5-1-2008

Optimal Bandwidth Choice for Interval Estimation in GMM Regression

Yixiao Sun

Peter C.B. Phillips

OPTIMAL BANDWIDTH CHOICE FOR INTERVAL ESTIMATION IN GMM REGRESSION

By

Yixiao Sun and Peter C.B. Phillips

May 2008

COWLES FOUNDATION DISCUSSION PAPER NO. 1661

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

http://cowles.econ.yale.edu/

Optimal Bandwidth Choice for Interval Estimation in GMM Regression*

Yixiao Sun
Department of Economics
University of California, San Diego

Peter C. B. Phillips
Cowles Foundation, Yale University,
University of Auckland & University of York

February 24, 2008

Abstract

In time series regression with nonparametrically autocorrelated errors, it is now standard empirical practice to construct confidence intervals for regression coefficients on the basis of nonparametrically studentized t-statistics. The standard error used in the studentization is typically estimated by a kernel method that involves some smoothing process over the sample autocovariances. The underlying parameter (M) that controls this tuning process is a bandwidth or truncation lag and it plays a key role in the finite sample properties of tests and the actual coverage properties of the associated confidence intervals. The present paper develops a bandwidth choice rule for M that optimizes the coverage accuracy of interval estimators in the context of linear GMM regression. The optimal bandwidth balances the asymptotic variance with the asymptotic bias of the robust standard error estimator. This approach contrasts with the conventional bandwidth choice rule for nonparametric estimation where the focus is the nonparametric quantity itself and the choice rule balances asymptotic variance with squared asymptotic bias. It turns out that the optimal bandwidth for interval estimation has a different expansion rate and is typically substantially larger than the optimal bandwidth for point estimation of the standard errors. The new approach to bandwidth choice calls for refined asymptotic measurement of the coverage probabilities, which are provided by means of an Edgeworth expansion of the finite sample distribution of the nonparametrically studentized t-statistic. This asymptotic expansion extends earlier work and is of independent interest. A simple plug-in procedure for implementing this optimal bandwidth is suggested and simulations confirm that the new plug-in procedure works well in finite samples. Issues of interval length and false coverage probability are also considered, leading to a secondary approach to bandwidth selection with similar properties.

JEL Classification: C13; C14; C22; C51

Keywords: Asymptotic expansion, Bias, Confidence interval, Coverage probability, Edgeworth expansion, Lag kernel, Long run variance, Optimal bandwidth, Spectrum.

*An earlier version of the paper was presented at Cowles 75th Anniversary Conference in June 2007. Sun acknowledges research support from the COR grant at UC San Diego. Phillips acknowledges partial research support from the Kelly Foundation and the NSF under Grant No. SES 06-47086.

1 Introduction

Robust inference in time series regression is typically accomplished by estimating the standard errors of the regression coefficients nonparametrically to allow for the effects of autocorrelation of unknown form by a kernel smoothing process. A critical element in achieving robustness is the bandwidth or truncation lag (M). Appropriate choice of M addresses the nonparametric autocorrelation but also affects other aspects of inference such as the coverage probability of associated confidence intervals. It seems sensible that the choice of M should take these other effects into account, for instance when confidence interval coverage probability is a primary concern. Conventional econometric approaches (e.g., Andrews, 1991; Newey and West, 1987, 1994) follow early practice in the time series spectral analysis literature by selecting M to minimize the asymptotic mean squared error (AMSE) of the nonparametric quantity itself, which in this context is the relevant long run variance (LRV). Such a choice of the smoothing parameter is designed to be optimal in the AMSE sense for the estimation of the LRV, but is not necessarily optimal and may not even be well suited for other purposes, as shown in Sun, Phillips and Jin (2007) in the context of hypothesis testing in a Gaussian location model.

The present paper pursues this theme of focused bandwidth choice in linear GMM regression by developing an approach to bandwidth selection that is based on minimizing the coverage probability error (CPE) of a relevant confidence interval in linear GMM regression. This approach to automated bandwidth selection requires asymptotic measurement of the coverage probabilities, which are provided by means of an Edgeworth expansion of the finite sample distribution of the nonparametrically studentized t-statistic. We show that the asymptotic coverage probability depends on the asymptotic bias and variance of the LRV estimator as well as other aspects of the data generating process. To minimize the coverage probability error, we would choose M to balance the asymptotic bias and variance. This selection process contrasts with the conventional MSE criterion that balances the *squared* asymptotic bias with variance. As a result, larger values of M are called for if coverage accuracy of confidence intervals is of primary concern. In particular, when second order kernels, such as the Parzen and Quadratic Spectral (QS) kernels are used, conventional wisdom and long historical practice in statistics suggests that M be of order $T^{1/5}$ as the sample size T increases. We show that if our goal is to achieve the best coverage accuracy of two-sided confidence intervals, then the optimal M should be of order $T^{1/3}$. Taking $M \sim T^{1/5}$ gives coverage errors of order $T^{-2/5}$ whereas $M \sim T^{1/3}$ gives coverage errors of order $T^{-2/3}$. Interestingly, the choice of M is not so critical in one-sided confidence intervals. As long as M increases faster than $T^{1/4}$ and more slowly than $T^{1/2}$, the dominant term in the Edgeworth expansion of the coverage error is of order $T^{-1/2}$ and does not depend on M. Again, if we use the MSE-optimal bandwidth $M \sim T^{1/5}$, the coverage error will of order $T^{-2/5}$, which is larger than $T^{-1/2}$ by an order of magnitude.

In addition to the difference in the rate of expansion, the CPE-optimal bandwidth differs from the MSE-optimal bandwidth in the following aspect: depending on the direction of the dominant asymptotic bias of the HAC estimator, the CPE-optimal bandwidth may trade the asymptotic bias with the asymptotic variance or zero out the asymptotic bias with the asymptotic variance. In the latter case, the coverage errors of two-sided confidence intervals will be of an order smaller than $T^{-2/3}$ when second order kernels are used. In the former case, we can use the Cornish-Fisher type of expansion and obtain high-order corrected critical values. These high-order corrected critical values are analogous to those obtained in Sun, Phillips and Jin (2007). The difference is that our correction here reflects both the asymptotic bias and variance while the correction in Sun, Phillips and Jin (2007) reflects only the asymptotic variance. With the high-order corrected critical values, the coverage error of two-sided confidence intervals will also be of an order smaller than $T^{-2/3}$. Therefore, the present paper makes two main innovations: the CPE-optimal bandwidth that minimizes the coverage error; and the high-order correction that further reduces the coverage error.

Another contribution of the present paper is to provide an automatic and data-dependent procedure to implement the CPE-optimal bandwidth. Following established statistical practice and the work of Andrews (1991), we use simple parametric models to capture the main features of the target vector process, that is, the product of the instruments with the regression error. This plug-in methodology allows us to gauge the values of the unknown parameters in the CPE-optimal bandwidth. The computational cost of our plug-in bandwidth procedure is the same as that of the conventional plug-in bandwidth so there is no increase in computation.

In a series of simulation experiments, we compare the coverage accuracy of conventional confidence intervals and new confidence intervals. We find that new confidence intervals outperform the conventional confidence intervals for all models considered, and often by a large margin. Prewhitening is shown to be effective in improving the coverage accuracy of both types of confidence intervals, especially for conventional confidence intervals. Nevertheless, new confidence intervals remain consistently more accurate in coverage probability than conventional confidence intervals.

The final contribution of the paper is to outline an alternative bandwidth choice rule that takes the length of the confidence interval into account. Two coverage types are considered in this approach – coverage of the true value (true coverage) and coverage of false values (false coverage). The bandwidth is selected to minimize the probability of false coverage after controlling for the probability of true coverage. The probability of false coverage indirectly measures the length of the confidence interval because the longer a confidence interval is, the more likely it is that the interval will cover false values. The optimal bandwidth formula from this approach turns out to be more complicated than the one that minimizes the absolute coverage error but the main conclusion remains valid: for confidence interval construction, it is generally advantageous to reduce bias in HAC estimation by undersmoothing.

Our theoretical development relies on the Edgeworth expansion of the nonparametrically studentized t-statistic. The Edgeworth expansion we obtain is of independent interest. For example, it may be used to search for the optimal kernel, if it exists, for the purpose of interval estimation. It may also be used to establish high-order refinements of the moving block bootstrap. Our Edgeworth expansion differs from the one obtained by Götze and Künsch (1996) in that they consider only nonparametrically studentized sample means and obtain the Edgeworth expansion with the remainder of order $o(T^{-1/2})$. The Edgeworth expansion they obtain is sufficient for proving the high order refinement of the moving block bootstrap for one-sided confidence intervals. In contrast, the Edgeworth expansion we obtain is for general linear GMM models with possible over-identification. To derive the CPE-optimal bandwidth, we have to establish the Edgeworth expansion with a remainder of order $o(M^{-q})$, where q is the so-called Parzen exponent of the kernel function used (Parzen (1957)). With a suitable choice of M, the remainder is smaller than $o(T^{-1/2})$ by an order of magnitude. This is also in contrast to the Edgeworth expansion established by Inoue and Shintani (2006) in that the remainder in their Edgeworth expansion is of the larger order $O(M^{-q})$. Therefore, this paper contributes to the statistical literature on Edgeworth expansions of nonparametrically studentized t-statistics. Nevertheless, our proofs are built upon those of Inoue and Shintani (2005, 2006), which in turn rely on Götze and Künsch (1996).

A paper with conceptual ideas related to those presented here is Hall and Sheather (1988). These authors considered interval estimation for a sample quantile where the asymptotic variance depends on the probability density function. As in the present paper, they used the absolute coverage error as the criterion to select the bandwidth for density estimation. The found that the optimal bandwidth should be of an order of magnitude smaller than is recommended by the square error theory. Their qualitative findings are analogous to ours although the problems considered and the technical machinery used are fundamentally different.

Other related papers include Kiefer, Vogelsang and Bunzel (2000), Kiefer and Vogelsang (2002a, 2002b and 2005). These papers considered alternative approximations to the finite sample distribution of the t-statistic for a given bandwidth. In constrast, we consider the conventional asymptotic normality approximation and choose the bandwidth to optimize the criteria that address the central concerns for interval estimation.

The plan of the paper is as follows. Section 2 describes the linear GMM model we consider and presents assumptions. Section 3 develops a high order Edgeworth expansion of the finite sample distribution of the t-statistic. This expansion is the basis for optimal bandwidth choice and high-order corrections. Sections 4 and 5 propose a selection rule for M that is suitable for confidence interval construction. Section 6 reports some simulation evidence on the performance of the new procedure. Section 7 outlines an alternative bandwidth choice rule. Section 8 concludes and discusses some possible extensions of the ideas and methods. For easy reference, notation is collected in the first subsection of the Appendix. Proofs and additional technical lemmas are given in the rest of the Appendix.

2 Model and Assumptions

We consider a linear regression model

$$y_t = x_t' \beta_0 + u_t, \tag{1}$$

where $x_t \in \mathbb{R}^{d_1}$ and u_t is a zero mean stationary process with a nonparametric autocorrelation structure. We assume that there exists a stochastic process $z_t \in \mathbb{R}^{d_2}$ such that the moment condition

$$E z_t u_t = 0 \tag{2}$$

holds. To identify the model parameter β_0, we assume $d_2 \geq d_1$. In the special case where $z_t = x_t$, the model reduces to conventional linear regression.

Given T_0 observations $(x_t', y_t', z_t')'$, $t = 1, 2, ..., T_0$, we are interested in inference about β_0. Let M be the bandwidth parameter used in heteroscedasticity-autocorrelation consistent (HAC) covariance matrix estimation and set $T = T_0 - M + 1$. Defining

$$G_T = \frac{1}{T} \sum_{t=1}^{T} z_t x_t', \ S_T = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_t u_t := \frac{1}{\sqrt{T}} \sum_{t=1}^{T} v_t, \tag{3}$$

the two step GMM estimator of β_0 based on the moment condition (2) satisfies

$$\sqrt{T} \left(\hat{\beta}_T - \beta_0 \right) = \left(G_T' \hat{\Omega}_T^{-1} G_T \right)^{-1} \left(G_T' \hat{\Omega}_T^{-1} S_T \right), \tag{4}$$

where $\hat{\Omega}_T$ is a consistent estimate of the long run variance matrix Ω_0 of v_t :

$$\Omega_0 = E z_t u_t^2 z_t' + \sum_{j=1}^{\infty} \left[z_{t+j} u_{t+j} u_t z_t' + z_t u_t u_{t+j} z_{t+j}' \right]. \tag{5}$$

It is standard empirical practice to estimate Ω_0 using kernel-based nonparametric estimators that smooth and truncate the sample autocovariance sequence. The resulting HAC estimate of

Ω_0 has the form[1]

$$\hat{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T}\left[z_t\hat{u}_t^2 z_t' + \sum_{j=1}^{M} k\left(\frac{j}{M}\right)\left(z_{t+j}\hat{u}_{t+j}\hat{u}_t z_t' + z_t\hat{u}_t\hat{u}_{t+j}z_{t+j}'\right)\right]. \tag{6}$$

In the above expression, $\hat{u}_t$ is the estimated residual $\hat{u}_t = y_t - x_t'\tilde{\beta}_T$ for some consistent initial estimate $\tilde{\beta}_T$, $k(\cdot)$ is the kernel function, and M is the bandwidth or truncation lag. Throughout the paper, we employ the first step GMM estimator as $\tilde{\beta}_T$:

$$\sqrt{T}\left(\tilde{\beta}_T - \beta_0\right) = \left(G_T' V_T G_T\right)^{-1}\left(G_T' V_T S_T\right), \tag{7}$$

where V_T is a weighting matrix.

When the model is just identified, we have

$$\sqrt{T}\left(\hat{\beta}_T - \beta_0\right) = G_T^{-1} S_T, \tag{8}$$

and the weighting matrix is irrelevant. When $z_t = x_t$, $\hat{\beta}_T$ reduces to the familiar OLS estimator. So, the analysis includes linear OLS regression as a special case.

Under some standard regularity conditions, $\sqrt{T}(\hat{\beta}_T - \beta_0)$ is known to be asymptotically normal with distribution

$$\sqrt{T}\left(\hat{\beta}_T - \beta_0\right) \to N(0, \Sigma_0) \tag{9}$$

where

$$\Sigma_0 = \left(G_0'\Omega_0^{-1}G_0\right)^{-1} \text{ and } G_0 = E\left(G_T\right). \tag{10}$$

This limit theorem provides the usual basis for robust testing about β_0. As in Rothenberg (1984), it is convenient to consider the distribution of the studentized statistic of a linear combination of the parameters as in the standardized expression

$$t_M = \sqrt{T}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\mathcal{R}'\left(\hat{\beta}_T - \beta_0\right), \tag{11}$$

where $\mathcal{R}$ is a $d_1 \times 1$ vector and $\hat{\Sigma}_T = (G_T'\hat{\Omega}_T^{-1}G_T)^{-1}$.

To establish an Edgeworth expansion for the distribution of this studentized statistic, we maintain the following conditions.

[1] This HAC estimator differs slightly from the typical formula in the literature in that the number of terms in the sample covariance sums is the same regardless of the order of the sample covariance. We use this modified HAC estimator in rigorously proving the validity of the Edgeworth expansion up to order $o(M/T)$. The technical modification is not necessary for a lower order Edgeworth expansion. Similar modifications have been employed to facilitate theoretical developments in the bootstrap literature, e.g. Hall and Horowitz (1996).

Assumption 1. **(a)** *There is a unique* $\beta_0 \in \mathbb{R}^{d_1}$ *such that* $E\left(z_t\left(y_t - x_t'\beta_0\right)\right) = 0$. **(b)** *The long run variance matrix* $\Omega_0 = \sum_{j=-\infty}^{\infty} E(z_0 u_0 u_j z_j')$ *is positive definite.*

Assumption 2. **(a)** $\{x_t', y_t', z_t'\}'$ *is strictly stationary and strongly mixing with mixing coefficients satisfying* $\alpha_m \leq d^{-1}\exp(-dm)$ *for some* $d > 0$.

(b) *Let* $R_t = \left((z_t u_t)', vec\left(z_t x_t'\right)'\right)'$. *Then* $E\left\|R_t\right\|^{r+\eta} < \infty$ *for* $r \geq 16$ *and some* $\eta > 0$.

Assumption 3. *Let* $\mathcal{F}_a^b$ *denote the sigma-algebra generated by* $R_a, R_{a+1}, ..., R_b$. *For all* $m, s, t = 1, 2, ...$ *and* $A \in \mathcal{F}_{t-s}^{t+s}$

$$E\left|P\left(A|\mathcal{F}_{-\infty}^{t-1} \cup \mathcal{F}_{t+1}^{\infty}\right) - P\left(A|\mathcal{F}_{t-s-m}^{t-1} \cup \mathcal{F}_{t+1}^{t+s+m}\right)\right| \leq d^{-1}\exp(-dm). \tag{12}$$

Assumption 4. *For all* $m, t = 1, 2, ...$ *and* $\theta \in \mathbb{R}^{d_2(d_1+1)}$ *such that* $1/d < m < t$ *and* $\|\theta\| \geq d$,

$$E\left|E\left\{\exp\left[i\theta'\left(\sum_{s=t-m}^{t+m}\left(R_s - ER_s\right)\right)\right]\middle|\mathcal{F}_{-\infty}^{t-1} \cup \mathcal{F}_{t+1}^{\infty}\right\}\right| \leq \exp\left(-d\right). \tag{13}$$

Assumption 5. $k\left(\cdot\right): \mathbb{R} \to [0, 1]$ *is an even function which satisfies: (i)* $k(x) = 1 - g_q\left|x\right|^q + O(\left|x\right|^{2q})$ *as* $x \to 0+$ *for some* $q \in \mathbb{Z}^+$ *and* $q \geq 1$; *(ii)* $k(x) = 0$ *for* $|x| \geq 1$; *and (iii)* $k(x)$ *is continuous at* 0 *and at all but a finite number of other points.*

Assumption 6. $M \sim CT^{1/(q+1)}$ for some constant $C > 0$ as $T \to \infty$.

Assumption 7. *The weighting matrix* V_T *converges to a positive definite matrix* V_0 *such that*

$$P\left(T^{q/(2q+1)}\left\|V_T - V_0\right\| > \varepsilon \log T\right) = o\left(T^{-\chi}\right) \tag{14}$$

for all $\varepsilon > 0$ *and some* $\chi \geq 2$.

Assumption 1(a) is a standard model identification condition. The condition holds when the rank of $E(z_t x_t')$ is at least d_1. Assumption 1(b) is also standard and ensures the limit distribution of the GMM estimator $\hat{\beta}_T$ is nondegenerate. The strong mixing condition in Assumption 2(a) is a convenient weak dependence condition but is stronger than necessary. It can be shown that all results in the paper hold provided that $(x_t', y_t', z_t')'$ can be approximated sufficiently well by a suitable strong mixing sequence. More specifically, the assumption can be replaced by a condition that ensures the existence of a strong mixing sequence of sub-σ fields $\{\mathcal{D}_t\}$ for which $(x_t', y_t', z_t')'$ can be approximated by a $\mathcal{D}_{t-m}^{t+m}$-measurable process $(x_{t,m}', y_{t,m}', z_{t,m}')'$ with

$$E\left\|(x_t', y_t', z_t')' - (x_{t,m}', y_{t,m}', z_{t,m}')'\right\| \leq d^{-1}\exp(-dm). \tag{15}$$

See Götze and Hipp (1983) for details. Assumption 2(b) is a convenient moment condition and is likely not to be the weakest possible. Lahiri (1993) provides a discussion on the validity of Edgeworth expansions under weaker moment conditions.

Assumption 3 is an approximate Markov-type property. It says that the conditional probability of an event $A \in \mathcal{F}_{t-s}^{t+s}$, given the larger σ-field $\mathcal{F}_{-\infty}^{t-1} \cup \mathcal{F}_{t+1}^{\infty}$, can be approximated with increasing accuracy when the conditioning σ-field $\mathcal{F}_{t-s-m}^{t-1} \cup \mathcal{F}_{t+1}^{t+s+m}$ grows with m. This assumption trivially holds when R_t is a finite order Markov process. Assumption 4 is a Cramér-type condition in the weakly dependent case. It imposes some restrictions on the joint distribution of R_t. As shown by Götze and Hipp (1983), the standard Cramér condition on the marginal distribution, namely that

$$\limsup_{\|\theta\| \to \infty} \left| E \exp\left(i\theta' R_t\right)\right| < 1, \tag{16}$$

is not enough to establish a "regular" Edgeworth expansion for the normalized sample mean. The Cramér-type condition given in Assumption 4 is analogous to those in Götze and Hipp (1983), Götze and Künsch (1996), and Inoue and Shintani (2006). Assumption 5 imposes some restrictions on the kernel function. The parameter q is the so-called Parzen exponent (e.g. Parzen (1957)). For first order asymptotics, it suffices to assume that $k(x) = 1 - g_q |x|^q + o(|x|^q)$, while for higher order asymptotics, we have to strengthen $o(|x|^q)$ to $O(|x|^{2q})$.

Assumption 6 is a rate condition on the bandwidth expansion. The given order $M \sim O\left(T^{1/q+1}\right)$ is the expansion rate for M such that the asymptotic bias and variance of the LRV estimator are of the same order of magnitude. Assumption 7 requires that that the weight matrix V_T converges to V_0 at a certain rate. It holds trivially in the special case where $V_T = (T^{-1}\sum_{t=1}^T z_t z_t')^{-1}$. When V_T is the inverse of a general nonparametric estimator of the HAC covariance matrix, Assumption 7 holds if the underlying bandwidth is proportional to $T^{1/(2q+1)}$ and if V_T has an Edgeworth expansion with an error term of order $o\left(T^{-\chi}\right)$. A direct implication of Assumption 7 is that the first step estimator $\tilde{\beta}_T$ can be approximated well by an estimator with finite high order moments. More specifically, let $\check{\beta}_T$ be defined according to

$$\sqrt{T}\left(\check{\beta}_T - \beta_0\right) = \left(G_0' V_0 G_0\right)^{-1}\left(G_0' V_0 S_T\right), \tag{17}$$

where $G_0 = E\left(G_T\right)$ is defined as before. Assumption 7 implies that the first step estimator $\tilde{\beta}_T$ satisfies

$$P\left(\sqrt{T}\left(\left\|\tilde{\beta}_T - \check{\beta}\right\|\right) > \varepsilon\left(\log^2 T\right) T^{-q/(2q+1)}\right) = o\left(T^{-\chi}\right), \tag{18}$$

for all $\varepsilon > 0$ and some $\chi \geq 2$, as shown in Lemma A.3(c) in the Appendix. While the consistency of the first step estimator is sufficient for first order asymptotic theory, further conditions are needed for higher-order analysis. In fact, the higher order properties of the studentized statistic depend crucially on how the first step estimator is constructed. Note that condition (18) is different from Assumption (i) in Inoue and Shintani (2006) where it is assumed that the r-th moment of $\sqrt{T}(\tilde{\beta}_T - \beta_0)$ is finite. The requirement of finite higher moments may be restrictive and there is some advantage in avoiding direct assumptions of this type.

3 Edgeworth Expansions for the Studentized Statistic

This section develops Edgeworth expansions for t_M, thereby extending the work of Götze and Künsch (1996) which gave an Edgeworth expansion for the studentized mean. It also provides a refinement of the results in Götze and Künsch (1996) and Inoue and Shintani (2006). The latter two papers do not include the asymptotic bias of the HAC variance estimator in their Edgeworth approximations while we explicitly take the bias into account. For our purpose of optimal bandwidth choice that is investigated here, it is necessary to deal explicitly with the bias and not to leave it as part of the remainder term in the Edgeworth expansion.

To establish the Edgeworth expansion for t_M, we first establish the validity of some stochastic approximations. Let

$$\bar{\Omega}_T = \sum_{j=-M}^{M} k(j/M)\Gamma_j, \ \Gamma_j = \begin{cases} Ev_{t+j}v_t', & j \geq 0, \\ Ev_t v_{t-j}', & j < 0, \end{cases} \tag{19}$$

and set

$$\begin{aligned} g_T &= \mathbf{a}'S_T + \mathbf{b}'\left[vec\left(G_T - G_0\right) \otimes S_T\right] + \mathbf{c}'\left[vec\left(\hat{\Omega}_T - \Omega_0\right) \otimes S_T\right] \\ &\quad + \mathbf{d}'\left[vec(\hat{\Omega}_T - \bar{\Omega}_T) \otimes vec(\hat{\Omega}_T - \bar{\Omega}_T) \otimes S_T\right] \end{aligned} \tag{20}$$

for vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{d}$ given in Lemma 1 below. Then it can be shown that

$$P\left(|t_M - g_T| > \eta_T\right) = o\left(\eta_T\right), \tag{21}$$

for

$$\eta_T = \max\left(M^{-q}/\log T, (M/(T\log T))\right). \tag{22}$$

This stochastic approximation implies that the Edgeworth expansion for t_M is the same as that for g_T up to a certain order. However, even with this approximation it is still difficult to obtain the Edgeworth expansion for g_T as it depends on the first step estimator whose moments may not exist. To overcome this difficulty, we establish a further stochastic approximation, this time for g_T. Let $\check{\Omega}_T$ be defined as $\hat{\Omega}_T$ but with the first step estimator $\tilde{\beta}_T$ replaced by $\check{\beta}_T$ (defined in (17)), then we can show that

$$P\left(|g_T - h_T| > \eta_T\right) = o\left(\eta_T\right), \tag{23}$$

where

$$\begin{aligned} h_T &= \mathbf{a}' S_T + \mathbf{b}' \left[vec\left(G_T - G_0\right) \otimes S_T\right] + \mathbf{c}' \left[vec\left(\check{\Omega}_T - \Omega_0\right) \otimes S_T\right] \\ &\quad + \mathbf{d}' \left[vec(\check{\Omega}_T - \bar{\Omega}_T) \otimes vec(\check{\Omega}_T - \bar{\Omega}_T) \otimes S_T\right]. \end{aligned} \tag{24}$$

That is to say, for the purposes of the present development we can replace $\hat{\Omega}_T$ whose high order moments may not be finite by $\check{\Omega}_T$ whose higher order moments are finite.

Let

$$\sigma_0 = \left(\mathcal{R}'\Sigma_0\mathcal{R}\right)^{1/2}, \Sigma_0 = \left(G_0'\Omega_0^{-1}G_0\right)^{-1}, \tag{25}$$

$$\Lambda_{10} = \Sigma_0\mathcal{R}, \Lambda_{20} = \Omega_0^{-1}G_0\Sigma_0, \tag{26}$$

and

$$\begin{aligned} \Theta_{10} &= \Omega_0^{-1}G_0\Sigma_0\mathcal{R}, \\ \Theta_{20} &= \Omega_0^{-1}, \\ \Theta_{30} &= \Omega_0^{-1}G_0\Sigma_0G_0'\Omega_0^{-1}, \\ \Theta_{40} &= \Theta_{30} - \Theta_{20} - \frac{1}{2\sigma_0^2}\Theta_{10}\Theta_{10}'. \end{aligned} \tag{27}$$

We formalize the result on the stochastic approximations in the following Lemma.

Lemma 1 *Let Assumptions 1-7 hold, then*

$$P\left(|t_M - h_T| > \eta_T\right) = o(\eta_T) \tag{28}$$

where

$$\mathbf{a} = vec\left(Q_a\right), \ \mathbf{b} = vec\left(Q_b\right), \ \mathbf{c} = vec\left(Q_c\right), \ \mathbf{d} = vec\left(Q_d\right), \tag{29}$$

$$Q_a = \frac{1}{\sigma_0}\Theta_{10}, \tag{30}$$

$$Q_b = \frac{1}{\sigma_0}\left[\Lambda_{10}' \otimes \left(\Theta_{20} - \Theta_{30}\right)\right] - \frac{1}{\sigma_0}\left(\Lambda_{20} \otimes \Theta_{10}'\right) + \frac{1}{\sigma_0^3}\left[\left(\Theta_{10}\Lambda_{10}'\right) \otimes \Theta_{10}'\right], \tag{31}$$

$$Q_c = \frac{1}{\sigma_0}\left(\Theta_{10}' \otimes \Theta_{40}\right), \tag{32}$$

and

$$\begin{aligned} Q_d &= \frac{1}{4\sigma_0^3}\left\{\left[vec\left(\Theta_{10}'\right) \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta_{10}'\right)\right]\right\} \\ &+\frac{1}{2\sigma_0}\left\{\left[vec\left(\Theta_{10}'\right) \otimes \Theta_{40} \otimes \left(\Theta_{30} - \Theta_{20}\right)\right]\right\} \\ &+\frac{1}{2\sigma_0}\left\{K_{d_2,d_2^2}\left[vec\left(\Theta_{40}\right) \otimes \left(\Theta_{30} - \Theta_{20}\right) \otimes \Theta_{10}'\right]\right\} \\ &-\frac{1}{2\sigma_0^2}\left\{vec\left(\Theta_{10}' \otimes \left[\frac{1}{2\sigma_0}\Theta_{40} - \frac{1}{2\sigma_0^3}\Theta_{10}\Theta_{10}'\right]\right)\left[\Theta_{10}' \otimes \Theta_{10}'\right]\right\}. \end{aligned} \tag{33}$$

Lemma 1 implies that t_M and h_T have the same Edgeworth expansion up to a certain order. Thus, to establish the Edgeworth expansion for t_M, it suffices to establish that for h_T. The most difficult part is to control the approximation error as h_T can not be written as a finite sum of means. In fact, $\check{\Omega}_T$ is the sum of M different means, namely, the sample covariance for lags $0, 1, ..., M-1$ with M increasing with the sample size T. Following recent studies by Götze and Künsch (1996) and Lahiri (1996), we are able to rigorously justify this expansion. Details are given in the Appendix.

Let $\tilde{\Omega}_T$ be the HAC estimator of Ω_0 based on $\{v_t\}$. Define

$$\rho_{1,\infty} = \lim_{T\to\infty} M^q E 2\mathbf{a}' S_T \mathbf{c}'[vec\left(\bar{\Omega}_T - \Omega_0\right) \otimes S_T], \tag{34}$$

$$\rho_{2,\infty} = \lim_{T\to\infty} \frac{T}{M} E 2\mathbf{a}' S_T \mathbf{c}'[vec\left(\check{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T], \tag{35}$$

and

$$\begin{aligned} \kappa_{1,\infty} &= \lim_{T\to\infty} \sqrt{T}\mathbf{b}'\left[E\left(vec\left(G_T - G_0\right) \otimes S_T\right)\right] \\ &+ \lim_{T\to\infty} \sqrt{T}\mathbf{c}'\left[E\left(vec\left(\check{\Omega}_T - \Omega_0\right) \otimes S_T\right)\right], \end{aligned} \tag{36}$$

$$\begin{aligned} \kappa_{2,\infty} &= 2\lim_{T\to\infty} \frac{T}{M} E\left(\mathbf{a}' S_T\right)\left\{\mathbf{d}'\left[vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T\right]\right\} \\ &+ \lim_{T\to\infty} \frac{T}{M} E\left[\mathbf{c}'\left(vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T\right)\right]^2, \end{aligned} \tag{37}$$

$$\begin{aligned} \kappa_{3,\infty} &= \lim_{T\to\infty} \sqrt{T} E\left(\mathbf{a}' S_T\right)^3 + 3\lim_{T\to\infty} \sqrt{T} E\left(\mathbf{a}' S_T\right)^2\left\{\mathbf{b}'\left[vec\left(G_T - G_0\right) \otimes S_T\right]\right\} \\ &+3\lim_{T\to\infty} \sqrt{T} E\left(\mathbf{a}' S_T\right)^2\left\{\mathbf{c}'[vec\left(\check{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T]\right\}, \end{aligned} \tag{38}$$

$$\begin{aligned} \kappa_{4,\infty} &= 4\lim_{T\to\infty} \frac{T}{M} E\left(\mathbf{a}' S_T\right)^3\left\{\mathbf{d}'\left[vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T\right]\right\} \\ &+6\lim_{T\to\infty} \frac{T}{M} E\left(\mathbf{a}' S_T\right)^2\left\{\mathbf{c}'[vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T]\right\}^2. \end{aligned} \tag{39}$$

The following theorem gives the Edgeworth expansion for t_M.

Theorem 2 *Let Assumptions 1-7 hold, then*

$$\sup_{x\in\mathbb{R}}\left|P(t_M<x)-\Phi_T(x)\right|=o(\eta_T) \tag{40}$$

where

$$\Phi_T(x)=\Phi(x)+\frac{1}{\sqrt{T}}p_1(x)\phi(x)+\frac{M}{T}p_2(x)\phi\left(x\right)+\frac{1}{M^q}p_3\left(x\right)\phi(x), \tag{41}$$

for polynomials $p_1(x)$, $p_2(x)$ *and* $p_3\left(x\right)$ *defined by*

$$\begin{aligned} p_1\left(x\right) &= -\kappa_{1,\infty}-\frac{1}{6}\left(\kappa_{3,\infty}-3\kappa_{1,\infty}\right)\left(x^2-1\right), \\ p_2\left(x\right) &= -\frac{1}{2}\left(\rho_{2,\infty}+\kappa_{2,\infty}\right)x-\frac{1}{24}\left(\kappa_{4,\infty}-6\kappa_{2,\infty}\right)(x^3-3x), \\ p_3(x) &= -\frac{1}{2}\rho_{1,\infty}x. \end{aligned} \tag{42}$$

As is clear from the proof of the theorem given in the Appendix, the coefficients for the polynomials $p_1(x)$, $p_2\left(x\right)$, $p_3\left(x\right)$ depend on the kernel function used. A direct implication is that higher order asymptotics are able to at least partially capture the effects of the kernel function. In contrast, the first order normal limit does not depend on the kernel function used. This is one of several reasons for poor finite sample performance in first order asymptotics.

The Edgeworth expansion in Theorem 2 consists of a term of order $1/\sqrt{T}$ plus a second "superimposed" series of terms of order M^{-q} and M/T. The term of order $(1/\sqrt{T})$ is typical in Edgeworth expansions. It is convenient to regard it as a contribution from the mean and skewness of the t-statistic. The M^{-q} term arises from the type I finite sample bias of the HAC estimator. The type I bias is the same as the nonparametric bias in spectral density estimation when the time series is observed and used in estimation. The M/T term arises from the type II finite sample bias of the HAC estimator, the variance of the HAC estimator, and the statistical dependence of the HAC estimator on $\mathbf{a}'S_T$, a quantity that largely captures the randomness of the numerator of the t-statistic. The type II bias, reflected in the term containing $\rho_{2,\infty}$, is due to the unobservability of v_t. This term can not be avoided because we have to estimate β first in order to estimate v_t and construct the HAC estimator.

Note that $p_1\left(x\right)$ is an even polynomial while $p_2(x)$ and $p_3\left(x\right)$ are odd polynomials. It follows immediately from Theorem 2 that for two-sided probabilities we have

$$\sup_{x\in\mathbb{R}^+}\left|P(|t_M|<x)-\Phi_T^{|\cdot|}(x)\right|=o(\eta_T), \tag{43}$$

where

$$\Phi_T^{|\cdot|}(x)=\Phi(x)-\Phi\left(-x\right)+\frac{2M}{T}p_2(x)\phi\left(x\right)+\frac{2}{M^q}p_3\left(x\right)\phi(x). \tag{44}$$

In general, the values of $\rho_{i,\infty}$ and $\kappa_{i,\infty}$ depend on the underlying data generating process in a complicated way. In the special case that v_t follows a linear Gaussian process, we can obtain analytical expressions for $\rho_{i,\infty}$, $i = 1, 2$ and $\kappa_{i,\infty}$, $i = 1, 2, 3, 4$. For our proposes here, it suffices to obtain the closed form solutions for $\rho_{1,\infty}, \rho_{2,\infty}, \kappa_{2,\infty}$ and $\kappa_{4,\infty}$ as the optimal bandwidth given in the next section depends only on these parameters. From now on we ignore the technical modification (c.f. footnote 1) and employ the conventional kernel-based HAC estimator. Let

$$\tilde{\Gamma}_j^c = \begin{cases} T^{-1}\sum_{t=1}^{T-j} v_{t+j}v_t', & j \geq 0, \\ T^{-1}\sum_{t=1}^{T+j} v_t v_{t-j}', & j < 0, \end{cases} \quad \check{\Gamma}_j^c = \begin{cases} T^{-1}\sum_{t=1}^{T-j} \check{v}_{t+j}\check{v}_t', & j \geq 0, \\ T^{-1}\sum_{t=1}^{T+j} \check{v}_t \check{v}_{t-j}', & j < 0, \end{cases} \tag{45}$$

$$\tilde{\Omega}_T^c = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\tilde{\Gamma}_j^c, \; \check{\Omega}_T^c = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)\check{\Gamma}_j^c, \text{ and } \bar{\Omega}_T^c = \sum_{j=-M}^{M} k\left(\frac{j}{M}\right)E\tilde{\Gamma}_j^c. \tag{46}$$

Define

$$\mu_1 = \int_{-\infty}^{\infty} k(x)dx, \; \mu_2 = \int_{-\infty}^{\infty} k^2(x)dx, \tag{47}$$

and

$$\digamma_0 = V_0G_0(G_0'V_0G_0)^{-1}G_0', \; \Omega_0^{(q)} = \sum_{j=-\infty}^{\infty} |j|^q \, \Gamma_j. \tag{48}$$

The next proposition gives closed-form expressions for $\rho_{1,\infty}, \rho_{2,\infty}, \kappa_{2,\infty}$ and $\kappa_{4,\infty}$ when $\check{\Omega}_T$, $\tilde{\Omega}_T$ and $\bar{\Omega}_T$ are replaced by $\check{\Omega}_T^c$, $\tilde{\Omega}_T^c$ and $\bar{\Omega}_T^c$ respectively.

Proposition 3 *Assume that v_t follows a linear Gaussian process:*

$$v_t = \sum_{s=0}^{\infty} \Psi_s e_{t-s}, \; t = 1, 2, ..., T \tag{49}$$

where $e_t \sim iidN(0, \Sigma_e)$ and Ψ_s satisfies $\sum_{s=1}^{\infty} s^4 \|\Psi_s\| < \infty$. Then

(a) $\rho_{1,\infty} = g_q \mathcal{R}'\Sigma_0 G_0'\Omega_0^{-1}\Omega_0^{(q)}\Omega_0^{-1}G_0\Sigma_0\mathcal{R}\left[\mathcal{R}'G_0^{-1}\Omega_0\left(G_0^{-1}\right)'\mathcal{R}\right]^{-1}$,

(b) $\rho_{2,\infty} = -2\mu_1\mathbf{c}'\left[vec\left(\Omega_0\digamma_0 + \digamma_0'\Omega_0 - \digamma_0'\Omega_0\digamma_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right],$
$+2\mu_1\mathbf{c}'\left\{I_{d_2^3} + K_{d_2^2,d_2}\left(I_{d_2}\otimes K_{d_2,d_2}\right)\right\}\left\{\left[\left(I-\digamma_0\right)'\left(\Omega_0\mathbf{a}\right)\right]\otimes vec\left[\Omega_0\left(I-\digamma_0\right)\right]\right\},$

(c) $\kappa_{2,\infty} = 2\mu_2 + 2\mu_2\left(d_2 - d_1\right),$

(d) $\kappa_{4,\infty} - 6\kappa_{2,\infty} = 6\mu_2.$

It is clear from Theorem 2 and Proposition 3 that the asymptotic expansion of the t-statistic depends on the first step estimator through the quantity $\digamma_0$, which in turn depends on V_0, the probability limit of the weighting matrix used in constructing the first step estimator. Although the first order asymptotics of the t-statistic does not reflect the estimation uncertainty in the first step GMM estimator, the higher order asymptotics do capture this uncertainty. Note that $\digamma_0$ becomes an identity matrix, which does not depend on the weighting matrix V_T, when the model is just identified. This is not surprising as in this case the weighting matrix is irrelevant.

