\lambda_s)^{-1} = n^{-1} \sum_s I_{ss}(\lambda_s) f_{uu}(\lambda_s)^{-1} + \sum_{k=2}^p (-1)^{k-1} n^{-1} \sum_s I_{ss}(\lambda_s) f_{uu}(\lambda_s)^{-k} (\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^{k-1} + (-1)^p n^{-1} \sum_s I_{ss}(\lambda_s) f_{uu}(\lambda_s)^{-p} (\hat{f}_{uu}(\lambda_s) - f_{uu}(\lambda_s))^{p}.
\]

We show that the leading term, \( n^{-1} \sum_s I_{ss}(\lambda_s) f_{uu}(\lambda_s)^{-1} \), converges weakly to \( [2 \pi f_{uu}(0)]^{-1} \int B_x \, dB_x \) and other terms are asymptotically zeros. Using the Fourier series representation of \( f_{uu}(\lambda_s)^{-1} \) (see, e.g., Naimark, 1960),

\[
f_{uu}(\lambda_s)^{-1} = \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g e^{ig\lambda_s},
\]

we have

\[
n^{-1} \sum_s I_{ss}(\lambda_s) f_{uu}(\lambda_s)^{-1} = \frac{1}{n} \sum_s \left[ \frac{1}{2\pi} \sum_r C_{ss}(r) e^{-i\lambda_s r} \right] \left[ \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g e^{ig\lambda_s} \right] = \frac{1}{n} \left( \frac{1}{2\pi} \right)^2 \sum_{g=-\infty}^{\infty} D_g \sum_r C_{ss}(r) \sum_s e^{i\lambda_s (g-r)}.
\]

Since \( \sum_s e^{i\lambda_s (g-r)} = 0 \) unless \( g - r = l \times n \), for \( l = 0, 1, 2, \ldots \), we have

\[
n^{-1} \sum_s I_{ss}(\lambda_s) f_{uu}(\lambda_s)^{-1} = \left( \frac{1}{2\pi} \right)^2 \sum_r C_{ss}(r) \sum_{|l|=0}^{\infty} D_{r+lsn} = \left( \frac{1}{2\pi} \right)^2 \sum_{|r| \leq M} C_{ss}(r) \sum_{|l|=0}^{\infty} D_{r+lsn} + \left( \frac{1}{2\pi} \right)^2 \sum_{|r| > M} C_{ss}(r) \sum_{|l|=0}^{\infty} D_{r+lsn}.
\]
Following a similar argument to that of Phillips (1991) and by summability of $D_r$, we have
\[
\left(\frac{1}{2\pi}\right)^2 \sum_{|r| \leq M} C_{xx}(r) \sum_{|l| = 0}^{\infty} D_{r+l} = \frac{1}{2\pi} \int B_x dB_u \left(\frac{1}{2\pi} \sum_{r = -\infty}^{\infty} D_r\right)
\]
\[= [2\pi f_{uu}(0)]^{-1} \int B_x dB_u.
\]
\[
\left(\frac{1}{2\pi}\right)^2 \sum_{|r| > M} C_{xx}(r) \sum_{|l| = 0}^{\infty} D_{r+l} \to 0.
\]

Thus
\[n^{-1} \sum_x I_x f_{xx}(\hat{\lambda}_x) f_{uu}(\hat{\lambda}_x)^{-1} = [2\pi f_{uu}(0)]^{-1} \int B_x dB_u. \quad (9.1)
\]

Notice that since $f_{uu}(\hat{\lambda}_x)$ is bounded away from zero and $\hat{f}_{uu}(\hat{\lambda}_x)$ is a uniformly consistent estimator,
\[
\left| n^{-1} \sum_x I_x f_{xx}(\hat{\lambda}_x) f_{uu}(\hat{\lambda}_x)^{-p} \hat{f}_{uu}(\hat{\lambda}_x)^{-1}(\hat{f}_{xx}(\hat{\lambda}_x) - f_{xx}(\lambda_x))^p \right|
\]
\[\leq c_1 n^{-1} \sum_x |I_x f_{xx}(\hat{\lambda}_x)| |f_{xx}(\hat{\lambda}_x)|^{-p-1}(\hat{f}_{xx}(\hat{\lambda}_x) - f_{xx}(\lambda_x))^p
\]
for some $c_1$, and
\[
|\hat{f}_{uu}(\hat{\lambda}_x) - f_{uu}(\hat{\lambda}_x)|^p \leq c_2 \{ |B_x|^p + |V_x|^p + |P_x|^p \}
\]
for some constant $c_2$. Under some regularity conditions, we can obtain uniform rates of convergence for $B_x$, $V_x$, and $P_x$. For $B_x$ and $V_x$, by standard results in spectral density estimation (see, e.g. Brillinger, 1980; Theorem 7.7.4), we have
\[
\sup_t |B_t| = O(M^{-q})
\]
\[
\sup_t |V_t| = O_p(m^{-1/2+\epsilon}) \quad \text{for any } \epsilon > 0.
\]

For $P_x$, under the assumption that $v_x$, $u_x$, and $x_0$ all have uniformly bounded moments, it can be verified that $C_{xx}(h)$ and $n^{-1}C_{xx}(h)$ are uniformly $O_p(1)$ quantities. Thus, the uniform rate of the term $P_t$ is obtained since
\[
P_t = -2(\hat{\beta}_{OLS} - \beta) \text{Re}(\hat{f}_{xx}(\hat{\lambda}_x)) + (\hat{\beta}_{OLS} - \beta)' \hat{f}_{xx}(\hat{\lambda}_x)(\hat{\beta}_{OLS} - \beta),
\]
\[n(\hat{\beta}_{OLS} - \beta) = O_p(1),
\]
where $\hat{\beta}_{OLS}$ is the OLS estimator, and
\[
\hat{f}_{xx}(\hat{\lambda}_x) = \frac{1}{2\pi} \sum_k \left( \frac{h}{M} \right) C_{xx}(h)e^{ih},
\]
\[
\hat{f}_{xx}(\hat{\lambda}_x) = \frac{1}{2\pi} \sum_k \left( \frac{h}{M} \right) C_{xx}(h)e^{ih}.
\]
We have
\[ \sup |f_{uu}(\lambda_s) - f_{uu}(\lambda_s)|^p = O_p(M^{-pq} + m^{-p/2+e}). \]

Giving the bandwidth choice and for \( p > 2 \),
\[ n^{-1} \sum_s I_u(\lambda_s) f_{uu}(\lambda_s)^{-p} \hat{f}_{uu}(\lambda_s)^{-1} f_{uu}(\lambda_s)^p = o_p(1). \]

