7-30-2018

Stein-rules and Testing in Generalized Mean Reverting Processes with Multiple Change-points

Kang Fu

University of Windsor

Stein-rules and Testing in Generalized Mean Reverting Processes with Multiple Change-points

by

Kang Fu

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

Stein-rules and Testing in Generalized Mean Reverting

Processes with Multiple Change-points

by

Kang Fu

APPROVED BY:

D. Li
Department of Economics

M. Hlynka
Department of Mathematics and Statistics

S. Nkurunziza, Advisor
Department of Mathematics and Statistics

July 26, 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: some parts of Chapters 3, 4, 5, 6, 7 and 8 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*
some parts of Chapters 3, 4, 5, 6, 7 and 8	Nkurunziza, S., and Fu, K., (2018). Improved inference in generalized mean-reverting processes with multiple change-points. *Electronic Journal of Statistics (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 thesis, we consider inference problems about the drift parameter vector in generalized mean reverting processes with multiple and unknown change-points. In particular, we study the case where the parameter may satisfy uncertain restrictions. As compared to the results in the literature, we generalize some findings in five ways. First, we consider a statistical model which incorporates uncertain prior information and the uncertain restriction includes as a special case the nonexistence of the change-points. Second, we derive the unrestricted estimator (UE) and the restricted estimator (RE), and we study their asymptotic properties. Specifically, in the context of a known number of change-points, we derive the joint asymptotic normality of the UE and the RE, under the set of local alternative hypotheses. Third, we derive a test for testing the hypothesized restriction and we derive its asymptotic local power. We also prove that the proposed test is consistent. Fourth, we construct a class of shrinkage type estimators (SEs) which includes as special cases the UE, RE, and classical SEs. Fifth, we derive the relative risk dominance of the proposed estimators. More precisely, we prove that the SEs dominate the UE. The novelty of the derived results consists in the fact that the dimensions of the proposed estimators are random variables. Finally, we present some simulation results which corroborate the established theoretical findings.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to whom supported and assisted me to finish this thesis during my master's study.

First, I would like to express my sincerest gratitude to my supervisor Dr. Sévérien Nkurunziza. He guided me patiently as I encountered problems. His serious attitude and precision taught me how to be an excellent scholar. His humor released the stress of work. As I was ill this winter, his encouragement and concern helped psychologically.

I would also express my gratitude to the department reader, Dr. Myron Hlynka, and the external reader, Dr. Dingding Li. I am thankful for their patient reading. Their valuable advice made a huge contribution in accomplishing this thesis. Also, I would like to thank to all the staff in the department. They helped me directly and indirectly during my master's study.

I would like to thank to my parents. They support me financially and psychologically in my study and daily life. I would like also to thank my girlfriend, my cousin and my aunt who accompanied me in my two years' graduate study. I will never forget their encouragement and assistence. Finally, I would like to thank to my co-worker and good friend, Lei Shen, who could always help me to find the key point to solve any problems.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction and contributions

Nowadays, Ornstein-Uhlenbeck (O-U) processes are applied in different fields, such as physical sciences (Lansky and Sacerdote (2001)) and biology (Rohlfs *et al.* (2010)). The O-U process is also called the mean reverting process since the mean reverting level is the component which has large effect on it. For the classical O-U processes, the mean reverting level is constant. However, the classical O-U processes do not fit well to data whose mean reverting level may change with the time. This is particularly the case for some phenomena which heavily depend on factors which change with the time. For instance, government policy is one factor which affects the stock price. Thus, if the government policies are changed in different time periods, the mean reverting level of the stock price may change. As a result, the stock price is changed. To overcome such a problem, Dehling *et al.* (2010) proposed a stochastic process which has a time-dependent periodic mean reverting function. This is the so called generalized Ornstein-Uhlenbeck process. Further, to take into account some unconventional shocks of the process, Dehling *et al.* (2014) and Nkurunziza and Zhang (2018) considered inference problems in generalized O-U processes. To give a closely related reference, we quote Chen *et al.* (2017) who proposed a method for detecting multiple change-points in generalized O-U process. In this thesis, we study the inference problem in generalized O-U processes with multiple unknown change-points where the drift parameter is suspected to satisfy some restrictions. We also revisit the conditions for the main results in Chen *et al.* (2017) to hold. In particular, we show that the results in Chen *et al.* (2017) hold without their Assumption 2. Nevertheless, the authors of the quoted paper omitted an important condition about the initial value of the SDE for their main results to hold. In the subsequent section, we highlight the main contribution of this thesis.

1.1 Main contributions

In this section, we present the main contributions of this thesis. Briefly, we generalize the methods in Chen *et al.* (2017) as well as that in Nkurunziza and Zhang (2018). In particular, the proposed method generalizes the work of Chen *et al.* (2017) in five ways.

1. We consider a statistical model which incorporates the uncertain prior knowledge.

2. We derive the unrestricted estimator (UE) and the restricted estimator (RE) for the drift parameter.

3. For a known number of change-points, we derive the joint asymptotic normality of the UE and the RE under the set of local alternative hypotheses.

4. We derive a test for testing the hypothesized restriction and we derive its asymptotic power. The proposed test is also useful for testing the absence of change

points.

5. We construct a class of shrinkage estimators (SEs) which includes as a special case the UE, the RE and classical SEs. The proposed SEs are expected to be robust with respect to the restriction.

The novelty of the derived results consists in the fact that the dimensions of the proposed estimators are random variables. To overcome the difficulty due to the randomness of the dimension, we establish two asymptotic results which are of interest on their own.

1.2 Organization of the thesis

The remainder of this thesis is organized as follows. In Chapter 2, we introduce the statistical model and assumptions. In Chapter 3, we study the joint asymptotic normality of the UE and the RE in the case of known change-points. In Chapter 4, we study the joint asymptotic normality of the UE and the RE in the case of unknown change-points. In Chapter 5, we present inference methods in the case of unknown change-points and unknown number of change-points. In Chapter 6, we construct a class of SEs and test the restriction. In Chapter 7, we compare the relative performance between estimators. In Chapter 8, we present some simulation results, and Chapter 9 gives some concluding remarks. Finally, for the convenience of the reader, some technical results and proofs are given in the Appendix A and B.

Chapter 2

Statistical model and regularity conditions

In this section, we present the statistical model of the generalized Ornstein-Uhlenbeck process which is mainly studied in this thesis. Two assumptions are presented. Under these assumptions, we derive the log-likelihood function. In Chapter 3 and 4, we use this log-likelihood function to derive the *Maximum Likelihood Estimator* (MLE) without restriction and with restriction.

The inference problem studied in this thesis was mainly inspired by the work in Chen *et al.* (2017) where the authors proposed a method for detecting multiple change-points in generalized O-U processes. To give some other references about inference problem in generalized O-U processes, we quote Dehling *et al.* (2010), Dehling *et al.* (2014), Nkurunziza and Zhang (2018). To introduce some notation, let $\{W_t; t \geq 0\}$ be a one-dimensional standard Brownian motion (Wiener process) defined on some probability space $(\Omega, \mathfrak{F}, P)$ and let $\sigma > 0$. The change points are denoted by $\tau_j = \phi_j T$, where $j = 1, \ldots, m$ and $0 < \phi_1 < \cdots < \phi_m < 1$. We let

$\tau_0 = 0$ and $\tau_{m+1} = T$ to simplify the notation. Let $\top$ denote the transpose of a matrix, let $\theta = (\theta_1^\top, \ldots, \theta_{m+1}^\top)^\top$ with $\theta_j = (\mu_{1,j}, \ldots, \mu_{p,j}, a_j)^\top$ for $\tau_{j-1} < t \leq \tau_j$ where, for $j = 1, \ldots, m+1$ and $k = 1, \ldots, p$, $\mu_{k,j}$ is real value and $a_j > 0$. Let $\varphi(t) = (\varphi_1(t), \varphi_2(t), \ldots, \varphi_p(t))$. Let $\mathbb{I}_{\{.\}}$ be an indicator function, and let I_p be the p-dimensional identity matrix. As in Chen *et al.* (2017), we consider the stochastic differential equation (SDE) given by

$$dX_t = S(\theta, t, X_t)dt + \sigma dW_t, \quad 0 \leq t \leq T \tag{2.1}$$

where the drift coefficient, $S(\theta, t, X_t)$, is as follows

$$S(\theta, t, X_t) = \sum_{j=1}^{m+1} \left(\sum_{k=1}^{p} \mu_{k,j}\varphi_k(t) - a_j X_t \right) \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}}. \tag{2.2}$$

In the SDE given in (2.1) and (2.2), m represents the number of unknown change-points ($m \geq 1$), while $\tau_1, \tau_2, \ldots, \tau_m$ are the locations of change-points. In this thesis, the parameter of interest is θ while m, τ_1, τ_2, $\ldots$, τ_m are the unknown nuisance parameters.

Sometimes, there exists a prior knowledge, called *prior information*, so that we might use both the *non-sample information* and the *sample information* to estimate the parameters. In this thesis, the prior information is considered as a form of a linear constraint on θ for a given m, τ_1, τ_2, $\ldots$, τ_m. Then, when τ_1, τ_2, $\ldots$, τ_m and m are known, the maximum likelihood estimator, which is derived based on linear restrictions, is called the *Restricted Maximum Likelihood Estimator* (RMLE). In particular, we consider the scenario where the target parameter may satisfy the restriction

$$H_0 : B\theta = r \tag{2.3}$$

where B is a known $q \times (m+1)(p+1)$ full rank matrix with $q < (m+1)(p+1)$, r is a known q-column vector, and θ is the vector of parameters. This restriction leads to the hypothesis testing problem

$$H_0 : B\theta = r \quad \text{vs} \quad H_1 : B\theta \neq r. \tag{2.4}$$

Particularly, if we choose $r = 0$ and

$$B = \begin{pmatrix} I_{p+1} & -I_{p+1} & 0 & \dots & 0 & 0 \\ 0 & I_{p+1} & -I_{p+1} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I_{p+1} & -I_{p+1} \end{pmatrix} = B_0,$$

the restriction in (2.3) corresponds to the case where there are no change points. Thus, the testing problem in (2.4) includes as a special case testing the absence of change points.

Assumption 1. *The distribution 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 of $\{W_t : t \geq 0\}$ and* $\mathrm{E}[|X_0|^d] < \infty$*, for some $d \geq 2$.*

Assumption 2. *For any $T > 0$, the basis functions $\{\varphi_k(t), k = 1, ..., p\}$ are Riemann-integrable on $[0, T]$ and satisfy*
(1) Periodicity: $\varphi(t + v) = \varphi(t)$ where v is the period in the data.
(2) Orthogonality in $L^2([0, v]\frac{1}{v}d\lambda)$: $\int_0^v \varphi(t)\varphi^\top(t)dt = vI_P$.

Remark 1. *Assumption 2 corresponds to a similar assumption in Chen et al. (2017). Assumption 1 is not explicitly given in Chen et al. (2017), but their results require the Assumption 1 to hold. For example, if* $\mathrm{E}[|X_0|^2] = \infty$*, the relation (3.8) in Chen et al. (2017) does not hold. Further, if the distribution of X_0 depends on θ, by Theorem 1.12 in Kutoyants (2004), the likelihood function given in Section 3.1 of Chen et al. (2017, see p. 2204) does not hold.*

Since, for $k = 1, ..., p$, $\varphi_k(t)$ is bounded on $[0, T]$ and is periodic, this implies that $\varphi_k(t)$ is bounded on R_+. Without loss of generality, as in Chen *et al.* (2017), we assume that $v = 1$.

The following proposition shows that the SDE (2.5) admits a strong and unique solution.

Proposition 2.1. *The SDE*

$$dX_t = \sum_{j=1}^{m+1} \left(\sum_{k=1}^{p} \mu_{k,j}\varphi_k(t) - a_j X_t \right) \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} dt + \sigma dW_t \tag{2.5}$$

$0 \leq t \leq T$ *admits a strong and unique solution that is* L^2*-bounded on* $[0, T]$*, i.e.,* $\sup_{0 \leq t \leq T} \mathrm{E}[X_t^2] < \infty$.

Proof. It suffices to check whether the coefficients of SDE satisfy both space-variable Lipschitz condition and the spatial growth condition. For more details, see the proof of Proposition 3.2 in Chen *et al.* (2017). □

Lemma 2.1. *The solution of SDE in (2.5) has the explicit representation*

$$X_t = \sum_{j=1}^{m+1} X_j(t)\mathbb{I}_{(\tau_{j-1} < t \leq \tau_j)}, \ X_j(t) = e^{-a_j t} X_0 + h_j(t) + z_j(t), \tag{2.6}$$

where

$$h_j(t) = e^{-a_j t} \sum_{k=1}^{p} \mu_{k,j} \int_0^t e^{a_j s} \varphi_k(s) ds, \ z_j(t) = \sigma e^{-a_j t} \int_0^t e^{a_j s} dW_s. \tag{2.7}$$

Further, $\sup_{t \geq 0} \mathrm{E}[|X_t|^2] < \infty$.

The proof is given in Appendix B.

Note that process $\{X_t\}_{\{\tau_{j-1} < t \leq \tau_j\}}$ is not stationary. Because of that we cannot apply the ergodic theorem for stationary processes. However, we can introduce some stationary stochastic processes associated to $\{X_t : t \geq 0\}$.

We define, for $\tau_{j-1} < t \leq \tau_j$, $j = 1, ..., m+1$,

$$\tilde{X}_j(t) = \tilde{h}_j(t) + \tilde{z}_j(t) \tag{2.8}$$

where

$$\tilde{h}_j(t) = e^{-a_j t} \sum_{k=1}^{p} \mu_{k,j} \int_{-\infty}^{t} e^{a_j s} \varphi_k(s) ds, \ \tilde{z}_j(t) = \sigma e^{-a_j t} \int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s, \tag{2.9}$$

where $\{\tilde{B}_s\}_{s\in\mathbb{R}}$ denotes a bilateral Brownian motion. i.e.

$$\tilde{B}_s = B_s \mathbb{I}_{\mathbb{R}+}(s) + \bar{B}_{-s} \mathbb{I}_{\mathbb{R}-}(s)$$

with $\{B_s\}_{s\geq 0}$ and $\{\bar{B}_{-s}\}_{s\geq 0}$ being two independent standard Brownian motions.

Let Σ_j be a $(p+1) \times (p+1)$ non-random matrix as, for $j = 1, ..., m+1$,

$$\Sigma_j = \begin{bmatrix} I_P & \Lambda_j \\ \Lambda_j^T & \omega_j \end{bmatrix} \tag{2.10}$$

where

$$\Lambda_j = -\int_0^1 \tilde{h}_j(t)\varphi(t)dt, \ \omega_j = \int_0^1 \tilde{h}_j^2(t)dt + \frac{\sigma^2}{2a_j},$$

with the function $\tilde{h}_j(t) \ : \ [0, \infty] \to \mathbb{R}$

$$\tilde{h}_j(t) = e^{-a_j t} \sum_{k=1}^{p} \mu_{k,j} \int_{-\infty}^{t} e^{a_j s} \varphi_k(s) ds.$$

Proposition 2.2. *Suppose that Assumptions 1-2 hold, then, for* $k = 0, 1, \ldots,$

(1) $\mathrm{E}[\tilde{X}_j(t+k)] = \tilde{h}_j(t)$;

(2) $Cov(\tilde{X}_j(t), \tilde{X}_j(t+k)) = e^{-a_j k} \dfrac{\sigma^2}{2a_j}$.

The proof is given in Appendix B. From Proposition 2.2, we derive the following lemma which shows that our introduced processes are stationary and ergodic.

Lemma 2.2. *For $t \in [0,1]$, for $j = 1, 2, \ldots, m$, the sequence of random variables $\{\tilde{X}_j(k+t)\}_{k\in\mathbb{N}_0}$ is stationary and ergodic.*

The proof is given in Appendix B.

Remark 2. *From Proposition 2.1, we have* $\mathrm{P}\left(\int_0^T S^2(\theta, t, X_t)dt < \infty\right) = 1$*, for all $0 < T < \infty$ and elements θ_j of θ involved in $S(\theta, t, X_t)$ given by equation (2.1). In passing, it should be noticed that this condition is given as a required assumption in Chen et al. (2017, Assumption 2). Thus, here we show that the results in Chen et al. (2017) hold without their Assumption 2.*

This condition is useful in deriving the likelihood function of the SDE in (2.1).

