การปรับของวิธีภาวะน่าจะเป็นสูงสุดสำหรับตัวแบบเฟ-เฮรอทแบบหลายตัวแปร

นายอรรณพ อังกูรสิทธิ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาคณิตศาสตร์ประยุกต์และวิทยาการคณนา

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2562

ADJUSTMENT OF MAXIMUM LIKELIHOOD METHOD FOR MULTIVARIATE FAY-HERRIOT MODEL

Mr. Annop Angkunsit

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Applied Mathematics and Computational Science

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2019

Thesis Title	ADJUSTMENT OF MAXIMUM LIKELIHOOD METHOD FOR MULTIVARIATE FAY-HERRIOT MODEL
By	Mr. Annop Angkunsit
Field of Study	Applied Mathematics and Computational Science
Thesis Advisor	Assistant Professor Jiraphan Suntornchost, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.... Dean of the Faculty of Science

(Professor Polkit Sangvanich, Ph.D.)

THESIS COMMITTEE

.... Chairman

(Professor Kritsana Neammanee, Ph.D.)

.... Thesis Advisor

(Assistant Professor Jiraphan Suntornchost, Ph.D.)

.... Examiner

(Associate Professor Nattakarn Chaidee, Ph.D.)

.... External Examiner

(Assistant Professor Thidaporn Supapakorn, Ph.D.)

อรรณพ อังกูรสิทธิ์ : การปรับของวิธีภาวะน่าจะเป็นสูงสุดสำหรับตัวแบบเฟ-เฮรอทแบบหลายตัวแปร. (ADJUSTMENT OF MAXIMUM LIKELIHOOD METHOD FOR MULTIVARIATE FAY-HERRIOT MODEL) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร.จิราพรรณ สุนทรโชติ, 169 หน้า.

ตัวแบบระดับพื้นที่ที่ใช้กันอย่างแพร่หลายในการประมาณค่าพื้นที่ขนาดเล็กคือตัวแบบเฟ-เฮรอท ซึ่งเสนอโดยเฟและเฮรอท ซึ่งถูกใช้ครั้งแรกในการประมาณรายได้เฉลี่ยต่อคนสำหรับสถานที่เล็กๆ (ประชากรน้อยกว่า 1,000 คน) ของประเทศสหรัฐอเมริกา ในบริบทของตัวแบบเฟ-เฮรอท วิธีการแบบดั้งเดิมในการได้รับการประมาณของค่าเฉลี่ยประชากรคือตัวประมาณไม่เอนเอียงเชิงเส้นที่ดีที่สุดเชิงประจักษ์ ค่าประมาณนี้สามารถแสดงเป็นผลรวมน้ำหนักถ่วงของตัวประมาณการสำรวจโดยตรงและตัวประมาณการถดถอย ปัญหาหนึ่งที่ได้รับความสนใจคือการประมาณค่าความแปรปรวนของผลกระทบแบบสุ่มของพื้นที่ในน้ำหนักถ่วงของตัวทำนายไม่เอนเอียงเชิงเส้นที่ดีที่สุดเชิงประจักษ์ อย่างไรก็ตามในบางกรณีน้ำหนักถ่วงของตัวประมาณการสำรวจโดยตรงมีค่าเป็นศูนย์และตัวทำนายไม่เอนเอียงเชิงเส้นที่ดีที่สุดเชิงประจักษ์จะลดลงเป็นตัวประมาณการถดถอย ซึ่งเป็นค่าประมาณที่ไม่น่าพอใจเนื่องจากจะละเลยกลุ่มตัวอย่างจากข้อมูลการสำรวจ ต่อมา ลิและลาฮิรินำเสนอตัวประมาณค่าความแปรปรวนต้องกันของภาวะน่าจะเป็นสูงสุดแบบปรับที่ประมาณเป็นบวก การปรับเหล่านี้จะป้องกันน้ำหนักถ่วงของตัวประมาณการสำรวจโดยตรงที่เป็นศูนย์ ในการศึกษานี้เราได้ขยายวิธีการเพื่อปรับวิธีภาวะน่าจะเป็นสูงสุดแบบปรับสำหรับตัวแบบเฟ-เฮรอทแบบหลายตัวแปร

ภาควิชา	คณิตศาสตร์และวิทยาการคอมพิวเตอร์	ลายมือชื่อนิสิต ....
		ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
สาขาวิชา	คณิตศาสตร์ประยุกต์และวิทยาการคณนา	
ปีการศึกษา	2562	

6172097523 : MAJOR APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE
KEYWORDS : SMALL AREA ESTIMATION / MULTIVARIATE FAY-HERRIOT MODEL / ADJUSTED MAXIMUM LIKELIHOOD METHOD

ANNOP ANGKUNSIT : ADJUSTMENT OF MAXIMUM LIKELIHOOD METHOD FOR MULTIVARIATE FAY-HERRIOT MODEL. ADVISOR : ASST. PROF. JIRAPHAN SUNTORNCHOST, Ph.D., 169 pp.

The most widely used area-level model in small area estimation is the Fay-Herriot model, proposed by Fay and Herriot. It was used first to estimate average per capita income for small places (population less than 1,000) of the USA. In the context of the Fay-Herriot model, the traditional method in obtaining estimation of the population mean is the empirical best linear unbiased prediction (EBLUP) estimator. The estimate can be expressed as a weighted sum of the direct survey estimator and regression estimator. One problem that has received attention is the estimation of variance of the area random effects in the weight of EBLUP. However, in some cases, the weight of the direct survey estimator is zero and the EBLUP reduces to regression estimator, which is undesirable estimates because it ignores the sample from survey data. Later on, Li and Lahiri proposed adjusted maximum likelihood consistent variance estimators with positive estimates. These adjustments prevent zero weight of the direct estimator. In this study, we extend their methods to adjust the adjusted maximum likelihood method for the multivariate Fay-Herriot model.

Department	: Mathematics and Computer Science	Student's Signature	....
		Advisor's Signature	....
Field of Study	: Applied Mathematics and Computational Science		
Academic Year	: 2019		

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my advisor, Assistant Professor Jiraphan Suntornchost for her support, patience, and encouragement throughout my graduate studies. She is always a great source of knowledge and inspiration for me. From her, I not only learn how to do good research but also how to be a good statistician. I am most grateful for her teaching and advice. I would not have achieved this far and this thesis would not have been completed without all the support that I have always received from her.

I would like to thank and deep appreciation to my thesis committees, Professor Kritsana Neammanee, Associate Professor Nattakarn Chaidee, and Assistant Professor Thidaporn Supapakorn for their suggestions and comments on my thesis. Moreover, I would like to thank all teachers who have instructed and taught me for valuable knowledge.

This thesis is funded by the Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, and "The His Royal Highness Crown Prince Maha Vajiralongkorn Scholarship from the Graduate School, Chulalongkorn University to commemorate the 72nd anniversary of his Majesty King Bhumibol Aduladej is gratefully acknowledged".

CONTENTS

LIST OF TABLES

Table | Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This chapter introduces the concept of small area estimation, multivariate small area estimation, problem in small area estimation and outline of this thesis.

1.1 Small Area Estimation

Survey data are now widely used in practice, which does not only provide the estimation for the total population but also for subpopulations (also known as domains). Domains can be defined by geographic areas or socio-demographic groups or other subpopulations. For example, a geographic domain (area) includes a state or province, school district, municipality, metropolitan area and country, and socio-demographic domain may be defined by an age-sex-race group within a large geographic area [34]. In the context of sample surveys, we refer to domain or area estimators that are obtained based only on the domain-specific sample data as “direct”. A direct (domain) estimator maybe use the known auxiliary information related to the variable of interest. A direct estimator is effective when the domain-specific sample sizes are large. However, in practice, when the domain-specific sample sizes are not large enough (known as small area) to provide a sufficiently precise direct estimator. In such situations, small area estimation (SAE) methods can be used to get precise estimates of the parameter of interest from related areas. It is necessary to use “indirect” (domain) estimators that “borrow strength” from the sample of other areas through appropriate linking models. For example, synthetic estimators are indirect estimators that are used under the assumption that small areas have the same characteristics as larger areas, and composite estimators are weighted averages of direct estimators and synthetic estimators. Indirect estimators provide better precision than direct estimators. However, both of these estimators suffer from design-bias, which does not necessarily decrease as the sample size increases.

Some SAE techniques make use of mixed models by incorporating random effects. There are two main kinds of mixed models which are used in small area estimation. The first type is the unit level model that can be used when the information at the unit level is available. The second type is the area-level model. The area-level models are developed using direct area-specific estimates available from a survey dataset and area-level covariates available from a census or administrative dataset. In 1979, Fay and Herriot [15] introduced a basic area-level model to obtain estimates for mean per capita income for small areas (with a population less than 1,000) in the USA. In the Fay-Herriot model, the direct survey estimates of a single response variable in each area are linear regression on area-specific covariates obtained from census data. The Fay-Herriot model is one of the most popular methods used in Small area estimation because of its flexibility in combining and explaining different sources of information. Moreover, it does not require unit-level data. Various extensions of the basic area-level model have been proposed to handle correlated sampling errors, the spatial dependence of the model errors, time series and cross-sectional data [34].

1.2 Multivarite Small Area Estimation

Statisticians use the Fay-Herriot model for one variable or separately for each variable. However, in some surveys, it may be desirable to consider two or more correlated response variables together. In such cases, the multivariate Fay-Herriot model (see [9] and [14]) can be used to incorporate the correlation among the response variables. When multiple dependent variables are correlated, multivariate Fay-Herriot models may produce better results than univariate Fay-Herriot models, but these models have not been received much attention.

Several applications of multivariate Fay-Herriot models have been introduced in the literature. For example, Datta et al. [10] applied a multivariate Fay-Herriot model to obtain hierarchical Bayes estimates of median income of four-person families for the US states; González and Manteiga et al. [16] studied a class of multivariate Fay-Herriot model with a common random effect for all the components of the target vector; Benavent and

Morales [1] studied a class of multivariate Fay-Herriot model with one random effect per components of the target vector and allowing for different covariance patterns between the components of the vector of random effects; Nesa [30] used a multivariate Fay-Herriot model to get improved estimates of health indicator; and Ubaidillah et al. [40] used a multivariate Fay-Herriot model to estimated household consumption per capita expenditure (HCPE) on food and HCPE of non-food in Indonesia.

1.3 Adjusted Method for Estimation of Variance Components

An important problem in the analysis of Fay-Herriot models is the estimation of the variance components. The standard methods for estimation of variance components have been considered in the literature. The methods include the Prasad-Rao simple method-of-moments [32], the Fay-Herriot method-of-moments [15], the profile maximum likelihood method [17] and the residual maximum likelihood method [31]. The estimators are all consistent, under certain regularity conditions. However, in some cases, all of the methods could produce zero estimates of variance components. Li and Lahiri [27] and Morris and Tang [29] proposed the adjusted maximum likelihood methods. These methods produce a strictly positive estimates of variance components. Later on, Yoshimiri and Lahiri [41] developed the adjusted maximum likelihood methods to obtain new adjusted maximum likelihood methods. In this thesis, we consider the variance components estimation methods for the multivariate Fay-Herriot model, such as the profile maximum likelihood and the residual maximum likelihood methods. These methods could produce zero estimates of variance components for some components. Then we propose the adjusted maximum likelihood method for the multivariate Fay-Herriot model.

1.4 Outline of the Thesis

In Chapter 2, we list all fundamental concepts of matrix algebra, matrix analysis, random vectors and matrices, mathematical statistics, mathematical analysis, and the concept of small area estimation. The concept of small area estimation include the Fay-Herriot model, empirical best linear unbiased prediction (EBLUP), variance components estimation and uncertainty of EBLUP.

In Chapter 3, we first review the multivariate Fay-Herriot model and variance components estimation for multivariate Fay-Herriot model, including the profile maximum likelihood (PML) method, residual maximum likelihood (REML) method and adjusted maximum likelihood method of Li and Lahiri (AML.LL) [27]. In this chapter, we propose an extension of the adjusted maximum likelihood (AML) method for multivariate Fay-Herriot model. Moreover, we derive the properties, including biases and consistency of the obtained estimators and compare the obtained estimators in simulation.

In Chapter 4, we present the uncertainty of EBLUP and prediction interval. The uncertainty of EBLUP follow Benavent and Morales [1], Datta et al. [8], and Datta and Lahiri [11]. We compare the uncertainty of EBLUP based on the PML, REML, and two AML estimators both in theory and simulation. The prediction interval include the Cox's empirical Bayes prediction interval in [5], traditional prediction interval, and parametric bootstrap prediction interval in [4]. We compare the prediction interval using EBLUP based on the PML, REML, and two AML estimators in terms of simulation.

In Chapter 5, we present a data analysis for the average household income and average household expenditure in Thailand data set. We apply the AML.LL and AML methods for this data set. The positiveness of the AML.LL and AML estimators play a vital role since, for this data set, the PML and REML methods could be zero. We observe that AML methods produce EBLUP's which generally put more weight on the direct survey estimates than the corresponding EBLUP's that use the PML and REML methods.

Finally, we give conclusions and future work of this thesis in Chapter 6.

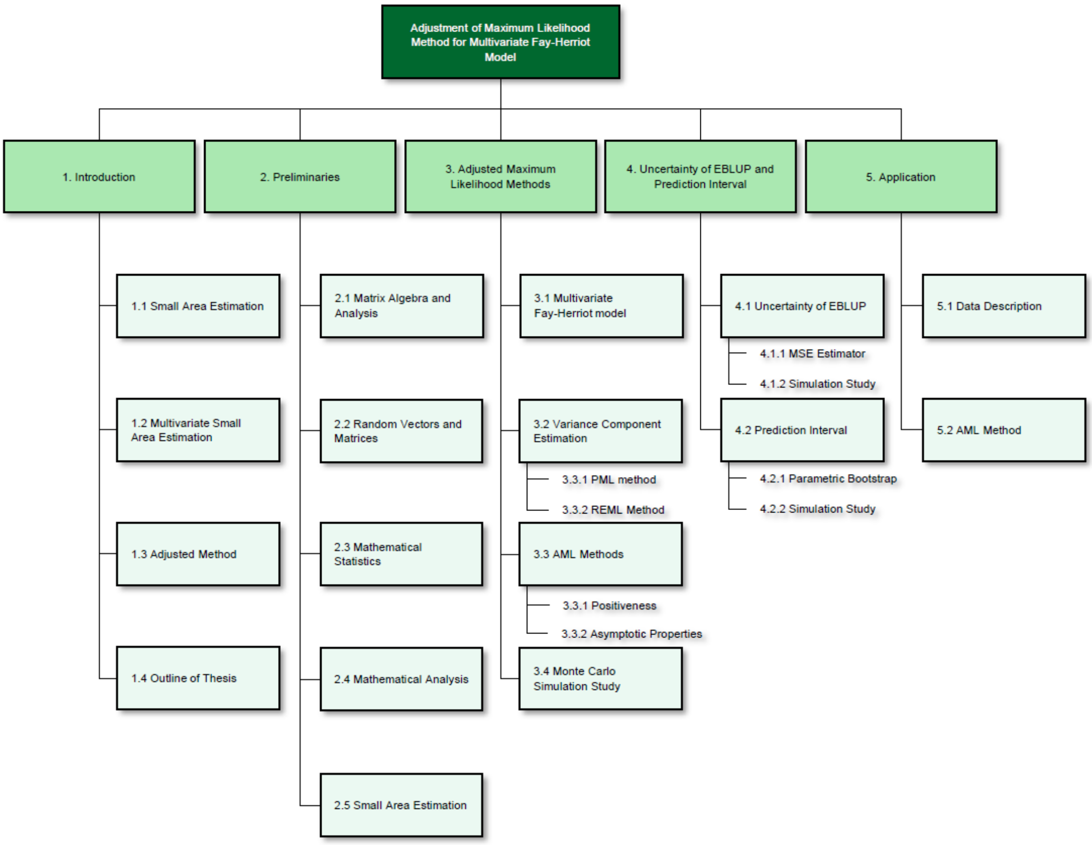

Figure 1.1: The diagram of thesis

CHAPTER II

PRELIMINARIES

In this chapter, we provide definitions and some theorems associated with matrix algebra, matrix analysis, random vector and matrices, mathematical statistic, mathematical analysis, and small area estimation which will be used in Chapter 3.

2.1 Matrix Algebra and Analysis

Since any linear model can be represented by a matrix form, knowledge in matrix algebra is needed in studying a linear model. In this section, we review some well-known definitions and theorems of matrix algebra and analysis needed in this study. More intensive reviews of these topics are provided in [2], [19], [35] and [37].

In this thesis, an $n \times n$ identity matrix, n zero vector and $n \times n$ zero matrix are denoted by $\mathbf{I}_n$, $\mathbf{0}_n$ and $\mathbf{0}_{n\times n}$, respectively.

2.1.1 Transpose

Definition 2.1. The **transpose** of a matrix $\mathbf{A} = (a_{ij})$ is obtained by interchanging the rows and columns of $\mathbf{A}$, denoted by $\mathbf{A}'$, that is $\mathbf{A}' = (a_{ij})' = (a_{ji})$.

Theorem 2.2. If $\mathbf{A}$ and $\mathbf{B}$ are both $m \times n$ matrices and $\mathbf{C}$ is an $n \times p$ matrix, then

(i) $(\mathbf{A}')' = \mathbf{A}$,

(ii) $(\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}'$,

(iii) $(\mathbf{AC})' = \mathbf{C}'\mathbf{A}'$.

Definition 2.3. A square matrix $\mathbf{A}$ is called **symmetric** if $\mathbf{A}' = \mathbf{A}$, or equivalently $(a_{ij}) = (a_{ji})$.

2.1.2 Rank

Definition 2.4. A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is said to be **linearly independent** if

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

implies that $c_i = 0$ for all i. Conversely, if $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are not linearly independent, they are said to be **linearly dependent**.

Definition 2.5. Let $\mathbf{A}$ be an $m \times n$ matrix, the **rank** of $\mathbf{A}$ is defined as

$$\begin{aligned} \text{rank}(\mathbf{A}) &= \text{number of linearly independent columns of } \mathbf{A} \\ &= \text{number of linearly independent rows of } \mathbf{A}. \end{aligned}$$

It is clear that

$$\text{rank}(\mathbf{A}) \leq \min(m, n).$$

If $\text{rank}(\mathbf{A}) = n$, where $n < m$, then $\mathbf{A}$ is said to be a **full rank** matrix.

2.1.3 Inverse

Definition 2.6. If $\mathbf{A}$ is a full-rank square matrix, then $\mathbf{A}$ is **nonsingular**. A nonsingular matrix $\mathbf{A}$ has a unique **inverse**, denoted by $\mathbf{A}^{-1}$, such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Theorem 2.7. If $\mathbf{A}$ and $\mathbf{B}$ are nonsingular, then

(i) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$,

(ii) $\mathbf{A}'$ is nonsingular and its inverse can be found as $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$,

(iii) $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

2.1.4 Determinant

Definition 2.8. The **determinant** of an $n \times n$ matrix $\mathbf{A}$ is a scalar function of $\mathbf{A}$ defined as the sum of all $n!$ possible products of n elements such that

(i) each product contains one element from every row and every column of $\mathbf{A}$.

(ii) the factors in each product are written so that the column subscripts appear in order of magnitude and each product is then preceded by a plus or minus sign according to whether the number of inversions in the row subscripts is even or odd. (An inversion occurs whenever a larger number precedes a smaller one.)

The determinant of $\mathbf{A}$ is denoted by $|\mathbf{A}|$ or $\det(\mathbf{A})$.

Theorem 2.9. (i) If $\mathbf{D} = \mathrm{diag}(d_1, d_2, \ldots, d_n)$, $|\mathbf{D}| = \prod_{i=1}^{n} d_i$.

(ii) The determinant of a triangular matrix is the product of the diagonal elements.

(iii) If $\mathbf{A}$ is singular matrix, $|\mathbf{A}| = 0$.

(iv) If $\mathbf{A}$ is nonsingular matrix, $|\mathbf{A}| \neq 0$.

(v) If $\mathbf{A}$ is positive definite matrix, $|\mathbf{A}| > 0$.

(vi) $|\mathbf{A}'| = |\mathbf{A}|$.

(vii) If $\mathbf{A}$ is nonsingular matrix, $|\mathbf{A}^{-1}| = \dfrac{1}{|\mathbf{A}|}$.

Theorem 2.10. If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ matrices, then the determinant of the product is the product of the determinants:

(i) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$,

(ii) $|\mathbf{AB}| = |\mathbf{BA}|$,

(iii) $|\mathbf{A}^2| = |\mathbf{A}|^2$.

Definition 2.11. Let $\mathbf{A}$ be an $n \times n$ matrix. The **minor**, M_{ij}, $i, j = 1, \ldots, n$ of the element a_{ij}, is the determinant of the matrix obtained by deleting the ith row vector and jth column vector of $\mathbf{A}$.

Definition 2.12. Let $\mathbf{A}$ be an $n \times n$ matrix. The **cofactor**, C_{ij}, $i, j = 1, \ldots, n$ of the element a_{ij}, is defined by

$$C_{ij} = (-1)^{i+j} M_{ij},$$

where M_{ij} is the minor of a_{ij}.

Theorem 2.13. Let $\mathbf{A}$ be an $n \times n$ matrix. If we multiply the elements in any row (or column) of $\mathbf{A}$ by their cofactors, then the sum of the resulting products is $|\mathbf{A}|$. That is,

(i) If we expand along row i,

$$|\mathbf{A}| = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = \sum_{k=1}^{n} a_{ik}C_{ik}.$$

(ii) If we expand along column j,

$$|\mathbf{A}| = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} = \sum_{k=1}^{n} a_{kj}C_{kj}.$$

Corollary 2.14. If the elements in the ith row (or column) of an $n \times n$ matrix $\mathbf{A}$ are multiplied by the cofactors of a different row (or column), then the sum of the resulting products is zero. That is,

(i) If we use the elements of row i and the cofactors of row j,

$$\sum_{k=1}^{n} a_{ik}C_{jk} = 0, \quad i \neq j.$$

(ii) If we use the elements of column i and the cofactors of column j,

$$\sum_{k=1}^{n} a_{ki}C_{kj} = 0, \quad i \neq j.$$

Definition 2.15. If every element in an $n \times n$ matrix $\mathbf{A}$ is replaced by its cofactor, the resulting matrix is called the **matrix of cofactors** and is denoted $\mathbf{M}_C$. The transpose of the matrix of cofactors, $\mathbf{M}'_C$, is called the **adjoint** of $\mathbf{A}$ and is denoted $\text{adj}(\mathbf{A})$. Thus, the elements of $\text{adj}(\mathbf{A})$ are

$$\text{adj}(\mathbf{A})_{ij} = C_{ji}.$$

Theorem 2.16. If $|\mathbf{A}| \neq 0$, then

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\,\text{adj}(\mathbf{A}).$$

2.1.5 Trace

Definition 2.17. The **trace** of an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is a scalar defined as the sum of the diagonal elements of $\mathbf{A}$:

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

Theorem 2.18. If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ matrices, $\mathbf{C}$ is an $n \times p$ matrix, and $\mathbf{D}$ is a $p \times n$ matrix, then

(i) $\text{tr}(\mathbf{A} \pm \mathbf{B}) = \text{tr}(\mathbf{A}) \pm \text{tr}(\mathbf{B})$,

(ii) $\text{tr}(c\mathbf{A}) = c\ \text{tr}(\mathbf{A})$, if c is scalar,

(iii) $\text{tr}(\mathbf{A}') = \text{tr}(\mathbf{A})$,

(iv) $\text{tr}(\mathbf{CD}) = \text{tr}(\mathbf{DC})$.

2.1.6 Positive Definite Matrices

Definition 2.19. The symmetric matrix $\mathbf{A}$ is said to be **positive definite** matrix if for all possible $\mathbf{y}$ except $\mathbf{y} = \mathbf{0}$,

$$\mathbf{y}'\mathbf{A}\mathbf{y} > 0.$$

Similarly, $\mathbf{A}$ is said to be **positive semidefinite** matrix if for all $\mathbf{y}$,

$$\mathbf{y}'\mathbf{A}\mathbf{y} \geq 0$$

and there is at least one $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{y}'\mathbf{A}\mathbf{y} = 0$.

Definition 2.20. The symmetric matrix $\mathbf{A}$ is said to be **negative definite** matrix if for all possible $\mathbf{y}$ except $\mathbf{y} = \mathbf{0}$,

$$\mathbf{y}'\mathbf{A}\mathbf{y} < 0.$$

Similarly, $\mathbf{A}$ is said to be **negative semidefinite** matrix if for all $\mathbf{y}$,

$$\mathbf{y}'\mathbf{A}\mathbf{y} \leq 0$$

and there is at least one $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{y}'\mathbf{A}\mathbf{y} = 0$.

Theorem 2.21. Let $\mathbf{P}$ be a nonsingular matrix.

(i) If $\mathbf{A}$ is positive definite, then $\mathbf{P}'\mathbf{A}\mathbf{P}$ is positive definite.

(ii) If $\mathbf{A}$ is positive semidefinite, then $\mathbf{P}'\mathbf{A}\mathbf{P}$ is positive semidefinite.

Corollary 2.22. Let $\mathbf{A}$ be an $n \times n$ positive definite matrix and let $\mathbf{B}$ be an $m \times n$ matrix of rank $m \leq n$. Then $\mathbf{B}\mathbf{A}\mathbf{B}'$ is positive definite.

Corollary 2.23. Let $\mathbf{A}$ be an $n \times n$ positive definite matrix and let $\mathbf{B}$ be an $m \times n$ matrix of rank $m > n$ or if $\text{rank}(\mathbf{B}) = r$, where $r < m$ and $r < n$. Then $\mathbf{B}\mathbf{A}\mathbf{B}'$ is positive semidefinite.

Theorem 2.24. A symmetric matrix $\mathbf{A}$ is positive definite if and only if there exists a nonsingular matrix $\mathbf{P}$ such that

$$\mathbf{A} = \mathbf{P}\mathbf{P}'.$$

Corollary 2.25. Any positive definite matrix is nonsingular.

Theorem 2.26. Let $\mathbf{A}$ be an $m \times n$ matrix.

(i) If $\text{rank}(\mathbf{A}) = n$, then $\mathbf{A}'\mathbf{A}$ is positive definite.

(ii) If $\text{rank}(\mathbf{A}) < n$, then $\mathbf{A}'\mathbf{A}$ is positive semidefinite.

Theorem 2.27. If $\mathbf{A}$ is positive definite matrix, then $\mathbf{A}^{-1}$ is positive definite matrix.

Theorem 2.28. If $\mathbf{A}$ is positive definite matrix, then there exists a unique $\mathbf{A}^{1/2}$ is positive definite matrix such that $(\mathbf{A}^{1/2})^2 = \mathbf{A}$.

2.1.7 Eigenvalues and Eigenvectors

Definition 2.29. For every square matrix $\mathbf{A}$, a scalar λ and a nonzero vector $\mathbf{x}$ can be found such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x},$$

where λ is an **eigenvalue** of $\mathbf{A}$ and $\mathbf{x}$ is an **eigenvector** of $\mathbf{A}$.

Theorem 2.30. If $\mathbf{A}$ is symmetric matrix, then the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $\mathbf{A}$ are real.

Notation 2.31. If the eigenvalues of $\mathbf{A}$ are real, then we index them from largest to smallest as follows:

$$\lambda_n \leq \cdots \leq \lambda_2 \leq \lambda_1.$$

In this case, we sometimes use the notation $\lambda_{\max}$ and $\lambda_{\min}$ to denote λ_1 and λ_n, respectively.

Theorem 2.32. If λ is an eigenvalue of $n \times n$ matrix $\mathbf{A}$ with corresponding eigenvector $\mathbf{x}$, then

(i) $c\lambda$ is an eigenvalue of $c\mathbf{A}$ and $\mathbf{x}$ is the corresponding eigenvector,

(ii) λ^k is an eigenvalue of $\mathbf{A}^k$ and $\mathbf{x}$ is the corresponding eigenvector,

(iii) $1/\lambda$ is an eigenvalue of $\mathbf{A}^{-1}$ and $\mathbf{x}$ is the corresponding eigenvector.

Corollary 2.33. If $\mathbf{A}$ is an $n \times n$ positive definite matrix, then

$$\lambda_{\min}(\mathbf{A}^{-1}) = \lambda_{\max}^{-1}(\mathbf{A})$$

and

$$\lambda_{\max}(\mathbf{A}^{-1}) = \lambda_{\min}^{-1}(\mathbf{A}).$$

Theorem 2.34. Let $\mathbf{A}$ be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

(i) If $\mathbf{A}$ is positive definite, then $\lambda_i > 0$ for $i = 1, 2, \ldots, n$.

(ii) If $\mathbf{A}$ is positive semidefinite, then $\lambda_i \geq 0$ for $i = 1, 2, \ldots, n$. The number of eigenvalues λ_i for which $\lambda_i > 0$ is the rank of $\mathbf{A}$.

Theorem 2.35. If $\mathbf{A}$ is an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

(i) $|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i$,

(ii) $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$.

Corollary 2.36. If $\mathbf{A}$ is an $n \times n$ matrix, then

$$\lambda_{\min}(\mathbf{A})\,\mathrm{tr}(\mathbf{I}) \leq \mathrm{tr}(\mathbf{A}) \leq \lambda_{\max}(\mathbf{A})\,\mathrm{tr}(\mathbf{I}).$$

Theorem 2.37. If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ symmetric matrices, then

$$\lambda_{\min}(\mathbf{A}) + \lambda_{\min}(\mathbf{B}) \leq \lambda_{\min}(\mathbf{A}+\mathbf{B}) \leq \lambda_{\max}(\mathbf{A}+\mathbf{B}) \leq \lambda_{\max}(\mathbf{A}) + \lambda_{\max}(\mathbf{B}).$$

Theorem 2.38. If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ matrices or if $\mathbf{A}$ is $n \times p$ matrix and $\mathbf{B}$ is $p \times n$ matrix, then the non-zero eigenvalues of $\mathbf{AB}$ are the same as those of $\mathbf{BA}$. If $\mathbf{x}$ is an eigenvector of $\mathbf{AB}$, then $\mathbf{Bx}$ is an eigenvector of $\mathbf{BA}$.

Theorem 2.39. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices. If $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite, then

$$\lambda_{\min}(\mathbf{A})\,\mathrm{tr}(\mathbf{B}) \leq \mathrm{tr}(\mathbf{AB}) \leq \lambda_{\max}(\mathbf{A})\,\mathrm{tr}(\mathbf{B}).$$

Theorem 2.40. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices. If $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite, then

$$\lambda_{\min}(\mathbf{A})\lambda_{\min}(\mathbf{B}) \leq \lambda_{\min}(\mathbf{AB}) \leq \lambda_{\max}(\mathbf{AB}) \leq \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{B}).$$

2.1.8 Idempotent

Definition 2.41. A square matrix $\mathbf{A}$ is said to be **idempotent** if

$$\mathbf{A}^2 = \mathbf{A}.$$

Theorem 2.42. If $\mathbf{A}$ is singular, symmetric, and idempotent matrix, then $\mathbf{A}$ is positive semidefinite.

Theorem 2.43. If $\mathbf{A}$ is an $n \times n$ symmetric idempotent matrix of rank $r \leq n$, then $\mathbf{A}$ has r eigenvalues equal to 1 and $n - r$ eigenvalues equal to 0.

Theorem 2.44. If $\mathbf{A}$ is an $n \times n$ symmetric idempotent matrix of rank $r \leq n$, then

$$\mathrm{tr}(\mathbf{A}) = \mathrm{rank}(\mathbf{A}) = r.$$

Theorem 2.45. If $\mathbf{A}$ is an $n \times n$ idempotent matrix and $\mathbf{P}$ is an $n \times n$ nonsingular matrix, then

(i) $\mathbf{I} - \mathbf{A}$ is idempotent,

(ii) $\mathbf{A}(\mathbf{I} - \mathbf{A}) = \mathbf{O}$ and $(\mathbf{I} - \mathbf{A})\mathbf{A} = \mathbf{O}$,

(iii) $\mathbf{P}^{-1}\mathbf{AP}$ is idempotent.

2.1.9 Kronecker Product

Definition 2.46. Let $\mathbf{A}$ and $\mathbf{B}$ be $m \times n$ and $p \times q$ matrices. The **Kronecker product** of $\mathbf{A}$ with $\mathbf{B}$, denoted by $\mathbf{A} \otimes \mathbf{B}$, is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix},$$

where $\mathbf{A} = [a_{ij}]_{i=1,\dots,m, j=1,\dots,n}$.

Theorem 2.47. Let $\mathbf{A}$ and $\mathbf{B}$ be both $m \times n$ matrices, $\mathbf{C}$ be a $p \times q$ matrix and α be a real number. Then

(i) $\mathbf{A} \otimes (\alpha\mathbf{B}) = (\alpha\mathbf{A}) \otimes \mathbf{B} = \alpha(\mathbf{A} \otimes \mathbf{B})$,

(ii) $(\mathbf{A} \otimes \mathbf{C})' = \mathbf{A}' \otimes \mathbf{C}'$,

(iii) $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$,

(iv) $\mathbf{C} \otimes (\mathbf{A} + \mathbf{B}) = \mathbf{C} \otimes \mathbf{A} + \mathbf{C} \otimes \mathbf{B}$.

Theorem 2.48. Let $\mathbf{A}$ be an $m \times n$ matrix, $\mathbf{B}$ be a $p \times k$ matrix, $\mathbf{C}$ be an $n \times k$ matrix and $\mathbf{D}$ be a $k \times q$ matrix. Then

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD}).$$

2.1.10 Derivatives of Function of Vectors and Matrices

Definition 2.49. Let $y = f(\mathbf{x})$ be a function of n variables of the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$. The vector of the derivatives of the functions f is defined as

$$\frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}.$$

Theorem 2.50. Let $y = \mathbf{a}'\mathbf{x} = \mathbf{x}'\mathbf{a}$, where $\mathbf{a}' = (a_1, a_2, \ldots, a_n)$ is a vector of constants. Then

$$\frac{\partial y}{\partial \mathbf{x}} = \frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial(\mathbf{x}'\mathbf{a})}{\partial \mathbf{x}} = \mathbf{a}.$$

Theorem 2.51. Let $y = \mathbf{x}'\mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a symmetric matrix of constants. Then

$$\frac{\partial y}{\partial \mathbf{x}} = \frac{\partial(\mathbf{x}'\mathbf{A}\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$

Theorem 2.52. Let $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ be vectors of length m, n and r, respectively, such that $\mathbf{z}$ is a function of $\mathbf{y}$, which is in turn a function of $\mathbf{x}$. Then the **chain rule** for vectors is

$$\frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \frac{\partial \mathbf{z}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{x}}.$$

Definition 2.53. Let $y = f(\mathbf{X})$ be a function of $n \times n$ variables of the matrix $\mathbf{X}$. The matrix of the derivatives of the function f is defined as

$$\frac{\partial y}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \cdots & \frac{\partial y}{\partial x_{1n}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{2n}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y}{\partial x_{n1}} & \frac{\partial y}{\partial x_{n2}} & \cdots & \frac{\partial y}{\partial x_{nn}} \end{bmatrix}.$$

Theorem 2.54. Let $\mathbf{X}$ be an $n \times n$ positive definite matrix and $\mathbf{A}$ is an $n \times n$ matrix of constants. Then

$$\frac{\partial}{\partial \mathbf{X}} \operatorname{tr}(\mathbf{XA}) = \mathbf{A} + \mathbf{A}' - \operatorname{diag}(\mathbf{A}).$$

Theorem 2.55. Let $\mathbf{X}$ be an $n \times n$ positive definite matrix. Then

$$\frac{\partial}{\partial \mathbf{X}} \log |\mathbf{X}| = 2\mathbf{X}^{-1} - \operatorname{diag}(\mathbf{X}^{-1}).$$

Theorem 2.56. Let $\mathbf{X}$ be an $n \times n$ positive definite matrix. Let $\mathbf{A}$ be a $p \times n$ matrix and $\mathbf{B}$ be an $n \times m$ matrix such that $\mathbf{AXB}$ is nonsingular. Then

$$\frac{\partial}{\partial \mathbf{X}} \log |\mathbf{AXB}| = \mathbf{B}(\mathbf{AXB})^{-1}\mathbf{A}.$$

Theorem 2.57. Let $\mathbf{X}$ be an $n \times n$ nonsingular matrix that are functions of a scalar x. Then

$$\frac{\partial}{\partial x} \mathbf{X}^{-1} = -\mathbf{X}^{-1} \frac{\partial \mathbf{X}}{\partial x} \mathbf{X}^{-1}.$$

Theorem 2.58. Let $\mathbf{X}$ be an $n \times n$ positive definite matrix that are functions of a scalar x. Then

$$\frac{\partial}{\partial x} \log |\mathbf{X}| = \operatorname{tr}\left(\mathbf{X}^{-1} \frac{\partial \mathbf{X}}{\partial x}\right).$$

Definition 2.59. Let $\mathbf{x} = (x_1, x_2, \ldots, x_m)'$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ be vectors of length m and n, respectively, where each element y_i is a function of $\mathbf{x}$, saying that $\mathbf{y}$ is a function of $\mathbf{x}$. The derivative of the vector $\mathbf{y}$ with respect to vector $\mathbf{x}$ is the $m \times n$ matrix defined as

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_1}{\partial x_m} & \frac{\partial y_2}{\partial x_m} & \cdots & \frac{\partial y_n}{\partial x_m} \end{bmatrix}.$$

2.1.11 Matrix Analysis

Definition 2.60. A **vector norm** on $\mathbb{R}^n$ is a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ that satisfies the following properties:

(i) $\|\mathbf{x}\| \geq 0$ for $\mathbf{x} \in \mathbb{R}^n$,

(ii) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

(iii) $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ for $\alpha \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n$.

For example, a useful class of vector norms are the $p-$norms defined by

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, p \geq 1.$$

The $1-$norm, $2-$norm and $\infty-$norm are the most important:

$$\begin{aligned}
\|\mathbf{x}\|_1 &= |x_1| + |x_2| + \cdots + |x_n|, \\
\|\mathbf{x}\|_2 &= \left(|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2\right)^{1/2} = (\mathbf{x}'\mathbf{x})^{1/2}, \\
\|\mathbf{x}\|_\infty &= \max_{1 \leq i \leq n} |x_i|.
\end{aligned}$$

Definition 2.61. A **matrix norm** on $\mathbb{R}^{m \times n}$ is a function $\|\cdot\| : \mathbb{R}^{m \times n} \to \mathbb{R}$ that satisfies the following properties:

(i) $\|\mathbf{A}\| \geq 0$ for all $\mathbf{A} \in \mathbb{R}^{m \times n}$,

(ii) $\|\mathbf{A}\| = 0$ if and only if $\|\mathbf{A}\| = \mathbf{0}$,

(iii) $\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\|$ for all $\alpha \in \mathbb{R}, \mathbf{A} \in \mathbb{R}^{m \times n}$,

(iv) $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ for all $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$,

(v) $\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$ for all $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$.

For example, the 1−matrix norm, 2−matrix norm, ∞−matrix norm and F−matrix norm are the most important:

$$\|\mathbf{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|,$$

$$\|\mathbf{A}\|_2 = [\lambda_{\max}(\mathbf{A}'\mathbf{A})]^{1/2},$$

$$\|\mathbf{A}\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|,$$

$$\|\mathbf{A}\|_F = \left[\mathrm{tr}(\mathbf{A}'\mathbf{A})\right]^{1/2}.$$

Theorem 2.62. Each norm $\|\cdot\|$ on $\mathbb{R}^n$ is a continuous function with respect to the metric $\rho(\mathbf{x},\mathbf{y}) = \max_{1\le i\le n} |x_i - y_i|$ on $\mathbb{R}^n$.

Theorem 2.63. ([13]) Let f is scalar function. Let $\mathbf{A}$ and $\mathbf{E}$ be $n\times n$ matrices. The Taylor series expansion of matrix function is

$$f(\mathbf{A}+\mathbf{E}) = \sum_{j=0}^{\infty} \frac{1}{j!} D_f^{[j]}(\mathbf{A},\mathbf{E}),$$

where

$$D_f^{[j]}(\mathbf{A},\mathbf{E}) = \left.\frac{d^j}{dt^j}\right|_{t=0} f(\mathbf{A}+t\mathbf{E}).$$

2.2 Random Vectors and Matrices

In this section, we will review briefly those statistical concepts and properties of random vector and matrices needed in this study. More intensive reviews of these topics are provided in [28], [35] and [37].

Definition 2.64. Let $y_1, y_2, \ldots, y_n$ be the random variables. Then the vector

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

is called a **random vector** of length n.