The analytical expression for $\rho_{2,\infty}$ can be greatly simplified if the model is just identified or the first step estimator is asymptotically efficient so that $V_0 = \Omega_0^{-1}$. Some algebraic manipulations show that in both cases $\rho_{2,\infty} = \mu_1$. Combining Proposition 3 with equation (43), we obtain

$$\begin{aligned} P(|t_M| &< x) = \Phi(x) - \Phi(-x) \\ &- \left[\frac{M}{T} \left\{ \mu_1 x + \frac{1}{2}\mu_2 \left[x^3 + x(4d_2 - 4d_1 + 1) \right] \right\} + \frac{1}{M^q}\rho_{1,\infty} x \right] \phi(x) + o(\eta_T), \end{aligned} \tag{50}$$

uniformly over $x \in \mathbb{R}^+$. When the model is just identified, we have $d_2 = d_1$ and thus

$$P(|t_M| < x) = \Phi(x) - \Phi(-x) - \left[\frac{M}{T} \left\{ \mu_1 x + \frac{1}{2}\mu_2 \left[x^3 + x \right] \right\} + \frac{1}{M^q}\rho_{1,\infty} x \right] \phi(x) + o(\eta_T), \tag{51}$$

where

$$\rho_{1,\infty} = g_q \frac{\mathcal{R}' G_0^{-1} \Omega_0^{(q)} \left(G_0^{-1} \right)' \mathcal{R}}{\mathcal{R}' G_0^{-1} \Omega_0 \left(G_0^{-1} \right)' \mathcal{R}}. \tag{52}$$

This asymptotic expansion coincides with the asymptotic expansion obtained by Velasco and Robinson (2001) and Sun, Phillips and Jin (2007) for Gaussian location models.

4 Optimal Bandwidth Choice for Interval Estimation

This section explores optimal bandwidth choices that minimize the coverage probability error of a confidence interval. Both one-sided and two-sided confidence intervals are considered.

One-sided confidence intervals are examined first. The coverage error of a one-sided confidence interval can be obtained directly from Theorem 2. Without loss of generality, we consider upper one-sided confidence intervals as the qualitative results are the same for lower one-sided confidence intervals. Let $z_\alpha = \Phi^{-1}(1-\alpha)$, then the coverage probability for the one-sided confidence interval $\mathcal{I}_T := (\hat{\beta}_{i,T} - (z_\alpha/\sqrt{T})\hat{\Sigma}_{i,i}^{1/2}, \infty)$ for the i-th component of β is given by

$$\begin{aligned} & P\left(\left(\hat{\Sigma}_{i,i} \right)^{-1/2} \sqrt{T} \left(\hat{\beta}_{i,T} - \beta_{i,0} \right) \leq z_\alpha \right) \\ = \quad & 1 - \alpha + \left(\frac{1}{\sqrt{T}} p_1(z_\alpha) + \frac{M}{T} p_2(z_\alpha) + \frac{1}{M^q} p_3(z_\alpha) \right) \phi(z_\alpha) + o(\eta_T), \end{aligned} \tag{53}$$

where $\hat{\Sigma}_{i,i}$ is the (i,i)-th element of matrix $\hat{\Sigma}$. The term

$$CPE = \left(\frac{1}{\sqrt{T}} p_1(z_\alpha) + \frac{M}{T} p_2(z_\alpha) + \frac{1}{M^q} p_3(z_\alpha) \right) \phi(z_\alpha) \tag{54}$$

provides an approximation to this coverage error.

If the order of magnitude of M lies between $T^{1/(2q)}$ and $T^{1/2}$, e.g. when $M = CT^{\gamma}$ for some $\gamma \in (1/(2q), 1/2)$, then the M^{-q} and M/T terms in (54) are negligible in comparison with the $1/\sqrt{T}$ term. However, if M is of larger order than $T^{1/2}$ or of smaller order than $T^{1/(2q)}$, then the coverage error will be larger than $1/\sqrt{T}$. For example, this increase in the coverage error probability occurs when we use the MSE-optimal bandwidth. In that case, $M \sim CT^{1/(2q+1)}$ and we have

$$CPE = \frac{1}{C^q T^{q/(2q+1)}} p_3(z_\alpha)\phi(z_\alpha) + o\left(\frac{1}{T^{q/(2q+1)}}\right). \tag{55}$$

For the widely used Bartlett kernel, $q = 1$, in which case, the coverage error for the MSE-optimal bandwidth is of order $1/T^{1/3}$. Therefore, for one-sided confidence intervals, the conventional MSE-optimal bandwidth choice is not optimal for minimizing the error in the coverage probability and actually inflates the coverage error by increasing its order of magnitude.

Our analysis suggests that the choice of M is not particularly important, provided that M increases faster than $T^{1/(2q)}$ and more slowly than $T^{1/2}$. For any such M, the coverage error will be dominated by a term of order $1/\sqrt{T}$. Although the expressions given in (36) and (38) may appear to suggest that this term depends on q, the proof in the appendix shows that the dependence actually manifests in higher order terms. In consequence, the $O(1/\sqrt{T})$ term cannot be removed by bandwidth adjustments.

We next consider two-sided confidence intervals, which are quite different from their one-sided counterparts. Let $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$, then the coverage probability for the two-sided confidence interval $\mathcal{I}_T := (\hat{\beta}_{i,T} - (z_{\alpha/2}/\sqrt{T})\hat{\Sigma}_{i,i}^{1/2}, \hat{\beta}_{i,T} + (z_{\alpha/2}/\sqrt{T})\hat{\Sigma}_{i,i}^{1/2})$ is

$$\begin{aligned} & P\left(-z_{\alpha/2} \leq (\hat{\Sigma}_{i,i})^{-1/2}\sqrt{T}(\hat{\beta}_{i,T} - \beta_{i,0}) \leq z_{\alpha/2}\right) \\ = \quad & 1 - \alpha + 2\left(\frac{M}{T} p_2(z_{\alpha/2}) + \frac{1}{M^q} p_3(z_{\alpha/2})\right)\phi(z_{\alpha/2}) + o(\eta_T). \end{aligned} \tag{56}$$

The approximate coverage error is now

$$CPE = 2\left(\frac{M}{T} p_2(z_{\alpha/2}) + \frac{1}{M^q} p_3(z_{\alpha/2})\right)\phi(z_{\alpha/2}). \tag{57}$$

In this expression, the term of order $1/\sqrt{T}$ vanishes, as usual. Minimizing the order of the coverage error is achieved by balancing the $O(M^{-q})$ term and the $O(M/T)$ term in (57).

The form of the optimal bandwidth depends on the signs of $p_2(z_{\alpha/2})$ and $p_3(z_{\alpha/2})$. If $p_2(z_{\alpha/2})$ and $p_3(z_{\alpha/2})$ are opposite in sign, then M can be chosen according to the rule

$$M = \left(\frac{-p_3(z_{\alpha/2})}{p_2(z_{\alpha/2})}\right)^{1/(q+1)} T^{1/(q+1)}, \tag{58}$$

in which case the two terms in (57) cancel and the coverage probability error becomes of smaller order, viz., $o(T^{-q/(q+1)})$. On the other hand, when $p_2(z_{\alpha/2})$ and $p_3(z_{\alpha/2})$ have the same sign, then we can choose M to minimize the absolute coverage error, leading to the rule

$$M = \left(\frac{q p_3(z_{\alpha/2})}{p_2(z_{\alpha/2})} \right)^{1/(q+1)} T^{1/(q+1)}. \tag{59}$$

For convenience, we call the bandwidth in (58) and (59) the CPE-optimal bandwidth. This bandwidth is quite different from the MSE-optimal bandwidth that minimizes the asymptotic mean square error of the HAC estimator. It can be shown that the bias of the HAC estimator is of order $O(M^{-q})$ and the variance is order $O(M/T)$. To minimize the mean squared error of the HAC estimator, we would choose the bandwidth to balance the variance and *squared* bias, resulting in a bandwidth choice of order $M \sim T^{1/(2q+1)}$ (see, for example, Andrews (1991)). However, to minimize the coverage error, we need to balance terms of order M/T and order M^{-q}, i.e., variance is balanced to bias, instead of squared bias. The optimal choice of M is then $O(T^{1/(q+1)})$ instead of $O\left(T^{1/(2q+1)}\right)$.

For this optimal choice of M, the resulting best rate of convergence of the coverage error to zero is $O\left(T^{-q/(q+1)}\right)$. This rate increases with q and can be arbitrarily close to $O\left(T^{-1}\right)$ if q is large and the autocovariances decay at an exponential rate. Examples of kernels with $q > 2$ include the familiar truncated kernel and the flat top kernel proposed by Politis and Romano (1995, 1998). These kernels are not positive semidefinite. However, we shall consider only the commonly-used positive semidefinite kernels for which $q \leq 2$ in this paper and leave the analysis of higher order kernels for future research.

We now turn to the case of just identification or the case where the first step estimator is asymptotically efficient. In these two cases, we can easily see the determinants of the optimal bandwidth. In particular, it follows from equation (50) that the optimal bandwidth is given by

$$M = \begin{cases} \left(-\frac{2\rho_{1,\infty}}{2\mu_1 + \mu_2\left(z_{\alpha/2}^2 + 4d_2 - 4d_1 + 1\right)} \right)^{1/(q+1)} T^{1/(q+1)}, & \rho_{1,\infty} < 0; \\ \left(\frac{2q\rho_{1,\infty}}{2\mu_1 + \mu_2\left(z_{\alpha/2}^2 + 4d_2 - 4d_1 + 1\right)} \right)^{1/(q+1)} T^{1/(q+1)}, & \rho_{1,\infty} > 0. \end{cases} \tag{60}$$

The above analytical expression provides some new insights. First, the optimal bandwidth depends on the kernel function not only through g_q and μ_2 but also through the parameter μ_1. This dependence contrasts with the MSE-optimal bandwidth which does not depend on μ_1. It is well known that the quadratic spectral kernel is the best with respect to the asymptotic truncated MSE in the class of positive definite kernels. This optimality property of the quadratic spectral kernel does not hold with respect to the coverage error for interval estimation. To see this, note that when $\rho_{1,\infty} > 0$ and the optimal bandwidth is used, the absolute coverage error is

$$\left(q^{1/(q+1)} + q^{-q/(q+1)}\right) \left\{ \rho_{1,\infty} g_q \left[\mu_1 + \frac{1}{2}\mu_2 \left(z_{\alpha/2}^2 + 4d_2 - 4d_1 + 1 \right) \right]^q \right\}^{1/(q+1)} T^{-q/(q+1)}. \quad (61)$$

For any given critical value $z_{\alpha/2}$, the optimal kernel in the class of positive-definite kernels with $k(0) = 1$ and $q = 2$ should minimize $g_2 \left(\mu_1 + \frac{1}{2}\mu_2 \left(z_{\alpha/2}^2 + 4d_2 - 4d_1 + 1 \right) \right)^2$. However, the quadratic spectral kernel is designed to minimize $g_2 \mu_2^2$ and may not be optimal any longer. The problem of selecting the optimal kernel in the present case is left for future research. Second, the optimal bandwidth depends on the design matrix G_0 and the coefficient considered via the restriction vector $\mathcal{R}$. This is again in contrast with the MSE-optimal bandwidth which is of course generally independent of both G_0 and $\mathcal{R}$. The MSE-optimal bandwidth does not depend on G_0 and $\mathcal{R}$ because Andrews (1991) focuses on the asymptotic truncated MSE of $\hat{\Omega}_T$ rather than of the HAC standard error of the regression coefficients. Andrews (1991) justified this approach by noting that the rate of convergence of G_T is faster than that of $\hat{\Omega}_T$. The faster convergence rate guarantees that the MSE of the HAC standard error is dominated by the asymptotic bias and variance of $\hat{\Omega}_T$ but it does not rule out that the MSE of the HAC standard error may depend on G_0, the limit of the design matrix, and $\mathcal{R}$, the restriction matrix and coefficient considered. Third, the optimal bandwidth depends on the relative bias of the HAC standard error. The quantity $\mathcal{R}' G_0^{-1} \Omega_0 \left(G_0^{-1} \right)' \mathcal{R}$, which is the denominator of $\rho_{1,\infty}$, is the true variance of the $\mathcal{R}' \sqrt{T} (\hat{\beta}_T - \beta_0)$ while $\mathcal{R}' G_0^{-1} \Omega_0^{(q)} \left(G_0^{-1} \right)' \mathcal{R}$, the numerator of $\rho_{1,\infty}$, can be regarded as the bias of the HAC estimator. Therefore, $\rho_{1,\infty}$ is proportional to the percentage bias. The higher the (absolute) percentage bias is, the larger is the bandwidth. Finally, the optimal bandwidth depends on the confidence level through the critical value z_α. The critical value increases as the confidence level increases. As a result, and with all else being equal, the higher the confidence level, the smaller the optimal bandwidth.

When $\rho_{1,\infty} > 0$ and the optimal bandwidth is used, we have

$$\begin{aligned} P(|t_M| &< x) = \Phi(x) - \Phi(-x) \\ &\quad -2\frac{q+1}{q} \left[\mu_1 x + \frac{1}{2}\mu_2 \left(x^3 + x\left(4d_2 - 4d_1 + 1\right) \right) \right] \phi(x) \frac{M}{T} + o(\eta_T) \end{aligned} \quad (62)$$

To reduce the coverage error of the two-sided confidence interval, we can remove the $O(M/T)$ term using a Cornish-Fisher type expansion. Let $z_{\alpha/2}$ be the critical value from the standard normal such that $\Phi\left(z_{\alpha/2}\right) = 1 - \alpha/2$ and

$$z_{\alpha/2}^* = z_{\alpha/2} + \frac{q+1}{q} \left\{ \frac{1}{2}\mu_1 z_{\alpha/2} + \frac{1}{4}\mu_2 \left[z_{\alpha/2}^3 + z_{\alpha/2}\left(4d_2 - 4d_1 + 1\right) \right] \right\} \frac{M}{T}. \quad (63)$$

Then

$$P(|t_M| > z_{\alpha/2}^*) = \alpha + o\left(\eta_T\right). \quad (64)$$

We call $z^*_{\alpha/2}$ the higher order corrected critical value. With this higher order correction, the coverage error is of order $o(T^{-q/(q+1)})$ regardless of whether $\rho_{1,\infty} > 0$ or not. The higher order correction is similar to that obtained by Sun, Phillips and Jin (2007). For a Gaussian location model, they established a higher order correction based on $p_2(x)$, the term that captures the asymptotic variance of the HAC estimator.

For illustrative purposes, we compute the optimal bandwidth for the Bartlett, Parzen, and Quadratic Spectral (QS) kernels for the Gaussian location model $y_t = \beta_0 + u_t$ where u_t follows an AR(1) process with autoregressive parameter ρ. Let

$$\alpha(1) = \frac{4\rho^2}{(1-\rho^2)^2} \text{ and } \alpha(2) = \frac{4\rho^2}{(1-\rho)^4}. \tag{65}$$

Standard calculations show that the MSE-optimal bandwidth is given by

$$\begin{array}{ll} \text{Bartlett Kernel:} & M = 1.1447\,[\alpha(1)]^{1/3}\,T^{1/3} \\ \text{Parzen Kernel:} & M = 2.6614\,[\alpha(2)]^{1/5}\,T^{1/5} \\ \text{Quadratic Spectral Kernel:} & M = 1.3221\,[\alpha(2)]^{1/5}\,T^{1/5} \end{array} \tag{66}$$

whereas the CPE-optimal bandwidth is given by

$$M = \begin{cases} \left(\dfrac{2g_q}{2\mu_1+\mu_2\left(z^2_{\alpha/2}+4d_2-4d_1+1\right)}\sqrt{|\alpha(q)|}\right)^{1/(q+1)} T^{1/(q+1)}, & \rho < 0 \\ \left(\dfrac{2qg_q}{2\mu_1+\mu_2\left(z^2_{\alpha/2}+4d_2-4d_1+1\right)}\sqrt{|\alpha(q)|}\right)^{1/(q+1)} T^{1/(q+1)}, & \rho > 0 \end{cases} \tag{67}$$

where the constants μ_1, μ_2, g_q and q are given in Table I below.

Table I. Values of the Constants for Different Kernels

	μ_1	μ_2	g_q	q
Bartlett Kernel	1.0000	0.6667	1.0000	1
Parzen Kernel	0.7500	0.5393	6.0000	2
QS Kernel	1.2500	1.0000	1.4212	2

Table II tabulates M under different criteria for the Bartlett and Parzen kernels. To save space, we omit the result for the QS kernel. For the CPE-optimal bandwidth, we consider two confidence levels, i.e. 90% and 95%. Some features of note in these calculations are as follows. First, as predicted by asymptotic theory, the CPE-optimal bandwidth is in general larger than the MSE-optimal bandwidth especially when T is large. Second, the CPE-optimal bandwidth for the 90% confidence interval is always larger than that for the 95% confidence interval. The difference is not very large, especially when the autoregressive parameter is not very large. Third, compared with the Bartlett kernel, the Parzen kernel requires larger bandwidths regardless of the criterion used. Of course, the optimal bandwidth given in Table II is not feasible. For this reason, calculations based on estimates of the unknown parameters are considered in the next section.

Table II. Asymptotically Optimal Bandwidth under Different Criteria
When $u_t = \rho u_{t-1} + \varepsilon_t$, $\varepsilon_t \sim iid\ N(0,1)$

	ρ -0.9	-0.5	-0.1	0.1	0.5	0.9	-0.9	-0.5	-0.1	0.1	0.5	0.9
T												
	Bartlett Kernel, MSE						Parzen Kernel, MSE					
128	25.8	7.0	2.0	2.0	7.0	25.8	5.3	5.1	3.4	4.0	12.2	56.1
256	32.5	8.8	2.5	2.5	8.8	32.5	6.1	5.8	3.9	4.6	14.0	64.4
512	41.0	11.1	3.2	3.2	11.1	41.0	7.0	6.7	4.5	5.3	16.1	74.0
1024	51.7	14.0	4.0	4.0	14.0	51.7	8.1	7.7	5.2	6.1	18.5	85.0
	Bartlett Kernel, CPE, 90%						Parzen Kernel, CPE, 90%					
128	23.3	8.7	3.4	3.4	8.7	23.3	6.0	5.8	4.2	6.0	15.2	54.1
256	32.9	12.4	4.8	4.8	12.4	32.9	7.6	7.3	5.3	7.6	19.2	68.1
512	46.6	17.5	6.8	6.8	17.5	46.6	9.6	9.2	6.6	9.5	24.1	85.8
1024	65.9	24.7	9.6	9.6	24.7	65.9	12.1	11.6	8.3	12.0	30.4	108.1
	Bartlett Kernel, CPE, 95%						Parzen Kernel, CPE, 95%					
128	21.5	8.1	3.1	3.1	8.1	21.5	5.7	5.5	4.0	5.7	14.4	51.2
256	30.5	11.4	4.4	4.4	11.4	30.5	7.2	6.9	5.0	7.2	18.1	64.6
512	43.1	16.2	6.3	6.3	16.2	43.1	9.1	8.7	6.3	9.0	22.9	81.3
1024	60.9	22.9	8.9	8.9	22.9	60.9	11.4	11.0	7.9	11.4	28.8	102.5

5 An Automatic Data-Driven Bandwidth

The optimal bandwidth in (58) and (59) involves unknown parameters $\rho_{1,\infty}$, $\rho_{2,\infty}$, $\kappa_{2,\infty}$, $\kappa_{4,\infty}$ which could be estimated nonparametrically (e.g. Newey and West (1994)) or by a standard plug-in procedure based on a simple model like a VAR (e.g. Andrews (1991)). Both methods achieve a valid order of magnitude and the procedure is analogous to conventional data-driven methods for HAC estimation.

We focus this discussion on the plug-in procedure, which involves the following steps. First, we estimate the model using the OLS or IV estimator, compute the residuals, and construct the sequence $\{\hat{v}_t\}$. Second, we specify a multivariate approximating parametric model and fit the model to $\{\hat{v}_t\}$ by standard methods. Third, we treat the fitted model as if it were the true model for the process $\{v_t\}$ and compute $\rho_{1,\infty}, \rho_{2,\infty}$, $\kappa_{2,\infty}$ and $\kappa_{4,\infty}$ as functions of the parameters of the parametric model. Plugging these estimates of $\rho_{1,\infty}$, $\rho_{2,\infty}$, $\kappa_{2,\infty}$ and $\kappa_{4,\infty}$ into (58) or (59) gives the automatic bandwidth $\hat{M}$.

In this paper, we assume that the approximating parametric model satisfies the assumptions of Proposition 3. As in the case of MSE-optimal bandwidth choice, the automatic bandwidth considered here deviates from the finite sample optimal one due to the error introduced by estimation, the use of approximating parametric models, and the approximation inherent in the asymptotic formula employed. It is hoped that in practical work the deviation is not large so that the resulting confidence interval still has a small coverage error. Some simulation evidence reported in the next section supports this argument.

Under the model given in Proposition 3, the CPE-optimal bandwidth depends only on $\Omega_0, \Omega_0^{(q)}, G_0, V_0$ and $\mathcal{R}$. In other words, we can write $M = M(\Omega_0, \Omega_0^{(q)}, G_0, V_0, \mathcal{R})$. Since G_0 and V_0 can be consistently estimated by G_T and V_T, we only need to estimate Ω_0 and $\Omega_0^{(q)}$. Suppose we use a VAR(1) as the approximating parametric model for v_t. Let $\hat{A}$ be the estimated parameter matrix and $\hat{\Sigma}$ be the estimated innovation covariance matrix, then the plug-in estimates of Ω_0 and $\Omega_0^{(q)}$ are

$$\hat{\Omega}_0 = (I_{d_2} - \hat{A})^{-1}\hat{\Sigma}(I_{d_2} - \hat{A}')^{-1}, \tag{68}$$

$$\begin{aligned} \hat{\Omega}_0^{(2)} &= (I_{d_2} - \hat{A})^{-3}\left(\hat{A}\hat{\Sigma} + \hat{A}^2\hat{\Sigma}\hat{A}' + \hat{A}^2\hat{\Sigma} - 6\hat{A}\hat{\Sigma}\hat{A}'\right. \\ &\quad \left. + \hat{\Sigma}(\hat{A}')^2 + \hat{A}\hat{\Sigma}(\hat{A}')^2 + \hat{\Sigma}\hat{A}'\right)(I_{d_2} - \hat{A}')^{-3}, \end{aligned} \tag{69}$$

$$\hat{\Omega}_0^{(1)} = \hat{H} + \hat{H}' \text{ with } \hat{H} = (I_{d_2} - \hat{A})^{-2}\hat{A}\Sigma_{j=0}^{\infty}\hat{A}^j\hat{\Sigma}(\hat{A}')^j. \tag{70}$$

For the plug-in estimates under a general VAR(p) model, the reader is referred to Andrews (1991) for the corresponding formulae. Given consistent estimates of G_0 and V_0 and plug-in estimates of Ω_0 and $\Omega_0^{(q)}$, the data-driven automatic bandwidth can be computed as

$$\hat{M} = \hat{M}(\hat{\Omega}_0, \hat{\Omega}_0^{(q)}, G_T, V_T, \mathcal{R}). \tag{71}$$

When the model is just identified or the initial weighting matrix is consistent for Ω_0^{-1}, the automatic bandwidth is given by

$$\hat{M} = \begin{cases} \left(-\frac{2\hat{\rho}_{1,\infty}}{2\mu_1 + \mu_2\left(z_{\alpha/2}^2 + 4d_2 - 4d_1 + 1\right)}\right)^{1/(q+1)} T^{1/(q+1)}, & \hat{\rho}_{1,\infty} < 0 \\ \left(\frac{2q\hat{\rho}_{1,\infty}}{2\mu_1 + \mu_2\left(z_{\alpha/2}^2 + 4d_2 - 4d_1 + 1\right)}\right)^{1/(q+1)} T^{1/(q+1)}, & \hat{\rho}_{1,\infty} > 0 \end{cases} \tag{72}$$

where

$$\hat{\rho}_{1,\infty} = g_q \frac{\mathcal{R}' \left(G_T' \hat{\Omega}_0 G_T \right)^{-1} G_T' \hat{\Omega}_0^{-1} \hat{\Omega}_0^{(q)} \hat{\Omega}_0^{-1} G_T \left(G_T' \hat{\Omega}_0 G_T \right)^{-1} \mathcal{R}}{\mathcal{R}' G_T^{-1} \hat{\Omega}_0 \left(G_T^{-1} \right)' \mathcal{R}}. \tag{73}$$

It should be pointed out that the computational cost involved in this automatic bandwidth is the same as that of the conventional plug-in bandwidth.

6 Simulation Evidence

This section provides some simulation evidence on the finite sample performance of these new procedures for confidence interval construction. The new confidence intervals are based on the CEP-optimal bandwidth and critical values that are possibly high-order corrected.

We consider several linear regression models in the experimental design of the form

$$y_t = \beta_1 + x_t \beta_2 + u_t, t = 1, 2, ..., T \tag{74}$$

each with a scalar regressor. In the first model, AR(1)-HOM, the regressor and errors are independent AR(1) processes with the same AR parameter ρ :

$$x_t = \rho x_{t-1} + \varepsilon_{x,t}, \quad u_t = \rho u_{t-1} + \varepsilon_{u,t} \tag{75}$$

where $\varepsilon_{x,t} \sim iidN(0,1)$, $\varepsilon_{u,t} \sim iidN(0,1)$ and $\{\varepsilon_{x,t}\}_{t=1}^T$ are independent of $\{\varepsilon_{u,t}\}_{t=1}^T$. The values considered for the AR(1) parameter ρ are $0.1, 0.3, 0.5, 0.7, 0.9, 0.95, -0.3$ and -0.5. In the second model, AR(1)-HET, we introduce multiplicative heteroscedasticity to the errors of the AR(1)-HOM model, leading to

$$x_t = \rho x_{t-1} + \varepsilon_{x,t}, \tilde{u}_t = \rho \tilde{u}_{t-1} + \varepsilon_{u,t}, \quad u_t = |x_t| \, \tilde{u}_t. \tag{76}$$

The same values of ρ are considered as in the AR(1)-HOM model. In the third model, MA(1)-HOM, the regressor and errors are independent MA(1) processes with the same MA parameter ψ :

$$x_t = \varepsilon_{x,t} + \psi \varepsilon_{x,t-1}, u_t = \varepsilon_{u,t} + \psi \varepsilon_{u,t-1}, \tag{77}$$

where $\varepsilon_{x,t} \sim iidN(0,1)$, $\varepsilon_{u,t} \sim iidN(0,1)$ and $\{\varepsilon_{x,t}\}_{t=1}^T$ are independent of $\{\varepsilon_{u,t}\}_{t=1}^T$. The values of ψ are taken to be $0.1, 0.3, 0.5, 0.7, 0.90, 0.99, -0.3$, and -0.7. These data generating processes are similar to those used in Andrews (1991).

We focus on constructing 90% and 95% two-sided confidence intervals for β_2. Since the coverage probabilities are invariant with respect to the regression coefficients β_1 and β_2, we set $\beta_1 = \beta_2 = 0$ and do so without losing generality. To compute the HAC standard error, we employ the three commonly-used positive definite kernels, i.e. Bartlett, Parzen and Quadratic Spectral kernels.

For comparative purposes, we use both MSE-optimal and CPE-optimal bandwidth selection rules. In both cases, the approximating parametric model is a VAR(1). The CPE-optimal bandwidth is given in equations (72) and (73). As in Andrews and Monahan (1992), we adjust the estimated VAR coefficient matrix $\hat{A}_{LS}$ before constructing $\hat{\Omega}_0^{(q)}$ and $\hat{\Omega}_0$. The adjustment is based on the singular value decomposition: $\hat{A}_{LS} = \hat{B}\hat{\Delta}_{LS}\hat{C}'$ where $\hat{B}$ and $\hat{C}$ are orthogonal matrices and $\hat{\Delta}_{LS}$ is a diagonal matrix. Let $\hat{\Delta}$ be the diagonal matrix constructed from Δ_{LS} by replacing any element of Δ_{LS} that exceeds 0.97 by 0.97 and any element that is less than -0.97 by -0.97. Then $\hat{A} = \hat{B}\hat{\Delta}\hat{C}'$. Given the adjusted estimate $\hat{A}$, $\hat{\Omega}_0^{(q)}$ and $\hat{\Omega}_0$ are computed using the formulae in the previous section. For completeness, we give the MSE-optimal bandwidth of Andrews (1991) below:

$$\begin{array}{ll} \text{Bartlett kernel:} & M = 1.1447\left[\hat{\alpha}\left(1\right)T\right]^{1/3} \\ \text{Parzen kernel:} & M = 2.6614\left[\hat{\alpha}\left(2\right)T\right]^{1/5} \\ \text{Quadratic Spectral kernel:} & M = 1.3221\left[\hat{\alpha}\left(2\right)T\right]^{1/5} \end{array} \tag{78}$$

where

$$\hat{\alpha}\left(q\right) = \frac{2vec(\hat{\Omega}_0^{(q)})'vec(\hat{\Omega}_0^{(q)})}{tr\left[\left(I_{d_2^2} + K_{d_2,d_2}\right)\left(\hat{\Omega}_0 \otimes \hat{\Omega}_0\right)\right]}. \tag{79}$$

It is well known that prewhitening can be used to reduce the finite sample bias of the HAC standard error estimator. In our simulation experiments, we combine prewhitening with both the conventional and new procedures. In the former case, Andrews and Monahan (1992) have established the consistency of the prewhitened HAC estimator and show via Monte Carlo experiments that prewhitening is effective in improving confidence interval coverage probabilities. In the Monte Carlo experiments here, we use VAR(1) prewhitening as in Andrews and Monahan (1992). The MSE-optimal bandwidth is based on the prewhitened error process $\hat{v}_t^*$ defined by

$$\hat{v}_t^* = \hat{v}_t - \hat{A}\hat{v}_{t-1}, \tag{80}$$

where $\hat{A}$ is the OLS estimates obtained from regressing $\hat{v}_t$ on $\hat{v}_{t-1}$. To compute the data-driven plug-in bandwidth, we fit another VAR(1) model to the prewhitened error process $\hat{v}_t^*$. Univariate AR(1) models have also been employed as approximating parametric models for each element of v_t^*, but the qualitative results are similar. Therefore, we focus on the VAR(1) plug-in estimate. Let $\hat{A}^*$ be the OLS estimate based on the following regression

$$\hat{v}_t^* = \hat{A}^*\hat{v}_{t-1} + error, \tag{81}$$

and $\hat{\Omega}_0^*$ and $\hat{\Omega}_0^{*(q)}$ be defined as in equations (68)–(70) but with $\hat{A}$ replaced by $\hat{A}^*$. Then the automatic MSE-optimal bandwidth is given in (78) but with $\hat{\Omega}_0$ and $\hat{\Omega}_0^{(q)}$ replaced by $\hat{\Omega}_0^*$ and $\hat{\Omega}_0^{*(q)}$. We note in passing that singular value adjustment has been made to both $\hat{A}$ and $\hat{A}^*$ so that the fitted parametric models are stationary.

Prewhitening can be combined with the new procedure in the same manner. To make a fair comparison, we employ a VAR(1) prewhitened HAC estimator as before. The point of departure is that the data-driven bandwidth is now based on the CPE criterion proposed above. Let

$$\hat{\Omega}_0 = \left(I - \hat{A}\right)^{-1} \hat{\Omega}_0^* \left(I - \hat{A}'\right)^{-1}, \ \hat{\Omega}_0^{(q)} = \left(I - \hat{A}\right)^{-1} \hat{\Omega}_0^{*(q)} \left(I - \hat{A}'\right)^{-1}. \tag{82}$$

Then the automatic CPE-optimal bandwidth is given in (72) and (73). This prewhitened bandwidth selection rule can be justified on the basis of the $\sqrt{T}$-consistency of $\hat{A}$. Due to the faster rate of convergence, the Edgeworth expansion of the two-sided probability will be not be affected by the estimation uncertainty of $\hat{A}$. Nevertheless, the estimation uncertainty may factor in the Edgeworth expansion of the one-sided probability and this consideration is left for future research as we concentrate on two-sided confidence intervals here.

For each parameter combination and HAC estimator, we construct two-sided symmetric confidence intervals of the form

$$\left[-z_{cv}\left(\left[\left(G_T\hat{\Omega}_0^{-1}G_T\right)^{-1}\right]_{22}\right)^{1/2}, \ z_{cv}\left(\left[\left(G_T\hat{\Omega}_0^{-1}G_T\right)^{-1}\right]_{22}\right)^{1/2}\right] \tag{83}$$

where z_{cv} is the critical value and $[\cdot]_{aa}$ stands for the (a,a)'th element of $[\cdot]$. For the conventional HAC procedure, we use critical values from the standard normal distribution, viz. 1.645 for the 90% confidence interval and 1.96 for the 95% confidence interval. For the new HAC procedure, we use the standard critical values if $\hat{\rho}_{1,\infty} \leq 0$ and the high-order corrected critical values given in (63) if $\hat{\rho}_{1,\infty} > 0$. The calculations reported below are for three sample sizes (100, 200 and 400) and use 10,000 simulation replications. For each scenario, we calculate the empirical coverage probability, i.e. the percentage of the replications for which the confidence interval contains the true parameter value.

Tables III-V provide a comparison of the two bandwidth choice rules when no prewhitening is used and $T = 100$ in the various models. The tables show that the confidence interval proposed in this paper has more accurate coverage than the conventional confidence interval. This is the case for both 90% and 95% confidence levels regardless of the kernel employed and the model considered here. The advantage of the new confidence interval becomes more apparent as temporal dependence in the regressor and the errors becomes stronger. Simulation results not reported here show that both the new bandwidth choice rule and the high order correction contribute to the improvement in coverage accuracy.

As in Andrews (1991), we find that the QS-based confidence intervals are fairly consistently the best among the conventional confidence intervals. The QS-based confidence intervals outperform other conventional confidence intervals in 40 out of 48 scenarios in Tables III-IV. However, the QS kernel does not deliver superior performance for the new confidence intervals. As a matter of fact, QS-based confidence intervals are the best in only 10 out the 48 scenarios. In these 10 scenarios, the QS kernel is either the same as or slightly better than the Parzen kernel. In contrast, the Bartlett kernel and Parzen kernel are very competitive with each one outperforming the other one in about half of the scenarios. More specifically, when the regressor and the errors are fairly persistent, the Bartlett kernel delivers confidence intervals with the best coverage. When the regressor and errors are less persistent, the Parzen kernel delivers confidence intervals with the best coverage.

These qualitative observations remain valid for sample sizes 200 and 400. Table VI presents the results for selected parameter combinations for sample size 400 when prewhitening is not used. As expected, the increase in sample size from 100 to 400 improves the performance of all confidence intervals. In all cases investigated, including those not reported here, the new procedure outperforms the conventional procedure.

Table VII presents selected results for sample size 100 when prewhitening is used. It is apparent that prewhitening is very effective in improving coverage probabilities for the conventional confidence interval. This is consistent with the findings in Andrews and Monahan (1992). However, the effectiveness of prewhitening is reduced for the new procedure. This is not surprising as the new procedure is expected to work the best when there is a substantial bias in the HAC estimator. Since prewhitening has already achieved considerable bias reduction, the room for further bias reduction is reduced. It is encouraging to note that, even with prewhitening, the new confidence intervals have consistently better performance than the conventional ones, although the margin is not very large.

7 Optimal Bandwidth: An Alternative Criterion

Previous sections considered the coverage accuracy of confidence intervals. Another performance criterion of an interval estimator is its length. In general, coverage probability and interval length tend to work against each other. Accordingly, it may be desirable to construct a loss function that takes both coverage probability and length into account and make a bandwidth choice to optimize this loss function. The challenge is to construct a satisfactory loss function that does not result in the paradoxical behavior described in Casella, Hwang and Robert (1993). Since there is no satisfactory solution to this problem, we use the probability of covering false values (or false coverage probability) as an indirect measure of the length of the confidence interval.

We then seek to minimize the probability of false coverage subject to the constraint that the probability of true coverage is bounded below by the nominal coverage probability.

Since the bandwidth is not very important for one sided confidence intervals, we focus the present discussion on two-sided confidence intervals. Given the confidence interval

$$\mathcal{I}_T := (\hat{\beta}_{i,T} - (z/\sqrt{T})\hat{\Sigma}_{i,i}^{1/2}, \hat{\beta}_{i,T} + (z/\sqrt{T})\hat{\Sigma}_{i,i}^{1/2}) \tag{84}$$

for the i-th element of β, the probability of covering the true value $\beta_{i,0}$ is $P_{\beta_0}\left\{\beta_{i,0} \in \mathcal{I}_T\right\}$ where P_{β_0} is the probability measure under $\beta = \beta_0$. To approximate the probability of false coverage, we consider local alternatives of the form $H_\delta : \beta_\delta = \left(\beta_{1,\delta}, ..., \beta_{d_1,\delta}\right)$ where

$$\beta_{i,\delta} = \beta_{i,0} + \delta/\sqrt{T} \times \left(\hat{\Sigma}_{i,i}\right)^{-1/2} \text{ for some } \delta \text{ and } \beta_{j,\delta} = \beta_{j,0} \text{ for } j \neq i. \tag{85}$$

Then, for each $\delta \neq 0$, the probability of false coverage is $P_{\beta_0}\left\{\beta_{i,\delta} \in \mathcal{I}_T\right\}$. Obviously, we can average this probability over a prior distribution on δ and obtain the average false coverage $E_\delta\left[P_{\beta_0}\left\{\beta_{i,\delta} \in \mathcal{I}_T\right\}\right]$, where E_δ is the expectation operator under some prior distribution for δ.

Our alternative bandwidth choice rule involves minimizing the average false coverage $E_\delta\left[P_{\beta_0}\left\{\beta_{i,\delta} \in \mathcal{I}_T\right\}\right]$ after controlling for the true coverage $P_{\beta_0}\left\{\beta_{i,0} \in \mathcal{I}_T\right\}$. Mathematically, we proceed to solve the following minimization problem:

$$\min_{M,z} E_\delta\left[P_{\beta_0}\left\{\beta_{i,\delta} \in \mathcal{I}_T\right\}\right] \text{ s.t. } P_{\beta_0}\left\{\beta_{i,0} \in \mathcal{I}_T\right\} \geq 1-\alpha \tag{86}$$

where $1-\alpha$ is the nominal coverage level of the confidence interval $\mathcal{I}_T$. Note that we choose the bandwidth and the critical value simultaneously. To the first order z is $z_{\alpha/2}$ but a higher order adjustment is possible, as described below. Confidence intervals that minimize the probability of false coverage are called Neyman shortest (Neyman (1937, page 371)). The fact that there is a length connotation to this name is somewhat justified by a theorem in Pratt (1961). Under some conditions, Pratt (1961) showed that the expected length of a confidence interval is equal to an integral of the probabilities of false coverage.

Our alternative approach to bandwidth choice requires improved measurements of the two coverage probabilities: the probability of true coverage and the probability of false coverage. Using the Edgeworth expansions established in this paper, we can obtain asymptotic approximations to the two coverage probabilities. The probability of true coverage satisfies

$$P_{\beta_0}\left\{\beta_{i,0} \in \mathcal{I}_T\right\} = \Phi_T(z) - \Phi_T(-z) + o(\eta_T), \tag{87}$$

while the probability of false coverage satisfies

$$P_{\beta_0}\left\{\beta_{i,\delta} \in \mathcal{I}_T\right\} = \Phi_T(z-\delta) - \Phi_T(-z-\delta) + o(\eta_T). \tag{88}$$

Hence, up to small order $o(\eta_T)$, the minimization problem reduces to

$$\min_{M,z} E_\delta \left[\Phi_T(z-\delta) - \Phi_T(-z-\delta)\right], \text{ s.t. } \Phi_T(z) - \Phi_T(-z) \geq 1-\alpha. \tag{89}$$

To gain further insight into this approach we may take an explicit prior distribution for δ of the form

$$\delta \sim \frac{1}{2}N(\mu, \omega^2) + \frac{1}{2}N(-\mu, \omega^2), \tag{90}$$

so that δ is distributed as a normal mixture. Other prior weights may be used, but the normal mixture is convenient because it leads to explicit analytic expressions and is rich enough to include certain uni-modal distribution, bimodal distributions, and discrete distribution as special cases. Let

$$\varphi_k(x,\omega) = \frac{1}{(1+\omega^2)^k\sqrt{1+\omega^2}} x^k \left[\phi\left(\frac{x}{\sqrt{1+\omega^2}}\right)\right], \quad k=0,1,2,3. \tag{91}$$

Then we can show that

$$\begin{aligned} & E_\delta \left[\Phi_T(z-\delta) - \Phi_T(-z-\delta)\right] \\ = {} & \Phi\left(\frac{z+\mu}{\sqrt{1+\omega^2}}\right) - \Phi\left(\frac{-z+\mu}{\sqrt{1+\omega^2}}\right) + 2\left[\frac{M}{T}p_2(z,\mu,\omega^2) + \frac{1}{M^q}p_3(z,\mu,\omega^2)\right], \end{aligned} \tag{92}$$

where

$$\begin{aligned} p_2\left(z,\mu,\omega^2\right) = {} & -\frac{1}{2}\left(\rho_{2,\infty} + \kappa_{2,\infty}\right) 0.5\left[\varphi_1(z-\mu,\omega) + \varphi_1(z+\mu,\omega)\right] \\ & -\frac{1}{24}\left(\kappa_{4,\infty} - 6\kappa_{2,\infty}\right) 0.5\left[\varphi_3(z-\mu,\omega) + \varphi_3(z+\mu,\omega)\right] \\ & +\frac{1}{24}\left(\kappa_{4,\infty} - 6\kappa_{2,\infty}\right)\frac{1.5}{1+\omega^2}\left[\varphi_1(z-\mu,\omega) + \varphi_1(z+\mu,\omega)\right], \end{aligned} \tag{93}$$

and

$$p_3\left(z,\mu,\omega^2\right) = -\frac{1}{2}\rho_{1,\infty} 0.5\left[\varphi_1(z-\mu,\omega) + \varphi_1(z+\mu,\omega)\right]. \tag{94}$$

For later use, we make a slight abuse of notation and write $p_j(z,0,0) = p_j(z)$ for $p_j(z)$ defined in theorem 2.

