Preliminary manuscript for external reviewers

Higher-order Expansion and Optimal HAR Inference in GMM with Over-identification*

Jungbin Hwang
Department of Economics,
University of Connecticut

Gonzalo Valdés
Departamento de Ingeniería Industrial y de Sistemas,
Universidad de Tarapacá

January, 2021

Abstract

In this paper we develop a higher-order expansion of test statistic for heteroskedasticity autocorrelation robust (HAR) inference in an over-identified two-step GMM framework. The new higher-order expansion rigorously shows higher-order accuracy of fixed-smoothing asymptotics in finite sample null rejection probability. We also provide an optimal choice of smoothing parameters that directly accounts for the degree of overidentification in HAR inferential problem.

*Email: jungbin.hwang@uconn.edu and gvaldes@uta.cl.

1 Introduction

Generalized method of moments (GMM) is one of the most widely used statistical methods for the analysis of economic data and financial data (Hall, 2005). In time-series data, with an unknown form of dependence, an important initial contribution is made by Newey and West (1987) to implement the feasible two-step efficient estimation of GMM proposed by Hansen (1982). The key step for the efficient GMM with over-identification procedure is an estimation of the optimal weighting matrix, which is referred to as the long-run variance (LRV) of the moment conditions. Newey and West (1987) and Andrews (1994) propose LRV using the nonparametric kernel method.

However, it is well known that the two-step GMM procedure shows poor performance in a finite sample context. See, for example, the special issue in the Journal of Business & Economic Statistics (Christiano, 1996). Making an accurate inference of the GMM estimator in the time series setting becomes more challenging due to the considerable estimation uncertainty embodied in the nonparametric LRV estimator. The standard first-order normal approximation completely ignores the estimation uncertainty in LRV and does not reflect the finite sample situations properly. Phillips (2005) proposes a new type of LRV estimators using orthonormal series basis functions. Müller (2007) and Sun (2011, 2013) develop its more accurate first-order approximation, so called fixed-smoothing asymptotics, and show its superior finite-sample performances.

In the two-step GMM with over-identification, Sun (2014) and Hwang and Sun (2017) develop fixed-smoothing asymptotics and show that improved performances relative to the normal approximation. Hwang and Valdés (2020) provide further finite-sample improvement by correcting the bias of variance two-step GMM estimator. The literature, however, show the finite-sample improvements using the fixed-smoothing asymptotics by Monte-Carlo simulations. To the best of our knowledge, there is no work in time-series GMM that rigorously prove the higher-order accuracy of the fixed-smoothing asymptotics that achieve more accurate finite sample null rejection in a precise technical sense. In this paper, we fill the technical gap in the literature and develop a higher-oder expansion of t-statistic of the two-step GMM. In addition to the higher-order expansion, we provide an optimal choice of smoothing parameters which directly focuses on heteroskedasticity autocorrelation robust (HAR) inferential problem in an over-identified GMM setting.

The higher-order properties of GMM estimator has been an active research area in econometrics. Hansen, Heaton, and Yaron (1996) first suggest continuously updating (CU) GMM to improve finite-sample performance of the two-step GMM. Its generalization to generalized empirical likelihood (GEL) method has been done in Smith (1997), Imbens, Spady, and Johnson (1998), Kitamura and Stutzer (1997), and Newey and Smith (2004). Importantly, Newey and Smith (2004) and Anatolyev (2005) prove improved higher-order bias and variance properties of GEL estimators in i.i.d setting and time-series setting, respectively. Still, the poor finite-sample performances of GEL remain issues, as pointed by extensive numerical work in Guggenberger and Hahn (2005) and Guggenberger (2008). Moreover, the results in Newey and Smith (2004) and Anatolyev (2005) do not lead to key inferential properties of GMM estimators, as they do not consider a higher-order expansion of finite sample null rejection probability of t-statistic.

In the context of bootstrapping GMM, Hall and Horowitz (1996) first prove the higher-order property of GMM inference via an Edgeworth expansion of the t-statistic, but they assume a less general finite dependence structure in time series. Inoue and Shintani (2005) relaxes a finite dependence structure in Hall and Horowitz (1996), but the proposed moving-block bootstrap method requires an additional choice of block length in practice. Sun and Phillips (2009) also pro-

vide an Edgeworth expansion in time-series GMM regression, but their approach is not applicable to the HAR inference with the OS-LRV considered in our paper.

Recently, a new type of inferential methods in time-series, so called HAR inference, develop the alternative first-order fixed-smoothing asymptotics (Phillips,2005; Müller, 2007; Sun et al.,2008). See also Lazarus et al. (2018) for its practical implementation. Sun et al. (2008), Sun (2011, 2013, 2014), and Zhang and Shao (2013) rigorously prove the higher-order accuracy of the alternative fixed-smoothing asymptotics in HAR inference. All of these papers, however, focus on the one-step GMM estimator, or exactly identified GMM model.

In this paper, we develop a higher-order expansion of the t-statistic considering the two-step GMM estimation in HAR inference. Technically, it is more challenging to account for the estimation uncertainty of the optimal weighting matrix in the two-step GMM framework. To derive the higher-order expansion, we first decompose the original t-statistic into an infeasible t-statistic and other error terms that do not involve the non-parametric estimation of the optimal weighting matrix. We then transform the higher-order approximation of the infeasible t-statistic into an inference in Gaussian over-identified location model. In the over-identified location model, we introduce a novel GLS-two-step t-statistic and show its asymptotic equivalence to the infeasible t-statistic. Using these findings, we are able to expand the finite-sample CDF of the original t-statistic around CDF of the chi-square distribution with one degree of freedom and characterize the leading higher-order terms.

Based on the higher-order expansion, we provide an optimal choice of smoothing parameters for HAR inference which directly focuses on the inferential problem in the over-identified GMM problem. Our formula generalizes what is previously suggested in Sun (2013) for the exact identified GMM, as it converges to the formula in Sun (2013) when there is no over-identification.

This paper generalizes a higher-order expansion in HAR inference pioneered in Sun, Phillips, and Jin (2008) and Sun (2011, 2013) to the over-identified two-step GMM framework. Our theoretical results show that the fixed-smoothing asymptotics in two-step GMM provides a higher-order refinement of the conventional GMM t-test using standard normal critical value. Our result also provides a new formula of optimal smoothing parameter in HAR inference in GMM, which directly focuses on inferential problems.

The rest of the paper is organized as follows. Section 2 describes a general two-step GMM in a time-series setting and derives its asymptotic property. Section 3 establishes a higher-order expansion for the test statistics in Gaussian overidentified location model. Section 4 consider selecting K to minimize the coverage probability error on the basis of this expansion in previous sections. Section 5 concludes. Proofs are presented in the Appendix.

2 Two-step GMM in Over-identified System

This paper considers the following over-identified GMM model in time-series and HAR inference, e.g., Sun (2014), Hwang and Sun (2017), and Hwang and Valdés (2020). We want to estimate a $d \times 1$ vector of parameter $\theta \in \Theta$ using a vector of observation $w_t \in \mathbb{R}^{d_w}$ at time t. The true parameter θ_0 is assumed to be an interior point of Θ. The moment condition is given as

$$Ef(w_t, \theta) = 0 \text{ iff } \theta = \theta_0,$$

where $f(w_t, \theta)$ is an $m \times 1$ vector of moment process which is linear in parameter vector θ. Leading example is a linear IV regression model in time series with $f(w_t, \theta) = z_t(y_t - x_t'\theta)$, where $w_t = (y_t, x_t', z_t')' \in \mathbb{R}^{d_w}$ consists of the observed data, e.g., Inoue and Shintani (2006) and Sun

and Phillips (2009). With the true parameter $\theta = \theta_0$, the process $f(w_t, \theta_0)$ is stationary with zero mean. We allow $f(w_t, \theta_0)$ to have general autocorrelation of unknown forms and satisfy $\sum_{j=-\infty}^{\infty} \|Ef(w_t, \theta_0) f(w_{t-j}, \theta_0)'\| < \infty$ and some mixing conditions for the time-series CLT as follows

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} f(w_t, \theta_0) \xrightarrow{d} N(0, \Sigma), \tag{1}$$

where $N(0, \Sigma)$ is a m-dimensional normal random vector. The $m \times m$ matrix $\Sigma = \Lambda\Lambda'$ is a positive definite long-run variance (LRV) of the moment process $f(w_t, \theta_0)$, i.e. $\Sigma = \sum_{j=-\infty}^{\infty} Ef(w_t, \theta_0) f(w_{t-j}, \theta_0)$. Also, we assume that $q = m - d > 0$, and the rank of $G = E[\partial f(w_t, \theta_0)/\partial \theta']$ is d. Thus the model is overidentified with a degree of over-identificaion q. Let $f_T(\theta) = \frac{1}{T} \sum_{s=1}^{T} f(w_s, \theta)$ and define a one-step GMM estimator as

$$\hat{\theta}_1 = \arg\min_{\theta \in \Theta} f_T(\theta)' W_T^{-1} f_T(\theta),$$

where W_T is an initial weight matrix. Using this one-step estimator, the feasible efficient two-step GMM estimator in Hansen(1982) is defined as

$$\hat{\theta}_2 = \arg\min_{\theta \in \Theta} f_T(\theta)' \left[\Sigma_T(\hat{\theta}_1)\right]^{-1} f_T(\theta),$$

where

$$\Sigma_T(\hat{\theta}_1) = \frac{1}{K} \sum_{j=1}^{K} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \phi_j(\frac{t}{T}) f(w_t, \theta) \right) \left(\frac{1}{\sqrt{T}} \sum_{s=1}^{T} \phi_j(\frac{s}{T}) f(w_s, \theta) \right)'$$

and $\{\phi_j(r)\}$ are orthonormal basis functions on $L^2[0,1]$ satisfying $\int_0^1 \phi_j(r)\, dr = 0$. We parametrize h in such a way that h indicates the level of smoothing for both types of LRV estimators.

By construction, $\Sigma_T(\hat{\theta}_1)$ is a quadratic heteroskedasticity autocorrelation robust (HAR) estimator for the LRV Ω. It is important to note that $\Sigma_T(\hat{\theta}_1)$ is a "centered" version of the LRV estimator, as it is based on the estimation of the demeaned moment process $f(w_t, \hat{\theta}_1) - f_T(\hat{\theta}_1)$. Under the conventional asymptotic theory, Hall (2000) shows that the demeaning procedure can potentially improve the power performance of the J-test using the HAR estimator. Also, this demeaned procedure plays an important role in fixed smoothing asymptotics as the random matrix limit of $\Sigma_T(\theta_0)$ is independent of the limiting distribution of $\sqrt{T} f_T(\theta_0)$ which is a normal distribution.

With $d = 1$, this paper considers the following GMM t statistic for testing $H_0 : \theta = \theta_0$:

$$\mathbb{T}_T := \mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1) = \frac{\sqrt{T}(\hat{\theta}_2 - \theta_0)}{\sqrt{[G_T' \Sigma_T^{-1}(\hat{\theta}_1) G_T]^{-1}}}.$$

The main contribution of this paper is to develop a higher-order expansion of the finite-sample probability of the GMM t-statistics

$$P(\mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1) \leq z),$$

as $T \to \infty$ such that $K/T \to 0$. The main object is the two-step GMM t-statistic $\mathbb{T}_T(\hat{\theta}_2; \hat{\theta}_1)$ that depends on both first-step and two-step estimators with non-linear interactions, especially through $\Sigma_T^{-1}(\hat{\theta}_1)$. It is technically challenging to obtain the valid Edgeworth expansion

for $\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)$. Sun and Phillips (2009) proves the same object with linear-GMM regression setting with kernel-LRV estimator for $\Sigma_T(\hat{\theta}_1)$. Sun (2013, 2014) develops the higher-order expansion of general GMM model, but is only applicable to the first-step GMM estimator $\hat{\theta}_1$.To confront these technical challenges and develop a high order expansion in the over-identified GMM setting, we first establish the following stochastic approximation in Lemma. We first introduce high level assumptions that are necessary for Lemma.

Assumption 1 *$\hat{\theta}_1 = \theta_0 + O_p(T^{-1/2})$ and $G_T = G + O_p(T^{-1/2})$, where the terms of order $O_p(T^{-1/2})$ does not depend on K.*

Assumption 2 *As $K \to \infty$ and $T \to \infty$ such that $K/T \to 0$, we have that*

$$\Sigma_T(\theta_0) = \Sigma + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right); \tag{2}$$

$$\Upsilon_T(\theta_0) + \Upsilon_T'(\theta_0) = \Upsilon + \Upsilon' + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) \tag{3}$$

where Σ is positive definite, and

$$\Upsilon_T(\theta) = \frac{1}{K}\sum_{j=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j\left(\frac{t}{T}\right)\frac{\partial f(w_t,\theta)}{\partial\theta}\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\phi_j\left(\frac{s}{T}\right)f(w_s,\theta)\right)'$$

.

Assumption 1 is made for convenience, and primitive sufficient conditions are available from the time-series standard GMM asymptotic theory, e.g. Sun (2013, 2014). One of the key requirements for the first condition is that the initial GMM weight matrix W_T do not depend on the unknown parameter value θ_0 and it converges to W with parametric rate of convergence of order $O_p(T^{-1/2})$. Assumption 2 is made for convenience, and sufficient conditions for (2) and (3) can be referred to HAR literature, as in Phillips (2005) and Sun (2013). They study the asymptotic results for the OS-LRV estimator assuming large $K \to \infty$ and $T \to \infty$ such that $K/T \to 0$.

Lemma 1 *Under Assumptions (1)–(2), we have that*

$$\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1) = \mathcal{T}_T(\theta_0) + \psi_T + \psi_T^*,$$

where

$$\begin{aligned}
\mathcal{T}_T(\theta_0) &= -\frac{\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1}G'\Sigma_T^{-1}(\theta_0)\sqrt{T}f_T(\theta_0)}{\sqrt{[G'\Sigma_T^{-1}(\theta_0)G]^{-1}}};\\
\psi_T &= O_p\left(\frac{1}{\sqrt{T}}\right);\\
\psi_T^* &= O_p\left(\frac{1}{\sqrt{T}}\left(\frac{1}{\sqrt{K}} + \frac{K^2}{T^2} + \frac{1}{\sqrt{T}}\right)\right),
\end{aligned}$$

and ψ_T does not depend on K.