Definition 2.65. The expected value of an $n \times 1$ random vector $\mathbf{y}$, denoted by $\boldsymbol{\mu}$, is defined as the vector of expected values of the n random variables $y_1, y_2, \ldots, y_n$ in $\mathbf{y}$:

$$\mathrm{E}[\mathbf{y}] = \mathrm{E}\left[\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}\right] = \begin{bmatrix} \mathrm{E}[y_1] \\ \mathrm{E}[y_2] \\ \vdots \\ \mathrm{E}[y_n] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} = \boldsymbol{\mu},$$

where $\mathrm{E}(y_i) = \mu_i$ for all $i = 1, 2, \ldots, n$.

Definition 2.66. Let $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ be a random vector of the n random variables $y_1, y_2, \ldots, y_n$ be such that

$$\mathrm{Var}[y_i] = \sigma_i^2$$

and

$$\mathrm{Cov}[y_i, y_j] = \sigma_{ij} \quad \text{for} \quad i \neq j.$$

Then the **covariance matrix**, denoted by $\mathrm{Cov}[\mathbf{y}]$, is defined as

$$\mathrm{Cov}[\mathbf{y}] = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}.$$

The covariance matrix $\boldsymbol{\Sigma}$ is symmetric because $\sigma_{ij} = \sigma_{ji}$.

Definition 2.67. Let $\mathbf{x} = (x_1, x_2, \ldots, x_m)'$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ be random vectors. We define their covariance matrix as

$$\operatorname{Cov}[\mathbf{x}, \mathbf{y}] = \begin{bmatrix} \operatorname{Cov}[x_1, y_1] & \operatorname{Cov}[x_1, y_2] & \cdots & \operatorname{Cov}[x_1, y_n] \\ \operatorname{Cov}[x_2, y_1] & \operatorname{Cov}[x_2, y_2] & \cdots & \operatorname{Cov}[x_2, y_n] \\ \vdots & \vdots & & \vdots \\ \operatorname{Cov}[x_m, y_1] & \operatorname{Cov}[x_m, y_2] & \cdots & \operatorname{Cov}[x_m, y_n] \end{bmatrix}.$$

Definition 2.68. The linear combination of the variables $y_1, y_2, \ldots, y_n$ from a random vector $\mathbf{y}$ using the $a_1, a_2, \ldots, a_n$ terms as coefficients is defined as

$$a_1 y_1 + a_2 y_2 + \cdots + a_n y_n = \mathbf{a}'\mathbf{y}.$$

Theorem 2.69. If $\mathbf{a}$ is an $n \times 1$ vector of constants and $\mathbf{y}$ is an $n \times 1$ random vector with mean vector $\boldsymbol{\mu}$, then

$$\operatorname{E}[\mathbf{a}'\mathbf{y}] = \mathbf{a}' \operatorname{E}[\mathbf{y}] = \mathbf{a}'\boldsymbol{\mu}.$$

Theorem 2.70. If $\mathbf{A}$ is an $m \times n$ matrix of constants and $\mathbf{y}$ is an $n \times 1$ random vector with mean vector $\boldsymbol{\mu}$, then

$$\operatorname{E}[\mathbf{A}\mathbf{y}] = \mathbf{A} \operatorname{E}[\mathbf{y}] = \mathbf{A}\boldsymbol{\mu}.$$

Corollary 2.71. If $\mathbf{A}$ is an $m \times n$ matrix of constants, $\mathbf{b}$ is an $m \times 1$ vector of constants and $\mathbf{y}$ is an $n \times 1$ random vector with mean vector $\boldsymbol{\mu}$, then

$$\operatorname{E}[\mathbf{A}\mathbf{y} + \mathbf{b}] = \mathbf{A}\boldsymbol{\mu} + \mathbf{b}.$$

Definition 2.72. An $n \times 1$ random vector $\mathbf{y}$ having density function given by

$$f(y) = \frac{1}{(\sqrt{2\pi})^n |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})' |\boldsymbol{\Sigma}|^{-1} (\mathbf{y} - \boldsymbol{\mu})\right)$$

which is the **multivariate normal density function** with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$, equivalently, we say that $\mathbf{y}$ is distributed as $N_n(\boldsymbol{\mu}, \mathbf{\Sigma})$ or

$$\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{\Sigma}).$$

The subscript n is the dimension of the n-variate normal distribution and indicates the number of variables.

Theorem 2.73. If $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{\Sigma})$ and $\mathbf{A}$ is an $n \times n$ symmetric matrix of constants, then

$$\mathrm{E}[\mathbf{y}'\mathbf{A}\mathbf{y}] = \mathrm{tr}(\mathbf{A}\mathbf{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$$

and

$$\mathrm{Var}[\mathbf{y}'\mathbf{A}\mathbf{y}] = 2\ \mathrm{tr}(\mathbf{A}\mathbf{\Sigma})^2 + 4\boldsymbol{\mu}'\mathbf{A}\mathbf{\Sigma}\mathbf{A}\boldsymbol{\mu}.$$

Theorem 2.74. ([38]) Let $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{A}$ and $\mathbf{B}$ be an $n \times n$ symmetric matrix of constants. Then

$$\begin{aligned}\mathrm{E}[(\mathbf{y}'\mathbf{A}\mathbf{y})(\mathbf{y}'\mathbf{B}\mathbf{y})] &= 2\,\mathrm{tr}(\mathbf{A}\mathbf{\Sigma}\mathbf{B}\mathbf{\Sigma}) + \mathrm{tr}(\mathbf{A}\mathbf{\Sigma})\,\mathrm{tr}(\mathbf{B}\mathbf{\Sigma}) + 4\boldsymbol{\mu}'\mathbf{A}\mathbf{\Sigma}\mathbf{B}\boldsymbol{\mu} \\ &\quad + \mathrm{tr}(\mathbf{A}\mathbf{\Sigma})\boldsymbol{\mu}'\mathbf{B}\boldsymbol{\mu} + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}\,\mathrm{tr}(\mathbf{B}\mathbf{\Sigma}) + (\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu})(\boldsymbol{\mu}'\mathbf{B}\boldsymbol{\mu}).\end{aligned}$$

Theorem 2.75. ([38]) Let $\mathbf{y} \sim N_n(\mathbf{0}, \mathbf{\Sigma})$, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be an $n \times n$ symmetric matrix of constants. Then

$$\begin{aligned}\mathrm{E}[(\mathbf{y}'\mathbf{A}\mathbf{y})(\mathbf{y}'\mathbf{B}\mathbf{y})(\mathbf{y}'\mathbf{C}\mathbf{y})] &= 4\,\mathrm{tr}(\mathbf{A}\mathbf{\Sigma}\mathbf{B}\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}) + 4\,\mathrm{tr}(\mathbf{A}\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}\mathbf{B}\mathbf{\Sigma}) \\ &\quad + 2\,\mathrm{tr}(\mathbf{A}\mathbf{\Sigma})\,\mathrm{tr}(\mathbf{B}\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}) + 2\,\mathrm{tr}(\mathbf{B}\mathbf{\Sigma})\,\mathrm{tr}(\mathbf{A}\mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma}) \\ &\quad + 2\,\mathrm{tr}(\mathbf{C}\mathbf{\Sigma})\,\mathrm{tr}(\mathbf{A}\mathbf{\Sigma}\mathbf{B}\mathbf{\Sigma}) + \mathrm{tr}(\mathbf{A}\mathbf{\Sigma})\,\mathrm{tr}(\mathbf{B}\mathbf{\Sigma})\,\mathrm{tr}(\mathbf{C}\mathbf{\Sigma}).\end{aligned}$$

2.3 Mathematical Statistics

In this section, we will introduce some important theories in statistics needed in this study. More intensive reviews of these topics are provided in [3] and [20].

Definition 2.76. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the population and $T(x_1, x_2, \ldots, x_n)$ be a real-valued function, then the random variable or random vector $Y = T(X_1, X_2, \ldots, X_n)$ is called a **statistic**.

Definition 2.77. Any function of a sample used to approximate a parameter is called a **point estimator**, that is, any statistic is a point estimator.

Definition 2.78. The **bias** of an estimator $\hat{\theta}$ of a parameter θ is the expectation of the difference between $\hat{\theta}$ and θ, that is $\mathrm{E}(\hat{\theta} - \theta)$. An estimator whose bias is equal to zero is called an **unbiased estimator**.

Definition 2.79. A sequence of random variables $X_1, X_2, \ldots$ **converges in probability** to a random variable X, written as $X_n \xrightarrow{p} X$, if

$$\lim_{n \to \infty} P(|X_n - X| > \epsilon) = 0,$$

for all $\epsilon > 0$.

Theorem 2.80. Assume that $\{X_n, n \geq 1\}$ are random variables and c is a constant such that

$$\mathrm{E}[X_n] \to c \quad \text{and} \quad \mathrm{Var}[X_n] \to 0,$$

then

$$X_n \xrightarrow{p} c.$$

Definition 2.81. A sequence of estimators, $\hat{\theta}_n$, is a **consistent sequence of estimators of the parameter** θ if for every $\epsilon > 0$,

$$\lim_{n \to \infty} P(|\hat{\theta}_n - \theta| > \epsilon) = 0,$$

or equivalently, $(\hat{\theta}_n)$ is said to be a consistent sequence of estimators of parameter θ if

$$\hat{\theta}_n \xrightarrow{p} \theta.$$

Theorem 2.82. If $(\hat{\theta}_n)$ is a sequence of estimators of parameter θ satisfying

(i) $\lim_{n\to\infty} \mathrm{Var}[\hat{\theta}_n] = 0$,

(ii) $\lim_{n\to\infty} \mathrm{E}[\hat{\theta}_n - \theta] = 0$.

Then, $(\hat{\theta}_n)$ is a consistent sequence of estimators of θ.

Definition 2.83. Let $\mathbf{X}$ be a random vector and $\mathbf{X}$ has probability function $f(\mathbf{X}|\boldsymbol{\theta})$ with parameter $\boldsymbol{\theta}$. Then

$$I(\boldsymbol{\theta}) = \mathrm{E}_{\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\boldsymbol{\theta}}\log f(\mathbf{X}|\boldsymbol{\theta})\right)^2\right]$$

is called the **information** (or the **Fisher information**).

Lemma 2.84. If $f(\mathbf{X}|\theta)$ is twice differentiable with respect to $\boldsymbol{\theta}$ and satisfies

$$\frac{d}{d\boldsymbol{\theta}}\mathrm{E}_{\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}}\log f(\mathbf{X}|\boldsymbol{\theta})\right] = \int\left[\left(\frac{\partial}{\partial\boldsymbol{\theta}}\log f(x|\boldsymbol{\theta})\right) f(x|\boldsymbol{\theta})\right] dx,$$

then

$$\mathrm{E}_{\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\boldsymbol{\theta}}\log f(\mathbf{X}|\boldsymbol{\theta})\right)^2\right] = -\,\mathrm{E}_{\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial\boldsymbol{\theta}^2}\log f(\mathbf{X}|\boldsymbol{\theta})\right].$$

Theorem 2.85 (Cauchy-Schwarz inequality). For any random variables X and Y, if X and Y have finite variances then

$$\mathrm{E}\,|XY| \leq \sqrt{\mathrm{E}[X^2]\,\mathrm{E}[Y^2]}.$$

Theorem 2.86 (Hölder inequality). For any random variables X and Y, then

$$\mathrm{E}\,|XY| \leq \left(\mathrm{E}[|X|^p]\right)^{1/p}\left(\mathrm{E}[|Y|^q]\right)^{1/q},$$

where $p, q \in (1,\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$.

Corollary 2.87. For any random variable X,

$$\mathrm{E}\,|X| \leq \left(\mathrm{E}[X^2]\right)^{1/2}.$$

2.4 Mathematical Analysis

In this section, we will introduce some important theories in the mathematical analysis needed in this study. More intensive reviews of these topics are provided in [6] and [24].

Theorem 2.88 (**Sequential Criterion for Continuity**)**.** A function $f : D \to \mathbb{R}$ is continuous at the point $c \in D$, where $D \subseteq \mathbb{R}$ if and only if for every sequence (x_n) in D that converges to c, the sequence $(f(x_n))$ converges to $f(c)$.

Theorem 2.89 (**Extreme Value Theorem**)**.** If f is a continuous function defined on a closed interval $[a, b]$, then the function attains its maximum value at some point c contained in the interval.

Theorem 2.90 (**Weierstrass Extreme Value Theorem**)**.** Let $X \subset \mathbb{R}^n$ be a close and bounded set and f be a continuous real-valued function on X. Then f attains a minimum and maximum on X.

Definition 2.91. Suppose f and g are two functions defined on some subsets of the real numbers, we write

$$f(x) = O(g(x)) \quad \text{as} \quad x \to \infty,$$

if and only if there exist a positive constant M and a real number N such that

$$\left|\frac{f(x)}{g(x)}\right| \leq M \quad \text{for all} \quad x \geq N.$$

Definition 2.92. Suppose f and g are two functions defined on some subsets of the real numbers, we write

$$f(x) = o(g(x)) \quad \text{as} \quad x \to \infty,$$

if and only if for every positive constant M there exists a real number N such that

$$\left|\frac{f(x)}{g(x)}\right| \leq M \quad \text{for all} \quad x \geq N.$$

Notation 2.93. Note that $F(x) = [O(f(x))]_{n\times n}$ for matrix-valued functions F and real function f such that $F(x)/f(x)$ is element-wise uniformly bounded $n \times n$ matrix, as $x \to \infty$.

Notation 2.94. Note that $F(x) = [o(f(x))]_{n\times n}$ for matrix-valued functions F and real function f such that $F(x)/f(x)$ is converge to an $n \times n$ zero matrix, as $x \to \infty$.

Theorem 2.95 (**Triangle inequality**)**.** For any real numbers x and y, then

$$|x + y| \leq |x| + |y|.$$

Theorem 2.96. For any real numbers $x_1, x_2, \ldots, x_n$, for $0 < p \leq 1$, then

$$\left|\sum_{i=1}^{n} x_i\right|^p \leq \sum_{i=1}^{n} |x_i|^p,$$

and for $p > 1$

$$\left|\sum_{i=1}^{n} x_i\right|^p \leq n^{p-1}\left(\sum_{i=1}^{n} |x_i|^p\right).$$

That is,

$$\left|\sum_{i=1}^{n} x_i\right|^p \leq \max(1, n^{p-1})\left(\sum_{i=1}^{n} |x_i|^p\right)$$

for $p \in (0, \infty)$.

2.5 Small Area Estimation: Univariate Fay-Herriot Models

In this section, we will review briefly the concepts of small area estimation. More intensive reviews of these topics are provided in [26] and [34].

2.5.1 Small Area Estimation

Small area estimation (SAE) is one of well-known statistical methods to estimate parameters when sample size is not large enough to provide reliable direct estimates. The concept of the small area estimation is to borrow strength from available information, such as administrative records and census data, through models, called model-based approach. Model-based methods can be conducted either area-level model or unit-level model based on data availability. When unit-level data are available, the unit level model can be used. However, most unit-level data is often inaccessible due to the security of the informant. Therefore, in this section, we will consider the area-level model.

2.5.2 Fay-Herriot Model

The most widely used area-level model in small area estimation is the Fay-Herriot model, proposed by Fay and Herriot in 1979 [15]. It was used first to estimate mean per capita income for small places in the USA. The Fay-Herriot model is a linear random effects model which links the small area mean θ_i in area i on the auxiliary variables $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})'$ through the linking model:

$$\theta_i = \mathbf{x}_i'\boldsymbol{\beta} + u_i, \quad i = 1, 2, \ldots, m,$$

where $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression coefficients and u_i are identically, independently and normally distributed area-specific random effects (also called the model error) with $\mathrm{E}[u_i] = 0$ and $\mathrm{Var}[u_i] = A$.

The sampling model:

$$y_i = \theta_i + e_i, \quad i = 1, 2, \ldots, m,$$

where y_i is a direct survey estimator of θ_i and e_i are independent and normally distributed sampling errors with $\mathrm{E}[e_i] = 0$ and $\mathrm{Var}[e_i] = D_i$. The sampling model indicates that the sample estimates are related to the unknown small area means and sampling errors e_i.

Combining the linking and sampling model, the form of the Fay-Herriot model is

$$y_i = \theta_i + e_i = \mathbf{x}_i'\boldsymbol{\beta} + u_i + e_i, i = 1, \ldots, m, \tag{2.1}$$

where u_i and e_i are mutually independent. The sampling variances D_i are assumed to be known. In practice, the parameters $\boldsymbol{\beta}$ and A of the linking model are generally unknown and are estimated from the available data.

We define $\mathbf{y} = (y_1, \ldots, y_m)'$, $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)'$, $\mathbf{u} = (u_1, \ldots, u_m)'$, and $\mathbf{e} = (e_1, \ldots, e_m)'$. The Fay-Herriot model can be rewritten as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} + \mathbf{e} \tag{2.2}$$

which is a special case of the general linear mixed model with block diagonal covariance structure. The covariance matrix of $\mathbf{y}$ is $\mathbf{V} = \underset{1 \le i \le m}{\text{diag}} (A + D_i)$.

2.5.3 Empirical Best Linear Unbiased Prediction

We are interested in estimating the ith small area means θ_i. It is well known that, among all linear unbiased predictors $\hat{\theta}_i$ of θ_i, the best linear unbiased predictor (BLUP) yields the minimum mean squared prediction error (MSE), which is defined as $\text{E}[(\hat{\theta}_i - \theta_i)^2]$. For known variance A, under model (2.1), the BLUP for θ_i is defined as

$$\tilde{\theta}_i = (1 - B_i)y_i + B_i \mathbf{x}_i' \tilde{\boldsymbol{\beta}},$$

where

$$B_i = \frac{D_i}{A + D_i}$$

and

$$\begin{aligned} \tilde{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ &= \left(\sum_{i=1}^{m} (A + D_i)^{-1} \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=1}^{m} (A + D_i)^{-1} \mathbf{x}_i \mathbf{y}_i \right). \end{aligned}$$

In the most realistic case when A is unknown. We can estimate A from the marginal distribution of y. An empirical BLUP or EBLUP estimator of θ_i is obtained by replacing

A in BLUP by an estimator $\hat{A}$ of A. The EBLUP estimator $\hat{\theta}_i$ is a weighted sum of a direct estimator y_i and a regression synthetic estimator $\mathbf{x}_i\hat{\boldsymbol{\beta}}$:

$$\hat{\theta}_i = (1 - \hat{B}_i)y_i + \hat{B}_i\mathbf{x}_i'\hat{\boldsymbol{\beta}},$$

where

$$\hat{B}_i = \frac{D_i}{\hat{A} + D_i}$$

and $\hat{\boldsymbol{\beta}}$ is $\tilde{\boldsymbol{\beta}}$ with A is replaced by $\hat{A}$. Hereafter, the estimator $\hat{A}$ also denotes an even translation invariant estimator for all $\boldsymbol{\beta}$ and $\mathbf{y}$ that achieve unbiasedness in the EBLUP, $\hat{A}(-\mathbf{y}) = \hat{A}(\mathbf{y})$ and $\hat{A}(\mathbf{y}-\mathbf{x}'\boldsymbol{\beta}) = \hat{A}(\mathbf{y})$, as in [22]. The weight $\hat{B}_i$ depends on the estimate of the ratio between sampling variance D_i and model variance A.

2.5.4 Variance Component Estimation

In the small area estimation with the Fay-Herriot model, accurate estimation of A is necessary in order to obtain an efficient EBLUP for the small area means θ_i. Jiang and Lahiri [21] and Rao and Molina [34] list several estimators of A satisfies the conditions. They include the Prasad-Rao simple method-of-moments estimator, the Fay-Herriot method-of-moments estimator, the profile maximum likelihood estimator and the residual maximum likelihood estimator. These estimators are all consistent for large m, under certain regularity conditions (i) and (ii), below:

(i) D_i are uniformly bounded;

(ii) $\sup_{1\le i\le m} \mathbf{x}_i' \left(\sum_{i=1}^{m} \mathbf{x}_i\mathbf{x}_i'\right)^{-1} \mathbf{x}_i = O(m^{-1})$.

However, a well-known problem associated with all of the above four variance component estimation methods is that all could yield a zero estimate, especially when the number of small areas is small. The zero estimate of A yields $\hat{B}_i = 0$ and consequently the EBLUP estimator of θ_i reduces to the regression estimator.

In this section, we first briefly review the four commonly used estimators of A. Then, we review the several adjusted method in [34], which produces strictly positive estimator of A. Thus, the resulting EBLUP of θ_i is never regression estimator and is always a weighted combination of the direct estimator and the regression estimator.

2.5.4.1 The Four Commonly Used Estimators of A

1. Prasad-Rao Method-of-Moments Estimator (PR)

In 1990, Prasad and Rao [32] proposed a simple method-of-moments to estimate A. The estimator is given by

$$\tilde{A}^{\text{PR}} = \frac{1}{m-p}[(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) - \sum_{i=1}^{m} D_i \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i],$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Note that the last equation could yield a negative estimate. In order to avoid the problem, they proposed the following estimator of A: $\hat{A}^{\text{PR}} = \max(\tilde{A}^{\text{PR}}, 0)$.

2. Fay-Herriot Method-of-Moments Estimator (FH)

The Fay-Herriot [15] estimator of A is based on the weighted least squares residual sum of squares. Using the best linear unbiased estimator of β, $\tilde{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$. The estimator of A is obtained by solving the following equation iteratively:

$$\sum_{i=1}^{m} \frac{(y_i - \mathbf{x}_i'\tilde{\boldsymbol{\beta}})^2}{\tilde{A}^{\text{FH}} + D_i} = m - p.$$

The left side of the last equation is the weighted residual sum of squares whose expectation, under the Fay-Herriot model, is identical to the right hand side. This is motivation for the Fay-Herriot estimator. Again the solution could be negative and so the following estimator is used in practice: $\hat{A}^{\text{FH}} = \max(\tilde{A}^{\text{FH}}, 0)$.

3. Profile Maximum Likelihood Estimator (PML)

In 1967, Hartley and Rao [17] proposed the maximum likelihood (ML) approach. The maximum log-likelihood under the Fay-Herriot model has the form

$$\ell_{\mathrm{M}}(\boldsymbol{\beta}, A) = c - \frac{1}{2}(\log|\mathbf{V}| + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \tag{2.3}$$

where c is a constant and $|\mathbf{V}|$ is the determinant of $\mathbf{V}$. By differentiating (2.3) with respect to $\boldsymbol{\beta}$ and A, we have

$$\begin{aligned}\frac{\partial}{\partial\beta}\ell_{\mathrm{M}}(\boldsymbol{\beta}, A) &= \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta},\\ \frac{\partial}{\partial A}\ell_{\mathrm{M}}(\boldsymbol{\beta}, A) &= \frac{1}{2}((\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{V}^{-2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) - \mathrm{tr}(\mathbf{V}^{-1})).\end{aligned}$$

From, letting $(\partial/\partial\boldsymbol{\beta})\ell_{\mathrm{M}}(\boldsymbol{\beta}, A) = 0$, we obtain $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$. Replacing $\boldsymbol{\beta}$ by $\tilde{\boldsymbol{\beta}}(A)$ in (2.3), we obtain the following profile log-likelihood:

$$\ell_{\mathrm{P}}(A) = c - \frac{1}{2}(\log|\mathbf{V}| + \mathbf{y}'\mathbf{P}\mathbf{y}),$$

where $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$. The first derivative of the last equation is given as

$$\frac{\partial}{\partial A}\ell_{\mathrm{P}}(A) = \frac{1}{2}(\mathbf{y}'\mathbf{P}^2\mathbf{y} - \mathrm{tr}(\mathbf{V}^{-1})).$$

The profile maximum likelihood or PML estimator of A is obtained as $\hat{A}^{\mathrm{P}} = \max(\tilde{A}^{\mathrm{P}}, 0)$ where $\tilde{A}^{\mathrm{P}}$ is a solution to $(\partial/\partial A)\ell_{\mathrm{P}}(A) = 0$.

4. Residual Maximum Likelihood Estimator (REML)

In 1971, Patterson and Thompson [31] proposed the restricted or residual maximum likelihood (REML) approach. The approach uses transformed data which do not include the inferences about the nuisance parameters $\boldsymbol{\beta}$. Under the Fay-Herriot model, the restricted log-likelihood has the form

$$\ell_{\mathrm{R}}(A) = c - \frac{1}{2}(\log|\mathbf{K}'\mathbf{V}\mathbf{K}| + \mathbf{y}'\mathbf{P}\mathbf{y}),$$

where c is a constant, $\mathbf{K}$ is an $m \times (m-p)$ matrix such that $\text{rank}(\mathbf{K}) = m - p$, $\mathbf{K}'\mathbf{X} = 0$, and $\mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}' = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1} = \mathbf{P}$. The first derivative of the last equation is given as

$$\frac{\partial}{\partial A}\ell_{\text{R}}(A) = \frac{1}{2}(\mathbf{y}'\mathbf{P}^2\mathbf{y} - \text{tr}(\mathbf{P})).$$

The residual maximum likelihood or REML estimator of A is obtained as $\hat{A}^{\text{R}} = \max(\tilde{A}^{\text{R}}, 0)$ where $\tilde{A}^{\text{R}}$ is a solution to $(\partial/\partial A)\ell_{\text{R}}(A) = 0$.

2.5.4.2 Adjusted Maximum Likelihood Methods

Section 2.5.4.1 considered the estimation of A, the variance of the model error u_i, and presented two methods of moment, PML and REML estimators. All these methods can lead to negative estimates $\tilde{A}$, especially for small sample size m, which are then truncated to zero: $\hat{A} = \max(\tilde{A}, 0)$. A drawback of this truncation is that the resulting EBLUP estimates, $\hat{\theta}_i$, will attach zero weight to all the direct survey estimates y_i regardless of the sample sizes when $\hat{A} = 0$. Then the EBLUP with zero weight of survey estimates is undesirable because it ignores the sample from survey data.

Several methods have been proposed to avoid a zero value for A. For example, in 2010, Lahiri and Li [25] defined the adjusted likelihood as

$$L_{\text{Adj}}(A) \propto h(A)L(A),$$

where $L(A)$ is either $L_{\text{P}}(A)$ or $L_{\text{R}}(A)$.

For Fay-Herriot model (2.2), the profile likelihood and the residual likelihood under normality are given by

$$L_{\text{P}}(A) \propto |\mathbf{V}|^{-1/2} \exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\}$$

and

$$L_{\text{R}}(A) \propto |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|^{-1/2}|\mathbf{V}|^{-1/2} \exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\},$$

respectively.

The adjustment factor $h(A)$ is chosen to ensure that the estimate maximizing $L_{\mathrm{Adj}}(A)$ with respect to A over $[0, \infty)$ This feature prevents zero weight of the EBLUP even for small m. A simple choice of $h(A)$ is $h(A) = A$ (see [27] and [29]). This choice gives a strictly positive estimate $\hat{A}_{LL}$. Then, in 2014, Yoshimori and Lahiri [41] proposed alternative choices of $h(A)$. This choice lead to $L_{\mathrm{Adj}}(A)$ closer to $L_{\mathrm{R}}(A)$ or $L_{\mathrm{P}}(A)$. The choice

$$h(A) = \left(\arctan\left(\sum_{i=1}^{m} \frac{A}{A+D_i}\right)\right)^{1/m}$$

satisfies these conditions, and we denote the resulting estimator of A as $\hat{A}_{YL}$.

2.5.5 Uncertainty of EBLUP

The MSE of the BLUP under the Fay-Herriot model can be derived as follows:

$$\mathrm{MSE}[\tilde{\theta}_i] \equiv \mathrm{E}[(\tilde{\theta}_i - \theta_i)^2] = g_{1i}(A) + g_{2i}(A),$$

where

$$g_{1i}(A) = \frac{AD_i}{A+D_i},$$

and

$$g_{2i}(A) = \frac{D_i^2}{(A+D_i)^2}\mathbf{x}_i' \left(\sum_{i=1}^{m}(A+D_i)^{-1}\mathbf{x}_i\mathbf{x}_i'\right)^{-1} \mathbf{x}_i.$$

Under the regularity conditions (i) and (ii), the first term of the MSE of the BLUP, $g_{1i}(A)$ is $O(1)$, whereas the second term, $g_{2i}(A)$, due to estimating $\boldsymbol{\beta}$, is $O(m^{-1})$ for large m. From the MSE of the BLUP, it is obvious that when the variance of the model error, A, is small relative to the total variance, $\tilde{\theta}_i$ is much more efficient than y_i which has variance D_i.

An important matter of SAE is to determine the accuracy of the predictors. A naive measure of uncertainty of an EBLUP is the MSE of the corresponding BLUP. However, in 1984, Kackar and Haville [23] showed that the MSE of BLUP is smaller than that of EBLUP and their difference depends on the variability of the estimator $\hat{A}$, which is of the order of $O(m^{-1})$ for large m. It is not accurate enough to be ignored for most small-area applications. In 1990, Prasad and Rao [32] provide an approximation of MSE of EBLUP under the Fay-Herriot model as

$$\mathrm{MSE}[\hat{\theta}_i] \equiv \mathrm{E}[(\hat{\theta}_i - \theta_i)^2] = g_{1i}(A) + g_{2i}(A) + g_{3i}(A),$$

where

$$g_{3i}(A) = \frac{D_i^2}{(A + D_i)^3}\,\mathrm{Var}(\hat{A}).$$

A second-order unbiased estimator of $\mathrm{MSE}[\hat{\theta}_i]$ is

$$\mathrm{mse}[\hat{\theta}_i] = g_{1i}(\hat{A}^{\mathrm{PR}}) + g_{2i}(\hat{A}^{\mathrm{PR}}) + 2g_{3i}(\hat{A}^{\mathrm{PR}}),$$

where

$$\mathrm{Var}[\hat{A}^{\mathrm{PR}}] = \frac{2}{m^2}\sum_{i=1}^{m}(A + D_i)^2.$$

In 2000, Datta and Lahiri [11] extended this to the case where the variance components are estimated by the profile maximum likelihood or residual maximum likelihood method. They obtained the second-order unbiased MSE estimator

$$\mathrm{mse}[\hat{\theta}_i] = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}) - \mathrm{bias}(\hat{A})\triangledown g_{1i}(\hat{A}),$$

with

$$\mathrm{Var}[\hat{A}^{\mathrm{ML}}] = \mathrm{Var}[\hat{A}^{\mathrm{RE}}] = 2\left(\sum_{i=1}^{m}\frac{1}{(A + D_i)^2}\right)^{-1},$$

where $\mathrm{bias}(\hat{A})$ is second-order unbiased estimator of $\mathrm{Bias}(\hat{A})$ and

$$\triangledown g_{1i}(A) = \left(\frac{A}{A + D_i}\right)^2.$$

CHAPTER III

ADJUSTED MAXIMUM LIKELIHOOD METHODS

In this chapter, we extend an adjusted maximum likelihood method to obtain a method for multivariate Fay-Herriot model. In Section 3.1, we give a review of the multivariate Fay-Herriot model. In Section 3.2, we discuss the variance components estimation for multivariate Fay-Herriot model (see [34] and [36]). Then, we propose a new adjusted maximum likelihood method for multivariate Fay-Herriot model and derive the properties of the obtained estimator in Section 3.3. Finally, we present results from a Monte Carlo simulation in Section 3.4.

3.1 Small Area Estimation: Multivariate Fay-Herriot Models

The multivariate Fay-Herriot (MFH) model is an extension of the Fay-Herriot model defined in (2.1) when we have multiple variables of interest, say R variables. Similar to the univariate Fay-Herriot model, the multivariate Fay-Herriot model contains of two models: sampling model and linking model. For the sampling model, let $\boldsymbol{\mu}_d = (\mu_{d1}, \dots, \mu_{dR})'$ be an $R \times 1$ vector of R characteristics of interest in the domain d, and $\mathbf{y}_d = (y_{d1}, \dots, y_{dR})'$ be an $R \times 1$ vector of direct survey estimators of $\boldsymbol{\mu}_d$. Then

$$\mathbf{y}_d = \boldsymbol{\mu}_d + \mathbf{e}_d, \quad d = 1, 2, \dots, D,$$

where the $\mathbf{e}_d = (e_{d1}, \dots, e_{dR})'$ is the vector of sampling errors which are independent and normally distributed with $\mathrm{E}[\mathbf{e}_d] = \mathbf{0}_R$ and known covariance matrix $\mathrm{Cov}[\mathbf{e}_d] = \mathbf{V}_{ed}$, with dimension $R \times R$. For the linking model, the population mean $\boldsymbol{\mu}_d$ is the linked to the auxiliary variables $\mathbf{X}_d$,

$$\boldsymbol{\mu}_d = \mathbf{X}_d\boldsymbol{\beta} + \mathbf{u}_d, \quad d = 1, 2, \ldots, D,$$

where $\mathbf{u}_d = (u_{d1}, \ldots, u_{dR})'$ is the vector of area specific random effects (also called the model errors) which are independent and normally distributed with mean $\mathrm{E}[\mathbf{u}_d] = \mathbf{0}_R$ and covariance matrix $\mathrm{Cov}[\mathbf{u}_d] = \mathbf{V}_{\mathrm{u}d}$, $\mathbf{X}_d$ is an $R \times pR$ matrix of auxiliary variables with rth row given by $(\mathbf{0}_p', \ldots, \mathbf{0}_p', \mathbf{x}_{dr}', \mathbf{0}_p', \ldots, \mathbf{0}_p')$ with $p < D$, and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_R')$ is a $pR \times 1$ vector of regression coefficients. Here, $\mathbf{x}_{dr}'$ occurs in the rth position of the row vector (r^{th} row).

Combining the sampling model and the linking model, we obtain a multivariate Fay-Herriot model

$$\mathbf{y}_d = \mathbf{X}_d\boldsymbol{\beta} + \mathbf{u}_d + \mathbf{e}_d, \quad d = 1, \ldots, D, \tag{3.1}$$

where $\mathbf{u}_d$ and $\mathbf{e}_d$ are mutually independent.

Let $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_D)'$ be the vector of direct survey estimators of $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_D)'$, $\mathbf{X} = \mathrm{col}_{1 \le d \le D}(\mathbf{X}_d)$, and define

$$\mathbf{u} = \underset{1 \le d \le D}{\mathrm{col}}(\mathbf{u}_d), \quad \mathbf{u}_d = \underset{1 \le r \le R}{\mathrm{col}}(u_{dr}), \quad \mathbf{V}_{\mathrm{u}} = \underset{1 \le d \le D}{\mathrm{diag}}(\mathbf{V}_{\mathrm{u}d}),$$
$$\mathbf{e} = \underset{1 \le d \le D}{\mathrm{col}}(\mathbf{e}_d), \quad \mathbf{e}_d = \underset{1 \le r \le R}{\mathrm{col}}(e_{dr}), \quad \mathbf{V}_{\mathrm{e}} = \underset{1 \le d \le D}{\mathrm{diag}}(\mathbf{V}_{\mathrm{e}d}),$$

where col is the matrix operator stacking columns of a matrix.

In the vector form, the multivariate Fay-Herriot model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} + \mathbf{e}, \tag{3.2}$$
$$= \mathbf{X}\boldsymbol{\beta} + \mathbf{w},$$

where $\mathbf{u}$ and $\mathbf{e}$ are independent random variables such that $\mathbf{u} \sim N(0, \mathbf{V}_{\mathrm{u}})$, $\mathbf{e} \sim N(0, \mathbf{V}_{\mathrm{e}})$, and $\mathbf{w} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} = \mathbf{u} + \mathbf{v} \sim N(0, \mathbf{V})$ with $\mathbf{V} = \mathbf{V}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}}$.

In this work, we consider Model 1 in [1] where the covariance matrix of model error and sampling error are

$$\mathbf{V}_{\mathrm{u}d} = \boldsymbol{\Sigma}(\boldsymbol{\theta}) = \underset{1 \le r \le R}{\operatorname{diag}}(\theta_r) \text{ and positive definite matrix } \mathbf{V}_{ed}, \tag{3.3}$$

respectively, for $d = 1, \ldots, D$. For notation simplicity, we denote $\boldsymbol{\Sigma}$ for $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ for the rest of this thesis book. The covariance matrix $\boldsymbol{\Sigma}$ depends on the vector of R unknown parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_R)'$.

The BLUP estimator of $\boldsymbol{\mu}$ is given by

$$\tilde{\boldsymbol{\mu}} = \mathbf{V}_{\mathrm{u}}(\mathbf{V}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}})^{-1}\mathbf{y} + \mathbf{V}_{\mathrm{e}}(\mathbf{V}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}})^{-1}\mathbf{X}\tilde{\boldsymbol{\beta}}, \tag{3.4}$$

where $\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$. The BLUP estimator depends on the unknown parameters $\boldsymbol{\theta}$. Substituting an estimator $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$, we obtain the empirical BLUP or EBLUP estimator $\hat{\boldsymbol{\mu}}$ of $\boldsymbol{\mu}$

$$\hat{\boldsymbol{\mu}} = \hat{\mathbf{V}}_{\mathrm{u}}(\hat{\mathbf{V}}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}})^{-1}\mathbf{y} + \mathbf{V}_{\mathrm{e}}(\hat{\mathbf{V}}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}})^{-1}\mathbf{X}\hat{\boldsymbol{\beta}}, \tag{3.5}$$

where $\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}) = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{y}$ and $\hat{\mathbf{V}} = \hat{\mathbf{V}}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}}$ with $\hat{\mathbf{V}}_{\mathrm{u}} = \mathbf{V}_{\mathrm{u}}(\hat{\boldsymbol{\theta}})$.

Following Benavent and Morales [1] and González-Menteiga et al. [16], we assume the following regularity conditions throughout the thesis:

1. $0 < p < D$, $0 < R < D$;

2. $\boldsymbol{\Sigma}$ and $\mathbf{V}_{\mathrm{e}d}, d = 1, \ldots, D$, are positive definite matrices with uniformly bounded elements. This implies that $\mathbf{V} = \mathbf{V}_{\mathrm{u}} + \mathbf{V}_{\mathrm{e}} = \underset{1 \le d \le D}{\operatorname{diag}}(\boldsymbol{\Sigma}) + \underset{1 \le d \le D}{\operatorname{diag}}(\mathbf{V}_{\mathrm{e}d})$ is positive definite matrix;

3. The elements of the covariate matrix $\mathbf{X}$, $|x_{drk}| \le x < \infty$, for some positive real number x, for all d, r, and k, $\mathbf{X}'\mathbf{X} = [O(D)]_{pR \times pR}$;

4. $\mathbf{X}'\mathbf{V}_{\mathrm{e}}^{-1}\mathbf{X} = [O(D)]_{pR \times pR}$, $\sum_{d=1}^{D} \mathbf{1}_R'\mathbf{V}_{\mathrm{e}d}\mathbf{1}_R = O(D)$;

5. $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} = [O(D^{-1})]_{pR \times pR}$;

6. $\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} = \left[O(D^{-1})\right]_{R\times R}$;

7. $\hat{\theta}_i = k + \mathbf{y}'\mathbf{C}_i\mathbf{y}$ is an unbiased, consistent and translation invariant estimator of θ_i, where $k = O(1)$ is constant and $\mathbf{C}_i = \operatorname{diag}([O(D^{-1})]_{R\times R}, \ldots, [O(D^{-1})]_{R\times R}) + [O(D^{-2})]_{DR\times DR}$, for all $i = 1, \ldots, R$;

8. $\limsup_{D\to\infty} \max_{1\le d\le D} \lambda_{\max}(\mathbf{V}_{ed}) < \infty$.

3.2 Variance Component Estimation for Multivariate Fay-Herriot Model

For the unknown parameters $\boldsymbol{\theta}$ in covariance matrix $\boldsymbol{\Sigma}$ in model (3.2), we consider two standard methods for the estimation of variance components $\boldsymbol{\theta}$: the profile maximum likelihood (PML) and the residual maximum likelihood (REML) methods.

3.2.1 Profile Maximum Likelihood Estimator (PML)

The profile log-likelihood function [17] is defined as

$$\ell_{\mathrm{P}}(\boldsymbol{\theta}) = c - \frac{1}{2}\left(\log|\mathbf{V}| + \mathbf{y}'\mathbf{P}\mathbf{y}\right), \tag{3.6}$$

where $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$. The first partial derivative of (3.6), for $i = 1, \ldots, R$, is given as

$$\begin{aligned}\frac{\partial \ell_{\mathrm{P}}(\boldsymbol{\theta})}{\partial \theta_i} &= \frac{1}{2}\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right) \\ &= \frac{1}{2}\left(\mathbf{w}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{w} - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right), \end{aligned} \tag{3.7}$$

where $\mathbf{w} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ and we use the fact that $\mathbf{P}\mathbf{X} = \mathbf{0}$ to obtain (3.7). The profile maximum likelihood or PML estimator of $\boldsymbol{\theta}$ is denoted as $\hat{\boldsymbol{\theta}}^{\mathrm{P}} = [\hat{\theta}_i^{\mathrm{P}}]_{i=1,\ldots,R}$, where $\hat{\theta}_i^{\mathrm{P}} = \max(\tilde{\theta}_i^{\mathrm{P}}, 0)$ such that $\tilde{\theta}_i^{\mathrm{P}}$ is the solution to $(\partial/\partial\theta_i)\ell_{\mathrm{P}}(\boldsymbol{\theta}) = 0$.