Let

$$z = z_{\alpha/2} + \frac{M}{T}c_1 + \frac{1}{M^q}c_2 \tag{95}$$

be the high-order corrected critical value, then

$$\begin{aligned} & E_\delta\left[\Phi(z-\delta) - \Phi(-z-\delta)\right] = \Phi\left(\frac{z_{\alpha/2}+\mu}{\sqrt{1+\omega^2}}\right) - \Phi\left(\frac{-z_{\alpha/2}+\mu}{\sqrt{1+\omega^2}}\right) \\ & + \left[\frac{M}{T}c_1 + \frac{1}{M^q}c_2\right]\left[\varphi_0\left(z_{\alpha/2}+\mu,\omega\right) + \varphi_0\left(-z_{\alpha/2}+\mu,\omega\right)\right] + o(\frac{M}{T}) + o(\frac{1}{M^q}). \end{aligned} \tag{96}$$

After dropping some constants and smaller order terms, the minimization problem is approximately

$$\min_{c_1,c_2,M}\left(\frac{M}{T}c_1+\frac{1}{M^q}c_2\right)\frac{1}{2}\left[\varphi_0\left(z_{\alpha/2}+\mu,\omega\right)+\varphi_0\left(-z_{\alpha/2}+\mu,\omega\right)\right]$$
$$+\left[\frac{M}{T}p_2(z_{\alpha/2},\mu,\omega^2)+\frac{1}{M^q}p_3\left(z_{\alpha/2},\mu,\omega^2\right)\right]\Bigg\}, \tag{97}$$

subject to

$$\left[\frac{M}{T}c_1+\frac{1}{M^q}c_2\right]\phi\left(z_{\alpha/2}\right)+\frac{M}{T}p_2(z_{\alpha/2},0,0)+\frac{1}{M^q}p_3\left(z_{\alpha/2},0,0\right)\geq 0. \tag{98}$$

Substituting the constraint into the objective function, we obtain the unconstrained minimization problem:

$$\min_{M\geq 0}\left(\frac{M}{T}p_2^*+\frac{1}{M^q}p_3^*\right) \tag{99}$$

where

$$p_j^*=p_j(z_{\alpha/2},0,0)-\kappa p_j(z_{\alpha/2},\mu,\omega^2), j=2,3, \tag{100}$$

and

$$\kappa=\frac{\left[\varphi_0\left(z_{\alpha/2}+\mu,\omega\right)+\varphi_0\left(-z_{\alpha/2}+\mu,\omega\right)\right]}{2\phi\left(z_{\alpha/2}\right)}. \tag{101}$$

Alternatively, we can let

$$c_1=-p_2(z_{\alpha/2},0,0)\text{ and }c_2=-p_3\left(z_{\alpha/2},0,0\right) \tag{102}$$

so that the constraint (98) is satisfied. This argument leads to exactly the same unconstrained minimization problem. The optimal bandwidth is now given by

$$M^*=\begin{cases}\left(\frac{qp_3^*}{p_2^*}\right)^{\frac{1}{q+1}}T^{\frac{1}{q+1}}, & \text{if } p_2^*>0\text{ and }p_3^*>0,\\ T/\log T, & \text{if } p_2^*<0\text{ and }p_3^*>0,\\ T/\log T\text{ or }\log T & \text{if } p_2^*<0\text{ and }p_3^*<0,\\ \log T, & \text{if } p_2^*>0\text{ and }p_3^*<0.\end{cases} \tag{103}$$

When $p_2^*>0$ and $p_3^*>0$, this optimal bandwidth choice rule is similar to what was obtained earlier in section 4. In particular, when $\kappa=0$, $p_j^*=p_j(z_{\alpha/2},0,0)=p_j(z_{\alpha/2})$ and the two bandwidth choice rules coincide. It is easy to see that $\kappa\to 0$ as $\mu\to\infty$ for any given finite ω^2. Intuitively, when the false value is very far away from the true value, the probability of false coverage is very small and becomes relatively unimportant. In this case, we choose the bandwidth just to maximize the probability of true coverage. This is asymptotically equivalent to maximizing the absolute coverage error when the probability of true coverage is smaller than the nominal coverage probability $1 - \alpha$.

When either p_2^* or p_3^* is negative, the optimal bandwidth formula is nonstandard. When $p_3^* < 0$, we can choose M to be as small as possible in order to maximize $|M^{-q} p_3^*|$. Similarly, when $p_2^* < 0$, we can choose M to maximize $|(M/T)\, p_2^*|$. Since the asymptotic expansion is obtained under the assumption that $M \to \infty$ and $M/T \to 0$, the choice of M is required to be compatible with these two rate conditions. These considerations lead to a choice of M of the form given in (103).

To implement this optimal bandwidth choice rule, we can follow exactly the same procedure as in section 5. We only need to specify the additional parameters μ and ω^2. The selection of μ may reflect a value of scientific interest or economic significance, while the selection of ω^2 can reflect the uncertainly about this value. In the absence of such a value, we recommend using the default values $\mu = 3$ and $\omega^2 = 1$. Such a choice will lead to confidence intervals that avoid covering false values three standard deviations away from the true value.

8 Conclusion

Automatic bandwidth choice is a long-standing practical issue in time series modeling when the autocorrelation is of unknown form. Existing automated methods all rely on early ideas from the time series literature which are based on minimizing the asymptotic mean square error of the long run variance (spectral) estimator, a criterion that is not directed at confidence interval construction. In constructing confidence intervals, the primary concern is often the coverage accuracy. This paper develops for the first time a theory of optimal bandwidth choice that optimizes the coverage accuracy of interval estimators. We show that optimal bandwidth selection for semiparametric interval estimation of the regression coefficients is possible and leads to results that differ from optimal bandwidth choices based on point estimation of the long-run variance. Semiparametric interval estimation along these lines actually undersmooths the long-run variance estimate to reduce bias and allows for greater variance in long-run variance estimation as it is manifested in the t-statistic by means of higher order adjustments to the nominal asymptotic critical values. A plug-in rule for the new optimal bandwidth choice is suggested and finite sample performances of this choice and the new confidence intervals are explored via Monte Carlo experiments. Overall, the results are encouraging for this new approach.

The theory developed in the paper suggests further areas of research. Our primary approach focuses on interval estimation for one model parameter or a linear combination of parameters. The basic ideas and methods explored here can be used to tackle the bandwidth choice problem for constructing multidimensional confidence regions. Relatedly, the methods can be used to select the optimal bandwidth that minimizes the size distortion of the over-identification test. In addition, we propose a secondary approach to bandwidth selection that takes false coverage probability into account. It seems desirable to further explore this approach and its finite sample performance.

Table III. Empirical Confidence Levels of Nominal 90% and 95% Confidence Intervals for AR(1)-HOM Model with Sample Size $T = 100$ with No Prewhitening

		ρ							
Kernel		0.1	0.3	0.5	0.7	0.9	0.95	-0.3	-0.5
						90%			
Bartlett	MSE	88.02	86.51	83.75	78.70	65.50	57.10	86.66	84.21
	CPE	88.46	88.43	87.38*	84.45*	76.07*	68.69*	88.39	87.68
Parzen	MSE	87.72	86.35	84.05	79.09	64.08	56.36	86.96	85.02
	CPE	88.71	88.67*	87.30	84.00	71.47	62.73	88.73*	88.16*
QS	MSE	87.89	86.52	84.38	79.71	64.21	55.28	87.09	85.04
	CPE	88.78*	88.64	87.14	83.76	71.58	62.55	88.72	88.00
						95%			
Bartlett	MSE	93.65	92.50	89.97	85.66	73.38	64.84	92.40	90.30
	CPE	94.14	93.64	92.81	90.72*	83.34*	76.30*	93.63	93.12
Parzen	MSE	93.45	92.17	90.07	85.67	71.58	63.59	92.57	90.79
	CPE	94.26*	93.85*	92.83*	90.00	78.80	70.72	93.86*	93.45*
QS	MSE	93.58	92.38	90.32	86.19	71.47	62.07	92.60	90.85
	CPE	94.22	93.76	92.75	89.94	79.01	70.37	93.80	93.43

The superscript * indicates the most accurate confidence interval for each scenario

Table IV. Empirical Confidence Levels of Nominal 90% and 95% Confidence Intervals for AR(1)-HET Model with Sample Size $T = 100$ with No Prewhitening

		ρ							
		0.1	0.3	0.5	0.7	0.9	0.95	-0.3	-0.5
					90%				
Bartlett	MSE	86.72	85.15	81.60	75.70	60.21	48.16	85.08	82.28
	CPE	87.23	86.77	84.84	81.49*	69.72*	58.15*	86.98	86.16
Parzen	MSE	86.50	85.18	82.27	76.68	61.66	51.12	85.35	83.55
	CPE	87.27	86.75	84.88	80.78	66.31	55.58	87.18*	86.65*
QS	MSE	86.76	85.30	82.40	77.20	62.51	51.53	85.43	83.46
	CPE	87.31*	86.81*	84.91*	80.53	66.49	55.53	87.12	86.60
					95%				
Bartlett	MSE	92.60	91.25	88.37	83.25	68.35	55.38	91.43	89.28
	CPE	93.06	92.33	91.05	88.41*	77.76*	66.31*	92.76	92.08
Parzen	MSE	92.33	91.07	88.59	83.97	69.55	58.32	91.70	90.06
	CPE	93.17*	92.59*	91.08*	87.70	74.22	63.34	92.96*	92.42*
QS	MSE	92.41	91.27	88.68	84.23	70.15	58.66	91.70	90.10
	CPE	93.17*	92.48	90.89	87.52	74.13	63.37	92.96*	92.34

The superscript * indicates the most accurate confidence interval for each scenario

Table V. Empirical Confidence Levels of Nominal 90% and 95% Confidence Intervals for MA(1)-HOM Model with Sample Size $T = 100$ with No Prewhitening

		ψ							
		0.1	0.3	0.5	0.7	0.9	0.99	-0.3	-0.7
					90%				
Bartlett	MSE	88.02	86.94	85.88	85.29	85.00	85.04	86.97	85.39
	CPE	88.48	88.50	88.53*	88.44*	88.48*	88.42*	88.62	88.61
Parzen	MSE	87.76	86.90	85.75	85.24	85.10	85.17	87.20	86.18
	CPE	88.71	88.81*	88.51	88.42	88.41	88.28	88.87	88.78
QS	MSE	87.93	87.12	86.11	85.63	85.43	85.36	87.27	86.26
	CPE	88.81*	88.73	88.50	88.35	88.29	88.19	88.88*	88.79*
					95%				
Bartlett	MSE	93.63	92.59	91.59	91.15	91.03	90.96	92.88	91.51
	CPE	94.15	93.73	93.59	93.75*	93.87*	93.90*	93.86	93.92
Parzen	MSE	93.48	92.48	91.61	91.14	90.99	91.03	93.04	91.97
	CPE	94.24*	93.89*	93.72	93.75*	93.75	93.74	93.97*	94.19*
QS	MSE	93.55	92.65	91.71	91.30	91.26	91.28	93.12	92.04
	CPE	94.23	93.87	93.74*	93.72	93.65	93.65	93.91	94.06

The superscript * indicates the most accurate confidence interval for each scenario

Table VI. Empirical Confidence Levels of Different Confidence Intervals with $T = 400$ with No Prewhitening

		AR(1)-HOM			AR(1)-HET			MA(1)-HOM		
	ρ or ψ	0.5	0.9	-0.5	0.5	0.9	-0.5	0.5	0.9	-0.7
			90%			90%			90%	
Bartlett	MSE	86.68	76.44	86.67	86.68	76.44	86.67	88.17	87.91	88.00
	CPE	88.49	82.40*	88.98	88.49	82.40*	88.98	89.67	89.73	89.81
Parzen	MSE	87.33	77.97	87.28	87.33	77.97	87.28	88.49	88.25	88.55
	CPE	88.62*	80.63	89.16	88.62*	80.63	89.16*	89.73*	89.72	89.89*
QS	MSE	87.30	78.22	87.21	87.30	78.22	87.21	88.70	88.34	88.59
	CPE	88.58	80.58	89.15	88.58	80.58	89.15	89.68	89.64	89.81
			95%			95%			95%	
Bartlett	MSE	92.87	85.56	92.91	92.18	83.80	92.70	93.67	93.36	93.55
	CPE	94.11	90.46*	94.67	93.94	88.96*	94.29	94.58	94.57	94.85
Parzen	MSE	93.20	85.74	93.46	92.65	85.00	93.26	93.88	93.68	93.92
	CPE	94.27*	88.83	94.83*	93.99*	87.62	94.42*	94.66*	94.58*	94.88
QS	MSE	93.24	86.24	93.36	92.69	85.26	93.14	93.91	93.76	94.00
	CPE	94.20	88.89	94.72	93.88	87.46	94.40	94.64	94.58*	94.91*

The superscript * indicates the most accurate confidence interval for each scenario

Table VII. Empirical Confidence Levels of Nominal 90% Confidence Intervals Under Different Models with Prewhitening

		AR(1)-HOM, $T = 100$							
	$\rho =$	0.1	0.3	0.5	0.7	0.9	0.95	-0.3	-0.5
Bartlett	MSE	87.98	87.94	86.92	84.14	76.97	71.15	87.91	87.06
	CPE	87.88	88.01	87.29	85.16	78.49	72.94	87.77	87.31
Parzen	MSE	87.78	87.94	86.90	84.38	77.17	71.45	87.74	87.02
	CPE	88.15	88.05	87.28	85.12	78.64	73.21	88.05	87.38
QS	MSE	87.90	87.96	87.00	84.36	77.15	71.49	87.81	87.12
	CPE	88.24	88.13	87.25	85.18	78.48	73.15	88.08	87.45
		AR(1)-HET, $T = 100$							
	$\rho =$	0.1	0.3	0.5	0.7	0.9	0.95	-0.3	-0.5
Bartlett	MSE	86.64	86.18	84.39	81.17	68.67	58.15	86.21	85.50
	CPE	86.57	86.27	84.83	82.07	70.57	60.61	85.99	85.26
Parzen	MSE	86.40	86.09	84.40	81.18	69.44	59.63	85.91	85.11
	CPE	86.61	86.31	84.78	81.97	70.54	60.92	86.28	85.58
QS	MSE	86.50	86.15	84.43	81.17	69.34	59.26	85.99	85.20
	CPE	86.74	86.35	84.79	82.02	70.22	60.70	86.39	85.56
		MA(1)-HOM, $T = 100$							
	$\psi =$	0.1	0.3	0.5	0.7	0.9	0.99	-0.3	-0.7
Bartlett	MSE	88.03	88.12	88.28	88.75	88.96	88.96	88.02	88.55
	CPE	87.94	88.22	88.59	89.10	89.39	89.35	88.08	88.47
Parzen	MSE	87.84	88.07	88.27	88.57	88.61	88.57	87.89	88.17
	CPE	88.21	88.33	88.62	89.04	89.25	89.20	88.27	88.64
QS	MSE	88.02	88.17	88.46	88.79	88.96	88.97	87.96	88.40
	CPE	88.29	88.44	88.73	89.11	89.38	89.33	88.37	88.83

Appendix

A.1 Notation

Much of the notation is conventional but for ease of reference is collected together in this subsection. Some definitions are repeated in order to enhance the readability of the proof.

For an $m\times n$ matrix $C=(C_{ij})$, $\|C\|^2=\sum_{i=1}^{m}\sum_{j=1}^{n}C_{ij}^2$, vec$(\cdot)$ is the column-by-column vectorization function, vech$(\cdot)$ is the column stacking operator that stacks the elements on and below the main diagonal, D_m is the $m^2\times m\,(m+1)/2$ duplication matrix such that for any symmetric $m\times m$ matrix C, $vec(C)=D_m vech(C)$, $K_{m,n}$ is the $mn\times mn$ commutation matrix such that for any $m\times n$ matrix C, $K_{m,n}vec(C)=vec(C')$. We use $\kappa_j(x)$ to denote the *j-th* cumulant of a random variable x and let $k_j:=k(j/M)$. $v_{t,c}$ is the c-th element of vector v_t.

Let $u_t=y_t-x_t'\beta_0$, $\hat{u}_t=y_t-x_t\tilde{\beta}_T$, $v_t=z_tu_t$, $\hat{v}_t=z_t\hat{u}_t$, $w_t=z_tx_t'$, and set

$$\hat{\Gamma}_j=\begin{cases}T^{-1}\sum_{t=1}^{T}\hat{v}_{t+j}\hat{v}_t', & j\geq 0,\\ T^{-1}\sum_{t=1}^{T}\hat{v}_t\hat{v}_{t-j}', & j<0,\end{cases}$$

$$\tilde{\Gamma}_j=\begin{cases}T^{-1}\sum_{t=1}^{T}v_{t+j}v_t', & j\geq 0,\\ T^{-1}\sum_{t=1}^{T}v_tv_{t-j}', & j<0,\end{cases}$$

$$\Gamma_j=\begin{cases}Ev_{t+j}v_t', & j\geq 0,\\ Ev_tv_{t-j}', & j<0,\end{cases}$$

$$\nabla\tilde{\Gamma}_j\delta=\begin{cases}T^{-1}\sum_{t=1}^{T}\left(v_{t+j}z_t'x_t'+z_{t+j}v_t'x_{t+j}'\right)\delta, & j\geq 0,\\ T^{-1}\sum_{t=1}^{T}\left(v_tz_{t-j}'x_{t-j}'+z_tv_{t-j}'x_t'\right)\delta, & j<0,\end{cases}$$

$$\delta'\nabla^2\tilde{\Gamma}_j\delta=\begin{cases}T^{-1}\sum_{t=1}^{T}\delta'\left(x_{t+j}z_{t+j}z_t'x_t'\right)\delta, & j\geq 0,\\ T^{-1}\sum_{t=1}^{T}\delta'\left(x_tz_tz_{t-j}'x_{t-j}'\right)\delta, & j<0,\end{cases}$$

$$\hat{\Omega}_T=\sum_{j=-M}^{M}k_j\hat{\Gamma}_j,\ \tilde{\Omega}_T=\sum_{j=-M}^{M}k_j\tilde{\Gamma}_j,$$

$$\bar{\Omega}_T=\sum_{j=-M}^{M}k_j\Gamma_j,\ \Omega_0=\sum_{j=-\infty}^{\infty}\Gamma_j,$$

$$\nabla\tilde{\Omega}_T\delta=\sum_{j=-M}^{M}k_j\nabla\tilde{\Gamma}_j\delta,\ \delta\nabla^2\tilde{\Omega}_T\delta=\sum_{j=-M}^{M}k_j\delta'\nabla^2\tilde{\Gamma}_j\delta,$$

where δ is a vector in R^{d_2}. Note that $\nabla\tilde{\Gamma}_j$ should be regarded as an operator, as $\nabla\tilde{\Gamma}_j$ is neither a vector nor a matrix. Define

$$\hat{\Omega}_T^+=\sum_{j=0}^{M}k_j\hat{\Gamma}_j,\hat{\Omega}_T^-=\sum_{j=-M}^{-1}k_j\hat{\Gamma}_j,$$

$$\tilde{\Omega}_T^+=\sum_{j=0}^{M}k_j\tilde{\Gamma}_j,\tilde{\Omega}_T^-=\sum_{j=-M}^{-1}k_j\tilde{\Gamma}_j,$$

$$\bar{\Omega}_T^{+} = \sum_{j=0}^{M} k_j \Gamma_j, \bar{\Omega}_T^{-} = \sum_{j=-M}^{-1} k_j \Gamma_j.$$

Let $\check{\beta}_T$ be an estimator of β_0 such that

$$\sqrt{T}\left(\check{\beta}_T - \beta_0\right) = \left(G_0' V_0 G_0\right)^{-1} \left(G_0' V_0 S_T\right),$$

where $G_0 = E\left(z_t x_t'\right)$. Denote $\check{u}_t = y_t - x_t'\check{\beta}_T$, $\check{v}_t = z_t \check{u}_t$ and

$$\check{\Omega}_T = \sum_{j=-M}^{M} k_j \check{\Gamma}_j,\ \check{\Gamma}_j = \begin{cases} \frac{1}{T}\sum_{t=1}^{T} \check{v}_{t+j}\check{v}_t', & j \geq 0, \\ \frac{1}{T}\sum_{t=1}^{T} \check{v}_t \check{v}_{t-j}', & j < 0. \end{cases}$$

Let

$$A_T = \left(A_{T,\ell}\right) = \left(S_T', vec(G_T)', vech(\hat{\Omega}_T)'\right)'$$

and

$$A_0 = \left(A_{0,\ell}\right) = \left(0, vec\left(G_0\right)', vech\left(\Omega_0\right)'\right)',$$

then t_M is a function of the vector A_T. Throughout the paper, we use the convention that

$$\frac{\partial t_M(A_0)}{\partial A_T'} = \frac{\partial t_M(A_T)}{\partial A_T'}|_{A_T = A_0}.$$

A.2 Lemmas

We use the theorem in Yokoyama (1980) repeatedly and it is presented here for reference.

Theorem A.1 (Yokoyama) *Let $\{X_t\}$ be a strictly stationary and strong mixing sequence with $EX_1 = 0$ and $E\left\|X_1\right\|^{r+\eta} < \infty$ for some $r \geq 2$ and $\eta > 0$. If the mixing coefficients $\{\alpha_i\}$ satisfy*

$$\sum_{i=1}^{\infty} (i+1)^{r/2-1} \alpha_i^{\eta/(r+\eta)} < \infty,$$

then

$$E\left(\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t\right\|^r\right) \leq C,$$

for some positive constant $C > 0$.

Lemma A.2 *The following moment bounds hold*

(a) $E\|S_T\|^r = O(1)$.

(b) $E\left\|T^{1/2} vec\left(G_T - G_0\right)\right\|^r = O\left(1\right)$.

(c) $E\left\|(T/M)^{1/2}\ vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right)\right\|^{r/2} = O\left(1\right)$.

(d) $E\left\|T^{1/2} vech\left(\check{\Omega}_T - \tilde{\Omega}_T\right)\right\|^{r/4} = O\left(1\right)$.

(e) $E\left\|T^{1/2} vech\left(\left[\nabla\tilde{\Omega}_T\left(\check{\beta}_T - \beta_0\right)\right]\right)\right\|^{r/4} = O(1)$.

Proof. Parts (a) & (b) Let $S_{T,i}$ be the i-th element of S_T, then

$$E\left\|S_T\right\|^r = E\left(\sum_{i=1}^{d_2} S_{T,i}^2\right)^{r/2} \leq E\left(\sum_{i=1}^{d_2}\left|S_{T,i}\right|\right)^r \leq \left[\sum_{i=1}^{d_2}\left(E\left|S_{T,i}\right|^r\right)^{1/r}\right]^r \tag{A.1}$$

where the first inequality follows from: $\sum_{i=1}^{d_2} S_{T,i}^2 \leq (\sum_{i=1}^{d_2}|S_{T,i}|)^2$ and the second inequality follows from the Minkowski inequality:

$$\left[E(\sum_{i=1}^{d_2}|S_{T,i}|)^r\right]^{1/r} \leq \sum_{i=1}^{d_2}\left(E\left|S_{T,i}\right|^r\right)^{1/r}.$$

It suffices to show that $E\left|S_{T,i}\right|^r < \infty$. In view of Assumption 2, this inequality follows immediately from Theorem A.1. Part (b) can be similarly proved.

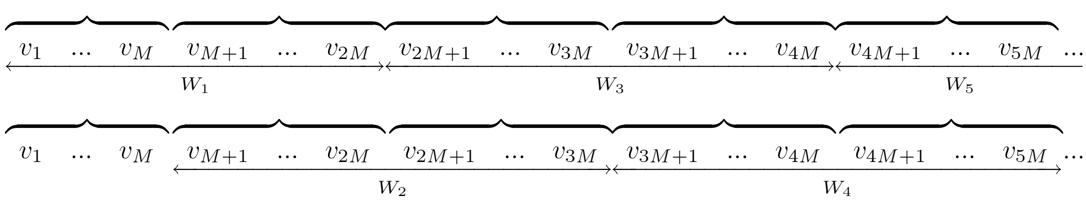
Part (c) We write

$$\begin{aligned}
&\left(\frac{T}{M}\right)^{1/2}\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \\
= &\left(\frac{T}{M}\right)^{1/2}\frac{1}{T}\sum_{t=1}^{T}\left\{\sum_{j=1}^{M}k_j\left(v_tv_{t+j}' - Ev_tv_{t+j}' + v_{t+j}'v_t - Ev_{t+j}'v_t\right) + v_tv_t' - Ev_tv_t'\right\} \\
= &\left(\frac{M}{T}\right)^{1/2}\sum_{i=1}^{[T/M]}\frac{1}{M}\sum_{t=(i-1)M+1}^{iM}\left\{\sum_{j=1}^{M}k_j\left(v_tv_{t+j}' - Ev_tv_{t+j}' + v_{t+j}'v_t - Ev_{t+j}'v_t\right) + v_tv_t' - Ev_tv_t'\right\} \\
= &\left(\frac{M}{T}\right)^{1/2}\sum_{i=1}^{[T/M]}W_i = \left(\frac{M}{T}\right)^{1/2}\left\{\sum_{i=0 \bmod 3}W_i + \sum_{i=1 \bmod 3}W_i + \sum_{i=2 \bmod 3}W_i\right\} \\
= &\left(\frac{M}{T}\right)^{1/2}\left(W^{(0)} + W^{(1)} + W^{(2)}\right),
\end{aligned}$$

where

$$\begin{aligned}
W_i &= \frac{1}{M}\sum_{t=(i-1)M+1}^{iM}\left\{\sum_{j=1}^{M}k_j\left(v_tv_{t+j}' - Ev_tv_{t+j}' + v_{t+j}'v_t - Ev_{t+j}'v_t\right) + v_tv_t' - Ev_tv_t'\right\} \\
W^{(0)} &= W_3 + W_6 + ..., \ W^{(1)} = W_1 + W_4 + ..., \text{ and } W^{(2)} = W_2 + W_5 + ...
\end{aligned}$$

The dependence of W_i on $\{v_t\}_{t=1}^T$ can be illustrated as follows:



This decomposition ensures that the summands in $W^{(j)}$, $j = 0, 1, 2$ are a strongly mixing sequence with mixing coefficients $\{\alpha_{iM}\}_{i=1}^{\infty}$.

Let $W_{k,\ell}^{(j)}$ be the (k,ℓ)-th element of $W^{(j)}$, then we have, using the Cauchy and Minkowski inequalities,

$$E\left\|\left(\frac{T}{M}\right)^{1/2}vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right)\right\|^{r/2} = E\left[\left(\frac{M}{T}\right)^{r/4}\left(\sum_{k,\ell}\left(W_{k,\ell}^{(0)} + W_{k,\ell}^{(1)} + W_{k,\ell}^{(2)}\right)^2\right)^{r/4}\right]$$

$$\leq \quad 3^{r/4} E\left[\left(\frac{M}{T}\right)^{r/4}\left(\sum_{k,\ell}\left[\left(W_{k,\ell}^{(0)}\right)^2+\left(W_{k,\ell}^{(1)}\right)^2+\left(W_{k,\ell}^{(2)}\right)^2\right]\right)^{r/4}\right] \quad \text{[Cauchy]}$$

$$\leq \quad 3^{r/4}\left\{\sum_{k,\ell}\left[E\left(\frac{M}{T}\right)^{r/4}\left(W_{k,\ell}^{(0)}\right)^{r/2}\right]^{4/r}+\sum_{k,\ell}\left[E\left(\frac{M}{T}\right)^{r/4}\left(W_{k,\ell}^{(1)}\right)^{r/2}\right]^{4/r} \right. \tag{A.2}$$

$$\left. +\sum_{k,\ell}\left[E\left(\frac{M}{T}\right)^{r/4}\left(W_{k,\ell}^{(2)}\right)^{r/2}\right]^{4/r}\right\}^{r/4}. \quad \text{[Minkowski]}$$

We proceed to evaluate each of the three sums. As a representative example, we consider the first term:

$$E\left[\sqrt{\frac{M}{T}}W_{k,\ell}^{(0)}\right]^{r/2}=E\left[\sqrt{\frac{M}{T}}(W_{3,k,\ell}+W_{6,k,\ell}+W_{9,k,\ell}+...\right]^{r/2},$$

where $W_{i,k,\ell}$ is the (k,ℓ)-th element of W_i. Note that the summands $W_3, W_6, W_9, ...$ form a strong mixing sequence with mixing coefficients $\{\alpha_{iM}\}_{i=1}^{\infty}$ and the number of summands is less than $\lfloor T/(3M)\rfloor+1$. It follows from the moment inequality in Theorem A.1 that $E\left[\sqrt{M/T}W_{k,\ell}^{(0)}\right]^{r/2}=O(1)$. Similarly, $E\left[\sqrt{M/T}W_{k,\ell}^{(j)}\right]^{r/2}=O(1)$ for $j=1$ and 2. Combining this with (A.2) yields the desired result.

Part (d) We first prove $E\left|\sqrt{T}\left(\check{\beta}_T-\beta_0\right)\right|^{r/2}=O(1)$, a result to be used below, which follows because

$$\begin{aligned} E\left|\sqrt{T}\left(\check{\beta}_T-\beta_0\right)\right|^{r/2} &= \left\|\left(G_0'V_0G_0\right)^{-1}\right\|^{r/2}E\left\|G_0'V_0S_T\right\|^{r/2} \\ &\leq \left\|\left(G_0'V_0G_0\right)^{-1}\right\|^{r/2}\left\|G_0'V_0\right\|^{r/2}\left(E\left\|S_T\right\|^r\right)^{1/2}=O(1) \end{aligned} \tag{A.3}$$

by parts (a) and (b). Here we have used the inequality: for any two compatible matrices A and B, $\|AB\|\leq\|A\|\,\|B\|$.

Next, we prove the lemma. It is easy to show that

$$\begin{aligned} & T^{1/2}\left(\check{\Omega}_T-\tilde{\Omega}_T\right) \\ = & -T^{1/2}\nabla\tilde{\Omega}_T\left(\check{\beta}_T-\beta_0\right)+T^{1/2}\left(\check{\beta}_T-\beta_0\right)\nabla^2\tilde{\Omega}_T\left(\check{\beta}_T-\beta_0\right) \\ = & -T^{1/2}\frac{1}{T}\sum_{t=1}^{T}\sum_{j=0}^{M}\left[k_jv_{t+j}z_t'x_t'(\check{\beta}_T-\beta_0)+k_jz_{t+j}v_t'x_{t+j}'(\check{\beta}_T-\beta_0)\right] \\ & +T^{1/2}\frac{1}{T}\sum_{t=1}^{T}\sum_{j=0}^{M}\left[k_j(\check{\beta}_T-\beta_0)'x_{t+j}z_{t+j}z_t'x_t'(\check{\beta}_T-\beta_0)\right] \\ & -T^{1/2}\frac{1}{T}\sum_{t=1}^{T}\sum_{j=-M}^{-1}k_jv_tz_{t-j}'x_{t-j}'(\check{\beta}_T-\beta_0)+k_jz_tv_{t-j}x_t'(\check{\beta}_T-\beta_0) \\ & +T^{1/2}\frac{1}{T}\sum_{t=1}^{T}\sum_{j=-M}^{-1}\left[k_j(\check{\beta}_T-\beta_0)'x_tz_tz_{t-j}'x_{t-j}'(\check{\beta}_T-\beta_0)\right]. \end{aligned} \tag{A.4}$$

We show that each of the above four terms has a finite $r/4$ moment. We focus on the first term and proofs for other terms are similar. In order to show

$$E\left(\left\|\sqrt{T}\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j vech\left[v_{t+j}z_t'x_t'(\check{\beta}_T-\beta_0)\right]\right\|^{r/4}\right)=O(1),$$

it suffices to prove each element of $\sqrt{T}\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j vech\left[v_{t+j}z_t' x_t'(\check{\beta}_T-\beta_0)\right]$ has a finite $r/4$ moment [c.f. (A.1)]. The typical element is of the form:

$$\sum_{a,b,c}\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j v_{t+j,\ell}z_{t,m}x_{t,a}\mathcal{C}_{a,b}^{-}\mathcal{C}_{b,c}\frac{1}{T^{1/2}}\sum_{p=1}^{T}v_{p,c},$$

where $\mathcal{C}^{-}=\left(\mathcal{C}_{a,b}^{-}\right)=\left(G_0'V_0G_0\right)^{-1}$ and $\mathcal{C}=\left(\mathcal{C}_{b,c}\right)=G_0'V_0$. It remains to show that

$$\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j v_{t+j,\ell}z_{t,m}x_{t,a}\mathcal{C}_{a,b}^{-}\mathcal{C}_{b,c}\frac{1}{T^{1/2}}\sum_{p=1}^{T}v_{p,c}$$

has a finite $r/4$ moment for each fixed $a,b,$ and c. Now, by the Hölder inequality and the inequality $(|x|+|y|)^{r/2}\leq 2^{r/2}\left(|x|^{r/2}+|y|^{r/2}\right)$,

$$\begin{aligned}
& E\left|\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j v_{t+j,\ell}z_{t,m}x_{t,a}\mathcal{C}_{a,b}^{-}\mathcal{C}_{b,c}\frac{1}{T^{1/2}}\sum_{p=1}^{T}v_{p,c}\right|^{r/4} \\
\leq\ & C\left(E\left(\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j v_{t+j,\ell}w_{t,m,a}\right)^{r/2}\right)^{1/2}\left(E\left(\frac{1}{T^{1/2}}\sum_{p=1}^{T}v_{p,c}\right)^{r/2}\right)^{1/2} \\
=\ & C\left\{E\left(\left|\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j\left(v_{t+j,\ell}w_{t,m,a}-Ev_{t+j,\ell}w_{t,m,a}\right)\right|\right)^{r/2}\right. \\
& \left.+\left(\left|\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j Ev_{t+j,\ell}w_{t,m,a}\right|\right)^{r/2}\right\}^{1/2} O(1),
\end{aligned}$$

where we have used Lemma A.2 (a) to obtain the $O(1)$ term. Following the same steps as in the proof of part (c), we can show that

$$E\left(\left|\frac{1}{M\sqrt{T}}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j\left(v_{t+j,\ell}w_{t,m,a}-Ev_{t+j,\ell}w_{t,m,a}\right)\right|\right)^{r/2}=O(1),$$

when $w_{t,m,a}=1$ and

$$E\left(\left|\frac{1}{\sqrt{MT}}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j\left(v_{t+j,\ell}w_{t,m,a}-Ev_{t+j,\ell}w_{t,m,a}\right)\right|\right)^{r/2}=O(1),$$

when $w_{t,m,a}\neq 1$

In addition, by the strong mixing assumption, we have

$$\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j Ev_{t+j,\ell}w_{t,m,a}=\sum_{j=0}^{M}k_j cov(v_{t+j,\ell},w_{t,m,a})=O(1).$$

Therefore, we have shown that

$$\begin{aligned}
& E\left|\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T}k_j v_{t+j,\ell}z_{t,m}x_{t,a}\mathcal{C}_{a,b}^{-}\mathcal{C}_{b,c}\frac{1}{T^{1/2}}\sum_{p=1}^{T}v_{p,c}\right|^{r/4} \\
= & O\left[\left(\frac{M}{\sqrt{T}}\right)^{r/2}+\left(\frac{1}{\sqrt{T}}\right)^{r/2}+O(1)\right]^{1/2} \\
= & O\left(1\right),
\end{aligned}$$

and thus the first term in (A.4) has a finite $r/4$ moment. Similarly, we can show that other terms have finite $r/4$ moments.

Part (e) This part has been proved in the proof of part (d) above. ■

Lemma A.3 *For all* $\varepsilon > 0$ *and some* $\chi \geq 2$,

(a) $P\left(\|S_T\| > \varepsilon \log T\right) = o(T^{-\chi})$,

(b) $P\left(\sqrt{T}\,\|vec\left(G_T - G_0\right)\| > \varepsilon \log T\right) = o(T^{-\chi})$,

(c) $P\left(\sqrt{T}\left\|\tilde{\beta}_T - \check{\beta}_T\right\| > \varepsilon\left(\log^2 T\right)T^{-q/(2q+1)}\right) = o(T^{-\chi})$,

(d) $P\left(\sqrt{T}\left\|\check{\beta}_T - \beta_0\right\| > \varepsilon \log T\right) = o(T^{-\chi})$,

(e) $P\left(\left\|\sqrt{T/M}vech\left(\hat{\Omega}_T - \check{\Omega}_T\right)\right\| > \varepsilon\left(\log^2 T\right)T^{-q/(2q+1)}\right) = o\left(\eta_T\right)$,

(f) $P\left(\left\|\sqrt{T/M}vech\left(\hat{\Omega}_T - \Omega_0\right)\right\| > \varepsilon \log^2 T\right) = o\left(\eta_T\right)$,

where

$$\eta_T = \frac{1}{\log T}\max\left(\frac{1}{M^q},\frac{M}{T}\right).$$

Proof. Parts (a) & (b) Note that

$$(S_T, vec(\sqrt{T}\left(G_T - G_0\right)))' = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(R_t - ER_t\right).$$

Parts (a) and (b) follow from Lemma 3(c) in Andrews (2002) (with his $\tilde{X}_i$ equal to R_t and his function $f(\cdot)$ equal to the identity function).

Part (c)

$$\begin{aligned}
& P\left(\sqrt{T}\left\|\tilde{\beta}_T - \check{\beta}_T\right\| > \varepsilon\left(\log^2 T\right)T^{-q/(2q+1)}\right) \\
= & P\left(\left[\left((G_T'V_TG_T)^{-1}\right)(G_T'V_T) - \left((G_0'V_0G_0)^{-1}\right)(G_0'V_0)\right]S_T > \varepsilon\left(\log^2 T\right)T^{-q/(2q+1)}\right) \\
\leq & P\left(T^{q/(2q+1)}\left\|V_T - V_0\right\| > C\log T\right) + P\left(T^{q/(2q+1)}\left\|G_T - G_0\right\| > C\log T\right) \\
& +P\left(\|S_T\| > C\log T\right) \\
= & o(T^{-\chi}),
\end{aligned}$$

by parts (a), (b) and Assumption 7.

Part (d) This part can be proved using the same arguments as those for part (c).