Assumption 3 *For sufficiently large $M > 0$, we have that $P(|\psi_T| > \log T/\sqrt{T}) = O(1/\sqrt{T})$ and $P\left(|\psi_T^*| > \delta_T/\log T\right) = o(\delta_T)$ for $\delta_T = 1/K + K^2/T^2$.*

Lemma 2 *Under Assumptions 1–3, and if $K \to \infty$ such that $K/T \to 0$, then we have that*

$$P\left(\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1) \leq z\right) = P\left(\mathcal{T}_T(\theta_0) \leq z\right) + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right).$$

Let $G = \underset{m\times m}{\mathbb{U}} \cdot \underset{m\times d}{\Psi} \cdot \underset{d\times d}{\mathbb{V}'}$ be a singular value decomposition (SVD) of G, where $\Psi' = \left(\underset{d\times d}{A}, \ \underset{d\times q}{O} \right)$, and A is a $d \times d$ diagonal matrix and O is a matrix of zeros. Also, we define the rotated moment conditions and the rotated long run variance

$$\begin{aligned} f^*(w_t,\theta_0) &: = \mathbb{U}' f(w_t,\theta_0) = (f_1^*(w_t,\theta_0), f_2^{*\prime}(w_t,\theta_0))' \\ \Omega &= \mathbb{U}'\Sigma\mathbb{U} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}, \end{aligned}$$

respectively, where $f_1^*(w_t,\theta_0) \in \mathbb{R}^d$ and $f_2^*(w_t,\theta_0) \in \mathbb{R}^q$. The corresponding long run variance estimator with the transformed moment condition is then naturally defined as

$$\begin{aligned} \hat{\Omega} &= \mathbb{U}'\Sigma_T(\theta_0)\mathbb{U} \\ &= \frac{1}{K}\sum_{j=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j(\frac{t}{T})f^*(w_t,\theta_0)\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\phi_j(\frac{s}{T})f^*(w_s,\theta)\right)' \\ &= \begin{pmatrix} \hat{\Omega}_{11} & \hat{\Omega}_{12} \\ \hat{\Omega}_{21} & \hat{\Omega}_{22} \end{pmatrix}. \end{aligned}$$

Lemma 3 *Let $G = \underset{m\times m}{\mathbb{U}} \cdot \underset{m\times d}{\Psi} \cdot \underset{d\times d}{\mathbb{V}'}$ be a singular value decomposition (SVD) of G, where $\Psi' = \left(\underset{d\times d}{A}, \ \underset{d\times q}{O} \right)$, and A is a $d \times d$ diagonal matrix and O is a matrix of zeros. Then, we have that*
(a) $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1}G'\Sigma_T^{-1}(\theta_0)f(w_t,\theta_0) = \frac{1}{\sqrt{T}}\mathbb{V}A^{-1}\sum_{t=1}^{T}\left[f_1^*(w_t,\theta_0) - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}f_2^*(w_t,\theta_0)\right]$;
(b) $G'\Sigma_T^{-1}(\theta_0)G = A^{-1}\mathbb{V}\hat{\Omega}_{11\cdot 2}A^{-1}\mathbb{V}$.

Let us define

$$s_{1t} = f_1^*(w_t,\theta_0) \text{ and } s_{2t} = f_2^*(w_t,\theta_0),$$

and $\bar{s}_{1T} = T^{-1}\sum_{t=1}^{T}s_{1t}$ and $\bar{s}_{2T} = T^{-1}\sum_{t=1}^{T}s_{2t}$. Then, the result in Lemma 3-(a) indicates that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1}G'\Sigma_T^{-1}(\theta_0)f(w_t,\theta_0) = \frac{1}{\sqrt{T}}\mathbb{V}A^{-1}\sum_{t=1}^{T}\left(s_{1t} - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}s_{2t}\right).$$

Using Lemma 3-(b), the infeasible t-statistic can be represented as

$$\mathcal{T}_T(\theta_0) = \frac{\sqrt{T}\left(\bar{s}_{1T} - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}\bar{s}_{2T}\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}}.$$

3 Over-identified Location Model

We consider the following over-identified location problem for testing H_0^*: $\theta_0^* = 0$ as below:

$$s_{1t} = \theta_0^* + v_{1t}; \tag{4}$$
$$s_{2t} = v_{2t} \tag{5}$$

for $t = 1, \ldots, T$, where $s_t = (s_{1t}, s_{2t})' \in \mathbb{R}^{d+q}$ and $v_t = (v_{1t}, v_{2t})' \in \mathbb{R}^{d+q}$. The corresponding LRV of the transformed moment process $\{v_t = (s_{1t} - \theta_0^*, s_{2t}')'\}$ is $\Omega = \sum_{j=-\infty}^{\infty} v_t v_{t-j}$. Using the original parameters, the parameters in the location model can be expresses as

$$\begin{aligned} \theta_0^* &= A\mathbb{V}'\theta_0; \\ \Omega &= \mathbb{U}'\Sigma\mathbb{U}. \end{aligned}$$

Let's denote $S_1 = (s_{11}, \ldots, s_{1T})$, and $S_2 = (s_{21}, \ldots, s_{2T})$. Also, $S = (s_1', s_2', \ldots, s_T') \in \mathbb{R}^{2T}$, with $s_t = (s_{1t}, s_{2t})'$, and $V = (v_1', v_2', \ldots, v_T)'$ with $v_t = (v_{1t}, v_{2t})'$. For simplicity, we assume $d = 1$. Define the one-step and the two-step GMM estimators in the over-identified location problems:

$$\begin{aligned} \hat{\theta}_1^* &= \bar{s}_{1T}; \\ \hat{\theta}_2^* &= T^{-1} \sum \left(s_{1t} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} s_{2t} \right) \\ &= \bar{s}_{1T} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\bar{s}_{2T}, \end{aligned}$$

where $\bar{s}_{1T} = T^{-1}\sum_{t=1}^{T} s_{1t}$ and $\bar{s}_{2T} = T^{-1}\sum_{t=1}^{T} s_{2t}$, and for $K \geq d + q$. Under $H_0 : \theta_0^* = 0$, our goal is equivalent to provide a finite-sample probability of the following two-step GMM statistics:

$$\begin{aligned} t(\hat{\theta}_2^*) &= \frac{\sqrt{T}\left(\hat{\theta}_2^* - \theta_0^*\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \\ &= \frac{\sqrt{T}\left(\bar{s}_{1T} - \hat{\Omega}_{12}[\hat{\Omega}_{22}]^{-1}\bar{s}_{2T}\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}}, \end{aligned}$$

which is equivalent to the infeasible t-statistic $\mathcal{T}_T(\theta_0)$ at the original problem. We want to first provide the exact expression for

$$P(t(\hat{\theta}_2^*) \leq z) = \text{Actual finite sample probability}, \tag{6}$$

as $K \to \infty$ such that $K/T \to 0$. The expression for (6) is given Sun and Phillips (2009) with a kernel long-run variance estimator. However, in the literature, the expression for (6) using the orthonormal-series (OS) long-run variance estimator has not found yet. So, we fill this gap in this note.

Recall that $S = (s_1', s_2', \ldots, s_T') \in \mathbb{R}^{2T}$, with $s_t = (s_{1t}, s_{2t})'$. We first define the GLS of the over-identified system in (4)–(5), i.e.

$$\begin{aligned} \hat{\mu}_{\text{GLS}} = \begin{pmatrix} \hat{\theta}_{1,\text{GLS}}^* \\ \bar{s}_{2T,\text{GLS}} \end{pmatrix} &= \left[(1_T \otimes I_2)' \left[\tilde{\Sigma}_{2T}\right]^{-1} (1_T \otimes I_2) \right]^{-1} \left[(1_T \otimes I_2)' \left[\tilde{\Sigma}_{2T}\right]^{-1} S \right] \in \mathbb{R}^2 \\ &= \left[L_T' \left[\tilde{\Sigma}_{2T}\right]^{-1} L_T \right]^{-1} \left[L_T' \left[\tilde{\Sigma}_{2T}\right]^{-1} S \right], \end{aligned}$$

where

$$L_T = (1_T \otimes I_2) = \begin{pmatrix} I_2 \\ \vdots \\ I_2 \end{pmatrix} \in \mathbb{R}^{2T \times 2};$$

$$\begin{aligned} \tilde{\Sigma}_{2T} &= Var((s_1', s_2', \ldots, s_T')) \\ &= Var(S) \in \mathbb{R}^{2T \times 2T}. \end{aligned}$$

We define a GLS version of the two-step GMM estimator as below:

$$\begin{aligned} \hat{\theta}^*_{2,\mathrm{GLS}} &= \hat{\theta}^*_{1,\mathrm{GLS}} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\bar{s}_{2T,\mathrm{GLS}} \\ &= \begin{pmatrix} 1 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\theta}^*_{1,\mathrm{GLS}} \\ \bar{s}_{2T,\mathrm{GLS}} \end{pmatrix} \\ &:= \hat{\Pi}\hat{\mu}_{\mathrm{GLS}}, \end{aligned}$$

where

$$\hat{\Pi} = \begin{pmatrix} 1 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{pmatrix} \in \mathbb{R}^{1 \times 2}.$$

Note that the original estimator of the two-step GMM estimator can be also written by using the OLS estimator $\hat{\mu}_{\mathrm{OLS}}$ as below:

$$\begin{aligned} \hat{\theta}^*_{2} &= \hat{\theta}^*_{1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\bar{s}_{2T} \\ &= \begin{pmatrix} 1 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\theta}^*_{1} \\ \bar{s}_{2T} \end{pmatrix} \\ &:= \hat{\Pi}\hat{\mu}_{\mathrm{OLS}} \end{aligned}$$

with

$$\begin{aligned} \hat{\mu}_{\mathrm{OLS}} &= \begin{pmatrix} \hat{\theta}^*_{1} \\ \bar{s}_{2T} \end{pmatrix} = \begin{pmatrix} \bar{s}_{1T} \\ \bar{s}_{2T} \end{pmatrix} \\ &= (L_T'L_T)^{-1}L_T'S. \end{aligned}$$

Then, it is straightforward to check that

$$\begin{aligned} \hat{\theta}^*_2 - \theta^*_0 &= \hat{\theta}^*_{2,\mathrm{GLS}} - \theta^*_0 + \hat{\Pi}\left[\underbrace{(1_T \otimes I_2)'}_{L_T'}(1_T \otimes I_2)\right]^{-1}\left[(1_T \otimes I_2)'\tilde{W}\right] \\ &= \hat{\Pi}(\hat{\mu}_{\mathrm{GLS}} - \mu_0) + \hat{\Pi}\left(L_T'L_T\right)^{-1}\left(L_T'\tilde{W}\right); \\ &= \hat{\Pi}\left[(\hat{\mu}_{\mathrm{GLS}} - \mu_0) + \left(L_T'L_T\right)^{-1}\left(L_T'\tilde{W}\right)\right] \\ &= \hat{\Pi}(\hat{\mu}_{\mathrm{GLS}} - \mu_0) + \hat{\Pi}\left(L_T'L_T\right)^{-1}\left(L_T'\tilde{W}\right) \end{aligned}$$

where

$$\begin{aligned} \tilde{W} &= \left(I_{2T} - (1_T \otimes I_2)\left[(1_T \otimes I_2)'\left[\tilde{\Sigma}_{2T}\right]^{-1}(1_T \otimes I_2)\right]^{-1}(1_T \otimes I_2)'\left[\tilde{\Sigma}_{2T}\right]^{-1}\right)S \\ &= \left(I_{2T} - L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}L_T\right]^{-1}L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}\right)S, \end{aligned}$$

which can be equivalently read as

$$\begin{aligned}\hat{\theta}_2^* - \theta_0^* &= \hat{\Pi}\hat{\mu}_{\text{OLS}} \\ &= \hat{\Pi}(\hat{\mu}_{\text{GLS}} + \left(L_T'L_T\right)^{-1} L'\tilde{W}) \\ &= (\hat{\theta}_{2,\text{GLS}}^* - \theta_0^*) + \hat{\Pi}\left(L_T'L_T\right)^{-1} L_T'\tilde{W}. \end{aligned} \tag{7}$$

Using $\hat{\theta}_{2,\text{GLS}}^*$, we define a GLS-version of t and F statistics:

$$t(\hat{\theta}_{2,\text{GLS}}^*) = \frac{\sqrt{T}(\hat{\theta}_{2,\text{GLS}}^* - \theta_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}};$$

Define

$$\Omega_T = \begin{pmatrix} \Omega_{T,11} & \Omega_{T,12} \\ \Omega_{T,21} & \Omega_{T,22} \end{pmatrix} = var\left[\sqrt{T}\left(\hat{\mu}_{\text{GLS}} - \mu_0\right)\right]$$

and

$$\begin{aligned}\Omega_{T,11\cdot 2} &= \Omega_{T,11} - \Omega_{T,12}\left[\Omega_{T,22}\right]^{-1}\Omega_{T,21}; \\ \Pi_T &= \begin{pmatrix} 1 & -\Omega_{T,12}[\Omega_{T,22}]^{-1} \end{pmatrix}.\end{aligned}$$

In Lemma 4, we prove the asymptotic equivalence of $\hat{\theta}_2^*$ and its GLS-modification $\hat{\theta}_{2,\text{GLS}}^*$, which is a key step to prove our main results.

Assumption 4 (Linear_Process) *Assume a linear process for $v_t = (v_{1t}, v_{2t}') \in \mathbb{R}^{1+q}$:*

$$v_t = C(L)\epsilon_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j},$$

where $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}')' \overset{i.i.d}{\sim} N(0, I_{q+1})$, and C_j's are $(q+1)\times(q+1)$ matrix such that

$$\begin{aligned} E||\epsilon_t||^{\nu} &< \infty \text{ for some } \nu \geq 4; \\ \sum_{j=0}^{\infty} j^a ||C_j|| &< \infty \text{ for } a > 3; \\ \Omega = C(1)C(1)' &> 0, \end{aligned}$$

where $C(1) = \sum_{j=0}^{\infty} C_j$.