Proposition 2.3. *If Assumption 1-2 hold, then the log likelihood function is*

$$\log L(\theta, X_t) = \frac{1}{\sigma^2}\sum_{j=1}^{m+1}\theta_j^\top \tilde{r}_{(\tau_{j-1},\tau_j)} - \frac{1}{2\sigma^2}\sum_{j=1}^{m+1}\theta_j^\top Q_{(\tau_{j-1},\tau_j)}\theta_j \tag{2.11}$$

where

$$\tilde{r}_{(\tau_{j-1},\tau_j)} = \left(\int_{\tau_{j-1}}^{\tau_j}\varphi_1(t)dX_t, ..., \int_{\tau_{j-1}}^{\tau_j}\varphi_p(t)dX_t, -\int_{\tau_{j-1}}^{\tau_j}X_t dX_t\right)^\top \tag{2.12}$$

and

$$Q_{(\tau_{j-1},\tau_j)} = \begin{bmatrix} \int_{\tau_{j-1}}^{\tau_j}\varphi_1^2(t)dt & \ldots & \int_{\tau_{j-1}}^{\tau_j}\varphi_1(t)\varphi_p(t)dt & -\int_{\tau_{j-1}}^{\tau_j}\varphi_1 X_t dt \\ \vdots & \vdots & \vdots & \vdots \\ -\int_{\tau_{j-1}}^{\tau_j}\varphi_1 X_t dt & \ldots & -\int_{\tau_{j-1}}^{\tau_j}\varphi_p X_t dt & \int_{\tau_{j-1}}^{\tau_j}X_t^2 dt \end{bmatrix}. \tag{2.13}$$

The proof is given in Appendix B. The following proposition shows that the matrix $Q(\phi, m)$ is positive definite. By using Proposition 2.3, we derive in the next section the UMLE for θ. To introduce some notation, let $\phi = (\phi_1, \ldots, \phi_m)^\top$. Let

$$\tilde{R}(\phi, m) = (\tilde{r}_{(0,\tau_1)}, ..., \tilde{r}_{(\tau_m,T)})^\top, \; Q(\phi, m) = \begin{bmatrix} Q_{(0,\tau_1)} & 0 & \ldots & 0 \\ 0 & Q_{(\tau_1,\tau_2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & Q_{(\tau_m,T)} \end{bmatrix}. \quad (2.14)$$

Proposition 2.4. *Suppose that Assumption 1-2 holds. Then, if* $T \geq \dfrac{1}{(\phi_j - \phi_{j-1})}$, $Q_{(\tau_{j-1},\tau_j)}$ *is positive definite for* $j = 1, \ldots, m+1$*. Further, if* $T \geq \dfrac{1}{\min\limits_{1\leq j\leq m+1}(\phi_j - \phi_{j-1})}$, $Q(\phi, m)$ *is a positive definite matrix.*

The proof of this proposition is similar to that of Proposition 3.2 of Shen (2018, p. 32). Further, for the convenience of the reader, we also give the proof in Appendix B.

Chapter 3

Estimation in the case of known change points

3.1 The unrestricted estimator

In this chapter, we assume that the change point $\tau_j = \phi_j T$ is known, $j = 1, ..., m$. Then, some preliminary results related to the *Maximum Likelihood Estimator* (MLE) of drift parameter are developed. In this chapter, all the estimation problems are studied on the basis of the sample information. Hence, the derived MLE is called the *Unrestricted Maximum Likelihood estimator* (UMLE). We also derive the asymptotic normality of the UMLE.

3.1.1 The UMLE $\hat{\theta}(\phi, m)$

The UMLE $\hat{\theta}(\phi, m)$ is derived based on Proposition 2.3 along with the positive definiteness of $Q(\phi, m)$.

By relation (2.11) in Proposition 2.3, we have

$$\log L(\theta, X_t) = \frac{1}{\sigma^2}\theta^\top \tilde{R}(\phi, m) - \frac{1}{2\sigma^2}\theta^\top Q(\phi, m)\theta. \tag{3.1}$$

Next, from Proposition 2.4, we derive the UMLE which is given in the following lemma.

Lemma 3.1. *Suppose that Assumptions 1-2 hold, and let $\tilde{R}(\phi, m)$ and $Q(\phi, m)$ be as defined in (2.14). Then the* UMLE *of θ is*

$$\hat{\theta}(\phi, m) = Q^{-1}(\phi, m)\tilde{R}(\phi, m).$$

.

The proof is given in Appendix B. Let

$$R(\phi, m) = (r_{(0,\tau_1)}, ..., r_{(\tau_m,T)})^\top \tag{3.2}$$

and

$$r(a, b) = \left(\int_a^b \varphi_1(t)dW_t, ..., \int_a^b \varphi_p(t)dW_t, -\int_a^b X_t dW_t \right)^\top$$

for $0 \leq a < b \leq T$, and $Q(\phi, m)$ defined in (2.14). From Lemma 3.1, we derive the following proposition which is useful in deriving the asymptotic normality of the UMLE.

Proposition 3.1. *Suppose that Assumptions 1-2 hold. The* UMLE *of θ can be rewritten as*

$$\hat{\theta}(\phi, m) = \theta + \sigma Q^{-1}(\phi, m)R(\phi, m). \tag{3.3}$$

The proof is given in Appendix B.

By Lemma 3.1, we can rewrite the UMLE of the drift parameter

$$\hat{\theta} = Q^{-1}(\phi, m)\tilde{R}(\phi, m) = TQ^{-1}(\phi, m)\frac{1}{T}\tilde{R}(\phi, m) = (\frac{1}{T}Q(\phi, m))^{-1}\frac{1}{T}\tilde{R}(\phi, m). \tag{3.4}$$

Thus, in order to study the convergence of $\hat{\theta}(\phi, m)$, we study first the convergence of $(\frac{1}{T}Q(\phi, m))^{-1}$. To introduce some notation, let $\xrightarrow[T\to\infty]{P}$, $\xrightarrow[T\to\infty]{d}$, $\xrightarrow[T\to\infty]{L^p}$, $\xrightarrow[T\to\infty]{a.s.}$ denote convergence in probability, in distribution , in L^p-space and almost surely respectively, as T tends to infinity.

Proposition 3.2. *If Assumption 2 holds, then, for* $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$, $\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \varphi(t)\varphi^\top(t)dt \xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})I_P$.

The proof is given in Appendix B.

Proposition 3.3. *Suppose that Assumptions 1-2 hold. Then,* $0 \leq \phi_{j-1} < \phi_j \leq 1$ *where* $j = 1, ..., m+1$,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t\varphi(t)dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\int_0^1 \tilde{h}_j(t)\varphi(t)dt.$$

The proof is given in Appendix B.

Proposition 3.4. *Suppose that Assumptions 1-2 hold. Then,* $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\left[\int_0^1 (\tilde{h}_j(t))^2 dt + \frac{\sigma^2}{2a_j}\right].$$

The proof is given in Appendix B. Let

$$\Sigma = \begin{bmatrix} \phi_1\Sigma_1 & 0 & \ldots & 0 \\ 0 & (\phi_2-\phi_1)\Sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & (1-\phi_m)\Sigma_{m+1} \end{bmatrix}.$$

Proposition 3.5. *Suppose that Assumption 2 holds. Then,* Σ_j *is a positive definite matrix for* $j = 1, ..., m+1$. *Further,* Σ *is a positive definite matrix.*

The proof is given in Appendix B. By combining Propositions 3.2-3.4, we derive the following proposition.

Proposition 3.6. *If Assumptions 1-2 hold, then, for* $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$, $TQ^{-1}_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{P} \frac{1}{\phi_j - \phi_{j-1}}(\Sigma_j)^{-1}$. *Further,* $TQ^{-1}(\phi, m) \xrightarrow[T\to\infty]{P} \Sigma^{-1}$.

The proof is given in Appendix B. It should be noticed that in Nkurunziza and Fu (2018), we prove a stronger result. Indeed, we prove that the above convergences hold almost surely.

3.1.2 Asymptotic normality of the UMLE $\hat{\theta}(\phi, m)$

In this subsection, we study the convergence of $\frac{1}{\sqrt{T}}R(\phi, m)$. Then, based on that convergence, we establish the asymptotic normality of the UMLE $\hat{\theta}(\phi, m)$.
The following proposition gives the limiting distribution of $\frac{1}{\sqrt{T}}R(\phi, m)$.

Proposition 3.7. *Suppose that Assumptions 1-2 hold. Then,*
$\frac{1}{\sqrt{T}}R(\phi, m) \xrightarrow[T\to\infty]{d} r^* \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma)$.

The proof is given in Appendix B. From Proposition 3.7, we derive below a proposition which gives the asymptotic normality of the UMLE. To simplify some mathematical expressions, let $\rho_T(\phi, m) = \sqrt{T}(\hat{\theta}(\phi, m) - \theta)$.

Proposition 3.8. *Suppose that Assumptions 1-2 hold. Then, the* UMLE $\hat{\theta}(\phi, m)$ *is asymptotically normal, i.e.,* $\rho_T(\phi, m) \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2\Sigma^{-1})$.

The proof is given in Appendix B.

3.2 The restricted estimator

In this section, we derive the restricted maximum likelihood estimator (RMLE).

Proposition 3.9. *Suppose that Assumptions 1-2 hold along with (2.3) and let* $G = Q^{-1}(\phi, m)B^\top(BQ^{-1}(\phi, m)B^\top)^{-1}$. *Then, the RMLE of* θ *is*

$$\tilde{\theta}(\phi, m) = \hat{\theta}(\phi, m) - G(B\hat{\theta}(\phi, m) - r). \tag{3.5}$$

The proof is given in Appendix B.

3.2.1 Asymptotic normality of the RMLE $\tilde{\theta}(\phi, m)$

In this subsection, we study the asymptotic property of the RMLE $\tilde{\theta}(\phi, m)$ based on the asymptotic normality of the UMLE $\hat{\theta}(\phi, m)$. Based on Proposition 3.9, we have

$$\begin{aligned}\sqrt{T}(\tilde{\theta}(\phi, m) - \theta) &= \sqrt{T}[Gr + (I_{(m+1)(p+1)} - GB)\hat{\theta}(\phi, m) - \theta]\\ &= \sqrt{T}(Gr - \theta) + \sqrt{T}(I_{(m+1)(p+1)} - GB)\hat{\theta}(\phi, m).\end{aligned}$$

This gives

$$\sqrt{T}(\tilde{\theta}(\phi, m) - \theta) = (I_{(m+1)(p+1)} - GB)\sqrt{T}(\hat{\theta}(\phi, m) - \theta) - \sqrt{T}G(B\theta - r).$$

Now, we define $\zeta_T(\phi, m) = \sqrt{T}(\tilde{\theta}(\phi, m) - \theta)$. We have

$$\zeta_T(\phi, m) = (I_{(m+1)(p+1)} - GB)\sqrt{T}(\hat{\theta}(\phi, m) - \theta) - \sqrt{T}G(B\theta - r). \tag{3.6}$$

Consider a continuous function $g(X) = XB^\top(BXB^\top)^{-1}$ where X is a positive definite matrix. We have

$$g(TQ^{-1}(\phi, m)) = G = TQ^{-1}(\phi, m)B^\top(BTQ^{-1}(\phi, m)B^\top)^{-1}.$$

By combining Proposition 3.6, and the continuous mapping theorem,

$$G \xrightarrow[T\to\infty]{P} G^* = \Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1},$$

and

$$I_{(m+1)(p+1)} - GB \xrightarrow[T\to\infty]{P} I_{(m+1)(p+1)} - G^*B. \tag{3.7}$$

To study the asymptotic normality of the RMLE $\tilde{\theta}(\phi, m)$, we consider the following set of local alternative restrictions,

$$H_{a,T} : B\theta - r = \frac{r_0}{\sqrt{T}},\ T > 0 \tag{3.8}$$

where r_0 is a fixed q-column vector. Then,

$$\sqrt{T}G(B\theta - r) = \sqrt{T}G\frac{r_0}{\sqrt{T}} = Gr_0 \xrightarrow[T\to\infty]{P} G^*r_0. \tag{3.9}$$

Proposition 3.10. *Suppose that Assumptions 1-2 hold along with the set of local alternatives in (3.8). Then* RMLE $\tilde{\theta}(\phi, m)$ *given in (3.5) is asymptotically normal, i.e.,* $\zeta_T(\phi, m) \xrightarrow[T\to\infty]{d} \zeta \sim \mathcal{N}_{(m+1)(p+1)}\big(- G^*r_0, \sigma^2(\Sigma^{-1} - G^*B\Sigma^{-1})\big)$.

The proof is given in Appendix B.

3.3 Joint asymptotic normality of $\hat{\theta}(\phi, m)$ and $\tilde{\theta}(\phi, m)$

In this section, we establish the joint asymptotic normality of UMLE $\hat{\theta}(\phi, m)$ and RMLE $\tilde{\theta}(\phi, m)$. This property is the foundation of developing a test for the testing problem in (2.4) as well as its power. The established result is also useful in constructing shrinkage estimators and their asymptotic efficiency. The following proposition presents the asymptotic property of
$(\rho_T^\top(\phi, m), \zeta_T^\top(\phi, m))^\top = \sqrt{T}\left((\hat{\theta}(\phi, m) - \theta)^\top, (\tilde{\theta}(\phi, m) - \theta)^\top\right)^\top$.

Proposition 3.11. *Suppose that Assumption 1-2 hold along with the set of local* *alternatives in (3.8). Then,* $(\rho_T^\top(\phi,m),\zeta_T^\top(\phi,m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top,\zeta^\top)^\top$, *where*

$$\begin{pmatrix}\rho\\ \zeta\end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix}0\\ -G^*r_0\end{pmatrix}, \sigma^2\begin{pmatrix}\Sigma^{-1} & \Sigma^{-1}-G^*B\Sigma^{-1}\\ \Sigma^{-1}-G^*B\Sigma^{-1} & \Sigma^{-1}-G^*B\Sigma^{-1}\end{pmatrix}\right).$$

Proof. We observe that

$$\begin{pmatrix}\sqrt{T}(\hat\theta(\phi,m)-\theta)\\ \sqrt{T}(\tilde\theta(\phi,m)-\theta)\end{pmatrix} = \begin{pmatrix}\sqrt{T}(\hat\theta(\phi,m)-\theta)\\ (I_{(m+1)(p+1)}-GB)\sqrt{T}(\hat\theta(\phi,m)-\theta)-\sqrt{T}G(B\theta-r)\end{pmatrix}$$
$$= \begin{pmatrix}I_{(m+1)(p+1)}\\ I_{(m+1)(p+1)}-GB\end{pmatrix}\sqrt{T}(\hat\theta(\phi,m)-\theta)+\begin{pmatrix}0\\ -Gr_0\end{pmatrix}.$$

From (3.7), we get

$$\begin{pmatrix}I_{(m+1)(p+1)}\\ I_{(m+1)(p+1)}-GB\end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix}I_{(m+1)(p+1)}\\ I_{(m+1)(p+1)}-G^*B\end{pmatrix} \tag{3.10}$$

where all the elements in (3.10) are non-random. Similarly, by (3.9), we have

$$\begin{pmatrix}0\\ -Gr_0\end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix}0\\ -G^*r_0\end{pmatrix}. \tag{3.11}$$

Then, by combining Proposition 3.8 and the relations (3.10) and (3.11) along with Slutsky's Theorem,

$$\begin{pmatrix}\sqrt{T}(\hat\theta(\phi,m)-\theta)\\ \sqrt{T}(\tilde\theta(\phi,m)-\theta)\end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix}I_{(m+1)(p+1)}\\ I_{(m+1)(p+1)}-G^*B\end{pmatrix}\rho+\begin{pmatrix}0\\ -G^*r_0\end{pmatrix}=\begin{pmatrix}\rho\\ \zeta\end{pmatrix}.$$

Then, by Proposition A.2 in Appendix A, $(\rho_T^\top(\phi,m),\zeta_T^\top(\phi,m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top,\zeta^\top)^\top$ with

$$\begin{pmatrix}\rho\\ \zeta\end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix}0\\ -G^*r_0\end{pmatrix}, \sigma^2\begin{pmatrix}I_{(m+1)(p+1)}\\ I_{(m+1)(p+1)}-G^*B\end{pmatrix}\Sigma^{-1}\begin{pmatrix}I_{(m+1)(p+1)}\\ I_{(m+1)(p+1)}-G^*B\end{pmatrix}^\top\right).$$

Note that

$$\sigma^2 \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^* B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^* B \end{pmatrix}^\top$$
$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - \Sigma^{-1} B^\top G^{*\top} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - \Sigma^{-1} B^\top G^{*\top} - G^* B \Sigma^{-1} + G^* B \Sigma^{-1} B^\top G^{*\top} \end{pmatrix}.$$

By the proof of Proposition 3.10, we know

$$G^* B \Sigma^{-1} B^\top G^{*\top} = \Sigma^{-1} B^\top G^{*\top}, \tag{3.12}$$

and, since $G^* = \Sigma^{-1} B^\top (B \Sigma^{-1} B^\top)^{-1}$,

$$\Sigma^{-1} B^\top G^{*\top} = \Sigma^{-1} B^\top (B \Sigma^{-1} B^\top)^{-1} B \Sigma^{-1} = G^* B \Sigma^{-1}. \tag{3.13}$$

Therefore, by (3.12) and (3.13),

$$\sigma^2 \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^* B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^* B \end{pmatrix}^\top$$
$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix}.$$

Finally, we have $(\rho_T^\top(\phi, m), \zeta_T^\top(\phi, m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top, \zeta^\top)^\top$, where

$$\begin{pmatrix} \rho \\ \zeta \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} 0 \\ -G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix} \right).$$

This completes the proof. □

Now, we define $\xi_T(\phi, m) = \sqrt{T}(\hat{\theta}(\phi, m) - \tilde{\theta}(\phi, m))$. Next, we study the asymptotic distribution of $(\rho_T^\top(\phi, m), \xi_T^\top(\phi, m))^\top = \sqrt{T}\left((\hat{\theta}(\phi, m) - \theta)^\top, (\hat{\theta}(\phi, m) - \tilde{\theta}(\phi, m))^\top\right)^\top$.