Part (e) Noting that $\hat{\Omega}_T - \check{\Omega}_T = \hat{\Omega}_T^+ - \check{\Omega}_T^+ + \hat{\Omega}_T^- - \check{\Omega}_T^-$, we start with the decomposition:

$$\begin{aligned}
\hat{\Omega}_T^+ - \check{\Omega}_T^+ &= \frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T} k_j z_{t+j}\left(u_{t+j} - x_{t+j}\left(\tilde{\beta}_T - \beta_0\right)\right)\left(u_t - x_t\left(\tilde{\beta}_T - \beta_0\right)\right) z_t' \\
&\quad -\frac{1}{T}\sum_{j=0}^{M}\sum_{t=1}^{T} k_j z_{t+j}\left(u_{t+j} - x_{t+j}\left(\check{\beta}_T - \beta_0\right)\right)\left(u_t - x_t\left(\check{\beta}_T - \beta_0\right)\right) z_t' \\
: &= B_1 + B_2 + B_3 + B_4,
\end{aligned}$$

where

$$\begin{aligned}
B_1 &= -\sum_{j=0}^{M}\frac{1}{T}\sum_{t=1}^{T} k_j z_{t+j} z_t' u_{t+j} x_t'\left(\tilde{\beta}_T - \check{\beta}_T\right), \\
B_2 &= -\sum_{j=0}^{M}\frac{1}{T}\sum_{t=1}^{T} k_j z_{t+j} z_t' u_t x_{t+j}'\left(\tilde{\beta}_T - \check{\beta}_T\right), \\
B_3 &= \sum_{j=0}^{M}\frac{1}{T}\sum_{t=1}^{T} k_j z_{t+j} z_t' x_{t+j}'\left(\tilde{\beta}_T - \beta_0\right) x_t'\left(\tilde{\beta}_T - \beta_0\right), \\
B_4 &= -\sum_{j=0}^{M}\frac{1}{T}\sum_{t=1}^{T} k_j z_{t+j} z_t' x_{t+j}'\left(\check{\beta}_T - \beta_0\right) x_t'\left(\check{\beta}_T - \beta_0\right).
\end{aligned}$$

We consider each of the four terms in the above equation. The typical element of B_1 is

$$-\sum_{c=1}^{d_1}\sum_{j=0}^{M}\frac{1}{T}\sum_{t=1}^{T} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\left(\tilde{\beta}_{T,c} - \check{\beta}_{T,c}\right).$$

Note that, for any constant $C > 0$

$$\begin{aligned}
& P\left(\sqrt{\frac{T}{M}}\left|\sum_{j=0}^{M}\frac{1}{T}\sum_{t=1}^{T} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\left(\tilde{\beta}_{T,c} - \check{\beta}_{T,c}\right)\right| > C\left(\log^3 T\right) T^{-q/(2q+1)}\right) \\
\le\ & P\left(\sqrt{\frac{T}{M}}\frac{1}{T}\sum_{t=1}^{T}\left|\sum_{j=0}^{M} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\left(\tilde{\beta}_{T,c} - \check{\beta}_{T,c}\right)\right| > C\left(\log^3 T\right) T^{-q/(2q+1)}\right) \\
=\ & P\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{\sqrt{M}}\sum_{j=0}^{M}\left[k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\sqrt{T}\left(\tilde{\beta}_{T,c} - \check{\beta}_{T,c}\right)\right]\right| > C\left(\log^3 T\right) T^{-q/(2q+1)}\right) \\
\le\ & \sum_{t=1}^{T} P\left(\left|\frac{1}{\sqrt{M}}\sum_{j=0}^{M} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right|\left|\sqrt{T}\left(\tilde{\beta}_{T,c} - \check{\beta}_{T,c}\right)\right| > C\left(\log^3 T\right) T^{-q/(2q+1)}\right)
\end{aligned}$$

which is bounded by

$$\begin{aligned}
& \sum_{t=1}^{T} P\left(\left|\frac{1}{\sqrt{M}}\sum_{j=0}^{M} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right| > \log T\right) \qquad \text{(A.5)} \\
& + \sum_{t=1}^{T} P\left(\left|\sqrt{T}\left(\tilde{\beta}_{T,c} - \check{\beta}_{T,c}\right)\right| > C\left(\log T\right)^2 T^{-q/(2q+1)}\right) \\
=\ & \sum_{t=1}^{T} P\left(\left|\frac{1}{\sqrt{M}}\sum_{j=0}^{M} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right| > \log T\right) + o(T^{-\chi+1}).
\end{aligned}$$

Here the equality follows from part (c). Now

$$
\begin{aligned}
& \sum_{t=1}^{T} P\left(\left|\frac{1}{\sqrt{M}} \sum_{j=0}^{M} k_j z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right| > \log T\right) \\
\leq & \sum_{t=1}^{T} P\left(\left|\frac{1}{\sqrt{M}} \sum_{j=0}^{M} k_j \left[z_{t+j,a} z_{t,b} u_{t+j} x_{t,c} - E\left(z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right)\right]\right| > \frac{1}{2}\log T\right) \\
& + \sum_{t=1}^{T} P\left(\left|\frac{1}{\sqrt{M}} \sum_{j=0}^{M} E\left(z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right)\right| > \frac{1}{2}\log T\right) \\
= & \ o\left(\eta_T\right),
\end{aligned}
$$

where the last equality holds because

$$
P\left(\left\|\frac{1}{\sqrt{M}} \sum_{j=0}^{M} k_j \left[z_{t+j,a} z_{t,b} u_{t+j} x_{t,c} - E z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right]\right\| > \frac{1}{2}\log T\right) = o\left(T^{-\chi}\right),
$$

by Lemma 3(c) in Andrews (2002) and

$$
\left\|\frac{1}{\sqrt{M}} \sum_{j=0}^{M} E\left(z_{t+j,a} z_{t,b} u_{t+j} x_{t,c}\right)\right\| = O\left(\frac{1}{\sqrt{M}}\right).
$$

Therefore, we have proved that

$$
P\left(\sqrt{\frac{T}{M}}\left\|B_1\right\| > \frac{\varepsilon}{8}\left(\log^3 T\right) T^{-q/(2q+1)}\right) = o\left(\eta_T\right) \text{ for all } \varepsilon > 0.
$$

Similarly, we can show that $P\left(\|B_i\| > (\varepsilon/8)\left(\log^2 T\right) T^{-q/(2q+1)}\right) = o(\eta_T)$ for $i = 2, 3, 4$. In consequence,

$$
P\left(\left\|\sqrt{\frac{T}{M}} vech\left(\hat{\Omega}_T^{+} - \check{\Omega}_T^{+}\right)\right\| > (\varepsilon/2)\left(\log^2 T\right) T^{-q/(2q+1)}\right) = o\left(\eta_T\right).
$$

By symmetric arguments, we can show that

$$
P\left(\left\|\sqrt{\frac{T}{M}} vech\left(\hat{\Omega}_T^{-} - \check{\Omega}_T^{-}\right)\right\| > (\varepsilon/2)\left(\log^2 T\right) T^{-q/(2q+1)}\right) = o\left(\eta_T\right).
$$

Combining the above two equations yields

$$
P\left(\left\|\sqrt{\frac{T}{M}} vech\left(\hat{\Omega}_T - \check{\Omega}_T\right)\right\| > \varepsilon\left(\log^2 T\right) T^{-q/(2q+1)}\right) = o\left(\eta_T\right),
$$

as desired.

Part (f) We write:

$$
\begin{aligned}
& P\left(\left\|\sqrt{T/M}\, vech\left(\hat{\Omega}_T - \Omega_0\right)\right\| > \varepsilon \log^2 T\right) \\
\leq & \ P\left(\left\|\sqrt{T/M}\, vech\left(\hat{\Omega}_T - \check{\Omega}_T\right)\right\| > \frac{1}{4}\varepsilon \log^2 T\right) \\
& + P\left(\left\|\sqrt{T/M}\, vech\left(\check{\Omega}_T - \tilde{\Omega}_T\right)\right\| > \frac{1}{4}\varepsilon \log^2 T\right) \\
& + P\left(\left\|\sqrt{T/M}\, vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right)\right\| > \frac{1}{4}\varepsilon \log^2 T\right) \\
& + P\left(\left\|\sqrt{T/M}\, vech\left(\bar{\Omega}_T - \Omega_0\right)\right\| > \frac{1}{4}\varepsilon \log^2 T\right).
\end{aligned}
$$

Consider each term in turn. First, it follows from part (e) that

$$P\left(\left\|\sqrt{T/M}vech\left(\hat{\Omega}_T-\check{\Omega}_T\right)\right\|>\frac{1}{4}\varepsilon\log^2 T\right)=o(\eta_T).$$

Second, using part (d) and a proof similar to that for part (c), we can show that

$$P\left(\left\|\sqrt{T/M}vech\left(\check{\Omega}_T-\tilde{\Omega}_T\right)\right\|>\frac{1}{4}\varepsilon\log^2 T\right)=o(\eta_T).$$

Third, using the same argument as in Velasco and Robinson (2001), we can establish an Edgeworth expansion of the form:

$$P\left(\left|\sqrt{T/M}\left(\tilde{\Omega}_T(i,j)-\bar{\Omega}_T(i,j)\right)\right|<x\right)=\Phi(\tilde{c}x)-\Phi(-\tilde{c}x)+\frac{M}{T}\phi(\tilde{c}x)p(\tilde{c}x)+o\left(\frac{M}{T}\right),$$

for some constant $\tilde{c}$ and polynomial $p(\cdot)$, where $\tilde{\Omega}_T(i,j)$ is the (i,j)-th element of $\tilde{\Omega}_T$ and $\bar{\Omega}_T(i,j)$ is the (i,j)-th element of $\bar{\Omega}_T$. Consequently,

$$\begin{aligned}&P\left(\left\|\sqrt{T/M}vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\right\|>(1/4)\,\varepsilon\log^2 T\right)\\=\;&2\Phi(-c\varepsilon\log^2 T)+\frac{M}{T}\phi(c\varepsilon\log^2 T)p(c\varepsilon\log^2 T)+o\left(\eta_T\right),\end{aligned}$$

where c is a constant. Using $\Phi(-z)\leq C\exp(-z^2/2)$ for some constant $C>0$ and $z>1$, we have

$$\Phi(-c\varepsilon\log^2 T)\leq C\exp\left(-\frac{1}{2}c^2\varepsilon^2\log^4 T\right)\leq C\exp\left(-\chi\log T\right)=O(T^{-\chi})$$

for sufficiently large T. The expression $\phi(c\varepsilon\log^2 T)p(c\varepsilon\log^2 T)$ is a finite sum of terms of the form $\left(c\varepsilon\log^2 T\right)^j\phi(c\varepsilon\log^2 T)$, which can be easily shown to be $o(\eta_T)$. Therefore, we have shown that

$$P\left(\left\|\sqrt{T/M}vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\right\|>\frac{1}{4}\varepsilon\log^2 T\right)=o\left(\eta_T\right).$$

Finally, $\sqrt{T/M}vech\left(\bar{\Omega}_T-\Omega_0\right)$ is asymptotically equivalent to $\sqrt{T/M^{2q+1}}C$ for some constant vector C. Therefore, under the rate condition in Assumption 6, $\sqrt{T/M}vech\left(\bar{\Omega}_T-\Omega_0\right)$ converges to zero. As a result,

$$P\left(\left\|\sqrt{T/M}vech\left(\bar{\Omega}_T-\Omega_0\right)\right\|>(1/4)\,\varepsilon\log^2 T\right)=0,$$

for sufficiently large T.

Combining the above results completes the proof of this part. ■

Lemma A.4 *Let*

$$\begin{aligned}g_T\;=\;&\mathbf{a}_h'S_T+\mathbf{b}_h'\left[vec\left(G_T-G_0\right)\otimes S_T\right]+\mathbf{c}_h'\left[vech\left(\hat{\Omega}_T-\Omega_0\right)\otimes S_T\right]\\&+\mathbf{d}_h'\left[vech(\hat{\Omega}_T-\bar{\Omega}_T)\otimes vech(\hat{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right],\end{aligned}$$

where

$$\mathbf{a}_h=\frac{\partial t_M(A_0)}{\partial S_T},\mathbf{b}_h=vec\left(\frac{\partial t_M(A_0)}{\partial S_T\partial vec(G_T)'}\right),\mathbf{c}_h=vec\left(\frac{\partial t_M(A_0)}{\partial S_T\partial vech(\hat{\Omega}_T)'}\right),$$

and

$$\mathbf{d}_h = \frac{1}{2} vec \left[\frac{\partial vec}{\partial \left[vech(\hat{\Omega}_T) \right]'} \left(\frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \right) \right].$$

Then $t_M = g_T + \xi_T$ *where* ξ_T *satisfies*

$$P\left(|\xi_T| > \eta_T \right) = o\left(\eta_T \right).$$

Proof. Note that

$$t_M(A_T) = \left(\mathcal{R}' \left(G_T' \hat{\Omega}_T^{-1} G_T \right)^{-1} \mathcal{R} \right)^{-1/2} \mathcal{R}' \left(G_T' \hat{\Omega}_T^{-1} G_T \right)^{-1} G_T' \hat{\Omega}_T^{-1} S_T.$$

Taking a Taylor series expansion of t_M around $A_T = A_0$, we have

$$\begin{aligned} t_M(A_T) \quad = \quad & \frac{\partial t_M(A_0)}{\partial S_T'} S_T + \frac{\partial t_M(A_0)}{\partial vec(G_T)'} vec\left(G_T - G_0 \right) + \frac{\partial t_M(A_0)}{\partial vech(\hat{\Omega}_T)'} vech(\hat{\Omega}_T - \Omega_0) \\ & + S_T' \frac{\partial t_M(A_0)}{\partial S_T \partial vec(G_T)'} vec\left(G_T - G_0 \right) + S_T' \frac{\partial t_M(A_0)}{\partial S_T \partial vech(\hat{\Omega}_T)'} vech(\hat{\Omega}_T - \Omega_0) \\ & + vech(\hat{\Omega}_T - \Omega_0)' \frac{\partial t_M(A_0)}{\partial vech(\hat{\Omega}_T) \partial vec(G_T)'} vec\left(G_T - G_0 \right) \\ & + I_1 + I_2 + I_3 + \frac{1}{24} Rem_T, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \sum_{i,j,p} \frac{\partial}{\partial vec(G_T)_p} \frac{\partial^2 t_M(A_0)}{\partial vech(\hat{\Omega}_T)_j \partial S_{T,i}} \left(vech(\hat{\Omega}_T - \Omega_0)_j \right) S_{T,i} \left(vec(G_T - G_0)_p \right), \\ I_2 &= \frac{1}{2} \sum_{i,j,p} \frac{\partial}{\partial vech(\hat{\Omega}_T)_p} \frac{\partial^2 t_M(A_0)}{\partial vech(\hat{\Omega}_T)_j \partial S_{T,i}} \left(vech(\hat{\Omega}_T - \Omega_0)_j \right) S_{T,i} \left(vech(\hat{\Omega}_T - \Omega_0)_p \right), \\ I_3 &= \frac{1}{2} \sum_{i,j,p} \frac{\partial}{\partial G_{T,p}} \frac{\partial^2 t_M(A_0)}{\partial vec(\hat{G}_T)_j \partial S_{T,i}} \left(vec(G_T - G_0)_j \right) S_{T,i} \left(vec(G_T - G_0)_p \right), \end{aligned}$$

and

$$Rem_T = \frac{1}{24} \sum_{i,j,p,q} \left(\frac{\partial^4 t_M(\tilde{A})}{\partial A_{T,i} \partial A_{T,j} \partial A_{T,p} \partial A_{T,q}} \prod_{\ell = i,j,p,q} \left(A_{T,\ell} - A_{0,\ell} \right) \right).$$

In the above expression $vec(U)_p$ denotes the p-th element of $vec(U)$.

Note that

$$\begin{aligned}
I_1 &= \sum_{i,j,k} \frac{\partial}{\partial vec(G_T)_k} \frac{\partial^2 t_M(A_0)}{\partial vech(\hat{\Omega}_T)_j \partial S_{T,i}} \left(vech(\hat{\Omega}_T - \Omega_0)_j \right) S_{T,i} \left(vec(G_T - G_0)_k \right) \\
&= \sum_k \frac{\partial}{\partial vec(G_T)_k} \left(S_T' \frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \left(vech(\hat{\Omega}_T - \Omega_0)_j \right) \right) \left(vec(G_T - G_0)_k \right) \\
&= \left(\left[vech(\hat{\Omega}_T - \Omega_0) \right]' \otimes S_T' \right) \sum_k \frac{\partial vec}{\partial vec(G_T)_k} \frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \left(vec(G_T - G_0)_k \right) \\
&= \left(\left[vech(\hat{\Omega}_T - \Omega_0) \right]' \otimes S_T' \right) \frac{\partial vec}{\partial \left[vec(G_T) \right]'} \left(\frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \right) vec\left(G_T - G_0 \right),
\end{aligned}$$

so

$$\begin{aligned}
I_1 &= \left[(vec(G_T - G_0))' \otimes \left(vech(\hat{\Omega}_T - \Omega_0) \right)' \otimes S_T' \right] \\
&\quad \times vec\left[\frac{\partial vec}{\partial vec(G_T)'} \left(\frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \right) \right] \\
&= vec\left[\frac{\partial vec}{\partial vec(G_T)'} \left(\frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \right) \right] \left[vec(G_T - G_0) \otimes vech(\hat{\Omega}_T - \Omega_0) \otimes S_T \right] \\
&= \mathbf{e}_h' \left[vec(G_T - G_0) \otimes vech(\hat{\Omega}_T - \Omega_0) \otimes S_T \right],
\end{aligned}$$

where

$$\mathbf{e}_h = vec\left[\frac{\partial vec}{\partial vec(G_T)'} \left(\frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vech\left(\hat{\Omega}_T \right) \right]'} \right) \right].$$

Similarly,

$$\begin{aligned}
I_2 &= \mathbf{d}_h' \left[vech(\hat{\Omega}_T - \Omega_0) \otimes vech(\hat{\Omega}_T - \Omega_0) \otimes S_T \right], \\
I_3 &= \mathbf{f}_h' \left[vech(G_T - G_0) \otimes vec(G_T - G_0) \otimes S_T \right],
\end{aligned}$$

where

$$\mathbf{f}_h = \frac{1}{2} vec\left[\frac{\partial vec}{\partial vec(G_T)'} \left(\frac{\partial^2 t_M(A_0)}{\partial S_T \partial \left[vec\left(G_T \right) \right]'} \right) \right].$$

Therefore, we have

$$\begin{aligned}
& t_M = \mathbf{a}_h' S_T + \mathbf{b}_h' [vec\left(G_T - G_0 \right) \otimes S_T] + \mathbf{c}_h' [vech\left(\hat{\Omega}_T - \Omega_0 \right) \otimes S_T] \\
& + \mathbf{d}_h' \left[vech(\hat{\Omega}_T - \Omega_0) \otimes vech(\hat{\Omega}_T - \Omega_0) \otimes S_T \right] \\
& + \mathbf{e}_h' \left[vec(G_T - G_0) \otimes vech(\hat{\Omega}_T - \Omega_0) \otimes S_T \right] \\
& + \mathbf{f}_h' \left[vech(G_T - G_0) \otimes vec(G_T - G_0) \otimes S_T \right] + Rem_T \\
= \; & g_T + \xi_T,
\end{aligned}$$

where

$$
\begin{aligned}
\xi_T &= Rem_T + \mathbf{d}_h' \left[vech(\hat{\Omega}_T - \bar{\Omega}_T) \otimes vech(\bar{\Omega}_T - \Omega_0) \otimes S_T \right] \\
&+ \mathbf{d}_h' \left[vech(\bar{\Omega}_T - \Omega_0) \otimes vech(\hat{\Omega}_T - \bar{\Omega}_T) \otimes S_T \right] \\
&+ \mathbf{d}_h' \left[vech(\bar{\Omega}_T - \Omega_0) \otimes vech(\bar{\Omega}_T - \Omega_0) \otimes S_T \right] \\
&+ \mathbf{e}_h' \left[vec(G_T - G_0) \otimes vech(\hat{\Omega}_T - \Omega_0) \otimes S_T \right] \\
&+ \mathbf{f}_h' \left[vech(G_T - G_0) \otimes vec(G_T - G_0) \otimes S_T \right].
\end{aligned} \tag{A.6}
$$

Let $N(A_0)$ be a small neighborhood around A_0, then, by Lemma A.3 (a), (b) and (f), we have

$$P\left(A_T \notin N(A_0) = o(\eta_T).\right.$$

It thus suffices to consider $P\left(|Rem_T| > \eta_T, A_T \in N(A_0)\right)$. Let C_1 and C_2 be some constants, then

$$
\begin{aligned}
& P\left(|Rem_T| > \eta_T, A_T \in N(A_0)\right) \\
\leq\ & P\left(\left| \sum_{i,j,p,q} \left\{ \sup_{A \in N(A_0)} \left| \frac{\partial^4 t_M(A)}{\partial A_{T,i} \partial A_{T,j} \partial A_{T,p} \partial A_{T,q}} \right| \prod_{\ell=i,j,p,q} |A_{T,\ell} - A_{0,\ell}| \right\} \right| > 24\eta_T \right) \\
\leq\ & C_2 P\left[\sum_{i,j,p,q} \left| \sup_{A \in N(A_0)} \left| \frac{\partial^4 t_M(A)}{\partial A_{T,i} \partial A_{T,j} \partial A_{T,p} \partial A_{T,q}} \right| \prod_{\ell=i,j,p,q} |A_{T,\ell} - A_{0,\ell}| \right| > C_1 \eta_T \right] \\
\leq\ & C_2 \sum_{i,j,p,q} P\left(\sup_{A \in N(A_0)} \left| \frac{\partial^4 t_M(A)}{\partial A_{T,i} \partial A_{T,j} \partial A_{T,p} \partial A_{T,q}} \right| > C_1 \log(T) \right) \\
& + C_2 P\left[\|S_T\| > C_1 \log T\right] + C_2 P\left[\|G_T - G_0\| > C_1 \left(\eta_T / \log^2 T\right)^{2/3} \right] \\
& + C_2 P\left[\left\|\hat{\Omega}_T - \Omega_0\right\| > C_1 \left(\eta_T / \log^2 T\right)^{1/3} \right] \\
=\ & o(\eta_T) + C_2 P\left[\|S_T\| > C_1 \log T\right] \\
& + C_2 P\left[\|G_T - G_0\| > C_1 \left(\eta_T / \log^2 T\right)^{2/3} \right] \\
& + C_2 P\left[\left\|\hat{\Omega}_T - \Omega_0\right\| > C_1 \left(\eta_T / \log^2 T\right)^{1/3} \right] \\
=\ & o(\eta_T),
\end{aligned}
$$

where the first $o(\eta_T)$ term holds because $\sup_{A \in N(A_0)} \left| \frac{\partial^4 t_M(A)}{\partial A_{T,i} \partial A_{T,j} \partial A_{T,p} \partial A_{T,q}} \right|$ is a bounded and deterministic constant. The second $o(\eta_T)$ term holds because of Lemma A.3(a), (b) and (f).

Using similar arguments, we can show that the other terms in (A.6) satisfy a similar inequality. Hence

$$P\left(|\xi_T| > \eta_T\right) = o(\eta_T),$$

as desired. ■

Lemma A.5 *Let* $\Sigma_0 = \left(G_0' \Omega_0^{-1} G_0\right)^{-1}$, $\sigma_0 = \left(\mathcal{R}' \Sigma_0 \mathcal{R}\right)^{1/2}$,

$$\Lambda_{10} = \Sigma_0 \mathcal{R},\ \ \Lambda_{20} = \Omega_0^{-1} G_0 \Sigma_0,$$

and

$$
\begin{aligned}
\Theta_{10} &= \Omega_0^{-1} G_0 \Sigma_0 \mathcal{R}, \Theta_{20} = \Omega_0^{-1}, \\
\Theta_{30} &= \Omega_0^{-1} G_0 \Sigma_0 G_0' \Omega_0^{-1}, \\
\Theta_{40} &= (\Theta_{30} - \Theta_{20} - \frac{1}{2\sigma_0^2} \Theta_{10} \Theta_{10}').
\end{aligned}
$$

Then

$$\mathbf{a}_h = vec\left(Q_a\right),\ \mathbf{b}_h = vec\left(Q_b\right),\ \mathbf{c}_h = vec\left(Q_c D_{d_2}\right),\ \mathbf{d}_h = vec\left(\left(D_{d_2}' \otimes I_{d_2}\right) Q_d D_{d_2}\right),$$

where

$$Q_a = \frac{1}{\sigma_0}\Theta_{10},$$

$$Q_b = \frac{1}{\sigma_0}\left[\Lambda_{10}' \otimes \left(\Theta_{20} - \Theta_{30}\right)\right] - \frac{1}{\sigma_0}\left(\Lambda_{20} \otimes \Theta_{10}'\right) + \frac{1}{\sigma_0^3}\left[\left(\Theta_{10}\Lambda_{10}'\right) \otimes \Theta_{10}'\right],$$

$$Q_c = \frac{1}{\sigma_0}\left[\Theta_{10}' \otimes \Theta_{40}\right],$$

and

$$\begin{aligned} Q_d &= \frac{1}{4\sigma_0^3}\left\{\left[vec\left(\Theta_{10}'\right) \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta_{10}'\right)\right]\right\} \\ &+ \frac{1}{2\sigma_0}\left\{\left[vec\left(\Theta_{10}'\right) \otimes \Theta_{40} \otimes \left(\Theta_{30} - \Theta_{20}\right)\right]\right\} \\ &+ \frac{1}{2\sigma_0}\left\{K_{d_2,d_2^2}\left[vec\left(\Theta_{40}\right) \otimes \left(\Theta_{30} - \Theta_{20}\right) \otimes \Theta_{10}'\right]\right\} \\ &- \frac{1}{2\sigma_0^2}\left\{vec\left(\Theta_{10}' \otimes \left[\frac{1}{2\sigma_0}\Theta_{40} - \frac{1}{2\sigma_0^3}\Theta_{10}'\Theta_{10}\right]\right)\left[\Theta_{10}' \otimes \Theta_{10}'\right]\right\}. \end{aligned}$$

Proof. First, it is easy to see that

$$\mathbf{a}_h = \Omega_0^{-1} G_0 \Sigma_0 \mathcal{R}\left(\mathcal{R}'\Sigma_0\mathcal{R}\right)^{-1/2} = \Theta_{10}/\sigma_0.$$

Second, we treat $\partial t_M(A_T)/\partial S_T$ as a function of G_T and compute its first order differential as follows (e.g. Abadir and Magnus (2005, Chapter 13)):

$$\begin{aligned} d\left[\partial t_M(A_T)/\partial S_T\right] &= d\left[\hat{\Omega}_T^{-1} G_T\left(G_T'\hat{\Omega}_T^{-1}G_T\right)^{-1}\mathcal{R}\left(\mathcal{R}'\left(G_T'\hat{\Omega}_T^{-1}G_T\right)^{-1}\mathcal{R}\right)^{-1/2}\right] \\ &= E_1 + E_2 + E_3 + E_4 + E_5, \end{aligned}$$

where

$$\begin{aligned} E_1 &= \hat{\Omega}_T^{-1}\left(dG_T\right)\left(G_T'\hat{\Omega}_T^{-1}G_T\right)^{-1}\mathcal{R}\left(\mathcal{R}'\left(G_T'\hat{\Omega}_T^{-1}G_T\right)^{-1}\mathcal{R}\right)^{-1/2} \\ &= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(\mathcal{R}'\hat{\Sigma}_T \otimes \hat{\Omega}_T^{-1}\right) dvec(G_T), \end{aligned}$$

$$\begin{aligned} E_2 &= -\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left[dG_T\right]\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} \\ &= -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\left(\mathcal{R}'\hat{\Sigma}_T\right) \otimes \left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] dvec(G_T), \end{aligned}$$

$$\begin{aligned} E_3 &= -\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\left[dG_T'\hat{\Omega}_T^{-1}G_T\right]\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} \\ &= -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left\{\left[\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right] \otimes \left[\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\right]\right\} dvec(G_T') \\ &= -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left\{\left[\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right] \otimes \left[\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\right]\right\} K_{d_2 d_1} dvec(G_T) \\ &= -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left\{\left[\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\right] \otimes \left[\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right]\right\} dvec(G_T), \end{aligned}$$

$$\begin{aligned} E_4 &= \frac{1}{2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\mathcal{R}'\hat{\Sigma}_T\left[G_T'\hat{\Omega}_T^{-1}dG_T\right]\hat{\Sigma}_T\mathcal{R} \\ &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left[\mathcal{R}'\hat{\Sigma}_T\otimes\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right]dvec(G_T), \end{aligned}$$

and

$$\begin{aligned} E_5 &= \frac{1}{2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\mathcal{R}'\hat{\Sigma}_T\left[dG_T'\hat{\Omega}_T^{-1}G_T\right]\hat{\Sigma}_T\mathcal{R} \\ &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left[\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_T\right]K_{d_2d_1}dvec(G_T) \\ &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left[\mathcal{R}'\hat{\Sigma}_T\otimes\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right]dvec(G_T). \end{aligned}$$

So

$$\begin{aligned} &\frac{\partial^2 t_M(A_T)}{\partial S_T\partial vec(G)'} = \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\mathcal{R}'\hat{\Sigma}_T\otimes\hat{\Omega}_T^{-1}\right] \\ &-\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left\{\left[\mathcal{R}'\hat{\Sigma}_T\right]\otimes\left[\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right]\right\} \\ &-\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left\{\left[\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\right]\otimes\left[\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right]\right\} \\ &+\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T\right)\otimes\left(\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right), \end{aligned}$$

and

$$\mathbf{b}_h = vec(Q_b).$$

Third, we treat $\partial t_M(A_T)/\partial S_T$ as a function of $\hat{\Omega}_T$ and compute its first order differential:

$$\begin{aligned} d\left[\partial t_M(A_T)/\partial S_T\right] &= d\left[\hat{\Omega}_T^{-1}G_T\left(G_T'\hat{\Omega}_T^{-1}G_T\right)^{-1}\mathcal{R}\left(\mathcal{R}'\left(G_T'\hat{\Omega}_T^{-1}G_T\right)^{-1}\mathcal{R}\right)^{-1/2}\right] \\ &= E_6+E_7+E_8, \end{aligned}$$

where

$$\begin{aligned} E_6 &= -\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} \\ &= -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right]dvec\left(\hat{\Omega}_T\right), \end{aligned}$$

$$\begin{aligned} E_7 &= \hat{\Omega}_T^{-1}G_T\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}d\hat{\Omega}_T\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} \\ &= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right]dvec\left(\hat{\Omega}_T\right), \end{aligned}$$

$$\begin{aligned} E_8 &= -\frac{1}{2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}d\hat{\Omega}_T\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R} \\ &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\left[\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right]dvec\left(\hat{\Omega}_T\right) \\ &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left[\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right)\right]dvec\left(\hat{\Omega}_T\right) \\ &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left[\left(\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_TG_T'\hat{\Omega}_T^{-1}\right)\right]K_{d_2,d_2}dvec\left(\hat{\Omega}_T\right). \end{aligned}$$

Therefore, using $dvec\left(\hat{\Omega}_T\right) = D_{d_2} dvech\left(\hat{\Omega}_T\right)$ and $K_{d_2,d_2} D_{d_2} = D_{d_2}$, we have

$$
\begin{aligned}
&\frac{\partial t_M(A_T)}{\partial S_T \partial vech(\hat{\Omega}_T)'} = -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \hat{\Omega}_T^{-1}\right] D_{d_2} \\
&+\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \left(\hat{\Omega}_T^{-1} G_T \hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] D_{d_2} \\
&-\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \left(\hat{\Omega}_T^{-1} G_T \hat{\Sigma}_T \mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] D_{d_2} \\
= \;& J_1 + J_2 + J_3.
\end{aligned}
$$

As a result,

$$
\begin{aligned}
\mathbf{c}_h &= -\frac{1}{\sigma_0} vec\left\{\left[\Theta_{10}'\otimes(\Theta_{20}-\Theta_{30})\right] D_{d_2}\right\} - \frac{1}{\sigma_0^3} vec\left\{\left(\Theta_{10}'\otimes \Theta_{10}\Theta_{10}'\right) D_{d_2}\right\} \\
&= -\frac{1}{\sigma_0} vec\left\{\Theta_{10}'\otimes\left(\Theta_{20}-\Theta_{30}+\frac{1}{2\sigma_0^2}\Theta_{10}\Theta_{10}'\right) D_{d_2}\right\} \\
&= \frac{1}{\sigma_0} vec\left[\left(\Theta_{10}'\otimes\Theta_{40}\right) D_{d_2}\right].
\end{aligned}
$$

Finally, we treat J_1, J_2 and J_3 as functions of $\hat{\Omega}_T$ and compute their differentials. Using the formula (e.g. Magnus and Neudecker (1999, theorem 3.10, page 47)

$$
vec(A\otimes B) = \left(I_n \otimes K_{q,m}\otimes I_p\right)\left(vec\left(A\right)\otimes vec\left(B\right)\right),
$$

where A is an $m\times n$ matrix and B is a $p \times q$ matrix, we can show that

$$
dvec\left(J_1\right) = J_{11} + J_{12} + J_{13} + J_{14}
$$

where

$$
\begin{aligned}
J_{11} &= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right] D_{d_2}\right\} \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right] \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)\left(I_{d_2}\otimes K_{d_2,1}\otimes I_{d_2}\right)\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes vec\left(\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\right] \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)\left\{vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \left[\left(\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right) vec\left(d\hat{\Omega}_T\right)\right]\right\} \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)\left\{vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes \left(\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right)\right\} vec\left(d\hat{\Omega}_T\right),
\end{aligned}
$$

$$
\begin{aligned}
J_{12} &= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} vec\left\{\left[\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\otimes \hat{\Omega}_T^{-1}\right] D_{d_2}\right\} \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) vec\left\{\left[\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\otimes \hat{\Omega}_T^{-1}\right]\right\} \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)\left\{\left[\left(\hat{\Omega}_T^{-1}\otimes \mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right) vec\left(d\hat{\Omega}_T\right)\otimes vec\left(\hat{\Omega}_T^{-1}\right)\right]\right\} \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) K_{d_2,d_2^2}\left[vec\left(\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}\otimes \mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right) vec\left(d\hat{\Omega}_T\right)\right] \\
&= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) K_{d_2,d_2^2}\left[vec\left(\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}\otimes \mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] vec\left(d\hat{\Omega}_T\right),
\end{aligned}
$$

$$\begin{aligned} & J_{13} \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left[\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right]D_{d_2} \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)K_{d_2,d_2^2}\left[vec\left(\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]vec\left(d\hat{\Omega}_T\right), \end{aligned}$$

and

$$\begin{aligned} & J_{14} \\ = & \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right)D_{d_2}\right]\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\right) \\ = & \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right)D_{d_2}\right]\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]vec\left(d\hat{\Omega}_T\right) \\ = & \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right)vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right) \\ & \times\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]vec\left(d\hat{\Omega}_T\right). \end{aligned}$$

Similarly,

$$dvec\left(J_2\right) = J_{21}+J_{22}+J_{23}+J_{24}+J_{25}+J_{26},$$

where

$$\begin{aligned} J_{21} = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\right]D_{d_2} \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]vec(d\Omega_T), \end{aligned}$$

$$\begin{aligned} J_{22} = & \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2} \\ = & \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\ & \times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right]vec(d\Omega_T), \end{aligned}$$

$$\begin{aligned} J_{23} = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2} \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}\right)\right]vec(d\Omega_T), \end{aligned}$$

$$\begin{aligned} J_{24} = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2}\right\} \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\ & \times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\otimes vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)K_{d_2,d_2^2}\left[vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\right] \\ = & -\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right)K_{d_2,d_2^2} \\ & \times\left[vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]vec(d\Omega_T), \end{aligned}$$

$$\begin{aligned} J_{25} &= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2} vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right.\right. \\ &\quad \left.\left.\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] D_{d_2}\right\} \\ &= \left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-1/2}\left(D_{d_2}'\otimes I_{d_2}\right) K_{d_2,d_2^2} \\ &\quad \times\left[vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] vec(d\Omega_T), \end{aligned}$$

and

$$\begin{aligned} J_{26} &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] D_{d_2}\right\} \\ &\quad \times\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\right) \\ &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] \\ &\quad \times\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] vec\left(d\hat{\Omega}_T\right). \end{aligned}$$

By the same argument, we have

$$dvec\left(J_3\right)=\sum_{i=1}^{7} J_{3i},$$

where

$$\begin{aligned} J_{31} &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} \\ &\quad \times vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\right] D_{d_2}\right\} \\ &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\ &\quad \times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\right] \\ &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\ &\quad \times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\hat{\Omega}_T^{-1}\otimes\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right] vec\left(d\hat{\Omega}_T\right), \end{aligned}$$

$$\begin{aligned} J_{32} &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} \\ &\quad \times vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] D_{d_2}\right\} \\ &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\ &\quad \times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] vec\left(d\hat{\Omega}_T\right), \end{aligned}$$

$$\begin{aligned} J_{33} &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} \\ &\quad \times vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] D_{d_2}\right\} \\ &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\ &\quad \times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] vec\left(d\hat{\Omega}_T\right), \end{aligned}$$

$$
\begin{aligned}
J_{34} &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2}\right\} \\
&= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\
&\quad\times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\hat{\Omega}_T^{-1}\right]vec\left(d\hat{\Omega}_T\right),
\end{aligned}
$$

$$
\begin{aligned}
J_{35} &= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2}\right\} \\
&= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\
&\quad\times\left[vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\otimes vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] \\
&= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right)K_{d_2,d_2^2} \\
&\quad\times\left[vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes vec\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}\right)\right] \\
&= \frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right)K_{d_2,d_2^2} \\
&\quad\times\left[vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right]vec\left(d\hat{\Omega}_T\right),
\end{aligned}
$$

$$
\begin{aligned}
J_{36} &= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2} \\
&\quad\times vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2}\right\} \\
&= -\frac{1}{2}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-3/2}\left(D_{d_2}'\otimes I_{d_2}\right) \\
&\quad\times\left[vec\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\otimes\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right]vec\left(d\hat{\Omega}_T\right),
\end{aligned}
$$

and

$$
\begin{aligned}
J_{37} &= \frac{3}{4}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-5/2} vec\left\{\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]D_{d_2}\right\} \\
&\quad\times\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\left(d\hat{\Omega}_T\right)\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\right) \\
&= \frac{3}{4}\left(\mathcal{R}'\hat{\Sigma}_T\mathcal{R}\right)^{-5/2}\left(D_{d_2}'\otimes I_{d_2}\right)vec\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\hat{\Omega}_T^{-1}G_T\hat{\Sigma}_T\mathcal{R}\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right] \\
&\quad\times\left[\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\otimes\left(\mathcal{R}'\hat{\Sigma}_T G_T'\hat{\Omega}_T^{-1}\right)\right]vec\left(d\hat{\Omega}_T\right).
\end{aligned}
$$

Summing up the above expressions for J_{pq} and evaluating the result at $A_T = A_0$, we obtain the stated formula for d_h. ■

Lemma A.6 *Let*

$$
\begin{aligned}
h_T &= \mathbf{a}_h' S_T + \mathbf{b}_h'\left[vec\left(G_T - G_0\right)\otimes S_T\right] + \mathbf{c}_h'\left[vech\left(\check{\Omega}_T - \Omega_0\right)\otimes S_T\right] \\
&\quad + \mathbf{d}_h'\left[vech(\check{\Omega}_T - \bar{\Omega}_T)\otimes vech(\check{\Omega}_T - \bar{\Omega}_T)\otimes S_T\right].
\end{aligned}
$$

Then

$$
P(\|g_T - h_T\| \geq \eta_T) = o\left(\eta_T\right).
$$

Proof. In view of

$$g_T - h_T = \mathbf{c}_h' R_{1T} + \mathbf{d}_h' R_{2T},$$

where

$$R_{1T} = vech\left(\hat{\Omega}_T - \check{\Omega}_T\right) \otimes S_T,$$

and

$$R_{2T} = \left[vech(\hat{\Omega}_T - \bar{\Omega}_T) \otimes vech(\hat{\Omega}_T - \bar{\Omega}_T) - vech(\check{\Omega}_T - \bar{\Omega}_T) \otimes vech(\check{\Omega}_T - \bar{\Omega}_T)\right] \otimes S_T,$$

it suffices to show

$$P(\|R_{iT}\| \geq \eta_T) = o\left(\eta_T\right) \text{ for } i = 1, 2.$$

We prove the above bound for $i = 1$ and the proof for $i = 2$ is similar.

$$\begin{aligned} P(\|R_{1T}\| &> \eta_T) = P\left(\left\|vech\left(\hat{\Omega}_T - \check{\Omega}_T\right) \otimes S_T\right\| > \eta_T\right) \\ &\leq P\left(\left\|vech\left(\hat{\Omega}_T - \check{\Omega}_T\right)\right\| \|S_T\| > \eta_T\right) \\ &\leq P\left(\left\|vech\left(\hat{\Omega}_T - \check{\Omega}_T\right)\right\| > \eta_T / \log\left(T\right)\right) + P\left(\|S_T\| > \log T\right) \\ &= P\left(\left\|vech\left(\hat{\Omega}_T - \check{\Omega}_T\right)\right\| > \eta_T / \log\left(T\right)\right) + o\left(\eta_T\right) \\ &= o(\eta_T), \end{aligned}$$

by Lemma A.3(e). ■

Lemma A.7 *The cumulants of h_T satisfy*

(a) $\kappa_1(h_T) = T^{-1/2}\kappa_{1,\infty} + O(T^{-1/2}M^{-q}) + o\left(M/T\right),$
(b) $\kappa_2(h_T) = 1 + \rho_{1,\infty} M^{-q} + \left(\rho_{2,\infty} + \kappa_{2,\infty}\right)\left(M/T\right) + o\left(M/T\right),$
(c) $\kappa_3(h_T) = \left(\kappa_{3,\infty} - 3\kappa_{1,\infty}\right)/\sqrt{T} + o(M^{-q}) + o\left(M/T\right),$
(d) $\kappa_4\left(h_T\right) = \left(\kappa_{4,\infty} - 6\kappa_{2,\infty}\right)\left(M/T\right) + o\left(M/T\right),$

where $\rho_{j,\infty}$ and $\kappa_{j,\infty}$ are finite constants defined by

$$\begin{aligned} \rho_{1,\infty} &= \lim_{T\to\infty} M^q E 2\mathbf{a}_h' S_T \mathbf{c}_h' [vech\left(\bar{\Omega}_T - \Omega_0\right) \otimes S_T], \\ \rho_{2,\infty} &= \lim_{T\to\infty} \frac{T}{M} E 2\mathbf{a}_h' S_T \mathbf{c}_h' [vech\left(\check{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T], \end{aligned} \tag{A.7}$$

$$\kappa_{1,\infty} = \lim_{T\to\infty} \sqrt{T}\mathbf{b}_h'\left[E\left(vec\left(G_T - G_0\right) \otimes S_T\right)\right] + \lim_{T\to\infty} \sqrt{T}\mathbf{c}_h'\left[E\left(vech\left(\check{\Omega}_T - \Omega_0\right) \otimes S_T\right)\right],$$

$$\begin{aligned} \kappa_{2,\infty} &= 2\lim_{T\to\infty} \frac{T}{M} E\left(\mathbf{a}_h' S_T\right)\left\{\mathbf{d}_h'\left[vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T\right]\right\} \\ &\quad + \lim_{T\to\infty} \frac{T}{M} E\left[\mathbf{c}_h'\left(vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T\right)\right]^2, \end{aligned}$$

$$\begin{aligned} \kappa_{3,\infty} &= \lim_{T\to\infty} \sqrt{T} E\left(\mathbf{a}_h' S_T\right)^3 + 3\lim_{T\to\infty} \sqrt{T} E\left(\mathbf{a}_h' S_T\right)^2 \left\{\mathbf{b}_h'\left[vec\left(G_T - G_0\right) \otimes S_T\right]\right\} \\ &\quad + 3\lim_{T\to\infty} \sqrt{T} E\left(\mathbf{a}_h' S_T\right)^2 \left\{\mathbf{c}_h'[vech\left(\check{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T]\right\}, \end{aligned}$$

$$
\begin{aligned}
\kappa_{4,\infty} &= 4\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}_h' S_T\right)^3\left\{\mathbf{d}_h'\left[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T\right]\right\} \\
&\quad +6\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}_h' S_T\right)^2\left\{\mathbf{c}_h'[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]\right\}^2.
\end{aligned}
$$

Proof. To prove the lemma, we need to show that: (i) $\rho_{j,\infty}$ and $\kappa_{j,\infty}$ are finite; and (ii) the asymptotic expansions of the cumulants hold.

