5-11-2018

Inference in Multivariate Generalized Ornstein-Uhlenbeck Processes with a Change-point

Lei Shen
University of Windsor

Inference in Multivariate Generalized Ornstein-Uhlenbeck Processes with a Change-point

by

Lei Shen

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Mathematics and Statistics
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2018

Inference in Multivariate Generalized Ornstein-Uhlenbeck

Processes with a Change-point

by

Lei Shen

APPROVED BY:

D. Li
Department of Economics

A. A. Hussein
Department of Mathematics and Statistics

S. Nkurunziza, Advisor
Department of Mathematics and Statistics

May 3, 2018

DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION

I. Co-Authorship

I hereby declare that this thesis incorporates material that is result of joint research, as follows: Chapter 4, 5, and 6 of the thesis were co-authored with professor Sévérien Nkurunziza. In all cases, the primary contributions, simulation, data analysis, interpretation, and writing were performed by the author, and the contribution of co-authors was primarily through the provision of some theoretical results.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work

II. Previous Publication

This thesis includes one original paper that has been previously published/submitted for publication in peer reviewed journals, as follows:

Thesis Chapter	Publication title/full citation	Publication status*
Chapter 4, 5, and 6	Nkurunziza, S., and Shen, L., (2018). Inference in a multivariate generalized mean-reverting process with a change-point. *Statistical Inference for Stochastic Processes (Submitted).*	Under review

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor.

III. General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

In this paper, we study inference problem about the drift parameter matrix in multivariate generalized Ornstein-Uhlenbeck processes with an unknown change-point. In particular, we study the case where the matrix parameter satisfies uncertain restriction. Thus, we generalize some recent findings about univariate generalized Ornstein-Uhlenbeck processes. First, we establish a weaker condition for the existence of the unrestricted estimator (UE) and we derive the unrestricted estimator and the restricted estimator. Second, we establish the joint asymptotic normality of the unrestricted estimator and the restricted estimator under the sequence of local alternatives. Third, we construct a test for testing the uncertain restriction. The proposed test is also useful for testing the absence of the change-point. Fourth, we derive the asymptotic power of the proposed test and we prove that it is consistent. Fifth, we propose the shrinkage estimators and we prove that shrinkage estimators dominate the unrestricted estimator. Finally, in order to illustrate the performance of the proposed methods in short and medium period of observations, we conduct a simulation study which corroborate our theoretical findings.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to whom directly or indirectly helped me and guided me work through this thesis.

I would like to first express my sincerest gratitude to my supervisor, Dr. Sévérien Nkurunziza. His encouragement, experience, advice, guide me through the hard time. His rigorous and conscientious attitude, and enthusiasm for teaching always inspire me deeply. All the things I learned from Dr. Sévérien Nkurunziza will not only benefit my academia, but for my entire life.

My great gratitude also goes to the departmental and external readers of my thesis, Dr. Abdulkadir Hussein from Mathematic and Statistic department and Dr. Dingding Li from Economic department, for patiently reading my work and for all the valuable advice. My great gratitude also goes to all the staffs in the department, all the teachers who use their experience to help me work through the thesis, and all the kind staffs assist me in many ways.

I would like to thank to all my family members: my parents, my girlfriend, and my grandma. They always stand by me and have faith in me. Without their support financially and mentally, I believe it would not have been possible to accomplish my studies successfully. Also, I would like to thank my coworker, Kang Fu, who always gives great ideas and helps me work through the challenging problems.

CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

The Ornstein-Uhlenbeck process (O-U) has been applied to model different phenomena in finance, physics, insurance among others. For instance, Vasicek (1977) applied univariate Ornstein-Uhlenbeck process to explain the mean reversion feature of bond yields, while Langetieg (1980) applied the multivariate Ornstein-Uhlenbeck process to analyse correlated economic factors. To give more applications of the Ornstein-Uhlenbeck (O-U) process, we also quote Erlwein *et al.* (2010) who used this process to study the electricity market. The O-U has also been used to analyse the insurance problems (see Liang *et al.*, 2011), the shipping industry (see Benth *et al.*, 2015) , and the survival data (see Aalen and Gjessing, 2004). However, the classical O-U process is suitable to model the dataset for which the mean reversion level does not depend on time. Thus, Dehling *et al.* (2010) introduced a generalized O-U process for which the mean reversion level is time-dependent. Further, Dehling *et al.* (2014) proposed a model which can capture possible unconventional shocks as well as the seasonality trend. For further details about the impact of change-point on statistical analysis, we quote Lu and Lund (2007), Gombay (2010) and Robbins *et al.* (2011) among others.

Just recently, Nkurunziza and Zhang (2018) studied inference problem in generalized O-U with an unknown change-point when the drift parameter is suspected to satisfy some restrictions. To give another recent reference about inference problem in generalized O-U, we also quote Chen *et al.* (2017) and the references therein.

To the best of our knowledge, there is no study about inference problem in context of multivariate periodic mean-reverting stochastic with a possible change-point. Nevertheless, as discussed in Pigorsch and Stelzer (2009), it is important to capture the individual dynamics of the model as well as the correlation structure and effects across different financial assets in a financial market. In this thesis, we hope to fill this gap by proposing inference methods about the drift parameter matrix in context of multivariate generalized O-U with an unknown change-point. The proposed model can capture the correlations between different factors, the seasonality trend as well as the possible unconventional shocks. The proposed inference incorporates also uncertain prior information about the drift parameter matrix. The uncertain prior information is given in form of linear restriction binding the columns or the rows of the drift parameter matrix. Such a restriction includes a special case of the nonexistence of the change-point as well as the absence of the seasonality factor in context of correlated stochastic processes.

1.1 Main contributions of the thesis

In this section, we highlight the important contributions of the thesis. As compared to the findings in literature, we generalize in five ways the results in Dehling *et al.* (2010, 2014), Nkurunziza and Zhang (2018) and Chen *et al.* (2017). First, we consider inference problem in multi-dimensional context and we establish a more general result underlying the existence of the unrestricted estimator (UE) and the restricted estimator (RE) of the drift parameter. We also derive the UE and the RE. Second, we establish the joint asymptotic normality of the UE and the RE under the sequence of local alternatives. Third, we construct a test for testing the uncertain restriction. The proposed test is also useful for testing the absence of the change-point as well as the nonexistence of the seasonality factor. Fourth, we derive the asymptotic power of the proposed test and we prove that it is consistent. Fifth, inspired by the work in James and Stein (1961), we develop some shrinkage estimators (SEs) and we prove that SEs dominate the UE.

1.2 Organization of the thesis

This thesis contains seven chapters including the introdution and the conclusion. The rest of this thesis is organized as follows: In Chapter 2, we introduce the statistical model and regularity conditions. We also present in this chapter some preliminary results on the no change-point case. In Chapter 3, we derive the unrestricted maximum likelihood estimator (UMLE) and restricted maximum likelihood estimator (RMLE) in the case of one known change-point. We also derive in this chapter the joint asymptotic normality of the UMLE and the RMLE. In Chapter 4, we derive the UE and RE in the case of one unknown change-point as well as their joint asymptotic normality. We also construct in this chapter a test for testing the uncertain restriction, and we introduce the SEs. In Chapter 5, we compute the asymptotic distributional risks (ADR) for the UE, RE, and SEs, and then, we compare the relative performance based on their ADRs. In Chapter 6, we carry out a simulation study. Chapter 7 is the conclusion. The theoretical background is provided in the Appendix A, and some proofs of the main results are provided in the Appendix B.

Chapter 2

Preliminary results

In this chapter, we present the statistical model and some preliminary results. We also present the main assumptions used to establish the proposed method. The chapter is organized in three sections. In Section 2.1, we introduce the multivariate generalized Ornstein-Uhlenbeck processes as well as some notations. In Section 2.2, we present the case where no change-point is involved as our preliminary result, and in Section 2.3, we derive some asymptotic properties of this case.

2.1 Statistical model

In this section, we present the model of multivariate generalized Ornstein-Uhlenbeck processes with a possible change-point, and then, we introduce some mathematical notations. Let $\mathbb{I}_A$ denote the indicator function of the event A. For $\gamma = \phi T$ and $\phi \in (0, 1)$, the statistical model of interest is

$$dX_t = \left[(\mu_1\varphi(t) - A_1X_t)\mathbb{I}_{\{t\leq\gamma\}} + (\mu_2\varphi(t) - A_2X_t)\mathbb{I}_{\{t>\gamma\}}\right] dt + \Sigma^{1/2}dW_t, \tag{2.1}$$

with $0 \leq t \leq T$, and $\{W_t, t \geq 0\}$ is a standard d-dimensional Brownian motion, i.e.

$$W_t = \begin{bmatrix} W_1(t) & W_2(t) & W_3(t) & \ldots & W_d(t) \end{bmatrix}',$$

$\{X_t, t \geq 0\}$ is the corresponding d-dimensional stochastic process, i.e.

$$X_t = \begin{bmatrix} X_1(t) & X_2(t) & X_3(t) & \ldots & X_d(t) \end{bmatrix}',$$

$\varphi(t)$ is $\mathbb{R}^p$-valued function on $[0, T]$, i.e.

$$\varphi(t) = \begin{bmatrix} \varphi_1(t) & \varphi_2(t) & \varphi_3(t) & \ldots & \varphi_p(t) \end{bmatrix}',$$

$\mu_1 \in \mathbb{R}^{d \times p}$, $\mu_2 \in \mathbb{R}^{d \times p}$, $A_1 \in \mathbb{R}^{d \times d}$, $A_2 \in \mathbb{R}^{d \times d}$ are the parameters of interest, i.e.

$$\mu_1 = \begin{bmatrix} \mu_{11}^{(1)} & \mu_{12}^{(1)} & \mu_{13}^{(1)} & \ldots & \mu_{1p}^{(1)} \\ \mu_{21}^{(1)} & \mu_{22}^{(1)} & \mu_{23}^{(1)} & \ldots & \mu_{2p}^{(1)} \\ \mu_{31}^{(1)} & \mu_{32}^{(1)} & \mu_{33}^{(1)} & \ldots & \mu_{3p}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{d1}^{(1)} & \mu_{d2}^{(1)} & \mu_{d3}^{(1)} & \ldots & \mu_{dp}^{(1)} \end{bmatrix}, \mu_2 = \begin{bmatrix} \mu_{11}^{(2)} & \mu_{12}^{(2)} & \mu_{13}^{(2)} & \ldots & \mu_{1p}^{(2)} \\ \mu_{21}^{(2)} & \mu_{22}^{(2)} & \mu_{23}^{(2)} & \ldots & \mu_{2p}^{(2)} \\ \mu_{31}^{(2)} & \mu_{32}^{(2)} & \mu_{33}^{(2)} & \ldots & \mu_{3p}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{d1}^{(2)} & \mu_{d2}^{(2)} & \mu_{d3}^{(2)} & \ldots & \mu_{dp}^{(2)} \end{bmatrix},$$

$$A_1 = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \ldots & a_{1d}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} & \ldots & a_{2d}^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} & \ldots & a_{3d}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{d1}^{(1)} & a_{d2}^{(1)} & a_{d3}^{(1)} & \ldots & a_{dd}^{(1)} \end{bmatrix}, A_2 = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \ldots & a_{1d}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} & \ldots & a_{2d}^{(2)} \\ a_{31}^{(2)} & a_{32}^{(2)} & a_{33}^{(2)} & \ldots & a_{3d}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{d1}^{(2)} & a_{d2}^{(2)} & a_{d3}^{(2)} & \ldots & a_{dd}^{(2)} \end{bmatrix},$$

$\Sigma = \text{diag}\ (\sigma_1^2, \sigma_2^2, ..., \sigma_d^2)$ is the diffusion parameter matrix of the stochastic process, which is assumed to be known, i.e.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & ... & 0 \\ 0 & \sigma_2^2 & 0 & ... & 0 \\ 0 & 0 & \sigma_3^2 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & \sigma_d^2 \end{bmatrix}.$$

Further, A_1, A_2, and Σ are assumed to be positive definite matrices in the mean-reverting process. Let $\theta_1 = \left[\begin{array}{c:c} \mu_1 & A_1 \end{array}\right]$ and $\theta_2 = \left[\begin{array}{c:c} \mu_2 & A_2 \end{array}\right]$. The parameter of interest is a $d \times 2(p+d)$-matrix given by

$$\theta = \left[\begin{array}{c:c} \theta_1 & \theta_2 \end{array}\right]. \tag{2.2}$$

Further, let $(\mu_1\varphi(t) - A_1X_t)\mathbb{I}_{\{t\leq\gamma\}} + (\mu_2\varphi(t) - A_2X_t)\mathbb{I}_{\{t>\gamma\}} = S(\theta, t, X_t)$. The SDE in (2.1) can be rewritten as $dX_t = S(\theta, t, X_t)d_t + \Sigma^{1/2}dW_t,\ 0 \leq t \leq T$. Let I_p be a p-dimensional identity matrix. In some situations, there is a prior information about the parameters, and hence the parameters might be estimated under certain constraints. In particular, we consider the case where the parameters may satisfy the restrictions: $L_1\theta = d_1$ and $\theta L_2 = d_2$. This restriction motivates the testing problem

$$H_0 : L_1\theta = d_1, \quad \theta L_2 = d_2 \quad \text{versus} \quad H_1 : L_1\theta \neq d_1, \text{ or } \theta L_2 \neq d_2, \tag{2.3}$$

where $L_1 \in \mathbb{R}^{q\times d}$, $L_2 \in \mathbb{R}^{2(p+d)\times n}$ are known full-rank matrices with $n < 2(p+d)$, $q \leq d$, and $d_1 \in \mathbb{R}^{q\times 2(p+d)}$, $d_2 \in \mathbb{R}^{d\times n}$ are known matrices. Furthermore, it should be noted that for a suitable choice of L_1, L_2, d_1, d_2, the testing problem can cover many interseting special cases. For instance, by taking $L_2 = \left[\begin{array}{c:c} I_{(p+d)} & -I_{(p+d)} \end{array}\right]'$ and $d_2 = 0$, one can test the nonexistence of the change-point with additional restrictions on the parameters given as $L_1\theta = d_1$. For instance, let $L_1 = \begin{bmatrix} 1 & -1 & 0 & ... & 0 \end{bmatrix}$ and $d_1 = 0_{1\times(p+d)}$ to reflect the highly positive correlation that is expected between $X_1(t)$ and $X_2(t)$ while we are testing the existence of the change-point. As another example, setting $L_2 = \left[\begin{array}{c|c|c|c} I_p & 0 & -2I_p & 0 \\ \hline 0 & I_d & 0 & -I_d \end{array}\right]'$ and $d_2 = 0$ gives a testing problem with $\mu_2 = 2\mu_1$ and $A_1 = A_2$ (i.e., coefficients of the base functions doubled after the change-point while other components of θ remain the same) with additional restrictions on the parameters given as $L_1\theta = d_1$.

In order to derive the proposed method, we require the following conditions.

Assumption 1. *The distributrion of the initial value, X_0, of the SDE in* (2.1) *does not depend on the drift parameter θ. Further, X_0 is independent to $\{W_t : t \geq 0\}$ and* $\mathrm{E}(\|X_0\|_2^m) < \infty$, *for some $m \geq 2$.*

Assumption 2. *For any $T > 0$, the base function $\{\varphi_i(t), i = 1, 2, ..., p\}$ is Riemann-integrable on [0,T] and possesses*

(i) Periodicity: $\varphi_i(t + v) = \varphi_i(t)$, for all i =1,2,...,p, where v is the period.

(ii) Orthogonality in $L^2([0, v], \frac{1}{v}d\lambda)$: $\displaystyle\int_0^v \varphi(t)\varphi'(t)dt = vI_p$.

Remark 1. *Since the base function $\varphi(t)$ is bounded on $[0, T]$ and v-periodic, this implies that $\varphi(t)$ is bounded on $\mathbb{R}_+$.*

To introduce some notations, let $(\Omega, \mathfrak{F}, \mathrm{P})$ be a probability space where $\mathfrak{F}$ is $\mathfrak{S}$-field on the sample space Ω, and P is a probability measure. Further, let L^p denote the space of measurable p-integrable functions, for some $p \geq 1$. For mathematical convenience, we suppose that $\mathfrak{F}$ is complete. We also denote $\xrightarrow[T\to\infty]{d}$, $\xrightarrow[T\to\infty]{L^p}$, $\xrightarrow[T\to\infty]{P}$ the convergence in distribution, in L^p-space, and in probability, respectively, as T tends to infinity. Also, let $O_p(a(T))$ stand for a random quantity such that $O_p(a(T))a^{-1}(T)$

is bounded in probability. Further, we say that a stochastic process $\{Y_t, t \geq 0\}$ is L^p-bounded if there exists $K > 0$ such that $\mathrm{E}(|Y_t|^p) < K$, for all $t \geq 0$, for some $p \geq 1$. We denote $\mathrm{Tr}(A)$ to stand for the trace function of a matrix A, and we denote $\mathrm{Vec}(A)$ to stand for the vectorizing operator of a matrix A, i.e., $\mathrm{Vec}(A)$ is obtained by stacking the columns of the matrix A on top of one another starting from the leftmost column. We define $\|.\|_2$ and $\|.\|_F$ to be the Euclidean norm and Frobenius norm respectively. Next, we introduce the following two definitions.

Definition 1. *The $p \times q$ random matrix $\boldsymbol{X}$ is said to follow a matrix-variate normal distribution with the $p \times q$ mean matrix M and the $pq \times pq$ covariance matrix Σ if* $\mathrm{Vec}(\boldsymbol{X}) \sim \mathcal{N}_{pq}(\mathrm{Vec}(M), \Sigma)$. *We denote it as $\boldsymbol{X} \sim \mathcal{N}_{p\times q}(M, \Sigma)$.*

Definition 2. *The matrix $\boldsymbol{W}$: $p \times p$ is said to be Wishart distributed if and only if $\boldsymbol{W} = \boldsymbol{X}\boldsymbol{X}'$, where $\boldsymbol{X} \sim \mathcal{N}_{p\times n}(\mu, I \otimes \Sigma), \Sigma \geq 0$. If $\mu = 0$, we have a central Wishart distribution which will be denoted by $\boldsymbol{W} \sim W_n(p, \Sigma)$, and if $\mu \neq 0$, we have a non-central Wishart distribution which will be denoted as $W_n(p, \Sigma, \Delta)$, where $\Delta = \mu\mu'$.*

2.2 Preliminary results: No change-point case

In this section, we study the case where there is no change-point. This case is studied as a preliminary step in order to facilitate the understanding of the proposed method. In no change-point case, the SDE in (2.1) can be written as

$$dX_t = (\mu\varphi(t) - AX_t)dt + \Sigma^{1/2} dW_t, \tag{2.4}$$

with $0 \leq t \leq T$, and $\mu \in \mathbb{R}^{d\times p}$, $A \in \mathbb{R}^{d\times d}$. In case of the statistical model in (2.4), the parameter of interest is $\theta = \left[\begin{array}{c:c} \mu & A \end{array}\right] \in \mathbb{R}^{d\times(p+d)}$. Thus, the drift coefficient is

$S(\theta, t, X_t) = \mu\varphi(t) - AX_t$. The following proposition shows that the SDE in (2.4) admits a unique and strong solution which is L^2-bounded on $[0, T]$.

Proposition 2.1. *Suppose that Assumption 1-2 hold. Then, the SDE in* (2.4) *admits a strong and unique solution that is L^2-bounded on $[0, T]$, i.e.* $\sup_{0 \leq t \leq T} \mathrm{E}(\|X_t\|_2^2) < \infty$.

The proof of this proposition is given in the Appendix B where a more general case is considered. Further, below we prove that $\{X_t, t \geq 0\}$ is uniformly L^2-bounded.

Remark 2. *From Proposition* 2.1*, one concludes that*

$$\mathrm{P}\left(\int_0^T \|S(\theta, t, X_t)\|_2^2 dt < \infty\right) = 1,$$

for all $0 < T < \infty$, for all $\theta \in \Theta$. This is a sufficient condition for the existence of the Radon-Nikodym derivative of a stochastic process.

Proposition 2.2. *The trajectory of the SDE in* (2.4) *is given by* $X_t = e^{-At}X_0 + e^{-At}\int_0^t e^{As}\mu\varphi(s)ds + e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s$. *Further,* $\sup_{t \geq 0} \mathrm{E}(\|X_t\|_2^2) < \infty$.

Proof. Let $g(x, t) = e^{At}x$, and apply Itô's formula to $g(x, t)$ with the process specified in (2.4), we get

$$dg(X_t, t) = e^{At}dX_t + e^{At}AX_t dt = e^{At}(\mu\varphi(t)dt + \Sigma^{1/2}dW_t). \tag{2.5}$$

Taking integral from 0 to t on both sides of (2.5), we get

$$e^{At}X_t = X_0 + \int_0^t e^{As}\mu\varphi(s)ds + \int_0^t e^{As}\Sigma^{1/2}dW_s. \tag{2.6}$$

Note that e^{At} is always invertible with $(e^{At})^{-1} = e^{-At}$, then mutiplying by e^{-At} on both sides of (2.6), we get

$$X_t = e^{-At}X_0 + e^{-At}\int_0^t e^{As}\mu\varphi(s)ds + e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s. \tag{2.7}$$

Further, using $(a+b+c)^2 \leq 3a^2+3b^2+3c^2$, we get

$$\mathrm{E}[\|X_t\|_2^2] \leq 3\|e^{-At}\|_F^2 \mathrm{E}(\|X_0\|_2^2) + 3\mathrm{E}\left(\left\|\int_0^t e^{-A(t-s)}\mu\varphi(s)ds\right\|_2^2\right) + 3\mathrm{E}\left(\left\|\int_0^t e^{-A(t-s)}\Sigma^{1/2}dW_s\right\|_2^2\right).$$

Then, by Itô's isometry, this gives

$$\mathrm{E}\left(\left\|\int_0^t e^{-A(t-s)}\Sigma^{1/2}dW_s\right\|_2^2\right) = \int_0^t \|e^{-A(t-s)}\Sigma^{1/2}\|_F^2 ds \leq \|\Sigma^{1/2}\|_F^2 \int_0^t \|e^{-A(t-s)}\|_F^2 ds.$$

Therefore, from Assumption 1, Proposition $A.3$, Remark 1, let $\|\mu\varphi(s)\|_2^2 \leq K_{\mu,\varphi}$, $\mathrm{E}(\|X_0\|_2^2) \leq K_0$, and λ_1 be the smallest eigenvalue of $A'+A$, we get

$$\mathrm{E}[\|X_t\|_2^2] \leq 3de^{-\lambda_1 t}K_0 + 3\left(K_{\mu,\varphi} + \|\Sigma^{1/2}\|_F^2\right)\left(\frac{d-de^{-\lambda_1 t}}{\lambda_1}\right),$$

which implies that $\sup_{t\geq 0} \mathrm{E}(\|X_t\|_2^2) < \infty$, this completes the proof. $\square$

In the sequel, let

$$X_t = e^{-At}X_0 + h(t) + Z_t, \quad 0 \leq t \leq T, \tag{2.8}$$

where

$$h(t) = e^{-At}\int_0^t e^{As}\mu\varphi(s)ds, \quad Z_t = e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s. \tag{2.9}$$

Notice that the process $\{X_t, t \geq 0\}$ is not stationary. Thus, to apply some limiting theorem such as Birkhoff's Ergodic Theorem, we introduce an auxiliary process

$$\widetilde{X}_t = \widetilde{h}(t) + \widetilde{Z}_t, \quad 0 \leq t \leq T, \tag{2.10}$$

where

$$\widetilde{h}(t) = e^{-At}\int_{-\infty}^t e^{As}\mu\varphi(s)ds, \quad \widetilde{Z}_t = e^{-At}\int_{-\infty}^t e^{As}\Sigma^{1/2}d\widetilde{W}_s, \tag{2.11}$$

where $\{\widetilde{W}_s, s \in \mathbb{R}\}$ denotes a d-dimensional bilateral Brownian motion, i.e.

$$\widetilde{W}_s = W_s^{(1)}\mathbb{I}_{\{s\in\mathbb{R}_+\}} + W_{-s}^{(2)}\mathbb{I}_{\{s\in\mathbb{R}_-\}}, \tag{2.12}$$

where $\{W_s^{(1)}, s \geq 0\}$ and $\{W_s^{(2)}, s \geq 0\}$ are two independent d-dimensional standard Brownian motions. Below, we prove that, for each $t \in [0, 1]$, $\{\widetilde{X}_{k+t},\ k \in \mathbb{N}_0\}$ is a stationary and ergodic process. As an intermediate result, we establish the following two propositions.

Proposition 2.3. *Suppose that Assumptions 1-2 hold. Then, for $t \in [0, 1]$, $k \in \mathbb{N}_0$, $\mathrm{E}(\widetilde{Z}_t\widetilde{Z}'_{t+k})$ does not depend on t.*

Proposition 2.4. *Suppose that Assumption 1-2 hold. Then, for $t \in [0, 1]$, the process $\{\widetilde{X}_{k+t},\ k \in \mathbb{N}_0\}$ is Gaussian.*

The proofs of these two propositions are given in Appendix B. By using Propositions 2.3-2.4, we prove the following proposition which shows that the auxiliary process $\{\widetilde{X}_{k+t},\ k \in \mathbb{N}_0\}$ is stationary and ergodic.

Proposition 2.5. *Suppose that Assumptions 1-2 hold. Then for $t \in [0, 1]$, the sequence of random vectors $\{\widetilde{X}_{k+t}\}_{k\in\mathbb{N}_0}$ is stationary and ergodic.*

The proof is given in the Appendix B.

2.3 Asymptotic properties

In this section, we provide some asymptotic properties of the process defined in (2.4). Also, in the rest of the thesis, we assume without loss of generality that the period $v = 1$ for the orthogonal set $\{\varphi_i(t), i = 1, 2, ..., p\}$.

Lemma 2.1. *Suppose that Assumptions* 1-2 *hold, let* $\phi_0 \in [0,1]$, *then*

$$\frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)X_t' dt \xrightarrow[T\to\infty]{P} 0.$$

Proof. It is sufficient to prove that $\left\| \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)X_t' dt \right\|_F \xrightarrow[T\to\infty]{L^1} 0.$

Note that

$$\begin{aligned}
&\left\| \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)X_t' dt \right\|_F = \left\| \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)(\widetilde{X}_t' - X_t') dt \right\|_F \\
&\le \frac{1}{T}\int_0^{\phi_0 T} \left\| \varphi(t)(\widetilde{X}_t' - X_t') \right\|_F dt \le \frac{1}{T}\int_0^{\phi_0 T} \|\varphi(t)\|_2 \left\| \widetilde{X}_t - X_t \right\|_2 dt.
\end{aligned}$$

According to the Remark 1, let $\|\varphi(t)\|_2 \le K_\varphi$ for all t, we have

$$\left\| \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} \varphi(t)X_t' dt \right\|_F \le \frac{K_\varphi}{T}\int_0^{\phi_0 T} \left\| \widetilde{X}_t - X_t \right\|_2 dt. \tag{2.13}$$

Note that from (2.8)-(2.11), we have

$$\begin{aligned}
\left\| \widetilde{X}_t - X_t \right\|_2 &= \left\| \widetilde{h}(t) + \widetilde{Z}_t - e^{-At}X_0 - h(t) - Z_t \right\|_2 \\
&= \left\| e^{-At}\int_{-\infty}^0 e^{As}\mu\varphi(s)ds + e^{-At}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)} - e^{-At}X_0 \right\|_2 \\
&= \left\| e^{-At}\int_{-\infty}^0 e^{As}\mu\varphi(s)ds + e^{-At}\int_0^{\infty} e^{-As}\Sigma^{1/2}dW_s^{(2)} - e^{-At}X_0 \right\|_2 \\
&\le \|e^{-At}\|_F \left\| \int_{-\infty}^0 e^{As}\mu\varphi(s)ds + \int_0^{\infty} e^{-As}\Sigma^{1/2}dW_s^{(2)} - X_0 \right\|_2.
\end{aligned} \tag{2.14}$$

Since A is positive definite, let λ_1 be the smallest eigenvalue of $A' + A$, then by Proposition $A.3$, we have

$$\begin{aligned}
\int_0^{\phi_0 T} \left\| \widetilde{X}_t - X_t \right\|_2 dt &\le \int_0^{\phi_0 T} \|e^{-At}\|_F \left\| \int_{-\infty}^0 e^{As}\mu\varphi(s)ds + \int_0^{\infty} e^{-As}\Sigma^{1/2}dW_s^{(2)} - X_0 \right\|_2 dt \\
&\le \left\| \int_{-\infty}^0 e^{As}\mu\varphi(s)ds + \int_0^{\infty} e^{-As}\Sigma^{1/2}dW_s^{(2)} - X_0 \right\|_2 \int_0^{\phi_0 T} \sqrt{de^{-t\lambda_1}}dt \\
&= \left\| \int_{-\infty}^0 e^{As}\mu\varphi(s)ds + \int_0^{\infty} e^{-As}\Sigma^{1/2}dW_s^{(2)} - X_0 \right\|_2 \frac{2\sqrt{d}}{\lambda_1}(1 - e^{-\frac{\lambda_1\phi_0}{2}T}) \\
&\le \left(\left\| \int_{-\infty}^0 e^{As}\mu\varphi(s)ds \right\|_2 + \|X_0\|_2 \right) \frac{2\sqrt{d}}{\lambda_1}(1 - e^{-\frac{\lambda_1\phi_0}{2}T}) \\
&\quad + \left\| \int_0^{\infty} e^{-As}\Sigma^{1/2}dW_s^{(2)} \right\|_2 \frac{2\sqrt{d}}{\lambda_1}(1 - e^{-\frac{\lambda_1\phi_0}{2}T}).
\end{aligned} \tag{2.15}$$

Now, by Remark 1 and Assumption 1, we can claim that $\|\mu\varphi(t)\|_2 \leq K_{\mu,\varphi}$ for all t and $\mathrm{E}(\|X_0\|_2) \leq K_0 < \infty$. Therefore,

$$
\begin{aligned}
&\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T}\varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T}\varphi(t)X_t' dt\right\|_F\right) \\
&\leq \frac{K_\varphi}{T}\left(\mathrm{E}\left(\left\|\int_{-\infty}^0 e^{As}\mu\varphi(s)ds\right\|_2\right) + \mathrm{E}\left(\|X_0\|_2\right)\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}) \\
&+ \frac{K_\varphi}{T}\mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}).
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T}\varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T}\varphi(t)X_t' dt\right\|_F\right) \\
&\leq \frac{K_\varphi}{T}\left(K_{\mu,\varphi}\left(\int_{-\infty}^0 \left\|e^{As}\right\|_F ds\right) + K_0\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}) \\
&+ \frac{K_\varphi}{T}\mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}) \\
&\leq \frac{K_\varphi}{T}\left(K_{\mu,\varphi}\frac{2\sqrt{d}}{\lambda_1} + K_0\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}) \\
&+ \frac{K_\varphi}{T}\mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}).
\end{aligned}
$$

Further, let $K_\varphi\left(K_{\mu,\varphi}\frac{2\sqrt{d}}{\lambda_1} + K_0\right)\frac{2\sqrt{d}}{\lambda_1} = K_1$, we have

$$
\begin{aligned}
&\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T}\varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T}\varphi(t)X_t' dt\right\|_F\right) \\
&\leq \frac{K_1}{T}(1-e^{-\frac{\lambda_1\phi_0}{2}T}) + \frac{K_\varphi}{T}\mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2\right)\frac{2\sqrt{d}}{\lambda_1}(1-e^{-\frac{\lambda_1\phi_0}{2}T}).
\end{aligned}
$$

From the proof of Proposition 2.5, we know that

$$
\mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2^2\right) \leq \frac{d\left\|\Sigma^{1/2}\right\|_F^2}{\lambda_1}. \tag{2.16}
$$

Therefore, by Cauchy Schwarz Inequality, we get

$$
\mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2\right) \leq \mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2^2\right)^{\frac{1}{2}} \leq \left(\frac{d\left\|\Sigma^{1/2}\right\|_F^2}{\lambda_1}\right)^{\frac{1}{2}},
$$

also, let $K_\varphi \left(\frac{d\left\|\Sigma^{1/2}\right\|_F^2}{\lambda_1}\right)^{\frac{1}{2}} \frac{2\sqrt{d}}{\lambda_1} = K_2$, we have

$$\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} \varphi(t) X_t' dt\right\|_F\right) \leq \frac{K_1 + K_2}{T}(1 - e^{-\frac{\lambda_1 \phi_0}{2} T}).$$

Therefore

$$\lim_{T\to\infty} \mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} \varphi(t) X_t' dt\right\|_F\right) = 0,$$

which completes the proof. □

Lemma 2.2. *Suppose that the conditions for Lemma* 2.1 *hold, then*

$$\frac{1}{T}\int_0^{\phi_0 T} \widetilde{X}_t \widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} X_t X_t' dt \xrightarrow[T\to\infty]{P} 0.$$

