MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

An asymptotic expansion of Wishart distribution when the population eigenvalues are infinitely dispersed

Yo Sheena and Akimichi Takemura

METR 2003–22 June 2003

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

An asymptotic expansion of Wishart distribution when the population eigenvalues are infinitely dispersed

Yo Sheena
Department of Mathematics and Statistics
Bowling Green State University
(On leave from Department of Economics, Shinshu University, Japan)
and
Akimichi Takemura
Department of Mathematical Informatics
University of Tokyo

Abstract
Takemura and Sheena (2002) derived the asymptotic joint distribution of the eigenvalues and the eigenvectors of Wishart matrix when the population eigenvalues become infinitely dispersed. They also showed necessary conditions for an estimator of the population covariance matrix to be minimax for typical loss functions by calculating the asymptotic risk of the estimator. In this paper, we further examine those distributions and risks by means of an asymptotic expansion. We focus on a limiting process where the population eigenvalues become linearly dispersed, which can be parametrized by one parameter. We obtain the asymptotic expansion of the distribution function of relevant elements of the sample eigenvalues and eigenvectors with respect to the parameter. We also derive the asymptotic expansion of the risk function of a scale and orthogonally equivariant estimator. As an application, we prove non-minimaxity of Stein's and Haff's estimators, which has been an open problem so far.

Keywords and phrases
asymptotic distribution, covariance matrix, minimax estimator, orthogonally equivariant, scale equivariant, Stein's loss, tail minimaxity.

1 Introduction and summary of results

Let $\boldsymbol{W} = (w_{ij})$ be distributed according to Wishart distribution $\boldsymbol{W}_p(n, \boldsymbol{\Sigma})$, where p is the dimension, n is the degrees of freedom and $\boldsymbol{\Sigma}$ is the covariance matrix. In this paper we consider asymptotics where the population eigenvalues become infinitely dispersed. Denote the spectral decompositions of $\boldsymbol{W}$ and $\boldsymbol{\Sigma}$ by

$$\boldsymbol{W} = \boldsymbol{G}\boldsymbol{L}\boldsymbol{G}', \qquad \boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}', \tag{1}$$

where

$$\boldsymbol{G},\ \boldsymbol{\Gamma} \in \mathcal{O}(p) = \{\boldsymbol{G} \mid \boldsymbol{G} : p \times p \text{ orthogonal matrix}\}$$

and

$$\boldsymbol{L} = \mathrm{diag}(l_1, \ldots, l_p),\ \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$$

are diagonal matrices with the eigenvalues $l_1 \geq \cdots \geq l_p > 0,\ \lambda_1 \geq \cdots \geq \lambda_p > 0$ of $\boldsymbol{W}$ and $\boldsymbol{\Sigma}$, respectively. We use the notations $\boldsymbol{l} = (l_1, \ldots, l_p)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$ hereafter. For the uniqueness of the spectral decomposition of $\boldsymbol{W}$, having chosen one decomposition of $\boldsymbol{\Sigma}$, we decide the sign of the columns of $\boldsymbol{G}$ so that $\boldsymbol{\Gamma}'\boldsymbol{G}$ (say $\widetilde{\boldsymbol{G}} = (\tilde{g}_{ij})$) has nonnegative diagonal elements. Then the uniqueness of $\boldsymbol{G}$ is guaranteed almost surely and $\widetilde{\boldsymbol{G}} \in \mathcal{O}^+(p)$, where

$$\mathcal{O}^+(p) = \{\boldsymbol{G} \in \mathcal{O}(p) \mid (\boldsymbol{G})_{ii} \geq 0,\ i = 1, \ldots, p\}.$$

Let

$$\boldsymbol{y} = (y_1, \ldots, y_{p-1}), \qquad y_j = \frac{\lambda_{j+1}}{\lambda_j}, \quad j = 1, \ldots, p-1.$$

We say that the population eigenvalues become *infinitely dispersed* when

$$\|\boldsymbol{y}\| \to 0. \tag{2}$$

Takemura and Sheena (2002) investigated the asymptotic distribution of the sample eigenvalues $(l_1, \ldots, l_p)$ and the sample eigenvectors $\boldsymbol{G}$ under the limiting process (2). They proved that $\widetilde{\boldsymbol{G}}$ converges to $\boldsymbol{I}_p$ in probability, i.e.; for any neighborhood $\mathcal{N}(\boldsymbol{I}_p)$ of $\boldsymbol{I}_p$,

$$P(\widetilde{\boldsymbol{G}} \in \mathcal{N}(\boldsymbol{I}_p)) \to 1, \qquad \text{as } \|\boldsymbol{y}\| \to 0.$$

They also proved that $\boldsymbol{f} = (f_i)_{1 \leq i \leq p}$ and $\boldsymbol{q} = (q_{ij})_{1 \leq j < i \leq p}$ defined by

$$\begin{aligned} f_i &= \frac{l_i}{\lambda_i}, \qquad 1 \leq i \leq p, \\ q_{ij} &= \tilde{g}_{ij} l_j^{\frac{1}{2}} \lambda_i^{-\frac{1}{2}} = \tilde{g}_{ij} f_j^{\frac{1}{2}} \lambda_j^{\frac{1}{2}} \lambda_i^{-\frac{1}{2}}, \qquad 1 \leq j < i \leq p. \end{aligned}$$

have the following asymptotic distributions;

$$f_i \sim \chi^2_{n-i+1}, \quad 1 \leq i \leq p, \qquad q_{ij} \sim N(0,1), \quad 1 \leq j < i \leq p, \tag{3}$$

and f_i $(1 \leq i \leq p)$, q_{ij} $(1 \leq j < i \leq p)$ are asymptotically mutually independently distributed. Notice that the asymptotic distributions do not depend on the specific form of the limiting process (2).

In this paper, we investigate these asymptotic distributions in more detail. For this purpose, we need to specify the form of the limiting process (2). We consider *linear convergence* of $\boldsymbol{y}$ defined by

$$y_j = \frac{\lambda_{j+1}}{\lambda_j} = a_j z, \quad a_j > 0, \quad z > 0, \quad 1 \leq j \leq p-1, \tag{4}$$

where a_j's are fixed and $z \to 0$. This linear convergence is basic and would be useful in investigating the behavior of $\boldsymbol{f}$ and $\boldsymbol{q}$ for other types of limiting processes. We derive asymptotic expansions of distribution functions and risk functions to the order $O(z)$. We say that a term of order $O(z)$ in our expansions is the first order term in z, although in usual large sample asymptotic expansions with respect to n, the terms of order $O(n^{-1/2})$ (or $O(n^{-1})$ depending on contexts) are referred to as the second order terms. In Section 3, we derive the asymptotic expansion of $P(\widetilde{\boldsymbol{G}} \in \mathcal{N}(\boldsymbol{I}_p))$ in Theorem 1 and the asymptotic expansion of the joint distribution function of $\boldsymbol{f}$ and $\boldsymbol{q}$ in Theorem 2 up to the first order in z. In particular in Theorem 1 we prove that $O(z)$ term vanishes and $P(\widetilde{\boldsymbol{G}} \in \mathcal{N}(\boldsymbol{I}_p)) = 1 + o(z)$. This guarantees that as far as the terms of order $O(z)$ are concerned we can concentrate on an arbitrary small neighborhood $\mathcal{N}(\boldsymbol{I}_p)$ of $\boldsymbol{I}_p$.

A recurring idea of the proof is that under (4) only the pairs of adjacent eigenvalues $(\lambda_j, \lambda_{j+1})$ contribute to the term of order $O(z)$ and the terms involving (λ_i, λ_j), $j \geq i+2$, are of the higher order.

In Section 4, we consider the estimation problem of $\boldsymbol{\Sigma}$ from decision theoretic point of view. Takemura and Sheena (2002) examined the asymptotic risk of estimators with respect to Stein's loss function

$$L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \mathrm{tr}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}) - \log|\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}| - p$$

and the quadratic loss function

$$L_2(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \mathrm{tr}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)^2.$$

They derived a necessary condition for an estimator to be *tail minimax*. We call an estimator $\widehat{\boldsymbol{\Sigma}}$ tail minimax with respect to L_i if it satisfies the condition

$$\exists \delta > 0, \quad \forall \|\boldsymbol{y}\| < \delta, \qquad E[L_i(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] \leq \bar{R}_i,$$

where $\bar{R}_i$ is the minimax risk for L_i, $i = 1, 2$. See Berger (1976) for the notion of tail minimaxity in the estimation of a location vector in a general multivariate location family. Obviously a minimax estimator is tail minimax.

Let $\widehat{\boldsymbol{\Sigma}} = \widehat{\boldsymbol{\Sigma}}(\boldsymbol{W}) = \widehat{\boldsymbol{\Sigma}}(\boldsymbol{G}, \boldsymbol{l})$ be an estimator of $\boldsymbol{\Sigma}$ and let

$$\widehat{\boldsymbol{\Sigma}}(\boldsymbol{G}, \boldsymbol{l}) = \boldsymbol{H}(\boldsymbol{G}, \boldsymbol{l})\boldsymbol{D}(\boldsymbol{G}, \boldsymbol{l})\boldsymbol{H}'(\boldsymbol{G}, \boldsymbol{l}) \tag{5}$$

be the spectral decomposition of $\widehat{\boldsymbol{\Sigma}}(\boldsymbol{G}, \boldsymbol{l})$, where $\boldsymbol{H}(\boldsymbol{G}, \boldsymbol{l}) \in \mathcal{O}(p)$ and

$$\boldsymbol{D}(\boldsymbol{G}, \boldsymbol{l}) = \mathrm{diag}(d_1(\boldsymbol{G}, \boldsymbol{l}), \ldots, d_p(\boldsymbol{G}, \boldsymbol{l})).$$

In accordance with the definition of $\boldsymbol{G}$, the sign of $\boldsymbol{H}(\boldsymbol{G},\boldsymbol{l})$ is determined by $(\boldsymbol{\Gamma}'\boldsymbol{H})_{ii} \geq 0,\ 1 \leq \forall i \leq p$. Hence $\boldsymbol{H} \in \mathcal{O}^+(p)$. Let

$$c_i(\boldsymbol{G},\boldsymbol{l}) = \frac{d_i(\boldsymbol{G},\boldsymbol{l})}{l_i}, \qquad i = 1,\dots,p.$$

An estimator of the form

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{G}\boldsymbol{\Psi}(\boldsymbol{L})\boldsymbol{G}', \qquad \boldsymbol{\Psi}(\boldsymbol{L}) = \mathrm{diag}(\psi_1(\boldsymbol{l}),\dots,\psi_p(\boldsymbol{l})). \tag{6}$$

is called *orthogonally equivariant*. For orthogonally equivariant estimators we have

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{G},\boldsymbol{l}) &= \boldsymbol{G}, \\ c_i(\boldsymbol{G},\boldsymbol{l}) &= c_i(\boldsymbol{l}) = \frac{\psi_i(\boldsymbol{l})}{l_i}, \qquad 1 \leq i \leq p. \end{aligned}$$

Let $\widehat{\boldsymbol{\Sigma}}^{JS}$ and $\widehat{\boldsymbol{\Sigma}}^{OS}$ denote the best triangularly equivariant estimators for L_1 and L_2, respectively. They are constant risk minimax estimators with respect to the corresponding loss functions.

Simple but important orthogonally equivariant estimators are those which have certain constant c_i's; $\widehat{\boldsymbol{\Sigma}}^{SDS}$ defined by $c_i = \delta_i^{JS}$ and $\widehat{\boldsymbol{\Sigma}}^{KG}$ defined by $c_i = \delta_i^{OS}$, where δ_i^{JS} and δ_i^{OS} are coefficients which respectively appear in $\widehat{\boldsymbol{\Sigma}}^{JS}$ and $\widehat{\boldsymbol{\Sigma}}^{OS}$. Actually

$$\delta_i^{JS} = \frac{1}{n+p+1-2i}, \qquad 1 \leq i \leq p, \tag{7}$$

and δ_i^{OS} $(1 \leq i \leq p)$ are given as the solution of a linear equation. $\widehat{\boldsymbol{\Sigma}}^{SDS}(\widehat{\boldsymbol{\Sigma}}^{KG})$ dominates $\widehat{\boldsymbol{\Sigma}}^{JS}(\widehat{\boldsymbol{\Sigma}}^{OS})$ with respect to L_1 (L_2), hence is a minimax estimator. For more details see Section 3 of Takemura and Sheena (2002). Original literature related to these two estimators include James and Stein (1961), Olkin and Selliah (1977), Sharma and Krishnamoorthy (1983), Dey and Srinivasan (1985), Krishnamoorthy and Gupta (1989) and Sheena (2002).

Let

$$\boldsymbol{x} = (x_1,\dots,x_{p-1}), \qquad x_j = \frac{l_{j+1}}{l_j}, \qquad 1 \leq j \leq p-1.$$

Takemura and Sheena (2002) showed that under some regularity conditions, any tail minimax estimator with respect to L_1 (L_2) converges to $\widehat{\boldsymbol{\Sigma}}^{SDS}$ $(\widehat{\boldsymbol{\Sigma}}^{KG})$ as $\|\boldsymbol{x}\| \to 0$. More specifically, for any tail minimax estimator, $\widehat{\boldsymbol{\Sigma}}(\boldsymbol{G},\boldsymbol{l})$, with the spectral decomposition (5)

$$\boldsymbol{H}(\boldsymbol{G},\boldsymbol{l}) \to \boldsymbol{G}, \tag{8}$$

$$c_i(\boldsymbol{G},\boldsymbol{l}) \to \delta_i^{JS(OS)}, \quad 1 \leq i \leq p, \tag{9}$$

as $\|\boldsymbol{x}\| \to 0$.

In Section 4 of this paper, we consider a scale and orthogonally equivariant estimator and its asymptotic risk with respect to Stein's loss. Note that any orthogonally equivariant estimator trivially satisfies (8). Therefore (9) is an important necessary condition for tail minimaxity of orthogonally equivariant estimators. We focus ourselves on orthogonally equivariant estimators which satisfy (9). We derive the expansion of their risks with respect to z up to the first order in Theorem 4. This result provides a second step test that orthogonally equivariant estimators must satisfy in order to be tail minimax after they pass the first test (9).

We apply the result to two scale and orthogonally equivariant estimators; Stein's estimator, $\widehat{\boldsymbol{\Sigma}}^S$, and Haff's estimator, $\widehat{\boldsymbol{\Sigma}}^H$. By an approximate minimization of the unbiased estimator of the risk function, Stein (1975) proposed an orthogonally equivariant estimator, which is given by (6) with

$$\psi_i(\boldsymbol{l}) = l_i \left\{ (n-p+1) + 2l_i \sum_{j \neq i} \frac{1}{l_i - l_j} \right\}^{-1}, \qquad i = 1, \ldots, p. \tag{10}$$

This estimator is called Stein's rough estimator. Since this estimator sometimes violates the condition

$$\psi_1 \geq \psi_2 \geq \cdots \geq \psi_p \geq 0, \tag{11}$$

isotonizing modification of the eigenvalues is carried out. We use the notation $\widehat{\boldsymbol{\Sigma}}^S$ for this modified estimator. See Sheena and Takemura (1992) for decision theoretic results on isotonizing the eigenvalues.

Haff (1991) studied the form of Bayes estimators and observed that a slight modification of $\widehat{\boldsymbol{\Sigma}}^S$ emerges as an approximation of the Bayes rule. We use the notation $\widehat{\boldsymbol{\Sigma}}^H$ for this estimator.

It is not easy to give an explicit form of these estimators for general p. (See Sugiura and Ishibayashi (1997) for their explicit forms when $p = 2$.) The calculation of their risks is also too complicated to handle. Therefore it has not been theoretically proved whether these estimators are minimax or not. According to Monte Carlo simulation carried out by Sugiura and Ishibayashi (1997), though they substantially outperform not only $\widehat{\boldsymbol{\Sigma}}^{JS}$ but also $\widehat{\boldsymbol{\Sigma}}^{SDS}$ when $\boldsymbol{\Sigma}$ is close to $\boldsymbol{I}_p$ (see also Lin and Perlman (1985)), their risks are over that of $\widehat{\boldsymbol{\Sigma}}^{JS}$ when $\|\boldsymbol{y}\|$ is small, that is, when the population eigenvalues are dispersed. Since the minimax estimator $\widehat{\boldsymbol{\Sigma}}^{JS}$ has a constant risk, this simulation result indicates that $\widehat{\boldsymbol{\Sigma}}^S, \widehat{\boldsymbol{\Sigma}}^H$ are not minimax.

Using the expansion of the risks of these estimators with respect to z, we prove in Corollary 5 that $\widehat{\boldsymbol{\Sigma}}^S$ and $\widehat{\boldsymbol{\Sigma}}^H$ are not minimax estimators.