Part (a) Using the relationship $\kappa_1(h_T) = Eh_T$, we have

$$
\begin{aligned}
&T^{1/2}\kappa_1(h_T) = \mathbf{b}_h'\left\{T^{1/2}E\left[vec\left(G_T-G_0\right)\otimes S_T\right]\right\}+\mathbf{c}_h'\left\{T^{1/2}E\left[vech\left(\check{\Omega}_T-\Omega_0\right)\otimes S_T\right]\right\} \\
&+\mathbf{d}_h'\left\{T^{1/2}E\left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right]\right\}.
\end{aligned}
$$

Consider each term in the above expression in turn. First

$$
\begin{aligned}
& T^{1/2}E\left[vec(G_T-G_0)\otimes S_T\right] \\
=\; & T^{1/2}vec\left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t'\right)\otimes\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T} v_s\right) \\
=\; & \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}E\left[vec(w_t)\otimes v_s\right]=\sum_{t=-\infty}^{\infty}E\left[vec(w_0)\otimes v_t\right]+O\left(\frac{1}{T}\right),
\end{aligned}
$$

where the infinite sum is finite because $(vec(w_t)', v_t')'$ is a strong mixing process. Second, we write

$$
\begin{aligned}
& \mathbf{c}_h'\left[T^{1/2}Evech\left(\check{\Omega}_T-\Omega_0\right)\otimes S_T\right] \\
=\; & \mathbf{c}_h'\left[T^{1/2}Evech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T\right]+\mathbf{c}_h'\left[T^{1/2}Evech\left(\check{\Omega}_T-\tilde{\Omega}_T\right)\otimes S_T\right]. \qquad \text{(A.8)}
\end{aligned}
$$

For the first term in (A.8), we have

$$
\begin{aligned}
& \left[T^{1/2}Evech\left(\tilde{\Omega}-\bar{\Omega}_T\right)\otimes S_T\right] \\
=\; & T^{1/2}E\sum_{j=0}^{M}k_j vech(\tilde{\Gamma}_j-\Gamma_j)\otimes S_T+T^{1/2}E\sum_{j=-M}^{-1}k_j vech(\tilde{\Gamma}_j-\Gamma_j)\otimes S_T.
\end{aligned}
$$

Note that

$$
\begin{aligned}
& T^{1/2}E\sum_{j=0}^{M}k_j vech(\tilde{\Gamma}_j-\Gamma_j)\otimes S_T \\
=\; & T^{1/2}E\sum_{j=0}^{M}k_j vech\left(\frac{1}{T}\sum_{t=1}^{T}v_{t+j}v_t'\right)\otimes\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}v_s\right) \\
=\; & \sum_{j=0}^{M}k_j\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left[vech\left(v_{t+j}v_t'\right)\otimes v_s\right] \\
=\; & \sum_{j=0}^{M}k_j\sum_{i=-T+1}^{T-1}Evech\left(v_0 v_{-j}'\right)\otimes v_i+O\left(\frac{M}{T}\right) \\
=\; & \sum_{j=0}^{M}\sum_{i=-T+1}^{T-1}Evech\left(v_0 v_{-j}'\right)\otimes v_i+O(M^{-q})+O\left(\frac{M}{T}\right) \\
=\; & \sum_{j=0}^{\infty}\sum_{i=-\infty}^{\infty}Evech\left(v_0 v_{-j}'\right)\otimes v_i+O(M^{-q})+O\left(\frac{M}{T}\right).
\end{aligned}
$$

By symmetric arguments,

$$\begin{aligned}
& T^{1/2} E \sum_{j=-M}^{-1} k_j vech(\tilde{\Gamma}_j - \Gamma_j) \otimes S_T \\
= & \sum_{j=-\infty}^{-1} \sum_{i=-\infty}^{\infty} Evech\left(v_0 v_{-j}'\right) \otimes v_i + O\left(\frac{M}{T}\right) + O\left(M^{-q}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
& \mathbf{c}_h' \left[T^{1/2} Evech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T\right] && \text{(A.9)} \\
= & \mathbf{c}_h' \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} Evech\left(v_0 v_{-j}'\right) \otimes v_i + O\left(\frac{M}{T}\right) + O\left(M^{-q}\right), && \text{(A.10)}
\end{aligned}$$

where the double infinite sum is finite because v_t is a strong mixing process.

For the second term in (A.8), we have

$$\begin{aligned}
& T^{1/2} E\left[vech\left(\check{\Omega}_T - \tilde{\Omega}_T\right) \otimes S_T\right] \\
= & T^{1/2} E\left[vech\left(\nabla \tilde{\Omega}_T \left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T\right] \\
& + T^{1/2} E\left[vech\left(\left(\check{\beta}_T - \beta_0\right)' \nabla^2 \tilde{\Omega}_T \left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T\right] \\
= & T^{1/2} E\left[vech\left(\nabla \tilde{\Omega}_T \left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T\right] + O\left(T^{-1/2}\right),
\end{aligned}$$

By Lemma A.2(e) and the Cauchy inequality, we know that $T^{1/2} E\left[vech\left(\nabla \tilde{\Omega}_T \left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T\right]$ is bounded. Following the same argument that leads to (A.10), we have

$$\begin{aligned}
& T^{1/2} E\left[vech\left(\nabla \tilde{\Omega}_T \left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T\right] \\
= & \lim_{T \to \infty} T^{1/2} E\left[vech\left(\nabla \tilde{\Omega}_T \left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T\right] + O(M^{-q}) + O\left(\frac{M}{T}\right).
\end{aligned}$$

Therefore, we have shown that

$$\begin{aligned}
& \mathbf{c}_h' \left[T^{1/2} Evech\left(\check{\Omega}_T - \Omega_0\right) \otimes S_T\right] \\
= & \lim_{T \to \infty} \mathbf{c}_h' \left[T^{1/2} Evech\left(\check{\Omega}_T - \Omega_0\right) \otimes S_T\right] + O(M^{-q}) + O\left(\frac{M}{T}\right) + O\left(\frac{1}{\sqrt{T}}\right).
\end{aligned}$$

Finally, using Lemma A.2 (c) and (d) and the Hölder inequality, we have

$$\begin{aligned}
& (T/M)\, Evech(\check{\Omega}_T - \bar{\Omega}_T) \otimes vech(\check{\Omega}_T - \bar{\Omega}_T) \otimes S_T \\
= & (T/M)\, Evech(\tilde{\Omega}_T - \bar{\Omega}_T) \otimes vech(\tilde{\Omega}_T - \bar{\Omega}_T) \otimes S_T + o(1) \\
= & CM^{-1} T^{-3/2} \sum_{s \le t \le u}^{T} \sum_{i=-M}^{M} \sum_{j=-M}^{M} E\left(vech\left(v_{s+i} v_s' - \Gamma_i\right) \otimes vech\left(v_{t+j} v_t' - \Gamma_j\right) \otimes v_u\right) + o\left(1\right).
\end{aligned}$$

To calculate the order of magnitude of the above quantity, we can assume, without loss of generality, that v_t is a scalar as the vector case can be reduced to the scalar case by considering each element of the vector. We split the sum over s, t, u into three sums over $\mathcal{S}_1, \mathcal{S}_2$ and $\mathcal{S}_3$ where

$$\mathcal{S}_1 = \left\{s, t, u : s \le t \le u, \;\; \max\left(t - s, u - t\right) \ge 3M\right\},$$

$$\mathcal{S}_2 = \left\{ s,t,u : s \le t \le u, \ \ \max(t-s, u-t) < 3\sqrt{M} \right\},$$

and

$$\mathcal{S}_3 = \left\{ s,t,u : s \le t \le u, \ 3\sqrt{M} < \max(t-s, u-t) < 3M \right\}.$$

To bound the first sum, we use the property of strong mixing processes that if $X \in \mathcal{F}_x$ and $Y \in \mathcal{F}_y$, then

$$|E(XY)| \le (EX)(EY) + 8\left[\alpha(\mathcal{F}_x, \mathcal{F}_y)\right]^{1/r} \left(E|X|^p\right)^{1/p} \left(E|Y|^q\right)^{1/q} \tag{A.11}$$

for any $p, q, r \ge 1$, and

$$p^{-1} + q^{-1} + r^{-1} = 1.$$

See, e.g. Doukhan (1995, Section 1.2.2). Let $p = q = 2+\epsilon$ and $r = 1 + 2/\epsilon$ for some small $\epsilon > 0$, then when $t - s \ge 3M$,

$$\begin{aligned}
& \left| E\left(\sum_{i=-M}^{M} (v_{s+i}v_s - \Gamma_i) \sum_{j=-M}^{M} (v_{t+j}v_t - \Gamma_j)\, v_u \right) \right| \\
\le\ & 8\alpha_M^{\epsilon/(2+\epsilon)} \left(E\left| \sum_{i=-M}^{M} (v_{s+i}v_s - \Gamma_i) \right|^p \right)^{1/p} \left(E\left| \sum_{j=-M}^{M} (v_{t+j}v_t - \Gamma_j)\, v_u \right|^q \right)^{1/q} \\
\le\ & 8\alpha_M^{\epsilon/(2+\epsilon)} \left[\sum_{i=-M}^{M} \left(E|v_{s+i}v_s - \Gamma_i|^p\right)^{1/p} \right] \left[\sum_{i=-M}^{M} \left(E|(v_{t+j}v_t - \Gamma_j)\, v_u|^q\right)^{1/q} \right] \\
\le\ & 8\alpha_M^{\epsilon/(2+\epsilon)} \left[\sum_{i=-M}^{M} \left(E|v_{s+i}v_s - \Gamma_i|^p\right)^{1/p} \right] \left[\sum_{i=-M}^{M} \left(E|(v_{t+j}v_t - \Gamma_j)|^q\right)^{1/q} \right] \left(E|v_u|^q\right)^{1/q} \\
=\ & O\left(\alpha_M^{\epsilon/(2+\epsilon)} M^2 \right)
\end{aligned}$$

where the last line follows because both sums are of order $O(M)$.

When $u - t \ge 3M$,

$$\begin{aligned}
& \left| E\left(\sum_{i=-M}^{M} (v_{s+i}v_s - \Gamma_i) \sum_{j=-M}^{M} (v_{t+j}v_t - \Gamma_j)\, v_u \right) \right| \\
\le\ & 8\alpha_{2M}^{\epsilon/(2+\epsilon)} \left(E\left| \sum_{i=-M}^{M} (v_{s+i}v_s - \Gamma_i) \sum_{j=-M}^{M} (v_{t+j}v_t - \Gamma_j) \right|^p \right)^{1/p} \left(E|v_u|^q\right)^{1/q} \\
\le\ & 8\alpha_{2M}^{\epsilon/(2+\epsilon)} \left(E\left| \sum_{i=-M}^{M} (v_{s+i}v_s - \Gamma_i) \right|^{2p} \right)^{1/p} \left(E|v_u|^q\right)^{1/q} \\
\le\ & 8\alpha_{2M}^{\epsilon/(2+\epsilon)} \left(\sum_{i=-M}^{M} \left(E|v_{s+i}v_s - \Gamma_i|^{2p} \right)^{1/(2p)} \right)^2 \left(E|v_u|^q\right)^{1/q} \\
=\ & O\left(\alpha_{2M}^{\epsilon/(2+\epsilon)} M^2 \right).
\end{aligned}$$

Therefore

$$
\begin{aligned}
&\sum_{s,t,u\in\mathcal{S}_1} E\left(\sum_{i=-M}^{M}\left(v_{s+i}v_s-\Gamma_i\right)\sum_{j=-M}^{M}\left(v_{t+j}v_t-\Gamma_j\right)v_u\right)\\
=\;& O\left(MT^{3/2}\alpha_M^{\epsilon/(2+\epsilon)}\right)=O\left(MT^{3/2}\exp\left(-\frac{\epsilon}{2+\epsilon}M\right)\right)=O(MT). \qquad \text{(A.12)}
\end{aligned}
$$

To bound the second sum, we note that the sum involves the forms:

$$
\sum_{s,t,u\in\mathcal{S}_2} E\left(\sum_{i=-M}^{M}\sum_{j=-M}^{M}\left(v_{s+i}v_s\right)\Gamma_j v_u\right) \text{ and } \sum_{s,t,u\in\mathcal{S}_2} E\left(\sum_{i=-M}^{M}\sum_{j=-M}^{M}\left(v_{s+i}v_s\right)\left(v_{t+j}v_t\right)v_u\right).
$$

Since each expectation is bounded and $\max\left(t-s,u-t\right)<3\sqrt{M}$, the sum over $\mathcal{S}_2$ satisfies:

$$
\sum_{s,t,u\in\mathcal{S}_2} E\left(\sum_{i=-M}^{M}\left(v_{s+i}v_s-\Gamma_i\right)\sum_{j=-M}^{M}\left(v_{t+j}v_t-\Gamma_j\right)v_u\right)\le T\left(3\sqrt{M}\right)^2=O(MT). \qquad \text{(A.13)}
$$

The third sum involves similar forms:

$$
\sum_{s,t,u\in\mathcal{S}_3} E\left(\sum_{i=-M}^{M}\sum_{j=-M}^{M}\left(v_{s+i}v_s\right)\Gamma_j v_u\right) \text{ and } \sum_{s,t,u\in\mathcal{S}_3} E\left(\sum_{i=-M}^{M}\sum_{j=-M}^{M}\left(v_{s+i}v_s\right)\left(v_{t+j}v_t\right)v_u\right). \qquad \text{(A.14)}
$$

For the first term in (A.14), we have

$$
\begin{aligned}
&\sum_{s,t,u\in\mathcal{S}_3} E\left(\sum_{i=-M}^{M}\sum_{j=-M}^{M}\left(v_{s+i}v_s\right)\Gamma_j v_u\right) \qquad \text{(A.15)}\\
=\;& CT\sum_{s,t\in\mathcal{S}_3} E\left(\sum_{i=-M}^{M}\left(v_{s+i}v_s\right)v_0\right)\\
=\;& CMT\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty} E\left(v_sv_tv_0\right)\left(1+o(1)\right)=O(MT).
\end{aligned}
$$

For the second term in (A.14), we need to consider the two cases: $3M>u-t>3\sqrt{M}$ or $3M>t-s>3\sqrt{M}$. We consider the first case here as the second case follows from the same argument. When $3M>u-t>3\sqrt{M}$, simple combinatorial argument shows that for any i and j, the largest gap between the five integers $s,s+i,t,t+j,u$ is greater than $\sqrt{M}$. Therefore

$$
\sum_{s,t,u\in\mathcal{S}_3} E\left(\sum_{i=-M}^{M}\sum_{j=-M}^{M}\left(v_{s+i}v_s\right)\left(v_{t+j}v_t\right)v_u\right)=O(MT)+O\left(\alpha_{\sqrt{M}}^{\epsilon/(2+\epsilon)}M^2\right) \qquad \text{(A.16)}
$$

for fixed $\epsilon>0$. Here we have used inequality (A.11) as before. In the above expression, the $O(MT)$ term arises because the term $(EX)(EY)$ in (A.11) leads to terms similar to that given in (A.15).

Combining (A.12), (A.13) and (A.16) yields

$$
(T/M)\,E\left(vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right)=O(T^{-1/2})+o(1)=o(1).
$$

We have thus proved part (a) of the lemma.

Part (b) Using the result in part (a), we get

$$\begin{aligned}
& (T/M)\left(\kappa_2(h_T)-1\right) \\
= \;& (T/M)\left(E\left(h_T^2\right)-1-\left[T^{-1/2}\kappa_{1,\infty}+O(T^{-1/2}M^{-q})+o\left(M/T\right)\right]^2\right) \\
= \;& (T/M)\left(E\left(h_T^2\right)-1\right)+O\left(1/M\right).
\end{aligned}$$

We now evaluate the order of magnitude of $E\left(h_T\right)^2-1$. In view of

$$E\mathbf{a}_h' S_T S_T' \mathbf{a}_h = 1+O\left(1/T\right),$$

we have, up to an order of $O\left(1/T\right)$,

$$\begin{aligned}
E\left(h_T^2\right)-1 \;=\;& 2E\mathbf{a}_h' S_T \mathbf{b}_h' vec\left(G_T-G_0\right)\otimes S_T \\
& +2E\mathbf{a}_h' S_T \mathbf{c}_h' [vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T \\
& +2E\mathbf{a}_h' S_T \mathbf{c}_h' [vech\left(\check{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T \\
& +2E\mathbf{a}_h' S_T \mathbf{d}_h' \left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right] \\
& +E\left[\mathbf{c}_h'\left(vech\left(\check{\Omega}_T-\Omega_0\right)\otimes S_T\right)\right]^2 \\
& +O\left\{E\left[\mathbf{b}_h'\left(vec\left(G_T-G_0\right)\otimes S_T\right)\right]^2\right\} \\
& +O\left\{E\left(\mathbf{d}_h'\left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right]\right)^2\right\}.
\end{aligned}$$

Now

$$\begin{aligned}
& E\left[\mathbf{b}_h'\left(vec\left(G_T-G_0\right)\otimes S_T\right)\right]^2 \\
= \;& E\mathbf{b}_h'\left[vec\left(G_T-G_0\right)\otimes S_T\right]\left[vec\left(G_T-G_0\right)'\otimes S_T'\right]\mathbf{b}_h \\
= \;& E\mathbf{b}_h'[\left(vec\left(G_T-G_0\right)vec\left(G_T-G_0\right)'\right)\otimes(S_T S_T')]\mathbf{b}_h \\
= \;& \frac{1}{T^3}\mathbf{b}_h' E\left\{\sum_{t,s,p,q}[vec(w_t-Ew_t)vec(w_s-Ew_s)']\otimes\left(v_p v_q'\right)\right\}\mathbf{b}_h \\
= \;& O\left(\frac{1}{T}\right),
\end{aligned}$$

and

$$\begin{aligned}
& E\left\{\mathbf{d}_h'\left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right]\right\}^2 \\
= \;& O\left[E\left\|vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\right\|^2 E\left\|S_T\right\|^2\right] \\
= \;& O\left[E\left\|vech(\check{\Omega}_T-\bar{\Omega}_T)\right\|^4 E\left\|S_T\right\|^2\right] \\
= \;& O\left(\left(\frac{M}{T}\right)^2\right),
\end{aligned}$$

using Lemma A.2(a), (c) and (d). Therefore

$$\begin{aligned}
E\left(h_T^2\right)-1 \;=\;& 2\mathbf{a}_h' S_T \mathbf{b}_h'\left(vec\left(G_T-G_0\right)\otimes S_T\right)+2\mathbf{a}_h' S_T \mathbf{c}_h'[vech\left(\check{\Omega}_T-\tilde{\Omega}_T\right)\otimes S_T] \\
& +2\mathbf{a}_h' S_T \mathbf{c}_h'[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]+2\mathbf{a}_h' S_T \mathbf{c}_h'[vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T] \\
& +2\mathbf{a}_h' S_T \mathbf{d}_h'\left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right] \\
& +E\left[\mathbf{c}_h'\left(vech\left(\check{\Omega}_T-\Omega_0\right)\otimes S_T\right)\right]^2+O\left(\frac{1}{T}\right)+O\left(\left(\frac{M}{T}\right)^2\right).
\end{aligned}$$

Next, we proceed to examine each term in the above equation. Consider the first term when both w and v are scalars, the vector case following by the similar argument. Note that

$$\begin{aligned}
& E\mathbf{a}_h' S_T \mathbf{b}_h' \left(vec\left(G_T - G_0\right) \otimes S_T\right) \\
= \ & O\left(\frac{1}{T^2}\sum_{i,j,\ell} \left|E\left[\left(w_i - Ew_i\right) v_j v_\ell\right]\right|\right) \\
= \ & O\left(\frac{1}{T}\sum \left|E\left[\left(w_0 - Ew_0\right) v_j v_{j+\ell}\right]\right| + \frac{1}{T}\sum \left|E\left[v_0 v_j \left(w_{j+\ell} - Ew_{j+\ell}\right)\right]\right|\right),
\end{aligned}$$

where the sum is over $j \geq 0, \ell \geq 0$ and $j + \ell \leq T$. Using the strong mixing property of $(w_t', v_t')'$, we have

$$\left|E\left(w_0 - Ew_0\right) v_j v_{j+\ell}\right| \leq C(\min(\alpha_j, \alpha_\ell)),$$

and

$$E\left[v_0 v_j \left(w_{j+\ell} - Ew_{j+\ell}\right)\right] \leq C(\min(\alpha_j, \alpha_\ell)),$$

for some constant $C > 0$. Without loss of generality, we can assume that α_i is decreasing (see Billingsley (1968, page 168)). Therefore, we find that

$$\begin{aligned}
& E\mathbf{a}_h' S_T \mathbf{b}_h' \left(vec\left(G_T - G_0\right) \otimes S_T\right) \\
= \ & O\left(\frac{1}{T}\sum_{j=0}^{T}\sum_{\ell=0}^{T} \min(\alpha_j, \alpha_\ell)\right) = O\left(\frac{1}{T}\sum_{j=0}^{T}\sum_{\ell=0}^{j} \min(\alpha_j, \alpha_\ell)\right) \\
= \ & O\left(\frac{1}{T}\sum_{j=0}^{T}\sum_{\ell=0}^{j} \alpha_j\right) = O\left(\frac{1}{T}\sum_{j=0}^{T} j\alpha_j\right) = O\left(\frac{1}{T}\right),
\end{aligned}$$

in view of using Assumption 2(a).

Secondly,

$$\check{\Omega}_T - \tilde{\Omega}_T = D_1 + D_2 + D_3 + D_4,$$

where

$$\begin{aligned}
D_1 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{j=0}^{M} \left[k_j v_{t+j} z_t' x_t' (\check{\beta}_T - \beta_0) + k_j z_{t+j} v_t' x_{t+j}' (\check{\beta}_T - \beta_0)\right], \\
D_2 &= \frac{1}{T}\sum_{t=1}^{T}\sum_{j=0}^{M} \left[k_j (\check{\beta}_T - \beta_0)' x_{t+j} z_{t+j} z_t' x_t' (\check{\beta}_T - \beta_0)\right], \\
D_3 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{j=-M}^{-1} k_j v_t z_{t-j}' x_{t-j}' (\check{\beta}_T - \beta_0) + k_j z_t v_{t-j} x_t' (\check{\beta}_T - \beta_0), \\
D_4 &= \frac{1}{T}\sum_{t=1}^{T}\sum_{j=-M}^{-1} \left[k_j (\check{\beta}_T - \beta_0)' x_t z_t z_{t-j}' x_{t-j}' (\check{\beta}_T - \beta_0)\right].
\end{aligned}$$

So

$$\begin{aligned}
& \frac{T}{M} E 2\mathbf{a}_h' S_T \mathbf{c}_h' [vech\left(\check{\Omega}_T - \tilde{\Omega}_T\right) \otimes S_T] \\
= \ & \sum_{i=1}^{4} \frac{T}{M} E 2\mathbf{a}_h' S_T \left[\mathbf{c}_h' \left(D_i \otimes S_T\right)\right] = \sum_{i=1}^{4} \lim_{T\to\infty} \frac{T}{M} E 2\mathbf{a}_h' S_T \left[\mathbf{c}_h' \left(D_i \otimes S_T\right)\right] + o(1).
\end{aligned}$$

The above limits are not zero in general, the reason being that

$$\lim_{T\to\infty}(T/M)\,Evech\left(\check{\Omega}_T-\tilde{\Omega}_T\right)\neq 0.$$

To see this, we consider the special case where $x_t = z_t = 1$ and take the following limit as an example:

$$\begin{aligned}
&\lim_{T\to\infty} E\frac{T}{M}\left[\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{j=0}^{M}\left(k_j v_{t+j} z_t' x_t'(\check{\beta}_T-\beta_0)\right)\right)\right]\\
=\;& C\lim_{T\to\infty}\frac{1}{M}\sum_{j=0}^{M}k_j E v_j \frac{1}{T}\sum_{i=1}^{T}\sum_{t=1}^{T}v_{i-t} = C\lim_{T\to\infty}\frac{1}{M}\sum_{j=0}^{M}k_j E v_j\sum_{r=-T+1}^{T-1}\left(1-\frac{r}{T}\right)v_r\\
=\;& C\lim_{T\to\infty}\frac{1}{M}\sum_{j=0}^{M}k_j\sum_{r=-T+1}^{T-1}\left(1-\frac{r}{T}\right)E\left(v_r v_j\right)=\int_0^1 k(x)dx\sum_{j=-\infty}^{\infty}cov(v_0,v_j)\neq 0.
\end{aligned}$$

The general case is more complicated but limits of the above type will show up, provided there is a regressor with a nonzero mean.

Thirdly,

$$\begin{aligned}
&E\frac{T}{M}\mathbf{a}_h' S_T\mathbf{c}_h'[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]\\
=\;&\frac{1}{MT}\sum_{j=-M}^{M}\sum_{t,\ell,m}k_j E\left\{\mathbf{a}_h' v_t\mathbf{c}_h'\left[vech\left(v_\ell v_{\ell-j}'-\Gamma_j\right)\otimes v_m\right]\right\}\\
=\;&\frac{1}{M}\sum_{j=-M}^{M}\sum_{\ell,m=-T+1}^{T-1}k_j E\left\{\mathbf{a}_h' v_0\mathbf{c}_h'\left[vech\left(v_\ell v_{\ell-j}'-\Gamma_j\right)\otimes v_m\right]\right\}+O\left(\frac{M}{T}\right)\\
=\;&\lim_{T\to\infty}\frac{1}{M}\sum_{j=-M}^{M}\sum_{\ell,m=-T+1}^{T-1}E\left\{\mathbf{a}_h' v_0\mathbf{c}_h'\left[vech\left(v_\ell v_{\ell-j}'-\Gamma_j\right)\otimes v_m\right]\right\}+O\left(\frac{M}{T}\right). \qquad \text{(A.17)}
\end{aligned}$$

Using Lemma A.2, the Hölder inequality and Assumption 2(a), we can show that

$$E\frac{T}{M}\mathbf{a}_h' S_T\mathbf{c}_h'[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]=O\left(1\right),$$

and the limit in (A.17) is finite.

Fourthly,

$$\begin{aligned}
&E\mathbf{a}_h' S_T\mathbf{c}_h'[vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T]\\
=\;&\frac{1}{T}\sum_{j=-M}^{M}\sum_{\ell=1}^{T}\sum_{m=1}^{T}\left(k_j-1\right)E\left(\mathbf{a}_h' v_\ell\mathbf{c}_h'\left[vech\left(\Gamma_j\right)\otimes v_m\right]\right)\\
=\;&M^{-q}\frac{1}{T}\sum_{j=-M}^{M}\sum_{\ell=1}^{T}\sum_{m=1}^{T}\left(\frac{k_j-1}{\left(|j|/M\right)^q}\right)|j|^q E\left(\mathbf{a}_h' v_\ell\mathbf{c}_h'\left[vech\left(\Gamma_j\right)\otimes v_m\right]\right)\\
=\;&-M^{-q}\lim_{T\to\infty}\frac{1}{T}\sum_{j=-M}^{M}\sum_{\ell=1}^{T}\sum_{m=1}^{T}g_q|j|^q E\left(\mathbf{a}_h' v_\ell\mathbf{c}_h'\left[vech\left(\Gamma_j\right)\otimes v_m\right]\right)+O(M^{-2q})\\
=\;&-M^{-q}\lim_{T\to\infty}\frac{1}{T}\sum_{\ell=1}^{T}\sum_{m=1}^{T}\sum_{j=-\infty}^{\infty}g_q|j|^q E\left(\mathbf{a}_h' v_\ell\mathbf{c}_h'\left[vech\left(\Gamma_j\right)\otimes v_m\right]\right)+O(M^{-2q}).
\end{aligned}$$

Fifthly,

$$
\begin{aligned}
& E\frac{T}{M}2\mathbf{a}_h' S_T \mathbf{d}_h' \left[vech(\check{\Omega}_T - \bar{\Omega}_T) \otimes vech(\check{\Omega}_T - \bar{\Omega}_T) \otimes S_T \right] \\
= \; & E\frac{T}{M}2\mathbf{a}_h' S_T \mathbf{d}_h' \left[vech(\tilde{\Omega}_T - \bar{\Omega}_T) \otimes vech(\tilde{\Omega}_T - \bar{\Omega}_T) \otimes S_T \right] + o(1) \\
= \; & \lim_{T\to\infty} \frac{1}{MT} \sum_{o,p=-M}^{M} \sum_{i,j,\ell=-T}^{T} 2k_o k_p E \mathbf{a}_h' v_0 \\
& \times \mathbf{d}_h' \left[vech(v_i v_{i-o}' - \Gamma_o) \otimes vech(v_j v_{j-p}' - \Gamma_p) \otimes v_\ell \right] + o(1) . \qquad \text{(A.18)}
\end{aligned}
$$

Finally,

$$
\begin{aligned}
& E\frac{T}{M} \left[\mathbf{c}_h' \left(vech\left(\check{\Omega}_T - \Omega_0 \right) \otimes S_T \right) \right]^2 \\
= \; & \frac{T}{M} E \left[\mathbf{c}_h' \left(vech\left(\check{\Omega}_T - \bar{\Omega}_T \right) \otimes S_T \right) \right]^2 + O\left\{ \frac{T}{M} E \left[\mathbf{c}_h' \left(vech\left(\bar{\Omega}_T - \Omega_0 \right) \otimes S_T \right) \right]^2 \right\} \\
= \; & \frac{T}{M} E \left[\mathbf{c}_h' \left(vech\left(\tilde{\Omega}_T - \bar{\Omega}_T \right) \otimes S_T \right) \right]^2 + o(1) \\
= \; & \lim_{T\to\infty} \frac{1}{MT} \sum_{o,p=-M}^{M} \sum_{i,j,\ell=-T}^{T} k_o k_p E \mathbf{c}_h' \left[vech\left(v_0 v_{-o}' - \Gamma_o \right) \otimes v_j \right] \\
& \times \mathbf{c}_h' \left[vech\left(v_i v_{i-p}' - \Gamma_p \right) \otimes v_\ell \right] + o(1). \qquad \text{(A.19)}
\end{aligned}
$$

Using Lemma A.2, the Hölder inequality and the mixing property, we can show that the limits in (A.18) and (A.19) are finite. Combining the above results yields

$$
\kappa_2(h_T) = 1 + \rho_{1,\infty} M^{-q} + \left(\rho_{2,\infty} + \kappa_{2,\infty} \right) (M/T) + o(M/T) .
$$

Part (c) We sketch the proof and omit the arguments for the finiteness of $\kappa_{3,\infty}$. By definition

$$
\begin{aligned}
& T^{1/2}\kappa_3(h_T) \\
= \; & T^{1/2} E((h_T)^3) - 3T^{1/2} E((h_T)^2) E(h_T) + 2T^{1/2} \left(Eh_T \right)^3 \\
= \; & T^{1/2} E((h_T)^3) - 3\left[1 - \frac{\rho_{1,\infty}}{M^q} + O\left(\frac{M}{T} \right) \right] \\
& \times \left[\kappa_{1,\infty} + O\left(\frac{1}{M^q} \right) + o\left(\frac{M}{\sqrt{T}} \right) \right] + O\left(\frac{1}{T} \right) \\
= \; & T^{1/2} E((h_T)^3) - 3\kappa_{1,\infty} + O\left(\frac{1}{M^q} \right) + o\left(\frac{M}{\sqrt{T}} \right),
\end{aligned}
$$

where we have used parts (a) and (b) of the lemma. It remains to show that

$$
T^{1/2} E((h_T)^3) = \kappa_{3,\infty} + O\left(\frac{1}{M^q} \right) + o\left(\frac{M}{\sqrt{T}} \right).
$$

Using Assumption 1, Holder's inequality and Lemma A.2, we have

$$
\begin{aligned}
E((h_T)^3) \; = \; & E\left(\mathbf{a}_h' S_T \right)^3 + 3\left(\mathbf{a}_h' S_T \right)^2 \mathbf{b}_h' [vec\left(G_T - G_0 \right) \otimes S_T] \\
& + 3\left(\mathbf{a}_h' S_T \right)^2 \mathbf{c}_h' [vech\left(\check{\Omega}_T - \tilde{\Omega}_T \right) \otimes S_T] \\
& + 3\left(\mathbf{a}_h' S_T \right)^2 \mathbf{c}_h' [vech\left(\tilde{\Omega}_T - \bar{\Omega}_T \right) \otimes S_T] \\
& + 3\left(\mathbf{a}_h' S_T \right)^2 \mathbf{c}_h' [vech\left(\bar{\Omega}_T - \Omega_0 \right) \otimes S_T] \\
& + o(M/T).
\end{aligned}
$$

We consider two of the above five terms here, the other terms following in the same way. First,

$$
\begin{aligned}
& 3\sqrt{T}E\left\{\left(\mathbf{a}_h' S_T\right)^2 \mathbf{c}_h'[vech\left(\check{\Omega}_T - \tilde{\Omega}_T\right) \otimes S_T]\right\} \\
= \; & 3\sqrt{T}E\left\{\left(\mathbf{a}_h' S_T\right)^2 \mathbf{c}_h'[vech\left(\nabla\tilde{\Omega}_T\left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T]\right\} + O\left(\frac{1}{\sqrt{T}}\right) \\
= \; & 3\lim_{T\to\infty}\sqrt{T}E\left\{\left(\mathbf{a}_h' S_T\right)^2 \mathbf{c}_h'[vech\left(\nabla\tilde{\Omega}_T\left(\check{\beta}_T - \beta_0\right)\right) \otimes S_T]\right\} + O\left(\frac{1}{\sqrt{T}}\right) + O\left(\frac{1}{M^q}\right).
\end{aligned}
$$

Second,

$$
\begin{aligned}
& 3\sqrt{T}\left(\mathbf{a}_h' S_T\right)^2 \mathbf{c}_h'[vech\left(\bar{\Omega}_T - \Omega_0\right) \otimes S_T] \\
= \; & \frac{3}{T}\sum_{o=1}^{T}\sum_{p=1}^{T}\sum_{i=1}^{T}\mathbf{a}_h' v_o \mathbf{a}_h' v_p \mathbf{c}_h'\left[\sum_{j=-M}^{M}\left(k_j - 1\right) vech(\Gamma_j) \otimes v_i\right] \\
= \; & 3\sum_{p=-T}^{T}\sum_{i=-T}^{T}\mathbf{a}_h' v_0 \mathbf{a}_h' v_p \mathbf{c}_h'\left[\sum_{j=-M}^{M}\left(k_j - 1\right) vech(\Gamma_j) \otimes v_i\right] + O\left(\frac{M}{T}\right) \\
= \; & -3g_q M^{-q}\sum_{p,i=-T}^{T}\mathbf{a}_h' v_0 \mathbf{a}_h' v_p \mathbf{c}_h'\left[\sum_{j=-M}^{M}|j|^q\, vech(\Gamma_j) \otimes v_i\right] + O\left(\frac{M}{T}\right) \\
= \; & O(M^{-q}) + O\left(\frac{M}{T}\right).
\end{aligned}
$$

Part (d) As before, we briefly sketch the proof and omit arguments relating to the finiteness of the limit quantities as these follow in the same way as earlier. Let $\lambda_j = E((h_T)^j)$, then

$$
\begin{aligned}
\kappa_4\left(h_T\right) &= \lambda_4 + 12\lambda_1^2\lambda_2 - 3\lambda_2^2 - 4\lambda_1\lambda_3 - 6\lambda_1^4 \\
&= \lambda_4 - 3\left[1 + \left(\frac{\rho_{1,\infty}}{M^q}\right) + \frac{M}{T}\left(\rho_{2,\infty} + \kappa_{2,\infty}\right)\right]^2 + O\left(\frac{1}{T}\right) \\
&= \lambda_4 - 3 - 6\left(\frac{\rho_{1,\infty}}{M^q}\right) - 6\frac{M}{T}\left(\rho_{2,\infty} + \kappa_{2,\infty}\right) + O\left(\frac{M^2}{T^2}\right) + O\left(\frac{1}{T}\right). \qquad \text{(A.20)}
\end{aligned}
$$