Now we show that terms like \( n^{-1} \sum_s I_u(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s, n^{-1} \sum_s I_u(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s \)
and \( n^{-1} \sum_s I_u(\lambda_s) f_{uu}(\lambda_s)^{-2} P_s \), etc., are \( o_p(1) \). By definition, \( V_s = \sum_j \alpha_{ij} [I_u(\lambda_j) - f_{uu}(\lambda_j)] = m^{-1} \sum_{\lambda_j \in B} K(\lambda_j - \lambda_s) [I_u(\lambda_j) - f_{uu}(\lambda_j)] \), and using the Fourier series representation of \( I_u(\lambda_j), I_u(\lambda_j), f_{uu}(\lambda_j) \), and \( f_{uu}(\lambda_j)^{-2} \), denoting the sample autocorrelation of \( u_t \) and the autocorrelation function of \( u_t \) as \( C_{ uu}(h) \) and \( \gamma_{ uu}(k) \) respectively, we have
\[
\begin{align*}
&n^{-1} \sum_s I_u(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s \\
&= n^{-1} \sum_s I_u(\lambda_s) f_{uu}(\lambda_s)^{-2} \sum_j \alpha_{ij} [I_u(\lambda_j) - f_{uu}(\lambda_j)] \\
&= n^{-1} \sum_s \left[ \frac{1}{2\pi} \sum_r C_{ uu}(r) e^{-irh} \right] \left[ \frac{1}{2\pi} g = -\infty \right. \\
&\left. \sum_{g = -\infty}^{\infty} B_g e^{ig\lambda_s} \right] \\
&\times \sum_j \alpha_{ij} \left[ \frac{1}{2\pi} \sum_h C_{ uu}(h) e^{-ihj} - \frac{1}{2\pi} \sum_{k = -\infty}^{\infty} \gamma_{ uu}(k) e^{-ikj} \right] \\
&= n^{-1} \sum_s \left[ \frac{1}{2\pi} \sum_r C_{ uu}(r) e^{-irh} \right] \left[ \frac{1}{2\pi} g = -\infty \right. \\
&\left. \sum_{g = -\infty}^{\infty} B_g e^{ig\lambda_s} \right] \\
&\times \frac{1}{2\pi} \sum_j \alpha_{ij} \sum_h [C_{ uu}(h) - \gamma_{ uu}(h)] e^{-ijh} \\
&- n^{-1} \sum_s \left[ \frac{1}{2\pi} \sum_r C_{ uu}(r) e^{-irh} \right] \left[ \frac{1}{2\pi} g = -\infty \right. \\
&\left. \sum_{g = -\infty}^{\infty} B_g e^{ig\lambda_s} \right] \\
&\times \sum_j \alpha_{ij} \left[ \frac{1}{2\pi} \sum_{|k| > n} \gamma_{ uu}(k) e^{-ijk} \right] \\
&= n^{-1} \left( \frac{1}{2\pi} \right)^3 \sum_s \sum_r \sum_{g = -\infty}^{\infty} B_g [C_{ uu}(h) - \gamma_{ uu}(h)] C_{ uu}(r) \sum_j \alpha_{ij} e^{i(g-r)j} e^{-ijh} \\
&- n^{-1} \left( \frac{1}{2\pi} \right)^3 \sum_r \sum_{|k| > n} \sum_{g = -\infty}^{\infty} B_g \gamma_{ uu}(k) C_{ uu}(r) \sum_j \alpha_{ij} e^{i(g-r)j} e^{-ijk}.
\end{align*}
\]
where \((1/2\pi)\sum_{g=-\infty}^{\infty} B_g e^{ig\hat{\omega}}\) is the Fourier series representation of \(f_{uu}(\hat{\omega})^{-2}\). Notice that
\[
\hat{\lambda}_j \in B(\hat{\omega}) = \left\{ \hat{\lambda}_j : \hat{\lambda}_s - \frac{\pi}{2M} < \hat{\lambda}_j \leq \hat{\lambda}_s + \frac{\pi}{2M} \right\} \\
= \left\{ \hat{\lambda}_s - \frac{\pi}{2M} + 2\pi t/n : t = 1, 2, \ldots, m = n/2M \right\}
\]
if we use the Daniell’s spectral window,
\[
\sum_{s} \sum_{j} \omega_{sj} e^{i(g-r)\hat{\omega}} e^{-i\hat{\lambda}_j h} \\
= m^{-1} \sum_{s=1}^{n} \sum_{\hat{\lambda}_j \in B(\hat{\omega})} K(\hat{\lambda}_j - \hat{\lambda}_s) e^{i(g-r)\hat{\omega}} e^{-i\hat{\lambda}_j h} \\
= (\pi m)^{-1} \sum_{s=1}^{n} \sum_{\hat{\lambda}_j \in B(\hat{\omega})} e^{i(g-r)\hat{\omega}} e^{-i\hat{\lambda}_j h} \\
= (\pi m)^{-1} \sum_{s=1}^{n} \sum_{\hat{\lambda}_j \in B(\hat{\omega})} e^{i(g-r)\hat{\omega}} e^{-i(\hat{\lambda}_s - n/2M + 2\pi t/n)h} \\
= (\pi m)^{-1} \sum_{s=1}^{n} \sum_{\hat{\lambda}_j \in B(\hat{\omega})} e^{i(g-r)\hat{\omega}} e^{-i(\hat{\lambda}_s - \pi/2M)h} m \sum_{t=1}^{m} e^{-2i\pi t h/n} \\
= e^{i(k\pi/2M} \left\{ \frac{1}{\pi m} \sum_{t=1}^{m} e^{-2i\pi t h/n} \right\} \left[ \sum_{s=1}^{n} e^{i(g-r-h)\hat{\omega}} \right].
\]
Since \(\sum_{s} e^{i(g-r-h)\hat{\omega}} = 0\) unless \(g - r - h = l = n\) (when \(\sum_{s} e^{i(g-r-h)\hat{\omega}} = n\)) for \(l = 0, \pm 1, \pm 2, \ldots \), we have
\[
n^{-1} \left( \frac{1}{2\pi} \right)^{3} \sum_{r} \sum_{h} \sum_{g=-\infty}^{\infty} B_g [C_{uu}(h) - \gamma_{uu}(h)] C_{xu}(r) \sum_{s} \sum_{j} \omega_{sj} e^{i(g-r-h)\hat{\omega}} e^{-i\hat{\lambda}_j h} \\
= \left( \frac{1}{2\pi} \right)^{3} \sum_{h} [C_{uu}(h) - \gamma_{uu}(h)] e^{i\hat{\omega} \pi/2M} \left[ \frac{1}{\pi m} \sum_{t=1}^{m} e^{-2i\pi t h/n} \right] \\
\left[ \sum_{r} \sum_{|l|=0}^{\infty} B_{r+h+l} C_{xu}(r) \right] \\
= o_p(1),
\]
since
\[
\frac{1}{\pi m} \sum_{t=1}^{m} e^{-2i\pi t h/n}, \sum_{r} \sum_{|l|=0}^{\infty} B_{r+h+l} C_{xu}(r) = O_p(1) \quad \text{for all } h.
\]
Similarly, for the second term,
\[
\begin{align*}
   n^{-1} \left( \frac{1}{2\pi} \right)^3 \sum_{r} \sum_{|k| > n} \sum_{g=-\infty}^{\infty} & B_g \gamma_{uw}(k) C_{xx}(r) \sum_{s} \omega_{s} e^{i(g-r)\lambda_s} e^{-i\lambda_s k} \\
   = \left( \frac{1}{2\pi} \right)^3 \sum_{|k| > n} \gamma_{uw}(k) e^{ik/2M} \left[ \frac{1}{\pi M} \sum_{r} \sum_{|l|=0}^{\infty} B_{r+k+l+u} C_{xx}(r) \right] \\
   = o_p(1).
\end{align*}
\]

The proofs for other terms are similar to that of (9.1). Thus we have
\[
n^{-1} \sum_{s} \omega_{s} \hat{f}_{uw}(\lambda_s)^{-1} \Rightarrow [2\pi f_{uu}(0)]^{-1} \int B_x dB_u.
\]

In a related fashion and as in Phillips (1991), we can show
\[
n^{-2} \sum_{s} \omega_{s} \hat{f}_{uu}(\lambda_s)^{-1} \frac{d}{d\lambda} \Rightarrow [2\pi f_{uu}(0)]^{-1} \int B_x B_u'.
\]

A.2. Proof of Lemma 2

The results for $B_t$ and $V_t$ come from standard analysis of the bias and variance of kernel spectral density estimates.

A.3. Proof of Lemma 3

Denote the preliminary OLS estimator for $\beta$ by $\hat{\beta}$, and notice that
\[
   P_t = -2(\hat{\beta} - \beta) \text{Re} \left( \sum_{s \neq t} \omega_{s} I_{uw}(\lambda_s) \right) + (\hat{\beta} - \beta)' \sum_{s \neq t} \omega_{s} I_{xx}(\lambda_s)(\hat{\beta} - \beta)
\]
\[
   = -2(\hat{\beta} - \beta) \text{Re}(\hat{f}_{uw}(\lambda_t)) + (\hat{\beta} - \beta)' \hat{f}_{xx}(\lambda_t)(\hat{\beta} - \beta)
\]
and
\[
n(\hat{\beta} - \beta) \Rightarrow \left[ \int B_x B_u' \right]^{-1} \int B_x dB_u.
\]

Following Phillips (1991), it can be verified that
\[
   \sum_{s} \omega_{s} I_{uw}(\lambda_s) = \hat{f}_{uw}(\lambda_t)
\]
\[
   = \frac{1}{2\pi} \sum_{m} k \left( \frac{h}{M} \right) C_{uw}(h) e^{i\lambda_s h}
\]
\[
   \sim \frac{1}{2\pi} \sum_{m} k \left( \frac{h}{M} \right) e^{i\lambda_s h} \int B_x dB_u.
\]

Notice that for the Daniell kernel
\[
   K(\lambda) = \frac{1}{2\pi M} \sum_{m} k \left( \frac{h}{M} \right) e^{i\lambda h} = \begin{cases} 
   \pi^{-1} & \text{for } |\lambda| \leq \pi/2, \\
   0 & \text{otherwise}.
\end{cases}
\]
Thus
\[ \sum_s \omega_s I_{xx}(\lambda_s) \sim M \cdot K(M\lambda) \int B_x \, dB_u \]
\[ = \begin{cases} O_p(M) & \text{if } |\lambda| \leq \pi/(2M), \\ 0 & \text{otherwise.} \end{cases} \]