3.2.2 Residual Maximum Likelihood Estimator (REML)

The residual log-likelihood function [31] is defined as

$$\ell_{\mathrm{R}}(\boldsymbol{\theta}) = c - \frac{1}{2}\left(\log|\mathbf{K}'\mathbf{V}\mathbf{K}| + \mathbf{y}'\mathbf{P}\mathbf{y}\right), \tag{3.8}$$

where c is a constant, $\mathbf{K}$ is a $DR \times (DR - pR)$ matrix such that $\operatorname{rank}(\mathbf{K}) = DR - pR$, $\mathbf{K}'\mathbf{X} = \mathbf{0}$, and $\mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}' = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1} = \mathbf{P}$. The first partial derivative of (3.8), for $i = 1, \ldots, R$, is given as

$$\begin{aligned}
\frac{\partial \ell_{\mathrm{R}}(\boldsymbol{\theta})}{\partial \theta_i} &= \frac{1}{2}\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \operatorname{tr}\left(\mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}'\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right) \\
&= \frac{1}{2}\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right) \\
&= \frac{1}{2}\left(\mathbf{w}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{w} - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right),
\end{aligned} \tag{3.9}$$

where $\mathbf{w} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ and we use the fact that $\mathbf{P}\mathbf{X} = \mathbf{0}$ to obtain (3.9). The residual maximum likelihood or REML estimator of $\boldsymbol{\theta}$ is denoted as $\hat{\boldsymbol{\theta}}^{\mathrm{R}} = [\hat{\theta}_i^{\mathrm{R}}]_{i=1,\ldots,R}$, where $\hat{\theta}_i^{\mathrm{R}} = \max(\tilde{\theta}_i^{\mathrm{R}}, 0)$ such that $\tilde{\theta}_i^{\mathrm{R}}$ is a solution to $(\partial/\partial\theta_i)\ell_{\mathrm{R}}(\boldsymbol{\theta}) = 0$.

3.2.3 Adjusted Maximum Likelihood Estimator of Li and Lahiri (AML.LL)

In this section, we consider a special case, where $\theta_i = \theta$ for all $i = 1, \ldots, R$. For this case, we apply the adjusted maximum likelihood method of Li and Lahiri, called AML.LL [27] to obtain the variance parameter θ. The adjusted likelihood function of θ is defined as

$$L_{\mathrm{Adj}}(\theta) = \theta \times L(\theta),$$

where $L(\theta)$ is the likelihood function of the parameter θ. The likelihood functions considered are either profile likelihood or residual likelihood.

The corresponding adjusted profile and residual log-likelihood function are respectively defined as

$$\ell_{\mathrm{AP.LL}}(\theta) = c - \frac{1}{2}\left(\log|\mathbf{V}| + \mathbf{y}'\mathbf{P}\mathbf{y}\right) + \log(\theta),$$

and

$$\ell_{\mathrm{AR.LL}}(\theta) = c - \frac{1}{2}\left(\log|\mathbf{K}'\mathbf{V}\mathbf{K}| + \mathbf{y}'\mathbf{P}\mathbf{y}\right) + \log(\theta).$$

The adjusted profile maximum likelihood (APML.LL) and adjusted residual maximum likelihood (AREML.LL) estimators of θ are denoted as $\hat{\theta}^{\mathrm{AP.LL}}$ and $\hat{\theta}^{\mathrm{AR.LL}}$, where $\hat{\theta}^{\mathrm{AP.LL}}$ is the solution to $(\partial/\partial\theta)\ell_{\mathrm{AP.LL}}(\theta) = 0$ and $\hat{\theta}^{\mathrm{AR.LL}}$ is the solution to $(\partial/\partial\theta)\ell_{\mathrm{AR.LL}}(\theta) = 0$, respectively.

Under the regularity conditions given in [27], we have

$$\mathrm{E}[\hat{\theta} - \theta] = \begin{cases} \dfrac{\mathrm{tr}(\mathbf{P} - \mathbf{V}^{-1}) + 2/\theta}{\mathrm{tr}(\mathbf{V}^{-2})} + o(D^{-1}) & \text{if } \hat{\theta} = \hat{\theta}^{\mathrm{AP.LL}} \\ \dfrac{2/\theta}{\mathrm{tr}(\mathbf{V}^{-2})} + o(D^{-1}) & \text{if } \hat{\theta} = \hat{\theta}^{\mathrm{AR.LL}} \end{cases}$$

and

$$\mathrm{E}[\hat{\theta} - \theta]^2 = \frac{2}{\mathrm{tr}(\mathbf{V}^{-2})} + o(D^{-1}).$$

In this section, we have discussed the estimation of parameters $\boldsymbol{\theta}$ in the covariance matrix $\boldsymbol{\Sigma}$, of the vector of random effects $\mathbf{u}_d$, $d = 1, \ldots, D$. The PML and REML methods of estimation of $\boldsymbol{\theta}$ can lead to negative estimates of variance for some cases of random effects u_{dr}, for $d = 1, \ldots, D$, $r = 1, \ldots, R$, which are then truncated at zero. For AML.LL methods, these methods can prevent the zero estimate of $\boldsymbol{\theta}$. However, the method is only applicable for the case where $\theta_i = \theta$ for all $i = 1, \ldots, R$.

3.3 Adjusted Maximum Likelihood Method (AML)

In this section, for covariance matrix $\boldsymbol{\Sigma}$ defined in Section 3.1, we propose the adjusted maximum likelihood method to obtain the strictly positive estimate of $\boldsymbol{\theta}$. The adjusted likelihood function of $\boldsymbol{\theta}$ is defined as

$$L_{\mathrm{Adj}}(\boldsymbol{\theta}) = |\boldsymbol{\Sigma}|^{1/D} \times L(\boldsymbol{\theta}),$$

where $L(\boldsymbol{\theta})$ is a standard likelihood function (profile and residual likelihood function).

From (3.6) and (3.8), the adjusted log-likelihood functions corresponding to profile and residual likelihood function are given by

$$\ell_{\mathrm{AP}}(\boldsymbol{\theta}) = c - \frac{1}{2}(\log|\mathbf{V}| + \mathbf{y}'\mathbf{P}\mathbf{y}) + \frac{1}{D}\log|\boldsymbol{\Sigma}|,$$

and

$$\ell_{\mathrm{AR}}(\boldsymbol{\theta}) = c - \frac{1}{2}(\log|\mathbf{K}'\mathbf{V}\mathbf{K}| + \mathbf{y}'\mathbf{P}\mathbf{y}) + \frac{1}{D}\log|\boldsymbol{\Sigma}|.$$

By solving $(\partial/\partial\boldsymbol{\theta})\ell_{\mathrm{AP}}(\boldsymbol{\theta}) = \mathbf{0}$ and $(\partial/\partial\boldsymbol{\theta})\ell_{\mathrm{AR}}(\boldsymbol{\theta}) = \mathbf{0}$, we obtain an adjusted profile maximum likelihood (APML) estimator, $\hat{\boldsymbol{\theta}}^{\mathrm{AP}}$, and adjusted residual maximum likelihood (AREML) estimator, $\hat{\boldsymbol{\theta}}^{\mathrm{AR}}$, respectively.

Both $\hat{\boldsymbol{\theta}}^{\mathrm{AP}}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{AR}}$ are even translation invariant estimators of $\boldsymbol{\theta}$. That is adjusted maximum likelihood estimators satisfy the following two conditions: (1) $\hat{\boldsymbol{\theta}}(-\mathbf{y}) = \hat{\boldsymbol{\theta}}(\mathbf{y})$ and (2) $\hat{\boldsymbol{\theta}}(\mathbf{y} - \mathbf{X}\mathbf{B}) = \hat{\boldsymbol{\theta}}(\mathbf{y})$ for all $\mathbf{y}$ and $\mathbf{B}$. This property follows from (3.7) and (3.9).

In the next sections, we consider positiveness and asymptotic properties of the adjusted maximum likelihood estimate $\hat{\boldsymbol{\theta}}$ in Section 3.3.1 and Section 3.3.2, respectively.

3.3.1 Positiveness of the AML Estimators

In this section, we will show that both $\hat{\boldsymbol{\theta}}^{\mathrm{AP}}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{AR}}$ are strictly positive.

Lemma 3.1. Let $f(\cdot,\cdot,\ldots,\cdot)$ be a continuous and positive-valued function of $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ where $x_i \geq 0$, for all $i = 1, \ldots, n$ and

$$\lim_{x_j\to\infty} \prod_{i=1}^{n} x_i^{\alpha} f(x_1, x_2, \ldots, x_n) = \lim_{x_i\to\infty} g(x_1, x_2, \ldots, x_n) = 0, \text{ for all } j = 1, \ldots, n$$

such that $\alpha > 0$. Then there exist $x_{01}, x_{02}, \ldots, x_{0n}$ such that

$$g(x_{01}, x_{02}, \ldots, x_{0n}) = \max_{x_1, x_2, \ldots, x_n} g(x_1, x_2, \ldots, x_n)$$

and $x_{0i} > 0$ for all $i = 1, 2, \ldots, n$.

Proof. Since $f(x_1, x_2, \ldots, x_n) > 0$, we have $g(x_1, x_2, \ldots, x_n) \geq 0$ when $\prod_{i=1}^{n} x_i \geq 0$. Since $\lim_{x_1\to\infty} \prod_{i=1}^{n} x_i^{\alpha} f(x_1, x_2, \ldots, x_n) = 0$ for any fixed $x_2, x_3, \ldots, x_n$, there exists $N_1 \in \mathbb{N}$,

$$\prod_{i=1}^{n} x_i^{\alpha} f(x_1, x_2, \ldots, x_n) < \epsilon_1$$

for all $x_1 > N_1$, where

$$\epsilon_1 = \prod_{i=2}^{n} x_i^{\alpha} f(1, x_2, x_3, \ldots, x_n) > 0.$$

Since $\lim_{x_2\to\infty} \prod_{i=2}^{n} x_i^{\alpha} f(1, x_2, \ldots, x_n) = 0$ for any fixed $x_3, x_4, \ldots, x_n$, there exists $N_2 \in \mathbb{N}$,

$$\prod_{i=2}^{n} x_i^{\alpha} f(1, x_2, \ldots, x_n) < \epsilon_2$$

for all $x_2 > N_2$, where

$$\epsilon_2 = \prod_{i=3}^{n} x_i^{\alpha} f(1, 1, x_3, x_4, \ldots, x_n) > 0.$$

Continue process. Since $\lim_{x_n\to\infty}\prod_{i=n-1}^{n} x_i^{\alpha} f(1,1,\ldots,1,x_n)=0$, there exists $N_n\in\mathbb{N}$,

$$\prod_{i=n-1}^{n} x_i^{\alpha} f(1,1,\ldots,1,x_n) < \epsilon_n$$

for all $x_n > N_n$, where

$$\epsilon_n = f(1,1,\ldots,1) > 0.$$

Let $N=\max(N_1,N_2,\ldots,N_n)$. For $(x_1,x_2,\ldots,x_n)\in(N,\infty)^n$,

$$\prod_{i=1}^{n} x_i^{\alpha} f(x_1,x_2,\ldots,x_n) < f(1,1,\ldots,1).$$

By the Extreme Value Weierstrass Theorem, there exist $x_{01},x_{02},\ldots,x_{0n}\in[0,N]^n$ such that

$$g(x_{01},x_{02},\ldots,x_{0n}) = \max_{x_1,x_2,\ldots,x_n\in[0,N]^n} g(x_1,x_2,\ldots,x_n).$$

If $g(x_{01},x_{02},\ldots,x_{0n})\geq f(1,1,\ldots,1)$, then

$$g(x_{01},x_{02},\ldots,x_{0n}) = \max_{x_1,x_2,\ldots,x_n} g(x_1,x_2,\ldots,x_n).$$

If $g(x_{01},x_{02},\ldots,x_{0n}) < f(1,1,\ldots,1)$, then

$$g(x_{01},x_{02},\ldots,x_{0n}) = \max_{x_1,x_2,\ldots,x_n} g(x_1,x_2,\ldots,x_n)$$

and $x_{01},x_{02},\ldots,x_{0n}=1$.

Since $g(x_{01},x_{02},\ldots,x_{0,i-1},0,x_{0,i+1},\ldots,x_n)=0$, g does not attain maximum at $(x_{01},x_{02},\ldots,x_{0n})$ with $x_{0i}=0$ for some $i=1,\ldots,n$. Hence g has maximum at $(x_{01},x_{02},\ldots,x_{0n})$ where $x_{0i}>0$ for all $i=1,\ldots,n$. $\square$

Theorem 3.2. The adjusted profile maximum likelihood (APML) and the adjusted residual maximum likelihood (AREML) estimators are strictly positive.

Proof. The positiveness of the APML and AREML estimator can be easily derived from Lemma 3.1. We first consider the APML estimator, the adjusted profile likelihood function is

$$\begin{aligned} L_{\mathrm{AP}}(\boldsymbol{\theta}) &= |\boldsymbol{\Sigma}|^{1/D} L_{\mathrm{P}}(\boldsymbol{\theta}) \\ &= c|\boldsymbol{\Sigma}|^{1/D}|\mathbf{V}|^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\}, \end{aligned}$$

where $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$. It is clearly, the profile likelihood function $L_{\mathrm{P}}(\boldsymbol{\theta})$ is a continuous and positive function of $\boldsymbol{\theta}$. Then, for $i = 1, \ldots, R$, we consider

$$\begin{aligned} &\lim_{\theta_i \to \infty} c|\boldsymbol{\Sigma}|^{1/D}|\mathbf{V}|^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\} \\ &= \lim_{\theta_i \to \infty} \frac{c\prod_{r=1}^{R}\theta_r^{1/D}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\}}{|\mathbf{V}|^{1/2}} \\ &= \lim_{\theta_i \to \infty} \frac{c\prod_{r=1}^{R}\theta_r^{1/D}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\}}{\prod_{d=1}^{D}|\mathbf{V}_d|^{1/2}} \\ &= \lim_{\theta_i \to \infty} \frac{c\prod_{r=1}^{R}\theta_r^{1/D}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\}}{\prod_{d=1}^{D}|\boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{ed}}|^{1/2}} \\ &= \lim_{\theta_i \to \infty} \frac{c\prod_{r=1}^{R}\theta_r^{1/D}}{\prod_{d=1}^{D}|\boldsymbol{\Sigma} + \mathbf{V}_{ed}|^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\} \\ &= \left(\lim_{\theta_i \to \infty} \frac{c\prod_{r=1}^{R}\theta_r^{1/D}}{\prod_{d=1}^{D}|\boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{ed}}|^{1/2}}\right)\left(\lim_{\theta_i \to \infty}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\}\right) \\ &= 0, \end{aligned}$$

where we use the fact that $\lim_{\theta_i \to \infty}\exp\left\{-\frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y}\right\} = 1$ for $i = 1, \ldots, R$.

By Lemma 3.1, there exist $\hat{\boldsymbol{\theta}} = (\theta_{01}, \ldots, \theta_{0R})'$ such that $L_{\mathrm{AP}}(\boldsymbol{\theta})$ attain maximum at $(\theta_{01}, \ldots, \theta_{0R})$ where $\theta_{0i} > 0$ for all $i = 1, \ldots, R$. The same technique can be applied to show that the AREML estimator is strictly positive. □

3.3.2 Asymptotic Properties of the AML Estimators

In this section, we will prove the asymptotic properties of the APML and AREML estimators. First, we obtain the asymptotic representations of $\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{AR}} - \boldsymbol{\theta}$ for the multivariate Fay-Herriot model. To prove this asymptotic representations of $\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{AR}} - \boldsymbol{\theta}$, we need the following lemmas and corollaries. The letter c appeared in this section stands for a constant which vary among difference places.

Lemma 3.3. ([7]) Let $\mathbf{Q}$ be a symmetric matrix, and $\boldsymbol{\xi} \sim N(\mathbf{0}, \mathbf{I})$. Then, for any $m \geq 2$, there is a constant c that only depends on m such that

$$\mathrm{E}\left|\boldsymbol{\xi}'\mathbf{Q}\boldsymbol{\xi} - \mathrm{E}\left[\boldsymbol{\xi}'\mathbf{Q}\boldsymbol{\xi}\right]\right|^m \leq c\,||\mathbf{Q}||_{\mathrm{F}}^m\,,$$

where $||\mathbf{Q}||_{\mathrm{F}}$ is defined as $(\mathrm{tr}(\mathbf{Q}'\mathbf{Q}))^{1/2}$.

Lemma 3.4. ([7]) Let $\ell(\boldsymbol{\theta})$ be the likelihood function of $\boldsymbol{\theta}$. For any $\hat{\boldsymbol{\theta}}$ which is obtained as a solution to a "score" equation of the form $\partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$, suppose that

1. $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}, \mathbf{y})$ is three times continuously differentiable with respect to $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_s)'$, where $\mathbf{y} = (y_1, \ldots, y_n)'$,

2. $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$, the interior of $\boldsymbol{\Theta}$,

3. $-\infty < \limsup_{n\to\infty} \lambda_{\max}(\mathbf{G}^{-1}\mathbf{A}\mathbf{G}^{-1}) < 0$, where λ_{max} means the largest eigenvalue, $\mathbf{A} = \mathrm{E}[\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2]$, and $\mathbf{G} = \mathrm{diag}(g_1, \ldots, g_s)$ with g_i, $1 \leq i \leq s$ such that $g_* = \min_{1\leq i\leq s} g_i \to \infty$ as $n \to \infty$,

4. the mth moments of the following are bounded ($m > 0$):

   (a) $\dfrac{1}{g_i}\left|\dfrac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_i}\right|$, $1 \leq i \leq s$,

   (b) $\dfrac{1}{\sqrt{g_i g_j}}\left|\dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j} - \mathrm{E}\left[\dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right]\right|$, $1 \leq i, j \leq s$,

   (c) $\dfrac{g_*}{g_i g_j g_k}\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\dfrac{\partial^3\ell(\tilde{\boldsymbol{\theta}})}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|$, $1 \leq i, j, k \leq s$,

   where $S_\delta(\boldsymbol{\theta}) = \{\tilde{\boldsymbol{\theta}} : \left|\tilde{\theta}_i - \theta_i\right| \leq \delta g_*/g_i, 1 \leq i \leq s\}$ for some $\delta > 0$.

Then there exists $\hat{\boldsymbol{\theta}}$ such that for any $0 < \rho < 1$, there is a set $\mathcal{B}$ satisfying for large n and on $\mathcal{B}$, $\hat{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$, $\partial\ell(\hat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta} = \mathbf{0}$, $|G(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})| < g_*^{1-\rho}$, and

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = -\mathbf{A}^{-1}\mathbf{a} + \mathbf{r},$$

where $\mathbf{a} = \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$, and $|\mathbf{r}| \leq g_*^{-2\rho}\eta$ with $\mathrm{E}(\eta^m)$ is bounded; and $P(\mathcal{B}^c) \leq c g_*^{-\tau m}$, where $\tau = (1/4) \wedge (1-\rho)$ and c is constant.

We now define an estimator $\hat{\eta}$ of $\eta(\boldsymbol{\theta})$ having the following property:

$$\mathrm{E}(\hat{\eta}) = \eta(\boldsymbol{\theta}) + o(g_*^{-2}). \tag{3.10}$$

It follows from (3.10) that the bias of $\hat{\eta}$ is $o(g_*^{-2})$. Let $\mathbf{a} = \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$, $\mathbf{A} = \mathrm{E}[\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2]$, $\mathbf{b} = \partial\eta(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = (b_i)$, $\mathbf{B} = \partial^2\eta(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2 = (B_{ij})$, $\mathbf{F} = \partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2$, $\mathbf{H}_i = \partial^3\ell(\boldsymbol{\theta})/\partial\theta_i\partial\boldsymbol{\theta}^2$, and $\mathbf{C} = (\mathbf{a}'\mathbf{A}^{-1}\mathbf{H}_i)_{i=1,\ldots,s}$, where s is the dimension of $\boldsymbol{\theta}$. Also, let $\mathbf{Q} = \mathbf{G}^{-1}\mathbf{A}\mathbf{G}^{-1}$, and $\mathbf{W} = \mathbf{Q}^{-1} = (w_{ij})$. Let $\mathbf{G}^{-1}\mathbf{a} = (\lambda_i), \mathbf{G}^{-1/2}(\mathbf{F} - \mathbf{A})\mathbf{G}^{-1/2} = (\lambda_{ij}), \mathbf{G}^{-1}\mathbf{H}_i\mathbf{G}^{-1} = (\lambda_{ijk})$. Then, we define the following vector, matrix and arrays: $\mathbf{U}_0 = (u_i), \mathbf{U}_1 = (u_{il}), \mathbf{U}_2 = (u_{jkl})$ and $\mathbf{U}_3 = (u_{jklmn})$, where $u_i = \mathrm{E}[\lambda_i]$, $u_{il} = \mathrm{E}[\lambda_i\lambda_l]$, $u_{jkl} = \mathrm{E}[\lambda_{jk}\lambda_l]$, and $u_{jklmn} = \mathrm{E}[\lambda_{jkm}\lambda_l\lambda_n]$. Note that all of these are functions of $\boldsymbol{\theta}$. For example, $\mathbf{A} = \mathbf{A}(\boldsymbol{\theta})$. The norm of a r-way array $(r \geq 3)$ $\mathbf{U}$, denoted by $||\mathbf{U}||$, is defined as the maximum of the absolute values of its elements. Recall that the norm of a matrix $\mathbf{M}$ is defined as $\|\mathbf{M}\| = [\lambda_{\max}(\mathbf{M}'\mathbf{M})]^{1/2}$. Define

$$\begin{aligned}
\Delta_0(\boldsymbol{\theta}) &= -2\mathbf{b}'\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{a}],\\
\Delta_1(\boldsymbol{\theta}) &= \mathbf{b}'\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{F}\mathbf{A}^{-1}\mathbf{a}],\\
\Delta_2(\boldsymbol{\theta}) &= \frac{1}{2}\,\mathrm{E}[\mathbf{a}'\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\mathbf{a}],\\
\Delta_3(\boldsymbol{\theta}) &= -\frac{1}{2}\mathbf{b}'\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{C}\mathbf{A}^{-1}\mathbf{a}].
\end{aligned}$$

Finally, we define

$$\hat{\eta} = \eta(\hat{\boldsymbol{\theta}}) - \sum_{j=0}^{3} \Delta_j(\hat{\boldsymbol{\theta}}),$$

provided that $|\hat{\eta}| \leq c_0 g_*^{\lambda}$; otherwise, let $\hat{\eta} = \eta(\boldsymbol{\theta}^*)$, where c_0 and λ are known positive constants, and $\boldsymbol{\theta}^*$ is a given point in $\boldsymbol{\Theta}$.

Lemma 3.5. ([7]) The estimator $\hat{\eta}$ of $\eta(\boldsymbol{\theta})$ given above satisfies the property (3.10) provide that

1. $\eta(\boldsymbol{\theta})$ is three times continuously differentiable and the following are bounded: $\eta(\boldsymbol{\theta})$, $|\mathbf{b}|$, $||\mathbf{B}||$ and
$$\sup_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta}_0)} \left| \frac{\partial^3 \eta(\tilde{\boldsymbol{\theta}})}{\partial \theta_i \partial \theta_j \partial \theta_k} \right|, \quad 1 \leq i, j, k \leq s,$$
where δ is positive number and $S_\delta(\boldsymbol{\theta}) = \{\tilde{\boldsymbol{\theta}} : |\tilde{\theta}_i - \theta_i| \leq \delta g_*/g_i, 1 \leq i \leq s\}$.

2. The conditions of Lemma 3.4 hold with $m > 8+4\lambda$ and $\ell(\boldsymbol{\theta})$ four times continuously differentiable with respect to $\boldsymbol{\theta}$.

3. The mth moments of the following are bounded:

   (a) $\dfrac{1}{\sqrt{g_j g_k}} \left| \dfrac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} - \mathrm{E}\left[\dfrac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} \right] \right|, 1 \leq i, j, k \leq s,$

   (b) $\dfrac{1}{g_j g_k} \left| \dfrac{\partial^3 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j \partial \theta_k} \right|, \ 1 \leq i, j, k \leq s,$

   (c) $\dfrac{g_*^2}{g_i g_j g_k g_l} \sup\limits_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta})} \left| \dfrac{\partial^4 \ell(\tilde{\boldsymbol{\theta}})}{\partial \theta_i \partial \theta_j \partial \theta_k \partial \theta_l} \right|, \ 1 \leq i, j, k, l \leq s.$

4. $\sup\limits_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta})} ||\mathbf{Q}(\tilde{\boldsymbol{\theta}}) - \mathbf{Q}(\boldsymbol{\theta})|| \to 0$, and $\sup\limits_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta})} ||\mathbf{U}_j(\tilde{\boldsymbol{\theta}}) - \mathbf{U}_j(\boldsymbol{\theta})|| \to 0$, $j = 1, 2, 3$, as $\delta \to 0$ uniformly in n.

5. $|\,\mathrm{E}[\mathbf{a}]|$ is bounded and $\sup\limits_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta})} \left| \mathrm{E}[\mathbf{a}]|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} - \mathrm{E}[\mathbf{a}] \right| \to 0$, as $\delta \to 0$ uniformly in n.

Lemma 3.6. Let $\mathbf{V} = \mathbf{I} \otimes \boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{e}}$ be the covariance matrix of $\mathbf{y}$ in multivariate Fay-Herriot model (3.2). We have the following properties:

(1) $\dfrac{\partial \mathbf{V}}{\partial \theta_i} = \mathbf{I}_D \otimes \mathbf{E}_i$, where $\mathbf{E}_i$ is the diagonal matrix of an $R \times 1$ identity vector $\mathbf{e}_i$,

(2) $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is symmetric,

(3) $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is idempotent,

(4) $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ has D eigenvalues equal to 1 and $D(R-1)$ eigenvalues equal to 0,

(5) $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is positive semidefinite,

(6) $\left(\dfrac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\dfrac{\partial \mathbf{V}}{\partial \theta_j}\right) = \mathbf{0}_{DR \times DR}$ for $i \neq j$.

Proof. For (1), note that

$$\begin{aligned}
\frac{\partial \mathbf{V}}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i}(\mathbf{I} \otimes \boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{e}}) \\
&= \frac{\partial}{\partial \theta_i}(\mathbf{I} \otimes \boldsymbol{\Sigma}) + \frac{\partial \mathbf{V}_{\mathrm{e}}}{\partial \theta_i} \\
&= \frac{\partial \mathbf{I}}{\partial \theta_i} \otimes \boldsymbol{\Sigma} + \mathbf{I} \otimes \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} + \mathbf{0}_{DR \times DR} \\
&= \mathbf{I} \otimes \mathbf{E}_i,
\end{aligned}$$

where we use the fact that $\boldsymbol{\Sigma} = \mathrm{diag}_{1 \leq r \leq R}(\theta_r)$ to obtain last equation and $\mathbf{E}_i = \mathrm{diag}(\mathbf{e}_i)$ is the diagonal matrix of an $R \times 1$ identity vector $\mathbf{e}_i$.

For (2), note that

$$\begin{aligned}
\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)' &= (\mathbf{I}_D \otimes \mathbf{E}_i)' \\
&= \mathbf{I}_D' \otimes \mathbf{E}_i' \\
&= \mathbf{I}_D \otimes \mathbf{E}_i,
\end{aligned}$$

where we use the fact that $\mathbf{E}_i' = (\text{diag}(\mathbf{e}_i))' = \text{diag}(\mathbf{e}_i) = \mathbf{E}_i$ to obtain the last equation. Hence $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is symmetric.

For (3), note that, using theorem 2.48,

$$\begin{aligned}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right) &= (\mathbf{I}_D \otimes \mathbf{E}_i)(\mathbf{I}_D \otimes \mathbf{E}_i) \\ &= (\mathbf{I}_D\mathbf{I}_D \otimes \mathbf{E}_i\mathbf{E}_i) \\ &= (\mathbf{I}_D \otimes \mathbf{E}_i) \qquad (3.11) \\ &= \frac{\partial \mathbf{V}}{\partial \theta_i},\end{aligned}$$

where we use the fact that $\mathbf{E}_i\mathbf{E}_i = (\text{diag}(\mathbf{e}_i))(\text{diag}(\mathbf{e}_i)) = \text{diag}(\mathbf{e}_i) = \mathbf{E}_i$ to obtain (3.11). Hence $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is idempotent.

For (4), Since $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is a symmetric and idempotent matrix of rank D, by Theorem 2.43, $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ has D eigenvalues equal to 1 and $D(R-1)$ eigenvalues equal to 0.

For (5), since $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ has the zero eigenvalue, it is singular. Since $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is singular, symmetric, and idempotent matrix, from Theorem 2.42, $\dfrac{\partial \mathbf{V}}{\partial \theta_i}$ is positive semidefinite.

For (6), note that

$$\begin{aligned}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right) &= (\mathbf{I}_D \otimes \mathbf{E}_i)(\mathbf{I}_D \otimes \mathbf{E}_j) \\ &= (\mathbf{I}_D\mathbf{I}_D \otimes \mathbf{E}_i\mathbf{E}_j) \\ &= (\mathbf{I}_D \otimes \mathbf{0}_{R\times R}) \\ &= \mathbf{0}_{DR\times DR}.\end{aligned}$$

Hence $\left(\dfrac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\dfrac{\partial \mathbf{V}}{\partial \theta_j}\right) = \mathbf{0}_{DR\times DR}$ for $i \neq j$. □

Lemma 3.7. Under the multivariate Fay-Herriot model (3.2), for $i = 1, \ldots, R$,

(1) $$0 \leq \mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \leq \frac{DR}{\lambda_{\min}(\mathbf{V})},$$

and

(2) $$0 \leq \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \leq \frac{DR - pR}{\lambda_{\min}(\mathbf{V})},$$

where $\mathbf{V} = \mathbf{I}_D \otimes \mathbf{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For (1), since $\mathbf{V}$ is positive definite, from Theorem 2.27, $\mathbf{V}^{-1}$ is positive definite. From Theorem 2.39, Lemma 3.6(5), and Corollary 2.36, we have

$$\begin{aligned}
\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) &= \mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\right) \\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathrm{tr}(\mathbf{V}^{-1}) \\
&= \mathrm{tr}(\mathbf{V}^{-1}) \\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}(\mathbf{I}_{DR}) \\
&= \frac{DR}{\lambda_{\min}(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 to obtain the last equation.

Then, from Theorem 2.39, Lemma 3.6(5) and Lemma 3.6(4), we have

$$\begin{aligned}
\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) &= \mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\right) \\
&\geq \lambda_{\min}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathrm{tr}(\mathbf{V}^{-1}) \\
&= 0.
\end{aligned}$$

Hence,

$$0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \leq \frac{DR}{\lambda_{\min}(\mathbf{V})},$$

for $i = 1, \ldots, R$.

For (2), let $\mathbf{B} = \mathbf{I}_{DR} - \mathbf{V}^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1/2}$. Then $\mathbf{B}$ is symmetric and idempotent. Therefore, the eigenvalues of $\mathbf{B}$ are either 0 or 1. Since $\mathbf{B}$ is symmetric and has non-negative eigenvalues, $\mathbf{B}$ is positive semidefinite. From Theorem 2.26(ii), we have $\mathbf{P} = (\mathbf{V}^{-1/2}\mathbf{B})(\mathbf{V}^{-1/2}\mathbf{B})'$ is positive semidefinite. From Theorem 2.39, we have

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) &= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}(\mathbf{V}^{-1/2}\mathbf{B})(\mathbf{V}^{-1/2}\mathbf{B})'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\operatorname{tr}\left((\mathbf{V}^{-1/2}\mathbf{B})(\mathbf{V}^{-1/2}\mathbf{B})'\right)\\
&= \operatorname{tr}(\mathbf{V}^{-1}\mathbf{B})\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}(\mathbf{B})\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}(\mathbf{I}_{DR} - \mathbf{V}^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1/2})\\
&= \lambda_{\max}(\mathbf{V}^{-1})\left[\operatorname{tr}(\mathbf{I}_{DR}) - \operatorname{tr}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1})\right]\\
&= \frac{DR - pR}{\lambda_{\min}(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 to obtain the last equation.

Then, from Theorem 2.39, Lemma 3.6(5) and Lemma 3.6(4), we have

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) &= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}(\mathbf{V}^{-1/2}\mathbf{B})(\mathbf{V}^{-1/2}\mathbf{B})'\right)\\
&\geq \lambda_{\min}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\operatorname{tr}\left((\mathbf{V}^{-1/2}\mathbf{B})(\mathbf{V}^{-1/2}\mathbf{B})'\right)\\
&= 0.
\end{aligned}$$

Hence

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \leq \frac{DR - pR}{\lambda_{\min}(\mathbf{V})},$$

for $i = 1, \ldots, R$. □

Lemma 3.8. Under the multivariate Fay-Herriot model (3.2), for $i, j = 1, \ldots, R$,

(1) $$0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \leq \frac{DR}{\lambda^2_{\min}(\mathbf{V})},$$

and

(2) $$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \leq \frac{DR - pR}{\lambda^2_{\min}(\mathbf{V})},$$

where $\mathbf{V} = \mathbf{I}_D \otimes \mathbf{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For (1), from the property that $(\partial \mathbf{V}/\partial \theta_i)$ is idempotent, Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) &= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right) \\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right) \\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right) \\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right) \\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \\
&\leq \frac{DR}{\lambda^2_{\min}(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 and Lemma 3.7(1) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial\mathbf{V}/\partial\theta_j) = 0$, we have

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) &= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\right) \\
&= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\left(\mathbf{V}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&\geq \lambda_{\min}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&= 0. \qquad (3.12)
\end{aligned}$$

Hence

$$0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR}{\lambda_{\min}^2(\mathbf{V})},$$

for $i, j = 1, \ldots, R$.

For (2), from the proof of Lemma 3.7, we know that $\mathbf{P} = \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}$ is positive semidefinite and $\mathbf{B}$ is positive semidefinite and idempotent. Then, from the property that $(\partial\mathbf{V}/\partial\theta_i)$ is idempotent, Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) &= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\right) \\
&= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&\leq \lambda_{\max}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) \\
&\leq \frac{DR - pR}{\lambda_{\min}^2(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 and Lemma 3.7(2) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial \mathbf{V}/\partial\theta_j) = 0$, we have

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right) &= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\right) \\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)'\right) \\
&\geq \lambda_{\min}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)'\right) \\
&= 0. \qquad (3.13)
\end{aligned}$$

Hence

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR - pR}{\lambda^2_{\min}(\mathbf{V})},$$

for $i, j = 1, \ldots, R$. □

Lemma 3.9. Under the multivariate Fay-Herriot model (3.2), for $i, j, k = 1, \ldots, R$,

(1)
$$\begin{aligned}
-\frac{2DR}{\lambda^3_{\min}(\mathbf{V})} \leq &\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\right) \\
&+ \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\right) \leq \frac{4DR}{\lambda^3_{\min}(\mathbf{V})},
\end{aligned}$$

and

(2)
$$\begin{aligned}
-\frac{2(DR - pR)}{\lambda^3_{\min}(\mathbf{V})} \leq &\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\right) \\
&+ \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\right) \leq \frac{4(DR - pR)}{\lambda^3_{\min}(\mathbf{V})},
\end{aligned}$$

where $\mathbf{V} = \mathbf{I}_D \otimes \boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For (1), from Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\\
&\leq \frac{DR}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned}\tag{3.14}$$

where we use Corollary 2.33 and Lemma 3.8(1) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial \mathbf{V}/\partial\theta_j) = 0$, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&\geq \lambda_{\min}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&= 0.
\end{aligned}\tag{3.15}$$

Thus, from (3.14) and (3.15),

$$0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR}{\lambda_{\min}^3(\mathbf{V})}, \tag{3.16}$$

for $i, j = 1, \ldots, R$.

From Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&=\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&=\operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\qquad\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&=\lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right.\\
&\qquad\qquad\left.+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&\leq \frac{4DR}{\lambda^3_{\min}(\mathbf{V})},
\end{aligned} \tag{3.17}$$

where we use Corollary 2.33 and Lemma 3.8(1) to obtain the last inequality.

By using the same technique as (3.15) and property that $\lambda_{\min}(\partial\mathbf{V}/\partial\theta_k)=0$, we have

$$\begin{aligned}
0\leq\ &\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right).
\end{aligned} \tag{3.18}$$

From (3.16), (3.17) and (3.18), we have

$$\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\leq\frac{4DR}{\lambda^3_{\min}(\mathbf{V})},$$

and

$$-\frac{2DR}{\lambda_{\min}^3(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right).$$

Hence,

$$-\frac{2DR}{\lambda_{\min}^3(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \leq \frac{4DR}{\lambda_{\min}^3(\mathbf{V})},$$

for $i, j, k = 1, \ldots, R$.

For (2), from Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\\
&\leq \frac{DR - pR}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 and Lemma 3.8(2) to obtain the last inequality.

By using the same technique as (3.15) and the property that $\lambda_{\min}(\partial \mathbf{V}/\partial \theta_j) = 0$, we have

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right).$$

Thus,

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \leq \frac{DR - pR}{\lambda_{\min}^3(\mathbf{V})}, \tag{3.19}$$

for $i, j = 1, \ldots, R$.

From Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \\
&= \operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\Bigg) \\
&= \operatorname{tr}\Bigg(\frac{\partial \mathbf{V}}{\partial \theta_k}\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right) \\
&\qquad \left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right)'\Bigg) \\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\operatorname{tr}\Bigg(\left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right) \\
&\qquad \left(\mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right)'\Bigg) \\
&= \operatorname{tr}\Bigg(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right) \\
&\qquad \left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right)'\Bigg) \\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\Bigg(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right) \\
&\qquad \left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B} + \mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{B}\right)'\Bigg) \\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \\
&\leq \frac{4(DR - pR)}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned} \tag{3.20}$$

where we use Corollary 2.33 and Lemma 3.8(2) to obtain the last inequality.

By using the same technique as (3.15) and the property that $\lambda_{\min}(\partial\mathbf{V}/\partial\theta_k) = 0$, we have

$$0 \le \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right). \tag{3.21}$$

From (3.19), (3.20) and (3.21), we have

$$\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) \le \frac{4(DR-pR)}{\lambda^3_{\min}(\mathbf{V})},$$

and

$$-\frac{2(DR-pR)}{\lambda^3_{\min}(\mathbf{V})} \le \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right),$$

Hence,

$$-\frac{2(DR-pR)}{\lambda^3_{\min}(\mathbf{V})} \le \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) \le \frac{4(DR-pR)}{\lambda^3_{\min}(\mathbf{V})},$$

for $i, j, k = 1, \ldots, R$. □

Lemma 3.10. Under the multivariate Fay-Herriot model (3.2), for $i, j, k, l = 1, \ldots, R$,

(1)
$$-\frac{2DR}{\lambda^4_{\min}(\mathbf{V})} \le \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) \le \frac{4DR}{\lambda^4_{\min}(\mathbf{V})},$$

and

(2)
$$-\frac{2(DR-pR)}{\lambda^4_{\min}(\mathbf{V})} \le \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) \le \frac{4(DR-pR)}{\lambda^4_{\min}(\mathbf{V})},$$

where $\mathbf{V} = \mathbf{I}_D \otimes \mathbf{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For (1), from Theorem 2.26(ii) and Theorem 2.39, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_l}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\\
&\leq \frac{DR}{\lambda_{\min}^4(\mathbf{V})},
\end{aligned}
\tag{3.22}
$$

where we use Corollary 2.33 and (3.16) to obtain last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial \mathbf{V}/\partial \theta_l) = 0$, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_l}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\geq \lambda_{\min}\left(\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&= 0.
\end{aligned}
\tag{3.23}
$$

Thus, from (3.22) and (3.23),

$$
0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right) \leq \frac{DR}{\lambda_{\min}^4(\mathbf{V})}, \tag{3.24}
$$

for $i, j, l = 1, \ldots, R$.