Proposition 3.12. *Suppose that Assumptions 1-2 hold along with the set of local alternatives in (3.8). Then,* $(\rho_T^\top(\phi,m), \xi_T^\top(\phi,m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top, \xi^\top)^\top$*, where*

$$\begin{pmatrix} \rho \\ \xi \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} 0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix}\right).$$

Proof. We have

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\phi,m)-\theta) \\ \sqrt{T}(\hat{\theta}(\phi,m)-\tilde{\theta}(\phi,m)) \end{pmatrix} = \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\phi,m)-\theta) \\ \sqrt{T}(\tilde{\theta}(\phi,m)-\theta) \end{pmatrix}.$$

We know

$$\begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix},$$

and, by Proposition 3.11 and Slutsky's Theorem, we have

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\phi,m)-\theta) \\ \sqrt{T}(\hat{\theta}(\phi,m)-\tilde{\theta}(\phi,m)) \end{pmatrix} = \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\phi,m)-\theta) \\ \sqrt{T}(\tilde{\theta}(\phi,m)-\theta) \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{d} \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \rho \\ \zeta \end{pmatrix} = \begin{pmatrix} \rho \\ \xi \end{pmatrix}.$$

Note that

$$\begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} 0 \\ -G^* r_0 \end{pmatrix} = \begin{pmatrix} 0 \\ G^* r_0 \end{pmatrix},$$

and

$$\begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix}$$

$$\times \begin{pmatrix} I_{(m+1)(p+1)} & I_{(m+1)(p+1)} \\ 0 & -I_{(m+1)(p+1)} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & 0 \end{pmatrix} \begin{pmatrix} I_{(m+1)(p+1)} & I_{(m+1)(p+1)} \\ 0 & -I_{(m+1)(p+1)} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix}.$$

Therefore, by Proposition A.2 in Appendix A,

$$\begin{pmatrix} \rho_T(\phi, m) \\ \xi_T(\phi, m) \end{pmatrix} \xrightarrow[T \to \infty]{d} \begin{pmatrix} \rho \\ \xi \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} 0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix} \right),$$

this completes the proof. □

From Proposition 3.12, we have following corollary.

Corollary 3.1. *Suppose that Assumptions 1-2 hold along with the set of local alternatives in (3.8). Then,* $\xi_T(\phi, m) \xrightarrow[T \to \infty]{d} \xi \sim \mathcal{N}_{(m+1)(p+1)}(G^* r_0, \sigma^2 G^* B \Sigma^{-1})$.

The proof follows directly from Proposition 3.12. Further, we study the asymptotic property of $(\zeta_T^\top(\phi, m), \xi_T^\top(\phi, m))^\top = \sqrt{T} \left((\tilde{\theta}(\phi, m) - \theta)^\top, (\hat{\theta}(\phi, m) - \tilde{\theta}(\phi, m))^\top \right)^\top$.

Proposition 3.13. *Suppose that Assumptions 1-2 hold along with the set of local alternatives in (3.8). Then,* $(\zeta_T^\top(\phi, m), \xi_T^\top(\phi, m))^\top \xrightarrow[T \to \infty]{d} (\zeta^\top, \xi^\top)^\top$*, where*

$$\begin{pmatrix} \zeta \\ \xi \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} -G^* r_0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & 0 \\ 0 & G^* B \Sigma^{-1} \end{pmatrix} \right).$$

Proof. We observe that

$$\begin{pmatrix} \sqrt{T}(\tilde{\theta}(\phi,m)-\theta) \\ \sqrt{T}(\hat{\theta}(\phi,m)-\tilde{\theta}(\phi,m)) \end{pmatrix} = \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\phi,m)-\theta) \\ \sqrt{T}(\tilde{\theta}(\phi,m)-\theta) \end{pmatrix}. \tag{3.14}$$

Further,

$$\begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix},$$

and

$$\begin{aligned} &\begin{pmatrix} \rho_T(\phi,m) \\ \zeta_T(\phi,m) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} \rho \\ \zeta \end{pmatrix} \\ &\sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} 0 \\ -G^*r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1}-G^*B\Sigma^{-1} \\ \Sigma^{-1}-G^*B\Sigma^{-1} & \Sigma^{-1}-G^*B\Sigma^{-1} \end{pmatrix} \right). \end{aligned} \tag{3.15}$$

Then, by combining (3.14) and (3.15) and Slutsky's Theorem, we get

$$\begin{pmatrix} \zeta_T(\phi,m) \\ \xi_T(\phi,m) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \rho \\ \zeta \end{pmatrix} = \begin{pmatrix} \zeta \\ \xi \end{pmatrix}.$$

Note that

$$\begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} 0 \\ -G^*r_0 \end{pmatrix} = \begin{pmatrix} -G^*r_0 \\ G^*r_0 \end{pmatrix},$$

and

$$\begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix}$$
$$\times \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix}$$
$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix}$$
$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & 0 \\ 0 & G^* B \Sigma^{-1} \end{pmatrix}.$$

Therefore, by Proposition A.2 in Appendix A, $(\zeta_T^\top(\phi, m), \xi_T^\top(\phi, m))^\top \xrightarrow[T\to\infty]{d} (\zeta^\top, \xi^\top)^\top$ with

$$\begin{pmatrix} \zeta \\ \xi \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} -G^* r_0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & 0 \\ 0 & G^* B \Sigma^{-1} \end{pmatrix} \right).$$

This completes the proof. □

Chapter 4

Estimation in the case of unknown change points

4.1 The unrestricted estimator

In the previous chapter, the locations of change-points, $\tau = (\tau_1, \ldots, \tau_m)^\top$, and the number of change points, m, are assumed to be known. Nevertheless, in practice, the change points are also unknown. Thus, the change points have to be estimated from the data. In this chapter, we assume that the number of change points, m, is known but the locations of change points are unknown. We show that the asymptotic property, in the case of known change points, holds when we replace change points by their consistent estimators. Let $\hat{\phi}_j$ be a consistent estimator of the parameter ϕ_j, $j = 1, ..., m$, and for convenience, let $\hat{\phi}_0 = 0$ and $\hat{\phi}_{m+1} = 1$. Let $\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_m)^\top$. First, for estimating the locations of change points, we recall the least sum of squared errors (LSSE) method, which is similar to that in Chen *et al.* (2017). We partition the time period $[0, T]$ into n parts, i.e., $0 = t_0 < \cdots < t_n = T$. The time increments,

$\Delta_t = t_{i+1} - t_i$, are exactly the same for $i = 0, ..., n-1$. Moreover, we define $Y_i = X_{t_{i+1}} - X_{t_i}$ and $z_i = (\varphi_1(t_i), ..., \varphi_p(t_i), -X_{t_i})\Delta_t$.

The exact value of the drift parameters θ may be different with the value of their MLE because of the uncertain location of estimated change points. For instance, if $\hat{\tau}_j > \tau_j$, then for all $t_i \in (\tau_j, \hat{\tau}_j]$, it is obvious that the corresponding true value of the drift parameters is $\theta^{(j+1)}$. However, in the same condition, the MLE of the drift parameters is $\hat{\theta}^{(j)}$ for all $t_i \in (\tau_j, \hat{\tau}_j]$. Therefore, we let $\theta_i = \sum_{j=1}^{m+1} \theta_j \mathbb{I}_{\{\tau_{j-1} < t_i \leq \tau_j\}}$ be the true value of the drift parameter at t_i. Also, $\hat{\theta}_i = \sum_{j=1}^{m+1} \hat{\theta}_j \mathbb{I}_{\{\hat{\tau}_{j-1} < t_i \leq \hat{\tau}_j\}}$, where $\hat{\theta}_j = Q^{-1}_{(\hat{\tau}_{j-1}, \hat{\tau}_j)} \tilde{r}_{(\hat{\tau}_{j-1}, \hat{\tau}_j)}$ for $j = 1, ..., m+1$, is set up to be the MLE of the drift parameters at t_i. By the Euler-Maruyama discretisation method, we have

$$Y_i = z_i\theta_i + \epsilon_i, \quad i = 1, ..., n \tag{4.1}$$

where ϵ_i is the error term $\sigma\sqrt{\Delta_t}\omega_i$, and ω_i is the ith independent draw from a standard normal variable. From (4.1), the estimators for the m change points, τ, are given by

$$\hat{\tau} = \arg\min_{\tau} \text{SSE}([0, T], \tau, \hat{\theta}(\tau)) \tag{4.2}$$

where

$$\text{SSE}([0, T], \tau, \hat{\theta}(\tau)) = \sum_{t_i \in [0,T]} (Y_i - z_i\hat{\theta}_i)^T(Y_i - z_i\hat{\theta}_i) \tag{4.3}$$

Assumption 3. *For every $j = 1, ..., m$, there exists an $L_0 > 0$ such that for all $L > L_0$ the minimum eigenvalues of $\frac{1}{L}\sum_{t_i \in (\tau_j, \tau_j+L]} z_i^T z_i$ and of $\frac{1}{L}\sum_{t_i \in (\tau_j - L, \tau_j]} z_i^T z_i$ as well as their respective continuous-time versions $\frac{1}{L}Q_{(\tau_j, \tau_j+L)}$ and $\frac{1}{L}Q_{(\tau_j - L, \tau_j)}$, are all bounded away from 0.*

Remark 3. *For the estimators of ϕ_j, we can directly obtain $\hat{\phi}_j = \frac{\hat{\tau}_j}{T}$, $j = 1, ..., m+1$. The consistency of $\hat{\phi}_j$ to ϕ_j is proved in Proposition B.1 in Appendix*

B. Clearly, the estimator $\hat{\phi}_j$ *is* $\mathfrak{F}_T$*-measurable and* $\hat{\phi}_j \in [0,1]$ *for* $j = 1, ..., m+1$. *Further, there exists* $\delta_0 > 0$ *such that* $\hat{\phi}_j - \phi_j = O_P(T^{-\delta_0})$ *for* $j = 1, ..., m$.

We introduce another method to estimate the locations of the change points. This is based on the Maximum log-likelihood. By Theorem 7.6 of Lipster and Shiryaev (2001), the log-likelihood function is given by

$$\log L(\tau, \theta) = \frac{1}{\sigma^2} \int_0^T S(\theta, t, X_t) dX_t - \frac{1}{2\sigma^2} \int_0^T S^2(\theta, t, X_t) dt,$$

where $\tau = (\tau_1, \tau_2, \ldots, \tau_m)$. Note that, by (B.16),

$$\begin{aligned}\frac{1}{\sigma^2} \int_0^T S(\theta, t, X_t) dX_t &= \frac{1}{\sigma^2} \int_0^T \sum_{j=1}^{m+1} \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right) \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} dX_t \\ &= \frac{1}{\sigma^2} \sum_{j=1}^{m+1} \int_0^T \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right) \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} dX_t.\end{aligned}$$

This gives

$$\frac{1}{\sigma^2} \int_0^T S(\theta, t, X_t) dX_t = \frac{1}{\sigma^2} \sum_{j=1}^{m+1} \int_{\tau_{j-1}}^{\tau_j} \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right) dX_t.$$

Further, by the proof of Proposition 2.3,

$$\begin{aligned}\frac{1}{2\sigma^2} \int_0^T S^2(\theta, t, X_t) dt &= \frac{1}{2\sigma^2} \int_0^T \left[\sum_{j=1}^{m+1} \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right) \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} \right]^2 dt \\ &= \frac{1}{2\sigma^2} \int_0^T \sum_{j=1}^{m+1} \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right)^2 \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} dt \\ &= \frac{1}{2\sigma^2} \sum_{j=1}^{m+1} \int_0^T \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right)^2 \mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} dt.\end{aligned}$$

This gives

$$\frac{1}{2\sigma^2} \int_0^T S^2(\theta, t, X_t) dt = \frac{1}{2\sigma^2} \sum_{j=1}^{m+1} \int_{\tau_j - 1}^{\tau_j} \left(\sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t \right)^2 dt.$$

Therefore, for the change points $\tau_1, \ldots, \tau_m$, the log-likelihood function for SDE (2.1) is given by

$$\log L(\tau, \theta) = \frac{1}{\sigma^2} \sum_{j=1}^{m+1} \int_{\tau_{j-1}}^{\tau_j} S(\theta_j, t, X_t) dX_t - \frac{1}{2\sigma^2} \sum_{j=1}^{m+1} \int_{\tau_{j-1}}^{\tau_j} S(\theta_j, t, X_t)^2 dt \qquad (4.4)$$

where $S(\theta_j, t, X_t) = \sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t$. From (4.4), when the number of change point, m, is known, the estimator of τ is

$$\hat{\tau} = \arg\max_{\tau} \log L(\tau, \hat{\theta}(\tau)) \qquad (4.5)$$

where $\hat{\theta}(\tau)$ is the MLE of θ by using the given change points τ. Auger and Lawrence (1989) introduced a numerical method to approximate the integrals inside the log-likelihood function. In this case, we use this method to calculate $\log L(\tau, \hat{\theta}(\tau))$ in (4.5). Divide $[0, T]$ into n parts, i.e. $0 = t_0^* < \cdots < t_n^* = T$ with $\Delta_t^* = t_{i+1}^* - t_i^*$. By the Riemann sum, the log-likelihood function in (4.5) is approximated as

$$\begin{aligned} \log L^*([0,T], \tau, \hat{\theta}(\tau)) = & \frac{1}{\sigma^2} \sum_{j=1}^{m+1} \sum_{t_i^* \in (\tau_{j-1}, \tau_j]} \hat{\theta}_j^\top V(t_i^*)(X_{t_{i+1}^*} - X_{t_i^*}) \\ & - \frac{1}{2\sigma^2} \sum_{j=1}^{m+1} \sum_{t_i^* \in (\tau_{j-1}, \tau_j]} \left(\hat{\theta}_j^\top V(t_i^*)\right)^2 \Delta_t^* \end{aligned} \qquad (4.6)$$

where $V(t) = (\varphi_1(t), \ldots, \varphi_p(t), -X_t)^\top$. Then, the estimator of τ is given by

$$\hat{\tau} = \arg\max_{\tau} \log L^*([0,T], \tau, \hat{\theta}(\tau)). \qquad (4.7)$$

For convenience, let $\hat{\phi}_0 = 0$ and $\hat{\phi}_{m+1} = 1$. In order to study the asymptotic properties of the estimators of θ, we derive first the following preliminary results.

Lemma 4.1. *Let $\{Y_t, t \geq 0\}$ be a stochastic process $\{\mathfrak{F}_t, t \geq 0\}$-adapted and L^2-bounded. Suppose that $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ are $\mathfrak{F}_T$-measurable and consistent estimators for ϕ_j and ϕ_{j-1} respectively, $j = 1, ..., m+1$, and $0 \leq \hat{\phi}_1 < ... < \hat{\phi}_m \leq 1$ a.s.. Then,*

$$\frac{1}{T} \int_{\hat{\phi}_{j-1} T}^{\hat{\phi}_j T} Y_t dt - \frac{1}{T} \int_{\phi_{j-1} T}^{\phi_j T} Y_t dt \xrightarrow[T \to \infty]{L^1} 0.$$

The proof is given in Appendix B.

Lemma 4.2. *Let $\{Y_t, t \geq 0\}$ be a R^p-valued deterministic and bounded function. Suppose that $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ are $\mathfrak{F}_T$-measurable and consistent estimators for ϕ_j and ϕ_{j-1} respectively, $j = 1, ..., m+1$, and $0 \leq \hat{\phi}_1 < ... < \hat{\phi}_m \leq 1$ a.s.. Then,*

$$\frac{1}{\sqrt{T}} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} Y_t dW_t - \frac{1}{\sqrt{T}} \int_{\phi_{j-1}T}^{\phi_j T} Y_t dW_t \xrightarrow[T\to\infty]{L^2} 0.$$

The proof is given in Appendix B. By using Lemma 4.1, we establish Propositions 4.1 and 4.2.

Proposition 4.1. *If Assumptions 1-3 hold, then, $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$,*

$$\frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \varphi(t)\varphi^\top(t) dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1}) I_P.$$

The proof is given in Appendix B.

Proposition 4.2. *Suppose that Assumptions 1-3 hold. Then, $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$,*

(i) $\displaystyle\frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t \varphi(t) dt - \frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t \varphi(t) dt \xrightarrow[T\to\infty]{P} 0;$

(ii) $\displaystyle\frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t^2 dt - \frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0.$

The proof is given in Appendix B. By using Proposition 4.2, we derive Propositions 4.3 and 4.4.