Proof. It is sufficient to prove that $\left\|\frac{1}{T}\int_0^{\phi_0 T} \widetilde{X}_t \widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} X_t X_t' dt\right\|_F \xrightarrow[T\to\infty]{L^1} 0$. Note that

$$\begin{aligned}\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T} \widetilde{X}_t \widetilde{X}_t' dt - \frac{1}{T}\int_0^{\phi_0 T} X_t X_t' dt\right\|_F\right) &= \mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T} (\widetilde{X}_t \widetilde{X}_t' - X_t X_t') dt\right\|_F\right) \\ &\leq \frac{1}{T}\int_0^{\phi_0 T} \mathrm{E}\left(\left\|\widetilde{X}_t \widetilde{X}_t' - X_t X_t'\right\|_F\right) dt.\end{aligned}$$

Notice that $\widetilde{X}_t \widetilde{X}_t' - X_t X_t' = \widetilde{X}_t(\widetilde{X}_t' - X_t') + (\widetilde{X}_t - X_t) X_t'$, and then, by Triangle Inequality, we get

$$\begin{aligned}\frac{1}{T}\int_0^{\phi_0 T} \mathrm{E}\left(\left\|\widetilde{X}_t \widetilde{X}_t' - X_t X_t'\right\|_F\right) dt &= \frac{1}{T}\int_0^{\phi_0 T} \mathrm{E}\left(\left\|\widetilde{X}_t(\widetilde{X}_t' - X_t') + (\widetilde{X}_t - X_t) X_t'\right\|_F\right) dt \\ &\leq \frac{1}{T}\int_0^{\phi_0 T} \mathrm{E}\left(\left\|\widetilde{X}_t(\widetilde{X}_t' - X_t')\right\|_F + \left\|(\widetilde{X}_t - X_t) X_t'\right\|_F\right) dt.\end{aligned}$$

By Cauchy Schwarz Inequality, we have

$$\mathrm{E}\left(\left\|\widetilde{X}_t(\widetilde{X}_t' - X_t')\right\|_F\right) \leq \mathrm{E}\left(\left\|\widetilde{X}_t\right\|_2^2\right)^{1/2} \mathrm{E}\left(\left\|\widetilde{X}_t - X_t\right\|_2^2\right)^{1/2},$$

$$\mathrm{E}\left(\left\|(\widetilde{X}_t - X_t) X_t'\right\|_F\right) \leq \mathrm{E}\left(\left\|\widetilde{X}_t - X_t\right\|_2^2\right)^{1/2} \mathrm{E}\left(\|X_t\|_2^2\right)^{1/2}.$$

Since $\mathrm{E}\left(\|X_t\|_2^2\right) < \infty$ as we showed in Proposition 2.2, let $\mathrm{E}(\|X_t\|)_2^2 \le K_x < \infty$. Also based on the proof of Proposition 2.5 (B.10)-(B.17), we have

$$\mathrm{E}(\|\widetilde{X}_t\|_2^2) \le 2\left(\left(\frac{2K_{\mu,\varphi}d}{\lambda_1}\right)^2 + \frac{d^2\|\Sigma^{1/2}\|_F^2}{\lambda_1}\right) < \infty.$$

Let $\sup_{t\ge 0}\{\mathrm{E}(\|X_t\|_2^2)^{\frac{1}{2}}, \mathrm{E}(\|\widetilde{X}_t\|_2^2)^{\frac{1}{2}}\} \le K < \infty$, we get

$$\frac{1}{T}\int_0^{\phi_0 T} \mathrm{E}(\|\widetilde{X}_t(\widetilde{X}_t' - X_t')\|_F + \|(\widetilde{X}_t - X_t)X_t'\|_F)dt \le \frac{2K}{T}\int_0^{\phi_0 T} \mathrm{E}(\|(\widetilde{X}_t - X_t)\|_2^2)^{1/2}dt.$$

By using $(a+b+c)^2 \le 3a^2 + 3b^2 + 3c^2$, we have $\mathrm{E}(\|(\widetilde{X}_t - X_t)\|_2^2)$ is equal to

$$\begin{aligned}
&\mathrm{E}(\|\widetilde{h}(t) + \widetilde{Z}_t - e^{-At}X_0 - h(t) - Z_t\|_2^2) \\
&\le \|e^{-At}\|_F^2 E\left(\left\|\int_{-\infty}^0 e^{As}\mu\varphi(s)ds + \int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)} - X_0\right\|_2^2\right) \\
&\le 3\|e^{-At}\|_F^2\left(\left\|\int_{-\infty}^0 e^{As}\mu\varphi(s)ds\right\|_2^2 + \mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2^2\right) + \mathrm{E}(\|X_0\|_2^2)\right).
\end{aligned}$$

Further, let $\|\mu\varphi(t)\|_2 \le K_{\mu,\varphi}$ for all t. Also, by Assumption 1, there exists $K_0 > 0$ such that $\mathrm{E}(\|X_0\|_2^2) \le K_0 < \infty$. Then, by Proposition $A.3$ and (2.16), we have

$$\begin{aligned}
&3\|e^{-At}\|_F^2\left(\left\|\int_{-\infty}^0 e^{As}\mu\varphi(s)ds\right\|_2^2 + \mathrm{E}\left(\left\|\int_0^\infty e^{-As}\Sigma^{1/2}dW_s^{(2)}\right\|_2^2\right) + \mathrm{E}(\|X_0\|_2^2)\right) \\
&\le 3de^{-\lambda_1 t}\left(\left(K_{\mu,\varphi}\frac{2\sqrt{d}}{\lambda_1}\right)^2 + K_0 + \frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1}\right).
\end{aligned}$$

Then, set $3d\left(\left(K_{\mu,\varphi}\frac{2\sqrt{d}}{\lambda_1}\right)^2 + K_0 + \frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1}\right) = K_1$. We have

$$\mathrm{E}(\|(\widetilde{X}_t - X_t)\|_2^2) \le K_1 e^{-\lambda_1 t}.$$

Therefore

$$\frac{2K}{T}\int_0^{\phi_0 T} \mathrm{E}(\|(\widetilde{X}_t - X_t)\|_2^2)^{1/2}dt \le \frac{2KK_1^{1/2}}{T}\int_0^{\phi_0 T} e^{-\frac{\lambda_1}{2}t}dt \le \frac{4KK_1^{1/2}}{\lambda_1 T}(1 - e^{-\frac{\lambda_1\phi_0}{2}T}).$$

Since $\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T}(\widetilde{X}_t\widetilde{X}_t' - X_tX_t')dt\right\|_F\right) \leq \frac{2K}{T}\int_0^{\phi_0 T}\mathrm{E}(\|(\widetilde{X}_t - X_t)\|_2^2)^{1/2}dt$, we get

$$\lim_{T\to\infty}\mathrm{E}\left(\left\|\frac{1}{T}\int_0^{\phi_0 T}(\widetilde{X}_t\widetilde{X}_t' - X_tX_t')dt\right\|_F\right) = 0,$$

which completes the proof. □

Lemma 2.3. *Suppose that the conditions for Lemma* 2.1 *hold, then*

$$\frac{1}{T}\int_{\phi_0 T}^{T}\varphi(t)\widetilde{X}_t'dt - \frac{1}{T}\int_{\phi_0 T}^{T}\varphi(t)X_t'dt \xrightarrow[T\to\infty]{P} 0,$$
$$\frac{1}{T}\int_{\phi_0 T}^{T}\widetilde{X}_t\widetilde{X}_t'dt - \frac{1}{T}\int_{\phi_0 T}^{T}X_tX_t'dt \xrightarrow[T\to\infty]{P} 0.$$

The proof of the first statement follows directly from Lemma 2.1. The proof of the second statement follows directly from Lemma 2.2

Proposition 2.6. *Suppose that the conditions for Lemma* 2.1 *hold, then*

$$\frac{1}{T}\int_0^{\phi_0 T}\varphi(t)X_t'dt \xrightarrow[T\to\infty]{P} \phi_0\int_0^1\varphi(t)\widetilde{h}'(t)dt.$$

The proof is provided in the Appendix B.

Now, let

$$V(k) = \mathrm{E}(\widetilde{Z}_0\widetilde{Z}_k'). \tag{2.17}$$

Proposition 2.7. *Suppose that A is a positive definite matrix and Σ is a symmetric and positive definite matrix. Then $V(0)$ is a positive definite matrix.*

The proof follows directly from algebraic computations.

Proposition 2.8. *Suppose that the conditions for Proposition* 2.6 *hold, then*

$$\frac{1}{T}\int_0^{\phi_0 T}X_tX_t'dt \xrightarrow[T\to\infty]{P} \phi_0\left\{\int_0^1\widetilde{h}(t)\widetilde{h}'(t)dt + V(0)\right\}.$$

The proof is provided in the Appendix B.

Proposition 2.9. *Suppose that the conditions for Proposition* 2.6 *hold, then*

$$\frac{1}{T}\int_{\phi_0 T}^{T} \varphi(t)X_t' dt \xrightarrow[T\to\infty]{P} (1-\phi_0)\int_0^1 \varphi(t)\widetilde{h}'(t)dt,$$
$$\frac{1}{T}\int_{\phi_0 T}^{T} X_t X_t' dt \xrightarrow[T\to\infty]{P} (1-\phi_0)\left\{\int_0^1 \widetilde{h}(t)\widetilde{h}'(t)dt + V(0)\right\}.$$

The proof of the first statement follows directly from Proposition 2.6 and the proof of the second statement follows directly from Proposition 2.8. Based on the Propositions 2.6-2.9, we have the following results, which are crucial in the rest of the Thesis. For $\phi_0 \in [0,1]$ and $\gamma = \phi_0 T$, let us define

$$O_\gamma = \begin{bmatrix} \int_0^{\phi_0 T} \varphi(t)\varphi'(t)dt & -\int_0^{\phi_0 T} \varphi(t)X_t' dt \\ -\int_0^{\phi_0 T} X_t \varphi'(t)dt & \int_0^{\phi_0 T} X_t X_t' dt \end{bmatrix}, \tag{2.18}$$

and let

$$\Sigma_a = \begin{bmatrix} I_p & -\int_0^1 \varphi(t)\widetilde{h}'(t)dt \\ -\int_0^1 \widetilde{h}(t)\varphi'(t)dt & \int_0^1 \widetilde{h}(t)\widetilde{h}'(t)dt + V(0) \end{bmatrix}. \tag{2.19}$$

Proposition 2.10. *Suppose that the conditions for Proposition* 2.8 *hold, then*

$$\frac{1}{T}O_\gamma \xrightarrow[T\to\infty]{P} \phi_0 \Sigma_a.$$

Proof. From Proposition 2.6 and Proposition 2.8, it is sufficient to show that

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\varphi'(t)dt = \phi_0 I_p.$$

Based on Assumption 2, we have

$$\begin{aligned}\frac{1}{T}\int_0^{\phi_0 T} \varphi(t)\varphi'(t)dt &= \frac{1}{T}\int_0^{\lfloor \phi_0 T\rfloor} \varphi(t)\varphi'(t)dt + \frac{1}{T}\int_{\lfloor \phi_0 T\rfloor}^{\phi_0 T} \varphi(t)\varphi'(t)dt \\ &= \frac{1}{T}\lfloor \phi_0 T\rfloor I_p + \frac{1}{T}\int_{\lfloor \phi_0 T\rfloor}^{\phi_0 T} \varphi(t)\varphi'(t)dt. \end{aligned} \tag{2.20}$$

Moreover

$$\left\|\int_{\lfloor\phi_0 T\rfloor}^{\phi_0 T}\varphi(t)\varphi'(t)dt\right\|_F \leq \int_{\lfloor\phi_0 T\rfloor}^{\phi_0 T}\|\varphi(t)\varphi'(t)\|_F dt \leq \int_{\lfloor\phi_0 T\rfloor}^{\lfloor\phi_0 T\rfloor+1}\|\varphi(t)\varphi'(t)\|_F dt$$
$$= \int_0^1 \|\varphi(t)\varphi'(t)\|_F dt = p.$$

Therefore

$$\lim_{T\to\infty}\frac{1}{T}\int_{\lfloor\phi_0 T\rfloor}^{\phi_0 T}\varphi(t)\varphi'(t)dt = 0. \tag{2.21}$$

Also, we have $0 \leq \phi_0 T - \lfloor\phi_0 T\rfloor \leq \lfloor\phi_0 T\rfloor + 1 - \lfloor\phi_0 T\rfloor$, then $0 \leq \frac{1}{T}(\phi_0 T - \lfloor\phi_0 T\rfloor) \leq \frac{1}{T}$, and then

$$\lim_{T\to\infty}\frac{\lfloor\phi_0 T\rfloor}{T} = \phi_0. \tag{2.22}$$

Therefore, by (2.20), (2.21), and (2.22), we get

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{\phi_0 T}\varphi(t)\varphi'(t)dt = \phi_0 I_p.$$

Combining Proposition 2.6 and Proposition 2.8, we complete the proof. □

Now, let us define

$$O_{\gamma,T} = O_T - O_\gamma = \begin{bmatrix} \int_{\phi_0 T}^{T}\varphi(t)\varphi'(t)dt & -\int_{\phi_0 T}^{T}\varphi(t)X_t'dt \\ -\int_{\phi_0 T}^{T}X_t\varphi'(t)dt & \int_{\phi_0 T}^{T}X_t X_t'dt \end{bmatrix}. \tag{2.23}$$

Proposition 2.11. *Suppose that the conditions for Proposition* 2.10 *hold, then*

$$\frac{1}{T}O_{\gamma,T} \xrightarrow[T\to\infty]{P} (1-\phi_0)\Sigma_a.$$

From Proposition 2.9, the proof is similar to that of Proposition 2.10.

Remark 3. *It is possible to derive stronger results than the ones given by Propositions* 2.10 *and* 2.11*. In particular, one can prove that* $\frac{1}{T}O_\gamma$ *and* $\frac{1}{T}O_{\gamma,T}$ *converge almost* *surely. For more details, we refer to Nkurunziza and Shen (2018). Nevertheless, the results given by Propositions* 2.10 *and* 2.11 *are sufficient for deriving the main results of this thesis.*

Chapter 3

Estimation method: the known change-point case

In this chapter, we present an estimation method in the case of a possible change-point. We assume that the change point $\gamma = \phi T$ is known. The chapter is subdivided into two sections. In Section 3.1, we derive the unrestricted maximum likelihood estimator (UMLE) and the restricted maximum likelihood estimator (RMLE). In Section 3.2, we derive the joint asymptotic normality of the UMLE and RMLE.

3.1 UMLE and RMLE

In this section, we derive the UMLE and the RMLE. In particular, the RMLE is obtained by using the method of Lagrange multipliers. To introduce some notations, let $\gamma = \phi T$ with $\phi \in (0, 1)$. Further, define

$$P_\gamma = \begin{bmatrix} \int_0^\gamma \varphi(t) dX_t' \\ -\int_0^\gamma X_t dX_t' \end{bmatrix} \in \mathbb{R}^{(p+d)\times d}, \; P_{\gamma,T} = \begin{bmatrix} \int_\gamma^T \varphi(t) dX_t' \\ -\int_\gamma^T X_t dX_t' \end{bmatrix} \in \mathbb{R}^{(p+d)\times d}, \tag{3.1}$$

and

$$Q_\gamma = \begin{bmatrix} \int_0^\gamma \varphi(t)\varphi'(t)dt & -\int_0^\gamma \varphi(t)X_t'dt \\ -\int_0^\gamma X_t\varphi'(t)dt & \int_0^\gamma X_tX_t'dt \end{bmatrix} \in \mathbb{R}^{(p+d)\times(p+d)}, \tag{3.2}$$

$$Q_{\gamma,T} = \begin{bmatrix} \int_\gamma^T \varphi(t)\varphi'(t)dt & -\int_\gamma^T \varphi(t)X_t'dt \\ -\int_\gamma^T X_t\varphi'(t)dt & \int_\gamma^T X_tX_t'dt \end{bmatrix} \in \mathbb{R}^{(p+d)\times(p+d)}. \tag{3.3}$$

Now define

$$P(\phi) = \left[\begin{array}{c:c} P_\gamma' & P_{\gamma,T}' \end{array}\right] \in \mathbb{R}^{d\times 2(p+d)}, \tag{3.4}$$

and

$$Q(\phi) = \begin{bmatrix} Q_\gamma & 0_{p+d} \\ 0_{p+d} & Q_{\gamma,T} \end{bmatrix} \in \mathbb{R}^{2(p+d)\times 2(p+d)}. \tag{3.5}$$

Proposition 3.1. *Suppose that the Assumptions 1-2 hold, then the likelihood function is given by* $L(\theta; X_{[0,T]}) = \exp\left[\operatorname{Tr}(\Sigma^{-1}\theta P'(\phi)) - \frac{1}{2}\operatorname{Tr}(\Sigma^{-1}\theta Q(\phi)\theta')\right]$.

Proof. By the Proposition 2.1 and Remark 2, one can apply Theorem 7.7 in Liptser and Shiryayev (2001). Thus, by this theorem, the Radon-Nikodym derivative of the measure induced by the SDE in (2.1) exists. Let $L(\theta; X_{[0,T]})$ be the likelihood function induced by the probability measure of the SDE in (2.1). Then,

$$L(\theta; X_{[0,T]}) = \exp\left\{\operatorname{Tr}\left[\Sigma^{-1}\int_0^T S(\theta,t,X_t)dX_t'\right] - \frac{1}{2}\operatorname{Tr}\left[\Sigma^{-1}\int_0^T S(\theta,t,X_t)S'(\theta,t,X_t)dt\right]\right\}.$$

Note that Q_γ and $Q_{\gamma,T}$ are real symmetric matrices. Further, since $\theta = \left[\begin{array}{c:c} \theta_1 & \theta_2 \end{array}\right]$ with $\theta_1 = \left[\begin{array}{c:c} \mu_1 & A_1 \end{array}\right]$ and $\theta_2 = \left[\begin{array}{c:c} \mu_2 & A_2 \end{array}\right]$, we have

$$\begin{aligned} \int_0^T S(\theta,t,X_t)dX_t' &= \int_0^\gamma (\mu_1\varphi(t) - A_1X_t)dX_t' + \int_\gamma^T (\mu_2\varphi(t) - A_2X_t)dX_t' \\ &= \theta_1 P_\gamma + \theta_2 P_{\gamma,T}. \end{aligned} \tag{3.6}$$

Note that $\mathbb{I}_{(t\leq\gamma)}\mathbb{I}_{\{t>\gamma\}}=0$ for all t, then we have

$$\begin{aligned}
&\int_0^T [(\mu_1\varphi(t)-A_1X_t)\mathbb{I}_{\{t\leq\gamma\}}+(\mu_2\varphi(t)-A_2X_t)\mathbb{I}_{\{t>\gamma\}}] \\
&\times [(\mu_1\varphi(t)-A_1X_t)\mathbb{I}_{\{t\leq\gamma\}}+(\mu_2\varphi(t)-A_2X_t)\mathbb{I}_{\{t>\gamma\}}]'dt \\
&=\int_0^\gamma (\mu_1\varphi(t)-A_1X_t)(\mu_1\varphi(t)-A_1X_t)'dt \\
&+\int_\gamma^T (\mu_2\varphi(t)-A_2X_t)(\mu_2\varphi(t)-A_2X_t)'dt.
\end{aligned}$$

This gives

$$\int_0^T S(\theta,t,X_t)S'(\theta,t,X_t)dt=\theta_1Q_\gamma\theta_1'+\theta_2Q_{\gamma,T}\theta_2'. \tag{3.7}$$

Combining (3.6) and (3.7), the likelihood function can be rewritten as

$$L(\theta;X_{[0,T]})=\exp\left\{\mathrm{Tr}\left[\Sigma^{-1}(\theta_1P_\gamma+\theta_2P_{\gamma,T})\right]-\frac{1}{2}\mathrm{Tr}\left[\Sigma^{-1}(\theta_1Q_\gamma\theta_1'+\theta_2Q_{\gamma,T}\theta_2')\right]\right\}.$$

Note that Q is a real symmetric matrix since Q_γ and $Q_{\gamma,T}$ are real symmetric matrices. Then, the likelihood function is

$$L(\theta;X_{[0,T]})=\exp\left[\mathrm{Tr}(\Sigma^{-1}\theta P'(\phi))-\frac{1}{2}\mathrm{Tr}(\Sigma^{-1}\theta Q(\phi)\theta')\right], \tag{3.8}$$

this completes the proof. $\square$

From Proposition 3.1, the log-likelihood function is

$$l(\theta;X_{[0,T]})=\ln(L(\theta;X_{[0,T]}))=\mathrm{Tr}(\Sigma^{-1}\theta P'(\phi))-\frac{1}{2}\mathrm{Tr}(\Sigma^{-1}\theta Q(\phi)\theta'). \tag{3.9}$$

Next, we present the positive definiteness of Q_γ and $Q_{\gamma,T}$. As a result, this implies that $Q(\phi)$ is also a positive definite matrix.

Proposition 3.2. *Suppose that Assumptions* 1-2 *hold, and let* $Q(\phi)$ *be defined as in* (3.5). *Then if* $T\geq\max(\frac{1}{\phi},\frac{2}{1-\phi})$, $Q(\phi)$ *is a positive definite matrix.*

The proof is given in the Appendix B. By Proposition 3.2, we have gave a sufficient condition for the matrix $Q(\phi)$ to be a positive definite matrix. The research is ongoing to derive a necessary and sufficient condition for $Q(\phi)$ to be a positive definite matrix in case T is not large. In the sequel, to simplify the presentation of this thesis, we suppose that the conditions are met for the matrix $Q(\phi)$ to be a positive definite matrix. Note that this assumption does not affect the asymptotic optimality of the proposed method. Indeed, if T is large, by the results in Dehling *et al.* (2010, 2014), one can prove that $Q(\phi)$ is a positive definite matrix. Further, let

$$J_1 = \Sigma L_1'(L_1\Sigma L_1')^{-1} \text{ and } J_2 = (L_2'Q^{-1}(\phi)L_2)^{-1}L_2'Q^{-1}(\phi), \tag{3.10}$$

and let $\widetilde{\theta}$ be the RMLE. Proposition 3.2 is crucial in deriving the existence of the UMLE and RMLE. Below, we present a result which gives the UMLE and RMLE.

Lemma 3.1. *Suppose that Assumptions 1-2 hold. Then, the UMLE of the parameter* θ *is* $\hat{\theta} = P(\phi)Q^{-1}(\phi)$. *Further, if* H_0 *in* (2.3) *holds, the RMLE is given by*
$\widetilde{\theta} = \hat{\theta} - J_1(L_1\hat{\theta} - d_1) + J_1L_1(\hat{\theta}L_2 - d_2)J_2 - (\hat{\theta}L_2 - d_2)J_2.$

The proof is given in the Appendix B.

3.2 Asymptotic normality

In this section, we first derive the asymptotic normality of the UMLE, then, by the relationship between UMLE and RMLE as stated in Lemma 3.1, we derive the joint asymptotic normality of the UMLE and RMLE.

3.2.1 Asymptotic normality of UMLE

In this subsection, we investigate the asymptotic normality of the UMLE given in Lemma 3.1. First, we derive the following proposition which is used as an intermediate result.

Proposition 3.3. *Suppose that the Assumptions 1-2 hold, the SDE in* (2.1) *has the solution:* $X_t = \{e^{-A_1 t}X_0 + h_1(t) + Z_1(t)\}\mathbb{I}_{\{0\leq t\leq\gamma\}} + \{e^{-A_2 t}X_0 + h_2(t) + Z_2(t)\}\mathbb{I}_{\{t\geq\gamma\}}$, *where, for* $k = 1, 2$,

$$h_k(t) = e^{-A_k t}\int_0^t e^{A_k s}\mu_k\varphi(s)ds, \quad Z_k(t) = e^{-A_k t}\int_0^t e^{A_k s}\Sigma^{1/2}dW_s. \tag{3.11}$$

Proof. Applying Ito's formula with $g(x,t) = e^{A_1 t}x$, $0 \leq t \leq \gamma$ and $g(x,t) = e^{A_2 t}x$, $\gamma \leq t \leq T$, and following the same procedure in (2.5)-(2.7), we get:

$$X_t = e^{-A_1 t}X_0 + h_1(t) + Z_1(t), \tag{3.12}$$

$0 \leq t \leq \gamma$, and

$$X_t = e^{-A_2 t}X_0 + h_2(t) + Z_2(t), \tag{3.13}$$

$\gamma \leq t \leq T$, this completes the proof. □

Obviously, the process from SDE (2.1) is not stationary and ergodic. In order to study the asymptotic behaviours of the $\hat{\theta}$, we define the following auxiliary processes. Let

$$\widetilde{X}_1(t) = \widetilde{h}_1(t) + \widetilde{Z}_1(t), \quad \widetilde{X}_2(t) = \widetilde{h}_2(t) + \widetilde{Z}_2(t), \quad 0 \leq t \leq T, \tag{3.14}$$

where, for $k = 1, 2$,

$$\widetilde{h}_k(t) = e^{-A_k t}\int_{-\infty}^t e^{A_k s}\mu_k\varphi(s)ds, \quad \widetilde{Z}_k(t) = e^{-A_k t}\int_{-\infty}^t e^{A_k s}\Sigma^{1/2}d\widetilde{W}_s, \tag{3.15}$$

where $\{\widetilde{W}_s, s \in \mathbb{R}\}$ denotes a d-dimensional bilateral Brownian motion as in (2.12). Further, let $\widetilde{X}_t = \widetilde{X}_1(t)\mathbb{I}_{\{t\leq\gamma\}} + \widetilde{X}_2(t)\mathbb{I}_{\{t>\gamma\}}$, $0 \leq t \leq T$. From (2.17), we denote $V_1(k) = \mathrm{E}(\widetilde{Z}_1(0)\widetilde{Z}_1'(k))$, $V_2(k) = \mathrm{E}(\widetilde{Z}_2(0)\widetilde{Z}_2'(k))$, and define

$$\Sigma_0 = \begin{bmatrix} I_p & -\int_0^1 \varphi(t)\widetilde{h}_1'(t)dt \\ -\int_0^1 \widetilde{h}_1(t)\varphi'(t)dt & \int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt + V_1(0) \end{bmatrix}, \tag{3.16}$$

and

$$\Sigma_1 = \begin{bmatrix} I_p & -\int_0^1 \varphi(t)\widetilde{h}_2'(t)dt \\ -\int_0^1 \widetilde{h}_2(t)\varphi'(t)dt & \int_0^1 \widetilde{h}_2(t)\widetilde{h}_2'(t)dt + V_2(0) \end{bmatrix}. \tag{3.17}$$

Proposition 3.4. *Suppose that Assumpitions 1-2 hold, then for $\phi \in (0, 1)$*

$$\frac{1}{T}Q_\gamma \xrightarrow[T\to\infty]{P} \phi\Sigma_0, \quad \text{and} \quad TQ_\gamma^{-1} \xrightarrow[T\to\infty]{P} \frac{1}{\phi}\Sigma_0^{-1}.$$

The proof is provided in the Appendix B. Analogically, by Proposition 2.11, we have the following result:

Proposition 3.5. *Suppose that Assumpitions 1-2 hold, then for $\phi \in (0, 1)$*

$$\frac{1}{T}Q_{\gamma,T} \xrightarrow[T\to\infty]{P} (1-\phi)\Sigma_1, \quad \text{and} \quad TQ_{\gamma,T}^{-1} \xrightarrow[T\to\infty]{P} \frac{1}{1-\phi}\Sigma_1^{-1}.$$

Proof. The proof of the first statement is similar to that given for Proposition 2.11. The proof of the second statement follows from the same technique as used in proof of Proposition 3.4 □

Now, denote

$$\Sigma_2 = \begin{bmatrix} \phi\Sigma_0 & 0_{p+d} \\ 0_{p+d} & (1-\phi)\Sigma_1 \end{bmatrix}, \tag{3.18}$$

where Σ_0 and Σ_1 are defined in (3.16) and (3.17) respectively, then we have

Proposition 3.6. *Suppose that Assumpitions 1-2 hold, then for $\phi \in (0,1)$*

$$\frac{1}{T}Q(\phi) \xrightarrow[T\to\infty]{P} \Sigma_2, \text{ and } TQ^{-1}(\phi) \xrightarrow[T\to\infty]{P} \Sigma_2^{-1}. \tag{3.19}$$

Proof. By Proposition 3.2, we have $\frac{1}{T}Q$ is positive definite and thus it is invertible, we have

$$\left(\frac{1}{T}Q(\phi)\right)^{-1} = TQ^{-1}(\phi) = \begin{bmatrix} TQ_\gamma^{-1} & 0_{p+d} \\ 0_{p+d} & TQ_{\gamma,T}^{-1} \end{bmatrix}.$$

By Proposition 3.4 and Proposition 3.5, we complete the proof. □

Proposition 3.7. *The UMLE $\hat{\theta}$ can be rewritten as*

$$\hat{\theta} = \theta + \Sigma^{1/2}\frac{1}{T}R_T(\phi)(TQ^{-1}(\phi)), \tag{3.20}$$

where

$$R_T'(\phi) = \int_0^T B'(t,\phi)dW_t', \tag{3.21}$$

and

$$B(t,\phi) = \begin{bmatrix} \varphi'(t)\mathbb{I}_{\{t\leq\gamma\}} & -X_t'\mathbb{I}_{\{t\leq\gamma\}} & \varphi'(t)\mathbb{I}_{\{t>\gamma\}} & -X_t'\mathbb{I}_{\{t>\gamma\}} \end{bmatrix} \in \mathbb{R}^{1\times 2(p+d)}. \tag{3.22}$$

The proof is provided in the Appendix B. By Proposition 3.7, we also have

$$\sqrt{T}(\hat{\theta}-\theta)' = (TQ^{-1}(\phi))\frac{1}{\sqrt{T}}R_T'(\phi)\Sigma^{1/2}.$$

To study the asymptotic normality of $\hat{\theta}$, we need to first explore the convergence of $\frac{1}{\sqrt{T}}R_T'$. In passing, by Cramer-Wold Theorem (Billingsley 1995), we have

$$\text{Vec}\left(\frac{1}{\sqrt{T}}R_T'(\phi)\right) \xrightarrow[T\to\infty]{d} M$$

if and only if

$$a'\mathrm{Vec}\left(\frac{1}{\sqrt{T}}R_T'(\phi)\right) \xrightarrow[T\to\infty]{d} a'M,$$

for all $a = \begin{bmatrix} a_1 & a_2 & \dots & a_{2d(p+d)} \end{bmatrix}' \in \mathbb{R}^{2d(p+d)}$. Therefore, we study the convergence of $a'\mathrm{Vec}\left(\frac{1}{\sqrt{T}}R_T'(\phi)\right)$ instead. Note that

$$a'\mathrm{Vec}\left(\frac{1}{\sqrt{T}}R_T'(\phi)\right) = \begin{bmatrix} a^{(1)} & a^{(2)} & \dots & a^{(d)} \end{bmatrix} \int_0^T dW_t \otimes C_T(t) = \sum_{i=1}^{d} \int_0^T a^{(i)} C_T(t) dW_t^{(i)},$$

where $a^{(i)}$ is a $2(p+d)$-row vector given by

$$a^{(i)} = \begin{bmatrix} a_{(i-1)2(p+d)+1} & a_{(i-1)2(p+d)+2} & \dots & a_{i2(p+d)} \end{bmatrix}, \tag{3.23}$$

and

$$C_T(t) = \begin{bmatrix} \frac{1}{\sqrt{T}}\varphi'(t)\mathbb{I}_{\{t\le\gamma\}} & -\frac{1}{\sqrt{T}}X_t'\mathbb{I}_{\{t\le\gamma\}} & \frac{1}{\sqrt{T}}\varphi'(t)\mathbb{I}_{\{t>\gamma\}} & -\frac{1}{\sqrt{T}}X_t'\mathbb{I}_{\{t>\gamma\}} \end{bmatrix}'. \tag{3.24}$$

Proposition 3.8. *Suppose that Assumptions 1-2 hold. Then for $T > 0$, $i = 1,2,\dots,d$,*

$$\mathrm{P}\left(\int_0^T (a^{(i)}C_T(t))^2 dt < \infty\right) = 1,$$

where $C_T(t)$ and $a^{(i)}$ are defined in (3.24) *and* (3.23)*.*

Proof. By Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathrm{E}\left(\int_0^T (a^{(i)}C_T(t))^2 dt\right) &\le \|a^{(i)}\|_2^2 \mathrm{E}\left(\int_0^T \|C_T(t)\|_2^2 dt\right) \\ &\le \|a^{(i)}\|_2^2 \mathrm{E}\left[\frac{1}{T}\left(\int_0^T \|\varphi(t)\mathbb{I}_{\{t\le\gamma\}}\|_2^2 dt + \int_0^T \|X_t\mathbb{I}_{\{t\le\gamma\}}\|_2^2 dt\right)\right] \\ &+ \|a^{(i)}\|_2^2 \mathrm{E}\left[\frac{1}{T}\left(\int_0^T \|\varphi(t)\mathbb{I}_{\{t>\gamma\}}\|_2^2 dt + \int_0^T \|X_t\mathbb{I}_{\{t>\gamma\}}\|_2^2 dt\right)\right]. \end{aligned}$$

Since $\|\varphi(t)\|_2^2$ and $\|X_t\|_2^2$ are non-negative, we have

$$\|\varphi(t)\mathbb{I}_{\{t\le\gamma\}}\|_2^2 = \|\varphi(t)\|_2^2\mathbb{I}_{\{t\le\gamma\}} \le \|\varphi(t)\|_2^2,$$

$$\|X_t\mathbb{I}_{\{t\le\gamma\}}\|_2^2 = \|X_t\|_2^2\mathbb{I}_{\{t\le\gamma\}} \le \|X_t\|_2^2.$$

Similarly, we have

$$\|\varphi(t)\mathbb{I}_{\{t>\gamma\}}\|_2^2 = \|\varphi(t)\|_2^2\mathbb{I}_{\{t>\gamma\}} \leq \|\varphi(t)\|_2^2,$$
$$\|X_t\mathbb{I}_{\{t>\gamma\}}\|_2^2 = \|X_t\|_2^2\mathbb{I}_{\{t>\gamma\}} \leq \|X_t\|_2^2.$$

Therefore

$$\mathrm{E}\left(\int_0^T (a^{(i)}C_T(t))^2 dt\right) \leq \|a^{(i)}\|_2^2 \left[\frac{2}{T}\left(\int_0^T \mathrm{E}(\|\varphi(t)\|_2^2)dt + \int_0^T \mathrm{E}(\|X_t\|_2^2)dt\right)\right].$$

From Remark 1 and Proposition 2.1, we have the boundedness of $\varphi(t)$ and X_t in L^2. Let $\mathrm{E}(\|\varphi(t)\|_2^2) < K_\varphi$ and $\mathrm{E}(\|X_t\|_2^2) < K_x$, we get

$$\mathrm{E}\left(\int_0^T (a^{(i)}C_T(t))^2 dt\right) < 2\|a^{(i)}\|_2^2(K_\varphi + K_x) < \infty.$$

Then, we have

$$\mathrm{P}\left(\int_0^T (a^{(i)}C_T(t)^2 dt < \infty)\right) = 1,$$

for all $i = 1, 2, ..., d$, which completes the proof. □

From Proposition 3.8, we establish below a proposition which gives the convergence in distribution of $\frac{1}{\sqrt{T}}R'_T(\phi)$. In short, we apply Proposition A.1 in the Appendix A, which is a special case of the proposition 1.21 in Kutoyants (2004) with $d_1 = 1$ and $d_2 = d$.

