Shrinkage Estimators for Large Covariance matrices in multivariate real and complex normal distributions under an invariant quadratic loss

Yoshihiko Konno *

Faculty of Sciences, Japan Women's University

April 2, 2008

Abstract

The problem of estimating large covariance matrices of multivariate real normal and complex normal distributions is considered when the dimension of the variables is larger than the number of sample size. The Stein-Haff identities and calculus on eigenstructures for singular Wishart matrices are developed for real and complex cases, respectively. By using these techniques, the unbiased risk estimates for certain class of estimators for the population covariance matrices under an invariant quadratic loss functions are obtained for real and complex cases, respectively. Based on the unbiased risk estimates, shrinkage estimators which are counterparts of the estimators due to Haff [1980, Ann. Statist. **8**

2000 *Mathematics Subject Classification.* Primary: 62H12, Secondary: 62F10.

Key words and phrases. Unbiased estimate of risk, integration by parts formula, singular Wishart distributions, Stein-Haff identity, calculus on eigenstructures.

586-697] are shown to improve upon the best scalar multiple of the empirical covariance matrix under the invariant quadratic loss functions for both real and complex multivariate normal distributions in the situation such that the dimension of the variables is larger than the number of sample size.

1 Introduction

Estimating a population covariance matrix is an important and difficult problem in the theory of the multivariate statistical analysis [29, 35]. It has been known that the empirical covariance matrix has unexpected feature for a finite sample situation, that is, the eigenvalues of the empirical covariance matrix spread out more than those of the population covariance matrix. Since James and Stein [16], there has been a lot of literature to explore better estimators for a population covariance matrix from a decision-theoretic perspective [7, 12, 13, 14, 24, 37] or from a Bayesian point of views [3, 4, 5, 42] in order to overcome the shortcoming of the empirical covariance matrix. Recently there has been an upsurge of investigation for procedures of the population large covariance matrix in the situation such that the dimension of variables, p, is larger than the number of observations, n. See [2, 15, 23].

In this article we consider the problem of estimating large covariance matrices in a decision-theoretic manner when the dimension of variables, p, is larger than the number of observations, n. Population distributions include not only real multivariate distributions but also complex multivariate distributions. We provide estimators that are better than the best scalar multiple of the empirical covariance matrix under an invariant quadratic loss function. Our approach to derive new estimators is the so-called 'unbiased risk estimate method' and calculus on the eigenstructures for singular Wishart matrices. Both methods for full-rank Wishart matrices have been well-established. See [6, 10, 11, 14, 37, 39] for the Stein-Haff identities for full-rank Wishart matrices and see [13, 14, 19, 21, 22, 25, 26, 27, 36, 37, 39] for calculus on eigenstructures for full-rank Wishart matrices. We extensively develop the Stein-Haff identities and calculus on eigenstructures for singular Wishart matrices, i.e., in the situation such that $p > n$, in order to obtain unbiased risk estimate for certain class of estimators which are analogues of estimators due to [12] for population covariance matrix in the situation such that $n > p$.

This paper is organized in the following way: In Section 2, situation for real singular Wishart matrices is considered. In Section 2.1, we derive integration by parts formula for singular real Wishart matrices in a matrix form. In Section 2.2, using calculus on eigenstructures for singular real Wishart matrices, we obtain unbiased risk estimate for certain class of estimators under an invariant loss function. In Section 2.3, we derive shrinkage estimators which are analogues of estimators due to Haff [12]. In Section 3, parallel results for singular complex Wishart matrices are explored. In Section 4, the proof of Theorems in Sections 2 and 3 are given.

2 Real case

Assume that $n < p$ and let $\boldsymbol{X}$ be an $n \times p$ random matrix having the multivariate real normal distribution $N_{n\times p}(\mathbf{0}_{n\times p}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a $p \times p$ positive-definite matrix. So the rows of the matrix $\boldsymbol{X}$ are mutually independent and have $p-$dimensional normal distribution with zero-mean vector and the covariance matrix $\boldsymbol{\Sigma}$. Put $\boldsymbol{S} = \boldsymbol{X}'\boldsymbol{X}$. Then $\boldsymbol{S}$ has a real Wishart distribution with the dimension p, the degrees of freedom n, and the scale parameter $\boldsymbol{\Sigma}$. We call $\boldsymbol{S}$ a singular real Wishart matrix. See Srivastava [34] for the density function of a partial block of singular real Wishart matrix with respect to Lebesgue measure.

2.1 The Stein-Haff identities and calculus on eigenstructures for singular real Wishart matrices

The Stein-Haff identity for singular real Wishart matrices was first established by Kubokawa and Srivastava [22]. Their derivation was based on the approach due to Sheena [32]. In this subsection, the Stein-Haff identity for singular Wishart matrices by [22] is generalized to a matrix form of the identity via a modification of an approach by [37].

To state our identity, let $\nabla_{\boldsymbol{X}} = (\partial/\partial x_{ij})_{\substack{i=1,2,\ldots,n\\ j=1,2,\ldots,p}}$ for $\boldsymbol{X} = (x_{ij})_{\substack{i=1,2,\ldots,n\\ j=1,2,\ldots,p}}$. For real numbers $a_1, a_2, \ldots, a_n$, we denote by $\mathbf{Diag}(a_1, a_2, \ldots, a_n)$ an $n \times n$ diagonal matrix with diagonal elements $a_1, a_2, \ldots, a_n$. Furthermore, set $\mathbb{R}^p_{\geq} = \{(a_1, a_2, \ldots, a_n) \in \mathbb{R}^p;\ a_1 \geq a_2 \geq \cdots \geq a_n > 0\}$.

Theorem 2.1. *Assume that an* $n \times p$ *real matrix* $\boldsymbol{X}$ *is distributed according to* $N_{n\times p}(\mathbf{0}_{n\times p}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$ *with a* $p \times p$ *positive-definite matrix* $\boldsymbol{\Sigma}$*. Assume that, for a* $p \times p$ *real random matrix* $\boldsymbol{G} := \boldsymbol{G}(\boldsymbol{S}) = (g_{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}}$*, each* g_{ij} *is a differentiable function of* $\boldsymbol{S} = \boldsymbol{X}'\boldsymbol{X}$ *and satisfies the following conditions;*

$$\mathbb{E}\left[|x^2_{j_1 i_1} g_{i_2 i_3}|\right] < \infty, \qquad \mathbb{E}\left[\left|x_{j_1 i_1} \frac{\partial g_{i_2 i_3}}{\partial x_{j_2 i_4}}\right|\right] < \infty$$

for $i_1, \ldots, i_4 = 1, 2, \ldots, p$ *and* $j_1, j_2 = 1, 2, \ldots, n$*. Then we have*

$$\mathbb{E}\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\boldsymbol{G}\right] = \mathbb{E}\left[n\boldsymbol{G} + (\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{G}\right], \tag{2.1}$$

where the superscript "′" stands for the transpose of a matrix. In particular,

$$\mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\boldsymbol{G})\right] = \mathbb{E}\left[n\mathrm{Tr}\,(\boldsymbol{G}) + \mathrm{Tr}\,(\boldsymbol{X}'\nabla_{\boldsymbol{X}}\boldsymbol{G}')\right].$$

The identity (2.1) appeared in the proof of the Wishart identity for nonsingular Wishart matrix in Loh [25]. Note that the identity (2.1) involves in a differential operator related to the multivariate normal random matrix $\boldsymbol{X}$ rather than an operator related to singular Wishart matrices. This is an ingredient to develop the Stein-Haff identity for singular Wishart matrices. Combining Theorem 2.1 with calculus on eigenstructures for the singular Wishart matrices in terms of the differential operator $\nabla_{\boldsymbol{X}}$, we give a matrix form of the Stein-Haff identity below. Another ingredient, i.e., calculus on eigenstructures for singular real Wishart matrices, is developed in Section 4.

Theorem 2.2. *Assume that $n < p$ and that an $n \times p$ real matrix $\boldsymbol{X}$ is distributed according to $N_{n\times p}(\boldsymbol{0}_{n\times p}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$ with a $p \times p$ positive-definite matrix $\boldsymbol{\Sigma}$. Decompose $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{O}_1\boldsymbol{L}\boldsymbol{O}_1'$, where $\boldsymbol{O}_1$ is a $p \times n$ semi-orthogonal matrix such that $\boldsymbol{O}_1'\boldsymbol{O}_1 = \boldsymbol{I}_n$. Let $\boldsymbol{\Psi} := \boldsymbol{\Psi}(\boldsymbol{L}) = \mathbf{Diag}(\psi_1(\boldsymbol{L}), \psi_2(\boldsymbol{L}), \ldots, \psi_n(\boldsymbol{L}))$, where $\psi_k := \psi_k(\boldsymbol{L})$ $(k = 1, 2, \ldots, n)$ is differentiable function from $\mathbb{R}^n_{\geq}$ to $\mathbb{R}$. If the conditions stated in* Theorem 2.1 *for $\boldsymbol{G} = \boldsymbol{O}_1\mathbf{Diag}(\ell_1^{-1}\psi_1, \ldots, \ell_n^{-1}\psi_n)\boldsymbol{O}_1'$ are satisfied, then we have the following identity;*

$$\mathbb{E}\,[\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'] = \mathbb{E}\,\left[\boldsymbol{O}_1\boldsymbol{\Psi}^{(1r)}\boldsymbol{O}_1' + \mathrm{Tr}\,(\boldsymbol{L}^{-1}\boldsymbol{\Psi})(\boldsymbol{I}_p - \boldsymbol{O}_1\boldsymbol{O}_1')\right],$$

where $\boldsymbol{\Psi}^{(1r)} = \mathbf{Diag}(\psi_1^{(1r)}, \psi_2^{(1r)}, \ldots, \psi_n^{(1r)})$ and, for $k = 1, 2, \ldots, n$,

$$\psi_k^{(1r)} = \sum_{b\neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b} + 2\frac{\partial\psi_k}{\partial\ell_k} - \frac{\psi_k}{\ell_k}.$$

In particular,

$$\mathbb{E}\,[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\}] = \mathbb{E}\,\left[\sum_{k=1}^{n}\left\{(p-n-1)\frac{\psi_k}{\ell_k} + 2\frac{\partial\psi_k}{\partial\ell_k} + \sum_{b\neq k}^{n}\frac{\psi_k - \psi_b}{\ell_k - \ell_b}\right\}\right]. \tag{2.2}$$

The identity (2.2) was given in [22] to develop unbiased risk estimate for orthogonally invariant estimators for precision matrices of the multivariate real normal distributions in the situation such that the number of samples n is less than the dimension p. Their approach to obtain the identity (2.2) is based on the arguments of Sheena [32]. It is interesting that the matrix form of the identity in Theorem 2.2 involves in the matrix $\boldsymbol{O}_2$. Here $\boldsymbol{O}_2$ is a $p \times (p-n)$ semi-orthogonal matrix such that a $p \times p$ matrix $[\boldsymbol{O}_1;\, \boldsymbol{O}_2]$ is orthogonal. This part causes essential difficulty in evaluation of risk for alternative estimators of the covariance matrix based on singular real Wishart matrices. Next theorem which plays an important role in derivation of an unbiased risk estimate under an invariant quadratic loss function can be obtained from an application of Theorem 2.2 and from calculus on the eigenstructures for singular Wishart matrices given in Section 4.

Theorem 2.3. *If the conditions stated in* Theorem 2.1 *for* $\boldsymbol{G} = \boldsymbol{O}_1 \mathbf{Diag}(\ell_1^{-1}\psi_1,\, \ldots,\, \ell_n^{-1}\psi_n)\boldsymbol{O}_1' \times \boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1 \mathbf{Diag}(\psi_1,\, \ldots,\, \psi_n)\boldsymbol{O}_1'$ *are satisfied, then we have*

$$\mathbb{E}[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\}] = \mathbb{E}[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\boldsymbol{\Psi}^{(1)}\boldsymbol{O}_1'\}],$$

where $\boldsymbol{\Psi}^{(1)} = \mathbf{Diag}(\psi_1^{(1)},\, \psi_2^{(1)},\, \ldots,\, \psi_n^{(1)})$ *with*

$$\psi_k^{(1)} = (p-n-1)\frac{\psi_k^2}{\ell_k} + 4\psi_k \cdot \frac{\partial \psi_k}{\partial \ell_k} + 2\psi_k \cdot \sum_{b \neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b}, \qquad k = 1,\, 2,\, \ldots,\, n. \tag{2.3}$$

2.2 Unbiased risk estimate for a class of invariant estimators

Consider the problem of estimating a covariance matrix $\boldsymbol{\Sigma}$ under a quadratic loss function

$$L(\widehat{\boldsymbol{\Sigma}},\, \boldsymbol{\Sigma}) = \mathrm{Tr}\,(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)^2, \tag{2.4}$$

where $\widehat{\boldsymbol{\Sigma}}$ is an estimator of $\boldsymbol{\Sigma}$. This loss function was used in [12, 30]. We denote by $R(\widehat{\boldsymbol{\Sigma}},\, \boldsymbol{\Sigma})$ the risk function of $\widehat{\boldsymbol{\Sigma}}$, i.e., the expected values of the loss function (2.4) with respect to the distributions of $\boldsymbol{X}$.

Recall that $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{O}_1\boldsymbol{L}\boldsymbol{O}_1'$, where $\boldsymbol{L} = \mathbf{Diag}(\ell_1, \ell_2, \ldots, \ell_n)$ and $\boldsymbol{O}_1$ is a $p \times n$ semi-orthogonal matrix such that $\boldsymbol{O}_1'\boldsymbol{O}_1 = \boldsymbol{I}_n$. Our class of estimators is of the form

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{O}_1\boldsymbol{\Phi}(\boldsymbol{L})\boldsymbol{O}_1', \tag{2.5}$$

where $\boldsymbol{\Phi} := \boldsymbol{\Phi}(\boldsymbol{L}) = \mathbf{Diag}(\varphi_1, \varphi_2, \ldots, \varphi_n)$, and $\varphi_k := \varphi_k(\boldsymbol{L})\,(k = 1, 2, \ldots, n)$ is a differentiable function from $\mathbb{R}^n_{\geq}$ to $\mathbb{R}$.

Theorem 2.4. *For the estimators of the form* (2.5) *that satisfies the regularity conditions stated in* Theorems 2.2 *and* 2.3, *we have*

$$R(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}) = \mathbb{E}\Big[\sum_{k=1}^{n}\Big\{(p-n-1)\big(\frac{\varphi_k^{(1)}}{\ell_k} - 2\frac{\varphi_k}{\ell_k}\big) + 2\big(\frac{\partial\varphi_k^{(1)}}{\partial\ell_k} - 2\frac{\partial\varphi_k}{\partial\ell_k}\big) + \sum_{b\neq k}^{n}\frac{(\varphi_k^{(1)} - 2\varphi_k) - (\varphi_b^{(1)} - 2\varphi_b)}{\ell_k - \ell_b}\Big\} + p\Big],$$

where $\varphi_k^{(1)} = (p-n-1)\varphi_k^2/\ell_k + 4\varphi_k(\partial\varphi_k/\partial\ell_k) + 2\varphi_k\sum_{b\neq k}^{n}(\varphi_k - \varphi_b)/(\ell_k - \ell_b)$ *for* $k = 1, 2, \ldots, n$.