The rest of the paper is organized as follows. In Section 2 we summarize some preliminary results from Takemura and Sheena (2002). In Section 3, after proving some relevant lemmas and Theorem 1, we derive asymptotic expansion of the distribution function of the eigenvalues and the eigenvectors in Theorem 2. In Section 4 we derive asymptotic expansion of risk functions of scale and orthogonally equivariant estimators. Theorem 4 is the main theorem of Section 4. As a corollary to this theorem, non-minimaxity of $\widehat{\boldsymbol{\Sigma}}^S, \widehat{\boldsymbol{\Sigma}}^H$ is proved in Corollary 5.

2 Preliminaries

We state some definitions and notations used in the subsequent sections. Most of them are in accordance with those in Takemura and Sheena (2002). We also briefly mention some related properties without proof. See Takemura and Sheena (2002) for the proof.

Most of the calculations hereafter involves the integral of a function of $\boldsymbol{G}$ on a neighborhood of $\boldsymbol{I}_p \in \mathcal{O}^+(p)$ with respect to the invariant probability measure $\mu(d\boldsymbol{G})$ on $\mathcal{O}(p)$. For the expansion with respect to z, it is convenient to specify the neighborhood. If we choose a small enough neighborhood of $\boldsymbol{I}_p$, we can use the lower triangular part of $\boldsymbol{G}$ as its coordinate; there exists a diffeomorphism, $\boldsymbol{G}(\boldsymbol{u}) = (g_{ij}(\boldsymbol{u}))$, from

$$\bar{U}^* = \bar{U}^*(\epsilon) = \left\{ \boldsymbol{u} = (u_{ij})_{1\le j<i\le p} \in R^{p'} \,\middle|\, \|\boldsymbol{u}\| = \left(\sum_{i>j} u_{ij}^2\right)^{\frac{1}{2}} \le \epsilon \right\}, \quad p' = \frac{p(p-1)}{2},$$

onto a neighborhood of $\boldsymbol{I}_p \in \mathcal{O}^+(p)$ (say $\boldsymbol{G}(\bar{U}^*)$) such that $\boldsymbol{G}(\boldsymbol{0}) = \boldsymbol{I}_p$, $\boldsymbol{0} = (0, \dots, 0)$. Furthermore real analytic functions $g_{ij}(\boldsymbol{u})$ $(1 \le i, j \le p)$ on $\bar{U}^*$ have the following expansions around the point $\boldsymbol{u} = \boldsymbol{0}$.

$$g_{ij}(\boldsymbol{u}) = \begin{cases} -u_{ji} - \sum_{k=1}^{i-1} u_{ik}u_{jk} + \sum_{k=i+1}^{j-1} u_{ki}u_{jk} - \sum_{k=j+1}^{p} u_{ki}u_{kj} + R_{ij} & \text{if } j > i, \\ 1 - \frac{1}{2}\sum_{k=1}^{i-1} u_{ik}^2 - \frac{1}{2}\sum_{k=i+1}^{p} u_{ki}^2 + R_{ii} & \text{if } j = i, \\ u_{ij} & \text{if } j < i, \end{cases} \tag{12}$$

where R_{ij} $(1 \le i \le j \le p)$ involves only higher order terms in $\boldsymbol{u}$ than 2. We choose ϵ so that

$$g_{ii}(\boldsymbol{u})^2 \ge 1/2, \quad i = 1, \dots, p. \tag{13}$$

In many integrals in the following sections, $\boldsymbol{G}(\bar{U}^*)$ or $\bar{U}^*$ will appear as the domain of the integral.

The measure on $\bar{U}^*$ induced from the invariant probability $\mu(d\boldsymbol{G})$ on $\mathcal{O}(p)$ has a density with respect to Lebesgue measure in $R^{p'}$. This is given as follows. Let p' pairs, (i, j), $1 \le j < i \le p$, be ordered by the lexicographical order; $(i_1, j_1) < (i_2, j_2)$ if and only if "$j_1 < j_2$" or "$j_1 = j_2$, $i_1 < i_2$". Consider 1-forms $\boldsymbol{g}_i' d\boldsymbol{g}_j = g_{1i}dg_{1j} + \cdots + g_{pi}dg_{pj}$, $1 \le j < i \le p$. Let $\boldsymbol{v}$ be the vector with these 1-forms as its elements lexicographically ordered, that is;

$$\boldsymbol{v} = (\boldsymbol{g}_2' d\boldsymbol{g}_1, \boldsymbol{g}_3' d\boldsymbol{g}_1, \dots, \boldsymbol{g}_p' d\boldsymbol{g}_1, \boldsymbol{g}_3' d\boldsymbol{g}_2, \dots, \boldsymbol{g}_p' d\boldsymbol{g}_{p-1})'.$$

Similarly define the vector $d\boldsymbol{u}$ as

$$d\boldsymbol{u} = (du_{21}, du_{31}, \dots, du_{p1}, du_{32}, \dots, du_{p,p-1})'.$$

We can express $\boldsymbol{v}$ as

$$\boldsymbol{v} = \Xi d\boldsymbol{u}, \qquad \Xi = (\xi_{st}(\boldsymbol{u}))_{1\le s,t\le p'},$$

where $\xi_{st}(\boldsymbol{u})$'s are real analytic functions of $\boldsymbol{u}$ on $\bar{U}^*$. Let $J^*(\boldsymbol{u}) = |\det \Xi|$. Then $\bar{c}J^*(\boldsymbol{u})$, with a normalizing constant $\bar{c}$, gives the density of the measure induced from $\mu(d\boldsymbol{G})$ on $\bar{U}^*$ with respect to Lebesgue measure in $R^{p'}$. Note that $J^*(\boldsymbol{u})$ is also a real analytic function on $\bar{U}^*$. For the purpose of the present paper we do not need an explicit expression of the normalizing constant $\bar{c}$.

Now consider an arbitrary function $x(\boldsymbol{G}, \boldsymbol{l}, \boldsymbol{\lambda})$ of $\boldsymbol{G}, \boldsymbol{l}$ and $\boldsymbol{\lambda}$. We give a formula on

$$E[I(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*))x(\boldsymbol{G}, \boldsymbol{l}, \boldsymbol{\lambda})],$$

where $I(\cdot)$ is the indicator function. Hereafter two notations for the indicator function

$$I(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*)), \quad I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})$$

will be used interchangeably.

We define a compound function, $x_{\mathbf{\Gamma}}(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$, as

$$x_{\mathbf{\Gamma}}(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) = x(\mathbf{\Gamma}\boldsymbol{G}(\boldsymbol{u}(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})),\boldsymbol{l}(\boldsymbol{f},\boldsymbol{\lambda}),\boldsymbol{\lambda}),$$

where

$$\begin{aligned} \boldsymbol{u}(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) &= (u_{ij}(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}))_{1\le j<i\le p} = (q_{ij}f_j^{-\frac{1}{2}}\lambda_j^{-\frac{1}{2}}\lambda_i^{\frac{1}{2}})_{1\le j<i\le p}, \\ \boldsymbol{l}(\boldsymbol{f},\boldsymbol{\lambda}) &= (l_1(\boldsymbol{f},\boldsymbol{\lambda}),\cdots,l_p(\boldsymbol{f},\boldsymbol{\lambda})) = (f_1\lambda_1,\ldots,f_p\lambda_p). \end{aligned} \tag{14}$$

Then

$$E[I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})x(\boldsymbol{G},\boldsymbol{l},\boldsymbol{\lambda})] = \int_{R^{p'}}\int_{R_+^p} x_{\mathbf{\Gamma}}(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})d\boldsymbol{f}d\boldsymbol{q}, \tag{15}$$

where $R_+ = (0,\ \infty)$ and

$$h(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) = h_0(\boldsymbol{f},\boldsymbol{q})h_1(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h_2(\boldsymbol{f},\boldsymbol{\lambda})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$$

with

$$\begin{aligned} h_0(\boldsymbol{f},\boldsymbol{q}) &= \tilde{c}\left(\prod_{i=1}^p f_i^{\frac{n-i-1}{2}}\right)\exp\left(-\frac{1}{2}\sum_{i>j}q_{ij}^2\right)\exp\left(-\frac{1}{2}\sum_{i=1}^p f_i\right), \\ h_1(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) &= \exp\left\{-\frac{1}{2}\left(\sum_{i=1}^p(g_{ii}^2(\boldsymbol{u})-1)f_i + \sum_{i<j}g_{ij}^2(\boldsymbol{u})f_j\frac{\lambda_j}{\lambda_i}\right)\right\}, && (16) \\ h_2(\boldsymbol{f},\boldsymbol{\lambda}) &= I\left(f_1\lambda_1>\cdots>f_p\lambda_p\right)\prod_{j<i}\left(1-\frac{f_i\lambda_i}{f_j\lambda_j}\right), && (17) \\ h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) &= I_{\bar{U}^*}(\boldsymbol{u})J^*(\boldsymbol{u}). && (18) \end{aligned}$$

The constant $\tilde{c}$ is the normalizing constant so that $h_0(\boldsymbol{f},\boldsymbol{q})$ is the density function of the asymptotic distribution of $(\boldsymbol{f},\boldsymbol{q})$ given by (3). We also often use the fact $h_2(\boldsymbol{f},\boldsymbol{q}), h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$ are bounded functions and the fact $h(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$ is dominated by the function

$$\bar{h}(\boldsymbol{f},\boldsymbol{q}) = K\left(\prod_{i=1}^p f_i^{\frac{n-i-1}{2}}\right)\exp\left(-\frac{1}{2}\sum_{i>j}q_{ij}^2\right)\exp\left(-\frac{1}{4}\sum_{i=1}^p f_i\right) \tag{19}$$

with some constant K.

We define $a_0 = a_p = 0$ for unified description of statements including a_i's.

3 Asymptotic expansion of eigenvalues and eigenvectors

We first prove some lemmas which are useful for the expansion w.r.t. z of the integrals of the type in (15). The following lemma evaluates the behavior of $J^*(\boldsymbol{u})$ around the origin.

Lemma 1 *On $\bar{U}^* = \bar{U}^*(\epsilon)$ with a small enough ϵ, $J^*(\boldsymbol{u})$ is given by*

$$J^*(\boldsymbol{u}) = 1 + \frac{1}{2}\sum_{i>j}(i-j)u_{ij}^2 + \sum_{(i,j)\in\mathcal{P}} J(\boldsymbol{u};(i,j))u_{i_1j_1}u_{i_2j_2}u_{i_3j_3}, \tag{20}$$

where $(i,j) = \{(i_1,j_1),(i_2,j_2),(i_3,j_3)\}$, $\mathcal{P} = \{(i,j) \mid 1 \le j_s < i_s \le p,\ s = 1,2,3\}$ and $J(\boldsymbol{u};(i,j))$, $(i,j) \in \mathcal{P}$ are bounded functions on $\bar{U}^*$.

Proof.

In the following argument, we use the notation $o(\|\boldsymbol{u}\|^r)$, $r = 1, 2$, as a power series in $\boldsymbol{u}$ which consists only of the terms $\prod_{s>t} u_{st}^{n_{st}}$ such that $\sum_{s>t} n_{st} \ge r+1$.

Differentiating (12), dg_{ij} is given by

$$dg_{ij}(\boldsymbol{u}) = \begin{cases} -du_{ji} - \sum_{k=1}^{i-1}(u_{jk}du_{ik} + u_{ik}du_{jk}) + \sum_{k=i+1}^{j-1}(u_{ki}du_{jk} + u_{jk}du_{ki}) \\ \qquad - \sum_{k=j+1}^{p}(u_{kj}du_{ki} + u_{ki}du_{kj}) + \boldsymbol{r}_{ij}d\boldsymbol{u} & \text{if } j > i, \\ -\sum_{k=1}^{i-1} u_{ik}du_{ik} - \sum_{k=i+1}^{p} u_{ki}du_{ki} + \boldsymbol{r}_{ii}d\boldsymbol{u} & \text{if } j = i, \\ du_{ij} & \text{if } j < i, \end{cases}$$

where for $1 \le i \le j \le p$,

$$\boldsymbol{r}_{ij}\, d\boldsymbol{u} = \sum_{l>m} \frac{\partial R_{ij}}{\partial u_{lm}}\, du_{lm}, \qquad \frac{\partial R_{ij}}{\partial u_{lm}} = o(\|\boldsymbol{u}\|).$$

Then for $1 \le j < i \le p$, we have

$$\boldsymbol{g}_i' d\boldsymbol{g}_j = \sum_{l<j} T_{1l} + T_2 + \sum_{j<l<i} T_{3l} + T_4 + \sum_{i<l} T_{5l} + \widetilde{R}_{ij}(d\boldsymbol{u}), \tag{21}$$

where

$$\begin{aligned} T_{1l} \;&=\; \left(u_{il} + \sum_{k=1}^{l-1} u_{lk}u_{ik} - \sum_{k=l+1}^{i-1} u_{kl}u_{ik} + \sum_{k=i+1}^{p} u_{kl}u_{ki}\right) du_{jl} \\ &\quad + \sum_{k=1}^{l-1}(u_{jk}u_{il}du_{lk} + u_{lk}u_{il}du_{jk}) - \sum_{k=l+1}^{j-1}(u_{kl}u_{il}du_{jk} + u_{jk}u_{il}du_{kl}) \\ &\quad + \sum_{k=j+1}^{p}(u_{kj}u_{il}du_{kl} + u_{kl}u_{il}du_{kj}), \\ T_2 \;&=\; \sum_{k=1}^{j-1} u_{jk}u_{ij}du_{jk} + \sum_{k=j+1}^{p} u_{kj}u_{ij}du_{kj}, \\ T_{3l} \;&=\; \left(-u_{il} - \sum_{k=1}^{l-1} u_{lk}u_{ik} + \sum_{k=l+1}^{i-1} u_{kl}u_{ik} - \sum_{k=i+1}^{p} u_{kl}u_{ki}\right) du_{lj}, \end{aligned}$$

$$
\begin{aligned}
T_4 &= \left(1 - \frac{1}{2}\sum_{k=1}^{i-1} u_{ik}^2 - \frac{1}{2}\sum_{k=i+1}^{p} u_{ki}^2\right) du_{ij},\\
T_{5l} &= u_{li} du_{lj},\\
\widetilde{R}_{ij}(d\boldsymbol{u}) &= \sum_{l>m} \tau_{lm}^{(ij)}(\boldsymbol{u}) du_{lm}, \qquad \tau_{lm}^{(ij)}(\boldsymbol{u}) = o(\|\boldsymbol{u}\|^2).
\end{aligned}
$$

From this expression, we notice that off-diagonal elements of Ξ do not contain a constant and the diagonal element $\xi_{tt},\ 1 \le t \le p' = p(p-1)/2$, is of the form

$$
\xi_{tt} = 1 + \frac{1}{2}\left(\sum_{l=1}^{j} u_{il}^2 - \sum_{l=j+1}^{i-1} u_{il}^2 - \sum_{l=i+1}^{p} u_{li}^2\right) + o(\|\boldsymbol{u}\|^2),
$$

where $t = (i, j)$ is the tth element in the lexicographical order.