Lemma 4 *Assume $V \sim N(0, \tilde{\Sigma}_{2T})$. Then, for $K \geq d+q$ and $d = 1$, we have that*

i) $P\left(|t(\hat{\theta}_{2,GLS}^*)| \leq z\right) = E\left[G_1\left(\Xi^{-1}\right)\right] + O\left(\frac{1}{T}\right)$;

ii) $P\left(|t(\hat{\theta}_2^*)| \leq z\right) = P\left(\left|t(\hat{\theta}_{2,GLS}^*)\right| \leq z\right) + O\left(\frac{1}{T}\right)$,

where

$$\Xi = \left(1 + \frac{\left(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)\Omega_{22}\left(\Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)'}{\Omega_{11\cdot 2}}\right) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right).$$

Theorem 5 *Let Assumption 4 hold. As $K \to \infty$ such that $K/T \to 0$, we have that*

$$P_{H_0^*}\left(|t(\hat{\theta}_2^*)| \leq z\right) = G_1(z) + G_1'(z)z\left(\frac{K^2}{T^2}\right)\tilde{B} - G_1'(z)z\left(\frac{2q}{K}\right) + G_1''(z)z^2\left(\frac{1}{K}\right) + o\left(\frac{1}{K}\right),$$

where

$$\tilde{B} = \frac{B_{11} - 2\Omega_{12}\Omega_{22}^{-1}B_{21} + \left(\Omega_{12}\Omega_{22}^{-1} \otimes \Omega_{12}\Omega_{22}^{-1}\right) vec(B_{22})}{\Omega_{11\cdot 2}}.$$

Also, we have that

$$P(\mathbb{F}_\infty \leq z) = G_1(z) - G_1'(z)\, z\left(\frac{2q}{K}\right) + G_1''(z)\, z^2\left(\frac{1}{K}\right) + o\left(\frac{1}{K}\right),$$

as $K \to \infty$.

Together with our theoretical results in Lemmas 1–3, the higher order expansion of the over-identified location model shown in Theorem 5 leads to the main result of this paper as follows.

Theorem 6 *Let Assumptions 1–4 hold. As $K \to \infty$ and $T \to$ such that $K/T \to 0$, we have that*

$$\begin{aligned} P_{H_0}\left(\left|\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)\right| \leq z\right) &= G_1(z) + G_1'(z)z\tilde{B}\left(\frac{K^2}{T^2}\right) - 2G_1'(z)zq\left(\frac{1}{K}\right) + G_1''(z)z^2\left(\frac{1}{K}\right) \\ &\quad + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right). \end{aligned}$$

Under alternative asymptotics which assumes that K is fixed when $T \to \infty$, Theorem 1 in Sun (2014) shows that

$$\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1) \xrightarrow{d} \mathbb{T} \stackrel{d}{=} \frac{Z_1 - \mathbb{S}_{12}\mathbb{S}_{22}^{-1}Z_2}{\sqrt{\mathbb{S}_{11\cdot 2}}}, \tag{8}$$

where $Z_1 \sim N(0,1)$, $Z_2 \sim N(0, I_q)$, $Z_1 \perp Z_2$, and $\mathbb{S}_{11}$, $\mathbb{S}_{12}$, and $\mathbb{S}_{22}$ are submatrices of $\mathbb{S} \sim K^{-1}\mathbb{W}_p(K, I_m)$ which is a scaled Wishart random matrix with degree of freedom equal to K, i.e.,

$$\mathbb{S} = \begin{pmatrix} \underset{1\times 1}{\mathbb{S}_{11}} & \underset{1\times q}{\mathbb{S}_{12}} \\ \underset{q\times 1}{\mathbb{S}_{21}} & \underset{q\times q}{\mathbb{S}_{22}} \end{pmatrix}.$$

Expanding the fixed-K limiting distribution, $\mathbb{T}$, by letting $K \to \infty$, Theorem 3-(c) in Sun (2014) implies that

$$P(|\mathbb{T}| > z) = 1 - G_1(z) + 2G_1'(z)zq\left(\frac{1}{K}\right) - G_1''(z)z^2\left(\frac{1}{K}\right) + o\left(\frac{1}{K}\right).$$

Thus, we obtain that

$$\alpha + o\left(\frac{1}{K}\right) = 1 - G_1\left(\mathbb{T}_{1-\alpha/2}\right) + 2G_1'(\mathbb{T}_{1-\alpha/2})\mathbb{T}_{1-\alpha/2}q\left(\frac{1}{K}\right) - G_1''(\mathbb{T}_{1-\alpha/2})\mathbb{T}_{1-\alpha/2}^2\left(\frac{1}{K}\right),$$

where $\mathbb{T}_{1-\alpha/2}$ is two-sided fixed-K asymptotic critical value such that $P(|\mathbb{T}| > \mathbb{T}_{1-\alpha/2}) = \alpha$. Then, using the result in Theorem 6, we have that

$$P_{H_0}\left(\left|\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)\right| > \mathbb{T}_{1-\alpha/2}\right) = \alpha + O\left(\frac{K^2}{T^2}\right) + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right), \tag{9}$$

whereas the null rejection probability using conventional normal critical value, $z_{1-\alpha/2}$, yields

$$P_{H_0}\left(\left|\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)\right| > z_{1-\alpha/2}\right) = \alpha + O\left(\frac{K^2}{T^2}\right) + O\left(\frac{(1+q)}{K}\right) + O\left(\frac{\log T}{\sqrt{T}}\right) + o\left(\frac{K^2}{T^2}\right) + o\left(\frac{1}{K}\right).$$

Therefore, using the alternative fixed-K critical value, we are able to remove the error of order $O\left((1+q)/K\right)$ embodied in the standard normal t-test. In this sense, our result in Theorem 6 is a key device to show the second-order correctness of the alternative fixed-K asymptotics in two-step GMM framework. In addition, one can implement Windmeijer-type correction in $\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)$, as in Hwang and Valdés (2021), and further correct the $O(\log T/\sqrt{T})$ error in (9).

4 Type-I Error Optimal K Choice in Two-step GMM

Suppose we are interested in testing the parameter θ_0. Using a conventional χ_1^α critical value, one can construct a two-sided confidence interval of the form $\{\theta : |\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)|^2 \leq \chi_1^\alpha\}$, and reject the null hypothesis $H_0 : \theta = \theta_0$ if and only if $\theta \notin \{\theta : |\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)|^2 \leq \chi_1^{1-\alpha}\}$, where $1 - G_1\left(\chi_1^{1-\alpha}\right) = \alpha$. Then up to small order terms, $o(K^{-1})$, its Type-I error, or false rejection probability, is

$$\begin{aligned}
& \left|P\left(|\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)|^2 > \chi_1^\alpha\right) - \alpha\right| \\
= & \left|1 - G_1\left(\chi_1^{1-\alpha}\right) - G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\left(\frac{K^2}{T^2}\right)\tilde{B}\right. \\
& \left. + G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\left(\frac{2q}{K}\right) - G_1''(\chi_1^{1-\alpha})(\chi_1^{1-\alpha})^2\left(\frac{1}{K}\right) - \alpha\right| \\
= & \left|G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\left(\frac{2q}{K}\right) - G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\left(\frac{K^2}{T^2}\right)\tilde{B} - G_1''(\chi_1^{1-\alpha})(\chi_1^{1-\alpha})^2\left(\frac{1}{K}\right)\right| \\
= & \left|\left(\frac{K^2}{T^2}\right)G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\tilde{B} + \left(\frac{1}{K}\right)G_1''(\chi_1^{1-\alpha})(\chi_1^{1-\alpha})^2 - \left(\frac{2q}{K}\right)G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\right| \\
= & \left|\left(\frac{K^2}{T^2}\right)G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\tilde{B} + \left(\frac{1}{K}\right)\left(G_1''(\chi_1^{1-\alpha})(\chi_1^{1-\alpha})^2 - 2qG_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\right)\right| \\
\leq & \left(\frac{K^2}{T^2}\right)\left|G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\tilde{B}\right| + \left(\frac{1}{K}\right)\left|\left(G_1''(\chi_1^{1-\alpha})(\chi_1^{1-\alpha})^2 - 2qG_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\right)\right|.,
\end{aligned}$$

To control the type-I error, we can choose optimal K^* to minimize the upper bound, and this leads us to F.O.C.

$$\left(\frac{2K^*}{T^2}\right)\left|G_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\tilde{B}\right| - \frac{1}{(K^*)^2}\left|\left(G_1''(\chi_1^{1-\alpha})(\chi_1^{1-\alpha})^2 - 2qG_1'(\chi_1^{1-\alpha})\chi_1^{1-\alpha}\right)\right| = 0;$$

and we have that

$$K^* = \left\lceil \left(\frac{\left|\left(G_1''(\chi_1^{1-\alpha})\chi_1^{1-\alpha} - 2qG_1'(\chi_1^{1-\alpha})\right)\right|}{2G_1'(\chi_1^{1-\alpha})\left|\tilde{B}\right|}\right)^{1/3} \cdot T^{2/3} \right\rceil,$$

where $\lceil \cdot \rceil$ is the ceiling function. When $q = 0$, i.e., the model is exactly identified with $\Omega_{12} = 0$ and $B_{21} = 0$, we have that

$$\begin{aligned} \tilde{B} &= \frac{B_{11} - 2\Omega_{12}\Omega_{22}^{-1}B_{21} + \left(\Omega_{12}\Omega_{22}^{-1} \otimes \Omega_{12}\Omega_{22}^{-1}\right) vec(B_{22})}{\Omega_{11\cdot 2}} \\ &= \frac{B_{11}}{\Omega_{11}}. \end{aligned}$$

Then, the optimal choice for K^* is equal to

$$K^* = \left\lceil \left(\frac{\left| \left(G_1''(\chi_1^{1-\alpha})\chi_1^{1-\alpha} \right) \right|}{2G_1'(\chi_1^{1-\alpha}) \left| B_{11}/\Omega_{11} \right|} \right)^{1/3} \cdot T^{2/3} \right\rceil,$$

which is exactly the result for CPE-optimized K obtained in Sun (2013). Therefore, our result is a generalization of the optimal K-formula when we consider an over-identified GMM model. The optimal formula for K^* establishes a relation between the value of q and other model parameters such as Ω_{12}, Ω_{22}, and B_{21} in the over-identified GMM system.

5 Conclusion

In this paper, we develop a higher-order expansion of test statistic for heteroskedasticity autocorrelation robust (HAR) inference in an over-identified two-step GMM framework. Our new higher-order expansion rigorously shows the higher-order accuracy of the fixed-smoothing asymptotics in finite sample null rejection probability. We also provide an optimal choice of smoothing parameters that directly accounts for the degree of over-identificaion in HAR inferential problem. Our formula for the optimal smoothing parameter generalizes the formula in Sun (2013) which works under the exact identified GMM, as both formulas coincide when there is no over-identification.

References

[1] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–854.

[2] Anatolyev, S. (2005). GMM, GEL, serial correlation, and asymptotic bias. Econometrica, 73(3), 983-1002.

[3] Christiano, L. J. (1996). Special Section on Small-Sample Properties of Generalized Method of Moments (GMM) Associate Editor's Introduction. Journal of Business & Economic Statistics, 14(3), 261-261.

[4] Hall, A. R. (2005). Generalized method of moments. Oxford university press.

[5] Hall, P., & Horowitz, J. L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. Econometrica: Journal of the Econometric Society, 891-916.

[6] Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-sample properties of some alternative GMM estimators. Journal of Business & Economic Statistics, 14(3), 262-280.

[7] Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica: Journal of the Econometric Society, 1269-1286.

[8] Hwang, J., & Sun, Y. (2017). Asymptotic F and t Tests in an Efficient GMM Setting. Journal of econometrics, 198(2), 277-295.

[9] Hwang, J., & Sun, Y. (2018). Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework. Journal of econometrics, 207(2), 381-405.

[10] Kiefer, N. M., & Vogelsang, T. J. (2002). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. Econometrica, 70(5), 2093-2095.

[11] Kiefer, N. M., & Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory, 21(6), 1130-1164.

[12] Kitamura, Y., & Stutzer, M. (1997). An information-theoretic alternative to generalized method of moments estimation. Econometrica: Journal of the Econometric Society, 861-874.

[13] Lazarus, E., Lewis, D. J., & Stock, J. H. (2019). The size-power tradeoff in HAR inference. Available at SSRN 3436372.

[14] Lazarus, E., Lewis, D. J., Stock, J. H., & Watson, M. W. (2018). HAR Inference: Recommendations for Practice. Journal of Business & Economic Statistics, 36(4), 541-559.

[15] Müller, U. K. (2007). A theory of robust long-run variance estimation. Journal of Econometrics, 141(2), 1331-1352.

[16] Müller, U. K., & Watson, M. W. (2018). Long-Run Covariability. Econometrica, 86(3), 775-804.

[17] Newey, W. K., & Smith, R. J. (2004). Higher order properties of GMM and generalized empirical likelihood estimators. Econometrica, 72(1), 219-255.

[18] Phillips, P. C. (2005). Automated discovery in econometrics. Econometric Theory, 21(1), 3-20.

[19] Sun, Y. (2013). A heteroskedasticity and autocorrelation robust F test using an orthonormal series variance estimator. The Econometrics Journal, 16(1), 1-26.

[20] Sun, Y. (2014a). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. Journal of Econometrics, 178, 659-677.

[21] Sun, Y. (2014b). Fixed-smoothing asymptotics in a two-step generalized method of moments framework. Econometrica, 82(6), 2327-2370.

[22] Sun, Y., Phillips, P. C., & Jin, S. (2008). Optimal bandwidth selection in heteroskedasticity—autocorrelation robust testing. Econometrica, 76(1), 175-194.

[23] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of econometrics, 126(1), 25-51.