Proof. Note that

$$\mathbb{E}[\mathrm{Tr}\,(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)^2] = \mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}})] - 2\mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}})] + p. \tag{2.6}$$

We first apply Theorem 2.3 to the first term in the right hand side of (2.6) to get

$$\mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}})] = \mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\boldsymbol{\Phi}^{(1)}\boldsymbol{O}_1')], \tag{2.7}$$

with $\boldsymbol{\Phi}^{(1)} = \mathbf{Diag}(\varphi_1^{(1)}, \varphi_2^{(1)}, \ldots, \varphi_n^{(1)})$. Then we apply Theorem 2.2 to the second term in the right hand side of (2.6) and the term in the right hand side of (2.7) to get the desired result. □

2.3 Alternative estimators

Proposition 2.1. *Consider the form of estimators* $\widehat{\boldsymbol{\Sigma}}_a = a\boldsymbol{S}$, *where a is a positive constant. Then the best constant is given by* $a = 1/(p+n+1)$ *under the loss function* (2.4).

Proof. Apply Theorem 2.4 with $\varphi_k = a\ell_k$ $(k = 1, 2, \ldots, n)$ to get that

$$\begin{aligned} R(\widehat{\boldsymbol{\Sigma}}_a, \boldsymbol{\Sigma}) &= np\{(p+n+1)a^2 - 2a\} + p \\ &= np(p+n+1)\big(a - \frac{1}{p+n+1}\big)^2 + \frac{p^2+p}{p+n+1}, \end{aligned}$$

which completes the proof. □

Proposition 2.2. *Let $a = 1/(p+n+1)$. Consider estimators of the form*

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}} &= \frac{1}{p+n+1}\boldsymbol{O}_1\mathbf{Diag}(\ell_1 + \frac{t}{\mathrm{Tr}\,\boldsymbol{S}^+}, \ell_2 + \frac{t}{\mathrm{Tr}\,\boldsymbol{S}^+}, \ldots, \ell_n + \frac{t}{\mathrm{Tr}\,\boldsymbol{S}^+})\boldsymbol{O}_1' \\ &= \frac{1}{p+n+1}\big(\boldsymbol{S} + \frac{t}{\mathrm{Tr}\,\boldsymbol{S}^+}\boldsymbol{O}_1\boldsymbol{O}_1'\big) \end{aligned}$$

where t is a positive constant and $\boldsymbol{S}^+$ is the Moore-Penrose inverse of $\boldsymbol{S}$. Then $\widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}}$ improves upon $\widehat{\boldsymbol{\Sigma}}_a$ if $0 < t < 2(n-1)(p-n-1)/\{(p-n+1)(p-n+3)\}$ under the loss function (2.4).

Proof. Apply Theorem 2.4 with $\varphi_k = a(\ell_k + t/\mathrm{Tr}\,\boldsymbol{S}^+)$ $(k = 1, 2, \ldots, n)$. Then we have, for $k = 1, 2, \ldots, n$,

$$\varphi_k^{(1)} = a\ell_k + 2a^2\big(\frac{p}{\mathrm{Tr}\,\boldsymbol{S}^+} + \frac{2}{\ell_k(\mathrm{Tr}\,\boldsymbol{S}^+)^2}\big)t + a^2\big(\frac{p-n-1}{\ell_k(\mathrm{Tr}\,\boldsymbol{S}^+)^2} + \frac{4}{\ell_k^2(\mathrm{Tr}\,\boldsymbol{S}^+)^3}\big)t^2.$$

Therefore, noting that

$$\sum_{k=1}^{n}\sum_{b\neq k}^{n}\frac{\ell_k^{-1} - \ell_b^{-1}}{\ell_k - \ell_b} < 0; \qquad \text{and} \qquad \sum_{k=1}^{n}\sum_{b\neq k}^{n}\frac{\ell_k^{-2} - \ell_b^{-2}}{\ell_k - \ell_b} < 0,$$

we have

$$\begin{aligned} &R(\widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}}, \boldsymbol{\Sigma}) - R(\widehat{\boldsymbol{\Sigma}}_a, \boldsymbol{\Sigma}) \\ &\quad < a\mathbb{E}\Big[(p-n-1)\big\{2a\big(p + \frac{2\mathrm{Tr}\,(\boldsymbol{S}^+)^2}{(\mathrm{Tr}\,\boldsymbol{S}^+)^2}\big)t + a\frac{(p-n+3)\mathrm{Tr}\,(\boldsymbol{S}^+)^2}{(\mathrm{Tr}\,\boldsymbol{S}^+)^2}t^2 - 2t\big\} \\ &\qquad + 2\big\{2a\frac{(p+2)\mathrm{Tr}\,(\boldsymbol{S}^+)^2}{(\mathrm{Tr}\,\boldsymbol{S}^+)^2}t + a\frac{(p-n+3)\mathrm{Tr}\,(\boldsymbol{S}^+)^2}{(\mathrm{Tr}\,\boldsymbol{S}^+)^2}t^2 - \frac{2\mathrm{Tr}\,(\boldsymbol{S}^+)^2}{(\mathrm{Tr}\,\boldsymbol{S}^+)^2}t\big\}\Big]. \end{aligned}$$

But the coefficients of $\{\mathrm{Tr}\,(\boldsymbol{S}^{+})^2/(\mathrm{Tr}\,\boldsymbol{S}^{+})^2\}t$ is evaluated as

$$\{2a(p-n-1)+4a(p+2)-4\}\frac{\mathrm{Tr}\,(\boldsymbol{S}^{+})^2}{(\mathrm{Tr}\,\boldsymbol{S}^{+})^2}t < 2a(p-n-1)t,$$

from which it follows that

$$R(\widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}},\,\boldsymbol{\Sigma}) - R(\widehat{\boldsymbol{\Sigma}}_{a},\,\boldsymbol{\Sigma}) < a^2\big\{(p-n+1)(p-n+3)t^2 - 2(n-1)(p-n-1)t\big\}.$$

This completes the proof. □

3 Complex case

Consider an $n \times p$ complex random matrix $\boldsymbol{Z}$ whose density function with respect to Lebesgue measure on $\mathbb{C}^{n\times p}$ is given by

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \pi^{-np}\mathrm{Det}\,(\boldsymbol{\Sigma})^{-n}\exp\{-\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{z}^{*}\boldsymbol{z})\}, \qquad \boldsymbol{z} \in \mathbb{C}^{n\times p},$$

where $\boldsymbol{\Sigma}$ is a $p\times p$ positive-definite Hermitian matrix. This is denoted by $\mathscr{L}(\boldsymbol{Z}) = \mathbb{C}N_{n\times p}(\boldsymbol{0},\,\boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$. See [1, 8, 18, 35] for multivariate complex normal distributions. Put $\boldsymbol{W} = \boldsymbol{Z}^{*}\boldsymbol{Z}$. Then the distribution of a $p \times p$ complex random matrix $\boldsymbol{W}$ is called a complex Wishart distribution with parameters $\boldsymbol{\Sigma}$, p, and n. This is denoted by $\mathscr{L}(\boldsymbol{W}) = \mathbb{C}W_p(\boldsymbol{\Sigma},\,n)$. The integers p and n are called the dimension and the degrees of freedom, respectively. The complex Wishart distributions were first explored by Goodman [8] and followed by [1, 9, 18, 33]. This model plays important roles in signal processing methods[17, 28, 41]. If $n < p$, then we call $\boldsymbol{W}$ a singular complex Wishart matrix, as the matrix $\boldsymbol{W}$ is singular. See Ratnaraja and Vaillancourt [31] for the density function of singular complex Wishart distribution with respect to Lebesgue measure on the set of $n \times p$ complex matrices $\mathbb{C}^{n\times p}$.

3.1 The Stein-Haff identities and calculus on eigenstructures for singular complex Wishart matrices

To describe integration by parts formula for the complex Wishart matrices, we introduce notion of a complex-valued function of complex variables. Recall that, for a complex number $z \in \mathbb{C}$, we write $z = \mathrm{Re}\, z + \sqrt{-1}\mathrm{Im}\, z$, where $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ are real numbers, and that we denote by $\bar{z}$ the complex conjugate of a complex number z, i.e., $\bar{z} = \mathrm{Re}\, z - \sqrt{-1}\mathrm{Im}\, z$. A continuous function $f : \mathbb{C} \to \mathbb{R}$ is called differentiable on $\mathbb{C}$ if $\partial f/\partial(\mathrm{Re}\, z)$ and $\partial f/\partial(\mathrm{Im}\, z)$ exist on $\mathbb{C}$. A function $f = u + \sqrt{-1}v$, where u, $v : \mathbb{C} \to \mathbb{R}$, is called differentiable if both u and v are differentiable. We define

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial(\mathrm{Rm}\, z)} - \sqrt{-1}\frac{\partial}{\partial(\mathrm{Im}\, z)}\right) \quad \text{and} \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial(\mathrm{Rm}\, z)} + \sqrt{-1}\frac{\partial}{\partial(\mathrm{Im}\, z)}\right).$$

For a differentiable function $f = u + \sqrt{-1}v : \mathbb{C} \to \mathbb{C}$, we set

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial u}{\partial(\mathrm{Re}\, z)} + \frac{\partial v}{\partial(\mathrm{Im}\, z)}\right) + \frac{\sqrt{-1}}{2}\left(\frac{\partial v}{\partial(\mathrm{Re}\, z)} - \frac{\partial u}{\partial(\mathrm{Im}\, z)}\right).$$

For an $n \times p$ complex matrix $\boldsymbol{Z} = (z_{ij})_{\substack{i=1,2,\dots,n\\ j=1,2,\dots,p}}$, we define an $n \times p$ matrix operator $\nabla_{\boldsymbol{Z}}$ as

$$\nabla_{\boldsymbol{Z}} = \left(\frac{\partial}{\partial z_{ij}}\right)_{\substack{i=1,2,\dots,n\\ j=1,2,\dots,p}} = \left(\frac{1}{2}\frac{\partial}{\partial(\mathrm{Re}\, z_{ij})} - \frac{\sqrt{-1}}{2}\frac{\partial}{\partial(\mathrm{Im}\, z_{ij})}\right)_{\substack{i=1,2,\dots,n\\ j=1,2,\dots,p}}.$$

For a $p \times q$ matrix $\boldsymbol{A} = (a_{ij})_{\substack{i=1,2,\dots,p\\ j=1,2,\dots,q}}$, whose (i, j) element a_{ij} is a differentiable function from $\mathbb{C}^{n\times p}$ to $\mathbb{C}$, we define the (i, j) element of a matrix $\nabla_{\boldsymbol{Z}}\boldsymbol{A}$ by

$$(\nabla_{\boldsymbol{Z}}\boldsymbol{A})_{ij} = \sum_{k=1}^{p}\frac{\partial a_{kj}}{\partial z_{ik}} \qquad \text{for } i = 1, 2, \dots, n;\ j = 1, 2, \dots, q.$$

Theorem 3.1. *Let $\boldsymbol{Z}$ be an $n \times p$ complex matrix with $\mathscr{L}(\boldsymbol{Z}) = \mathbb{C}N_{n\times p}(\boldsymbol{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a $p \times p$ positive-definite Hermitian matrix. Assume that, for a $p \times p$ complex random matrix*

$\boldsymbol{G} := \boldsymbol{G}(\boldsymbol{W}) = (g_{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}}$, *the* (i, j) *element* g_{ij} *is a differentiable function of* $\boldsymbol{Z}$ *through* $\boldsymbol{W} = \boldsymbol{Z}^*\boldsymbol{Z}$ *and satisfies the following conditions;*

$$\mathbb{E}\left[|z_{j_1 i_1}^2 g_{i_2 i_3}|\right] < \infty, \qquad \mathbb{E}\left[\left|z_{j_1 i_1}\frac{\partial g_{i_2 i_3}}{\partial z_{j_2 i_4}}\right|\right] < \infty$$

for $i_1, \ldots, i_4 = 1, 2, \ldots, p$; $j_1, j_2 = 1, 2, \ldots, n$. *Then we have*

$$\mathbb{E}\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{G}\right] = \mathbb{E}\left[n\boldsymbol{G} + (\boldsymbol{Z}'\nabla_{\boldsymbol{Z}})'\boldsymbol{G}\right], \tag{3.1}$$

where "$'$" stands for the transpose of a matrix. In particular,

$$\mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{G})\right] = \mathbb{E}\left[n\mathrm{Tr}\,(\boldsymbol{G}) + \mathrm{Tr}\,(\boldsymbol{Z}'\nabla_{\boldsymbol{Z}}\boldsymbol{G}')\right].$$

Remark 3.1. *Assume that* $n \geq p$. *Hence* $\boldsymbol{W}$ *is invertible with probability one. Let*

$$D_{\boldsymbol{W}} = \left\{\frac{1+\delta_{ij}}{2}\left(\frac{\partial}{\partial(\mathrm{Re}\, w_{ij})} + (1-\delta_{ij})\sqrt{-1}\frac{\partial}{\partial(\mathrm{Im}\, w_{ij})}\right)\right\}_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}}.$$

Note that the operator above is slightly different from that in Svensson [39] *so that the expressions below are changed correspondingly* [*see also* [21]]. *From* (3.1) *and the fact that* $\nabla_{\boldsymbol{Z}} = \overline{\boldsymbol{Z}}D_{\boldsymbol{W}}$ *and that* $\mathrm{Tr}\,\{D_{\boldsymbol{W}}\boldsymbol{W}\boldsymbol{G}\} = p\mathrm{Tr}\,(\boldsymbol{G}) + \mathrm{Tr}\,(\overline{\boldsymbol{W}}D_{\boldsymbol{W}}\boldsymbol{G}')$, *we can see that*

$$\mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{G})\right] = \mathbb{E}\left[(n-p)\mathrm{Tr}\,(\boldsymbol{G}) + \mathrm{Tr}\,\{D_{\boldsymbol{W}}(\boldsymbol{W}\boldsymbol{G})\}\right].$$

Replacing $\boldsymbol{G}$ *with* $\boldsymbol{W}^{-1}\boldsymbol{G}$, *we obtain that*

$$\mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{G})\right] = \mathbb{E}\left[(n-p)\mathrm{Tr}\,(\boldsymbol{G}\boldsymbol{W}^{-1}) + \mathrm{Tr}\,(D_{\boldsymbol{W}}\boldsymbol{G})\right], \tag{3.2}$$

which was obtained by Svensson [39].