From Theorem 2.26(ii) and Theorem 2.39, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\\
&=\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_l}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\le \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.\\
&\qquad\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&=\operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\le \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.\\
&\qquad\qquad\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&=\lambda_{\max}\left(\mathbf{V}^{-1}\right)\\
&\quad\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right.\\
&\qquad\left.+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&\le \frac{4DR}{\lambda_{\min}^4(\mathbf{V})},
\end{aligned}
\tag{3.25}
$$

where we use Corollary 2.33 and Lemma 3.9(1) to obtain the last inequality.

By using the same technique as (3.23) and the property that $\lambda_{\min}(\partial \mathbf{V}/\partial \theta_l) = 0$, we have

$$
\begin{aligned}
0 \le\; &\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right).
\end{aligned}
\tag{3.26}
$$

From (3.24), (3.25) and (3.26), we have

$$\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)$$
$$\leq \frac{4DR}{\lambda^4_{\min}(\mathbf{V})},$$

and

$$\begin{aligned}-\frac{2DR}{\lambda^4_{\min}(\mathbf{V})} \leq &\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)\\ &+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right).\end{aligned}$$

Hence, for $i, j, k, l = 1, \ldots, R$,

$$\begin{aligned}-\frac{2DR}{\lambda^4_{\min}(\mathbf{V})} \leq &\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)\\ &+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) \leq \frac{4DR}{\lambda^4_{\min}(\mathbf{V})}.\end{aligned}$$

For (2), by using the same technique as (3.25) and (3.26), we have

$$\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) \leq \frac{4(DR-pR)}{\lambda^4_{\min}(\mathbf{V})},$$

and

$$-\frac{2(DR-pR)}{\lambda^4_{\min}(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right).$$

Hence, for $i, j, k, l = 1, \ldots, R$,

$$\begin{aligned}-\frac{2(DR-pR)}{\lambda^4_{\min}(\mathbf{V})} \leq &\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)\\ &+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) \leq \frac{4(DR-pR)}{\lambda^4_{\min}(\mathbf{V})}.\end{aligned}$$

□

Lemma 3.11. Under the multivariate Fay-Herriot model (3.2), for $i, j, k, l, m = 1, \ldots, R$,

$$(1) \qquad -\frac{2DR}{\lambda_{\min}^5(\mathbf{V})} \le \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) \le \frac{4DR}{\lambda_{\min}^5(\mathbf{V})},$$

and

$$(2) \qquad -\frac{2(DR-pR)}{\lambda_{\min}^5(\mathbf{V})} \le \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) \le \frac{4(DR-pR)}{\lambda_{\min}^5(\mathbf{V})},$$

where $\mathbf{V} = \mathbf{I}_D \otimes \boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For (1), from Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&\le \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&\le \frac{DR}{\lambda_{\min}^4(\mathbf{V})}, \qquad (3.27)
\end{aligned}$$

where we use Corollary 2.33 and (3.16) to obtain the last inequality.

Using Theorem 2.39, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&\geq \lambda_{\min}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\right)'\right)\\
&= \lambda_{\min}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\\
&\geq 0, \qquad (3.28)
\end{aligned}
$$

where we use (3.14) to obtain the last inequality.

Thus, from (3.27) and (3.28),

$$
0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR}{\lambda_{\min}^4(\mathbf{V})}, \qquad (3.29)
$$

for $i, j = 1, \ldots, R$.

From Theorem 2.26(ii) and Theorem 2.39, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\\
&= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_m}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\\
&\leq \frac{DR}{\lambda_{\min}^5(\mathbf{V})}, \qquad (3.30)
\end{aligned}
$$

where we use Corollary 2.33 and (3.27) to obtain last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial\mathbf{V}/\partial\theta_m) = 0$, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\\
&= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_m}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&\geq \lambda_{\min}\left(\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\right)'\right)\\
&= 0. \qquad (3.31)
\end{aligned}$$

Thus, from (3.30) and (3.31),

$$0 \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right) \leq \frac{DR}{\lambda_{\min}^5(\mathbf{V})}, \qquad (3.32)$$

for $i, j, m = 1, \ldots, R$.

From Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\\
&+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\\
&= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_m}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial\mathbf{V}}{\partial\theta_m}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)'\right)
\end{aligned}$$

$$
\begin{aligned}
&\le \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad\qquad\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right.\\
&\qquad\qquad + \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\\
&\qquad\qquad + \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\\
&\qquad\qquad \left.+ \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\right)\\
&\le \frac{4DR}{\lambda_{\min}^5(\mathbf{V})}, \qquad (3.33)
\end{aligned}
$$

where we use Corollary 2.33, Lemma 3.10(1) and (3.29) to obtain the last inequality.

By using the same technique as (3.31) and property that $\lambda_{\min}(\partial \mathbf{V}/\partial\theta_m) = 0$, we have

$$
\begin{aligned}
0 \le\ &\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_m}\right) \qquad (3.34)\\
&+ \mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_m}\right)\\
&+ \mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_m}\right)\\
&+ \mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_m}\right).
\end{aligned}
$$

From (3.32), (3.33) and (3.34), we have

$$
\begin{aligned}
&\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_m}\right)\\
&+ \mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_m}\right)\\
&\le \frac{4DR}{\lambda_{\min}^5(\mathbf{V})},
\end{aligned}
$$

and

$$-\frac{2DR}{\lambda_{\min}^5(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right).$$

Hence,

$$-\frac{2DR}{\lambda_{\min}^5(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) \leq \frac{4DR}{\lambda_{\min}^5(\mathbf{V})},$$

for $i, j, k, l, m = 1, \ldots, R$.

For (2), by using the same technique as (3.30) and (3.31),

$$\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) \leq \frac{4(DR - pR)}{\lambda_{\min}^5(\mathbf{V})},$$

and

$$-\frac{2(DR - pR)}{\lambda_{\min}^5(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right).$$

Hence,

$$-\frac{2(DR - pR)}{\lambda_{\min}^5(\mathbf{V})} \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) \leq \frac{4(DR - pR)}{\lambda_{\min}^5(\mathbf{V})},$$

for $i, j, k, l, m = 1, \ldots, R$. $\square$

Lemma 3.12. Under the multivariate Fay-Herriot model (3.2), for $i, j = 1, \ldots, R$,

(1) $$0 \le \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \le \frac{DR - pR}{\lambda_{\min}^4(\mathbf{V})}$$

and

(2) $$0 \le \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \le \frac{DR - pR}{\lambda_{\min}^4(\mathbf{V})},$$

where $\mathbf{V} = \mathbf{I}_D \otimes \boldsymbol{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For (1), from the proof of Lemma 3.7, we know that $\mathbf{P} = \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}$ is positive semidefinite and $\mathbf{B}$ is positive semidefinite and idempotent. From Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \\
&= \operatorname{tr}\left(\mathbf{B}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right) \\
&\le \lambda_{\max}(\mathbf{B})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right) \\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right) \\
&\le \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right) \\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \\
&\le \frac{DR - pR}{\lambda_{\min}^4(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 and (3.19) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\mathbf{B}) = 0$, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \operatorname{tr}\left(\mathbf{B}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\geq \lambda_{\min}(\mathbf{B})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\geq 0.
\end{aligned}$$

Hence,

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \leq \frac{DR - pR}{\lambda^4_{\min}(\mathbf{V})},$$

for $i, j = 1, \ldots, R$.

For (2), from Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\operatorname{tr}\left(\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\\
&\leq \frac{DR - pR}{\lambda^4_{\min}(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 and (3.19) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial \mathbf{V}/\partial\theta_j) = 0$, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)'\right)\\
&\geq \lambda_{\min}\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\operatorname{tr}\left(\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)'\right)\\
&= 0.
\end{aligned}$$

Hence,

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR - pR}{\lambda^4_{\min}(\mathbf{V})},$$

for $i, j = 1, \ldots, R$. □

Lemma 3.13. Under the multivariate Fay-Herriot model (3.2), for $i, j, k = 1, \ldots, R$,

(1) $$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\right) \leq \frac{DR - pR}{\lambda^6_{\min}(\mathbf{V})}$$

and

(2) $$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR - pR}{\lambda^6_{\min}(\mathbf{V})},$$

where $\mathbf{V} = \mathbf{I}_D \otimes \mathbf{\Sigma} + \mathbf{V}_{\mathrm{e}}$, $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$.

Proof. For(1), from the proof of Lemma 3.7, we know that $\mathbf{P} = \mathbf{V}^{-1/2}\mathbf{B}\mathbf{V}^{-1/2}$ is positive semidefinite and $\mathbf{B}$ is positive semidefinite and idempotent. From Theorem 2.26(ii) and Theorem 2.39, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&= \operatorname{tr}\left(\mathbf{B}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{B})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\operatorname{tr}\left(\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\Bigg(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\\
&\qquad\qquad\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\Bigg)\\
&= \lambda_{\max}^2(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&\leq \frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})},
\end{aligned}
$$

where we use Corollary 2.33 and Lemma 3.12(1) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\mathbf{B}) = 0$, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&= \operatorname{tr}\left(\mathbf{B}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&\geq \lambda_{\min}(\mathbf{B})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{B}\right)'\right)\\
&= 0.
\end{aligned}
$$

Hence,

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \leq \frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})},$$

for $i, j, k = 1, \ldots, R$.

For (2), from Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\operatorname{tr}\left(\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\mathbf{B}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\lambda_{\max}\left(\mathbf{B}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&= \lambda_{\max}^2\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&\leq \lambda_{\max}^2\left(\mathbf{V}^{-1}\right)\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\operatorname{tr}\left(\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&= \lambda_{\max}^2\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&\leq \lambda_{\max}^2\left(\mathbf{V}^{-1}\right)\lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\left(\mathbf{B}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)'\right)\\
&= \lambda_{\max}^3(\mathbf{V}^{-1})\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\\
&\leq \frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})},
\end{aligned}$$

where we use Corollary 2.33 and (3.19) to obtain the last inequality.

Using Theorem 2.39 and the property that $\lambda_{\min}(\partial\mathbf{V}/\partial\theta_k) = 0$, we have

$$\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\\
&= \operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)'\right)\\
&\geq \lambda_{\min}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\operatorname{tr}\left(\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)'\right)\\
&= 0.
\end{aligned}$$

Hence,

$$0 \leq \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) \leq \frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})},$$

for $i, j, k = 1, \ldots, R$. □

Remark 3.14. Under the multivariate Fay-Herriot model (3.2), for fixed p and R, for $i, j, k, l, m = 1, \ldots, R$, from Lemma 3.7 – Lemma 3.11

(1) $\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) = O(D)$,

(2) $\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) = O(D)$,

(3) $\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) = O(D)$,

(4) $\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) = O(D)$,

(5) $\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k} + \mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) = O(D)$,

(6) $\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right) = O(D)$,

(7) $\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) = O(D)$,

(8) $\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right) = O(D)$,

(9) $\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right)$
$+\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) = O(D),$

(10) $\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_m}\right) = O(D).$

Lemma 3.15. Under the multivariate Fay-Herriot model (3.2), for fixed p and R, for $i, j, k, l = 1, \ldots, R$,

(1) $\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right| = O(1),$

(2) $\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| = O(1),$

(3) $\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\right.$
$\left.- \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k} + \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\right| = O(1),$

(4) $\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\right.$
$\left.- \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l} + \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_l}\right)\right|$
$= O(1).$

Proof. For (1), let $\mathbf{C} = \mathbf{V}^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1/2}$. Since $\mathbf{V}^{-1}$ is positive definite and $\mathbf{X}$ is full rank matrix, by Corollary 2.22 and Theorem 2.27, $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ is positive definite. By applying Corollary 2.23, $\mathbf{C}$ is positive semidefinite. Therefore, from Theorem 2.39,

$$\begin{aligned}
\operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) &= \operatorname{tr}\left(\mathbf{V}^{-1}\mathbf{C}\right) \\
&\leq \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}(\mathbf{C}) \\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}(\mathbf{V}^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1/2}) \\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}) \\
&= \frac{pR}{\lambda_{\min}(\mathbf{V})},
\end{aligned} \tag{3.35}$$

where we use Corollary 2.33 to obtain the last inequality.

Since $\mathbf{V}^{-1/2}$ is nonsingular, from Theorem 2.21(2), $\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}$ is positive semidefinite. Moreover, from the definition of $\mathbf{C}$ and $\mathbf{P}$ can be alternatively written as $\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}$. From Theorem 2.26(ii) and Theorem 2.39, we have

$$\begin{aligned}
\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right| &= \left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i} - \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right| \\
&= \left|\operatorname{tr}\left(-\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right| \\
&= \left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\right| \\
&= \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&= \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\leq \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\leq \frac{pR}{\lambda_{\min}(\mathbf{V})},
\end{aligned}$$

where we use (3.35) to obtain the last inequality.

Hence, for p and R are fixed,

$$\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right| = O(1),$$

for $i = 1, \ldots, R$.

For (2), using the triangle inequality, we have

$$\begin{aligned}
&\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| \\
&= \left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j} - \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| \\
&= \left|\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j} - \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j} - \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right.\right. \\
&\qquad \left.\left.+\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j} - \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right|
\end{aligned}$$

$$
\begin{aligned}
&= \left| -\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \right. \\
&\quad \left. + \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| \\
&\le \left| \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \right| \\
&\quad + \left| \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| \\
&= \left| \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\quad + \left| \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&= \left| \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right. \\
&\quad \left. - \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\quad + \left| \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\le \left| \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\quad + \left| \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| + \left| \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\quad + \left| \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|. \qquad (3.36)
\end{aligned}
$$

For the first term of (3.36), from Theorem 2.26(ii), Theorem 2.39, we have

$$
\begin{aligned}
&\left| \operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&= \left| \operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C} + \mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\right.\right. \\
&\qquad \left.\left. \left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C} + \mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\right| \\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&\le \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right. \\
&\qquad \left. + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2} + \frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\le \lambda_{\max}(\mathbf{V}^{-1})\left(\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathrm{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) + \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathrm{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right)\\
&\le \frac{2pR}{\lambda^2_{\min}(\mathbf{V})}, \qquad (3.37)
\end{aligned}
$$

where we use Corollary 2.33 and (3.35) to obtain the last inequality.

For the second and third terms of (3.36), from Theorem 2.26(ii) and Theorem 2.39, we have

$$
\begin{aligned}
&\left|\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\\
&= \left|\mathrm{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\right|\\
&= \mathrm{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&\le \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&= \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\le \lambda_{\max}(\mathbf{V}^{-1})\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathrm{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\le \frac{pR}{\lambda^2_{\min}(\mathbf{V})}, \qquad (3.38)
\end{aligned}
$$

where we use Corollary 2.33 and (3.35) to obtain the last inequality.

For the fourth term of (3.36), from Theorem 2.26(ii) and Theorem 2.39, we have

$$
\begin{aligned}
&\left|\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\\
&= \left|\mathrm{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right)\right|\\
&= \mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right)\\
&\le \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathrm{tr}\left(\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right) \\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right) \\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\leq \frac{pR}{\lambda^2_{\min}(\mathbf{V})},
\end{aligned}
\tag{3.39}
$$

where we use Corollary 2.33 and (3.35) to obtain the last inequality.

Thus, from (3.37) – (3.39),

$$
\begin{aligned}
\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| &\leq \frac{2pR}{\lambda^2_{\min}(\mathbf{V})} + \frac{2pR}{\lambda^2_{\min}(\mathbf{V})} + \frac{pR}{\lambda^2_{\min}(\mathbf{V})} \\
&= \frac{5pR}{\lambda^2_{\min}(\mathbf{V})}.
\end{aligned}
$$

Hence, for p and R are fixed,

$$
\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right| = O(1),
$$

for $i, j = 1, \ldots, R$.

For (3), we consider

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \\
&= \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k} - \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \operatorname{tr}\Bigg(-\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k} - \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad - \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad - \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\Bigg) \\
&= -\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) - \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\quad - \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \\
&\quad - \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right). \qquad (3.40)
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \\
&= -\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) - \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\quad - \operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) \\
&\quad - \operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right). \qquad (3.41)
\end{aligned}
$$

Apply the triangle inequality to the sum of (3.40) and (3.41), we obtain

$$
\begin{aligned}
&\Bigg|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\right)+\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\right)\\
&\quad-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\right)-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\right)\Bigg|\\
&=\Bigg|-\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}+\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\quad-\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}+\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\quad-\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}+\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\quad+\operatorname{tr}\Bigg(\mathbf{V}^{-1}\Bigg(\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\\
&\qquad\qquad+\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\Bigg)\Bigg)\\
&\quad+\operatorname{tr}\Bigg(\mathbf{V}^{-1}\Bigg(\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\\
&\qquad\qquad+\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\Bigg)\Bigg)\\
&\quad+\operatorname{tr}\Bigg(\mathbf{V}^{-1}\Bigg(\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\\
&\qquad\qquad+\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\Bigg)\Bigg)\\
&\quad-\operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\\
&\qquad\qquad-\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\Bigg)\Bigg|\\
&\le\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}+\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\\
&\quad+\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}+\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\\
&\quad+\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}+\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\\
&\quad+\Bigg|\operatorname{tr}\Bigg(\mathbf{V}^{-1}\Bigg(\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\\
&\qquad\qquad+\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\Bigg)\Bigg)\Bigg|\\
&\quad+\Bigg|\operatorname{tr}\Bigg(\mathbf{V}^{-1}\Bigg(\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_k}\\
&\qquad\qquad+\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial\theta_i}\Bigg)\Bigg)\Bigg|
\end{aligned}
$$

$$+\left|\mathrm{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\right.\right.\right.$$
$$\left.\left.\left.+\;\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\right)\right|$$
$$+\left|\mathrm{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\right.\right.$$
$$\left.\left.+\;\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\right|. \tag{3.42}$$

For the first three terms of (3.42), from Theorem 2.26(ii) and Theorem 2.39,

$$\begin{aligned}
&\left|\mathrm{tr}\left(\left(\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&=\left|\mathrm{tr}\left(\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\right| \\
&=\mathrm{tr}\left(\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&=\mathrm{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&\le\lambda_{\max}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\mathrm{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&=\mathrm{tr}\left(\mathbf{V}^{-1}\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&\le\lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right) \\
&=\lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\le\frac{pR}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned} \tag{3.43}$$

where we use Corollary 2.33 and (3.38) to obtain the last inequality.

From Theorem 2.26(2), Theorem 2.39, we have

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\right.\right.\\
&\qquad\left.\left.\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\right|\\
&=\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\right.\\
&\qquad\left.\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&=\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\right.\\
&\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&\le \lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\operatorname{tr}\left(\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\right.\\
&\qquad\qquad\left.\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&=\operatorname{tr}\left(\mathbf{V}^{-2}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&\le \lambda_{\max}(\mathbf{V}^{-2})\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\\
&=\lambda_{\max}^2(\mathbf{V}^{-1})\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right.\\
&\qquad\qquad\left.+\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&=\lambda_{\max}^2(\mathbf{V}^{-1})\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\\
&\le \lambda_{\max}^2(\mathbf{V}^{-1})\left(\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)+\lambda_{\max}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right)\\
&\le \frac{2pR}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned}
\tag{3.44}
$$

where we use Corollary 2.33 and (3.35) to obtain the last inequality.

From (3.43) and (3.44), we have

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&=\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right.\right.\right. \\
&\qquad\left.\left.+\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\quad\left.-\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\le\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right.\right.\right. \\
&\qquad\left.\left.\left.+\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\quad+\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\le\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)\right.\right. \\
&\qquad\left.\left.\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\right)'\right)\right| \\
&\quad+\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\quad+\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right| \\
&\le\frac{2pR}{\lambda^3_{\min}(\mathbf{V})}+\frac{pR}{\lambda^3_{\min}(\mathbf{V})}+\frac{pR}{\lambda^3_{\min}(\mathbf{V})} \\
&=\frac{4pR}{\lambda^3_{\min}(\mathbf{V})}.
\end{aligned}
\tag{3.45}
$$

Similarly, we have

$$
\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\le\frac{4pR}{\lambda^3_{\min}(\mathbf{V})},
\tag{3.46}
$$

and

$$
\left|\operatorname{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right|\le\frac{4pR}{\lambda^3_{\min}(\mathbf{V})}.
\tag{3.47}
$$

For the second three terms of (3.42), from Theorem 2.26(ii) and Theorem 2.39, we consider

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\right| \\
&= \left|\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right)\right| \\
&= \operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&\le \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\right) \\
&= \lambda_{\max}(\mathbf{V}^{-1})\operatorname{tr}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right) \\
&\le \frac{pR}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned}
\tag{3.48}
$$

where we use Corollary 2.33 and (3.39) to obtain the last inequality.

From Theorem 2.26(ii) and Theorem 2.39 and triangle inequality, we have

$$
\begin{aligned}
&\Bigg|\operatorname{tr}\Bigg(\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) \\
&\qquad\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\Bigg)\Bigg| \\
&= \operatorname{tr}\Bigg(\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) \\
&\qquad\left(\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\Bigg) \\
&= \operatorname{tr}\Bigg(\mathbf{V}^{-1}\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) \\
&\qquad\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\Bigg) \\
&\le \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\Bigg(\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) \\
&\qquad\left(\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}+\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)'\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right.\\
&\qquad + \frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\\
&\qquad + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\\
&\qquad \left.+ \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\\
&= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\left(\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right.\\
&\qquad \left.+\operatorname{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)\right)\\
&\le \frac{2pR}{\lambda_{\min}^3(\mathbf{V})},
\end{aligned}
\tag{3.49}
$$

where we use Corollary 2.33 and (3.39) to obtain the last inequality.

Apply the triangle inequality, from (3.48) and (3.49), we have

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right.\right.\right.\\
&\qquad \left.\left.\left.+ \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\right)\right|\\
&= \left|\operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right.\right.\right.\\
&\qquad + \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\\
&\qquad + \frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\\
&\qquad \left.\left.+ \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right)\\
&\quad - \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right.\right.\\
&\qquad \left.\left.\left.+ \frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right)\right|\\
&\le \left|\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} + \mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right.\right.\\
&\qquad \left.\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} + \mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right)\right|\\
&\quad + \left|\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right.\right.\\
&\qquad \left.\left.+ \mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \left|\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} + \mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right.\right.\\
&\qquad \left.\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} + \mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)'\right)\right|\\
&\quad + \left|\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\right|\\
&\quad + \left|\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right|\\
&\leq \frac{2pR}{\lambda^3_{\min}(\mathbf{V})} + \frac{pR}{\lambda^3_{\min}(\mathbf{V})} + \frac{pR}{\lambda^3_{\min}(\mathbf{V})}\\
&\leq \frac{4pR}{\lambda^3_{\min}(\mathbf{V})}. \qquad (3.50)
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right.\right.\right.\\
&\qquad \left.\left.\left. + \frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right)\right| \leq \frac{4pR}{\lambda_{\min}(\mathbf{V})}, \qquad (3.51)
\end{aligned}
$$

and

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right.\right.\right.\\
&\qquad \left.\left.\left. + \frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right)\right| \leq \frac{4pR}{\lambda_{\min}(\mathbf{V})}. \qquad (3.52)
\end{aligned}
$$

For the last term of (3.42), from Theorem 2.26(ii) and Theorem 2.39, we consider

$$
\begin{aligned}
&\left|\operatorname{tr}\left(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|\\
&= \left|\operatorname{tr}\left(\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\right|\\
&= \operatorname{tr}\left(\mathbf{C}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&\leq \lambda_{\max}(\mathbf{C})\operatorname{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\\
&= \operatorname{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)
\end{aligned}
$$

$$\leq \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)$$
$$= \lambda_{\max}(\mathbf{V}^{-1})\,\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\right)$$
$$\leq \frac{pR}{\lambda_{\min}^3(\mathbf{V})}, \tag{3.53}$$

where we use Corollary 2.33 and (3.39) to obtain the last inequality.

From Theorem 2.26(ii), Theorem 2.39 and the triangle inequality, we have

$$\Bigg|\mathrm{tr}\left(\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.$$
$$\left.\left(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)\Bigg|$$
$$= \mathrm{tr}\left(\mathbf{C}\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.$$
$$\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)$$
$$\leq \lambda_{\max}\left(\mathbf{C}\right)\mathrm{tr}\left(\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.$$
$$\left.\left(\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)$$
$$= \mathrm{tr}\left(\mathbf{V}^{-1}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.$$
$$\left.\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)$$
$$\leq \lambda_{\max}\left(\mathbf{V}^{-1}\right)\mathrm{tr}\left(\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.$$
$$\left.\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}+\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)'\right)$$
$$= \lambda_{\max}\left(\mathbf{V}^{-1}\right)\left(\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right.$$
$$+\,\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)$$
$$+\,\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)$$
$$\left.+\,\mathrm{tr}\left(\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\right)$$
$$\leq \frac{2pR}{\lambda_{\min}^3(\mathbf{V})}, \tag{3.54}$$

where we use Corollary 2.33 and (3.39) to obtain the last inequality.

From (3.53) and (3.54), we have

$$
\begin{aligned}
&\Bigg|- \operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\Bigg)\Bigg| \\
&= \Bigg|- \operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\Bigg) \\
&\quad + \operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i} \\
&\qquad + \mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\Bigg)\Bigg| \\
&\leq \Bigg|\operatorname{tr}\Bigg(\Bigg(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j} + \mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\Bigg) \\
&\qquad \Bigg(\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j} + \mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\Bigg)'\Bigg)\Bigg| \\
&\quad + \Bigg|\operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_i}\Bigg)\Bigg| \\
&\quad + \Bigg|\operatorname{tr}\Bigg(\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1/2}\mathbf{C}\mathbf{V}^{-1/2}\frac{\partial \mathbf{V}}{\partial \theta_k}\Bigg)\Bigg| \\
&\leq \frac{4pR}{\lambda_{\min}^3(\mathbf{V})}. \qquad (3.55)
\end{aligned}
$$

Thus, from (3.45), (3.46), (3.47), (3.50), (3.51), (3.52) and (3.55), we have

$$
\begin{aligned}
&\Bigg|\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) \\
&- \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\Bigg| \\
&\leq \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} + \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} + \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} + \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} + \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} + \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} + \frac{4pR}{\lambda_{\min}^3(\mathbf{V})} \\
&\leq \frac{28pR}{\lambda_{\min}^3(\mathbf{V})}.
\end{aligned}
$$

Hence, for p and R are fixed,

$$\begin{aligned}&\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\right.\\&\left.-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\right|\\&=O(1),\end{aligned}$$

for $i, j, k = 1, \ldots, R$.

For (4), by the same technique, we can show that, for p and R are fixed,

$$\begin{aligned}&\left|\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)\right.\\&\left.-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_l}\right)\right|\\&=O(1),\end{aligned}$$

for $i, j, k, l = 1, \ldots, R$. □

Corollary 3.16. Under the multivariate Fay-Herriot model (3.2), for any positive integer m and $i = 1, \ldots, R$,

$$\mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right|^m\right]\le c\left(\frac{DR-pR}{\lambda^2_{\min}(\mathbf{V})}\right)^{m/2},$$

where c is a constant.

Proof. Since $\mathbf{V}_\mathrm{u}$ and $\mathbf{V}_\mathrm{e}$ are positive definite matrices, there exist nonsingular matrices $\mathbf{L}_\mathrm{u}$ and $\mathbf{L}_\mathrm{e}$ such that $\mathbf{V}_\mathrm{u} = \mathbf{L}_\mathrm{u}\mathbf{L}'_\mathrm{u}$ and $\mathbf{V}_\mathrm{e} = \mathbf{L}_\mathrm{e}\mathbf{L}'_\mathrm{e}$, respectively. Let $\mathbf{W} = [\mathbf{L}_\mathrm{u}, \mathbf{L}_\mathrm{e}]$. Thus, $\mathbf{w} = \mathbf{u} + \mathbf{e} = \mathbf{W}\boldsymbol{\xi}$, where $\boldsymbol{\xi} \sim N(\mathbf{0}, \mathbf{I}_{2DR})$. Since $\mathbf{PX} = \mathbf{0}$, we have

$$\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}=(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})=\mathbf{w}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{w}=\boldsymbol{\xi}'\mathbf{W}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{W}\boldsymbol{\xi}. \quad (3.56)$$

Using Theorem 2.73, and the fact that $\mathbf{PVP} = \mathbf{P}$ and $\mathbf{PX} = \mathbf{0}$, we have

$$\mathrm{E}\left[\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y}\right] = \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{V}\right) = \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right). \tag{3.57}$$

For any $m \geq 2$, using (3.56), (3.57), Lemma 3.3 and the fact that $\mathbf{WW}' = \mathbf{V}$ and $\mathbf{PVP} = \mathbf{P}$, we have

$$\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^m\right] \\
&= \mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \mathrm{E}\left[\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y}\right]\right|^m\right] \\
&= \mathrm{E}\left[\left|\boldsymbol{\xi}'\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\boldsymbol{\xi} - \mathrm{E}\left[\boldsymbol{\xi}'\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\boldsymbol{\xi}\right]\right|^m\right] \\
&\leq c\left\|\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\right\|_{\mathrm{F}}^m \\
&= c\left(\mathrm{tr}\left(\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\right)\right)^{m/2} \\
&= c\left(\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right)^{m/2} \\
&\leq c\left(\frac{DR - pR}{\lambda_{\min}^2(\mathbf{V})}\right)^{m/2},
\end{aligned} \tag{3.58}$$

where we use Lemma 3.8(2) to obtain the last inequality.

For $m = 1$, using Corollary 2.87, we have

$$\begin{aligned}
\mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|\right] &\leq \left(\mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^2\right]\right)^{1/2} \\
&\leq \left(c\left(\frac{DR - pR}{\lambda_{\min}^2(\mathbf{V})}\right)^{2/2}\right)^{1/2} \\
&= c\left(\frac{DR - pR}{\lambda_{\min}^2(\mathbf{V})}\right)^{1/2},
\end{aligned} \tag{3.59}$$

where we use (3.58) with $m = 2$ to obtain (3.59). □

Corollary 3.17. Under the multivariate Fay-Herriot model (3.2), for any positive integer m, and $i, j = 1, \ldots, R$,

$$\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y} - 2\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^m\right] \leq c\left(\frac{DR - pR}{\lambda^4_{\min}(\mathbf{V})}\right)^{m/2},$$

where c is a constant.

Proof. Since $\mathbf{V}_{\mathrm{u}}$ and $\mathbf{V}_{\mathrm{e}}$ are positive definite matrices, there exist nonsingular matrices $\mathbf{L}_{\mathrm{u}}$ and $\mathbf{L}_{\mathrm{e}}$ such that $\mathbf{V}_{\mathrm{u}} = \mathbf{L}_{\mathrm{u}}\mathbf{L}'_{\mathrm{u}}$ and $\mathbf{V}_{\mathrm{e}} = \mathbf{L}_{\mathrm{e}}\mathbf{L}'_{\mathrm{e}}$, respectively. Let $\mathbf{W} = [\mathbf{L}_{\mathrm{u}}, \mathbf{L}_{\mathrm{e}}]$. Thus, $\mathbf{w} = \mathbf{u} + \mathbf{e} = \mathbf{W}\boldsymbol{\xi}$, where $\boldsymbol{\xi} \sim N(\mathbf{0}, \mathbf{I}_{2DR})$. Since $\mathbf{PX} = \mathbf{0}$, we have

$$\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} = \mathbf{w}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{w} = \boldsymbol{\xi}'\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\boldsymbol{\xi}. \tag{3.60}$$

Using Theorem 2.73, and the fact that $\mathbf{PVP} = \mathbf{P}$ and $\mathbf{PX} = \mathbf{0}$, we have

$$\mathrm{E}\left[\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y}\right] = \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{V}\right) = \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right). \tag{3.61}$$

For any $m \geq 2$, using (3.60), (3.61), Lemma 3.3 and the fact that $\mathbf{WW}' = \mathbf{V}$ and $\mathbf{PVP} = \mathbf{P}$, we have

$$\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y} - 2\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^m\right] \\
&= \mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad \left.\left. - \mathrm{E}\left[\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right]\right|^m\right] \\
&= \mathrm{E}\left[\left|\boldsymbol{\xi}'\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\boldsymbol{\xi}\right.\right. \\
&\qquad \left.\left. - \mathrm{E}\left[\boldsymbol{\xi}'\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\boldsymbol{\xi}\right]\right|^m\right] \\
&\leq c\left\|\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\right\|^m_{\mathrm{F}} \\
&= c\left(\mathrm{tr}\left(\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\right.\right. \\
&\qquad \left.\left.\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\right)\mathbf{W}\right)\right)^{m/2}
\end{aligned}$$

$$
\begin{aligned}
&= c\left(\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right. \\
&\qquad \left. + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right)^{m/2} \\
&\leq c\left(\frac{DR - pR}{\lambda_{\min}^4(\mathbf{V})}\right)^{m/2}, \qquad (3.62)
\end{aligned}
$$

where we use Lemma 3.12 to obtain the last inequality.

For $m = 1$, using Corollary 2.87, we have

$$
\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y} - 2\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|\right] \\
&\leq \left(\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y} - 2\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^2\right]\right)^{1/2} \\
&\leq \left(c\left(\frac{DR - pR}{\lambda_{\min}^4(\mathbf{V})}\right)^{2/2}\right)^{1/2} \qquad (3.63) \\
&= c\left(\frac{DR - pR}{\lambda_{\min}^4(\mathbf{V})}\right)^{1/2},
\end{aligned}
$$

where we use (3.62) with $m = 2$ to obtain (3.63). □

Corollary 3.18. Under the multivariate Fay-Herriot model (3.2), for any positive integer m, and $i, j, k = 1, \ldots, R$,

$$
\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad \left.\left. -\ 3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^m\right] \leq c\left(\frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})}\right)^{m/2},
\end{aligned}
$$

where c is a constant.

Proof. Since $\mathbf{V}_{\mathrm{u}}$ and $\mathbf{V}_{\mathrm{e}}$ are positive definite matrices, there exist nonsingular matrices $\mathbf{L}_{\mathrm{u}}$ and $\mathbf{L}_{\mathrm{e}}$ such that $\mathbf{V}_{\mathrm{u}} = \mathbf{L}_{\mathrm{u}}\mathbf{L}_{\mathrm{u}}'$ and $\mathbf{V}_{\mathrm{e}} = \mathbf{L}_{\mathrm{e}}\mathbf{L}_{\mathrm{e}}'$, respectively. Let $\mathbf{W} = [\mathbf{L}_{\mathrm{u}}, \mathbf{L}_{\mathrm{e}}]$. Thus, $\mathbf{w} = \mathbf{u} + \mathbf{e} = \mathbf{W}\boldsymbol{\xi}$, where $\boldsymbol{\xi} \sim N(\mathbf{0}, \mathbf{I}_{2DR})$. Since $\mathbf{PX} = \mathbf{0}$, we have

$$
\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} = \boldsymbol{\xi}'\mathbf{W}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{W}\boldsymbol{\xi}. \qquad (3.64)
$$

Using Theorem 2.73, and the fact that $\mathbf{PVP} = \mathbf{P}$ and $\mathbf{PX} = \mathbf{0}$, we have

$$\mathrm{E}\left[\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y}\right] = \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right). \tag{3.65}$$

For any $m \geq 2$, using (3.64), (3.65), Lemma 3.3 and the fact that $\mathbf{WW}' = \mathbf{V}$ and $\mathbf{PVP} = \mathbf{P}$, we have

$$\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right.\right.\\
&\qquad \left.\left. -\ 3\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^m\right]\\
&= \mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right.\right.\\
&\qquad \left.\left. -\ \mathrm{E}\left[\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right]\right|^m\right]\\
&= \mathrm{E}\left[\left|\boldsymbol{\xi}'\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\boldsymbol{\xi}\right.\right.\\
&\qquad \left.\left. -\ \mathrm{E}\left[\boldsymbol{\xi}'\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\boldsymbol{\xi}\right]\right|^m\right]\\
&\leq c\left\|\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{W}\right\|_{\mathrm{F}}^m\\
&= c\left(\mathrm{tr}\left(\mathbf{W}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{V}\right.\right.\\
&\qquad \left.\left.\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\right)\mathbf{W}\right)\right)^{m/2}\\
&= c\left(\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right.\\
&\qquad + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right)\\
&\qquad + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\\
&\qquad + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\right) + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\\
&\qquad \left. + \mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right)^{m/2}\\
&\leq c\left(\frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})}\right)^{m/2},
\end{aligned} \tag{3.66}$$

where we use Lemma 3.13 to obtain the last inequality.

For $m = 1$, using Corollary 2.87, we have

$$\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\quad \left.\left. -\ 3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|\right] \\
&\le \left(\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P} + \mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\right)\mathbf{y}\right.\right.\right. \\
&\qquad \left.\left.\left. -\ 3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^2\right]\right)^{1/2} \\
&\le \left(c\left(\frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})}\right)^{2/2}\right)^{1/2} \qquad (3.67) \\
&= c\left(\frac{DR - pR}{\lambda_{\min}^6(\mathbf{V})}\right)^{1/2},
\end{aligned}$$

where we use (3.66) with $m = 2$ to obtain (3.67). □

Theorem 3.19. Under the multivariate Fay-Herriot model (3.2), we have

$$\begin{aligned}
(1)\quad \hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} = &-\left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right.\right. \\
&\left.\left. -\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} \\
&\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i=1,\dots,R} + \mathbf{r}_{\mathrm{AP}}, \\
(2)\quad \hat{\boldsymbol{\theta}}^{\mathrm{AR}} - \boldsymbol{\theta} = &-\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} \\
&\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i=1,\dots,R} + \mathbf{r}_{\mathrm{AR}},
\end{aligned}$$

where $|\mathbf{r}_{\mathrm{AP}}| \le D^{-\rho}\eta$ and $|\mathbf{r}_{\mathrm{AR}}| \le D^{-\rho}\eta$ with $\mathrm{E}(\eta^m)$ bounded for any fixed $0 < \rho < 1$ and $m > 0$.

Proof. We prove this theorem by verifying the four conditions of Lemma 3.4.

First, consider (1), the adjusted profile maximum likelihood estimate $\hat{\boldsymbol{\theta}}^{\text{AP}}$. For condition 1, note that $\hat{\boldsymbol{\theta}}^{\text{AP}}$ is the solution of $\partial\ell_{\text{AP}}(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$, where

$$\ell_{\text{AP}}(\boldsymbol{\theta}) = c - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y} + \frac{1}{D}\log|\boldsymbol{\Sigma}|.$$

Then, by differentiating $\ell_{\text{AP}}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ three times, we have

$$\begin{aligned}
\frac{\partial\ell_{\text{AP}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \left[\frac{\partial\ell_{\text{AP}}(\boldsymbol{\theta})}{\partial\theta_i}\right]_{i=1,\ldots,R} \\
&= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\ldots,R}, \qquad (3.68)\\
\frac{\partial^2\ell_{\text{AP}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} &= \left[\frac{\partial^2\ell_{\text{AP}}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right]_{i,j=1,\ldots,R} \\
&= \left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}\right.\\
&\qquad \left. - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}, \qquad (3.69)\\
\frac{\partial^3\ell_{\text{AP}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^3} &= \left[\frac{\partial^3\ell_{\text{AP}}(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right]_{i,j,k=1,\ldots,R} \\
&= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right.\\
&\qquad - \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\\
&\qquad + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}\\
&\qquad + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y}\\
&\qquad + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y}\\
&\qquad + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\\
&\qquad \left. + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i,j,k=1,\ldots,R}. \qquad (3.70)
\end{aligned}$$

Since the third derivative of adjusted profile log-likelihood function is continuous in term of $\boldsymbol{\theta}$, the first conditions holds.