Proposition 4.3. *Suppose that Assumptions 1-3 hold. Then, for $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$,* $\displaystyle\frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t \varphi(t) dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1}) \int_0^1 \tilde{h}_j(t) \varphi(t) dt.$

Proof. We have

$$\begin{aligned}\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t\varphi(t)dt &= \left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t\varphi(t)dt\right)\\ &+ \left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t\varphi(t)dt\right)\\ &+ \left(\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t\varphi(t)dt\right)\\ &+ \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t\varphi(t)dt.\end{aligned}$$

By Proposition 4.2

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \tag{4.8}$$

By Lemma 4.1,

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{L^1} 0,$$

which implies that

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \tag{4.9}$$

As in the proof of Proposition 3.3, we have

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \tag{4.10}$$

By Proposition 3.3,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t\varphi(t)dt \xrightarrow[T\to\infty]{P} (\phi_j-\phi_{j-1})\int_0^1 \tilde{h}_j(t)\varphi(t)dt. \tag{4.11}$$

Finally, combining (4.8), (4.9), (4.10) and (4.11),

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t\varphi(t)dt \xrightarrow[T\to\infty]{P} (\phi_j-\phi_{j-1})\int_0^1 \tilde{h}_j(t)\varphi(t)dt,$$

and this completes the proof. □

Proposition 4.4. *Suppose that Assumptions 1-3 hold. Then, for* $0 \le \phi_{j-1} < \phi_j \le 1$, $j = 1, ..., m+1$, $\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\left[\int_0^1 (\tilde{h}_j(t))^2 dt + \frac{\sigma^2}{2a_j}\right]$.

Proof. We have

$$\begin{aligned}\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t^2 dt &= \left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t^2 dt\right) \\ &+ \left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt\right) \\ &+ \left(\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt\right) \\ &+ \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt.\end{aligned}$$

By Proposition 4.2

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{4.12}$$

By Lemma 4.1,

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{L^1} 0,$$

which implies that

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{4.13}$$

As in the proof of Proposition 3.4, we have

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{4.14}$$

By Proposition 3.4,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\left[\int_0^1 (\tilde{h}_j(t))^2 dt + \frac{\sigma^2}{2a_j}\right]. \tag{4.15}$$

Finally, combining (4.12), (4.13), (4.14) and (4.15),

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\left[\int_0^1 (\tilde{h}_j(t))^2 dt + \frac{\sigma^2}{2a_j}\right],$$

this completes the proof. □

By combining Proposition 4.1, Proposition 4.3 and Proposition 4.4, we derive the following propositions which give the same results as in Proposition 3.6 in case the change points ϕ_j and ϕ_{j-1} are replaced by their consistent estimators $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ respectively.

First, we define

$$Q_{(\hat{\tau}_{j-1},\hat{\tau}_j)} = \begin{bmatrix} \int_{\hat{\tau}_{j-1}}^{\hat{\tau}_j} \varphi_1^2(t)dt & \dots & \int_{\hat{\tau}_{j-1}}^{\hat{\tau}_j} \varphi_1(t)\varphi_p(t)dt & -\int_{\hat{\tau}_{j-1}}^{\hat{\tau}_j} \varphi_1 X_t dt \\ \vdots & \vdots & \vdots & \vdots \\ -\int_{\hat{\tau}_{j-1}}^{\hat{\tau}_j} \varphi_1 X_t dt & \dots & -\int_{\hat{\tau}_{j-1}}^{\hat{\tau}_j} \varphi_p X_t dt & \int_{\hat{\tau}_{j-1}}^{\hat{\tau}_j} X_t^2 dt \end{bmatrix} \quad (4.16)$$

where $\hat{\tau}_j = \hat{\phi}_j T$, $\hat{\tau}_{j-1} = \hat{\phi}_{j-1} T$ for $j = 1, ..., m+1$.

Proposition 4.5. *Suppose that Assumptions 1-3 hold. Then, for* $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$, $TQ^{-1}_{(\hat{\tau}_{j-1},\hat{\tau}_j)} \xrightarrow[T\to\infty]{P} \frac{1}{\phi_j - \phi_{j-1}}(\Sigma_j)^{-1}$.

Proof. By Proposition 4.1, Proposition 4.3 and Proposition 4.4, we have

$$\frac{1}{T} Q_{(\hat{\tau}_{j-1},\hat{\tau}_j)} \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\Sigma_j, \ j = 1, ..., m+1.$$

By combining Propositions 2.4, 3.5, and the continuous mapping theorem, we get $TQ^{-1}_{(\hat{\tau}_{j-1},\hat{\tau}_j)} \xrightarrow[T\to\infty]{P} \frac{1}{\phi_j - \phi_{j-1}}(\Sigma_j)^{-1}$. This completes the proof. □

Now, we define

$$Q(\hat{\phi}, m) = \begin{bmatrix} Q_{(0,\hat{\tau}_1)} & 0 & \dots & 0 \\ 0 & Q_{(\hat{\tau}_1,\hat{\tau}_2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_{(\hat{\tau}_m,T)} \end{bmatrix}. \quad (4.17)$$

Proposition 4.6. *Suppose Assumptions 1-3 hold. Then, for* $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$, $TQ^{-1}(\hat{\phi}, m) \xrightarrow[T\to\infty]{P} \Sigma^{-1}$.

The proof is given in Appendix B.

Next, we derive some results which generalize Proposition 3.7 and Proposition 3.8. The results show that similar propositions hold with the change-points replaced by their consistent estimators. Let $\hat{\tau}_j = \hat{\phi}_j T$, $j = 1, ..., m+1$, and let

$$R(\hat{\phi}, m) = (r_{(0,\hat{\tau}_1)}, ..., r_{(\hat{\tau}_m,T)})^\top, \tag{4.18}$$

$$r(a, b) = \left(\int_a^b \varphi_1(t)dW_t, ..., \int_a^b \varphi_p(t)dW_t, -\int_a^b X_t dW_t \right)^\top,$$

for $0 \leq a < b \leq T$.

4.1.1 Asmptotic normality of the UE $\hat{\theta}(\hat{\phi}, m)$

In deriving the asymptotic normality of the UE, we use the following lemma.

Lemma 4.3. *Let* $\{Y_t, t \geq 0\}$ *be a solution of SDE*

$$dY_t = \sum_{k=1}^{m+1} f(\mu_k, Y_t)\mathbb{I}_{\{\tau_{j-1} < t \leq \tau_j\}} dt + \sigma dW_t, \ 0 \leq t \leq T \tag{4.19}$$

where $f(\theta, x)$ *is a real-valued function such that the processes* $\{Y_t, t \geq 0\}$ *and* $\{f(\theta, Y_t), t \geq 0\}$ *are* L^2*-bounded. Suppose that* $\hat{\phi}_j$ *and* $\hat{\phi}_{j-1}$ *are* $\mathfrak{F}_T$*-measurable and consistent estimators for* ϕ_j *and* ϕ_{j-1} *respectively,* $j = 1, ..., m+1$, *and* $0 \leq \hat{\phi}_1 < ... < \hat{\phi}_m \leq 1$ *a.s.. Further, assume there exists* $\delta_0 > \frac{1}{2}$ *such that*

$\max_{1\le j\le m}(|\hat{\phi}_j - \phi_j|) = O_P(T^{-\delta_0})$. *Then, for* $0 \le \phi_{j-1} < \phi_j \le 1$, $j = 1, \ldots, m+1$, $\frac{1}{\sqrt{T}}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} Y_t dW_t - \frac{1}{\sqrt{T}}\int_{\phi_{j-1}T}^{\phi_j T} Y_t dW_t \xrightarrow[T\to\infty]{P} 0.$

The proof is given in Appendix B. Lemma 4.3 yields Lemma 3.3 in Nkurunziza and Zhang (2018) for which $m = 1$.

Proposition 4.7. *Suppose that Assumptions 1-3 hold. Then, for* $0 \le \phi_{j-1} < \phi_j \le 1$, $j = 1, ..., m+1$, $\frac{1}{\sqrt{T}}\Big(R(\hat{\phi}, m) - R(\phi, m)\Big) \xrightarrow[T\to\infty]{P} 0$, *where* $R(\hat{\phi}, m)$ *is defined in (4.18) and* $R(\phi, m)$ *is defined in (3.2).*

The proof is given in Appendix B. By using Proposition 4.7, we derive the following proposition which shows the limiting distribution of $\frac{1}{\sqrt{T}}R(\hat{\phi}, m)$.

Proposition 4.8. *If the conditions in Proposition 4.7 hold, then, for* $0 \le \phi_{j-1} < \phi_j \le 1$, $j = 1, ..., m+1$, $\frac{1}{\sqrt{T}}R(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} r^* \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma)$.

Proof. By Proposition 3.7 and Proposition 4.7, we have

$$\frac{1}{\sqrt{T}}R(\phi, m) \xrightarrow[T\to\infty]{d} r^* \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma)$$

and

$$\frac{1}{\sqrt{T}}\Big(R(\hat{\phi}, m) - R(\phi, m)\Big) \xrightarrow[T\to\infty]{P} 0.$$

Hence, by Slutsky's Theorem,

$$\frac{1}{\sqrt{T}}R(\hat{\phi}, m) = \frac{1}{\sqrt{T}}\Big(R(\hat{\phi}, m) - R(\phi, m)\Big) + \frac{1}{\sqrt{T}}R(\phi, m) \xrightarrow[T\to\infty]{d} r^* \sim \mathbb{N}_{(m+1)(p+1)}(0, \Sigma).$$

This completes the proof. □

Now, let $\hat{\tau}_j = \hat{\phi}_j T$, $j = 1, ..., m+1$, and

$$\tilde{R}(\hat{\phi}, m) = (\tilde{r}_{(0,\hat{\tau}_1)}, ..., (\tilde{r}_{(\hat{\tau}_m,T)})^\top, \tag{4.20}$$

$$\tilde{r}(a,b) = \left(\int_a^b \varphi_1(t)dX_t, ..., \int_a^b \varphi_p(t)dX_t, -\int_a^b X_t dX_t \right)^\top,$$

for $0 \leq a < b \leq T$.

Further, let $\hat{\theta}(\hat{\phi}, m) = Q^{-1}(\hat{\phi}, m)\tilde{R}(\hat{\phi}, m)$ be the plug-in estimator where $Q(\hat{\phi}, m)$ and $\tilde{R}(\hat{\phi}, m)$ are defined in (4.17) and (4.20). Now, we define $\rho_T(\hat{\phi}, m) = \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta)$. By combining Propositions 4.6 and 4.8, we derive the asymptotic normality of UE $\hat{\theta}(\hat{\phi}, m)$ in the following proposition.

Corollary 4.1. *Suppose that the conditions in Proposition 4.7 hold. Then,*
$\rho_T(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2\Sigma^{-1})$.

The proof is given in Appendix B.

4.2 The restricted estimator

In the previous chapter, we studied the asymptotic property of the *Restricted Estimator* (RE) in context of known multiple change points. In this section, we give the *Restricted Estimator* (RE) in context of unknown multiple change-points. The estimator of the rate of change point $\hat{\phi}_j$, which is consistent, will be involved instead of the known rate of change point ϕ_j, $j = 1, ..., m$.

Proposition 4.9. *Suppose that Assumptions 1-3 hold along with (2.3) and let* $J = Q^{-1}(\hat{\phi}, m)B^\top(BQ^{-1}(\hat{\phi}, m)B^\top)^{-1}$. *Then, the* RE *of* θ *is*

$$\tilde{\theta}(\hat{\phi}, m) = \hat{\theta}(\hat{\phi}, m) - J(B\hat{\theta}(\hat{\phi}, m) - r). \tag{4.21}$$

This proof follows from the similar steps of the proof of Proposition 3.9. Since we consider the case of unknown change points, we just replace $\tilde{R}(\phi, m)$ and $Q(\phi, m)$ by $\tilde{R}(\hat{\phi}, m)$ and $Q(\hat{\phi}, m)$, respectively.

4.2.1 Asymptotic normality of the RE $\tilde{\theta}(\hat{\phi}, m)$

In this section, we derive the asymptotic normality of the RE $\tilde{\theta}(\hat{\phi}, m)$ based on the asymptotic normality of the *Unrestricted Estimator* (UE) $\hat{\theta}(\hat{\phi}, m)$. We show that the RE $\tilde{\theta}(\hat{\phi}, m)$ has the same limiting distribution as the RMLE $\tilde{\theta}(\phi, m)$. The established result is similar to that in Perron and Qu (2006), Chen and Nkurunziza (2015), Nkurunziza and Zhang (2018) among others.

Based on Proposition 4.9, we have

$$\begin{aligned}\sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) &= \sqrt{T}[Jr + (I_{(m+1)(p+1)} - JB)\hat{\theta}(\hat{\phi}, m) - \theta] \\ &= \sqrt{T}(Jr - \theta) + \sqrt{T}(I_{(m+1)(p+1)} - JB)\hat{\theta}(\hat{\phi}, m),\end{aligned}$$

this gives

$$\sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) = (I_{(m+1)(p+1)} - JB)\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) - \sqrt{T}J(B\theta - r).$$

Now, we define $\zeta_T(\hat{\phi}, m) = \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta)$. Then,

$$\sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) = (I_{(m+1)(p+1)} - JB)\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) - \sqrt{T}J(B\theta - r). \quad (4.22)$$

Consider a continuous function $g(X) = XB^\top(BXB^\top)^{-1}$ where X is a positive definite matrix. We have $g(TQ^{-1}(\hat{\phi}, m)) = J = TQ^{-1}(\hat{\phi}, m)B^\top(BTQ^{-1}(\hat{\phi}, m)B^\top)^{-1}$.

By combining Proposition 4.6 and the continuous mapping theorem, we get

$$J = TQ^{-1}(\hat{\phi}, m)B^\top(BTQ^{-1}(\hat{\phi}, m)B^\top)^{-1} \xrightarrow[T\to\infty]{P} \Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1} = G^*,$$

and

$$I_{(m+1)(p+1)} - JB \xrightarrow[T\to\infty]{P} I_{(m+1)(p+1)} - G^*B. \quad (4.23)$$

Under the set of local alternatives in (3.8), we have

$$\sqrt{T}J(B\theta - r) = \sqrt{T}J\frac{r_0}{\sqrt{T}} = Jr_0 \xrightarrow[T\to\infty]{P} G^*r_0. \quad (4.24)$$

Corollary 4.2 which shows the asymptotic normality of the RE $\tilde{\theta}(\hat{\phi}, m)$ is given in the next section.

4.3 Joint asymptotic normality of $\hat{\theta}(\hat{\phi}, m)$ and $\tilde{\theta}(\hat{\phi}, m)$

In this section, we present the joint asymptotic normality of UE $\hat{\theta}(\hat{\phi}, m)$ and RE $\tilde{\theta}(\hat{\phi}, m)$. First, we study the asymptotic distribution of $(\rho_T^\top(\hat{\phi}, m), \zeta_T^\top(\hat{\phi}, m))^\top = \sqrt{T}\left((\hat{\theta}(\hat{\phi}, m) - \theta)^\top, (\tilde{\theta}(\hat{\phi}, m) - \theta)^\top\right)^\top$.

Proposition 4.10. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). Then, if* $r_0 \neq 0$, $(\rho_T^\top(\hat{\phi}, m), \zeta_T^\top(\hat{\phi}, m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top, \zeta^\top)^\top$ *and if* $r_0 = 0$, $(\rho_T^\top(\hat{\phi}, m), \zeta_T^\top(\hat{\phi}, m))^\top \xrightarrow[T\to\infty]{d} (\rho_0^\top, \zeta_0^\top)^\top$, *where*

$$\begin{pmatrix}\rho \\ \zeta\end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix}0 \\ -G^* r_0\end{pmatrix}, \sigma^2 \begin{pmatrix}\Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1}\end{pmatrix}\right),$$

$$\begin{pmatrix}\rho_0 \\ \zeta_0\end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix}0 \\ 0\end{pmatrix}, \sigma^2 \begin{pmatrix}\Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1}\end{pmatrix}\right).$$

Proof. We can observe that

$$\begin{aligned}
\begin{pmatrix}\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta)\end{pmatrix} &= \begin{pmatrix}\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ (I_{(m+1)(p+1)} - JB)\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) - \sqrt{T}J(B\theta - r)\end{pmatrix} \\
&= \begin{pmatrix}\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ (I_{(m+1)(p+1)} - JB)\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) - Jr_0\end{pmatrix} \\
&= \begin{pmatrix}I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - JB\end{pmatrix}\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) + \begin{pmatrix}0 \\ -Jr_0\end{pmatrix}.
\end{aligned}$$

By (4.23), we know

$$I_{(m+1)(p+1)} - JB \xrightarrow[T\to\infty]{P} I_{(m+1)(p+1)} - G^* B.$$

Then,

$$\begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - JB \end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^*B \end{pmatrix} \tag{4.25}$$

where all the elements in (4.25) are non-random. Similarly, by (4.24), we have

$$\begin{pmatrix} 0 \\ -Jr_0 \end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix} 0 \\ -G^*r_0 \end{pmatrix}. \tag{4.26}$$

By combining Corollary 4.1 and the relations (4.25), (4.26) along with Slutsky's Theorem,

$$\begin{pmatrix} \rho_T(\hat{\phi}, m) \\ \zeta_T(\hat{\phi}, m) \end{pmatrix} = \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - JB \end{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) + \begin{pmatrix} 0 \\ -Jr_0 \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{d} \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^*B \end{pmatrix} \rho + \begin{pmatrix} 0 \\ -G^*r_0 \end{pmatrix} = \begin{pmatrix} \rho \\ \zeta \end{pmatrix}$$

Then, by Proposition A.2 in Appendix A,

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} \rho \\ \zeta \end{pmatrix}$$

$$\sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} 0 \\ -G^*r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^*B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^*B \end{pmatrix}^\top \right).$$

Note that, from the proof of Proposition 3.11, we get

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} \rho \\ \zeta \end{pmatrix}$$

$$\sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} 0 \\ -G^*r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^*B\Sigma^{-1} \\ \Sigma^{-1} - G^*B\Sigma^{-1} & \Sigma^{-1} - G^*B\Sigma^{-1} \end{pmatrix} \right),$$

and this completes the proof. □

By using Proposition 4.10, we derive the following corollary which shows that the RE is asymptotically normal.