Using Lemma A.2 and Hölder's inequality, we have

$$
\begin{aligned}
\lambda_4 = \; & E\left(\mathbf{a}_h' S_T\right)^4 + 4E(\mathbf{a}_h' S_T)^3 \mathbf{b}_h'[vec\left(G_T - G_0\right) \otimes S_T] \\
& + 4E(\mathbf{a}_h' S_T)^3 \mathbf{c}_h'[vech\left(\check{\Omega}_T - \tilde{\Omega}_T\right) \otimes S_T] \\
& + 4E(\mathbf{a}_h' S_T)^3 \mathbf{c}_h'[vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T] \\
& + 4E(\mathbf{a}_h' S_T)^3 \mathbf{c}_h'[vech\left(\bar{\Omega}_T - \Omega_0\right) \otimes S_T] \qquad \text{(A.21)} \\
& + 4E(\mathbf{a}_h' S_T)^3 \mathbf{d}_h'\left[vech(\check{\Omega}_T - \bar{\Omega}_T) \otimes vech(\check{\Omega}_T - \bar{\Omega}_T) \otimes S_T\right] \\
& + 6E(\mathbf{a}_h' S_T)^2\left\{\mathbf{c}_h'[vech\left(\check{\Omega}_T - \Omega_0\right) \otimes S_T]\right\}^2 + o\left(\frac{M}{T}\right) + o\left(\frac{1}{M^q}\right).
\end{aligned}
$$

Combining equations (A.20) and (A.21), we obtain:

$$
\begin{aligned}
& \frac{T}{M}\left[\kappa_4\left(h_T\right)+6\frac{M}{T}\kappa_{2,\infty}\right] \\
= \; & 4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{c}_h'[vech\left(\check{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]-12\frac{T}{M}E(\mathbf{a}_h'S_T)\mathbf{c}_h'[vech\left(\check{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T] \\
& +4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{c}_h'[vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T]-12\frac{T}{M}E(\mathbf{a}_h'S_T)\mathbf{c}_h'[vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T] \\
& +4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{d}_h'\left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right] \\
& +6\frac{T}{M}E(\mathbf{a}_h'S_T)^2\left\{\mathbf{c}_h'[vech\left(\check{\Omega}_T-\Omega_0\right)\otimes S_T]\right\}^2 \\
& +o(1), \qquad\qquad (A.22)
\end{aligned}
$$

where we use

$$E\left(\mathbf{a}_h'S_T\right)^4-3=O(1/T), \qquad (A.23)$$

and

$$4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{b}_h'[vec\left(G_T-G_0\right)\otimes S_T]=o(1).$$

Consider each term in equation (A.22) in turn. First, construct the orthogonal matrix $[a_h, a_\perp]$ and write $S_t = a_h a_h' S_T + a_\perp a_\perp' S_T$. Using (A.23), we have

$$E\left\{\left(\mathbf{a}_h'S_T\right)^3S_T-3\left(\mathbf{a}_h'S_T\right)S_T\right\}=E\left\{\left[\left(\mathbf{a}_h'S_T\right)^4-3\left(\mathbf{a}_h'S_T\right)^2\right]a_h+\left[\left(\mathbf{a}_h'S_T\right)^3-3\mathbf{a}_h'S_T\right]a_\perp a_\perp'S_T\right\}=O(1/T),$$

and it can then be shown that

$$4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{c}_h'[vech\left(\check{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]=12\frac{T}{M}E(\mathbf{a}_h'S_T)\mathbf{c}_h'[vech\left(\check{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]+o(1),$$

and

$$4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{c}_h'[vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T]=12\frac{T}{M}E(\mathbf{a}_h'S_T)\mathbf{c}_h'[vech\left(\bar{\Omega}_T-\Omega_0\right)\otimes S_T]+o(1).$$

Second,

$$
\begin{aligned}
& 4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{c}_h'[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T] \\
= \; & 4\frac{1}{MT^2}\sum_{i,j,\ell,m,p}\sum_{n=-M}^{M}k_nE\mathbf{a}_h'v_i\mathbf{a}_h'v_j\mathbf{a}_h'v_\ell\mathbf{c}_h'\left[\left(vech(v_mv_{m-n}')-\Gamma_n\right)\otimes v_p\right] \\
= \; & 4\frac{1}{MT}\sum_{j,\ell,m,p}\sum_{n=-M}^{M}k_nE\mathbf{a}_h'v_0\mathbf{a}_h'v_j\mathbf{a}_h'v_\ell\mathbf{c}_h'\left[\left(vech(v_mv_{m-n}')-\Gamma_n\right)\otimes v_p\right] \\
= \; & \lim_{T\to\infty}4\frac{1}{MT}\sum_{j,\ell,m,p}\sum_{n=-M}^{M}k_nE\mathbf{a}_h'v_0\mathbf{a}_h'v_j\mathbf{a}_h'v_\ell\mathbf{c}_h'\left[\left(vech(v_mv_{m-n}')-\Gamma_n\right)\otimes v_p\right]+o(1).
\end{aligned}
$$

Third,

$$
\begin{aligned}
& 4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{d}_h'\left[vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes vech(\check{\Omega}_T-\bar{\Omega}_T)\otimes S_T\right] \\
= \; & 4\frac{T}{M}E(\mathbf{a}_h'S_T)^3\mathbf{d}_h'\left[vech(\tilde{\Omega}-\bar{\Omega}_T)\otimes vech(\tilde{\Omega}-\bar{\Omega}_T)\otimes S_T\right]+o\left(1\right) \\
= \; & \lim_{T\to\infty}\frac{4}{MT^2}\sum_{i,j,k,m,o=-T}^{T}\sum_{l,n=-M}^{M}Ek_\ell k_n\mathbf{a}_h'v_0\mathbf{a}_h'v_i\mathbf{a}_h'v_j\times \\
& \mathbf{d}_h'\left[vech(v_kv_{k-\ell}'-\Gamma_\ell)\otimes vech\left(v_mv_{m-n}'-\Gamma_n\right)\otimes v_o\right]+o\left(1\right).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
& 6\frac{T}{M}E(\mathbf{a}_h'S_T)^2\left\{\mathbf{c}_h'[vech\left(\check{\Omega}_T-\Omega_0\right)\otimes S_T]\right\}^2 \\
= \; & 6\frac{T}{M}E(\mathbf{a}_h'S_T)^2\left\{\mathbf{c}_h'[vech\left(\tilde{\Omega}_T-\Omega_T\right)\otimes S_T]\right\}^2+o\left(1\right) \\
= \; & \lim_{T\to\infty}\frac{6}{MT^3}\sum_{i,j,\ell,m,o=-T}^{T}\sum_{k,n=-M}^{M}Ek_\ell k_n\mathbf{a}_h'v_0\mathbf{a}_h'v_i\times \\
& \mathbf{c}_h'\left[vech(v_jv_{j-k}'-\Gamma_k)\otimes v_\ell\right]\mathbf{c}_h'[vech\left(v_mv_{m-n}'-\Gamma_n\right)\otimes v_o].
\end{aligned}
$$

Combining the above results completes the proof of this part. ■

Lemma A.8 *Define*

$$
\begin{aligned}
\varsigma_T\left(\theta\right) = \; & \frac{1}{\sqrt{T}}\left[\kappa_{1,\infty}\left(i\theta\right)+\frac{1}{6}\left(\kappa_{3,\infty}-3\kappa_{1,\infty}\right)\left(i\theta\right)^3\right]+\frac{1}{2}\frac{1}{M^q}\rho_{1,\infty}\left(i\theta\right)^2 \\
& +\frac{M}{T}\left[\frac{(i\theta)^2}{2}(\rho_{2,\infty}+\kappa_{2,\infty})+\frac{(i\theta)^4}{24}\left(\kappa_{4,\infty}-6\kappa_{2,\infty}\right)\right],
\end{aligned}
$$

$$
\bar{\phi}_h(\theta)=\exp\left(-\theta^2/2\right)\left(1+\varsigma_T\left(\theta\right)\right),
$$

$$
f_h(x)=\frac{1}{2\pi}\int\exp(-i\theta x)\bar{\phi}_h(\theta)d\theta,\;\; \Phi_T(x)=\int_{-\infty}^{x}f_h(u)du.
$$

We have

$$
\begin{aligned}
\Phi_T(x) = \; & \Phi(x)-\frac{1}{\sqrt{T}}\left[\kappa_{1,\infty}+\frac{1}{6}\left(\kappa_{3,\infty}-3\kappa_{1,\infty}\right)\left(x^2-1\right)\right]\phi(x)-\frac{1}{2}\frac{1}{M^q}\rho_{1,\infty}x\phi(x) \\
& -\frac{1}{2}\frac{M}{T}\left[\left(\rho_{2,\infty}+\kappa_{2,\infty}\right)x+\frac{1}{12}\left(\kappa_{4,\infty}-6\kappa_{2,\infty}\right)\left(x^3-3x\right)\right]\phi\left(x\right).
\end{aligned}
$$

Proof. Let $\phi\left(x\right)$ be the pdf of a standard normal random variable, then

$$
\frac{d^n}{dx^n}\phi(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-i\theta x\right)\left(-i\theta\right)^n\exp\left(-\theta^2/2\right)d\theta.
$$

So

$$
\begin{aligned}
f_h(x) = \; & \frac{1}{2\pi}\int\exp(-i\theta x)\bar{\phi}_h(\theta)d\theta \\
= \; & \phi(x)+\frac{1}{\sqrt{T}}\left[\left(-\kappa_{1,\infty}\right)\phi'(x)-\frac{1}{6}\left(\kappa_{3,\infty}-3\kappa_{1,\infty}\right)\phi^{(3)}(x)\right]+\frac{1}{2}\frac{1}{M^q}\left[\rho_{1,\infty}\phi''(x)\right] \\
& +\frac{M}{T}\left[\frac{\rho_{2,\infty}+\kappa_{2,\infty}}{2}\phi''(x)+\frac{1}{24}\left(\kappa_{4,\infty}-6\kappa_{2,\infty}\right)\phi^{(4)}(x)\right],
\end{aligned}
$$

and thus

$$\begin{aligned}\Phi_T(x) &= \Phi(x) + \frac{1}{\sqrt{T}}\left[\left(-\kappa_{1,\infty}\right)\phi(x) - \frac{1}{6}\left(\kappa_{3,\infty} - 3\kappa_{1,\infty}\right)\phi''(x)\right] + \frac{1}{2}\frac{1}{M^q}\left[\rho_{1,\infty}\phi'(x)\right] \\ &\quad + \frac{M}{T}\left[\frac{\rho_{2,\infty}+\kappa_{2,\infty}}{2}\phi'(x) + \frac{1}{24}\left(\kappa_{4,\infty} - 6\kappa_{2,\infty}\right)\phi^{(3)}(x)\right] \\ &= \Phi(x) - \frac{1}{\sqrt{T}}\left[\kappa_{1,\infty} + \frac{1}{6}\left(\kappa_{3,\infty} - 3\kappa_{1,\infty}\right)\left(x^2-1\right)\right]\phi(x) - \frac{1}{2}\frac{1}{M^q}\rho_{1,\infty}x\phi(x) \\ &\quad - \frac{1}{2}\frac{M}{T}\left[\left(\rho_{2,\infty}+\kappa_{2,\infty}\right)x + \frac{1}{12}\left(\kappa_{4,\infty} - 6\kappa_{2,\infty}\right)\left(x^3-3x\right)\right]\phi\left(x\right).\end{aligned}$$

■

Lemma A.9 *For some constant* $C > 0$, *we have*

$$\sup_{x\in\mathbb{R}}\left|P(t_M < x) - \Phi_T(x)\right| \le C\int_{|\theta|\le M^{-1}T\log T}\left|\phi_h(\theta) - \bar{\phi}_h(\theta)\right||\theta|^{-1}\,d\theta + o\left(\frac{M}{T}\right).$$

Proof. It follows from Lemmas A.4 and A.6 that

$$t_M = h_T + \xi_T^*,$$

for ξ_T^* satisfying

$$P\left(\left|\xi_T^*\right| > \eta_T\right) = o\left(\eta_T\right).$$

Using Lemma 6 in Andrews (2002, page 1064), we then obtain

$$\sup_{x\in\mathbb{R}}\left|P\left(t_M < x\right) - P\left(h_T < x\right)\right| = o\left(\frac{M}{T}\right). \tag{A.24}$$

Note that $\bar{\phi}_h(\theta) = \int \exp(i\theta x)d\Phi_T(x)$. It follows from the smoothing lemma (see, for example, Kolassa (1997), Lemma 2.5.2, page 17) that, for some constant $C > 0$

$$\begin{aligned}&\sup_{x\in\mathbb{R}}\left|P(h_T < x) - \Phi_T(x)\right| \\ \le\; & C\int_{|\theta|\le M^{-1}T\log T}\left|\phi_h(\theta) - \bar{\phi}_h(\theta)\right||\theta|^{-1}\,d\theta + O\left(\frac{M}{T\log T}\right) \\ =\; & C\int_{|\theta|\le M^{-1}T\log T}\left|\phi_h(\theta) - \bar{\phi}_h(\theta)\right||\theta|^{-1}\,d\theta + o\left(\frac{M}{T}\right).\end{aligned} \tag{A.25}$$

Combining (A.24) and (A.25) yields the desired result. ■

Following Götze and Künsch (1996), define a truncation function by

$$\tau(x) = T^{\gamma}xf\left(T^{-\gamma}\left\|x\right\|\right)/\left\|x\right\|,$$

where $\gamma \in (2/r, 1/2)$ and $f \in C^{\infty}(0,\infty)$ satisfies (i) $f(x) = x$ for $x \le 1$; (ii) f is increasing; (iii) $f(x) = 2$ for $x \ge 2$. Figure 1 gives an example of such a truncation function.

Define

$$\tilde{R}_t = \tau\left(R_t\right) := \begin{cases} R_t, & \text{if } \left\|R_t\right\| \le T^{\gamma}, \\ T^{\gamma}R_t f\left(T^{-\gamma}\left\|R_t\right\|\right)/\left\|R_t\right\|, & \text{if } T^{\gamma} < \left\|R_t\right\| \le 2T^{\gamma}, \\ 2T^{\gamma}R_t/\left\|R_t\right\|, & \text{if } \left\|R_t\right\| > 2T^{\gamma}, \end{cases}$$

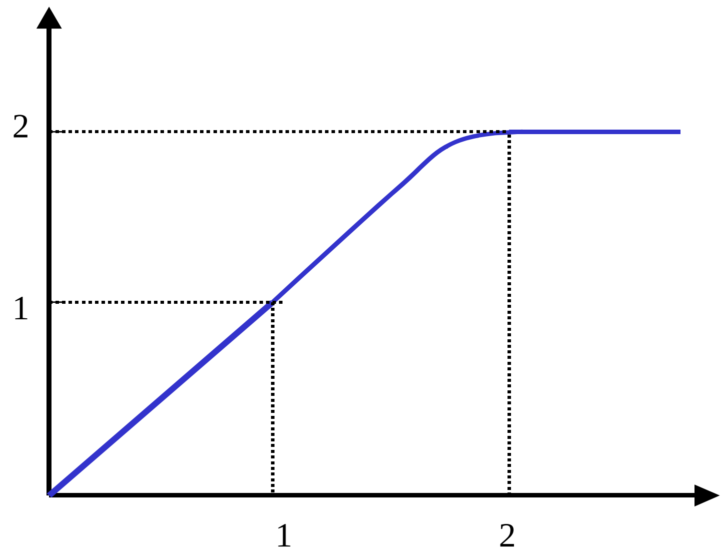


Figure 1: Graph of the Truncation Function

which satisfies $\left\|\tilde{R}_t\right\| \leq 2T^{\gamma}$. Let $\tilde{h}_T$ be defined as h_T but with R_t replaced $\tilde{R}_t$ and let $\phi_{\tilde{h}}(\theta)$ be its characteristic function. It is not difficult to show that

$$\begin{aligned} & \int_{|\theta| \leq M^{-1} T \log T} \left|\phi_h(\theta) - \bar{\phi}_h(\theta)\right| |\theta|^{-1} d\theta \\ = & \int_{|\theta| \leq M^{-1} T \log T} \left|\phi_{\tilde{h}}(\theta) - \bar{\phi}_h(\theta)\right| |\theta|^{-1} d\theta + o\left(\frac{M}{T}\right). \end{aligned}$$

We will assume in the sequel that such a truncation transformation has been made. For ease of exposition, we will drop the '~' notation and identify $\tilde{h}_T$ with h_T.

Lemma A.10 *For* $0 \leq \varepsilon < 1/\left[7(q+1)\right]$, *we have*

$$\int_{|\theta| \leq T^{\varepsilon}} \left|\phi_h(\theta) - \bar{\phi}_h(\theta)\right| |\theta|^{-1} d\theta = o\left(M/T\right).$$

Proof. Following the same proof as that for Lemma A.7, we can show that $\kappa_5\left(h_T\right) = o(M/T)$. Therefore,

$$\log \phi_h\left(\theta\right) = \sum_{j=1}^{4} \kappa_j(h_T) \frac{(i\theta)^j}{j!} + o\left(\frac{M}{T} |\theta|^5\right)$$

uniformly over θ and T. Using Lemma A.7, we have

$$\begin{aligned} \log \phi_h\left(\theta\right) &= -\frac{\theta^2}{2} + \varsigma_T\left(\theta\right) + (i\theta)\left[O(T^{-1/2} M^{-q}) + o\left(M/T\right)\right] \\ &\quad + (i\theta)^2 o\left(M/T\right) + (i\theta)^3\left[o(M^{-q}) + o\left(M/T\right)\right] \\ &\quad + (i\theta)^4 o\left(M/T\right) + o\left(\frac{M}{T} |\theta|^5\right) \\ : &= -\frac{\theta^2}{2} + \varsigma_T\left(\theta\right) + \theta \frac{M}{T} \Delta_2(\theta, T) \end{aligned}$$

where $\sup_{|\theta|\le T^{\varepsilon}} |\Delta_2(\theta,T)| \to 0$ as $T\to\infty$.

Now

$$\begin{aligned}
\left|\phi_h(\theta)-\bar{\phi}_h(\theta)\right| &= \left|\exp(\log\phi_h(\theta)) - \exp\left(-\frac{\theta^2}{2}\right)(1+\varsigma_T(\theta))\right| \\
&= \exp\left(-\frac{\theta^2}{2}\right)\left|\exp\left[\left(\log\phi_h(\theta)+\frac{\theta^2}{2}\right)\right]-(1+\varsigma_T(\theta))\right| \\
&= \exp\left(-\frac{\theta^2}{2}\right)\left|\exp\left(\varsigma_T(\theta)+\theta\frac{M}{T}\Delta_2(\theta,T)\right)-(1+\varsigma_T(\theta))\right| \\
&\le \exp\left(-\frac{\theta^2}{2}\right)\left[\frac{M}{T}\left|\theta\Delta_2(\theta,T)\right|+\frac{1}{2}\left|\varsigma_T(\theta)+\theta\frac{M}{T}\Delta_2(\theta,T)\right|^2\right] \\
&\quad \times\exp\left(\left|\varsigma_T(\theta)+\theta\frac{M}{T}\Delta_2(\theta,T)\right|\right),
\end{aligned}$$

where we have used the result that for any two complex numbers u and v,

$$|\exp(u)-(1+v)| \le \left[|u-v|+\frac{|u|^2}{2}\right]\exp(|u|),$$

(see Lemma A2.2 in Severini (2005, page 481)).

In view of the definition of $\varsigma_T(\theta)$, we have, when $|\theta|\le T^{\varepsilon}$,

$$\left|\varsigma_T(\theta)+\theta\frac{M}{T}\Delta_2(\theta,T)\right| \le c\theta+\delta\theta^2,$$

for some constant $c>0$ and $\delta\in(0,1/4]$ and

$$\begin{aligned}
\sup_{\|\theta\|\le T^{\varepsilon}}\left|\varsigma_T^2(\theta)\theta^{-1}\right| &\le O\left\{\sup_{\|\theta\|\le T^{\varepsilon}}\left[\left(\frac{M}{T}\right)^2|\theta|^7+\frac{1}{T}|\theta|^5\right]\right\} \\
&= o\left[\frac{M}{T}\left(\frac{M}{T}\right)T^{1/[(q+1)]}\right]+o\left[\frac{1}{T}T^{5/[7(q+1)]}\right] \\
&= o\left(\frac{M}{T}T^{(1-q)/(1+q)}\right)+o\left[\frac{1}{T}T^{5/[5(q+1)]}\right] \\
&= o\left(\frac{M}{T}T^{(1-q)/(1+q)}\right)+o\left(\frac{M}{T}\right)=o\left(\frac{M}{T}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\int_{|\theta|\le T^{\varepsilon}}\left|\phi_h(\theta)-\bar{\phi}_h(\theta)\right||\theta|^{-1}d\theta \\
\le\ &\frac{M}{T}\int_{|\theta|\le T^{\varepsilon}}|\Delta_2(\theta,T)|\exp\left(-\frac{\theta^2}{2}+c\theta+\delta\theta^2\right)d\theta \\
&+\int_{|\theta|\le T^{\varepsilon}}\left|\theta^{-1}\varsigma_T^2(\theta)\right|\exp\left(-\frac{\theta^2}{2}+c\theta+\delta\theta^2\right)d\theta \\
&+\left(\frac{M}{T}\right)^2\int_{|\theta|\le T^{\varepsilon}}\theta\left[\Delta_2(\theta,T)\right]^2\exp\left(-\frac{\theta^2}{2}+c\theta+\delta\theta^2\right)d\theta \\
=\ &\left[\frac{M}{T}\sup_{|\theta|\le T^{\varepsilon}}|\Delta_2(\theta,T)|(1+o(1))+\sup_{\|\theta\|\le T^{\varepsilon}}\left|\varsigma_T^2(\theta)\theta^{-1}\right|\right]\int_{-\infty}^{\infty}\exp\left(-\frac{\theta^2}{2}+c\theta+\delta\theta^2\right)d\theta \\
=\ &o\left(\frac{M}{T}\right),
\end{aligned}$$

as desired. ■

Lemma A.11 *For* $0 \leq \varepsilon < 1/[7(q+1)]$, *we have*

$$\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\phi_h(\theta)-\bar{\phi}_h(\theta)\right||\theta|^{-1}d\theta=o(M/T).$$

Proof. Since

$$\begin{aligned}&\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\phi_h(\theta)-\bar{\phi}_h(\theta)\right||\theta|^{-1}d\theta\\ \leq\ &\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\phi_h(\theta)\right||\theta|^{-1}d\theta+\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\bar{\phi}_h(\theta)\right||\theta|^{-1}d\theta\end{aligned}$$

and

$$\begin{aligned}\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\bar{\phi}_h(\theta)\right||\theta|^{-1}d\theta &\leq T^{-\varepsilon}\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\bar{\phi}_h(\theta)\right|d\theta\\ &\leq T^{-\varepsilon}\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\exp(-\frac{1}{4}\theta^2)d\theta=o\left(\frac{M}{T}\right),\end{aligned}$$

it suffices to show that

$$\int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T}\left|\phi_h(\theta)\right||\theta|^{-1}d\theta=o(M/T).$$

We follow the same steps as in Götze and Künsch (1996, pages 1927-1930). Let $m=\mathcal{K}\log T$ for some $\mathcal{K}>0$ and $N=\left[\left(T/\theta^2+1\right)m^2\right]$ for $T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T$. Define

$$\begin{aligned}S_N &= \frac{1}{\sqrt{T}}\sum_{t=1}^{N}v_t,\ S_{T-N}=\frac{1}{\sqrt{T}}\sum_{t=N+1}^{T}v_t,\\ G_N &= \frac{1}{T}\sum_{t=1}^{N}w_t,\ G_{T-N}=\frac{1}{T}\sum_{t=N+1}^{T}w_t.\end{aligned}$$

Let

$$\tilde{\Omega}_N=\sum_{j=-M}^{M}k_j\tilde{\Gamma}_{j,N},\tilde{\Omega}_{T-N}=\sum_{j=-M}^{M}k_j\tilde{\Gamma}_{j,T-N},$$

where

$$\begin{aligned}\tilde{\Gamma}_{j,N} &= \begin{cases}\frac{1}{T}\sum_{t=1}^{N}v_{t+j}v_t', & j\geq 0,\\ \frac{1}{T}\sum_{t=1}^{N}v_tv_{t-j}', & j<0,\end{cases}\\ \tilde{\Gamma}_{j,T-N} &= \begin{cases}\frac{1}{T}\sum_{t=N+1}^{T}v_{t+j}v_t', & j\geq 0,\\ \frac{1}{T}\sum_{t=N+1}^{T}v_tv_{t-j}', & j<0,\end{cases}\end{aligned}$$

and $\check{\Omega}_N$, $\check{\Omega}_{T-N}$ are similarly defined. Given these definitions, we have

$$S_T=S_N+S_{T-N},\ G_T=G_N+G_{T-N},\ \check{\Omega}_T=\check{\Omega}_N+\check{\Omega}_{T-N},$$

and

$$\begin{aligned}h_T &= \mathbf{a}'S_T+Q^0\left(G_T-G_0,S_T,\check{\Omega}_T-\bar{\Omega}_T,\bar{\Omega}_T-\Omega_0\right)\\ : &= \mathbf{a}'S_T+Q\left(G_T,S_T,\check{\Omega}_T,\bar{\Omega}_T\right)\end{aligned}$$

where $Q(\cdot)$ is a polynomial in its arguments.

When $v_t = 0, w_t = 0$, for $t = 1, 2, ..., N$, we have $S_N = 0, G_N = 0$ and

$$\check{\Omega}_N = \tilde{\Omega}_N - \nabla\tilde{\Omega}_N \left(\check{\beta}_T - \beta_0\right) + \left(\check{\beta}_T - \beta_0\right) \nabla^2\tilde{\Omega}_N \left(\check{\beta}_T - \beta_0\right) = 0.$$

Expanding $Q\left(G_T, S_T, \check{\Omega}_T, \bar{\Omega}_T\right)$ around $v_t = 0, w_t = 0$, for $t = 1, 2, ..., N$ yields

$$\begin{aligned} Q\left(G_T, S_T, \check{\Omega}_T, \bar{\Omega}_T\right) &= Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right) \\ &\quad + \sum_{\mu,\nu} c^*_{\mu,\nu} v^\mu w^\nu Q^*_{\mu\nu}\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right) \end{aligned}$$

where $\left\{Q^*_{\mu\nu}\right\}$ denotes an appropriate set of polynomials,

$$\mu = (\mu_1, ..., \mu_N, 0, ..., 0), \nu = (\nu_1, ..., \nu_N, 0, ..., 0),$$

and $|\mu| + |\nu| \leq 9$. Here, for two compatible vectors $a = (a_1, a_2, ..., a_T)$ and $b = (b_1, b_2, ..., b_T)$, we use the following convention

$$a^b = \prod_{i=1}^T a_i^{b_i}, \ ab = \sum_{i=1}^T a_i b_i.$$

Using the above expansion and the result that

$$\exp(i\theta) = \sum_{j=0}^{r-1} \frac{(it)^j}{j!} + O\left(|\theta|^r\right)$$

(see Lemma A2.1 in Severini (2005, page 480), we obtain

$$\begin{aligned} \phi_h(\theta) &= E\exp(i\theta h_T) = E\left[\exp(i\theta\mathbf{a}' S_T) + i\theta Q\left(G_T, S_T, \check{\Omega}_T, \bar{\Omega}_T\right)\right] \\ &= E\exp\left[i\theta\mathbf{a}' S_T + i\theta Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right] \\ &\quad \times \exp\left[i\theta\sum_{\mu,\nu} c^*_{\mu,\nu} v^\mu w^\nu Q^*_{\mu\nu}\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right] \\ &= E\exp\left[i\theta\mathbf{a}' S_T + i\theta Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right] \\ &\quad \times \sum_{\mu,\nu} c_{\mu,\nu}(\theta) v^\mu w^\nu Q_{\mu\nu}\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right) \\ &\quad + O\left(|\theta|^r E\left|Q\left(G_T, S_T, \check{\Omega}_T, \bar{\Omega}_T\right) - Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right|^r\right), \end{aligned} \tag{A.26}$$

where $Q_{\mu\nu}$ is another set of polynomials, $c_{\mu,\nu}(\theta)$ is a set of coefficients that depend on θ, and $|\mu| + |\nu| \leq 9(r-1)$.

Let us estimate first the last remainder term in (A.26). Note that

$$\begin{aligned} & Q\left(G_T, S_T, \check{\Omega}_T, \bar{\Omega}_T\right) - Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right) \\ = \;& Q^0\left(G_T - G_0, S_T, \check{\Omega}_T - \bar{\Omega}_T, \bar{\Omega}_T - \Omega_0\right) \\ & - Q^0\left(G_{T-N} - G_0, S_{T-N}, \check{\Omega}_{T-N} - \bar{\Omega}_{T-N}, \bar{\Omega}_T - \Omega_0\right) \\ : \;& = Q_N^0(G_N, S_N, \check{\Omega}_N - \bar{\Omega}_N, \bar{\Omega}_T - \Omega_0) \end{aligned}$$

where $Q_N^0(\cdot)$ is a polynomial in its arguments with bounded coefficients. It follows from Lemma A.2 that

$$\begin{aligned} E\left\|S_N\right\|^r &= O\left(\left(\frac{N}{T}\right)^{r/2}\right), \\ E\left(\left\|G_N\right\|\right)^r &= O\left(\left(\frac{N}{T}\right)^{r/2}\left(\frac{1}{T}\right)^{r/2}+\left(\frac{N}{T}\right)^r\right), \\ E\left(\left\|vech(\check{\Omega}_N-\bar{\Omega}_N)\right\|\right)^r &= O\left(\left(\frac{N}{T}\right)^{r/2}\left(\frac{M}{T}\right)^{r/2}\right), \\ \left\|\bar{\Omega}_T-\Omega_0\right\|^r &= O\left(\left(\frac{M}{T}\right)^r\right). \end{aligned}$$

As a result,

$$\begin{aligned} & |\theta|^r E\left|Q\left(G_T,S_T,\check{\Omega}_T,\bar{\Omega}_T\right)-Q\left(G_{T-N},S_{T-N},\check{\Omega}_{T-N},\bar{\Omega}_T\right)\right|^r \\ = & O\left(|\theta^r|\left(\frac{N}{T}\right)^{r/2}\left(\frac{M}{T}\right)^{r/2}\right)+O\left(|\theta^r|\left(\frac{N}{T}\right)^{r/2}\left(\frac{N}{T}\right)^r\right) \end{aligned}$$

using the fact that $Q_N^0(\cdot)$ does not constant and linear terms. But for some $\varpi>0$,

$$\begin{aligned} & \left(|\theta^r|\left(\frac{N}{T}\right)^{r/2}\left(\frac{M}{T}\right)^{r/2}\right)=O\left(|\theta^r|\,N^{r/2}M^{r/2}T^{-r}\right) \\ = & \begin{cases} O\left(M^{r/2}m^rT^{-r/2}\right), & \text{for } T^{\varepsilon}\leq|\theta|\leq\sqrt{T} \\ O\left(|\theta^r|\,M^{r/2}m^rT^{-r}\right), & \text{for } \sqrt{T}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon} \end{cases} \qquad \text{(A.27)} \\ = & \begin{cases} O\left(M/T^{1+\varpi}\right), & \text{for } T^{\varepsilon}\leq|\theta|\leq\sqrt{T} \\ O\left(T^{-r\varepsilon}m^r\right), & \text{for } \sqrt{T}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon} \end{cases} \qquad \text{(A.28)} \\ = & O(M/T^{1+\varpi}), \end{aligned}$$

provided that $\varepsilon r\geq 1+\varpi$. Here we use the fact that $N\theta^2=O(Tm^2)$ when $|\theta|\leq\sqrt{T}$ and $N=O(m^2)$ when $|\theta|\geq\sqrt{T}$. Similarly,

$$\begin{aligned} & |\theta^r|\left(\frac{N}{T}\right)^{r/2}\left(\frac{N}{T}\right)^r \\ = & \begin{cases} O\left(\left(Tm^2\right)^{r/2}T^{-r/2}\right)\left(\frac{N}{T}\right)^r, & \text{for } T^{\varepsilon}\leq|\theta|\leq\sqrt{T} \\ O\left(|\theta^r|\left(\frac{m^2}{T}\right)^{r/2}\left(\frac{m^2}{T}\right)^r\right), & \text{for } \sqrt{T}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon} \end{cases} \\ = & \begin{cases} O\left(m^r\right)\left(\frac{Tm^2}{T}\frac{1}{T^{2\varepsilon}}\right)^r, & \text{for } T^{\varepsilon}\leq|\theta|\leq\sqrt{T} \\ O\left(|\theta^r|\left(\frac{m^2}{T}\right)^{r/2}\left(\frac{m^2}{T}\right)^r\right), & \text{for } \sqrt{T}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon} \end{cases} \\ = & \begin{cases} O\left(m^{3r}T^{-2\varepsilon r}\right), & \text{for } T^{\varepsilon}\leq|\theta|\leq\sqrt{T} \\ O\left(|\theta^r|\left(\frac{m^2}{T}\right)^{r/2}\left(\frac{m^2}{T}\right)^r\right), & \text{for } \sqrt{T}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon} \end{cases} \\ = & o\left(\frac{M}{T}\right). \end{aligned}$$

We have therefore proved that

$$|\theta|^r E\left|Q\left(G_T,S_T,\check{\Omega}_T,\bar{\Omega}_T\right)-Q\left(G_{T-N},S_{T-N},\check{\Omega}_{T-N},\bar{\Omega}_T\right)\right|^r=o\left(M/T\right).$$

In order to evaluate the expansion terms in (A.26), we proceed as in Götze and Künsch (1996, page 1929). Define

$$\left\{t_1^0, ..., t_{9(r-1)}^0\right\} = \{t : \mu_t > 0 \text{ or } \nu_t > 0\},$$

and

$$\mathbb{T} = \left\{t : t \in \{1, 2, ..., N - m\} \text{ and } \left|t - t_j^0\right| \geq 3m, \text{ for } j = 1, 2, ..., 9(r-1)\right\}.$$

Define $\{t_1, ..., t_J\}$ as follows:

$$t_1 = \inf(\mathbb{T}), t_{j+1} = \inf\{j : t \geq t_j + 7m \text{ and } t \in \mathbb{T}\}.$$

and J denotes the smallest integer for which the inf is undefined. Let

$$\begin{aligned}
A_j &= \prod\left\{\exp i\theta \mathbf{a}' v_t/\sqrt{T},\ \ |t - t_j| \leq m, t \in \mathbb{T}\right\},\ j = 1, 2, ..., J \\
B_j &= \prod\left\{\exp i\theta \mathbf{a}' v_t/\sqrt{T},\ t_j + m + 1 \leq t \leq t_{j+1} - m - 1, t \in \mathbb{T}\right\},\ j = 1, 2, ..., J - 1 \\
B_J &= \prod\left\{\exp i\theta \mathbf{a}' v_t/\sqrt{T},\ t \geq t_J + m + 1, t \in \mathbb{T}\right\}, \\
L_T &= \prod\left\{\exp i\theta \mathbf{a}' v_t/\sqrt{T}, t \notin \mathbb{T}\right\}\left\{\exp\left[i\theta Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right]\right\} v^\mu w^\nu \\
&\quad \times Q_{\mu\nu}\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right).
\end{aligned}$$

Then

$$\begin{aligned}
&E\exp\left[i\theta \mathbf{a}' S_T + i\theta Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right] \\
&\times \sum_{\mu,\nu} c_{\mu,\nu}(\theta)\, v^\mu w^\nu Q_{\mu\nu}\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right) = \sum_{\mu,\nu} c_{\mu,\nu}(\theta)\, E\prod_{j=1}^{J} A_j B_j L_T,
\end{aligned}$$

where $|A_j| \leq 1, |B_j| \leq 1$ and $|L_T| = O\left(T^{C_0}\right)$ for some constant $C_0 > 0$. The above multiplicative decomposition is illustrated in Figure 2.

It follows from Assumption 3 that

$$\begin{aligned}
&\left|E\prod_{j=1}^{J} A_j B_j L_T - E\prod_{j=1}^{J} E\left(A_j | \mathcal{F}_t : |t - t_j| \leq 3m\right) B_j L_T\right| \\
\leq\ & \sum_{j=1}^{J}\left|E\prod_{k=1}^{j-1} A_k B_k \left\{A_j - E\left(A_j | \mathcal{F}_t : |t - t_j| \leq 3m\right)\right\} B_j \prod_{\ell=j+1}^{J} E\left(A_\ell | \mathcal{F}_t : |t - t_j| \leq 3m\right) B_\ell L_T\right| \\
\leq\ & \sum_{j=1}^{J}\left|E\prod_{k=1}^{j-1} A_k B_k \left\{A_j - E\left(A_j | \mathcal{F}_t : t \neq j\right)\right\} B_j \prod_{\ell=j+1}^{J} E\left(A_\ell | \mathcal{F}_t : |t - t_j| \leq 3m\right) B_\ell L_T\right| \\
& + \sum_{j=1}^{J}\left|E\prod_{k=1}^{j-1} A_k B_k \left\{E\left(A_j | \mathcal{F}_t : t \neq j\right) - E\left(A_j | \mathcal{F}_t : |t - t_j| \leq 3m\right)\right\} B_j\right. \\
& \left.\times \prod_{\ell=j+1}^{J} E\left(A_\ell | \mathcal{F}_t : |t - t_j| \leq 3m\right) B_\ell L_T\right| \\
\leq\ & \sum_{j=1}^{J}\left|E\prod_{k=1}^{j-1} A_k B_k \left\{A_j - E\left(A_j | \mathcal{F}_t : t \neq j\right)\right\} B_j \prod_{\ell=j+1}^{J} E\left(A_\ell | \mathcal{F}_t : |t - t_j| \leq 3m\right) B_\ell L_T\right| \\
& + d^{-1} T^{C_1} \exp(-dm) \\
=\ & o(T^{-C_2}),
\end{aligned}$$

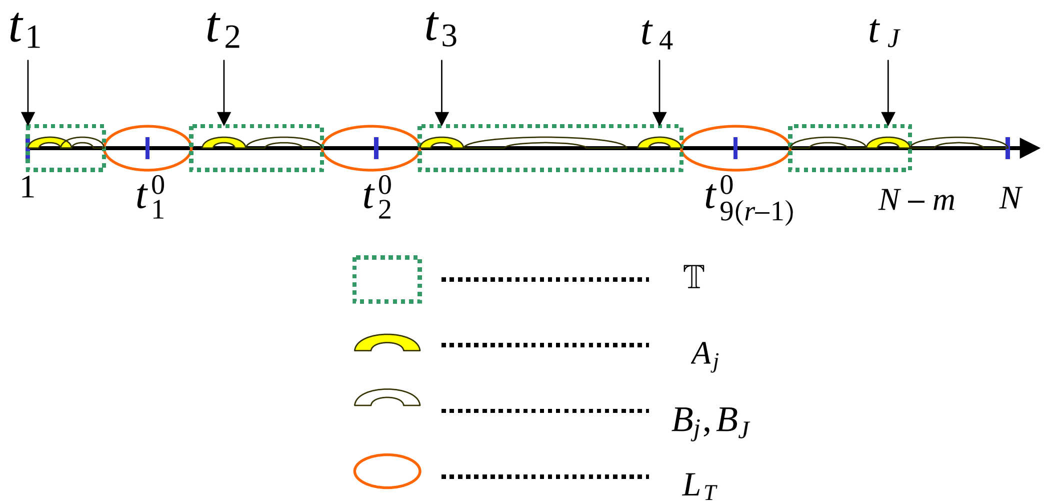


Figure 2: Illustration of the Decomposition

for some $C_1 > 0$ and arbitrary $C_2 > 0$ by choosing $\mathcal{K}$ in the definition of m sufficiently large.

Next, repeated applications of the mixing inequality in (A.11) yields:

$$\begin{aligned}
& \left| E \prod_{j=1}^{J} E\left(A_j | \mathcal{F}_t : |t - t_j| \le 3m\right) B_j L_T \right| \\
\le \quad & T^{C_3} \left| E \prod_{j=1}^{J} \left| E\left(A_j | \mathcal{F}_t : |t - t_j| \le 3m\right) \right| \right| \\
\le \quad & T^{C_3} \left| \prod_{j=1}^{J} E \left| E\left(A_j | \mathcal{F}_t : |t - t_j| \le 3m\right) \right| \right| + O\left(T^{C_3} J d^{-1} \exp\left(-dm\right)\right),
\end{aligned}$$

for some $C_3 > 0$.