6 Appendix of Proofs

Proof of Lemma 1. Recall that

$$\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1)=\frac{\sqrt{T}(\hat{\theta}_2-\theta_0)}{\sqrt{[G_T'[\Sigma_T(\hat{\theta}_1)]^{-1}G_T]^{-1}}}.$$

To investigate the asymptotic behavior of $\hat{\theta}_2$, we look at F.O.C for $\hat{\theta}_2$ which is given by :

$$G_T'[\Sigma_T(\hat{\theta}_1)]^{-1}f_T(\hat{\theta}_2)=0,$$

and thus we have that

$$\sqrt{T}(\hat{\theta}_2-\theta_0)=-(G_T'[\Sigma_T(\hat{\theta}_1)]^{-1}G_T')^{-1}G_T'[\Sigma_T(\hat{\theta}_1)]^{-1}\sqrt{T}f_T(\theta_0),$$

by the linearity of moment process. From equation (10) in Hwang and Valdés (2020), we have that

$$\sqrt{T}(\hat{\theta}_2-\theta_0)=-(G_T'[\Sigma_T(\theta_0)]^{-1}G_T')^{-1}G_T'[\Sigma_T(\theta_0)]^{-1}\sqrt{T}f_T(\theta_0)+D_{\theta_0,\Sigma_T(\theta_0)}\sqrt{T}(\hat{\theta}_1-\theta_0)+o_p\left(\frac{1}{\sqrt{T}}\right),$$

where

$$\begin{aligned}
D_{\theta_0,\Sigma_T(\theta_0)} &= \left.\frac{\partial-(G_T'\Sigma_T^{-1}(\theta)G_T)^{-1}G_T'\Sigma_T^{-1}(\theta)\sqrt{T}f_T(\theta_0)}{\partial\theta'}\right|_{\theta=\theta_0}\\
&= \underbrace{-(G_T'\Sigma_T^{-1}(\theta_0)G_T)^{-1}G_T'\Sigma_T^{-1}(\theta_0)\left.\frac{\partial\Sigma_T(\theta)}{\partial\theta}\right|_{\theta=\theta_0}\Sigma_T^{-1}(\theta_0)G_T'(G_T'\Sigma_T^{-1}(\theta_0)G_T)^{-1}G_T'\Sigma_T^{-1}(\theta_0)f_T(\theta_0)}_{:=D_I}\\
&\quad\underbrace{+(G_T'\Sigma_T^{-1}(\theta_0)G_T)^{-1}G_T'\Sigma_T^{-1}(\theta_0)\left.\frac{\partial\Sigma_T(\theta)}{\partial\theta}\right|_{\theta=\theta_0}\Sigma_T^{-1}(\theta_0)f_T(\theta_0)}_{:=D_{II}},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial\Sigma_T(\theta)}{\partial\theta} &= \Upsilon_T(\theta)+\Upsilon_T(\theta)';\\
\Upsilon_T(\theta) &= \frac{1}{K}\sum_{j=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j\left(\frac{t}{T}\right)\frac{\partial f(w_t,\theta)}{\partial\theta}\right)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\phi_j\left(\frac{s}{T}\right)f(w_s,\theta)\right)'.
\end{aligned}$$

By Assumptions 1 and 2, we have that

$$\begin{aligned}
G_T &= G+O_p\left(\frac{1}{\sqrt{T}}\right);\\
\Sigma_T^{-1}(\theta_0) &= \Sigma^{-1}+O_p\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}\right);\\
\left.\frac{\partial\Sigma_T(\theta)}{\partial\theta}\right|_{\theta=\theta_0} &= \Upsilon+\Upsilon'+O_p\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}\right),
\end{aligned}$$

and using these equations, we obtain that

$$\begin{aligned} G_T'\Sigma_T^{-1}(\theta_0)G_T &= \left(G' + O_p\left(\frac{1}{\sqrt{T}}\right)\right)\left(\Sigma^{-1} + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right)\right)\left(G + O_p\left(\frac{1}{\sqrt{T}}\right)\right) \\ &= (G'\Sigma^{-1}G) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) + O_p\left(\frac{1}{T}\right), \end{aligned}$$

which implies that

$$\left(G_T'\Sigma_T^{-1}(\theta_0)G_T\right)^{-1} = (G'\Sigma^{-1}G)^{-1} + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) + O_p\left(\frac{1}{T}\right).$$

Similarly, we have that

$$\begin{aligned} & G_T'\Sigma_T^{-1}(\theta_0)\left.\frac{\partial\Sigma_T(\theta)}{\partial\theta}\right|_{\theta=\theta_0}\Sigma_T^{-1}(\theta_0)G_T' \\ = \; & G'\Sigma^{-1}\left(\Upsilon + \Upsilon'\right)\Sigma^{-1}G + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) + O_p\left(\frac{1}{T}\right) \end{aligned}$$

Also, we have that

$$\begin{aligned} G_T'\Sigma_T^{-1}(\theta_0)f_T(\theta_0) &= \left(G' + O_p\left(\frac{1}{\sqrt{T}}\right)\right) \times \left(\Sigma^{-1} + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right)\right) O_p\left(\frac{1}{\sqrt{T}}\right) \\ &= O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}} + \frac{1}{\sqrt{T}}\right)\right) + O_p\left(\frac{1}{T}\right). \end{aligned}$$

Careful inspections of the above equations show that the terms $O_p(T^{-1/2})$ does not depend on K. Therefore, we have that

$$\begin{aligned} D_I &= -(G_T'\Sigma_T^{-1}(\theta_0)G_T)^{-1}\left\{G_T'\Sigma_T^{-1}(\theta_0)\left.\frac{\partial\Sigma_T(\theta)}{\partial\theta}\right|_{\theta=\theta_0}\Sigma_T^{-1}(\theta_0)G_T'\right\}(G_T'\Sigma_T^{-1}(\theta_0)G_T)^{-1}G_T'\Sigma_T^{-1}(\theta_0)f_T(\theta_0). \\ &= -\left\{(G'\Sigma^{-1}G)^{-1} + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) + O_p\left(\frac{1}{T}\right)\right\} \\ &\quad \times\left\{G'\Sigma^{-1}\left(\Upsilon + \Upsilon'\right)\Sigma^{-1}G + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) + O_p\left(\frac{1}{T}\right)\right\} \\ &\quad \times\left\{(G'\Sigma^{-1}G)^{-1} + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}}\right) + +O_p\left(\frac{1}{T}\right)\right\} \\ &\quad \times\left(O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}} + \frac{1}{\sqrt{T}}\right)\right) + O_p\left(\frac{1}{T}\right)\right) \\ &= O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}} + \frac{1}{\sqrt{T}}\right), \end{aligned}$$

where the term inside $O_p(T^{-1/2})$ does not depend on K. Similarly, we have that the second term admits the same order of the expansion derived in the last equation, i.e.

$$\begin{aligned} D_{II} &= -(G'\Sigma^{-1}G)^{-1}G'\Sigma^{-1}\left(\Upsilon + \Upsilon'\right)\Sigma^{-1}f_T(\theta_0) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}} + \frac{1}{\sqrt{T}}\right)\right) \\ &= O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2} + \frac{1}{\sqrt{K}} + \frac{1}{\sqrt{T}}\right)\right). \end{aligned}$$

Using similar expansions, we can also show that

$$\begin{aligned}
\left(G_T\Sigma_T^{-1}(\hat{\theta}_1)G_T\right) &= (G'\Sigma_T^{-1}(\theta_0)G) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}+\frac{1}{\sqrt{T}}\right)\right); \\
G_T'[\Sigma_T(\theta_0)]^{-1}\sqrt{T}f_T(\theta_0) &= G'[\Sigma_T(\theta_0)]^{-1}\sqrt{T}f_T(\theta_0) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}+\frac{1}{\sqrt{T}}\right)\right).
\end{aligned}$$

Combining these results together, we have that

$$\begin{aligned}
\sqrt{T}(\hat{\theta}_2-\theta_0) &= -(G'[\Sigma_T(\theta_0)]^{-1}G')^{-1}G'[\Sigma_T(\theta_0)]^{-1}\sqrt{T}f_T(\theta_0) \\
&\quad + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}+\frac{1}{\sqrt{T}}\right)\right),
\end{aligned}$$

where the term of order $O_p(T^{-1/2})$ does not depend on K. The results imply that

$$\begin{aligned}
\mathbb{T}_T(\hat{\theta}_2;\hat{\theta}_1) &= \frac{\sqrt{T}(\hat{\theta}_2-\theta_0)}{\sqrt{[G_T'\Sigma_T^{-1}(\hat{\theta}_1)G_T]^{-1}}} \\
&= \left(-\left(G'\Sigma_T^{-1}(\theta_0)G\right)^{-1}G'\Sigma_T^{-1}(\theta_0)\sqrt{T}f_T(\theta_0) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}+\frac{1}{\sqrt{T}}\right)\right)\right) \\
&\quad \times \left((G'\Sigma_T^{-1}(\theta_0)G)^{1/2} + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{T}}\left(\frac{K^2}{T^2}+\frac{1}{\sqrt{K}}+\frac{1}{\sqrt{T}}\right)\right)\right) \\
&= \mathcal{T}_T(\theta_0) + \psi_T + \psi_T^*,
\end{aligned}$$

where $\psi_T = O_p(T^{-1/2})$ does not depend on K, and $\psi_T^* = O_p\left(T^{-1/2}\left(K^2/T^2 + K^{-1/2} + T^{-1/2}\right)\right)$. This completes the proof. ■

Proof of Lemma 3. We first prove Part (b). Using the SVD $\mathbb{U}\Psi\mathbb{V}'$ of G, we have

$$\begin{aligned}
G'\Sigma_T^{-1}(\theta_0)G &= \mathbb{V}\Psi'\mathbb{U}'\Sigma_T(\theta_0)^{-1}\mathbb{U}\Psi\mathbb{V}' \\
&= \mathbb{V}\Psi'(\mathbb{U}'\Sigma_T(\theta_0)\mathbb{U})^{-1}\Psi\mathbb{V}' \\
&= \mathbb{V}A\begin{pmatrix} I_d, & O \end{pmatrix}\hat{\Omega}^{-1}\begin{pmatrix} I_d, & O \end{pmatrix}'A\mathbb{V}' \\
&= \mathbb{V}A\hat{\Omega}_{11\cdot 2}^{-1}A\mathbb{V}',
\end{aligned}$$

which is the desired result. In the last equation, we have used the following matrix inverse formula

$$\begin{aligned}
\hat{\Omega}^{-1} &= \left[\begin{pmatrix} \hat{\Omega}_{11} & \hat{\Omega}_{12} \\ \hat{\Omega}_{21} & \hat{\Omega}_{22} \end{pmatrix}\right]^{-1} \\
&= \begin{pmatrix} \hat{\Omega}_{11\cdot 2}^{-1} & -\hat{\Omega}_{11\cdot 2}^{-1}\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \\ -\hat{\Omega}_{22}^{-1}\hat{\Omega}_{21}\hat{\Omega}_{11\cdot 2}^{-1} & \left(\hat{\Omega}_{22}^{-1} + \hat{\Omega}_{22}^{-1}\hat{\Omega}_{21}\hat{\Omega}_{11\cdot 2}^{-1}\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right) \end{pmatrix}.
\end{aligned}$$

To prove Part (a), we have that

$$\begin{aligned}
& G'\Sigma_T^{-1}(\theta_0)f(w_t,\theta_0) \\
= \;& \mathbb{V}\Psi'(\mathbb{U}_T'\Sigma_T(\theta_0)\mathbb{U})^{-1}\mathbb{U}'f(w_t,\theta_0) \\
= \;& \mathbb{V}A\left(\begin{array}{cc} I_d, & O \end{array}\right)\hat{\Omega}^{-1}f^*(w_t,\theta_0) \\
= \;& \mathbb{V}A\left(\begin{array}{cc} I_d, & O \end{array}\right)\left(\begin{array}{cc} \hat{\Omega}_{11\cdot 2}^{-1} & -\hat{\Omega}_{11\cdot 2}^{-1}\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \\ -\hat{\Omega}_{22}^{-1}\hat{\Omega}_{21}\hat{\Omega}_{11\cdot 2}^{-1} & \left(\hat{\Omega}_{22}^{-1}+\hat{\Omega}_{22}^{-1}\hat{\Omega}_{21}\hat{\Omega}_{11\cdot 2}^{-1}\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right) \end{array}\right)f^*(w_t,\theta_0) \\
= \;& \mathbb{V}A\left(\begin{array}{cc} \hat{\Omega}_{11\cdot 2}^{-1}, & -\hat{\Omega}_{11\cdot 2}^{-1}\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{array}\right)f^*(w_t,\theta_0) \\
= \;& \mathbb{V}A\hat{\Omega}_{11\cdot 2}^{-1}\left[f_1^*(w_t,\theta_0)-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}f_2^*(w_t,\theta_0)\right].
\end{aligned}$$

Combining this result into the one in Part (b), we have that

$$\begin{aligned}
& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[G'\Sigma_T^{-1}(\theta_0)G\right]^{-1}G'\Sigma_T^{-1}(\theta_0)f(v_t,\theta_0) \\
= \;& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\mathbb{V}A\hat{\Omega}_{11\cdot 2}^{-1}A\mathbb{V}'\right]^{-1}\mathbb{V}A\hat{\Omega}_{11\cdot 2}^{-1} \\
& \times\left[f_1^*(w_t,\theta_0)-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}f_2^*(w_t,\theta_0)\right] \\
= \;& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbb{V}A^{-1}\left[f_1^*(w_t,\theta_0)-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}f_2^*(w_t,\theta_0)\right], \qquad (10)
\end{aligned}$$

as desired. ■

6.1 Proof of Lemma 4

Before we prove Lemma 4, we first provide some technical results in the following Lemma.

Lemma 7 *Let $\hat{\Pi}$, $\tilde{W}$, and $\hat{\mu}_{GLS}$ be defined as before, and Assumption 4 holds, so that $V \sim N(0,\tilde{\Sigma}_{2T})$. Then, we have*

i) $\hat{\Omega}$ and $\hat{\mu}_{GLS}$ are independent.

ii) $\hat{\Pi} = \left(\begin{array}{cc} 1 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{array}\right)$ is independent with $\hat{\mu}_{GLS}$

iii) $\hat{\mu}_{GLS}$ and $\tilde{W}$ are independent.