For conditon 3, from (3.69), Theorem 2.73 and the fact that $\mathbf{PX} = \mathbf{0}$, we have

$$\begin{aligned}\mathrm{E}\left[\frac{\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right] &= \left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right.\\ &\quad \left. -\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R}. \end{aligned} \tag{3.71}$$

If we take $g_i = \sqrt{D}$, for $i = 1, \dots, R$ and $\mathbf{G} = \operatorname{diag}_{1\le i\le R}(g_i)$, then

$$\mathbf{G}^{-1}\,\mathrm{E}\left[\frac{\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right]\mathbf{G}^{-1} = \frac{1}{D}\,\mathrm{E}\left[\frac{\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right].$$

Note that $-\,\mathrm{E}\left[\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2\right]$ is the Fisher information matrix. From the property that any Fisher information matrix is positive semidefinite, we have $\mathbf{G}^{-1}\,\mathrm{E}\left[\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2\right]\mathbf{G}^{-1}$ is negative semidefinite. Since $\mathrm{E}\left[\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2\right]$ is invertible, $\mathbf{G}^{-1}\,\mathrm{E}\left[\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2\right]\mathbf{G}^{-1}$ is invertible. Then every eigenvalues of $\mathbf{G}^{-1}\,\mathrm{E}\left[\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2\right]\mathbf{G}^{-1}$ is nonzero. Thus, the third condition holds: that is,

$$-\infty < \limsup_{D\to\infty} \lambda_{\max}\left(\mathbf{G}^{-1}\,\mathrm{E}\left[\frac{\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right]\mathbf{G}^{-1}\right) < 0.$$

For condition 4 (a), from (3.68), Lemma 3.15, Theorem 2.96 and Corollary 3.16, we have

$$\begin{aligned}&\mathrm{E}\left[\left(\frac{1}{g_i}\left|\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_i}\right|\right)^m\right]\\ &= \mathrm{E}\left[\left(\frac{1}{\sqrt{D}}\left|\frac{\partial \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \theta_i}\right|\right)^m\right]\\ &= \frac{1}{D^{m/2}}\,\mathrm{E}\left[\left|-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right|^m\right]\\ &= \frac{1}{(4D)^{m/2}}\,\mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + O(1) + O(D^{-1})\right|^m\right]\\ &\le \frac{\max(1, 2^{m-1})}{(4D)^{m/2}}\,\mathrm{E}\left[\left|\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right|^m\right] + \frac{\max(1, 2^{m-1})}{(4D)^{m/2}}O(1)\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{\max(1,2^{m-1})}{(4D)^{m/2}}\left(c\left(\frac{DR-pR}{\lambda^2_{\min}(\mathbf{V})}\right)^{m/2}\right)+O(D^{-m/2})\\
&= c\left(\frac{R}{\lambda^2_{\min}(\mathbf{V})}\right)^{m/2}\left(1-\frac{p}{D}\right)^{m/2}+O(D^{-m/2}).
\end{aligned}
$$

Thus, $\mathrm{E}\left[\left(\frac{1}{g_i}\left|\frac{\partial \ell(\boldsymbol{\theta})}{\partial\theta_i}\right|\right)^m\right]$ is bounded for any $m>0$, for $i=1,\dots,R$.

For condition 4 (b), from (3.69), we have

$$
\begin{aligned}
\mathrm{E}\left[\frac{\partial^2\ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right] &= \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\\
&\quad-\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right).
\end{aligned}
\tag{3.72}
$$

From (3.69), (3.72) and Corollary 3.17, we have

$$
\begin{aligned}
&\mathrm{E}\left[\left(\frac{1}{\sqrt{g_jg_i}}\left|\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}-\mathrm{E}\left[\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right]\right|\right)^m\right]\\
&=\mathrm{E}\left[\left(\frac{1}{\sqrt{D}}\left|\frac{\partial^2\ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}-\mathrm{E}\left[\frac{\partial^2\ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right]\right|\right)^m\right]\\
&=\frac{1}{D^{m/2}}\mathrm{E}\left[\left|-\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}-\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y}+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right|^m\right]\\
&=\frac{1}{(4D)^{m/2}}\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\right)\mathbf{y}-2\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right|^m\right]\\
&\leq\frac{1}{(4D)^{m/2}}\left(c\left(\frac{DR-pR}{\lambda^4_{\min}(\mathbf{V})}\right)^{m/2}\right)\\
&=c\left(\frac{R}{\lambda^4_{\min}(\mathbf{V})}\right)^{m/2}\left(1-\frac{p}{D}\right)^{m/2}.
\end{aligned}
$$

Thus, $\mathrm{E}\left[\left(\frac{1}{\sqrt{g_jg_i}}\left|\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}-\mathrm{E}\left[\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right]\right|\right)^m\right]$ is bounded for any $m>0$, for $i,j=1,\dots,R$.

For condition 4 (c), when $\delta=\min(\boldsymbol{\theta})/2$, let $\tilde{\boldsymbol{\theta}}$ such that

$$
-\frac{\min(\boldsymbol{\theta})}{2}+\theta_i\leq\tilde{\theta}_i\leq\frac{\min(\boldsymbol{\theta})}{2}+\theta_i,\ \text{for all } i=1,\dots,R.
$$

For $i, j, k = 1, \ldots, R$, from the fact that $\tilde{\mathbf{P}}\mathbf{X} = \mathbf{0}$, the same idea of the proof of Lemma 3.10 with the rank one positive semidefinite matrix $\mathbf{w}\mathbf{w}'$, we have

$$\begin{aligned}
&\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\mathbf{y} \\
&= \mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{w} + \mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\mathbf{w} \\
&= \operatorname{tr}\left(\mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{w} + \mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\mathbf{w}\right) \\
&\leq \frac{4\operatorname{tr}(\mathbf{w}'\mathbf{w})}{\lambda^4_{\min}(\tilde{\mathbf{V}})} \\
&= \frac{4\mathbf{w}'\mathbf{w}}{\lambda^4_{\min}(\tilde{\mathbf{V}})}
\end{aligned}$$

and

$$-\frac{2\mathbf{w}'\mathbf{w}}{\lambda^4_{\max}(\tilde{\mathbf{V}})} \leq \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\mathbf{y},$$

where $\tilde{\mathbf{P}}$, $\tilde{\mathbf{V}}$ and $\widetilde{\partial \mathbf{V}}/\partial\theta_i$ are obtained when $\boldsymbol{\theta}$ is replaced by $\tilde{\boldsymbol{\theta}}$ in $\mathbf{P}$, $\mathbf{V}$ and $\partial\mathbf{V}/\partial\theta_i$, respectively. From the last inequality, the fact that $\mathbf{w} \sim N(\mathbf{0}, \mathbf{V})$, Theorem 2.73 and Corollary 2.36, we have

$$\begin{aligned}
\operatorname{E}\left[\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\mathbf{y}\right|\right]^m &\leq c\left(\frac{4\operatorname{E}[\mathbf{w}'\mathbf{w}]}{\lambda^4_{\min}(\tilde{\mathbf{V}})}\right)^m \\
&= c\left(\frac{4\operatorname{tr}(\mathbf{V})}{\lambda^4_{\min}(\tilde{\mathbf{V}})}\right)^m \\
&\leq c\left(\frac{4\lambda_{\max}(\mathbf{V})\operatorname{tr}(\mathbf{I}_{DR})}{\lambda^4_{\min}(\tilde{\mathbf{V}})}\right)^m \\
&= c\left(\left(\frac{4DR\lambda_{\max}(\mathbf{V})}{\lambda^4_{\min}(\tilde{\mathbf{V}})}\right)\right)^m.
\end{aligned} \tag{3.73}$$

Then, from (3.70), Lemma 3.15, Theorem 2.96, Lemma 3.9 and (3.73), we have

$$
\begin{aligned}
&\mathrm{E}\left[\left(\frac{g_*}{g_k g_j g_i}\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\frac{\partial^3\ell(\tilde{\boldsymbol{\theta}})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right] \\
&= \mathrm{E}\left[\left(\frac{1}{D}\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\frac{\partial^3\ell_{\mathrm{AP}}(\tilde{\boldsymbol{\theta}})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right] \\
&= \frac{1}{D^m}\mathrm{E}\left[\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left| -\frac{1}{2}\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\right) - \frac{1}{2}\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\right)\right.\right. \\
&\qquad + \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} + \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad + \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} + \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad + \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} + \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad + \frac{1}{D}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right) \\
&\qquad \left.\left. + \frac{1}{D}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right|^m\right] \\
&= \frac{1}{(2D)^m}\mathrm{E}\left[\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left| -\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\right) - \mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\right)\right.\right. \\
&\qquad + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad \left.\left. + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} + O(1)\right|^m\right] \\
&\le \frac{\max(1, 5^{m-1})}{(2D)^m} \\
&\quad \left(\mathrm{E}\left[\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left(\left|\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\right) + \mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\right)\right|^m\right.\right.\right. \\
&\qquad + \left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\qquad + \left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\qquad \left.\left.\left. + \left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}\right|^m + O(1)\right)\right]\right)
\end{aligned}
$$

$$\le \frac{\max(1,5^{m-1})}{(2D)^m}\left(\left|\frac{4(DR-pR)}{\lambda^3_{\min}(\tilde{\mathbf{V}})}\right|^m + 6\left[c\left(\frac{4DR\lambda_{\max}(\mathbf{V})}{\lambda^4_{\min}(\tilde{\mathbf{V}})}\right)^m\right]\right) + O(D^{-m})$$
$$= c\left(\left(\frac{R}{\lambda^3_{\min}(\tilde{\mathbf{V}})}\right)\left(1-\frac{p}{D}\right)\right)^m + c\left(\left(\frac{R\lambda_{\max}(\mathbf{V})}{\lambda^4_{\min}(\tilde{\mathbf{V}})}\right)\right)^m + O(D^{-m}),$$

where $S_\delta(\boldsymbol{\theta}) = \{\tilde{\boldsymbol{\theta}} : \left|\tilde{\theta}_i - \theta_i\right| \le \min(\boldsymbol{\theta})/2, 1 \le i \le R\}$, is bounded for any $m > 0$.

Since all the four conditions of Lemma 3.4 are satisfied, from (3.71), we have

$$\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} = -\left(\mathrm{E}\left[\frac{\partial^2 l_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2}\right]\right)^{-1}\frac{\partial\ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} + \mathbf{r}_{\mathrm{AP}}$$
$$= -\left(\left[\frac{1}{2}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) - \mathrm{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right.\right.$$
$$\left.\left. - \frac{1}{D}\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1}$$
$$\left[-\frac{1}{2}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\ldots,R} + \mathbf{r}_{\mathrm{AP}}$$

where $|\mathbf{r}_{\mathrm{AP}}| \le D^{-\rho}\eta$ with $\mathrm{E}(\eta^m)$ bounded for any fixed $0 < \rho < 1$ and $m > 0$.

For (2), we consider the adjusted residual maximum likelihood estimation. Note that $\hat{\boldsymbol{\theta}}^{\mathrm{AR}}$ is the solution of $\partial\ell_{\mathrm{AR}}(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$, where

$$\ell_{\mathrm{AR}}(\boldsymbol{\theta}) = c - \frac{1}{2}\log|\mathbf{K}'\mathbf{V}\mathbf{K}| - \frac{1}{2}\mathbf{y}'\mathbf{P}\mathbf{y} + \frac{1}{D}\log|\boldsymbol{\Sigma}|.$$

Then, by differentiating with respect to $\boldsymbol{\theta}$ three times, we have

$$\frac{\partial\ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \left[\frac{\partial\ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial\theta_i}\right]_{i=1,\ldots,R}$$
$$= \left[-\frac{1}{2}\,\mathrm{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\ldots,R}, \tag{3.74}$$

$$\frac{\partial^2\ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} = \left[\frac{\partial^2\ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right]_{i,j=1,\ldots,R}$$
$$= \left[\frac{1}{2}\,\mathrm{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}\right.$$
$$\left. - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} - \frac{1}{D}\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}, \tag{3.75}$$

$$
\begin{aligned}
\frac{\partial^3 \ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^3} &= \left[\frac{\partial^3 \ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial \theta_k \partial \theta_j \partial \theta_i}\right]_{i,j,k=1,\ldots,R} \\
&= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right. \\
&\quad + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} \\
&\quad + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\mathbf{y} + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\mathbf{y} \\
&\quad + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\mathbf{y} + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\mathbf{y} \\
&\quad + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right) \\
&\quad \left. + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j,k=1,\ldots,R}.
\end{aligned}
$$

So, the first condition holds: ℓ_{AR} is three times continuously differentiable.

As for the last two conditions, comparing the derivatives of ℓ_{AP} with ℓ_{AR}, we can see that the differences are only

$$
\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_{k_1}}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_{k_2}}\cdots\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_{k_n}}\right) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_{k_1}}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_{k_2}}\cdots\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_{k_n}}\right),
$$

for $k_1, \ldots, k_n = 1, \ldots, R$ and $n = 1, 2, 3$. By Lemma 3.15, these differences are bounded when p and R are fixed. Thus, following the previous proof for ℓ_{AP}, we can show that ℓ_{AR} also satisfies the conditions of Lemma 3.4. From (3.75), Theorem 2.73 and the fact that $\mathbf{PX} = \mathbf{0}$, we have

$$
\begin{aligned}
\mathrm{E}\left[\frac{\partial^2 \ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right] &= \left[\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right. \\
&\quad \left. - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R} \\
&= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}. \qquad (3.76)
\end{aligned}
$$

Thus, from (3.74) and (3.76), we have

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{\mathrm{AR}} - \boldsymbol{\theta} &= -\left(\mathrm{E}\left[\frac{\partial^2 l_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right]\right)^{-1} \frac{\partial \ell_{\mathrm{AR}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \mathbf{r}_{\mathrm{AR}} \\
&= -\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&\quad \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i=1,\ldots,R} + \mathbf{r}_{\mathrm{AR}}
\end{aligned}$$

where $|\mathbf{r}_{\mathrm{AR}}| \leq D^{-\rho}\eta$ with $\mathrm{E}(\eta^m)$ bounded for any fixed $0 < \rho < 1$ and $m > 0$. □

Next, we will prove the asymptotic properties of the APML and AREML estimators. To prove this asymptotic properties, we need the following lemma.

Lemma 3.20. Under the multivariate Fay-Herriot model (3.2), for $i, j = 1, \ldots, R$, for large D,

$$\begin{aligned}
&\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + O(1)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \Big[\, o(D^{-1}) \,\Big]_{R\times R}.
\end{aligned}$$

Proof. Note that $-\mathrm{E}\left[\frac{\partial^2 \ell_P}{\partial \boldsymbol{\theta}^2}\right] = \left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}$ is Fisher information matrix, which is invertible. Then $\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}$ is invertible.

From Theorem 2.63, the first-order Taylor Series approximation of matrix function with $f(\mathbf{X}) = \mathbf{X}^{-1}$ is

$$\begin{aligned}
(\mathbf{A}+\mathbf{E})^{-1} &= \mathbf{A}^{-1} + \frac{d}{dt}\bigg|_{t=0} (\mathbf{A}+t\mathbf{E})^{-1} \\
&= \mathbf{A}^{-1} - (\mathbf{A}+t\mathbf{E})^{-1}(\mathbf{E})(\mathbf{A}+t\mathbf{E})^{-1}\bigg|_{t=0} \\
&= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{E}\mathbf{A}^{-1}. \qquad (3.77)
\end{aligned}$$

Applying (3.77) with $\mathbf{A} = \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}$ and $\mathbf{E} = \left[\, O(1)\,\right]_{R\times R}$,

$$\begin{aligned}
&\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) + O(1)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&\quad - \left(\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1}\left[\, O(1)\,\right]_{R\times R}\right. \\
&\qquad \left.\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1}\right) \\
&= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&\quad + \left[\, O(D^{-1})\,\right]_{R\times R}\left[\, O(1)\,\right]_{R\times R}\left[\, O(D^{-1})\,\right]_{R\times R} \\
&= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \left[\, o(D^{-1})\,\right]_{R\times R}.
\end{aligned}$$

□

Theorem 3.21. Under the multivariate Fay-Herriot model (3.2), both adjusted profile maximum likelihood estimator, $\hat{\boldsymbol{\theta}}^{\mathrm{AP}}$, and adjusted residual maximum likelihood estimator, $\hat{\boldsymbol{\theta}}^{\mathrm{AR}}$, are consistent estimators of $\boldsymbol{\theta}$. In addition, we have

$$\mathrm{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)'\right] = 2\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \left[\, o(D^{-1})\,\right]_{R\times R}.$$

Proof. From Theorem 3.19, we have

$$
\begin{aligned}
&\mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}}-\boldsymbol{\theta}\right)'\right] \\
&= \mathrm{E}\Bigg[\Bigg(-\Bigg(\Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) \\
&\qquad\qquad -\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\Bigg]_{i,j=1,\ldots,R}\Bigg)^{-1} \\
&\qquad \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\ldots,R}+\mathbf{r}_{\mathrm{AP}}\Bigg) \\
&\qquad \Bigg(\mathbf{r}'_{\mathrm{AP}}-\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]'_{i=1,\ldots,R} \\
&\qquad \Bigg(\Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right) \\
&\qquad\qquad -\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right)\Bigg]_{i,j=1,\ldots,R}\Bigg)^{-1}\Bigg)\Bigg] \\
&= \mathrm{E}\left[\mathbf{r}_{\mathrm{AP}}\mathbf{r}'_{\mathrm{AP}}\right] \\
&\quad +\Bigg(\Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\Bigg]_{i,j=1,\ldots,R}\Bigg)^{-1} \\
&\quad \mathrm{E}\Bigg[\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\ldots,R} \\
&\qquad \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]'_{i=1,\ldots,R}\Bigg] \\
&\quad \Bigg(\Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right)\Bigg]_{i,j=1,\ldots,R}\Bigg)^{-1} \\
&\quad -\Bigg(\Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\Bigg]_{i,j=1,\ldots,R}\Bigg)^{-1} \\
&\qquad \mathrm{E}\left[\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\ldots,R}\mathbf{r}'_{\mathrm{AP}}\right] \\
&\quad -\mathrm{E}\left[\mathbf{r}_{\mathrm{AP}}\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]'_{i=1,\ldots,R}\right] \\
&\qquad \Bigg(\Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\right)\Bigg]_{i,j=1,\ldots,R}\Bigg)^{-1}.
\end{aligned}
\tag{3.78}
$$

We have four terms of $\mathrm{E}[(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta})']$. For the first term of (3.78), from the condition in Theorem 3.23 with $\rho = 3/4$, we have

$$\begin{aligned} \mathrm{E}[\mathbf{r}_{\mathrm{AP}}\mathbf{r}'_{\mathrm{AP}}] &= \mathrm{E}[[r_i r_j]_{i,j=1,\dots,R}] \\ &= \Big[\, o(D)^{-1} \,\Big]_{R\times R}, \end{aligned} \tag{3.79}$$

where $\mathbf{r}_{\mathrm{AP}} = [r_i]_{i=1,\dots,R}$.

For the second term of (3.78), from Lemma 3.15(2), we have

$$\begin{aligned} &\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R} \\ &= \left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right. \\ &\qquad \left. - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R} \\ &= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + O(1) + O(D^{-1})\right]_{i,j=1,\dots,R} \\ &= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + O(1)\right]_{i,j=1,\dots,R}. \end{aligned} \tag{3.80}$$

From (3.80), we have

$$\begin{aligned} &\left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} \\ &= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + O(1)\right]_{i,j=1,\dots,R}\right)^{-1} \\ &= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} + \Big[\, o(D^{-1}) \,\Big]_{R\times R}, \end{aligned} \tag{3.81}$$

where we use Lemma 3.20 to obtain the last equation.

Using Theorem 2.73 and the fact that $\mathbf{PX}=\mathbf{0}$,

$$\mathrm{E}\left[\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}\right]=\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right). \tag{3.82}$$

Using Theorem 2.74 and the fact that $\mathbf{PX}=\mathbf{0}$,

$$\mathrm{E}\left[\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}\right)\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{Py}\right)\right]=2\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)+\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right). \tag{3.83}$$

From (3.82) and (3.83), we have

$$\begin{aligned}
&\mathrm{E}\left[\left[-\frac{1}{2}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}+\frac{1}{D}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\dots,R}\right.\\
&\qquad\left.\left[-\frac{1}{2}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}+\frac{1}{D}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\right]'_{i=1,\dots,R}\right]\\
&=\mathrm{E}\left[\left[\frac{1}{4}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)-\frac{1}{4}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{Py}\right)\right.\right.\\
&\qquad-\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_j}\right)-\frac{1}{4}\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}\right)\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&\qquad+\frac{1}{4}\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}\right)\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{Py}\right)+\frac{1}{2D}\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{Py}\right)\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_j}\right)\\
&\qquad-\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)+\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\left(\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{Py}\right)\\
&\qquad\left.\left.+\frac{1}{D^2}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_j}\right)\right]_{i,j=1,\dots,R}\right]\\
&=\left[\frac{1}{4}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)-\frac{1}{4}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\right.\\
&\qquad-\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_j}\right)-\frac{1}{4}\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&\qquad+\frac{1}{2}\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)+\frac{1}{4}\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&\qquad+\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_j}\right)-\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)\\
&\qquad\left.+\frac{1}{2D}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\right)+\frac{1}{D^2}\,\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_i}\right)\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial\theta_j}\right)\right]_{i,j=1,\dots,R}
\end{aligned}$$

$$
\begin{aligned}
&= \Bigg[\frac{1}{4}\left(\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right)\left(\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right) \\
&\quad + \frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \frac{1}{2D}\left(\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right)\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\right) \\
&\quad - \frac{1}{2D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\left(\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right) \\
&\quad + \frac{1}{D^2}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\right)\Bigg]_{i,j=1,\ldots,R} \\
&= \left[O(1) + \frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + O(D^{-1}) + O(D^{-1}) + O(D^{-2})\right]_{i,j=1,\ldots,R} \\
&= \left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + O(1)\right]_{i,j=1,\ldots,R}, \qquad (3.84)
\end{aligned}
$$

where we use Lemma 3.15 to obtain the last two equations.

Thus, from (3.81) and (3.84),

$$
\begin{aligned}
&\left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&\quad \mathrm{E}\Bigg[\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i=1,\ldots,R} \\
&\qquad \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y} + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]'_{i=1,\ldots,R}\Bigg] \\
&\quad \left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) - \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} \\
&= \left(\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \left[o(D^{-1})\right]_{R\times R}\right) \\
&\quad \left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right) + O(1)\right]_{i,j=1,\ldots,R} \\
&\quad \left(\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \left[o(D^{-1})\right]_{R\times R}\right) \\
&= 2\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \left[o(D^{-1})\right]_{R\times R}. \qquad (3.85)
\end{aligned}
$$

For the third term of (3.78), by the Cauchy-Schwartz inequality for vector, (3.79) and (3.84), we have

$$\begin{aligned}
&\mathrm{E}\left[\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right)r_j\right]\\
&\leq\left(\mathrm{E}\left[\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right)\right.\right.\\
&\qquad\left.\left.\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right)\right]\right)^{1/2}\left(\mathrm{E}\left[r_jr_j\right]\right)^{1/2}\\
&=\left(\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+O(1)\right)^{1/2}\left(o(D)^{-1}\right)^{1/2}\\
&=\left(O(D)\right)^{1/2}\left(o(D)^{-1}\right)^{1/2} &(3.86)\\
&=o(1), &(3.87)
\end{aligned}$$

where we use Remark 3.14 to obtain (3.86).

From (3.81) and (3.87) and Remark 3.14, we have

$$\begin{aligned}
&\left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)-\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1}\\
&\mathrm{E}\left[\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right]_{i=1,\dots,R}\mathbf{r}'_{\mathrm{AP}}\right]\\
&=\left(\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1}+\left[\,o(D^{-1})\,\right]_{R\times R}\right)\\
&\quad\left[\mathrm{E}\left[\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\right)\right)r_j\right]\right]_{i,j=1,\dots,R}\\
&=\left[\,o(1)\,\right]_{R\times R}\left[\,O(D^{-1})\,\right]_{R\times R}+\left[\,o(1)\,\right]_{R\times R}\left[\,o(D^{-1})\,\right]_{R\times R}\\
&=\left[\,o(D)^{-1}\,\right]_{R\times R}. &(3.88)
\end{aligned}$$

For the fourth term of (3.78), by the Cauchy-Schwartz inequality for vector, similar (3.87), we have

$$\begin{aligned}
&\mathrm{E}\left[r_i\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\right)\right)\right]\\
&\leq (\mathrm{E}\left[r_i r_i\right])^{1/2}\left(\mathrm{E}\left[\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\right)\right)\right.\right.\\
&\qquad\left.\left.\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\right)\right)\right]\right)^{1/2}\\
&=\left(o(D)^{-1}\right)^{1/2}\left(\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)+O(1)\right)^{1/2}\\
&=\left(o(D)^{-1}\right)^{1/2}(O(D))^{1/2}\\
&=o(1).
\end{aligned}\tag{3.89}$$

From (3.81) and (3.89), we have

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{r}_{\mathrm{AP}}\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\right]'_{i=1,\ldots,R}\right]\\
&\left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\right)\right]_{i,j=1,\ldots,R}\right)^{-1}\\
&=\left[\mathrm{E}\left[r_i\left(-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)+\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y}+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\right)\right)\right]\right]_{i,j=1,\ldots,R}\\
&\quad\left(\left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1}+\left[\,o(D^{-1})\,\right]_{R\times R}\right)\\
&=\left[\,o(1)\,\right]_{R\times R}\left[\,O(D^{-1})\,\right]_{R\times R}+\left[\,o(1)\,\right]_{R\times R}\left[\,o(D^{-1})\,\right]_{R\times R}\\
&=\left[\,o(D)^{-1}\,\right]_{R\times R}.
\end{aligned}\tag{3.90}$$

Thus, from (3.79), (3.85), (3.88) and (3.90),

$$\mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}}-\boldsymbol{\theta}\right)'\right]=2\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1}+\left[\,o(D^{-1})\,\right]_{R\times R}$$

Similarly, by the same technique, we can show that

$$\mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AR}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}^{\mathrm{AR}}-\boldsymbol{\theta}\right)'\right]=2\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1}+\left[o(D^{-1})\right]_{R\times R}.$$

□

Remark 3.22. From Theorem 3.21, we can notice that APML and AREML estimators have the same asymptotic variances as the PML and REML estimators.

Theorem 3.23. Under the multivariate Fay-Herriot model (3.2), we have

$$\begin{aligned}\mathrm{Bias}(\hat{\boldsymbol{\theta}}^{\mathrm{AP}})&=\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1}\\&\qquad\left[-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i=1,\dots,R}+\left[o(D^{-1})\right]_{R\times1},\\\mathrm{Bias}(\hat{\boldsymbol{\theta}}^{\mathrm{AR}})&=\left[o(D^{-1})\right]_{R\times1}.\end{aligned}$$

Proof. We prove this theorem by verifying the five conditions of Lemma 3.5. We first consider the η function. Define $\eta_r(\boldsymbol{\theta})=\theta_r$ for $r=1,\dots,R$. Then, by differentiating with respect to $\boldsymbol{\theta}$ three times, we have, for $i,j,k=1,\dots,R$,

$$\frac{\partial\eta_r(\boldsymbol{\theta})}{\partial\theta_j}=\begin{cases}1 & \text{if } j=r,\\ 0 & \text{otherwise}\end{cases},\quad\frac{\partial^2\eta_r(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}=0,\quad\frac{\partial^3\eta_r(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}=0.$$

It is clearly, $\eta_r(\boldsymbol{\theta})$, $|\partial\eta_r(\boldsymbol{\theta})/\partial\boldsymbol{\theta}|$, $||\partial^2\eta_r(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2||$, $\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\partial^3\eta_r(\boldsymbol{\theta})/\partial\theta_k\partial\theta_j\partial\theta_i\right|$, for all $i,j,k=1,\dots,R$, for $r=1,\dots,R$, are bounded. So the first condition of Lemma 3.19 holds.

Since the condition of prove of Lemma 3.19 hold for any $m>0$, the conditions of Theorem 3.5 hold with $m>8+4\lambda$. The first-third order derivatives of $\ell(\boldsymbol{\theta})$ are given respectively in (3.68) – (3.70). The fourth order derivative is given as follows

$$
\begin{aligned}
\frac{\partial^4 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^4} &= \left[\frac{\partial^4 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_l \partial\theta_k \partial\theta_j \partial\theta_i}\right]_{i,j,k,l=1,\ldots,R} \\
&= \Bigg[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\quad + \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\quad + \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\quad + \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\quad + \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\quad + \frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} \\
&\quad - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_l}\mathbf{P}\mathbf{y}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_l}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_l}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_l}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_l}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_l}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\\
&\left.-\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_l}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\right]_{i,j,k,l=1,\ldots,R}.
\end{aligned}
\tag{3.91}
$$

Since the fourth derivative of adjusted profile log-likelihood function is continuous in term of $\boldsymbol{\theta}$, the second condition holds.

Next, we verify condition 3 of Lemma 3.5. To show that the mth moments of those three terms are bounded. If we take $g_i = \sqrt{D}$ for $i = 1, \ldots, R$, for condition 3 (a), from (3.65), we have

$$
\begin{aligned}
\mathrm{E}\left[\frac{\partial^3 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right] = &-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\\
&-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\\
&+\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)+\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\\
&+\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)+\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)\\
&+\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\right)+\frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\\
&+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\\
&+\frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_k}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right).
\end{aligned}
\tag{3.92}
$$

From (3.70), (3.92), Theorem 2.96 and Corollary 3.18, we have

$$
\begin{aligned}
&\mathrm{E}\left[\left(\frac{1}{\sqrt{g_j g_i}}\left|\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i} - \mathrm{E}\left(\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right)\right|\right)^m\right] \\
&= \mathrm{E}\left[\left(\frac{1}{\sqrt{D}}\left|\frac{\partial^3 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i} - \mathrm{E}\left[\frac{\partial^3 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right]\right|\right)^m\right] \\
&= \frac{1}{D^{m/2}}\,\mathrm{E}\left[\left|\frac{1}{2}\mathbf{y}'\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad + \frac{1}{2}\mathbf{y}'\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\right)\mathbf{y} \\
&\qquad \left.\left. - \frac{3}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) - \frac{3}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\right)\right|^m\right] \\
&\le \frac{\max(1,2^{m-1})}{(4D)^{m/2}} \\
&\quad \mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad \left.\left. -3\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right|^m\right] \\
&\quad + \frac{\max(1,2^{m-1})}{(4D)^{m/2}} \\
&\qquad \mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P} + \mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_k}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad\quad \left.\left. -3\operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_\mathrm{i}}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_\mathrm{j}}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_\mathrm{k}}\right)\right|^m\right] \\
&\le \frac{2\max(1,2^{m-1})}{(4D)^{m/2}}\left(c\left(\frac{DR-pR}{\lambda_{\min}^6(\mathbf{V})}\right)^{m/2}\right) \\
&= c\left(\left(\frac{R}{\lambda_{\min}^6(\mathbf{V})}\right)\left(1-\frac{p}{D}\right)\right)^{m/2}.
\end{aligned}
$$

Thus, $\mathrm{E}\left[\left(\frac{1}{\sqrt{g_j g_i}}\left|\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i} - \mathrm{E}\left(\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right)\right|\right)^m\right]$ is bounded for any $m > 0$, for $i, j, k = 1, \ldots, R$.

For condition 3 (b), from (3.70), Lemma 3.15, Theorem 2.96, Corollary 3.18 and Lemma 3.9, we have

$$
\begin{aligned}
&\mathrm{E}\left[\left(\frac{1}{g_j g_i}\left|\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right] \\
&= \mathrm{E}\left[\left(\frac{1}{D}\left|\frac{\partial^3 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right] \\
&= \frac{1}{D^m}\,\mathrm{E}\left[\left|-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\right.\right. \\
&\qquad\qquad +\frac{1}{2}\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\right)\mathbf{y} \\
&\qquad\qquad +\frac{1}{2}\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\right)\mathbf{y} \\
&\qquad\qquad \left.\left.+\ \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial\theta_i}\right)+\frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial\theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial\theta_i}\right)\right|^m\right]. \\
&= \frac{1}{(2D)^m}\,\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad\qquad +\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\right)\mathbf{y} \\
&\qquad\qquad -3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)+2\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right) \\
&\qquad\qquad \left.\left.-3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)+2\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)+O(1)\right|^m\right]. \\
&\le \frac{\max(1,4^{m-1})}{(2D)^m} \\
&\quad \left(\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\right)\mathbf{y}\right.\right.\right. \\
&\qquad\qquad \left.\left.-3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\right|^m\right] \\
&\quad +\mathrm{E}\left[\left|\mathbf{y}'\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}+\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\right)\mathbf{y}\right.\right. \\
&\qquad\qquad \left.\left.-3\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\right|^m\right] \\
&\quad \left.+\mathrm{E}\left[\left|2\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)+2\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial\theta_i}\right)\right|^m\right]+O(1)\right) \\
&\le \frac{\max(1,2^{m-1})}{(2D)^m}\left(2c\left(\frac{DR-pR}{\lambda^6_{\min}(\mathbf{V})}\right)^{m/2}+2\left|\frac{4(DR-pR)}{\lambda^3_{\min}(\mathbf{V})}\right|^m\right)+O(D^{-m}) \\
&= c\left(\frac{R^{1/2}}{\lambda^3_{\min}(\mathbf{V})}\right)^m\left(\frac{1}{D}-\frac{p}{D^2}\right)^{m/2}+c\left(\frac{R}{\lambda^3_{\min}(\mathbf{V})}\right)^m\left(1-\frac{p}{D}\right)^m+O(D^{-m}).
\end{aligned}
$$

Thus, $\mathrm{E}\left[\left(\frac{1}{g_j g_i}\left|\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right]$ is bounded for any $m > 0$, for $i, j, k = 1, \ldots, R$.

For condition 3 (c), when $\delta = \min(\boldsymbol{\theta})/2$, let $\tilde{\boldsymbol{\theta}}$ such that

$$-\frac{\min(\boldsymbol{\theta})}{2} + \theta_i \leq \tilde{\theta}_i \leq \frac{\min(\boldsymbol{\theta})}{2} + \theta_i, \text{ for all } i = 1, \dots, R.$$

For $i, j, k, l = 1, \dots, R$, from the fact that $\tilde{\mathbf{P}}\mathbf{X} = \mathbf{0}$, the same idea of the proof of Lemma 3.11 with the rank one positive semidefinite matrix $\mathbf{w}\mathbf{w}'$, we have

$$\begin{aligned}
&\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\mathbf{y} \\
&= \mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{w} + \mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\mathbf{w} \\
&= \operatorname{tr}\left(\mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{w} + \mathbf{w}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\mathbf{w}\right) \\
&\leq \frac{4\operatorname{tr}(\mathbf{w}'\mathbf{w})}{\lambda^5_{\min}(\tilde{\mathbf{V}})} \\
&= \frac{4\mathbf{w}'\mathbf{w}}{\lambda^5_{\min}(\tilde{\mathbf{V}})},
\end{aligned}$$

and

$$-\frac{2\mathbf{w}'\mathbf{w}}{\lambda^5_{\max}(\tilde{\mathbf{V}})} \leq \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\mathbf{y},$$

where $\tilde{\mathbf{P}}$, $\tilde{\mathbf{V}}$ and $\widetilde{\partial \mathbf{V}}/\partial \theta_i$ are obtained when $\boldsymbol{\theta}$ is replaced by $\tilde{\boldsymbol{\theta}}$ in $\mathbf{P}$, $\mathbf{V}$ and $\partial \mathbf{V}/\partial \theta_i$, respectively.

From the last inequality, the fact that $\mathbf{w} \sim N(\mathbf{0}, \mathbf{V})$, Theorem 2.73 and Corollary 2.36, we have

$$\begin{aligned}
&\mathrm{E}\left[\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\mathbf{y} + \mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial \mathbf{V}}}{\partial \theta_l}\tilde{\mathbf{P}}\mathbf{y}\right|\right]^m \\
&\leq c\left(\frac{4\,\mathrm{E}\left[\mathbf{w}'\mathbf{w}\right]}{\lambda^5_{\min}(\tilde{\mathbf{V}})}\right)^m
\end{aligned}$$

$$
\begin{aligned}
&= c\left(\frac{4\,\mathrm{tr}\,(\mathbf{V})}{\lambda_{\min}^{5}(\tilde{\mathbf{V}})}\right)^{m} \\
&\leq c\left(\frac{4\lambda_{\max}(\mathbf{V})\,\mathrm{tr}(\mathbf{I}_{DR})}{\lambda_{\min}^{5}(\tilde{\mathbf{V}})}\right)^{m} \\
&= c\left(\frac{4DR\lambda_{\max}(\mathbf{V})}{\lambda_{\min}^{5}(\tilde{\mathbf{V}})}\right)^{m}. \qquad (3.93)
\end{aligned}
$$

Then, from (3.91), Lemma 3.15, Theorem 2.96, Lemma 3.10 and (3.93), we have

$$
\begin{aligned}
&\mathrm{E}\left[\left(\frac{g_*^2}{g_l g_k g_j g_i}\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\frac{\partial^4\ell(\tilde{\boldsymbol{\theta}})}{\partial\theta_l\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right] \\
&= \mathrm{E}\left[\left(\frac{1}{D}\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\frac{\partial^4\ell_{\mathrm{AP}}(\tilde{\boldsymbol{\theta}})}{\partial\theta_l\partial\theta_k\partial\theta_j\partial\theta_i}\right|\right)^m\right] \\
&= \frac{1}{D^m}\,\mathrm{E}\left[\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\frac{1}{2}\,\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\right.\right. \\
&\qquad + \frac{1}{2}\,\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) \\
&\qquad + \frac{1}{2}\,\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) \\
&\qquad + \frac{1}{2}\,\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) \\
&\qquad + \frac{1}{2}\,\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) \\
&\qquad + \frac{1}{2}\,\mathrm{tr}\left(\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{V}}^{-1}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) \\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} \\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}-\frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\widetilde{\partial\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&-\frac{1}{D}\operatorname{tr}\left(\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_l}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_k}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_j}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_k}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_l}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_j}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_k}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_j}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_l}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_l}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_j}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_k}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_j}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_l}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_k}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_i}\right)\\
&-\frac{1}{D}\operatorname{tr}\left(\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_j}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_k}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_l}\tilde{\mathbf{\Sigma}}^{-1}\frac{\partial\tilde{\mathbf{\Sigma}}}{\partial\theta_i}\right)\Bigg|^{m}\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{D^m}\,\mathrm{E}\left[\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\left|\frac{1}{2}\operatorname{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) + \frac{1}{2}\operatorname{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\right.\right.\\
&\qquad + \frac{1}{2}\operatorname{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) + \frac{1}{2}\operatorname{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\\
&\qquad + \frac{1}{2}\operatorname{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right) + \frac{1}{2}\operatorname{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y} - \frac{1}{2}\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\\
&\qquad \left.\left. + O(1)\right|^m\right]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\max(1,16^{m-1})}{(2D)^m} \\
&\Bigg(\mathrm{E}\Bigg[\sup_{\tilde{\boldsymbol{\theta}}\in S_\delta(\boldsymbol{\theta})}\Bigg(\Bigg|\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)+\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\Bigg|^m \\
&\quad+\left|\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)+\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\right|^m \\
&\quad+\left|\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)+\mathrm{tr}\left(\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\right)\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+\left|\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\mathbf{y}+\mathbf{y}'\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_l}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_k}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_i}\tilde{\mathbf{P}}\frac{\partial\widetilde{\mathbf{V}}}{\partial\theta_j}\tilde{\mathbf{P}}\mathbf{y}\right|^m \\
&\quad+|O(1)|^m\Bigg)\Bigg]\Bigg) \\
&\leq \frac{\max(1,16^{m-1})}{(2D)^m}\left(3\left|\frac{4(DR-pR)}{\lambda_{\min}^4(\tilde{\mathbf{V}})}\right|^m+12\left[c\left(\frac{4DR\lambda_{\max}(\mathbf{V})}{\lambda_{\min}^5(\tilde{\mathbf{V}})}\right)^m\right]\right)+O(D^{-m})
\end{aligned}
$$

$$\leq c\left(\left(\frac{R}{\lambda_{\min}^{4}(\tilde{\mathbf{V}})}\right)\left(1-\frac{p}{D}\right)\right)^{m}+c\left(\frac{R\lambda_{\max}(\mathbf{V})}{\lambda_{\min}^{5}(\tilde{\mathbf{V}})}\right)^{m}+O(D^{-m}),$$

where $S_\delta(\boldsymbol{\theta}) = \{\tilde{\boldsymbol{\theta}} : \left|\tilde{\theta}_i - \theta_i\right| \leq \min(\boldsymbol{\theta})/2, 1 \leq i \leq R\}$, is bounded for any $m > 0$.

For condition 4 of Lemma 3.5, we consider

$$\mathbf{Q} = \mathbf{G}^{-1}\mathbf{A}\mathbf{G}^{-1},$$

$$\mathbf{U}_0 = (u_i) = (\mathrm{E}[\lambda_i]),$$

$$\mathbf{U}_1 = (u_{il}) = (\mathrm{E}[\lambda_i\lambda_l]),$$

$$\mathbf{U}_2 = (u_{jkl}) = (\mathrm{E}[\lambda_{jk}\lambda_l]),$$

$$\mathbf{U}_3 = (u_{jklmn}) = (\mathrm{E}[\lambda_{jkm}\lambda_l\lambda_n]),$$

where $\mathbf{A} = \mathrm{E}[\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2]$, $\mathbf{G} = \mathrm{diag}(g_1, \ldots, g_R)$ with $g_i = \sqrt{D}$, for all $i = 1, \ldots, R$, $\mathbf{G}^{-1}\mathbf{a} = (\lambda_i)$, $\mathbf{G}^{-1/2}(\mathbf{F} - \mathbf{A})\mathbf{G}^{-1/2} = (\lambda_{ij})$, and $\mathbf{G}^{-1}\mathbf{H}_i\mathbf{G}^{-1} = (\lambda_{ijk})$.