Corollary 4.2. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). Then,* $\zeta_T(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} \zeta \sim \mathcal{N}_{(m+1)(p+1)}\big(-G^* r_0, \sigma^2(\Sigma^{-1} - G^* B \Sigma^{-1})\big)$.

The proof follows from Proposition 4.10.

Now, we define $\xi_T(\hat{\phi}, m) = \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m))$. Next, we study the asymptotic distribution of $(\rho_T^\top(\hat{\phi}, m), \xi_T^\top(\hat{\phi}, m))^\top = \sqrt{T}\left((\hat{\theta}(\hat{\phi}, m) - \theta)^\top, (\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m))^\top\right)^\top$.

Proposition 4.11. *Suppose that Assumption 1-3 hold along with the set of local alternatives in (3.8). Then, if* $r_0 \neq 0$, $(\rho_T^\top(\hat{\phi}, m), \xi_T^\top(\hat{\phi}, m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top, \xi^\top)^\top$ *and if* $r_0 = 0$, $(\rho_T^\top(\hat{\phi}, m), \xi_T^\top(\hat{\phi}, m))^\top \xrightarrow[T\to\infty]{d} (\rho_0^\top, \xi_0^\top)^\top$, *where*

$$\begin{pmatrix} \rho \\ \xi \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} 0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix}\right),$$

$$\begin{pmatrix} \rho_0 \\ \xi_0 \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix}\right).$$

Proof. We have

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m)) \end{pmatrix} = \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix}.$$

We know

$$\begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix},$$

and

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} \rho \\ \zeta \end{pmatrix}$$

$$\sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} 0 \\ -G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix} \right).$$

Then, by Slutsky's Theorem,

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m)) \end{pmatrix} = \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{d} \begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \rho \\ \zeta \end{pmatrix}.$$

Note that

$$\begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} 0 \\ -G^* r_0 \end{pmatrix} = \begin{pmatrix} 0 \\ G^* r_0 \end{pmatrix},$$

and

$$\begin{pmatrix} I_{(m+1)(p+1)} & 0 \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix}$$

$$\times \begin{pmatrix} I_{(m+1)(p+1)} & I_{(m+1)(p+1)} \\ 0 & -I_{(m+1)(p+1)} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & 0 \end{pmatrix} \begin{pmatrix} I_{(m+1)(p+1)} & I_{(m+1)(p+1)} \\ 0 & -I_{(m+1)(p+1)} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix}.$$

Therefore, by Proposition A.2 in Appendix A, $(\rho_T^\top(\hat{\phi}, m), \xi_T^\top(\hat{\phi}, m))^\top \xrightarrow[T\to\infty]{d} (\rho^\top, \xi^\top)^\top$ with $(\rho^\top, \xi^\top)^\top \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} 0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & G^* B \Sigma^{-1} \end{pmatrix}\right)$. This completes the proof. □

From Proposition 4.11, we derive the following result which gives the limiting distribution of $\xi_T(\hat{\phi}, m)$.

Corollary 4.3. *If Assumptions 1-3 hold along with the set of local alternatives in (3.8), then,* $\xi_T(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} \xi \sim \mathcal{N}_{(m+1)(p+1)}(G^* r_0, \sigma^2 G^* B \Sigma^{-1})$.

The proof follows from Proposition 4.11. We also derive the asymptotic distribution of $(\zeta_T^\top(\hat{\phi}, m), \xi_T^\top(\hat{\phi}, m))^\top = \sqrt{T}\left((\tilde{\theta}(\hat{\phi}, m) - \theta)^\top, (\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m))^\top\right)^\top$.

Proposition 4.12. *Suppose that Assumption 1-3 hold along with the set of local alternatives in (3.8). Then,* $(\zeta_T^\top(\hat{\phi}, m), \xi_T^\top(\hat{\phi}, m))^\top \xrightarrow[T\to\infty]{d} (\zeta^\top, \xi^\top)^\top$, *where*

$$\begin{pmatrix} \zeta \\ \xi \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} -G^* r_0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & 0 \\ 0 & G^* B \Sigma^{-1} \end{pmatrix}\right).$$

Proof. We have

$$\begin{pmatrix} \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m)) \end{pmatrix} = \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix}.$$

Further, we have

$$\begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \xrightarrow[T\to\infty]{P} \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix},$$

and

$$\begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} \rho \\ \zeta \end{pmatrix}$$

$$\sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} 0 \\ -G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix} \right).$$

Then, by Slutsky's Theorem,

$$\begin{pmatrix} \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m)) \end{pmatrix} = \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{d} \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} \rho \\ \zeta \end{pmatrix}.$$

Note that

$$\begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \begin{pmatrix} 0 \\ -G^* r_0 \end{pmatrix} = \begin{pmatrix} -G^* r_0 \\ G^* r_0 \end{pmatrix},$$

and

$$\begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix} \sigma^2 \begin{pmatrix} \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix}$$

$$\times \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & \Sigma^{-1} - G^* B \Sigma^{-1} \\ G^* B \Sigma^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} & -I_{(m+1)(p+1)} \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & 0 \\ 0 & G^* B \Sigma^{-1} \end{pmatrix}.$$

Therefore, by Proposition A.2 in Appendix A,

$$\begin{pmatrix} \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta) \\ \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m)) \end{pmatrix} \xrightarrow[T\to\infty]{d} \begin{pmatrix} \zeta \\ \xi \end{pmatrix}$$
$$\sim \mathcal{N}_{2(m+1)(p+1)} \left(\begin{pmatrix} -G^* r_0 \\ G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \Sigma^{-1} - G^* B \Sigma^{-1} & 0 \\ 0 & G^* B \Sigma^{-1} \end{pmatrix} \right).$$

This completes the proof. □

Chapter 5

The case of an unknown number of change points

In Chapters 3 and 4, we suppose that the number of change points, m, is known. However, for some data sets, m is unknown. Thus, in this chapter, we solve a more general problem where the nuisance parameter m, τ_1, τ_2, ..., τ_m are unknown.

5.1 Estimating the number of change points

In this section, we consider detecting the unknown number of change points. We use similar methodology as introduced by Chen *et al.* (2017). In Chen *et al.* (2017), they treated estimating the number of change points as selecting the best fitting model. Thus, for models with different possible numbers of change points, we choose the model which fits the data best. To choose the best fitting model, we are looking for the one which minimizes the log-likelihood-based information criterion

$$\mathrm{IC}(m) = -2\log L(\tau, \hat{\theta}) + (m+1)h(p)\Upsilon(T) + \lambda^\top (B\hat{\theta} - r) \tag{5.1}$$

where $\log L(\tau, \hat{\theta})$ is defined in (4.4); $\hat{\tau}$ is established by (4.5) corresponding to each m; $h(p) = p + 1$ if there is no change in σ or $h(p) = p + 2$ if there is a change in σ; $\Upsilon(T)$ is a non-decreasing function of T, the total time period of the data set; and m is the potential number of change points to be set; B and r are defined in (2.3).

From the asymptotic property of Riemann sum approximation of $\log L(\tau, \hat{\theta})$, the information criterion is given by

$$\mathrm{IC}(m) = -2 \log L^*([0, T], \tau, \hat{\theta}(\tau)) + (m + 1)h(p)\Upsilon(T) + \lambda^\top (B\hat{\theta} - r) \tag{5.2}$$

where $\log L^*([0, T], \tau, \hat{\theta}(\tau))$ is defined in (4.6); and $\hat{\tau}$ is established by (4.7) corresponding to each m.

It should be noticed that the term $(m+1)h(p)\Upsilon(T)$ is fixed when the number of change points is known. Then, the approach involving (5.2) is the same as the maximum log-likelihood method introduced in Section 4.1. It is obvious that (5.2) represents the well-known information criterion called Akaike information criterion (AIC) Akaike (1973) when $\Upsilon(T) = 2$. However, as mentioned in Chen *et al.* (2017), due to the problem of the consistency of AIC, one uses the Schwarz information criterion (SIC) as proposed in Schwarz (1978). In SIC, $\Upsilon(T)$ is set as the logarithm of the sample size. In Schwarz (1978), the authors used the SIC successfully in change-point analysis.

By Proposition B.2 in Appendix B, as T is large, IC(m) given in (5.2) reaches its minimum value when $m = m^0$ where m^0 is the exact value of the number of change points. Hence, detecting m^0 is the same as finding the IC(m) in (5.2) which reaches its minimum. Then, its corresponding m is the number of change points we would like to estimate.

5.2 Computational algorithms

In this section, we introduce an algorithm which is useful in finding θ, τ and m. In particular, the algorithm is based on (5.2). Let $\hat{m}$ be the estimator of m, let $\hat{\tau}(\hat{m})$ be the estimator of $\tau(m)$. For estimating $\tau(m)$, we apply the LSSE method or the Maximum log-likelihood method in Section 4. Note that some steps of the algorithm are based on the dynamic programming algorithm from Bai and Perron (1998), Perron and Qu (2006).

Algorithm. Let $H_1(r, T_r)$ be either $H_1(r, T_r) = \min_\tau \mathrm{SSE}([0, T_r], \tau, \hat{\theta}(\tau))$, the least sum squared error for (4.2) or $H_1(r, T_r) = \max_\tau \log L^*([0, T_r], \tau, \hat{\theta}(\tau))$, the maximum Riemann sum approximation of log-likelihood for (4.7) computed based on the optimal partition of time interval $[0, T_r]$ that contains r change points. Also, let $H_2(a, b)$ be the SSE for (4.2) or Riemann sum approximation of log-likelihood for (4.7) computed based on a time regime $(a, b]$. Further, let $h = \epsilon T$ be the minimal permissible length of a time regime. Then, (4.2) or (4.7) with m change points can be computed as follows.

Step 1: Compute and save $H_2(a, b)$ for all time periods $(a, b]$ that satisfy $b - a \geq h$.

Step 2: Compute and save $H_1(1, T_1)$ for all $T_1 \in [2h, T - (m-1)h]$ by solving the optimization problem

$$H_1(1, T_1) = \begin{cases} \min_{a \in [h, T_1 - h]} [H_2(0, a) + H_2(a, T_1)] & \text{for (4.2)} \\ \max_{a \in [h, T_1 - h]} [H_2(0, a) + H_2(a, T_1)] & \text{for (4.7).} \end{cases}$$

Step 3: Sequentially compute and save

$$H_1(r, T_r) = \begin{cases} \min_{a\in[rh,T_r-h]}[H_1(r-1,a) + H_2(a,T_r)] & \text{for (4.2)} \\ \max_{a\in[rh,T_r-h]}[H_1(r-1,a) + H_2(a,T_r)] & \text{for (4.7).} \end{cases}$$

for $r = 2, \ldots, m-1$, and $T_r \in [(r+1)h, T-(m-r)h]$.

Step 4: Finally, the estimated change points are obtained by solving

$$H_1(m, T) = \begin{cases} \min_{a\in[mh,T-h]}[H_1(m-1,a) + H_2(a,T)] & \text{for (4.2)} \\ \max_{a\in[mh,T-h]}[H_1(m-1,a) + H_2(a,T)] & \text{for (4.7),} \end{cases}$$

and $H_1(m-1, a) = H_2(0, a)$ if $m = 1$.

Step 5: Follow steps 1-4 to search for the optimal locations of the m estimated change points then store the computed value of (5.2) for $m = 0, 1, 2$. Note that the results of $H_2(a, b)$ for all $(a, b]$ such that $a - b \geq h$ as well as the optimization results of $H_1(r, T_r)$ for all $r = 1, \ldots, m$ and $T_r \in [(r+1)h, T-(m-r)h]$ need to be stored for future use.

Step 6: For $m = 3, \ldots, m_{\max}$, first let $r = m-1$ and $T_r \in [(r+1)h, T-(m-r)h]$ then compute and store $H_1(r, T_r)$. Next let $r = m$ and the estimated change points are obtained by solving $H_1(m, T)$, where $H_1(r, T_r)$ and $H_1(m, T)$. Finally, based on the estimated m change points, compute and store $\text{IC}(m)$.

Step 7: $\hat{m}$ is obtained from $m = 1, \ldots, m_{\max}$ that returns the smallest value of (5.2).

To find $\hat{m}$, at first, we need to find the range of m, $0 < m \leq m_{\max}$ where $0 \leq m_{\max} \leq \lceil [T/h] \rceil$. The $m_{\max}$ can be determined by observing and analyzing the given process. By Proposition B.2 in Appendix B, $\hat{m}$ is a consistent estimator provided $m^0 \in [0, m_{\max}]$.

5.3 Asymptotic properties of the UE and the RE

In this section, we derive some asymptotic properties of $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$. As compared to the results of Chapter 3-4, the problem studied here is very challenging. The main difficulty consists in the fact that the dimensions of $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$ depend on $\hat{m}$ which is random variable. To overcome this difficulty, we establish a lemma and a proposition which are of interest in their own.

Let $\hat{m}$ be a consistent estimator for m. The UE and RE are obtained as in Section 4, by plug-in i.e. by replacing , in $\hat{\theta}(\hat{\phi}, m)$ and $\tilde{\theta}(\hat{\phi}, m)$, m by $\hat{m}$. Thus, the UE is given by $\hat{\theta}(\hat{\phi}, \hat{m})$ and the RE is given by $\tilde{\theta}(\hat{\phi}, \hat{m})$. Below, we derive a result which is useful in establishing a test for the testing problem in (2.4), as well as in studying the relative efficiency of the UE and the RE. As a preliminary result, we prove the following lemma.

Lemma 5.1. *Let $\hat{m}$ be non-negative integer valued random variable and let m be a nonrandom integer number such that $\hat{m} \xrightarrow[T\to\infty]{P} m$. Let $X_T(\hat{m})$, $X_T(m)$ and $X(m)$ be q-column random vectors such that $X_T(m) \xrightarrow[T\to\infty]{d} X(m)$. Then, $X_T(\hat{m}) \xrightarrow[T\to\infty]{d} X(m)$.*

Proof. For the sake of simplicity, for q-column vectors a and b, we write $a \leq b$ to stand for $a_i \leq b_i$, $i = 1, 2, \ldots, q$. Let x be a point of continuity of the cdf of $X(m)$. We have

$$\lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leq x) = \lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leq x, \hat{m} = m) + \lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leq x, \hat{m} \neq m).$$

$$\lim_{T\to\infty} \mathrm{P}(X_T(m) \leq x) = \lim_{T\to\infty} \mathrm{P}(X_T(m) \leq x, \hat{m} = m) + \lim_{T\to\infty} \mathrm{P}(X_T(m) \leq x, \hat{m} \neq m).$$

Since $\lim_{T\to\infty} \mathrm{P}(\hat{m} = m) = 1$, then,

$$\lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leq x) = \lim_{T\to\infty} \mathrm{P}(X_T(m) \leq x, \hat{m} = m), \tag{5.3}$$

$$\lim_{T\to\infty} \mathrm{P}(X_T(m) \le x) = \lim_{T\to\infty} \mathrm{P}(X_T(m) \le x, \hat{m} = m). \tag{5.4}$$

By combining (5.3), (5.4) and $\lim_{T\to\infty} \mathrm{P}(X_T(m) \le x) = \mathrm{P}(X(m) \le x)$, we have $\lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \le x) = \mathrm{P}(X(m) \le x)$. This complete the proof. □

By combining this lemma with Proposition 4.10 and Proposition 4.11., we establish the following proposition. Let $\rho_T(\hat{\phi}, \hat{m}) = \sqrt{T}(\hat{\theta}(\hat{\phi}, \hat{m}) - \theta)$, let
$\zeta_T(\hat{\phi}, \hat{m}) = \sqrt{T}(\tilde{\theta}(\hat{\phi}, \hat{m}) - \theta)$ and let $\xi_T(\hat{\phi}, \hat{m}) = \sqrt{T}(\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m}))$.
In the following proposition, let $g : \mathbb{R}^{(m+1)(p+1)} \times \mathbb{R}^{(m+1)(p+1)} \to \mathbb{R}^q$ be a continuous function, where q does not depend on m.