Proof of Lemma 7. Part i) Note that we can represent $\hat{\Omega}$ as:

$$\hat{\Omega} = \frac{1}{K}\sum_{j=1}^{K}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j(\frac{t}{T})(v_t-\bar{v}_T)\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\phi_j(\frac{t}{T})(v_t-\bar{v}_T)\right)',$$

where

$$v_t-\bar{v}_T = \left(\begin{array}{c} v_{1t}-\bar{v}_{1T} \\ v_{1t}-\bar{v}_{2T} \end{array}\right) = \left(\begin{array}{c} s_{1t}-\bar{s}_{1T} \\ s_{2t}-\bar{s}_{2T} \end{array}\right).$$

Let $\hat{V} = (v_1'-\bar{v}_T',\ldots,v_T'-\bar{v}_T')' \in \mathbb{R}^{2T}$, then it is easy to check that

$$\hat{V} = \left(I_{2T}-L_T(L_T'L_T)^{-1}L_T'\right)V$$

$$\hat{\mu}_{\text{GLS}}-\mu_0 = \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}L_T\right]^{-1}L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}V$$

with $V = (v_1', v_2', \ldots, v_T)'$ and $v_t = (v_{1t}, v_{2t})'$. Thus

$$\begin{aligned}
&E\left[(\hat{\mu}_{\mathrm{GLS}} - \mu_0)\hat{V}'\right] \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} \times \underbrace{E[VV']}_{:=\tilde{\Sigma}_{2T}} \times \left(I_{2T} - L_T(L_T'L_T)^{-1}L_T'\right) \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T' \times \left(I_{2T} - L_T(L_T'L_T)^{-1}L_T'\right) \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} \left(L_T' - L_T'\right) = 0,
\end{aligned}$$

which implies that $\hat{\mu}_{\mathrm{GLS}} - \mu_0$ and $\hat{V}$ are independent. Since $\hat{\Omega}$ can be represented as a quadratic form of $\hat{V}$, the result also indicates that $\hat{\mu}_{\mathrm{GLS}} - \mu_0$ and $\hat{\Omega}$ are independent. Part ii) Since $\hat{\Pi}$ is a function of the matrix $\hat{\Omega}$, we have that $\hat{\mu}_{\mathrm{GLS}} - \mu_0$ and $\hat{\Pi}$ are independent, which is the desired result. Part iii) Recall that

$$\tilde{W} = \left(I_{2T} - L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}\right) V,$$

because $S = L_T\mu_0 + V$ and

$$\begin{aligned}
&\left(I_{2T} - L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}\right) L_T\mu_0 \\
&= L_T\mu_0 - L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\mu_0 \\
&= L_T\mu_0 - L_T\mu_0 \\
&= 0.
\end{aligned}$$

Using this, we have

$$\begin{aligned}
&E\left[(\hat{\mu}_{\mathrm{GLS}} - \mu_0)\tilde{W}'\right] \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} \times \underbrace{E[VV']}_{:=\tilde{\Sigma}_{2T}} \times \left(I_{2T} - L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}\right)' \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T' \times \left[I_{2T} - \left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\right] \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T' - \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T' \\
&= \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T' - \left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1} L_T\right]^{-1} L_T' \\
&= 0,
\end{aligned}$$

which leads to the desired result. ■

Proof of Lemma 4. Part i) With $d = 1$, we first write the statistic $F(\hat{\theta}_{2,\mathrm{GLS}})$ as

$$\begin{aligned}
F(\hat{\theta}^*_{2,\mathrm{GLS}}) &= \left[t(\hat{\theta}^*_{2,\mathrm{GLS}})\right]^2 \\
&= \left(\frac{\sqrt{T}(\hat{\theta}^*_{2,\mathrm{GLS}} - \theta^*_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}}\right)^2 \\
&= \left(\frac{\sqrt{T}\hat{\Pi}\left(\hat{\mu}_{\mathrm{GLS}} - \mu_0\right)}{\sqrt{\hat{\Omega}_{11\cdot 2}}}\right)^2 \\
&= \left(\frac{\sqrt{T}\left\{\Pi_T\left(\hat{\mu}_{\mathrm{GLS}} - \mu_0\right) + \left(\hat{\Pi} - \Pi_T\right)\left(\hat{\mu}_{\mathrm{GLS}} - \mu_0\right)\right\}}{\sqrt{\Omega_{T,11\cdot 2}}}\right)^2 \times \left(\frac{\Omega_{T,11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right)
\end{aligned}$$

Using the result in Grenader and Rosenblatt (1957), we have that

$$\Omega_T = \Omega\left(1 + O\left(\frac{1}{T}\right)\right), \tag{11}$$

and thus

$$\Omega_{T,11\cdot 2} = \Omega_{11\cdot 2}\left(1 + O\left(\frac{1}{T}\right)\right); \tag{12}$$

$$\Pi_T = \Pi\left(1 + O\left(\frac{1}{T}\right)\right). \tag{13}$$

Therefore, we obtain

$$F(\hat{\theta}_{2,\mathrm{GLS}}) = \left(\frac{\Pi_T\sqrt{T}\left(\hat{\mu}_{\mathrm{GLS}} - \mu_0\right)}{\sqrt{\Omega_{T,11\cdot 2}}} + \frac{\left(\hat{\Pi} - \Pi\right)\sqrt{T}\left(\hat{\mu}_{\mathrm{GLS}} - \mu_0\right)}{\sqrt{\Omega_{T,11\cdot 2}}}\right)^2 \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) + O\left(\frac{1}{T}\right).$$

To further proceed the above expression, we define

$$\begin{aligned}
A_T &= \frac{\Pi_T\sqrt{T}(\hat{\mu}_{\mathrm{GLS}} - \mu_0)}{\sqrt{\Omega_{T,11\cdot 2}}}; \\
B_T &= \frac{\left(\hat{\Pi} - \Pi\right)\sqrt{T}(\hat{\mu}_{\mathrm{GLS}} - \mu_0)}{\sqrt{\Omega_{T,11\cdot 2}}},
\end{aligned}$$

and want to show that

$$A_T \sim N(0,1) \text{ and } B_T \sim N\left(0, \Phi_T\right); \tag{14}$$

$$A_T \text{ and } B_T \text{ are independent,} \tag{15}$$

where

$$\Phi_T = \frac{\delta_T\Omega_{T,22}\delta_T'}{\Omega_{T,11\cdot 2}}.$$

We first prove $A_T \sim N(0,1)$ in (14). Note that we assumed that $V \sim N(0, \tilde{\Sigma}_{2T})$, and

$$\sqrt{T}(\hat{\mu}_{\mathrm{GLS}} - \mu_0) = \sqrt{T}\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}L_T\right]^{-1}\left[L_T'\left[\tilde{\Sigma}_{2T}\right]^{-1}V\right]$$

, which indicates that

$$\sqrt{T}(\hat{\mu}_{\text{GLS}} - \mu_0) \sim N(0, \Omega_T),$$

and thus $\sqrt{T}\Pi_{T,\text{GLS}}(\hat{\mu}_{\text{GLS}} - \mu_0) \sim N(0, \Psi_T)$, where

$$\begin{aligned}
\Psi_T &= \Pi_T \Omega_T \Pi_T' \\
&= \begin{pmatrix} 1 & -\Omega_{T,12}[\Omega_{T,22}]^{-1} \end{pmatrix} \begin{pmatrix} \Omega_{T,11} & \Omega_{T,12} \\ \Omega_{T,21} & \Omega_{T,22} \end{pmatrix} \begin{pmatrix} 1 \\ -[\Omega_{T,22}]^{-1}\Omega_{T,21} \end{pmatrix} \\
&= \begin{pmatrix} \Omega_{T,11} - \Omega_{T,12}[\Omega_{T,22}]^{-1}\Omega_{T,21} & \Omega_{T,12} - \Omega_{T,12}[\Omega_{T,22}]^{-1}\Omega_{T,22} \end{pmatrix} \\
&\times \begin{pmatrix} 1 \\ -[\Omega_{T,22}]^{-1}\Omega_{T,21} \end{pmatrix} \\
&= \begin{pmatrix} \Omega_{T,11\cdot 2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -[\Omega_{T,22}]^{-1}\Omega_{T,21} \end{pmatrix} \\
&= \Omega_{T,11\cdot 2}.
\end{aligned}$$

The result leads us to have

$$\frac{\Pi_T \sqrt{T}(\hat{\mu}_{\text{GLS}} - \mu_0)}{\sqrt{\Omega_{T,11\cdot 2}}} \sim N(0,1),$$

as desired. To prove the second result in (14), note that

$$\begin{aligned}
\hat{\Pi} - \Pi &= \begin{pmatrix} 1 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{pmatrix} - \begin{pmatrix} 1 & -\Omega_{12}\Omega_{22}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} 0 & -\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} + \Omega_{12}\Omega_{22}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} 0 & \Omega_{12}\Omega_{22}^{-1} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1} \end{pmatrix} \\
&:= \begin{pmatrix} 0 & \delta_T \end{pmatrix}.
\end{aligned}$$

Then, we have that

$$\begin{aligned}
B_T &= \frac{\left(\hat{\Pi} - \Pi\right)\sqrt{T}\left(\hat{\mu}_{\text{GLS}} - \mu_0\right)}{\sqrt{\Omega_{T,11\cdot 2}}} \\
&= \begin{pmatrix} 0 & \delta_T \end{pmatrix} \sqrt{T} \begin{pmatrix} \bar{s}_{2T,\text{GLS}} - \theta_0 \\ \bar{s}_{2T,\text{GLS}} \end{pmatrix} / \sqrt{\Omega_{T,11\cdot 2}} \\
&= \frac{\delta_T \sqrt{T}\bar{s}_{2T,\text{GLS}}}{\sqrt{\Omega_{T,11\cdot 2}}}.
\end{aligned}$$

Since $\sqrt{T}\bar{s}_{2T,\text{GLS}} \sim N(0, \Omega_{T,22})$, this implies that

$$B_T \sim N\left(0, \Phi_T\right) \text{ with } \Phi_T = \frac{\delta_T \Omega_{T,22} \delta_T'}{\Omega_{T,11\cdot 2}}.$$

We now prove the independence between A_T and B_T in (15) which is equivalent to prove that

$$\sqrt{T}\bar{s}_{2T,\text{GLS}} \text{ and } \Pi_T\sqrt{T}(\hat{\mu}_{\text{GLS}} - \mu_0) \text{ are independent.} \tag{16}$$

We have that

$$\begin{aligned}
&\Pi_{T,\text{GLS}}\sqrt{T}(\hat{\mu}_{\text{GLS}} - \mu_0) = \sqrt{T}(\bar{s}_{1T,\text{GLS}} - \theta_0) - \Omega_{T,12}\Omega_{T,22}^{-1}\sqrt{T}\bar{s}_{2T,\text{GLS}}; \\
&\sqrt{T}\begin{pmatrix} \bar{s}_{1T,\text{GLS}} - \theta_0 \\ \bar{s}_{2T,\text{GLS}} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega_{T,11} & \Omega_{T,12} \\ \Omega_{T,21} & \Omega_{T,22} \end{pmatrix}\right),
\end{aligned}$$

which directly implies (16). Summing up so far, we have proved that

$$F(\hat{\theta}_{2,\mathrm{GLS}}) = (A_T + B_T)^2 \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) + O\left(\frac{1}{T}\right)$$
$$A_T \sim N(0,1) \;\perp\; B_T \sim N\left(0, \Phi_T\right),$$

and this implies that

$$A_T + B_T \sim N(0, 1+\Phi_T) \stackrel{d}{=} N(0,1) \times \sqrt{1+\Phi_T}$$

So we have that

$$F(\hat{\theta}_{2,\mathrm{GLS}}) = \Theta \times (1+\Phi_T) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) + O\left(\frac{1}{T}\right);$$
$$\Theta \sim \chi^2(1) \text{ and } \Phi_T = \frac{\delta_T \Omega_{T,22} \delta_T'}{\Omega_{T,11\cdot 2}}.$$

Here Θ is independent of $\hat{\Omega}_{11\cdot 2}$ and Φ_T by the results in Lemma 7 (i) and (ii). From (11), we have that

$$\begin{aligned}
(1+\Phi_T) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) &= \left(1 + \frac{\delta_T \Omega_{T,22} \delta_T'}{\Omega_{T,11\cdot 2}}\right) \times \left(\frac{\Omega_{T,11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) \\
&= \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) \times \left(1 + O\left(\frac{1}{T}\right)\right) + \frac{\delta_T \Omega_{T,22} \delta_T'}{\hat{\Omega}_{11\cdot 2}} \\
&= \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) \times \left(1 + O\left(\frac{1}{T}\right)\right) + \delta_T \Omega_{T,22} \delta_T' \times \left[\hat{\Omega}_{11\cdot 2}\right]^{-1} \\
&= \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) \times \left(1 + O\left(\frac{1}{T}\right)\right) + \delta_T \Omega_{22} \left(1 + O\left(\frac{1}{T}\right)\right) \delta_T' \times \left[\hat{\Omega}_{11\cdot 2}\right]^{-1} \\
&= \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) \times \left(1 + O\left(\frac{1}{T}\right)\right) + \frac{\delta_T \Omega_{22} \delta_T'}{\hat{\Omega}_{11\cdot 2}} \times \left(1 + O_p\left(\frac{1}{T}\right)\right) \\
&= \left(1 + \frac{\delta_T \Omega_{22} \delta_T'}{\Omega_{11\cdot 2}}\right) \times \left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11\cdot 2}}\right) \times \left(1 + O_p\left(\frac{1}{T}\right)\right) \\
&= \Xi \left(1 + O_p\left(\frac{1}{T}\right)\right),
\end{aligned}$$

where Ξ and Θ are independent. Hence, we have that

$$\begin{aligned}
P\left(F(\hat{\theta}_{2,\mathrm{GLS}}) \leq z\right) &= P\left(\Theta\Xi + O_p\left(\frac{1}{T}\right) \leq z\right) \\
&= P\left(\Theta\Xi \leq z\right) + O\left(\frac{1}{T}\right) \\
&= EG_1\left(\Xi^{-1}\right) + O\left(\frac{1}{T}\right),
\end{aligned}$$

which is the desired. Part ii) We want to prove that

$$P\left(F(\hat{\theta}_2^*) \leq z\right) = P\left(F(\hat{\theta}_{2,\mathrm{GLS}}^*) \leq z\right) + O\left(\frac{1}{T}\right),$$

where

$$F(\hat{\theta}_2^*) = \left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \right)^2 \text{ and } F(\hat{\theta}_{2,\text{GLS}}^*) = \left(\frac{\sqrt{T}(\hat{\theta}_{2,\text{GLS}}^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \right)^2.$$