For integers n, p such that $p > n \geq 1$, we denote by $\mathbb{C}V_{p,n}$ the set of all $p \times n$ semi-unitary matrices $\boldsymbol{U}_1$ such that $\boldsymbol{U}_1^*\boldsymbol{U}_1 = \boldsymbol{I}_n$, i.e., $\mathbb{C}V_{p,n} = \{\boldsymbol{U}_1 \in \mathbb{C}^{p\times n};\ \boldsymbol{U}_1^*\boldsymbol{U}_1 = \boldsymbol{I}_n\}$. Next theorem gives the Stein-Haff identity in a matrix form. Its proof is a combination of an application of Theorem 3.1 with calculus on the eigenstructures related to the singular real Wishart matrices given in Section 4.

Theorem 3.2. *Assume that* $n < p$ *and that* $\mathscr{L}(\boldsymbol{Z}) = \mathbb{C}N_{n\times p}(\boldsymbol{0},\, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$*, where* $\boldsymbol{\Sigma}$ *is a* $p \times p$ *positive-definite Hermitian matrix. Decompose* $\boldsymbol{Z}^*\boldsymbol{Z} = \boldsymbol{U}_1\boldsymbol{L}\boldsymbol{U}_1^*$*, where* $\boldsymbol{U}_1$ *is a* $p \times n$ *semi-unitary matrix such that* $\boldsymbol{U}_1^*\boldsymbol{U}_1 = \boldsymbol{I}_n$*. Let* $\boldsymbol{\Psi} := \boldsymbol{\Psi}(\boldsymbol{L}) = \mathbf{Diag}(\psi_1(\boldsymbol{L}),\, \psi_2(\boldsymbol{L}),\, \ldots,\, \psi_n(\boldsymbol{L}))$*, where* $\psi_k := \psi_k(\boldsymbol{L})\,(k = 1,\, 2,\, \ldots,\, n)$ *is differentiable function from* $\mathbb{R}^n_{\geq}$ *to* $\mathbb{R}$*. If the conditions stated in* Theorem 3.1 *for* $\boldsymbol{G} = \boldsymbol{U}_1\mathbf{Diag}(\ell_1^{-1}\psi_1,\, \ldots,\, \ell_n^{-1}\psi_n)\boldsymbol{U}_1^*$ *are satisfied, then we have the following identity;*

$$\mathbb{E}\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\right] = \mathbb{E}\left[\boldsymbol{U}_1\boldsymbol{\Psi}^{(1c)}\boldsymbol{U}_1^* + \mathrm{Tr}\,(\boldsymbol{L}^{-1}\boldsymbol{\Psi})(\boldsymbol{I}_p - \boldsymbol{U}_1\boldsymbol{U}_1^*)\right],$$

where $\boldsymbol{\Psi}^{(1c)} = \mathrm{Diag}(\psi_1^{(1c)}, \psi_2^{(1c)}, \ldots, \psi_n^{(1c)})$ *and, for* $k = 1,\, 2,\, \ldots,\, n$*,*

$$\psi_k^{(1c)} = \sum_{b\neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b} + \frac{\partial \psi_k}{\partial \ell_k}.$$

In particular,

$$\mathbb{E}\left[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\}\right] = \mathbb{E}\left[\sum_{k=1}^{n}\left\{(p-n)\frac{\psi_k}{\ell_k} + \frac{\partial \psi_k}{\partial \ell_k} + \sum_{b\neq k}^{n}\frac{\psi_k - \psi_b}{\ell_k - \ell_b}\right\}\right]. \tag{3.3}$$

Remark 3.2. *Combining* Theorem 3.2 *with the result obtained by* Svensson [39][*see also* [21, 40]]*, we can see that, under suitable conditions,*

$$\mathbb{E}\left[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\}\right] = \mathbb{E}\left[\sum_{k=1}^{\min(n,\,p)}\left\{|p-n|\frac{\psi_k}{\ell_k} + \frac{\partial \psi_k}{\partial \ell_k} + \sum_{b\neq k}^{\min(n,\,p)}\frac{\psi_k - \psi_b}{\ell_k - \ell_b}\right\}\right],$$

where $\boldsymbol{\Psi} = \mathrm{Diag}(\psi_1,\, \psi_2,\, \ldots,\, \psi_{\min(n,\,p)})$*. Here we decompose* $\boldsymbol{Z}^*\boldsymbol{Z}$ *as* $\boldsymbol{Z}^*\boldsymbol{Z} = \boldsymbol{U}_1\boldsymbol{L}\boldsymbol{U}_1^*$*, where* $\boldsymbol{L} = \mathrm{Diag}(\ell_1,\, \ldots,\, \ell_{\min(n,\,p)})$*,* $\boldsymbol{U}_1$ *belongs to* $\mathbb{C}V_{n,\,p}$ *if* $p > n$*, and* $\boldsymbol{U}_1$ *is a* $p \times p$ *unitary matrix if* $n > p$*.*

Next theorem is a complex analog of Theorem 2.3.

Theorem 3.3. *Under regularity conditions which are similar to those stated in* Theorem 2.3*, we have*

$$\mathbb{E}[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\}] = \mathbb{E}[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\widetilde{\boldsymbol{\Psi}}^{(1)}\boldsymbol{U}_1^*\}],$$

where $\widetilde{\boldsymbol{\Psi}}^{(1)} = \mathbf{Diag}(\tilde{\psi}_1^{(1)},\, \tilde{\psi}_2^{(1)},\, \ldots,\, \tilde{\psi}_n^{(1)})$ *with*

$$\tilde{\psi}_k^{(1)} = (p-n)\frac{\psi_k^2}{\ell_k} + 2\psi_k \cdot \frac{\partial \psi_k}{\partial \ell_k} + 2\psi_k \cdot \sum_{b\neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b}, \qquad k = 1,\, 2,\, \ldots,\, n. \tag{3.4}$$

3.2 Unbiased risk estimate for a class of invariant estimators

Consider the problem of estimating a covariance matrix $\boldsymbol{\Sigma}$ under the loss function (2.4), where $\widehat{\boldsymbol{\Sigma}}$ is an estimator of $\boldsymbol{\Sigma}$ based on $\boldsymbol{W}$. We denote by $R(\widehat{\boldsymbol{\Sigma}},\, \boldsymbol{\Sigma})$ the risk function of $\widehat{\boldsymbol{\Sigma}}$, i.e., the expected values of the loss function (2.4) with respect to the distribution of $\boldsymbol{Z}$.

Recall that $\boldsymbol{Z}^*\boldsymbol{Z} = \boldsymbol{U}_1\boldsymbol{L}\boldsymbol{U}_1^*$, where $\boldsymbol{L} = \mathbf{Diag}(\ell_1,\, \ell_2,\, \ldots,\, \ell_n)$ and $\boldsymbol{U}_1$ is a $p\times n$ semi-unitary matrix such that $\boldsymbol{U}_1^*\boldsymbol{U}_1 = \boldsymbol{I}_n$. Our class of estimators are of the form

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{U}_1\boldsymbol{\Phi}(\boldsymbol{L})\boldsymbol{U}_1^*, \tag{3.5}$$

where $\boldsymbol{\Phi} := \boldsymbol{\Phi}(\boldsymbol{L}) = \mathbf{Diag}(\varphi_1,\, \varphi_2,\, \ldots,\, \varphi_n)$ and $\varphi_k := \varphi_k(\boldsymbol{L})(k = 1,\, 2,\, \ldots,\, n)$ is a differentiable function from $\mathbb{R}^n_{\geq}$ to $\mathbb{R}$.

Theorem 3.4. *For the estimators of the form* (3.5) *that satisfies the regularity conditions stated in* Theorems 3.2 *and* 3.3*, we have*

$$\begin{aligned} R(\widehat{\boldsymbol{\Sigma}},\, \boldsymbol{\Sigma}) \;=\; \mathbb{E}\Bigg[\sum_{k=1}^{n}\Bigg\{&(p-n)\Big(\frac{\widetilde{\varphi}_k^{(1)}}{\ell_k} - 2\frac{\varphi_k}{\ell_k}\Big) + \Big(\frac{\partial \widetilde{\varphi}_k^{(1)}}{\partial \ell_k} - 2\frac{\partial \varphi_k}{\partial \ell_k}\Big) \\ &+ \sum_{b\neq k}^{n} \frac{(\widetilde{\varphi}_k^{(1)} - 2\varphi_k) - (\widetilde{\varphi}_b^{(1)} - 2\varphi_b)}{\ell_k - \ell_b}\Bigg\} + p\Bigg], \end{aligned}$$

where $\widetilde{\varphi}_k^{(1)} = (p-n)\varphi_k^2/\ell_k + 2\varphi_k(\partial\varphi_k/\partial\ell_k) + 2\varphi_k \sum_{b\neq k}^n (\varphi_k - \varphi_b)/(\ell_k - \ell_b)$ *for* $k = 1, 2, \ldots, n$.

Proof. Note that

$$\mathbb{E}[\mathrm{Tr}\,(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} - \boldsymbol{I}_p)^2] = \mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}})] - 2\mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}})] + p. \tag{3.6}$$

We first apply Theorem 3.3 to the first term in the right hand side of (3.6) to get

$$\mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\Sigma}})] = \mathbb{E}[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\widetilde{\boldsymbol{\Phi}}^{(1)}\boldsymbol{U}_1^*)], \tag{3.7}$$

where $\widetilde{\boldsymbol{\Phi}}^{(1)} = \mathbf{Diag}(\tilde{\varphi}_1^{(1)}, \tilde{\varphi}_2^{(1)}, \ldots, \tilde{\varphi}_n^{(1)})$. Then apply Theorem 3.2 to the second term in the right hand side of (3.6) and the term in the right hand side of (3.7) to get the desired result. □

3.3 Alternative estimators

Proposition 3.1. *Consider the form of estimators* $\widehat{\boldsymbol{\Sigma}}_a = a\boldsymbol{W}$, *where* a *is a positive constant. Then the best constant is given by* $a = 1/(p+n)$ *under the loss function* (2.4).

Proof. Apply Theorem 3.4 with $\varphi_k = a\ell_k$ $(k = 1, 2, \ldots, n)$ to get that

$$\begin{aligned} R(\widehat{\boldsymbol{\Sigma}}_a, \boldsymbol{\Sigma}) &= np\{(p+n)a^2 - 2a\} + p \\ &= np(p+n)\big(a - \frac{1}{p+n}\big)^2 + \frac{p^2}{p+n}, \end{aligned}$$

which completes the proof. □

Proposition 3.2. *Put* $a = 1/(p+n)$ *and consider estimators of the form*

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}} &= \frac{1}{p+n}\boldsymbol{U}_1\mathbf{Diag}(\ell_1 + \frac{t}{\mathrm{Tr}\,\boldsymbol{W}^+}, \ell_2 + \frac{t}{\mathrm{Tr}\,\boldsymbol{W}^+}, \ldots, \ell_n + \frac{t}{\mathrm{Tr}\,\boldsymbol{W}^+})\boldsymbol{U}_1^* \\ &= \frac{1}{p+n}\big(\boldsymbol{W} + \frac{t}{\mathrm{Tr}\,\boldsymbol{W}^+}\boldsymbol{U}_1\boldsymbol{U}_1^*\big) \end{aligned}$$

where t *is a positive constant and* $\boldsymbol{W}^+$ *is the Moore-Penrose inverse of* $\boldsymbol{W}$. *Then* $\widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}}$ *improves upon* $\widehat{\boldsymbol{\Sigma}}_a$ *if* $0 < t < 2(n-1)(p-n)/\{(p-n+1)(p-n+2)\}$ *under the loss function* (2.4).

Proof. Apply Theorem 3.4 with $\varphi_k = a(\ell_k + t/\mathrm{Tr}\,\boldsymbol{W}^+)$ $(k = 1, 2, \ldots, n)$. Then we have, for $k = 1, 2, \ldots, n$,

$$\varphi_k^{(1)} = a\ell_k + 2a^2\Big(\frac{p}{\mathrm{Tr}\,\boldsymbol{W}^+} + \frac{1}{\ell_k(\mathrm{Tr}\,\boldsymbol{W}^+)^2}\Big)t + a^2\Big(\frac{p-n}{\ell_k(\mathrm{Tr}\,\boldsymbol{W}^+)^2} + \frac{2}{\ell_k^2(\mathrm{Tr}\,\boldsymbol{W}^+)^3}\Big)t^2.$$

After a calculation similar to that in the proof of Proposition 2.2 we have

$$\begin{aligned} &R(\widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}}, \boldsymbol{\Sigma}) - R(\widehat{\boldsymbol{\Sigma}}_a, \boldsymbol{\Sigma}) \\ &\quad < a\mathbb{E}\Big[(p-n)\big\{2a\big(p + \frac{2\mathrm{Tr}\,(\boldsymbol{W}^+)^2}{(\mathrm{Tr}\,\boldsymbol{W}^+)^2}\big)t + a\frac{(p-n+2)\mathrm{Tr}\,(\boldsymbol{W}^+)^2}{(\mathrm{Tr}\,\boldsymbol{W}^+)^2}t^2 - 2t\big\} \\ &\qquad + \big\{2a\frac{(p+1)\mathrm{Tr}\,(\boldsymbol{W}^+)^2}{(\mathrm{Tr}\,\boldsymbol{W}^+)^2}t + a\frac{(p-n+2)\mathrm{Tr}\,(\boldsymbol{W}^+)^2}{(\mathrm{Tr}\,\boldsymbol{W}^+)^2}t^2 - \frac{2\mathrm{Tr}\,(\boldsymbol{W}^+)^2}{(\mathrm{Tr}\,\boldsymbol{W}^+)^2}t\big\}\Big]. \end{aligned}$$

But the coefficients of $\{\mathrm{Tr}\,(\boldsymbol{W}^+)^2/(\mathrm{Tr}\,\boldsymbol{W}^+)^2\}t$ is evaluated as

$$\{2a(p-n) + 2a(p+1) - 2\}\frac{\mathrm{Tr}\,(\boldsymbol{W}^+)^2}{(\mathrm{Tr}\,\boldsymbol{W}^+)^2}t < 2a(p-n)t,$$

from which it follows that

$$R(\widehat{\boldsymbol{\Sigma}}_{\mathrm{HF}}, \boldsymbol{\Sigma}) - R(\widehat{\boldsymbol{\Sigma}}_a, \boldsymbol{\Sigma}) < a^2\big\{(p-n+1)(p-n+2)t^2 - 2(n-1)(p-n-1)t\big\}.$$