Similarly
\[ \sum_s \omega_s I_{xx}(\lambda_s) = \hat{f}_{xx}(\lambda) \]
\[ = \frac{n}{2\pi} \sum_h k \left( \frac{h}{M} \right) \left[ C_{xx}(h) \right] \frac{e^{ih}}{n} \]
\[ \sim \frac{n}{2\pi} \sum_h k \left( \frac{h}{M} \right) e^{ih} \int B_x B'_x \]
\[ = M n K(M\lambda) \int B_x B'_x \]
\[ = \begin{cases} O_p(nM) & \text{if } |\lambda| \leq \pi/(2M), \\ 0 & \text{otherwise.} \end{cases} \]

Thus, for $|\lambda| > \pi/(2M)$,
\[ P_t = -2(\hat{\beta} - \beta) \sum_{x \neq t} \omega_s I_{xx}(\lambda_s) + (\hat{\beta} - \beta)' \sum_{x \neq t} \omega_s I_{xx}(\lambda_s)(\hat{\beta} - \beta) \]
\[ = O_p(m^{-1}), \]
and for $|\lambda| \leq \pi/(2M)$,
\[ P_t = -2(\hat{\beta} - \beta) \sum_{x \neq t} \omega_s I_{xx}(\lambda_s) + (\hat{\beta} - \beta)' \sum_{x \neq t} \omega_s I_{xx}(\lambda_s)(\hat{\beta} - \beta) \]
\[ \sim -\frac{1}{2m} K(M\lambda) \int dB_u B'_u \left[ \int B_x B'_x \right]^{-1} \int B_x dB_u. \]

A.4. Proof of Lemma 5

Since $g(\lambda)$ has uniformly and absolutely convergent Fourier representation, we can write the sum as
\[ n^{-2} \sum_j I_{xx}(\lambda_j) g(\lambda_j) = n^{-2} \sum_j \int I_{xx}(\lambda_j) \left[ \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g e^{i\lambda j} \right] \]
\[ = n^{-2} \sum_j \left[ \frac{1}{2\pi} \sum_r C_{xx}(r) e^{-i\lambda j} \right] \left[ \frac{1}{2\pi} \sum_{g=-\infty}^{\infty} D_g e^{i\lambda j} \right] \]
\[ = \frac{1}{4\pi^2} n^{-2} \sum_g D_g \sum_r C_{xx}(r) \sum_j e^{i(g-r)\lambda_j} \].
Note that the summations may be interchanged because the infinite series in \( h \) and \( g \) converge absolutely and the sum over \( j \) is finite. Notice that \( \sum e^{i\theta(g-r)} = 0 \) unless \( g - r = n*l \) for some \( l \). Thus,

\[
n^{-2} \sum_j I_{xx}(\lambda_j)g(\lambda_j) = \frac{1}{4\pi^2} n^{-2} \sum_r C_{xx}(r) \sum_{|l|=0}^\infty D_{r+n*l}
\]

\[
\rightarrow d \frac{1}{2\pi} g(0) \int B_x B_x',
\]

**A.5. Lemma A.1**

\( Z_0 = O_p(1), \quad Z_{V1} = O_p(m^{-1/2}), \quad Z_{V2} = O_p(m^{-1}), \quad Z_{B1} = O_p(M^{-q}), \quad Z_{B2} = O_p(M^{-2q}), \quad Z_{P1} = O_p(m^{-1}), \quad Z_{P2} = O_p(M^{-q)} \)

and

\( D_{V1} = O_p(m^{-1/2}), \quad D_{V2} = O_p(m^{-1}), \quad D_{B1} = O_p(M^{-q}), \quad D_{B2} = O_p(M^{-2q}), \quad D_{P1} = O_p(m^{-1}) \).

**Proof.** First, we show that \( D_{V1} = O_p(m^{-1/2}) \) by verifying that \( E[D_{V1}]^2 = O_p(m^{-1}) \).

Notice that

\[
D_{V1} = \frac{1}{n^2} \sum_t I_{xx}(\hat{\lambda}_t) f_{uu}(\hat{\lambda}_t) - \sum_{s \neq t} \omega_{ts} [I_{uu}(\hat{\lambda}_s) - f_{uu}(\hat{\lambda}_s)]
\]

\[
= \frac{1}{n} \sum_s \varphi_s [I_{uu}(\hat{\lambda}_s) - f_{uu}(\hat{\lambda}_s)],
\]

where

\[
\varphi_s = \sum_t \omega_{ts} \left( \frac{1}{n} I_{xx}(\hat{\lambda}_t) f_{uu}(\hat{\lambda}_t) - 1 \right)
\]

\[
\sim \frac{1}{n} f_{uu}(\hat{\lambda}_s)^{-2} \sum_t \omega_{ts} I_{xx}(\hat{\lambda}_t)
\]

\[
\sim \frac{1}{n} f_{uu}(\hat{\lambda}_s)^{-2} nM \cdot K(M_{\hat{\lambda}_s}) \int B_x^2
\]

\[
\sim \begin{cases} 
  Mf_{uu}(\hat{\lambda}_s)^{-2} K(0) \int B_x^2 & \text{if } |\hat{\lambda}_s| \leq \pi/(2M), \\
  0 & \text{otherwise.}
\end{cases}
\]

Under the regularity conditions given in Section 2, \( E[I_{uu}(\hat{\lambda}_s) - f_{uu}(\hat{\lambda}_s)] = O(n^{-1}) \), uniformly in \( \hat{\lambda}_s \), (see, e.g., Brillinger, 1980, Theorem 5.22), and it can be shown that
\[ ED_{V1} = O(n^{-1}). \] For the second moment, notice that \( Var(I_{uu}(\lambda_t)) \sim f_{uu}(\lambda_t)^2, \) and

\[
E[D_{V1}]^2 \sim \frac{1}{n^2} \sum_{|\lambda_s| \leq \pi(2M)} M^2 f_{uu}(\lambda_s)^{-4} K(0)^2 E\left( \int B_x^2 \right)^2 f_{uu}(\lambda_s)^2 \\
+ \frac{1}{n^2} \sum_{s \neq t} \phi_s \phi_t \, Cov[I_{uu}(\lambda_s), I_{uu}(\lambda_t)].
\]

Since \( \lambda_s \) and \( \lambda_t \) are fundamental frequencies, \( Cov[I_{uu}(\lambda_s), I_{uu}(\lambda_t)] \) is of order \( O(n^{-1}) \), uniformly in \( s \neq t \). Thus,

\[
E[D_{V1}]^2 \sim \frac{1}{(2m)^2} K(0)^2 \sum_{|\lambda_s| \leq \pi(2M)} E\left( \int B_x^2 \right)^2 f_{uu}(\lambda_s)^{-2} \\
\sim \frac{1}{4m} K(0)^2 E\left( \int B_x^2 \right)^2 f_{uu}(0)^{-2}.
\]

As a result, \( D_{V1} = O_p(m^{-1/2}) \). Next,

\[
ED_{V2} = E \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3} V_t^2 \\
\sim E \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-3} \frac{1}{m} f_{uu}(\lambda_t)^2 \\
\sim \frac{1}{m} E \left[ \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-1} \right] \\
\sim \frac{1}{2 \pi m} f_{uu}(0)^{-1} E \int B_x^2 \\
= O_p(m^{-1}).
\]

For \( D_{B1} \), by an application of Lemma 5,

\[
D_{B1} = \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} B_t \\
\sim \frac{1}{n^2} \sum_t I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} M^{-q} k_q f_q(\lambda_t) \\
\sim M^{-q} k_q \frac{1}{2 \pi} f_{uu}(0)^{-2} f_q(0) \int B_x^2 \\
= O_p(M^{-q}),
\]

and similarly for \( D_{B2} \) and

\[
D_{P1} \sim \frac{1}{n^2} \sum_{|\lambda_s| \leq \pi(2M)} I_{xx}(\lambda_t) f_{uu}(\lambda_t)^{-2} \left[ \frac{1}{2m \pi} \left( \int dB_x B_x^2 \right)^2 \left( \int B_x^2 \right)^{-1} \right] \\
= O_p(m^{-1}).
\]
Lemma 6 (see (9.5) below), we have

\[ E \left| Z_0 \right|^2 = E \frac{1}{n^2} \sum_{t} I_u \lambda_{t} H_{w_{2}}^{-1} \right( \lambda_{t} f_{w_{2}}(\lambda_{t}) \right) \]