By Assumption 4, for $|\theta| > d$, we have

$$E\left|E\left(A_j | \mathcal{F}_t, t \neq j\right)\right| \le \exp(-d).$$

Therefore, using Lemma 3.2 of Götze and Hipp (1983) and Assumption 4, we have

$$\begin{aligned}
& E\left|E\left(A_j | \mathcal{F}_t : |t - t_j| \le 3m\right)\right| \\
\le \quad & E\left|E\left(A_j | \mathcal{F}_t : t \neq t_j\right)\right| + O(d^{-1}\exp(-dm)) \\
\le \quad & \max(\exp(-d), \exp(-d\theta^2/T) + O(d^{-1}\exp(-dm)),
\end{aligned}$$

and thus

$$E \prod_{j=1}^{J} A_j B_j L_T = T^{C} \left\{ \max\left[\exp(-d), \exp\left(-d\frac{\theta^2}{T}\right) \right] \right\}^{\frac{N}{m}} + O\left(T^{-C_4}\right) = O\left(T^{-C_5}\right), \tag{A.29}$$

for arbitrary constant $C_5 > 0$ provided $\mathcal{K}$ in the definition of m is chosen large enough. This implies that

$$E \exp\left[i\theta \mathbf{a}' S_T + i\theta Q\left(G_{T-N}, S_{T-N}, \check{\Omega}_{T-N}, \bar{\Omega}_T\right)\right] = O\left(T^{-C_5}\right).$$

Note that when $\varepsilon < 1/[7(q+1)]$, we have $M^{-1/2}T^{1-\varepsilon} \geq M^{-1}T\log T$ when T is sufficiently large. Combining (A.27) and (A.29) yields

$$\begin{aligned} & \int_{T^{\varepsilon}\leq|\theta|\leq M^{-1}T\log T} |\phi_h(\theta)| \, |\theta|^{-1} \, d\theta \leq \int_{T^{\varepsilon}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon}} |\phi_h(\theta)| \, |\theta|^{-1} \, d\theta \\ = \; & O\left(\frac{M}{T^{1+\varpi}}\right) \int_{T^{\varepsilon}\leq|\theta|\leq M^{-1/2}T^{1-\varepsilon}} |\theta|^{-1} \, d\theta = O\left(\frac{M}{T^{1+\varpi}} \log T^{\varepsilon}\right) = o\left(\frac{M}{T}\right), \end{aligned}$$

which completes the proof of the lemma. ■

Lemma A.12 *The following results hold:*
(a) $2\mu_2 \left\{\mathbf{d}' \left[\left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a}))\right]\right\} = \frac{3}{4}\mu_2 + \frac{\mu_2}{2}(d_2 - d_1)$,
(b) $2\mu_2 \left\{\mathbf{d}' \left(K_{d_2,d_2^2} \otimes I_{d_2^2}\right) (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a}))\right\} = \frac{3}{4}\mu_2 + \frac{\mu_2}{2}(d_2 - d_1)$,
(c) $\mu_2 c' (\Omega_0 \otimes \Omega_0 \otimes \Omega_0) c = \frac{1}{4}\mu_2 + \mu_2 (d_2 - d_1)$,
(d) $\mu_2 c' (K_{d_2,d_2} \otimes I_{d_2}) (\Omega_0 \otimes \Omega_0 \otimes \Omega_0) c = \frac{\mu_2}{4}$,
(e) $12\mu_2 c' [\Omega_0 \otimes \Omega_0 \otimes (\Omega_0 \mathbf{a}\mathbf{a}' \Omega_0)] c = 6\mu_2$
(f) $12\mu_2 c' (K_{d_2,d_2} \otimes I_{d_2}) [\Omega_0 \otimes \Omega_0 \otimes (\Omega_0 \mathbf{a}\mathbf{a}' \Omega_0)] c = 6\mu_2$.

Proof. Part (a) Using the definition of d in Lemma 1, we have

$$\begin{aligned} & 2\mu_2 \left\{\mathbf{d}' \left[\left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a}))\right]\right\} \\ = \; & \frac{\mu_2}{2\sigma_0^3} vec\,[(\Theta_{10}) \otimes (\Theta_{20} - \Theta_{30}) \otimes (\Theta_{10}\Theta_{10}')]' \left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a})) \\ & + \frac{\mu_2}{\sigma_0} vec\,[(\Theta_{10}) \otimes \Theta_{40} \otimes (\Theta_{30} - \Theta_{20})]' \left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a})) \\ & + \frac{\mu_2}{\sigma_0} vec\left(K_{d_2,d_2^2} \left[vec\,(\Theta_{40}) \otimes (\Theta_{30} - \Theta_{20}) \otimes \Theta_{10}'\right]\right)' \left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) \\ & \times (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a})) \\ & - \frac{\mu_2}{2\sigma_0^3} vec\,\{vec\,(\Theta_{10}' \otimes [\Theta_{30} - \Theta_{20}]) [\Theta_{10}' \otimes \Theta_{10}']\} \left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) \\ & \times (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a})) \\ & + \frac{3\mu_2}{4\sigma_0^5} vec\,\{vec\,(\Theta_{10}' \otimes \Theta_{10}\Theta_{10}') (\Theta_{10}' \otimes \Theta_{10}')\} \left(I_{d_2} \otimes K_{d_2,d_2} \otimes I_{d_2^2}\right) \\ & \times (vec\,(\Omega_0) \otimes vec\,(\Omega_0) \otimes (\Omega_0 \mathbf{a})) \\ : \; & = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{A}_4 + \mathcal{A}_5. \end{aligned}$$

Let e_i be the i-th column of I_m and u_j be j-th column of I_n. Then $E_{ij} = e_i u_j'$ is the $m \times n$ matrix with 1 in its (i,j)-th position and zeros elsewhere and the commutation matrix can be written as

$$K_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} E_{ij} \otimes E_{ij}'.$$

See Abadir and Magnus (2005, Exercise 11.8). Using this representation, we have

$$\begin{aligned}
& \mathcal{A}_1 \\
= \; & \frac{2\mu_2}{4\sigma_0^3} vec\left[(\Theta_{10})\otimes(\Theta_{20}-\Theta_{30})\otimes(\Theta_{10}\Theta_{10}')\right]'\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)(vec(\Omega_0)\otimes vec(\Omega_0)\otimes(\Omega_0\mathbf{a})) \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}\left(vec(\Omega_0)'\otimes vec(\Omega_0)'\otimes(\Omega_0\mathbf{a})'\right)\left(I_{d_2}\otimes E_{ij}'\otimes E_{ij}\otimes I_{d_2^2}\right) vec\left[(\Theta_{10})\otimes(\Theta_{20}-\Theta_{30})\otimes(\Theta_{10}\Theta_{10}')\right] \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}\left(vec(\Omega_0)'\otimes vec(\Omega_0)'\otimes(\Omega_0\mathbf{a})'\right) vec\left\{\left[(E_{ij}\Theta_{10})\otimes(\Theta_{20}-\Theta_{30})\otimes(\Theta_{10}\Theta_{10}')\right](I_{d_2}\otimes E_{ij})\right\} \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}\left(vec(\Omega_0)'\otimes vec(\Omega_0)'\otimes(\Omega_0\mathbf{a})'\right) vec\left[(E_{ij}\Theta_{10})\otimes(\Theta_{20}-\Theta_{30})\otimes(\Theta_{10}\Theta_{10}')E_{ij}\right] \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}\left[vec(\Omega_0)'\otimes(\Omega_0\mathbf{a})'\right]\left[(E_{ij}\Theta_{10})\otimes(\Theta_{20}-\Theta_{30})\otimes(\Theta_{10}\Theta_{10}')E_{ij}\right] vec(\Omega_0) \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}\left[vec(\Omega_0)'\otimes(\Omega_0\mathbf{a})'\right] vec\left\{(\Theta_{10}\Theta_{10}')E_{ij}\Omega_0\left[(E_{ij}\Theta_{10})'\otimes(\Theta_{20}-\Theta_{30})'\right]\right\} \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}(\Omega_0\mathbf{a})'(\Theta_{10}\Theta_{10}')E_{ij}\Omega_0\left[(E_{ij}\Theta_{10})'\otimes(\Theta_{20}-\Theta_{30})'\right] vec(\Omega_0) \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}\mathbf{a}'\Omega_0\Theta_{10}\Theta_{10}'E_{ij}\Omega_0\left[(E_{ij}\Theta_{10})'\otimes(\Theta_{20}-\Theta_{30})'\right] vec(\Omega_0) \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}(\mathbf{a}'\Omega_0\Theta_{10})\Theta_{10}'E_{ij}\left[\Omega_0(\Theta_{20}-\Theta_{30})\Omega_0\right]E_{ij}\Theta_{10}
\end{aligned}$$

$$\begin{aligned}
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}(\mathbf{a}'\Omega_0\Theta_{10})\Theta_{10}'e_iu_j'\left[\Omega_0(\Theta_{20}-\Theta_{30})\Omega_0\right]e_iu_j'\Theta_{10} \\
= \; & \frac{\mu_2}{2\sigma_0^3}\sum_{i,j}(\mathbf{a}'\Omega_0\Theta_{10})u_j'\Theta_{10}\Theta_{10}'e_i\left(u_j'\left[\Omega_0(\Theta_{20}-\Theta_{30})\Omega_0\right]e_i\right) \\
= \; & \frac{\mu_2}{\sigma_0}\sum_{i,j}(\Omega_0\mathbf{a})'(\Theta_{30}-\Theta_{20})e_ie_j'\Omega_0\Theta_{40}\Omega_0e_ie_j'\Theta_{10}=0.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathcal{A}_2 \; = \; & \frac{2\mu_2}{2\sigma_0} vec\left[(\Theta_{10})\otimes\Theta_{40}\otimes(\Theta_{30}-\Theta_{20})\right]'\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)(vec(\Omega_0)\otimes vec(\Omega_0)\otimes(\Omega_0\mathbf{a})) \\
= \; & \frac{\mu_2}{\sigma_0}\sum_{i,j}(\Omega_0\mathbf{a})'(\Theta_{30}-\Theta_{20})E_{ij}\Omega_0\left[(E_{ij}\Theta_{10})'\otimes(\Theta_{40})'\right] vec(\Omega_0) \\
= \; & \frac{\mu_2}{\sigma_0}\sum_{i,j}(\Omega_0\mathbf{a})'(\Theta_{30}-\Theta_{20})e_ie_j'\Omega_0\Theta_{40}\Omega_0e_ie_j'\Theta_{10} \\
= \; & \frac{\mu_2}{\sigma_0}\sum_{i,j}\mathbf{a}'\Omega_0(\Theta_{30}-\Theta_{20})e_ie_j'\Theta_{10}\left(e_j'\Omega_0\Theta_{40}\Omega_0e_i\right) \\
= \; & 0.
\end{aligned}$$

Next,

$$
\begin{aligned}
& \mathcal{A}_3 \\
= & \frac{\mu_2}{\sigma_0} vec\left(K_{d_2,d_2^2}\left[vec\left(\Theta_{40}\right)\otimes\left(\Theta_{30}-\Theta_{20}\right)\otimes\Theta_{10}'\right]\right)'\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right) \\
= & \frac{\mu_2}{\sigma_0} vec\left[\left(\Theta_{30}-\Theta_{20}\right)\otimes\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right]'\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left(vec\left(\Omega_0\right)'\otimes vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right)\left(I_{d_2}\otimes E_{ij}'\otimes E_{ij}\otimes I_{d_2^2}\right) vec\left[\left(\Theta_{30}-\Theta_{20}\right)\otimes\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left(vec\left(\Omega_0\right)'\otimes vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right) vec\left\{\left(E_{ij}\otimes I_{d_2^2}\right)\left[\left(\Theta_{30}-\Theta_{20}\right)\otimes\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right]\left(I_{d_2}\otimes E_{ij}\right)\right\} \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]\left\{E_{ij}\left(\Theta_{30}-\Theta_{20}\right)\otimes\left[\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right]E_{ij}\right\} vec\left(\Omega_0\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j} vec\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]' vec\left\{\left[\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right]E_{ij}\Omega_0\left(\Theta_{30}-\Theta_{20}\right)E_{ij}'\right\} \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j} vec\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]' vec\left\{\left[\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right]e_ie_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_je_i'\right\} \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right] vec\left\{\left[\Theta_{10}'\otimes vec\left(\Theta_{40}\right)\right]e_ie_i'\right\} e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right] vec\left\{\left[\Theta_{10}'e_i\otimes vec\left(\Theta_{40}\right)\right]e_i'\right\} e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right] vec\left\{vec\left(\Theta_{40}\right)e_i'\right\}\Theta_{10}'e_ie_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]\left[e_i\otimes vec\left(\Theta_{40}\right)\right]\Theta_{10}'e_ie_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j
\end{aligned}
$$

Let $\Theta_{40}=\sum_{\ell=1}^{d_2} v_\ell\Upsilon_\ell'\Upsilon_\ell$ be the spectral representation of Θ_{40} with ν_ℓ and Υ_ℓ being the corresponding eigenvalues and eigenvectors. Then

$$
\begin{aligned}
\mathcal{A}_3 = & \frac{\mu_2}{\sigma_0}\sum_{i,j,\ell}\nu_\ell\left[vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]\left[e_i\otimes\Upsilon_\ell\otimes\Upsilon_\ell\right]\Theta_{10}'e_ie_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,\ell}\nu_\ell\left[vec\left(\Omega_0\right)'\left(e_i\otimes\Upsilon_\ell\right)\right]\left(\mathbf{a}'\Omega_0\Upsilon_\ell\right)\left(\Theta_{10}'e_i\right)\left[e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,\ell}\nu_\ell\left[\left(e_i'\otimes\Upsilon_\ell'\right)vec\left(\Omega_0\right)\right]'\left(\mathbf{a}'\Omega_0\Upsilon_\ell\right)\left(\Theta_{10}'e_i\right)\left[e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,\ell}\nu_\ell\left[\Upsilon_\ell'\Omega_0e_i\right]'\left(\mathbf{a}'\Omega_0\Upsilon_\ell\right)\left(\Theta_{10}'e_i\right)\left[e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,\ell}\nu_\ell\left[e_i'\Omega_0\Upsilon_\ell\right]\left(\Upsilon_\ell'\Omega_0\mathbf{a}\right)\left(\Theta_{10}'e_i\right)\left[e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j} e_i'\Omega_0\left(\left(\Theta_{30}-\Theta_{20}\right)\right)\Omega_0\mathbf{a}\left(\Theta_{10}'e_i\right)\left[e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j\right] \\
& -\frac{\mu_2}{2\sigma_0^3}\sum_{i,j} e_i'\Omega_0\Theta_{10}\Theta_{10}'\Omega_0\mathbf{a}\Theta_{10}'e_i\left[e_j'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)e_j\right] \\
= & -\frac{\mu_2}{2}tr\left[\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\right]=\frac{\mu_2}{2}\left(d_2-d_1\right).
\end{aligned}
$$

Similarly, letting $\Theta_{30}-\Theta_{20}=\sum_{\ell=1}^{d_2}\upsilon_\ell\Xi_\ell\Xi_\ell'$ be the spectral representation of $\Theta_{30}-\Theta_{20}$, we have

$$\begin{aligned}
\mathcal{A}_4 &= -\frac{\mu_2}{2\sigma_0^3}vec\left\{vec\left(\Theta_{10}'\otimes\left[\Theta_{30}-\Theta_{20}\right]\right)\left[\Theta_{10}'\otimes\Theta_{10}'\right]\right\}'\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&= -\frac{\mu_2}{2\sigma_0^3}\left(\Theta_{10}'\otimes\Theta_{10}'\otimes\Theta_{10}'\otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right)\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&= -\frac{\mu_2}{2\sigma_0^3}\left(\Theta_{10}'\otimes\left(\Theta_{10}'\otimes\Theta_{10}'\right)K_{d_2,d_2}\otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&= -\frac{\mu_2}{2\sigma_0^3}\left(\Theta_{10}'\otimes\left(\Theta_{10}'\otimes\Theta_{10}'\right)\otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&= -\frac{\mu_2}{2\sigma_0^3}\sum_\ell\upsilon_\ell\left[\Theta_{10}'\otimes\Theta_{10}'\right]vec\left(\Omega_0\right)\left(\Theta_{10}'\otimes\Xi_\ell'\otimes\Xi_\ell'\right)\left[vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right]\\
&= -\frac{\mu_2}{2\sigma_0^3}\sum_\ell\upsilon_\ell\Theta_{10}'\Omega_0\Theta_{10}\Theta_{10}'\Omega_0\Xi_\ell\Xi_\ell'\Omega_0\mathbf{a}\\
&= -\frac{\mu_2}{2\sigma_0^3}\Theta_{10}'\Omega_0\Theta_{10}\Theta_{10}'\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0\mathbf{a}=0.
\end{aligned}$$

Finally,

$$\begin{aligned}
\mathcal{A}_5 &= \frac{3\mu_2}{4\sigma_0^5}\left[\left(vec\left(\Omega_0\right)'\otimes vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right)\left(K_{d_2,d_2^2}'\otimes I_{d_2^2}\right)\right]vec\left[vec\left[\Theta_{10}'\otimes\left(\Theta_{10}\Theta_{10}'\right)\right]\left(\Theta_{10}'\otimes\Theta_{10}'\right)\right]\\
&= \frac{3\mu_2}{4\sigma_0^5}\left[\left(vec\left(\Omega_0\right)'\otimes vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right)\left(K_{d_2,d_2^2}'\otimes I_{d_2^2}\right)\right]\left(\Theta_{10}\otimes\Theta_{10}\right)\otimes vec\left[\Theta_{10}'\otimes\left(\Theta_{10}\Theta_{10}'\right)\right]\\
&= \frac{3\mu_2}{4\sigma_0^5}\left[\left(vec\left(\Omega_0\right)'\otimes vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right)\left(K_{d_2,d_2^2}'\otimes I_{d_2^2}\right)\right]\left(\Theta_{10}\otimes\Theta_{10}\right)\otimes\Theta_{10}\otimes\Theta_{10}\otimes\Theta_{10}\\
&= \frac{3\mu_2}{4\sigma_0^5}\left[\left(vec\left(\Omega_0\right)'\otimes vec\left(\Omega_0\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right)\right]\left(\Theta_{10}\otimes\Theta_{10}\right)\otimes\Theta_{10}\otimes\Theta_{10}\otimes\Theta_{10}\\
&= \frac{3}{4}\mu_2.
\end{aligned}$$

Therefore,

$$2\mu_2\left\{\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\right]\right\}=\frac{3}{4}\mu_2+\frac{\mu_2}{2}\left(d_2-d_1\right),$$

as desired.

Part (b). We sketch the proof, starting with

$$\begin{aligned}
&2\mu_2\left\{\mathbf{d}'\left[\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\right]\right\}\\
=\ &\frac{\mu_2}{2\sigma_0^3}vec\left[\left(\Theta_{10}\right)\otimes\left(\Theta_{20}-\Theta_{30}\right)\otimes\left(\Theta_{10}\Theta_{10}'\right)\right]'\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&+\frac{\mu_2}{\sigma_0}vec\left[\left(\Theta_{10}\right)\otimes\Theta_{40}\otimes\left(\Theta_{30}-\Theta_{20}\right)\right]'\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&+\frac{\mu_2}{\sigma_0}vec\left(K_{d_2,d_2^2}\left[vec\left(\Theta_{40}\right)\otimes\left(\Theta_{30}-\Theta_{20}\right)\otimes\Theta_{10}'\right]\right)'\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&-\frac{\mu_2}{2\sigma_0^3}vec\left\{vec\left(\Theta_{10}'\otimes\left[\Theta_{30}-\Theta_{20}\right]\right)\left[\Theta_{10}'\otimes\Theta_{10}'\right]\right\}\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
&+\frac{3\mu_2}{4\sigma_0^5}vec\left\{vec\left(\Theta_{10}'\otimes\Theta_{10}\Theta_{10}'\right)\left(\Theta_{10}'\otimes\Theta_{10}'\right)\right\}\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\\
=\ &:\mathcal{B}_1+\mathcal{B}_2+\mathcal{B}_3+\mathcal{B}_4+\mathcal{B}_5.
\end{aligned}$$

Now, with some abuse of notation, we write $K_{d_2,d_2^2} = \sum_{i,j} \left(E_{ij} \otimes E'_{ij}\right)$. So

$$
\begin{aligned}
& \mathcal{B}_1 \\
= \;& \frac{\mu_2}{2\sigma_0^3} \left(vec\left(\Omega_0\right)' \otimes vec\left(\Omega_0\right)' \otimes \left(\Omega_0 \mathbf{a}\right)'\right) \left(K_{d_2^2,d_2} \otimes I_{d_2^2}\right) vec\left[\left(\Theta_{10}\right) \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta'_{10}\right)\right] \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left(vec\left(\Omega_0\right)' \otimes vec\left(\Omega_0\right)' \otimes \left(\Omega_0 \mathbf{a}\right)'\right) \left(E_{ij} \otimes E'_{ij} \otimes I_{d_2^2}\right) vec\left[\left(\Theta_{10}\right) \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta'_{10}\right)\right] \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left[vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0\right)' \left(E'_{ij} \otimes I_{d_2}\right) \otimes \left(\Omega_0 \mathbf{a}\right)'\right] vec\left[\left(\Theta_{10}\right) \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta'_{10}\right)\right] \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left[vec\left(\Omega_0\right)' E_{ij} \otimes \left[\left(E_{ij} \otimes I_{d_2}\right) vec\left(\Omega_0\right)\right]' \otimes \left(\Omega_0 \mathbf{a}\right)'\right] vec\left[\left(\Theta_{10}\right) \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta'_{10}\right)\right] \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left[vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E'_{ij}\right)' \otimes \left(\Omega_0 \mathbf{a}\right)'\right] vec\left[\Theta_{10} \otimes \left(\Theta_{20} - \Theta_{30}\right) \otimes \left(\Theta_{10}\Theta'_{10}\right)\right] \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left[vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E'_{ij}\right)' \otimes \left(\Omega_0 \mathbf{a}\right)'\right] \left(I_{d_2} \otimes K_{d_2,d_2^2} \otimes I_{d_2}\right) \\
& \times \left\{vec\left[\Theta_{10} \otimes \left(\Theta_{20} - \Theta_{30}\right)\right] \otimes \Theta_{10} \otimes \Theta_{10}\right\} \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left[vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E'_{ij}\right)'\right] \\
& \times \left\{\left(I_{d_2} \otimes K_{d_2,d_2^2}\right) \left(vec\left[\Theta_{10} \otimes \left(\Theta_{20} - \Theta_{30}\right)\right] \otimes \Theta_{10}\right)\right\} \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& \frac{\mu_2}{2\sigma_0^3} \sum_{i,j} \left[vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E'_{ij}\right)' K_{d_2,d_2^2}\right] \\
& \times \left(K_{d_2,d_2} \otimes I_{d_2} \otimes I_{d_2}\right) \left[\Theta_{10} \otimes vec\left(\Theta_{20} - \Theta_{30}\right) \otimes \Theta_{10}\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right).
\end{aligned}
$$

Plugging $\Theta_{30} - \Theta_{20} = \sum_{\ell=1}^{d_2} \upsilon_\ell \Xi_\ell \Xi'_\ell$ into the above equation, we have

$$
\begin{aligned}
& \mathcal{B}_1 \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E'_{ij}\right)' K_{d_2,d_2^2}\right] \left[\Xi_\ell \otimes \Theta_{10} \otimes \Xi_\ell \otimes \Theta_{10}\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[vec\left(\Omega_0\right)' E_{ij}\Xi_\ell \otimes vec\left(\Omega_0 E'_{ij}\right)' K_{d_2,d_2^2} \left\{\Theta_{10} \otimes \Xi_\ell \otimes \Theta_{10}\right\}\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[vec\left(\Omega_0\right)' E_{ij}\Xi_\ell\right] \left[vec\left(\Omega_0 E'_{ij}\right)' \left\{\Theta_{10} \otimes \Theta_{10} \otimes \Xi_\ell\right\}\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[vec\left(\Omega_0\right)' E_{ij}\Xi_\ell\right] \left[\Xi'_\ell \Omega_0 E'_{ij} \left(\Theta_{10} \otimes \Theta_{10}\right)\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[\Xi'_\ell E'_{ij} vec\left(\Omega_0\right)\right] \left[\Xi'_\ell \Omega_0 u_j e'_i \left(\Theta_{10} \otimes \Theta_{10}\right)\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[\Xi'_\ell u_j e'_i vec\left(\Omega_0\right)\right] \left[\Xi'_\ell \Omega_0 u_j\right] e'_i \left(\Theta_{10} \otimes \Theta_{10}\right) \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,\ell} \upsilon_\ell \left[u'_j \left(\Xi_\ell \Xi'_\ell \Omega_0\right) u_j\right] \left[e'_i \left(\Theta_{10} \otimes \Theta_{10}\right) vec\left(\Omega_0\right)' e_i\right] \left(\mathbf{a}'\Omega_0\Theta_{10}\right) \\
= \;& \frac{\mu_2}{2\sigma_0^3} tr\left(\left(\Theta_{20} - \Theta_{30}\right) \Omega_0\right) \; \left(\Theta'_{10}\Omega_0\Theta_{10}\right) \left(\mathbf{a}'\Omega_0\Theta_{10}\right) = \frac{\mu_2}{2} \left(d_2 - d_1\right).
\end{aligned}
$$

Next, some tedious calculations show that

$$
\begin{aligned}
& \mathcal{B}_2 \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)' E_{ij}\otimes vec\left(\Omega_0 E_{ij}'\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right] vec\left[\Theta_{10}\otimes\Theta_{40}\otimes\left(\Theta_{30}-\Theta_{20}\right)\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j}\left[vec\left(\Omega_0\right)' E_{ij}\otimes vec\left(\Omega_0 E_{ij}'\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]\left(I_{d_2}\otimes K_{d_2,d_2^2}\otimes I_{d_2}\right)\left(K_{d_2,d_2}\otimes I_{d_2}\otimes I_{d_2}\right) \\
& \times\left[\Theta_{10}\otimes vec\left(\Theta_{40}\right)\otimes vec\left(\Theta_{30}-\Theta_{20}\right)\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left[vec\left(\Omega_0\right)' E_{ij}\otimes\left[vec\left(\Omega_0 E_{ij}'\right)' K_{d_2,d_2^2}\right]\otimes\left(\Omega_0\mathbf{a}\right)'\right] \\
& \times\left\{\left(K_{d_2,d_2}\otimes I_{d_2}\right)\left[\Theta_{10}\otimes vec\left(\Theta_{40}\right)\right]\otimes\Xi_k\otimes\Xi_k\right\} \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left(vec\left(\Omega_0\right)' E_{ij}\otimes\left[vec\left(\Omega_0 E_{ij}'\right)' K_{d_2,d_2^2}\right]\right)\left\{\left(K_{d_2,d_2}\otimes I_{d_2}\right)\left[\Theta_{10}\otimes\Upsilon_\ell\otimes\Upsilon_\ell\right]\otimes\Xi_k\right\}\left(\mathbf{a}'\Omega_0\Xi_k\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left[vec\left(\Omega_0\right)' E_{ij}\otimes\left[vec\left(\Omega_0 E_{ij}'\right)' K_{d_2,d_2^2}\right]\right]\left\{\left(\Upsilon_\ell\otimes\Theta_{10}\otimes\Upsilon_\ell\right)\otimes\Xi_k\right\}\left(\mathbf{a}'\Omega_0\Xi_k\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left[vec\left(\Omega_0\right)' E_{ij}\Upsilon_\ell\right] vec\left(\Omega_0 E_{ij}'\right)' K_{d_2,d_2^2}\left\{\Theta_{10}\otimes\Upsilon_\ell\otimes\Xi_k\right\}\left(\mathbf{a}'\Omega_0\Xi_k\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left[vec\left(\Omega_0\right)' E_{ij}\Upsilon_\ell\right] vec\left(\Omega_0 E_{ij}'\right)'\left\{\Xi_k\otimes\Theta_{10}\otimes\Upsilon_\ell\right\}\left(\mathbf{a}'\Omega_0\Xi_k\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left[vec\left(\Omega_0\right)' E_{ij}\Upsilon_\ell\right]\Upsilon_\ell'\Omega_0 E_{ij}'\left(\Xi_k\otimes\Theta_{10}\right)\left(\mathbf{a}'\Omega_0\Xi_k\right)
\end{aligned}
$$

$$
\begin{aligned}
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\nu_\ell\upsilon_k\left[vec\left(\Omega_0\right)' E_{ij}\Upsilon_\ell\right]\Upsilon_\ell'\Omega_0 E_{ij}'\left(\Xi_k\Xi_k'\Omega_0\mathbf{a}\otimes\Theta_{10}\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j} vec\left(\Omega_0\right)' E_{ij}\Theta_{40}\Omega_0 E_{ij}'\left[\left(\Theta_{30}-\Theta_{20}\right)\Omega_0\mathbf{a}\otimes\Theta_{10}\right)\right] = 0,
\end{aligned}
$$

$$
\begin{aligned}
& \mathcal{B}_3 \\
= & \frac{\mu_2}{\sigma_0} vec\left(K_{d_2,d_2^2}\left[vec\left(\Theta_{40}\right)\otimes\left(\Theta_{30}-\Theta_{20}\right)\otimes\Theta_{10}'\right]\right)'\left(\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\upsilon_k\nu_\ell\left[vec\left(\Omega_0\right)' E_{ij}\otimes vec\left(\Omega_0 E_{ij}'\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left[\Xi_k\otimes\Xi_k\otimes\Theta_{10}\otimes\Upsilon_\ell\otimes\Upsilon_\ell\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\upsilon_k\nu_\ell\left[vec\left(\Omega_0\right)' E_{ij}\otimes vec\left(\Omega_0 E_{ij}'\right)'\otimes\left(\Omega_0\mathbf{a}\right)'\right]\left[\Xi_k\otimes\Theta_{10}\otimes\Xi_k\otimes\Upsilon_\ell\otimes\Upsilon_\ell\right] \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\upsilon_k\nu_\ell\left[\mathbf{a}'\Omega_0\Upsilon_\ell\right]\left[vec\left(\Omega_0\right)' E_{ij}X_k\right] vec\left(\Omega_0 E_{ij}'\right)'\left(\Theta_{10}\otimes X_k\otimes Y_\ell\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\upsilon_k\nu_\ell\left[\mathbf{a}'\Omega_0\Upsilon_\ell\right]\left[vec\left(\Omega_0\right)' E_{ij}\Xi_k\right]\left(\Theta_{10}'\otimes\Xi_k'\otimes\Upsilon_\ell'\right) vec\left(\Omega_0 E_{ij}'\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\upsilon_k\nu_\ell\left[\mathbf{a}'\Omega_0\Upsilon_\ell\right]\left[vec\left(\Omega_0\right)' E_{ij}\Xi_k\right]\Upsilon_\ell'\Omega_0 E_{ij}'\left(\Theta_{10}\otimes\Xi_k\right) \\
= & \frac{\mu_2}{\sigma_0}\sum_{i,j,k,\ell}\upsilon_k\nu_\ell\left[\mathbf{a}'\Omega_0\Upsilon_\ell\right]\left[vec\left(\Omega_0\right)' e_i u_j'\Xi_k\right]\Upsilon_\ell'\Omega_0 u_j e_i'\left(\Theta_{10}\otimes\Xi_k\right) \\
= & \frac{\mu_2}{\sigma_0}\mathbf{a}'\Omega_0\Theta_{40}\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0\Theta_{10} = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
& \mathcal{B}_4 \\
= & -\frac{\mu_2}{2\sigma_0^3} vec\left\{vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)[\Theta_{10}' \otimes \Theta_{10}']\right\}\left(\left(K_{d_2,d_2^2} \otimes I_{d_2^2}\right)\right)\left(vec\left(\Omega_0\right) \otimes vec\left(\Omega_0\right) \otimes \left(\Omega_0 \mathbf{a}\right)\right) \\
= & -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,k}\left\{vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E_{ij}'\right)' \otimes \left(\Omega_0 \mathbf{a}\right)'\right\} vec\left\{vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)[\Theta_{10}' \otimes \Theta_{10}']\right\} \\
= & -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,k} \upsilon_k\left\{vec\left(\Omega_0\right)' E_{ij} \otimes vec\left(\Omega_0 E_{ij}'\right)' \otimes \left(\Omega_0 \mathbf{a}\right)'\right\}\left\{\Theta_{10} \otimes \Theta_{10} \otimes \Theta_{10} \otimes \Xi_k \otimes \Xi_k\right\} \\
= & -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,k} \upsilon_k vec\left(\Omega_0\right)' E_{ij} \Theta_{10} vec\left(\Omega_0 E_{ij}'\right)'\left(\Theta_{10} \otimes \Theta_{10} \otimes \Xi_k\right)\left(\mathbf{a}' \Omega_0 \Xi_k\right) \\
= & -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,k} \upsilon_k vec\left(\Omega_0\right)' E_{ij} \Theta_{10}\left[\left(\Xi_k' \Omega_0 E_{ij}'\left(\Theta_{10} \otimes \Theta_{10}\right)\right)\right]\left(\mathbf{a}' \Omega_0 \Xi_k\right) \\
= & -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j,k} \upsilon_k vec\left(\Omega_0\right)' e_i u_j' \Theta_{10}\left[\left(\Xi_k' \Omega_0 u_j e_i'\left(\Theta_{10} \otimes \Theta_{10}\right)\right)\right]\left(\mathbf{a}' \Omega_0 \Xi_k\right) \\
= & -\frac{\mu_2}{2\sigma_0^3} \sum_{i,j} vec\left(\Omega_0\right)' e_i u_j' \Theta_{10}\left[e_i'\left(\Theta_{10} \otimes \Theta_{10}\right)\right] \mathbf{a}' \Omega_0\left(\Theta_{30}-\Theta_{20}\right) \Omega_0 u_j=0.
\end{aligned}
$$

The same derivation leading to $A_5 = \frac{3}{4}\mu_2$ gives us $B_5 = \frac{3}{4}\mu_2$. Combining the above results, we obtain

$$
2\mu_2\left\{\mathbf{d}'\left(K_{d_2,d_2^2} \otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right) \otimes vec\left(\Omega_0\right) \otimes \left(\Omega_0 \mathbf{a}\right)\right)\right\}=\frac{3}{4}\mu_2+\frac{\mu_2}{2}\left(d_2-d_1\right)
$$

as stated.

Part (c) Using the definition of c in Lemma 1, we can write

$$
\begin{aligned}
& \mu_2 \mathbf{c}'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) \mathbf{c} \\
= & \frac{1}{4\sigma_0^6} \mu_2 vec\left(\Theta_{10}' \otimes \left(\Theta_{10} \Theta_{10}'\right)\right)'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \left(\Theta_{10} \Theta_{10}'\right)\right) \\
& +\frac{1}{\sigma_0^2} \mu_2\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \Theta_{30}-\Theta_{20}\right) \\
& -\frac{1}{\sigma_0^4} \mu_2\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \Theta_{10} \Theta_{10}'\right) \\
: & =\mathcal{C}_1+\mathcal{C}_2+\mathcal{C}_3,
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{C}_1 & =\frac{\mu_2}{4\sigma_0^6} vec\left(\Theta_{10}' \otimes \Theta_{10} \Theta_{10}'\right)'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \Theta_{10} \Theta_{10}'\right) \\
& =\frac{\mu_2}{4\sigma_0^6}\left(\Theta_{10}' \otimes \Theta_{10}' \otimes \Theta_{10}'\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\left(\Theta_{10} \otimes \Theta_{10} \otimes \Theta_{10}\right) \\
& =\frac{\mu_2}{4\sigma_0^6}\left(\Theta_{10}' \Omega_0 \Theta_{10}\right)^3 \\
& =\frac{\mu_2}{4},
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{C}_2 = \frac{\mu_2}{\sigma_0^2}\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]' \left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\left(\Theta_{10} \otimes vec\left(\Theta_{30}-\Theta_{20}\right)\right) \\
= \frac{\mu_2}{\sigma_0^2}\left(\Theta_{10}'\Omega_0^{1/2} \otimes vec\left[\Omega_0^{1/2}\left(\Theta_{30}-\Theta_{20}\right)\Omega_0^{1/2}\right]'\right)\left(\Omega_0^{1/2}\Theta_{10} \otimes vec\left[\Omega_0^{1/2}\left(\Theta_{30}-\Theta_{20}\right)\Omega_0^{1/2}\right]\right) \\
= \frac{\mu_2}{\sigma_0^2}\Theta_{10}'\Omega_0\Theta_{10} vec\left[\Omega_0^{1/2}\left(\Theta_{30}-\Theta_{20}\right)\Omega_0^{1/2}\right]' vec\left[\Omega_0^{1/2}\left(\Theta_{30}-\Theta_{20}\right)\Omega_0^{1/2}\right] \\
= \frac{\mu_2}{\sigma_0^2}\Theta_{10}'\Omega_0\Theta_{10} tr\left\{\left[\Omega_0^{1/2}\left(\Theta_{30}-\Theta_{20}\right)\Omega_0^{1/2}\right]\left[\Omega_0^{1/2}\left(\Theta_{30}-\Theta_{20}\right)\Omega_0^{1/2}\right]\right\} \\
= \frac{\mu_2}{\sigma_0^2}\Theta_{10}'\Omega_0\Theta_{10} tr\left\{\left(\Theta_{30}-\Theta_{20}\right)\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0\right\} \\
= \frac{\mu_2}{\sigma_0^2}\Theta_{10}'\Omega_0\Theta_{10}\left[tr\left(\Sigma_0 G_0'\Omega_0^{-1}G_0\right) - 2tr\left(\Omega_0^{-1}G_0\Sigma_0 G_0'\right) + tr(I_{d_2})\right] \\
= \frac{\mu_2}{\sigma_0^2}\Theta_{10}'\Omega_0\Theta_{10}\left[d_1 - 2d_1 + d_2\right] \\
= \mu_2\left(d_2 - d_1\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{C}_3 = -\frac{1}{\sigma_0^4}\mu_2\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]'\left(\Omega_0\Theta_{10} \otimes \Omega_0\Theta_{10} \otimes \Omega_0\Theta_{10}\right) \\
= -\frac{1}{\sigma_0^4}\mu_2\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]'\left(G_0\Sigma_0 R \otimes G_0\Sigma_0 R \otimes G_0\Sigma_0 R\right) \\
= -\frac{1}{\sigma_0^4}\mu_2 vec(I_{d_2})'\left\{\left[\Theta_{10}' \otimes \left(\Theta_{30}-\Theta_{20}\right)\right] \otimes I_{d_2}\right\}\left(G_0\Sigma_0 R \otimes G_0\Sigma_0 R \otimes G_0\Sigma_0 R\right) = 0.
\end{aligned}
$$

Therefore

$$
\mu_2 \mathbf{c}'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c} = \frac{1}{4}\mu_2 + \mu_2\left(d_2 - d_1\right).
$$

Part (d) We write

$$
\begin{aligned}
& \mu_2 \mathbf{c}'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c} \\
= & \frac{1}{4\sigma_0^6}\mu_2 vec\left(\Theta_{10}' \otimes \left(\Theta_{10}\Theta_{10}'\right)\right)'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \left(\Theta_{10}\Theta_{10}'\right)\right) \\
& + \frac{1}{\sigma_0^2}\mu_2\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \Theta_{30}-\Theta_{20}\right) \\
& - \frac{1}{\sigma_0^4}\mu_2\left[vec\left(\Theta_{10}' \otimes [\Theta_{30}-\Theta_{20}]\right)\right]'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{10}' \otimes \Theta_{10}\Theta_{10}'\right) \\
: & = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{D}_3,
\end{aligned}
$$

where it is easy to show that

$$
\mathcal{D}_1 = \mu_2/4,
$$

$$
\begin{aligned}
\mathcal{D}_3 &= -\frac{1}{\sigma_0^4}\mu_2 vec(I_{d_2})'\left\{\left[\Theta_{10}' \otimes \left(\Theta_{30}-\Theta_{20}\right)\right] \otimes I_{d_2}\right\}\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(G_0\Sigma_0 R \otimes G_0\Sigma_0 R \otimes G_0\Sigma_0 R\right) \\
&= -\frac{1}{\sigma_0^4}\mu_2 vec(I_{d_2})'\left\{\left[\Theta_{10}' \otimes \left(\Theta_{30}-\Theta_{20}\right)\right] \otimes I_{d_2}\right\}\left(G_0\Sigma_0 R \otimes G_0\Sigma_0 R \otimes G_0\Sigma_0 R\right) = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{D}_2 &= \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}' \otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right]\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\left(\Theta_{10} \otimes vec\left(\Theta_{30}-\Theta_{20}\right)\right) \\
&= \sum_{i,j} \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}' \otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right]\left(E_{ij} \otimes E_{ij}' \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\left(\Theta_{10} \otimes vec\left(\Theta_{30}-\Theta_{20}\right)\right) \\
&= \sum_{i,j} \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}' \otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right]\left[\left(E_{ij}\Omega_0\Theta_{10}\right) \otimes \left(E_{ij}'\Omega_0 \otimes \Omega_0\right) vec\left(\Theta_{30}-\Theta_{20}\right)\right] \\
&= \sum_{i,j} \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}' \otimes vec\left(\Theta_{30}-\Theta_{20}\right)'\right]\left[\left(E_{ij}\Omega_0\Theta_{10}\right) \otimes vec\left(\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0 E_{ij}\right)\right] \\
&= \sum_{i,j} \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}'\left(E_{ij}\Omega_0\right)\Theta_{10}\right] vec\left(\Theta_{30}-\Theta_{20}\right)' vec\left(\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0 E_{ij}\right) \\
&= \sum_{i,j} \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}'\left(E_{ij}\Omega_0\right)\Theta_{10}\right] tr\left[\left(\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0 E_{ij}\left(\Theta_{30}-\Theta_{20}\right)'\right)\right] \\
&= \sum_{i,j} \frac{\mu_2}{\sigma_0^2}\left[\Theta_{10}'\left(E_{ij}\Omega_0\right)\Theta_{10}\right] tr\left[\left(\Theta_{30}-\Theta_{20}\right)\Omega_0\left(\Theta_{30}-\Theta_{20}\right)\Omega_0 E_{ij}\right] = 0.
\end{aligned}
$$

Hence,

$$
\mu_2 \mathbf{c}'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c} = \frac{\mu_2}{4}.
$$

Parts (e) and (f) can be proved analogously and the details are omitted. ■

A.3 Proofs of the Main Results

Proof of Theorem 2. The theorem follows from Lemmas A.9 to A.11. Note that some notational changes are made to simplify the presentation. Let $\mathbf{a} = vec\left(Q_a\right)$, $\mathbf{b} = vec\left(Q_b\right)$, $\mathbf{c} = vec\left(Q_c\right)$, $\mathbf{d} = vec\left(Q_d\right)$. Then $\rho_{i,\infty}$ and $\kappa_{i,\infty}$ will not change if $\mathbf{a}_h, \mathbf{b}_h, \mathbf{c}_h, \mathbf{d}_h$ are replaced by $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ and the 'vech' operator is replaced by 'vec' operator. In fact, $\mathbf{a} = \mathbf{a}_h$ and $\mathbf{b} = \mathbf{b}_h$. ■

Proof of Proposition 3. Part (a) Some algebraic manipulations yield:

$$
\begin{aligned}
\rho_{1,\infty} &= \frac{2}{\sigma_0} \lim_{T\to\infty} E\mathbf{a}' S_T S_T' \left[\Theta_{10}' \otimes \Theta_{40}\right] \lim_{T\to\infty} M^q vec\left(\bar{\Omega}_T - \Omega_0\right)\right] \\
&= \frac{2}{\sigma_0}\mathbf{a}'\Omega_0\left[\Theta_{10}' \otimes \Theta_{40}\right] \lim_{T\to\infty} M^q vec\left(\bar{\Omega}_T - \Omega_0\right)\right] \\
&= -\frac{2g_q}{\sigma_0}\mathbf{a}'\Omega_0\left[\Theta_{10}' \otimes \Theta_{40}\right] vec(\Omega_0^{(q)}) \\
&= -2g_q \Theta_{10}'\Omega_0\Theta_{40}\Omega_0^{(q)}\Theta_{10}/\sigma_0^2 \\
&= g_q \Theta_{10}'\Omega_0\Theta_{10}\Theta_{10}'\Omega_0^{(q)}\Theta_{10}/\sigma_0^4 \\
&= g_q \frac{\mathcal{R}'\Sigma_0 G_0'\Omega_0^{-1}\Omega_0^{(q)}\Omega_0^{-1}G_0\Sigma_0\mathcal{R}}{\mathcal{R}'G_0^{-1}\Omega_0\left(G_0^{-1}\right)'\mathcal{R}},
\end{aligned}
$$

as stated.