Consider

$$
\det \Xi = \sum_{\sigma} \operatorname{sign}(\boldsymbol{\sigma})\, \xi_{1,\sigma(1)} \cdots \xi_{p',\sigma(p')},
$$

where $\boldsymbol{\sigma} = (\sigma(1), \ldots, \sigma(p'))$ is a permutation of $(1, \ldots, p')$, and $\operatorname{sign}(\boldsymbol{\sigma}) = 1$ if $\boldsymbol{\sigma}$ is an even permutation; -1 if $\boldsymbol{\sigma}$ is an odd permutation. Note that in the right side, the terms of order at most two in $\boldsymbol{u}$ appear only when $\boldsymbol{\sigma}$ is the identity permutation, $\sigma(t) = t,\ 1 \le t \le p'$ with $\operatorname{sign}(\boldsymbol{\sigma}) = 1$, or when $\boldsymbol{\sigma}$ is a transposition, i.e. for some $s < t$,

$$
\begin{cases}
\sigma(s) &= t,\\
\sigma(t) &= s,\\
\sigma(w) &= w, \quad w \ne s, t.
\end{cases}
$$

with $\operatorname{sign}(\boldsymbol{\sigma}) = -1$. Consequently $\det \Xi$, a real analytic function on $\bar{U}^*$ can be expanded around $\mathbf{0}$ as

$$
\det \Xi = 1 + \frac{1}{2}\sum_{i>j}\left(\sum_{l=1}^{j} u_{il}^2 - \sum_{l=j+1}^{i-1} u_{il}^2 - \sum_{l=i+1}^{p} u_{li}^2\right) - \sum_{s<t} \xi_{st}\xi_{ts} + o(\|\boldsymbol{u}\|^2). \tag{22}
$$

Now we examine the term $\sum_{s<t} \xi_{st}\xi_{ts}$. Suppose that $s = (i_1, j_1)$ and $t = (i_2, j_2)$ are respectively the sth and tth in the lexicographical order. Then ξ_{st} is the coefficient of $du_{i_2 j_2}$ in the expression of $\boldsymbol{g}_{i_1}' d\boldsymbol{g}_{j_1}$. Note that "$j_1 < j_2,\ i_1 > j_1,\ i_2 > j_2$" or "$j_1 = j_2,\ i_1 < i_2,\ i_1 > j_1,\ i_2 > j_2$". From (21), we have

$$
\boldsymbol{g}_{i_1}' d\boldsymbol{g}_{j_1} = du_{i_1 j_1} + \sum_{l=1}^{j_1-1} u_{i_1 l} du_{j_1 l} - \sum_{l=j_1+1}^{i_1-1} u_{i_1 l} du_{l j_1} + \sum_{l=i_1+1}^{p} u_{l i_1} du_{l j_1} + \sum_{m>k} \tau_{mk}^{(i_1 j_1)}(\boldsymbol{u}) du_{mk},
$$

where $\tau_{mk}^{(i_1 j_1)}(\boldsymbol{u}) = o(\|\boldsymbol{u}\|)$. Therefore we have

$$
\xi_{st} = \begin{cases} u_{i_2 i_1} + o(\|\boldsymbol{u}\|) & \text{if } j_1 = j_2,\\ o(\|\boldsymbol{u}\|) & \text{otherwise.} \end{cases}
$$

Similarly we have

$$\xi_{ts} = \begin{cases} -u_{i_2 i_1} + o(\|\boldsymbol{u}\|) & \text{if } j_1 = j_2, \\ u_{i_2 j_1} + o(\|\boldsymbol{u}\|) & \text{if } i_1 = j_2, \\ o(\|\boldsymbol{u}\|) & \text{otherwise.} \end{cases}$$

Therefore

$$\sum_{s<t} \xi_{st}\xi_{ts} = -\sum_{j_1=j_2<i_1<i_2} u_{i_2 i_1}^2 + o(\|\boldsymbol{u}\|^2) = -\sum_{i_1<i_2} (i_1-1)u_{i_2 i_1}^2 + o(\|\boldsymbol{u}\|^2). \tag{23}$$

From (22) and (23),

$$\begin{aligned} \det \Xi &= 1 + \frac{1}{2}\sum_{i>j}\left(\sum_{l=1}^{j} u_{il}^2 - \sum_{l=j+1}^{i-1} u_{il}^2 - \sum_{l=i+1}^{p} u_{li}^2\right) + \sum_{j_1=j_2<i_1<i_2} u_{i_2 i_1}^2 + o(\|\boldsymbol{u}\|^2) \\ &= 1 + \frac{1}{2}\sum_{i>j\geq l} u_{il}^2 - \frac{1}{2}\sum_{i>l>j} u_{il}^2 - \frac{1}{2}\sum_{l>i>j} u_{li}^2 + \sum_{i>l>j_1=j_2} u_{il}^2 + o(\|\boldsymbol{u}\|^2) \\ &= 1 + \frac{1}{2}\sum_{i>l}(i-l)u_{il}^2 - \frac{1}{2}\sum_{i>l}(l-1)u_{il}^2 - \frac{1}{2}\sum_{l>i}(i-1)u_{li}^2 + \sum_{i>l}(l-1)u_{il}^2 + o(\|\boldsymbol{u}\|^2) \\ &= 1 + \frac{1}{2}\sum_{i>l}(i-l)u_{il}^2 + o(\|\boldsymbol{u}\|^2). \end{aligned}$$

Since $J^*(\boldsymbol{u}) = |\det \Xi|$, if we choose $\bar{U}^* = \bar{U}^*(\epsilon)$ with a small enough ϵ, we have

$$J^*(\boldsymbol{u}) = 1 + \frac{1}{2}\sum_{i>l}(i-l)u_{il}^2 + o(\|\boldsymbol{u}\|^2). \tag{24}$$

On the other hand, Taylor expansion of $J^*(\boldsymbol{u})$ around $\boldsymbol{0}$ is given by

$$\begin{aligned} J^*(\boldsymbol{u}) &= J^*(\boldsymbol{0}) + \sum_{i>j}\frac{\partial J^*}{\partial u_{ij}}(\boldsymbol{0})\, u_{ij} + \frac{1}{2}\sum_{i_1>j_1, i_2>j_2}\frac{\partial^2 J^*}{\partial u_{i_1 j_1}\partial u_{i_2 j_2}}(\boldsymbol{0})\, u_{i_1 j_1}u_{i_2 j_2} \\ &\quad + \frac{1}{3!}\sum_{(i,j)\in\mathcal{P}}\frac{\partial^3 J^*}{\partial u_{i_1 j_1}\partial u_{i_2 j_2}\partial u_{i_3 j_3}}(\theta(\boldsymbol{u})\boldsymbol{u})\, u_{i_1 j_1}u_{i_2 j_2}u_{i_3 j_3}, \end{aligned} \tag{25}$$

where $\theta(\boldsymbol{u})$ is a function of $\boldsymbol{u}$ such that $0 < \theta(\boldsymbol{u}) < 1$. Note that

$$\frac{1}{3!}\frac{\partial^3 J^*}{\partial u_{i_1 j_1}\partial u_{i_2 j_2}\partial u_{i_3 j_3}}(\theta(\boldsymbol{u})\boldsymbol{u})$$

is bounded on the compact set $\bar{U}^*$. Using the notation $J(\boldsymbol{u};(i,j))$ for this function and comparing (24) with (25), we have the result. ∎

For the next two lemmas, we define $h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})$ as

$$h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^{p} f_i^{\alpha_i} \prod_{i>j} q_{ij}^{\beta_{ij}} \exp\left(-a\sum_{i=1}^{p} f_i - b\sum_{i>j} q_{ij}\right), \tag{26}$$

where $a > 0,\ b > 0$ and

$$\begin{aligned}\boldsymbol{\alpha} &= (\alpha_1, \ldots, \alpha_p),\\ \boldsymbol{\beta} &= (\beta_{21}, \beta_{31}, \ldots, \beta_{p1}, \beta_{32}, \ldots, \beta_{pp-1}), \qquad \beta_{ij} = 0, 1, \ldots\ (1 \le j < i \le p).\end{aligned}$$

Let $\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})$ denote an arbitrary bounded function, that is,

$$\exists M > 0, \qquad |\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})| < M.$$

Using Lemma 1 we can approximate integrals involving $h_3(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})$ in (18) in the following lemma.

Lemma 2 *Suppose that $\alpha_i > 1,\ i = 1, \ldots, p$. Then*

$$\begin{aligned}&\int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, h_3(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}\\ &\quad = \int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \Big(1 + \frac{1}{2} \sum_{i>j} (i-j) u_{ij}^2\Big)\, d\boldsymbol{f} d\boldsymbol{q} + o(z).\end{aligned}$$

Proof.

From (14) and (4),

$$\begin{aligned}&\int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \Big(\sum_{(i,j)\in\mathcal{P}} J(\boldsymbol{u}; (i,j)) \prod_{s=1}^{3} u_{i_s j_s}\Big)\, I_{\bar{U}^*}(\boldsymbol{u})\, d\boldsymbol{f} d\boldsymbol{q}\\ &\quad = \int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\\ &\qquad\qquad \times \Big(\sum_{(i,j)\in\mathcal{P}} J(\boldsymbol{u}; (i,j)) \prod_{s=1}^{3} q_{i_s j_s} f_{j_s}^{-\frac{1}{2}} \big(\prod_{m=j_s}^{i_s-1} a_m^{\frac{1}{2}}\big) z^{(i_s - j_s)/2}\Big) I_{\bar{U}^*}(\boldsymbol{u}) d\boldsymbol{f} d\boldsymbol{q}.\end{aligned}$$

Since $\alpha_i > 1,\ i = 1, \ldots, p$, and $\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}),\ J(\boldsymbol{u}; (i,j)) I_{\bar{U}^*}(\boldsymbol{u})$ $((i,j) \in \mathcal{P})$ are all bounded, the integral of the right side converges. Taking into account that $\sum_{s=1}^{3} (i_s - j_s)/2 \ge 3/2$, we notice the right side, hence the left side equals $o(z)$. Therefore it suffices to show

$$\begin{aligned}&\int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \Big(1 + \frac{1}{2} \sum_{i>j} (i-j) u_{ij}^2\Big)\, I_{\bar{U}^*}(\boldsymbol{u}) d\boldsymbol{f} d\boldsymbol{q}\\ &\quad = \int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \Big(1 + \frac{1}{2} \sum_{i>j} (i-j) u_{ij}^2\Big)\, d\boldsymbol{f} d\boldsymbol{q} + o(z).\end{aligned}$$

Since $I_{\bar{U}^*}(\boldsymbol{u}) = 1 - I_{\bar{U}^{*C}}(\boldsymbol{u})$, the left side equals $I_1 - I_2$, where

$$\begin{aligned}I_1 &= \int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \Big(1 + \frac{1}{2} \sum_{i>j} (i-j) u_{ij}^2\Big) d\boldsymbol{f} d\boldsymbol{q},\\ I_2 &= \int_{R^{p'}} \int_{R^p_+} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\\ &\qquad \times \Big(1 + \frac{1}{2} \sum_{i>j} (i-j) u_{ij}^2\Big) I\Big(\sum_{i>j} u_{ij}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) > \epsilon^2\Big) d\boldsymbol{f} d\boldsymbol{q}.\end{aligned}$$

We will prove $I_2 = o(z)$. The following inequalities hold. Note that every integral converges since $\alpha_i > 1,\ i = 1, \ldots, p$.

$$\begin{aligned} I_2 &\le \epsilon^{-2} \sum_{s>t} \int_{R^{p'}} \int_{R_+^p} u_{st}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) |\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})| \\ &\qquad \times \Big(1 + \frac{1}{2} \sum_{i>j} (i-j) u_{ij}^2\Big) I\Big(\sum_{i>j} u_{ij}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) > \epsilon^2\Big) d\boldsymbol{f} d\boldsymbol{q} \\ &= \epsilon^{-2} \sum_{s>t} z^{s-t} \Big(\prod_{m=t}^{s-1} a_m\Big) \int_{R^{p'}} \int_{R_+^p} f_t^{-1} q_{st}^2 |\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})| \\ &\qquad \times \Big\{1 + \frac{1}{2} \sum_{i>j} (i-j) q_{ij}^2 f_j^{-1} \Big(\prod_{m=j}^{i-1} a_m\Big) z^{i-j}\Big\} I\Big(\sum_{i>j} u_{ij}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) > \epsilon^2\Big) d\boldsymbol{f} d\boldsymbol{q} \\ &= \epsilon^{-2} \sum_{t=1}^{p-1} z a_t \int_{R^{p'}} \int_{R_+^p} f_t^{-1} q_{t+1,t}^2 |\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})| I\Big(\sum_{i>j} u_{ij}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) > \epsilon^2\Big) d\boldsymbol{f} d\boldsymbol{q} \\ &\qquad + o(z). \end{aligned}$$

It now remains to prove

$$\lim_{z \to 0} \int_{R^{p'}} \int_{R_+^p} f_t^{-1} q_{t+1,t}^2 |\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})| I\Big(\sum_{i>j} u_{ij}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) > \epsilon^2\Big) d\boldsymbol{f} d\boldsymbol{q} = 0.$$

However this is obvious from the dominated convergence theorem and the fact

$$\begin{aligned} \lim_{z \to 0} I\Big(\sum_{i>j} u_{ij}^2(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) > \epsilon^2\Big) &= \lim_{z \to 0} I\Big(\sum_{i>j} q_{ij}^2 f_j^{-1} \Big(\prod_{m=j}^{i-1} a_m\Big) z^{i-j} > \epsilon^2\Big) \\ &= 0. \end{aligned}$$

∎

We need another lemma on the approximation of integrals involving $h_2(\boldsymbol{f}, \boldsymbol{\lambda})$ in (17). We use the same notation as in Lemma 2.

Lemma 3 *Suppose that $\alpha_i > 1,\ i = 1, \ldots, p$. Then*

$$\begin{aligned} &\int_{R^{p'}} \int_{R_+^p} \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, h_2(\boldsymbol{f}, \boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q} \\ &\quad = \int_{R^{p'}} \int_{R_+^p} \Big\{1 - \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\}\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} + o(z). \end{aligned}$$

Proof.

Let $\mathcal{F}(z) = \{\boldsymbol{f} \mid f_1 \lambda_1 > \cdots > f_p \lambda_p\}$ and

$$\begin{aligned} g(\boldsymbol{f}, \boldsymbol{\lambda}) &= h_2(\boldsymbol{f}, \boldsymbol{\lambda}) - I_{\mathcal{F}(z)}(\boldsymbol{f}) \times \Big(1 - \sum_{i>j} \frac{f_i}{f_j} \frac{\lambda_i}{\lambda_j}\Big) \\ &= I_{\mathcal{F}(z)}(\boldsymbol{f}) \prod_{j<i} \left(1 - \frac{f_i}{f_j} \frac{\lambda_i}{\lambda_j}\right) - I_{\mathcal{F}(z)}(\boldsymbol{f}) \times \Big(1 - \sum_{i>j} \frac{f_i}{f_j} \frac{\lambda_i}{\lambda_j}\Big). \end{aligned}$$

Notice that $g(\boldsymbol{f}, \boldsymbol{\lambda})$ is a finite sum of terms, each of which has the form

$$I_{\mathcal{F}(z)}(\boldsymbol{f})(-1)^t \prod_{s=1}^{t} \frac{f_{i_s}}{f_{j_s}} \frac{\lambda_{i_s}}{\lambda_{j_s}}$$

with $i_s > j_s,\ s = 1, \ldots, t,\ t \geq 2$. Since $f_{i_s} f_{j_s}^{-1} \lambda_{i_s} \lambda_{j_s}^{-1} < 1$ on $\mathcal{F}(z)$, we have

$$\begin{aligned}
&\Big|\int_{R^{p'}} \int_{R_+^p} I_{\mathcal{F}(z)}(\boldsymbol{f}) \Big(\prod_{s=1}^{t} \frac{f_{i_s}}{f_{j_s}} \frac{\lambda_{i_s}}{\lambda_{j_s}}\Big) \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q}\Big| \\
&\quad \leq \int_{R^{p'}} \int_{R_+^p} I_{\mathcal{F}(z)}(\boldsymbol{f}) \Big(\prod_{s=1}^{2} \frac{f_{i_s}}{f_{j_s}} \frac{\lambda_{i_s}}{\lambda_{j_s}}\Big) \Big|\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\Big|\, d\boldsymbol{f} d\boldsymbol{q} \\
&\quad = \prod_{s=1}^{2} \Big(\prod_{m=j_s}^{i_s-1} a_m\Big) z^{i_s - j_s} \int_{R^{p'}} \int_{R_+^p} I_{\mathcal{F}(z)}(\boldsymbol{f}) \Big(\prod_{s=1}^{2} \frac{f_{i_s}}{f_{j_s}}\Big) \Big|\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\Big|\, d\boldsymbol{f} d\boldsymbol{q}.
\end{aligned}$$

Since $\alpha_i > 1,\ i = 1, \ldots, p$, the right side integral converges. Therefore the left side integral is $o(z)$. Consequently

$$\int_{R^{p'}} \int_{R_+^p} g(\boldsymbol{f}, \boldsymbol{\lambda})\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} = o(z). \tag{27}$$

Besides

$$\begin{aligned}
&\sum_{j+1<i} \int_{R^{p'}} \int_{R_+^p} \frac{f_i}{f_j} \frac{\lambda_i}{\lambda_j} I_{\mathcal{F}(z)}(\boldsymbol{f})\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} \\
&\quad = \sum_{j+1<i} \Big(\prod_{m=j}^{i-1} a_m\Big) z^{i-j} \int_{R^{p'}} \int_{R_+^p} \frac{f_i}{f_j} I_{\mathcal{F}(z)}(\boldsymbol{f})\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} \qquad (28) \\
&\quad = o(z).
\end{aligned}$$

From (27) and (28), we have

$$\begin{aligned}
&\int_{R^{p'}} \int_{R_+^p} h_2(\boldsymbol{f}, \boldsymbol{\lambda})\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} \\
&\quad = \int_{R^{p'}} \int_{R_+^p} \Big\{1 - \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\} I_{\mathcal{F}(z)}(\boldsymbol{f})\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} + o(z).
\end{aligned}$$