+ \frac{1}{n^2} \sum_{s \neq t} w_{s}(\lambda_{s})^* w_{s}(\lambda_{s}) w_{u}(\lambda_{s})^* H_{w_{2}}^{-1} \right( \lambda_{s} f_{w_{2}}(\lambda_{s}) \right) \right) \]

\[ = A + B \]

\[ \sim E \frac{1}{n^2} \sum_{t} I_u \lambda_{t} H_{w_{2}}^{-1} \right( \lambda_{t} f_{w_{2}}(\lambda_{t}) \right) \]

Under some regularity conditions and if the serial correlation in \( u_t \) is weak, the second term \( B \) is seen to be of smaller order of magnitude than the leading term. For example, if we assume that \( u_t \) is the linear process \( u_t = C(L)e_t \), with \( C(L) = \sum_{j=0}^{\infty} c_j L^j \), and where the \( e_t \) are iid—see Assumption A, below) then, as discussed in the proof of Lemma 6 (see (9.5) below),

\[ w_u(\lambda_s) = C(e^{i\lambda_s}) w_u(\lambda_s) + r(\lambda_s), \]

where the \( w_u(\lambda_s) \) are uncorrelated and the remainder term \( r(\lambda_s) \) is small. Under the summability condition \( \sum_{j=1}^{\infty} j^\rho |c_j| < \infty \), for \( \rho > 1 \), an explicit expression for \( r(\lambda_s) \) is available from the spectral BN decomposition (see (9.5) below). It is then easily seen that the order of magnitude of \( r(\lambda_s) \) is \( O_p(n^{-1/2}) \) uniformly in \( \lambda_s \) (c.f. Priestley, 1981, Theorem 6.2.1) and that \( E|w_u(\lambda_s)r(\lambda_s)| = O(1/\sqrt{n}) \), \( E|r(\lambda_s)|^2 = O(1/n) \), and \( E|r(\lambda_s)|^2 |r^*(\lambda_s)|^2 = O(1/n^2) \), uniformly in \( \lambda_s, \lambda_t \). Then, writing

\[ B = E \frac{1}{n^2} \sum_{s \neq t} w_s(\lambda_s)^* w_s(\lambda_s) H_{w_s}(\lambda_s) r^*(\lambda_s) f_{w_s}(\lambda_s)^{-1} f_{w_s}(\lambda_s)^{-1} \]

\[ + E \frac{1}{n^2} \sum_{s \neq t} w_s(\lambda_s)^* w_s(\lambda_s) H_{w_s}(\lambda_s) r^*(\lambda_s) H_{w_s}(\lambda_s)^{-1} f_{w_s}(\lambda_s)^{-1} \]

\[ + E \frac{1}{n^2} \sum_{s \neq t} w_s(\lambda_s)^* w_s(\lambda_s) H_{w_s}(\lambda_s) r^*(\lambda_s) f_{w_s}(\lambda_s)^{-1} f_{w_s}(\lambda_s)^{-1} \]

\[ = B_1 + B_2 + B_3, \]

we have

\[ E \frac{1}{n^2} \sum_{s \neq t} w_s(\lambda_s)^* w_s(\lambda_s) H_{w_s}(\lambda_s) r^*(\lambda_s) f_{w_s}(\lambda_s)^{-1} f_{w_s}(\lambda_s)^{-1} \]

\[ \leq \max_{s,t} \{ |C(e^{i\lambda_s})|E\{ |w_s(\lambda_s)||r^*(\lambda_s)| \} \} \]

\[ \times E \left| I_{xx}(\lambda_s) I_{xx}(\lambda_t) \right|^{1/2} f_{w_s}(\lambda_s)^{-1} f_{w_s}(\lambda_s)^{-1}, \]
and

\[ E \frac{1}{n^2} \sum_{s \neq t} w_s(\lambda_s)^2 w_t(\lambda_t) r(\lambda_s) r^*(\lambda_t) f_{uw}(\lambda_s)^{-1} f_{uw}(\lambda_t)^{-1} \]

\[ \leq \max_{s,t} \{ E[r(\lambda_s)^2] |r^*(\lambda_t)|^2 \} \frac{1}{n^2} \sum_{s \neq t} \{ E[I_{xx}(\lambda_s) E[I_{xx}(\lambda_t)]^{-1} \} f_{uw}(\lambda_s)^{-1} f_{uw}(\lambda_t)^{-1}, \]

so that each of the terms \( B_1, B_2, \) and \( B_3, \) tends to zero. It follows that

\[ E[|Z_0|^2] \sim E \frac{1}{n^2} \sum_{t} I_{xx}(\lambda_t) f_{uw}(\lambda_t)^{-1} \sim E \frac{1}{2\pi} f_{uw}(0)^{-1} \int B_x^2. \]

Similarly

\[ E[|Z_1|^2] \sim E \frac{1}{n^2} \sum_{t} I_{xx}(\lambda_t) f_{uw}(\lambda_t)^{-2} \]

\[ \sim \frac{1}{m} E \frac{1}{n^2} \sum_{t} I_{xx}(\lambda_t) f_{uw}(\lambda_t)^{-1} \]

\[ \sim \frac{1}{m} \frac{1}{2\pi} f_{uw}(0)^{-1} E \int B_x^2, \]

\[ E[|Z_2|^2] \sim E \frac{1}{n^2} \sum_{|\lambda_t| \geq \pi/(2M)} I_{xx}(\lambda_t) I_{uw}(\lambda_t) f_{uw}(\lambda_t)^{-2} \]

\[ \times \left[ \frac{1}{2m\pi} \left( \int dB_u B_s \right)^2 \left( \int B_x^2 \right)^{-1} \right]^2 \]

\[ \sim E \frac{1}{2m^2 \pi^3} f_{uw}(0)^{-1} \left( \frac{1}{2\pi} \left[ \int dB_u B_s \right] \right)^2 \left( \int B_x^2 \right)^{-1} \]

\[ = O_p(m^{-2}), \]

\[ E[|Z_3|^2] \sim \frac{1}{m^2} E \frac{1}{n^2} \sum_{t} I_{xx}(\lambda_t) f_{uw}(\lambda_t)^{-1} = O_p(m^{-2}), \]

\[ E[|Z_{B1}|^2] \sim E \frac{1}{n^2} \sum_{t} I_{xx}(\lambda_t) f_{uw}(\lambda_t)^{-3} M^{-2q} k^2 \]

\[ \sim M^{-2q} k^2 \frac{1}{2\pi} f_{uw}(0)^{-3} f_q(0)^{2} E \int B_x^2 \]

\[ = O_p(M^{-2q}). \]

Similar arguments give the result for \( Z_{B2}. \)

\[ A.6. \text{ Proof of Theorem 1} \]

Notice that

\[ W = Z^* \hat{H}^{-1} Z, \quad (9.2) \]
and

\[ \hat{H}^{-1} = H^{-1} - H^{-2}D + H^{-3}D^2 - R_2, \]

\[ Z = Z_0 - Z_{V1} + Z_{V2} - Z_{B1} + Z_{B2} - Z_{P1} + R_3, \]

\[ D = -D_{V1} + D_{V2} - D_{B1} + D_{B2} - D_{P1} + R_4. \]

The terms in these expansions are analyzed in Lemma A.1. Drop the higher order terms and substitute the truncated versions of \( Z, D, \) and \( \hat{H}^{-1} \) into (9.2), giving

\[ W \sim H^{-1}Z_0^2 + H^{-1}Z_{V1}^2 - 2H^{-1}Z_0Z_{P1} + H^{-2}D_{P1}Z_0^*Z_0^* + H^{-3}D_{V1}Z_0^2 \]

\[ + H^{-1}Z_{B1}^2 + H^{-3}D_{B1}^2Z_0^2 - 2H^{-2}D_{B1}Z_0Z_{B1} - H^{-2}D_{B2}Z_0^2 \]

\[ - 2H^{-1}Z_0Z_{B1} + H^{-2}D_{B1}Z_0^2 \]

\[ + H^{-2}D_{V1}Z_0^2 - 2H^{-1}Z_0Z_{V1} \]

\[ + 2H^{-1}Z_0Z_{V2} - H^{-2}D_{V2}Z_0^2 \]

\[ + 2H^{-1}Z_0^2Z_{B2} \]

\[ + 2H^{-1}Z_{V1}Z_{B1} - 2H^{-2}D_{B1}Z_0Z_{V1}. \]

It can be verified that, besides the leading term that is of order \( \mathcal{O}(q) \), the above approximation contains terms of order \( m^{-1}, m^{-q}, \) and \( M^{-2q} \), etc. The bias effect of order \( M^{-q} \) dominates the \( \mathcal{O}(M^{-2q}) \) terms, giving the results in Theorem 1.

