ON THE LOG PERIODOGRAM REGRESSION ESTIMATOR OF THE MEMORY PARAMETER IN LONG MEMORY STOCHASTIC VOLATILITY MODELS

ROHIT S. DEO AND CLIFFORD M. HURVICH
New York University

We consider semiparametric estimation of the memory parameter in a long memory stochastic volatility model. We study the estimator based on a log periodogram regression as originally proposed by Geweke and Porter-Hudak (1983, *Journal of Time Series Analysis* 4, 221–238). Expressions for the asymptotic bias and variance of the estimator are obtained, and the asymptotic distribution is shown to be the same as that obtained in recent literature for a Gaussian long memory series. The theoretical result does not require omission of a block of frequencies near the origin. We show that this ability to use the lowest frequencies is particularly desirable in the context of the long memory stochastic volatility model.

1. INTRODUCTION

There is a rapidly expanding empirical literature that has found evidence of slowly decaying auto-correlations in the volatility of financial time series. This literature includes Ding, Granger, and Engle (1993), de Lima and Crato (1993), Bollerslev and Mikkelsen (1996), Lobato and Savin (1996), and Breidt, Crato, and de Lima (1998) for stock returns; Lobato (1999), Andersen and Bollerslev (1997a, 1997b), and Henry and Payne (1998) for foreign exchange rates. One of the models to account for this persistence is the long memory stochastic volatility (LMSV) model, which was proposed independently by Breidt et al. (1998) and Harvey (1993). Breidt et al. (1998) found that the LMSV model provided a better match than a nearly integrated GARCH model (Bollerslev, 1986) for the sample autocorrelations of the log squared returns. Breidt et al. (1998) also argued that the LMSV model has certain advantages over the observation driven models such as the fractionally integrated GARCH (FIGARCH) (Baillie, Bollerslev, and Mikkelsen, 1996) and fractionally integrated exponential GARCH (FIEGARCH) (Bollerslev and Mikkelsen, 1996).

Address correspondence to: Rohit Deo, 8-57 KMEC, 44 West 4th Street, New York, NY 10012, USA; e-mail: rdeo@stern.nyu.edu.

The LMSV model is a particular case of the class of stochastic volatility (SV) models (see Taylor, 1986) of the form

$$\varepsilon_t = \sigma \exp(Y_t/2)e_t, \tag{1}$$

where $\{Y_t\}$ is a stationary Gaussian process independent of the process $\{e_t\}$ that is a sequence of independent and identically distributed (i.i.d.) random variables with zero mean. The persistence in the volatility of $\{\varepsilon_t\}$ is determined by the persistence in the correlation structure of $\{Y_t\}$. The LMSV model assumes that $\{Y_t\}$ has long memory so that the autocorrelations of $\{Y_t\}$ at lag k decay at the rate k^{2d-1}, where $d \in (0, 0.5)$ is the memory parameter. In contrast, the autocorrelations of a short memory process typically decay at an exponential rate.

Most of the currently used estimation techniques for SV models exploit the fact that the model (1) can be linearized by using the transformation

$$Z_t = \log(\varepsilon_t^2) = \log \sigma^2 + Y_t + \log e_t^2$$

$$= (\log \sigma^2 + E[\log e_t^2]) + Y_t + (\log e_t^2 - E[\log e_t^2])$$

$$= \mu + Y_t + u_t, \tag{2}$$

where $\{u_t\}$ is i.i.d. with mean zero and variance σ_u^2. If, for example, e_t is standard normal, then $\log e_t^2$ is distributed as log of a χ_1^2 variable, $E(\log e_t^2) = \Psi(\frac{1}{2}) + \log 2$ and $\sigma_u^2 = \pi^2/2$, where $\Psi(\cdot)$ is the digamma function. It is evident that the autocovariances of $\{Z_t\}$ are identical to those of $\{Y_t\}$ except at lag zero because $\{u_t\}$ is an independent process.

Assuming that $\{Y_t\}$ obeys a fully specified parametric model, it is known that the frequency domain quasi-maximum likelihood (QML) estimates based on $\{Z_t\}$ are consistent and asymptotically normal when $\{Y_t\}$ exhibits short memory (Dunsmuir, 1979) and also when $\{Y_t\}$ has long memory (Hosoya, 1997). However, if the model is misspecified, then the parameter estimates can be inconsistent. Because d is a measure of persistence as described previously, it may be appealing to estimate d semiparametrically from $\{Z_t\}$ without relying on a parametric specification for $\{Y_t\}$.

Several authors (Breidt et al., 1998; Andersen and Bollerslev, 1997a) have estimated d semiparametrically using either squared returns or some transformation such as absolute or log squared returns. The semiparametric estimator they have used is the Geweke and Porter-Hudak (1983) estimator based on a log periodogram regression. However, the asymptotic behavior of this estimator is known only under the assumption that the observations used in computing the periodogram are Gaussian. See Robinson (1995a) and Hurvich, Deo, and Brodsky (1998). As Andersen and Bollerslev (1997a) point out in their analysis using the absolute returns, this assumption will be clearly violated in their case. Also, as can be seen from (2), the log squared returns in an LMSV model are not Gaussian, and hence the asymptotic theory for the Geweke and Porter-

Hudak estimator does not follow immediately from currently known results in this case. In this paper, we derive the limiting distribution of the Geweke and Porter-Hudak estimator of d in the LMSV model based on the log periodogram of $\{Z_t\}$.

Another popular semiparametric estimator of d is the Gaussian semiparametric (GSE) estimator proposed by Künsch (1987). The limiting distribution of the GSE has been established under general conditions by Robinson (1995b) and Henry and Robinson (1997) without assuming Gaussianity. Both these papers assume that the observed data is generated by a process that is linear in martingale differences. However, the $\{Z_t\}$ process in the LMSV model does not possess such a representation. Moreover, it seems that none of the other currently existing models of long memory conditional heteroskedasticity would satisfy these assumptions either.

In the next section, we define the LMSV model and state our result on the Geweke and Porter-Hudak estimator of d. Section 3 presents results from a simulation study, and we conclude with a technical Appendix containing the proofs of our results.

2. ASYMPTOTIC THEORY

In addition to (1), we assume that the spectral density of $\{Y_t\}$ is

$$f_Y(\lambda) = \left| 2 \sin\left(\frac{\lambda}{2}\right) \right|^{-2d} g^*(\lambda), \tag{3}$$

where $d \in (0, 0.5)$ is the memory parameter and $g^*(\cdot)$ is a spectral density continuous on $[-\pi, \pi]$ bounded above and bounded away from zero, differentiable twice with the second derivative bounded in a neighborhood of zero. By symmetry of g^* around the origin, it follows that $g^{*\prime}(0) = 0$. In particular, if g^* were taken to be the spectral density of a stationary ARMA(p, q) model, then $\{Y_t\}$ would obey the ARFIMA(p, d, q) specification considered by Breidt et al. (1998). We also assume that $\{u_t\}$ has a finite eighth moment. A sufficient condition for this is that the density of $\{e_t\}$ be bounded at the origin and obey a power law decay in the tails as would occur, for example, in all t and stable distributions.

Letting σ_u^2 denote the variance of $\{u_t\}$ and $f_Z(\lambda)$ and $f_Y(\lambda)$ denote the spectral densities of $\{Z_t\}$ and $\{Y_t\}$, respectively, we have

$$f_Z(\lambda) = f_Y(\lambda) + \frac{\sigma_u^2}{2\pi} = \left| 2 \sin\left(\frac{\lambda}{2}\right) \right|^{-2d} g^*(\lambda) + \frac{\sigma_u^2}{2\pi}$$

$$= \left| 2 \sin\left(\frac{\lambda}{2}\right) \right|^{-2d} f^*(\lambda), \tag{4}$$

where

$$f^*(\lambda) = g^*(\lambda) + \left| 2 \sin\left(\frac{\lambda}{2}\right) \right|^{2d} \frac{\sigma_u^2}{2\pi}. \tag{5}$$

Note that because f^* is bounded above and bounded away from zero, the ratio f_Z/f_Y converges to 1 as λ goes to zero, so that the spectral density of $\{Z_t\}$ behaves like that of $\{Y_t\}$ at low frequencies. This motivates the use of the Geweke and Porter-Hudak estimator based on $\{Z_t\}$ described as follows.

Define the periodogram of the observations $Z_0, Z_1, \ldots, Z_{n-1}$ at the jth Fourier frequency $\omega_j = 2\pi j/n$ by

$$I_{jZ} = \frac{1}{2\pi n} \left| \sum_{t=0}^{n-1} Z_t e^{it\omega_j} \right|^2.$$

The Geweke and Porter-Hudak estimator of d using the first m Fourier frequencies may be written as

$$\hat{d} = -\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log I_{jZ},$$

where $a_j = X_j - \bar{X}$, $X_j = \log|2\sin(\omega_j/2)|$, $\bar{X} = m^{-1} \sum_{j=1}^{m} X_j$, and $S_{xx} = \sum_{j=1}^{m} a_j^2$. Note that we do not require a lower truncation on the value of j in computing the estimator $\hat{d}$. Such a truncation was originally required in Robinson (1995a) but not in Hurvich, Deo, and Brodsky (1998). Freedom from lower truncation is of particular importance in the LMSV context because the presence of the noise spectrum $\sigma_u^2/(2\pi)$ will place a potentially stringent upper bound on m, particularly when d is small.

The following theorem provides expressions for the asymptotic bias and variance of $\hat{d}$.