Note that

$$\begin{aligned} P\left(F(\hat{\theta}_2^*) \leq z\right) &= P\left(\left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \right)^2 \leq z \right) \\ &= P\left(\left| \frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \right| \leq \tilde{z} \right) \\ &= P\left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq \tilde{z} \right) - P\left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq -\tilde{z} \right), \end{aligned}$$

where $\tilde{z} = \sqrt{z} > 0$. Thus, the result in Part ii) holds when we prove that

$$P\left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq \tilde{z} \right) = P\left(\frac{\sqrt{T}(\hat{\theta}_{2,\text{GLS}}^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq \tilde{z} \right) + O\left(\frac{1}{T}\right); \tag{17}$$

$$P\left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq -\tilde{z} \right) = P\left(\frac{\sqrt{T}(\hat{\theta}_{2,\text{GLS}}^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq -\tilde{z} \right) + O\left(\frac{1}{T}\right). \tag{18}$$

We only prove (17) because the proof of (18) can be done in the same way. From (7), we have that

$$\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} = \frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}} - \mu_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} + \frac{\sqrt{T}\hat{\Pi}\left(L_T' L_T\right)^{-1} L_T' \tilde{W}}{\sqrt{\hat{\Omega}_{11\cdot 2}}}.$$

The results in Lemma 7 imply that $\hat{\mu}_{\text{GLS}} - \mu_0$ is independent with $\hat{\Omega}$ and $\tilde{W}$, and thus it is independent with $\hat{\Omega}_{11\cdot 2}$, $\hat{\Pi}$, and $\tilde{W}$. Thus, we have that

$$\begin{aligned} P\left(\frac{\sqrt{T}(\hat{\theta}_2^* - \theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq \tilde{z} \right) &= P\left(\frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}} - \mu_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq \tilde{z} - \frac{\sqrt{T}\hat{\Pi}\left(L_T' L_T\right)^{-1} L_T' \tilde{W}}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \right) \\ &= E\left[P\left(\frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}} - \mu_0)}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \leq \tilde{z} - \frac{\sqrt{T}\hat{\Pi}\left(L_T' L_T\right)^{-1} L_T' \tilde{W}}{\sqrt{\hat{\Omega}_{11\cdot 2}}} \middle| \hat{\Omega}, \tilde{W} \right) \right] \\ &= E\left[P\left(\frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}} - \mu_0)}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot 2}\hat{\Pi}'}} \leq \left(\frac{\hat{\Omega}_{11\cdot 2}}{\hat{\Pi}\Omega_{T,11\cdot 2}\hat{\Pi}'} \right)^{1/2} \tilde{z} - \frac{\sqrt{T}\hat{\Pi}\left(L_T' L_T\right)^{-1} L_T' \tilde{W}}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot 2}\hat{\Pi}'}} \middle| \hat{\Omega}, \hat{\Pi} \right) \right] \end{aligned}$$

Note that $\Omega_{T,11\cdot 2}$ is a non-random constant, and we have that

$$\frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}} - \mu_0)}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot 2}\hat{\Pi}'}} \sim N(0,1);$$

$$\sqrt{T}\left(\hat{\mu}_{\text{GLS}} - \mu_0\right) \text{ is independent with } \hat{\Omega}_{11\cdot 2}, \hat{\Pi}, \text{and } \tilde{W}.$$

So, the last probability equals

$$E\left[\Psi\left(\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}-\frac{\sqrt{T}\hat{\Pi}\left(L_T'L_T\right)^{-1}L_T'\tilde{W}}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}}\right)\right],$$

where $\Psi(\cdot)$ is the cdf of standard normal distribution, and the expectation is taken with respect to $\hat{\Omega}_{11\cdot2},\hat{\Pi}$,and $\tilde{W}$. Let

$$\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}=C \text{ and } \frac{\sqrt{T}\hat{\Pi}\left(L_T'L_T\right)^{-1}L_T'\tilde{W}}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}}=X,$$

Then, using a Taylor expansion around $X=0$,

$$\begin{aligned}
&E\left[\Psi\left(\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}x-\frac{\sqrt{T}\hat{\Pi}\left(L_T'L_T\right)^{-1}L_T'\tilde{W}}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}}\right)\right]\\
&=E\left[\Psi\left(C-X\right)\right]\\
&=E\left[\Psi\left(C\right)\right]-E\left[\Psi'\left(C\right)\right]\times\left(X\right)+O(X^2)\\
&=E\left[\Psi\left(\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}\right)\right]\\
&-E\left[\varphi\left(\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}\right)\times\left(\frac{\sqrt{T}\hat{\Pi}\left(L_T'L_T\right)^{-1}L_T'\tilde{W}}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}}\right)\right]+O\left(\frac{1}{T}\right)\\
&=E\left[\Psi\left(\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}\right)\right]+O\left(\frac{1}{T}\right),
\end{aligned}$$

where $\varphi(\cdot)$ is the pdf of standard normal random variable. Here, the last equation holds because $\hat{\Omega}$ ($\hat{\Omega}_{11\cdot2}$ and $\hat{\Pi}$) is quadratic in $\tilde{W}$, as argued in Sun, Phillips and Jin (2008, pp 186). This result indicates that

$$\begin{aligned}
P\left(\frac{\sqrt{T}(\hat{\theta}_2^*-\theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot2}}}\leq\tilde{z}\right)&=E\left[\Psi\left(\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}\right)\right]+O\left(\frac{1}{T}\right)\\
&=P\left(\frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}}-\mu_0)}{\sqrt{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}}\leq\left(\frac{\hat{\Omega}_{11\cdot2}}{\hat{\Pi}\Omega_{T,11\cdot2}\hat{\Pi}'}\right)^{1/2}\tilde{z}\right)+O\left(\frac{1}{T}\right)\\
&=P\left(\frac{\sqrt{T}\hat{\Pi}(\hat{\mu}_{\text{GLS}}-\mu_0)}{\sqrt{\hat{\Omega}_{11\cdot2}}}\leq\tilde{z}\right)+O\left(\frac{1}{T}\right)\\
&=P\left(\frac{\sqrt{T}(\hat{\theta}_{2,\text{GLS}}^*-\theta_0^*)}{\sqrt{\hat{\Omega}_{11\cdot2}}}\leq\tilde{z}\right)+O\left(\frac{1}{T}\right),
\end{aligned}$$

which is the desired result. ■

6.2 Proof of Theorem 5

Before we prove Theorem 5, we first provide some technical results in the following series of Lemmas.

Lemma 8 (Beveridge and Nelson, 1981) *Under Assumption 4, the process v_t admits the following decomposition*

$$v_t = C(1)\epsilon_t + \tilde{v}_{t-1} - \tilde{v}_t$$

for $\tilde{v}_t = \sum_{j=0}^{\infty} \tilde{C}_j \epsilon_{t-j}$, $\tilde{C}_j = \sum_{s=j+1}^{\infty} C_s$, *where* $\sum_{j=0}^{\infty} ||\tilde{C}_j||^2 < \infty$

Lemma 9 *Let Assumption 4 hold. As $K \to \infty$ such that $K/T \to 0$, we have that*

i) $E[\hat{\Omega}] - \Omega = \left(\frac{K^2}{T^2}\right) B + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$;

ii) $E[(\hat{\Omega}_{11} - \Omega_{11})^2] = \frac{2\Omega_{11}^2}{K}(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$;

iii) $E\left\{(vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))(vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))'\right\} = \frac{1}{K}(\Omega_{22} \otimes \Omega_{22})\left(I_{q^2} + \mathbb{K}_{qq}\right)(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$,

where

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}' & B_{22} \end{bmatrix} := -\frac{\pi^2}{6} \begin{bmatrix} \sum_{h=-\infty}^{\infty} h^2 \Gamma_{11,u}(h) & \sum_{h=-\infty}^{\infty} h^2 \Gamma_{12,u}(h) \\ \sum_{h=-\infty}^{\infty} h^2 \Gamma_{12,u}'(h) & \sum_{h=-\infty}^{\infty} h^2 \Gamma_{22,u}(h) \end{bmatrix},$$

and $\Gamma_u(h) = E v_t v_{t-h}'$.

Proof of Lemma 9. The proofs for lemma can be proved in the same way as in Theorem 2 in Sun (2011) and Theorem 1 in Phillips (2005), and we omit the details. ■

Lemma 10 *Let Assumption 4 hold. As $K \to \infty$ such that $K/T \to 0$, we have that*

i) $E\left[(\hat{\Omega}_{11} - \Omega_{11})(\hat{\Omega}_{12} - \Omega_{12})\right] = \frac{2\Omega_{11}\Omega_{12}}{K}(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$;

ii) $E\left[(\hat{\Omega}_{21} - \Omega_{21})(\hat{\Omega}_{12} - \Omega_{12})\right] = \frac{(\Omega_{11}\Omega_{22} + \Omega_{21}\Omega_{12})}{K}(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$;

iii) $E\left[(\hat{\Omega}_{21} - \Omega_{21})\left(vec(\hat{\Omega}_{22}) - vec(\Omega_{22})\right)'\right] = \frac{1}{K}(\Omega_{22} \otimes \Omega_{12})\left(I_{q^2} + \mathbb{K}_{qq}\right)(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$;

iv) $E\left[(\hat{\Omega}_{11} - \Omega_{11})\left(vec(\hat{\Omega}_{22}) - vec(\Omega_{22})\right)'\right] = \frac{2}{K}(\Omega_{12} \otimes \Omega_{12})(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$;

Proof of Lemma 10. To save pages, we omit the proof, and it is available upon request. ■

Proof of Theorem 5.

Part i) Let

$$\hat{\Lambda} = \left(\hat{\Omega}_{11}, \hat{\Omega}_{12}', vec(\hat{\Omega}_{22})'\right)' \in \mathbb{R}^{1+q+q^2};$$
$$\Lambda = \left(\Omega_{11}, \Omega_{12}', vec(\Omega_{22})'\right)' \in \mathbb{R}^{1+q+q^2}.$$

Note that

$$\Xi^{-1}\left(\hat{\Lambda}\right) := \Xi^{-1}\left(\left[\hat{\Omega}_{11}, \hat{\Omega}_{12}, vec(\hat{\Omega}_{22})'\right]'\right)$$

where

$$\Xi\left(\left[\hat{\Omega}_{11},\hat{\Omega}_{12},vec(\hat{\Omega}_{22})'\right]'\right)=\left[\left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\hat{\Omega}_{21}}\right)+\left(\frac{\left(\Omega_{12}\Omega_{22}^{-1}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)\Omega_{22}\left(\Omega_{12}\Omega_{22}^{-1}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\right)'}{\Omega_{11\cdot 2}}\right)\times\left(\frac{\Omega_{11\cdot 2}}{\hat{\Omega}_{11}-\hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\hat{\Omega}_{21}}\right)\right].$$

We want to first derive a Taylor expansion of $\Xi^{-1}(\hat{\Lambda})$ at Λ up to the second order, i.e.

$$\Xi^{-1}(\hat{\Lambda})\approx\Xi^{-1}(\Lambda)+\underbrace{\nabla'\Xi^{-1}(\hat{\Lambda})\Big|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda)}_{:=L}+\underbrace{\frac{1}{2}(\hat{\Lambda}-\Lambda)'\left\{\nabla^2\Xi^{-1}(\hat{\Lambda})\Big|_{\hat{\Lambda}=\Lambda}\right\}(\hat{\Lambda}-\Lambda)}_{:=Q}. \tag{19}$$

Here, we have

$$\begin{aligned}\Xi^{-1}(\Lambda)&=\left[\left(\frac{\Omega_{11\cdot 2}}{\Omega_{11\cdot 2}}\right)+\left(\frac{\left(\Omega_{12}\Omega_{22}^{-1}-\Omega_{12}\Omega_{22}^{-1}\right)\Omega_{22}\left(\Omega_{12}\Omega_{22}^{-1}-\Omega_{12}\Omega_{22}^{-1}\right)'}{\Omega_{11\cdot 2}}\right)\left(\frac{\Omega_{11\cdot 2}}{\Omega_{11\cdot 2}}\right)\right]^{-1}\\&=1.\end{aligned}$$

We compute the first derivative term evaluated at Λ in (19). With a considerable amount of algebra, which we omit to save pages, we can obtain that

$$\begin{aligned}L=&\ \nabla'\Xi^{-1}(\hat{\Lambda})\Big|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda)\\=&\ \frac{1}{\Omega_{11\cdot 2}}\times\left[1,-2\left(\Omega_{21}'\Omega_{22}^{-1}\right),\left(\Omega_{21}'\Omega_{22}^{-1}\otimes\Omega_{21}'\Omega_{22}^{-1}\right)\right]\\&\times\left(\hat{\Omega}_{11}-\Omega_{11},\hat{\Omega}_{21}'-\Omega_{21}',vec(\hat{\Omega}_{22})'-vec(\Omega_{22})'\right)'.\end{aligned}$$