### A.7. Lemma A.2

1. \( mE[H^{-1}Z_0^2Z_{V2}] \to 1; \)
2. \( mE[H^{-2}D_{V2}Z_0^2] \to 1; \)
3. \( mE[H^{-3}D_{V1}^2Z_0^2] \to 1; \)
4. \( mE[H^{-1}Z_{V1}^2Z_{V1}] \to 1; \)
5. \( mE[H^{-1}Z_0^2Z_{P1}] \to 1; \)
6. \( mE[H^{-2}D_{P1}Z_0^3] \to 0; \)
7. \( M^q E[H^{-1}Z_0^2Z_{B1}] \to -k_q f_{\text{ua}}(0)^{-1} f_q(0); \)
8. \( M^q E[H^{-2}D_{B1}Z_0^2Z_0] \to -k_q f_{\text{ua}}(0)^{-1} f_q(0); \)
9. \( M^{2q} E[H^{-1}Z_{B1}^2Z_{B1}] \to k_q^2 f_{\text{ua}}(0)^{-2} f_q(0)^2; \)
10. \( M^{2q} E[H^{-3}D_{B1}^2Z_0^2Z_0] \to k_q^2 f_{\text{ua}}(0)^{-2} f_q(0)^2; \)
11. \( M^{2q} E[H^{-1}Z_0^2Z_{B2}] \to k_q^2 f_{\text{ua}}(0)^{-2} f_q(0)^2; \)
12. \( M^{2q} E[H^{-2}D_{B1}Z_0^3Z_{B1}] \to k_q^2 f_{\text{ua}}(0)^{-2} f_q(0)^2; \)
13. \( M^{2q} E[H^{-2}D_{B2}Z_0^3Z_0] \to k_q^2 f_{\text{ua}}(0)^{-2} f_q(0)^2. \)
Proof.

\[
E[H^{-1}Z_0^2] = E\left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]^{-1} \left[\frac{1}{n^2} \sum_s I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]^2 \\
\sim E\left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]^{-1} \left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-2}\right] \\
\rightarrow 1,
\]

and

\[
H^{-1}Z_1^2 = \left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]^{-1} \left[\frac{1}{n^2} \sum_s I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{2} V_s\right].
\]

Since the expectation of this term is zero, we check the order of the second moment. Notice that

\[
H \rightarrow_d \frac{1}{2\pi f_{uu}(0)} \int B_x^2,
\]

by Lemma 5, and

\[
E[Z_1^2] = E\left[\frac{1}{n^2} \sum_s \sum_t I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s I_{uu}(\lambda_t) f_{uu}(\lambda_t)^{-2} V_t\right] \\
\sim E \left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-4} I_{uu}(\lambda_s)V_s^2\right].
\]

Since the leave-one out estimator is used,

\[
E[Z_1^2] \sim \frac{1}{n^2} E[I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-4} E[I_{uu}(\lambda_s)]E[V_s]^2] \\
\sim E \left[\frac{1}{m n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right] \\
\sim E \left\{ \frac{1}{m} \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right\},
\]

\[
mE[H^{-1}Z_1^2] = mE\left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]^{-1} \left[\frac{1}{n^2} \sum_s I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s\right]^2 \\
\sim mE \left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right]^{-1} \left[\frac{1}{n^2} \sum_s I_{xx}(\lambda_s) I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s^2\right] \\
\sim a = 1,
\]

and so

\[
mE[Z_{12}^*H^{-1}Z_0] \sim mE H^{-1} \left[\frac{1}{\sqrt{n}} \sum_s I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-1}\right] \left[\frac{1}{\sqrt{n}} \sum_s I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s\right] \\
\sim mE H^{-1} \left[\frac{1}{n} \sum_s I_{xx}(\lambda_s) I_{uu}(\lambda_s) f_{uu}(\lambda_s)^{-4} V_s^2\right] \\
\sim a = 1.
\]
Similarly,

\[ mE[Z_n^* H^{-2} D V_2 Z_0] \sim a = 1, \]

\[ mE[H^{-3} D V_1 Z_n^* Z_0] \sim m \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right]^{-3} \]

\[ \frac{1}{4m\pi^2} \left( \int B_x^2 \right)^2 f_{uu}(0)^{-2} \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right] \sim 1. \]

E[H^{-1} Z_n^* Z_{P1}]

\[ \sim -E \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x B_x' \right]^{-1} \left[ \frac{1}{n} \sum I_{xx}(\lambda_x) f_{uu}(\lambda_x)^{-1} \right] \times \left[ \frac{1}{n^2} \sum |\lambda_x| \leq n(2M) I_{xx}(\lambda_x) f_{uu}(\lambda_x)^{-2} \right] \]

\[ \times \left( \frac{1}{2m\pi} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_u \right) \]

\[ \sim -E \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right]^{-1} \left[ \frac{1}{n^2} \sum |\lambda_x| \leq n(2M) I_{xx}(\lambda_x) f_{uu}(0)^{-2} + o(1) \right] \]

\[ \times \left( \frac{1}{2m\pi} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_u \right) \]

\[ \sim -E \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right]^{-1} \left[ \frac{m}{n^2} \left( \frac{1}{m} \sum |\lambda_x| \leq n(2M) I_{xx}(\lambda_x) \right) f_{uu}(0)^{-2} \right] \]

\[ \times \left( \frac{1}{2m\pi} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_u \right) \]

\[ \sim -E \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right]^{-1} \left[ \frac{m}{n^2} \hat{I}_{xx}(0) f_{uu}(0)^{-2} \right] \]

\[ \times \left( \frac{1}{2m\pi} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_u \right) \]
\[
\]
\[
\sim -E \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} \int B_x^2 \right]^{-1} \left[ \frac{1}{2\pi} f_{uu}(0)^{-2} \int B_x^2 \right] \\
\times \left( \frac{1}{2\pi} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_x \right) \\
\sim -\frac{1}{m} E \left( \frac{1}{2\pi} f_{uu}(0)^{-1} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_x \right).
\]

Notice that \( B_u(r) = BM(2\pi f_{uu}(0)) = d \left[ 2\pi f_{uu}(0) \right] ^ {1/2} W_u(r) \), where \( W_u(r) \) is a standard Brownian motion independent of \( B_x(r) \). Thus,
\[
E \left( \frac{1}{2\pi} f_{uu}(0)^{-1} \int dB_x B_x \left[ \int B_x^2 \right]^{-1} \int B_x dB_x \right) = 1,
\]
and
\[
E[H^{-1}Z_{P1}^* Z_1] \sim -m^{-1}.
\]
\[
E[Z_0^* H^{-2} D_{B1} Z_0] \sim EH^{-2} D_{B1} \left[ \frac{1}{\sqrt{n}} \sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-1} \right] \\
\times \left[ \frac{1}{\sqrt{n}} \sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-1} \right] \\
\sim -M^{-q} k_q f_{uu}(0)^{-1} f_q(0),
\]
\[
E[Z_{B1}^* H^{-1} Z_0] = EH^{-1} \left[ \frac{1}{\sqrt{n}} \sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_x \right] \left[ \frac{1}{\sqrt{n}} \sum I_{ux}(\lambda_s) f_{uu}(\lambda_s)^{-1} \right] \\
\sim -M^{-q} k_q f_{uu}(0)^{-1} f_q(0),
\]
\[
E[Z_{B1}^* H^{-1} Z_{B1}] = EH^{-1} k_q^2 M^{-2q} \frac{1}{n^2} \sum I_{sx}(\lambda_s) f_{uu}(\lambda_s)^2 f_q(\lambda_s)^3 f_u(\lambda_s)^{-4} \\
\sim EH^{-1} k_q^2 M^{-2q} \frac{1}{n^2} \sum I_{sx}(\lambda_s) f_q(\lambda_s)^2 f_u(\lambda_s)^{-3} \\
= EH^{-1} k_q^2 M^{-2q} \frac{1}{n^2} \sum I_{sx}(\lambda_s) f_q(\lambda_s)^2 f_u(\lambda_s)^{-3} \\
= EM^{-2q} k_q^2 H^{-1} \frac{1}{n^2} \sum \left[ \frac{1}{2\pi} \sum C_{sx}(r)e^{-ir\lambda_s} \right] \\
\times \left[ \frac{1}{2\pi} \sum C_{u}(l) C_{uu}(j) |l|^q |l|^q e^{-it(l+j)\lambda_s} \right] \\
\times \left[ \frac{1}{2\pi} \sum D_{g,h,p} e^{i(g+h+p)\lambda_s} \right] 
\]
\[
\begin{align*}
&= EM^{-2q}k_q^2H^{-1} \frac{1}{n^2} \frac{1}{(2\pi)^3} \sum_{g,h,p,l,j} c_{uw}(j)c_{uw}(l)|j|^q|l|^qD_qD_pD_lD_p