Proposition 3.9. *Suppose that the conditions for Proposition* 3.6 *hold. Then* $\frac{1}{\sqrt{T}}R'_T(\phi) \xrightarrow[T\to\infty]{d} R \sim \mathcal{N}_{2(p+d)\times d}(0, I_d \otimes \Sigma_2)$*, where* Σ_2 *is defined in* (3.18).

The proof is provided in Appendix B. From Proposition 3.9, we derive below the asymptotic normality of the UMLE.

Proposition 3.10. *Suppose that the conditions for Proposition* 3.6 *hold. Then the UMLE $\hat{\theta}$ is asymptotically normal. More precisely*

$$\rho_T = \sqrt{T}(\hat{\theta} - \theta)' \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{2(p+d)\times d}(0, \Sigma \otimes \Sigma_2^{-1}).$$

The proof is provided in Appendix B.

3.2.2 Joint asymptotic normality of MLE and RMLE

In this subsection, we derive the joint asymptotic properties of the UMLE, RMLE and some other estimators. To avoid asymptotic degeneracy, we consider the following set of local alternatives:

$$K_T : L_1\theta = d_1 \quad \text{and} \quad \theta L_2 = d_2 + \frac{r_2}{\sqrt{T}}, \quad T > 0, \tag{3.25}$$

where $r_2 \in \mathbb{R}^{d\times n}$ is a fixed matrix. Also, we assume that $0 < \|r_2\| < \infty$. Define $\zeta_T = \sqrt{T}(\widetilde{\theta} - \theta)'$, according to Lemma 3.1, we have

$$\begin{aligned}
\sqrt{T}(\widetilde{\theta} - \theta) &= \sqrt{T}(\hat{\theta} - \theta) - J_1L_1\sqrt{T}(\hat{\theta} - \theta) \\
&\quad + J_1L_1(\sqrt{T}(\hat{\theta} - \theta)L_2 + r_2)J_2 - (\sqrt{T}(\hat{\theta} - \theta)L_2 + r_2)J_2 \\
&= \sqrt{T}(\hat{\theta} - \theta) - J_1L_1\sqrt{T}(\hat{\theta} - \theta) - r_2J_2 \\
&\quad + J_1L_1\sqrt{T}(\hat{\theta} - \theta)L_2J_2 + J_1L_1r_2J_2 - \sqrt{T}(\hat{\theta} - \theta)L_2J_2.
\end{aligned}$$

Then

$$\sqrt{T}(\widetilde{\theta} - \theta) = (I_d - J_1L_1)\sqrt{T}(\hat{\theta} - \theta)(I_{2(p+d)} - L_2J_2) + J_1L_1r_2J_2 - r_2J_2. \tag{3.26}$$

Further, let $f(\mathbf{X}^{-1}) = (L_2'\mathbf{X}^{-1}L_2)^{-1}L_2'\mathbf{X}^{-1}$ for a positive definite matrix $\mathbf{X}$. Then we have

$$J_2 = f(Q^{-1}(\phi)) = (L_2'Q^{-1}(\phi)L_2)^{-1}L_2'Q^{-1}(\phi) = [L_2'(TQ^{-1}(\phi))L_2]^{-1}L_2'(TQ^{-1}(\phi)).$$

By Proposition 3.2, we have

$$TQ^{-1}(\phi) \xrightarrow[T\to\infty]{P} \Sigma_2^{-1}.$$

Therefore, by the continuous mapping theorem, we have

$$J_2 \xrightarrow[T\to\infty]{P} (L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1} = J_3. \tag{3.27}$$

Similarly, we have

$$J_4 = I_{2(p+d)} - L_2J_2 \xrightarrow[T\to\infty]{P} I_{2(p+d)} - L_2J_3 = J_5, \tag{3.28}$$

$$J_6 = J_1L_1r_2J_2 - r_2J_2 \xrightarrow[T\to\infty]{P} J_1L_1r_2J_3 - r_2J_3 = J_7. \tag{3.29}$$

Further, to simplify some notations, denote $J = I_d - J_1L_1$. Note that

$$\begin{aligned} J\Sigma J' = (I_d - J_1L_1)\Sigma(I_d - J_1L_1)' &= (\Sigma - J_1L_1\Sigma)(I_d - J_1L_1)' \\ &= \Sigma - \Sigma L_1'J_1' - J_1L_1\Sigma + J_1L_1\Sigma L_1'J_1'. \end{aligned}$$

Further, since Σ is symmetric, by (3.10), we have $J_1 = \Sigma L_1'(L_1\Sigma L_1')^{-1}$, therefore

$$\Sigma L_1'J_1' = \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma = J_1L_1\Sigma, \tag{3.30}$$

and $J_1L_1\Sigma L_1'J_1' = \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma$. Then,

$$J_1L_1\Sigma L_1'J_1' = \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma = J_1L_1\Sigma. \tag{3.31}$$

Therefore, by (3.30) and (3.31), we get

$$J\Sigma J' = \Sigma - \Sigma L_1'J_1' = \Sigma - J_1L_1\Sigma = J\Sigma. \tag{3.32}$$

Further, we have

$$J_5'\Sigma_2^{-1}J_5 = (I_{2(p+d)} - J_3'L_2')\Sigma_2^{-1}(I_{2(p+d)} - L_2J_3) = (\Sigma_2^{-1} - J_3'L_2'\Sigma_2^{-1})(I_{2(p+d)} - L_2J_3).$$

By (3.27), we have $J_3 = (L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1}$, and since Σ_2^{-1} is symmetric, we get

$$J_3'L_2'\Sigma_2^{-1} = \Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1} = \Sigma_2^{-1}L_2J_3, \tag{3.33}$$

and $J_3'L_2'\Sigma_2^{-1}L_2J_3 = \Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1}$, and then

$$J_3'L_2'\Sigma_2^{-1}L_2J_3 = \Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1} = J_3'L_2'\Sigma_2^{-1}. \tag{3.34}$$

Then by (3.33) and (3.34), we get

$$J_5'\Sigma_2^{-1}J_5 = \Sigma_2^{-1} - \Sigma_2^{-1}L_2J_3 - J_3'L_2'\Sigma_2^{-1} + J_3'L_2'\Sigma_2^{-1}L_2J_3 = \Sigma_2^{-1} - \Sigma_2^{-1}L_2J_3 = \Sigma_2^{-1}J_5. \tag{3.35}$$

The asymptotic normality of RMLE follows from the following proposition which gives the joint asymptotic distribution of $\begin{bmatrix}\rho_T & \zeta_T\end{bmatrix}$.

Proposition 3.11. *Suppose that the conditions of Propositions* 3.6 *hold along with the set of local alternatives* K_T *in* (3.25)*, then*

$$\begin{bmatrix}\rho_T & \zeta_T\end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix}\rho & \zeta\end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d}\left(\begin{bmatrix}0 & J_7'\end{bmatrix}, \begin{bmatrix}\Omega_{22} & \Omega_{22}-\Omega_{11} \\ \Omega_{22}-\Omega_{11} & \Omega_{22}-\Omega_{11}\end{bmatrix}\right),$$

where $\Omega_{11} = \Sigma\otimes\Sigma_2^{-1} - (J\Sigma)\otimes(\Sigma_2^{-1}J_5)$, $\Omega_{22} = \Sigma\otimes\Sigma_2^{-1}$.

The proof is provided in the Appendix B.

Corollary 3.1. *Suppose that the conditions of Propositions* 3.6 *hold along with the set of local alternatives* K_T *in* (3.25)*. Then,* $\zeta_T \xrightarrow[T\to\infty]{d} \zeta \sim \mathcal{N}_{2(p+d)\times d}(J_7', \Omega_{22}-\Omega_{11})$.

The proof follows from Proposition 3.11. Define $\xi_T = \sqrt{T}(\hat{\theta}-\widetilde{\theta})'$. From Proposition 3.11, we derive the asymptotic distribution of $\begin{bmatrix}\rho_T & \xi_T\end{bmatrix}$.

Proposition 3.12. *Suppose that the conditions of Propositions* 3.11 *hold, then*

$$\begin{bmatrix}\rho_T & \xi_T\end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix}\rho & \xi\end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d}\left(\begin{bmatrix}0 & J_7'\end{bmatrix}, \begin{bmatrix}\Omega_{22} & \Omega_{11} \\ \Omega_{11} & \Omega_{11}\end{bmatrix}\right).$$

Proof. Observe that

$$\begin{bmatrix} \rho_T & \xi_T \end{bmatrix} = \begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} \begin{bmatrix} I_d & I_d \\ 0 & -I_d \end{bmatrix}.$$

Using vectorization, we get

$$\operatorname{Vec} \begin{bmatrix} \rho_T & \xi_T \end{bmatrix} = \left(\begin{bmatrix} I_d & 0 \\ I_d & -I_d \end{bmatrix} \otimes I_{2(p+d)} \right) \operatorname{Vec} \begin{bmatrix} \rho_T & \zeta_T \end{bmatrix}.$$

From Proposition 3.11, we have

$$\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix} \rho & \zeta \end{bmatrix},$$

where

$$\begin{bmatrix} \rho & \zeta \end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d} \left(\begin{bmatrix} 0 & J_7' \end{bmatrix}, \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \\ (J\Sigma) \otimes (\Sigma_2^{-1} J_5) & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \end{bmatrix} \right).$$

Therefore, by Slutsky's theorem, we have

$$\operatorname{Vec} \begin{bmatrix} \rho_T & \xi_T \end{bmatrix} \xrightarrow[T\to\infty]{d} \left(\begin{bmatrix} I_d & 0 \\ I_d & -I_d \end{bmatrix} \otimes I_{2(p+d)} \right) \operatorname{Vec} \begin{bmatrix} \rho & \zeta \end{bmatrix}. \tag{3.36}$$

Note that

$$\begin{aligned} \left(\begin{bmatrix} I_d & 0 \\ I_d & -I_d \end{bmatrix} \otimes I_{2(p+d)} \right) \operatorname{Vec} \left(\begin{bmatrix} 0 & J_7' \end{bmatrix} \right) &= \operatorname{Vec} \left(\begin{bmatrix} 0 & J_7' \end{bmatrix} \begin{bmatrix} I_d & I_d \\ 0 & -I_d \end{bmatrix} \right) \\ &= \operatorname{Vec} \begin{bmatrix} 0_{2(p+d)\times d} \\ -J_7' \end{bmatrix}. \end{aligned} \tag{3.37}$$

Moreover, we have

$$\begin{bmatrix} I_d & 0 \\ I_d & -I_d \end{bmatrix} \otimes I_{2(p+d)} = \begin{bmatrix} I_d \otimes I_{2(p+d)} & 0 \\ I_d \otimes I_{2(p+d)} & -I_d \otimes I_{2(p+d)} \end{bmatrix} = \begin{bmatrix} I_{2d(p+d)} & 0 \\ I_{2d(p+d)} & -I_{2d(p+d)} \end{bmatrix}.$$

Therefore, for the covariance term, we get

$$\begin{bmatrix} I_{2d(p+d)} & 0 \\ I_{2d(p+d)} & -I_{2d(p+d)} \end{bmatrix} \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \\ (J\Sigma) \otimes (\Sigma_2^{-1} J_5) & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \end{bmatrix} \begin{bmatrix} I_{2d(p+d)} & 0 \\ I_{2d(p+d)} & -I_{2d(p+d)} \end{bmatrix}'$$

$$= \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & \Sigma \otimes \Sigma_2^{-1} - (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \\ \Sigma \otimes \Sigma_2^{-1} - (J\Sigma) \otimes (\Sigma_2^{-1} J_5) & \Sigma \otimes \Sigma_2^{-1} - (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \end{bmatrix}. \tag{3.38}$$

By combining (3.36), (3.37), and (3.38), we complete the proof. □

From Proposition 3.11, we also derive the asymptotic distribution of $\begin{bmatrix} \zeta_T & \xi_T \end{bmatrix}$.

Proposition 3.13. *Suppose that the conditions of Propositions* 3.11 *hold, then*

$$\begin{bmatrix} \zeta_T & \xi_T \end{bmatrix} \xrightarrow[T \to \infty]{d} \begin{bmatrix} \zeta & \xi \end{bmatrix} \sim \mathcal{N}_{2(p+d) \times 2d} \left(\begin{bmatrix} J_7' & -J_7' \end{bmatrix}, \begin{bmatrix} \Omega_{22} - \Omega_{11} & 0 \\ 0 & \Omega_{11} \end{bmatrix} \right).$$

Proof. Observe that

$$\begin{bmatrix} \zeta_T & \xi_T \end{bmatrix} = \begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} \begin{bmatrix} 0 & I_d \\ I_d & -I_d \end{bmatrix}.$$

Using vectorization

$$\operatorname{Vec} \begin{bmatrix} \zeta_T & \xi_T \end{bmatrix} = \left(\begin{bmatrix} 0 & I_d \\ I_d & -I_d \end{bmatrix} \otimes I_{2(p+d)} \right) \operatorname{Vec} \begin{bmatrix} \rho_T & \zeta_T \end{bmatrix}.$$

From Proposition 3.11, we have

$$\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} \xrightarrow[T \to \infty]{d} \begin{bmatrix} \rho & \zeta \end{bmatrix},$$

where

$$\begin{bmatrix} \rho & \zeta \end{bmatrix} \sim \mathcal{N}_{2(p+d) \times 2d} \left(\begin{bmatrix} 0 & J_7' \end{bmatrix}, \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \\ (J\Sigma) \otimes (\Sigma_2^{-1} J_5) & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \end{bmatrix} \right).$$

Therefore, by Slutsky's theorem, we have

$$\text{Vec}\begin{bmatrix}\zeta_T & \xi_T\end{bmatrix} \xrightarrow[T\to\infty]{d} \left(\begin{bmatrix}0 & I_d \\ I_d & -I_d\end{bmatrix} \otimes I_{2(p+d)}\right) \text{Vec}\begin{bmatrix}\rho & \zeta\end{bmatrix}. \tag{3.39}$$

Note that

$$\left(\begin{bmatrix}0 & I_d \\ I_d & -I_d\end{bmatrix} \otimes I_{2(p+d)}\right) \text{Vec}(\begin{bmatrix}0 & J_7'\end{bmatrix}) = \text{Vec}\begin{bmatrix}0 & J_7'\end{bmatrix}\left(\begin{bmatrix}0 & I_d \\ I_d & -I_d\end{bmatrix}\right) = \text{Vec}\begin{bmatrix}J_7' & -J_7'\end{bmatrix}. \tag{3.40}$$

Moreover, we have

$$\begin{bmatrix}0 & I_d \\ I_d & -I_d\end{bmatrix} \otimes I_{2(p+d)} = \begin{bmatrix}0 & I_d \otimes I_{2(p+d)} \\ I_d \otimes I_{2(p+d)} & -I_d \otimes I_{2(p+d)}\end{bmatrix} = \begin{bmatrix}0 & I_{2d(p+d)} \\ I_{2d(p+d)} & -I_{2d(p+d)}\end{bmatrix}.$$

Therefore, for the covariance term, we get

$$\begin{aligned}&\begin{bmatrix}0 & I_{2d(p+d)} \\ I_{2d(p+d)} & -I_{2d(p+d)}\end{bmatrix}\begin{bmatrix}\Sigma \otimes \Sigma_2^{-1} & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \\ (J\Sigma) \otimes (\Sigma_2^{-1} J_5) & (J\Sigma) \otimes (\Sigma_2^{-1} J_5)\end{bmatrix}\begin{bmatrix}0 & I_{2d(p+d)} \\ I_{2d(p+d)} & -I_{2d(p+d)}\end{bmatrix}' \\ &= \begin{bmatrix}(J\Sigma) \otimes (\Sigma_2^{-1} J_5) & 0 \\ 0 & \Sigma \otimes \Sigma_2^{-1} - (J\Sigma) \otimes (\Sigma_2^{-1} J_5)\end{bmatrix}.\end{aligned} \tag{3.41}$$

By combining (3.39), (3.40), and (3.41), we complete the proof. □

Chapter 4

Inference in case of unknown change-point

In this chapter, we present the proposed inference method in the case of unknown change-point. This chapter is subdivided into 4 sections. In Section 4.1, we derive the unrestricted estimator (UE) and the restricted estimator (RE). Briefly, the UE and the RE are obtained from the UMLE and RMLE along with plug-in method. In Section 4.2, we establish the joint asymptotic normality of the UE and RE. Further, in Section 4.3, we address the testing problem in (2.3), and in Section 4.4, we introduce the shrinkage estimators.

4.1 The UE and the RE

In this section, we derive the UE and RE by using plug-in method. Let $\hat{\phi}$ be a $\mathfrak{F}_T$-measurable and a consistent estimator of the change-point. To introduce some notations, let $Q(\hat{\phi}) = \begin{bmatrix} Q_{\hat{\phi}T} & 0_{p+d} \\ 0_{p+d} & Q_{\hat{\phi}T,T} \end{bmatrix}$, where

$$Q_{\hat{\phi}T} = \begin{bmatrix} \int_0^{\hat{\phi}T} \varphi(t)\varphi'(t)dt & -\int_0^{\hat{\phi}T} \varphi(t)X_t'dt \\ -\int_0^{\hat{\phi}T} X_t\varphi'(t)dt & \int_0^{\hat{\phi}T} X_tX_t'dt \end{bmatrix}, \tag{4.1}$$

$$Q_{\hat{\phi}T,T} = \begin{bmatrix} \int_{\hat{\phi}T}^{T} \varphi(t)\varphi'(t)dt & -\int_{\hat{\phi}T}^{T} \varphi(t)X_t'dt \\ -\int_{\hat{\phi}T}^{T} X_t\varphi'(t)dt & \int_{\hat{\phi}T}^{T} X_tX_t'dt \end{bmatrix}. \tag{4.2}$$

According to Proposition 3.2, one can verify the positive definitness of $\frac{1}{T}Q_{\hat{\phi}T}$, and $\frac{1}{T}Q_{\hat{\phi}T,T}$. Let $J_2(\hat{\phi})$ and $P(\hat{\phi})$ be as J_2 and $P(\phi)$ by replacing ϕ by $\hat{\phi}$. Then, the plug-in UMLE and plug-in RMLE are given by

$$\hat{\theta}(\hat{\phi}) = P(\hat{\phi})Q^{-1}(\hat{\phi}), \tag{4.3}$$

$$\widetilde{\theta}(\hat{\phi}) = \hat{\theta}(\hat{\phi}) - J_1(L_1\hat{\theta}(\hat{\phi}) - d_1) + J_1(L_1\hat{\theta}(\hat{\phi}) - d_1)L_2J_2(\hat{\phi}) - (\hat{\theta}L_2 - d_2)J_2(\hat{\phi}). \tag{4.4}$$

Remark 4. *A consistent estimator can be obtained using a method based on that given in Chen and Nkurunziza (2015).*

Assumption 3. *The estimator $\hat{\phi}$ is $\mathfrak{F}_T$-measurable, valued on [0,1]. Further, there exists $\delta_0 > 0$ such that $\hat{\phi} - \phi = O_p(T^{-\delta_0})$.*

Remark 5. *This Assumption is similar to the Assumption 3 in Nkurunziza and Zhang (2018). It is used to derive the asymptotic behaviours of $\hat{\theta}(\hat{\phi})$ and $\widetilde{\theta}(\hat{\phi})$.*

Proposition 4.1. *Suppose that the conditions for Proposition 3.6 hold as well as Assumption 3, then*

$$(i)\ \frac{1}{T}\int_0^{\hat{\phi}T} \varphi(t)X_t'dt \xrightarrow[T\to\infty]{P} \phi\int_0^1 \varphi(t)\widetilde{h}_1'(t)dt,$$

$$(ii)\ \frac{1}{T}\int_{\hat{\phi}T}^{T} \varphi(t)X_t'dt \xrightarrow[T\to\infty]{P} (1-\phi)\int_0^1 \varphi(t)\widetilde{h}_2'(t)dt.$$

Proof. From Remark 1 and Proposition 2.2 we have the boundedness of $\varphi(t)$ and X_t in L^2. Let $\|\varphi(t)\|_2^2 \leq K_\varphi$ and $\mathrm{E}(\|X_t\|_2^2) \leq \sup\limits_{t\geq 0} \mathrm{E}(\|X_t\|_2^2) \leq K_x$ for all t, then we have

$$\mathrm{E}(\|\varphi(t)X_t'\|_F^2) \leq \|\varphi(t)\|_2^2 \mathrm{E}(\|X_t\|_2^2) \leq K_\varphi K_x < \infty.$$

Therefore, by Lemma $A.2$ in the Appendix A, we have

$$\frac{1}{T}\int_0^{\hat{\phi}T} \varphi(t)X_t'dt - \frac{1}{T}\int_0^{\phi T} \varphi(t)X_t'dt \xrightarrow[T\to\infty]{L^1} 0, \tag{4.5}$$

$$\frac{1}{T}\int_{\hat{\phi}T}^{T} \varphi(t)X_t'dt - \frac{1}{T}\int_{\phi T}^{T} \varphi(t)X_t'dt \xrightarrow[T\to\infty]{L^1} 0. \tag{4.6}$$

From Proposition 3.6, we have

$$\frac{1}{T}\int_0^{\phi T} \varphi(t)X_t'dt \xrightarrow[T\to\infty]{P} \phi\int_0^1 \varphi(t)\widetilde{h}_1'(t)dt, \tag{4.7}$$

$$\frac{1}{T}\int_{\phi T}^{T} \varphi(t)X_t'dt \xrightarrow[T\to\infty]{P} (1-\phi)\int_0^1 \varphi(t)\widetilde{h}_2'(t)dt, \tag{4.8}$$

which completes the proof. $\square$

Proposition 4.2. *Suppose that the conditions for Proposition* 4.1 *hold, then*

$$(i)\ \frac{1}{T}\int_0^{\hat{\phi}T} X_tX_t'dt \xrightarrow[T\to\infty]{P} \phi\left\{\int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt + V_1(0)\right\},$$

$$(ii)\ \frac{1}{T}\int_{\hat{\phi}T}^{T} X_tX_t'dt \xrightarrow[T\to\infty]{P} (1-\phi)\left\{\int_0^1 \widetilde{h}_2(t)\widetilde{h}_2'(t)dt + V_2(0)\right\}.$$

The proof is provided in the Appendix B. From Popositions 4.1-4.2, we derive the following proposition which is useful in obtaining the joint asymptotic normality of the UE and RE.

Proposition 4.3. *Suppose that the conditions for Proposition* 4.1 *hold, then* $\frac{1}{T}Q(\hat{\phi}) \xrightarrow[T\to\infty]{P} \Sigma_2$, and $TQ^{-1}(\hat{\phi}) \xrightarrow[T\to\infty]{P} \Sigma_2^{-1}$, *with* Σ_2 *defined in* (3.18).

Proof. From Popositions 4.1-4.2, we have $\frac{1}{T}Q(\hat{\phi}) \xrightarrow[T\to\infty]{P} \Sigma_2$. Further, let $g(\mathbf{X}) = \mathbf{X}^{-1}$ for a positive definite matrix $\mathbf{X}$. By the continuous mapping theorem, we have from the first statement

$$g\left(\frac{1}{T}Q_{\hat{\phi}T}\right) = TQ_{\hat{\phi}T}^{-1} \xrightarrow[T\to\infty]{P} g(\phi\Sigma_0) = \frac{1}{\phi}\Sigma_0^{-1},$$

and

$$g\left(\frac{1}{T}Q_{\hat{\phi}T,T}\right) = TQ_{\hat{\phi}T,T}^{-1} \xrightarrow[T\to\infty]{P} g(\phi\Sigma_1) = \frac{1}{1-\phi}\Sigma_1^{-1},$$

which completes the proof. □

Now define

$$R_T'(\hat{\phi}) = \int_0^T B'(\hat{\phi}, t)dW_t', \tag{4.9}$$

where

$$B(\hat{\phi}, t) = \begin{bmatrix} \varphi'(t)\mathbb{I}_{\{t\leq\hat{\phi}T\}} & -X_t'\mathbb{I}_{\{t\leq\hat{\phi}T\}} & \varphi'(t)\mathbb{I}_{\{t>\hat{\phi}T\}} & -X_t'\mathbb{I}_{\{t>\hat{\phi}T\}} \end{bmatrix}. \tag{4.10}$$

Proposition 4.4. *Suppose that the conditions for Proposition* 4.1 *hold as well as Assumption* 3 *with* $\delta_0 > \frac{1}{2}$*, then* $\frac{1}{\sqrt{T}}(R_T'(\hat{\phi}) - R_T'(\phi)) \xrightarrow[T\to\infty]{P} 0$*, where* $R_T'(\phi)$ *is defined in* (3.21)*.*

Proof. From Remark 1 and Proposition 2.2, we have the boundedness of $\varphi(t)$ and X_t in L^2, also. Let $f(\mu, A, X_t) = \mu\varphi(t) - AX_t$, by the Triangle Inequality, we have

$$\begin{aligned} \mathrm{E}(\|f(\mu, A, X_t)\|_2^2) &= \mathrm{E}(\|\mu\varphi(t) - AX_t\|_2^2) \leq \mathrm{E}[(\|\mu\varphi(t)\|_2 - \|AX_t\|_2)^2] \\ &\leq 2\mathrm{E}(\|\mu\varphi(t)\|_2^2) + 2\mathrm{E}(\|AX_t\|_2^2) \leq 2\|\mu\varphi(t)\|_2^2 + 2\|A\|_F^2(\mathrm{E}\|X_t\|_2^2) < \infty, \end{aligned}$$

for $\mu = \mu_1, \mu_2$, and $A = A_1, A_2$. Then, by Lemma 3.3 in Nkurunziza and Zhang (2018), we get

$$\begin{aligned}
\frac{1}{\sqrt{T}}\int_0^{\hat{\phi}T} X_t dW_t' - \frac{1}{\sqrt{T}}\int_0^{\phi T} X_t dW_t' &\xrightarrow[T\to\infty]{P} 0,\\
\frac{1}{\sqrt{T}}\int_{\hat{\phi}T}^{T} X_t dW_t' - \frac{1}{\sqrt{T}}\int_{\hat{\phi}T}^{T} X_t dW_t' &\xrightarrow[T\to\infty]{P} 0,\\
\frac{1}{\sqrt{T}}\int_0^{\hat{\phi}T} \varphi(t) dW_t' - \frac{1}{\sqrt{T}}\int_0^{\phi T} \varphi(t) dW_t' &\xrightarrow[T\to\infty]{P} 0,\\
\frac{1}{\sqrt{T}}\int_{\hat{\phi}T}^{T} \varphi(t) dW_t' - \frac{1}{\sqrt{T}}\int_{\hat{\phi}T}^{T} \varphi(t) dW_t' &\xrightarrow[T\to\infty]{P} 0,
\end{aligned}$$

which completes the proof. □

Proposition 4.5. *Suppose that the conditions for Proposition* 4.4 *hold, then*

$$\frac{1}{\sqrt{T}} R_T'(\hat{\phi}) \xrightarrow[T\to\infty]{d} R \sim \mathcal{N}_{2(p+d)\times d}(0, I_d \otimes \Sigma_2),$$

Corollary 4.1. *Suppose that the conditions for Proposition* 4.4 *hold, then*

$$\rho_T(\hat{\phi}) = \sqrt{T}(\hat{\theta}(\hat{\phi}) - \theta)' \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{2(p+d)\times d}(0, \Sigma \otimes \Sigma_2^{-1}).$$

Proof. From Proposition 3.7, one can get

$$\hat{\theta}(\hat{\phi}) = \theta + \Sigma^{1/2}\frac{1}{T} R_T(\hat{\phi})(TQ^{-1}(\hat{\phi})),$$

where $R_T(\hat{\phi})$ is defined in (4.9). Therefore

$$\sqrt{T}(\hat{\theta}(\hat{\phi}) - \theta) = \Sigma^{1/2}\frac{1}{\sqrt{T}} R_T(\hat{\phi})(TQ^{-1}(\hat{\phi})).$$

By Propositions 4.3 and 4.5, along with Slutsky's Theorem, we complete the proof.

□

4.2 Joint asymptotic normality

In this section, we present the joint asymptotic normality of the UE and the RE: $\widetilde{\theta}(\hat{\phi})$ and $\hat{\theta}(\hat{\phi})$. First of all , we study the asymptotic property of $\begin{bmatrix}\rho_T(\hat{\phi}) & \zeta_T(\hat{\phi})\end{bmatrix}$. To introduce some notations, from (3.27), (3.28), (3.29) and by Proposition 4.3, we get

$$J_2(\hat{\phi}) = [L_2'(TQ^{-1}(\hat{\phi}))L_2]^{-1}L_2'(TQ^{-1}(\hat{\phi})) \xrightarrow[T\to\infty]{P} (L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1} = J_3. \quad (4.11)$$

Also

$$J_4(\hat{\phi}) = I_{2(p+d)} - L_2J_2(\hat{\phi}) \xrightarrow[T\to\infty]{P} I_{2(p+d)} - L_2J_3 = J_5, \quad (4.12)$$

$$J_6(\hat{\phi}) = J_1L_1r_2J_2(\hat{\phi}) - r_2J_2(\hat{\phi}) \xrightarrow[T\to\infty]{P} J_1L_1r_2J_3 - r_2J_3 = J_7. \quad (4.13)$$

From (3.26) and (4.4), one can verify that

$$\sqrt{T}(\widetilde{\theta}(\hat{\phi}) - \theta) = (I_d - J_1L_1)\rho_T'(\hat{\phi})(I_{2(p+d)} - L_2J_2(\hat{\phi})) + J_1L_1r_2J_2(\hat{\phi}) - r_2J_2(\hat{\phi}). \quad (4.14)$$

Then

$$\begin{bmatrix}\sqrt{T}(\hat{\theta}(\hat{\phi}) - \theta)\\ \sqrt{T}(\widetilde{\theta}(\hat{\phi}) - \theta)\end{bmatrix} = \begin{bmatrix}\rho_T'(\hat{\phi})\\ J\rho_T'(\hat{\phi})J_4(\hat{\phi}) + J_6(\hat{\phi})\end{bmatrix}$$
$$= \begin{bmatrix}I_d\\ 0_d\end{bmatrix}\rho_T'(\hat{\phi}) + \begin{bmatrix}0_d\\ J\end{bmatrix}\rho_T'(\hat{\phi})J_4(\hat{\phi}) + \begin{bmatrix}0_{d\times 2(p+d)}\\ J_6(\hat{\phi})\end{bmatrix}, \quad (4.15)$$

where $J_4(\hat{\phi})$ and $J_6(\hat{\phi})$ are defined in (4.12) and (4.13). Denote

$$I^{(3)}(\hat{\phi}) = \begin{bmatrix}0_{d\times 2(p+d)}\\ J_6(\hat{\phi})\end{bmatrix} \in \mathbb{R}^{2d\times 2(p+d)}, \quad (4.16)$$

we have

$$\begin{bmatrix}\rho_T(\hat{\phi}) & \zeta_T(\hat{\phi})\end{bmatrix} = \begin{bmatrix}\sqrt{T}(\hat{\theta}(\hat{\phi}) - \theta)\\ \sqrt{T}(\widetilde{\theta}(\hat{\phi}) - \theta)\end{bmatrix}' = \rho_T(\hat{\phi})I^{(1)'} + J_4'(\hat{\phi})\rho_T(\hat{\phi})I^{(2)'} + I^{(3)}(\hat{\phi})',$$

where $I^{(1)}$ and $I^{(2)}$ are defined in (B.36). Further, using vectorization, we get

$$\mathrm{Vec}\begin{bmatrix}\rho_T(\hat{\phi}) & \zeta_T(\hat{\phi})\end{bmatrix} = (I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_4'(\hat{\phi}))\mathrm{Vec}(\rho_T(\hat{\phi})) + \mathrm{Vec}(I^{(3)}(\hat{\phi})').$$

By (4.13), we have

$$I^{(3)}(\hat{\phi}) = \begin{bmatrix} 0_{d\times 2(p+d)} \\ J_6(\hat{\phi}) \end{bmatrix} \xrightarrow[T\to\infty]{P} \begin{bmatrix} 0_{d\times 2(p+d)} \\ J_7 \end{bmatrix} = I^{(4)}. \tag{4.17}$$

Proposition 4.6. *Suppose that the conditions for Proposition* 4.4 *along with the set of local alternatives* K_T *in* (3.25)*. Then*

$$\begin{bmatrix}\rho_T(\hat{\phi}) & \zeta_T(\hat{\phi})\end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix}\rho & \zeta\end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d}\left(\begin{bmatrix}0 & J_7'\end{bmatrix}, \begin{bmatrix}\Omega_{22} & \Omega_{22}-\Omega_{11} \\ \Omega_{22}-\Omega_{11} & \Omega_{22}-\Omega_{11}\end{bmatrix}\right),$$

where $\Omega_{11} = \Sigma\otimes\Sigma_2^{-1} - (J\Sigma)\otimes(\Sigma_2^{-1}J_5)$, $\Omega_{22} = \Sigma\otimes\Sigma_2^{-1}$.