This completes the proof. □

4 Proofs

4.1 Proof of Theorem 2.1

Write $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}'$, where $\boldsymbol{A}$ is a $p\times p$ nonsingular matrix, and put $\widetilde{\boldsymbol{X}} = (\tilde{x}_{ij})_{\substack{i=1,2,\ldots,n\\ j=1,2,\ldots,p}} = \boldsymbol{X}(\boldsymbol{A}')^{-1}$. Then $\widetilde{\boldsymbol{X}}$ is distributed according to $N_{n\times p}(\boldsymbol{0}, \boldsymbol{I}_n \otimes \boldsymbol{I}_p)$. Furthermore, put $\boldsymbol{H} = \boldsymbol{A}'\boldsymbol{G}(\boldsymbol{A}')^{-1}$ $= (h_{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}}$. We regard h_{ij} as a differentiable real-valued functions of $\widetilde{\boldsymbol{X}}$. Define $\widetilde{\boldsymbol{S}} =$

$(\tilde{s}_{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}} = \widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}}$. So $\widetilde{\boldsymbol{S}} = \boldsymbol{A}^{-1}\boldsymbol{S}(\boldsymbol{A}')^{-1}$. Since

$$\frac{\partial}{\partial \tilde{x}_{ij}} \exp\left(-\frac{1}{2}\mathrm{Tr}\,(\widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}})\right) = -\tilde{x}_{ij}\exp\left(-\frac{1}{2}\mathrm{Tr}\,(\widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}})\right),$$

for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$. Therefore, for a differentiable function $h : \mathbb{R}^{n\times p} \to \mathbb{R}$, $k = 1, 2, \ldots, n$ and $i, j = 1, 2, \ldots, p$, we have

$$\frac{\partial}{\partial \tilde{x}_{ki}} \left[\tilde{x}_{kj} h \exp\left(-\frac{1}{2}\mathrm{Tr}\,(\widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}})\right)\right] = \left\{\delta_{ij}h + \tilde{x}_{kj}\frac{\partial h}{\partial \tilde{x}_{ki}} - \tilde{x}_{ki}\tilde{x}_{kj}h\right\}\exp\left(-\frac{1}{2}\mathrm{Tr}\,(\widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}})\right),$$

from which it follows that

$$\mathbb{E}\,[\tilde{x}_{ki}\tilde{x}_{kj}h] = \mathbb{E}\,\left[\delta_{ij}h + \tilde{x}_{kj}\cdot\frac{\partial h}{\partial \tilde{x}_{ki}}\right].$$

Here δ_{ij} is Kronecker's delta, i.e., $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$ for integers i, j. Putting $h = h_{jl}\,(j, l = 1, 2, \ldots, p)$ and summing over j from 1 to p and over k from 1 to n, we obtain

$$\mathbb{E}\,\left[\sum_{j=1}^{p}\tilde{s}_{ij}h_{jl}\right] = \mathbb{E}\,\left[nh_{il} + \sum_{j=1}^{p}\sum_{k=1}^{n}\tilde{x}_{kj}\cdot\frac{\partial h_{jl}}{\partial \tilde{x}_{ki}}\right].$$

Thus, putting $\boldsymbol{H} = (h_{jl})_{\substack{j=1,2,\ldots,p\\ l=1,2,\ldots,p}}$, we get

$$\mathbb{E}\,[\widetilde{\boldsymbol{S}}\boldsymbol{H}] = \mathbb{E}\,\left[n\boldsymbol{H} + (\widetilde{\boldsymbol{X}}'\nabla_{\widetilde{\boldsymbol{X}}})'\boldsymbol{H}\right]. \tag{4.1}$$

Finally, by the definition of $\boldsymbol{H}$, we have $\mathbb{E}\,[\widetilde{\boldsymbol{S}}\boldsymbol{H}] = \mathbb{E}\,[\boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{G}(\boldsymbol{A}')^{-1}]$ while, since $\nabla_{\widetilde{\boldsymbol{X}}} = \nabla_{\boldsymbol{X}}\boldsymbol{A}$, we have $\mathbb{E}\,[(\widetilde{\boldsymbol{X}}'\nabla_{\widetilde{\boldsymbol{X}}})'\boldsymbol{H}] = \mathbb{E}\,[\boldsymbol{A}'(\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{G}(\boldsymbol{A}')^{-1}]$. Putting these two equations into (4.1) and multiplying by $(\boldsymbol{A}')^{-1}$ from the left and by $\boldsymbol{A}'$ from the right, we get $\mathbb{E}\,[(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{S}\boldsymbol{G}] = \mathbb{E}\,[n\boldsymbol{G} + (\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{G}]$, which completes the proof of (2.1). □

4.2 Proof of Theorem 2.2

To prove Theorem 2.2, we need the following lemma which states the partial derivatives of the eigenvalues and the elements of eigenvectors of the singular real Wishart matrix $\boldsymbol{S} = \boldsymbol{X}'\boldsymbol{X}$

with respect to the elements of the matrix $\boldsymbol{X}$. For full-rank real Wishart matrices, partial derivatives which play a similar role to those in the next lemma appeared in Stein [37].

In the rest of the paper, we denote by $\{\boldsymbol{AB}\}_{ij}$ the $(i,\,j)$ element of product of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

Lemma 4.1. *Assume that $p > n$. Let $\boldsymbol{X} = (x_{ij})_{\substack{i=1,2,\dots,n\\ j=1,2,\dots,p}}$ and decompose a $p \times p$ matrix $\boldsymbol{X}'\boldsymbol{X}$ as $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{O}_1\boldsymbol{L}\boldsymbol{O}_1'$, where $\boldsymbol{O}_1 \in V_{p,n} = \{\boldsymbol{O}_1 \in \mathbb{R}^{p\times n};\, \boldsymbol{O}_1'\boldsymbol{O}_1 = \boldsymbol{I}_n\}$ and $\boldsymbol{L} = \mathbf{Diag}\,(\ell_1,\,\ell_2,\,\dots,\,\ell_n)$ is an $n\times n$ diagonal matrix with $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_n > 0$. Furthermore, let $\boldsymbol{O}_2 = (o_{ij})_{\substack{i=1,2,\dots,p\\ j=n+1,2,\dots,p}} \in V_{p,\,p-n}$ be a $p\times(p-n)$ semi-orthogonal matrix such that $\boldsymbol{O} = [\boldsymbol{O}_1;\,\boldsymbol{O}_2]$ is a $p\times p$ orthogonal matrix. If $\ell_1 > \ell_2 > \cdots > \ell_n > 0$, then we have, for $i,\,k,\,m = 1,\,2,\,\dots,\,n$ and $a,\,j = 1,\,2,\,\dots,\,p$,*

$$\begin{aligned}
\frac{\partial \ell_m}{\partial x_{ij}} &= 2\sum_{c_1=1}^{p} o_{c_1 m} x_{ic_1} o_{jm};\\
\frac{\partial o_{ak}}{\partial x_{ij}} &= \sum_{b\neq k}^{n}\sum_{c_1=1}^{p} \frac{o_{ab}\{o_{jb}o_{c_1k} + o_{c_1b}o_{jk}\}x_{ic_1}}{\ell_k-\ell_b} + \sum_{b=n+1}^{p}\sum_{c_1=1}^{p}\frac{o_{ab}\{o_{jb}o_{c_1k} + o_{c_1b}o_{jk}\}x_{ic_1}}{\ell_k},
\end{aligned}$$

for $a \neq k$, and $\partial o_{kk}/\partial x_{ij} = 0$ for $k = 1,\,2,\,\dots,\,n$.

Proof. Taking differentials of

$$\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{O}\begin{bmatrix} \boldsymbol{L} & \mathbf{0}_{n\times(p-n)} \\ \mathbf{0}_{(p-n)\times n} & \mathbf{0}_{(p-n)\times(p-n)} \end{bmatrix}\boldsymbol{O}' = [\boldsymbol{O}_1;\boldsymbol{O}_2]\begin{bmatrix} \boldsymbol{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{O}_1' \\ \boldsymbol{O}_2'\end{bmatrix},$$

and using the fact that $\boldsymbol{O}'(d\boldsymbol{O}) + (d\boldsymbol{O}')\boldsymbol{O} = \mathbf{0}_{p\times p}$, we have, for $a,\,k,\,m = 1,\,2,\,\dots,\,n$ such that $a\neq k$,

$$\{\boldsymbol{O}_1'(d\boldsymbol{O}_1)\}_{ak} = \frac{1}{\ell_k-\ell_a}\{\boldsymbol{O}_1'((d\boldsymbol{X}')\boldsymbol{X} + \boldsymbol{X}'(d\boldsymbol{X}))\boldsymbol{O}_1\}_{ak}; \tag{4.2}$$

$$(d\boldsymbol{L})_{mm} = \{\boldsymbol{O}_1'((d\boldsymbol{X}')\boldsymbol{X} + \boldsymbol{X}'(d\boldsymbol{X}))\boldsymbol{O}_1\}_{mm}; \tag{4.3}$$

and, for $a = n+1, 2, \ldots, p$ and $k = 1, 2, \ldots, n$,

$$\{\boldsymbol{O}_2'(d\boldsymbol{O}_1)\}_{ak} = \frac{1}{\ell_k}\{\boldsymbol{O}_2'((d\boldsymbol{X}')\boldsymbol{X} + \boldsymbol{X}'(d\boldsymbol{X}))\boldsymbol{O}_1\}_{ak}. \tag{4.4}$$

From the fact that dx_{ij} is the dual basis of $\partial/\partial x_{ij}$ for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$, we have

$$dx_{ij}\left(\frac{\partial}{\partial x_{kl}}\right) = \delta_{ik}\delta_{jl} \tag{4.5}$$

for $k = 1, 2, \ldots, n$ and $l = 1, 2, \ldots, p$. Using (4.2) and (4.5), we get that, for $i, m = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$,

$$\begin{aligned}
\frac{\partial \ell_m}{\partial x_{ij}} &= (d\boldsymbol{L})_{mm}\left(\frac{\partial}{\partial x_{ij}}\right) \\
&= \sum_{c_1, c_3=1}^{p}\sum_{c_2=1}^{n} o_{c_1 m} dx_{c_2 c_1}\left(\frac{\partial}{\partial x_{ij}}\right) x_{c_2 c_3} o_{c_3 m} + \sum_{c_1, c_3=1}^{p}\sum_{c_2=1}^{n} o_{c_1 m} x_{c_2 c_1} dx_{c_2 c_3}\left(\frac{\partial}{\partial x_{ij}}\right) o_{c_3 m} \\
&= 2\sum_{c_1}^{p} o_{c_1 m} x_{i c_1} o_{jm},
\end{aligned}$$

which competes the first part of the lemma.

Using (4.2) and (4.4), we get that, for $i,\, k = 1,\, 2,\, \ldots,\, n$ and $a,\, j = 1,\, 2,\, \ldots,\, p$,

$$
\begin{aligned}
\frac{\partial o_{ak}}{\partial x_{ij}} &= \sum_{b=1}^{p}\sum_{c_1=1}^{p} o_{ab}o_{c_1 b}\frac{\partial o_{c_1 k}}{\partial x_{ij}} \\
&= \sum_{b=1}^{n} o_{ab}\{\boldsymbol{O}_1'(d\boldsymbol{O}_1)\}_{bk}\left(\frac{\partial}{\partial x_{ij}}\right) + \sum_{b=n+1}^{p} o_{ab}\{\boldsymbol{O}_2'(d\boldsymbol{O}_1)\}_{bk}\left(\frac{\partial}{\partial x_{ij}}\right) \\
&= \sum_{b\neq k}^{n} o_{ab}\frac{1}{\ell_k-\ell_b}\{\boldsymbol{O}_1'((d\boldsymbol{X}')\boldsymbol{X}+\boldsymbol{X}'(d\boldsymbol{X}))\boldsymbol{O}_1\}_{bk}\left(\frac{\partial}{\partial x_{ij}}\right) \\
&\quad + \sum_{b=n+1}^{p} o_{ab}\frac{1}{\ell_k}\{\boldsymbol{O}_2'((d\boldsymbol{X}')\boldsymbol{X}+\boldsymbol{X}'(d\boldsymbol{X}))\boldsymbol{O}_1\}_{bk}\left(\frac{\partial}{\partial x_{ij}}\right) \\
&= \sum_{b\neq k}^{n}\sum_{c_1,\, c_3=1}^{p}\sum_{c_2=1}^{n}\frac{o_{ab}o_{c_1 b}o_{c_3 k}}{\ell_k-\ell_b}\{(dx_{c_2c_1})x_{c_2c_3}+x_{c_2c_1}(dx_{c_2c_3})\}\left(\frac{\partial}{\partial x_{ij}}\right) \\
&\quad + \sum_{b=n+1}^{p}\sum_{c_1,\, c_3=1}^{p}\sum_{c_2=1}^{n}\frac{o_{ab}o_{c_1 b}o_{c_3 k}}{\ell_k}\{(dx_{c_2c_1})x_{c_2c_3}+x_{c_2c_1}(dx_{c_2c_3})\}\left(\frac{\partial}{\partial x_{ij}}\right) \\
&= \sum_{b\neq k}^{n}\sum_{c_1=1}^{p}\frac{o_{ab}o_{jb}o_{c_1 k}x_{ic_1}}{\ell_k-\ell_b} + \sum_{b\neq k}^{n}\sum_{c_1=1}^{p}\frac{o_{ab}o_{c_1 b}o_{jk}x_{ic_1}}{\ell_k-\ell_b} + \sum_{b=n+1}^{p}\sum_{c_1=1}^{p}\frac{o_{ab}o_{jb}o_{c_1 k}x_{ic_1}}{\ell_k} \\
&\qquad + \sum_{b=n+1}^{p}\sum_{c_1=1}^{p}\frac{o_{ab}o_{c_1 b}o_{jk}x_{ic_1}}{\ell_k},
\end{aligned}
$$

which completes the proof of the lemma. □

Proof of Theorem 2.2. We adapt the notation in Theorem 2.1 and Lemma 4.1. Apply Theorem 2.1 with $\boldsymbol{G} = \boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'$ to get

$$
\mathbb{E}\,[\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'] = \mathbb{E}\,[n\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1' + \{(\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\}]. \tag{4.6}
$$

In order to apply Lemma 4.1 to the second term inside the right expectation, we first note that, for $i,\, j = 1,\, 2,\, \ldots,\, p$, the $(i,\, j)$ element of $(\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'$ is given by

$$
\begin{aligned}
\{(\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\}_{ij} &= \sum_{c_1,\, c_3=1}^{n}\sum_{c_2=1}^{p} x_{c_1c_2}\frac{\partial}{\partial x_{c_1 i}}\{o_{c_2c_3}\psi_{c_3}o_{jc_3}\} \\
&= \sum_{c_1,\, c_3=1}^{n}\sum_{c_2=1}^{p} x_{c_1c_2}\psi_{c_3}o_{jc_3}\frac{\partial o_{c_2c_3}}{\partial x_{c_1 i}} + \sum_{c_1,\, c_3=1}^{n}\sum_{c_2=1}^{p} x_{c_1c_2}\psi_{c_3}o_{c_2c_3}\frac{\partial o_{jc_3}}{\partial x_{c_1 i}} \\
&\quad + \sum_{c_1,\, c_3,\, m=1}^{n}\sum_{c_2=1}^{p} x_{c_1c_2}o_{c_2c_3}o_{jc_3}\frac{\partial \psi_{c_3}}{\partial \ell_m}\cdot\frac{\partial \ell_m}{\partial x_{c_1 i}} =: A_1 + A_2 + A_3.
\end{aligned} \tag{4.7}
$$