Now it suffices to show

$$I = \int_{R^{p'}} \int_{R_+^p} \Big\{1 - \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\} I_{\mathcal{F}(z)^C}(\boldsymbol{f})\, \eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, d\boldsymbol{f} d\boldsymbol{q} = o(z),$$

where $\mathcal{F}(z)^C$ is the complement of $\mathcal{F}(z)$. From the decomposition,

$$\mathcal{F}(z)^C = \bigcup_{i=1}^{p-1} \mathcal{F}_i(z),$$

where

$$\mathcal{F}_i(z) = \{\boldsymbol{f} \mid 1 \le \frac{f_{i+1}}{f_i}\frac{\lambda_{i+1}}{\lambda_i}\} = \{\boldsymbol{f} \mid 1 \le \frac{f_{i+1}}{f_i} a_i z\}, \quad i = 1, \ldots, p-1,$$

we notice $|I| \le \sum_{i=1}^{p-1} I_i$ with

$$I_i = \int_{R^{p'}} \int_{R_+^p} \Big\{1 + \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\} I_{\mathcal{F}_i}(\boldsymbol{f}) \left|\eta(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})\, h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\right| d\boldsymbol{f} d\boldsymbol{q}$$

for $1 \le i \le p-1$. With some constant M we have

$$\begin{aligned} I_i &\le M \int_{R^{p'}} \int_{R_+^p} \Big\{1 + \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\} I_{\mathcal{F}_i}(\boldsymbol{f}) \left|h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\right| d\boldsymbol{f} d\boldsymbol{q} \\ &\le M a_i z \int_{R^{p'}} \int_{R_+^p} \Big\{1 + \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\} \frac{f_{i+1}}{f_i}\, I_{\mathcal{F}_i}(\boldsymbol{f}) \left|h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\right| d\boldsymbol{f} d\boldsymbol{q}. \end{aligned}$$

By the dominated convergence theorem, of which use is guaranteed by $\alpha_i > 1,\ 1 \le i \le p$, we have

$$\begin{aligned} &\lim_{z \to 0} \int_{R^{p'}} \int_{R_+^p} \Big\{1 + \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\} \frac{f_{i+1}}{f_i}\, I_{\mathcal{F}_i}(\boldsymbol{f}) \left|h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\right| d\boldsymbol{f} d\boldsymbol{q} \\ &\quad = \int_{R^{p'}} \int_{R_+^p} \lim_{z \to 0} \Big[\Big\{1 + \Big(\sum_{j=1}^{p-1} \frac{f_{j+1}}{f_j} a_j\Big) z\Big\}\, I_{\mathcal{F}_i}(\boldsymbol{f})\Big] \frac{f_{i+1}}{f_i} \left|h^*(\boldsymbol{f}, \boldsymbol{q}; \boldsymbol{\alpha}, \boldsymbol{\beta})\right| d\boldsymbol{f} d\boldsymbol{q} \\ &\quad = 0. \end{aligned}$$

Therefore $I_i = o(z),\ i = 1, \ldots, p-1$. This means $I = o(z)$. ∎

Now we state the main result of this section. The next theorem tells us that $P(\widetilde{\boldsymbol{G}} \in \mathcal{N}(\boldsymbol{I}_p)^C)$ vanishes faster than z. This guarantees that as far as the terms of order $O(z)$ are concerned, we can concentrate on an arbitrary small neighborhood $\mathcal{N}(\boldsymbol{I}_p)$ of $\boldsymbol{I}_p$. In the proof of the theorem we need approximations of integrals involving $h_1(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})$ in (16).

Theorem 1 *Suppose that $n > p + 4$. For any open neighborhood $\mathcal{N}(\boldsymbol{I}_p)$ of $\boldsymbol{I}_p \in \mathcal{O}^+(p)$,*

$$P(\widetilde{\boldsymbol{G}} \in \mathcal{N}(\boldsymbol{I}_p)) = 1 + o(z),$$

or equivalently

$$P(\widetilde{\boldsymbol{G}} \in \mathcal{N}(\boldsymbol{I}_p)^C) = o(z).$$

Proof.

It suffices to show that $P(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*)^C) = o(z)$ for any small enough ϵ. Let $\tau = z^{\frac{1}{2}}$ and

$$\boldsymbol{u}(\tau) = \Big(u_{ij}(\tau)\Big)_{i>j} = \Big(q_{ij} f_j^{-\frac{1}{2}} \Big(\prod_{m=j}^{i-1} a_m^{\frac{1}{2}}\Big) \tau^{i-j}\Big)_{i>j}.$$

Then we have

$$\begin{aligned} h_1(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) &= \exp\Big[-\frac{1}{2}\Big\{\sum_{i=1}^{p}(g_{ii}^2(\boldsymbol{u})-1)f_i + \sum_{i<j} g_{ij}^2(\boldsymbol{u})f_j\lambda_j\lambda_i^{-1}\Big\}\Big] \\ &= t(\tau), \end{aligned}$$

where

$$t(\tau) = \exp\Big[-\frac{1}{2}\Big\{\sum_{i=1}^{p}(g_{ii}^2(\boldsymbol{u}(\tau))-1)f_i + \sum_{i<j} g_{ij}^2(\boldsymbol{u}(\tau))f_j\Big(\prod_{m=i}^{j-1} a_m\Big)\tau^{2(j-i)}\Big\}\Big]. \qquad (29)$$

Then Taylor expansion of $t(\tau)$ around 0 is given by

$$t(\tau) = t(0) + \tau t'(0) + \frac{\tau^2}{2}t''(0) + \frac{\tau^3}{3!}t'''(\theta(\tau)),$$

where $0 < \theta(\tau) < \tau$.

First notice that $t(0) = 1$. Next straightforward calculation shows that $t'''(\tau)/t(\tau)$ can be expressed as a finite sum (denoted by $\sum'$ hereafter) of terms each of which is of the form

$$c_0 M(\boldsymbol{u}(\tau)) \prod_{i=1}^{p} f_i^{\alpha_i} \prod_{i>j} q_{ij}^{\beta_{ij}}\ \tau^{\gamma}, \qquad c_0\text{: constant},$$

where

$$\alpha_i \geq -\frac{3}{2}\ (1 \leq i \leq p-1), \quad \alpha_p \geq 0, \quad \beta_{ij} = 0,1,\ldots\ (1 \leq j < i \leq p), \quad \gamma = 0,1,\ldots,$$

and $M(\boldsymbol{u})$ is a multiple of some (possibly higher order) derivatives of $g_{ij}(\boldsymbol{u})$ $(1 \leq i \leq j \leq p)$ w.r.t $\boldsymbol{u}$ (the order may vary from zero to three), hence a bounded function on $\bar{U}^*$.

Now we calculate $t'(0)$ and $t''(0)$. From (12), the real analytic function $g_{ij}(\boldsymbol{u}(\tau))$ of τ is expanded as

$$\begin{aligned} g_{ii}(\boldsymbol{u}(\tau)) &= 1 - \frac{1}{2}\tau^2(q_{i,i-1}^2 f_{i-1}^{-1}a_{i-1} + q_{i+1,i}^2 f_i^{-1}a_i) + \widetilde{g}_{ii}(\tau), \qquad 1 \leq i \leq p, \\ g_{ij}(\boldsymbol{u}(\tau)) &= \begin{cases} -\tau(q_{i+1,i}f_i^{-\frac{1}{2}}a_i^{\frac{1}{2}}) + \widetilde{g}_{i,i+1}(\tau) & \text{if } j = i+1, \\ -\tau^2(q_{i+2,i}f_i^{-\frac{1}{2}}a_i^{\frac{1}{2}}a_{i+1}^{\frac{1}{2}}) + \tau^2(q_{i+1,i}q_{i+2,i+1}f_i^{-\frac{1}{2}}f_{i+1}^{-\frac{1}{2}}a_i^{\frac{1}{2}}a_{i+1}^{\frac{1}{2}}) + \widetilde{g}_{i,i+2}(\tau) & \text{if } j = i+2, \\ \widetilde{g}_{ij}(\tau) & \text{if } j \geq i+3, \end{cases} \end{aligned}$$

where $\widetilde{g}_{ij}(\tau)$ $(1 \leq i \leq j \leq p)$ contains only the terms in τ of order higher than 2. Remember the conventional definition, $a_0 = a_p = 0$. From this, we have the following expansions around the point $\tau = 0$.

$$(g_{ii}^2(\boldsymbol{u}(\tau))-1)f_i = -\tau^2(q_{i,i-1}^2\frac{f_i}{f_{i-1}}a_{i-1} + q_{i+1,i}^2 a_i) + \xi_{ii}(\tau), \qquad 1 \leq i \leq p,$$

$$g_{ij}^2(\boldsymbol{u}(\tau))f_j\Big(\prod_{m=i}^{j-1} a_m\Big)\tau^{2(j-i)} = \xi_{ij}(\tau), \qquad 1 \leq i < j \leq p,$$

where $\xi_{ij}(\tau)$ $(1 \le i \le j \le p)$ contains only the terms of order higher than 2. Therefore

$$t(\tau) = \exp\Big\{\frac{1}{2}\tau^2 \sum_{i=1}^{p}(q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i) - \frac{1}{2}\sum_{i \le j} \xi_{ij}(\tau)\Big\}.$$

From this we have $t'(0) = 0$ and

$$t''(0) = \sum_{i=1}^{p}(q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i).$$

Consequently $h_1(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})$ in (16) can be expressed as

$$\begin{aligned} h_1(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) &= 1 + \frac{z}{2}\sum_{i=1}^{p}(q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i) \\ &\quad + \frac{z^{\frac{3}{2}}}{3!} t(\theta(\tau)) {\sum}' c_0 M(\boldsymbol{u}(\theta(\tau))) \prod_{i=1}^{p} f_i^{\alpha_i} \prod_{i>j} q_{ij}^{\beta_{ij}} \, (\theta(\tau))^{\gamma}. \end{aligned} \tag{30}$$

Now we calculate $P(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*))$. As in (15),

$$\begin{aligned} P(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*)) &= E[I(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*))] \\ &= \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q}) \, h_1(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) \, h_2(\boldsymbol{f}, \boldsymbol{\lambda}) \, h_3(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) \, d\boldsymbol{f} d\boldsymbol{q}. \end{aligned}$$

Substituting $h_1(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda})$ in the right side with (30), we have

$$P(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*)) = I_1 + \frac{z}{2} I_2 + \frac{z^{\frac{3}{2}}}{3!} I_3,$$

where

$$\begin{aligned} I_1 &= \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q}) \, h_2(\boldsymbol{f}, \boldsymbol{\lambda}) \, h_3(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) \, d\boldsymbol{f} d\boldsymbol{q}, \\ I_2 &= \sum_{i=1}^{p} \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q}) \, (q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i) \, h_2(\boldsymbol{f}, \boldsymbol{\lambda}) \, h_3(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) \, d\boldsymbol{f} d\boldsymbol{q}, \\ I_3 &= {\sum}' \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q}) \, t(\theta(\tau)) \, c_0 \, M(\boldsymbol{u}(\theta(\tau))) \prod_{i=1}^{p} f_i^{\alpha_i} \prod_{i>j} q_{ij}^{\beta_{ij}} \, (\theta(\tau))^{\gamma} \\ &\qquad \times h_2(\boldsymbol{f}, \boldsymbol{\lambda}) \, h_3(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) d\boldsymbol{f} d\boldsymbol{q}. \end{aligned}$$

First we prove $z^{\frac{3}{2}} I_3 = o(z)$. Note that if $\boldsymbol{u} = \boldsymbol{u}(\tau) \in \bar{U}^*(\epsilon)$, that is, $\sum_{s>t} u_{st}^2(\tau) \le \epsilon^2$, then from $\theta(\tau) < \tau$, we have

$$\sum_{s>t} u_{st}^2(\theta(\tau)) = \sum_{s>t} q_{st}^2 f_t^{-1} \Big(\prod_{m=t}^{s-1} a_m\Big)(\theta(\tau))^{2(s-t)} < \sum_{s>t} u_{st}^2(\tau) \le \epsilon^2.$$

This means $\boldsymbol{u}(\theta(\tau)) \in \bar{U}^*$. Since $M(\boldsymbol{u})$ is bounded on $\bar{U}^*$, $M(\boldsymbol{u}(\theta(\tau)))$ is also bounded. Besides from (13), $g_{ii}^2(\boldsymbol{u}(\theta(\tau))) \geq 1/2$ $(1 \leq i \leq p)$. Therefore $h_0(\boldsymbol{f},\boldsymbol{q})\, t(\theta(\tau)) \leq \bar{h}_0(\boldsymbol{f},\boldsymbol{q})$, where

$$\bar{h}_0(\boldsymbol{f},\boldsymbol{q}) = \tilde{c}\left(\prod_{i=1}^{p} f_i^{\frac{n-i-1}{2}}\right)\exp\left(-\frac{1}{2}\sum_{i>j} q_{ij}^2\right)\exp\left(-\frac{1}{4}\sum_{i=1}^{p} f_i\right). \tag{31}$$

$h_2(\boldsymbol{f},\boldsymbol{q})$ and $h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$ are also bounded functions. Consequently there exists some $K>0$ such that for any $z<1$

$$\begin{aligned} |I_3| &\leq \sum{}' \int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\, t(\theta(\tau))\, |c_0\, M(\boldsymbol{u}(\theta(\tau)))| \prod_{i=1}^{p} f_i^{\alpha_i} \prod_{i>j} |q_{ij}^{\beta_{ij}}|\, (\theta(\tau))^{\gamma} \\ &\qquad \times h_2(\boldsymbol{f},\boldsymbol{\lambda})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) d\boldsymbol{f} d\boldsymbol{q} \\ &\leq K\sum{}' \int_{R^{p'}}\int_{R_+^p} \bar{h}_0(\boldsymbol{f},\boldsymbol{q}) \prod_{i=1}^{p} f_i^{\alpha_i} \prod_{i>j} |q_{ij}^{\beta_{ij}}|\, d\boldsymbol{f} d\boldsymbol{q}. \end{aligned}$$

The condition $n-p-4>0$ guarantees the convergence of the right side integral. This means

$$z^{\frac{3}{2}} I_3 = o(z). \tag{32}$$

Now we consider zI_2. Since $h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$ is bounded and $n-p-4>0$, the conditions on α_i $(1 \leq i \leq p)$ in Lemma 3 is satisfied. Therefore we have

$$\begin{aligned} zI_2 &= z\sum_{i=1}^{p}\int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\, (q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i)\, h_2(\boldsymbol{f},\boldsymbol{\lambda})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q} \\ &= z\sum_{i=1}^{p}\left(I_{21}^{(i)} - zI_{22}^{(i)}\right) + o(z^2), \end{aligned}$$

where for $1 \leq i \leq p$

$$\begin{aligned} I_{21}^{(i)} &= \int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\, (q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i)\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}, \\ I_{22}^{(i)} &= \int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\, (q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i)\left(\sum_{j=1}^{p-1}\frac{f_{j+1}}{f_j} a_j\right) h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}. \end{aligned}$$

Since $z^2 I_{22}^{(i)} = o(z)$, we have

$$zI_2 = z\sum_{i=1}^{p} I_{21}^{(i)} + o(z).$$

Noticing again that $n-p-4>0$ guarantees the conditions on α_i $(1 \leq i \leq p)$ in Lemma 2, we have

$$\begin{aligned} zI_{21}^{(i)} &= z\int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\, (q_{i,i-1}^2 \frac{f_i}{f_{i-1}} a_{i-1} + q_{i+1,i}^2 a_i)\left(1 + \frac{1}{2}\sum_{i>j}(i-j)u_{ij}^2\right) d\boldsymbol{f} d\boldsymbol{q} + o(z^2) \\ &= zI_{211}^{(i)} + zI_{212}^{(i)}, \end{aligned}$$

where

$$I_{211}^{(i)} = \int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\,(q_{i,i-1}^2\frac{f_i}{f_{i-1}}a_{i-1} + q_{i+1,i}^2 a_i)\,d\boldsymbol{f}d\boldsymbol{q},$$

$$I_{212}^{(i)} = \frac{1}{2}\sum_{s>t}(s-t)\Big(\prod_{m=t}^{s-1}a_m\Big)z^{s-t}\int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\,(q_{i,i-1}^2\frac{f_i}{f_{i-1}}a_{i-1} + q_{i+1,i}^2 a_i)\,q_{st}^2 f_t^{-1}\,d\boldsymbol{f}d\boldsymbol{q}.$$

Obviously $zI_{212}^{(i)} = o(z)$. Consequently

$$zI_2 = z\sum_{i=1}^{p} I_{211}^{(i)} + o(z).$$

Using the fact that $h_0(\boldsymbol{f},\boldsymbol{q})$ is the density function of the asymptotic distribution of $(\boldsymbol{f},\boldsymbol{q})$ given by (3), we have

$$\begin{aligned}\sum_{i=1}^{p} I_{211}^{(i)} &= \sum_{i=2}^{p} a_{i-1}E[q_{i,i-1}^2\frac{f_i}{f_{i-1}}] + \sum_{i=1}^{p-1} a_i E[q_{i+1,i}^2]\\ &= \sum_{i=2}^{p} a_{i-1}E[\chi^2(1)]\,E[(\chi^2(n-i+2))^{-1}]\,E[\chi^2(n-i+1)] + \sum_{i=1}^{p-1} a_i E[\chi^2(1)]\\ &= \sum_{i=2}^{p} a_{i-1}\frac{n-i+1}{n-i} + \sum_{i=1}^{p-1} a_i\\ &= \sum_{i=1}^{p-1} a_i\{\frac{n-i}{n-i-1}+1\}.\end{aligned}$$