Proof. The proof follows from Corollary 4.1, and using the same method as in the proof of Proposition 3.11. □

Corollary 4.1. *Suppose that the conditions for Proposition* 4.6 *hold . Then, the RE* $\widetilde{\theta}(\hat{\phi})$ *given in* (4.4) *is asymptotically normal. More precisely,*

$$\zeta_T(\hat{\phi}) = \sqrt{T}(\widetilde{\theta}(\hat{\phi}) - \theta)' \xrightarrow[T\to\infty]{d} \zeta \sim \mathcal{N}_{2(p+d)\times d}(J_7', \Omega_{22}-\Omega_{11}).$$

The proof follows from the Proposition 4.6. From Proposition 4.6, we also derive the asymptotic distribution of both $\begin{bmatrix}\rho_T(\hat{\phi}) & \xi_T(\hat{\phi})\end{bmatrix}$, $\begin{bmatrix}\zeta_T(\hat{\phi}) & \xi_T(\hat{\phi})\end{bmatrix}$.

Proposition 4.7. *Suppose that the conditions for Proposition* 4.6 *hold. Then*

$$\begin{bmatrix}\rho_T(\hat{\phi}) & \xi_T(\hat{\phi})\end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix}\rho & \xi\end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d}\left(\begin{bmatrix}0 & J_7'\end{bmatrix}, \begin{bmatrix}\Omega_{22} & \Omega_{11} \\ \Omega_{11} & \Omega_{11}\end{bmatrix}\right).$$

Proof. Observe that

$$\begin{bmatrix} \rho_T(\hat{\phi}) & \xi_T(\hat{\phi}) \end{bmatrix} = \begin{bmatrix} \rho_T(\hat{\phi}) & \zeta_T(\hat{\phi}) \end{bmatrix} \begin{bmatrix} I_d & I_d \\ 0 & -I_d \end{bmatrix}.$$

By Proposition 4.6 and by using the same method as in Proposition 3.12, we complete the proof. □

Proposition 4.8. *Suppose that the conditions for Proposition* 4.6 *hold. Then*

$$\begin{bmatrix} \zeta_T(\hat{\phi}) & \xi_T(\hat{\phi}) \end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix} \zeta & \xi \end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d}\left(\begin{bmatrix} J_7' & -J_7' \end{bmatrix}, \begin{bmatrix} \Omega_{22}-\Omega_{11} & 0 \\ 0 & \Omega_{11} \end{bmatrix} \right).$$

Proof. Observe that

$$\begin{bmatrix} \zeta_T(\hat{\phi}) & \xi_T(\hat{\phi}) \end{bmatrix} = \begin{bmatrix} \rho_T(\hat{\phi}) & \zeta_T(\hat{\phi}) \end{bmatrix} \begin{bmatrix} 0 & I_d \\ I_d & -I_d \end{bmatrix}.$$

The proof follows from Proposition 4.6 and by using the same method as in Proposition 3.13. □

4.3 Testing the restriction

In this section, we give a test for the hypotheses in problem in (2.3) based on the properties of the joint asymptotic normality of the estimators. By using Propositions 4.6-4.8, we establish below a corollary which can be used for testing the restriction in (2.3), and for deriving the proposed shrinkage estimators. To introduce some notations, let $W_d(n,\Sigma)$ be a random matrix in $\mathbb{R}^{n\times n}$, whose distribution is Wishart with parameter Σ and degrees of freedom d. Also, let $W_d(n,\Sigma,\Lambda)$ be a random matrix in $\mathbb{R}^{n\times n}$, whose distribution is Wishart with parameter Σ, with degrees of freedom d and non-centrality parameter Λ, and let $\chi^2_q(\lambda)$ be a chi-square random variable with q degrees of freedom, and non-centrality parameter λ. It should be noted that in continuous times observation, the diffusion parameter Σ is assumed to be known and equals to the quadratic variation. However, in realistic case, the data are always collected in discrete times and therefore it needs to be estimated through the discrete observations. Thus, let $\hat{\Sigma}$ be a consistent estimator of Σ. Moreover, let $\Xi = L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'$ and $\Delta = \mathrm{Tr}(J_7\Xi J_7'\Sigma^{-1})$, where J_7 is defined in (3.29).

Corollary 4.1. *Suppose that the conditions for Proposition* 4.6 *hold, then*

$$\xi_T'(\hat{\phi})L_2(L_2'TQ^{-1}(\hat{\phi},T)L_2)^{-1}L_2'\xi_T(\hat{\phi}) \xrightarrow[T\to\infty]{d} W_n(d,\Sigma,J_7\Xi J_7'),\ \text{and}$$

$$\mathrm{Tr}(\xi_T'(\hat{\phi})L_2(L_2'TQ^{-1}(\hat{\phi},T)L_2)^{-1}L_2'\xi_T(\hat{\phi})\hat{\Sigma}^{-1}) \xrightarrow[T\to\infty]{d} \psi \sim \chi^2_{nd}(\Delta).$$

Proof. Note that from Propositions 4.3, 4.6 and 4.7 along with Slutsky's Theorem, we have

$$\xi_T'(\hat{\phi})L_2(L_2'TQ^{-1}(\hat{\phi},T)L_2)^{-1}L_2'\xi_T(\hat{\phi}) \xrightarrow[T\to\infty]{d} \xi'\Xi\xi,$$

where $\Xi = L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'$ and

$$\xi \sim \mathcal{N}_{2(p+d)\times d}\left(-J_7', \Sigma\otimes\Sigma_2^{-1} - (J\Sigma)\otimes(\Sigma_2^{-1}J_5)\right).$$

Further, notice that $(L_2'\Sigma_2^{-1}L_2)^{-1}$ is positive definite since Σ_2^{-1} is positive definite and L_2 is a full rank matrix. Then, let $P = (L_2'\Sigma_2^{-1}L_2)^{-1/2}L_2'$. Obviously, $\xi'P'P\xi = \xi'\Xi\xi$, therefore, we study the distribution of $P\xi$. Taking vectorization, we have

$$\mathrm{Vec}(P\xi) = (I_d\otimes P)\mathrm{Vec}(\xi),$$

then

$$\mathrm{Vec}(P\xi) \sim (I_d\otimes P)\mathcal{N}_{2d(p+d)}\left(-\mathrm{Vec}(J_7'), \Sigma\otimes\Sigma_2^{-1} - (J\Sigma)\otimes(\Sigma_2^{-1}J_5)\right).$$

To simplify the covariance term, we have that the covariance is equal to

$$(I_d \otimes P)(\Sigma \otimes \Sigma_2^{-1} - (J\Sigma) \otimes (\Sigma_2^{-1} J_5))(I_d \otimes P'). \tag{4.18}$$

We have

$$\Sigma \otimes (P\Sigma_2^{-1}P') = \Sigma \otimes ((L_2'\Sigma_2^{-1}L_2)^{-1/2}L_2'\Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1/2}).$$

Then

$$(I_d \otimes P)(\Sigma \otimes \Sigma_2^{-1})(I_d \otimes P') = \Sigma \otimes I_n. \tag{4.19}$$

Further,

$$(I_d \otimes P)[(J\Sigma) \otimes (\Sigma_2^{-1}J_5)](I_d \otimes P') = (J\Sigma) \otimes P(\Sigma_2^{-1}J_5)P'.$$

Since $J_5 = I_{2(p+d)} - L_2J_3$, we have

$$P(\Sigma_2^{-1}J_5)P' = P(\Sigma_2^{-1}(I_{2(p+d)} - L_2J_3))P' = P\Sigma_2^{-1}P' - P\Sigma_2^{-1}L_2J_3P'.$$

Notice that

$$P\Sigma_2^{-1}L_2J_3P' = (L_2'\Sigma_2^{-1}L_2)^{-1/2}L_2'\Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'\Sigma_2^{-1}L_2(L_2'\Sigma_2^{-1}L_2)^{-1/2} = I_n,$$

combining with (4.19), we get

$$(I_d \otimes P)[(J\Sigma) \otimes (\Sigma_2^{-1}J_5)](I_d \otimes P') = (J\Sigma) \otimes (P\Sigma_2^{-1}P' - P\Sigma_2^{-1}L_2J_3P') = 0. \tag{4.20}$$

Therefore, from (4.18), (4.19)and (4.20), we have

$$(I_d \otimes P)(\Sigma \otimes \Sigma_2^{-1} - (J\Sigma) \otimes (\Sigma_2^{-1}J_5))(I_d \otimes P') = \Sigma \otimes I_n.$$

Moreover, we have $-(I_d \otimes P)\text{Vec}(J_7') = -\text{Vec}(PJ_7')$, therefore

$$P\xi \sim \mathcal{N}_{n\times d}(-PJ_7', \Sigma \otimes I_n).$$

Hence, by the definition of Wishart distribution, we get

$$\xi'\Xi\xi = \xi'P'P\xi \sim W_n(d, \Sigma, J_7P'PJ_7') = W_n(d, \Sigma, J_7\Xi J_7'),$$

which completes the first statement of the proposition. Further, we have

$$\mathrm{Tr}(\hat{\Sigma}^{-1/2}\xi_T'(\hat{\phi})L_2(L_2'TQ^{-1}(\hat{\phi},T)L_2)^{-1}L_2'\xi_T(\hat{\phi})\hat{\Sigma}^{-1/2}) \xrightarrow[T\to\infty]{d} \mathrm{Tr}(\Sigma^{-1/2}\xi'\Xi\xi\Sigma^{-1/2}),$$

and, from previous result, we have

$$\Sigma^{-1/2}\xi'\Xi\xi\Sigma^{-1/2} \sim W_n(d, \Sigma^{-1/2}\Sigma\Sigma^{-1/2}, \Sigma^{-1/2}J_7\Xi J_7'\Sigma^{-1/2}) = W_n(d, I_d, \Sigma^{-1/2}J_7\Xi J_7'\Sigma^{-1/2}).$$

Then, by Corollary 2.4.2.2. in Kollo and Rosen (2011), we have

$$\mathrm{Tr}(\Sigma^{-1/2}\xi'\Xi\xi\Sigma^{-1/2}) = \mathrm{Tr}(\xi'\Xi\xi\Sigma^{-1}) \sim \chi^2_{nd}(\Delta),$$

where $\Delta = \mathrm{Tr}(J_7\Xi J_7'\Sigma^{-1})$, which completes the proof. $\square$

Note that if r_2 is a zero-matrix, then $J_7 = J_1L_1r_2J_3 - r_2J_3$ is also a zero-matrix and $\Delta = 0$, we have $\psi \sim \chi^2_{nd}$. From this corollary, one constructs a test for testing the restriction in (2.3). Let $\chi^2_{\alpha;nd}$ denote the αth-quantile of a χ^2_{nd}, for a given $0 < \alpha \leq 1$. For the testing problem in (2.3), we suggest to use the following test

$$\kappa(\phi) = \mathbb{I}_{\{\psi_T > \chi^2_{\alpha;nd}\}}, \tag{4.21}$$

where $\psi_T = \mathrm{Tr}(\xi_T'(\hat{\phi})L_2(L_2'TQ^{-1}(\hat{\phi},T)L_2)^{-1}L_2'\xi_T(\hat{\phi})\hat{\Sigma}^{-1})$.

Corollary 4.2. *Suppose that the conditions for Corollary* 4.1 *hold, then the asymptotic power function of the test in* (4.21) *is given by* $\Pi(\Delta) = \mathrm{P}\left(\chi^2_{nd}(\Delta) \geq \chi^2_{\alpha;nd}\right)$.

The proof follows from Corollary 4.1.

4.4 The Shrinkage Estimators

In this section, we present the proposed shrinkage estimators (SEs). First, note that, generally, the RE performs much better than the UE if the restriction holds, and the RE performs much worse if the restriction is seriously violated. To address this problem, we consider an intermediate case where the prior information is nearly correct. The proposed method combines the sample information and the prior information. Thus, the method is more flexible as it should preserve a good performance in case the prior holds or in case the prior does not hold. Following Sen and Saleh (1987), Nkurunziza (2012), Saleh (2006), Nkurunziza and Ahmed (2011) among others, we consider two Stein-rule (or shrinkage) estimators of the matrix parameter. The shrinkage estimator (SE) $\hat{\theta}^S$ is defined as

$$\hat{\theta}^S = \widetilde{\theta}(\hat{\phi}) + [1 - (nd-2)\psi_T^{-1}](\hat{\theta}(\hat{\phi}) - \widetilde{\theta}(\hat{\phi})), \tag{4.22}$$

where we assume $nd > 2$, and $\psi_T = \mathrm{Tr}(\xi_T'(\hat{\phi})L_2(L_2'TQ^{-1}(\hat{\phi},T)L_2)^{-1}L_2'\xi_T(\hat{\phi})\hat{\Sigma}^{-1})$. Following Nkurunziza (2012), the random quantity ψ_T captures the information from the sample as well as the prior information. Further, by Nkurunziza and Ahmed (2011) among others, the estimator $\hat{\theta}^S$ is not a convex combination of the UE and RE since $1-(nd-2)\psi_T^{-1} < 0$ whenever $\psi_T < (nd-2)$. So it may change the sign of UE $\hat{\theta}(\hat{\phi})$ and may cause an over-shrinking problem. To aviod the problem, let $a^+ = \max\{0, a\}$. We consider the *positive-part shrinkage estimator* (PSE) which is defined as

$$\hat{\theta}^{S+} = \widetilde{\theta}(\hat{\phi}) + [1 - (nd-2)\psi_T^{-1}]^+(\hat{\theta}(\hat{\phi}) - \widetilde{\theta}(\hat{\phi})). \tag{4.23}$$

Chapter 5

Relative efficiency of estimators

In this chapter, we first present the asymptotic distributional risk (ADR) of the proposed estimators and we study the risk performance of these estimators. The chapter is organized in two sections. Section 5.1 presents the ADR of the UE, RE, and the ADR of SEs. In Section 5.2, we compare the relative performance among these estimators via their ADRs.

5.1 Asymptotic distributional risk

In order to evaluate the performance of the proposed estimators, it is convenient to compare their asymptotic distributional risks (ADR). For more details about the ADR, we refer to Sen and Saleh (1987), Saleh (2006) among others. For an estimator $\hat{\theta}^{\star}$ of θ, we consider a quadratic loss function of the form

$$L(\hat{\theta}^{\star}, \theta; W) = \mathrm{Tr}\left[\sqrt{T}(\hat{\theta}^{\star} - \theta)W\sqrt{T}(\hat{\theta}^{\star} - \theta)'\right], \tag{5.1}$$

where W is a $2(p+d) \times 2(p+d)$ symmetric positive semi-definite weighting matrix, and $\hat{\theta}^{\star}$ refers to $\hat{\theta}(\hat{\phi})$, $\widetilde{\theta}(\hat{\phi})$, $\hat{\theta}^{S}$, and $\hat{\theta}^{S+}$. Further, let ϵ be the random matrix such that $\sqrt{T}(\hat{\theta}^{\star}-\theta)' \xrightarrow[T\to\infty]{d} \epsilon$. Following Nkurunziza and Ahmed (2011) and references therein, the ADR is defined as

$$\mathrm{ADR}(\hat{\theta}^{\star},\theta,W)=\mathrm{E}(\mathrm{Tr}(\epsilon' W\epsilon)). \tag{5.2}$$

The following theorem gives the ADR of the UE and RE.

Theorem 5.1. *Suppose that the conditions for Proposition* 4.6 *hold. Then*

$\mathrm{ADR}(\hat{\theta}(\hat{\phi}),\theta,W)=\mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1})$ and

$$\begin{aligned}\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W)&=\mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1})-\mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)-\mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1})\\&\quad+\mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)+\mathrm{Tr}(J_7WJ_7'),\end{aligned}$$

where J_3, J_7 *are defined in* (3.27) *and* (3.29) *respectively.*

The proof is provided in the Appendix B. We also derive the following theorem which gives the ADR of SEs.

Theorem 5.2. *Suppose that the conditions for Proposition* 4.6 *hold. Then*

$$\begin{aligned}\mathrm{ADR}(\hat{\theta}^{S},\theta,W)&=\mathrm{ADR}(\hat{\theta}(\hat{\phi}),\theta,W)-\mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\\&\quad-(nd-2)(2\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)]-(nd-2)\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)])\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\&\quad+\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma)\\&\quad+((nd)^2-4)\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7WJ_7');\end{aligned}$$

$$\begin{aligned}\mathrm{ADR}(\hat{\theta}^{S+},\theta,W)&=\mathrm{ADR}(\hat{\theta}^{S},\theta,W)\\&\quad+2\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)<nd-2\}}]\,\mathrm{Tr}(J_7WJ_7')\\&\quad-\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)<nd-2\}}]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma)\\&\quad-\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)<nd-2\}}]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\&\quad-\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)<nd-2\}}]\,\mathrm{Tr}(J_7WJ_7').\end{aligned}$$

The proof is provided in the Appendix B.

5.2 Risk analysis

In the previous section, we have obtained the ADRs of the proposed estimators. In this section, we compare the relative performance of these estimators via their ADRs.

5.2.1 Comparison between UE and RE

In this subsection, we derive a result which shows that near the null hypothesis, the RE dominates the UE. The derived result also shows that the UE dominates the RE as one moves away from the null hypothesis.

Proposition 5.1. *Suppose that the conditions of Theorem* 5.1 *hold and let* $W = L_2 C L_2'$ *such that the matrix* C *is a* $n \times n$ *real positive semidefinite symmetric matrix, then* $\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}), \theta, W) \leq \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W)$ *if* $\Delta \leq \frac{\mathrm{Tr}(\Sigma \otimes (AC))}{\lambda_{max}(\Sigma \otimes (AC))}$, *where* $A = L_2' \Sigma_2^{-1} L_2$.

Proof. From Theorem 5.6, we have

$$\begin{aligned}\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}), \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) = & - \mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W \Sigma_2^{-1} L_2 J_3) + \mathrm{Tr}(J_7 W J_7') \\ & + \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W \Sigma_2^{-1}(L_2 J_3 - I)).\end{aligned}$$

One can verify $\mathrm{Tr}(W \Sigma_2^{-1}(L_2 J_3 - I)) = 0$ and $\mathrm{Tr}(W \Sigma_2^{-1} L_2 J_3) = \mathrm{Tr}(CA)$. Thus, $\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}), \theta, W) \leq \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W)$ whenever $-\mathrm{Tr}(\Sigma)\,\mathrm{Tr}(CA) + \mathrm{Tr}(J_7 W J_7') \leq 0$. Further, note that $\Xi = L_2 (L_2' \Sigma_2^{-1} L_2)^{-1} L_2'$ and $\Delta = \mathrm{Tr}(J_7 \Xi J_7' \Sigma^{-1})$, we get

$$\mathrm{Tr}(J_7 L_2 A^{-1} L_2' J_7' \Sigma^{-1}) = \mathrm{Vec}(L_2' J_7')'(I_d \otimes A^{-1})(\Sigma^{-1} \otimes I_n)\mathrm{Vec}(L_2' J_7').$$

Then, we have

$$\Delta = \mathrm{Vec}(L_2' J_7')'(\Sigma^{-1} \otimes A^{-1})\mathrm{Vec}(L_2' J_7'). \tag{5.3}$$

Also, note that $\mathrm{Tr}(J_7 W J_7') = \mathrm{Tr}(J_7 L_2 C L_2' J_7')$. Similarly, we get

$$\mathrm{Tr}(J_7 W J_7') = \mathrm{Vec}(L_2' J_7')'(I_d \otimes C)\mathrm{Vec}(L_2' J_7'). \tag{5.4}$$

Since $(\Sigma^{-1} \otimes A^{-1})^{-1}(I_d \otimes C) = \Sigma \otimes (AC)$, let $\lambda_{max}(M)$ and $\lambda_{min}(M)$ represent the largest and smallest eigenvalues of a matrix M respectively. Then

$$\lambda_{min}(\Sigma \otimes (AC)) \le \frac{\mathrm{Tr}(J_7 W J_7')}{\Delta} \le \lambda_{max}(\Sigma \otimes (AC)). \tag{5.5}$$

Thus, we get $-\mathrm{Tr}(\Sigma)\,\mathrm{Tr}(CA) + \mathrm{Tr}(J_7 W J_7') \le 0$ if $\Delta \le \frac{\mathrm{Tr}(\Sigma \otimes (AC))}{\lambda_{max}(\Sigma \otimes (AC))}$, which completes the proof. $\square$

5.2.2 Comparison between UE and SEs

In this subsection, we present a result which shows that $\hat{\theta}^{S+}$ dominates $\hat{\theta}^{S}$, and thus also dominates the UE. Thus, the derived result also shows that as one moves far away from the null hypothesis, the SEs dominate the RE.

Proposition 5.2. *Suppose that the conditions of Theorem* 5.1 *hold and let* $W = L_2 C L_2'$ *such that the matrix* C *is a* $n \times n$ *positive semidefinite symmetric matrix that satisfies* $\frac{\lambda_{max}(\Sigma \otimes (AC))}{\mathrm{Tr}(\Sigma \otimes (AC))} \le \frac{2}{nd+2}$, *where* $A = L_2' \Sigma_2^{-1} L_2$. *Then,*
$\mathrm{ADR}(\hat{\theta}^{S+}, \theta, W) \le \mathrm{ADR}(\hat{\theta}^{S}, \theta, W) \le \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W)$, *for all* $\Delta \ge 0$.

Proof. From Theorem 5.2 and, we have

$$\begin{aligned}
\mathrm{ADR}(\hat{\theta}^{S}, \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) &= -\mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3)) \\
&\quad - (nd-2)(2\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)] - (nd-2)\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)])\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \\
&\quad + \mathrm{E}[(1 - (nd-2)\chi^{-2}_{nd}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3))\,\mathrm{Tr}(\Sigma) \\
&\quad + ((nd)^2 - 4)\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7'),
\end{aligned}$$

by the identity in Saleh (2006, p. 32), we have

$\mathrm{E}[\chi_{nd+2}^{-2}(\Delta)] = \Delta \mathrm{E}[\chi_{nd+4}^{-4}(\Delta)] + (nd-2)\mathrm{E}[\chi_{nd+2}^{-4}(\Delta)]$, we get

$$
\begin{aligned}
\mathrm{ADR}(\hat{\theta}^S, \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) = & - \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3)) \\
& - (nd-2)(2\Delta \mathrm{E}[\chi_{nd+4}^{-4}(\Delta)] + (nd-2)\mathrm{E}[\chi_{nd+2}^{-4}(\Delta)])\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \\
& + \mathrm{E}[(1-(nd-2)\chi_{nd}^{-2}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3))\,\mathrm{Tr}(\Sigma) \\
& + ((nd)^2 - 4)\mathrm{E}[\chi_{nd+4}^{-4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7'). \qquad (5.6)
\end{aligned}
$$

Notice that $\Delta = \mathrm{Tr}(J_7 \Xi J_7' \Sigma^{-1}) = \mathrm{Tr}(\Sigma^{-1/2} J_7 \Xi J_7' \Sigma^{-1/2}) \geq 0$ since $(L_2' \Sigma_2^{-1} L_2)^{-1}$ is a positive definite matrix, therefore $\mathrm{Tr}(\Sigma^{-1/2} J_7 \Xi J_7' \Sigma^{-1/2}) \geq 0$ with equality holding if and only if $\Sigma^{-1/2} J_7 L_2 = 0$. Also, noting that $\Sigma_2^{-1} L_2 J_3$ and W are symmetric positive semidefinite matrices, we have $\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) = \mathrm{Tr}(W^{1/2}\Sigma_2^{-1} L_2 J_3 W^{1/2}) \geq 0$. Moreover, note that $\mathrm{E}[\chi_{nd+2}^{-4}(\Delta)] \geq 0$, $\mathrm{E}[\chi_{nd+4}^{-4}(\Delta)] \geq 0$ and $nd > 2$. Further, notice that whenever the weighting matrix $W = L_2 C L_2'$ with C an $n \times n$ real symmetric matrix, then we get

$$
W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3) = L_2 C L_2' (\Sigma_2^{-1} - \Sigma_2^{-1} L_2 (L_2' \Sigma_2^{-1} L_2)^{-1} L_2' \Sigma_2^{-1}) = 0. \qquad (5.7)
$$

Therefore, for $\Delta = 0$, we have $J_7 L_2 = 0$ since Σ is positive definite, thus, $\mathrm{Tr}(J_7 W J_7') = \mathrm{Tr}(J_7 L_2 C L_2' J_7') = 0$ and by combining (5.6) and (5.7), we get

$$
\begin{aligned}
\mathrm{ADR}(\hat{\theta}^S, \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) = & \\
& - (nd-2)^2 \mathrm{E}[\chi_{nd+2}^{-4}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \leq 0.
\end{aligned}
$$

For $\Delta > 0$, we have

$$\begin{aligned}
\mathrm{ADR}(\hat{\theta}^S, \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) &= - \operatorname{Tr}(J_1 L_1 \Sigma) \operatorname{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3)) \\
&+ \mathrm{E}[(1-(nd-2)\chi_{nd}^{-2}(\Delta))^2] \operatorname{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3)) \operatorname{Tr}(\Sigma) \\
-2(nd-2)\Delta \operatorname{Tr}(W\Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma) \mathrm{E}[\chi_{nd+4}^{-4}(\Delta)] & \left(1 - \frac{(nd+2)\operatorname{Tr}(J_7 W J_7')}{2\Delta \operatorname{Tr}(W\Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma)}\right) \\
&- (nd-2)^2 \mathrm{E}[\chi_{nd+2}^{-4}(\Delta)] \operatorname{Tr}(W\Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma). \qquad (5.8)
\end{aligned}$$

Note that

$$-2(nd-2)\Delta \operatorname{Tr}(W\Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma) \mathrm{E}[\chi_{nd+4}^{-4}(\Delta)] \left(1 - \frac{(nd+2)\operatorname{Tr}(J_7 W J_7')}{2\Delta \operatorname{Tr}(W\Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma)}\right) \leq 0,$$

whenever

$$1 - \frac{(nd+2)\operatorname{Tr}(J_7 W J_7')}{2\Delta \operatorname{Tr}(W\Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma)} \geq 0. \qquad (5.9)$$

Therefore by combining (5.7), (5.8), and (5.9), we get

$$\mathrm{ADR}(\hat{\theta}^S, \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) \leq 0,$$

if

$$1 - \frac{(nd+2)\operatorname{Tr}(J_7 W J_7')}{2\Delta \operatorname{Tr}(L_2 C L_2' \Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma)} \geq 0. \qquad (5.10)$$

Let $\lambda_{max}(M)$ and $\lambda_{min}(M)$ represent the largest and smallest eigenvalues of a matrix M respectively. Note that $(\Sigma^{-1} \otimes A^{-1})^{-1}(I_d \otimes C) = \Sigma \otimes (AC)$. From (5.3), (5.4) and Theorem A.2 in the Appendix, we get

$$\lambda_{min}(\Sigma \otimes (AC)) \leq \frac{\operatorname{Tr}(J_7 W J_7')}{\Delta} \leq \lambda_{max}(\Sigma \otimes (AC)).$$

Also, we have $\operatorname{Tr}(L_2 C L_2' \Sigma_2^{-1} L_2 J_3) = \operatorname{Tr}(L_2 C L_2' \Sigma_2^{-1}) = \operatorname{Tr}(AC)$. Then, we get

$$1 - \frac{(nd+2)\operatorname{Tr}(J_7 W J_7')}{2\Delta \operatorname{Tr}(L_2 C L_2' \Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma)} \geq 1 - \frac{(nd+2)\lambda_{max}(\Sigma \otimes (AC))}{2 \operatorname{Tr}(AC) \operatorname{Tr}(\Sigma)}.$$

Since $\operatorname{Tr}(AC)\operatorname{Tr}(\Sigma) = \operatorname{Tr}(\Sigma \otimes (AC))$. By (5.10), we have

$$\mathrm{ADR}(\hat{\theta}^S, \theta, W) - \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) \leq 0,$$

if

$$1 - \frac{(nd+2)\lambda_{max}(\Sigma \otimes (AC))}{2\operatorname{Tr}(\Sigma \otimes (AC))} \geq 0 \Leftrightarrow \frac{\lambda_{max}(\Sigma \otimes (AC))}{\operatorname{Tr}(\Sigma \otimes (AC))} \leq \frac{2}{nd+2}. \tag{5.11}$$

Further, note that from Theorem 5.2, we have

$$\begin{aligned}
\mathrm{ADR}(\hat{\theta}^{S+}, \theta, W) - \mathrm{ADR}(\hat{\theta}^S, \theta, W) &= 2\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)<nd-2\}}]\operatorname{Tr}(J_7 W J_7') \\
&\quad - \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)<nd-2\}}]\operatorname{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1}L_2J_3))\operatorname{Tr}(\Sigma) \\
&\quad - \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)<nd-2\}}]\operatorname{Tr}(W\Sigma_2^{-1}L_2J_3)\operatorname{Tr}(\Sigma) \\
&\quad - \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)<nd-2\}}]\operatorname{Tr}(J_7 W J_7').
\end{aligned}$$

In order to study the risks of $\hat{\theta}^S$ and $\hat{\theta}^{S+}$, we study the sign of each term in the equation above. Note that W is symmetric and positive semidefinite, then it can be rewritten as $W = PP'$ for some P, and $\Sigma_2^{-1}L_2J_3$ is also symmetric and positive semidefinite, therefore, $\operatorname{Tr}(W\Sigma_2^{-1}L_2J_3) = \operatorname{Tr}(P'\Sigma_2^{-1}L_2J_3P) \geq 0$. Also, $\operatorname{Tr}(J_7WJ_7') \geq 0$ due to W being symmetric and positive semidefinite, and $\operatorname{Tr}(\Sigma) > 0$ since Σ is positive definite. Moreover, since

$$\begin{aligned}
(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)<nd-2\}} &\geq 0, \\
(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)<nd-2\}} &\geq 0, \\
(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)<nd-2\}} &\geq 0.
\end{aligned}$$

One can verify that

$$\begin{aligned}
\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)<nd-2\}}] &\geq 0, \\
\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)<nd-2\}}] &\geq 0, \\
\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)<nd-2\}}] &\geq 0.
\end{aligned}$$

For a given choice of the weighting matrix W, we have

$$
\begin{aligned}
&- \mathrm{E}[(1-(nd-2)\chi_{nd}^{-2}(\Delta))^2 \mathbb{I}_{\{\chi_{nd}^2(\Delta)<nd-2\}}] \operatorname{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3)) \operatorname{Tr}(\Sigma) \\
&- \mathrm{E}[(1-(nd-2)\chi_{nd+2}^{-2}(\Delta))^2 \mathbb{I}_{\{\chi_{nd+2}^2(\Delta)<nd-2\}}] \operatorname{Tr}(W \Sigma_2^{-1} L_2 J_3) \operatorname{Tr}(\Sigma) \\
&- \mathrm{E}[(1-(nd-2)\chi_{nd+4}^{-2}(\Delta))^2 \mathbb{I}_{\{\chi_{nd+4}^2(\Delta)<nd-2\}}] \operatorname{Tr}(J_7 W J_7') \leq 0. \qquad (5.12)
\end{aligned}
$$