Note that $\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta}) = \{\tilde{\boldsymbol{\theta}} : \left|\tilde{\theta}_i - \theta_i\right| \leq \delta, 1 \leq i \leq R\}$, then $\tilde{\boldsymbol{\theta}} \to \boldsymbol{\theta}$ as $\delta \to 0$, we have

$$\mathbf{Q}(\tilde{\boldsymbol{\theta}}) \to \mathbf{Q}(\boldsymbol{\theta}) \quad \text{and} \quad \mathbf{U}_j(\tilde{\boldsymbol{\theta}}) \to \mathbf{U}_j(\boldsymbol{\theta}), \quad \text{for } j = 1, 2, 3$$

as $\delta \to 0$.

Since matrix norm is a continuous function, we have

$$\sup_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta}_0)} ||\mathbf{Q}(\tilde{\boldsymbol{\theta}}) - \mathbf{Q}(\boldsymbol{\theta}_0)|| \to 0,$$

and

$$\sup_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta}_0)} ||\mathbf{U}_j(\tilde{\boldsymbol{\theta}}) - \mathbf{U}_j(\boldsymbol{\theta}_0)|| \to 0 \quad j = 1, 2, 3,$$

as $\delta \to 0$.

Finally, we need to verify the last condition. From (3.68) and Lemma 3.15(1),

$$\begin{aligned}
\mathrm{E}\left[\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] &= \mathrm{E}\left[\frac{\partial \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] \\
&= \left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{2}\operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right) + \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right)\right]_{i=1,\dots,R} \\
&= \left[O(1) + O(D^{-1})\right]_{R\times 1} \\
&= \left[O(1)\right]_{R\times 1}.
\end{aligned}$$

Thus, $|\,\mathrm{E}\,[\mathbf{a}]\,|$ is bounded.

Next, since $\tilde{\boldsymbol{\theta}} \to \boldsymbol{\theta}$ as $\delta \to 0$, we have $\mathrm{E}\,[\mathbf{a}]\,|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \to \mathrm{E}\,[\mathbf{a}]$ as $\delta \to 0$. Since matrix norm is a continuous function, we get

$$\sup_{\tilde{\boldsymbol{\theta}} \in S_\delta(\boldsymbol{\theta})} \left|\mathrm{E}[\mathbf{a}]|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} - \mathrm{E}[\mathbf{a}]\right| \to 0, \text{ as } \delta \to 0.$$

Remark 3.24. Note that

$$\begin{aligned}
tv_i &= \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right), \\
tv_{ij} &= \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right), \\
tv_{ijk} &= \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right).
\end{aligned}$$

Remark 3.25. Note that

$$\begin{aligned}
tp_i &= \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right), \\
tp_{ij} &= \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right), \\
tp_{ijk} &= \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\right).
\end{aligned}$$

Remark 3.26. Note that

$$
\begin{aligned}
Y_i &= \mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y},\\
Y_{ij} &= \mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y},\\
Y_{ijk} &= \mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_k}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{P}\mathbf{y}.
\end{aligned}
$$

Remark 3.27. Note that

$$
\begin{aligned}
ts_i &= \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right),\\
ts_{ij} &= \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right),\\
ts_{ijk} &= \frac{1}{D}\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_k}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_j}\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i}\right).
\end{aligned}
$$

From Remark 3.24 – Remark 3.27, we have

$$
\begin{aligned}
\mathbf{b}_i &= \frac{\partial \eta_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{e}_i, \text{ where } \mathbf{e}_i \text{ is the identity vector },\qquad (3.94)\\
\mathbf{a} &= \frac{\partial \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]_{i=1\ldots,R},\qquad (3.95)\\
\mathrm{E}[\mathbf{a}] &= \mathrm{E}\left[\frac{\partial \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] = \left[-\frac{1}{2}tv_i + \frac{1}{2}tp_i + ts_i\right]_{i=1\ldots,R},\qquad (3.96)\\
\mathbf{B}_i &= \frac{\partial^2 \eta_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} = \mathbf{0}_{R\times R},\qquad (3.97)\\
\mathbf{F} &= \frac{\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} = \left[\frac{1}{2}tv_{ij} - \frac{1}{2}Y_{ij} - \frac{1}{2}Y_{ji} - ts_{ij}\right]_{i,j=1,\ldots,R},\qquad (3.98)\\
\mathbf{A} &= \mathrm{E}\left[\frac{\partial^2 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right] = \left[\frac{1}{2}tv_{ij} - tp_{ij} - ts_{ij}\right]_{i,j=1,\ldots,R}\qquad (3.99)\\
&= \left[-\frac{1}{2}tv_{ij} + O(1)\right]_{i,j=1\ldots,R},\qquad (3.100)\\
\mathbf{H}_i &= \frac{\partial^3 \ell_{\mathrm{AP}}(\boldsymbol{\theta})}{\partial \theta_i \partial \boldsymbol{\theta}^2} = \Bigg[-\frac{1}{2}tv_{jki} - \frac{1}{2}tv_{jik} + \frac{1}{2}Y_{jki} + \frac{1}{2}Y_{jik} + \frac{1}{2}Y_{ijk} + \frac{1}{2}Y_{kji}\\
&\qquad + \frac{1}{2}Y_{kij} + \frac{1}{2}Y_{ikj} + ts_{jki} + ts_{jik}\Bigg]_{i,j=1,\ldots,R}.\qquad (3.101)
\end{aligned}
$$

Then, from (3.95), (3.99) and (3.101),

$$
\begin{aligned}
\mathbf{C} &= [\mathbf{a}'\mathbf{A}^{-1}\mathbf{H}_i]_{i=1,\dots,R} \\
&= \Bigg[\mathbf{a}'\mathbf{A}^{-1}\Bigg[-\frac{1}{2}tv_{jki} - \frac{1}{2}tv_{jik} + \frac{1}{2}Y_{jki} + \frac{1}{2}Y_{jik} + \frac{1}{2}Y_{ijk} + \frac{1}{2}Y_{kji} \\
&\qquad + \frac{1}{2}Y_{kij} + \frac{1}{2}Y_{ikj} + ts_{jki} + ts_{jik}\Bigg]_{j,k=1,\dots,R}\Bigg]_{i=1,\dots,R} \\
&= \Bigg[\left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]'_{i=1\dots,R} \\
&\quad \Bigg[\sum_{l=1}^{R}(A^{-1})_{jl}\Bigg(-\frac{1}{2}tv_{lki} - \frac{1}{2}tv_{lik} + \frac{1}{2}Y_{lki} + \frac{1}{2}Y_{lik} + \frac{1}{2}Y_{ilk} + \frac{1}{2}Y_{kli} \\
&\qquad + \frac{1}{2}Y_{kil} + \frac{1}{2}Y_{ikl} + ts_{lki} + ts_{lik}\Bigg)\Bigg]_{j,k=1,\dots,R}\Bigg]_{i=1,\dots,R} \\
&= \Bigg[\sum_{j=1}^{R}\sum_{l=1}^{R}\left(-\frac{1}{2}tv_j + \frac{1}{2}Y_j + ts_j\right)(A^{-1})_{jl} \\
&\qquad \Bigg(-\frac{1}{2}tv_{lki} - \frac{1}{2}tv_{lik} + \frac{1}{2}Y_{lki} + \frac{1}{2}Y_{lik} + \frac{1}{2}Y_{ilk} + \frac{1}{2}Y_{kli} \\
&\qquad + \frac{1}{2}Y_{kil} + \frac{1}{2}Y_{ikl} + ts_{lki} + ts_{lik}\Bigg)\Bigg]_{i,k=1,\dots,R}. \qquad (3.102)
\end{aligned}
$$

From the proof of Theorem 3.21, we have

$$
\begin{aligned}
&\left(\left[\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) - \operatorname{tr}\left(\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{P}\frac{\partial\mathbf{V}}{\partial\theta_i}\right) - \frac{1}{D}\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_j}\mathbf{\Sigma}^{-1}\frac{\partial\mathbf{\Sigma}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} \\
&= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} + \left[\,o(D^{-1})\,\right]_{R\times R}.
\end{aligned}
$$

That is,

$$
\begin{aligned}
\mathbf{A}^{-1} &= \left(\left[-\frac{1}{2}\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_j}\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\theta_i}\right)\right]_{i,j=1,\dots,R}\right)^{-1} + \left[\,o(D^{-1})\,\right]_{R\times R} \\
&= \left(\left[-\frac{1}{2}tv_{ij}\right]_{i,j=1,\dots,R}\right)^{-1} + \left[\,o(D^{-1})\,\right]_{R\times R} \qquad (3.103) \\
&= \left[\,O(D^{-1})\,\right]_{R\times R} + \left[\,o(D^{-1})\,\right]_{R\times R} \\
&= \left[\,O(D^{-1})\,\right]_{R\times R}. \qquad (3.104)
\end{aligned}
$$

Since all the five conditions of Lemma 3.5 are satisfied, we have

$$\begin{aligned}
\mathrm{E}(\hat{\theta}_i^{\mathrm{AP}}) - \theta_i &= \sum_{j=0}^{3} \mathrm{E}[\Delta_{ij}(\hat{\boldsymbol{\theta}}^{\mathrm{AP}})] + o(g_*^{-2}) \\
&= \sum_{j=0}^{3} \Delta_{ij}(\boldsymbol{\theta}) + o(D^{-1}),
\end{aligned}$$

where

$$\begin{aligned}
\Delta_{i0}(\boldsymbol{\theta}) &= -2\mathbf{b}_i'\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{a}], \\
\Delta_{i1}(\boldsymbol{\theta}) &= \mathbf{b}_i'\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{F}\mathbf{A}^{-1}\mathbf{a}], \\
\Delta_{i2}(\boldsymbol{\theta}) &= \frac{1}{2}\,\mathrm{E}[\mathbf{a}'\mathbf{A}^{-1}\mathbf{B}_i\mathbf{A}^{-1}\mathbf{a}], \\
\Delta_{i3}(\boldsymbol{\theta}) &= -\frac{1}{2}\mathbf{b}_i'\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{C}\mathbf{A}^{-1}\mathbf{a}].
\end{aligned}$$

For the first term, from (3.94) and (3.96), we have

$$\begin{aligned}
\Delta_{i0}(\boldsymbol{\theta}) &= -2\mathbf{b}_i\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{a}] \\
&= -2\mathbf{b}_i\mathbf{A}^{-1}\left[-\frac{1}{2}tv_i + \frac{1}{2}tp_i + ts_i\right]_{i=1\dots,R} \\
&= -2\mathbf{b}_i\left[\sum_{j=1}^{R}(A^{-1})_{ij}\left(-\frac{1}{2}tv_j + \frac{1}{2}tp_j + ts_j\right)\right]_{i=1,\dots,R} \\
&= \sum_{j=1}^{R}(A^{-1})_{ij}\left(tv_j - tp_j - 2ts_j\right). \qquad (3.105)
\end{aligned}$$

For the second term, from (3.95), (3.98), Theorem 2.73 and Theorem 2.74, we have

$$
\begin{aligned}
&\mathrm{E}[\mathbf{F}\mathbf{A}^{-1}\mathbf{a}] \\
&= \mathrm{E}\left[\left[\frac{1}{2}tv_{ij} - \frac{1}{2}Y_{ij} - \frac{1}{2}Y_{ji} + ts_{ij}\right]_{i,j=1\ldots,R} \mathbf{A}^{-1}\left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]_{i=1\ldots,R}\right] \\
&= \mathrm{E}\left[\left[\sum_{k=1}^{R}\left(\frac{1}{2}tv_{ik} - \frac{1}{2}Y_{ik} - \frac{1}{2}Y_{ki} + ts_{ik}\right)(A^{-1})_{kj}\right]_{i,j=1,\ldots,R}\left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]_{i=1\ldots,R}\right] \\
&= \mathrm{E}\left[\left[\sum_{j=1}^{R}\left(\sum_{k=1}^{R}\left(\frac{1}{2}tv_{ik} - \frac{1}{2}Y_{ik} - \frac{1}{2}Y_{ki} + ts_{ik}\right)(A^{-1})_{kj}\right)\left(-\frac{1}{2}tv_j + \frac{1}{2}Y_j + ts_j\right)\right]_{i=1\ldots,R}\right] \\
&= \mathrm{E}\left[\left[\sum_{j=1}^{R}\sum_{k=1}^{R}(A^{-1})_{kj}\left(-\frac{1}{4}tv_{ik}tv_j + \frac{1}{4}Y_{ik}tv_j + \frac{1}{4}Y_{ki}tv_j + -\frac{1}{2}ts_{ik}tv_j + \frac{1}{4}tv_{ik}Y_j - \frac{1}{4}Y_{ik}Y_j\right.\right.\right. \\
&\qquad\left.\left.\left. -\frac{1}{4}Y_{ki}Y_j + \frac{1}{2}ts_{ik}Y_j + \frac{1}{2}tv_{ik}ts_j - \frac{1}{2}Y_{ik}ts_j - \frac{1}{2}Y_{ki}ts_j + ts_{ik}ts_j\right)\right]_{i=1\ldots,R}\right] \\
&= \left[\sum_{j=1}^{R}\sum_{k=1}^{R}(A^{-1})_{kj}\left(-\frac{1}{4}tv_{ik}tv_j + \frac{1}{4}tp_{ik}tv_j + \frac{1}{4}tp_{ki}tv_j - \frac{1}{2}ts_{ik}tv_j + \frac{1}{4}tv_{ik}tp_j\right.\right. \\
&\qquad - \frac{1}{4}\left(2tp_{jik} + tp_{ik}tp_j\right) - \frac{1}{4}\left(2tp_{jki} + tp_{ki}tp_j\right) + \frac{1}{2}ts_{ik}tp_j \\
&\qquad\left.\left. + \frac{1}{2}tv_{ik}ts_j - \frac{1}{2}tp_{ik}ts_j - \frac{1}{2}tp_{ki}ts_j + ts_{ik}ts_j\right)\right]_{i=1\ldots,R} \\
&= \left[\sum_{j=1}^{R}\sum_{k=1}^{R}(A^{-1})_{kj}\left(-\frac{1}{4}tv_{ik}\left(tv_j - tp_j\right) + \frac{1}{2}tp_{ik}\left(tv_j - tp_j\right) - \frac{1}{2}tp_{jik} - \frac{1}{2}tp_{jki}\right.\right. \\
&\qquad\left.\left. -\frac{1}{2}ts_{ik}\left(tv_j - tp_j\right) + \frac{1}{2}\left(tv_{ik} - tp_{ik}\right)ts_j - \frac{1}{2}tv_{ik}ts_j + ts_{ik}ts_j\right)\right]_{i=1\ldots,R} \\
&= \left[\sum_{j=1}^{R}\sum_{k=1}^{R}(A^{-1})_{kj}\left(-\frac{1}{2}\left(tv_{ik} - tp_{ik}\right)\left(tv_j - tp_j\right) + \frac{1}{4}tv_{ik}\left(tv_j - tp_j\right) - \frac{1}{2}tp_{jik} - \frac{1}{2}tp_{jki}\right.\right. \\
&\qquad\left.\left. - O(D^{-1})O(1) - O(D^{-1})O(1) - \frac{1}{2}tv_{ik}ts_j + O(D^{-1})O(D^{-1})\right)\right]_{i=1\ldots,R} \\
&= \left[\sum_{j=1}^{R}\sum_{k=1}^{R}(A^{-1})_{kj}\left(\frac{1}{4}tv_{ik}\left(tv_j - tp_j - 2ts_j\right) - \frac{1}{2}tp_{jik} - \frac{1}{2}tp_{jki} + O(1)\right)\right]_{i=1\ldots,R},
\end{aligned}
\tag{3.106}
$$

where we use Lemma 3.15 to obtain the last two equations.

Thus, from (3.94), (3.103) and (3.106), we have

$$
\begin{aligned}
&\Delta_{i1}(\boldsymbol{\theta}) \\
&= \mathbf{b}_i \mathbf{A}^{-1}\,\mathrm{E}[\mathbf{F}\mathbf{A}^{-1}\mathbf{a}] \\
&= \mathbf{b}_i \mathbf{A}^{-1} \left[\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{kj} \left(\frac{1}{4} tv_{ik}\,(tv_j - tp_j - 2ts_j) - \frac{1}{2} tp_{jik} - \frac{1}{2} tp_{jki} + O(1) \right) \right]_{i=1\dots,R} \\
&= \mathbf{b}_i \left[\sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj} \left(\frac{1}{4} tv_{lk}\,(tv_j - tp_j - 2ts_j) - \frac{1}{2} tp_{jlk} - \frac{1}{2} tp_{jkl} + O(1) \right) \right]_{i=1\dots,R} \\
&= \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj} \left(\frac{1}{4} tv_{lk}\,(tv_j - tp_j - 2ts_j) - \frac{1}{2} tp_{jlk} - \frac{1}{2} tp_{jkl} + O(1) \right) \\
&= -\frac{1}{2} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj}\,(tp_{jlk} + tp_{jkl}) \\
&\quad + \frac{1}{4} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj} tv_{lk}\,(tv_j - tp_j - 2ts_j) + o(D^{-1}) \\
&= -\frac{1}{2} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj}\,(tp_{jlk} + tp_{jkl}) \\
&\quad + \frac{1}{4} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}|\mathbf{A}|^{-1} C_{jk} tv_{lk}\,(tv_j - tp_j - 2ts_j) + o(D^{-1}) && (3.107) \\
&= -\frac{1}{2} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj}\,(tp_{jlk} + tp_{jkl}) \\
&\quad - \frac{1}{2} \sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R} (A^{-1})_{il}|\mathbf{A}|^{-1} C_{jk} A_{lk}\,(tv_j - tp_j - 2ts_j) + o(D^{-1}) && (3.108) \\
&= -\frac{1}{2} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj}\,(tp_{jlk} + tp_{jkl}) \\
&\quad - \frac{1}{2} \sum_{k=1}^{R}\sum_{l=1}^{R} (A^{-1})_{il}|\mathbf{A}|^{-1} C_{lk} A_{lk}\,(tv_l - tp_l + 2ts_l) \\
&\quad - \frac{1}{2} \sum_{k=1}^{R}\sum_{j \neq l}^{R} (A^{-1})_{il}|\mathbf{A}|^{-1} C_{jk} A_{lk}\,(tv_j - tp_j - 2ts_j) + o(D^{-1}) \\
&= -\frac{1}{2} \sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R} (A^{-1})_{il}(A^{-1})_{kj}\,(tp_{jlk} + tp_{jkl}) \\
&\quad - \frac{1}{2} \sum_{l=1}^{R} (A^{-1})_{il}\,(tv_l - tp_l - 2ts_l) + o(D^{-1}) && (3.109)
\end{aligned}
$$

$$= -\frac{1}{2}\sum_{l=1}^{R}\sum_{j=1}^{R}\sum_{k=1}^{R}(A^{-1})_{il}(A^{-1})_{kj}\left(tp_{jlk} + tp_{jkl}\right) - \frac{1}{2}\Delta_{i0} + o(D^{-1}), \qquad (3.110)$$

where we use Theorem 2.16 to obtain (3.107), use (3.100) to obtain (3.108) and use Theorem 2.13 and Corollary 2.14 to obtain (3.109).

For the third term, from (3.95), (3.99) and (3.97), we have

$$\begin{aligned}
\Delta_{i2}(\boldsymbol{\theta}) &= \frac{1}{2}\,\mathrm{E}[\mathbf{a}'\mathbf{A}^{-1}\mathbf{B}_i\mathbf{A}^{-1}\mathbf{a}] \\
&= \frac{1}{2}\,\mathrm{E}\left[\left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]'_{i=1\dots,R}\mathbf{A}^{-1}\mathbf{0}_{R\times R}\mathbf{A}^{-1}\left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]_{i=1\dots,R}\right] \\
&= 0. \qquad (3.111)
\end{aligned}$$

For the fourth term, from (3.95), (3.102), Theorem 2.73, Theorem 2.74 and Theorem 2.75, we have

$$\begin{aligned}
&\mathrm{E}[\mathbf{C}\mathbf{A}^{-1}\mathbf{a}] \\
&= \mathrm{E}\Bigg[\Bigg[\sum_{j=1}^{R}\sum_{l=1}^{R}\left(-\frac{1}{2}tv_j + \frac{1}{2}Y_j + ts_j\right)(A^{-1})_{jl} \\
&\qquad \Bigg(-\frac{1}{2}tv_{lki} - \frac{1}{2}tv_{lik} + \frac{1}{2}Y_{lki} + \frac{1}{2}Y_{lik} + \frac{1}{2}Y_{ilk} + \frac{1}{2}Y_{kli} \\
&\qquad + \frac{1}{2}Y_{kil} + \frac{1}{2}Y_{ikj} + ts_{lki} + ts_{lik}\Bigg)\Bigg]_{i,k=1,\dots,R} \\
&\quad \mathbf{A}^{-1}\left[-\frac{1}{2}tv_i + \frac{1}{2}Y_i + ts_i\right]_{i=1\dots,R}\Bigg] \\
&= \mathrm{E}\Bigg[\Bigg[\sum_{j=1}^{R}\sum_{l=1}^{R}\left(-\frac{1}{2}tv_j + \frac{1}{2}Y_j + ts_j\right)(A^{-1})_{jl} \\
&\qquad \Bigg(-\frac{1}{2}tv_{lki} - \frac{1}{2}tv_{lik} + \frac{1}{2}Y_{lki} + \frac{1}{2}Y_{lik} + \frac{1}{2}Y_{ilk} + \frac{1}{2}Y_{kli} \\
&\qquad + \frac{1}{2}Y_{kil} + \frac{1}{2}Y_{ikl} + ts_{lki} + ts_{lik}\Bigg)\Bigg]_{i,k=1,\dots,R} \\
&\quad \left[\sum_{m=1}^{R}(A^{-1})_{ij}\left(-\frac{1}{2}tv_m + \frac{1}{2}Y_m + ts_m\right)\right]_{i=1\dots,R}\Bigg]
\end{aligned}$$

$$= \mathrm{E}\left[\left[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\left(-\frac{1}{2}tv_j+\frac{1}{2}Y_j+ts_j\right)\right.\right.$$

$$\left(-\frac{1}{2}tv_{lki}-\frac{1}{2}tv_{lik}+\frac{1}{2}Y_{lki}+\frac{1}{2}Y_{lik}+\frac{1}{2}Y_{ilk}+\frac{1}{2}Y_{kli}\right.$$

$$\left.+\frac{1}{2}Y_{kil}+\frac{1}{2}Y_{ikl}+ts_{lki}+ts_{lik}\right)$$

$$\left.\left.\left(-\frac{1}{2}tv_m+\frac{1}{2}Y_m+ts_m\right)\right]_{i=1,\ldots,R}\right]$$

$$= \frac{1}{8}\mathrm{E}\left[\left[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\right.\right.$$

$$(-tv_jtv_{lki}tv_m+Y_jtv_{lki}tv_m+2ts_jtv_{lki}tv_m-tv_jtv_{lik}tv_m+Y_jtv_{lik}tv_m+2ts_jtv_{lik}tv_m$$

$$+tv_jY_{lki}tv_m-Y_jY_{lki}tv_m-2ts_jY_{lki}tv_m+tv_jY_{lik}tv_m-Y_jY_{lik}tv_m-2ts_jY_{lik}tv_m$$

$$+tv_jY_{ilk}tv_m-Y_jY_{ilk}tv_m-2ts_jY_{ilk}tv_m+tv_jY_{kli}tv_m-Y_jY_{kli}tv_m-2ts_jY_{kli}tv_m$$

$$+tv_jY_{kil}tv_m-Y_jY_{kil}tv_m-2ts_jY_{kil}tv_m+tv_jY_{ikl}tv_m-Y_jY_{ikl}tv_m-2ts_jY_{ikl}tv_m$$

$$+2tv_jts_{lki}tv_m-2Y_jts_{lki}tv_m-4ts_jts_{lki}tv_m$$

$$+2tv_jts_{lik}tv_m-2Y_jts_{lik}tv_m-4ts_jts_{lik}tv_m$$

$$+tv_jtv_{lki}Y_m-Y_jtv_{lki}Y_m-2ts_jtv_{lki}Y_m+tv_jtv_{lik}Y_m-Y_jtv_{lik}Y_m-2ts_jtv_{lik}Y_m$$

$$-tv_jY_{lki}Y_m+Y_jY_{lki}Y_m+2ts_jY_{lki}Y_m-tv_jY_{lik}Y_m+Y_jY_{lik}Y_m+2ts_jY_{lik}Y_m$$

$$-tv_jY_{ilk}Y_m+Y_jY_{ilk}Y_m+2ts_jY_{ilk}Y_m-tv_jY_{kli}Y_m+Y_jY_{kli}Y_m+2ts_jY_{kli}Y_m$$

$$-tv_jY_{kil}Y_m+Y_jY_{kil}Y_m+2ts_jY_{kil}Y_m-tv_jY_{ikl}Y_m+Y_jY_{ikl}Y_m+2ts_jY_{ikl}Y_m$$

$$-2tv_jts_{lki}Y_m+2Y_jts_{lki}Y_m+4ts_jts_{lki}Y_m-2tv_jts_{lik}Y_m+2Y_jts_{lik}Y_m+4ts_jts_{lik}Y_m$$

$$+2tv_jtv_{lki}ts_m-2Y_jtv_{lki}ts_m-4ts_jtv_{lki}ts_m+2tv_jtv_{lik}ts_m-2Y_jtv_{lik}ts_m-4ts_jtv_{lik}ts_m$$

$$-2tv_jY_{lki}ts_m+2Y_jY_{lki}ts_m+4ts_jY_{lki}ts_m-2tv_jY_{lik}ts_m+2Y_jY_{lik}ts_m+4ts_jY_{lik}ts_m$$

$$-2tv_jY_{ilk}ts_m+2Y_jY_{ilk}ts_m+4ts_jY_{ilk}ts_m-2tv_jY_{kli}ts_m+2Y_jY_{kli}ts_m+4ts_jY_{kli}ts_m$$

$$-2tv_jY_{kil}ts_m+2Y_jY_{kil}ts_m+4ts_jY_{kil}ts_m-2tv_jY_{ikl}ts_m+2Y_jY_{ikl}ts_m+4ts_jY_{ikl}ts_m$$

$$-4tv_jts_{lki}ts_m+4Y_jts_{lki}ts_m+8ts_jts_{lki}ts_m$$

$$\left.\left.-4tv_jts_{lik}ts_m+4Y_jts_{lik}ts_m+8ts_jts_{lik}ts_m)\right]_{i=1,\ldots,R}\right]$$

$$= \frac{1}{8}\left[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\right.$$

$$(-tv_jtv_{lki}tv_m + tp_jtv_{lki}tv_m$$

$$+ 2ts_jtv_{lki}tv_m - tv_jtv_{lik}tv_m + tp_jtv_{lik}tv_m + 2ts_jtv_{lik}tv_m$$

$$+ tv_jtp_{lki}tv_m - (2tp_{lkij} + tp_jtp_{lki})tv_m - 2ts_jtp_{lki}tv_m$$

$$+ tv_jtp_{lik}tv_m - (2tp_{likj} + tp_jtp_{lik})tv_m - 2ts_jtp_{lik}tv_m$$

$$+ tv_jtp_{ilk}tv_m - (2tp_{ilkj} + tp_jtp_{ilk})tv_m - 2ts_jtp_{ilk}tv_m$$

$$+ tv_jtp_{kli}tv_m - (2tp_{klij} + tp_jtp_{kli})tv_m - 2ts_jtp_{kli}tv_m$$

$$+ tv_jtp_{kil}tv_m - (2tp_{kilj} + tp_jtp_{kil})tv_m - 2ts_jtp_{kil}tv_m$$

$$+ tv_jtp_{ikl}tv_m - (2tp_{iklj} + tp_jtp_{ikl})tv_m - 2ts_jtp_{ikl}tv_m$$

$$+ 2tv_jts_{lki}tv_m - 2tp_jts_{lki}tv_m - 4ts_jts_{lki}tv_m$$

$$+ 2tv_jts_{lik}tv_m - 2tp_jts_{lik}tv_m - 4ts_jts_{lik}tv_m$$

$$+ tv_jtv_{lki}tp_m - tv_{lki}(2tp_{mj} + tp_jtp_m) - 2ts_jtv_{lki}tp_m$$

$$+ tv_jtv_{lik}tp_m - tv_{lik}(2tp_{mj} + tp_jtp_m) - 2ts_jtv_{lik}tp_m$$

$$- tv_j(2tp_{mlki} + tp_{lki}tp_m) + 2ts_j(2tp_{mlki} + tp_{lki}tp_m)$$

$$+ (4tp_{mlkij} + 4tp_{lkimj} + 2tp_jtp_{mlki} + 2tp_{lki}tp_{mj} + 2tp_mtp_{lkij} + tp_jtp_{lki}tp_m)$$

$$- tv_j(2tp_{mlik} + tp_{lik}tp_m) + 2ts_j(2tp_{mlik} + tp_{lik}tp_m)$$

$$+ (4tp_{mlikj} + 4tp_{likmj} + 2tp_jtp_{mlik} + 2tp_{lik}tp_{mj} + 2tp_mtp_{likj} + tp_jtp_{lik}tp_m)$$

$$- tv_j(2tp_{milk} + tp_{ilk}tp_m) + 2ts_j(2tp_{milk} + tp_{ilk}tp_m)$$

$$+ (4tp_{milkj} + 4tp_{ilkmj} + 2tp_jtp_{milk} + 2tp_{ilk}tp_{mj} + 2tp_mtp_{ilkj} + tp_jtp_{ilk}tp_m)$$

$$- tv_j(2tp_{mkli} + tp_{kli}tp_m) + 2ts_j(2tp_{mkli} + tp_{kli}tp_m)$$

$$+ (4tp_{mklij} + 4tp_{klimj} + 2tp_jtp_{mkli} + 2tp_{kli}tp_{mj} + 2tp_mtp_{klij} + tp_jtp_{kli}tp_m)$$

$$- tv_j(2tp_{mkil} + tp_{kil}tp_m) + 2ts_j(2tp_{mkil} + tp_{kil}tp_m)$$

$$+ (4tp_{mkilj} + 4tp_{kilmj} + 2tp_jtp_{mkil} + 2tp_{kil}tp_{mj} + 2tp_mtp_{kilj} + tp_jtp_{kil}tp_m)$$

$$- tv_j(2tp_{mikl} + tp_{ikl}tp_m) + 2ts_j(2tp_{mikl} + tp_{ikl}tp_m)$$
$$+ (4tp_{miklj} + 4tp_{iklmj} + 2tp_jtp_{mikl} + 2tp_{ikl}tp_{mj} + 2tp_mtp_{iklj} + tp_jtp_{ikl}tp_m)$$
$$- 2tv_jts_{lki}tp_m + 2ts_{lki}(2tp_{mj} + tp_jtp_m) + 4ts_jts_{lki}tp_m$$
$$- 2tv_jts_{lik}tp_m + 2ts_{lik}(2tp_{mj} + tp_jtp_m) + 4ts_jts_{lik}tp_m$$
$$+ 2tv_jtv_{lki}ts_m - 2tp_jtv_{lki}ts_m - 4ts_jtv_{lki}ts_m$$
$$+ 2tv_jtv_{lik}ts_m - 2tp_jtv_{lik}ts_m - 4ts_jtv_{lik}ts_m$$
$$- 2tv_jtp_{lki}ts_m + 2(2tp_{lkij} + tp_jtp_{lki})ts_m + 4ts_jtp_{lki}ts_m$$
$$- 2tv_jtp_{lik}ts_m + 2(2tp_{likj} + tp_jtp_{lik})ts_m + 4ts_jtp_{lik}ts_m$$
$$- 2tv_jtp_{ilk}ts_m + 2(2tp_{ilkj} + tp_jtp_{ilk})ts_m + 4ts_jtp_{ilk}ts_m$$
$$- 2tv_jtp_{kli}ts_m + 2(2tp_{klij} + tp_jtp_{kli})ts_m + 4ts_jtp_{kli}ts_m$$
$$- 2tv_jtp_{kil}ts_m + 2(2tp_{kilj} + tp_jtp_{kil})ts_m + 4ts_jtp_{kil}ts_m$$
$$- 2tv_jtp_{ikl}ts_m + 2(2tp_{iklj} + tp_jtp_{ikl})ts_m + 4ts_jtp_{ikl}ts_m$$
$$- 4tv_jts_{lki}ts_m + 4tp_jts_{lki}ts_m + 8ts_jts_{lki}ts_m$$
$$-4tv_jts_{lik}ts_m + 4tp_jts_{lik}ts_m + 8ts_jts_{lik}ts_m)]_{i=1,\ldots,R}$$

$$
\begin{aligned}
&= \frac{1}{8}\left[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\right.\\
&(-tv_j tv_{lki} tv_m + tp_j tv_{lki} tv_m + 2ts_j tv_{lki} tv_m\\
&- tv_j tv_{lik} tv_m + tp_j tv_{lik} tv_m + 2ts_j tv_{lik} tv_m\\
&+ 3tv_j tp_{lki} tv_m - 3tp_j tp_{lki} tv_m - 6ts_j tp_{lki} tv_m\\
&+ 3tv_j tp_{lik} tv_m - 3tp_j tp_{lik} tv_m - 6ts_j tp_{lik} tv_m\\
&- 2(tp_{lkij} + tp_{likj} + tp_{ilkj} + tp_{klij} + tp_{kilj} + tp_{iklj}) tv_m\\
&+ 2tv_j ts_{lki} tv_m - 2tp_j ts_{lki} tv_m - 4ts_j ts_{lki} tv_m\\
&+ 2tv_j ts_{lik} tv_m - 2tp_j ts_{lik} tv_m - 4ts_j ts_{lik} tv_m\\
&+ tv_j tv_{lki} tp_m - 2tv_{lki} tp_{mj} - tv_{lki} tp_j tp_m - 2ts_j tv_{lki} tp_m\\
&+ tv_j tv_{lik} tp_m - 2tv_{lik} tp_{mj} - tv_{lik} tp_j tp_m - 2ts_j tv_{lik} tp_m\\
&- 3tv_j tp_{lki} tp_m + 6ts_j tp_{lki} tp_m + 3tp_j tp_{lki} tp_m + 6tp_{lki} tp_{mj}\\
&- 3tv_j tp_{lik} tp_m + 6ts_j tp_{lik} tp_m + 3tp_j tp_{lik} tp_m + 6tp_{lik} tp_{mj}\\
&- 2tv_j (tp_{mlki} + tp_{mlik} + tp_{milk} + tp_{mkli} + tp_{mkil} + tp_{mikl})\\
&+ 4ts_j (tp_{mlki} + tp_{mlik} + tp_{milk} + tp_{mkli} + tp_{mkil} + tp_{mikl})\\
&+ 4(tp_{mlkij} + tp_{mlikj} + tp_{milkj} + tp_{mklij} + tp_{mkilj} + tp_{miklj})\\
&+ 4(tp_{lkimj} + tp_{likmj} + tp_{ilkmj} + tp_{klimj} + tp_{kilmj} + tp_{iklmj})\\
&+ 2tp_j (tp_{mlki} + tp_{mlik} + tp_{milk} + tp_{mkli} + tp_{mkil} + tp_{mikl})\\
&+ 2tp_m (tp_{lkij} + tp_{likj} + tp_{ilkj} + tp_{klij} + tp_{kilj} + tp_{iklj})\\
&- 2tv_j ts_{lki} tp_m + 4ts_{lki} tp_{mj} + 2ts_{lki} tp_j tp_m + 4ts_j ts_{lki} tp_m\\
&- 2tv_j ts_{lik} tp_m + 4ts_{lik} tp_{mj} + 2ts_{lik} tp_j tp_m + 4ts_j ts_{lik} tp_m\\
&+ 2tv_j tv_{lki} ts_m - 2tp_j tv_{lki} ts_m - 4ts_j tv_{lki} ts_m\\
&+ 2tv_j tv_{lik} ts_m - 2tp_j tv_{lik} ts_m - 4ts_j tv_{lik} ts_m\\
&- 6tv_j tp_{lki} ts_m + 6tp_j tp_{lki} ts_m + 12ts_j tp_{lki} ts_m\\
&- 6tv_j tp_{lik} ts_m + 6tp_j tp_{lik} ts_m + 12ts_j tp_{lik} ts_m\\
&+ 4(tp_{lkij} + tp_{likj} + tp_{ilkj} + tp_{klij} + tp_{kilj} + tp_{iklj})\\
&- 4tv_j ts_{lki} ts_m + 4tp_j ts_{lki} ts_m + 8ts_j ts_{lki} ts_m\\
&\left.- 4tv_j ts_{lik} ts_m + 4tp_j ts_{lik} ts_m + 8ts_j ts_{lik} ts_m)\right]_{i=1,\ldots,R}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{8}\Bigg[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km} \\
&\quad (-(tp_j - tv_j)(tp_m - tv_m)tv_{lki} - 2(tp_m - tv_m)ts_j tv_{lki} \\
&\quad - (tp_j - tv_j)(tp_m - tv_m)tv_{lik} - 2(tp_m - tv_m)ts_j tv_{lik} \\
&\quad + 3(tp_j - tv_j)(tp_m - tv_m)tp_{lki} + 6(tp_m - tv_m)ts_j tp_{lki} \\
&\quad + 3(tp_j - tv_j)(tp_m - tv_m)tp_{lik} + 6(tp_m - tv_m)ts_j tp_{lik} \\
&\quad + 2(tp_m - tv_m)(tp_{lkij} + tp_{likj} + tp_{ilkj} + tp_{klij} + tp_{kilj} + tp_{iklj}) \\
&\quad - 2(tp_j - tv_j)ts_{lki}tv_m + 4(tp_m - tv_m)ts_j ts_{lki} \\
&\quad - 2(tp_j - tv_j)ts_{lik}tv_m + 4(tp_m - tv_m)ts_j ts_{lik} \\
&\quad + 2(tp_{lki} - tv_{lki})tp_{mj} + 4tp_{lki}tp_{mj} + 2(tp_{lik} - tv_{lik})tp_{mj} + 4tp_{lik}tp_{mj} \\
&\quad + 2(tp_j - tv_j)(tp_{mlki} + tp_{mlik} + tp_{milk} + tp_{mkli} + tp_{mkil} + tp_{mikl}) \\
&\quad + 4ts_j(tp_{mlki} + tp_{mlik} + tp_{milk} + tp_{mkli} + tp_{mkil} + tp_{mikl}) \\
&\quad + 4(tp_{mlkij} + tp_{mlikj} + tp_{milkj} + tp_{mklij} + tp_{mkilj} + tp_{miklj}) \\
&\quad + 4(tp_{lkimj} + tp_{likmj} + tp_{ilkmj} + tp_{klimj} + tp_{kilmj} + tp_{iklmj}) \\
&\quad + 2(tp_j - 2tv_j)ts_{lki}tp_m + 4ts_{lki}tp_{mj} + 2(tp_j - 2tv_j)ts_{lik}tp_m + 4ts_{lik}tp_{mj} \\
&\quad - 2(tp_j - 2tv_j)tv_{lki}ts_m - 4ts_j tv_{lki}ts_m - 2(tp_j - 2tv_j)tv_{lik}ts_m - 4ts_j tv_{lik}ts_m \\
&\quad + 6(tp_j - tv_j)tp_{lki}ts_m + 12ts_j tp_{lki}ts_m + 6(tp_j - tv_j)tp_{lik}ts_m + 12ts_j tp_{lik}ts_m \\
&\quad + 4(tp_{lkij} + tp_{likj} + tp_{ilkj} + tp_{klij} + tp_{kilj} + tp_{iklj}) \\
&\quad + 4(tp_j - tv_j)ts_{lki}ts_m + 8ts_j ts_{lki}ts_m + 4(tp_j - tv_j)ts_{lik}ts_m \\
&\quad + 8ts_j ts_{lik}ts_m)\Big]_{i=1,\ldots,R} \\
&= \frac{1}{8}\Bigg[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\left(4tp_{lki}tp_{mj} + 4tp_{lik}tp_{mj} + O(D) + O(1)\right. \\
&\qquad\qquad \left. + O(D^{-1}) + O(D^{-2}) + O(D^{-3})\right)\Bigg]_{i=1,\ldots,R} \qquad (3.112) \\
&= \frac{1}{8}\left[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\left(4tp_{lki}tp_{mj} + 4tp_{lik}tp_{mj} + O(D)\right)\right]_{i=1,\ldots,R}, \qquad (3.113)
\end{aligned}
$$

where we use Lemma 3.15 and Remark 3.14 to obtain (3.112).