Consequently

$$\frac{z}{2}I_2 = \frac{z}{2}\sum_{i=1}^{p-1} a_i\{\frac{n-i}{n-i-1}+1\} + o(z). \qquad (33)$$

Last we evaluate I_1. Using again Lemma 3, we have $I_1 = I_{11} - zI_{12} + o(z)$, where

$$I_{11} = \int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\,h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\,d\boldsymbol{f}d\boldsymbol{q}$$

$$I_{12} = \sum_{j=1}^{p-1} a_j \int_{R^{p'}}\int_{R_+^p} \frac{f_{j+1}}{f_j}\,h_0(\boldsymbol{f},\boldsymbol{q})\,h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\,d\boldsymbol{f}d\boldsymbol{q}.$$

Furthermore by Lemma 2, we have $I_{11} = I_{111} + I_{112} + o(z)$, $I_{12} = I_{121} + I_{122} + o(z)$, where

$$I_{111} = \int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})\,d\boldsymbol{f}d\boldsymbol{q},$$

$$I_{112} = \frac{1}{2}\sum_{s>t}(s-t)\Big(\prod_{m=t}^{s-1}a_m\Big)\,z^{s-t}\int_{R^{p'}}\int_{R_+^p} h_0(\boldsymbol{f},\boldsymbol{q})q_{st}^2\,f_t^{-1}\,d\boldsymbol{f}d\boldsymbol{q},$$

$$I_{121} = \sum_{j=1}^{p-1} a_j \int_{R^{p'}}\int_{R_+^p} \frac{f_{j+1}}{f_j}\,h_0(\boldsymbol{f},\boldsymbol{q})\,d\boldsymbol{f}d\boldsymbol{q},$$

$$I_{122} = \frac{1}{2}\sum_{j=1}^{p-1} a_j\sum_{s>t}(s-t)\Big(\prod_{m=t}^{s-1}a_m\Big)\,z^{s-t}\int_{R^{p'}}\int_{R_+^p}\frac{f_{j+1}}{f_j}q_{st}^2\,f_t^{-1}\,h_0(\boldsymbol{f},\boldsymbol{q})\,d\boldsymbol{f}d\boldsymbol{q}.$$

Note that $I_{111} = 1$, $zI_{122} = o(z)$ and

$$\begin{aligned} I_{112} &= \frac{z}{2}\sum_{t=1}^{p-1} a_t \int_{R^{p'}}\int_{R^p_+} h_0(\boldsymbol{f},\boldsymbol{q})\, q^2_{t+1,t} f_t^{-1}\, d\boldsymbol{f} d\boldsymbol{q} + o(z) \\ &= \frac{z}{2}\sum_{t=1}^{p-1} a_t\, E[\chi^2(1)]\, E[(\chi^2(n-t+1))^{-1}] + o(z) \\ &= \frac{z}{2}\sum_{t=1}^{p-1} a_t\, \frac{1}{n-t-1} + o(z), \end{aligned}$$

$$zI_{121} = z\sum_{j=1}^{p-1} a_j\, E[\chi^2(n-j)]\, E[(\chi^2(n-j+1))^{-1}] = z\sum_{j=1}^{p-1} a_j\, \frac{n-j}{n-j-1}.$$

Consequently

$$I_1 = 1 + \frac{z}{2}\sum_{t=1}^{p-1} a_t \frac{1}{n-t-1} - z\sum_{j=1}^{p-1} a_j \frac{n-j}{n-j-1} + o(z). \tag{34}$$

From (32), (33) and (34),

$$\begin{aligned} P(\widetilde{\boldsymbol{G}} \in \boldsymbol{G}(\bar{U}^*)) &= 1 + \frac{z}{2}\Big[\sum_{i=1}^{p-1} a_i\Big\{-\frac{n-i}{n-i-1} + \frac{1}{n-i-1} + 1\Big\}\Big] + o(z) \\ &= 1 + o(z). \end{aligned}$$

∎

Next theorem gives the expansion of the asymptotic distribution function of $(\boldsymbol{f},\boldsymbol{q})$ w.r.t. z up to the first order.

Theorem 2 *Suppose that $n-p-4>0$. Then*

$$\begin{aligned} &P\Big(f_i \le \alpha_i,\ 1\le \forall i \le p, \quad q_{st} \le \alpha_{st},\ 1 \le \forall t < \forall s \le p\Big) \\ &\quad = \prod_{i=1}^{p} F_{n-i+1}(\alpha_i) \prod_{s>t} \Phi(\alpha_{st}) \\ &\qquad + \frac{1}{2} z \sum_{j=1}^{p-1} a_j \Big(\prod_{(s,t)\neq(j+1,j)} \Phi(\alpha_{st})\Big)\Big(\prod_{i\neq j,j+1} F_{n-i+1}(\alpha_i)\Big)\zeta_j(\alpha_j,\alpha_{j+1},\alpha_{j+1,j}) + o(z), \\ &P(f_i \le \alpha_i) \\ &\quad = F_{n-i+1}(\alpha_i) - \frac{1}{2} z\Big\{a_i F_{n-i-1}(\alpha_i) - \Big(a_{i-1}\frac{n-i+1}{n-i} + a_i\Big)F_{n-i+1}(\alpha_i) \\ &\qquad\qquad + a_{i-1}\frac{n-i+1}{n-i} F_{n-i+3}(\alpha_i)\Big\} + o(z), \qquad 1 \le i \le p, \\ &P(q_{st} \le \alpha_{st}) \\ &\quad = \begin{cases} \Phi(\alpha_{st}) - z\dfrac{n-t}{n-t-1} a_t \Phi'(\alpha_{st})\alpha_{st} + o(z) & \text{if } s = t+1, \\ \Phi(\alpha_{st}) + o(z) & \text{if } t+2 \le s \le p, \end{cases} \end{aligned}$$

where $F_n(\alpha) = P(\chi^2(n) \le \alpha), \quad \Phi(\alpha) = P(N(0,1) \le \alpha)$ *and*

$$\begin{aligned}\zeta_j(a,b,c) &= \frac{1}{n-j-1} F_{n-j-1}(a) F_{n-j}(b) \Big(\Phi(c) - \Phi'(c)c\Big) \\ &\quad -2\frac{n-j}{n-j-1} F_{n-j-1}(a) F_{n-j+2}(b) \Phi(c) \\ &\quad + F_{n-j+1}(a) F_{n-j}(b) \Big(\Phi(c) - \Phi'(c)c\Big) \\ &\quad + \frac{n-j}{n-j-1} F_{n-j-1}(a) F_{n-j+2}(b) \Big(\Phi(c) - \Phi'(c)c\Big).\end{aligned}$$

Proof.

Let A denote the event

$$f_i \le \alpha_i,\ 1 \le \forall i \le p, \text{ and } q_{st} \le \alpha_{st},\ 1 \le \forall t < \forall s \le p.$$

We have

$$P(A) = E[I_A\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] + E[I_A\, I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})].$$

Since

$$E[I_A\, I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] \le E[I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})]$$

and the right side is $o(z)$ by Theorem 1,

$$E[I_A\, I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] = o(z).$$

We consider $E[I_A\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})]$. Substitute $I_A\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})$ with $I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})$ in the proof of Theorem 1. Then since

$$I_A\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}}) \le I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}}),$$

the integrals of order $o(z)$ in the proof are still $o(z)$ after the substitution. Therefore proceeding as in the proof of Theorem 1 we have

$$P(A) = \prod_{i=1}^{p} F_{n-i+1}(\alpha_i) \prod_{s>t} \Phi(\alpha_{st}) + \frac{1}{2} z \sum_{m=1}^{4} I_m + o(z), \tag{35}$$

where

$$\begin{aligned}I_1 &= \sum_{j=1}^{p-1} a_j \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q})\, q_{j+1,j}^2\, f_j^{-1}\, I_A\, d\boldsymbol{f} d\boldsymbol{q}, \\ I_2 &= -2 \sum_{j=1}^{p-1} a_j \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q}) \frac{f_{j+1}}{f_j}\, I_A\, d\boldsymbol{f} d\boldsymbol{q}, \\ I_3 &= \sum_{j=1}^{p-1} a_j \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q})\, q_{j+1,j}^2\, I_A\, d\boldsymbol{f} d\boldsymbol{q}, \\ I_4 &= \sum_{j=1}^{p-1} a_j \int_{R^{p'}} \int_{R_+^p} h_0(\boldsymbol{f}, \boldsymbol{q})\, q_{j+1,j}^2\, \frac{f_{j+1}}{f_j}\, I_A\, d\boldsymbol{f} d\boldsymbol{q}.\end{aligned}$$

If we use the formulas

$$
\begin{aligned}
E[\chi^2(n)\, I(\chi^2(n)\le\alpha)] &= nF_{n+2}(\alpha),\\
E[(\chi^2(n))^{-1}\, I(\chi^2(n)\le\alpha)] &= \frac{1}{n-2}F_{n-2}(\alpha),\\
E[(N(0,1))^2\, I(N(0,1)\le\alpha)] &= \Phi(\alpha)-\Phi'(\alpha)\alpha,
\end{aligned}
$$

we have

$$
\begin{aligned}
I_1 &= \sum_{j=1}^{p-1} a_j\Big(\prod_{(s,t)\neq(j+1,j)}\Phi(\alpha_{st})\Big)\Big(\prod_{i\neq j}F_{n-i+1}(\alpha_i)\Big)\\
&\qquad\times\Big(\Phi(\alpha_{j+1,j})-\Phi'(\alpha_{j+1,j})\alpha_{j+1,j}\Big)\frac{1}{n-j-1}F_{n-j-1}(\alpha_j),\\
I_2 &= -2\sum_{j=1}^{p-1} a_j\Big(\prod_{s>t}\Phi(\alpha_{st})\Big)\Big(\prod_{i\neq j,j+1}F_{n-i+1}(\alpha_i)\Big)\\
&\qquad\times\frac{n-j}{n-j-1}F_{n-j+2}(\alpha_{j+1})F_{n-j-1}(\alpha_j),\\
I_3 &= \sum_{j=1}^{p-1} a_j\Big(\prod_{(s,t)\neq(j+1,j)}\Phi(\alpha_{st})\Big)\Big(\prod_{i=1}^{p}F_{n-i+1}(\alpha_i)\Big)\Big(\Phi(\alpha_{j+1,j})-\Phi'(\alpha_{j+1,j})\alpha_{j+1,j}\Big),\\
I_4 &= \sum_{j=1}^{p-1} a_j\Big(\prod_{(s,t)\neq(j+1,j)}\Phi(\alpha_{st})\Big)\Big(\prod_{i\neq j,j+1}F_{n-i+1}(\alpha_i)\Big)\\
&\qquad\times\frac{n-j}{n-j-1}F_{n-j+2}(\alpha_{j+1})F_{n-j-1}(\alpha_j)\Big(\Phi(\alpha_{j+1,j})-\Phi'(\alpha_{j+1,j})\alpha_{j+1,j}\Big).
\end{aligned}
$$

If we substitute these into (35), we have the result.

We can prove the statement for $P(f_i\le\alpha_i)$ $(1\le i\le p)$ or $P(q_{st}\le\alpha_{st})$ $(1\le t<s\le p)$ completely similarly. Note that the proof is essentially the same as putting ∞ in all the elements other than $\alpha_i(\alpha_{st})$ in the result on $P\Big(f_i\le\alpha_i,\ 1\le\forall i\le p,\quad q_{st}\le\alpha_{st},\ 1\le\forall t<\forall s\le p\Big)$.

∎

The next lemma, which will be used in the next section, is of interest by itself. Takemura and Sheena (2002) showed $\boldsymbol{x}$ converges in probability to $\mathbf{0}$; that is,

$$P(\boldsymbol{x}\in\mathcal{N}(\mathbf{0})^C)\to 0$$

as $\|\boldsymbol{y}\|\to 0$ for any open neighborhood $\mathcal{N}(\mathbf{0})$ of $\mathbf{0}$ in $[0,1]^{p-1}$. The following result tells that $P(\boldsymbol{x}\in\mathcal{N}(\mathbf{0})^C)$ vanishes more rapidly than z.

Lemma 4 *Suppose that $n-p-1>0$. For any open neighborhood $\mathcal{N}(\mathbf{0})$ of $\mathbf{0}$ in $[0,1]^{p-1}$,*

$$P(\boldsymbol{x}\in\mathcal{N}(\mathbf{0})^C)=o(z).$$

Proof. Without loss of generality, we can assume that $\mathbf{\Sigma} = \mathbf{\Lambda}$ is diagonal.

It suffices to prove the statement for the case

$$\mathcal{N}(\mathbf{0}) = [0, \epsilon)^{p-1}$$

with arbitrarily small $\epsilon > 0$. Note that

$$\mathcal{N}(\mathbf{0})^C = \bigcup_{1 \le i \le p-1} \{\boldsymbol{x} \mid x_i \ge \epsilon\},$$

hence that

$$P_{\mathbf{\Lambda}}(\boldsymbol{x} \in \mathcal{N}(\mathbf{0})^C) \le \sum_{i=1}^{p-1} E_{\mathbf{\Lambda}}[I(x_i \ge \epsilon)],$$

where the subscript $\mathbf{\Lambda}$ of P and E refers to the probability and the expected value under $\mathbf{\Sigma} = \mathbf{\Lambda}$. We will prove $E_{\mathbf{\Lambda}}[I(x_i \ge \epsilon)] = o(z)$. Since the following inequality holds,

$$E_{\mathbf{\Lambda}}[I(x_i \ge \epsilon)] \le \epsilon^{-1} E_{\mathbf{\Lambda}}[x_i I(x_i \ge \epsilon)] = \epsilon^{-1} a_i z E_{\mathbf{\Lambda}}[\frac{f_{i+1}}{f_i} I(x_i \ge \epsilon)],$$

it suffices to show

$$\lim_{z \to 0} E_{\mathbf{\Lambda}}[\frac{f_{i+1}}{f_i} I(x_i \ge \epsilon)] = 0.$$

By Lemma 5 in Appendix, we have

$$\begin{aligned} E_{\mathbf{\Lambda}}[\frac{f_{i+1}}{f_i} I(x_i \ge \epsilon)] &\le E_{\mathbf{\Lambda}}\Big[\Big(\sum_{j=i+1}^{p} \tilde{w}_{jj}\Big)\Big(\sum_{j=1}^{i} \tilde{w}^{jj}\Big) I\Big(x_i(\boldsymbol{W}) \ge \epsilon\Big)\Big] \\ &= E_{\boldsymbol{I}}\Big[\Big(\sum_{j=i+1}^{p} w_{jj}\Big)\Big(\sum_{j=1}^{i} w^{jj}\Big) I\Big(x_i(\mathbf{\Lambda}^{\frac{1}{2}} \boldsymbol{W} \mathbf{\Lambda}^{\frac{1}{2}}) \ge \epsilon\Big)\Big]. \end{aligned}$$

By the dominated convergence theorem (note $n - p - 1 > 0$) and Lemma 6 in Appendix,

$$\begin{aligned} &\lim_{z \to 0} E_{\boldsymbol{I}}\Big[\Big(\sum_{j=i+1}^{p} w_{jj}\Big)\Big(\sum_{j=1}^{i} w^{jj}\Big) I\Big(x_i(\mathbf{\Lambda}^{\frac{1}{2}} \boldsymbol{W} \mathbf{\Lambda}^{\frac{1}{2}}) \ge \epsilon\Big)\Big] \\ &= E_{\boldsymbol{I}}\Big[\Big(\sum_{j=i+1}^{p} w_{jj}\Big)\Big(\sum_{j=1}^{i} w^{jj}\Big) \lim_{z \to 0} I\Big(x_i(\mathbf{\Lambda}^{\frac{1}{2}} \boldsymbol{W} \mathbf{\Lambda}^{\frac{1}{2}}) \ge \epsilon\Big)\Big] \\ &= 0 \end{aligned}$$

∎

4 Asymptotic expansion of risk

In this section we consider the asymptotic expansion of the risk of a scale and orthogonally equivariant estimator with respect to Stein's loss. It is possible to apply directly the lemmas and theorems in the previous section to the calculation of the risk, but it seems somewhat cumbersome. We take another approach which uses the unbiased estimator of risk.