For the sign of $2\mathrm{E}[(1-(nd-2)\chi_{nd+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{nd+2}^2(\Delta)<nd-2\}}] \operatorname{Tr}(J_7 W J_7')$, note that

$$
(1-(nd-2)\chi_{nd+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{nd+2}^2(\Delta)<nd-2\}} \leq 0,
$$

then we have

$$
\mathrm{E}[(1-(nd-2)\chi_{nd+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{nd+2}^2(\Delta)<nd-2\}}] \leq 0.
$$

Therefore,

$$
2\mathrm{E}[(1-(nd-2)\chi_{nd+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{nd+2}^2(\Delta)<nd-2\}}] \operatorname{Tr}(J_7 W J_7') \leq 0. \qquad (5.13)
$$

Combining (5.12) and (5.13), we have

$$
\operatorname{ADR}(\hat{\theta}^{S+}, \theta, W) - \operatorname{ADR}(\hat{\theta}^{S}, \theta, W) \leq 0,
$$

for all $\Delta \geq 0$, which completes the proof. $\square$

Chapter 6

Numerical study

In this chapter, we examine the performance of the estimators $\hat{\theta}(\hat{\phi})$, $\widetilde{\theta}(\hat{\phi})$, $\hat{\theta}^S$, and $\hat{\theta}^{S+}$ in case of a 4-dimensional stochastic process. Firstly, we use Euler-Maruyama discretization to generate the stochastic process in (2.1), then we calculate the weighted squared error of each estimator based on different non-centrality parameter Δ with the weighting matrix $W = L_2(L_2'\Sigma_2^{-1}L_2)^{-1}L_2'$. By 1000 replications, we compute the ADR of each estimator as well as the empirical relative mean squared efficiencies (RMSE), which is defined as

$$\text{RMSE}(\widetilde{\theta}^\star) = \text{ADR}(\hat{\theta}(\hat{\phi}), \theta, W)/\text{ADR}(\widetilde{\theta}^\star, \theta, W)$$

where $\widetilde{\theta}^\star$ represents for different estimators. Thus, RMSE shows a degree of superiority of the estimator over UE, a gold standard. In this simulation study, we define the increment of time in the interval $[0, T]$ as $v = 0.01$, and T is choosen as $T = 50$ and $T = 100$ for two scenarios. Also, we choose a 2-dimensional periodic incomplete orthogonal set of functions $[1, \sqrt{2}\cos(\frac{\pi t}{v})]$, $t \in [0, T]$ as our base functions $\varphi(t)$. The true parameter θ is set as:

$$\theta = \left[\begin{array}{cccccc|cccccc} 4 & 1 & 6 & 4 & 3 & 1 & 12 & 2 & 6 & 4 & 3 & 1 \\ 9 & 2 & 4 & 5 & 4 & 1 & 27 & 4 & 4 & 5 & 4 & 1 \\ 6 & 3 & 3 & 3 & 4 & 2 & 18 & 6 & 3 & 3 & 4 & 2 \\ 5 & 4 & 5 & 2 & 2 & 3 & 15 & 8 & 5 & 2 & 2 & 3 \end{array}\right].$$
$$\qquad \underbrace{\quad}_{\mu_1} \quad \underbrace{\qquad\qquad}_{A_1} \quad \underbrace{\quad}_{\mu_2} \quad \underbrace{\qquad\qquad}_{A_2}$$

Thus, $A_1 = A_2$ are positive-definite matrices, we have the parameter μ which changes after the change-point (i.e. the coefficient for the first element of the base functions $\varphi(t)$ tripled, and the coefficient for the second element of the base functions $\varphi(t)$ doubled) and the parameter A remains the same. For simplicity, we choose $\Sigma = I_4$. We also choose $\phi = 0.4$. Let $0 < t_0 < ... < t_n = T$ be a partition on a given time period $[0, T]$ with a constant increment $\tau = t_{i+1} - t_i$, then $\hat{\Sigma} = \mathrm{diag}(\widehat{\sigma_1}^2, \widehat{\sigma_2}^2, \widehat{\sigma_3}^2, \widehat{\sigma_4}^2)$ is a strongly consistent estimator for $\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)$, where $\widehat{\sigma_i}^2 = \frac{1}{T}\sum_{j=1}^{n}(X_i(j\tau) - X_i((j-1)\tau))^2$. For the change-point, we use the method similar to that given in Chen and Nkurunziza (2015). Let $Y_i = X_{t_{i+1}} - X_{t_i}$ and $Z_i = (1, \sqrt{2}\cos(\frac{\pi t}{v}), -X'_{t_i})(\tau)$. The consistent estimator for the change-point ϕ is obtained by $\hat{\phi} = \underset{\phi}{\mathrm{argmin}}\ \mathrm{SSE}(\phi)$, where $\mathrm{SSE}(\phi) = \sum_{t_i \in [0,T]} (Y_i - \hat{\theta}(\phi)Z_i)'(Y_i - \hat{\theta}(\phi)Z_i)$ and $\hat{\theta}(\phi) = \mathbb{I}_{\{\frac{t_i}{T} \leq \phi\}}\widehat{\theta_1} + \mathbb{I}_{\{\frac{t_i}{T} > \phi\}}\widehat{\theta_2}$, where $\left[\begin{array}{c:c} \widehat{\theta_1} & \widehat{\theta_2} \end{array}\right]$ forms the MLE $\hat{\theta}$ with the change-point given by ϕ. We compute the estimates of the rate of the change-point, and below, we present histograms in Figure 6.1-6.3 , which show that the method used locates very well the change-point. Indeed, the histograms show that the pick of the estimates corresponds to the exact value of $\phi = 0.4$. The distribution of the obtained estimates are unimodal and symmetric with respect to the exact value of $\phi = 0.4$. For the linear restrictions, we choose $L_1 = (1, -1, 0, 0)$ and $d_1 = L_1\theta$, also we choose $L_2 = \left[\begin{array}{c:c} 2I_6 & -I_6 \end{array}\right]'$. Under the null hypothesis, d_2 is calculated as θL_2, i.e.,

$$d_2 = \begin{bmatrix} -4 & 0 & 6 & 4 & 3 & 1 \\ -9 & 0 & 4 & 5 & 4 & 1 \\ -6 & 0 & 3 & 3 & 4 & 2 \\ -5 & 0 & 5 & 2 & 2 & 3 \end{bmatrix}.$$

Under the alternative hypothesis defined in (3.25), let $r_2 = kd_2$, where $k = 1, .., 6$. From previous sections, we know that non-centrality parameter $\Delta = \mathrm{Tr}(J_7 \Xi J_7' \Sigma^{-1})$ depends on r_2 since $J_7 = J_1 L_1 r_2 J_3 - r_2 J_3$. Thus, different values of r_2 corresponds to different levels of Δ. For $T = 50$ and $T = 100$, we plot respectively the RMSEs of the proposed estimators versus Δ in the Figures 6.4 and 6.5.

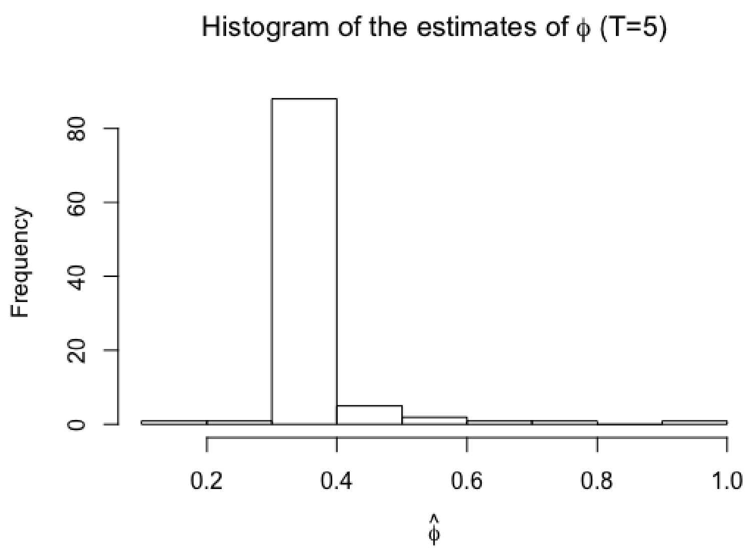

Figure 6.1: Histogram of the estimates of ϕ for T=5

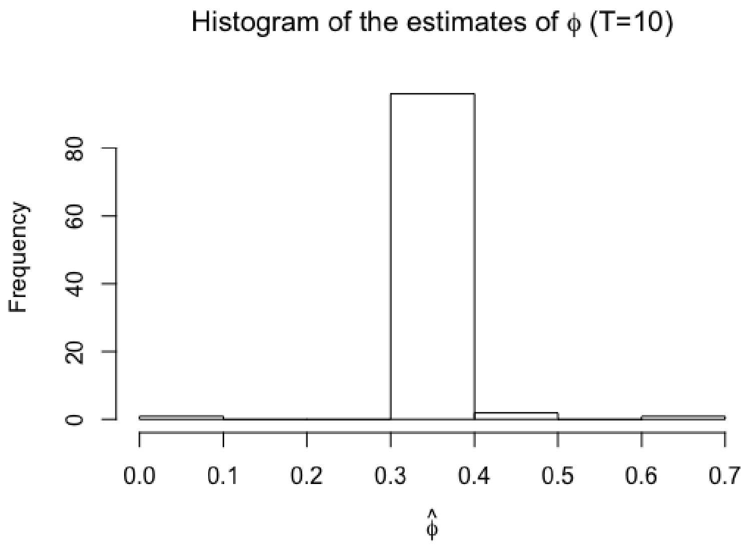

Figure 6.2: Histogram of the estimates of ϕ for T=10

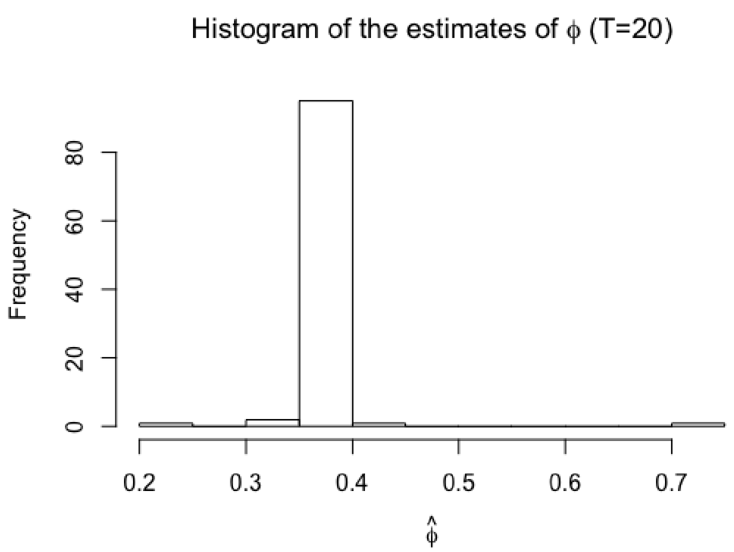

Figure 6.3: Histogram of the estimates of ϕ for T=20

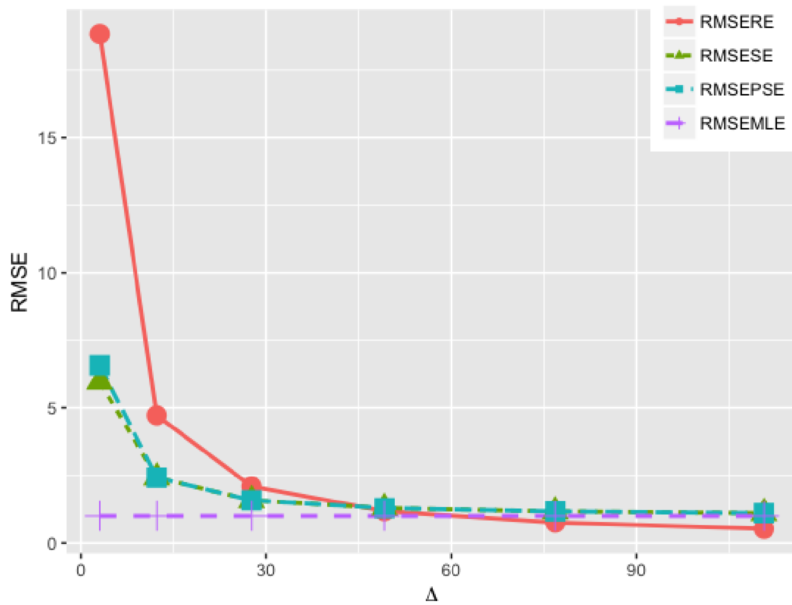

Figure 6.4: RMSE of RE, SE, PSE versus Δ (T = 50)

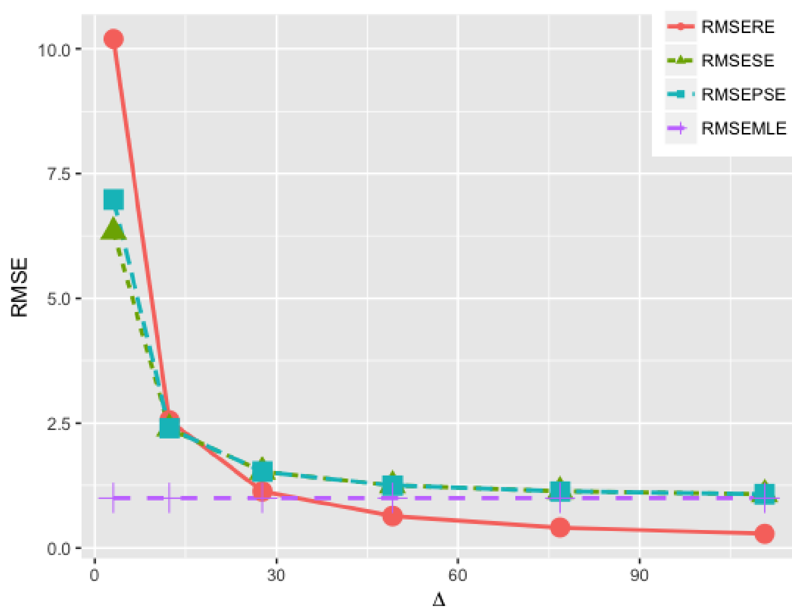

Figure 6.5: RMSE of RE, SE, PSE versus Δ (T = 100)

Further, by setting $d_2 = 0$ and $L_2 = \left[\begin{array}{c:c} I_6 & -I_6 \end{array}\right]'$, we simulate the case with the absence of the change-point for T=20 and T=100. We plot the RMSEs in the following diagrams.

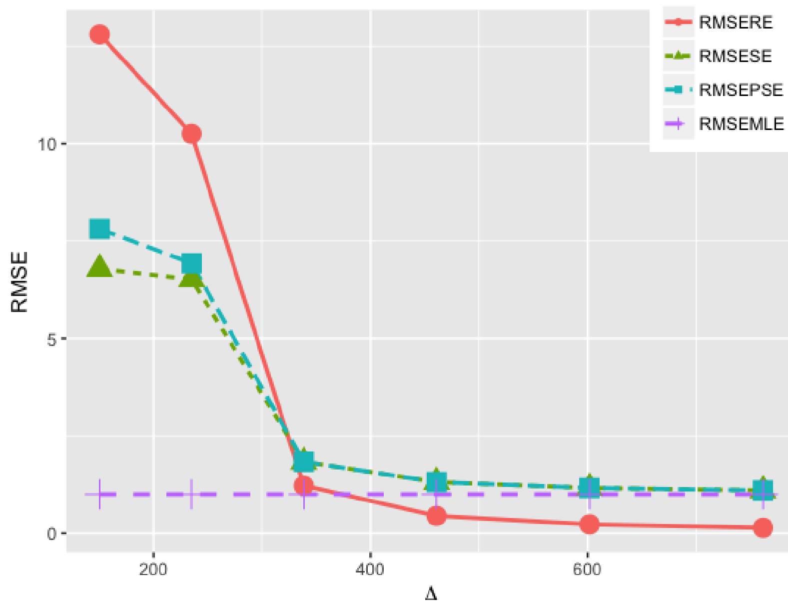

Figure 6.6: RMSEs versus Δ (T = 20)

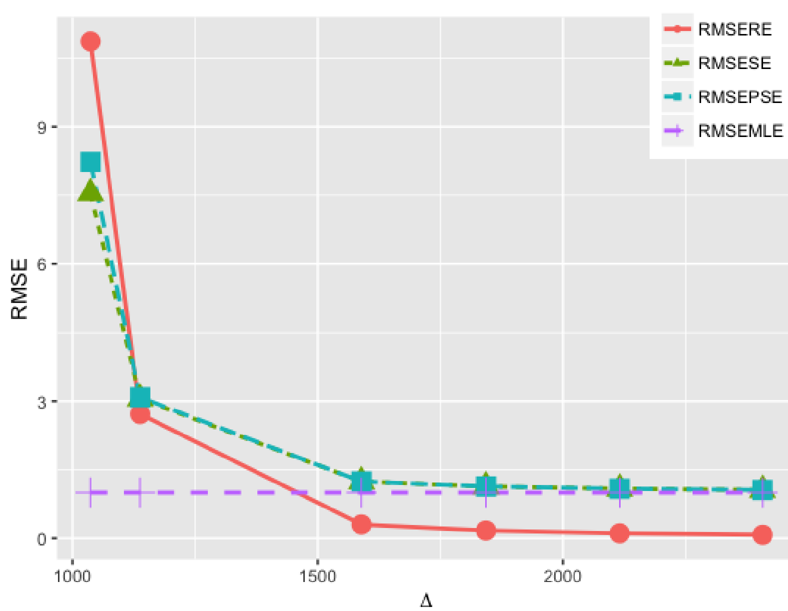

Figure 6.7: RMSEs versus Δ (T = 100)

According to Figure 6.4 - 6.7, it is clear that the shrinkage estimators outperform over the UE. In addition, the positive shrinkage estimator dominates the shrinkage estimator. These simulation results coincide with the theoretical results that are estabilshed in this thesis. Also, around a neighbourhood of the hypothesized restriction, the RE dominates any other estimators; however, it performs much worse as the hypothesized constraint is severely violated. Further, for the test of (2.3), we simulate the empirical power of the test versus Δ and T, and the results are presented in the Figures 6.8 - 6.10 .

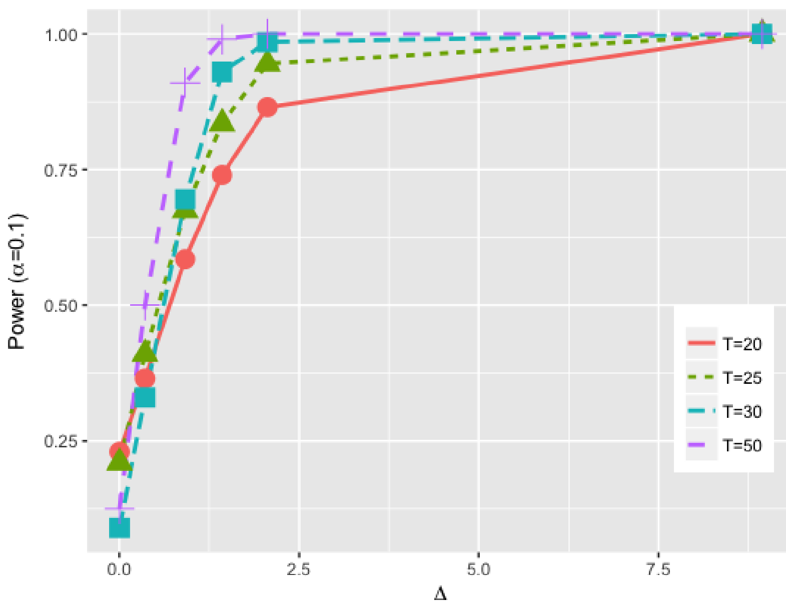

Figure 6.8: Empirical power of the test $\alpha = 0.1$

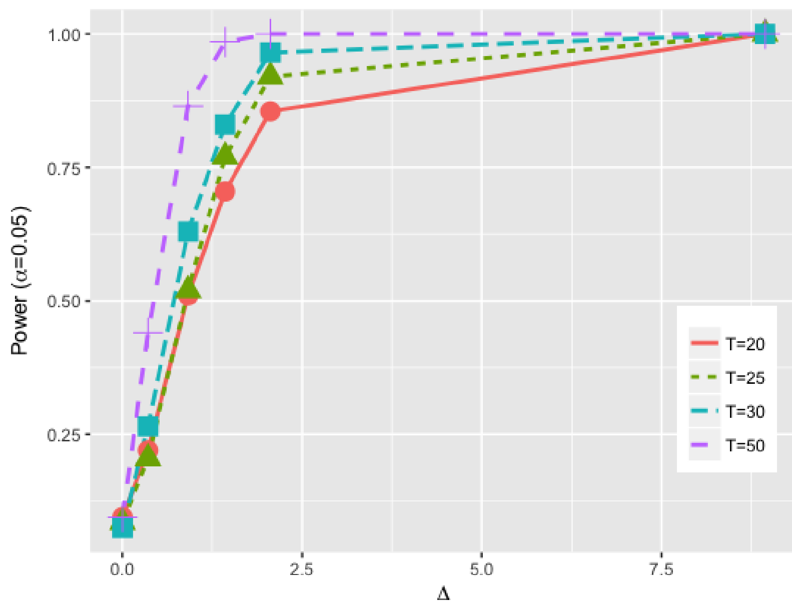

Figure 6.9: Empirical power of the test $\alpha = 0.05$

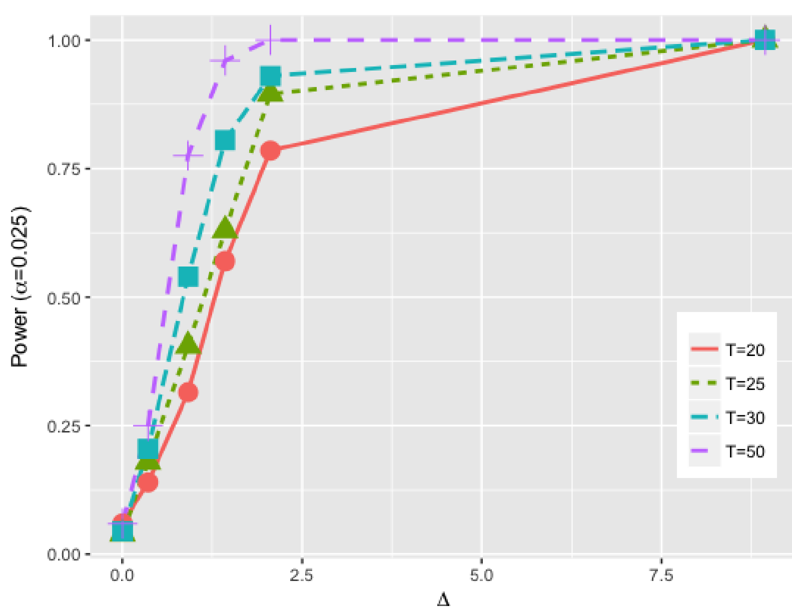

Figure 6.10: Empirical power of the test $\alpha = 0.025$

Figures 6.8 - 6.10 confirm the established theoretical result given in Section 4.3. In particular, Figures 6.8 - 6.10 show that the proposed test is consistent.

Conclusion

This thesis generalizes in five ways some results in Dehling *et al.* (2010, 2014), Chen *et al.* (2017) as well as that in Nkurunziza and Zhang (2018). First, we propose inference methods in the context of multivariate generalized O-U processes. Thus, the target parameter is a matrix. As a preliminary step, we present some results in the no change-point case. Second, we extend the results to the case of a known change-point. In particular, we prove the existence of the UMLE and RMLE, also, we present the joint asymptotic normality of the UMLE and RMLE. Third, we present the UE, RE, and SEs as well as their joint asymptotic normality in the case of the unknown change-point. Forth, we propose a test for testing the hypothesized restriction. The proposed test includes some special cases such as testing the absence of a change-point and testing the nonexistence of the seasonality factor. Fifth, we derive the asymptotic local power and prove that the proposed test is consistent. Sixth, we propose SEs and we derive the ADRs of the UE, RE and SEs. We also compare the relative efficiency of the proposed estimators via their ADRs. By theoretical approach and by the simulation study, our findings show that for a suitable choice of the weighting matrix W, the PSE dominates the SE, and SE dominates the UE. Also, the RE is the best in the neighborhood of the null hypothesis, but it performs poorly as one moves far away from the hypothesized restriction.

BIBLIOGRAPHY

[1] Aalen, O. & Gjessing, H. (2004). Survival models based on the Ornstein-Uhlenbeck process, Lifetime Data Analysis. 10(4), pp. 407-423.

[2] Ahmed, S. E., Nkurunziza, S., & Liu, S. (2009). Improved Estimation Strategy in Multi-Factor Vasicek Model. In Statistical Inference, Econometric Analysis and Matrix Algebra (pp. 255-270). Physica-Verlag, Heidelberg.

[3] Benth, F., Koekebakker, S., & Taib, C. (2015). Stochastic dynamical modelling of spot freight rates, IMA Journal of Management Mathematics. 26(3), pp. 273-297.

[4] Billingsley, Patrick (1995). Probability and Measure (3 ed.). John Wiley & Sons. ISBN 978-0-471-00710-4, pp. 383.

[5] Chen, F., & Nkurunziza, S. (2015). Optimal method in Multiple Regression with Structural Changes. Bernoulli, Volume 21, Number 4, pp. 2217-2241.

[6] Chen, F., Mamon, R., & Nkurunziza, S. (2017). Inference for a change-point problem under a generalised Ornstein-Uhlenbeck setting. Annals of the Institute of Statistical Mathematics (Accepted).

[7] Dehling, H., Franke, B., & Kott, T. (2010). Drift estimation for a periodic mean reversion process. Stat Inference Stoch Process, 13, pp. 175-192.

[8] Dehling, H., Franke, B., Kott, T., & Kulperger, R. (2014). Change point testing for the drift parameters of a periodic mean reversion process. Stat Inference Stoch

Process, 17(1), pp. 1-18.

[9] Erlwein,C., Benth, F., & Mamon, R.S. (2010). HMM filtering and parameter estimation of an electricity spot price model, Energy Economics, 32(5), pp. 1034-1043.

[10] Gombay, E. (2010). Change detection in linear regression with time series errors. Can J Stat 38(1): pp. 65-79

[11] Horn, R. A. & Johnson, C. R. (1994). Topics in Matrix Analysis. Cambridge University Press, Cambridge, England, pp. 208.

[12] James, W., & Stein, C. (1961, June). Estimation with quadratic loss. In Proceedings of the fourth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 1961, pp. 361-379).

[13] Klenke, A. (2013). Probability theory: a comprehensive course. Springer-Verlag, London.

[14] Kollo, T., & Rosen, D. V. (2011). Advanced Multivariate Statistics with Matrices (Vol. 579). Springer.

[15] Kutoyants, Y. A. (2004). Statistical inference for ergodic diffusion processes. Springer-Verlag, London.

[16] Langetieg, T.C. (1980). A multivariate model of the term structure. The Journal of Finance, 35(1): pp. 71-97.

[17] Liang, Z., Yuen, K., & Guo, J. (2011). Optimal proportional reinsurance and investment in a stock market with Ornstein-Uhlenbeck process, Insurance: Mathematics and Economics. 49(2), pp. 207-215.

[18] Liptser, R. S., & Shiryayev, A. N. (2001). Statistics of Random Processes I: I. General Theory (Vol. 1). Springer-Verlag, Berlin Heidelberg.

[19] Lu, Q. & Lund, B.R. (2007) Simple linear regression with multiple level shifts. Can J Stat 35(3): pp. 447-458

[20] Mathai, A. M., & Provost, S. B. (1992). Quadratic forms in random variables: theory and applications. CRC Press, New York.

[21] Nkurunziza, S., & Ahmed, S. E. (2010). Shrinkage drift parameter estimation for multi-factor Ornstein-Uhlenbeck processes. Applied Stochastic Models in Business and Industry, 26(2), pp. 103-124.

[22] Nkurunziza, S., & Ahmed, S. E. (2011). Estimation strategies for the regression coefficient parameter matrix in multivariate multiple regression. Statistica Neerlandica 65(4): pp. 387-406.

[23] Nkurunziza, S. (2012). Shrinkage Strategies In Some Multiple Multi-factor Dynamical Systems. ESAIM: PS, 16, pp. 139-150.

[24] Nkurunziza, S., & Zhang, P.P.(2018). Estimation and testing in generalized mean-reverting processes with change-point. SISP, 21, pp. 191-215.

[25] Nkurunziza, S., & Shen, L. (2018). Inference in a multivariate generalized mean-reverting process with a change-point. Statistical Inference for Stochastic Processes (Submitted).

[26] Saleh, A. M. E. (2006). Theory of preliminary test and Stein-type estimation with applications (Vol. 517). John Wiley & Sons, New Jersey.

[27] Sen, P. K. & Saleh, A. M. E. (1987). On Preliminary Test and Shrinkage M-estimation in Linear Models. The Annals of Statistics, 15(4): pp. 1580-1592.

[28] Pigorsch, C. & Stelzer, R. (2009). A multivariate Ornstein-Uhlenbeck type stochastic volatility model. Available at http://www-m4.ma.tum.de.

[29] Robbins, M.W., Lund, B.R., Gallagher, C.M., & Lu, Q. (2011) Change points in the North Atlantic tropical cyclone record. JASA 106(493): pp. 89-99.

[30] Vasicek, O. (1977) An equilibrium characterization of the term structure. Journal of Financial Economics 5(2): pp. 177-188.

APPENDICES

A Theoretical background

Theorem A.1. $(\Omega, \mathcal{A}, \mathrm{P}, \tau)$ *is ergodic if and only if for all* A *,* $B \in \mathcal{A}$*, the measure preserving transformation* τ*is weakly-mixing.*

The proof is referred to Klenke (2013, Theorem 20.23, p.450).

Theorem A.2. *(Mathai and Provost, 1992, Theorem 2.4.7). Let* B *be any* $n \times n$ *positive definite matrix and* A *be an* $n \times n$ *symmetric matrix. Let* $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ *be the eigenvalues of* $B^{-1}A$ *with eigenvectors* q_1*,* q_2*, ...,* q_n *respectively. Then,* $\sup_x \left(\frac{x'Ax}{x'Bx}\right) = \lambda_n$*, and* $\inf_x \left(\frac{x'Ax}{x'Bx}\right) = \lambda_1$*, where* λ_1 *and* λ_n *are respectively the largest and smallest eigenvalues of* $B^{-1}A$*.*

Proposition A.1 (Proposition 1.21 Kutoyants, 2004)**.** *Let every* $T > 0$*,* $\theta \in \Theta$*, and i=1,...,*d_1*, j=1,...,*d_2*, define*

$$I_T(\theta) = \left(I_T^{(1)}(\theta), ..., I_T^{(d_1)}(\theta)\right)', \; I_T^{(i)}(t, \theta) = \sum_{j=1}^{d_2} \int_0^T h_T^{(i,j)}(\theta, t, \omega) dB_t^{(j)},$$

where $\mathrm{P}\left(\int_0^T \left(h_T^{(i,j)}(\theta, t, \omega)\right)^2 dt < \infty\right) = 1$*, for all i,j and* $\{B_t^{(1)}, ..., B_t^{(d_2)}, 0 \leq t \leq T\}$ *are* d_2 *independent Wiener processes. Suppose that there exists a (non-random) positive definite matrix* $\Sigma(\theta) = \left(\Sigma^{(i,m)}(\theta)\right)_{d_1 \times d_2}$ *such that*

$$\sum_{l=1}^{d_2}\int_0^T h_T^{(i,l)}(\theta,t,\omega)h_T^{(m,l)}(\theta,t,\omega)dt \xrightarrow[T\to\infty]{\mathrm{P}} \Sigma^{(i,m)}(\theta),$$ *uniformly with respect to* $\theta \in \Theta$, *then*

$$I_T(\theta) \xrightarrow[T\to\infty]{D} \mathcal{N}\big(0, \Sigma(\theta)\big)$$

uniformly with respect to $\theta \in \Theta$ *too.*

The proof is referred to Kutoyants (2004 Proposition 1.21, p.46).

Proposition A.2. *Let* A *and* B *be constant matrices of proper sizes. Then*

$$\frac{\partial(AXB)}{\partial X} = B \otimes A',$$
$$\frac{\partial(AYB)}{\partial X} = \frac{\partial Y}{\partial X}(B \otimes A').$$

The proof is referred to Kollo and Rosen (Proposition 1.4.4, p.129).