Thus, from (3.94), (3.103) and (3.113), we have

$$
\begin{aligned}
\Delta_{i3}(\boldsymbol{\theta}) &= -\frac{1}{2}\mathbf{b}_i\mathbf{A}^{-1}\,\mathrm{E}[\mathbf{C}\mathbf{A}^{-1}\mathbf{a}] \\
&= -\frac{1}{16}\mathbf{b}_i\mathbf{A}^{-1}\left[\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{jl}(A^{-1})_{km}\right. \\
&\qquad\qquad \left.(4tp_{lki}tp_{mj} + 4tp_{lik}tp_{mj} + O(D))\right]_{i=1,\dots,R} \\
&= -\frac{1}{4}\mathbf{b}_i\left[\sum_{n=1}^{R}\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{in}(A^{-1})_{jl}(A^{-1})_{km}\right. \\
&\qquad\qquad \left.(tp_{lkn}tp_{mj} + tp_{lnk}tp_{mj} + O(D))\right]_{i=1,\dots,R} \\
&= -\frac{1}{4}\sum_{n=1}^{R}\sum_{k=1}^{R}\sum_{j=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{in}(A^{-1})_{jl}(A^{-1})_{km} \\
&\qquad\qquad (tp_{lkn}tp_{mj} + tp_{lnk}tp_{mj} + O(D)) \\
&= -\frac{1}{4}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}\sum_{k=1}^{R}\sum_{j=1}^{R}(A^{-1})_{in}(A^{-1})_{jl} \\
&\qquad\qquad \left((A^{-1})_{km}tv_{lkn}tv_{mj} + (A^{-1})_{km}tv_{lnk}tv_{mj}\right) + o(D^{-1}) \\
&= -\frac{1}{4}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}\sum_{k=1}^{R}\sum_{j=1}^{R}(A^{-1})_{in}(A^{-1})_{jl} \qquad (3.114) \\
&\qquad\qquad \left(|\mathbf{A}|^{-1}C_{mk}tv_{lkn}tv_{mj} + |\mathbf{A}|^{-1}C_{mk}tv_{lnk}tv_{mj}\right) + o(D^{-1}) \\
&= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{in}\sum_{k=1}^{R}\sum_{j=1}^{R}(A^{-1})_{jl} \qquad (3.115) \\
&\qquad\qquad \left(|\mathbf{A}|^{-1}C_{mk}tv_{lkn}A_{mj} + |\mathbf{A}|^{-1}C_{mk}tv_{lnk}A_{mj}\right) + o(D^{-1}) \\
&= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{in}\left(\sum_{k=1}^{R}\sum_{j=1}^{R}(A^{-1})_{jl}|\mathbf{A}|^{-1}C_{mk}tv_{lkn}A_{mj}\right. \\
&\qquad\qquad \left. + \sum_{k=1}^{R}\sum_{j=1}^{R}(A^{-1})_{jl}|\mathbf{A}|^{-1}C_{mk}tv_{lnk}A_{mj}\right) + o(D^{-1}) \\
&= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{m=1}^{R}(A^{-1})_{in} \\
&\quad \left(\left(\sum_{j=1}^{R}(A^{-1})_{jl}|\mathbf{A}|^{-1}C_{mj}tv_{ljn}A_{mj} + \sum_{k\neq j}^{R}(A^{-1})_{jl}|\mathbf{A}|^{-1}C_{mk}tv_{lkn}\mathbf{A}_{mj}\right)\right. \\
&\quad \left. + \left(\sum_{j=1}^{R}(A^{-1})_{jl}|\mathbf{A}|^{-1}C_{mj}tv_{lnj}A_{mj} + \sum_{k\neq j}^{R}(A^{-1})_{jl}|\mathbf{A}|^{-1}C_{mk}tv_{lnk}A_{mj}\right)\right) + o(D^{-1})
\end{aligned}
$$

$$= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}(A^{-1})_{in}$$
$$\Bigg(\Bigg(\sum_{j=1}^{R}(A^{-1})_{jl}tv_{ljn}|\mathbf{A}|^{-1}\sum_{m=1}^{R}C_{mj}A_{mj} + \sum_{k\neq j}^{R}(A^{-1})_{jl}tv_{lkn}|\mathbf{A}|^{-1}\sum_{m=1}^{R}C_{mk}A_{mj}\Bigg)$$
$$+\Bigg(\sum_{j=1}^{R}(A^{-1})_{jl}tv_{lnj}|\mathbf{A}|^{-1}\sum_{m=1}^{R}C_{mj}A_{mj} + \sum_{k\neq j}^{R}(A^{-1})_{jl}tv_{lnk}|\mathbf{A}|^{-1}\sum_{m=1}^{R}C_{mk}A_{mj}\Bigg)\Bigg)$$
$$+ o(D^{-1})$$
$$= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}(A^{-1})_{in}\left(\sum_{j=1}^{R}(A^{-1})_{jl}tv_{ljn} + \sum_{j=1}^{R}(A^{-1})_{jl}tv_{lnj}\right) + o(D^{-1}) \tag{3.116}$$
$$= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{j=1}^{R}(A^{-1})_{in}(A^{-1})_{lj}\left(tv_{ljn} + tv_{lnj}\right) + o(D^{-1})$$
$$= \frac{1}{2}\sum_{n=1}^{R}\sum_{l=1}^{R}\sum_{j=1}^{R}(A^{-1})_{in}(A^{-1})_{lj}\left(tp_{ljn} + tp_{lnj}\right) + o(D^{-1}) \tag{3.117}$$
$$= -\Delta_{i1}(\boldsymbol{\theta}) - \frac{1}{2}\Delta_{i0}(\boldsymbol{\theta}), \tag{3.118}$$

where we use Theorem 2.16 to obtain (3.114), use (3.100) to obtain (3.115), use Theorem 2.13 and Corollary 2.14 to obtain (3.116) and use Lemma 3.15 to obtain (3.117).

Then, from (3.105), (3.110), (3.111) and (3.118),

$$\begin{aligned}
\mathrm{E}[\hat{\theta}_i^{\mathrm{AP}}] - \theta_i &= \sum_{j=0}^{3}\Delta_{ij}(\boldsymbol{\theta}) + o(D^{-1}),\\
&= \Delta_{i0}(\boldsymbol{\theta}) + \Delta_{i1}(\boldsymbol{\theta}) + \Delta_{i2}(\boldsymbol{\theta}) + \Delta_{i3}(\boldsymbol{\theta}) + o(D^{-1})\\
&= \Delta_{i0}(\boldsymbol{\theta}) + \Delta_{i1}(\boldsymbol{\theta}) + 0 - \Delta_{i1}(\boldsymbol{\theta}) - \frac{1}{2}\Delta_{i0}(\boldsymbol{\theta}) + o(D^{-1})\\
&= \frac{1}{2}\Delta_{i0}(\boldsymbol{\theta}) + o(D^{-1})\\
&= \frac{1}{2}\sum_{j=1}^{R}(A^{-1})_{ij}\left(tv_j - tp_j - 2ts_j\right) + o(D^{-1}).
\end{aligned}$$

Thus,

$$\mathrm{Bias}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] = \mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta}$$

$$
\begin{aligned}
&= \left[\frac{1}{2} \sum_{j=1}^{R} (A^{-1})_{ij} \left(tv_j - tp_j - 2ts_j \right) + o(D^{-1}) \right]_{i=1,\ldots,R} \\
&= \frac{1}{2} \left(\left[-\frac{1}{2} tv_{ij} \right]_{i,j=1,\ldots,R} + \Big[\; O(1) \;\Big]_{R\times R} \right)^{-1} [tv_i - tp_i - 2ts_i]_{i=1,\ldots,R} + [o(D^{-1})]_{R\times 1} \\
&= \left([tv_{ij}]_{i,j=1,\ldots,R} + \Big[\; O(1) \;\Big]_{R\times R} \right)^{-1} [-tv_i + tp_i + 2ts_i]_{i=1,\ldots,R} + [o(D^{-1})]_{R\times 1} \\
&= \left(\left(\left[\operatorname{tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_j} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} + \Big[\; o(D^{-1}) \;\Big]_{R\times R} \right) \\
&\quad \left[-\operatorname{tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) + \operatorname{tr} \left(\mathbf{P} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) + \frac{2}{D} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \right) \right]_{i=1,\ldots,R} + [o(D^{-1})]_{R\times 1} \\
&= \left(\left[\operatorname{tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_j} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} \\
&\quad \left[-\operatorname{tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) + \operatorname{tr} \left(\mathbf{P} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i=1,\ldots,R} + [o(D^{-1})]_{R\times 1}.
\end{aligned}
$$

Next we consider the adjusted residual maximum likelihood estimation. As for the conditions of Lemma 3.5, comparing the derivatives of ℓ_{AP} with ℓ_{AR}, we can see that the differences are only

$$
\operatorname{tr} \left(\mathbf{P} \frac{\partial \mathbf{V}}{\partial \theta_{k_1}} \mathbf{P} \frac{\partial \mathbf{V}}{\partial \theta_{k_2}} \cdots \mathbf{P} \frac{\partial \mathbf{V}}{\partial \theta_{k_n}} \right) - \operatorname{tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_{k_1}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_{k_2}} \cdots \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_{k_n}} \right),
$$

for $k_1, \ldots, k_n = 1, \ldots, R$ and $n = 1, 2, 3, 4$. By Lemma 3.15, these differences are bounded when p and R are fixed. Thus, following the previous proof for ℓ_{AP}, we can show that ℓ_{AR} also satisfies the conditions of Lemma 3.5. Then we have

$$
\begin{aligned}
\operatorname{Bias}[\hat{\boldsymbol{\theta}}^{\mathrm{AR}}] &= \left(\left[\operatorname{tr} \left(\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_j} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} \left[\frac{2}{D} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \theta_i} \right) \right]_{i=1,\ldots,R} \\
&\quad + \Big[\; o(D^{-1}) \;\Big]_{R\times 1} \\
&= \Big[\; O(D^{-1}) \;\Big]_{R\times R} \Big[\; O(D^{-1}) \;\Big]_{R\times 1} \\
&= \Big[\; o(D^{-1}) \;\Big]_{R\times 1}.
\end{aligned}
$$

□

3.4 Monte Carlo Simulation Study

In this section, using a Monte Carlo simulation, we investigate the performances of the various variance component estimators. First, we consider the special case of covariance matrix $\mathbf{\Sigma}$ and apply the AML.LL method for multivariate Fay-Herriot model with $R = 2$ or bivariate Fay-Herriot model. Second, we use the AML method for bivariate Fay-Herriot model for the covariance matrix $\mathbf{\Sigma}$.

First, for the special case of covariance matrix $\mathbf{\Sigma}$ when $\theta_i = \theta$ for all $i = 1, \ldots, R$. We apply the AML.LL method to obtain the APML.LL and AREML.LL estimators. The settings of simulation follow González-Menteiga et al. [16] and Li and Lahiri [27]. From the multivariate Fay-Herriot model defined in (3.1),

$$\mathbf{y}_d = \mathbf{X}_d\boldsymbol{\beta} + \mathbf{u}_d + \mathbf{e}_d, \quad d = 1, \ldots, D.$$

We simulate the matrix of covariates $\mathbf{X}_d = (\mathbf{x}_{d1}, \mathbf{x}_{d2})'$, where the two covariates $\mathbf{x}_{d1}$ and $\mathbf{x}_{d2}$ are generated from a bivariate normal distribution with means $\mu_{x1} = \mu_{x2} = 10$, variances $\sigma^2_{x1} = 1$ and $\sigma^2_{x2} = 2$, and correlation $\rho_x = 0.5$. The vectors of regression coefficients are $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = (1, 1)'$. We generate the random effects $\mathbf{u}_d \sim N_2\left(\mathbf{0}_2, \mathbf{\Sigma}\right)$ where $\mathbf{\Sigma} = \text{diag}(\theta_1, \theta_2)$ with $\theta_1 = \theta_2 = \theta = 2$ and the sampling errors $\mathbf{e}_d \sim N_2\left(\mathbf{0}_2, \mathbf{V}_{ed}\right)$. To study different situations of sampling errors, we let $\mathbf{V}_{ed} = (v_{dij})_{ij=1,2}$, where $v_{dij} = r_{ij}\sqrt{w_d}$ and w_d are the heteroscedasticity weights. We assume $r_{11} = 1$, $r_{22} = 2$ and $r_{12} = r_{21} = \rho_e\sqrt{r_{11}r_{22}}$ with $\rho_e = 0.5$. We consider five scenarios based on heteroscedasticity and relation between model variance and sampling variance.

Scenario (a): $w_d = 1$ representing homoscedastic model when sampling variances are smaller than model variance [16].

Scenario (b): $w_d = 4$ representing homoscedastic model when sampling variances are the same as model variance.

Scenario (c): $w_d = \max_{1 \le i \le d}\left(\sqrt{x^2_{i1} + x^2_{i2}}\right)$ representing homoscedastic model when sampling variances are larger than model variance.

Scenario (d): $w_d = \sqrt{x_{d1}^2 + x_{d2}^2}$ representing heteroscedastic model when sampling variances vary according to regressors [16].

Scenario (e): $\mathbf{V}_{ed} = \mathbf{L}_d\mathbf{L}_d'$, where $\mathbf{L}_d = (L_{dij})_{i,j=1,2}$ with $L_{d11} = L_{d22} = \sqrt{\ell_n}$, $L_{d12} = 0$ and $L_{d21} = 0.5\sqrt{\ell_n}$. There are five groups, specifically, $\ell_n = 8.0$ if n in first group; $\ell_n = 4.0$ if n in second group; $\ell_n = 2.0$ if n in third group; $\ell_n = 1.0$ if n in fourth group; $\ell_n = 0.5$ if n in fifth group. This case represents heteroscedastic model with different relations between sampling variances and model variance [27].

The different estimators are compared using absolute bias and mean squared error. The steps of the simulation are as follows.

1. For each case of sampling covariave matrix, repeat $K = 10000$ times. That is, for $k = 1, \dots, K$,
   (a) Generate $\{e_{dr}^{(k)}, u_{dr}^{(k)}, y_{dr}^{(k)}, \mathbf{x}_{dr}\}$, $d = 1, \dots, D$, $r = 1, 2$. We take $D = 15, 30$;
   (b) Calculate the variance estimator, $\hat{\theta}^{(k)}$ and EBLUP, $\hat{\boldsymbol{\mu}}^{(k)}$ based on PML, APML.LL, REML and AREML.LL methods.
2. Calculate the absolute bias and mean squared error of $\hat{\theta}$,

$$\text{AbBias} = \frac{1}{K}\sum_{k=1}^{K}|\hat{\theta}^{(k)} - \theta|, \quad \text{MSE} = \frac{1}{K}\sum_{k=1}^{K}(\hat{\theta}^{(k)} - \theta)^2.$$

3. Calculate the average of absolute relative error ($\overline{\text{ARE}}$) and average of mean squared error ($\overline{\text{MSE}}$) of $\hat{\mu}_{dr}$,

$$\overline{\text{ARE}}_r = \frac{1}{D}\sum_{d=1}^{D}\frac{1}{K}\sum_{k=1}^{K}\left|\frac{\hat{\mu}_{dr}^{(k)} - \mu_{dr}^{(k)}}{\mu_{dr}^{(k)}}\right|, \quad \overline{\text{MSE}}_r = \frac{1}{D}\sum_{d=1}^{D}\frac{1}{K}\sum_{k=1}^{K}\left(\hat{\mu}_{dr}^{(k)} - \mu_{dr}^{(k)}\right)^2,$$

for $r = 1, 2$.

Note that, in the Tables 3.1 - 3.3, the names of APML.LL and AREML.LL methods are denoted by APML and AREML, respectively.

Table 3.1: The percentages of zero estimates of θ for different estimation methods

Scenario	D	PML	APML	REML	AREML
(a)	15	0.37	0	0.14	0
(b)	15	3.88	0	1.83	0
(c)	15	18.06	0	9.14	0
(d)	15	15.10	0	7.64	0
(e)	15	2.49	0	0.60	0
(a)	30	0.01	0	0.01	0
(b)	30	0.12	0	0.06	0
(c)	30	4.49	0	2.43	0
(d)	30	3.28	0	1.67	0
(e)	30	0.08	0	0	0

Table 3.2: The absolute biases and mean squared errors of different estimators of θ

Scenario	D	PML	APML	REML	AREML
Absolute Bias					
(a)	15	0.7168	0.6485	0.7092	0.8098
(b)	15	0.9125	0.7912	0.9214	1.0865
(c)	15	1.2274	1.0660	1.2992	1.6902
(d)	15	1.1709	1.0066	1.2267	1.5576
(e)	15	0.8486	0.7482	0.8267	0.9829
(a)	30	0.4865	0.4613	0.4799	0.5079
(b)	30	0.6305	0.5828	0.6274	0.6699
(c)	30	0.8990	0.7952	0.9080	0.9876
(d)	30	0.8513	0.7567	0.8571	0.9268
(e)	30	0.5737	0.5352	0.5580	0.5992
Mean Squared Error					
(a)	15	0.7502	0.6658	0.8125	1.1374
(b)	15	1.2036	1.0408	1.3738	2.1109
(c)	15	2.0997	2.1575	2.7311	5.2049
(d)	15	1.9240	1.8773	2.4301	4.4271
(e)	15	1.0330	0.8946	1.1116	1.7174
(a)	30	0.3567	0.3325	0.3633	0.4214
(b)	30	0.5979	0.5378	0.6205	0.7435
(c)	30	1.1902	1.0555	1.2894	1.6751
(d)	30	1.0733	0.9435	1.1509	1.4662
(e)	30	0.4898	0.4458	0.4899	0.5942

Table 3.3: The averages of absolute relative errors and mean squared errors for EBLUPs based on different methods

Parameter		θ_1				θ_2			
Scenario	D	PML	APML	REML	AREML	PML	APML	REML	AREML
Average of Absolute Relative Errors									
(a)	15	0.0355	0.0351	0.0352	0.0351	0.0454	0.0450	0.0452	0.0451
(b)	15	0.0455	0.0444	0.0449	0.0447	0.0557	0.0550	0.0554	0.0554
(c)	15	0.0568	0.0554	0.0563	0.0564	0.0676	0.0669	0.0675	0.0681
(d)	15	0.0550	0.0535	0.0544	0.0544	0.0656	0.0649	0.0655	0.0659
(e)	15	0.0444	0.0436	0.0438	0.0438	0.0472	0.0465	0.0467	0.0467
(a)	30	0.0335	0.0334	0.0334	0.0334	0.0426	0.0425	0.0426	0.0425
(b)	30	0.0420	0.0417	0.0419	0.0417	0.0509	0.0507	0.0508	0.0508
(c)	30	0.0514	0.0507	0.0512	0.0509	0.0596	0.0593	0.0596	0.0596
(d)	30	0.0500	0.0493	0.0498	0.0496	0.0583	0.0580	0.0582	0.0582
(e)	30	0.0413	0.0411	0.0411	0.0411	0.0438	0.0436	0.0436	0.0436
Average of Mean Squared Errors									
(a)	15	0.7009	0.6844	0.6888	0.6849	1.1485	1.1282	1.1353	1.1346
(b)	15	1.1502	1.0999	1.1237	1.1134	1.7273	1.6859	1.7112	1.7162
(c)	15	1.7988	1.7158	1.7722	1.7807	2.5624	2.5182	2.5618	2.6089
(d)	15	1.6898	1.6075	1.6600	1.6602	2.4187	2.3709	2.4134	2.4486
(e)	15	1.2095	1.1726	1.1847	1.1850	1.3513	1.3149	1.3269	1.3297
(a)	30	0.6486	0.6450	0.6457	0.6447	1.0495	1.0446	1.0459	1.0454
(b)	30	1.0207	1.0062	1.0127	1.0082	1.4964	1.4841	1.4910	1.4903
(c)	30	1.5285	1.4842	1.5135	1.5014	2.0582	2.0333	2.0535	2.0562
(d)	30	1.4505	1.4108	1.4360	1.4243	1.9717	1.9481	1.9664	1.9676
(e)	30	1.0586	1.0507	1.0525	1.0522	1.1743	1.1661	1.1679	1.1678

From Tables 3.1 – 3.3, we conclude the following:

1. Table 3.1 reports the percentages of zero estimates for the four variance component estimation methods. We can see that the percentages of zero estimates decrease when sample size increases. For example, in scenario (a), the percentages of zero estimates of PML method in cases $D = 15$ and $D = 30$ are 0.37 and 0.01, respectively. In scenarios (a), (b) and (c), we can see that the percentages of zero estimates of PML and REML increase when sampling variances increase. For heteroscedastic models in scenarios (d) and (e), we can see that percentages of zero estimates are high for small sample sizes $D = 15$. For all scenarios, the APML.LL

and AREML.LL methods can prevent the zero estimate of θ, because their positiveness is guaranteed in theory [27].

2. Table 3.2 reports the absolute bias and mean squared error of four estimators of θ. We can see that absolute bias and mean squared error decrease when sample size increases, or equivalently when sampling variance decreases. For example, in scenario (a), the absolute biases of PML estimator in cases $D = 15$ and $D = 30$ are consecutively 0.7168 and 0.4865. Another example, in case $D = 15$, the absolute biases of PML estimator in scenarios (a) and (c) are 0.7168 and 1.2274, respectively. The absolute bias and mean squared error of the APML.LL method are less than that PML method for all scenarios and all sample sizes ($D = 15, 30$), except the mean squared error the case when $D = 15$ in scenarios (c). However, the absolute bias and mean squared error of the AREML.LL method are greater than that REML method for all scenarios and all sample sizes ($D = 15, 30$). For example, in scenario (a) and case $D = 15$, the absolute biases of APML.LL and PML estimators are 0.6485 and 0.7168, respectively.

3. Table 3.3 reports the average of absolute relative errors and mean squared errors of four methods of EBLUPs $\hat{\boldsymbol{\mu}}$. We can see that the average of absolute relative errors and mean squared errors decrease when sample size increases, or equivalently when sampling variance decreases. For example, for the first parameter θ_1, in scenario (a), the averages of absolute relative errors of PML estimator in cases $D = 15$ and $D = 30$ are consecutively 0.0355 and 0.0335. Another example, for the first parameter θ_1, in case $D = 15$, the averages of absolute relative errors of PML estimator in scenarios (a) and (c) are 0.0355 and 0.0568, respectively. The average of absolute relative errors and mean squared errors of the APML.LL method are less than or equal to PML method for all parameters, all scenarios and all sample sizes ($D = 15, 30$). The average of absolute relative errors of the AREML.LL method are less than or equal to REML method for all parameters, all scenarios and all sample sizes ($D = 15, 30$), except the case $D = 15$ in scenario (c) for first parameter and scenarios (c) and (d) for second parameter. The average of mean squared errors of the AREML.LL method are less than or equal to REML method

for all parameters, all scenarios and all sample sizes ($D = 15, 30$), except the case $D = 15$ in scenario (c) and (d) for first parameter and scenarios (c), (d) and (e) for second parameter, and the case $D = 30$ in scenarios (c) and (d) for second parameter.

Second, we use the AML method for bivariate Fay-Herriot model with covariance matrix $\mathbf{\Sigma}$ defined in (3.3). The settings of simulation follow Li and Lahiri [27]. From the multivariate Fay-Herriot model defined in (3.1),

$$\mathbf{y}_d = \mathbf{X}_d\boldsymbol{\beta} + \mathbf{u}_d + \mathbf{e}_d, \quad d = 1, \ldots, D.$$

The setting of simulation is same the previous simulation. We simulate the random effects $\mathbf{u}_d \sim N_2(\mathbf{0}_2, \mathbf{\Sigma})$, where $\mathbf{\Sigma} = \text{diag}\{\theta_1, \theta_2\}$ with $\theta_1 = 2$ and $\theta_2 = 4$ and the sampling errors $\mathbf{e}_d \sim N_2(\mathbf{0}_2, \mathbf{V}_{\text{ed}})$. Let $\mathbf{V}_{ed} = \mathbf{L}_d\mathbf{L}_d^{'}$, where $\mathbf{L}_d = (L_{dij})_{i,j=1,2}$ with $L_{d11} = L_{d22} = \sqrt{\ell_n}$, $L_{d12} = 0$, $L_{d21} = 0.5\sqrt{\ell_n}$.

We consider the unbalanced case that adjust to the sampling variance pattern of [12] for $D = 15, 30$. For $D = 15$, we consider three groups of small areas say $G = (G_1, G_2, G_3)$, specifically, $\ell_n = G_1$ if n in first group; $\ell_n = G_2$ if n in second group; $\ell_n = G_3$ if n in third group. Each groups are five small area, such that sampling variances $\mathbf{V}_{ed}$ are the same within a given group. For $D = 30$, we simply add five more small areas in each group. We consider the following four scenarios of the sampling variances:

Scenario (a): $G = (0.4, 0.6, 0.8)$, representing the case where all small area sampling variances less than the model variance.

Scenario (b): $G = (6, 0.6, 0.8)$, representing the case where sampling variances of one out of three small area is greater than the model variance.

Scenario (c): $G = (6, 6, 0.8)$, representing the case where sampling variances of two out of three small areas are greater than the model variance.

Scenario (d): $G = (6, 6, 6)$, representing the case where sampling variances of all small areas are greater than the model variance.

Next, we compare following estimators of $\boldsymbol{\theta}$: (1) profile maximum likelihood (PML) estimator, (2) adjusted profile maximum likelihood (APML) estimator, (3) residual maximum likelihood (REML) estimator and (4) adjusted residual maximum likelihood (AREML) estimator.

The steps of simulation are as follows.

1. For each case of sampling covariave matrix, repeat $K = 10000$ times. That is, for $k = 1, \ldots, K$,
   (a) Generate $\{(e_{dr}^{(k)}, u_{dr}^{(k)}, y_{dr}^{(k)}, \mathbf{x}_{dr}) : d = 1, \ldots, D, r = 1, 2\}$;
   (b) Calculate $\{\theta_r^{(k)} : r = 1, 2\}$ and $\{\text{mse}_{dr}^{(k)} : d = 1, \ldots, D, r = 1, \ldots, R\}$;
   (c) Calculate EBLUP, $\hat{\boldsymbol{\mu}}_d^{(k)}$ based on PML, APML.LL, REML and AREML.LL methods.
2. Calculate the absolute bias and mean squared error of $\hat{\theta}_r$,

$$\text{AbBias}(\hat{\theta}_r) = \frac{1}{K}\sum_{k=1}^{K} |\hat{\theta_r}^{(k)} - \theta_r|, \quad \text{MSE}(\hat{\theta}_r) = \frac{1}{K}\sum_{k=1}^{K} (\hat{\theta_r}^{(k)} - \theta_r)^2$$

   for $r = 1, 2$.

3. Calculate the average of absolute relative error ($\overline{\text{ARE}}$) and average of mean squared error ($\overline{\text{MSE}}$) of $\hat{\mu}_{dr}$,

$$\overline{\text{ARE}}_r = \frac{1}{D}\sum_{d=1}^{D} \text{ARE}_{dr}, \text{ where } \text{ARE}_{dr} = \frac{1}{K}\sum_{k=1}^{K} \left| \frac{\hat{\mu}_{dr}^{(k)} - \mu_{dr}^{(k)}}{\mu_{dr}^{(k)}} \right|,$$

$$\overline{\text{MSE}}_r = \frac{1}{D}\sum_{d=1}^{D} \text{MSE}_{dr}, \text{ where } \text{MSE}_{dr} = \frac{1}{K}\sum_{k=1}^{K} (\hat{\mu}_{dr}^{(k)} - \mu_{dr}^{(k)})^2,$$

   for $r = 1, 2$.

We present the percentages of the situations where either zero estimates of θ_1 or θ_2 and the absolute biases and mean squared errors of different estimators of θ_1 and θ_2 for the four variance component estimation methods in Tables 3.4 and 3.5, respectively. The averages of absolute relative errors and mean squared errors for EBLUPs based on different methods are presented in Table 3.6.

Table 3.4: The percentages of the situations where either zero estimates of θ_1 or θ_2 for different estimation methods

Scenario	D	PML	APML	REML	AREML
(a)	15	3.80	0	2.05	0
(b)	15	18.92	0	11.75	0
(c)	15	23.33	0	12.79	0
(d)	15	40.46	0	28.06	0
(a)	30	3.65	0	1.73	0
(b)	30	7.43	0	4.86	0
(c)	30	15.41	0	9.78	0
(d)	30	24.12	0	17.29	0

Table 3.5: The absolute biases and mean squared errors of different estimators for θ_1 and θ_2

Parameter		θ_1				θ_2			
Scenario	D	PML	APML	REML	AREML	PML	APML	REML	AREML
Absolute Bias									
(a)	15	0.7974	0.7636	0.8068	0.7935	1.4870	1.4271	1.4885	1.4859
(b)	15	1.1078	0.9504	1.0872	0.9946	1.9235	1.7027	1.9133	1.7873
(c)	15	1.2952	1.2271	1.3750	1.3183	2.2424	2.1931	2.3267	2.2982
(d)	15	1.7501	1.5581	1.9869	1.8108	2.8987	2.6685	3.1694	2.9497
(a)	30	0.5744	0.5397	0.5628	0.5477	1.0279	0.9854	1.0098	1.0005
(b)	30	0.7457	0.6747	0.7335	0.6889	1.3053	1.1843	1.2754	1.2020
(c)	30	1.0252	0.9047	1.0052	0.9246	1.6969	1.5231	1.6686	1.5483
(d)	30	1.3985	1.3262	1.4743	1.4046	2.1332	2.0606	2.2139	2.1413
Mean Squared Error									
(a)	15	0.9534	0.8637	1.0543	1.0255	3.3762	3.0007	3.6049	3.5883
(b)	15	1.8936	1.3235	1.9586	1.6467	6.1279	4.3048	6.5104	5.2617
(c)	15	2.3665	2.1821	3.0699	2.9884	7.2697	6.9354	8.8123	8.6786
(d)	15	4.2731	3.6783	6.3219	5.9592	11.8134	10.2517	15.6958	14.4103
(a)	30	0.5454	0.4441	0.5247	0.4799	1.6397	1.4700	1.6256	1.5834
(b)	30	0.8895	0.6883	0.8849	0.7592	2.8407	2.1173	2.8149	2.3053
(c)	30	1.6423	1.2415	1.6562	1.4050	4.8865	3.5073	4.9277	3.8307
(d)	30	2.8161	2.5928	3.3523	3.1559	6.7940	6.3066	7.6489	7.1764

Table 3.6: The averages of absolute relative errors and mean squared errors for EBLUPs based on different methods

Parameter		θ_1				θ_2			
Scenario	D	PML	APML	REML	AREML	PML	APML	REML	AREML
Average of Absolute Relative Errors									
(a)	15	0.0287	0.0281	0.0282	0.0278	0.0326	0.0324	0.0323	0.0322
(b)	15	0.0410	0.0392	0.0399	0.0388	0.0482	0.0475	0.0475	0.0471
(c)	15	0.0502	0.0495	0.0497	0.0492	0.0625	0.0619	0.0617	0.0611
(d)	15	0.0609	0.0600	0.0613	0.0605	0.0766	0.0755	0.0764	0.0753
(a)	30	0.0283	0.0276	0.0279	0.0275	0.0323	0.0321	0.0321	0.0321
(b)	30	0.0385	0.0375	0.0380	0.0374	0.0463	0.0460	0.0461	0.0459
(c)	30	0.0478	0.0468	0.0472	0.0466	0.0593	0.0589	0.0590	0.0586
(d)	30	0.0569	0.0565	0.0569	0.0565	0.0722	0.0718	0.0720	0.0716
Average of Mean Squared Errors									
(a)	15	0.5220	0.4940	0.5002	0.4855	0.6599	0.6522	0.6477	0.6436
(b)	15	1.1447	1.0470	1.0917	1.0335	1.6114	1.5688	1.5717	1.5459
(c)	15	1.6757	1.6366	1.6500	1.6181	2.6119	2.5615	2.5498	2.5062
(d)	15	2.3262	2.2574	2.3544	2.2963	3.6067	3.4976	3.5942	3.4898
(a)	30	0.4975	0.4657	0.4775	0.4636	0.6287	0.6216	0.6229	0.6198
(b)	30	0.9616	0.9100	0.9371	0.9059	1.4149	1.3968	1.4034	1.3922
(c)	30	1.4466	1.3838	1.4141	1.3740	2.2306	2.1957	2.2044	2.1798
(d)	30	1.9531	1.9248	1.9531	1.9274	3.0888	3.0532	3.0722	3.0365

From Tables 3.4 – 3.6, we conclude the following:

1. From Table 3.4, we can see that the percentages of zero estimates of PML and REML decrease when sample size increases, or equivalently when sampling variance decreases. For example, in scenario (a), the percentages of zero estimates of PML method in case $D = 15$ and $D = 30$ are consecutively 3.80 and 3.65. Another example, in case $D = 15$, the percentages of zero estimates of PML method in scenarios (a) and (d) are 3.42 and 40.46, respectively. For all cases, the APML and AREML methods can prevent the zero estimates of θ_1 and θ_2, The simulation results agree to Theorem 3.1.

2. From Table 3.5, we can see that absolute bias and mean squared error decrease when sample size increases, or equivalently when sampling variance decreases. For

example, for the first parameter θ_1, in scenario (a), the absolute biases of PML estimator in cases $D = 15$ and $D = 30$ are consecutively 0.7974 and 0.5744. Another example, for the first parameter θ_1, in case $D = 15$, the absolute biases of PML estimator in scenarios (a) and (d), are 0.7974 and 1.7501, respectively. The absolute bias and mean squared error of the APML and AREML methods are less than that PML and REML methods for all parameters, all scenarios and all sample sizes ($D = 15, 30$). For example, for the first parameter θ_1, in scenario (a) and case $D = 15$, the absolute biases of APML and PML estimators are 0.7636 and 0.7974, respectively.

3. From Table 3.6, we can see that the average of absolute relative errors and mean squared errors decrease when sample size increases, or equivalently when sampling variance decreases. For example, for the first parameter θ_1, in scenario (a), the averages of absolute relative errors of PML estimator in cases $D = 15$ and $D = 30$ are consecutively 0.0287 and 0.0281. Another example, for the first parameter θ_1, in case $D = 15$, the averages of absolute relative errors of PML estimators in scenarios (a) and (d) are 0.0287 and 0.0609, respectively. The average of absolute relative errors and mean squared errors of the APML and AREML methods method are less than or equal to the PML and REML methods for all parameters, all scenarios and all sample sizes ($D = 15, 30$). For example, for the first parameter θ_1, in scenario (a) and case $D = 15$, the averages of absolute relative errors of APML and PML estimators are 0.0281 and 0.0287, respectively.

CHAPTER IV

UNCERTAINTY OF EBLUP AND PREDICTION INTERVAL

In this chapter, we will discuss the mean squared error approximation of the EBLUP and mean squared error estimator of the EBLUP in Section 4.1. Then we review the prediction interval in Section 4.2.

4.1 Uncertainty of EBLUP

In this section, under the regularity conditions defined in Section 3.1, we can measure the uncertainty of the EBLUP by its mean squared error (MSE), defined as $\mathrm{MSE}[\hat{\boldsymbol{\mu}}] = \mathrm{E}[\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}]^2$. Note that,

$$\mathrm{MSE}[\hat{\boldsymbol{\mu}}] = \mathrm{E}[\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}]^2 + \mathrm{E}[\hat{\boldsymbol{\mu}} - \tilde{\boldsymbol{\mu}}]^2 + 2\,\mathrm{E}[\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}][\hat{\boldsymbol{\mu}} - \tilde{\boldsymbol{\mu}}]$$

where $\tilde{\boldsymbol{\mu}}$ is the BLUP of $\boldsymbol{\mu}$ and $\hat{\boldsymbol{\mu}}$ is the EBLUP of $\boldsymbol{\mu}$ defined in (3.4) and (3.5), respectively. In 1984, Kackar and Harville [23] showed that the cross-product term of last equation vanishes, under the normality of $\mathbf{u}$ and $\mathbf{e}$, and provided that the variance estimator $\hat{\boldsymbol{\theta}}$ is a translation invariant and even function. We have already seen in Chapter 3 that the PML, REML, APML, and AREML estimators of $\boldsymbol{\theta}$ are all translation invariant and even functions follow from (3.7) and (3.9). Therefore, the MSE of EBLUP is

$$\begin{aligned}\mathrm{MSE}[\hat{\boldsymbol{\mu}}] &= \mathrm{E}[\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}]^2 + \mathrm{E}[\hat{\boldsymbol{\mu}} - \tilde{\boldsymbol{\mu}}]^2 \\ &= \mathrm{MSE}[\tilde{\boldsymbol{\mu}}] + \mathrm{E}[\hat{\boldsymbol{\mu}} - \tilde{\boldsymbol{\mu}}]^2. \qquad (4.1)\end{aligned}$$

The first term of (4.1) is the MSE of the BLUP estimator given as (see [18] and [34])

$$\text{MSE}[\tilde{\boldsymbol{\mu}}] = G_1(\boldsymbol{\theta}) + G_2(\boldsymbol{\theta}),$$

where

$$G_1(\boldsymbol{\theta}) = (\mathbf{V}_u^{-1} + \mathbf{V}_e^{-1})^{-1}$$

and

$$G_2(\boldsymbol{\theta}) = (\mathbf{V}_u^{-1} + \mathbf{V}_e^{-1})^{-1} + (\mathbf{V}_u^{-1} + \mathbf{V}_e^{-1})^{-1}\mathbf{V}_u^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_u^{-1}(\mathbf{V}_u^{-1} + \mathbf{V}_e^{-1})^{-1}.$$

The second term of (4.1) is the uncertainty due to the estimation of $\boldsymbol{\theta}$. This term has no closed-form. Following Benavent and Morales [1], we obtain the approximation of the second term,

$$\begin{aligned} \text{E}[\hat{\boldsymbol{\mu}} - \tilde{\boldsymbol{\mu}}]^2 &= \sum_{i=1}^{R}\sum_{j=1}^{R} \text{cov}[\hat{\theta}_i, \hat{\theta}_j]\mathbf{L}^{(i)}\mathbf{V}\mathbf{L}^{(j)'} + [o(D^{-1})]_{DR\times DR} \\ &= G_3(\boldsymbol{\theta}) + [o(D^{-1})]_{DR\times DR}, \end{aligned}$$

where $\mathbf{L}^{(i)} = (\mathbf{I} - \mathbf{V}_u\mathbf{V}^{-1})\frac{\partial \mathbf{V}}{\partial \theta_i}\mathbf{V}^{-1}$.

Thus, the MSE of EBLUP is

$$\text{MSE}[\hat{\boldsymbol{\mu}}] = G_1(\boldsymbol{\theta}) + G_2(\boldsymbol{\theta}) + G_3(\boldsymbol{\theta}) + [o(D^{-1})]_{DR\times DR}.$$

Therefore, the second-order approximation of MSE of EBLUP is

$$\text{MSE}[\hat{\boldsymbol{\mu}}] = G_1(\boldsymbol{\theta}) + G_2(\boldsymbol{\theta}) + G_3(\boldsymbol{\theta}).$$

4.1.1 Estimation of Mean Squared Error

Note that the second-order approximation of MSE of EBLUP involves the unknown variance components $\boldsymbol{\theta}$. It cannot directly applied to assess the uncertainty of EBLUP for a given data set. We follow Datta et al. [8] and Datta and Lahiri [11] to obtain the second-order unbiased estimator of MSE of EBLUP, defined as

$$\mathrm{mse}[\hat{\boldsymbol{\mu}}] = G_1(\hat{\boldsymbol{\theta}}) + G_2(\hat{\boldsymbol{\theta}}) + 2G_3(\hat{\boldsymbol{\theta}}) - \widehat{\mathrm{Bias}}(\hat{\boldsymbol{\theta}})' \nabla G_1(\hat{\boldsymbol{\theta}}),$$

where $\widehat{\mathrm{Bias}}(\hat{\boldsymbol{\theta}})$ is the second-order unbiased estimator of $\mathrm{Bias}(\hat{\boldsymbol{\theta}})$ and

$$\nabla G_1(\boldsymbol{\theta}) = \left[\frac{\partial G_1(\boldsymbol{\theta})}{\partial \theta_i}\right]'_{i=1,\ldots,R}$$

is the vector of the first-order derivatives of $G_1(\hat{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\theta}$, defined as

$$\frac{\partial G_1(\boldsymbol{\theta})}{\partial \theta_i} = (\mathbf{V} - \mathbf{V}_u)\mathbf{L}^{(i)\prime},$$

for $i = 1, \ldots, R$.