The unbiased estimator of risk was given independently by Stein (1977) and Haff (1979). Sheena (1995) gave an alternative short derivation for the case of orthogonally equivariant estimators. Kubokawa and Srivastava (1999) generalized the unbiased estimator to elliptically contoured distributions.

Suppose that an orthogonally equivariant estimator given by (6) has C^1 functions $\psi_i(\boldsymbol{l}),\ i = 1,\ldots,p$. Let

$$\hat{R}(\widehat{\boldsymbol{\Sigma}}) = (n-p-1)\sum_{i=1}^{p}\frac{\psi_i(\boldsymbol{l})}{l_i} + 2\sum_{i=1}^{p}\frac{\partial\psi_i(\boldsymbol{l})}{\partial l_i} + 2\sum_{i<j}\frac{\psi_i(\boldsymbol{l})-\psi_j(\boldsymbol{l})}{l_i-l_j}.$$

Then

$$E[\hat{R}(\widehat{\boldsymbol{\Sigma}})] = E[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1})], \quad \forall\boldsymbol{\Sigma} > 0. \tag{36}$$

Note that if an orthogonally equivariant estimator (6) is also scale equivariant, $c_i(\boldsymbol{l}),\ 1 \le i \le p$, depends on $\boldsymbol{l}$ only through $x_i = l_{i+1}/l_i,\ 1 \le i \le p-1$. Hence we use the notation $c_i(\boldsymbol{x}),\ 1 \le i \le p$. Using (36) for this estimator, we obtain

$$\begin{aligned}
&E[L_1(\widehat{\boldsymbol{\Sigma}},\boldsymbol{\Sigma})] \\
&= E\Big[(n-p+1)\sum_{i=1}^{p}c_i(\boldsymbol{x}) + 2\sum_{i=1}^{p}l_i\frac{\partial c_i(\boldsymbol{x})}{\partial l_i} + 2\sum_{i<j}\frac{c_i(\boldsymbol{x})l_i - c_j(\boldsymbol{x})l_j}{l_i-l_j} \\
&\qquad\qquad - \sum_{i=1}^{p}\log c_i(\boldsymbol{x}) - \log|\boldsymbol{W}\boldsymbol{\Sigma}^{-1}| - p\Big] \\
&= E\Big[\sum_{i=1}^{p}(n+p+1-2i)c_i(\boldsymbol{x}) + 2\sum_{i<j}(c_i(\boldsymbol{x})-c_j(\boldsymbol{x}))\frac{l_j}{l_i-l_j} + 2\sum_{i=1}^{p}l_i\frac{\partial c_i(\boldsymbol{x})}{\partial l_i} \\
&\qquad\qquad - \sum_{i=1}^{p}\log c_i(\boldsymbol{x}) - \log|\boldsymbol{W}\boldsymbol{\Sigma}^{-1}| - p\Big] \\
&= E\Big[\sum_{i=1}^{p}(n+p+1-2i)c_i(\boldsymbol{x}) + 2\sum_{i<j}(c_i(\boldsymbol{x})-c_j(\boldsymbol{x}))\frac{\prod_{m=i}^{j-1}x_m}{1-\prod_{m=i}^{j-1}x_m} \\
&\qquad -2\sum_{i=1}^{p-1}\frac{\partial c_i(\boldsymbol{x})}{\partial x_i}x_i + 2\sum_{i=2}^{p}\frac{\partial c_i(\boldsymbol{x})}{\partial x_{i-1}}x_{i-1} - \sum_{i=1}^{p}\log c_i(\boldsymbol{x}) - \log|\boldsymbol{W}\boldsymbol{\Sigma}^{-1}| - p\Big].
\end{aligned} \tag{37}$$

If we use this formula, as we will see later, we only have to evaluate $E[x_i],\ 1 \le i \le p-1$, for the expansion of the asymptotic risk w.r.t. z up to the first order.

Theorem 3 *Suppose that* $n-p-4 > 0$. *Then*

$$E[x_i] = a_i z\frac{n-i}{n-i-1} + o(z), \qquad 1 \le i \le p-1.$$

Proof.

Note

$$E[x_i] = E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] + E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})].$$

Since

$$E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] \le E[I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] = o(z)$$

by Theorem 1, we have

$$E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] = o(z).$$

Therefore we only have to prove

$$E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] = a_i z \frac{n-i}{n-i-1} + o(z), \qquad 1 \le i \le p-1. \tag{38}$$

As in (15),

$$E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] = a_i z \int_{R^{p'}} \int_{R^p_+} \frac{f_{i+1}}{f_i}\, h_0(\boldsymbol{f},\boldsymbol{q})\, h_1(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h_2(\boldsymbol{f},\boldsymbol{\lambda})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}. \tag{39}$$

The argument hereafter is almost the same as that in the proof of Theorem 1. The slight difference is that here we only need lower terms in the expansions of the related functions or integrals since z already exists in (39).

First consider the expansion of h_1. If we use $t(\tau)$ in (29) and its Taylor expansion,

$$t(\tau) = t(0) + \tau t'(0) + \frac{\tau^2}{2} t''(\theta(\tau)), \qquad 0 < \theta(\tau) < \tau,$$

by similar argument for the proof of (30), we can prove

$$h_1(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda}) = 1 + z\, t(\theta(\tau)) {\sum}'' \widetilde{M}(\boldsymbol{f},\boldsymbol{q},\tau) \prod_{j=1}^{p} f_j^{\alpha_j} \prod_{s>t} q_{st}^{\beta_{st}}; \tag{40}$$

Here ${\sum}''$ is a finite sum of terms of the form $\widetilde{M}(\boldsymbol{f},\boldsymbol{q},\tau) \prod_{j=1}^{p} f_j^{\alpha_j} \prod_{s>t} q_{st}^{\beta_{st}}$, where $\alpha_j \ge -1$ $(1 \le j \le p-1)$, $\alpha_p \ge 0$, $\beta_{st} = 0, 1, \ldots$ $(1 \le t < s \le p)$ and $\widetilde{M}(\boldsymbol{f},\boldsymbol{q},\tau)$ is bounded on $R^p_+ \times R^{p'} \times [0,1]$.

Substituting (40) into (39), we have

$$E[x_i\, I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] = a_i z I_1 + a_i z^2 I_2,$$

where

$$\begin{aligned} I_1 &= \int_{R^{p'}} \int_{R^p_+} \frac{f_{i+1}}{f_i}\, h_0(\boldsymbol{f},\boldsymbol{q})\, h_2(\boldsymbol{f},\boldsymbol{\lambda})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}, \\ I_2 &= {\sum}'' \int_{R^{p'}} \int_{R^p_+} \frac{f_{i+1}}{f_i}\, h_0(\boldsymbol{f},\boldsymbol{q})\, t(\theta(\tau))\, \widetilde{M}(\boldsymbol{f},\boldsymbol{q},\tau) \prod_{j=1}^{p} f_j^{\alpha_j} \prod_{s>t} q_{st}^{\beta_{st}} \\ &\qquad \times h_2(\boldsymbol{f},\boldsymbol{\lambda})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}. \end{aligned}$$

Using boundedness of $\widetilde{M}(\boldsymbol{f},\boldsymbol{q},\tau)h_2(\boldsymbol{f},\boldsymbol{\lambda})h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})$ and the inequality $h_0(\boldsymbol{f},\boldsymbol{q})t(\theta(\tau)) \leq \bar{h}_0(\boldsymbol{f},\boldsymbol{q})$, where $\bar{h}_0(\boldsymbol{f},\boldsymbol{q})$ is given in (31) we can easily prove $z^2 I_2 = o(z)$ under the condition $n-p-4>0$.

Now we focus on I_1. We apply to I_1 the following modified versions of Lemma 2 and 3. Using the same notation as these lemmas, we can easily check that if $\alpha_i > 0,\ i=1,\dots,p$, then

$$\begin{aligned}\int_{R^{p'}}\int_{R_+^p} &\eta(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h^*(\boldsymbol{f},\boldsymbol{q};\boldsymbol{\alpha},\boldsymbol{\beta})\, h_3(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q} \\ &= \int_{R^{p'}}\int_{R_+^p} \eta(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h^*(\boldsymbol{f},\boldsymbol{q};\boldsymbol{\alpha},\boldsymbol{\beta})\ d\boldsymbol{f} d\boldsymbol{q} + o(1), \qquad (41)\end{aligned}$$

$$\begin{aligned}\int_{R^{p'}}\int_{R_+^p} &\eta(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h^*(\boldsymbol{f},\boldsymbol{q};\boldsymbol{\alpha},\boldsymbol{\beta})\, h_2(\boldsymbol{f},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q} \\ &= \int_{R^{p'}}\int_{R_+^p} \eta(\boldsymbol{f},\boldsymbol{q},\boldsymbol{\lambda})\, h^*(\boldsymbol{f},\boldsymbol{q};\boldsymbol{\alpha},\boldsymbol{\beta})\ d\boldsymbol{f} d\boldsymbol{q} + o(1). \qquad (42)\end{aligned}$$

Using (41), we have

$$a_i z I_1 = a_i z I_3 + o(z),$$

where

$$I_3 = \int_{R^{p'}}\int_{R_+^p} \frac{f_{i+1}}{f_i}\, h_0(\boldsymbol{f},\boldsymbol{q})\, h_2(\boldsymbol{f},\boldsymbol{\lambda})\, d\boldsymbol{f} d\boldsymbol{q}.$$

Furthermore by (42), we have

$$a_i z I_3 = a_i z I_4 + o(z),$$

where

$$\begin{aligned} I_4 &= \int_{R^{p'}}\int_{R_+^p} \frac{f_{i+1}}{f_i}\, h_0(\boldsymbol{f},\boldsymbol{q})\, d\boldsymbol{f} d\boldsymbol{q} \\ &= E[\chi^2(n-i)]E[(\chi^2(n-i+1))^{-1}] \\ &= \frac{n-i}{n-i-1}. \end{aligned}$$

This completes the proof of (38). ∎

From Theorem 3, we derive two corollaries we will use later.

Corollary 1 *Suppose that $n-p-4>0$. If nonnegative integers $\alpha_i,\ i=1,\dots,p-1$, satisfy $\sum_{i=1}^{p-1}\alpha_i \geq 2$, then*

$$E[\prod_{i=1}^{p-1} x_i^{\alpha_i}] = o(z).$$

Proof.

There exist s,t $(1 \leq s,t \leq p-1)$ such that

$$\prod_{i=1}^{p-1} x_i^{\alpha_i} \leq x_s x_t.$$

Therefore it suffices to prove

$$E[x_i x_j] = o(z), \qquad 1 \le i, j \le p-1.$$

Note that

$$E[x_i x_j] = E[x_i x_j I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] + E[x_i x_j I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})].$$

Since

$$E[x_i x_j I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] \le E[I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})]$$

and the right side is $o(z)$ by Theorem 1,

$$E[x_i x_j I_{\boldsymbol{G}(\bar{U}^*)^C}(\widetilde{\boldsymbol{G}})] = o(z).$$

By (15), we have

$$\begin{aligned} E[x_i x_j I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] &= z^2 a_i a_j \int_{R^{p'}} \int_{R_+^p} \frac{f_{i+1}}{f_i} \frac{f_{j+1}}{f_j} h(\boldsymbol{f}, \boldsymbol{q}, \boldsymbol{\lambda}) \, d\boldsymbol{f} d\boldsymbol{q} \\ &\le z^2 a_i a_j \int_{R^{p'}} \int_{R_+^p} \frac{f_{i+1}}{f_i} \frac{f_{j+1}}{f_j} \bar{h}(\boldsymbol{f}, \boldsymbol{q}) \, d\boldsymbol{f} d\boldsymbol{q}, \end{aligned}$$

where $\bar{h}(\boldsymbol{f}, \boldsymbol{q})$ is given by (19). The last integral converges if $n - p - 2 > 0$. Therefore

$$E[x_i x_j I_{\boldsymbol{G}(\bar{U}^*)}(\widetilde{\boldsymbol{G}})] = o(z).$$

This completes the proof. ∎

Corollary 2 *Suppose that $n - p - 4 > 0$ and nonnegative integers $\alpha_j, \beta_j,\ j = 1, \ldots, p-1$, satisfy $\sum_{j=1}^{p-1} \alpha_j \ge 1, \quad \sum_{j=1}^{p-1} \beta_j \ge 1$, then*

$$E\left[\frac{\prod_{j=1}^{p-1} x_j^{\alpha_j}}{1 - \prod_{j=1}^{p-1} x_j^{\beta_j}}\right] = \begin{cases} a_i z \dfrac{n-i}{n-i-1} + o(z) & \text{if } \alpha_i = 1,\ \alpha_j = 0,\ \forall j \ne i, \\ o(z) & \text{if } \sum_{j=1}^{p-1} \alpha_j \ge 2. \end{cases}$$

Proof.

First consider the case $\alpha_i = 1,\ \alpha_j = 0,\ \forall j \ne i$. Let $y = \prod_{j=1}^{p-1} x_j^{\beta_j}$. Then

$$\frac{x_i}{1 - \prod_{j=1}^{p-1} x_j^{\beta_j}} = x_i \sum_{m=0}^{\infty} y^m = x_i + \lim_{t \to \infty} S_t,$$

where $S_t = x_i \sum_{m=1}^{t} y^m$. By Theorem 3, $E[x_i] = a_i z \frac{n-i}{n-i-1} + o(z)$. Therefore we only have to show

$$\lim_{z \to 0} z^{-1} E[\lim_{t \to \infty} S_t] = 0. \tag{43}$$

Since $0 \le S_1 \le S_2 \le \cdots$, we have, by the monotone convergence theorem,

$$z^{-1}E[\lim_{t\to\infty} S_t] = z^{-1}\lim_{t\to\infty} E[S_t] = \sum_{m=1}^{\infty} a_m(z),$$

where $a_m(z) = E[x_i y^m]z^{-1},\ m \ge 1$. By Fatou's lemma for a counting measure, we have

$$0 \le \sum_{m=1}^{\infty} \liminf_{z\to 0} a_m(z) \le \liminf_{z\to 0} \sum_{m=1}^{\infty} a_m(z) \le \limsup_{z\to 0} \sum_{m=1}^{\infty} a_m(z) \le \sum_{m=1}^{\infty} \limsup_{z\to 0} a_m(z).$$

Since $\sum_{j=1}^{p-1} \beta_j \ge 1$, by Corollary 1, we have

$$\limsup_{z\to 0} a_m(z) = \lim_{z\to 0} a_m(z) = 0.$$

Therefore $\lim_{z\to 0}\sum_{m=1}^{\infty} a_m(z) = 0$. This completes the proof of (43).

For the case $\sum_{j=1}^{p-1} \alpha_j \ge 2$, we can prove the statement similarly. We omit the proof. ∎

Finally we give the result on the expansion of the risk w.r.t. z up to the first order. Here we consider a scale and orthogonally equivariant estimator that satisfies the condition (9), which is a necessary condition for minimaxity. Let the estimator be given by (6) with $c_i(\boldsymbol{x}) = \psi_i(\boldsymbol{l})/l_i,\ i = 1, \ldots, p$. We state three conditions: First two conditions are rather technical, while the last one is the equivalence of (9).

1. $c_i(\boldsymbol{x})$ $(1 \le \forall i \le p)$ is a C^2 function of $\boldsymbol{x}$ on $[0,1]^{p-1}$.
2. There exists some lower bound $\underline{c}$ such that $0 < \underline{c} \le c_i(\boldsymbol{x}),\ 1 \le \forall i \le p$.
3. $c_i(\boldsymbol{0}) = \delta_i^{JS},\ i = 1, \ldots, p$, where δ_i^{JS} is given by (7).

Theorem 4 *Suppose that $n-p-4 > 0$. If a scale and orthogonally equivariant estimator satisfies the above three conditions, then*

$$E[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] = \bar{R} - z\Big[2\sum_{i=1}^{p-1} a_i \frac{n-i}{n-i-1}\{2\delta_i^{JS}\delta_{i+1}^{JS} + \eta_{ii} - \eta_{i+1,i}\}\Big] + o(z),$$

where

$$\eta_{ij} = \frac{\partial c_i}{\partial x_j}(\boldsymbol{0}), \quad 1 \le i \le p, \quad 1 \le j \le p-1,$$

and $\bar{R}$ is the minimax risk.

Proof.