Part (b) We start with the identity:

$$
\begin{aligned}
\rho_{2,\infty} &= 2\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)\left\{\mathbf{c}'[vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T]\right\} \\
&\quad +2\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)\left\{\mathbf{c}'[vec\left(\check{\Omega}_T - \tilde{\Omega}_T\right) \otimes S_T]\right\}. \qquad \text{(A.30)}
\end{aligned}
$$

Using the BN decomposition as in Phillips and Solo (1992), we write

$$v_t = \Psi e_t + \tilde{v}_{t-1} - \tilde{v}_t,$$

where $\Psi = \sum_{s=0}^{\infty} \Psi_s$ and $\tilde{v}_t$ is a stationary Gaussian process. Based on this decomposition, we can show that the first term in (A.30) is

$$\begin{aligned}
& 2 \lim_{T\to\infty} \frac{T}{M} E\left(\mathbf{a}' S_T\right) \left\{ \mathbf{c}' [vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right) \otimes S_T] \right\} \\
= \; & 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{t,m,n=1}^{T} E\mathbf{c}'[vec\left(v_t v_t' - E v_t v_t'\right) \otimes \left(v_m v_n' \mathbf{a}\right)] \\
& +2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t,m,n=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'[vec\left(v_s v_t'\right) \otimes \left(v_m v_n' \mathbf{a}\right)] \\
= \; & 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t,m,n=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'[v_t \otimes v_s \otimes \left(v_m v_n' \mathbf{a}\right)] \\
= \; & 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t,m,n=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left[\Psi e_t \otimes \Psi e_s \otimes \left(\Psi e_m e_n' \Psi' \mathbf{a}\right)\right] \\
= \; & 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left[\Psi e_t \otimes \Psi e_s \otimes \left(\Psi e_s e_t' \Psi' \mathbf{a}\right)\right] \\
& +2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left[\Psi e_t \otimes \Psi e_s \otimes \left(\Psi e_t e_s' \Psi' \mathbf{a}\right)\right] \\
= \; & \mathbb{A}_1 + \mathbb{A}_2,
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{A}_1 &= 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left(\Psi e_t \otimes \Psi e_s \otimes \Psi e_s e_t' \Psi' \mathbf{a}\right) \\
&= 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left[\Psi e_t e_t' \Psi' \mathbf{a} \otimes \Psi e_s \otimes \Psi e_s\right] \\
&= \left(2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} k\left(\frac{s-t}{M}\right)\right) E\mathbf{c}'\left[\Psi e_t e_t' \Psi' \mathbf{a} \otimes vec\left(\Psi e_s e_s' \Psi\right)\right] \\
&= 2\mu_1 \mathbf{c}'\left[\Omega_0 \mathbf{a} \otimes vec\left(\Omega_0\right)\right],
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{A}_2 &= 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left[\left(\Psi e_t\right) \otimes \left(\Psi e_s\right) \otimes \left(\Psi e_t e_s' \Psi' \mathbf{a}\right)\right] \\
&= 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}'\left[\left(\Psi e_t\right) \otimes \left(\Psi e_s e_s' \Psi' \mathbf{a}\right) \otimes \left(\Psi e_t\right)\right] \\
&= 2 \lim_{T\to\infty} \frac{1}{MT} \sum_{s,t=1}^{T} 1\left\{s \neq t\right\} k\left(\frac{s-t}{M}\right) E\mathbf{c}' K_{d_2^2,d_2}\left[\left(\Psi e_s e_s' \Psi' \mathbf{a}\right) \otimes vec\left(\Psi e_t e_t' \Psi\right)\right] \\
&= 2\mu_1 \mathbf{c}' K_{d_2^2,d_2}\left[\left(\Omega_0 \mathbf{a}\right) \otimes vec\left(\Omega_0\right)\right].
\end{aligned}$$

Therefore

$$2\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)\left\{\mathbf{c}'[vec\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]\right\}=2\mu_1\mathbf{c}'\left(I_{d_2^3}+K_{d_2^2,d_2}\right)\left[\left(\Omega_0\mathbf{a}\right)\otimes vec\left(\Omega_0\right)\right]. \quad \text{(A.31)}$$

In the rest of the derivations, we employ the BN decomposition as before. For notational convenience, we simply identify v_i with Ψe_i. For the second term in (A.30), we note that

$$\begin{aligned}
\check{\Omega}_T &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}k\left(\frac{t-s}{M}\right)\check{v}_t\left(\check{v}_s\right)' \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}k\left(\frac{t-s}{M}\right)\left[v_t-w_t(G_0'V_0G_0)^{-1}G_0'V_0\frac{1}{T}\sum_{p=1}^{T}v_p\right] \\
&\quad\times\left[v_s-w_s(G_0'V_0G_0)^{-1}G_0'V_0\frac{1}{T}\sum_{q=1}^{T}v_q\right]' \\
&= \tilde{\Omega}_T+\check{B}_1+\check{B}_2+\check{B}_3,
\end{aligned}$$

where

$$\begin{aligned}
\check{B}_1 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}v_tv_s'V_0G_0(G_0'V_0G_0)^{-1}\left\{\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)w_j'\right\}, \\
\check{B}_2 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{j-s}{M}\right)w_j\right]\left[(G_0'V_0G_0)^{-1}G_0'V_0\right]v_tv_s', \\
\check{B}_3 &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T}k\left(\frac{i-j}{M}\right)w_i'(G_0'V_0G_0)^{-1}G_0'V_0v_tv_s'V_0G_0(G_0'V_0G_0)^{-1}w_j.
\end{aligned}$$

It is easy to show that

$$\begin{aligned}
\check{B}_1 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}v_tv_s'V_0G_0(G_0'V_0G_0)^{-1}\left\{\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)Ew_j'\right\} \\
&\quad -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}v_tv_s'V_0G_0(G_0'V_0G_0)^{-1}\left\{\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)\left(w_j'-Ew_j'\right)\right\} \\
&= -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}v_tv_s'V_0G_0(G_0'V_0G_0)^{-1}\left\{\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)Ew_j'\right\}+O_{q,m}\left(\frac{1}{\sqrt{T}}\right),
\end{aligned}$$

where for a random matrix U, $U=O_{q,m}\left(1/\sqrt{T}\right)$ stands for $\left(E\left\|U\right\|^2\right)^{1/2}=1/\sqrt{T}$. Let

$$F_0=V_0G_0(G_0'V_0G_0)^{-1}G_0'.$$

Then

$$\begin{aligned}
\check{B}_1 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} v_t v_s' F_0 \left\{ \frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right) \right\} + O_{q,m}\left(1/\sqrt{T}\right), \\
\check{B}_2 &= -\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} \left[\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{j-s}{M}\right) \right] F_0' v_t v_s' + O_{q,m}\left(1/\sqrt{T}\right), \\
\check{B}_3 &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} \frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T} k\left(\frac{i-j}{M}\right) F_0' v_t v_s' F_0 + O_{q,m}\left(1/\sqrt{T}\right).
\end{aligned}$$

Using the above results, we can write

$$2\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)\left\{\mathbf{c}'[vec\left(\check{\Omega}_T-\tilde{\Omega}_T\right)\otimes S_T]\right\} := \mathbb{B}_1+\mathbb{B}_2+\mathbb{B}_3,$$

where

$$\mathbb{B}_1 = -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s,m,n} E\mathbf{c}'\left[\left(vec\left(v_t v_s' F_0\right)\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right)\right)\otimes\left(v_m v_n'\mathbf{a}\right)\right],$$

$$\mathbb{B}_2 = -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s,m,n} E\mathbf{c}'\left[\left(vec\left(F_0' v_t v_s'\right)\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right)\right)\otimes\left(v_m v_n'\mathbf{a}\right)\right],$$

and

$$\mathbb{B}_3 = 2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s,m,n} E\mathbf{c}'\left[\left(vec\left(F_0' v_t v_s' F_0\right)\frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right)\right)\otimes\left(v_m v_n'\mathbf{a}\right)\right].$$

We proceed to compute B_1, B_2 and B_3, starting with

$$\begin{aligned}
\mathbb{B}_1 &= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s,m,n}\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right) E\mathbf{c}'\left[\left(vec\left(v_t v_s' F_0\right)\right)\otimes\left(v_m v_n'\mathbf{a}\right)\right] \\
&= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s,m,n}\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right) E\mathbf{c}'\left[\left(F_0' v_s\otimes v_t\right)\otimes\left(v_m v_n'\mathbf{a}\right)\right] \\
&= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right) E\mathbf{c}'\left[\left(F_0' v_t\otimes v_t\right)\otimes\left(v_n v_n'\mathbf{a}\right)\right] \\
&\quad -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right) E\mathbf{c}'\left[\left(F_0' v_s\otimes v_t\right)\otimes\left(v_t v_s'\mathbf{a}\right)\right] \\
&\quad -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T} k\left(\frac{t-j}{M}\right) E\mathbf{c}'\left[\left(F_0' v_s\otimes v_t\right)\otimes\left(v_s v_t'\mathbf{a}\right)\right] \\
:&= \mathbb{B}_{11}+\mathbb{B}_{12}+\mathbb{B}_{13},
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{B}_{11} &= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[(F_0'v_t\otimes v_t)\otimes(v_nv_n'\mathbf{a})\right] \\
&= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[vec\left(v_tv_t'F_0\right)\otimes(v_nv_n'\mathbf{a})\right] \\
&= -2\mu_1\mathbf{c}'\left[vec\left(\Omega_0F_0\right)\otimes(\Omega_0\mathbf{a})\right],
\end{aligned}$$

$$\begin{aligned}
\mathbb{B}_{12} &= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left(F_0'v_sv_s'\mathbf{a}\otimes v_t\right)\otimes v_t] \\
&= -2\lim_{T\to\infty}\left[\frac{1}{MT}\sum_{t,j}k\left(\frac{t-j}{M}\right)\right]\mathbf{c}'\left[(F_0'\Omega_0\mathbf{a})\otimes vec(\Omega_0)\right] \\
&= -2\mu_1\mathbf{c}'\left[(F_0'\Omega_0\mathbf{a})\otimes vec(\Omega_0)\right],
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{B}_{13} &= -2\lim_{T\to\infty}\left[\frac{1}{MT}\sum_{t,j}k\left(\frac{t-j}{M}\right)\right]\mathbf{c}'K_{d_2^2,d_2}\left[(\Omega_0\mathbf{a})\otimes vec\left(F_0\Omega_0\right)\right] \\
&= -2\mu_1\mathbf{c}'K_{d_2^2,d_2}\left[(\Omega_0\mathbf{a})\otimes vec\left(F_0'\Omega_0\right)\right].
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{B}_1 &= -2\mu_1\mathbf{c}'\left[vec\left(\Omega_0F_0\right)\otimes\Omega_0\mathbf{a}\right]-2\mu_1\mathbf{c}'\left[F_0'\Omega_0\mathbf{a}\otimes vec(\Omega_0)\right] \\
&\quad -2\mu_1\mathbf{c}'K_{d_2^2,d_2}\left[(\Omega_0\mathbf{a})\otimes vec\left(F_0'\Omega_0\right)\right].
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{B}_2 &= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[(v_t\otimes F_0'v_t)\otimes(v_nv_n'\mathbf{a})\right] \\
&\quad -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[(v_s\otimes F_0'v_t)\otimes(v_tv_s'\mathbf{a})\right] \\
&\quad -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[(v_s\otimes F_0'v_t)\otimes(v_sv_t'\mathbf{a})\right] \\
&= -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[vec\left(F_0'v_tv_t'\right)\otimes(v_nv_n'\mathbf{a})\right] \\
&\quad -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[(v_sv_s'\mathbf{a}\otimes F_0'v_t)\otimes v_t\right] \\
&\quad -2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[(v_s\otimes F_0'v_tv_t'\mathbf{a})\otimes v_s\right] \\
&= -2\mu_1\mathbf{c}'\left[vec\left(F_0'\Omega_0\right)\otimes(\Omega_0\mathbf{a})\right]-2\mu_1\mathbf{c}'\left[(\Omega_0\mathbf{a})\otimes vec(\Omega_0F_0)\right] \\
&\quad -2\mu_1\mathbf{c}'K_{d_2^2,d_2}\left[(F_0'\Omega_0\mathbf{a})\otimes vec\left(\Omega_0\right)\right],
\end{aligned}$$

$$
\begin{aligned}
\mathbb{B}_3 &= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s,m,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\left\{\mathbf{c}'\left[\left(F_0'v_s\otimes F_0'v_t\right)\otimes\left(v_m v_n'\mathbf{a}\right)\right]\right\}\\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[\left(F_0'v_t\otimes F_0'v_t\right)\otimes\left(v_n v_n'\mathbf{a}\right)\right]\\
&\quad +2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[\left(F_0'v_s\otimes F_0'v_t\right)\otimes\left(v_t v_s'\mathbf{a}\right)\right]\\
&\quad +2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[\left(F_0'v_s\otimes F_0'v_t\right)\otimes\left(v_s v_t'\mathbf{a}\right)\right]\\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,n}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[vec\left(F_0'v_t v_t'F_0\right)\otimes\left(v_n v_n'\mathbf{a}\right)\right]\\
&\quad +2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left[\left(F_0'v_s v_s'\mathbf{a}\otimes F_0'v_t\right)\otimes v_t\right]\\
&\quad +2\lim_{T\to\infty}\frac{1}{MT}\sum_{t,s}\frac{1}{T}\sum_{j=1}^{T}k\left(\frac{t-j}{M}\right)E\mathbf{c}'\left(F_0'v_s\otimes F_0'v_t v_t'\mathbf{a}\right)\otimes v_s]\\
&= 2\mu_1\mathbf{c}'\left[vec\left(F_0'\Omega_0F_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right]+2\mu_1\mathbf{c}'\left[F_0'\Omega_0\mathbf{a}\otimes vec(\Omega_0F_0)\right]\\
&\quad +2\mu_1\mathbf{c}'K_{d_2^2,d_2}\left[\left(F_0'\Omega_0\mathbf{a}\right)\otimes vec\left(F_0'\Omega_0\right)\right].
\end{aligned}
$$

Combining the above results yields

$$
\begin{aligned}
&\rho_{2,\infty}\\
=\ & -2\mu_1\mathbf{c}'\left[vec\left(\Omega_0F_0+F_0'\Omega_0-F_0'\Omega_0F_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right]\\
&+2\mu_1\mathbf{c}'\left[\left(\Omega_0\mathbf{a}\right)\otimes vec\left(\Omega_0\right)+\left(F_0'\Omega_0\mathbf{a}\right)\otimes vec(\Omega_0F_0)\right]\\
&-2\mu_1\mathbf{c}'\left[\left(\Omega_0\mathbf{a}\right)\otimes vec(\Omega_0F_0)+\left(F_0'\Omega_0\mathbf{a}\right)\otimes vec(\Omega_0)\right]\\
&+2\mu_1\mathbf{c}'K_{d_2^2,d_2}\left[\left(\Omega_0\mathbf{a}\right)\otimes vec\left(\Omega_0\right)+\left(F_0'\Omega_0\mathbf{a}\right)\otimes vec\left(F_0'\Omega_0\right)\right]\\
&-2\mu_1\mathbf{c}'K_{d_2^2,d_2}\left[\left(\Omega_0\mathbf{a}\right)\otimes vec\left(F_0'\Omega_0\right)+\left(F_0'\Omega_0\mathbf{a}\right)\otimes vec\left(\Omega_0\right)\right].
\end{aligned}
$$

After some algebraic manipulations, we obtain

$$
\begin{aligned}
\rho_{2,\infty} &= -2\mu_1\mathbf{c}'\left[vec\left(\Omega_0F_0+F_0'\Omega_0-F_0'\Omega_0F_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right]\\
&\quad +2\mu_1\mathbf{c}'\left\{I_{d_2^3}+K_{d_2^2,d_2}\left(I_{d_2}\otimes K_{d_2,d_2}\right)\right\}\left\{\left[\left(I-F_0\right)'\left(\Omega_0\mathbf{a}\right)\right]\otimes vec\left[\Omega_0\left(I-F_0\right)\right]\right\}.
\end{aligned}
$$

Part (c) We first show that

$$
\begin{aligned}
\kappa_{2,\infty} &= 2\mu_2\left\{\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\right]\right\}\\
&\quad +2\mu_2\left\{\mathbf{d}'\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\left(\Omega_0\mathbf{a}\right)\right)\right\}\\
&\quad +\mu_2\mathbf{c}'\left(\Omega_0\otimes\Omega_0\otimes\Omega_0\right)\mathbf{c}+\mu_2\mathbf{c}'\left(K_{d_2,d_2}\otimes I_{d_2}\right)\left(\Omega_0\otimes\Omega_0\otimes\Omega_0\right)\mathbf{c}. \qquad \text{(A.32)}
\end{aligned}
$$

The first term in $\kappa_{2,\infty}$ is

$$\begin{aligned}
& 2\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)\left\{\mathbf{d}'\left[vec\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes vec\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T\right]\right\} \\
= \ & 2\lim_{T\to\infty}\frac{1}{MT^2}\sum_{s\neq t}\sum_{i\neq j}\sum_{\ell,m}k_{s-t}k_{i-j}E\left\{\mathbf{d}'\left[vec(v_s v_t')\otimes vec(v_i v_j')\otimes v_\ell\right]v_m'\mathbf{a}\right\} \\
= \ & 2\lim_{T\to\infty}\frac{1}{MT^2}\sum_{s\neq t}\sum_{i\neq j}\sum_{\ell,m}k_{s-t}k_{i-j}E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_j\otimes v_i\otimes v_\ell\right]v_m'\mathbf{a}\right\} \\
= \ & 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_t\otimes v_s\otimes v_m\right]v_m'\mathbf{a}\right\} \\
& +2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_s\otimes v_t\otimes v_m\right]v_m'\mathbf{a}\right\} \\
= \ & \mathbb{C}_1+\mathbb{C}_2,
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{C}_1 &= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes(v_s\otimes v_t)\otimes v_s\otimes v_m\right]v_m'\mathbf{a}\right\} \\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes K_{d_2,d_2}\left(v_t\otimes v_s\right)\otimes v_s\otimes v_m v_m'\mathbf{a}\right]\right\} \\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[I_{d_2}v_t\otimes K_{d_2,d_2}\left(v_t\otimes v_s\right)\otimes v_s\otimes v_m v_m'\mathbf{a}\right]\right\} \\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\right)\left(v_t\otimes v_t\otimes v_s\right)\otimes v_s\otimes v_m v_m'\mathbf{a}\right]\right\} \\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(v_t\otimes v_t\otimes v_s\otimes v_s\otimes v_m v_m'\mathbf{a}\right)\right]\right\} \\
&= \left(2\int_{-1}^{1}k^2(x)dx\right)\left\{\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\Omega_0\mathbf{a}\right)\right]\right\},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{C}_2 &= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_s\otimes v_t\otimes v_m\right]v_m'\mathbf{a}\right\} \\
&= 2\lim_{T\to\infty}\frac{1}{MT}\sum_{s\neq t}k_{s-t}^2E\left\{\mathbf{d}'\left[K_{d_2,d_2^2}\left(vec\left(\Omega_0\right)\otimes v_t\right)\otimes v_t\otimes\Omega_0\mathbf{a}\right]\right\} \\
&= \left(2\int_{-1}^{1}k^2(x)dx\right)\left\{\mathbf{d}'\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\Omega_0\mathbf{a}\right)\right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& 2\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)\left\{\mathbf{d}'\left[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T\right]\right\} \\
= \ & \left(2\int_{-1}^{1}k^2(x)dx\right)\left\{\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\Omega_0\mathbf{a}\right)\right]\right\} \\
& +\left(2\int_{-1}^{1}k^2(x)dx\right)\left\{\mathbf{d}'\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\Omega_0\mathbf{a}\right)\right\}. \qquad \text{(A.33)}
\end{aligned}$$

The second term in $\kappa_{2,\infty}$ is

$$\begin{aligned}
& \lim_{T\to\infty} \frac{T}{M} E\left[\mathbf{c}'\left(vec\left(\tilde{\Omega}_T - \bar{\Omega}_T\right)\otimes S_T\right)\right]^2 \\
= & \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} \sum_{\ell,m,n} k_{h-i}k_{\ell-m} E\left\{\mathbf{c}'\left[v_h \otimes v_i \otimes v_j\right]\left[v_\ell' \otimes v_m' \otimes v_n'\right]\mathbf{c}\right\} \\
= & \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} k_{h-i}^2 E\left\{\mathbf{c}'\left[v_h \otimes v_i \otimes v_j\right]\left[v_h' \otimes v_i' \otimes v_j'\right]\mathbf{c}\right\} \\
& + \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} k_{h-i}^2 E\left\{\mathbf{c}'\left[v_h \otimes v_i \otimes v_j\right]\left[v_i' \otimes v_h' \otimes v_j'\right]\mathbf{c}\right\} \\
= & \mathbb{D}_1 + \mathbb{D}_2
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{D}_1 &= \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} k_{h-i}^2 E\left\{\mathbf{c}'\left[v_h \otimes v_i \otimes v_j\right]\left[v_h' \otimes v_i' \otimes v_j'\right]\mathbf{c}\right\} \\
&= \left(\int_{-1}^{1} k^2(x)dx\right)\mathbf{c}'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c}
\end{aligned}$$

$$\begin{aligned}
\mathbb{D}_2 &= \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} k_{h-i}^2 E\left\{\mathbf{c}'\left[v_h \otimes v_i \otimes v_j\right]\left[v_i' \otimes v_h' \otimes v_j'\right]\mathbf{c}\right\} \\
&= \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} k_{h-i}^2 E\left\{\mathbf{c}'\left[K_{d_2,d_2}\left(v_i \otimes v_h\right) \otimes v_j\right]\left[v_i' \otimes v_h' \otimes v_j'\right]\mathbf{c}\right\} \\
&= \lim_{T\to\infty} \frac{1}{MT^2} \sum_{h,i,j} k_{h-i}^2 E\left\{\mathbf{c}'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\left(v_i \otimes v_h\right) \otimes v_j\right)\left[v_i' \otimes v_h' \otimes v_j'\right]\mathbf{c}\right\} \\
&= \left(\int_{-1}^{1} k^2(x)dx\right)\mathbf{c}'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c}
\end{aligned}$$

Therefore

$$\begin{aligned}
& \lim_{T\to\infty} \frac{T}{M} E\left[\mathbf{c}'\left(vech\left(\tilde{\Omega}_T - \bar{\Omega}_T\right)\otimes S_T\right)\right]^2 \\
= & \left(\int_{-1}^{1} k^2(x)dx\right)\mathbf{c}'\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c} \\
& + \left(\int_{-1}^{1} k^2(x)dx\right)\mathbf{c}'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c}. \qquad \text{(A.34)}
\end{aligned}$$

Combining (A.31), (A.33), and (A.34) yields equation (A.32). This equation, combined with Lemma A.12, leads to the stated result.

Part (d) In view of Lemma A.12, it suffices to show that

$$\begin{aligned}
\kappa_{4,\infty} - 6\kappa_{2,\infty} &= 12\mu_2\mathbf{c}'\left[\Omega_0 \otimes \Omega_0 \otimes \left(\Omega_0\mathbf{aa}'\Omega_0\right)\right]\mathbf{c} \\
&\quad + 12\mu_2\mathbf{c}'\left(K_{d_2,d_2} \otimes I_{d_2}\right)\left[\Omega_0 \otimes \Omega_0 \otimes \left(\Omega_0\mathbf{aa}'\Omega_0\right)\right]\mathbf{c}. \qquad \text{(A.35)}
\end{aligned}$$

First,

$$
\begin{aligned}
& 4\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)^3\left\{\mathbf{d}'\left[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T\right]\right\} \\
= & \lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{h,i,j}\sum_{\ell,m,n}k_{s-t}k_{i-j}E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_j\otimes v_i\otimes v_h\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_n'\mathbf{a}\right\} \\
= & \lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{h}\sum_{\ell,m,n}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_t\otimes v_s\otimes v_h\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_n'\mathbf{a}\right\} \\
& +\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{h}\sum_{\ell,m,n}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_s\otimes v_t\otimes v_h\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_n'\mathbf{a}\right\} \\
= & \mathbb{E}_1+\mathbb{E}_2,
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{E}_1 = & 4\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{\ell,m}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_t\otimes v_s\otimes v_\ell\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_m'\mathbf{a}\right\} \\
& +4\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{\ell,m}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_t\otimes v_s\otimes v_m\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_\ell'\mathbf{a}\right\} \\
& +4\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{m,n}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_t\otimes v_s\otimes v_n\right]v_m'\mathbf{a}v_m'\mathbf{a}v_n'\mathbf{a}\right\} \\
= & 12\mu_2\mathbf{d}'\left[\left(I_{d_2}\otimes K_{d_2,d_2}\otimes I_{d_2^2}\right)vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\Omega_0\mathbf{a}\right]\left(\mathbf{a}'\Omega_0\mathbf{a}\right),
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_2 = & 4\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{\ell,m}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_s\otimes v_t\otimes v_\ell\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_m'\mathbf{a}\right\} \\
& +4\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{\ell,m}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_s\otimes v_t\otimes v_m\right]v_\ell'\mathbf{a}v_m'\mathbf{a}v_\ell'\mathbf{a}\right\} \\
& +4\lim_{T\to\infty}\frac{1}{MT^3}\sum_{s,t}\sum_{m,n}k_{s-t}^2E\left\{\mathbf{d}'\left[v_t\otimes v_s\otimes v_s\otimes v_t\otimes v_n\right]v_m'\mathbf{a}v_m'\mathbf{a}v_n'\mathbf{a}\right\} \\
= & 12\mu_2\mathbf{d}'\left[\left(K_{d_2,d_2^2}\otimes I_{d_2^2}\right)\left(vec\left(\Omega_0\right)\otimes vec\left(\Omega_0\right)\otimes\Omega_0\mathbf{a}\right)\right]\left(\mathbf{a}'\Omega_0\mathbf{a}\right).
\end{aligned}
$$

Second,

$$
\begin{aligned}
& 6\lim_{T\to\infty}\frac{T}{M}E\left(\mathbf{a}'S_T\right)^2\left\{\mathbf{c}'[vech\left(\tilde{\Omega}_T-\bar{\Omega}_T\right)\otimes S_T]\right\}^2 \\
= & 6\lim_{T\to\infty}\frac{1}{MT^3}\sum_{o,p}\sum_{h,i,j,\ell,m,n}k_{h-i}k_{\ell-m}E\left\{\mathbf{c}'\left[v_h\otimes v_i\otimes v_j\right]\left[v_\ell'\otimes v_m'\otimes v_n'\right]\mathbf{c}\right\}\left(\mathbf{a}'v_o\right)\left(\mathbf{a}'v_p\right) \\
= & \mathbb{F}_1+\mathbb{F}_2,
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{F}_1 = & 6\lim_{T\to\infty}\frac{1}{MT^3}\sum_{o,p}\sum_{h,i,j,n}k_{h-i}^2E\left\{\mathbf{c}'\left[v_h\otimes v_i\otimes v_j\right]\left[v_h'\otimes v_i'\otimes v_n'\right]\mathbf{c}\right\}\left(\mathbf{a}'v_o\right)\left(\mathbf{a}'v_p\right) \\
= & 6\lim_{T\to\infty}\frac{1}{MT^3}\sum_{o,p}\sum_{h,i,j,n}k_{h-i}^2E\left\{\mathbf{c}'\left(\Omega_0\otimes\Omega_0\otimes\Omega_0\right)\mathbf{c}\right\}\left(\mathbf{a}'\Omega_0\mathbf{a}\right) \\
& +12\lim_{T\to\infty}\frac{1}{MT^3}\sum_{h,i,j,n}k_{h-i}^2E\left\{\mathbf{c}'\left[v_h\otimes v_i\otimes v_jv_j'\mathbf{a}\right]\left[v_h'\otimes v_i'\otimes\mathbf{a}'v_nv_n'\right]\mathbf{c}\right\} \\
= & 6\mu_2\mathbf{c}'\left(\Omega_0\otimes\Omega_0\otimes\Omega_0\right)\mathbf{c}\left(\mathbf{a}'\Omega_0\mathbf{a}\right)+12\mu_2\mathbf{c}'\left[\Omega_0\otimes\Omega_0\otimes\left(\Omega_0\mathbf{a}\mathbf{a}'\Omega_0\right)\right]\mathbf{c},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{F}_2 &= 6 \lim_{T\to\infty} \frac{1}{MT^3} \sum_{o,p} \sum_{h,i,j,n} k_{h-i}^2 E\left\{\mathbf{c}' \left[v_h \otimes v_i \otimes v_j\right]\left[v_i' \otimes v_h' \otimes v_n'\right]\mathbf{c}\right\}\left(\mathbf{a}'v_o\right)\left(\mathbf{a}'v_p\right) \\
&= 6 \lim_{T\to\infty} \frac{1}{MT^3} \sum_{o,p} \sum_{h,i,j,n} k_{h-i}^2 E\left\{\mathbf{c}' \left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\left(v_i \otimes v_h\right) \otimes v_j\right)\left[v_i' \otimes v_h' \otimes v_n'\right]\mathbf{c}\right\}\left(\mathbf{a}'v_o\right)\left(\mathbf{a}'v_p\right) \\
&= 6\mu_2 \mathbf{c}' \left(K_{d_2,d_2} \otimes I_{d_2}\right)\left(\Omega_0 \otimes \Omega_0 \otimes \Omega_0\right)\mathbf{c}\left(\mathbf{a}'\Omega_0\mathbf{a}\right) \\
&\quad + 12\mu_2 \mathbf{c}' \left(K_{d_2,d_2} \otimes I_{d_2}\right)\left[\Omega_0 \otimes \Omega_0 \otimes \left(\Omega_0 \mathbf{a}\mathbf{a}'\Omega_0\right)\right]\mathbf{c}.
\end{aligned}
$$

Combining the above results completes the proof. ■

References

[1] Abadir, K.M. and Magnus, J.R. (2005). *Matrix Algebra,* Econometric Exercise 1, Cambridge University Press, New York: NY.

[2] Andrews, D.W.K. (1991). "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica,* 59, 817–858.

[3] Andrews, D.W.K. (2002). "Equivalence of the Higher-order Asymptotic Efficiency of k-step and Extremum Statistics." *Econometric Theory*, 18, 1040–1085.

[4] Andrews, D.W.K. and J. C. Monahan (1992). "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator." *Econometrica,* 60, 953–966.

[5] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York: NY.

[6] Casella, G., J.T.G. Hwang and C. Robert (1993). "A Paradox in Decision-Theoretic Interval Estimation." *Statistica Sinica*, 3, 141–155.

[7] Doukhan, P. (1995). *Mixing: Properties and Examples.* Springer-Verlag, New York:NY

[8] de Jong, R.M. and J. Davidson (2000). "Consistency of Kernel Estimators of Heteroskedastic and Autocorrelated Covariance Matrices." *Econometrica,* 68, 407–424.

[9] Götze, F. and C. Hipp (1983). "Asymptotic Expansions for Sums of Weakly Dependent Random Vectors." *Probability Theory and Related Fields*, 64, 211–239.

[10] Götze, F. and H.R. Künsch (1996). "Second-order Correctness of the Blockwise Bootstrap for Stationary Observations." *Annals of Statistics*, 24, 1914–1933.

[11] Grenander, U. and M. Rosenblatt (1957). *Statistical Analysis of Stationary Time Series.* Wiley, New York: NY.

[12] Hall, P. and J.L. Horowitz (1996). "Bootstrap Critical Values for Tests Based on Generalized-method-of-moments Estimators." *Econometrica*, 64, 891–916.

[13] Hall, P. and S.J. Sheather (1988). "On the Distribution of a Studentized Quantile." *Journal of the Royal Statistical Society*, Series B, 50, 381–391.

[14] Hannan, E.J. (1970). *Multiple Time Series.* Wiley, New York: NY.

[15] Inoue, A. and M. Shintani (2005). "Technical Appendix for 'Bootstrapping GMM Estimators for Time Series." Working paper, Department of Economics, Vanderbilt University.

[16] Inoue, A. and M. Shintani (2006). "Bootstrapping GMM Estimators for Time Series." *Journal of Econometrics*, 133(2), 531–555.

[17] Kolassa J.E. (1997). *Series Approximation Methods in Statistics.* Springer, New York: NY.

[18] Lahiri, S.N. (1993). "Refinements in the Asymptotic Expansions for Sums of Weakly Dependent Random Vectors." *Annals of Probability,* 21, 791–799.

[19] Lahiri, S.N. (1996). "On Edgeworth Expansion and Moving Block Bootstrap for Studentized M-estimators in Multiple Linear Regression Models." *Journal of Multivariate Analysis*, 56, 42–59.

[20] Kiefer, N.M., T.J. Vogelsang and H. Bunzel (2000): "Simple Robust Testing of Regression Hypotheses," *Econometrica,* 68, 695–714.

[21] Kiefer, N.M. and T.J. Vogelsang (2002a): "Heteroskedasticity-autocorrelation Robust Testing Using Bandwidth Equal to Sample Size," *Econometric Theory,* 18, 1350–1366.

[22] ——— (2002b): "Heteroskedasticity-autocorrelation Robust Standard Errors Using the Bartlett Kernel without Truncation," *Econometrica,* 70, 2093–2095.

[23] ——— (2005): "A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests," *Econometric Theory* 21, 1130–1164.

[24] Magnus, J.R. and H. Neudecker (1988). *Matrix Differential Calculus.* Revised Edition. Wiley, New York: NY.

[25] Newey, W.K. and K.D. West (1987). "A Simple, Positive semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55, 703–708.

[26] ——— (1994). "Automatic Lag Selection in Covariance Estimation." *Review of Economic Studies*, 61, 631–654.

[27] Neyman, J. (1937). "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability." *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences,* 236 (767), 333–380.

[28] Parzen, E. (1957). "On the Consistent Estimates of the Spectrum of a Stationary Time Series." *Annals of Mathematical Statistics*, 28, 329–348.

[29] Phillips, P.C.B. and V. Solo (1992). "Asymptotics for Linear Processes." *Annals of Statistics*, 20, 971–1001.

[30] Politis, D.N. and J.P. Romano (1995). "Bias Corrected Nonparametric Spectral Density Estimator." *Journal of Time Series Analysis,* 16, 67–103.

[31] Politis, D.N. and J.P. Romano (1998). "Multivariate Density Estimation with General Flat-top Kernels of Infinite Order." *Journal of Multivariate Analysis*, 68, 1–25.

[32] Pratt, J.W. (1961). "Length of Confidence Intervals." *Journal of American Statistics Association*, 56, 549-567.

[33] Rothenberg, T.J. (1984). "Approximating the Distributions of Econometric Estimators and Test Statistics." In: Griliches, Z., Intriligator, M.D. (Eds.), Handbook of Econometrics, vol. 2. North-Holland, Amsterdam, pp. 881–935.

[34] Severini, T.A. (2005). *Elements of Distribution Theory.* Cambridge University Press, New York: NY.

[35] Sun, Y., P.C.B. Phillips and S. Jin (2007). "Optimal Bandwidth Selection in Heteroskedasticity-Autocorrelation Robust Testing." Forthcoming, *Econometrica.*

[36] Taniguchi, M and M.L. Puri (1996). "Valid Edgeworth Expansions of M-estimators in Regression Models with Weakly Dependent Residuals." *Econometric Theory,* 12, 331–346.

[37] Velasco, C. and P.M. Robinson (2001). "Edgeworth Expansions for Spectral Density Estimates and Studentized Sample Mean." *Econometric Theory*, 17, 497–539.

[38] Yokoyama, R. (1980). "Moment Bounds for Stationary Mixing Sequences." *Probability Theory and Related Fields*, 52(1), 45–57.