Proposition 5.1. *Suppose that Assumption 1-3 hold along with the set of local alternatives in (3.8). Then, if* $r_0 \neq 0$, $g(\rho_T(\hat{\phi}, \hat{m}), \zeta_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho, \zeta)$, *and* $g(\rho_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho, \xi)$, *where* ρ, ζ *and* ξ *are defined in Proposition 4.10 and Proposition 4.11. Moreover, if* $r_0 = 0$, $g(\rho_T(\hat{\phi}, \hat{m}), \zeta_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho_0, \zeta_0)$, *and* $g(\rho_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho_0, \xi_0)$, *where* ρ_0, ζ_0 *and* ξ_0 *are defined in Proposition 4.10 and Proposition 4.11.*

The proof follows directly from Proposition 4.10 and Proposition 4.11. Proposition 5.1 is useful in constructing a test statistic for the testing problem in (2.4). It is also used to derive the local power as well as the asymptotic distribution risk of the proposed estimators.

Chapter 6

Shrinkage estimators

In this chapter, we construct a test for testing the restriction and derive a class of shrinkage estimators which includes as special cases the UE, the RE, the shrinkage estimator (SE) and positive-part shrinkage estimator (PSE) for θ. As compared to the results in statistical literature, the novelty of the established results consists in the fact that, the distributions of the RE and the UE are random, as they are functions of $\hat{m}$.

6.1 Testing the restriction

In this section, we develop a test for testing $H_0 : B\theta = r$ versus $H_a : B\theta \neq r$. First, note that, in the continuous time observation, the diffusion coefficient (i.e. σ^2) is considered as known as it is equal to the quadratic variation of the process. Let $\hat{\sigma}^2$ be the discretized version of quadratic variation of the process, and note that $\hat{\sigma}^2$ is a consistent estimator for σ^2. Let $\chi^2_q(\lambda)$ be the chi-square random variable (r.v.) with q-degrees of freedom (df), and non-centrality parameter λ; let χ^2_q be the chi-square r.v. with q df. Also, define $\Delta = \frac{1}{\sigma^2} r_0^\top (B\Sigma^{-1}B^\top)^{-1} r_0$ where r_0 is given as in (3.8), and let $\hat{\Gamma} = \frac{1}{\hat{\sigma}^2} B^\top (BTQ^{-1}(\hat{\phi}, \hat{m})B^\top)^{-1}B$, $\Gamma = \frac{1}{\sigma^2} B^\top (B\Sigma^{-1}B^\top)^{-1}B$.

From Proposition 5.1, we derive the following corollary. This corollary is the foundation to test $H_0 : B\theta = r$ versus $H_a : B\theta \neq r$. Let $\psi_T(\hat{m}) = \xi_T(\hat{\phi}, \hat{m})^\top \hat{\Gamma} \xi_T(\hat{\phi}, \hat{m})$, let $\psi(m) = \xi^\top \Gamma \xi$, and let $\psi_0(m) = \xi_0^\top \Gamma \xi_0$.

Corollary 6.1. *Suppose that the conditions of Proposition 5.1 hold. Then, if* $r_0 \neq 0$, $\psi_T(\hat{m}) \xrightarrow[T\to\infty]{d} \psi(m) \sim \chi_q^2(\Delta)$. *Moreover, if* $r_0 = 0$, *then* $\psi_T(\hat{m}) \xrightarrow[T\to\infty]{d} \psi_0(m) \sim \chi_q^2$.

Proof. We first give the proof for the case when $r_0 \neq 0$. By Proposition 5.1, we have $g(\rho_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho, \xi)$, where $(\rho^\top, \xi^\top)^\top$ are given by Proposition 4.11, for any function $g : \mathbb{R}^{(m+1)(p+1)} \times \mathbb{R}^{(m+1)(p+1)} \to \mathbb{R}^q$.

Take $g(x, y) = y^\top \Gamma y$. We get

$$\xi_T^\top(\hat{\phi}, \hat{m}) \Gamma \xi_T(\hat{\phi}, \hat{m}) \xrightarrow[T\to\infty]{d} \xi^\top \Gamma \xi. \tag{6.1}$$

Further, from lemma 5.1, we have

$$\hat{\Gamma} = \frac{1}{\hat{\sigma}^2} B^\top (BTQ^{-1}(\hat{\phi}, \hat{m})B^\top)^{-1}B \xrightarrow[T\to\infty]{P} \Gamma = \frac{1}{\sigma^2} B^\top (B\Sigma^{-1}B^\top)^{-1}B. \tag{6.2}$$

Then, combining (6.1) and (6.2) with Slutsky's Theorem,

$$\psi_T(\hat{m}) = \xi_T(\hat{\phi}, \hat{m})^\top \hat{\Gamma} \xi_T(\hat{\phi}, \hat{m}) \xrightarrow[T\to\infty]{d} \psi(m) = \xi^\top \Gamma \xi.$$

It suffices to apply Theorem 5.1.3 in Mathai and Provost (1992) (see also Theorem A.5 in the Appendix A) to prove that $\xi^\top \Gamma \xi \sim \chi_q^2(\Delta)$. Namely, it suffices to show that

(i) $\mathrm{trace}(\Gamma \sigma^2 G^* B \Sigma^{-1}) = q$ and $(G^* r_0)^\top \Gamma G^* r_0 = \Delta$

(ii) $\sigma^2 G^* B \Sigma^{-1} \Gamma \sigma^2 G^* B \Sigma^{-1} \Gamma \sigma^2 G^* B \Sigma^{-1} = \sigma^2 G^* B \Sigma^{-1} \Gamma \sigma^2 G^* B \Sigma^{-1}$

(iii) $(G^* r_0)^\top \Gamma \sigma^2 G^* B \Sigma^{-1} \Gamma G^* r_0 = (G^* r_0)^\top \Gamma G^* r_0$

(iv) $(G^*r_0)^\top(\Gamma\sigma^2G^*B\Sigma^{-1})^2 = (G^*r_0)^\top\Gamma\sigma^2G^*B\Sigma^{-1}$

For the statment in (i), since we defined $G^* = \Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}$,

$$
\begin{aligned}
(G^*r_0)^\top\Gamma G^*r_0 &= r_0^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}r_0\\
&= \frac{1}{\sigma^2}r_0^\top(B\Sigma^{-1}B^\top)^{-1}r_0 = \Delta,
\end{aligned}
$$

and

$$
\begin{aligned}
\Gamma\sigma^2G^*B\Sigma^{-1} &= \frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B\sigma^2\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\\
&= B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1},
\end{aligned}
$$

which implies that

$$
\begin{aligned}
\operatorname{trace}(\Gamma\sigma^2G^*B\Sigma^{-1}) &= \operatorname{trace}(B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1})\\
&= \operatorname{trace}((B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}B^\top) = q,
\end{aligned}
$$

this proves the statement in (i).

For the statement in (ii), we have

$$
\sigma^2G^*B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1} = \sigma^2\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B
$$
$$
\times\,\sigma^2\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1}.
$$

This gives

$$
\begin{aligned}
&\sigma^2G^*B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1} = \sigma^2\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1}\\
&= \sigma^2G^*B\Sigma^{-1}\Gamma\sigma^2G^*B\Sigma^{-1},
\end{aligned}
$$

this proves the statement in (ii).

For the statement in (iii), we have

$$
\Gamma\sigma^2G^*B\Sigma^{-1}\Gamma = \frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B\sigma^2\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B,
$$

This gives

$$\Gamma\sigma^2 G^* B\Sigma^{-1}\Gamma = \frac{1}{\sigma^2} B^\top (B\Sigma^{-1}B^\top)^{-1} B = \Gamma,$$

which implies that

$$(G^* r_0)^\top \Gamma\sigma^2 G^* B\Sigma^{-1}\Gamma G^* r_0 = (G^* r_0)^\top \Gamma G^* r_0,$$

this proves the statement in (iii).

For the statement in (iv), since we have $\Gamma\sigma^2 G^* B\Sigma^{-1}\Gamma = \Gamma$,

$$(G^* r_0)^\top (\Gamma\sigma^2 G^* B\Sigma^{-1})^2 = (G^* r_0)^\top \Gamma\sigma^2 G^* B\Sigma^{-1},$$

this proves the statement in (iv).

Similarly, in the case of $r_0 = 0$, we have

$$\psi_T(\hat{m}) \xrightarrow[T\to\infty]{D} \psi_0(m) = \xi_0^\top \Gamma \xi_0 \sim \chi_q^2,$$

this completes the proof. □

Then, let $\chi^2_{\alpha;q}$ be the αth-quantile of a χ^2_q where $0 < \alpha \leq 1$. From Corollary 6.1, we propose a test for the hypothesis testing problem in (2.3). We suggest

$$\kappa(\hat{\phi}, T) = \mathbb{I}_{\{\psi_T(\hat{m}) > \chi^2_{\alpha;q}\}}. \tag{6.3}$$

The following corollary shows that the test $\kappa(\hat{\phi}, T)$ is consistent.

Corollary 6.2. *Suppose that the conditions of Corollary 6.1 hold. Then, the asymptotic power function of the test in (6.3) is given by* $\Pi(\Delta) = \mathrm{P}\left(\chi_q^2(\Delta) \geq \chi^2_{\alpha;q}\right)$.

The proof follows directly from Corollary 6.1.

It is obvious that $r_0 = 0$ under the null hypothesis in (2.3). It implies that $\Delta = 0$. Then, by Corollary 6.2, the asymptotic power of the test is equal to α. Moreover, the asymptotic power tends to 1 as Δ tends to infinity.

6.2 A class of shrinkage estimators

Usually, the RE should dominate the UE if the restriction holds. In contrast, when the restriction is wrong, the UE is more efficient than the RE. As a comprising estimation method, we construct shrinkage estimators (SEs) by combining the RE and the UE in the optimal way. To this end, by following Nkurunziza (2012b), we consider the following class of shrinkage type estimators

$$\hat{\theta}^s(h) = \tilde{\theta}(\hat{\phi}, \hat{m}) + h(\|\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})\|_{\hat{\Gamma}})(\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})), \tag{6.4}$$

where $\|x\|_A = x^\top Ax$, h is continuous real-valued function on $(0, +\infty)$. It should be noticed that although (6.4) looks like some existing results in literature, this is not the case. Indeed, the dimensions of the random vectors in (6.4) are random, as they depend on $\hat{m}$. Because of that, the derivation of the asymptotic distributional risk does not follow from the results in literature. In particular, if $h(x) = (1 - \frac{q-2}{x}),\ x > 0$, we get the shrinkage estimator (SE) given by

$$\hat{\theta}^s = \tilde{\theta}(\hat{\phi}, \hat{m}) + [1 - (q-2)\psi_T(\hat{m})^{-1}](\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})), \tag{6.5}$$

where $2 < q = \text{rank}(B) < (m+1)(p+1)$, and $\psi_T(\hat{m})$ is given as in Corollary 6.1. Further, let $a^+ = \max\{0, a\}$, and let $h(x) = (1 - \frac{q-2}{x})^+,\ x > 0$. We get the positive-part shrinkage estimator (PSE) given by

$$\hat{\theta}^{s+} = \tilde{\theta}(\hat{\phi}, \hat{m}) + [1 - (q-2)\psi_T(\hat{m})^{-1}]^+(\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})). \tag{6.6}$$

Note that the proposed class of estimators includes also the UE and the RE by taking $h \equiv 1$ and $h \equiv 0$, respectively. Further, note that the SEs in (6.5) and (6.6) have the same form as that in Saleh (2006), Sen and Saleh (1987) among others. Nevertheless, the dimensions of $\hat{\theta}^s$ and $\hat{\theta}^{s+}$ are random variables. Thus, the derivation of the relative efficiency does not follow from the results in literature.

Chapter 7

Comparison between estimators

In this chapter, we derive the asymptotic distributional risk (ADR) functions of the proposed class of estimators as well as that of SEs, UE and RE. We also compare the performance of these estimators.

7.1 Asymptotic distributional risk

In this section, we derive the ADR functions based on Theorem 2.1-2.3 of Nkurunziza (2012) along with Proposition 5.1. Let Ω be the $(m+1)(p+1)\times(m+1)(p+1)$ positive symmetric semi-definite weighting matrix. First of all, we introduce the *quadratic loss function* in the form of

$$L\left(\hat{\theta}_0, \theta; \Gamma\right) = T(\hat{\theta}_0 - \theta)^\top \Omega(\hat{\theta}_0 - \theta), \tag{7.1}$$

where $\hat{\theta}_0$ represents an estimator such as $\hat{\theta}^s$, $\hat{\theta}^{s+}$, $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$.

The ADR of an estimator $\hat{\theta}_0$ is defined as

$$\text{ADR}\left(\hat{\theta}_0, \theta, \Omega\right) = \text{E}[\varepsilon^\top \Omega \varepsilon], \tag{7.2}$$

where ε is the random vector such that $T(\hat{\theta}_0 - \theta)^\top \Omega(\hat{\theta}_0 - \theta) \xrightarrow[T\to\infty]{d} \varepsilon^\top \Omega \varepsilon$.

Proposition 7.1. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). Then, the following two conditions hold.*

(i) $(\sigma^2 G^* B \Sigma^{-1})\Xi$ *is an idempotent matrix;*

(ii) $\Xi(\sigma^2 G^* B \Sigma^{-1})\Xi G^* r_0 = \Xi G^* r_0$.

Proof. To prove the statement in (i), we observe that

$$(\sigma^2 G^* B \Sigma^{-1})\Xi = (\sigma^2 G^* B \Sigma^{-1})\frac{1}{\sigma^2}\Sigma = G^* B.$$

Then, we have

$$(\sigma^2 G^* B \Sigma^{-1})\Xi(\sigma^2 G^* B \Sigma^{-1})\Xi = G^* B G^* B.$$

Note that, since $G^* = \Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1}$,

$$G^* B G^* B = \Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1} B \Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1} B$$
$$= \Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1} B = G^* B.$$

This gives

$$(\sigma^2 G^* B \Sigma^{-1})\Xi(\sigma^2 G^* B \Sigma^{-1})\Xi = (\sigma^2 G^* B \Sigma^{-1})\Xi,$$

this proves the statement in (i). To prove the statement in (ii), we have

$$\Xi(\sigma^2 G^* B \Sigma^{-1})\Xi G^* r_0 = \frac{1}{\sigma^2}\Sigma(\sigma^2 \Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1} B \Sigma^{-1})\frac{1}{\sigma^2}\Sigma\Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1} r_0$$
$$= \frac{1}{\sigma^2}\Sigma\Sigma^{-1} B^\top (B\Sigma^{-1}B^\top)^{-1} r_0 = \Xi G^* r_0,$$

this proves the statement in (ii). □

Let $\Lambda_{22} = \Sigma^{-1} - G^* B \Sigma^{-1}$.

Theorem 7.1. *Suppose that Assumptions 1-3 hold. Then,*

$$\mathrm{ADR}(\hat{\theta}^s(h), \theta, \Omega) = \sigma^2 \mathrm{trace}(\Omega\Lambda_{22}) + r_0^\top G^{*\top}\Omega G^* r_0 - 2\mathrm{E}[h(\chi^2_{q+2}(\Delta))] r_0^\top G^{*\top}\Omega G^* r_0$$
$$+ \sigma^2 \mathrm{E}[h^2(\chi^2_{q+2}(\Delta))]\mathrm{trace}(\Omega(\Sigma^{-1} - \Lambda_{22})) + \mathrm{E}[h^2(\chi^2_{q+4}(\Delta))] r_0^\top G^{*\top}\Omega G^* r_0.$$

The proof follows from Proposition 5.1 and Theorems 2.1-2.3 of Nkurunziza (2012b) by taking $L_1 \equiv B$, $L_2 \equiv 1$, $\Xi_1 \equiv \frac{1}{\sigma^2} B^\top (B\Sigma^{-1}B^\top)^{-1}B$, $\delta \equiv G^* r_0$, $\Sigma^* \equiv \sigma^2(\Sigma^{-1} - \Lambda_{22})$, $p \equiv 1$. By using Theorem 7.1, we derive the ADR functions of UE and RE in Theorem 7.2 and Theorem 7.3.

Theorem 7.2. *Suppose that Assumptions 1-3 hold. Then, the* ADR *of the UE* $\hat{\theta}(\hat{\phi}, \hat{m})$ *is given by* $\mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \sigma^2 \mathrm{trace}(\Omega\Sigma^{-1})$.

The proof follows from Theorem 7.1 by taking $h = 1$. We also give an alternative proof in the Appendix B.