From the second equation of Lemma 4.1 and the fact that $\boldsymbol{O}_1'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{O}_1 = \boldsymbol{L}$, $\boldsymbol{O}_1'\boldsymbol{O}_1 = \boldsymbol{I}_n$ and that $\boldsymbol{O}_2'\boldsymbol{O}_1 = \mathbf{0}_{(p-n)\times n}$, we get

$$A_1 = \sum_{c_3=1}^{n} o_{ic_3} o_{jc_3} \sum_{b\neq c_3}^{n} \frac{\ell_b \psi_{c_3}}{\ell_{c_3} - \ell_b}.$$

We use Lemma 4.1 and perform a calculation which is similar to that above to get that

$$A_2 = \sum_{b=1}^{n} o_{ib} o_{jb} \sum_{c_3\neq b}^{n} \frac{\ell_{c_3}\psi_{c_3}}{\ell_{c_3} - \ell_b} + \sum_{b=n+1}^{p} o_{ib} o_{jb} \sum_{c_3=1}^{n} \psi_{c_3}.$$

Furthermore, from the first equation of Lemma 4.1, we get that

$$A_3 = 2 \sum_{c_3=1}^{n} o_{ic_3} o_{jc_3} \ell_{c_3} \frac{\partial \psi_{c_3}}{\partial \ell_{c_3}}.$$

Putting these three expressions into (4.7), we obtain that

$$\begin{aligned} \{(\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\}_{ij} = & \sum_{c_3=1}^{n} o_{ic_3} o_{jc_3} \left\{ 2\ell_{c_3} \frac{\partial \psi_{c_3}}{\partial \ell_{c_3}} + \sum_{b\neq c_3}^{n} \frac{\ell_b\psi_{c_3} - \ell_b\psi_b}{\ell_{c_3} - \ell_b} \right\} \\ & + \sum_{b=n+1}^{p} o_{ib} o_{jb} \sum_{c_3=1}^{n} \psi_{c_3}. \end{aligned} \tag{4.8}$$

Putting (4.8) into (4.6), we get that

$$\mathbb{E}\left[\{\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\boldsymbol{O}_1\boldsymbol{\Psi}\boldsymbol{O}_1'\}_{ij}\right] = \mathbb{E}\left[\sum_{k=1}^{n} o_{ik} o_{jk} \left\{ n\psi_k + 2\ell_k \frac{\partial \psi_k}{\partial \ell_k} + \sum_{b\neq k}^{n} \frac{\ell_b\psi_k - \ell_b\psi_b}{\ell_k - \ell_b} \right\} + \sum_{b=n+1}^{p} o_{ib} o_{jb} \sum_{k=1}^{n} \psi_k \right].$$

Finally, changing ψ_k into $\ell_k^{-1}\psi_k$, noting that $\boldsymbol{O}_2\boldsymbol{O}_2' = \boldsymbol{I}_p - \boldsymbol{O}_1\boldsymbol{O}_1'$ and that

$$\sum_{b\neq k}^{n} \frac{\ell_b\ell_k^{-1}\psi_k - \ell_b\ell_b^{-1}\psi_b}{\ell_k - \ell_b} = \sum_{b\neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b} - (n-1)\frac{\psi_k}{\ell_k}, \tag{4.9}$$

we can complete the proof of this theorem. □

4.3 Proof of Theorem 2.3

Write

$$\begin{aligned} \boldsymbol{F} &= (f_{ij})_{\substack{i=1,2,\dots,p \\ j=1,2,\dots,p}} = \boldsymbol{O}_1 \mathbf{Diag}\Big(\frac{\psi_1}{\ell_1}, \frac{\psi_2}{\ell_2}, \dots, \frac{\psi_n}{\ell_n}\Big)\boldsymbol{O}_1'; \\ \widetilde{\boldsymbol{F}} &= (\tilde{f}_{ij})_{\substack{i=1,2,\dots,p \\ j=1,2,\dots,p}} = \boldsymbol{O}_1 \mathbf{Diag}\big(\psi_1, \psi_2, \dots, \psi_n\big)\boldsymbol{O}_1'. \end{aligned}$$

First apply Theorem 2.1 with $\boldsymbol{G} = \boldsymbol{F}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}}$ to get that

$$\mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}})\right] = \mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\boldsymbol{F}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}})\right] =: \mathbb{E}[n\Delta_1 + \Delta_2], \tag{4.10}$$

where $\Delta_1 = \mathrm{Tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\mathbf{Diag}\big(\psi_1^2/\ell_1,\, \psi_2^2/\ell_2,\, \ldots,\, \psi_n^2/\ell_n\big)\boldsymbol{O}_1'\right\}$ and $\Delta_2 = \mathrm{Tr}\left\{(\boldsymbol{X}'\nabla_{\boldsymbol{X}})'\boldsymbol{F}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}}\right\}$. We evaluate the expectation of Δ_2 in (4.10). Since $\boldsymbol{F}$, $\widetilde{\boldsymbol{F}}$, and $\boldsymbol{\Sigma}^{-1} = (\sigma^{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}}$ are symmetric matrices, we see that the expectation of Δ_2 is given by

$$\begin{aligned}
\mathbb{E}[\Delta_2] &= \mathbb{E}[\mathrm{Tr}\,\{\boldsymbol{X}'\nabla_{\boldsymbol{X}}\widetilde{\boldsymbol{F}}\boldsymbol{\Sigma}^{-1}\boldsymbol{F}\}] \\
&= \mathbb{E}\Bigg[\sum_{c_1=1}^{n}\sum_{c_2,\,c_3,\,c_4,\,i=1}^{p} x_{c_1 i}\frac{\partial}{\partial x_{c_1c_2}}\big(\widetilde{f}_{c_2c_3}\sigma^{c_3c_4}f_{c_4i}\big)\Bigg] \\
&=: \mathbb{E}\Bigg[\sum_{c_3,\,c_4,\,i=1}^{p}\sigma^{c_3c_4}f_{c_4i}T_{ic_3}^{(1)} + \sum_{c_3,\,c_4=1}^{p}\sigma^{c_3c_4}T_{c_4c_3}^{(2)}\Bigg],
\end{aligned} \tag{4.11}$$

where, for i, c_3, $c_4 = 1, 2, \ldots, p$,

$$T_{ic_3}^{(1)} = \sum_{c_1=1}^{n}\sum_{c_2=1}^{p} x_{c_1i}\frac{\partial\,\widetilde{f}_{c_2c_3}}{\partial x_{c_1c_2}}; \qquad \text{and} \qquad T_{c_4c_3}^{(2)} = \sum_{c_1=1}^{n}\sum_{c_2,\,i=1}^{p} x_{c_1i}\widetilde{f}_{c_2c_3}\frac{\partial f_{c_4i}}{\partial x_{c_1c_2}}.$$

Next, using Lemma 4.1, we evaluate $T_{ic_3}^{(1)}$ and $T_{c_4c_3}^{(2)}$, respectively. To evaluate $T_{ic_3}^{(1)}$, recall that $\widetilde{\boldsymbol{F}} = \boldsymbol{O}_1\mathbf{Diag}(\psi_1,\, \psi_2,\, \ldots,\, \psi_n)\boldsymbol{O}_1'$. Then we have

$$\begin{aligned}
T_{ic_3}^{(1)} &= \sum_{c_1=1}^{n}\sum_{c_2=1}^{p} x_{c_1i}\frac{\partial}{\partial x_{c_1c_2}}\left(\sum_{c_5=1}^{n} o_{c_2c_5}\psi_{c_5}o_{c_3c_5}\right) \\
&= \sum_{c_1,\,c_5=1}^{n}\sum_{c_2=1}^{p} x_{c_1i}\psi_{c_5}o_{c_3c_5}\frac{\partial o_{c_2c_5}}{\partial x_{c_1c_2}} + \sum_{c_1,\,c_5=1}^{n}\sum_{c_2=1}^{p} x_{c_1i}\psi_{c_5}o_{c_2c_5}\frac{\partial o_{c_3c_5}}{\partial x_{c_1c_2}} \\
&\quad + \sum_{c_1,\,c_5,\,m=1}^{n}\sum_{c_2=1}^{p} x_{c_1i}o_{c_2c_5}o_{c_3c_5}\frac{\partial\psi_{c_5}}{\partial\ell_m}\cdot\frac{\partial\ell_m}{\partial x_{c_1c_2}} =: T_{ic_3}^{(11)} + T_{ic_3}^{(12)} + T_{ic_3}^{(13)}.
\end{aligned} \tag{4.12}$$

Applying Lemma 4.1 and using the fact that $\boldsymbol{O}_1'\boldsymbol{O}_1 = \boldsymbol{I}_n$ and that $\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{O}_1\boldsymbol{L}\boldsymbol{O}_1'$, we have

$$\begin{aligned}
(T_{ic_3}^{(11)})_{\substack{i=1,2,\ldots,p\\ c_3=1,2,\ldots,p}} &= \boldsymbol{O}_1\mathbf{Diag}\Big(\sum_{b\neq1}\frac{\ell_1\psi_1}{\ell_1-\ell_b},\, \sum_{b\neq2}\frac{\ell_2\psi_2}{\ell_2-\ell_b},\, \ldots,\, \sum_{b\neq n}\frac{\ell_n\psi_n}{\ell_n-\ell_b}\Big)\boldsymbol{O}_1' \\
&\quad + (p-n)\boldsymbol{O}_1\mathbf{Diag}\big(\psi_1,\, \psi_2,\, \ldots,\, \psi_n\big)\boldsymbol{O}_1'.
\end{aligned}$$

We use Lemma 4.1 and the fact that $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{O}_2'\boldsymbol{O}_2 = \mathbf{0}_{p\times(p-n)}$ to perform a calculation which is similar to that above so that we have

$$(T^{(12)}_{ic_3})_{\substack{i=1,\,2,\,\dots,\,p\\ c_3=1,\,2,\,\dots,\,p}} \quad = \quad \boldsymbol{O}_1\mathbf{Diag}\big(\sum_{b\neq 1}\frac{\ell_1\psi_b}{\ell_b-\ell_1},\,\sum_{b\neq 2}\frac{\ell_2\psi_b}{\ell_b-\ell_2},\,\dots,\,\sum_{b\neq n}\frac{\ell_n\psi_b}{\ell_b-\ell_n}\big)\boldsymbol{O}_1'.$$

Finally, by Lemma 4.1, we have

$$(T^{(13)}_{ic_3})_{\substack{i=1,\,2,\,\dots,\,p\\ c_3=1,\,2,\,\dots,\,p}} \quad = \quad \boldsymbol{O}_1\mathbf{Diag}\big(2\ell_1\frac{\partial\psi_1}{\partial\ell_1},\,2\ell_2\frac{\partial\psi_2}{\partial\ell_2},\,\dots,\,2\ell_n\frac{\partial\psi_n}{\partial\ell_n}\big)\boldsymbol{O}_1'.$$

Putting these three expressions into (4.12), we get that

$$\mathbb{E}\Big[\sum_{c_3,\,c_4,\,i=1}^{p}\sigma^{c_3c_4}f_{c_4i}T^{(1)}_{ic_3}\Big] = \mathbb{E}\big[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1\mathbf{Diag}\big(\psi_1^{(1a)},\,\psi_2^{(1a)},\,\dots,\,\psi_n^{(1a)}\big)\boldsymbol{O}_1'\}\big], \tag{4.13}$$

where, for $i = 1,\,2,\,\dots,\,n$,

$$\psi_k^{(1a)} = \psi_k\sum_{b\neq k}^{n}\frac{\psi_k-\psi_b}{\ell_k-\ell_b} + 2\psi_k\frac{\partial\psi_k}{\partial\ell_k} + (p-n)\frac{\psi_k^2}{\ell_k}.$$

Since $\boldsymbol{F} = \boldsymbol{O}_1\mathbf{Diag}\big(\ell_1^{-1}\psi_1,\,\ell_2^{-1}\psi_2,\,\dots,\,\ell_n^{-1}\psi_n\big)\boldsymbol{O}_1'$, we have

$$\begin{aligned}
T^{(2)}_{c_4c_3} &= \sum_{c_1,\,c_5=1}^{n}\sum_{c_2,\,i=1}^{p}x_{c_1i}\widetilde{f}_{c_2c_3}\frac{\partial}{\partial x_{c_1c_2}}\big(o_{c_4c_5}\frac{\psi_{c_5}}{\ell_{c_5}}o_{ic_5}\big)\\
&= \sum_{c_1,\,c_5=1}^{n}\sum_{c_2,\,i=1}^{p}x_{c_1i}\widetilde{f}_{c_2c_3}o_{ic_5}\frac{\psi_{c_5}}{\ell_{c_5}}\cdot\frac{\partial o_{c_4c_5}}{\partial x_{c_1c_2}} + \sum_{c_1,\,c_5=1}^{n}\sum_{c_2,\,i=1}^{p}x_{c_1i}\widetilde{f}_{c_2c_3}o_{c_4c_5}\frac{\psi_{c_5}}{\ell_{c_5}}\cdot\frac{\partial o_{ic_5}}{\partial x_{c_1c_2}}\\
&\quad + \sum_{c_1,\,c_5,\,m=1}^{n}\sum_{c_2,\,i=1}^{p}x_{c_1i}\widetilde{f}_{c_2c_3}o_{c_4c_5}o_{ic_5}\frac{\partial\ell_m}{\partial x_{c_1c_2}}\cdot\frac{\partial}{\partial\ell_m}\left(\frac{\psi_{c_5}}{\ell_{c_5}}\right)\\
&=: T^{(21)}_{c_4c_3} + T^{(23)}_{c_4c_3} + T^{(23)}_{c_4c_3}.
\end{aligned} \tag{4.14}$$