&\times \sum_r C_{xx}(r) \left( \sum_{t=-n/2+1}^{n/2} e^{ig+h+p-j-l-r} \right)

&= EM^{-2q}k_q^2H^{-1} \frac{1}{n^2} \left[ \frac{1}{(2\pi)^3} \sum_{g,h,p,l,j} c_{uw}(j)c_{uw}(l)|j|^q|l|^qD_qD_pD_lD_pC_{xx}(r) \right]

&\sim M^{-2q}k_q^2H^{-1} \frac{1}{2\pi} f_{uu}(0)^{-3} f_q(0)^2 \int B_xB_x'

&\sim M^{-2q}k_q^2 f_{uu}(0)^{-2} f_q(0)^2.
\end{align*}
\]

Similarly,
\[
\begin{align*}
E[Z_0^*H^{-3}D_B^2Z_0] &\sim EH^{-3}k_q^2M^{-2q} \left[ \frac{1}{n^2} \sum_j I_{xx}(\lambda_j) f_q(\lambda_j) f_u(\lambda_j) \right]^2

&\times \left[ \frac{1}{n^2} \sum_j I_{xx}(\lambda_j) f_u(\lambda_j) \right]^{-1}

&\sim EH^{-4}k_q^2M^{-2q} \left[ \frac{1}{2\pi} f_{uu}(0)^{-1} f_q(0) \int B_x^2 \right]^2

&\times \left[ \frac{1}{2\pi} f_{uu}(0) \int B_x^2 \right]^{-1}

&= k_q^2M^{-2q} f_{uu}(0)^{-2} f_q(0)^2,
\end{align*}
\]

and
\[
\begin{align*}
E[Z_B^*H^{-2}D_B^2Z_0] &\sim EH^{-2} \left[ \frac{1}{n} \sum_j I_{xx}(\lambda_j) f_u(\lambda_j) \right]^2

&\times \left[ \frac{1}{n} \sum_j I_{xx}(\lambda_j) f_u(\lambda_j) \right]^{-1}

&\sim k_q^2M^{-2q} f_{uu}(0)^{-2} f_q(0)^2.
\end{align*}
\]

Finally,
\[
\begin{align*}
E[Z_0^*H^{-2}D_B^2Z_0] &\sim k_q^2M^{-2q} f_{uu}(0)^{-2} f_q(0)^2,
\end{align*}
\]

and
\[
\begin{align*}
E[Z_0^*H^{-2}D_B^2Z_0] &\sim k_q^2M^{-2q} f_{uu}(0)^{-2} f_q(0)^2.
\end{align*}
\]
A.8. Proof of Lemma 6

We have
\[
\sqrt{2T}(\hat{\beta}_B - \beta) = \left[ \frac{1}{2T} \sum_{\lambda \in B} I_{xx}(\lambda) \hat{f}_{w}(\lambda) \right]^{-1} \left[ \frac{1}{\sqrt{2T}} \sum_{\lambda \in B} I_{w}(\lambda) \hat{f}_{w}(\lambda) \right],
\]
and, following the proof of Lemma 1, we perform a geometric expansion on \( \hat{f}_{w}(\lambda) \), and analyze the leading terms. In particular, we need
\[
\frac{1}{2} \sum_{\lambda \in B} I_{xx}(\lambda) \hat{f}_{w}(\lambda) \rightarrow_d \text{fuu}(\lambda),
\]  
(9.3)
and we also need the following central limit theorem:
\[
\frac{1}{\sqrt{2T}} \sum_{\lambda \in B} I_{xx}(\lambda) \hat{f}_{w}(\lambda) \rightarrow_d \text{MN}(0, \text{fuu}(\lambda)).
\]  
(9.4)

We start with (9.4) and prove (9.3) en route. For convenience in proving this CLT, we assume that \( u_t \) follows a linear process (Phillips and Solo, 1992):

**Assumption A.1.** \( u_t \) is independent with \( v_t \) and has a linear process representation:
\[
u_t = C(L)u_t, \quad C(L) = \sum_{j=0}^{\infty} c_j L^j,
\]
where \( L \) is the lag operator defined as \( L u_t = v_{t-1} \), \( C(1) \neq 0 \), \( \sum_{j=1}^{\infty} j^j |c_j| < \infty \), for some \( \rho \geq 1 \). \( \bar{v}_t \) is an i.i.d. process with zero mean and finite 4th moments.

The linear process assumption is satisfied by a wide class of time series models and has a long history in frequency-domain time series analysis—see, for example, Hannan (1960), Walker (1965) and Priestley (1981). Under the summability condition in Assumption A.1, it is known (e.g. Phillips and Solo, 1992) that partial sums of \( u_t \) satisfy the invariance principle
\[
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t \rightarrow_d B_\rho(r) \equiv BM(2\pi f_w(0)).
\]
In addition, we have the following spectral decomposition of the operator:
\[
C(L) = C(e^{i\omega}) + \tilde{C}_\omega(L) (L^e - 1),
\]
where
\[
\tilde{C}_\omega(L) = \sum_{j=0}^{\infty} \tilde{c}_j L^j, \quad \tilde{c}_j = e^{-ij\omega} \left( \sum_{l=j+1}^{\infty} c_l e^{il\omega} \right).
\]

Thus, as in Phillips and Solo (1992), we get the decomposition
\[
u_t = C(L)u_t = C(e^{i\omega})u_t + e^{-i\omega} \tilde{v}_{t-1} - \tilde{v}_t,
\]
where
\[
\tilde{v}_t = \tilde{C}_\omega(L)u_t.
\]
Then,
\[
\tilde{w}_t(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} u_t e^{it\lambda} = C(e^{i\omega})w_t(\lambda) + \frac{1}{\sqrt{2\pi T}} (\tilde{v}_{t,0} - e^{-i\omega \tilde{v}_{t,n}}
\]
\[
= C(e^{i\omega})w_t(\lambda) + r_n(\lambda),
\]  
(9.5)
Notice that the $\varepsilon_t$ are i.i.d. $(0, \sigma^2)$, and the $w_t(\lambda_s)$ form an uncorrelated sequence. It is readily seen from (9.5) that $r_n(\lambda_s)$ is $O_p(1/\sqrt{n})$, uniformly in $\lambda_s$.

By definition, the discrete transform of the $I(1)$ process $x_t$ is

$$w_t(\lambda_s) = n^{-1/2} \sum_{t=1}^{n} x_{t-1} e^{i\lambda_s t} + w_t(\lambda_s)$$

$$= e^{i\lambda_s}[w_t(\lambda_s) - n^{-1/2}(e^{i\lambda_s n} x_n - x_0)] + w_t(\lambda_s),$$

so that, as shown in Corbae et al. (2002),

$$w_t(\lambda_s) = \frac{1}{1 - e^{i\lambda_s}} w_t(\lambda_s) - \frac{e^{i\lambda_s}}{1 - e^{i\lambda_s}} \frac{x_n - x_0}{n^{1/2}}.$$  \hspace{1cm} (9.6)

Using this result and the continuity of $f_{uu}$ it is a straightforward matter to show that (9.3) holds. We also have

$$\frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_u(\lambda_s) f_{uu}(\lambda_s)^{-1}$$

$$= \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} f_{uu}(\lambda_s)^{-1} \left[ \frac{1}{1 - e^{i\lambda_s}} w_t(\lambda_s) w_n^*(\lambda_s) \right]$$

$$- \frac{x_n - x_0}{n^{1/2}} \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} f_{uu}(\lambda_s)^{-1} \left[ \frac{e^{-i\lambda_s}}{1 - e^{i\lambda_s}} \right] w_n^*(\lambda_s).$$