THEOREM 1. *Let $n \to \infty$, $m \to \infty$, and $n^{-2d}m^{2d}\log^2 m \to 0$. Then*

$$E(\hat{d} - d) = -(2\pi)^{2d} \frac{\sigma_u^2}{2\pi g^*(0)} \frac{d}{(2d+1)^2} \left(\frac{m^{2d}}{n^{2d}}\right) + O\left(\frac{\log^3 m}{m}\right) + o\left(\frac{m^{2d}}{n^{2d}}\right)$$

and

$$Var(\hat{d}) = \frac{\pi^2}{24m} + o(m^{-1}) + O\left(\frac{m^{4d}}{n^{4d}} \log^2 m\right).$$

Theorem 1 implies that $\hat{d}$ is consistent for d if $m = Kn^\delta$ for any $0 < \delta < 1$. The first term in the preceding expression for bias is due to the presence of the noise $\{u_t\}$ in equation (2). The second term is due to the bias from the long memory in $\{Y_t\}$. The first term is dominant if and only if $\delta > (1 + 2d)^{-1}2d$. Hence, $\hat{d}$ will tend to have an increasingly negative bias as m becomes sufficiently large. The quantity $\sigma_u^2/(2\pi g^*(0))$ is a measure of the relative importance of the noise

term $\{u_t\}$ compared to the short memory component of $\{Y_t\}$. Hence, for a given $\{Y_t\}$ process, as σ_u^2 is increased, the bias in $\hat{d}$ increases. The first two terms in the expression for the variance are identical to those given by Hurvich, Deo, and Brodsky (1998), whereas the remaining term is due to the presence of the noise $\{u_t\}$. The first term will dominate if and only if $\delta < (1 + 4d)^{-1}4d$.

The condition that g^* be twice differentiable may be relaxed as in Robinson (1995a). This would result in a somewhat more complicated expression for the bias in Theorem 1. The next theorem shows that under stronger conditions on m, we obtain asymptotic normality for $\hat{d}$.

THEOREM 2. *Let* $n \to \infty$, $m \to \infty$, $n^{-4d}m^{4d+1}\log^2 m \to 0$, *and* $\log^2 n = o(m)$. *Then*

$$m^{1/2}(\hat{d} - d) \xrightarrow{D} N\left(0, \frac{\pi^2}{24}\right).$$

Theorem 2 shows that the limiting distribution of $\hat{d}$ remains unchanged compared to the Gaussian case considered by Robinson (1995a) and Hurvich, Deo, and Brodsky (1998). However, the conditions on m here are much stronger. The limiting distribution will hold if and only if $\delta < (1 + 4d)^{-1}4d$, which can be arbitrarily small if d is sufficiently close to zero. Because this limitation on m depends on the unknown value of d, the construction of confidence intervals for d based upon Theorem 2 may be problematic.

Remark 1. Robinson (1995a) obtained the limiting distribution of $\hat{d}$ for a Gaussian long memory process $\{Z_t\}$ under the assumption that its spectral density was of the form

$$f_Z(\lambda) = C\lambda^{-2d}(1 + O(\lambda^\alpha)),$$

where $C > 0$ and $0 < \alpha \le 2$. He showed that the feasible range of values for m required to obtain asymptotic normality for $\hat{d}$ depended on α. More specifically, Robinson showed that the condition on m was of the form $m^{2\alpha+1}/n^{2\alpha} \to 0$. From equations (4) and (5), it can be seen that in our context $f_Z(\lambda) = \lambda^{-2d}g^*(0)[1 + O(\lambda^{2d})]$. This is analogous to Robinson's formulation with $\alpha = 2d$. Hence, it is not surprising that the conditions on m imposed in Theorem 2 depend on d.

Remark 2. We believe that our methodology of proof could be combined with the results of Robinson (1995a) to obtain, in a fairly straightforward manner, the limiting distribution of the Geweke and Porter-Hudak estimator in a multivariate LMSV model. We have presented the results here only for the univariate case for the sake of notational simplicity.

3. SIMULATION RESULTS

In this section, we report the results from a Monte Carlo study on the performance of the Geweke and Porter-Hudak estimator for LMSV models. We gen-

erated 500 replications of time series of length $n = 6{,}144$ from the model (1), where $\sigma = 1$ and $\{Y_t\}$ is an ARFIMA$(0, d, 0)$ given by

$$(1 - B)^d Y_t = \eta_t,$$

where $\{\eta_t\}$ is Gaussian white noise with variance σ_η^2 and independent of $\{e_t\}$, which is Gaussian white noise with unit variance. The sample size of $n = 6{,}144$ was used in the Monte Carlo study of Breidt et al. (1998). We considered three values of d (0.1, 0.3, and 0.47) and two values of σ_η^2 (0.37 and 0.8). The parameter configuration of $d = 0.47$ and $\sigma_\eta^2 = 0.37$ was also considered by Breidt et al. (1998). The ratio $\sigma_u^2/(2\pi g^*(0))$ takes the values 13.37 and 6.17 when σ_η^2 is 0.37 and 0.8, respectively. The estimator $\hat{d}$ was computed without any lower truncation using $m = n^{0.3}$, $n^{0.4}$, and $n^{0.5}$. To study the effects of lower truncation, we also computed $\hat{d}$ for $m = n^{0.5}$ excluding the first two frequencies, as was done in Breidt et al. (1998).

Table 1 presents the simulation means and standard deviations of the estimates. As predicted by Theorem 1, the bias is negative in all cases and becomes more negative as m increases. This phenomenon was also noted by Breidt et al. (1998). Note that when $m = n^{0.5}$, the truncated estimator has a greater bias compared to the nontruncated version. This seems to suggest that any low frequency bias due to long memory is more than compensated for by the resulting increase in bias due to the noise term at high frequencies.

We now consider the variances of the estimators. In Table 2, we present the theoretical standard errors of the estimators as obtained from Theorem 2 and given by $SE_1 = \pi[24(m - l)]^{-1/2}$, where $l = 0$ or 2 for the nontruncated and truncated version, respectively. Also presented in Table 2 are the values $SE_2 =$

TABLE 1. Average estimates of d over 500 replications[a]

	$m = n^{0.3}$	$m = n^{0.4}$	$m = n^{0.5}$	$m = n^{0.5}$, Trunc
$d = 0.1$, $\sigma_\eta^2 = 0.37$	0.018	0.016	0.016	0.014
	(0.245)	(0.137)	(0.083)	(0.094)
$d = 0.1$, $\sigma_\eta^2 = 0.8$	0.032	0.031	0.028	0.025
	(0.249)	(0.134)	(0.084)	(0.096)
$d = 0.3$, $\sigma_\eta^2 = 0.37$	0.200	0.175	0.145	0.132
	(0.247)	(0.135)	(0.083)	(0.092)
$d = 0.3$, $\sigma_\eta^2 = 0.8$	0.238	0.222	0.197	0.188
	(0.244)	(0.138)	(0.084)	(0.094)
$d = 0.47$, $\sigma_\eta^2 = 0.37$	0.433	0.408	0.362	0.344
	(0.247)	(0.142)	(0.085)	(0.094)
$d = 0.47$, $\sigma_\eta^2 = 0.8$	0.451	0.438	0.412	0.402
	(0.251)	(0.143)	(0.084)	(0.092)

[a]Associated simulation standard deviations are given in parentheses.

TABLE 2. Theoretical standard errors followed by approximate standard errors based on standard regression theory

	$m = n^{0.3}$	$m = n^{0.4}$	$m = n^{0.5}$	$m = n^{0.5}$, Trunc
SE_1	0.178	0.113	0.073	0.074
SE_2	0.242	0.135	0.080	0.095

$\pi[24\sum_{j=l}^{m}(X_j - \bar{X})^2]^{-1/2}$. These are the standard errors for the estimated slope in a regression with $0.5X_j$ as the regressors and assuming that the errors are independent with mean zero and variance $\pi^2/6$. The approximation SE_2 was originally suggested by Geweke and Porter-Hudak (1983). From Table 1 it is seen that the variances of the estimates decrease as m increases. Comparison of the simulation standard deviations from Table 1 and the theoretical standard errors from Table 2 indicates that SE_2 provides a much better approximation than SE_1 to the simulation standard deviations. Furthermore, SE_1 understates the standard deviation of the estimates in all cases, and it also understates the inflation of variance that would be incurred by truncation. The inflation in variance due to truncation can be quite substantial and is apparently due to the high leverage of the initial values of X_j. In view of the increases in both the bias and the variance, truncation is particularly inadvisable for estimation of d in the LMSV model.

REFERENCES

Andersen, T. & T. Bollerslev (1997a) Heterogenous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance* 52, 975–1005.

Andersen, T. & T. Bollerslev (1997b) Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance* 4, 115–158.

Baillie, R., T. Bollerslev, & H. Mikkelsen (1996) Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 74, 3–30.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307–327.

Bollerslev, T. & H. Mikkelsen (1996) Modelling and pricing long memory in stock market volatility. *Journal of Econometrics* 73, 151–184.

Breidt, F.J., N. Crato, & P. De Lima (1998) The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83, 325–348.

Courant, R. & F. John (1965) *Introduction to Calculus and Analysis*. New York: Wiley.

de Lima, P. & N. Crato (1993) Long range dependence in the conditional variance of stock returns. *Proceedings of the Business and Economics Statistics Section*, Joint Statistical Meetings.

Ding, Z., C.W.J. Granger, & R. Engle (1993) A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83–106.

Dunsmuir, W. (1979) A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *Annals of Statistics* 7, 490–506.

Geweke, J. & S. Porter-Hudak (1983) The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4, 221–238.

Harvey, A.C. (1993) Long memory in stochastic volatility. In J. Knight & S. Satchell (eds.), *Forecasting Volatility in the Financial Markets*. London: Butterworth Heinemann.