By Lemma 4.1, we have

$$\begin{aligned}
(T^{(21)}_{c_4c_3})_{\substack{c_4=4,\,2,\,\dots,\,p\\ c_3=1,\,2,\,\dots,\,p}} &= \boldsymbol{O}_1\mathbf{Diag}\big(\sum_{b\neq 1}\frac{\psi_1\psi_b}{\ell_b-\ell_1},\,\sum_{b\neq 2}\frac{\psi_2\psi_b}{\ell_b-\ell_2},\,\dots,\,\sum_{b\neq n}\frac{\psi_n\psi_b}{\ell_b-\ell_n}\big)\boldsymbol{O}_1';\\
(T^{(22)}_{c_4c_3})_{\substack{c_4=4,\,2,\,\dots,\,p\\ c_3=1,\,2,\,\dots,\,p}} &= \boldsymbol{O}_1\mathbf{Diag}\big(\sum_{b\neq 1}\frac{\ell_b\psi_1^2}{(\ell_1-\ell_b)\ell_1},\,\sum_{b\neq 2}\frac{\ell_b\psi_2^2}{(\ell_2-\ell_b)\ell_2},\,\dots,\,\sum_{b\neq n}\frac{\ell_b\psi_n^2}{(\ell_n-\ell_b)\ell_n}\big)\boldsymbol{O}_1';\\
(T^{(23)}_{c_4c_3})_{\substack{c_4=4,\,2,\,\dots,\,p\\ c_3=1,\,2,\,\dots,\,p}} &= \boldsymbol{O}_1\mathbf{Diag}\big(2\ell_1\psi_1\frac{\partial}{\partial\ell_1}\left(\frac{\psi_1}{\ell_1}\right),\,2\ell_2\psi_2\frac{\partial}{\partial\ell_2}\left(\frac{\psi_2}{\ell_2}\right),\,\dots,\,2\ell_n\psi_n\frac{\partial}{\partial\ell_n}\left(\frac{\psi_n}{\ell_n}\right)\big)\boldsymbol{O}_1'.
\end{aligned}$$

Putting the above three equations into (4.14), we see that the second term in the right hand side of (4.11) is given by

$$\mathbb{E}\left[\sum_{c_3,\, c_4=1}^{p} \sigma^{c_3 c_4} T^{(2)}_{c_4 c_3}\right] = \mathbb{E}\left[\mathrm{Tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1 \mathbf{Diag}(\psi_1^{(1b)},\, \psi_2^{(1b)},\, \ldots,\, \psi_n^{(1b)})\boldsymbol{O}_1'\right\}\right], \tag{4.15}$$

where, for $k = 1,\, 2,\, \ldots,\, n$,

$$\psi_k^{(1b)} = -(n+1)\frac{\psi_k^2}{\ell_k} + 2\psi_k \cdot \frac{\partial \psi_k}{\partial \ell_k} + \psi_k \sum_{b \neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b}.$$

Putting (4.13) and (4.15) into (4.11), we see that the expectation of $n\Delta_1 + \Delta_2$ is given by

$$\mathbb{E}[n\Delta_1 + \Delta_2] \;=\; \mathbb{E}\left[\mathrm{Tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{O}_1 \mathbf{Diag}(\psi_1^{(1)},\, \psi_2^{(1)},\, \ldots,\, \psi_n^{(1)})\boldsymbol{O}_1'\right\}\right],$$

where the $\psi_k^{(1)}$ $(k = 1,\, 2,\, \ldots,\, n)$ is given by (2.3). This completes the proof of this theorem. □

4.4 Proof of Theorem 3.1

Write $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^*$, where $\boldsymbol{A}$ is a $p \times p$ nonsingular complex matrix, and put $\widetilde{\boldsymbol{Z}} = (\tilde{z}_{ij})_{\substack{i=1,2,\ldots,n \\ j=1,2,\ldots,p}} = \boldsymbol{Z}(\boldsymbol{A}^*)^{-1}$. Then $\mathscr{L}(\widetilde{\boldsymbol{Z}}) = \mathbb{C}N_{n\times p}(\boldsymbol{0},\, \boldsymbol{I}_n \otimes \boldsymbol{I}_p)$. Furthermore, put $\boldsymbol{H} = (h_{ij})_{\substack{i=1,2,\ldots,p \\ j=1,2,\ldots,p}} = \boldsymbol{A}^*\boldsymbol{G}(\boldsymbol{A}^*)^{-1}$. We regard h_{ij} as a differentiable functions of $\widetilde{\boldsymbol{Z}}$. Define $\widetilde{\boldsymbol{W}} = \widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}}$. Since

$$\frac{1}{2}\frac{\partial}{\partial(\mathrm{Re}\, \tilde{z}_{ij})} \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}})\right) = -(\mathrm{Re}\, \tilde{z}_{ij}) \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}})\right)$$

and

$$-\frac{\sqrt{-1}}{2}\frac{\partial}{\partial(\mathrm{Im}\, \tilde{z}_{ij})} \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}})\right) = \sqrt{-1}(\mathrm{Im}\, \tilde{z}_{ij}) \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}})\right)$$

for $i = 1,\, 2,\, \ldots,\, n$ and $j = 1,\, 2,\, \ldots,\, p$, we have

$$\frac{\partial}{\partial \tilde{z}_{ij}} \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}})\right) = -\overline{\tilde{z}_{ij}} \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^*\widetilde{\boldsymbol{Z}})\right).$$

Therefore, for a differentiable function $h : \mathbb{C}^{n\times p} \to \mathbb{C}$ and $k = 1, 2, \ldots, n$; $i, j = 1, 2, \ldots, p$, we have

$$\frac{\partial}{\partial \tilde{z}_{ki}} \left[\tilde{z}_{kj} h \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^* \widetilde{\boldsymbol{Z}})\right)\right] = \left\{\delta_{ij} h + \tilde{z}_{kj} \frac{\partial h}{\partial \tilde{z}_{ki}} - \overline{\tilde{z}}_{ki} \tilde{z}_{kj} h\right\} \exp\left(-\mathrm{Tr}\,(\widetilde{\boldsymbol{Z}}^* \widetilde{\boldsymbol{Z}})\right),$$

from which it follows that

$$\mathbb{E}\,[\overline{\tilde{z}_{ki}} \tilde{z}_{kj} h] = \mathbb{E}\left[\delta_{ij} h + \tilde{z}_{kj} \cdot \frac{\partial h}{\partial \tilde{z}_{ki}}\right].$$

Putting $h = h_{jl}$ ($j, l = 1, 2, \ldots, p$) and summing over j from 1 to p and over k from 1 to n, we obtain

$$\mathbb{E}\left[\sum_{j=1}^{p} \tilde{w}_{ij} h_{jl}\right] = \mathbb{E}\left[n h_{il} + \sum_{j=1}^{p} \sum_{k=1}^{n} \tilde{z}_{kj} \cdot \frac{\partial h_{jl}}{\partial \tilde{z}_{ki}}\right].$$

Thus, putting $\boldsymbol{H} = (h_{jl})_{\substack{j=1,2,\ldots,p\\ l=1,2,\ldots,p}}$, we get

$$\mathbb{E}\,[\widetilde{\boldsymbol{W}} \boldsymbol{H}] = \mathbb{E}\left[n\boldsymbol{H} + (\widetilde{\boldsymbol{Z}}' \nabla_{\widetilde{\boldsymbol{Z}}})' \boldsymbol{H}\right]. \tag{4.16}$$

Finally, by the definition of $\boldsymbol{H}$, we have $\mathbb{E}\,[\widetilde{\boldsymbol{W}} \boldsymbol{H}] = \mathbb{E}\,[\boldsymbol{A}^{-1} \boldsymbol{W} \boldsymbol{G} (\boldsymbol{A}^*)^{-1}]$ while, since $\nabla_{\widetilde{\boldsymbol{Z}}} = \nabla_{\boldsymbol{Z}} \overline{\boldsymbol{A}}$, we have $\mathbb{E}\,[(\widetilde{\boldsymbol{Z}}' \nabla_{\widetilde{\boldsymbol{Z}}})' \boldsymbol{H}] = \mathbb{E}\,[\boldsymbol{A}^* (\boldsymbol{Z}' \nabla_{\boldsymbol{Z}})' \boldsymbol{G} (\boldsymbol{A}^*)^{-1}]$. Putting these two equations above into (4.16) and multiplying by $(\boldsymbol{A}^*)^{-1}$ from the left and by $\boldsymbol{A}^*$ from the right, we get $\mathbb{E}\,[(\boldsymbol{A}\boldsymbol{A}^*)^{-1} \boldsymbol{W} \boldsymbol{G}] = \mathbb{E}\,[n\boldsymbol{G} + (\boldsymbol{Z}' \nabla_{\boldsymbol{Z}})' \boldsymbol{G}]$, which completes the proof of (3.1). □

4.5 Proof of Theorem 3.2

To prove Theorem 3.2, we need the following lemma which states the partial derivatives of the eigenvalues and the elements of eigenvectors of the singular complex Wishart matrix $\boldsymbol{W} = \boldsymbol{Z}^* \boldsymbol{Z}$ with respect to the elements of the matrix $\boldsymbol{Z}$. For full-rank complex Wishart matrices, partial derivatives which play a similar role to those in the next lemma appear in [21, 39].

Lemma 4.2. *Assume that* $p > n$. *Let* $\boldsymbol{Z} = (z_{ij})_{\substack{i=1,2,\dots,n\\ j=1,2,\dots,p}}$ *and decompose a* $p \times p$ *matrix* $\boldsymbol{Z}^*\boldsymbol{Z}$ *as* $\boldsymbol{Z}^*\boldsymbol{Z} = \boldsymbol{U}_1\boldsymbol{L}\boldsymbol{U}_1^*$, *where* $\boldsymbol{U}_1 \in \mathbb{C}V_{p,n}$ *and* $\boldsymbol{L} = \mathbf{Diag}\,(\ell_1, \ell_2, \dots, \ell_n)$ *is an* $n \times n$ *diagonal matrix with* $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_n > 0$. *Furthermore, let* $\boldsymbol{U}_2 = (u_{ij})_{\substack{i=1,2,\dots,p\\ j=n+1,2,\dots,p}} \in \mathbb{C}V_{p,p-n}$ *be a* $p \times (p-n)$ *semi-unitary matrix such that* $\boldsymbol{U} = [\boldsymbol{U}_1; \boldsymbol{U}_2]$ *is a* $p \times p$ *unitary matrix. If* $\ell_1 > \ell_2 > \cdots > \ell_n > 0$, *then we have, for* $i, k, m = 1, 2, \dots, n$ *and* $a, j = 1, 2, \dots, p$,

$$
\begin{aligned}
\frac{\partial \ell_m}{\partial z_{ij}} &= \sum_{c_1=1}^{p} \overline{u}_{c_1 m}\overline{z}_{ic_1}u_{jm};\\
\frac{\partial u_{ak}}{\partial z_{ij}} &= \sum_{b\neq k}^{n}\sum_{c_1=1}^{p} \frac{u_{ab}\overline{u}_{c_1 b}u_{jk}\overline{z}_{ic_1}}{\ell_k - \ell_b} + \sum_{b=n+1}^{p}\sum_{c_1=1}^{p} \frac{u_{ab}\overline{u}_{c_1 b}u_{jk}\overline{z}_{ic_1}}{\ell_k};\\
\frac{\partial \overline{u}_{ak}}{\partial z_{ij}} &= \sum_{b\neq k}^{n}\sum_{c_1=1}^{p} \frac{\overline{u}_{ab}u_{jb}\overline{u}_{c_1 k}\overline{z}_{ic_1}}{\ell_k - \ell_b} + \sum_{b=n+1}^{p}\sum_{c_1=1}^{p} \frac{\overline{u}_{ab}u_{jb}\overline{u}_{c_1 k}\overline{z}_{ic_1}}{\ell_k},
\end{aligned}
$$

for $a \neq k$, *and* $\partial u_{kk}/\partial x_{ij} = 0$ *and* $\partial \overline{u}_{kk}/\partial x_{ij} = 0$ *for* $k = 1, 2, \dots, n$.

Proof. Take differentials of

$$
\boldsymbol{Z}^*\boldsymbol{Z} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\boldsymbol{U}^* = [\boldsymbol{U}_1; \boldsymbol{U}_2]\begin{bmatrix} \boldsymbol{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{U}_1^* \\ \boldsymbol{U}_2^* \end{bmatrix},
$$

and multiply $\boldsymbol{U}^*$ and $\boldsymbol{U}$ from the left side and from the right side, respectively. Then we obtain that

$$
\begin{aligned}
&\boldsymbol{U}^*\{(d\boldsymbol{Z}^*)\boldsymbol{Z} + \boldsymbol{Z}^*(d\boldsymbol{Z})\}\boldsymbol{U}\\
&\qquad = \boldsymbol{U}^*(d\boldsymbol{U})\begin{bmatrix} \boldsymbol{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}(d\boldsymbol{U}^*)\boldsymbol{U} + \begin{bmatrix} d\boldsymbol{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.
\end{aligned} \tag{4.17}
$$

But, using that $\boldsymbol{U}^*(d\boldsymbol{U}) + (d\boldsymbol{U}^*)\boldsymbol{U} = \mathbf{0}_{p\times p}$, and comparing the left and right hand sides of (4.17) blockwisely, we get

$$
\boldsymbol{U}_1^*\{(d\boldsymbol{Z}^*)\boldsymbol{Z} + \boldsymbol{Z}^*(d\boldsymbol{Z})\}\boldsymbol{U}_1 = \boldsymbol{U}_1^*(d\boldsymbol{U}_1)\boldsymbol{L} - \boldsymbol{L}\boldsymbol{U}_1^*(d\boldsymbol{U}_1) + d\boldsymbol{L} \tag{4.18}
$$

and

$$\boldsymbol{U}_2^*\{(d\boldsymbol{Z}^*)\boldsymbol{Z} + \boldsymbol{Z}^*(d\boldsymbol{Z})\}\boldsymbol{U}_1 = \boldsymbol{U}_2^*(d\boldsymbol{U}_1)\boldsymbol{L}. \tag{4.19}$$

From (4.18), we get, for $m = 1, 2, \ldots, n$,

$$(d\boldsymbol{L})_{mm} = \{\boldsymbol{U}_1^*((d\boldsymbol{Z}^*)\boldsymbol{Z} + \boldsymbol{Z}^*(d\boldsymbol{Z}))\boldsymbol{U}_1\}_{mm}. \tag{4.20}$$

From (4.19), we get that, for $a = 1, 2, \ldots, n$ and $k = 1, 2, \ldots, n$ such that $a \neq k$,

$$\{\boldsymbol{U}_1^*(d\boldsymbol{U}_1)\}_{ak} = \frac{1}{\ell_k - \ell_a}\{\boldsymbol{U}_1^*((d\boldsymbol{Z}^*)\boldsymbol{Z} + \boldsymbol{Z}^*(d\boldsymbol{Z}))\boldsymbol{U}_1\}_{ak}. \tag{4.21}$$

From (4.20), we get that, for $a = n+1, 2, \ldots, 1$ and $k = 1, 2, \ldots, n$,

$$\{\boldsymbol{U}_2^*(d\boldsymbol{U}_1)\}_{ak} = \frac{1}{\ell_k}\{\boldsymbol{U}_2^*((d\boldsymbol{Z}^*)\boldsymbol{Z} + \boldsymbol{Z}^*(d\boldsymbol{Z}))\boldsymbol{U}_1\}_{ak}. \tag{4.22}$$

From the fact that $d(\operatorname{Re} z_{ij})$ and $d(\operatorname{Im} z_{ij})$ are the dual basis of $\partial/\partial(\operatorname{Re} z_{ij})$ and $\partial/\partial(\operatorname{Im} z_{ij})$ for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$, we have

$$dz\left(\frac{\partial}{\partial z_{kl}}\right) = d(\operatorname{Re} z_{ij}) + \sqrt{-1}d(\operatorname{Im} z_{ij})\}\left(\frac{\partial}{\partial z_{kl}}\right) = \delta_{ik}\delta_{jl} \quad \text{and} \quad d\bar{z}\left(\frac{\partial}{\partial z_{kl}}\right) = 0 \tag{4.23}$$

for $k = 1, 2, \ldots, n$ and $l = 1, 2, \ldots, p$. Using (4.20) and (4.23), we get that, for $i, m = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$,

$$\begin{aligned}
\frac{\partial \ell_m}{\partial z_{ij}} &= (d\boldsymbol{L})_{mm}\left(\frac{\partial}{\partial z_{ij}}\right) \\
&= \sum_{c_1, c_3=1}^{p}\sum_{c_2=1}^{n} \overline{u}_{c_1 m}\{d(\operatorname{Re} z_{c_2 c_1}) - \sqrt{-1}d(\operatorname{Im} z_{c_2 c_1})\}\left(\frac{\partial}{\partial z_{ij}}\right) z_{c_2 c_3} u_{c_3 m} \\
&\quad + \sum_{c_1, c_3=1}^{p}\sum_{c_2=1}^{n} \overline{u}_{c_1 m}\overline{z}_{c_2 c_1}\{d(\operatorname{Re} z_{c_2 c_3}) + \sqrt{-1}d(\operatorname{Im} z_{c_2 c_3})\}\left(\frac{\partial}{\partial z_{ij}}\right) u_{c_3 m} \\
&= \sum_{c_1}^{p} \overline{u}_{c_1 m}\overline{z}_{i c_1} u_{jm},
\end{aligned}$$

which completes the first part of the lemma.