From (9.5)

$$\frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} I_u(\lambda_s) f_{uu}(\lambda_s)^{-1}$$

$$= \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} f_{uu}(\lambda_s)^{-1} w_t(\lambda_s) w_n^*(\lambda_s) C(e^{i\lambda_s})^*$$

$$+ \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} f_{uu}(\lambda_s)^{-1} w_t(\lambda_s) r_n^*(\lambda_s),$$  \hspace{1cm} (9.7)

and by (9.6),

$$w_t(\lambda_s) r_n^*(\lambda_s) = \left[ \frac{1}{1 - e^{i\lambda_s}} w_t(\lambda_s) r_n^*(\lambda_s) - \frac{x_n - x_0}{n^{1/2}} \frac{e^{-i\lambda_s}}{1 - e^{i\lambda_s}} r_n^*(\lambda_s) \right].$$

Note that $w_t(\lambda_s) r_n^*(\lambda_s)$ has zero mean, variance of $O(n^{-1})$ and is asymptotically uncorrelated over frequencies because $w_t(\lambda_s)$ is asymptotically uncorrelated across frequencies and is independent of $r_n^*(\lambda_s)$ for all $\lambda_s$. It follows that

$$E \left( \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} f_{uu}(\lambda_s)^{-1} w_t(\lambda_s) r_n^*(\lambda_s) \right)^2 = O\left( \frac{1}{n} \right).$$

Also, since $E|r_n^*(\lambda_s)| = O(n^{-1/2})$ uniformly in $\lambda_s$ we have

$$E \left| \frac{x_n - x_0}{n^{1/2}} \frac{1}{\sqrt{2\ell}} \sum_{\lambda_s \in B} f_{uu}(\lambda_s)^{-1} \frac{e^{-i\lambda_s}}{1 - e^{i\lambda_s}} r_n^*(\lambda_s) \right| = O\left( \sqrt{\frac{1}{n}} \right) = o(1).$$
since $\ell/n \to 0$, and we deduce that
\[ \frac{1}{\sqrt{2\ell}} \sum_{\lambda \in B} I_{\omega}(\lambda) f_{\omega}(\lambda)^{-1} = o_p(1). \]
Thus,
\[ \frac{1}{\sqrt{2\ell}} \sum_{\lambda \in B} I_{\omega}(\lambda) f_{\omega}(\lambda)^{-1} = \frac{1}{\sqrt{2\ell}} \sum_{\lambda \in B} f_{\omega}(\lambda)^{-1} w_{\lambda}(\lambda) w_{\lambda}^*(\lambda) C(e^{i\lambda})^* + o_p(1) \]
\[ = f_{\omega}(\omega)^{-1} \left[ \frac{1}{\sqrt{2\ell}} \sum_{\lambda \in B} w_{\lambda}(\lambda) w_{\lambda}^*(\lambda) C(e^{i\lambda})^* \right] + o_p(1) \]
\[ = f_{\omega}(\omega)^{-1} \left[ \frac{2}{\sqrt{2\ell}} \text{Re} \left\{ \sum_{\lambda \in B} w_{\lambda}(\lambda) w_{\lambda}^*(\lambda) C(e^{i\lambda})^* \right\} \right] + o_p(1). \]
Since $B_0$ is a shrinking band, the $w_{\lambda}(\lambda)$ are asymptotically independent \(^2\) $N(0, \sigma^2/\ell)$ and are independent of $w_{\lambda}(\lambda)$. Thus,
\[ \frac{1}{\sqrt{2\ell}} \sum_{\lambda \in B_0} w_{\lambda}(\lambda) w_{\lambda}^*(\lambda) C(e^{i\lambda})^* \to_d MN(0, \sigma^2 C(e^{i\omega})^*) \]
\[ = MN(0, h_{\lambda}(\omega) f_{\omega}(\omega)), \]
and since the real and imaginary parts of this limit distribution have conditional variance
\[ \frac{1}{\ell} h_{\lambda}(\omega) f_{\omega}(\omega), \]
it follows that
\[ \sqrt{\frac{2}{\ell}} \text{Re} \left\{ \sum_{\lambda \in B_0} w_{\lambda}(\lambda) w_{\lambda}^*(\lambda) C(e^{i\lambda})^* \right\} \to_d MN(0, h_{\lambda}(\omega) f_{\omega}(\omega)), \]
giving the required result (9.4).

By an analysis similar to that of Lemmas A1 and A2, we obtain the following results.

A.9. Lemma A.3

\[ A_{F1} = \frac{1}{2\ell} \sum_{\lambda \in B} I_{\omega}(\lambda) f_{\omega}(\lambda)^{-1} V_{\lambda} = O_p(m^{-1/2}), \]

\(^2\) Hannan (1973, Theorem 3) showed that a finite collection of $w_{\lambda}(\lambda)$ are asymptotically independent and satisfy a central limit theorem. Phillips (2000, Theorem 3.2) recently showed that an asymptotically infinite collection of $w_{\lambda}(\lambda)$ satisfy a central limit theorem and are asymptotically independent for frequencies $\lambda$ in the neighbourhood of the origin provided the number of frequencies $m = o(n^{-1/2-1/p})$ where $E(|\xi|^p) < \infty$ for some $p > 2$, i.e. provided the number of frequencies does not go to infinity too fast. The Phillips result can be extended to frequency bands away from the origin, although a proof was not given in that paper. In the present context, the collection $\{w_{\lambda}(\lambda); \lambda \in B_0\}$ is asymptotically infinite and the number of frequencies in $B_0$ is $\ell$. We need to choose $\ell$ such that $1/\ell + \ell/n^{1/4} \to 0$ in the present case since 4th moments of $\xi$ are finite.
\[ \Delta V_2 = \frac{1}{2T} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s^2 = \mathcal{O}_p(m^{-1}), \]
\[ \Delta B_1 = \frac{1}{2T} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s = \mathcal{O}_p(M^{-q}), \]
\[ \Delta B_2 = \frac{1}{2T} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 = \mathcal{O}_p(M^{-2q}), \]

and

\[ \Psi_0 = \frac{1}{\sqrt{2T}} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-1} = \mathcal{O}_p(1), \]
\[ \Psi_{V1} = \frac{1}{\sqrt{2T}} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} V_s = \mathcal{O}_p(m^{-1/2}), \]
\[ \Psi_{V2} = \frac{1}{\sqrt{2T}} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} V_s^2 = \mathcal{O}_p(m^{-1}), \]
\[ \Psi_{B1} = \frac{1}{\sqrt{2T}} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-2} B_s = \mathcal{O}_p(M^{-q}), \]
\[ \Psi_{B2} = \frac{1}{\sqrt{2T}} \sum_{\lambda_s \in B} I_{xx}(\lambda_s) f_{uu}(\lambda_s)^{-3} B_s^2 = \mathcal{O}_p(M^{-2q}), \]

A.10. Lemma A.4

(1) \( m \mathbb{E}[H_{B}^{-1} \Psi_{0}^{*} \Psi_{V2}] \to 1, \)
(2) \( m \mathbb{E}[H_{B}^{-2} \Delta V_{2} \Psi_{0}^{*} \Psi_{0}] \to 1, \)
(3) \( m \mathbb{E}[H_{B}^{-1} \Psi_{0}^{*} \Psi_{0}] \to 1, \)
(4) \( m \mathbb{E}[H_{B}^{-3} \Delta V_{1} \Psi_{0}^{*} \Psi_{0}] \to 1, \)
(5) \( M_{q} H_{B}^{-1} \Psi_{0}^{*} \Psi_{B1} \xrightarrow{d} - k_{q} f_{uu}(\omega)^{-1} f_{q}(\omega), \)
(6) \( M_{q} H_{B}^{-2} A_{B1} \Psi_{0}^{*} \Psi_{0} \xrightarrow{d} - k_{q} f_{uu}(\omega)^{-1} f_{q}(\omega). \)