Theorem 7.3. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). Then, the* ADR *of the RE* $\tilde{\theta}(\hat{\phi}, \hat{m})$ *is*

$$\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) + r_0^\top G^{*\top} \Omega G^* r_0.$$

The proof follows from Theorem 7.1 by taking $h = 0$. We also give an alternative proof in the Appendix B.

In this section, the ADR function of the shrinkage estimator is derived based on the Theorem 3.1 in Nkurunziza (2012b) (see also Theorem A.8 in Appendix A). First, we prove the conditions in Theorem 2.2 and 2.3 in Nkurunziza (2012b) (see also Theorem A.6 and Theorem A.7 in Appendix A).

By Proposition 5.1, we have $g(\zeta_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\zeta, \xi)$, with

$$\begin{pmatrix} \xi \\ \zeta \end{pmatrix} \sim \mathcal{N}_{2(m+1)(p+1)}\left(\begin{pmatrix} G^* r_0 \\ -G^* r_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} G^* B \Sigma^{-1} & 0 \\ 0 & \Sigma^{-1} - G^* B \Sigma^{-1} \end{pmatrix}\right).$$

Now, we define $\Xi = \dfrac{1}{\sigma^2}\Sigma$. By the way, by Proposition 3.5, one can prove that it is positive definite. Now, the ADR of the shrinkage estimator is shown in the following theorem. Before that, we recall that $\Delta = \dfrac{1}{\sigma^2} r_0^\top (B\Sigma^{-1}B^\top)^{-1} r_0$.

Theorem 7.4. *Suppose that the conditions of Proposition 7.1 hold. Then, the* ADR *of the shrinkage estimator* $\hat{\theta}^s$ *is*

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) &= \mathrm{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right) + (q+2)(q-2)r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] \\ &\quad - (q-2)\sigma^2 \mathrm{trace}(\Omega G^* B\Sigma^{-1})(2\mathrm{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]).\end{aligned}$$

Proof. Let $L_1 = B$, $L_2 = 1$, $k = 1$, $\Xi_1 = \frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B$, $\delta = G^* r_0$ and $\Sigma^* = \sigma^2 G^* B\Sigma^{-1}$. Since the conditions (i) and (ii) in Proposition 7.1 hold, we can apply Theorem 3.1 in Nkurunziza (2012b) by taking the measurable function $h(x) = \left[1 - \frac{q-2}{x}\right],\ x > 0$. We have

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) &= \mathrm{trace}\left[\Omega(\sigma^2\Sigma^{-1} - \sigma^2 G^* B\Sigma^{-1})\right] + r_0^\top G^{*\top}\Omega G^* r_0 \\ &\quad - 2r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[1 - (q-2)\chi_{q+2}^{-2}(\Delta)] \\ &\quad + \sigma^2 \mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[(1 - (q-2)\chi_{q+2}^{-2}(\Delta))^2] \\ &\quad + r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[(1 - (q-2)\chi_{q+4}^{-2}(\Delta))^2].\end{aligned}$$

Then, we have

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) &= \sigma^2\mathrm{trace}(\Omega\Sigma^{-1}) - \sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1}) + r_0^\top G^{*\top}\Omega G^* r_0 \\ &\quad - 2r_0^\top G^{*\top}\Omega G^* r_0 + 2(q-2)r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[\chi_{q+2}^{-2}(\Delta)] \\ &\quad + \sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1}) - 2(q-2)\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[\chi_{q+2}^{-2}(\Delta)] \\ &\quad + (q-2)^2\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[\chi_{q+2}^{-4}(\Delta)] \\ &\quad + r_0^\top G^{*\top}\Omega G^* r_0 - 2(q-2)r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[\chi_{q+4}^{-2}(\Delta)] \\ &\quad + (q-2)^2 r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[\chi_{q+4}^{-4}(\Delta)].\end{aligned}$$

This gives that

$$\begin{aligned}\text{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) &= \sigma^2\text{trace}(\Omega\Sigma^{-1}) + 2(q-2)r_0^\top G^{*\top}\Omega G^* r_0(\text{E}[\chi_{q+2}^{-2}(\Delta)] - \text{E}[\chi_{q+4}^{-2}(\Delta)])\\ &- (q-2)\sigma^2\text{trace}(\Omega G^* B\Sigma^{-1})(2\text{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\text{E}[\chi_{q+2}^{-4}(\Delta)])\\ &+ (q-2)^2 r_0^\top G^{*\top}\Omega G^* r_0 \text{E}[\chi_{q+4}^{-4}(\Delta)].\end{aligned}$$

Since

$$\text{E}[\chi_{q+2}^{-2}(\Delta)] - \text{E}[\chi_{q+4}^{-2}(\Delta)] = 2\text{E}[\chi_{q+4}^{-4}(\Delta)],$$

then,

$$\begin{aligned}\text{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) &= \text{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right) + (q+2)(q-2)r_0^\top G^{*\top}\Omega G^* r_0\text{E}[\chi_{q+4}^{-4}(\Delta)]\\ &- (q-2)\sigma^2\text{trace}(\Omega G^* B\Sigma^{-1})(2\text{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\text{E}[\chi_{q+2}^{-4}(\Delta)]),\end{aligned}$$

this completes the proof. □

The following theorem shows the ADR function of the positive-part shrinkage estimator.

Theorem 7.5. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). Then, the* ADR *of the positive-part shrinkage estimator* $\hat{\theta}^{s+}$ *is*

$$\begin{aligned}\text{ADR}\left(\hat{\theta}^{s+},\theta,\Omega\right) &= \text{ADR}\left(\hat{\theta}^s,\theta,\Omega\right)\\ &+ 2r_0^\top G^{*\top}\Omega G^* r_0\text{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}]\\ &- \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1})\text{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}]\\ &- r_0^\top G^{*\top}\Omega G^* r_0\text{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)<q-2\}}].\end{aligned}$$

Proof. To begin this proof, we redefine the measurable function

$$h(x) = \left[1 - \frac{q-2}{x}\right]\mathbb{I}_{\{x\geq q-2\}},\ x>0.$$

Then, by Theorem 3.1 in Nkurunziza (2012b), we have

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right) &= \mathrm{trace}\left[\Omega(\sigma^2\Sigma^{-1} - \sigma^2 G^* B \Sigma^{-1})\right] + r_0^\top G^{*\top}\Omega G^* r_0 \\ &- 2r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)\geq q-2\}}] \\ &+ \sigma^2 \mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)\geq q-2\}}] \\ &+ r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)\geq q-2\}}].\end{aligned}$$

Note that

$$\begin{aligned}\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)\geq q-2\}}] &= \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))] \\ &- \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)< q-2\}}],\end{aligned}$$

$$\begin{aligned}\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)\geq q-2\}}] &= \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2] \\ &- \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)< q-2\}}],\end{aligned}$$

$$\begin{aligned}\mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)\geq q-2\}}] &= \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2] \\ &- \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)< q-2\}}].\end{aligned}$$

This gives

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right) &= \mathrm{trace}\left[\Omega(\sigma^2\Sigma^{-1} - \sigma^2 G^* B \Sigma^{-1})\right] + r_0^\top G^{*\top}\Omega G^* r_0 \\ &- 2r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[1-(q-2)\chi_{q+2}^{-2}(\Delta)] \\ &+ \sigma^2 \mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2] \\ &+ r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2] \\ &+ 2r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)< q-2\}}] \\ &- \sigma^2 \mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)< q-2\}}] \\ &- r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)< q-2\}}],\end{aligned}$$

which implies that

$$\begin{aligned}\text{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right) &= \text{ADR}\left(\hat{\theta}^{s}, \theta, \Omega\right) \\ &+ 2r_0^\top G^{*\top}\Omega G^* r_0 \text{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}] \\ &- \sigma^2 trace(\Omega G^* B \Sigma^{-1})\text{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}] \\ &- r_0^\top G^{*\top}\Omega G^* r_0 \text{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)<q-2\}}].\end{aligned}$$

This completes the proof. □

7.2 Risk Analysis

In this section, by using the ADR function derived in Section 7.1, we compare the performance of the proposed estimators. To this end, let λ_1 denote the smallest eigenvalue of the matrix $\left[(G^{*\top}\Gamma G^*)^{-1}G^{*\top}\Omega G^*\right]$ and let λ_n denote the largest eigenvalue of it. First, we compare the relative efficiency of the UE and the RE by the following proposition.

Proposition 7.2. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). If $\Delta \le \left(\sigma^2\text{trace}(\Omega G^* B\Sigma^{-1})\right)/\lambda_n$, then the $\tilde{\theta}(\hat{\phi}, \hat{m})$ dominates the $\hat{\theta}(\hat{\phi}, \hat{m})$, and if $\Delta \ge \left(\sigma^2\text{trace}(\Omega G^* B\Sigma^{-1})\right)/\lambda_1$, then the $\hat{\theta}(\hat{\phi}, \hat{m})$ dominates the $\tilde{\theta}(\hat{\phi}, \hat{m})$.*

Proof. By Theorem 7.3, we have

$$\text{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1}) + r_0^\top G^{*\top}\Omega G^* r_0,$$

which implies that

$$\text{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = r_0^\top G^{*\top}\Omega G^* r_0 - \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1}).$$

We observe that, since $G^* = \Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}$ and $\Gamma = \frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B$,

$$G^{*\top}\Gamma G^* = (B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}\frac{1}{\sigma^2}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1} = \frac{1}{\sigma^2}(B\Sigma^{-1}B^\top)^{-1},$$

which is positive definite for $\sigma > 0$.

Then, by Theorem 2.4.7 in Mathai and Provost (1992), we have

$$\lambda_1 \leq \frac{r_0^\top G^{*\top}\Omega G^* r_0}{r_0^\top G^{*\top}\Gamma G^* r_0} \leq \lambda_n.$$

Note that, since $\Delta = \dfrac{1}{\sigma^2}r_0^\top(B\Sigma^{-1}B^\top)^{-1}r_0$, $(B\Sigma^{-1}B^\top)^{-1}$ is positive definite, $\Delta = \dfrac{1}{\sigma^2}r_0^\top(B\Sigma^{-1}B^\top)^{-1}r_0 \geq 0$.

If $\Delta > 0$, from the proof of Corollary 6.1, we have $\Delta = r_0^\top G^{*\top}\Gamma G^* r_0$, then

$$\lambda_1\Delta - \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1}) \leq \text{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right), \quad (7.3)$$

and

$$\text{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) \leq \lambda_n\Delta - \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1}). \quad (7.4)$$

By (7.3), if

$$\lambda_1\Delta - \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1}) \geq 0,$$

then, $\text{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) \geq \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right)$.

Similarly, by (7.4), if

$$\lambda_n\Delta - \sigma^2\text{trace}(\Omega G^* B\Sigma^{-1}) \leq 0,$$

then, $\text{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) \leq \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right)$. This completes the proof. □

Next, we present the following proposition to show the dominance between $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\hat{\theta}^s$.

Proposition 7.3. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). If* $\sigma^2\text{trace}(\Omega G^* B\Sigma^{-1})/\lambda_n \geq (q+2)/2$ *with* $q \in (2, (m+1)(p+1))$*, then the shrinkage estimator* $\hat{\theta}^s$ *dominates the UE* $\hat{\theta}(\hat{\phi}, \hat{m})$*.*

Proof. By Theorem 7.4, we have

$$\text{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) = \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) + (q+2)(q-2) r_0^\top G^{*\top} \Omega G^* r_0 \text{E}[\chi_{q+4}^{-4}(\Delta)]$$
$$- (q-2)\sigma^2 \text{trace}(\Omega G^* B \Sigma^{-1})(2\text{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\text{E}[\chi_{q+2}^{-4}(\Delta)]).$$

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

$$\Delta \text{E}[\chi_{q+4}^{-4}(\Delta)] = \text{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\text{E}[\chi_{q+2}^{-4}(\Delta)],$$

this gives

$$\text{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) = \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) + (q+2)(q-2) r_0^\top G^{*\top} \Omega G^* r_0 \text{E}[\chi_{q+4}^{-4}(\Delta)]$$
$$- (q-2)\sigma^2 \text{trace}(\Omega G^* B \Sigma^{-1})(2\Delta \text{E}[\chi_{q+4}^{-4}(\Delta)] + (q-2)\text{E}[\chi_{q+2}^{-4}(\Delta)]).$$

Note that, since $\Delta = \dfrac{1}{\sigma^2} r_0^\top (B\Sigma^{-1}B^\top)^{-1} r_0$, $(B\Sigma^{-1}B^\top)^{-1}$ is positive definite and, $\Delta = \dfrac{1}{\sigma^2} r_0^\top (B\Sigma^{-1}B^\top)^{-1} r_0 \geq 0$. Then, $\Delta = 0$ if an only if $r_0 = 0$.
If $\Delta = 0$, we have

$$\text{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) = \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - (q-2)^2 \sigma^2 \text{trace}(\Omega G^* B \Sigma^{-1}) \text{E}[\chi_{q+2}^{-4}],$$

this gives

$$\text{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) - \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = -(q-2)^2 \sigma^2 \text{trace}(\Omega G^* B \Sigma^{-1}) \text{E}[\chi_{q+2}^{-4}].$$

Since $\Omega G^* B \Sigma^{-1}$ is positive definite, we have

$$\text{trace}(\Omega G^* B \Sigma^{-1}) \geq 0. \tag{7.5}$$

Further, since χ_{q+2}^{-4} is a non-negative random variable, by (7.5), we have

$$\text{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) - \text{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = -(q-2)^2 \sigma^2 \text{trace}(\Omega G^* B \Sigma^{-1}) \text{E}[\chi_{q+2}^{-4}] \leq 0,$$

which implies that $\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) \leq \mathrm{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right)$. If $\Delta > 0$, we have

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) &= \mathrm{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right) + (q+2)(q-2)r_0^\top G^{*\top}\Omega G^* r_0 \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] \\ &\quad - (q-2)\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})(2\Delta\mathrm{E}[\chi_{q+4}^{-4}(\Delta)] + (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]),\end{aligned}$$

then,

$$\begin{aligned}&\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) - \mathrm{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right) = -(q-2)\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1}) \\ &\times \left[2\Delta\mathrm{E}[\chi_{q+4}^{-4}(\Delta)]\left(1 - \frac{(q+2)r_0^\top G^{*\top}\Omega G^* r_0}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})}\right) + (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]\right].\end{aligned}$$

Now, let

$$H = \left(1 - \frac{(q+2)r_0^\top G^{*\top}\Omega G^* r_0}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})}\right),$$

we have

$$\begin{aligned}&\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) - \mathrm{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right) \\ &= -(q-2)\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})\left[2\Delta\mathrm{E}[\chi_{q+4}^{-4}(\Delta)]H + (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]\right].\end{aligned}$$

From (7.5) with the fact that $\chi_{q+2}^{-4}(\Delta)$ and $\chi_{q+4}^{-4}(\Delta)$ are non-negative random variables, we have, $\mathrm{ADR}\left(\hat{\theta}^s,\theta,\Omega\right) - \mathrm{ADR}\left(\hat{\theta}(\hat{\phi},\hat{m}),\theta,\Omega\right) \leq 0$ for all $\Delta > 0$ given that $q > 2$ and

$$H = 1 - \frac{(q+2)r_0^\top G^{*\top}\Omega G^* r_0}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})} \geq 0. \tag{7.6}$$

Then, by Theorem 2.4.7 in Mathai and Provost (1992), we have

$$\lambda_1 \leq \frac{r_0^\top G^{*\top}\Omega G^* r_0}{r_0^\top G^{*\top}\Gamma G^* r_0} \leq \lambda_n.$$

This gives, from the proof of Corollary 6.1, we know that $\Delta = r_0^\top G^{*\top}\Gamma G^* r_0$,

$$1 - \frac{(q+2)\lambda_n\Delta}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})} \leq 1 - \frac{(q+2)r_0^\top G^{*\top}\Omega G^* r_0}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})},$$

and

$$1-\frac{(q+2)\lambda_1\Delta}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})}\geq 1-\frac{(q+2)r_0^\top G^{*\top}\Omega G^* r_0}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})}.$$

Then, the relation in (7.6) holds if

$$1-\frac{(q+2)\lambda_n\Delta}{2\Delta\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})}\geq 0,$$

and this holds if and only if

$$\frac{\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})}{\lambda_n}\geq\frac{q+2}{2},$$

this completes the proof. □

Finally, we compare the relative performance between $\hat{\theta}^s$ and $\hat{\theta}^{s+}$ in the proposition below.