Proposition A.3. *Let* A *be any positive definite matrix, and let* λ_1 *and* λ_d *be the smallest and largest eigenvalues of* $A' + A$ *respectively. Then* $\sqrt{de^{-t\lambda_d}} \leq \|e^{-At}\|_F \leq \sqrt{de^{-t\lambda_1}}$, *for all* $t > 0$, *and* $\sqrt{de^{-t\lambda_1}} \leq \|e^{-At}\|_F \leq \sqrt{de^{-t\lambda_d}}$, *for all* $t < 0$, *and thus*

$$\lim_{t\to+\infty} e^{-At} = 0.$$

Proof. It is sufficient to prove that $\lim_{t\to\infty}\|e^{-At}\|_F = 0$, where $\|.\|_F$ denotes Frobenius norm, notice that

$$\|e^{-At}\|_F = \sqrt{\mathrm{Tr}(e^{-A't}e^{-At})} = \sqrt{\mathrm{Tr}(e^{-(A'+A)t})} = \sqrt{\mathrm{Tr}\left(\sum_{k=0}^{\infty}\frac{1}{k!}(-t)^k(A'+A)^k\right)}.$$

By sub-multipicative property of the Frobenius norm. i.e. $\|AB\|_F \leq \|A\|_F\|B\|_F$, we have:

$$\sum_{k=0}^{\infty}\frac{1}{k!}\|(-t)^k(A'+A)^k\|_F \leq \sum_{k=0}^{\infty}\frac{1}{k!}(t^2)^k\|(A'+A)\|_F^k = e^{t^2\|A'+A\|_F} < \infty.$$

Therefore

$$\|e^{-At}\|_F = \sqrt{\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \operatorname{Tr}[(A' + A)^k]}.$$

Moreover, since $A' + A$ is real symmertic, it can be diagonalized as $L\Lambda L'$, where $LL' = I$, and Λ is a diagonal matrix with diagonal entries equal to the eigenvalues of $A' + A$, we have

$$\begin{aligned}
\|e^{-At}\|_F &= \sqrt{\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \operatorname{Tr}[(L\Lambda L')^k]} = \sqrt{\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \operatorname{Tr}(L\Lambda^k L')} \\
&= \sqrt{\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \operatorname{Tr}(\Lambda^k)} = \sqrt{\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \sum_{j=1}^{d} \lambda_j^k}.
\end{aligned}$$

Since A is a positive definite matrix, we have $A' + A$ is also a positive definite matrix. Therefore, all the eigenvalues of $A' + A$ are strictly greater than 0, then $\left|\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \sum_{j=1}^{d} \lambda_j^k\right| \leq \sum_{k=0}^{\infty} \frac{1}{k!}|t|^k \sum_{j=1}^{d} \lambda_j^k = \sum_{j=1}^{d} \left(\sum_{k=0}^{\infty} \frac{(|t|\lambda_j)^k}{k!}\right) = \sum_{j=1}^{d} e^{\lambda_j |t|} < \infty, \forall t \in \mathbb{R}$. This gives

$$\|e^{-At}\|_F = \sqrt{\sum_{j=1}^{d}\sum_{k=0}^{\infty} \frac{1}{k!}(-t)^k \lambda_j^k} = \sqrt{\sum_{j=1}^{d} e^{-t\lambda_j}}$$

Now, let λ_1 be the smallest eigenvalue of $A' + A$, and let λ_d be the largest eigenvalue of $A' + A$, we have $e^{-\lambda_d t} \leq e^{-\lambda_j t} \leq e^{-\lambda_1 t}$, $\forall t > 0$. Then

$$\sqrt{de^{-\lambda_d t}} \leq \|e^{-At}\|_F \leq \sqrt{de^{-t\lambda_1}}. \tag{A.1}$$

Similarly, we have $e^{-\lambda_1 t} \leq e^{-\lambda_j t} \leq e^{-\lambda_d t}$, $\forall t < 0$, this proves the inequalities stated. Further, by taking limits both sides, we have $\lim_{t\to+\infty} \|e^{-At}\|_F = 0$, which completes the proof. $\square$

Proposition A.4 (Nkurunziza, 2012). *Suppose that the conditions of Corollary* (4.1) *hold and let W be nonnegative definite matrix. Then, for any real number c, we have*

$$\begin{aligned}
\mathrm{E}\{\mathrm{Tr}[(1-c\psi^{-1})^2\xi' W\xi]\} &= \mathrm{E}[(1-c\chi_{nd}^{-2}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma)\\
&\quad + \mathrm{E}[(1-c\chi_{nd+2}^{-2}(\Delta))^2]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\
&\quad + \mathrm{E}[(1-c\chi_{nd+4}^{-2}(\Delta))^2]\,\mathrm{Tr}(J_7WJ_7');\\
\mathrm{E}[(1-c\psi^{-1})\zeta' W\xi] &= -\mathrm{E}[(1-c\chi_{nd+2}^{-2}(\Delta))]J_7WJ_7'.
\end{aligned}$$

For the proof, we refer to Theorem 2.3 in Nkurunziza (2012).

Lemma A.1. *(Bessel's Inequality). Let H be a Hilbert space. If $\{\varphi_i : i = 1, ..., p\}$ is a finite orthonormal set in H, then for any $x \in H$, $\sum_{i=1}^{p} |\langle x, \varphi_i\rangle|^2 \leq \|x\|^2$.*

Lemma A.2. *Let $\{Y_t, t \geq 0\}$ be a d-dimensional stochastic process, $\{F_t, t \geq 0\}$ adapted and L^2 bounded. Suppose that $\hat{\phi}$ is F_t-measurable, valued on [0,1] and a consistent estimator for ϕ, then,*

(i) $\frac{1}{T}\displaystyle\int_0^{\hat{\phi}T} Y_t dt - \frac{1}{T}\int_0^{\phi T} Y_t dt \xrightarrow[T\to\infty]{L^1} 0,$

(ii) $\frac{1}{T}\displaystyle\int_{\hat{\phi}T}^{T} Y_t dt - \frac{1}{T}\int_{\phi T}^{T} Y_t dt \xrightarrow[T\to\infty]{L^1} 0.$

The proof follows from the similar derivation as used in Lemma 3.1 of Nkurunziza and Zhang (2018).

Lemma A.3. *Let $f(\theta, x)$ be a $\mathbb{R}^d$-valued function, and let $\{Y_t, t \geq 0\}$ be a d-dimensional stochastic process which is a solution of the SDE,*

$$dY_t = f(\mu_1, Y_t)\mathbb{I}_{\{t\leq\gamma\}}dt + f(\mu_2, Y_t)\mathbb{I}_{\{t>\gamma\}}dt + \sigma dW_t, \tag{A.2}$$

where $f(\theta, x)$ is such that the processes $\{Y_t, t \geq 0\}$ and $\{f(\theta, Y_t), t \geq 0\}$ are L^2 bounded. If Assumption 3 holds with $\delta_0 > \frac{1}{2}$, then,

(i) $\frac{1}{\sqrt{T}}\displaystyle\int_0^{\hat{\phi}T} Y_t dW_t - \frac{1}{\sqrt{T}}\int_0^{\phi T} Y_t dW_t \xrightarrow[T\to\infty]{\mathrm{P}} 0,$

(ii) $\frac{1}{\sqrt{T}}\displaystyle\int_{\hat{\phi}T}^{T} Y_t dW_t - \frac{1}{\sqrt{T}}\int_{\phi T}^{T} Y_t dW_t \xrightarrow[T\to\infty]{\mathrm{P}} 0.$

The proof follows from the similar techniques as used in Lemma 3.3 of Nkurunziza and Zhang (2018).

Corollary A.1. *Let* $W \sim W_n(p, kI, \Delta)$*, then* $\frac{1}{k}\operatorname{Tr}(W) \sim \chi^2_{pn}(\operatorname{Tr}(\Delta))$.

For the proof, we refer to Corollary 2.4.2.2. in Kollo and Rosen (2011, p.238).

B Proof of important results

Proof of Proposition 2.1. First, we verify space-variable lipshitz condition. By Triangle Inequality, we get:

$$\|S(t,x) - S(t,y)\|_2^2 + \|\Sigma(t,x)^{1/2} - \Sigma(t,y)^{1/2}\|_F^2 = \|S(t,x) - S(t,y)\|_2^2$$

$$= \|(\mu_1\varphi(t) - A_1x)\mathbb{I}_{\{t\leq\gamma\}} + (\mu_2\varphi(t) - A_2x)\mathbb{I}_{\{t>\gamma\}} -$$

$$[(\mu_1\varphi(t) - A_1y)\mathbb{I}_{\{t\leq\gamma\}} + (\mu_2\varphi(t) - A_2y)\mathbb{I}_{\{t>\gamma\}}]\|_2^2$$

$$= \|(A_1(y-x))\mathbb{I}_{\{t\leq\gamma\}} + (A_2(y-x))\mathbb{I}_{\{t>\gamma\}}\|_2^2.$$

Note that $\mathbb{I}_{\{t\leq\gamma\}}\mathbb{I}_{\{t>\gamma\}} = 0$ for all t. Also since $\|A_1(y-x)\|_2^2 \geq 0$ and $\|A_2(y-x)\|_2^2 \geq 0$, we have

$$\begin{aligned}\|(A_1(y-x))\mathbb{I}_{\{t\leq\gamma\}} + (A_2(y-x))\mathbb{I}_{\{t>\gamma\}}\|_2^2 &\leq \|(A_1(y-x))\mathbb{I}_{\{t\leq\gamma\}}\|_2^2 + \|(A_2(y-x))\mathbb{I}_{\{t>\gamma\}}\|_2^2 \\ &\leq \|A_1(y-x)\|_2^2\mathbb{I}_{\{t\leq\gamma\}} + \|A_2(y-x)\|_2^2\mathbb{I}_{\{t>\gamma\}} \\ &\leq \|A_1(y-x)\|_2^2 + \|A_2(y-x)\|_2^2 \\ &\leq \|A_1\|_F^2\|y-x\|_2^2 + \|A_2\|_F^2\|y-x\|_2^2.\end{aligned}$$

Let $\|A_1\|_F^2 + \|A_2\|_F^2 \leq K_A$, we have

$$\|S(t,x) - S(t,y)\|_2^2 + \|\Sigma(t,x)^{1/2} - \Sigma(t,y)^{1/2}\|_F^2 \leq K_A\|y-x\|_2^2.$$

Second, we verify spatial growth condition. Note that from Assumption 2, we have the boundedness of $\varphi(t)$. Therefore, by Triangle Inequality and $(a+b)^2 \le 2a^2 + 2b^2$, we have

$$
\begin{aligned}
&\|(\mu_1\varphi(t) - A_1x)\mathbb{I}_{\{t\le\gamma\}} + (\mu_2\varphi(t) - A_2x)\mathbb{I}_{\{t>\gamma\}}\|_2^2 + \|\Sigma^{1/2}\|_F^2 \\
&\le \|(\mu_1\varphi(t) - A_1x)\mathbb{I}_{\{t\le\gamma\}}\|_2^2 + \|(\mu_2\varphi(t) - A_2x)\mathbb{I}_{\{t>\gamma\}}\|_2^2 + \|\Sigma^{1/2}\|_F^2 \\
&\le \|\mu_1\varphi(t) - A_1x\|_2^2 + \|\mu_2\varphi(t) - A_2x\|_2^2 + \|\Sigma^{1/2}\|_F^2 \\
&\le (\|\mu_1\varphi(t)\|_2 + \|A_1x\|_2)^2 + (\|\mu_2\varphi(t)\|_2 + \|A_2x\|_2)^2 + \|\Sigma^{1/2}\|_F^2 \\
&\le 2\|\mu_1\varphi(t)\|_2^2 + 2\|A_1x\|_2^2 + 2\|\mu_2\varphi(t)\|_2^2 + 2\|A_2x\|_2^2 + \|\Sigma^{1/2}\|_F^2 \\
&\le 2\|\mu_1\varphi(t)\|_2^2 + 2\|A_1\|_F^2\|x\|_2^2 + 2\|\mu_2\varphi(t)\|_2^2 + 2\|A_2\|_F^2\|x\|_2^2 + \|\Sigma^{1/2}\|_F^2,
\end{aligned}
$$

then $\|S(t,x)\|_2^2 + \|\Sigma(t,x)^{1/2}\|_F^2 \le G(1 + \|x\|_2^2)$ for some constant G. Further, let $G' = \max(G, K_A)$, we have

$$
\begin{aligned}
&\|S(t,x) - S(t,y)\|_2^2 + \|\Sigma(t,x)^{1/2} - \Sigma(t,y)^{1/2}\|_F^2 \le G'\|y - x\|_2^2 \\
&\|S(t,x)\|_2^2 + \|\Sigma(t,x)^{1/2}\|_F^2 \le G'(1 + \|x\|_2^2),
\end{aligned}
$$

which completes the proof. $\square$

Proof of Proposition 2.3. By the independence of $W_s^{(1)}$ and $W_{-s}^{(2)}$, we get

$$
\begin{aligned}
\operatorname{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t}) &= \operatorname{Cov}\left(e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}, e^{-A(k+t)}\int_0^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right) \\
&+ \operatorname{Cov}\left(e^{-At}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}, e^{-A(k+t)}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right) \\
&= \operatorname{Cov}\left(e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}, e^{-A(k+t)}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}\right) \\
&+ \operatorname{Cov}\left(e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}, e^{-A(k+t)}\int_t^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right) \\
&+ \operatorname{Cov}\left(e^{-At}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}, e^{-A(k+t)}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right).
\end{aligned}
$$

By the independence of increments of wiener process, we have

$$\mathrm{Cov}\left(e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}, e^{-A(k+t)}\int_t^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right) = 0.$$

Then, we get

$$\begin{aligned}
\mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t}) &= \mathrm{Cov}\left(e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}, e^{-A(k+t)}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}\right) \\
&\quad + \mathrm{Cov}\left(e^{-At}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}, e^{-A(k+t)}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right) \\
&= \left[\mathrm{Var}\left(\int_0^t e^{-A(t-s)}\Sigma^{1/2}dW_s^{(1)}\right) + \mathrm{Var}\left(\int_{-\infty}^0 e^{-A(t-s)}\Sigma^{1/2}dW_{-s}^{(2)}\right)\right] e^{-A'k}.
\end{aligned} \tag{B.1}$$

Since the It$\hat{o}$'s integral $\int_0^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}$ is a martingale, we get $\mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t}) = \mathrm{E}(\widetilde{Z}_t\widetilde{Z}'_{k+t})$. Also, using It$\hat{o}$'s isometry, we get

$$\begin{aligned}
\mathrm{Var}\left(\int_0^t e^{-A(t-s)}\Sigma^{1/2}dW_s^{(1)}\right) &= \int_0^t e^{-A(t-s)}\Sigma^{1/2}\Sigma^{1/2\prime}e^{-A'(t-s)}ds \\
&= \int_0^t e^{-A(t-s)}\Sigma e^{-A'(t-s)}ds.
\end{aligned} \tag{B.2}$$

Furthermore, we have

$$\mathrm{Var}\left(\int_{-\infty}^0 e^{-A(t-s)}\Sigma^{1/2}dW_{-s}^{(2)}\right) = \mathrm{Var}\left(\int_0^\infty e^{-A(t+s)}\Sigma^{1/2}dW_s^{(2)}\right).$$

Let $I_L = \int_0^L e^{-A(t+s)}\Sigma^{1/2}dW_s^{(2)}$. As verified later in (B.15), we have $I_L \xrightarrow[L\to\infty]{L^2} I_\infty$, which implies that $\lim_{L\to\infty}\mathrm{Var}(I_L) = \mathrm{Var}(I_\infty)$, therefore

$$\mathrm{Var}\left(\int_0^\infty e^{-A(t+s)}\Sigma^{1/2}dW_s^{(2)}\right) = \lim_{L\to\infty}\mathrm{Var}(I_L) = \lim_{L\to\infty}\mathrm{Var}\left(\int_0^L e^{-A(t+s)}\Sigma^{1/2}dW_s^{(2)}\right). \tag{B.3}$$

Using It$\hat{o}$'s isometry, we get

$$\mathrm{Var}\left(\int_0^L e^{-A(t+s)}\Sigma^{1/2}dW_s^{(2)}\right) = \int_0^L e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds. \tag{B.4}$$

Combining (B.3) and (B.4), we get

$$\mathrm{Var}\left(\int_0^\infty e^{-A(t+s)}\Sigma^{1/2}dW_s^{(2)}\right) = \int_0^\infty e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds. \tag{B.5}$$

Combining (B.1), (B.2), and (B.5), we have

$$\mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t}) = \left(\int_0^t e^{-A(t-s)}\Sigma e^{-A'(t-s)}ds + \int_0^\infty e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds\right) e^{-A'k}.$$

In order to get the explicit form of $\mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t})$, let us consider the vectorization of $\mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t})$. Using $\mathrm{Vec}(ABC) = (C' \otimes A)\mathrm{Vec}(B)$ where "$\otimes$" denotes the Kronecker product, we get

$$\begin{aligned}
\mathrm{Vec}(\mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t})) &= \mathrm{Vec}\left(\left(\int_0^t e^{-A(t-s)}\Sigma e^{-A'(t-s)}ds + \int_0^\infty e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds\right) e^{-A'k}\right)\\
&= (e^{-Ak} \otimes I_d)\mathrm{Vec}\left(\int_0^t e^{-A(t-s)}\Sigma e^{-A'(t-s)}ds\right)\\
&+ (e^{-Ak} \otimes I_d)\mathrm{Vec}\left(\int_0^\infty e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds\right),
\end{aligned} \tag{B.6}$$

where I_d is a d-dimensional indentity matrix. Note that

$$\begin{aligned}
\mathrm{Vec}\left(\int_0^t e^{-A(t-s)}\Sigma e^{-A'(t-s)}ds\right) &= \int_0^t \mathrm{Vec}(e^{-A(t-s)}\Sigma e^{-A'(t-s)})ds\\
&= \int_0^t (e^{-A(t-s)} \otimes e^{-A(t-s)})\mathrm{Vec}(\Sigma)ds.
\end{aligned}$$

Using $e^A \otimes e^B = e^{A \oplus B}$ (Horn and Johnson, 1994), where "$\oplus$" denotes Kronecker sum (i.e. $A \oplus B = A \otimes I_m + I_n \otimes B$ for A, B square matrices of order n, m respectively), we get

$$\int_0^t e^{-(A\oplus A)(t-s)}\mathrm{Vec}(\Sigma)ds = \left[(A \oplus A)^{-1}e^{-(A\oplus A)(t-s)}\mathrm{Vec}(\Sigma)\right]_0^t.$$

Then, we get

$$\mathrm{Vec}\left(\int_0^t e^{-A(t-s)}\Sigma e^{-A'(t-s)}ds\right) = (A \oplus A)^{-1}\mathrm{Vec}(\Sigma) - (A \oplus A)^{-1}e^{-(A\oplus A)t}\mathrm{Vec}(\Sigma). \tag{B.7}$$

Similarly, we have

$$\begin{aligned}\operatorname{Vec}\left(\int_0^\infty e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds\right) &= \int_0^\infty \operatorname{Vec}(e^{-A(t+s)}\Sigma e^{-A'(t+s)})ds \\ &= \left[-(A\oplus A)^{-1}e^{-(A\oplus A)(t+s)}\operatorname{Vec}(\Sigma)\right]_0^\infty .\end{aligned}$$

Since A is positive definite, $A \oplus A$ is also positive definite, then by Proposition $A.1$, we get

$$\operatorname{Vec}\left(\int_0^\infty e^{-A(t+s)}\Sigma e^{-A'(t+s)}ds\right) = (A\oplus A)^{-1}e^{-(A\oplus A)t}\operatorname{Vec}(\Sigma). \tag{B.8}$$

Combining (B.6), (B.7), and (B.8), we have

$$\begin{aligned}\operatorname{Vec}(\operatorname{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t})) = (e^{-Ak}\otimes I_d)[&(A\oplus A)^{-1}\operatorname{Vec}(\Sigma) - (A\oplus A)^{-1}e^{-(A\oplus A)t}\operatorname{Vec}(\Sigma) \\ &+ (A\oplus A)^{-1}e^{-(A\oplus A)t}\operatorname{Vec}(\Sigma)],\end{aligned}$$

then

$$\operatorname{Vec}(\operatorname{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t})) = (e^{-Ak}\otimes I_d)(A\oplus A)^{-1}\operatorname{Vec}(\Sigma), \tag{B.9}$$

this completes the proof. $\square$

Proof of Proposition 2.4. Note that for every $t \in [0,1]$ and $k \in \mathbb{N}_0$, we have $\widetilde{X}_{k+t} = \widetilde{h}(t) + \widetilde{Z}_{k+t}$. Thus, it suffices to prove that $\{\widetilde{Z}_{k+t}\}_{k\in\mathbb{N}_0}$ is a Gaussian process. Further, we have

$$\widetilde{Z}_{k+t} = e^{-A(k+t)}\int_0^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)} + e^{-A(k+t)}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}.$$

let $Z_{k+t} = e^{-A(k+t)}\int_0^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}$, and $\bar{Z}_{k+t} = e^{-A(k+t)}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}$.

Taking any partition of k, i.e. k=1,2,...,n, we have

$$\begin{bmatrix} \widetilde{Z}_{1+t} \\ \widetilde{Z}_{2+t} - \widetilde{Z}_{1+t} \\ \vdots \\ \widetilde{Z}_{n+t} - \widetilde{Z}_{(n-1)+t} \end{bmatrix} = \begin{bmatrix} Z_{1+t} \\ Z_{2+t} - Z_{1+t} \\ \vdots \\ Z_{n+t} - Z_{(n-1)+t} \end{bmatrix} + \begin{bmatrix} \bar{Z}_{1+t} \\ \bar{Z}_{2+t} - \bar{Z}_{1+t} \\ \vdots \\ \bar{Z}_{n+t} - \bar{Z}_{(n-1)+t} \end{bmatrix}.$$

By the independence of increments of wiener process, we have

$$\begin{bmatrix} Z'_{1+t} & Z'_{2+t} - Z'_{1+t} & ... & Z'_{n+t} - Z'_{(n-1)+t} \end{bmatrix}'$$

follows multivariate normal distribution. Further, we have

$$\begin{bmatrix} \bar{Z}_{1+t} \\ \bar{Z}_{2+t} - \bar{Z}_{1+t} \\ \vdots \\ \bar{Z}_{n+t} - \bar{Z}_{(n-1)+t} \end{bmatrix} = \begin{bmatrix} e^{-A(1+t)} \\ e^{-A(2+t)} - e^{-A(1+t)} \\ \vdots \\ e^{-A(n+t)} - e^{-A(n-1+t)} \end{bmatrix} \int_{-\infty}^{0} e^{As} \Sigma^{1/2} dW^{(2)}_{-s},$$

which also follows multivariate normal distribution.

By the independence of $W^{(1)}_s$ and $W^{(2)}_{-s}$, we have

$$\begin{bmatrix} \widetilde{Z}'_{1+t} & \widetilde{Z}'_{2+t} - \widetilde{Z}'_{1+t} & ... & \widetilde{Z}'_{n+t} - \widetilde{Z}'_{(n-1)+t} \end{bmatrix}'$$

follows multivariate normal distribution. Therefore,

$$\begin{bmatrix} \widetilde{Z}_{1+t} \\ \widetilde{Z}_{2+t} \\ \vdots \\ \widetilde{Z}_{n+t} \end{bmatrix} = \begin{bmatrix} I_d & 0 & 0 & ... & 0 \\ I_d & I_d & 0 & ... & 0 \\ I_d & I_d & I_d & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_d & I_d & I_d & ... & I_d \end{bmatrix} \begin{bmatrix} \widetilde{Z}_{1+t} \\ \widetilde{Z}_{2+t} - \widetilde{Z}_{1+t} \\ \vdots \\ \widetilde{Z}_{n+t} - \widetilde{Z}_{(n-1)+t} \end{bmatrix}$$

follows multivariate Gaussian distribution and this proves that $\{\widetilde{X}_{k+t}\}_{k\in\mathbb{N}_0}$ is a Gaussian process. $\square$

Proof of Proposition 2.5. First, let us prove that for all $k \in \mathbb{N}_0$ and $t \in [0,1]$, $\mathrm{E}[\|\widetilde{X}_{k+t}\|_2^2] < \infty$. By Triangle Inequality and the fact $(a+b)^2 \le 2a^2 + 2b^2$, we have

$$\begin{aligned}\mathrm{E}[\|\widetilde{X}_{k+t}\|_2^2] = \mathrm{E}[\|\widetilde{h}(k+t) + \widetilde{Z}_{k+t}\|_2^2] &\le \mathrm{E}[(\|\widetilde{h}(k+t)\|_2 + \|\widetilde{Z}_{k+t}\|_2)^2] \\ &\le 2\mathrm{E}[\|\widetilde{h}(k+t)\|_2^2] + 2\mathrm{E}[\|\widetilde{Z}_{k+t}\|_2^2].\end{aligned}$$

Let $\|\mu\varphi(t)\|_2 \le K_{\mu,\varphi}$ for all t, we have

$$\begin{aligned}\mathrm{E}\left[\left\|\widetilde{h}(k+t)\right\|_2^2\right] &= \mathrm{E}\left[\left\|\int_{-\infty}^{k+t} e^{-A(k+t-s)}\mu\varphi(s)ds\right\|_2^2\right] \\ &\le K_{\mu,\varphi}^2 \int_{-\infty}^{k+t} \|e^{-A(k+t-s)}\|_F^2 ds.\end{aligned}$$

From Proposition $A.3$, and let λ_1 be the smallest eigenvalue of $A' + A$, we get

$$\mathrm{E}[\|\widetilde{h}(k+t)\|_2^2] \le K_{\mu,\varphi}^2 d \int_{-\infty}^{k+t} e^{-\lambda_1(k+t-s)}ds \le K_{\mu,\varphi}^2 \frac{d}{\lambda_1} < \infty. \tag{B.10}$$

Further, by the independence of $W_s^{(1)}$ and $W_{-s}^{(2)}$, we have

$$\begin{aligned}\mathrm{E}[\|\widetilde{Z}_{k+t}\|_2^2] &= \mathrm{E}\left[\left\|e^{-A(k+t)}\int_{-\infty}^{k+t} e^{As}\Sigma^{1/2}d\widetilde{W}_s\right\|_2^2\right] \\ &= \mathrm{E}\left[\left\|e^{-A(k+t)}\int_{0}^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)} + e^{-A(k+t)}\int_{-\infty}^{0} e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right\|_2^2\right] \\ &= \mathrm{E}\left[\left\|e^{-A(k+t)}\int_{0}^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right\|_2^2\right] + \mathrm{E}\left[\left\|e^{-A(k+t)}\int_{-\infty}^{0} e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right\|_2^2\right] \\ &+ 2\mathrm{E}\left(e^{-A(k+t)}\int_{0}^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right)' \mathrm{E}\left(e^{-A(k+t)}\int_{-\infty}^{0} e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right).\end{aligned}$$

Since the Itô's integral $\int_0^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}$ is a martingale, therefore

$$\mathrm{E}\left(\int_{0}^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right) = 0. \tag{B.11}$$

Then

$$\mathrm{E}[\|\widetilde{Z}_{k+t}\|_2^2] = \mathrm{E}\left[\left\|e^{-A(k+t)}\int_{0}^{k+t} e^{As}\Sigma^{1/2}dW_s^{(1)}\right\|_2^2\right] + \mathrm{E}\left[\left\|e^{-A(k+t)}\int_{-\infty}^{0} e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right\|_2^2\right]. \tag{B.12}$$

Moreover, let

$$e^{-A(k+t-s)}\Sigma^{1/2} = \begin{bmatrix} a_{11}(s) & a_{12}(s) & a_{13}(s) & ... & a_{1d}(s) \\ a_{21}(s) & a_{22}(s) & a_{23}(s) & ... & a_{2d}(s) \\ a_{31}(s) & a_{32}(s) & a_{33}(s) & ... & a_{3d}(s) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{d1}(s) & a_{d2}(s) & a_{d3}(s) & ... & a_{dd}(s) \end{bmatrix},$$

and $W_s^{(1)} = \begin{bmatrix} W_s^1 & W_s^2 & W_s^3 & ... & W_s^d \end{bmatrix}'$, we have

$$\begin{aligned}
&\mathrm{E}\left(\left\|\int_0^{k+t} e^{-A(k+t-s)}\Sigma^{1/2} dW_s^{(1)}\right\|_2^2\right) = \mathrm{E}\left(\left\|\int_0^{k+t} e^{-A(k+t-s)}\Sigma^{1/2} dW_s^{(1)}\right\|_2^2\right) \\
&= \mathrm{E}\left(\sum_{i=1}^d \left(\sum_{j=1}^d \int_0^{k+t} a_{ij}(s) dW_s^j\right)^2\right) = \sum_{i=1}^d \mathrm{E}\left(\sum_{j=1}^d \int_0^{k+t} a_{ij}(s) dW_s^j\right)^2 \\
&= \sum_{i=1}^d \mathrm{E}\left(\sum_{j=1}^d \left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)^2 + \sum_{j\neq k}\left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)\left(\int_0^{k+t} a_{ik}(s) dW_s^k\right)\right) \\
&= \sum_{i=1}^d \mathrm{E}\left(\sum_{j=1}^d \left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)^2\right) + \mathrm{E}\left(\sum_{j\neq k}\left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)\left(\int_0^{k+t} a_{ik}(s) dW_s^k\right)\right).
\end{aligned}$$

By the independence of components of the standard Brownian motion, we have

$$\begin{aligned}
&\mathrm{E}\left(\left\|\int_0^{k+t} e^{-A(k+t-s)}\Sigma^{1/2} dW_s^{(1)}\right\|_2^2\right) \\
&= \sum_{i=1}^d \mathrm{E}\left(\sum_{j=1}^d \left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)^2\right) + \sum_{j\neq k} \mathrm{E}\left(\int_0^{k+t} a_{ij}(s) dW_s^j\right) \mathrm{E}\left(\int_0^{k+t} a_{ik}(s) dW_s^k\right).
\end{aligned}$$

Since $\mathrm{E}\left(\int_0^{k+t} a_{ij}(s) dW_s^j\right) = 0$ for all i, j, we have

$$\mathrm{E}\left(\left\|\int_0^{k+t} e^{-A(k+t-s)}\Sigma^{1/2} dW_s^{(1)}\right\|_2^2\right) = \sum_{i=1}^d \sum_{j=1}^d \mathrm{E}\left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)^2.$$

By It$\hat{o}$'s isometry, this gives

$$\mathrm{E}\left(\int_0^{k+t} a_{ij}(s) dW_s^j\right)^2 = \int_0^{k+t} a_{ij}^2(s) ds.$$

Therefore, we get

$$\mathrm{E}\left(\left\|\int_0^{k+t} e^{-A(k+t-s)}\Sigma^{1/2}dW_s^{(1)}\right\|_2^2\right) = \sum_{i=1}^{d}\sum_{j=1}^{d}\left(\int_0^{k+t} a_{ij}^2(s)ds\right)$$
$$= \int_0^{k+t} \|e^{-A(k+t-s)}\Sigma^{1/2}\|_F^2 ds \leq \|\Sigma^{1/2}\|_F^2 \int_0^{k+t} \|e^{-A(k+t-s)}\|_F^2 ds.$$

From Proposition $A.3$, and let λ_1 be the smallest eigenvalue of $A' + A$, we get

$$\mathrm{E}\left(\left\|\int_0^{k+t} e^{-A(k+t-s)}\Sigma^{1/2}dW_s^{(1)}\right\|_2^2\right) \leq \frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1}\left(1 - e^{-(k+t)\lambda_1}\right). \tag{B.13}$$

Meanwhile, let $l = -s$. This gives

$$\mathrm{E}\left(\left\|\int_{-\infty}^{0} e^{-A(t+k-s)}\Sigma^{1/2}dW_{-s}^{(2)}\right\|_2^2\right) = \mathrm{E}\left(\left\|\int_0^{\infty} e^{-A(t+k+l)}\Sigma^{1/2}dW_l^{(2)}\right\|_2^2\right).$$

Also, one can verify that for all $L_1 \geq 0$, we have

$$\mathrm{E}\left(\left\|\int_0^{L_1} e^{-A(t+k+l)}\Sigma^{1/2}dW_l^{(2)}\right\|_2^2\right) \leq d\|\Sigma^{1/2}\|_F^2 \int_0^{L_1} e^{-(k+t+l)\lambda_1} dl$$
$$\leq e^{-A(t+k)\lambda_1}\frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1} \leq \frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1} < \infty. \tag{B.14}$$

Now, by L^2-Bounded Martingale Convergence Theorem, we have

$$I_{L_1} \xrightarrow[L_1\to\infty]{L^2} I_\infty = \int_0^{\infty} e^{-A(t+k+l)}\Sigma^{1/2}dW_l^{(2)}. \tag{B.15}$$

Therefore, we have

$$\mathrm{E}\left[\left\|e^{-A(k+t)}\int_{-\infty}^{0} e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right\|_2^2\right] \leq e^{-(k+t)\lambda_1}\frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1}. \tag{B.16}$$

Combining (B.12), (B.13), and (B.16), we have

$$\mathrm{E}[\|\widetilde{Z}_{k+t}\|_2^2] \leq \frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1}\left(1 - e^{-(k+t)\lambda_1}\right) + e^{-(k+t)\lambda_1}\frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1} = \frac{d\|\Sigma^{1/2}\|_F^2}{\lambda_1} < \infty. \tag{B.17}$$

Combining (B.10) and (B.17), one concludes that $\mathrm{E}[\|\widetilde{X}_{k+t}\|_2^2] < \infty$. Second, let us prove that $\mathrm{E}[\widetilde{X}_{k+t}]$ is a constant vector. We have

$$\begin{aligned}\mathrm{E}[\widetilde{X}_{k+t}] &= \mathrm{E}[\widetilde{h}(k+t)] + \mathrm{E}[\widetilde{Z}_{k+t}] \\ &= e^{-A(k+t)} \int_{-\infty}^{k+t} e^{As}\mu\varphi(s)ds + \mathrm{E}\left[e^{-A(k+t)} \int_{-\infty}^{k+t} e^{As}\Sigma^{1/2} d\widetilde{W}_s\right]. \end{aligned} \tag{B.18}$$