Taylor expansion of $c_i(\boldsymbol{x}),\ i = 1, \ldots, p$, around $\boldsymbol{0}$ is given by

$$c_i(\boldsymbol{x}) = c_i(\boldsymbol{0}) + \sum_{j=1}^{p-1} x_j \eta_{ij} + \frac{1}{2}\sum_{1\le s,t\le p-1} \tilde{c}_{st}^{(i)}(\theta\boldsymbol{x}) x_s x_t, \tag{44}$$

where

$$\tilde{c}_{st}^{(i)}(\boldsymbol{x}) = \frac{\partial^2 c_i}{\partial x_s \partial x_t}(\boldsymbol{x}), \qquad 0 < \theta = \theta(\boldsymbol{x}) < 1.$$

Note that since $c_i(\boldsymbol{x})$ is a C^2 function on $[0,1]^{p-1}$, $\tilde{c}_{st}^{(i)}(\boldsymbol{x})$ is continuous and hence bounded on $[0,1]^{p-1}$. Similarly we have

$$\frac{\partial c_i}{\partial x_i}(\boldsymbol{x}) = \eta_{ii} + \sum_{s=1}^{p-1} \tilde{c}_{is}^{(i)}(\theta\boldsymbol{x})x_s, \qquad 1 \le i \le p-1, \tag{45}$$

$$\frac{\partial c_i}{\partial x_{i-1}}(\boldsymbol{x}) = \eta_{i,i-1} + \sum_{s=1}^{p-1} \tilde{c}_{i-1,s}^{(i)}(\theta\boldsymbol{x})x_s, \qquad 2 \le i \le p, \tag{46}$$

$$\log c_i(\boldsymbol{x}) = \log c_i(0) + \sum_{j=1}^{p-1} \frac{x_j}{c_i(\mathbf{0})}\eta_{ij} + \frac{1}{2}\sum_{1\le s,t\le p-1} \bar{c}_{st}^{(i)}(\theta\boldsymbol{x})x_s x_t, \quad 1 \le i \le p, \tag{47}$$

where

$$\bar{c}_{st}^{(i)}(\boldsymbol{x}) = \frac{\partial}{\partial x_s \partial x_t}\log c_i(\boldsymbol{x})$$

and from the condition 2 this is also bounded. Note that we use notations sparingly, i.e., $\theta = \theta(\boldsymbol{x})$ is different from one equation to another. We use the same abbreviated notation below.

Substituting (44), (45), (46) and (47) into (37), we have

$$\begin{aligned}
&E[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] \\
&= -\sum_{i=1}^{p} \log \delta_i^{JS} - E[\log|\boldsymbol{W}\boldsymbol{\Sigma}^{-1}|] \\
&\quad + E\Big[\sum_{i=1}^{p}(n+p+1-2i)\Big\{\sum_{j=1}^{p-1} x_j\eta_{ij} + \frac{1}{2}\sum_{1\le s,t\le p-1}\bar{c}_{st}^{(i)}(\theta\boldsymbol{x})x_s x_t\Big\} \\
&\qquad + 2\sum_{i<j}\frac{\prod_{m=i}^{j-1} x_m}{1-\prod_{m=i}^{j-1} x_m} \\
&\qquad \times\Big\{(\delta_i^{JS} - \delta_j^{JS}) + \sum_{s=1}^{p-1} x_s(\eta_{is} - \eta_{js}) + \frac{1}{2}\sum_{1\le s,t\le p-1}(\tilde{c}_{st}^{(i)}(\theta\boldsymbol{x}) - \tilde{c}_{st}^{(j)}(\theta\boldsymbol{x}))x_s x_t\Big\} \\
&\qquad - 2\sum_{i=1}^{p-1}\eta_{ii}x_i + 2\sum_{i=2}^{p}\eta_{i,i-1}x_{i-1} - 2\sum_{i=1}^{p-1} x_i \sum_{s=1}^{p-1}\tilde{c}_{is}^{(i)}(\theta\boldsymbol{x})x_s + 2\sum_{i=2}^{p} x_{i-1}\sum_{s=1}^{p-1}\tilde{c}_{i-1,s}^{(i)}(\theta\boldsymbol{x})x_s \\
&\qquad - \sum_{i=1}^{p}\sum_{j=1}^{p-1} x_j(\delta_i^{JS})^{-1}\eta_{ij} - \frac{1}{2}\sum_{i=1}^{p}\sum_{1\le s,t\le p-1}\bar{c}_{st}^{(i)}(\theta\boldsymbol{x})x_s x_t\Big].
\end{aligned}$$

Apply Theorem 3, Corollary 1,2 to the right side. By straightforward calculation, using the boundedness of $\tilde{c}_{st}^{(i)}(\theta\boldsymbol{x})$, $\bar{c}_{st}^{(i)}(\theta\boldsymbol{x})$ $(1 \le i \le p,\ 1 \le s,t \le p-1)$ and the fact

$$\bar{R} = -\sum_{i=1}^{p}\log\delta_i^{JS} - E[\log|\boldsymbol{W}\boldsymbol{\Sigma}^{-1}|]$$

(see James and Stein (1961)), we have

$$E[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] = \bar{R} - z\Big[2\sum_{i=1}^{p-1} a_i \frac{n-i}{n-i-1}\{2\delta_i^{JS}\delta_{i+1}^{JS} + \eta_{ii} - \eta_{i+1,i}\}\Big] + o(z).$$

∎

The conditions 1–3 of Theorem 4 can be changed as follows.

1* $c_i(\boldsymbol{x})$ $(1 \le \forall i \le p)$ is a C^2 function of $\boldsymbol{x}$ on some open neighborhood of $\mathbf{0}$, say $\mathcal{N}(\mathbf{0})$, in $[0,1]^{p-1}$.

2* There exists some bounds such that $0 < \underline{c} < c_i(\boldsymbol{x}) < \bar{c}$, $1 \le \forall i \le p$.

3* $c_i(\mathbf{0}) = \delta_i^{JS}, \qquad i = 1, \dots, p.$

Corollary 3 *Suppose that $n-p-4>0$. If a scale and orthogonally equivariant estimator satisfies the above three conditions, the result of Theorem 4 holds.*

Proof.

Since the risk of an orthogonally equivariant estimator depends on $\boldsymbol{\Sigma}$ only through $\boldsymbol{\Lambda}$, we can assume $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}$ without loss of generality.

First we prove that any scale and orthogonally invariant estimator that satisfies the condition 2* has the following property;

$$E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Lambda}) I(\boldsymbol{x} \in \widetilde{\mathcal{N}}(\mathbf{0})^C)] = o(z) \tag{48}$$

for any open neighborhood $\widetilde{\mathcal{N}}(\mathbf{0})$ of $\mathbf{0}$ in $[0,1]^{p-1}$.

Let

$$L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Lambda}) = \widetilde{L}(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Lambda}) - \log|\boldsymbol{\Lambda}^{-1}\boldsymbol{W}|,$$

where

$$\widetilde{L}(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Lambda}) = \mathrm{tr}(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Lambda}^{-1}) - \sum_{i=1}^{p} \log c_i(\boldsymbol{x}) - p.$$

Since

$$\begin{aligned}
|\widetilde{L}(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Lambda})| &\le \sum_{i=1}^{p}\sum_{j=1}^{p} g_{ij}^2 c_j(\boldsymbol{x}) l_j \lambda_i^{-1} + \sum_{i=1}^{p}\Big|\log c_i(\boldsymbol{x})\Big| + p \\
&\le \bar{c}\sum_{i=1}^{p}\sum_{j=1}^{p} g_{ij}^2 l_j \lambda_i^{-1} + p\max\Big(\Big|\log \bar{c}\Big|, \Big|\log \underline{c}\Big|\Big) + p \\
&= \bar{c}\,\mathrm{tr}(\boldsymbol{W}\boldsymbol{\Lambda}^{-1}) + p\max\Big(\Big|\log \bar{c}\Big|, \Big|\log \underline{c}\Big|\Big) + p,
\end{aligned}$$

there exist α, M such that

$$0 < \alpha < \frac{1}{2}, \qquad M > 0, \qquad M\exp\Big(\mathrm{tr}\,\alpha\boldsymbol{W}\boldsymbol{\Lambda}^{-1}\Big) > \Big|\widetilde{L}(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Lambda})\Big|.$$

Therefore

$$
\begin{aligned}
\Big|E_{\boldsymbol{\Lambda}}[\widetilde{L}(\widehat{\boldsymbol{\Sigma}},\boldsymbol{\Lambda})I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})]\Big| &< ME_{\boldsymbol{\Lambda}}[\exp\Big(\operatorname{tr}\alpha\boldsymbol{W}\boldsymbol{\Lambda}^{-1}\Big)I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})] \\
&= \frac{M}{2^{np/2}\Gamma_p(n/2)}|\boldsymbol{\Lambda}|^{-n/2} \\
&\quad \times\int_{W>0}|\boldsymbol{W}|^{(n-p-1)/2}\exp\Big(-\operatorname{tr}\widetilde{\alpha}\boldsymbol{W}\boldsymbol{\Lambda}^{-1}\Big)I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})d\boldsymbol{W}, \quad \widetilde{\alpha}=\frac{1}{2}-\alpha \\
&= \widetilde{\alpha}^{-np/2}ME_{\boldsymbol{\Lambda}}[I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})].
\end{aligned}
$$

By Lemma 4, $E_{\boldsymbol{\Lambda}}[I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})]=o(z)$. Therefore

$$
E_{\boldsymbol{\Lambda}}[\widetilde{L}(\widehat{\boldsymbol{\Sigma}},\boldsymbol{\Lambda})I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})]=o(z). \tag{49}
$$

We can also prove

$$
E_{\boldsymbol{\Lambda}}[\log|\boldsymbol{\Lambda}^{-1}\boldsymbol{W}|\;I_{\widetilde{\mathcal{N}}(\mathbf{0})^C}(\boldsymbol{x})]=o(z) \tag{50}
$$

similarly to the proof of Lemma 4. We briefly describe the proof. First note it suffices to give a proof for the case when

$$
\widetilde{\mathcal{N}}(\mathbf{0})=[0,\epsilon)^{p-1}
$$

with arbitrarily small positive number ϵ. By Lemma 5, the following inequality holds.

$$
\begin{aligned}
&E_{\boldsymbol{\Lambda}}\Big[\Big|\log|\boldsymbol{\Lambda}^{-1}\boldsymbol{W}|\Big|I\Big(\boldsymbol{x}(\boldsymbol{W})\in\widetilde{\mathcal{N}}(\mathbf{0})^C\Big)\Big] \\
&\le z\epsilon^{-1}\sum_{i=1}^{p-1}a_iE_{\boldsymbol{\Lambda}}\Big[\Big|\log|\boldsymbol{\Lambda}^{-1}\boldsymbol{W}|\Big|\Big(\sum_{j=i+1}^{p}\widetilde{w}_{jj}\Big)\Big(\sum_{j=1}^{i}\widetilde{w}^{jj}\Big)I\Big(x_i(\boldsymbol{W})\ge\epsilon\Big)\Big] \\
&= z\epsilon^{-1}\sum_{i=1}^{p-1}a_iE_I\Big[\Big|\log|\boldsymbol{W}|\Big|\Big(\sum_{j=i+1}^{p}w_{jj}\Big)\Big(\sum_{j=1}^{i}w^{jj}\Big)I\Big(x_i(\boldsymbol{\Lambda}^{1/2}\boldsymbol{W}\boldsymbol{\Lambda}^{1/2})\ge\epsilon\Big)\Big].
\end{aligned}
$$

The condition $n-p-4>0$ is sufficient for the convergence of

$$
E_I\Big[\Big|\log|\boldsymbol{W}|\Big|\Big(\sum_{j=i+1}^{p}w_{jj}\Big)\Big(\sum_{j=1}^{i}w^{jj}\Big)\Big].
$$

Therefore by Lemma 6 and the dominated convergence theorem,

$$
\lim_{z\to0}E_I\Big[\Big|\log|\boldsymbol{W}|\Big|\Big(\sum_{j=i+1}^{p}w_{jj}\Big)\Big(\sum_{j=1}^{i}w^{jj}\Big)I\Big(x_i(\boldsymbol{\Lambda}^{1/2}\boldsymbol{W}\boldsymbol{\Lambda}^{1/2})\ge\epsilon\Big)\Big]=0.
$$

This completes the proof of (50). From (49) and (50), it is obvious that (48) holds true.

We can construct another scale and orthogonally equivariant estimator $\widehat{\boldsymbol{\Sigma}}^*$,

$$
\widehat{\boldsymbol{\Sigma}}^*=\boldsymbol{G}\boldsymbol{\Psi}^*\boldsymbol{G}', \qquad \boldsymbol{\Psi}^*=\operatorname{diag}(\psi_1^*(\boldsymbol{l}),\ldots,\psi_p^*(\boldsymbol{l})), \qquad c_i^*(\boldsymbol{x})=\frac{\psi_i^*(\boldsymbol{l})}{l_i}, \quad 1\le i\le p
$$

such that $c_i^*(\boldsymbol{x})$ $(1 \le i \le p)$ satisfy the conditions $1,\ 2^*,\ 3^*$ and on some open neighborhood of $\mathbf{0}$ (say $\mathcal{N}^*$), $c_i^*(\boldsymbol{x}) = c_i(\boldsymbol{x}),\ 1 \le i \le p$. Then it follows that by (48) and Theorem 4

$$\begin{aligned} E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] &= E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) I_{\mathcal{N}^*}(\boldsymbol{x})] + o(z) \\ &= E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}^*, \boldsymbol{\Sigma}) I_{\mathcal{N}^*}(\boldsymbol{x})] + o(z) \\ &= E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}^*, \boldsymbol{\Sigma})] + o(z) \\ &= \bar{R} - z\Big[2\sum_{i=1}^{p-1} a_i \frac{n-i}{n-i-1}\{2\delta_i^{JS}\delta_{i+1}^{JS} + \eta_{ii} - \eta_{i+1,i}\}\Big] + o(z). \end{aligned}$$

∎

Corollary 3 still requires not only the local conditions (1^* and 3^*) but also the global condition (2^*). Next corollary gives necessary conditions for an estimator to be tail minimax. It is relatively easy to apply this corollary to a particular estimator, since it requires only local conditions.

Corollary 4 *Suppose that $n - p - 4 > 0$. If a scale and orthogonally equivariant estimator that satisfies the condition 1^* is tail minimax, then*

$$c_i(\mathbf{0}) = \delta_i^{JS}, \qquad 1 \le i \le p, \tag{51}$$

and

$$2\delta_i^{JS}\delta_{i+1}^{JS} + \eta_{ii} - \eta_{i+1,i} \ge 0, \qquad 1 \le i \le p-1. \tag{52}$$

Proof.

(51) is due to Theorem 4 of Takemura and Sheena (2002). Now suppose (51) is satisfied but (52) is violated. Then there exist some positive numbers, $a_i,\ i = 1, \dots p-1$, such that

$$\sum_{i=1}^{p-1} a_i \frac{n-i}{n-i-1}\{2\delta_i^{JS}\delta_{i+1}^{JS} + \eta_{ii} - \eta_{i+1,i}\} < 0.$$

We consider the linear convergence of $\|y\|$ to $\mathbf{0}$ with these a_i's. Note that we can construct another scale and orthogonally equivariant estimator $\widehat{\boldsymbol{\Sigma}}^*$,

$$\widehat{\boldsymbol{\Sigma}}^* = \boldsymbol{G}\boldsymbol{\Psi}^*\boldsymbol{G}', \qquad \boldsymbol{\Psi}^* = \mathrm{diag}(\psi_1^*(\boldsymbol{l}), \dots, \psi_p^*(\boldsymbol{l})), \qquad c_i^*(\boldsymbol{x}) = \psi_i^*(\boldsymbol{l})/l_i, \quad 1 \le i \le p$$

such that $c_i^*(\boldsymbol{x})$ $(1 \le i \le p)$ satisfy the conditions 1^*–3^* and on some open neighborhood of $\mathbf{0}$ (say $\mathcal{N}^*$), $c_i^*(\boldsymbol{x}) = c_i(\boldsymbol{x}),\ 1 \le i \le p$. Then the following inequality holds

$$\begin{aligned} E_{\boldsymbol{\Sigma}}[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] &= E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma})] \\ &\ge E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) I(\boldsymbol{x} \in \mathcal{N}^*)] \\ &= E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}^*, \boldsymbol{\Sigma}) I(\boldsymbol{x} \in \mathcal{N}^*)] \\ &= E_{\boldsymbol{\Lambda}}[L_1(\widehat{\boldsymbol{\Sigma}}^*, \boldsymbol{\Sigma})] + o(z) \qquad \text{(by (48))} \\ &= \bar{R} - z\Big[2\sum_{i=1}^{p-1} a_i \frac{n-i}{n-i-1}\{2\delta_i^{JS}\delta_{i+1}^{JS} + \eta_{ii} - \eta_{i+1,i}\}\Big] + o(z). \end{aligned}$$

The right side is strictly larger than $\bar{R}$ for small enough z. ∎

Now we apply Corollary 4 to Stein's and Haff's estimators. Both of them are known to perform well numerically. (see Lin and Perlman (1985), Sugiura and Ishibayashi (1997)), but their minimaxity has been an open problem.