With a considerable amount of algebra, which we omit to save pages, we can obtain that

$$\begin{aligned}Q=&-\frac{2\left(\hat{\Omega}_{12}-\Omega_{12}\right)\Omega_{22}^{-1}\left(\hat{\Omega}_{21}-\Omega_{21}\right)}{\Omega_{11\cdot 2}}\\&+\frac{4\left(vec(\hat{\Omega}_{22})'-vec(\Omega_{22})'\right)\left(\Omega_{22}^{-1}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left(\hat{\Omega}_{21}-\Omega_{21}\right)}{\Omega_{11\cdot 2}}\\&-\left(vec(\hat{\Omega}_{22})'-vec(\Omega_{22})'\right)\times\left\{\frac{\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)\left\{\left(I_q\otimes\Omega_{12}\Omega_{22}^{-1}\right)+\left(\Omega_{12}\Omega_{22}^{-1}\otimes I_q\right)\right\}}{\Omega_{11\cdot 2}}\right\}\\&\times\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)\\:=&\ Q_1+Q_2+Q_3,\end{aligned}$$

Also, we have that

$$\begin{aligned}
L &= \nabla'\Xi^{-1}(\hat{\Lambda})\Big|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda) \\
&= \frac{1}{\Omega_{11\cdot 2}} \times \left[1, -2\left(\Omega_{21}'\Omega_{22}^{-1}\right), \left(\Omega_{21}'\Omega_{22}^{-1}\otimes\Omega_{21}'\Omega_{22}^{-1}\right)\right] \\
&\times \left(\hat{\Omega}_{11}-\Omega_{11}, \hat{\Omega}_{21}'-\Omega_{21}', vec(\hat{\Omega}_{22})'-vec(\Omega_{22})'\right)' \\
&= \frac{(\hat{\Omega}_{11}-\Omega_{11})}{\Omega_{11\cdot 2}} - \frac{2\Omega_{12}\Omega_{22}^{-1}(\hat{\Omega}_{21}-\Omega_{21})}{\Omega_{11\cdot 2}} + \frac{\left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)}{\Omega_{11\cdot 2}}.
\end{aligned}$$

Now, we need to compute

$$\begin{aligned}
E[L+Q] &= E[L]+E[Q] \\
E[(L+Q)^2] &= E[L^2] + \text{small orders.}
\end{aligned}$$

First, we have that

$$\begin{aligned}
E[L] &= \frac{E[(\hat{\Omega}_{11}-\Omega_{11})]}{\Omega_{11\cdot 2}} - \frac{2\Omega_{12}\Omega_{22}^{-1}E[(\hat{\Omega}_{21}-\Omega_{21})]}{\Omega_{11\cdot 2}} \\
&\quad + \frac{\left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)E\left[\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)\right]}{\Omega_{11\cdot 2}}.
\end{aligned}$$

From Lemma 9, we have that

$$\begin{aligned}
E[\hat{\Omega}_{11}-\Omega_{11}] &= \frac{K^2}{T^2}B_{11} + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
E[\hat{\Omega}_{21}-\Omega_{21}] &= \frac{K^2}{T^2}B_{21} + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
E\left\{vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right\} &= \frac{K^2}{T^2}vec(B_{22}) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right),
\end{aligned}$$

So that, we have that

$$\begin{aligned}
E[L] &= \frac{K^2}{T^2}\frac{B_{11}}{\Omega_{11\cdot 2}} - \frac{K^2}{T^2}\times\left(\frac{2\Omega_{12}\Omega_{22}^{-1}B_{21}}{\Omega_{11\cdot 2}}\right) + \frac{K^2}{T^2}\frac{\left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)vec(B_{22})}{\Omega_{11\cdot 2}} + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= \frac{K^2}{T^2}\frac{1}{\Omega_{11\cdot 2}}\left(B_{11} - 2\Omega_{12}\Omega_{22}^{-1}B_{21} + \left(\Omega_{12}\Omega_{22}^{-1}\otimes\Omega_{12}\Omega_{22}^{-1}\right)vec(B_{22})\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&:= \frac{K^2}{T^2}\tilde{B} + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)
\end{aligned}$$

and

$$\begin{aligned}
vec(B_{22})'\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right) &= \left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\Omega_{21}\right)'vec(B_{22}) \\
&= \left((\Omega_{22}^{-1}\Omega_{21})'\otimes\Omega_{12}\Omega_{22}^{-1}\right)vec(B_{22}) \\
&= vec(\Omega_{12}\Omega_{22}^{-1}B_{22}\Omega_{22}^{-1}\Omega_{21}) \\
&= \mathrm{tr}(\Omega_{12}\Omega_{22}^{-1}B_{22}\Omega_{22}^{-1}\Omega_{21}) \\
&= \mathrm{tr}(B_{22}\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\Omega_{22}^{-1}),
\end{aligned}$$

which leads to

$$E[L] = \frac{K^2}{T^2}\frac{1}{\Omega_{11\cdot 2}}\left(B_{11} - 2\Omega_{12}\Omega_{22}^{-1}B_{21} + \text{tr}(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\Omega_{22}^{-1}B_{22})\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).$$

We now express $E[Q]$. Recall that

$$\begin{aligned}
Q = & -\frac{2\left(\hat{\Omega}_{12} - \Omega_{12}\right)\Omega_{22}^{-1}\left(\hat{\Omega}_{21} - \Omega_{21}\right)}{\Omega_{11\cdot 2}} \\
& + \frac{4\left(vec(\hat{\Omega}_{22})' - vec(\Omega_{22})'\right)\left(\Omega_{22}^{-1} \otimes \Omega_{22}^{-1}\Omega_{21}\right)\left(\hat{\Omega}_{21} - \Omega_{21}\right)}{\Omega_{11\cdot 2}} \\
& - \left(vec(\hat{\Omega}_{22})' - vec(\Omega_{22})'\right) \times \left\{\frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right)\left\{\left(I_q \otimes \Omega_{12}\Omega_{22}^{-1}\right) + \left(\Omega_{12}\Omega_{22}^{-1} \otimes I_q\right)\right\}}{\Omega_{11\cdot 2}}\right\} \\
& \times \left(vec(\hat{\Omega}_{22}) - vec(\Omega_{22})\right) \\
& := Q_1 + Q_2 + Q_3,
\end{aligned}$$

So that

$$\begin{aligned}
E[Q] = & -\frac{2E\left[\left(\hat{\Omega}_{12} - \Omega_{12}\right)\Omega_{22}^{-1}\left(\hat{\Omega}_{21} - \Omega_{21}\right)\right]}{\Omega_{11\cdot 2}} \\
& + \frac{4E\left[\left(vec(\hat{\Omega}_{22})' - vec(\Omega_{22})'\right)\left(\Omega_{22}^{-1} \otimes \Omega_{22}^{-1}\Omega_{21}\right)\left(\hat{\Omega}_{21} - \Omega_{21}\right)\right]}{\Omega_{11\cdot 2}} \\
& - E\left[\left\{vec(\hat{\Omega}_{22})' - vec(\Omega_{22})'\right\}\underbrace{\frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right)\left\{\left(I_q \otimes \Omega_{12}\Omega_{22}^{-1}\right) + \left(\Omega_{12}\Omega_{22}^{-1} \otimes I_q\right)\right\}}{\Omega_{11\cdot 2}}}_{:=H}\right. \\
& \left. \times \left\{vec(\hat{\Omega}_{22}) - vec(\Omega_{22})\right\}\right] \\
& := E[Q_1] + E[Q_2] + E[Q_3],
\end{aligned}$$

where

$$\begin{aligned}
E[Q_1] &= -\text{tr}2E\left[\left(\hat{\Omega}_{12} - \Omega_{12}\right)\frac{\Omega_{22}^{-1}}{\Omega_{11\cdot 2}}\left(\hat{\Omega}_{21} - \Omega_{21}\right)\right] \\
&= -\frac{2\text{tr}\left[\Omega_{22}^{-1}E\left[\left(\hat{\Omega}_{21} - \Omega_{21}\right)\left(\hat{\Omega}_{12} - \Omega_{12}\right)\right]\right]}{\Omega_{11\cdot 2}},
\end{aligned}$$

and

$$\begin{aligned}
E[Q_2] &= \text{tr}4E\left[\left(vec(\hat{\Omega}_{22})' - vec(\Omega_{22})'\right)\frac{\left(\Omega_{22}^{-1} \otimes \Omega_{22}^{-1}\Omega_{21}\right)}{\Omega_{11\cdot 2}}(\hat{\Omega}_{21} - \Omega_{21})\right] \\
&= 4\text{tr}\left[\frac{\left(\Omega_{22}^{-1} \otimes \Omega_{22}^{-1}\Omega_{21}\right)E\left[\left(\hat{\Omega}_{21} - \Omega_{21}\right)\left(vec(\hat{\Omega}_{22}) - vec(\Omega_{22})\right)'\right]}{\Omega_{11\cdot 2}}\right],
\end{aligned}$$

and

$$E[Q_3] = -\text{tr}\left[HE\left\{(vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))(vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))'\right\}\right].$$

By Lemma 9, we can express the last term as

$$\begin{aligned}
&E\left\{(vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))(vec(\hat{\Omega}_{22}) - vec(\Omega_{22}))'\right\} \\
&= \frac{1}{K}(\Omega_{22} \otimes \Omega_{22})\left(I_{q^2} + \mathbb{K}_{qq}\right)(1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).
\end{aligned}$$

So that

$$\begin{aligned}
E[Q_3] &= -\frac{1}{K}\text{tr}\left[H(\Omega_{22} \otimes \Omega_{22})\left(I_{q^2} + \mathbb{K}_{qq}\right)\right] + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= -\frac{1}{K}\text{tr}\left[\tilde{H}\left(I_{q^2} + \mathbb{K}_{qq}\right)\right](1 + o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)
\end{aligned}$$

where

$$\begin{aligned}
\tilde{H} &= H(\Omega_{22} \otimes \Omega_{22}) \\
&= \frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right)\left\{\left(I_q \otimes \Omega_{12}\Omega_{22}^{-1}\right) + \left(\Omega_{12}\Omega_{22}^{-1} \otimes I_q\right)\right\}}{\Omega_{11\cdot 2}} \times (\Omega_{22} \otimes \Omega_{22}) \\
&= \frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right) \times \{(\Omega_{22} \otimes \Omega_{12}) + (\Omega_{12} \otimes \Omega_{22})\}}{\Omega_{11\cdot 2}}.
\end{aligned}$$

$$\begin{aligned}
\text{tr}\left[\tilde{H}\left(I_{q^2} + \mathbb{K}_{qq}\right)\right] &= \text{tr}(\tilde{H}) + \text{tr}(\tilde{H}\mathbb{K}_{qq}) \\
&= \text{tr}(\tilde{H}) + \text{tr}(\mathbb{K}_{qq}\tilde{H}),
\end{aligned}$$

and

$$\begin{aligned}
\text{tr}(\tilde{H}) &= \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right) \times \{(\Omega_{22} \otimes \Omega_{12}) + (\Omega_{12} \otimes \Omega_{22})\}}{\Omega_{11\cdot 2}}\right] \\
&= \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right)(\Omega_{22} \otimes \Omega_{12})}{\Omega_{11\cdot 2}}\right] + \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right)(\Omega_{12} \otimes \Omega_{22})}{\Omega_{11\cdot 2}}\right] \\
&= \text{tr}\left[\frac{(\Omega_{22} \otimes \Omega_{12})\left(\Omega_{22}^{-1}\Omega_{21} \otimes \Omega_{22}^{-1}\right)}{\Omega_{11\cdot 2}}\right] + \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12} \otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&= \text{tr}\left[\frac{(\Omega_{21} \otimes \Omega_{12}\Omega_{22}^{-1})}{\Omega_{11\cdot 2}}\right] + \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12} \otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&= \text{tr}\left[\frac{(\Omega_{21}\Omega_{12}\Omega_{22}^{-1})}{\Omega_{11\cdot 2}}\right] + \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12} \otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&= \frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}} + \text{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12} \otimes I_q\right)}{\Omega_{11\cdot 2}}\right],
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{tr}(\mathbb{K}_{qq}\tilde{H}) &= \operatorname{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)\times\left\{(\Omega_{22}\otimes\Omega_{12})+(\Omega_{12}\otimes\Omega_{22})\right\}}{\Omega_{11\cdot 2}}\right] \\
&= \operatorname{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)(\Omega_{22}\otimes\Omega_{12})}{\Omega_{11\cdot 2}}\right]+\operatorname{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)(\Omega_{12}\otimes\Omega_{22})}{\Omega_{11\cdot 2}}\right] \\
&= \operatorname{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)(\Omega_{22}\otimes\Omega_{12})}{\Omega_{11\cdot 2}}\mathbb{K}_{qq}\right]+\operatorname{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&= \operatorname{tr}\left[\frac{(\Omega_{22}\otimes\Omega_{12})}{\Omega_{11\cdot 2}}\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)\right]+\operatorname{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)}{\Omega_{11\cdot 2}}\right]
\end{aligned}$$

By Magnus and Neudecker (1988), we have that

$$\mathbb{K}_{mn}(A\otimes B)\mathbb{K}_{st}=(B\otimes A)$$

for $A\in\mathbb{R}^{n\times s}$ and $B\in\mathbb{R}^{m\times t}$. Using this property, we obtain that

$$\begin{aligned}
I_q(\Omega_{22}\otimes\Omega_{12})\mathbb{K}_{qq} &= \mathbb{K}_{1q}(\Omega_{22}\otimes\Omega_{12})\mathbb{K}_{qq} \\
&= (\Omega_{12}\otimes\Omega_{22}),
\end{aligned}$$

since $\mathbb{K}_{q1}=\mathbb{K}_{1q}=I_q$ by Theorem 3.1 in Magnus and Neudecker (1988) . Thus, we have that

$$\begin{aligned}
&\operatorname{tr}\left[\frac{(\Omega_{22}\otimes\Omega_{12})}{\Omega_{11\cdot 2}}\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)\right] \\
&= \operatorname{tr}\left[\frac{(\Omega_{12}\otimes\Omega_{22})\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)}{\Omega_{11\cdot 2}}\right] \\
&= \operatorname{tr}\frac{(\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\otimes I_q)}{\Omega_{11\cdot 2}} \\
&= \operatorname{tr}\left[\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\times I_q}{\Omega_{11\cdot 2}}\right]=\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}q
\end{aligned}$$

which implies that

$$\begin{aligned}
\operatorname{tr}(\mathbb{K}_{qq}\tilde{H}) &= \operatorname{tr}\left[\frac{(\Omega_{22}\otimes\Omega_{12})}{\Omega_{11\cdot 2}}\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\otimes\Omega_{22}^{-1}\right)\right]+\operatorname{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&= \operatorname{tr}\frac{(\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\otimes I_q)}{\Omega_{11\cdot 2}}+\operatorname{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)\mathbb{K}_{qq}}{\Omega_{11\cdot 2}}\right].
\end{aligned}$$