In order to obtain the second-order unbiased estimator of MSE of EBLUP, we will approximate the covariance of variance components $\mathrm{cov}[\hat{\theta}_i, \hat{\theta}_j]$, $i, j = 1, \ldots, R$ for the adjusted profile maximum likelihood (APML) estimator and adjusted residual maximum likelihood (AREML) estimator.

Theorem 4.1. Under the multivariate Fay-Herriot model (3.2), the covariance matrices of the adjusted profile maximum likelihood estimator, $\hat{\boldsymbol{\theta}}^{\mathrm{AP}}$, and the adjusted residual maximum likelihood estimator, $\hat{\boldsymbol{\theta}}^{\mathrm{AR}}$, satisfy the following approximation:

$$\mathrm{cov}[\hat{\boldsymbol{\theta}}] = 2\left(\left[\mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i}\right)\right]_{i,j=1,\ldots,R}\right)^{-1} + \left[\, o(D^{-1})\,\right]_{R\times R}.$$

Proof. From Theorem 3.23,

$$
\begin{aligned}
\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} &= \left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} \\
&\quad \left[-\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} \right) + \operatorname{tr}\left(\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i=1,\ldots,R} + \left[o(D^{-1}) \right]_{R\times 1}. \\
&= \left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} \left[O(1) \right]_{R\times 1} + \left[o(D^{-1}) \right]_{R\times 1} \\
&= \left[O(D)^{-1} \right]_{R\times R} \left[O(1) \right]_{R\times 1} + \left[o(D^{-1}) \right]_{R\times 1} \\
&= \left[O(D)^{-1} \right]_{R\times 1}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\operatorname{cov}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] &= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] \right) \left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] \right)' \right] \\
&= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right) \left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right)' \right] - \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right) \left(\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} \right)' \right] \\
&\quad - \mathrm{E}\left[\left(\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} \right) \left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right)' \right] + \mathrm{E}\left[\left(\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} \right) \left(\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} \right)' \right] \\
&= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right) \left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right)' \right] - \left(\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} \right) \left(\mathrm{E}[\hat{\boldsymbol{\theta}}^{\mathrm{AP}}] - \boldsymbol{\theta} \right)' \\
&= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right) \left(\hat{\boldsymbol{\theta}}^{\mathrm{AP}} - \boldsymbol{\theta} \right)' \right] + \left[O(D^{-1}) \right]_{R\times R} \left[O(D^{-1}) \right]_{R\times R} \\
&= 2\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} + \left[o(D^{-1}) \right]_{R\times R}.
\end{aligned}
$$

Similarly, we can show that

$$
\operatorname{cov}[\hat{\boldsymbol{\theta}}^{\mathrm{AR}}] = 2\left(\left[\operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_j}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \theta_i} \right) \right]_{i,j=1,\ldots,R} \right)^{-1} + \left[o(D^{-1}) \right]_{R\times R}.
$$

□

4.1.2 Monte Carlo Simulation Study

In this section, the setting of simulation is the same as the simulation in Chapter 3. In 2010, Li and Lahiri [27] proposed second-order unbiased estimators of MSE of EBLUP based on AML.LL estimators of θ. However, these estimators are not guaranteed to be positive. Therefore, in this section, we study only the performances of estimators of MSE of EBLUP based on the APML and AREML methods in term of relative bias (RB) and relative empirical root mean squared error (RERMSE) of the second-order unbiased estimators of MSE of EBLUP: (see [1]).

$$\mathrm{RB}_r = \frac{\frac{1}{D}\sum_{d=1}^{D}\frac{1}{K}\sum_{k=1}^{K}(\mathrm{mse}_{dr}^{(k)} - \mathrm{MSE}_{dr})}{\frac{1}{D}\sum_{d=1}^{D}\frac{1}{K}\sum_{k=1}^{K}(\mathrm{MSE}_{dr})},$$

$$\mathrm{RERMSE}_r = \frac{\sqrt{\frac{1}{D}\sum_{d=1}^{D}\frac{1}{K}\sum_{k=1}^{K}(\mathrm{mse}_{dr}^{(k)} - \mathrm{MSE}_{dr})^2}}{\frac{1}{D}\sum_{d=1}^{D}\frac{1}{K}\sum_{k=1}^{K}(\mathrm{MSE}_{dr})},$$

where $\mathrm{mse}_{dr}^{(k)}$ is the rth element of $\underset{1\le d\le D}{\mathrm{diag}}\,(\mathrm{mse}_{d}^{(k)})$, for $r = 1, 2$.

We present the relative biases and relative empirical root mean squared errors of the second-order unbiased estimators of MSE of EBLUP based on different methods in Table 4.1.

Table 4.1: The relative biases and relative empirical root mean squared errors of different methods of the second-order unbiased estimators of MSE of EBLUPs

Parameter		θ_1				θ_2			
Scenario	D	PML	APML	REML	AREML	PML	APML	REML	AREML
Relative Biases									
(a)	15	-0.0687	-0.0015	-0.0402	0.0005	-0.0235	-0.0084	-0.0156	-0.0058
(b)	15	-0.1633	-0.0327	-0.1091	-0.0123	-0.0599	-0.0416	-0.0410	-0.0202
(c)	15	-0.0774	-0.0173	-0.0491	0.0160	-0.1132	-0.0749	-0.0793	-0.0369
(d)	15	0.3853	0.4757	0.2391	0.3369	0.1742	0.2415	0.0925	0.1687
(a)	30	-0.0836	0.0005	-0.0388	0.0014	-0.0152	-0.0007	-0.0073	-0.0003
(b)	30	-0.1004	-0.0113	-0.0632	-0.0036	-0.0231	-0.0149	-0.0163	-0.0092
(c)	30	-0.1155	-0.0248	-0.0746	-0.0075	-0.0348	-0.0325	-0.0266	-0.0186
(d)	30	0.1265	0.1722	0.0808	0.1287	0.0327	0.0583	0.0146	0.0437
Relative Empirical Root Mean Squared Errors									
(a)	15	0.1314	0.0889	0.1243	0.0946	0.0657	0.0631	0.0675	0.0656
(b)	15	0.3425	0.2641	0.3255	0.2665	0.2385	0.2232	0.2290	0.2200
(c)	15	0.4191	0.3965	0.4081	0.3836	0.3554	0.3325	0.3375	0.3122
(d)	15	0.4449	0.5182	0.3762	0.4249	0.2659	0.3042	0.2693	0.2822
(a)	30	0.1632	0.0722	0.1257	0.0731	0.0518	0.0490	0.0508	0.0495
(b)	30	0.2888	0.2113	0.2706	0.2128	0.1718	0.1640	0.1671	0.1634
(c)	30	0.3848	0.3208	0.3726	0.3217	0.2616	0.2474	0.2540	0.2433
(d)	30	0.3562	0.3599	0.3788	0.3705	0.2725	0.2629	0.2924	0.2770

From Table 4.1, we can see that the relative biases of second-order unbiased estimator of MSE of EBLUPs are close to zero and the relative root mean squared errors of second-order unbiased estimator of MSE of EBLUPs decrease when sample size increases, or equivalently when sampling variance decreases. Unless the scenario (d), the relative biases of second-order unbiased estimator of MSE of EBLUPs from the APML and AREML methods are closer to zero than that from the PML and REML methods for all parameters, all scenarios and all sample sizes ($D = 15, 30$). The root mean squared errors of second-order unbiased estimators of MSE of EBLUPs from the APML and AREML methods are less than that from PML and REML methods for all parameters, all scenarios and all sample sizes ($D = 15, 30$), except the case in scenario (d). For example, for the first parameter θ_1, in case $D = 15$ and scenario (a), the relative biases of second-order unbiased estimator of MSE of EBLUPs from APML and PML methods are consecutively -0.0015 and -0.0687. Another example, for the first parameter θ_1, in case $D = 15$ and scenario (a), the root mean squared errors of second-order unbiased estimator of MSE of EBLUPs from APML and PML methods are 0.0889 and 0.1314.

4.2 Prediction interval

In this section, we review the prediction interval of Li and Lahiri [27]. In the small area context, prediction intervals are often produced using the standard rule, which is EBLUP $\pm z_{\alpha/2}\sqrt{\text{mse}}$, where $z_{\alpha/2}$ is the upper $100(1-\alpha/2)\%$ point of the standard normal distribution. These prediction intervals are asymptotically correct when sample size is large. However, when sample size is small, it is not efficient prediction interval and depends on the choice of mse. In this section, we consider the three prediction intervals: (1) Cox's empirical Bayes prediction interval, using $G_1(\hat{\boldsymbol{\theta}})$ as the mse; (2) Traditional prediction interval, using the second-order unbiased estimator of MSE of EBLUP mse$[\hat{\boldsymbol{\mu}}]$ as the mse; (3) Parametric bootstrap prediction interval presenting in the next section.

4.2.1 Parametric Bootstrap Prediction Interval

In 2008, Chatterjee et al. [4] proposed a parametric bootstrap method to obtain a prediction interval directly from the bootstrap histogram of the pivot

$$\text{mse}^{-1/2}(\text{true mean} - \text{EBLUP}).$$

Their methods using ordinary least square estimators of $\boldsymbol{\beta}$ and PML or REML estimators of the variance components. In 2010, Li and Lahiri [27] extended the method using weighted least squares estimator of $\boldsymbol{\beta}$ and APL.LL or AREML.LL estimators of the variance components. They showed that the coverage accuracy of this prediction interval is in the order of $O(D^{-3/2})$.

Follow Li and Lahiri [27], a prediction interval of μ_{dr} can be constructed based on $\mathcal{L}_{dr}$, the distribution of $(G_{1dr}(\hat{\boldsymbol{\theta}}))^{-1/2}(\mu_{dr} - \hat{\mu}_{dr})$. They approximate $\mathcal{L}_{dr}$ using a parametric bootstrap method. Let

$$\mathbf{y}_d^* = \mathbf{X}_d\hat{\boldsymbol{\beta}} + \mathbf{u}_d^* + \mathbf{e}_d^*,$$

where $\mathbf{u}_d^* \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}))$ and $\mathbf{e}_d^* \overset{ind}{\sim} N(\mathbf{0}, \mathbf{V}_{ed})$, for $d = 1, \ldots, D$. Let $\hat{\boldsymbol{\beta}}^*$, $\boldsymbol{\Sigma}^*(\hat{\boldsymbol{\theta}})$, $\boldsymbol{\mu}_d^*$ and $G_1^*(\hat{\boldsymbol{\theta}})$ be computed from $\mathbf{y}_d^*$ and $\boldsymbol{\mu}_d^* = \mathbf{X}_d\hat{\boldsymbol{\beta}}^* + \mathbf{u}_d^*$. The distribution of $(G_{1dr}^*(\hat{\boldsymbol{\theta}}))^{-1/2}(\mu_{dr}^* - \hat{\mu}_{dr}^*)$, conditional on the data $\boldsymbol{\mu}$, is the parametric bootstrap approximation $\mathcal{L}_{dr}^*$ of $\mathcal{L}_{dr}$. The parametric bootstrap prediction interval is given in the following theorem.

Theorem 4.2. ([27]) With $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$, for a preassigned $\alpha \in (0, 1)$ and arbitrary $d = 1, \ldots, D$ and $r = 1, \ldots, R$, let q_1 and q_2 be real numbers such that

$$\mathcal{L}_{dr}^*(q_2) - \mathcal{L}_{dr}^*(q_1) = 1 - \alpha.$$

Then, under the regularity conditions given in [27], we have

$$\text{Prob}\left[\hat{\mu}_{dr} - q_1 G_{1dr}(\hat{\boldsymbol{\theta}}) \le \mu_{dr} \le \hat{\mu}_{dr} - q_2 G_{1dr}(\hat{\boldsymbol{\theta}})\right] = 1 - \alpha + O(D^{-3/2}).$$

4.2.2 Monte Carlo Simulation Study

In order to study the performances of prediction interval, we compare twelve prediction intervals of μ_{dr} using coverage probabilities (with a nominal coverage of 0.95) and average lengths. Different twelve prediction intervals include the four Cox's empirical Bayes prediction intervals, four traditional prediction intervals of the form $\hat{\mu}_{dr} \pm z_{\alpha/2}\sqrt{\text{mse}_{dr}}$ and four prediction intervals based on the parameter bootstrap methods. All the results are based on 10000 simulation runs. For each parameter bootstrap method, we consider 200 bootstrap samples and the shortest length prediction intervals.

Table 4.2: Average coverages of twelve prediction intervals

Parameter		θ_1				θ_2			
Scenario	D	PML	APML	REML	AREML	PML	APML	REML	AREML
Cox's Empirical Bayes Prediction Interval									
(a)	15	0.8801	0.9120	0.9084	0.9269	0.9235	0.9264	0.9330	0.9348
(b)	15	0.7288	0.8521	0.8047	0.8891	0.8690	0.8860	0.8977	0.9090
(c)	15	0.6650	0.7555	0.7549	0.8309	0.7485	0.8010	0.8102	0.8582
(d)	15	0.4823	0.7142	0.5761	0.7895	0.5893	0.7402	0.6611	0.8070
(a)	30	0.9039	0.9366	0.9251	0.9408	0.9385	0.9408	0.9424	0.9435
(b)	30	0.8525	0.9130	0.8847	0.9248	0.9243	0.9273	0.9322	0.9341
(c)	30	0.7596	0.8552	0.8191	0.8855	0.8838	0.8930	0.9036	0.9112
(d)	30	0.6517	0.7556	0.7117	0.8047	0.7696	0.8192	0.8073	0.8542
Traditional Prediction Interval									
(a)	15	0.9373	0.9476	0.9408	0.9479	0.9463	0.9484	0.9473	0.9485
(b)	15	0.9047	0.9371	0.9130	0.9399	0.9349	0.9405	0.9382	0.9430
(c)	15	0.9102	0.9262	0.9114	0.9309	0.9115	0.9233	0.9156	0.9303
(d)	15	0.9759	0.9811	0.9648	0.9738	0.9608	0.9674	0.9500	0.9609
(a)	30	0.9328	0.9493	0.9407	0.9493	0.9474	0.9493	0.9484	0.9493
(b)	30	0.9159	0.9426	0.9234	0.9435	0.9437	0.9457	0.9450	0.9464
(c)	30	0.8952	0.9277	0.9027	0.9308	0.9326	0.9360	0.9350	0.9387
(d)	30	0.9492	0.9562	0.9391	0.9492	0.9397	0.9455	0.9347	0.9425
Parametric Bootstrap Prediction Interval									
(a)	15	0.9541	0.9816	0.9464	0.9585	0.9906	0.9834	0.9766	0.9706
(b)	15	0.8297	0.9828	0.8879	0.9499	0.9770	0.9867	0.9857	0.9808
(c)	15	0.8076	0.9588	0.8645	0.9399	0.8892	0.9746	0.9242	0.9810
(d)	15	0.5827	0.9324	0.6660	0.9325	0.7092	0.9588	0.7632	0.9648
(a)	30	0.9361	0.9693	0.9470	0.9623	0.9744	0.9727	0.9657	0.9653
(b)	30	0.9220	0.9790	0.9247	0.9576	0.9914	0.9822	0.9776	0.9690
(c)	30	0.8551	0.9820	0.8898	0.9433	0.9817	0.9881	0.9854	0.9834
(d)	30	0.7676	0.9437	0.8037	0.9165	0.8879	0.9727	0.9084	0.9791

From Table 4.2, we can see that the average coverages of prediction interval is close to a nominal coverage of 0.95 when sample size increases, or equivalently when sampling variance decreases. Cox's prediction interval methods give average coverage below 0.95. Therefore it is undercoverage. The traditional prediction intervals give average coverage below 0.95, except in case $D = 15$ in scenario (d). It is higher than the average coverage of Cox's prediction intervals. Therefore, it is undercoverage. The parametric bootstrap prediction interval from PML and REML methods give average coverage below 0.95 for

θ_1 and above 0.95 for θ_2. However, the average coverage of the parametric bootstrap prediction intervals from APML and AREML are close to 0.95 similar the average coverage of the traditional prediction intervals. For all cases, the average coverages of prediction intervals from the APML and AREML methods are close to 0.95 than that based on PML and REML method.

Table 4.3: Average lengths of twelve prediction intervals

Parameter		θ_1				θ_2			
Scenario	D	PML	APML	REML	AREML	PML	APML	REML	AREML
Cox's Empirical Bayes Prediction Interval									
(a)	15	2.3597	2.4432	2.4718	2.5222	2.8862	2.8970	2.9496	2.9586
(b)	15	2.5927	2.9852	2.9487	3.2279	3.7798	3.8133	3.9861	4.0283
(c)	15	2.8999	3.2387	3.4401	3.7548	4.0600	4.3015	4.5575	4.7942
(d)	15	2.6631	3.6801	3.3260	4.3307	3.9624	4.7820	4.6329	5.4706
(a)	30	2.4482	2.5328	2.5242	2.5648	2.9594	2.9660	2.9876	2.9915
(b)	30	3.0262	3.2251	3.1982	3.3321	4.0344	4.0333	4.1258	4.1289
(c)	30	3.2875	3.6449	3.6360	3.8931	4.8204	4.8147	5.0158	5.0312
(d)	30	3.3874	3.8169	3.8151	4.2118	4.9985	5.2557	5.3773	5.6335
Traditional Prediction Interval									
(a)	15	2.7112	2.7337	2.6930	2.7120	3.1237	3.1291	3.1066	3.1121
(b)	15	3.6292	3.7532	3.6331	3.7487	4.5125	4.5096	4.4923	4.5104
(c)	15	4.6636	4.7758	4.6619	4.8032	5.6843	5.7698	5.6971	5.8068
(d)	15	7.0033	7.1276	6.6367	6.8257	8.0268	8.1365	7.7014	7.8688
(a)	30	2.6195	2.6579	2.6317	2.6530	3.0624	3.0674	3.0603	3.0635
(b)	30	3.4621	3.5597	3.4857	3.5610	4.3455	4.3419	4.3420	4.3436
(c)	30	4.2200	4.3859	4.2621	4.4026	5.5138	5.4916	5.5050	5.5080
(d)	30	5.7441	5.8291	5.5992	5.7042	6.9305	6.9867	6.8328	6.9073
Parametric Bootstrap Prediction Interval									
(a)	15	4.5627	3.6217	3.2005	3.0202	6.9642	4.8019	3.9438	3.5999
(b)	15	19.4376	6.7437	14.6885	4.7906	26.8797	10.4030	20.9599	6.1641
(c)	15	27.1996	11.4028	21.6528	7.8520	41.3225	17.0033	32.2858	10.3392
(d)	15	35.5610	10.3064	35.1631	9.0612	53.4344	14.6791	57.0920	12.1327
(a)	30	2.9633	3.0447	2.8822	2.9119	3.7379	3.6521	3.4305	3.4131
(b)	30	7.7219	4.6489	4.3831	4.0118	10.6830	6.6309	5.9016	5.0618
(c)	30	22.3324	8.5641	12.7904	5.7790	35.0799	14.9128	18.2328	7.7692
(d)	30	31.5128	11.4712	24.9931	9.0046	51.5909	18.1979	39.2564	12.5033

From Table 4.3, we can see that the average lengths of Cox's empirical Bayes prediction intervals increase when sample size increases or equivalently when sampling variance increases. The average lengths of Traditional and parametric bootstrap prediction intervals decrease when sample size increases, or equivalently when sampling variance decreases. The average lengths of Cox's empirical Bayes and traditional prediction intervals from APML and AREML methods are higher than that from PML and REML methods. The average lengths of parametric bootstrap prediction intervals from APML and AREML methods are less than the average lengths of parametric bootstrap prediction intervals from PML and REML methods. However, the average lengths of Cox's empirical Bayes prediction intervals are shorter than average lengths of traditional prediction intervals. The parametric bootstrap prediction interval gives the longest average lengths. Due to the approximated pivots $(G^*_{1dr}(\hat{\boldsymbol{\theta}}))^{-1/2}(\mu^*_{dr} - \hat{\mu}^*_{dr})$ are undefined when the estimators of θ_r are zero, we replaced 0.01 in those zero estimates. Thus, the pivot are large, which provide the large lengths and overcoverages of the prediction intervals.

CHAPTER V

APPLICATION

In this chapter, we apply the adjusted maximum likelihood (AML) method in Chapter 3 to study the average household income and average household expenditure data in Thailand. Then we apply the adjusted maximum likelihood method of Li and Lahiri (AML.LL) [27]. In Section 5.1, we describe the data used in this thesis. Then, we provide the estimators of variance components from the PML, REML, APML and AREML methods and EBLUP from the APML and AREML methods for this data in Section 5.2.

5.1 Data Description

In this section, we consider the average household income and average household expenditure data in Thailand. We use this data set in 2017 from the Household Socio-Economic Survey (SES) 2017, which is designed with stratified two-stage sampling (see [39]). Next, we present the details of this data set in Table 5.1.

Table 5.1: Sample size, mean and standard deviation of the average household income and average household expenditure of SES 2017

SES 2017	Region	Size	Municipal area		Non-municipal area	
			Mean	SD	Mean	SD
average household income (Unit: 10,000 Baht)	Central	18	3.1229	0.6693	2.7230	0.6699
	East	7	2.9715	0.3799	2.4834	0.2609
	North	17	2.3638	0.4822	1.7062	0.2673
	Northeast	20	2.3727	0.3414	1.8042	0.3550
	South	14	2.9824	0.7153	2.4504	0.8221
	Total	**76**	**2.7158**	**0.6321**	**2.1815**	**0.6731**
average household expenditure (Unit: 10,000 Baht)	Central	18	2.3147	0.5168	2.1496	0.5613
	East	7	2.2101	0.1766	2.0382	0.2323
	North	17	1.7734	0.3014	1.4013	0.2293
	Northeast	20	1.8855	0.2631	1.5568	0.2566
	South	14	2.3561	0.4689	1.9602	0.5073
	Total	**76**	**2.0787**	**0.4455**	**1.7811**	**0.4890**

In Table 5.1, the data from SES 2017 consist of 5 regions, including Central (C), East (E), North (N), Northeast (NE), and South (S). Each province (except Bangkok) is divided into two parts: municipal area and non-municipal area. From the table, the means of the average household income are higher than the means of the average household expenditure. The means of the average household income and average household expenditure of municipal area are higher than those of non-municipal area. However, the standard deviations of the average household income and average household expenditure different of both municipal area and non-municipal area are not significantly.

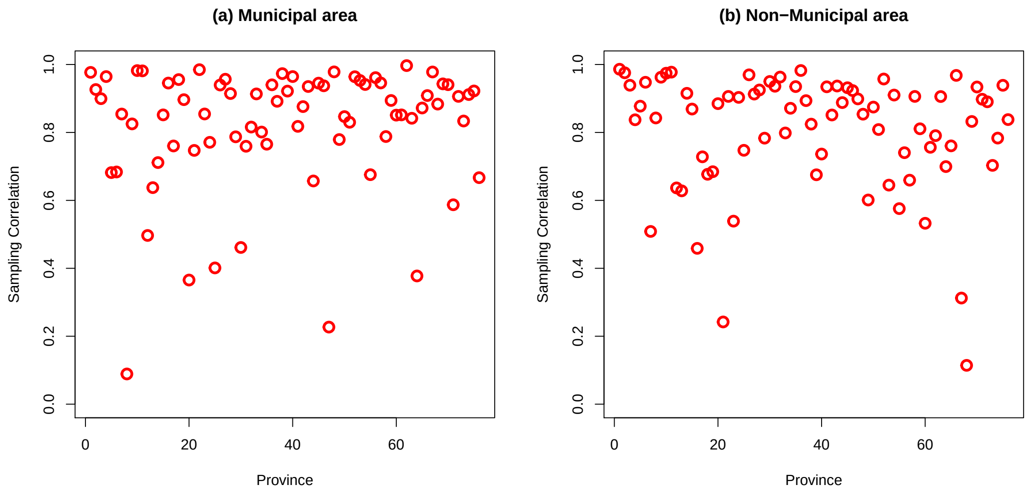

Figure 5.1: The correlation of the average household income and average household expenditure

From Figure 5.1, we can see that the sample correlations of the average household income and average household expenditure in both areas are close to 1. Therefore, a bivariate model is more appropriate than a univariate model for this data set.

For explanatory variables, we select four explanatory variables, including proportion of households that cement or brick dwellings, proportion of households that own land, proportion of households using gas for cooking, and average population per private household. These four variables are collected from the Population and Housing Census 2010.

5.2 AML Method of Average Household Income and Average Household Expenditure

In this section, we first consider the special case of the covariance matrix. That is when $\theta_i = \theta$ for all $i = 1, \ldots, R$ by applying the AML.LL method defined in Section 3.2.3. Next, we consider a more general situation, where the covariance matrix is defined in (3.3).

First, for the special case of covariance matrix, where $\theta_i = \theta$ for all $i = 1, \ldots, R$, we follow the multivariate Fay-Herriot model with $R = 2$, or bivariate Fay-Herriot model, with the average household income and average household expenditure as response variables and four explanatory variables. The variance component θ are estimated by the PML, REML, APML.LL and AREML.LL methods using **optim** function in R [33].

Table 5.2: The value of estimate of θ for standard methods and adjusted methods of Li and Lahiri

Region	D	Municipal area				Non-municipal area			
		PML	APML	REML	AREML	PML	APML	REML	AREML
C	18	0.0076	0.0123	0.0182	0.0276	0	0.0042	0.0062	0.0180
E	7	0	0.0028	0	0.0309	0	0.0015	0	0.0140
N	17	0.0258	0.0343	0.0550	0.0702	0.0056	0.0079	0.0122	0.0176
NE	20	0.0124	0.0161	0.0230	0.0298	0.0145	0.0166	0.0205	0.0239
S	14	0.0083	0.0164	0.0468	0.0894	0	0.0230	0.0543	0.0800

Table 5.2 presents the PML, APML.LL, REML, and AREML.LL estimates of θ. From the table, we can see that the PML and REML estimates of θ are zeros in some cases. For example, the REML estimates are zeros for both municipal area and non-municipal area in the east region. For all cases, the APML.LL and AREML.LL methods prevent the zero estimates of θ.

Next, we apply the AML.LL estimates to produce EBLUP estimates of the average household income and average household expenditure and present in Table 5.3.

Table 5.3: Mean and standard deviation of EBLUP based on AML.LL methods of the average household income and average household expenditure

Region	D	Municipal area				Non-municipal area			
		APML.LL		AREML.LL		APML.LL		AREML.LL	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
average household income									
C	18	2.7665	0.4774	2.8171	0.4608	2.5160	0.5112	2.5479	0.5210
E	7	2.8493	0.2042	2.8642	0.2056	2.4534	0.2279	2.4649	0.2338
N	17	2.2072	0.3337	2.2443	0.3494	1.6450	0.2037	1.6603	0.2076
NE	20	2.2570	0.2820	2.2791	0.2788	1.7297	0.2602	1.7381	0.2658
S	14	2.7241	0.4238	2.8139	0.4766	2.1259	0.3766	2.1913	0.4445
Total	**76**	**2.5071**	**0.4505**	**2.5511**	**0.4655**	**2.0366**	**0.4967**	**2.0629**	**0.5164**
average household expenditure									
C	18	2.2053	0.4130	2.2219	0.4167	2.0534	0.4590	2.0695	0.4679
E	7	2.1980	0.1926	2.1998	0.1912	2.0385	0.2432	2.0379	0.2382
N	17	1.7096	0.2126	1.7292	0.2346	1.3541	0.1829	1.3668	0.1909
NE	20	1.8378	0.2395	1.8480	0.2410	1.5236	0.2242	1.5289	0.2295
S	14	2.2567	0.3477	2.2978	0.3811	1.8548	0.3547	1.8975	0.4140
Total	**76**	**2.0065**	**0.3727**	**2.0252**	**0.3849**	**1.7196**	**0.4175**	**1.7355**	**0.4329**

From Table 5.3, we can see that the means of average household income is higher than the means of average household expenditure. For example, in municipal area of the central region, the means of average household income and expenditure are 27,665 and 22,053 Bahts, respectively. The means of the average household income and means of the average household expenditure of municipal area are higher than that non-municipal area. The standard deviations of the average household income and average household expenditures have similar values for both municipal area and non-municipal area.

We apply the AML method defined in Section 3.3 to the model with the covariance matrix defined in Section 3.1. Similar to the previous study, we follow the bivariate Fay-Herriot model, with the average household income and average household expenditure as response variables and four explanatory variables. The variance components $\boldsymbol{\theta} = (\theta_1, \theta_2)$ are estimated by PML, REML, APML and AREML methods using **optim** function in R [33].

Table 5.4: The value of estimate of θ_1 and θ_2 for standard methods and adjusted methods

Region	D		Municipal area				Non-municipal area			
			PML	APML	REML	AREML	PML	APML	REML	AREML
C	18	θ_1	0.0986	0.0010	0	0.0060	0	0.0004	0.0072	0.0090
		θ_2	0	0.0116	0.0197	0.0219	0.0127	0.0003	0.0061	0.0070
E	7	θ_1	0.0004	0.0006	0.0004	0.0037	0.0004	0.0004	0.0088	0.0033
		θ_2	0	0.0004	0	0.0036	0	0.0002	0	0.0014
N	17	θ_1	0.1853	0.0406	0.0818	0.0842	0.0036	0.0041	0.0137	0.0146
		θ_2	0	0.0219	0.0405	0.0415	0.0057	0.0059	0.0119	0.0124
NE	20	θ_1	0.0067	0.0075	0.0237	0.0248	0.0044	0.0048	0.0105	0.0110
		θ_2	0.0127	0.0130	0.0229	0.0234	0.0172	0.0174	0.0237	0.0240
S	14	θ_1	0.1299	0.0231	0.0774	0.0889	0.3099	0.0007	0.3578	0.0766
		θ_2	0	0.0006	0.0167	0.0210	0	0.0046	0	0.0522

From Table 5.4, we can see that the PML and REML estimates of θ_1 or θ_2 are zeros in some cases. For example, the REML estimates are zeros in municipal area of the central region, non-municipal area of the south region, and both municipal area and non-municipal area of the east region. For all cases, the APML and AREML methods prevent the zero estimates of θ_1 and θ_2.

Next, we apply the AML estimates to produce EBLUP estimates of the average household income and average household expenditure and present in Table 5.5.

Table 5.5: Mean and standard deviation of EBLUP based on AML methods of the average household income and average household expenditure

Region	D	Municipal area				Non-municipal area			
		APML.		AREML		APML		AREML	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
average household income									
C	18	2.7385	0.4804	2.7780	0.4650	2.4957	0.5097	2.5276	0.5133
E	7	2.8488	0.2018	2.8497	0.2048	2.4512	0.2264	2.4553	0.2299
N	17	2.2041	0.3326	2.2414	0.3482	1.6370	0.2034	1.6552	0.2060
NE	20	2.2405	0.2884	2.2715	0.2793	1.7101	0.2419	1.7235	0.2509
S	14	2.7318	0.4268	2.8042	0.4685	2.0765	0.3358	2.1845	0.4324
Total	**76**	**2.4969**	**0.4524**	**2.5361**	**0.4594**	**2.0156**	**0.4871**	**2.0510**	**0.5085**
average household expenditure									
C	18	2.1993	0.4124	2.2124	0.4152	2.0425	0.4576	2.0588	0.4608
E	7	2.1972	0.1928	2.1982	0.1925	2.0385	0.2446	2.0388	0.2434
N	17	1.7048	0.2045	1.7242	0.2256	1.3478	0.1806	1.3623	0.1872
NE	20	1.8314	0.2437	1.8443	0.2397	1.5155	0.2221	1.5228	0.2272
S	14	2.2448	0.3388	2.2791	0.3601	1.8104	0.3093	1.8894	0.4005
Total	**76**	**2.0001**	**0.3702**	**2.0173**	**0.3773**	**1.7053**	**0.4081**	**1.7289**	**0.4277**

From Table 5.5, we can see that the means of average household income is higher than the means of average household expenditures. For example, in municipal area of the central region, the means of average household income and expenditure are 27,385 and 21,993 Bahts, respectively. The means of the average household income and means of average household expenditure of municipal area are higher than that non-municipal area. The standard deviations of the average household income and standard deviation of the average household expenditures have similar values both both municipal area and non-municipal area.

CHAPTER VI

CONCLUSIONS AND FUTURE WORK

In this thesis, we have extended the adjusted maximum likelihood method proposed of Li and Lahiri [27] to multivariate Fay-Herriot models. We proposed a new adjusted maximum likelihood method, as an alternative method to produced the positive estimator. From our work, under the regularity conditions, the adjusted maximum likelihood estimators of $\boldsymbol{\theta}$ are consistent. Furthermore, under the same asymptotic setting the biases and mean squared errors of the adjusted and original maximum likelihood estimator of $\boldsymbol{\theta}$ are all equivalent. In terms of EBLUP and mean squared error estimator of EBLUP, the simulation support the usage of new adjusted maximum likelihood estimator of $\boldsymbol{\theta}$. For the application, the zero estimates of $\boldsymbol{\theta}$ is available for the average household income and average household expenditure data. Then the adjusted maximum likelihood method can prevent that zero estimates and produce the EBLUP with nonzero weight.

For interesting work in the future, we expect that our proposed method can be applied to obtain the adjusted maximum likelihood method for other models, for instance AR(1) and HAR(1) Fay-Herriot models, which are found in [1].

REFERENCES

[1] R. Benavent and D. Morales. Multivariate Fay-Herriot models for small area estimation. *Computational Statistics and Data Analysis*, 94:372–390, 2016.

[2] D. S. Bernstein. *Matrix Mathematics: theory, facts, and formulas.* Princeton University Press, New Jersey, 2 edition, 2009.

[3] G. Casella and R. L. Berger. *Statistical Inference.* Duxbury, California, 2 edition, 2002.

[4] A. Chatterjee, P. Lahiri, and H. Li. Parametric bootstrap approximation to the distribution of eblup, and related prediction intervals in linear mixed models. *The Annals of Statistics*, 36:1221–1245, 2008.

[5] D. R. Cox. Prediction intervals and empirical Bayes confidence intervals. *Perspectives in Probability and Statistics, Papers in Honor of M.S. Bartlett, Academic Press*, 12:47–55, 1975.

[6] S. Dan. The Calculus of Functions of Several Variables, 2001. `https://onlinebooks.library.upenn.edu/webbin/book/lookupid?key=olbp48995`.

[7] K. Das, J. Jiang, and J. N. K. Rao. Mean squared error of empirical predictor. *The Annals of Statistics*, 32:818–840, 2004.

[8] G. S. Datta, B. Day, and I. Basawa. Empirical best linear unbiased and empirical bayes prediction in multivariate small area estimation. *Journal of Statistical Planning and Inference*, 75:269–279, 1999.

[9] G. S. Datta, R. E. Fay, and M. Ghosh. Hierarchical and empirical Bayes multivariate analysis in small area estimation. In *Bureau of the Census 1991 Annual Research Conference*, pages 63–79, 1991.

[10] G. S. Datta, M. Ghosh, N. Nangia, and K. Natarajan. Estimation of median income of four-person families: a Bayesian approach. *In: Berry, D.A., Chaloner, K.M., Geweke, J.M. (Eds.). Bayesian Analysis in Statistics and Econometrics*, pages 129–140, 1996.

[11] G. S. Datta and P. Lahiri. A unified measure of uncertainty of estimated best linear unbiased predictions in small area estimation problems. *Statistica sinica*, 10:613–627, 2000.

[12] G. S. Datta, J. N. K. Rao, and D. D. Smith. On measuring the variability of small area estimators under a basic area level model. *Biometrika*, 92:183–196, 2005.

[13] E. Deadman and S. Relton. Taylor's theorem for matrix functions with applications to condition number estimation. *Linear Algebra and its Application*, 504:354–371, 2016.

[14] R. E. Fay. Application of multivariate regression of small domain estimation. *In: Platek, R., Rao, J.N.K., Särndal, C.E., Singh, M.P. (Eds.), Small Area Statistics*, pages 91–102, 1987.

[15] R. E. Fay and R. A. Herriot. Estimates of income for small places: An application of James-Stein procedure to census data. *Journal of the American Statistical Association*, 74(366):269–277, 1979.

[16] W. González-Menteiga, M. J. Lombardiáa, I. Molina, D. Morales, and L. Santamaría. Analytic and bootstrap approximations of prediction errors under a multivariate Fay-Herriot model. *Computational Statistics and Data Analysis*, 52(12):5242–5252, 2008.

[17] H. O. Hartley and J. N. K. Rao. Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54(1/2):93–108, 1967.

[18] C. R. Henderson. Best linear unbiased estimation and prediction under selection model. *Biometrics*, 31(2):423–427, 1975.

[19] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 2 edition, 2013.

[20] J. Jiang. *Linear and Generalized Linear Mixed Models and Their Applications.* Springer, New York, 2007.

[21] J. Jiang and P. Lahiri. Mixed model prediction and small area estimation (with discussions). *Test*, 15(1):1–96, 2006.

[22] R. Kackar and D. A. Harville. Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. *Communications in Statistics—Theory and Methods, Series A*, 10:1249–1261, 1981.

[23] R. Kackar and D. A. Harville. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Journal of the American Statistical Association*, 79:853–862, 1984.

[24] E. Kreyszig. *Introductory Functional Analysis with Applications.* John Wiley and Sons, Inc, New Jersey, 1978.

[25] P. Lahiri and H. Li. Generalized maximum likelihood method in linear mixed models with an application in small area estimation. In *the Federal Committee on Statistical Methodology Research Conference*, 2009.

[26] H. Li. *Small Area Estimation: An Empirical Best Linear Unbiased Prediction Approach.* Ph.D. Thesis, Department of Mathematics, University of Maryland, College Park, 2007.

[27] H. Li and P. Lahiri. An adjusted maximum likelihood method for solving small area estimation problems. *Journal of Multivariate Analysis*, 101(4):882–892, 2010.

[28] A. M. Mathai and S. B. Provost. *Quadratic Forms in Random Variables: theory and applications.* Marcel dekker, Inc., New York, 1992.

[29] C. Morris and R. Tang. Estimating random effects via adjustment for density maximization. *Statistical Science*, 26(2):271–287, 2011.

[30] M. K. Nesa. *Multivariate Small Area Estimation for Heath Indicators.* Ph.D. Thesis, School of Mathematics and Applied Statistics, University of Wollongong, 2017.

[31] H. D. Patterson and R. Thompson. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *Biometrika*, 58:545–554, 1971.

[32] N. G. N. Prasad and J. N. K. Rao. The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85(409):163–171, 1990.

[33] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[34] J. N. K. Rao and I. Molina. *Small Area Estimation*. John Wiley and Sons, Inc, New York, 2 edition, 2015.

[35] A. C. Rencher and G. B. Schaalje. *Linear Models in Statistics*. John Wiley and Sons, Inc., New Jersey, 2 edition, 2008.

[36] S. R. Searle, G. Casella, and C. E. McCulloch. *Sampling Theory of Surveys with Applications*. John Wiley and Sons, Inc, New York, 1992.

[37] G. A. F. Seber. *A Matrix Handbook for Statisticians*. John Wiley and Sons, Inc, New Jersey, 2008.

[38] V. K. Srivastava and R. Tiwari. Evaluation of expectations of products of stochastic matrices. *Scandinavian Journal of Statistics*, 3(3):135–138, 1976.

[39] P. V. Sukhatme. *Sampling Theory of Surveys with Applications*. The Iowa State University Press, Ames, Iowa, 1984.

[40] A. Ubaidillah, K. A. Notodiputro, A. Kurnia, and I. W. Mangku. Multivariate Fay-Herriot models for small area estimation with application to household consumption per capita expenditure in indonesia. *Journal of Applied Statistics*, 46(15):2845–2861, 2019.

[41] M. Yoshimori and P. Lahiri. A new adjusted maximum likelihood method for the Fay-Herriot small area model. *Journal of Multivariate Analysis*, 124:281–294, 2014.

BIOGRAPHY

Name	Mr. Annop Angkunsit
Date of Birth	September 7, 1995
Place of Birth	Bangkok, Thailand
Education	B.Sc. (Mathematics, First Class Honours), Chulalongkorn University, 2017
Scholarship	The His Royal Highness Crown Prince Maha Vajiralongkorn Scholarship from the Graduate School, Chulalongkorn University to commemorate the 72^{nd} anniversary of his Majesty King Bhumibol Aduladej is gratefully acknowledged

Publications

- A. Angkunsit and J. Suntornchost. Variable selection in multivariate linear regression models subject to sampling errors. *Chamchuri Journal of Mathematics*, 10:28–52, 2018.
- A. Angkunsit and J. Suntornchost. Bivariate Fay-Herriot Models with Application to Thai Socio-Economic Data. *Naresuan University Journal: Science and Technology (NUJST)*, 29(1):36–49, 2021.