For $k \in \mathbb{N}_0$, let $r = s-k \in (-\infty, t)$, and by the periodicity of $\varphi(t)$, i.e. $\varphi(r+k) = \varphi(r)$, we have

$$\begin{aligned} e^{-A(k+t)} \int_{-\infty}^{k+t} e^{As}\mu\varphi(s)ds &= e^{-At} \int_{-\infty}^{k+t} e^{-A(k-s)}\mu\varphi(s)ds \\ &= e^{-At} \int_{-\infty}^{t} e^{Ar}\mu\varphi(r)dr = \widetilde{h}(t), \end{aligned} \tag{B.19}$$

which does not depend on k and is a constant for every $t \in [0,1]$. Furthermore, we have

$$\begin{aligned} &\mathrm{E}\left[e^{-A(k+t)} \int_{-\infty}^{k+t} e^{As}\Sigma^{1/2} d\widetilde{W}_s\right] \\ &= \mathrm{E}\left[e^{-A(k+t)} \int_{0}^{k+t} e^{As}\Sigma^{1/2} dW_s^{(1)}\right] + \mathrm{E}\left[e^{-A(k+t)} \int_{-\infty}^{0} e^{As}\Sigma^{1/2} dW_{-s}^{(2)}\right] \\ &= e^{-A(k+t)}\left[\mathrm{E}\left(\int_{0}^{k+t} e^{As}\Sigma^{1/2} dW_s^{(1)}\right) + \mathrm{E}\left(\int_{-\infty}^{0} e^{As}\Sigma^{1/2} dW_{-s}^{(2)}\right)\right]. \end{aligned} \tag{B.20}$$

From (B.15), we have $I_{k+t} \xrightarrow[k\to\infty]{L^2} I_\infty = \int_0^\infty e^{-Al}\Sigma^{1/2} dW_l^{(2)}$. This implies that

$$\mathrm{E}\left[\int_{-\infty}^{0} e^{As}\Sigma^{1/2} dW_{-s}^{(2)}\right] = \mathrm{E}\left[\int_{0}^{\infty} e^{-As}\Sigma^{1/2} dW_{s}^{(2)}\right] = \lim_{k\to\infty} \mathrm{E}[I_{k+t}].$$

Since $\mathrm{E}[I_{k+t}] = 0$ for all $k+t \geq 0$, we have

$$\mathrm{E}\left[\int_{-\infty}^{0} e^{As}\Sigma^{1/2} dW_{-s}^{(2)}\right] = 0. \tag{B.21}$$

Combining (B.11), (B.19), (B.20), and (B.21), one concludes that $\mathrm{E}[\widetilde{X}_{k+t}] = \widetilde{h}(t)$ for $k \in \mathbb{N}_0$, for all $t \in [0,1]$. Further, since $\widetilde{h}(t)$ is non-random, we have

$$\mathrm{Cov}(\widetilde{X}_t, \widetilde{X}_{k+t}) = \mathrm{Cov}(\widetilde{h}(t) + \widetilde{Z}_t, \widetilde{h}(k+t) + \widetilde{Z}_{k+t}) = \mathrm{Cov}(\widetilde{Z}_t, \widetilde{Z}_{k+t}).$$

Therefore, from Proposition 2.3, one concludes that $\mathrm{Cov}(\widetilde{X}_t, \widetilde{X}_{k+t})$ is a function of k only. Further, by Proposition 2.4, the stochastic process $\{\widetilde{X}_{k+t}\}_{k\in\mathbb{N}_0}$ is also Gaussian. Then, for any $t \in [0,1]$, $\{\widetilde{X}_{k+t}\}_{k\in\mathbb{N}_0}$ is a weakly stationary process. This implies that the process $\{\widetilde{X}_{k+t}\}_{k\in\mathbb{N}_0}$ is also strictly stationary. Further, for $t \in [0,1]$ and $k \in \mathbb{N}_0$, the correlation coefficient function is defined as:

$$R_k = \mathrm{Var}(\widetilde{X}_t)^{-1/2}\mathrm{Cov}(\widetilde{X}_t, \widetilde{X}_{k+t})\mathrm{Var}(\widetilde{X}_{k+t})^{-1/2}.$$

Taking vectorization, we get

$$\mathrm{Vec}(R_k) = [(\mathrm{Var}(\widetilde{X}_{k+t})^{-1/2})' \otimes \mathrm{Var}(\widetilde{X}_t)^{-1/2}]\mathrm{Vec}(\mathrm{Cov}(\widetilde{X}_t, \widetilde{X}_{k+t})).$$

Note that $\mathrm{Var}(\widetilde{X}_{k+t})^{-1/2}$ is symmetric, and from (B.9), we have

$$(\mathrm{Var}(\widetilde{X}_{k+t})^{-1/2})' \otimes \mathrm{Var}(\widetilde{X}_t)^{-1/2} = \mathrm{Var}(\widetilde{X}_t)^{-1/2} \otimes \mathrm{Var}(\widetilde{X}_t)^{-1/2},$$

which does not depends on k. Also

$$\mathrm{Vec}(\mathrm{Cov}(\widetilde{X}_t, \widetilde{X}_{k+t})) = (e^{-Ak} \otimes I_d)(A \oplus A)^{-1}\mathrm{Vec}(\Sigma).$$

By $A.1$, we get $\lim_{k\to\infty} \mathrm{Vec}(\mathrm{Cov}(\widetilde{X}_t, \widetilde{X}_{k+t})) = 0$. Therefore

$$\lim_{k\to\infty} \mathrm{Vec}(R_k) = 0.$$

Hence, $\{\widetilde{X}_{k+t}\}_{k\in\mathbb{N}_0}$ is ergodic for any $t \in [0,1]$, which completes the proof. $\square$

Proof of Proposition 2.6. By Lemma 2.1, it suffices to prove

$$\frac{1}{T}\int_0^{\phi T} \varphi(t)\widetilde{X}_t' dt \xrightarrow[T\to\infty]{P} \phi \int_0^1 \varphi(t)\widetilde{h}'(t)dt.$$

We have

$$\frac{1}{T}\int_0^{\phi T} \varphi(t)\widetilde{X}_t' dt = \phi\frac{1}{\phi T}\int_0^{\phi T} \varphi(t)\widetilde{X}_t' dt = \phi\frac{1}{\phi T}\sum_{k=1}^{\lfloor \phi T\rfloor}\int_{k-1}^{k} \varphi(t)\widetilde{X}_t' dt + \phi\frac{1}{\phi T}\int_{\lfloor \phi T\rfloor}^{\phi T} \varphi(t)\widetilde{X}_t' dt.$$

Let $Y_k = \int_{k-1}^{k} \varphi(t)\widetilde{X}_t' dt$, and $r = t-k+1 \in [0,1]$. By the periodicity of $\varphi(t)$, we have

$$Y_k = \int_0^1 \varphi(r+k-1)\widetilde{X}_{r+k-1}' dr = \int_0^1 \varphi(r)\widetilde{X}_{r+k-1}' dr.$$

According to Proposition 2.5, for $r \in [0,1]$, $\{\widetilde{X}_{r+k-1}\}_{k\in N}$ is a stationary and ergodic process with $r+k-1 \in [0, \phi T]$. Thus, Y_k is a measurable function of the stationary and ergodic process $\{\widetilde{X}_{r+k-1}\}_{k\in N}$. Thus, $\{Y_k\}_{k\in N}$ is stationary and ergodic, and then by Birkhoff Ergodic Theorem, we get

$$\frac{\lfloor \phi T \rfloor}{\phi T} \phi \frac{1}{\lfloor \phi T \rfloor} \sum_{k=1}^{\lfloor \phi T \rfloor} Y_k \xrightarrow[T\to\infty]{a.s.} \phi \mathrm{E}\left(\int_0^1 \varphi(t)\widetilde{X}_t' dt \right).$$

Moreover, $\|\varphi(t)\|_2^2 \le K_\varphi$. Then, by Triangle Inequality, Jensen's Inequality, and Cauchy Schwarz Inequality, we have

$$\mathrm{E}\left(\left\| \phi \frac{1}{\phi T} \int_{\lfloor \phi T \rfloor}^{\phi T} \varphi(t)\widetilde{X}_t' dt \right\|_F \right) \le \phi \frac{1}{\phi T} \int_{\lfloor \phi T \rfloor}^{\phi T} \mathrm{E}(\|\varphi(t)\widetilde{X}_t'\|_F) dt \le \phi \frac{1}{\phi T} K_\varphi \int_{\lfloor \phi T \rfloor}^{\phi T} \mathrm{E}(\|\widetilde{X}_t\|_2^2)^{1/2} dt.$$

From (B.10) and (B.17), we have $\widetilde{X}_t$ is uniformly bounded in L^2. Let
$\mathrm{E}(\|\widetilde{X}_t\|_2^2) \le K' < \infty$, this implies

$$\phi \frac{1}{\phi T} \int_{\lfloor \phi T \rfloor}^{\phi T} \varphi(t)\widetilde{X}_t' dt \xrightarrow[T\to\infty]{L_1} 0.$$

Therefore, since $\mathrm{E}(\widetilde{X}_t) = \widetilde{h}(t)$, from (B.18)-(B.21), we have $\mathrm{E}(\varphi(t)\widetilde{X}_t') = \varphi(t)\widetilde{h}'(t)$, which completes the proof. $\square$

Proof of Proposition 2.8. By Lemma 2.2, it suffices to prove that

$$\frac{1}{T}\int_0^{\phi T} \widetilde{X}_t \widetilde{X}_t' dt \xrightarrow[T\to\infty]{P} \phi \left\{ \int_0^1 \widetilde{h}(t)\widetilde{h}'(t) dt + V(0) \right\}.$$

We have

$$\begin{aligned} \frac{1}{T}\int_0^{\phi T} \widetilde{X}_t \widetilde{X}_t' dt &= \phi \frac{1}{\phi T} \int_0^{\phi T} \widetilde{X}_t \widetilde{X}_t' dt \\ &= \phi \frac{1}{\phi T} \sum_{k=1}^{\lfloor \phi T \rfloor} \int_{k-1}^{k} \widetilde{X}_t \widetilde{X}_t' dt + \phi \frac{1}{\phi T} \int_{\lfloor \phi T \rfloor}^{\phi T} \widetilde{X}_t \widetilde{X}_t' dt. \end{aligned}$$

Since $\{\widetilde{X}_t\}$ is stationary and ergodic, we have $\{\widetilde{X}_t\widetilde{X}_t'\}$ is also stationary and ergodic. Let $Y_k = \int_{k-1}^k \widetilde{X}_t\widetilde{X}_t' dt$, and $r = t - k + 1 \in [0, 1]$, we have

$$Y_k = \int_0^1 \widetilde{X}_{r+k-1}\widetilde{X}_{r+k-1}' dr.$$

According to Proposition 2.5, for $r \in [0, 1]$, $\{\widetilde{X}_{r+k-1}\widetilde{X}_{r+k-1}'\}_{k\in N}$ is a stationary and ergodic process with $r + k - 1 \in [0, \phi T]$. Thus, Y_k is a measurable function of the stationary and ergodic process $\{\widetilde{X}_{r+k-1}\widetilde{X}_{r+k-1}'\}_{k\in N}$. Then, $\{Y_k\}_{k\in N}$ is stationary and ergodic, and then, by Birkhoff Ergodic Theorem, we get

$$\frac{\lfloor\phi T\rfloor}{\phi T}\phi\frac{1}{\lfloor\phi T\rfloor}\sum_{k=1}^{\lfloor\phi T\rfloor}\int_{k-1}^k \widetilde{X}_t\widetilde{X}_t' dt \xrightarrow[T\to\infty]{a.s.} \phi\mathrm{E}\left(\int_0^1 \widetilde{X}_t\widetilde{X}_t' dt\right).$$

Further, by Jensen's Inequality, we get

$$\mathrm{E}\left(\left\|\phi\frac{1}{\phi T}\int_{\lfloor\phi T\rfloor}^{\phi T}\widetilde{X}_t\widetilde{X}_t' dt\right\|_F\right) \le \phi\frac{1}{\phi T}\int_{\lfloor\phi T\rfloor}^{\phi T}\mathrm{E}(\|\widetilde{X}_t\widetilde{X}_t'\|_F)dt \le \phi\frac{1}{\phi T}\int_{\lfloor\phi T\rfloor}^{\phi T}\mathrm{E}(\|\widetilde{X}_t\|_2^2)dt.$$

From (B.10) and (B.17), we have $\widetilde{X}_t$ is uniformly bounded in L^2. Let $\mathrm{E}(\|\widetilde{X}_t\|_2^2) \le K' < \infty$, this implies

$$\phi\frac{1}{\phi T}\int_{\lfloor\phi T\rfloor}^{\phi T}\widetilde{X}_t\widetilde{X}_t' dt \xrightarrow[T\to\infty]{L_1} 0.$$

Further, we have

$$\phi\mathrm{E}\left(\int_0^1 \widetilde{X}_t\widetilde{X}_t' dt\right) = \phi\int_0^1 \mathrm{E}(\widetilde{X}_t\widetilde{X}_t')dt = \phi\int_0^1 \mathrm{E}[(\widetilde{h}(t) + \widetilde{Z}_t)(\widetilde{h}'(t) + \widetilde{Z}_t')]dt.$$

Note that for all $t > 0$, we have

$$\begin{aligned}
\mathrm{E}(\widetilde{Z}_t) &= \mathrm{E}\left[e^{-At}\int_{-\infty}^t e^{As}\Sigma^{1/2}d\widetilde{W}_s\right] \\
&= \mathrm{E}\left[e^{-At}\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}\right] + \mathrm{E}\left[e^{-At}\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right] \\
&= e^{-A(t)}\left[\mathrm{E}\left(\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}\right) + \mathrm{E}\left(\int_{-\infty}^0 e^{As}\Sigma^{1/2}dW_{-s}^{(2)}\right)\right].
\end{aligned} \tag{B.22}$$

Obviously, $\mathrm{E}\left(\int_0^t e^{As}\Sigma^{1/2}dW_s^{(1)}\right) = 0$ as this is Itô's integral which is a zero mean martingale. Further, by (B.21), we get $\mathrm{E}(\widetilde{Z}_t) = 0$. Therefore

$$\phi\mathrm{E}\left(\int_0^1 \widetilde{X}_t\widetilde{X}_t'dt\right) = \phi\int_0^1 [\widetilde{h}(t)\widetilde{h}'(t) + \mathrm{E}(\widetilde{Z}_t\widetilde{Z}_t')]dt.$$

From (B.9), $\mathrm{E}(\widetilde{Z}_t\widetilde{Z}_t')$ does not depend on t. Thus, letting $V(0) = \mathrm{E}(\widetilde{Z}_t\widetilde{Z}_t')$, we complete the proof. □

Proof of Proposition 3.2. For any $T > 0$

$$\frac{1}{T}Q_\gamma = \begin{bmatrix} \frac{1}{T}\int_0^{\phi T}\varphi(t)\varphi'(t)dt & -\frac{1}{T}\int_0^{\phi T}\varphi(t)X_t'dt \\ -\frac{1}{T}\int_0^{\phi T}X_t\varphi'(t)dt & \frac{1}{T}\int_0^{\phi T}X_tX_t'dt \end{bmatrix}.$$

Let $a = \begin{bmatrix} a_{(1)}' & a_{(2)}' \end{bmatrix}$ with $a_{(1)}$ a p-column vector, and $a_{(2)}$ a d-column vector. Then $aQ_\gamma a' = \int_0^{\phi T}\left\|\begin{bmatrix} a_{(1)}' & a_{(2)}' \end{bmatrix}\begin{bmatrix} \varphi'(t) & -X_t' \end{bmatrix}'\right\|_2^2 dt \geq 0$, and the equality hold if and only if $\left\|\begin{bmatrix} a_{(1)}' & a_{(2)}' \end{bmatrix}\begin{bmatrix} \varphi'(t) & -X_t' \end{bmatrix}'\right\|_2^2 = 0$ almost everywhere on $[0, \phi T]$, which is the same as $\begin{bmatrix} a_{(1)}' & a_{(2)}' \end{bmatrix}\begin{bmatrix} \varphi'(t) & -X_t' \end{bmatrix}' = 0$ almost everywhere on $[0, \phi T]$. Then, we have $a_{(1)}'\varphi(t) - a_{(2)}'\mathrm{E}(X_t) = 0$ and $\mathrm{Var}(a_{(2)}'X_t) = 0$ $\forall t \in [0, \phi T]$. Since $\exists t_0 \in [0, \phi T]$, such that $\mathrm{Var}(X_{t_0})$ is a positive definite matrix, then $a_{(2)}' = 0$. Then $a_{(1)}'\varphi(t) = 0$ $\forall t \in [0, \phi T]$. Since $\{\varphi_1(t), \varphi_2(t), ..., \varphi_p(t)\}$ is linearly independent on $[0, 1]$. Suppose now that $T \geq \frac{1}{\phi}$, we have $[0, 1] \subset [0, \phi T]$, then this implies $a_{(1)}' = 0$. Thus, Q_γ is a positive definite matrix. Similarly, one can verify that if $T \geq \frac{2}{1-\phi}$, then $Q_{\gamma,T}$ is a positive definite matrix. Therefore, if $T \geq \max(\frac{1}{\phi}, \frac{2}{1-\phi})$, we have $Q(\phi)$ is a positive definite matrix, this completes the proof. □

Proof of Lemma 3.1. Taking derivative of the log-likelihood function $l(\theta; X_{[0,T]})$ in (3.9) with respect to θ, since Σ and $Q(\phi)$ are symmetric matrices, we have

$$\frac{\partial l(\theta; X_{[0,T]})}{\partial\theta} = \Sigma^{-1}P(\phi) - \Sigma^{-1}\theta Q(\phi), \tag{B.23}$$

and setting this last term to be equal to 0, we get

$$\hat{\theta} = P(\phi)Q^{-1}(\phi). \tag{B.24}$$

Now, taking the second derivative of the log-likelihood function $l(\theta; X_{[0,T]})$ with respect to θ', we get

$$\frac{\partial(\Sigma^{-1}P(\phi) - \Sigma^{-1}\theta Q(\phi))}{\partial\theta'} = -\frac{\partial(\Sigma^{-1}\theta Q(\phi))}{\partial\theta'} = -(Q(\phi) \otimes \Sigma^{-1}).$$

From Proposition 3.2, we know that $Q(\phi)$ is a positive definite matrix, and since Σ is a positive definite matrix, we have Σ^{-1} is also a positive definite matrix, hence $Q(\phi) \otimes \Sigma^{-1}$ is a positive definite matrix, which complete the proof of the first statement. Moreover, from (3.9), we have

$$l(\theta; X_{[0,T]}) = \mathrm{Tr}(P(\phi)\theta'\Sigma^{-1}) - \frac{1}{2}\mathrm{Tr}(\theta'\Sigma^{-1}\theta Q(\phi)),$$

applying Lagrangian method with $\lambda_1 \in \mathbb{R}^{2(p+d)\times q}$, $\lambda_2 \in \mathbb{R}^{n\times d}$, let the lagrangian

$$l_\lambda(\theta, \lambda_1, \lambda_2; X_{[0,T]}) = l(\theta; X_t^T) + \mathrm{Tr}[\lambda_1(L_1\theta - d_1)] + \mathrm{Tr}[\lambda_2(\theta L_2 - d_2)].$$

Taking derivatives with respect to λ_1 and λ_2 and set to 0, we get

$$\frac{dl_\lambda(\theta, \lambda_1, \lambda_2; X_{[0,T]})}{d\lambda_1} = L_1\widetilde{\theta} - d_1 = 0, \tag{B.25}$$

$$\frac{dl_\lambda(\theta, \lambda_1, \lambda_2; X_{[0,T]})}{d\lambda_2} = \widetilde{\theta}L_2 - d_2 = 0, \tag{B.26}$$

and taking derivative with respect to θ and set to 0, we get

$$\frac{dl_{new}(\theta, \lambda_1, \lambda_2; X_{[0,T]})}{d\theta} = \Sigma^{-1}P(\phi) - \Sigma^{-1}\widetilde{\theta}Q(\phi) + L_1'\lambda_1' + \lambda_2'L_2' = 0_{d\times 2(p+d)},$$

$$P(\phi)Q^{-1}(\phi) - \widetilde{\theta} + \Sigma L_1'\lambda_1'Q^{-1}(\phi) + \Sigma\lambda_2'L_2'Q^{-1}(\phi) = 0_{d\times 2(p+d)},$$

since $\hat{\theta} = P(\phi)Q^{-1}(\phi)$, we have

$$\hat{\theta} - \widetilde{\theta} + \Sigma L_1' \lambda_1' Q^{-1}(\phi) + \Sigma \lambda_2' L_2' Q^{-1}(\phi) = 0_{d\times 2(p+d)}. \tag{B.27}$$

Then, L_1 times equation (B.27) from the left side gives

$$L_1\hat{\theta} - L_1\widetilde{\theta} + L_1\Sigma L_1' \lambda_1' Q^{-1}(\phi) + L_1\Sigma \lambda_2' L_2' Q^{-1}(\phi) = 0_{q\times 2(p+d)}.$$

By (B.25), we get

$$L_1\hat{\theta} - d_1 + L_1\Sigma L_1' \lambda_1' Q^{-1}(\phi) + L_1\Sigma \lambda_2' L_2' Q^{-1}(\phi) = 0_{q\times 2(p+d)}. \tag{B.28}$$

From equation (B.27), by multiplying each term by L_2, we get

$$\hat{\theta}L_2 - \widetilde{\theta}L_2 + \Sigma L_1' \lambda_1' Q^{-1}(\phi)L_2 + \Sigma \lambda_2' L_2' Q^{-1}(\phi)L_2 = 0_{d\times n}.$$

By (B.26), we get

$$\hat{\theta}L_2 - d_2 + \Sigma L_1' \lambda_1' Q^{-1}(\phi)L_2 + \Sigma \lambda_2' L_2' Q^{-1}(\phi)L_2 = 0_{d\times n}. \tag{B.29}$$

From equation (B.28) and (B.29), we notice that

$$(L_1\hat{\theta} - d_1)L_2 = L_1(\hat{\theta}L_2 - d_2).$$

Further, we have $L_1\Sigma L_1'$ and $L_2'Q^{-1}(\phi)L_2$ are positive definite matrices, and therefore, the inverses exist. Moreover, $(L_1\Sigma L_1')^{-1}$ times equation (B.28) from left side and equation (B.28) times $Q(\phi)$ from right side, we get

$$(L_1\Sigma L_1')^{-1}(L_1\hat{\theta} - d_1)Q(\phi) + \lambda_1' + (L_1\Sigma L_1')^{-1}(L_1\Sigma\lambda_2')L_2' = 0,$$

therefore

$$\lambda_1' = -(L_1\Sigma L_1')^{-1}(L_1\Sigma\lambda_2')L_2' - (L_1\Sigma L_1')^{-1}(L_1\hat{\theta} - d_1)Q(\phi). \tag{B.30}$$

Substituting (B.30) back into equation (B.27), we get

$$\hat{\theta} - \widetilde{\theta} + \Sigma L_1'[-(L_1\Sigma L_1')^{-1}(L_1\Sigma\lambda_2')L_2'$$

$$- (L_1\Sigma L_1')^{-1}(L_1\hat{\theta} - d_1)Q(\phi)]Q^{-1}(\phi) + \Sigma\lambda_2'L_2'Q^{-1}(\phi) = 0,$$

$$\hat{\theta} - \widetilde{\theta} - \Sigma L_1'(L_1\Sigma L_1')^{-1}(L_1\Sigma\lambda_2')L_2'Q^{-1}(\phi)$$

$$- \Sigma L_1'(L_1\Sigma L_1')^{-1}(L_1\hat{\theta} - d_1) + \Sigma\lambda_2'L_2'Q^{-1}(\phi) = 0,$$

$$\hat{\theta} - \widetilde{\theta} - \Sigma L_1'(L_1\Sigma L_1')^{-1}(L_1\hat{\theta} - d_1)$$

$$+ [\Sigma - \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma]\lambda_2'L_2'Q^{-1}(\phi) = 0. \tag{B.31}$$

In order to find the expression for $[\Sigma - \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma]\lambda_2'$, we subsititute equation (B.30) back into equation (B.29), then

$$\hat{\theta}L_2 - d_2 + \Sigma L_1'[-(L_1\Sigma L_1')^{-1}(L_1\Sigma\lambda_2')L_2' - (L_1\Sigma L_1')^{-1}(L_1\hat{\theta} - d_1)Q(\phi)]Q^{-1}(\phi)L_2$$

$$+ \Sigma\lambda_2'L_2'Q^{-1}(\phi)L_2 = 0.$$

Note that $d_1L_2 = L_1d_1$. This gives

$$[\Sigma - \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1\Sigma]\lambda_2'$$

$$= \Sigma L_1'(L_1\Sigma L_1')^{-1}L_1(\hat{\theta}L_2 - d_2)(L_2'Q^{-1}(\phi)L_2)^{-1} - (\hat{\theta}L_2 - d_2)(L_2'Q^{-1}(\phi)L_2)^{-1}. \tag{B.32}$$

Let $J_1 = \Sigma L_1'(L_1\Sigma L_1')^{-1} \in \mathbb{R}^{d\times q}$ and $J_2 = (L_2'Q^{-1}(\phi)L_2)^{-1}L_2'Q^{-1}(\phi) \in \mathbb{R}^{n\times 2(p+d)}$, and we subsititute equation (B.32) back into equation (B.31), then

$$\hat{\theta} - \widetilde{\theta} - J_1(L_1\hat{\theta} - d_1) + J_1L_1(\hat{\theta}L_2 - d_2)J_2 - (\hat{\theta}L_2 - d_2)J_2 = 0,$$

$$\widetilde{\theta} = \hat{\theta} - J_1(L_1\hat{\theta} - d_1) + J_1L_1(\hat{\theta}L_2 - d_2)J_2 - (\hat{\theta}L_2 - d_2)J_2,$$

this completes the proof. □

Remark 6. *$L_1 \Sigma L_1'$ and $L_2' Q^{-1}(\phi) L_2$ are positive definite matrices since L_1 and L_2 are full rank matrices and from Proposition* 3.2, *we know that Σ and $Q(\phi)$ are positive definite matrices.*

Proof of Proposition 3.4. Note that $X_t = X_1(t)\mathbb{I}_{\{t\leq\gamma\}} + X_2(t)\mathbb{I}_{\{t>\gamma\}}$, $0 \leq t \leq T$ where

$$X_1(t) = h_1(t) + Z_1(t), \quad X_2(t) = h_2(t) + Z_2(t), \quad 0 \leq t \leq T, \tag{B.33}$$

with h_1, h_2, Z_1, Z_2 defined in (3.11). By Assumption 1, we have the distribution of X_0 does not depend on $\theta = \left[\begin{array}{c:c} \theta_1 & \theta_2 \end{array}\right]$. Since $X_1(t) = X_1(t)\mathbb{I}_{\{t\leq\gamma\}} + X_1(t)\mathbb{I}_{\{t>\gamma\}}$, we know that the distribution of $X_1(0)$ is the same as the distribution of X_0, which does not depend on θ_1. As a result, $\mathrm{E}(\|X_1(0)\|_2^m) = \mathrm{E}(\|X_0\|_2^m) < \infty$. Then the result follows from the Proposition 2.10, which completes the proof. Moreover, from Proposition 3.2 and Proposition 3.4, it is sufficient to prove that Σ_0 is a positive definite matrix. First, by Schur Complement Theorem, we have Σ_0 is positive definite if and only if $\int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt + V_1(0) - \int_0^1 \widetilde{h}_1(t)\varphi'(t)dt \int_0^1 \varphi(t)\widetilde{h}_1'(t)dt$ is positive definite. Further, let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_d$ be the eigenvalues of

$$\int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt - \int_0^1 \widetilde{h}_1(t)\varphi'(t)dt \int_0^1 \varphi(t)\widetilde{h}_1'(t)dt.$$

By Theorem A.2 in Appendix A, we have

$$\begin{aligned}
\lambda_d &= \min_{y\in\mathbb{R}^d:\|y\|_2=1} y'\left(\int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt - \int_0^1 \widetilde{h}_1(t)\varphi'(t)dt \int_0^1 \varphi(t)\widetilde{h}_1'(t)dt\right) y \\
&= \min_{y\in\mathbb{R}^d:\|y\|_2=1} \left(\int_0^1 (y'\widetilde{h}_1(t))(\widetilde{h}_1'(t)y)dt - \int_0^1 (y'\widetilde{h}_1(t))\varphi'(t)dt \int_0^1 \varphi(t)(\widetilde{h}_1'(t)y)dt\right) \\
&= \min_{y\in\mathbb{R}^d:\|y\|_2=1} \left(\int_0^1 (y'\widetilde{h}_1(t))(\widetilde{h}_1'(t)y)dt - \sum_{i=1}^p \left(\int_0^1 (y'\widetilde{h}_1(t))\varphi_i(t)dt\right)^2\right) \\
&= \min_{y\in\mathbb{R}^d:\|y\|_2=1} \left(\int_0^1 (y'\widetilde{h}_1(t))(\widetilde{h}_1'(t)y)dt - \sum_{i=1}^p \left(\int_0^1 (y'\widetilde{h}_1(t))\frac{\varphi_i(t)}{\|\varphi_i(t)\|}\|\varphi_i(t)\|dt\right)^2\right).
\end{aligned}$$

Since $\|\varphi_i(t)\|^2 = \int_0^1 (\varphi_i(t))^2 dt = 1$, by Bessel's inequality, we get

$$\int_0^1 (y'\widetilde{h}_1(t))(\widetilde{h}_1'(t)y)dt - \sum_{i=1}^{p} \left(\int_0^1 (y'\widetilde{h}_1(t))\varphi_i(t)dt \right)^2 \geq 0.$$

Thus, since the matrix is symmetric with all the eigenvalues are nonnegative, we have $\int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt - \int_0^1 \widetilde{h}_1(t)\varphi'(t)dt \int_0^1 \varphi(t)\widetilde{h}_1'(t)dt$ is a positive semi-definite matrix. Moreover, by Proposition 2.7, $V_1(0)$ is a positive definite matrix. Therefore $\int_0^1 \widetilde{h}_1(t)\widetilde{h}_1'(t)dt + V_1(0) - \int_0^1 \widetilde{h}_1(t)\varphi'(t)dt \int_0^1 \varphi(t)\widetilde{h}_1'(t)dt$ is positive definite, which implies that Σ_0 is a positive definite matrix. Further, let $g(\mathbf{X}) = \mathbf{X}^{-1}$ for a positive definite matrix $\mathbf{X}$. Therefore, by the continuous mapping theorem, we have

$$g\left(\frac{1}{T}Q_\gamma\right) = TQ_\gamma^{-1} \xrightarrow[T\to\infty]{P} g(\phi\Sigma_0) = \frac{1}{\phi}\Sigma_0^{-1},$$

which completes the proof. $\square$

Proof of Proposition 3.7. From the SDE in (2.1), we have

$$\begin{aligned}\int_0^T dX_t B(t,\phi) = &\int_0^T [(\mu_1\varphi(t) - A_1X_t)\mathbb{I}_{\{t\leq\gamma\}} + (\mu_2\varphi(t) - A_2X_t)\mathbb{I}_{\{t>\gamma\}}]B(t,\phi)dt \\ &+ \int_0^T \Sigma^{1/2} dW_t B(t,\phi).\end{aligned}$$

Further, using the notations defined in (3.1) and (3.4), we have

$$\begin{aligned}&\int_0^T dX_t \begin{bmatrix} \varphi'(t)\mathbb{I}_{\{t\leq\gamma\}} & -X_t'\mathbb{I}_{\{t\leq\gamma\}} & \varphi'(t)\mathbb{I}_{\{t>\gamma\}} & -X_t'\mathbb{I}_{\{t>\gamma\}} \end{bmatrix} \\ &= \begin{bmatrix} \int_0^\gamma dX_t\varphi'(t) & -\int_0^\gamma dX_tX_t' & \int_\gamma^T dX_t\varphi'(t) & -\int_\gamma^T dX_tX_t' \end{bmatrix}.\end{aligned}$$

Then

$$\int_0^T dX_t B(t,\phi) = \begin{bmatrix} P_\gamma' & P_{\gamma,T}' \end{bmatrix} = P(\phi). \tag{B.34}$$

Note that $\mathbb{I}_{\{t\leq\gamma\}}\mathbb{I}_{\{t>\gamma\}} = 0$ for all t, then

$$\int_0^T [(\mu_1\varphi(t) - A_1X_t)\mathbb{I}_{\{t\leq\gamma\}} + (\mu_2\varphi(t) - A_2X_t)\mathbb{I}_{\{t>\gamma\}}]B(t,\phi)dt$$

can be expressed as

$$\begin{bmatrix}\mu_1 & A_1 & \mu_2 & A_2\end{bmatrix}\begin{bmatrix}\int_0^\gamma \varphi(t)\varphi'(t)dt & -\int_0^\gamma \varphi(t)X_t'dt & 0 & 0\\ -\int_0^\gamma X_t\varphi'(t)dt & \int_0^\gamma X_tX_t'dt & 0 & 0\\ 0 & 0 & \int_\gamma^T \varphi(t)\varphi'(t)dt & -\int_\gamma^T \varphi(t)X_t'dt\\ 0 & 0 & -\int_\gamma^T X_t\varphi'(t)dt & \int_\gamma^T X_tX_t'dt\end{bmatrix},$$