Henry, M. & R. Payne (1998) An Investigation of Long Range Dependence in Intra-day Foreign Exchange Rate Volatility. Preprint.

Henry, M. & P. Robinson (1997) Long Memory Conditional Heteroscedasticity in Semiparametric Estimation of Long Memory. Preprint.

Hosoya, Y. (1997) A limit theory for long range dependence and statistical inference on related models. *Annals of Statistics* 25, 105–137.

Hurvich, C.M., R. Deo, & J. Brodsky (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter in a long-memory time series. *Journal of Time Series Analysis* 19, 19–46.

Johnson, N., S. Kotz, & N. Balakrishnan (1995) *Continuous Univariate Distributions*, vol. 2. New York: Wiley.

Künsch, H. (1987) Statistical aspects of self similar processes. *Proceedings of the First World Congress of the Bernoulli Society* pp. 67–74. VNU Science Press.

Lobato, I.N. (1999) A semiparametric two step estimator in a multivariate long memory model. *Journal of Econometrics* 90, 129–153.

Lobato, I.N. & N. Savin (1996) Real and Spurious Long Memory Properties of Stock Market Data. Preprint.

Rao, C.R. (1973) *Linear Statistical Inference and Its Applications*, 2nd ed. New York: Wiley.

Robinson, P. (1995a) Log periodogram regression of time series with long range dependence. *Annals of Statistics* 23, 1048–1072.

Robinson, P. (1995b) Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* 23, 1630–1661.

Taylor, S. (1986) *Modelling Financial Time Series*. New York: Wiley.

Zwillinger, D., ed. (1996) *Standard Mathematical Tables and Formulae*, 30th edition. Boca Raton, Florida: CRC Press.

TECHNICAL APPENDIX

In all the proofs, we will use the following notation. Let

$$I_{jZ} = A_{jZ}^2 + B_{jZ}^2,$$

where

$$A_{jZ} = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} Z_t \cos \omega_j t = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} Y_t \cos \omega_j t + \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} u_t \cos \omega_j t$$

$$\equiv A_{jY} + A_{jU}$$

and

$$B_{jZ} = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} Z_t \sin \omega_j t = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} Y_t \sin \omega_j t + \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} u_t \sin \omega_j t \equiv B_{jY} + B_{jU}.$$

Proof of Theorem 1. We will prove the expression for bias first. In this proof, we will use the facts that $S_{xx} = m + o(m)$, $a_j = O(\log m)$, and $a_j = \log(j/m) + 1 + o(1)$ (see Hurvich, Deo, and Brodsky, 1998, pp. 22, 38). We have

$$\hat{d} = -\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log I_{jZ} = -\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log \frac{2I_{jZ}}{f_{jY}} - \frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log f_{jY}$$

$$= -\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) - \frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log f_{jY},$$

where $C = -\Psi(1) - \log 2$ and we have used the fact that $\sum_{j=1}^{m} a_j = 0$. Thus,

$$\hat{d} = d - \frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log g^*(\omega_j) - \frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right)$$

$$= d + O\left(\frac{m^2}{n^2}\right) - \frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) \tag{A.1}$$

by Lemma 1 of Hurvich, Deo, and Brodsky (1998). Now

$$E\left\{ \frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) \right\}$$

$$= E\left\{ \frac{1}{2S_{xx}} \sum_{j=1}^{\log^2 m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) \right\}$$

$$+ E\left\{ \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) \right\}$$

$$= O\left(\frac{\log^3 m}{m}\right) + E\left\{ \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) \right\} \tag{A.2}$$

by Lemma 1. By Lemmas 2 and 3, it follows that

$$E\left\{ \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} a_j \left(\log \frac{2I_{jZ}}{f_{jY}} + C \right) \right\} = \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} a_j f_{jY}^{-1} \frac{\sigma_u^2}{2\pi}$$

$$+ O\left(\frac{m^{4d}}{n^{4d}} \log^2 m\right) + O\left(\frac{\log^3 m}{m}\right). \tag{A.3}$$

But

$$\frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} a_j f_{jY}^{-1} \frac{\sigma_u^2}{2\pi}$$

$$= \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} \{\log(j/m) + 1 + o(1)\} f_{jY}^{-1} \frac{\sigma_u^2}{2\pi}$$

$$= \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} \{\log(j/m) + 1\} f_{jY}^{-1} \frac{\sigma_u^2}{2\pi} + o\left(\frac{m^{2d}}{n^{2d}}\right)$$

$$= \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} \{\log(j/m) + 1\} \frac{|2\sin\omega_j/2|^{2d}}{g^*(\omega_j)} \frac{\sigma_u^2}{2\pi} + o\left(\frac{m^{2d}}{n^{2d}}\right).$$

By a Taylor expansion and using the fact that $g^{*\prime}(0) = 0$, we have

$$\frac{1}{g^*(\omega_j)} = \frac{1}{g^*(0)} + O(\omega_j^2)$$

and also

$$|2\sin\omega_j/2|^{2d} = \omega_j^{2d} + O(\omega_j^{2d+1}).$$

Hence,

$$\frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} \{\log(j/m) + 1\} \frac{|2\sin\omega_j/2|^{2d}}{g^*(\omega_j)} \frac{\sigma_u^2}{2\pi}$$

$$= \frac{1}{2S_{xx}} \sum_{j=\log^2 m+1}^{m} \{\log(j/m) + 1\} \frac{\omega_j^{2d}}{g^*(0)} \frac{\sigma_u^2}{2\pi} + O\left(\frac{m^{2d+1}}{n^{2d+1}} \log m\right)$$

$$= \frac{m^{2d}}{n^{2d}} \frac{\sigma_u^2}{2\pi g^*(0)} \frac{1}{2} \int_0^1 \{\log x + 1\}(2\pi x)^{2d} dx + O\left(\frac{1}{m}\right) + o\left(\frac{m^{2d}}{n^{2d}}\right)$$

$$+ O\left(\frac{m^{2d+1}}{n^{2d+1}} \log m\right)$$

$$= (2\pi)^{2d} \frac{\sigma_u^2}{2\pi g^*(0)} \frac{d}{(2d+1)^2} \left(\frac{m^{2d}}{n^{2d}}\right) + O\left(\frac{1}{m}\right) + o\left(\frac{m^{2d}}{n^{2d}}\right) + O\left(\frac{m^{2d+1}}{n^{2d+1}} \log m\right).$$

$$(A.4)$$

The expression for the bias follows from equations (A.1)–(A–4). We now derive the variance of $\hat{d}$. We have

$$\text{Var}(\hat{d}) = \text{Var}\left(-\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log I_{jZ}\right) = \text{Var}\left(-\frac{1}{2S_{xx}} \sum_{j=1}^{m} a_j \log \frac{I_{jZ}}{f_{jY}}\right)$$

$$= \text{Var}(T_1) + \text{Var}(T_2) + 2\text{Cov}(T_1, T_2),$$

where $T_1 = -(2S_{xx})^{-1} \sum_{j=1}^{\log^8 m} a_j \log(I_{jZ}/f_{jY})$ and $T_2 = -(2S_{xx})^{-1} \sum_{j=\log^8 m+1}^{m} a_j \times \log(I_{jZ}/f_{jY})$. We can write

$$T_2 = -(2S_{xx})^{-1} \sum_{j=\log^8 m+1}^{m} a_j \log \frac{I_{jY}}{f_{jY}} - (2S_{xx})^{-1} \sum_{j=\log^8 m+1}^{m} a_j \left(\log \frac{I_{jZ}}{f_{jY}} - \log \frac{I_{jY}}{f_{jY}}\right)$$

$$\equiv T_{21} + T_{22}.$$

By Lemma 5 of Hurvich et al. (1998),

$$\text{Var}(T_{21}) = \frac{\pi^2}{24m} + o(m^{-1}).$$

$$(A.5)$$

Letting $\mu_{jZ} = E\log(f_{jY}^{-1}I_{jZ})$ and $\mu_{jY} = E\log(f_{jY}^{-1}I_{jY})$ and applying Lemma 4, we also have

$$
\begin{aligned}
\text{Var}(T_{22}) &= E\left[(2S_{xx})^{-1}\sum_{j=\log^8 m+1}^{m} a_j\left(\log\frac{I_{jZ}}{f_{jY}} - \log\frac{I_{jY}}{f_{jY}}\right)\right]^2 \\
&\quad - \left[(2S_{xx})^{-1}\sum_{j=\log^8 m+1}^{m} a_j(\mu_{jZ} - \mu_{jY})\right]^2 \\
&= O\left(\frac{m^{4d}}{n^{4d}}\log^2 m\right).
\end{aligned}
\tag{A.6}
$$

The expressions in (A.5) and (A.6) in conjunction with the Cauchy–Schwarz inequality give

$$
\begin{aligned}
\text{Var}(T_2) &= \frac{\pi^2}{24m} + o(m^{-1}) + O\left(\frac{m^{4d}}{n^{4d}}\log^2 m\right) + O\left(\frac{m^{2d}}{m^{1/2}n^{2d}}\log m\right) \\
&= \frac{\pi^2}{24m} + o(m^{-1}) + O\left(\frac{m^{4d}}{n^{4d}}\log^2 m\right).
\end{aligned}
\tag{A.7}
$$

From Lemma 1 we also get

$$
\text{Var}(T_1) = O\left(\frac{\log^{18} m}{m^2}\right),
\tag{A.8}
$$

whereas the Cauchy–Schwarz inequality gives

$$
|\text{Cov}(T_1, T_2)| = O\left(m^{-1}\log^9 m\left[O(m^{-1}) + O\left(\frac{m^{4d}}{n^{4d}}\log^2 m\right)\right]^{1/2}\right),
\tag{A.9}
$$

and the expression for the variance of $\hat{d}$ follows from (A.7)–(A.9).