A.11. Proof of Theorem 4

Notice that

\[ W = Z^{*} \tilde{H}^{-1} R' (J_{0}^{-1} - J_{0}^{-1}(J_{0} - J_{0})J_{0}^{-1} + J_{0}^{-1}(J_{0} - J_{0})J_{0}^{-1}) \]
\[ (J_{0} - J_{0})J_{0}^{-1} + R_{0} \] \[ \tilde{H}^{-1} Z, \]

and \( \tilde{H}^{-1} = H^{-1} - H^{-1} D H^{-1} + H^{-1} D H^{-1} D H^{-1} + R_{1}. \) Denote \( H^{-1} D H^{-1} = A_{1}, \)
\( H^{-1} D H^{-1} D H^{-1} = A_{2}, \) \( R H^{-1} D H^{-1} R' = J_{1}, \) and \( R H^{-1} D H^{-1} D H^{-1} R' = J_{2}, \) and drop higher order terms in the covariance matrix expansion, giving the following
approximation:
\[
W \sim Z^*(H^{-1} - A_1 + A_2) R'[J_0^{-1} + J_0^{-1} J_0^{-1} - J_0^{-1} J_0^{-1} + J_0^{-1} J_0^{-1} J_0^{-1}]^{-1} R(H^{-1} - A_1 + A_2) Z
\sim Z^* H^{-1} \mathcal{R}' J_0^{-1} R^{-1} H^{-1} Z + Z^* H^{-1} \mathcal{R}' J_0^{-1} J_0^{-1} R H^{-1} Z - 2 Z^* H^{-1} \mathcal{R}' J_0^{-1} \mathcal{R} A_1 Z
\sim Z^* H^{-1} \mathcal{R}' J_0^{-1} J_0^{-1} R H^{-1} Z + Z^* H^{-1} \mathcal{R}' J_0^{-1} J_0^{-1} J_0^{-1} R H^{-1} Z
\sim 2 Z^* H^{-1} \mathcal{R}' J_0^{-1} J_0^{-1} \mathcal{R} A_1 Z + 2 Z^* H^{-1} \mathcal{R}' J_0^{-1} \mathcal{R} A_1 Z + Z^* A_1 \mathcal{R}' J_0^{-1} \mathcal{R} A_1 Z.
\]

\(D\) and \(Z\) follow similar expansions as (4.6), (4.7), Substituting these preliminary expansions into the above formula, and collecting terms up to \(O_p(m^{-1} + M^{-q})\), we obtain the following expansion for the expected Wald statistic.

\[
E[W] \sim E[Z_0^* H^{-1} \mathcal{R}'(R H^{-1} \mathcal{R}')^{-1} R H^{-1} Z_0] + E[Z_0^{*1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_{v1}]
+ E[Z_0^* H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0]
- E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0]
+ E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0]
- 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0]
+ 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0]
+ E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v2} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0]
+ 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_{v2}] - 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v2} H^{-1} Z_0]
+ 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0] - 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0]
- E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0]
- 2Z_{v1}^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0 + 2E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0]
- E[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0].
\]

A.12. Lemma A.5

\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_{v1}] \rightarrow p,\]
\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_{v2}] \rightarrow p,\]
\[mE[Z_0^* H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0] \rightarrow p,\]
\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} Z_0] \rightarrow p,\]
\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0] \rightarrow p,\]
\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0] \rightarrow p,\]
\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0] \rightarrow p,\]
\[mE[Z_0^* H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} \mathcal{R}' J_0^{-1} R H^{-1} D_{v1} H^{-1} Z_0] \rightarrow p,\]
Similarly, the results for 

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{V2} H^{-1} R J_0^{-1} R H^{-1} Z_0] \rightarrow p, \]

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{V2} H^{-1} Z_0] \rightarrow p, \]

\[ M^q E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{B1} H^{-1} Z_0] \rightarrow -p k_q f_{uw}(0)^{-1} f_q(0), \]

\[ M^q E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} Z_0] \rightarrow -p k_q f_{uw}(0)^{-1} f_q(0), \]

\[ M^q E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{B1} H^{-1} R J_0^{-1} R H^{-1} Z_0] \rightarrow -p k_q f_{uw}(0)^{-1} f_q(0). \]

**Proof.** Proofs of the asymptotic forms for these terms are similar, and we will derive the results for \( Z_{V1}' H^{-1} R J_0^{-1} R H^{-1} Z_{V1} \) and \( Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{B1} H^{-1} Z_0 \) here. First,

\[
E[Z_{V1}' H^{-1} R J_0^{-1} R H^{-1} Z_{V1}] = E \left[ \frac{1}{n} \sum_s I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-2} V_s \right] H^{-1} R J_0^{-1} R H^{-1} \left[ \frac{1}{n} \sum_s I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-2} V_s \right] \\
\sim E \left[ \frac{1}{n^2} \sum_s w_s(\lambda_s)^4 I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-4} V_s^2 H^{-1} R J_0^{-1} R H^{-1} w_s(\lambda_s) \right] \\
= E \left[ tr \left( \frac{1}{n^2} \sum_s w_s(\lambda_s)^4 I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-4} V_s^2 H^{-1} R J_0^{-1} R H^{-1} w_s(\lambda_s) \right) \right] \\
= E \left[ tr \left( H^{-1} R J_0^{-1} R H^{-1} \frac{1}{n^2} \sum_s w_s(\lambda_s)^4 I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-4} V_s^2 \right) \right] \\
= E \left( tr \left[ H^{-1} R J_0^{-1} R H^{-1} \frac{1}{n^2} \sum_s I_s(\lambda_s) f_{uw}(\lambda_s)^{-4} V_s^2 \right] \right) \\
\sim \frac{1}{m} E \left[ tr \left( H^{-1} R J_0^{-1} R H^{-1} \frac{1}{n^2} \sum_s I_s(\lambda_s) f_{uw}(\lambda_s)^{-4} V_s^2 \right) \right] \\
\sim \frac{1}{m} E \left[ tr \left( H^{-1} R J_0^{-1} R \right) \right] = \frac{1}{m} E \left[ tr \left( H^{-1} R J_0^{-1} R \right) \right] = \frac{p}{m}.
\]

Similarly,

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} Z_{V2}] \sim \frac{p}{m}, \]

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{V2} H^{-1} R J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m}, \]

\[ E[Z_0^* H^{-1} D_{V1} H^{-1} R J_0^{-1} R H^{-1} D_{V1} H^{-1} Z_0] \sim \frac{p}{m}, \]

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{V1} H^{-1} R J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m}, \]

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{V1} H^{-1} R J_0^{-1} R H^{-1} D_{V1} H^{-1} R J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m}, \]

\[ E[Z_0^* H^{-1} R J_0^{-1} R H^{-1} D_{V1} H^{-1} R J_0^{-1} R H^{-1} D_{V1} H^{-1} R J_0^{-1} R H^{-1} Z_0] \sim \frac{p}{m}. \]
\[
E[Z_n H^{-1} R' J_0^{-1} R H^{-1} D_{V1} H^{-1} D_{V1} H^{-1} Z_0] \sim \frac{P}{m},
\]
\[
E[Z_n H^{-1} R' J_0^{-1} R H^{-1} D_{V2} H^{-1} Z_0] \sim \frac{P}{m},
\]
and
\[
E[Z_n H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} Z_0]
= E \left[ \frac{1}{n} \sum_s I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-1} \right] H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1}
\times \left[ \frac{1}{n} \sum_s I_{uw}(\lambda_s) f_{uw}(\lambda_s)^{-1} \right] \sim E \left[ \frac{1}{n^2} \sum_s w_s(\lambda_s)^* I_{uw}(\lambda_s) f_{uw}(\lambda_s) - 2 H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} w_s(\lambda_s) \right]
\]
\[
\sim E \left\{ tr \left[ \frac{1}{n^2} \sum_s w_s(\lambda_s)^* I_{uw}(\lambda_s) f_{uw}(\lambda_s) - 2 H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} w_s(\lambda_s) \right] \right\}
\]
\[
\sim E \left\{ tr \left[ H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} \left( - \frac{1}{n^2} \sum_s I_{sx}(\lambda_s) f_q(\lambda_s) f_{uw}(\lambda_s)^{-2} M^{-q} k_q \right) \right] \right\}
\]
\[
\sim - \frac{k_q}{M^q} f_{uw}(0)^{-1} f_q(0) E \left\{ tr \left[ \left( \int B_x B_x' \right)^{-1} R' \left( \int B_x B_x' \right)^{-1} R' \right] \right\}
\times R \left( \int B_x B_x' \right)^{-1} \left( \int B_x B_x' \right) \}
\]
\[
\sim - \frac{k_q}{M^q} f_{uw}(0)^{-1} f_q(0) E \left\{ tr \left( R \left( \int B_x B_x' \right)^{-1} R' \left( \int B_x B_x' \right)^{-1} R' \right) \right\}
\]
\[
\sim - \frac{p k_q}{M^q} f_{uw}(0)^{-1} f_q(0),
\]
and
\[
E[Z_{B1} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim - \frac{p k_q}{M^q} f_{uw}(0)^{-1} f_q(0),
\]
\[
E[Z_n H^{-1} R' J_0^{-1} R H^{-1} D_{B1} H^{-1} R' J_0^{-1} R H^{-1} Z_0] \sim - \frac{p k_q}{M^q} f_{uw}(0)^{-1} f_q(0). \]
References