Then, by combining, (2.2), (3.5), and (B.34), we get

$$P(\phi) = \theta Q(\phi) + \int_0^T \Sigma^{1/2} dW_t B(t,\phi),$$
$$P(\phi)Q^{-1}(\phi) = \theta + \int_0^T \Sigma^{1/2} dW_t B(t,\phi)Q^{-1}(\phi).$$

Then, from (B.24), we get

$$\hat{\theta} - \theta = \Sigma^{1/2}\int_0^T dW_t B(t,\phi)Q^{-1}(\phi).$$

Then, letting $R_T'(\phi) = \displaystyle\int_0^T B'(t,\phi)dW_t'$, we complete the proof. $\square$

Proof of Proposition 3.9. To prove this proposition, we directly apply Proposition 1.21 in Kutoyants (2004) with $d_1 = 1$ and $d_2 = d$. First, in Proposition 3.8, we have verified the conditions to apply Proposition 1.21 in Kutoyants (2004), i.e. we have $\mathrm{P}(\int_0^T (a^{(i)}C_T(t)^2 dt < \infty) = 1$. We have

$$\sum_{i=1}^d \int_0^T (a^{(i)}C_T(t))^2 dt = \int_0^T \sum_{i=1}^d (a^{(i)}C_T(t))^2 dt.$$

Note that since $a = \begin{bmatrix}a^{(1)} & a^{(2)} & a^{(3)} & \ldots & a^{(d)}\end{bmatrix}$, we have

$$\sum_{i=1}^d (a^{(i)}C_T(t))^2 = a'(I_d \otimes C_T(t))(I_d \otimes C_T'(t))a.$$

Therefore

$$\sum_{i=1}^{d}\int_0^T (a^{(i)}C_T(t))^2 dt = \int_0^T a'(I_d \otimes C_T(t))(I_d \otimes C_T'(t))a dt$$
$$= \int_0^T a'(I_d \otimes C_T(t)C_T'(t))a dt = a'\left(I_d \otimes \int_0^T C_T(t)C_T'(t)dt\right)a.$$

Since $\mathbb{I}_{\{t\leq\gamma\}}\mathbb{I}_{\{t>\gamma\}} = 0$ for all t, we have

$$\int_0^T \frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{t\leq\gamma\}}\frac{1}{\sqrt{T}}X_t'\mathbb{I}_{\{t>\gamma\}}dt = 0,\ \int_0^T \frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{t\leq\gamma\}}\frac{1}{\sqrt{T}}X_t'\mathbb{I}_{\{t>\gamma\}}dt = 0,$$
$$\int_0^T \frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{t\leq\gamma\}}\frac{1}{\sqrt{T}}\varphi'(t)\mathbb{I}_{\{t>\gamma\}}dt = 0,\ \int_0^T \frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{t\leq\gamma\}}\frac{1}{\sqrt{T}}\varphi'(t)\mathbb{I}_{\{t>\gamma\}}dt = 0.$$

Also, one can easily verify that $\int_0^T C_T(t)C_T'(t)dt = \frac{1}{T}Q(\phi)$, we get

$$a'\left(I_d \otimes \int_0^T C_T(t)C_T'(t)dt\right)a = a'\left(I_d \otimes \frac{1}{T}Q(\phi)\right)a$$

where $Q(\phi)$ is defined in (3.5). From Proposition 3.5, we have

$$\frac{1}{T}Q(\phi) \xrightarrow[T\to\infty]{P} \Sigma_2.$$

Therefore,

$$\sum_{i=1}^{d}\int_0^T (a^{(i)}C_T(t))^2 dt \xrightarrow[T\to\infty]{P} a'(I_d \otimes \Sigma_2)a.$$

By Proposition 1.21 in Kutoyants (2004), we have

$$a'\mathrm{Vec}\left(\frac{1}{\sqrt{T}}R_T'(\phi)\right) \xrightarrow[T\to\infty]{d} a'\mathcal{N}_{2(p+d)d}(0, I_d \otimes \Sigma_2).$$

By Cramer-Wold Theorem, we get

$$\mathrm{Vec}\left(\frac{1}{\sqrt{T}}R_T'(\phi)\right) \xrightarrow[T\to\infty]{d} \mathcal{N}_{2(p+d)d}(0, I_d \otimes \Sigma_2),$$

which completes the proof. □

Proof of Proposition 3.10. By combining Proposition 3.6, Proposition 3.9, Proposition 3.7 and Slutsky's theorem, we get

$$\sqrt{T}(\hat{\theta}-\theta)' = (TQ^{-1}(\phi))\frac{1}{\sqrt{T}}R_T'(\phi)\Sigma^{1/2} \xrightarrow[T\to\infty]{d} \Sigma_2^{-1}R\Sigma^{1/2}.$$

Note that $\Sigma^{1/2}$ and Σ_2^{-1} are non-random and symmetric matrices, we get

$$\Sigma_2^{-1}R\Sigma^{1/2} \sim \mathcal{N}_{2(p+d)\times d}(0, (\Sigma^{1/2}I_d\Sigma^{1/2}) \otimes (\Sigma_2^{-1}\Sigma_2\Sigma_2^{-1})) = \mathcal{N}_{2(p+d)\times d}(0, \Sigma \otimes \Sigma_2^{-1}),$$

which completes the proof. □

Proof of Proposition 3.11

Proof. From (3.26), we have

$$\begin{aligned}
\begin{bmatrix} \sqrt{T}(\hat{\theta}-\theta) \\ \sqrt{T}(\widetilde{\theta}-\theta) \end{bmatrix} &= \begin{bmatrix} \sqrt{T}(\hat{\theta}-\theta) \\ J\sqrt{T}(\hat{\theta}-\theta)J_4 + J_6 \end{bmatrix} \\
&= \begin{bmatrix} I_d \\ 0_d \end{bmatrix}\sqrt{T}(\hat{\theta}-\theta) + \begin{bmatrix} 0_d \\ J \end{bmatrix}\sqrt{T}(\hat{\theta}-\theta)J_4 + \begin{bmatrix} 0_{d\times 2(p+d)} \\ J_6 \end{bmatrix}, \qquad \text{(B.35)}
\end{aligned}$$

where $J = I_d - J_1L_1$, J_4 and J_6 are defined in (3.28) and (3.29). Further, denote

$$I^{(1)} = \begin{bmatrix} I_d \\ 0_d \end{bmatrix} \in \mathbb{R}^{2d\times d},\ I^{(2)} = \begin{bmatrix} 0_d \\ J \end{bmatrix} \in \mathbb{R}^{2d\times d},\ \text{and } I^{(3)} = \begin{bmatrix} 0_{d\times 2(p+d)} \\ J_6 \end{bmatrix} \in \mathbb{R}^{2d\times 2(p+d)}. \qquad \text{(B.36)}$$

From (B.35) and (B.36), we get

$$\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} = \begin{bmatrix} \sqrt{T}(\hat{\theta}-\theta) \\ \sqrt{T}(\widetilde{\theta}-\theta) \end{bmatrix}' = \rho_T I^{(1)'} + J_4'\rho_T I^{(2)'} + I^{(3)'}. \qquad \text{(B.37)}$$

Using vectorization, we get

$$\begin{aligned}
\mathrm{Vec}\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} &= (I^{(1)} \otimes I_{2(p+d)})\mathrm{Vec}(\rho_T) + (I^{(2)} \otimes J_4')\mathrm{Vec}(\rho_T) + \mathrm{Vec}(I^{(3)}) \\
&= (I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_4')\mathrm{Vec}(\rho_T) + \mathrm{Vec}(I^{(3)'}).
\end{aligned}$$

By (3.28) and (3.29), we have

$$J_4 = I_{2(p+d)} - L_2 J_2 \xrightarrow[T\to\infty]{P} I_{2(p+d)} - L_2 J_3 = J_5,$$

$$J_6 = J_1 L_1 r_2 J_2 - r_2 J_2 \xrightarrow[T\to\infty]{P} J_1 L_1 r_2 J_3 - r_2 J_3 = J_7.$$

Therefore

$$I^{(3)} = \begin{bmatrix} 0_{d\times 2(p+d)} \\ J_6 \end{bmatrix} \xrightarrow[T\to\infty]{P} \begin{bmatrix} 0_{d\times 2(p+d)} \\ J_7 \end{bmatrix} = I^{(4)}. \tag{B.38}$$

By (B.37), we know

$$\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} = \begin{bmatrix} \sqrt{T}(\hat{\theta}-\theta) \\ \sqrt{T}(\widetilde{\theta}-\theta) \end{bmatrix}' = \rho_T I^{(1)'} + J_4' \rho_T I^{(2)'} + I^{(3)'}.$$

Using vectorization, we get

$$\begin{aligned} \mathrm{Vec}\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} &= (I^{(1)} \otimes I_{2(p+d)})\mathrm{Vec}(\rho_T) + (I^{(2)} \otimes J_4')\mathrm{Vec}(\rho_T) + \mathrm{Vec}(I^{(3)}) \\ &= (I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_4')\mathrm{Vec}(\rho_T) + \mathrm{Vec}(I^{(3)'}), \end{aligned}$$

where J_4 and J_6 are defined in (3.28) and (3.29), $I^{(1)}$, $I^{(2)}$ and $I^{(3)}$ are defined in (B.36). Also by Proposition 3.10, we have

$$\mathrm{Vec}(\rho_T) \xrightarrow[T\to\infty]{d} \mathcal{N}_{2d(p+d)}(0, \Sigma \otimes \Sigma_2^{-1}). \tag{B.39}$$

Therefore, combining (3.28) and (B.38), by Slutsky's Theorem, we have

$$\begin{bmatrix} \rho_T & \zeta_T \end{bmatrix} \xrightarrow[T\to\infty]{d} \begin{bmatrix} \rho & \zeta \end{bmatrix}, \text{ where}$$

$$\begin{bmatrix} \rho & \zeta \end{bmatrix} \sim \mathcal{N}_{2(p+d)\times 2d}(I^{(4)'}, (I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_5')(\Sigma \otimes \Sigma_2^{-1})(I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_5')').$$

To simplify the covariance term, we have

$$I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_5' = \begin{bmatrix} I_d \\ 0_d \end{bmatrix} \otimes I_{2(p+d)} + \begin{bmatrix} I_d \\ J \end{bmatrix} \otimes J_5' = \begin{bmatrix} I_d \otimes I_{2(p+d)} \\ J \otimes J_5' \end{bmatrix} = \begin{bmatrix} I_{2d(p+d)} \\ J \otimes J_5' \end{bmatrix}.$$

Therefore $(I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_5')(\Sigma \otimes \Sigma_2^{-1})(I^{(1)} \otimes I_{2(p+d)} + I^{(2)} \otimes J_5')'$

$$
\begin{aligned}
&= \begin{bmatrix} I_{2d(p+d)} \\ J \otimes J_5' \end{bmatrix} (\Sigma \otimes \Sigma_2^{-1}) \begin{bmatrix} I_{2d(p+d)} \\ J \otimes J_5' \end{bmatrix}' \\
&= \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} \\ (J \otimes J_5')(\Sigma \otimes \Sigma_2^{-1}) \end{bmatrix} \begin{bmatrix} I_{2d(p+d)} \\ J \otimes J_5' \end{bmatrix}' \\
&= \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & (\Sigma \otimes \Sigma_2^{-1})(J' \otimes J_5) \\ (J \otimes J_5')(\Sigma \otimes \Sigma_2^{-1}) & (J \otimes J_5')(\Sigma \otimes \Sigma_2^{-1})(J' \otimes J_5) \end{bmatrix} \\
&= \begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & (\Sigma J') \otimes (\Sigma_2^{-1} J_5) \\ (J\Sigma) \otimes (J_5' \Sigma_2^{-1}) & (J \Sigma J') \otimes (J_5' \Sigma_2^{-1} J_5) \end{bmatrix}.
\end{aligned}
$$

From (3.32), we know that $J\Sigma J' = J\Sigma = \Sigma J'$. Also, from (3.35), we know that $J_5'\Sigma_2^{-1}J_5 = \Sigma_2^{-1}J_5 = J_5'\Sigma_2^{-1}$. Therefore, the covariance term is

$$
\begin{bmatrix} \Sigma \otimes \Sigma_2^{-1} & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \\ (J\Sigma) \otimes (\Sigma_2^{-1} J_5) & (J\Sigma) \otimes (\Sigma_2^{-1} J_5) \end{bmatrix},
$$

which completes the proof. □

Proof of Proposition 4.2. From Proposition 3.6 we have

$$
\frac{1}{T}\int_0^{\phi T} X_t X_t' dt \xrightarrow[T\to\infty]{P} \phi \left\{ \int_0^1 \widetilde{h}_1(t) \widetilde{h}_1'(t) dt + V_1(0) \right\}.
$$

Therefore, it sufficies to prove that

$$
\frac{1}{T}\int_0^{\hat{\phi} T} X_t X_t' dt - \frac{1}{T}\int_0^{\phi T} X_t X_t' dt \xrightarrow[T\to\infty]{P} 0.
$$

First, let $0 < \delta < \frac{\phi}{2}$. We have

$$
\lim_{T\to\infty} \mathrm{P}(|\hat{\phi} - \phi| > \delta) = 0. \tag{B.40}
$$

Further, we have

$$
\begin{aligned}
&\mathrm{P}\left(\left\|\frac{1}{T}\int_0^{\hat{\phi}T} X_tX_t'dt - \frac{1}{T}\int_0^{\phi T} X_tX_t'dt\right\|_F > \epsilon\right) \\
&= \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{\hat{\phi}T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \epsilon, |\hat{\phi}-\phi| > \delta\right) \\
&+ \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{\hat{\phi}T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \epsilon, |\hat{\phi}-\phi| \le \delta\right) \\
&\le \mathrm{P}(|\hat{\phi}-\phi| > \delta) + \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{\hat{\phi}T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \epsilon, |\hat{\phi}-\phi| \le \delta\right).
\end{aligned}
$$

By (B.40), it is suffices to prove that

$$
\lim_{T\to\infty} \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{\hat{\phi}T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \epsilon, |\hat{\phi}-\phi| \le \delta\right).
$$

Note that $\{|\hat{\phi}-\phi| \le \delta\}$ is the same as $\{(\phi-\delta) \le \hat{\phi} \le (\phi+\delta)\}$. We have

$$
\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_0^{(\phi-\delta)T} X_tX_t'dt + \int_{(\phi-\delta)T}^{\hat{\phi}T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \epsilon, |\hat{\phi}-\phi| \le \delta\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\left\|\int_{(\phi-\delta)T}^{\hat{\phi}T} X_tX_t'dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right) \\
&+ \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{(\phi-\delta)T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_tX_t'\|_F dt > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right) \\
&+ \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{(\phi-\delta)T} X_tX_t'dt - \int_0^{\phi T} X_tX_t'dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right),
\end{aligned}
$$

then

$$
\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_0^{\hat{\phi}T} X_t X_t' dt - \int_0^{\phi T} X_t X_t' dt\right\|_F > \epsilon, |\hat{\phi}-\phi| \le \delta\right) \\
&= \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{(\phi-\delta)T} X_t X_t' dt + \int_{(\phi-\delta)T}^{\hat{\phi}T} X_t X_t' dt - \int_0^{\phi T} X_t X_t' dt\right\|_F > \epsilon, |\hat{\phi}-\phi| \le \delta\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right) \\
&+ \mathrm{P}\left(\frac{1}{T}\left\|\int_0^{(\phi-\delta)T} X_t X_t' dt - \int_0^{\phi T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right). \qquad \text{(B.41)}
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_0^{\phi T} X_t X_t' dt - \int_{(\phi-\delta)T}^{\phi T} X_t X_t' dt - \int_0^{\phi T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right) \\
&= \mathrm{P}\left(\frac{1}{T}\left\|\int_{(\phi-\delta)T}^{\phi T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right),
\end{aligned}
$$

then

$$
\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_0^{(\phi-\delta)T} X_t X_t' dt - \int_0^{\phi T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}, |\hat{\phi}-\phi| \le \delta\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right). \qquad \text{(B.42)}
\end{aligned}
$$

Thus, from (B.41) and (B.42), it is suffices to prove that

$$
\lim_{T\to\infty} \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right) = 0.
$$

Now, by Markov Inequality, we have

$$
\begin{aligned}
\mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right) &\le \frac{2\mathrm{E}(\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt)}{\epsilon T} \\
&= \frac{2\int_{(\phi-\delta)T}^{(\phi+\delta)T} \mathrm{E}(\|X_t\|_2^2) dt}{\epsilon T} \le \frac{4K_x \delta T}{\epsilon T} = \frac{4K_x\delta}{\epsilon}.
\end{aligned} \qquad \text{(B.43)}
$$

Note that $K_x < \infty$ and we can choose δ arbitrarily small, which completes the proof of part (i). For part (ii), using the same method as we did in Part (i), and note that

$$\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_{\hat{\phi}T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \epsilon, |\hat{\phi} - \phi| \le \delta\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\int_{\hat{\phi}T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right) + \mathrm{P}\left(\frac{1}{T}\left\|\int_{(\phi+\delta)T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right) + \mathrm{P}\left(\frac{1}{T}\left\|\int_{(\phi+\delta)T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}\right).
\end{aligned}$$

Also, we have

$$\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_{(\phi+\delta)T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}\right) \\
&= \mathrm{P}\left(\frac{1}{T}\left\|\int_{\phi T}^{T} X_t X_t' dt - \int_{\phi T}^{(\phi+\delta)T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}\right) \\
&= \mathrm{P}\left(\frac{1}{T}\left\|\int_{\phi T}^{(\phi+\delta)T} X_t X_t' dt\right\|_F > \frac{\epsilon}{2}\right) \le \mathrm{P}\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right).
\end{aligned}$$

This implies the fact that

$$\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_{\hat{\phi}T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \epsilon, |\hat{\phi} - \phi| \le \delta\right) \\
&\le 2P\left(\frac{1}{T}\int_{(\phi-\delta)T}^{(\phi+\delta)T} \|X_t X_t'\|_F dt > \frac{\epsilon}{2}\right).
\end{aligned}$$

Note that

$$\begin{aligned}
&\mathrm{P}\left(\frac{1}{T}\left\|\int_{\hat{\phi}T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \epsilon\right) \\
&\le \mathrm{P}\left(\frac{1}{T}\left\|\int_{\hat{\phi}T}^{T} X_t X_t' dt - \int_{\phi T}^{T} X_t X_t' dt\right\|_F > \epsilon, |\hat{\phi} - \phi| \le \delta\right) + \mathrm{P}\left(|\hat{\phi} - \phi| > \delta\right).
\end{aligned}$$

By (B.40) and (B.43), we complete the proof. □

Proof of Proposition 4.5. Since

$$\frac{1}{\sqrt{T}} R_T'(\hat{\phi}) = \frac{1}{\sqrt{T}}(R_T'(\hat{\phi}) - R_T'(\phi)) + \frac{1}{\sqrt{T}} R_T'(\phi).$$

From Proposition 3.9, Proposition 4.4, and Slutsky's Theorem, we complete the proof.

$\square$

Proof of Theorem 5.1. From (5.2), we have

$$\mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) = \mathrm{E}(\mathrm{Tr}(\rho' W \rho)).$$

From Corollary 4.1, we have

$$\rho \sim \mathcal{N}_{2(p+d)\times d}(0, \Sigma \otimes \Sigma_2^{-1}),$$

then $\mathrm{Vec}(\rho) \sim \mathcal{N}_{2d(p+d)}(0, \Sigma \otimes \Sigma_2^{-1})$, we get

$$\mathrm{E}(\mathrm{Vec}(\rho)\mathrm{Vec}(\rho)') = \Sigma \otimes \Sigma_2^{-1},$$

$$(I_d \otimes W)\mathrm{E}(\mathrm{Vec}(\rho)\mathrm{Vec}(\rho)') = (I_d \otimes W)(\Sigma \otimes \Sigma_2^{-1}).$$

Since $(I_d \otimes W)\mathrm{Vec}(\rho) = \mathrm{Vec}(W\rho)$ and $(I_d \otimes W)(\Sigma \otimes \Sigma_2^{-1}) = \Sigma \otimes W\Sigma_2^{-1}$, we have

$$\mathrm{E}(\mathrm{Vec}(W\rho)\mathrm{Vec}(\rho)') = \Sigma \otimes W\Sigma_2^{-1},$$

$$\mathrm{E}(\mathrm{Tr}(\mathrm{Vec}(\rho)'\mathrm{Vec}(W\rho))) = \mathrm{Tr}(\Sigma \otimes W\Sigma_2^{-1}).$$

Using $\mathrm{Tr}(AB) = (\mathrm{Vec}(A'))'\mathrm{Vec}(B)$, and $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\,\mathrm{Tr}(B)$, we get

$$\mathrm{E}(\mathrm{Tr}(\mathrm{Vec}(\rho)'\mathrm{Vec}(W\rho))) = \mathrm{E}(\mathrm{Tr}(\rho' W\rho)) \text{ and } \mathrm{Tr}(\Sigma \otimes W\Sigma_2^{-1}) = \mathrm{Tr}(\Sigma)Tr(W\Sigma_2^{-1}).$$

This gives the ADR of the UE. Further, from (5.2), we have

$$\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}), \theta, W) = \mathrm{E}(\mathrm{Tr}(\zeta' W \zeta)).$$

From Corollary 4.1, we have $\mathrm{Vec}(\zeta) \sim \mathcal{N}_{2d(p+d)}(\mathrm{Vec}(J_7'), (J\Sigma) \otimes (\Sigma_2^{-1}J_5))$. Then

$$\mathrm{E}(\mathrm{Vec}(\zeta)\mathrm{Vec}(\zeta)') = (J\Sigma) \otimes (\Sigma_2^{-1}J_5) + \mathrm{Vec}(J_7')\mathrm{Vec}(J_7')'.$$

Using $\mathrm{Tr}(AB) = (\mathrm{Vec}(A'))'\mathrm{Vec}(B)$, we have

$$\mathrm{Tr}(\zeta' W \zeta) = \mathrm{Vec}(\zeta)'\mathrm{Vec}(W\zeta) = \mathrm{Vec}(\zeta)'(I_d \otimes W)\mathrm{Vec}(\zeta),$$

then

$$\mathrm{Vec}(\zeta)'(I_d \otimes W)\mathrm{Vec}(\zeta) = \mathrm{Tr}(\mathrm{Vec}(\zeta)'(I_d \otimes W)\mathrm{Vec}(\zeta)) = \mathrm{Tr}((I_d \otimes W)\mathrm{Vec}(\zeta)\mathrm{Vec}(\zeta)').$$

Therefore, we have

$$\begin{aligned}
\mathrm{E}(\mathrm{Tr}(\zeta' W \zeta)) &= \mathrm{Tr}[(I_d \otimes W)\mathrm{E}(\mathrm{Vec}(\zeta)\mathrm{Vec}(\zeta)')] \\
&= \mathrm{Tr}[(I_d \otimes W)((J\Sigma) \otimes (\Sigma_2^{-1} J_5) + \mathrm{Vec}(J_7')\mathrm{Vec}(J_7')')] \\
&= \mathrm{Tr}[(J\Sigma) \otimes (W\Sigma_2^{-1} J_5)] + \mathrm{Tr}[(I_d \otimes W)\mathrm{Vec}(J_7')\mathrm{Vec}(J_7')'].
\end{aligned}$$

Note that $\mathrm{Tr}[(J\Sigma) \otimes (W\Sigma_2^{-1} J_5)] = \mathrm{Tr}(J\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} J_5)$, and

$$\begin{aligned}
\mathrm{Tr}[(I_d \otimes W)\mathrm{Vec}(J_7')\mathrm{Vec}(J_7')'] &= \mathrm{Vec}(J_7')'(I_d \otimes W)\mathrm{Vec}(J_7') \\
&= \mathrm{Vec}(J_7')'\mathrm{Vec}(W J_7') = \mathrm{Tr}(J_7 W J_7'),
\end{aligned}$$

Since $J = I_d - J_1 L_1$ and $J_5 = I_{2(p+d)} - L_2 J_3$ with J_3 defined in (4.11), we get

$$\begin{aligned}
\mathrm{E}(\mathrm{Tr}(\zeta' W \zeta)) &= \mathrm{Tr}(J\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} J_5) + \mathrm{Tr}(J_7 W J_7') \\
&= \mathrm{Tr}((I_d - J_1 L_1)\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}(I_{2(p+d)} - L_2 J_3)) + \mathrm{Tr}(J_7 W J_7') \\
&= \mathrm{Tr}(\Sigma - J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} - W\Sigma_2^{-1} L_2 J_3) + \mathrm{Tr}(J_7 W J_7') \\
&= \mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) - \mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) - \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) \\
&+ \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) + \mathrm{Tr}(J_7 W J_7'),
\end{aligned}$$

which completes the proof. □

Proof of Theorem 5.2. Note that

$$\begin{aligned}\mathrm{ADR}(\hat{\theta}^S,\theta,W) &= \mathrm{E}[\mathrm{Tr}((\zeta+[1-(nd-2)\psi^{-1}]\xi)'W(\zeta+[1-(nd-2)\psi^{-1}]\xi))] \\ &= \mathrm{E}[\mathrm{Tr}(\zeta'W\zeta)] + \mathrm{E}[\mathrm{Tr}(\zeta'W[1-(nd-2)\psi^{-1}]\xi)] \\ &+ \mathrm{E}[\mathrm{Tr}([1-(nd-2)\psi^{-1}]\xi'W\zeta)] \\ &+ \mathrm{E}[\mathrm{Tr}([1-(nd-2)\psi^{-1}]^2\xi'W\xi)],\end{aligned}$$

then

$$\begin{aligned}\mathrm{ADR}(\hat{\theta}^S,\theta,W) = {}& \mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W) + 2\mathrm{E}[\mathrm{Tr}(\zeta'W[1-(nd-2)\psi^{-1}]\xi)] \\ &+ \mathrm{E}[\mathrm{Tr}([1-(nd-2)\psi^{-1}]^2\xi'W\xi)]\end{aligned}$$

From Proposition 4.8 and Proposition $A.4$ in the Appendix A, we get

$$\begin{aligned}&\mathrm{E}[\mathrm{Tr}([1-(nd-2)\psi^{-1}]^2\xi'W\xi)] \\ &= \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma) \\ &+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma) \\ &+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2]\,\mathrm{Tr}(J_7WJ_7'),\end{aligned} \tag{B.44}$$

also, we have

$$\mathrm{E}[\zeta'W[1-(nd-2)\psi^{-1}]\xi] = -\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))]J_7WJ_7', \tag{B.45}$$

where $\Delta = \mathrm{Tr}(J_7\Xi J_7'\Sigma^{-1})$. From (B.44) and (B.45), we get

$$\begin{aligned}\mathrm{ADR}(\hat{\theta}^S,\theta,W) &= \mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W) - 2\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))]\,\mathrm{Tr}(J_7WJ_7') \\ &+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma) \\ &+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma) \\ &+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2]\,\mathrm{Tr}(J_7WJ_7').\end{aligned}$$

To further simplify the terms, note that

$$\begin{aligned}\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma) &= \mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\ &\quad -2(nd-2)\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\ &\quad +(nd-2)^2\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma),\end{aligned}$$

also

$$\begin{aligned}\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2]\,\mathrm{Tr}(J_7WJ_7') &= \mathrm{Tr}(J_7WJ_7')\\ &\quad -2(nd-2)\mathrm{E}[\chi^{-2}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7WJ_7')\\ &\quad +(nd-2)^2\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7WJ_7').\end{aligned}$$

Note that from Theorem 5.1, we have

$$\begin{aligned}\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W) &= \mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) - \mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3) - \mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1})\\ &\quad + \mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3) + \mathrm{Tr}(J_7WJ_7'),\end{aligned}$$

also, note that $\mathrm{ADR}(\hat{\theta}(\hat{\phi}),\theta,W) = \mathrm{Tr}(\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1})$, we get

$$\begin{aligned}&\mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W) - 2Tr(J_7WJ_7') + \mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma) + Tr(J_7WJ_7')\\ &= \mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W) - Tr(J_7WJ_7') + \mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\ &= \mathrm{ADR}(\hat{\theta}(\hat{\phi}),\theta,W) - \mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) + \mathrm{Tr}(J_1L_1\Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3).\end{aligned}$$

Then, using the identity $\mathrm{E}[\chi^{-2}_{nd+4}(\Delta)] = \mathrm{E}[\chi^{-2}_{nd+2}(\Delta)] - 2\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]$, we get

$$
\begin{aligned}
&\mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) - \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) + \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) \\
&+ 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') - 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \\
&+ (nd-2)^2\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) - 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') \\
&+ (nd-2)^2\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') \\
&= \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) - \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) + \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) \\
&+ 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') + 4(nd-2)\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') \\
&- 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \\
&+ (nd-2)^2\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) - 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') \\
&+ (nd-2)^2\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7'),
\end{aligned}
$$

then, we have $\mathrm{ADR}(\hat{\theta}^S, \theta, W)$ is equal to

$$
\begin{aligned}
&\mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) - \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) + \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) \\
&+ 4(nd-2)\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') - 2(nd-2)\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \\
&+ (nd-2)^2\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)]\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) + (nd-2)^2\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7') \\
&+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3))\,\mathrm{Tr}(\Sigma) \\
&= \mathrm{ADR}(\hat{\theta}(\hat{\phi}), \theta, W) - \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1}) + \mathrm{Tr}(J_1 L_1 \Sigma)\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3) \\
&- (nd-2)(2\mathrm{E}[\chi^{-2}_{nd+2}(\Delta)] - (nd-2)\mathrm{E}[\chi^{-4}_{nd+2}(\Delta)])\,\mathrm{Tr}(W\Sigma_2^{-1} L_2 J_3)\,\mathrm{Tr}(\Sigma) \\
&+ \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2]\,\mathrm{Tr}(W(\Sigma_2^{-1} - \Sigma_2^{-1} L_2 J_3))\,\mathrm{Tr}(\Sigma) \\
&+ ((nd)^2 - 4)\mathrm{E}[\chi^{-4}_{nd+4}(\Delta)]\,\mathrm{Tr}(J_7 W J_7').
\end{aligned}
$$

This gives the ADR of the SE. Further, note that $\psi > 0$ and $nd - 2 > 0$, then $1 - (nd-2)\psi^{-1} \geq 0$ if and only if $\psi \geq nd - 2$. Following the same steps above, we get

$$\begin{aligned}
\mathrm{ADR}(\hat{\theta}^{S+},\theta,W) &= \mathrm{ADR}(\widetilde{\theta}(\hat{\phi}),\theta,W)\\
&\quad - 2\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)\geq nd-2\}}]\,\mathrm{Tr}(J_7 W J_7')\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)\geq nd-2\}}]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma)\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)\geq nd-2\}}]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)\geq nd-2\}}]\,\mathrm{Tr}(J_7 W J_7').
\end{aligned}$$

Also, note that

$$\begin{aligned}
\mathrm{E}[1-(nd-2)\chi^{-2}_{nd+2}(\Delta)] &= \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)\geq nd-2\}}]\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)< nd-2\}}],\\
\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2] &= \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)\geq nd-2\}}]\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)< nd-2\}}],\\
\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2] &= \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)\geq nd-2\}}]\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)< nd-2\}}],\\
\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2] &= \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)\geq nd-2\}}]\\
&\quad + \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)< nd-2\}}].
\end{aligned}$$

Therefore, we have $\mathrm{ADR}(\hat{\theta}^{S+},\theta,W)$ is equal to

$$\begin{aligned}
&\mathrm{ADR}(\hat{\theta}^{S},\theta,W) + 2\mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)< nd-2\}}]\,\mathrm{Tr}(J_7 W J_7')\\
&- \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd}(\Delta)< nd-2\}}]\,\mathrm{Tr}(W(\Sigma_2^{-1}-\Sigma_2^{-1}L_2J_3))\,\mathrm{Tr}(\Sigma)\\
&- \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+2}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+2}(\Delta)< nd-2\}}]\,\mathrm{Tr}(W\Sigma_2^{-1}L_2J_3)\,\mathrm{Tr}(\Sigma)\\
&- \mathrm{E}[(1-(nd-2)\chi^{-2}_{nd+4}(\Delta))^2\mathbb{I}_{\{\chi^2_{nd+4}(\Delta)< nd-2\}}]\,\mathrm{Tr}(J_7 W J_7'),
\end{aligned}$$

which completes the proof. □

VITA AUCTORIS

NAME:	Lei Shen
PLACE OF BIRTH:	Ningbo, Zhejiang, China
YEAR OF BIRTH:	1993
EDUCATION:	Cambridge International Centre of Shanghai Normal University, Shanghai, China, 2008-2011
	University of Waterloo, B.Sc., Waterloo, ON, 2011-2015
	University of Windsor, M.Sc., Windsor, ON, 2016-2018