Then,

$$\begin{aligned}
\mathrm{tr}\left[\tilde{H}\left(I_{q^2}+\mathbb{K}_{qq}\right)\right] &= \mathrm{tr}(\tilde{H})+\mathrm{tr}(\mathbb{K}_{qq}\tilde{H}) \\
&= \frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}+\mathrm{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&+\mathrm{tr}\frac{\left(\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\otimes I_q\right)}{\Omega_{11\cdot 2}}+\mathrm{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)\mathbb{K}_{qq}}{\Omega_{11\cdot 2}}\right] \\
&= \frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}(1+q)+\mathrm{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)\left(I_{q^2}+\mathbb{K}_{qq}\right)}{\Omega_{11\cdot 2}}\right] \\
&= \mathrm{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)}{\Omega_{11\cdot 2}}\right]+\mathrm{tr}\left[\frac{\mathbb{K}_{qq}\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)}{\Omega_{11\cdot 2}}\right] \\
&= 2(1+q)\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}},
\end{aligned}$$

where the last equation follows from Theorem 3.1 in Magnus and Neudecker (1979). Thus

$$\begin{aligned}
E[Q_3] &= -\frac{1}{K}\mathrm{tr}\left[\tilde{H}\left(I_{q^2}+\mathbb{K}_{qq}\right)\right](1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= -\frac{(1+q)}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}(1+o(1)) \\
&- \frac{1}{K}\mathrm{tr}\left[\frac{\left(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\otimes I_q\right)\left(I_{q^2}+\mathbb{K}_{qq}\right)}{\Omega_{11\cdot 2}}\right](1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= -\frac{2(1+q)}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}(1+o(1))+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right).
\end{aligned}$$

Now we express $E[Q_1]$ and $E[Q_2]$. Recall that

$$\begin{aligned}
E[Q_1] &= -\mathrm{tr}2E\left[\left(\hat{\Omega}_{12}-\Omega_{12}\right)\frac{\Omega_{22}^{-1}}{\Omega_{11\cdot 2}}\left(\hat{\Omega}_{21}-\Omega_{21}\right)\right] \\
&= -\frac{2\mathrm{tr}\left[\Omega_{22}^{-1}E\left[\left(\hat{\Omega}_{21}-\Omega_{21}\right)\left(\hat{\Omega}_{12}-\Omega_{12}\right)\right]\right]}{\Omega_{11\cdot 2}},
\end{aligned}$$

and

$$\begin{aligned}
E[Q_2] &= \mathrm{tr}4E\left[\left(vec(\hat{\Omega}_{22})'-vec(\Omega_{22})'\right)\frac{\left(\Omega_{22}^{-1}\otimes\Omega_{22}^{-1}\Omega_{21}\right)}{\Omega_{11\cdot 2}}(\hat{\Omega}_{21}-\Omega_{21})\right] \\
&= 4\mathrm{tr}\left[\frac{\left(\Omega_{22}^{-1}\otimes\Omega_{22}^{-1}\Omega_{21}\right)E\left[\left(\hat{\Omega}_{21}-\Omega_{21}\right)\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)'\right]}{\Omega_{11\cdot 2}}\right],
\end{aligned}$$

From Lemma 9,

$$E\left[\left(\hat{\Omega}_{21}-\Omega_{21}\right)\left(\hat{\Omega}_{12}-\Omega_{12}\right)\right]=\frac{\left(\Omega_{11}\Omega_{22}+\Omega_{21}\Omega_{12}\right)}{K}\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right);$$

$$E\left[\left(\hat{\Omega}_{21}-\Omega_{21}\right)\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)'\right]=\frac{1}{K}\left(\Omega_{22}\otimes\Omega_{12}\right)\left(I_{q^2}+\mathbb{K}_{qq}\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right).$$

So we have

$$\begin{aligned}
E[Q_1] &= -\frac{2\text{tr}\left[\Omega_{22}^{-1}E\left[\left(\hat{\Omega}_{21}-\Omega_{21}\right)\left(\hat{\Omega}_{12}-\Omega_{12}\right)\right]\right]}{\Omega_{11\cdot 2}} \\
&= -\frac{2\text{tr}\left[\Omega_{22}^{-1}\left[\frac{\left(\Omega_{11}\Omega_{22}+\Omega_{21}\Omega_{12}\right)}{K}\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)\right]\right]}{\Omega_{11\cdot 2}} \\
&= -\frac{2}{K}\frac{\text{tr}\left[\Omega_{22}^{-1}\left[\left(\Omega_{11}\Omega_{22}+\Omega_{21}\Omega_{12}\right)\right]\right]}{\Omega_{11\cdot 2}}\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= -\frac{2}{K}\frac{\text{tr}\left[\Omega_{22}^{-1}\Omega_{11}\Omega_{22}\right]}{\Omega_{11\cdot 2}}\left(1+o(1)\right)-\frac{2}{K}\frac{\text{tr}\left[\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\right]}{\Omega_{11\cdot 2}}\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= -\frac{2}{K}\frac{\text{tr}[\Omega_{11}I_q]}{\Omega_{11\cdot 2}}\left(1+o(1)\right)-\frac{2}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= -\frac{2}{K}\frac{\left(q\Omega_{11}+\Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right)}{\Omega_{11\cdot 2}}\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right),
\end{aligned}$$

and

$$\begin{aligned}
E[Q_2] &= 4\text{tr}\left[\frac{\left(\Omega_{22}^{-1}\otimes\Omega_{22}^{-1}\Omega_{21}\right)E\left[\left(\hat{\Omega}_{21}-\Omega_{21}\right)\left(vec(\hat{\Omega}_{22})-vec(\Omega_{22})\right)'\right]}{\Omega_{11\cdot 2}}\right] \\
&= 4\text{tr}\left[\frac{\left(\Omega_{22}^{-1}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left(\frac{1}{K}\left(\Omega_{22}\otimes\Omega_{12}\right)\left(I_{q^2}+\mathbb{K}_{qq}\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right)\right)}{\Omega_{11\cdot 2}}\right] \\
&= \frac{4}{K}\text{tr}\left[\frac{\left(\Omega_{22}^{-1}\otimes\Omega_{22}^{-1}\Omega_{21}\right)\left(\Omega_{22}\otimes\Omega_{12}\right)\left(I_{q^2}+\mathbb{K}_{qq}\right)}{\Omega_{11\cdot 2}}\right]\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= \frac{4}{K}\text{tr}\left[\frac{\left(I_q\otimes\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\right)\left(I_{q^2}+\mathbb{K}_{qq}\right)}{\Omega_{11\cdot 2}}\right]\left(1+o(1)\right)+o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right) \\
&= \frac{4}{K}\text{tr}\left[\frac{\left(I_q\otimes\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\right)}{\Omega_{11\cdot 2}}\right]\left(1+o(1)\right)+\frac{4}{K}\text{tr}\left[\frac{\left(I_q\otimes\Omega_{22}^{-1}\Omega_{21}\Omega_{12}\right)\mathbb{K}_{qq}}{\Omega_{11\cdot 2}}\right]\left(1+o(1)\right) \\
&\quad +o\left(\frac{K^2}{T^2}\right)+O\left(\frac{1}{T}\right).
\end{aligned}$$

By Theorem 3.1 - (xiii) in Magnus and Neudeker (1979), we have that

$$\text{tr}(\mathbb{K}_{mn}(A'\otimes B))=\text{tr}(A'B),$$

for $A, B \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$.Using this result, we have that

$$\begin{aligned}
\mathrm{tr}\left[\frac{(I_q \otimes \Omega_{22}^{-1}\Omega_{21}\Omega_{12})\mathbb{K}_{qq}}{\Omega_{11\cdot 2}}\right] &= \mathrm{tr}\left[\frac{\mathbb{K}_{qq}(I_q \otimes \Omega_{22}^{-1}\Omega_{21}\Omega_{12})}{\Omega_{11\cdot 2}}\right] \\
&= \frac{1}{\Omega_{11\cdot 2}}\mathrm{tr}\left[I_q \times \Omega_{22}^{-1}\Omega_{21}\Omega_{12})\right] \\
&= \frac{1}{\Omega_{11\cdot 2}}\mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}\Omega_{12}) \\
&= \frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}.
\end{aligned}$$

Also, we have that

$$\begin{aligned}
\mathrm{tr}\left[\frac{(I_q \otimes \Omega_{22}^{-1}\Omega_{21}\Omega_{12})}{\Omega_{11\cdot 2}}\right] &= \frac{\mathrm{tr}(I_q)\mathrm{tr}(\Omega_{22}^{-1}\Omega_{21}\Omega_{12})}{\Omega_{11\cdot 2}} \\
&= q\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}.
\end{aligned}$$

Thus, we have that

$$\begin{aligned}
E[Q_2] &= \frac{4}{K}\mathrm{tr}\left[\frac{(I_q \otimes \Omega_{22}^{-1}\Omega_{21}\Omega_{12})}{\Omega_{11\cdot 2}}\right](1+o(1)) + \frac{4}{K}\mathrm{tr}\left[\frac{(I_q \otimes \Omega_{22}^{-1}\Omega_{21}\Omega_{12})\mathbb{K}_{qq}}{\Omega_{11\cdot 2}}\right](1+o(1)) \\
&\quad + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right); \\
&= \frac{4}{K}\left(\frac{q\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}} + \frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}\right)(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right); \\
&= \frac{4(q+1)}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).
\end{aligned}$$

Combining these results, we have that

$$\begin{aligned}
E[Q] &= E[Q_1] + E[Q_2] + E[Q_3] \\
&= -\frac{2}{K}\frac{(q\Omega_{11} + \Omega_{12}\Omega_{22}^{-1}\Omega_{21})}{\Omega_{11\cdot 2}}(1+o(1)) + \frac{4(q+1)}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}(1+o(1)) \\
&\quad - \frac{2(1+q)}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}} + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= \left\{-\frac{2}{K}\frac{(q\Omega_{11} + \Omega_{12}\Omega_{22}^{-1}\Omega_{21})}{\Omega_{11\cdot 2}} + \frac{2(q+1)}{K}\frac{\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}\right\}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= -\frac{2}{K}\left\{\frac{(q\Omega_{11} + \Omega_{12}\Omega_{22}^{-1}\Omega_{21}) - (q+1)\Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}\right\}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= -\frac{2}{K}\left\{\frac{(q(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}) + \Omega_{12}\Omega_{22}^{-1}\Omega_{21} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}}{\Omega_{11\cdot 2}}\right\}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\
&= -\frac{2q}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).
\end{aligned}$$

We have that

$$E[L^2] = \frac{2\Omega_{11\cdot 2}^2}{\Omega_{11\cdot 2}^2}\frac{1}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right)$$
$$= \frac{2}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right).$$

Summing up the results so far, we have that

$$\begin{aligned} L &= \nabla'\Xi^{-1}(\hat{\Lambda})\Big|_{\hat{\Lambda}=\Lambda}(\hat{\Lambda}-\Lambda); \\ Q &= \frac{1}{2}(\hat{\Lambda}-\Lambda)'\left\{\nabla^2\Xi^{-1}(\hat{\Lambda})\Big|_{\hat{\Lambda}=\Lambda}\right\}(\hat{\Lambda}-\Lambda), \end{aligned}$$

and

$$\begin{aligned} E[L] &= \frac{K^2}{T^2}\tilde{B} + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right); \\ E[Q] &= -\frac{2q}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right), \end{aligned}$$

and

$$\begin{aligned} E[L^2] &= \frac{2}{K}(1+o(1)) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right); \\ E[Q^2] &= o\left(\frac{1}{K}+\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right), \end{aligned}$$

and thus

$$\Xi^{-1}\left(\hat{\Lambda}\right) = 1 + L + Q + o_p\left(\frac{1}{K}+\frac{K^2}{T^2}\right) + O_p\left(\frac{1}{T}\right).$$

Using this results, we have that

$$P\left(\left|t(\hat{\theta}_2)\right| \le z\right) = EG_1\left(z\Xi^{-1}\left(\hat{\Lambda}\right)\right) + O\left(\frac{1}{T}\right)$$
$$= EG_1\left(z\left(1+L+Q\right)\right) + o\left(\frac{1}{K}+\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right),$$

and we use a Talyor expansion of $EG_1\left(z\left(1+L+Q\right)\right)$ around z,

$$\begin{aligned} &EG_1\left(z\left(1+L+Q\right)\right) + o\left(\frac{1}{K}+\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\ &= G_1(z) + G_1'(z)zE[(z\left(1+L+Q\right)-z)] + \frac{1}{2}G_1''(z)E[(z\left(1+L+Q\right)-z)^2] \\ &+ o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right), \\ &= G_1(z) + G_1'(z)zE[L] + G_1'(z)zE[Q] \\ &+ \frac{1}{2}G_1''(z)E[z^2(L+Q)^2] + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right) \\ &= G_1(z) + G_1'(z)zE[L] + G_1'(z)zE[Q] \\ &+ \frac{1}{2}G_1''(z)z^2E[L^2] + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right), \end{aligned}$$

and this leads to the desired result

$$\begin{aligned} P\left(\left|t(\hat{\theta}_2^*)\right| \leq z\right) &= G_1(z) + G_1'(z)z\left(\frac{K^2}{T^2}\right)\tilde{B} - G_1'(z)z\left(\frac{2q}{K}\right) \\ &\quad + G_1''(z)z^2 \cdot \left(\frac{1}{K}\right) + o\left(\frac{1}{K}\right) + o\left(\frac{K^2}{T^2}\right) + O\left(\frac{1}{T}\right), \end{aligned}$$

which completes the proof. ■