Stein's estimator $\widehat{\boldsymbol{\Sigma}}^S$ is an isotonized modification of the rough estimator given by (10). The algorithm is depicted in Lin and Perlman (1985). After the modification ψ_i $(1 \le i \le p)$ is given as the ratio of pooled adjacent l_i's and c_i^{-1}'s of the rough estimator. Since the pooling depends on the value of c_i's, it gets complicated especially when p is large. However for the application of Corollary 4, we only have to know its behavior in a neighborhood of the point $\boldsymbol{x} = \mathbf{0}$. When l_i's are fully dispersed, that is, $\|\boldsymbol{x}\|$ is small enough, the ψ_i's in (10) keep the order (11), hence do not need any modification. Then c_i's of $\widehat{\boldsymbol{\Sigma}}^S$ are given by

$$c_i(\boldsymbol{x}) = \Big[(n-p+1) + 2\sum_{i<j}\Big(\frac{1}{1-\prod_{m=i}^{j-1} x_m}\Big) + 2\sum_{i>j}\Big(1 - \frac{1}{1-\prod_{m=j}^{i-1} x_m}\Big)\Big]^{-1}, \qquad 1 \le i \le p,$$

in a small enough neighborhood of $\mathbf{0}$. It is obvious that this $c_i(\boldsymbol{x})$ satisfies the condition 1^*. Besides we can easily check

$$c_i(\mathbf{0}) = \delta_i^{JS}, \qquad 1 \le i \le p,$$

and

$$\eta_{ii} = \frac{\partial c_i}{\partial x_i}(\mathbf{0}) = -2(\delta_i^{JS})^2, \qquad \eta_{i+1,i} = \frac{\partial c_{i+1}}{\partial x_i}(\mathbf{0}) = 2(\delta_{i+1}^{JS})^2, \qquad 1 \le i \le p-1.$$

Consequently the left side of (52) equals to

$$2\Big\{\delta_i^{JS}\delta_{i+1}^{JS} - (\delta_i^{JS})^2 - (\delta_{i+1}^{JS})^2\Big\} = -2\Big\{\Big(\delta_i^{JS} - \delta_{i+1}^{JS}\Big)^2 + \delta_i^{JS}\delta_{i+1}^{JS}\Big\}. \tag{53}$$

The right side is obviously negative. Therefore from Corollary 4, Stein's estimator, $\widehat{\boldsymbol{\Sigma}}^S$ is not tail minimax, hence not minimax when $n-p-4 > 0$. More strictly speaking, since (53) is negative for $1 \le \forall i \le p-1$, we can say for any linear convergence of $\|\boldsymbol{y}\|$, the risk of $\widehat{\boldsymbol{\Sigma}}^S$ is over $\bar{R}$ for small enough z.

The same argument holds for $\widehat{\boldsymbol{\Sigma}}^H$. Haff's estimator is derived through a formal Bayes rule, but in the end it emerges as another type of modification of Stein's rough estimator (10) under the constraint (11). If the rough estimator already keeps the order (11), it is completely the same as $\widehat{\boldsymbol{\Sigma}}^S$. Therefore the above argument for $c_i(\boldsymbol{x})$ on a small enough neighborhood of $\mathbf{0}$ still holds good. Consequently $\widehat{\boldsymbol{\Sigma}}^H$ is not minimax either. We summarize the results in the following corollary.

Corollary 5 *Stein's estimator $\widehat{\boldsymbol{\Sigma}}^S$ and Haff's estimator $\widehat{\boldsymbol{\Sigma}}^H$ are not minimax.*

Another interesting estimator of which minimaxity or non-minimaxity has not been proved theoretically is the reference prior Bayes estimator derived by Yang and Berger (1994). This estimator is given as $\{E[\boldsymbol{\Sigma}^{-1}]\}^{-1}$, where the expectation is taken with respect to the posterior density of $\boldsymbol{\Sigma}$

$$c|\boldsymbol{\Sigma}|^{-(n/2+1)}\Big\{\prod_{i<j}(\lambda_i - \lambda_j)\Big\}^{-1}\exp\Big(-\frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\Big)$$

with the normalizing constant c. Though Sugiura and Ishibayashi (1997) gives some concrete expressions of this estimator using Legendre polynomials or hypergeometric functions when $p = 2$, it seems difficult to give explicit forms for general p. According to the simulation study by Sugiura and Ishibayashi (1997), its risk approaches to the minimax risk $\bar{R}$ from upper side as $\|y\| \to 0$ like $\widehat{\boldsymbol{\Sigma}}^S$ and $\widehat{\boldsymbol{\Sigma}}^H$. This indicates that the reference prior estimator is not minimax either. To prove this using Theorem 4 or its corollary, we need to evaluate $c_i(\mathbf{0})$ and η_{ij}. For this purpose, the methods of Takemura and Sheena (2002) might be useful, since the form of the posterior density resembles that of Wishart distribution.

5 Appendix

The next result was essentially proved in the proof of Lemma 1 of Takemura and Sheena (2002). We state it as an independent lemma and give a proof.

Lemma 5 *The following inequalities hold for every fixed* $\boldsymbol{W} = (w_{ij}) > 0$.

$$f_i = l_i \lambda_i^{-1} \le \sum_{j=i}^{p} \tilde{w}_{jj}$$

$$f_i^{-1} = l_i^{-1} \lambda_i \le \sum_{j=1}^{i} \tilde{w}^{jj},$$

where

$$(\tilde{w}_{ij}) = \widetilde{\boldsymbol{W}} = \boldsymbol{\Lambda}^{-\frac{1}{2}} \boldsymbol{W} \boldsymbol{\Lambda}^{-\frac{1}{2}}, \qquad (\tilde{w}^{ij}) = \widetilde{\boldsymbol{W}}^{-1} = \boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{W}^{-1} \boldsymbol{\Lambda}^{\frac{1}{2}}.$$

Proof.

Let $\boldsymbol{W}_{(i)}$ be the $(p-i+1) \times (p-i+1)$ submatrix of $\boldsymbol{W}$ made by deleting the first $i-1$ rows and columns. Let $\bar{l}_{(i)}$ denote the largest eigenvalue of $\boldsymbol{W}_{(i)}$. Then using a theorem on the magnitude of the eigenvalues of a submatrix (see, e.g. (6) in p227 of Marshal and Olkin (1979)), $l_i \le \bar{l}_{(i)}$. Therefore

$$l_i \lambda_i^{-1} \le \bar{l}_{(i)} \lambda_i^{-1} \le \sum_{j=1}^{p-i+1} (\boldsymbol{W}_{(i)})_{jj} \lambda_i^{-1} \le \sum_{j=1}^{p-i+1} (\boldsymbol{W}_{(i)})_{jj} \lambda_{i+j-1}^{-1} = \sum_{j=i}^{p} \tilde{w}_{jj}.$$

Now let $\boldsymbol{W}^{(i)}$ be the $i \times i$ submatrix of $\boldsymbol{W}^{-1}$ composed of the first i rows and columns, and $\bar{l}^{(i)}$ be its largest eigenvalue. Then we have $l_i^{-1} \le \bar{l}^{(i)}$ and

$$l_i^{-1} \lambda_i \le \bar{l}^{(i)} \lambda_i \le \sum_{j=1}^{i} (\boldsymbol{W}^{(i)})_{jj} \lambda_i \le \sum_{j=1}^{i} (\boldsymbol{W}^{(i)})_{jj} \lambda_j = \sum_{j=1}^{i} \tilde{w}^{jj}.$$

∎

Let $\boldsymbol{L}(\boldsymbol{W}) = \mathrm{diag}(l_1(\boldsymbol{W}), \ldots, l_p(\boldsymbol{W}))$ be the functions of $\boldsymbol{W}$ defined by the spectral decomposition of $\boldsymbol{W}$ in (1).

Lemma 6 *For every fixed $\boldsymbol{W} = (w_{ij}) > 0$, it holds that*

$$\lim_{\|\boldsymbol{y}\|\to 0} \frac{l_{i+1}(\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{W}\boldsymbol{\Lambda}^{\frac{1}{2}})}{l_i(\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{W}\boldsymbol{\Lambda}^{\frac{1}{2}})} = 0, \qquad 1 \le i \le p-1.$$

Proof.

We define the notations;

$\boldsymbol{A} = (a_{ij}) = \boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{W}\boldsymbol{\Lambda}^{\frac{1}{2}}$,

$\boldsymbol{A}(i_1,\dots,i_s)$: principle submatrix composed of the i_jth row and column, $j = 1,\dots,s$,

$l_1(\boldsymbol{A}(i_1,\dots,i_s)) \ge \cdots \ge l_s(\boldsymbol{A}(i_1,\dots,i_s))$: the ordered eigenvalues of $\boldsymbol{A}(i_1,\dots,i_s)$.

Suppose that there exist some positive numbers $\gamma_s, M_{s1}, M_{s2},\ s = 1,\dots,p$, such that

$$\|\boldsymbol{y}\| < \gamma_s \ \Rightarrow\ M_{s1} > \frac{l_s(\boldsymbol{A})}{a_{ss}} > M_{s2}. \tag{54}$$

Then for $\|\boldsymbol{y}\| < \min(\gamma_{s+1}, \gamma_s)$

$$\frac{l_{s+1}(\boldsymbol{A})}{l_s(\boldsymbol{A})} = \frac{l_{s+1}(\boldsymbol{A})}{a_{s+1,s+1}}\frac{a_{ss}}{l_s(\boldsymbol{A})}\frac{a_{s+1,s+1}}{a_{ss}} \le M_{s+1,1}M_{s2}^{-1}w_{s+1,s+1}w_{ss}^{-1}\lambda_{s+1}\lambda_s^{-1}.$$

This implies $l_{s+1}(\boldsymbol{A})/l_s(\boldsymbol{A}) \to 0$ as $\|\boldsymbol{y}\| \to 0$. Therefore we only have to prove the existence of γ_s, M_{s1}, M_{s2} $(1 \le s \le p)$ that satisfy (54). We prove it inductively with respect to p. If $p = 1$, it is obvious since $l_1(\boldsymbol{A})/a_{11} = 1$. Suppose the existence of γ_s, M_{s1}, M_{s2} $(1 \le s \le p)$ satisfying (54) when $p = k-1$. Now we consider the case $p = k$. Using the formula on the eigenvalues of a submatrix, we have

$$\begin{aligned} l_1(\boldsymbol{A}) \ge l_1(\boldsymbol{A}(1,\dots,k-1)) \ge l_2(\boldsymbol{A}) \ge l_2(\boldsymbol{A}(1,\dots,k-1)) \ge \cdots \\ \ge l_{k-1}(\boldsymbol{A}(1,\dots,k-1)) \ge l_k(\boldsymbol{A}), \end{aligned}$$

$$l_1(\boldsymbol{A}) \ge l_1(\boldsymbol{A}(2,\dots,k)) \ge l_2(\boldsymbol{A}) \ge l_2(\boldsymbol{A}(2,\dots,k)) \ge \cdots \ge l_{k-1}(\boldsymbol{A}(2,\dots,k)) \ge l_k(\boldsymbol{A}).$$

Therefore for $2 \le s \le k-1$,

$$\frac{l_{s-1}(\boldsymbol{A}(2,\dots,k))}{a_{ss}} \ge \frac{l_s(\boldsymbol{A})}{a_{ss}} \ge \frac{l_s(\boldsymbol{A}(1,\dots,k-1))}{a_{ss}}.$$

By the assumption for the case $p = k-1$, there exist γ, M_1, M_2 such that

$$\|\boldsymbol{y}\| < \gamma \ \Rightarrow\ M_1 > l_{s-1}(\boldsymbol{A}(2,\dots,k))a_{ss}^{-1}, \quad l_s(\boldsymbol{A}(1,\dots,k-1))a_{ss}^{-1} > M_2.$$

Combining these inequalities, we have

$$\|\boldsymbol{y}\| < \gamma \ \Rightarrow\ M_1 > l_s(\boldsymbol{A})a_{ss}^{-1} > M_2.$$

This completes the proof for the existence of γ_s, M_{s1}, M_{s2} when $s = 2,\dots,k-1$.

Now we consider the cases $s = 1$ and $s = k = p$. Since

$$a_{11}^{-1}\boldsymbol{A} \to \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0}' & \mathbf{0} \end{pmatrix}$$

as $\|\boldsymbol{y}\| \to 0$,

$$\lim_{\|\boldsymbol{y}\|\to 0} \frac{l_1(\boldsymbol{A})}{a_{11}} = l_1\Big(\lim_{\|\boldsymbol{y}\|\to 0} a_{11}{}^{-1}\boldsymbol{A}\Big) = 1.$$

This guarantees the existence of the required γ_1, M_{11}, M_{12}.

Note that

$$l_p^{-1}(\boldsymbol{A})a_{pp} = l_1(\boldsymbol{A}^{-1})a_{pp} = l_1(\lambda_p\boldsymbol{A}^{-1})w_{pp}.$$

Since as $\|\boldsymbol{y}\| \to 0$,

$$\lambda_p\boldsymbol{A}^{-1} \to \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0}' & w^{pp} \end{pmatrix}, \qquad w^{pp} = (\boldsymbol{W}^{-1})_{pp},$$

$l_1(\lambda_p\boldsymbol{A}^{-1}) \to w^{pp}$. Therefore

$$\lim_{\|\boldsymbol{y}\|\to 0} \frac{l_p(\boldsymbol{A})}{a_{pp}} = \frac{1}{w_{pp}w^{pp}}.$$

This guarantees the existence of the required γ_p, M_{p1}, M_{p2}. ∎

Acknowledgements

The first author is deeply grateful to the Department of Mathematics and Statistics, Bowling Green State University for providing comfortable research facilities.

References

[1] Berger, O. B. (1976). Tail minimaxity in location vector problems and its applications. *Ann. Statist.*, **4**, 33–50.

[2] Dey, D. K. and Srinivasan, C. (1985). Estimation of a covariance matrix under Stein's loss. *Ann. Statist.*, **13**, 1581–1591.

[3] Haff, L. R. (1979). An identity for the Wishart distribution with applications. *Journal of Multivariate Analysis*, **9**, 531–544.

[4] Haff, L. R. (1991). The variational form of certain Bayes estimators. *Ann. Statisit.*, **19**, 1163–1190.

[5] James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, **1**, 361–380. Univ. of California Press.

[6] Krishnamoorthy, K. and Gupta, A. K. (1989). Improved minimax estimation of a normal precision matrix. *Canad. J. Statist.*, **17**, 91–102.

[7] Kubokawa, T. and Srivastava, M. S. (1999). Robust improvement in estimation of a covariance matrix in an elliptically contoured distribution. *Ann. Statist.*, **27**, 600–609.

[8] Lin, S. P. and Perlman, M. D. (1985). A Monte Carlo comparison of four estimators for a covariance matrix. *Multivariate Analysis*, **6**, 411–429, ed. P. R. Krishnaiah, North Holland, Amsterdam.

[9] Marshal, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, San Diego.

[10] Olkin, I. and Selliah, J. B. (1977). Estimating covariances in a multivariate normal distribution. *Statistical Decision Theory and Related Topics*, **2**, 313-326, ed. Gupta, S. S. and Moore, D., Academic Press, New York.

[11] Sharma, D. and Krishnamoorthy, K. (1983). Orthogonal equivariant minimax estimators of bivariate normal covariance matrix and precision matrix. *Calcutta Statist. Ass. Bull.*, **32**, 23–45.

[12] Sheena, Y. (1995). Unbiased estimator of risk for an orthogonally invariant estimator of a covariance matrix. *J. Japan Statist. Soc.*, **25**, 35–48.

[13] Sheena, Y. (2002). On minimaxity of some orthogonally invariant estimators of bivariate normal dispersion matrix. *J. Japan Statist. Soc.*, **32**, 193–207.

[14] Sheena, Y. and Takemura, A. (1992). Inadmissibility of non-order-preserving orthogonally invariant estimators of the covariance matrix in the case of Stein's loss. *Journal of Multivariate Analysis*, **41**, 117–131.

[15] Stein, C. (1975). Estimation of a covariance matrix. Rietz Lecture, 1975 Annual Meeting of American Statistical Association, Atlanta Ga.

[16] Stein, C. (1977). Estimating the covariance matrix. Unpublished manuscript.

[17] Sugiura, N. and Ishibayashi, H. (1997). Reference prior Bayes estimator for bivariate normal covariance matrix with risk comparison. *Comm. Statist.–Theory Meth.*, **26**, 2203-2221.

[18] Takemura, A. and Sheena, Y. (2002). Distribution of eigenvalues and eigenvectors of Wishart matrix when the population eigenvectors are infinitely dispersed. *Mathematical Engineering Section Technical Report 2002-13.* Tokyo University.

[19] Yang, R. and Berger, J. O. (1994). Estimation of a covariance matrix using the reference prior. *Ann. Statisit.*, **22**, 1195-1211.