Proof of Theorem 2. Let $\gamma = -\Psi(1) = 0.577216\ldots$ denote Euler's constant. We have

$$
\begin{aligned}
m^{1/2}(\hat{d} - d) &= -\frac{m^{1/2}}{2S_{xx}}\sum_{j=1}^{m} a_j\log g_j^* - \frac{m^{1/2}}{2S_{xx}}\sum_{j=1}^{m} a_j\left(\log\frac{I_{jZ}}{f_{jY}} + \gamma\right) \\
&= o(1) - \frac{m}{2S_{xx}}\frac{1}{m^{1/2}}\sum_{j=1}^{m} a_j\left(\log\frac{I_{jZ}}{f_{jY}} + \gamma\right),
\end{aligned}
\tag{A.10}
$$

where the last step follows from Lemma 1 of Hurvich, Deo, and Brodsky. Now consider

$$
\begin{aligned}
-\frac{1}{m^{1/2}}\sum_{j=1}^{m} a_j\left(\log\frac{I_{jZ}}{f_{jY}} + \gamma\right) &= -\frac{1}{m^{1/2}}\sum_{j=1}^{\log^8 m} a_j\left(\log\frac{I_{jZ}}{f_{jY}} + \gamma\right) \\
&\quad -\frac{1}{m^{1/2}}\sum_{j=1+\log^8 m}^{m} a_j\left(\log\frac{I_{jZ}}{f_{jY}} + \gamma\right) \equiv T_1 + T_2.
\end{aligned}
\tag{A.11}
$$

By Lemma 1, Chebyshev's inequality, and the fact that $a_j = O(\log m)$ (see Hurvich, Deo, and Brodsky, 1998, p. 22), we have

$$T_1 = o_p(1). \tag{A.12}$$

Using (A.10)–(A.12), it remains to show that

$$T_2 \xrightarrow{D} N\left(0, \frac{\pi^2}{6}\right). \tag{A.13}$$

Let

$$T_3 = \frac{1}{m^{1/2}} \sum_{j=1+\log^8 m}^{m} a_j \left\{ \log \frac{I_{jY}}{f_{jY}} + \gamma \right\}.$$

From the proof of Theorem 2 of Hurvich, Deo, and Brodsky, it is known that

$$T_3 \xrightarrow{D} N\left(0, \frac{\pi^2}{6}\right). \tag{A.14}$$

Hence, to show (A.13) it is sufficient to prove that

$$T_2 - T_3 = o_p(1). \tag{A.15}$$

But

$$T_2 - T_3 = \frac{1}{m^{1/2}} \sum_{j=1+\log^8 m}^{m} a_j \left\{ \log \frac{I_{jZ}}{f_{jY}} - \log \frac{I_{jY}}{f_{jY}} \right\}. \tag{A.16}$$

Using Lemma 4, we have

$$E\left(\frac{1}{m^{1/2}} \sum_{j=1+\log^8 m}^{m} a_j \left\{ \log \frac{I_{jZ}}{f_{jY}} - \log \frac{I_{jY}}{f_{jY}} \right\} \right)^2$$

$$= O\left(\frac{m^{4d+1} \log^2 m}{n^{4d}} \right).$$

Thus,

$$\frac{1}{m^{1/2}} \sum_{j=1+\log^8 m}^{m} a_j \left\{ \log \frac{I_{jZ}}{f_{jY}} - \log \frac{I_{jY}}{f_{jY}} \right\} = o_p(1), \tag{A.17}$$

and (A.15) follows from (A.16) and (A.17).

LEMMA 1. $\lim_{n\to\infty} \sup_{1\le j\le m} E(\{\log(I_{jZ}/\omega_j^{-2d})\}^2) < \infty.$

Proof of Lemma 1. There exist constants C_1 and C_2 such that for all $x \neq 0$,

$$|\log|x|| \leq C_1|x|^{-1/8} + C_2|x|^{1/2}.$$

Thus,

$$E\left(\left\{\log \frac{I_{jZ}}{\omega_j^{-2d}}\right\}^2\right) \leq C_1^2 E\left(\left|\frac{I_{jZ}}{\omega_j^{-2d}}\right|^{-1/4}\right) + C_2^2 E\left(\left|\frac{I_{jZ}}{\omega_j^{-2d}}\right|\right)$$

$$+ 2C_1 C_2 \left\{E\left(\left|\frac{I_{jZ}}{\omega_j^{-2d}}\right|^{-1/4}\right)\right\}^{1/2} \left\{E\left(\left|\frac{I_{jZ}}{\omega_j^{-2d}}\right|\right)\right\}^{1/2}. \tag{A.18}$$

From Theorem 2 part (a) of Robinson (1995a), it follows that

$$\lim_{n \to \infty} \sup_{1 \leq j \leq m} E\left(\left|\frac{I_{jZ}}{\omega_j^{-2d}}\right|\right) < \infty. \tag{A.19}$$

From (A.18) it follows that the lemma is proved if

$$\lim_{n \to \infty} \sup_{1 \leq j \leq m} E\left(\left|\frac{I_{jZ}}{\omega_j^{-2d}}\right|^{-1/4}\right) < \infty, \tag{A.20}$$

which we now proceed to show. Let $\mathbf{\Sigma}_{jn}$ be the 2×2 covariance matrix of the conditional distribution of $\mathbf{g}_j = (A_{jZ}/\omega_j^{-d}, B_{jZ}/\omega_j^{-d})'$ given $\mathbf{U} = (u_1, \ldots, u_n)$. Note that because $\mathbf{Y} = (Y_1, \ldots, Y_n)$ and $\mathbf{U}$ are independent, $\mathbf{\Sigma}_{jn}$ is deterministic and does not depend on $\mathbf{U}$, and hence any function of $\mathbf{\Sigma}_{jn}$ will also be free from $\mathbf{U}$. There exists an orthonormal matrix $\mathbf{P}_{jn} = (\mathbf{p}_1', \mathbf{p}_2')'$ such that $\mathbf{P}_{jn} \mathbf{\Sigma}_{jn} \mathbf{P}_{jn}' = \text{diag}(\lambda_{jn1}, \lambda_{jn2})$, where $0 < \lambda_{jn1} \leq \lambda_{jn2} < \infty$ are the eigenvalues of $\mathbf{\Sigma}_{jn}$. Furthermore,

$$I_{jZ}/\omega_j^{-2d} = \mathbf{g}_j'\mathbf{g}_j = \mathbf{g}_j'\mathbf{P}_{jn}'\mathbf{P}_{jn}\mathbf{g}_j = \mathbf{h}_j'\mathbf{h}_j,$$

where the conditional distribution of $\mathbf{h}_j = \mathbf{P}_{jn}\mathbf{g}_j = (h_{j1}, h_{j2})'$ given $\mathbf{U}$ is bivariate normal with mean $\mathbf{P}_{jn}(A_{jU}/\omega_j^{-d}, B_{jU}/\omega_j^{-d})'$ and covariance matrix $\text{diag}(\lambda_{jn1}, \lambda_{jn2})$. Also,

$$\lambda_{jn2} \geq \max\left\{E\left(\frac{A_{jY}^2}{\omega_j^{-2d}}\right), E\left(\frac{B_{jY}^2}{\omega_j^{-2d}}\right)\right\} \geq \frac{1}{2} E\left(\frac{I_{jY}}{\omega_j^{-2d}}\right)$$

(see Rao, 1973, equation 1f.2.1) and hence by Lemma 4 of Hurvich, Deo, and Brodsky,

$$\lim_{n \to \infty} \inf_{1 \leq j \leq m} \lambda_{jn2} > 0. \tag{A.21}$$

Thus,

$$E\left(\frac{I_{jZ}}{\omega_j^{-2d}}\right)^{-1/4} = E\left\{E\left(\frac{I_{jZ}}{\omega_j^{-2d}}\right)^{-1/4} \middle| \mathbf{U}\right\} = E\{E(\mathbf{h}_j'\mathbf{h}_j)^{-1/4} | \mathbf{U}\}$$

$$\leq E\{E(h_{j2}^2)^{-1/4} | \mathbf{U}\} = E\{E|h_{j2}|^{-1/2} | \mathbf{U}\}. \tag{A.22}$$

The conditional distribution of h_{j2} given $\mathbf{U}$ is normal with mean $\mu_n \equiv \mathbf{p}'_2(A_{jU}/\omega_j^{-d}, B_{jU}/\omega_j^{-d})'$ and variance λ_{jn2}. Furthermore, from (A.21) there exists some $C_1 > 0$ such that $\inf_n \inf_{1 \le j \le m} \lambda_{jn2} > C_1$. Hence

$$E\{|h_{j2}|^{-1/2}|\mathbf{U}\} = \int_{-\infty}^{\infty} |x|^{-1/2}(\sqrt{2\pi\lambda_{jn2}})^{-1} \exp(-0.5\lambda_{jn2}^{-1}\{x - \mu_n\}^2)\, dx$$

$$= \int_{|x|<1} |x|^{-1/2}(\sqrt{2\pi\lambda_{jn2}})^{-1} \exp(-0.5\lambda_{jn2}^{-1}\{x - \mu_n\}^2)\, dx$$

$$+ \int_{|x|\ge 1} |x|^{-1/2}(\sqrt{2\pi\lambda_{jn2}})^{-1} \exp(-0.5\lambda_{jn2}^{-1}\{x - \mu_n\}^2)\, dx$$

$$\equiv T_{1n} + T_{2n}.$$

Using the fact that $\exp(-a) < 1$ for $a > 0$, we get

$$T_{1n} \le \int_{|x|<1} |x|^{-1/2}(\sqrt{2\pi C_1})^{-1}\, dx < K_1 < \infty \qquad \text{for all } n.$$

Using $|x|^{-1/2} < 1$ for $|x| > 1$, we get

$$T_{2n} \le \int_{|x|\ge 1} (\sqrt{2\pi\lambda_{jn2}})^{-1} \exp(-0.5\lambda_{jn2}^{-1}\{x - \mu_n\}^2)\, dx$$

$$\le \int_{-\infty}^{\infty} (\sqrt{2\pi\lambda_{jn2}})^{-1} \exp(-0.5\lambda_{jn2}^{-1}\{x - \mu_n\}^2)\, dx = 1.$$

Hence,

$$E\{|h_{j2}|^{-1/2}|\mathbf{U}\} < K < \infty, \tag{A.23}$$

where K is a constant that does not depend on $\mathbf{U}$, and hence (A.20) follows from (A.22) and (A.23).