Using (4.21) and (4.23), we get that, for i, $k = 1, 2, \ldots, n$ and a, $j = 1, 2, \ldots, p$,

$$
\begin{aligned}
\frac{\partial u_{ak}}{\partial z_{ij}} &= \sum_{b=1}^{p}\sum_{c_1=1}^{p} u_{ab}\overline{u}_{c_1 b}\frac{\partial u_{c_1 k}}{\partial z_{ij}} \\
&= \sum_{b=1}^{n} u_{ab}\{\boldsymbol{U}_1^*(d\boldsymbol{U}_1)\}_{bk}\left(\frac{\partial}{\partial z_{ij}}\right) + \sum_{b=n+1}^{p} u_{ab}\{\boldsymbol{U}_2^*(d\boldsymbol{U}_1)\}_{bk}\left(\frac{\partial}{\partial z_{ij}}\right) \\
&= \sum_{b\neq k}^{n} u_{ab}\frac{1}{\ell_k-\ell_b}\{\boldsymbol{U}_1^*((d\boldsymbol{Z}^*)\boldsymbol{Z}+\boldsymbol{Z}^*(d\boldsymbol{Z}))\boldsymbol{U}_1\}_{bk}\left(\frac{\partial}{\partial z_{ij}}\right) \\
&\quad + \sum_{b=n+1}^{p} u_{ab}\frac{1}{\ell_k}\{\boldsymbol{U}_2^*((d\boldsymbol{Z}^*)\boldsymbol{Z}+\boldsymbol{Z}^*(d\boldsymbol{Z}))\boldsymbol{U}_1\}_{bk}\left(\frac{\partial}{\partial z_{ij}}\right) \\
&= \sum_{b\neq k}^{n}\sum_{c_1,c_3=1}^{p}\sum_{c_2=1}^{n}\frac{u_{ab}\overline{u}_{c_1 b}u_{c_3 k}}{\ell_k-\ell_b}\{(d\overline{z}_{c_2 c1})z_{c_2 c_3}+\overline{z}_{c_2 c_1}(dz_{c_2 c_3})\}\left(\frac{\partial}{\partial z_{ij}}\right) \\
&\quad + \sum_{b=n+1}^{p}\sum_{c_1,c_3=1}^{p}\sum_{c_2=1}^{n}\frac{u_{ab}\overline{u}_{c_1 b}u_{c_3 k}}{\ell_k}\{(d\overline{z}_{c_2 c1})z_{c_2 c_3}+\overline{z}_{c_2 c_1}(dz_{c_2 c_3})\}\left(\frac{\partial}{\partial z_{ij}}\right) \\
&= \sum_{b\neq k}^{n}\sum_{c_1=1}^{p}\frac{u_{ab}\overline{u}_{c_1 b}u_{jk}\overline{z}_{ic_1}}{\ell_k-\ell_b} + \sum_{b=n+1}^{p}\sum_{c_1=1}^{p}\frac{u_{ab}\overline{u}_{c_1 b}u_{jk}\overline{z}_{ic_1}}{\ell_k},
\end{aligned}
$$

which completes the proof of the second part of the lemma.

Taking the complex conjugate of (4.21) and (4.22), we proceed in a similar way to the lines above to get the third equation of the lemma. □

Proof of Theorem 3.2. We adapt the notation in Theorem 3.1 and Lemma 4.2. Put $\boldsymbol{G} = \boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*$ and apply Theorem 3.1 to get

$$
\mathbb{E}\,[\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*] = \mathbb{E}\,[n\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^* + (\boldsymbol{Z}'\nabla_{\boldsymbol{Z}})'\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*]. \tag{4.24}
$$

In order to apply Lemma 4.2 to the second term inside the right expectation, we first note that, for $i,\, j = 1,\, 2,\, \ldots,\, p$, the $(i,\, j)$ element of $(\boldsymbol{Z}'\nabla_{\boldsymbol{Z}})'\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*$ is given by

$$\begin{aligned} \{(\boldsymbol{Z}'\nabla_{\boldsymbol{Z}})'\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\}_{ij} &= \sum_{c_1,\, c_3=1}^{n}\sum_{c_2=1}^{p} z_{c_1c_2}\frac{\partial}{\partial z_{c_1i}}\{u_{c_2c_3}\psi_{c_3}\overline{u}_{jc_3}\} \\ &= \sum_{c_1,\, c_3=1}^{n}\sum_{c_2=1}^{p} z_{c_1c_2}\psi_{c_3}\overline{u}_{jc_3}\frac{\partial u_{c_2c_3}}{\partial z_{c_1i}} + \sum_{c_1,\, c_3=1}^{n}\sum_{c_2=1}^{p} z_{c_1c_2}\psi_{c_3}u_{c_2c_3}\frac{\partial \overline{u}_{jc_3}}{\partial z_{c_1i}} \\ &\quad + \sum_{c_1,\, c_3,\, m=1}^{n}\sum_{c_2=1}^{p} z_{c_1c_2}u_{c_2c_3}\overline{u}_{jc_3}\frac{\partial \psi_{c_3}}{\partial \ell_m}\cdot\frac{\partial \ell_m}{\partial z_{c_1i}} =: A_4 + A_5 + A_6. \end{aligned} \tag{4.25}$$

From the third equation of Lemma 4.2 and the fact that $\boldsymbol{Z}'\overline{\boldsymbol{Z}} = \overline{\boldsymbol{U}}_1\boldsymbol{L}\boldsymbol{U}_1'$, $\boldsymbol{U}_1'\overline{\boldsymbol{U}}_1 = \boldsymbol{I}_n$, and that $\boldsymbol{U}_2'\overline{\boldsymbol{U}}_1 = \boldsymbol{0}$, we get

$$A_4 = \sum_{c_3=1}^{n} u_{ic_3}\overline{u}_{jc_3}\sum_{b\neq c_3}^{n}\frac{\ell_b\psi_{c_3}}{\ell_{c_3}-\ell_b}.$$

From the second equation of Lemma 4.2 and a similar argument, we get that

$$A_5 = \sum_{b=1}^{n} u_{ib}\overline{u}_{jb}\sum_{c_3\neq b}\frac{\ell_{c_3}\psi_{c_3}}{\ell_{c_3}-\ell_b} + \sum_{b=n+1}^{p} u_{ib}\overline{u}_{jb}\sum_{c_3=1}^{n}\psi_{c_3}.$$

Furthermore, from the first equation of Lemma 4.2, we get that

$$A_6 = \sum_{c_3=1}^{n} u_{ic_3}\overline{u}_{jc_3}\ell_{c_3}\frac{\partial \psi_{c_3}}{\partial \ell_{c_3}}.$$

Putting these three expressions into (4.25), we obtain that

$$\begin{aligned} \{(\boldsymbol{Z}'\nabla_{\boldsymbol{Z}})'\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\}_{ij} &= \sum_{c_3=1}^{n} u_{ic_3}\overline{u}_{jc_3}\left\{\ell_{c_3}\frac{\partial \psi_{c_3}}{\partial \ell_{c_3}} + \sum_{b\neq c_3}^{n}\frac{\ell_b\psi_{c_3}-\ell_b\psi_b}{\ell_{c_3}-\ell_b}\right\} \\ &\quad + \sum_{b=n+1}^{p} u_{ib}\overline{u}_{jb}\sum_{c_3=1}^{n}\psi_{c_3}. \end{aligned}$$

Putting this expression into (4.24), we get that

$$\begin{aligned} &\mathbb{E}\left[\{\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{U}_1\boldsymbol{\Psi}\boldsymbol{U}_1^*\}_{ij}\right] \\ &\quad = \mathbb{E}\left[\sum_{k=1}^{n} u_{ik}\overline{u}_{jk}\left\{n\psi_k + \ell_k\frac{\partial \psi_k}{\partial \ell_k} + \sum_{b\neq k}^{n}\frac{\ell_b\psi_k-\ell_b\psi_b}{\ell_k-\ell_b}\right\} + \sum_{b=n+1}^{p} u_{ib}\overline{u}_{jb}\sum_{k=1}^{n}\psi_k\right]. \end{aligned}$$

Finally, changing ψ_k into $\ell_k^{-1}\psi_k$, noting that $\boldsymbol{U}_2\boldsymbol{U}_2^* = \boldsymbol{I}_p - \boldsymbol{U}_1\boldsymbol{U}_1^*$, and using (4.9), we can complete the proof of this theorem. □

4.6 Proof of Theorem 3.3

Write

$$\begin{aligned}\boldsymbol{F} &= (f_{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}} = \boldsymbol{U}_1\mathbf{Diag}\big(\frac{\psi_1}{\ell_1},\, \frac{\psi_2}{\ell_2},\, \ldots,\, \frac{\psi_n}{\ell_n}\big)\boldsymbol{U}_1^*;\\ \widetilde{\boldsymbol{F}} &= (\tilde{f}_{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}} = \boldsymbol{U}_1\mathbf{Diag}\big(\psi_1,\, \psi_2,\, \ldots,\, \psi_n\big)\boldsymbol{U}_1^*.\end{aligned}$$

Apply Theorem 3.1 with $\boldsymbol{G} = \boldsymbol{F}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}}$ to get that

$$\mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}})\right] = \mathbb{E}\left[\mathrm{Tr}\,(\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\boldsymbol{F}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}})\right] =: \mathbb{E}[n\Delta_3 + \Delta_4], \tag{4.26}$$

where $\Delta_3 = \mathrm{Tr}\left\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\mathbf{Diag}\big(\psi_1^2/\ell_1,\, \psi_2^2/\ell_2,\, \ldots,\, \psi_n^2/\ell_n\big)\boldsymbol{U}_1^*\right\}$ and $\Delta_4 = \mathrm{Tr}\left\{(\boldsymbol{Z}'\nabla_{\boldsymbol{Z}})'\boldsymbol{F}\boldsymbol{\Sigma}^{-1}\widetilde{\boldsymbol{F}}\right\}$. We evaluate the expectation of Δ_4 in (4.26). For $\boldsymbol{\Sigma}^{-1} = (\sigma^{ij})_{\substack{i=1,2,\ldots,p\\ j=1,2,\ldots,p}}$, we see that the expectation of Δ_4 is given by

$$\begin{aligned}\mathbb{E}[\Delta_4] &= \mathbb{E}\Big[\sum_{c_1=1}^{n}\sum_{c_2,\,c_3,\,c_4,\,i=1}^{p} z_{c_1c_2}\frac{\partial}{\partial z_{c_1 i}}\big(f_{c_2c_3}\sigma^{c_3c_4}\widetilde{f}_{c_4 i}\big)\Big]\\ &=: \mathbb{E}\Big[\sum_{c_2,\,c_3,\,c_4,=1}^{p}\sigma^{c_3c_4}f_{c_2c_3}T^{(3)}_{c_4c_2} + \sum_{c_3,\,c_4=1}^{p}\sigma^{c_3c_4}T^{(4)}_{c_4c_3}\Big],\end{aligned} \tag{4.27}$$

where, for $c_2,\, c_3,\, c_4 = 1,\, 2,\, \ldots,\, p$,

$$T^{(3)}_{c_4c_2} = \sum_{c_1=1}^{n}\sum_{i=1}^{p} z_{c_1c_2}\frac{\partial\,\widetilde{f}_{c_4 i}}{\partial z_{c_1 i}}; \qquad \text{and} \qquad T^{(4)}_{c_4c_3} = \sum_{c_1=1}^{n}\sum_{c_2,\,i=1}^{p} z_{c_1c_2}\widetilde{f}_{c_4 i}\frac{\partial f_{c_2c_3}}{\partial z_{c_1 i}}.$$

Next we evaluate $T^{(3)}_{c_4c_2}$ and $T^{(4)}_{C_4c_3}$, respectively. Since $\widetilde{\boldsymbol{F}} = \boldsymbol{U}_1\mathbf{Diag}(\psi_1,\, \psi_2,\, \ldots,\, \psi_n)\boldsymbol{U}_1^*$, we have

$$\begin{aligned}T^{(3)}_{c_4c_2} &= \sum_{c_1=1}^{n}\sum_{i=1}^{p} z_{c_1c_2}\frac{\partial}{\partial z_{c_1 i}}\Big(\sum_{c_5=1}^{n} u_{c_4c_5}\psi_{c_5}\overline{u}_{ic_5}\Big)\\ &= \sum_{c_1,\,c_5=1}^{n}\sum_{i=1}^{p} z_{c_1c_2}\psi_{c_5}\overline{u}_{ic_5}\frac{\partial u_{c_4c_5}}{\partial z_{c_1 i}} + \sum_{c_1,\,c_5=1}^{n}\sum_{i=1}^{p} z_{c_1c_2}\psi_{c_5}u_{c_4c_5}\frac{\partial \overline{u}_{ic_5}}{\partial z_{c_1 i}}\\ &\quad + \sum_{c_1,\,c_5,\,m=1}^{n}\sum_{i=1}^{p} z_{c_1c_2}u_{c_4c_5}\overline{u}_{ic_5}\frac{\partial\psi_{c_5}}{\partial\ell_m}\cdot\frac{\partial\ell_m}{\partial z_{c_1 i}} =: T^{(31)}_{c_4c_2} + T^{(32)}_{c_4c_2} + T^{(33)}_{c_4c_2}.\end{aligned} \tag{4.28}$$