LEMMA 2. *For $\log^2 m \le j \le m$,*

$$E\left(\log \frac{2I_{jZ}}{f_{jY}}\right) = (1 + O(j^{-1}\log j))E(E\{\log \chi_2^2(\lambda)|\mathbf{U}\}) + O(j^{-1}\log j),$$

where

$$\lambda = \frac{2I_{jU}}{f_{jY}}$$

is the noncentrality parameter of the χ_2^2 variable that is independent of $\mathbf{U}$.

Proof of Lemma 2. The conditional distribution of $\mathbf{X} = (f_{jY}^{-1/2}\sqrt{2}A_{jZ}, f_{jY}^{-1/2}\sqrt{2}B_{jZ})'$ given $\mathbf{U} = (u_1, u_2, \ldots, u_n)$ is bivariate normal with mean vector

$$\boldsymbol{\mu}_U = (f_{jY}^{-1/2}\sqrt{2}A_{jU}, f_{jY}^{-1/2}\sqrt{2}B_{jU})'$$

and a conditional variance covariance matrix $\mathbf{\Sigma}$ that is free from $\mathbf{U}$. From Theorem 2 of Robinson (1995a), using the fact that if $\mathbf{\Sigma} = \mathbf{I} + \mathbf{A}$ then $\mathbf{\Sigma}^{-1} = \mathbf{I} - (\mathbf{I} + \mathbf{A})^{-1}\mathbf{A}$, it follows that

$$\mathbf{\Sigma}^{-1} = \mathbf{I} + \mathbf{R}, \tag{A.24}$$

where every element of the matrix $\mathbf{R}$ is $O(b_j)$ where $b_j = j^{-1}\log j$. Hence, letting $\mathbf{x} = (x_1, x_2)'$, $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}_U$, and $g(\mathbf{x}) = \log(x_1^2 + x_2^2)$, we get

$$E\left(\log \frac{2I_{jZ}}{f_{jY}}\middle|\mathbf{U}\right) = (2\pi)^{-1}|\mathbf{\Sigma}|^{-1/2}\iint g(\mathbf{x})\exp\left(-\frac{\mathbf{z}'\mathbf{\Sigma}^{-1}\mathbf{z}}{2}\right)d\mathbf{x}$$

$$= (2\pi)^{-1}|\mathbf{\Sigma}|^{-1/2}\iint g(\mathbf{x})\exp\left(-\frac{\mathbf{z}'\mathbf{z}}{2}\right)d\mathbf{x}$$

$$+ (2\pi)^{-1}|\mathbf{\Sigma}|^{-1/2}\iint g(\mathbf{x})\left\{\exp\left(-\frac{\mathbf{z}'\mathbf{\Sigma}^{-1}\mathbf{z}}{2}\right) - \exp\left(-\frac{\mathbf{z}'\mathbf{z}}{2}\right)\right\}d\mathbf{x}$$

$$= (2\pi)^{-1}\iint g(\mathbf{x})\exp\left(-\frac{\mathbf{z}'\mathbf{z}}{2}\right)d\mathbf{x}$$

$$+ (2\pi)^{-1}\{|\mathbf{\Sigma}|^{-1/2} - 1\}\iint g(\mathbf{x})\exp\left(-\frac{\mathbf{z}'\mathbf{z}}{2}\right)d\mathbf{x}$$

$$+ (2\pi)^{-1}|\mathbf{\Sigma}|^{-1/2}\iint g(\mathbf{x})\exp\left(-\frac{\mathbf{z}'\mathbf{z}}{2}\right)\left[\exp\left(-\frac{\mathbf{z}'\mathbf{R}\mathbf{z}}{2}\right) - 1\right]d\mathbf{x}$$

$$\equiv L_1(\mathbf{U}) + L_2(\mathbf{U}) + L_3(\mathbf{U}). \tag{A.25}$$

It is clear that

$$L_1(\mathbf{U}) = E\{\log \chi_2^2(\lambda)|\mathbf{U}\}, \tag{A.26}$$

where $\lambda = 2f_{jY}^{-1}I_{jU}$ is the noncentrality parameter of a χ_2^2 variable that is independent of $\mathbf{U}$. It also follows from equation (A.24) that

$$L_2(\mathbf{U}) = O(j^{-1}\log j)E\{\log \chi_2^2(\lambda)|\mathbf{U}\}. \tag{A.27}$$

We now bound $L_3(\mathbf{U})$. By the mean value theorem, we have $|\exp x - 1 - x| \leq 0.5x^2\exp|x|$ for any x. Letting K_1 denote the largest absolute entry of $\mathbf{R}$ and letting $\mathbf{V}$ denote a bivariate standard normal vector that is independent of $\mathbf{U}$, we get

$$L_3(\mathbf{U}) = -|\mathbf{\Sigma}|^{-1/2}E_{\mathbf{V}}\left\{\log[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)]\frac{\mathbf{V}'\mathbf{R}\mathbf{V}}{2}\right\}$$

$$+ O\left\{K_1^2\iint g(\mathbf{x})\|\mathbf{z}\|^4\exp\left(-\frac{(1 - 4K_1)\mathbf{z}'\mathbf{z}}{2}\right)d\mathbf{x}\right\}$$

$$\equiv L_{31}(\mathbf{U}) + L_{32}(\mathbf{U}), \tag{A.28}$$

where $|\mathbf{\Sigma}|^{-1/2} = O(1)$. By the Cauchy–Schwarz inequality,

$$\left| E_{\mathbf{V}} \left\{ \log[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)] \frac{\mathbf{V}'\mathbf{R}\mathbf{V}}{2} \right\} \right|$$

$$\leq \sqrt{E_{\mathbf{V}} \log^2[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)]} \sqrt{E_{\mathbf{V}} \left(\frac{\mathbf{V}'\mathbf{R}\mathbf{V}}{2} \right)^2}. \tag{A.29}$$

But there exist finite constants C_1 and C_2 such that for all x,

$$|\log^2|x|| \leq C_1 |x|^{-1/4} + C_2 |x|^{1/2}.$$

Hence,

$$E_{\mathbf{V}} \log^2[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)] \leq C_1 E_{\mathbf{V}}[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)]^{-1/4}$$

$$+ C_2 E_{\mathbf{V}}[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)]^{1/2}. \tag{A.30}$$

But

$$C_1 E_{\mathbf{V}}[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)]^{-1/4} \leq C_1 E_{\mathbf{V}}[(V_1 + \mu_{U1})^2]^{-1/4} \leq C_3, \tag{A.31}$$

where $\mathbf{V} = (V_1, V_2)'$, $\boldsymbol{\mu}_U = (\mu_{U1}, \mu_{U2})'$, and C_3 does not depend on $\boldsymbol{\mu}_U$. Also, by Minkowski's inequality,

$$E_{\mathbf{V}}[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)]^{1/2} \leq E_{\mathbf{V}}[|V_1 + \mu_{U1}|] + E_{\mathbf{V}}[|V_2 + \mu_{U2}|]$$

$$\leq C_4 + |\boldsymbol{\mu}_{U1}| + |\boldsymbol{\mu}_{U2}|, \tag{A.32}$$

and C_4 does not depend on $\mathbf{U}$. From equations (A.30)–(A.32) we get

$$E_{\mathbf{V}} \log^2[(\mathbf{V} + \boldsymbol{\mu}_U)'(\mathbf{V} + \boldsymbol{\mu}_U)] \leq C_5 + C_2|\boldsymbol{\mu}_{U1}| + C_2|\boldsymbol{\mu}_{U2}|, \tag{A.33}$$

where C_5 does not depend on $\mathbf{U}$. Because $E_{\mathbf{V}}(0.5\mathbf{V}'\mathbf{R}\mathbf{V})^2 = O(b_j^2)$ and does not depend on $\mathbf{U}$, equations (A.29) and (A.33) yield

$$|L_{31}(\mathbf{U})| \leq O(b_j) \sqrt{C_5 + C_2|\boldsymbol{\mu}_{U1}| + C_2|\boldsymbol{\mu}_{U2}|}. \tag{A.34}$$

A similar argument shows that

$$|L_{32}(\mathbf{U})| \leq O(b_j^2) \sqrt{C_5 + C_2|\boldsymbol{\mu}_{U1}| + C_2|\boldsymbol{\mu}_{U2}|}. \tag{A.35}$$

From equations (A.28), (A.34), and (A.35) and the Cauchy–Schwarz inequality it follows that

$$E|L_3(\mathbf{U})| \leq M_2 = O(b_j) \tag{A.36}$$

and M_2 is free from $\mathbf{U}$. The lemma now follows from (A.25)–(A.27) and (A.36).

LEMMA 3. *Let X be a noncentral chi-square random variable with two degrees of freedom and with noncentrality parameter* $\lambda = 2f_{jY}^{-1}I_{jU}$. *If X and* **U** *are independent, then for* $\log^2 m \leq j \leq m$,

(a) $E(\log X) = \Psi(1) + \log 2 + \sigma_u^2/(2\pi f_{jY}) + O(\omega_j^{4d})$
(b) $E(\log X)^2 = E(\log \chi_2^2)^2 + O(\omega_j^{2d})$

where χ_2^2 *is a central chi-square random variable with two degrees of freedom.*

Proof of Lemma 3. We will demonstrate only part (a) because the proof of (b) is similar. From Johnson, Kotz, and Balakrishnan (1995, p. 448) we have for integer $r > 0$

$$E\left\{\left(\log\left(\frac{X}{2}\right)\right)^r \Big| \mathbf{U}\right\} = \exp\left(-\frac{\lambda}{2}\right) \sum_{j=0}^{\infty} \left[\frac{(0.5\lambda)^j}{j!}\right] E\left\{\left(\log\left(\frac{\chi_{2+2j}^2}{2}\right)\right)^r\right\},$$

where $E(\log \chi_j^2) = \Psi(j) + \log 2$. Thus,

$$E\{\log X | \mathbf{U}\} = \exp\left(-\frac{\lambda}{2}\right)[E\log\chi_2^2 + 0.5\lambda E\log\chi_4^2]$$

$$+ \exp\left(-\frac{\lambda}{2}\right) \sum_{j=2}^{\infty} \left[\frac{(0.5\lambda)^j}{j!}\right] E\log\chi_{2+2j}^2$$

$$:= \exp\left(-\frac{\lambda}{2}\right)[E\log\chi_2^2 + 0.5\lambda E\log\chi_4^2] + R_1$$

$$= E\log\chi_2^2 + 0.5\lambda E\log\chi_4^2 + \left[\exp\left(-\frac{\lambda}{2}\right) - 1\right][E\log\chi_2^2 + 0.5\lambda E\log\chi_4^2]$$

$$+ R_1$$

$$= E\log\chi_2^2 + 0.5\lambda E\log\chi_4^2 + \left[-\frac{\lambda}{2} + O(\lambda^2)\right][E\log\chi_2^2 + 0.5\lambda E\log\chi_4^2]$$

$$+ R_1$$

$$= E\log\chi_2^2 + 0.5\lambda(E\log\chi_4^2 - E\log\chi_2^2) + O(\lambda^2) + R_1$$

$$= \Psi(1) + \log 2 + 0.5\lambda + O(\lambda^2) + R_1, \tag{A.37}$$

where we have used the fact that $\Psi(2) - \Psi(1) = 1$. Now, using the fact that $|\Psi(j)| = O(\log j)$ (see Zwillinger, 1996, p. 496), we have for some finite K,

$$|R_1| \leq K \exp\left(-\frac{\lambda}{2}\right) \sum_{j=2}^{\infty} \left[\frac{(0.5\lambda)^j}{j!}\right] \log j$$

$$\leq K(0.5\lambda)^2 \exp\left(-\frac{\lambda}{2}\right) \sum_{j=2}^{\infty} \left[\frac{(0.5\lambda)^{j-2}}{j!}\right] \log j.$$

But the function $x^r \exp(-x)$ has a maximum of $r^r \exp(-r)$ for $r > 0$. Furthermore, by Stirling's approximation (see Courant and John, 1965, p. 504),

$$r^r \exp(-r) < r!(2\pi)^{-1/2} r^{-1/2}.$$

Thus, noting that $\lambda > 0$, we have

$$|R_1| \leq K(0.5\lambda)^2 \exp\left(-\frac{\lambda}{2}\right) \frac{\log 2}{2} + K(0.5\lambda)^2 \sum_{j=3}^{\infty} (j-2)! (2\pi)^{-1/2} (j-2)^{-1/2} \frac{1}{j!} \log j$$

$$= O(\lambda^2). \tag{A.38}$$

Part (a) of the lemma now follows from (A.37), (A.38), and the facts that $E(0.5\lambda) = \sigma_u^2/(2\pi f_{jY})$ and that $E(I_{jU}^2)$ is bounded.

LEMMA 4. *For* $\log^8 m \leq j \leq k \leq m$,

(a)

$$E\left[\left(\log \frac{I_{jZ}}{f_{jY}} - \log \frac{I_{jY}}{f_{jY}}\right)\left(\log \frac{I_{kZ}}{f_{kY}} - \log \frac{I_{kY}}{f_{kY}}\right)\right] = O(\omega_k^{4d}) + O\left(\omega_k^{2d} \frac{\log k}{j}\right)$$

(b)

$$E\log \frac{I_{kZ}}{f_{kY}} = E\log \frac{I_{kY}}{f_{kY}} + O(\omega_k^{2d}).$$

Proof of Lemma 4. We will prove only (a) because (b) follows along similar lines. To demonstrate (a), we expand the product $E[(\log(I_{jZ}/f_{jY}) - \log(I_{jY}/f_{jY}))(\log(I_{kZ}/f_{kY}) - \log(I_{kY}/f_{kY}))]$ into its four components and obtain expressions for each of them. Throughout the rest of this proof, we will assume that $\log^8 m \leq j < k \leq m$. The proof for $\log^8 m \leq j = k \leq m$ follows along similar lines but is simpler. Let

$$\mathbf{X} = \left(\frac{A_{jZ}}{f_{jY}^{1/2}}, \frac{B_{jZ}}{f_{jY}^{1/2}}, \frac{A_{kZ}}{f_{kY}^{1/2}}, \frac{B_{kZ}}{f_{kY}^{1/2}}\right)' \equiv (X_1, X_2, X_3, X_4)'$$

and let $\mathbf{U} = (u_1, u_2, \ldots, u_n)'$. The conditional distribution of $\mathbf{X}$ given $\mathbf{U}$ is multivariate normal with mean

$$\boldsymbol{\mu}_U = \left(\frac{A_{jU}}{f_{jY}^{1/2}}, \frac{B_{jU}}{f_{jY}^{1/2}}, \frac{A_{kU}}{f_{kY}^{1/2}}, \frac{B_{kU}}{f_{kY}^{1/2}}\right)'$$

and variance covariance matrix $\boldsymbol{\Sigma}$. From Robinson (1995, Theorem 2, parts (a)–(d)), we have

$$\boldsymbol{\Sigma} = \frac{1}{2}\mathbf{I}_4 + o(1) \tag{A.39}$$

uniformly in $\log^8 m \leq j < k \leq m$. Define $\boldsymbol{\Psi} = \boldsymbol{\Sigma}^{-1}$. Now

$$E\log \frac{I_{jZ}}{f_{jY}} \log \frac{I_{kZ}}{f_{kY}} = E\log \frac{A_{jZ}^2 + B_{jZ}^2}{f_{jY}} \log \frac{A_{kZ}^2 + B_{kZ}^2}{f_{kY}}$$

$$= E\left[E\left\{\log \frac{A_{jZ}^2 + B_{jZ}^2}{f_{jY}} \log \frac{A_{kZ}^2 + B_{kZ}^2}{f_{kY}} \Big| \mathbf{U}\right\}\right]. \tag{A.40}$$

We now work with the conditional expectation in (A.40). For $\mathbf{x} = (x_1, x_2, x_3, x_4)'$, define

$$g(\mathbf{x}) = (2\pi)^{-2} |\mathbf{\Psi}|^{1/2} \log(x_1^2 + x_2^2) \log(x_3^2 + x_4^2).$$

Then

$$E\left\{ \log \frac{A_{jZ}^2 + B_{jZ}^2}{f_{jY}} \log \frac{A_{kZ}^2 + B_{kZ}^2}{f_{kY}} \,\bigg|\, \mathbf{U} \right\}$$

$$= \int g(\mathbf{x}) \exp\left(-\frac{(\mathbf{x} - \boldsymbol{\mu}_U)' \mathbf{\Psi} (\mathbf{x} - \boldsymbol{\mu}_U)}{2} \right) d\mathbf{x}$$

$$= \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \mathbf{\Psi} \mathbf{x}}{2} \right) \exp(\mathbf{x}' \mathbf{\Psi} \boldsymbol{\mu}_U) \exp\left(-\frac{\boldsymbol{\mu}_U' \mathbf{\Psi} \boldsymbol{\mu}_U}{2} \right) d\mathbf{x}$$

$$= \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \mathbf{\Psi} \mathbf{x}}{2} \right) d\mathbf{x}$$

$$+ \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \mathbf{\Psi} \mathbf{x}}{2} \right) \left[\exp(\mathbf{x}' \mathbf{\Psi} \boldsymbol{\mu}_U) \exp\left(-\frac{\boldsymbol{\mu}_U' \mathbf{\Psi} \boldsymbol{\mu}_U}{2} \right) - 1 \right] d\mathbf{x}$$

$$= E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} + \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \mathbf{\Psi} \mathbf{x}}{2} \right)$$

$$\times \left[\exp(\mathbf{x}' \mathbf{\Psi} \boldsymbol{\mu}_U) \exp\left(-\frac{\boldsymbol{\mu}_U' \mathbf{\Psi} \boldsymbol{\mu}_U}{2} \right) - 1 \right] d\mathbf{x}.$$

Let

$$R(\boldsymbol{\mu}_U) = \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \mathbf{\Psi} \mathbf{x}}{2} \right) \left[\exp(\mathbf{x}' \mathbf{\Psi} \boldsymbol{\mu}_U) \exp\left(-\frac{\boldsymbol{\mu}_U' \mathbf{\Psi} \boldsymbol{\mu}_U}{2} \right) - 1 \right] d\mathbf{x}.$$