Applying Lemma 4.2 and using the fact that $\boldsymbol{U}_1^*\boldsymbol{U}_1 = \boldsymbol{I}_n$ and that $\boldsymbol{Z}^*\boldsymbol{Z} = \boldsymbol{U}_1\boldsymbol{L}\boldsymbol{U}_1^*$, we have

$$(T^{(31)}_{c_4c_2})_{\substack{c_4=1,2,\ldots,p\\ c_2=1,2,\ldots,p}} = \boldsymbol{U}_1\mathbf{Diag}\big(\sum_{b\neq 1}\frac{\ell_1\psi_b}{\ell_b-\ell_1},\, \sum_{b\neq 2}\frac{\ell_2\psi_b}{\ell_b-\ell_2},\, \ldots,\, \sum_{b\neq n}\frac{\ell_n\psi_b}{\ell_b-\ell_n}\big)\boldsymbol{U}_1^*.$$

From Lemma 4.2, the fact that $\boldsymbol{Z}^*\boldsymbol{Z}\boldsymbol{U}_2^*\boldsymbol{U}_2 = \mathbf{0}_{p\times(p-n)}$, and a similar argument, we have

$$(T^{(32)}_{c_4c_2})_{\substack{c_4=4,2,\dots,p\\c_2=1,2,\dots,p}} = \boldsymbol{U}_1\mathbf{Diag}\Big(\sum_{b\neq 1}\frac{\ell_1\psi_1}{\ell_1-\ell_b},\ \sum_{b\neq 2}\frac{\ell_2\psi_2}{\ell_2-\ell_b},\ \dots,\ \sum_{b\neq n}\frac{\ell_n\psi_n}{\ell_n-\ell_b}\Big)\boldsymbol{U}_1^*$$

$$+\,(p-n)\boldsymbol{U}_1\mathbf{Diag}(\psi_1,\,\psi_2,\,\dots,\,\psi_n)\boldsymbol{U}_1^*.$$

Finally, by Lemma 4.2, we have

$$(T^{(33)}_{c_4c_2})_{\substack{c_4=1,2,\dots,p\\c_2=1,2,\dots,p}} \;=\; \boldsymbol{U}_1\mathbf{Diag}\Big(\ell_1\frac{\partial\psi_1}{\partial\ell_1},\,\ell_2\frac{\partial\psi_2}{\partial\ell_2},\,\dots,\,\ell_n\frac{\partial\psi_n}{\partial\ell_n}\Big)\boldsymbol{U}_1^*.$$

Putting these three expressions into (4.28), we get that

$$\mathbb{E}\Big[\sum_{c_2,c_3,c_4=1}^{p}\sigma^{c_3c_4}f_{c_2c_3}T^{(3)}_{c_4c_2}\Big] = \mathbb{E}\big[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\mathbf{Diag}\big(\widetilde{\psi}_1^{(1a)},\,\widetilde{\psi}_2^{(1a)},\,\dots,\,\widetilde{\psi}_n^{(1a)}\big)\boldsymbol{U}_1^*\}\big]. \tag{4.29}$$

where, for $i = 1, 2, \dots, n$,

$$\widetilde{\psi}_k^{(1a)} = \psi_k\sum_{b\neq k}^{n}\frac{\psi_k-\psi_b}{\ell_k-\ell_b} + \psi_k\frac{\partial\psi_k}{\partial\ell_k} + (p-n)\frac{\psi_k^2}{\ell_k}.$$

Since $\boldsymbol{F} = \boldsymbol{U}_1\mathbf{Diag}\big(\ell_1^{-1}\psi_1,\,\ell_2^{-1}\psi_2,\,\dots,\,\ell_n^{-1}\psi_n\big)\boldsymbol{U}_1^*$, we have

$$\begin{aligned}
T^{(4)}_{c_4c_3} &= \sum_{c_1,c_5=1}^{n}\sum_{c_2,i=1}^{p} z_{c_1c_2}\widetilde{f}_{c_4i}\frac{\partial}{\partial z_{c_1i}}\big(u_{c_2c_5}\frac{\psi_{c_5}}{\ell_{c_5}}\overline{u}_{c_3c_5}\big)\\
&= \sum_{c_1,c_5=1}^{n}\sum_{c_2,i=1}^{p} z_{c_1c_2}\widetilde{f}_{c_4i}\overline{u}_{c_3c_5}\frac{\psi_{c_5}}{\ell_{c_5}}\cdot\frac{\partial u_{c_2c_5}}{\partial z_{c_1i}} + \sum_{c_1,c_5=1}^{n}\sum_{c_2,i=1}^{p} z_{c_1c_2}\widetilde{f}_{c_4i}u_{c_2c_5}\frac{\psi_{c_5}}{\ell_{c_5}}\cdot\frac{\partial\overline{u}_{c_3c_5}}{\partial z_{c_1i}}\\
&\quad+\sum_{c_1,c_5,m=1}^{n}\sum_{c_2,i=1}^{p} z_{c_1c_2}\widetilde{f}_{c_4i}u_{c_2c_5}\overline{u}_{c_3c_5}\frac{\partial\ell_m}{\partial z_{c_1i}}\cdot\frac{\partial}{\partial\ell_m}\left(\frac{\psi_{c_5}}{\ell_{c_5}}\right)\\
&=: T^{(41)}_{c_4c_3} + T^{(43)}_{c_4c_3} + T^{(43)}_{c_4c_3}.
\end{aligned} \tag{4.30}$$

By Lemma 4.2, we have

$$\begin{aligned}
(T^{(41)}_{c_4c_3})_{\substack{c_4=4,2,\dots,p\\c_3=1,2,\dots,p}} &= \boldsymbol{U}_1\mathbf{Diag}\Big(\sum_{b\neq 1}\frac{\ell_b\psi_1^2}{(\ell_1-\ell_b)\ell_1},\,\sum_{b\neq 2}\frac{\ell_b\psi_2^2}{(\ell_2-\ell_b)\ell_2},\,\dots,\,\sum_{b\neq n}\frac{\ell_b\psi_n^2}{(\ell_n-\ell_b)\ell_n}\Big)\boldsymbol{U}_1^*;\\
(T^{(42)}_{c_4c_3})_{\substack{c_4=4,2,\dots,p\\c_3=1,2,\dots,p}} &= \boldsymbol{U}_1\mathbf{Diag}\Big(\sum_{b\neq 1}\frac{\psi_1\psi_b}{\ell_b-\ell_1},\,\sum_{b\neq 2}\frac{\psi_2\psi_b}{\ell_b-\ell_2},\,\dots,\,\sum_{b\neq n}\frac{\psi_n\psi_b}{\ell_b-\ell_n}\Big)\boldsymbol{U}_1^*;\\
(T^{(43)}_{c_4c_3})_{\substack{c_4=4,2,\dots,p\\c_3=1,2,\dots,p}} &= \boldsymbol{U}_1\mathbf{Diag}\Big(\ell_1\psi_1\frac{\partial}{\partial\ell_1}\left(\frac{\psi_1}{\ell_1}\right),\,\ell_2\psi_2\frac{\partial}{\partial\ell_2}\left(\frac{\psi_2}{\ell_2}\right),\,\dots,\,\ell_n\psi_n\frac{\partial}{\partial\ell_n}\left(\frac{\psi_n}{\ell_n}\right)\Big)\boldsymbol{U}_1^*.
\end{aligned}$$

Putting the above three equations into (4.30), we see that the second term in the right hand side of (4.27) is given by

$$\mathbb{E}\left[\sum_{c_3,\,c_4=1}^{p} \sigma^{c_3c_4} T^{(4)}_{c_4c_3}\right] = \mathbb{E}\big[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\mathbf{Diag}(\widetilde{\psi}_1^{(1b)},\,\widetilde{\psi}_2^{(1b)},\,\ldots,\,\widetilde{\psi}_n^{(1b)})\boldsymbol{U}_1^*\}\big], \tag{4.31}$$

where, for $k = 1,\,2,\,\ldots,\,k$,

$$\widetilde{\psi}_k^{(1b)} = -n\frac{\psi_k^2}{\ell_k} + \psi_k \cdot \frac{\partial \psi_k}{\partial \ell_k} + \psi_k \sum_{b\neq k}^{n} \frac{\psi_k - \psi_b}{\ell_k - \ell_b}.$$

Putting (4.29) and (4.31) into (4.26), we see that the expectation of $n\Delta_3 + \Delta_4$ is given by

$$\mathbb{E}[n\Delta_3 + \Delta_4] \;=\; \mathbb{E}\big[\mathrm{Tr}\,\{\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1\mathbf{Diag}(\widetilde{\psi}_1^{(1)},\,\widetilde{\psi}_2^{(1)},\,\ldots,\,\widetilde{\psi}_n^{(1)})\boldsymbol{U}_1^*\}\big], \tag{4.32}$$

where $\widetilde{\psi}_k^{(1)}$'s, $k = 1,\,2,\,\ldots,\,n$, are given by (3.4). This completes the proof of this theorem. □

Acknowledgements

This work was in part supported by the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research (C) (No.17500185).

References

[1] Andersen, H. H., Højbjerre, M., Sørensen, D. and Eriksen, P. S. (1995). Linear and Graphical Models. Springer-Verlag, New York.

[2] Bickel, P. J. and Levina, L. (2008). Regularized estimation of large covariance matrices. Ann. Statist. **36** 199-227.

[3] Consonni, G. and Veronese, P. (2003). Enriched conjugate and reference priors for the Wishart family on symmetric cones. Ann. Statist. **31** 1491-1516.

[4] Daniels, M. J. and Kass, R. E. (1999). Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models. J. Amer. Statist. Assoc. **94** 1254–1263.

[5] Daniels, M. J. and Kass, R. E. (2001). Shrinkage estimators for covariance matrices. Biometrics **57** 1173-1184.

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

[7] Efron, B. and Morris, C. (1976). Multivariate empirical Bayes and estimation of covariance matrices. Ann. Statist. **4** 22–32.

[8] Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction). Ann. Math. Statist. **34** 152-176.

[9] Graczyk, P., Letac, G. and Massam, H. (2003). The complex Wishart distribution and symmetric group. Ann. Statist. **31** 287-309.

[10] Haff, L. R. (1979a). An identity for the Wishart distribution with application. J. Multivariate Anal. **9** 531–542.

[11] Haff, L. R. (1979b). Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity. Ann. Statist. **7** 1264-1276.

[12] Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. Ann. Statist. **8** 586-697.

[13] Haff, L. R. (1988). The variational form of certain Bayes estimators. Unpublished manuscript.

[14] Haff, L. R. (1991). The variational form of certain Bayes estimators. Ann. Statist. **19** 1163-1190.

[15] Huang, L., Liu, N., Pourahamadi, M. and Liu, L. (2006). Covariance matrix selection and estimation via penalized normal likelihood. Biometrika **93** 83-98.

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

[17] Kay, S. M. (1993). Fundamentals of statistical signal processing: Estimation theory. Prentice Hall PTR.

[18] Khatri, C. G. (1965). Classical statistical analysis based on a certain multivariate complex Gaussian distribution. Ann. Math. Statist. **36** 98-114.

[19] Konno, Y. (1992). Improved estimation of matrix of normal mean and eigenvalues in the multivariate F-distributions. Doctoral dissertation, University of Tsukuba.

[20] Konno, Y. (2007a). Estimation of normal covariance matrices parametrized by irreducible symmetric cones under Stein's loss. J. Multivariate Anal. **98** 295-316.

[21] Konno, Y. (2007b). Improving on the sample covariance matrix for a complex elliptically contoured distribution. J. Statist. Plan. Infer. **137** 2475-2486.

[22] Kubokawa, T. and Srivastava, M. (2008). Estimation of the precision matrix of a singular wishart distribution and its application in high dimensional data. J. Multivariate Anal. (In press).

[23] Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. J. Multivariate Anal. **88** 365-411.

[24] Lin, S. P. and Perlman, M. D. (1985). A Monte Carlo comparison of four estimators for a covariance matrix. Multivariate Analysis VI (P.R. Krishnaiah, ed.) 411-426, North-Holland, Amsterdam.

[25] Loh, W. L. (1988). Estimating covariance matrices. Ph.D. thesis, Stanford University.

[26] Loh, W. L. (1991a). Estimating covariance matrices. Ann. Statist. **19** 283–296.

[27] Loh, W. L. (1991b). Estimating covariance matrices II. J. Multivariate Anal. **36** 263–174.

[28] Maiward, D. and Kraus, D. (2000). Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices. Radar, Sonar and Navigation, IEEE Proceedings **147** 162 - 168

[29] Muirhead, R. J. (1982). Aspects of Multivariate Statistical Analysis. John Wiley & Sons, Inc.

[30] Olkin, I. and Selliah, J. B. (1977). Estimating covariances in a multivariate normal distribution. Statistical decision theory and related topics II 313–326. Academic Press, New York.

[31] Ratnaraja, T. and Vaillancourt, R. (2005). Complex singular Wishart matrices and applications. Comput. Math. Appl. **50** 399-411.

[32] Sheena, Y. (1995). Unbiased estimator of risk for an orthogonally invariant estimator of a covariance matrix. J. JAPAN STATIST. SOC. **25** 35-48.

[33] Srivastava, M. S. (1965). On the complex Wishart distribution. ANN. MATH. STATIST. **36** 313-315.

[34] Srivastava, M. (2003). Singular Wishart and multivariate beta distributions. ANN. STATIST. **31** 1537-1560.

[35] Srivastava, M. and Khatri, C. G. (1979). AN INTRODUCTION TO MULTIVARIATE STATISTICS. North-Holland, New York.

[36] Stein, C. (1973). Estimation of the mean of a multivariate normal distribution. PROC. PRAGUE SYMP. ASYMPTOTIC STATIST. 35–381.

[37] Stein, C. (1977). Lectures on the theory of estimation of many parameters. STUDIES IN THE STATISTICAL THEORY OF ESTIMATION I (I.A. Ibragimov and M.S. Nikulin, eds.), Proceedings of Scientific Seminars of the Steklov Institute, Leningrad Division **74**, 4-65 (In Russia). English translation of this article is available in J. SOV. MATH. **34** (1986) 1373-1403.

[38] Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. ANN. STATIST. **9** 1135–1151.

[39] Svensson, L. (2004). A useful identity for complex Wishart forms. Technical reports, Department of Signals and Systems, Chalmers University of Technology, December 2004.

[40] Svensson, L. and Lundberg, M. (2004). Estimating complex covariance matrix. Signals, Systems and Computers, Conference Record of the Thirty-Eighth Asilomar Conference on 7-10, 2151-2154.

[41] Tague, J. A. and Caldwell, C. (1994). Expectations of useful complex Wishart forms. Multidimens, Syst. Signal Process. **5** 263-279.

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