Thus,

$$E \log \frac{I_{jZ}}{f_{jY}} \log \frac{I_{kZ}}{f_{kY}} = E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} + ER(\boldsymbol{\mu}_U), \tag{A.41}$$

and we now proceed to work with $ER(\boldsymbol{\mu}_U)$. Let

$$\boldsymbol{\mu}_{jU} = \left(\frac{A_{jU}}{f_{jY}^{1/2}}, \frac{B_{jU}}{f_{jY}^{1/2}} \right)', \qquad \boldsymbol{\mu}_{kU} = \left(\frac{A_{kU}}{f_{kY}^{1/2}}, \frac{B_{kU}}{f_{kY}^{1/2}} \right)'$$

and partition $\mathbf{\Psi}$ as

$$\mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_{jj} & \mathbf{\Psi}_{jk} \\ \mathbf{\Psi}_{jk}' & \mathbf{\Psi}_{kk} \end{pmatrix}.$$

Let $v = \mathbf{x}' \mathbf{\Psi} \boldsymbol{\mu}_U - \boldsymbol{\mu}_U' \mathbf{\Psi} \boldsymbol{\mu}_U / 2$. By a Taylor series expansion, we have $\exp(v) - 1 = \sum_{i=1}^3 v^i / i! + (v^4/4!) \exp(\theta v)$, where $0 < \theta < 1$. This implies that

$$\exp(v) - 1 = \sum_{i=1}^3 v^i / i! + O\left(v^4 I(v < 0) + v^4 \exp\left(\mathbf{x}' \mathbf{\Psi} \boldsymbol{\mu}_U - \frac{\boldsymbol{\mu}_U' \mathbf{\Psi} \boldsymbol{\mu}_U}{2} \right) I(v > 0) \right)$$

and hence

$$ER(\boldsymbol{\mu}_U) = \sum_{i=1}^{4} R_{1i}, \tag{A.42}$$

where

$$R_{1i} = E\left\{ \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}}{2} \right) v^i/i! \, d\mathbf{x} \right\} \qquad i = 1,2,3$$

and

$$R_{14} = O\left(E\left\{ \int |g(\mathbf{x})| \exp\left(-\frac{\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}}{2} \right) v^4 d\mathbf{x} \right\} \right)$$

$$+ O\left(E\left\{ \int |g(\mathbf{x})| \exp\left(-\frac{\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}}{2} \right) v^4 \exp\left(\mathbf{x}'\boldsymbol{\Psi}\boldsymbol{\mu}_U - \frac{\boldsymbol{\mu}_U'\boldsymbol{\Psi}\boldsymbol{\mu}_U}{2} \right) d\mathbf{x} \right\} \right)$$

$$\equiv R_{14,1} + R_{14,2}.$$

Let $g_1(\mathbf{x}) = g(\mathbf{x}) \exp(-(\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}/2))$. Because $g_1(\mathbf{x})$ is an even function and $\mathbf{x}'\boldsymbol{\Psi}\boldsymbol{\mu}_U$ is an odd function of $\mathbf{x}$, it follows that

$$\int g_1(\mathbf{x})(\mathbf{x}'\boldsymbol{\Psi}\boldsymbol{\mu}_U) \, d\mathbf{x} = 0 \tag{A.43}$$

for all $\boldsymbol{\mu}_U$. Hence, we have

$$R_{11} = -E\left(\frac{\boldsymbol{\mu}_U'\boldsymbol{\Psi}\boldsymbol{\mu}_U}{2} \right) E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}}$$

$$= \left[-E\left(\frac{\boldsymbol{\mu}_{jU}'\boldsymbol{\Psi}_{jj}\boldsymbol{\mu}_{jU}}{2} \right) - E\left(\frac{\boldsymbol{\mu}_{kU}'\boldsymbol{\Psi}_{kk}\boldsymbol{\mu}_{kU}}{2} \right) - E(\boldsymbol{\mu}_{jU}'\boldsymbol{\Psi}_{jk}\boldsymbol{\mu}_{kU}) \right] E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}}$$

$$= \left[-E\left(\frac{\boldsymbol{\mu}_{jU}'\boldsymbol{\Psi}_{jj}\boldsymbol{\mu}_{jU}}{2} \right) - E\left(\frac{\boldsymbol{\mu}_{kU}'\boldsymbol{\Psi}_{kk}\boldsymbol{\mu}_{kU}}{2} \right) \right] E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}}. \tag{A.44}$$

Applying (A.43) again, we get

$$R_{12} = 0.5E\left\{ \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}}{2} \right) (\mathbf{x}'\boldsymbol{\Psi}\boldsymbol{\mu}_U)^2 d\mathbf{x} \right\}$$

$$+ 8^{-1}E\left\{ \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}}{2} \right) (\boldsymbol{\mu}_U'\boldsymbol{\Psi}\boldsymbol{\mu}_U)^2 d\mathbf{x} \right\}$$

$$= 0.5E\left\{ \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}'\boldsymbol{\Psi}\mathbf{x}}{2} \right) (\mathbf{x}'\boldsymbol{\Psi}\boldsymbol{\mu}_U)^2 d\mathbf{x} \right\} + 8^{-1}E(\boldsymbol{\mu}_U'\boldsymbol{\Psi}\boldsymbol{\mu}_U)^2 E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}}. \tag{A.45}$$

From equation (A.39), it follows that there exist finite positive constants K_1 and K_2 such that

$$0 < K_1 < \lambda_{\min} < \lambda_{\max} < K_2 < \infty \tag{A.46}$$

uniformly in $\log^8 m \le j < k \le m$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest characteristic roots, respectively, of $\boldsymbol{\Psi}$. Hence, we have

$$\boldsymbol{\mu}'_U \boldsymbol{\Psi} \boldsymbol{\mu}_U \le \lambda_{\max} \boldsymbol{\mu}'_U \boldsymbol{\mu}_U \le K_2 \boldsymbol{\mu}'_U \boldsymbol{\mu}_U, \tag{A.47}$$

giving

$$E(\boldsymbol{\mu}'_U \boldsymbol{\Psi} \boldsymbol{\mu}_U)^2 = O(\omega_k^{4d}).$$

Also

$$\left| E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} \right| \le \left(E \left\{ \log \frac{I_{jY}}{f_{jY}} \right\}^2 \right)^{1/2} \left(E \left\{ \log \frac{I_{kY}}{f_{kY}} \right\}^2 \right)^{1/2} = O(1)$$

from Lemma 5 of Hurvich, Deo, and Brodsky. Thus, the second term on the right in (A.45) is $O(\omega_k^{4d})$. We now tackle the first term in (A.45). Noting that $E(\boldsymbol{\mu}'_{jU} \boldsymbol{\mu}_{kU}) = 0$ and letting $\mathbf{x}_1 = (x_1, x_2)'$ and $\mathbf{x}_2 = (x_3, x_4)'$, we get

$$0.5E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}' \boldsymbol{\Psi} \boldsymbol{\mu}_U)^2 d\mathbf{x}$$

$$= 0.5E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}'_1 \boldsymbol{\Psi}_{jj} \boldsymbol{\mu}_{jU})^2 d\mathbf{x}$$

$$+ 0.5E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}'_2 \boldsymbol{\Psi}_{kk} \boldsymbol{\mu}_{kU})^2 d\mathbf{x}$$

$$+ 0.5E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}'_1 \boldsymbol{\Psi}_{jk} \boldsymbol{\mu}_{kU})^2 d\mathbf{x}$$

$$+ 0.5E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}'_2 \boldsymbol{\Psi}'_{jk} \boldsymbol{\mu}_{jU})^2 d\mathbf{x}$$

$$+ E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}'_1 \boldsymbol{\Psi}_{jj} \boldsymbol{\mu}_{jU})(\mathbf{x}'_2 \boldsymbol{\Psi}'_{jk} \boldsymbol{\mu}_{jU}) d\mathbf{x}$$

$$+ E \int g(\mathbf{x}) \exp\left(-\frac{\mathbf{x}' \boldsymbol{\Psi} \mathbf{x}}{2} \right) (\mathbf{x}'_1 \boldsymbol{\Psi}_{jk} \boldsymbol{\mu}_{kU})(\mathbf{x}'_2 \boldsymbol{\Psi}_{kk} \boldsymbol{\mu}_{kU}) d\mathbf{x}$$

$$\equiv \sum_{i=1}^{6} R_{12,i}.$$

Using the Cauchy–Schwarz inequality and the fact that $\mathbf{y}'\mathbf{A}'\mathbf{A}\mathbf{y} \leq \text{trace}(\mathbf{A}'\mathbf{A})\mathbf{y}'\mathbf{y}$, we get

$$R_{12,5} \leq K\{\text{trace}(\mathbf{\Psi}'_{jk}\mathbf{\Psi}_{jk})\}^{1/2}E(\boldsymbol{\mu}'_{jU}\boldsymbol{\mu}_{jU}) \int |g(\mathbf{x})|\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)(\mathbf{x}'_1\mathbf{x}_1)^{1/2}(\mathbf{x}'_2\mathbf{x}_2)^{1/2}d\mathbf{x}$$

$$= O\left(\omega_j^{2d}\frac{\log k}{j}\right),$$

where we have used the fact that $\mathbf{\Psi}_{jk} = O(\log k/j)$. See Theorem 2 of Robinson (1995a). A similar argument shows that

$$R_{12,6} = O\left(\omega_k^{2d}\frac{\log k}{j}\right).$$

Thus, we conclude that

$$R_{12} = \sum_{i=1}^{4} R_{12,i} + O\left(\omega_j^{2d}\frac{\log k}{j}\right) + O\left(\omega_k^{2d}\frac{\log k}{j}\right) + O(\omega_k^{4d}). \tag{A.48}$$

To bound R_{13}, we note that because $g_1(\mathbf{x})$ is an even function,

$$\int g_1(\mathbf{x})(\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U)^3 d\mathbf{x} = 0. \tag{A.49}$$

Applying (A.43) and (A.49), we see that

$$R_{13} = -E\left\{\int g(\mathbf{x})\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)(\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U)^2(\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U)\,d\mathbf{x}\right\}$$

$$- 6^{-1}E(\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U)^3 \int g(\mathbf{x})\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)d\mathbf{x}.$$

Applying the Cauchy–Schwarz inequality and (A.47), we get

$$E\left\{\int g(\mathbf{x})\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)(\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U)^2(\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U)\,d\mathbf{x}\right\}$$

$$\leq K_2^3 E(\boldsymbol{\mu}'_U\boldsymbol{\mu}_U)^2 \int |g(\mathbf{x})|\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)(\mathbf{x}'\mathbf{x})\,d\mathbf{x}$$

$$= O(\omega_k^{4d}),$$

and hence

$$R_{13} = O(\omega_k^{4d}) + O(\omega_k^{6d}). \tag{A.50}$$

We now finally bound R_{14}. Arguments similar to the ones presented previously show that

$$R_{14,1} = O(\omega_k^{4d}), \tag{A.51}$$

and hence we concentrate on $R_{14,2}$. From the inequality $(a + b)^4 \leq K(|a|^4 + |b|^4)$, we get

$$E\left\{\int |g(\mathbf{x})|\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)v^4\exp\left(\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U - \frac{\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U}{2}\right)d\mathbf{x}\right\}$$

$$\leq KE\left\{\int |g(\mathbf{x})|\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)|\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U|^4\exp\left(\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U - \frac{\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U}{2}\right)d\mathbf{x}\right\}$$

$$+ KE\left\{\int |g(\mathbf{x})|\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)(\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U)^4\exp\left(\mathbf{x}'\mathbf{\Psi}\boldsymbol{\mu}_U - \frac{\boldsymbol{\mu}'_U\mathbf{\Psi}\boldsymbol{\mu}_U}{2}\right)d\mathbf{x}\right\}$$

$$\equiv R_{14,2,1} + R_{14,2,2}. \tag{A.52}$$

Letting $\mathbf{a}_U = \mathbf{\Psi}^{1/2}\boldsymbol{\mu}_U$ and using the definition of $g(\cdot)$ we have

$$R_{14,2,1} \leq KE\left\{\int \log^2(\mathbf{x}'\mathbf{x})\exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right)|\mathbf{x}'\mathbf{\Psi}^{1/2}\mathbf{a}_U|^4\exp\left(\mathbf{x}'\mathbf{\Psi}^{1/2}\mathbf{a}_U - \frac{\mathbf{a}'_U\mathbf{a}_U}{2}\right)d\mathbf{x}\right\}$$

$$\leq KE\left\{\int \log^2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x})\exp\left(-\frac{\mathbf{x}'\mathbf{x}}{2}\right)|\mathbf{x}'\mathbf{a}_U|^4\exp\left(\mathbf{x}'\mathbf{a}_U - \frac{\mathbf{a}'_U\mathbf{a}_U}{2}\right)d\mathbf{x}\right\}$$

$$\leq KE\left\{\int |\mathbf{x}'\mathbf{a}_U|^4\log^2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x})\exp\left(-\frac{(\mathbf{x} - \mathbf{a}_U)'(\mathbf{x} - \mathbf{a}_U)}{2}\right)d\mathbf{x}\right\}$$

$$\leq KE\left[(\mathbf{a}'_U\mathbf{a}_U)^2\int (\mathbf{x}'\mathbf{x})^2\log^2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x})\exp\left(-\frac{(\mathbf{x} - \mathbf{a}_U)'(\mathbf{x} - \mathbf{a}_U)}{2}\right)d\mathbf{x}\right]. \tag{A.53}$$

But using (A.46) and the fact that $\log^2|x| \leq C_1|x|^{-2} + C_2|x|^2$, we have

$$\int (\mathbf{x}'\mathbf{x})^2\log^2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x})\exp\left(-\frac{(\mathbf{x} - \mathbf{a}_U)'(\mathbf{x} - \mathbf{a}_U)}{2}\right)d\mathbf{x}$$

$$\leq \int (\mathbf{x}'\mathbf{x})^2\{C_1(\mathbf{x}'\mathbf{\Sigma}\mathbf{x})^{-2} + C_2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x})^2\}\exp\left(-\frac{(\mathbf{x} - \mathbf{a}_U)'(\mathbf{x} - \mathbf{a}_U)}{2}\right)d\mathbf{x}$$

$$\leq \int (\mathbf{x}'\mathbf{x})^2\{C_1(\mathbf{x}'\mathbf{x})^{-2} + C_2(\mathbf{x}'\mathbf{x})^2\}\exp\left(-\frac{(\mathbf{x} - \mathbf{a}_U)'(\mathbf{x} - \mathbf{a}_U)}{2}\right)d\mathbf{x}$$

$$= O(1) + O((\mathbf{a}'_U\mathbf{a}_U)^2) + O((\mathbf{a}'_U\mathbf{a}_U)^4). \tag{A.54}$$

Equations (A.53) and (A.54) imply that

$$R_{14,2,1} = O(\omega_k^{4d}). \tag{A.55}$$

To bound $R_{14,2,2}$ we note that

$$R_{14,2,2} \leq KE \left\{ \int \log^2(\mathbf{x}'\mathbf{x}) \exp\left(-\frac{\mathbf{x}'\mathbf{\Psi}\mathbf{x}}{2}\right) |\mathbf{a}_U'\mathbf{a}_U|^4 \exp\left(\mathbf{x}'\mathbf{\Psi}^{1/2}\mathbf{a}_U - \frac{\mathbf{a}_U'\mathbf{a}_U}{2}\right) d\mathbf{x} \right\}$$

$$\leq KE\left[(\mathbf{a}_U'\mathbf{a}_U)^4 \int \log^2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x}) \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x} \right]. \qquad \textbf{(A.56)}$$

But using (A.46) and the fact that $\log^2|x| \leq C_1|x|^{-1/4} + C_2|x|$, we have

$$\int \log^2(\mathbf{x}'\mathbf{\Sigma}\mathbf{x}) \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x}$$

$$\leq \int \{C_1(\mathbf{x}'\mathbf{x})^{-1/4} + C_2|\mathbf{x}'\mathbf{x}|\} \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x}, \qquad \textbf{(A.57)}$$

and we now bound this integral. Letting $A = \{\mathbf{x}'\mathbf{x} \leq 1\}$, we have

$$\int (\mathbf{x}'\mathbf{x})^{-1/4} \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x}$$

$$= \int_A (\mathbf{x}'\mathbf{x})^{-1/4} \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x}$$

$$+ \int_{A^c} (\mathbf{x}'\mathbf{x})^{-1/4} \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x}$$

$$\leq \int_A (\mathbf{x}'\mathbf{x})^{-1/4} d\mathbf{x} + \int_{A^c} \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x} = O(1).$$

Also

$$\int |\mathbf{x}'\mathbf{x}| \exp\left(-\frac{(\mathbf{x}-\mathbf{a}_U)'(\mathbf{x}-\mathbf{a}_U)}{2}\right) d\mathbf{x} = O(1) + O(\mathbf{a}_U'\mathbf{a}_U).$$

Thus, equations (A.56) and (A.57) imply that

$$R_{14,2,2} = O(\omega_k^{4d}). \qquad \textbf{(A.58)}$$

From (A.51), (A.55), and (A.58), we get

$$R_{14} = O(\omega_k^{4d}). \qquad \textbf{(A.59)}$$

From (A.41), (A.42), (A.44), (A.48), (A.50), and (A.59) we have

$$E \log \frac{I_{jZ}}{f_{jY}} \log \frac{I_{kZ}}{f_{kY}} = E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} \left[1 - E\left(\frac{\boldsymbol{\mu}_{jU}'\mathbf{\Psi}_{jj}\boldsymbol{\mu}_{jU}}{2}\right) - E\left(\frac{\boldsymbol{\mu}_{kU}'\mathbf{\Psi}_{kk}\boldsymbol{\mu}_{kU}}{2}\right) \right]$$

$$+ \sum_{i=1}^{4} R_{12,i} + O\left(\omega_k^{2d} \frac{\log k}{j}\right) + O(\omega_k^{4d}). \qquad \textbf{(A.60)}$$

Similar arguments establish that

$$E \log \frac{I_{jZ}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} = E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} \left[1 - E\left(\frac{\boldsymbol{\mu}_{jU}' \boldsymbol{\Psi}_{jj} \boldsymbol{\mu}_{jU}}{2} \right) \right]$$

$$+ R_{12,1} + R_{12,4} + O\left(\omega_k^{2d} \frac{\log k}{j} \right) + O(\omega_k^{4d}) \qquad \textbf{(A.61)}$$

and

$$E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kZ}}{f_{kY}} = E \log \frac{I_{jY}}{f_{jY}} \log \frac{I_{kY}}{f_{kY}} \left[1 - E\left(\frac{\boldsymbol{\mu}_{kU}' \boldsymbol{\Psi}_{kk} \boldsymbol{\mu}_{kU}}{2} \right) \right]$$

$$+ R_{12,2} + R_{12,3} + O\left(\omega_k^{2d} \frac{\log k}{j} \right) + O(\omega_k^{4d}). \qquad \textbf{(A.62)}$$

Using (A.60)–(A.62) we get the desired result.