Proposition 7.4. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.8). Then, the positive-part shrinkage estimator $\hat{\theta}^{s+}$ dominates the shrinkage estimator $\hat{\theta}^s$.*

Proof. By Theorem 7.5

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}^{s+},\theta,\Omega\right)&=\mathrm{ADR}\left(\hat{\theta}^{s},\theta,\Omega\right)\\&+2r_0^\top G^{*\top}\Omega G^* r_0\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}]\\&-\sigma^2\mathrm{trace}(\Omega G^* B\Sigma^{-1})\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}]\\&-r_0^\top G^{*\top}\Omega G^* r_0\mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^2(\Delta)<q-2\}}].\end{aligned}$$

This gives

$$\begin{aligned}
&\mathrm{ADR}\left(\hat{\theta}^{s}, \theta, \Omega\right) - \mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right) \\
&= -2 r_0^{\top} G^{*\top} \Omega G^{*} r_0 \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}}] \\
&+ \sigma^2 \mathrm{trace}(\Omega G^{*} B \Sigma^{-1}) \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}}] \\
&+ r_0^{\top} G^{*\top} \Omega G^{*} r_0 \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^{2}(\Delta)<q-2\}}].
\end{aligned}$$

Note that Ω is positive semi-definite matrix, $r_0^{\top} G^{*\top} \Omega G^{*} r_0$ is a non-negative real number. Also, we observe that $(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}} < 0$, then $\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}}] < 0$. Hence,

$$-2 r_0^{\top} G^{*\top} \Omega G^{*} r_0 \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}}] \geq 0. \tag{7.7}$$

Since $\Omega G^{*} B \Sigma^{-1}$ is positive definite, $\mathrm{trace}(\Omega G^{*} B \Sigma^{-1})$ is a non-negative real number. Also, we observe that $(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}} \geq 0$, then $\mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}}] \geq 0$. Hence,

$$\sigma^2 \mathrm{trace}(\Omega G^{*} B \Sigma^{-1}) \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+2}^{2}(\Delta)<q-2\}}] \geq 0. \tag{7.8}$$

We observe that $(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^{2}(\Delta)<q-2\}} \geq 0$, then $\mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^{2}(\Delta)<q-2\}}] \geq 0$. Hence, since $r_0^{\top} G^{*\top} \Omega G^{*} r_0$ is a non-negative real number, we have

$$r_0^{\top} G^{*\top} \Omega G^{*} r_0 \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2\mathbb{I}_{\{\chi_{q+4}^{2}(\Delta)<q-2\}}] \geq 0. \tag{7.9}$$

Therefore, by (7.7)-(7.9), we establish that

$$\mathrm{ADR}\left(\hat{\theta}^{s}, \theta, \Omega\right) - \mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right) \geq 0,$$

this gives

$$\mathrm{ADR}\left(\hat{\theta}^{s}, \theta, \Omega\right) \geq \mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right),$$

this completes the proof. □

Chapter 8

Numerical study

In previous chapter, we derived the UE, RE, SE and PSE. In this chapter, we estimate the number of change points by using the algorithm introduced in Section 5.2. We also estimate the positions of change points, and compare the relative performance of these estimators.

8.1 Simulation study

We illustrate the performance of the proposed method by using the simulation studies. We use Monte-Carlo simulation to generate the generalized O-U process. Two cases are reported here: 1. The case of two change points; 2. The case of three change points. For both cases, we generate the O-U process with a periodic two-dimensional incomplete set of basis functions $\left\{1, \sqrt{2}\cos\left(\frac{2\pi t}{\Delta}\right)\right\}$ where $\Delta = t_{i+1} - t_i$ is the time increment in time period $[0, T]$. Thus, the process is given as

$$dX_t = \sum_{j=1}^{m} \left(\mu_{1,j} + \mu_{2,j}\sqrt{2}\cos\left(\frac{2\pi t}{\Delta_t}\right) - \alpha_j X_t\right) \mathbb{I}_{(\tau_{j-1},\tau_j)} dt + \sigma dW_t \tag{8.1}$$

where $j = 1, \ldots, m$ (m is the number of change points), $\phi_{j-1}T < t \leq \phi_j T$ and $X_0 = 0.05$. To simplify, we take $\sigma = 1$. In each case, 500 iterations are performed. In each iteration, the positions of change points and the number of change points are estimated. Moreover, we take $\Omega = \mathrm{I}_{(p+1)(m+1)\times(p+1)(m+1)}$, and we also compare the relative performance of estimators via empirical ADR. To estimate σ^2, we use $\hat{\sigma}^2 = \frac{1}{T}\sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2$.

8.2 Performance comparison

First, we consider the case of two change points so that we let $m = 2$, with $\phi_1 = 0.35$ and $\phi_2 = 0.7$. In order to evaluate the effect of time period T, we generate the O-U process with $T = 20$ and $T = 50$, with the time increment of $\Delta = 0.001$. Table 8.1 shows the value of coefficients which are used to generate the process. To set a linear restriction, we take the matrix B which is given as

$$B = [(I_3, 0)^\top, (-I_3, I_3)^\top, (0, -I_3)^\top]. \tag{8.2}$$

Table 8.1: Two change points ($\phi_1 = 0.35, \phi_2 = 0.7$)

coefficient	$j = 1$	$j = 2$	$j = 3$
$\mu_{1,j}$	10	5	15
$\mu_{2,j}$	5	2	8
α_j	3	1	4

We also consider the case of three change points. Let $\phi_1 = 0.25$, $\phi_2 = 0.5$ and $\phi_3 = 0.75$. The value of coefficients is given in the Table 8.2. In this case, we choose the linear restriction as

$$B = [(I_3, 0, 0)^\top, (-I_3, I_3, 0)^\top, (0, -I_3, I_3)^\top, (0, 0, -I_3)^\top]. \tag{8.3}$$

Table 8.2: Three change points ($\phi_1 = 0.25, \phi_2 = 0.5, \phi_3 = 0.75$)

coefficient	$j=1$	$j=2$	$j=3$	$j=4$
$\mu_{1,j}$	10	5	15	20
$\mu_{2,j}$	5	2	7	10
α_j	3	1	3	5

For the two cases considered, we estimate the number of change points based on the algorithm in Section 5.2. To estimate the number of change points, we take $m_{max} = 6$. From 500 iterations, the cumulative frequency (CF) and the relative frequency (RF) are shown in Table 8.3. The CF and RF are defined as CF $= \sum_{i=1}^{500} \mathbb{I}_{(\hat{m}_i = m)}$ and RF $= \frac{1}{500} \sum_{i=1}^{500} \mathbb{I}_{(\hat{m}_i = m)} \times 100\%$.

Table 8.3: Cumulative frequency and relative frequency of 500 iterations

	$T=20$	$T=20$	$T=50$	$T=50$
case	CF	RF	CF	RF
$m=2$	497	99.4%	500	100%
$m=3$	492	98.4%	500	100%

From Table 8.3, the cumulative frequency and relative frequency become larger when we change T from 20 to 50. Thus, it seems accurate to estimate the number of change points when T is large.

From 500 iterations, we also estimate the locations of change points based on LSSE method in (4.2). The mean of these locations are recorded in Table 8.4 and Table 8.5.

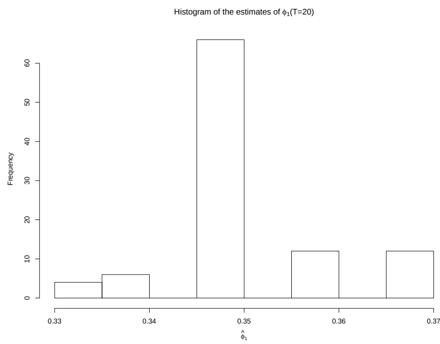
Table 8.4: Mean of estimates of ϕ_1, ϕ_2 $(m = 2)$

	$T = 20$	$T = 50$
$\hat{\phi}_1$	0.3522	0.3492
$\hat{\phi}_2$	0.6996	0.7

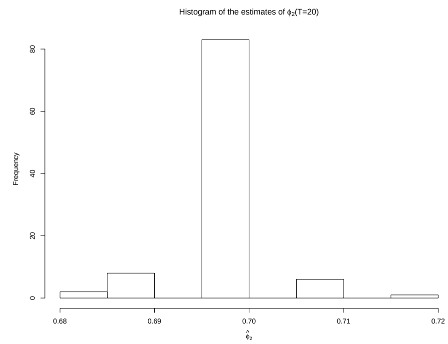
Table 8.5: Mean of estimates of ϕ_1, ϕ_2, ϕ_3 $(m = 3)$

	$T = 20$	$T = 50$
$\hat{\phi}_1$	0.2519	0.2501
$\hat{\phi}_2$	0.4995	0.5002
$\hat{\phi}_3$	0.7497	0.7502

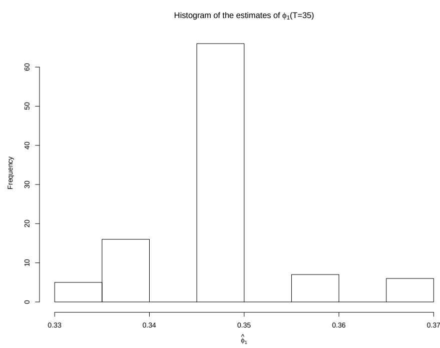
From Tables 8.4 and 8.5, it is obvious that, as T becomes large, the estimated locations of change points are closer to the pre-assigned values. In other words, the method is more accurate as T increases. Further, under the case of two change points, we estimated $\hat{\phi}_1$ and $\hat{\phi}_2$ in 100 replicates as $T = 20, 35, 50$ where $\phi = (0.35, 0.7)$. In Figure 8.1, all the histograms are quite symmetric and unimodal with the mode which corresponds to the exact value. As T increases, the estimates become closer to the pre-assigned values. In the case of three change points, we also estimated $\hat{\phi}$ as $T = 20, 35, 50$ in 100 replicates where $\phi = 0.25, 0.5, 0.7$. From Figure 8.2, we observe the similar results as in the case of two change points.



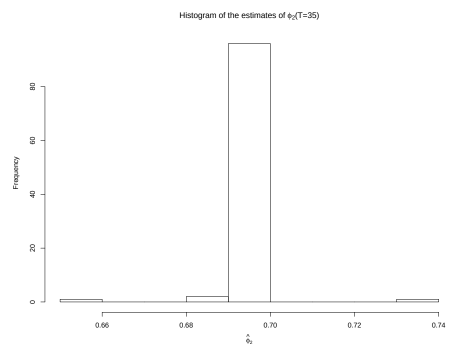
(a) Histogram of $\hat{\phi}_1$ $(T = 20)$



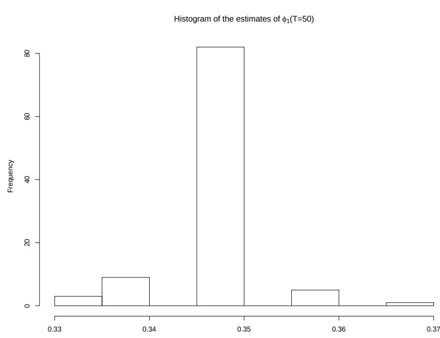
(b) Histogram of $\hat{\phi}_2$ $(T = 20)$



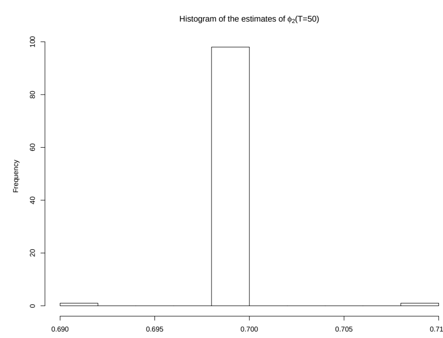
(c) Histogram of $\hat{\phi}_1$ $(T = 35)$



(d) Histogram of $\hat{\phi}_2$ $(T = 35)$

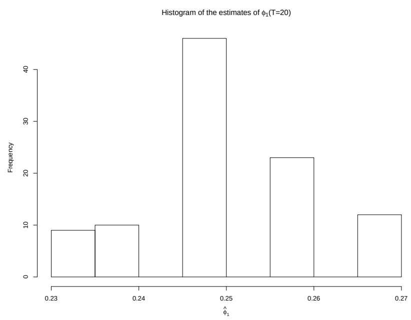


(e) Histogram of $\hat{\phi}_1$ $(T = 50)$

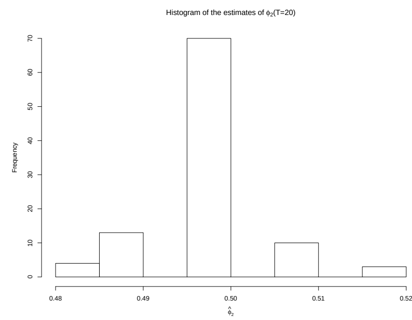


(f) Histogram of $\hat{\phi}_2$ $(T = 50)$

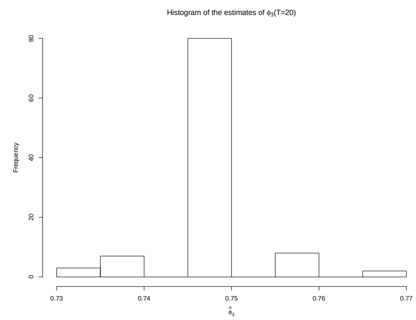
Figure 8.1: Histogram of estimates of $\hat{\phi}$, $m = 2$, $T = (20, 35, 50)$, $\phi = (0.35, 0.7)$



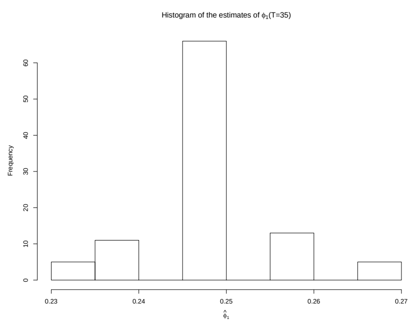
(a) Histogram of $\hat{\phi}_1$ $(T = 20)$



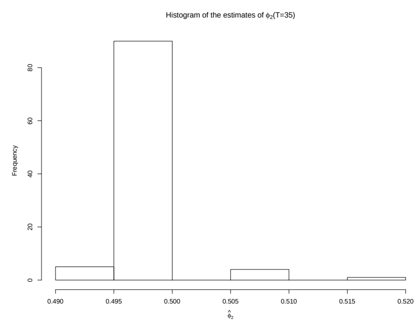
(b) Histogram of $\hat{\phi}_2$ $(T = 20)$



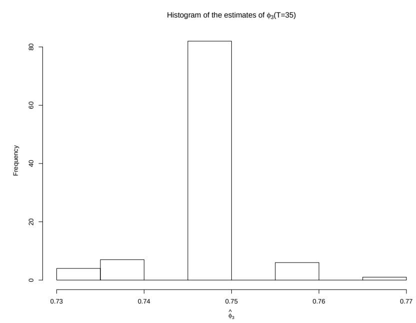
(c) Histogram of $\hat{\phi}_3$ $(T = 20)$



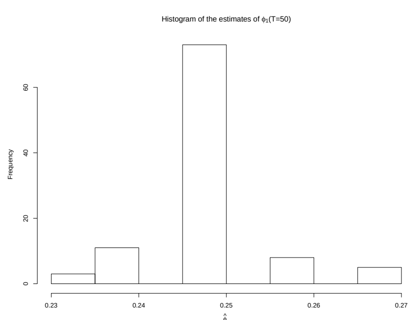
(d) Histogram of $\hat{\phi}_1$ $(T = 35)$



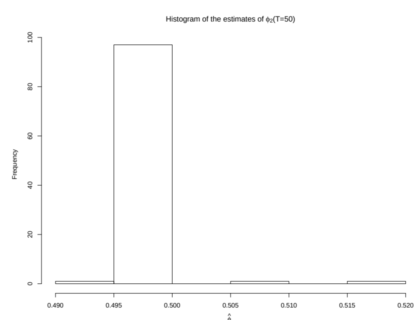
(e) Histogram of $\hat{\phi}_2$ $(T = 35)$



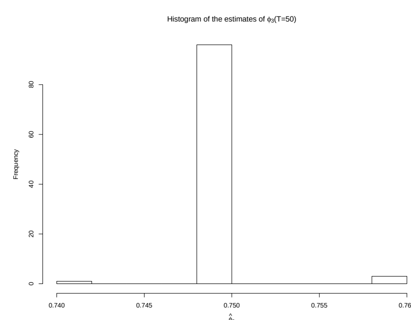
(f) Histogram of $\hat{\phi}_3$ $(T = 35)$



(g) Histogram of $\hat{\phi}_1$ $(T = 50)$



(h) Histogram of $\hat{\phi}_2$ $(T = 50)$



(i) Histogram of $\hat{\phi}_3$ $(T = 50)$

Figure 8.2: Histogram of estimates of $\hat{\phi}$, $m = 3$, $T = (20, 35, 50)$, $\phi = (0.25, 0.5, 0.75)$

As in Nkurunziza and Zhang (2018), we compute the relative mean squared efficiency (RMSE) by

$$\mathrm{RMSE}(\hat{\theta}_0) = \mathrm{ADR}(\hat{\theta}(\hat{\phi}, \hat{m}), \theta; \Omega)/\mathrm{ADR}(\hat{\theta}_0, \theta; \Omega) \tag{8.4}$$

where $\hat{\theta}_0$ represents an estimator such as $\hat{\theta}^s$, $\hat{\theta}^{s+}$, $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$. We compute Δ by using $\Delta = \frac{1}{\sigma^2} r_0^\top (B\Sigma^{-1}B^\top)^{-1} r_0$. We take $r_0 = 0.5nr,\ n = 1, 2, 3, 4, 5, 6$.

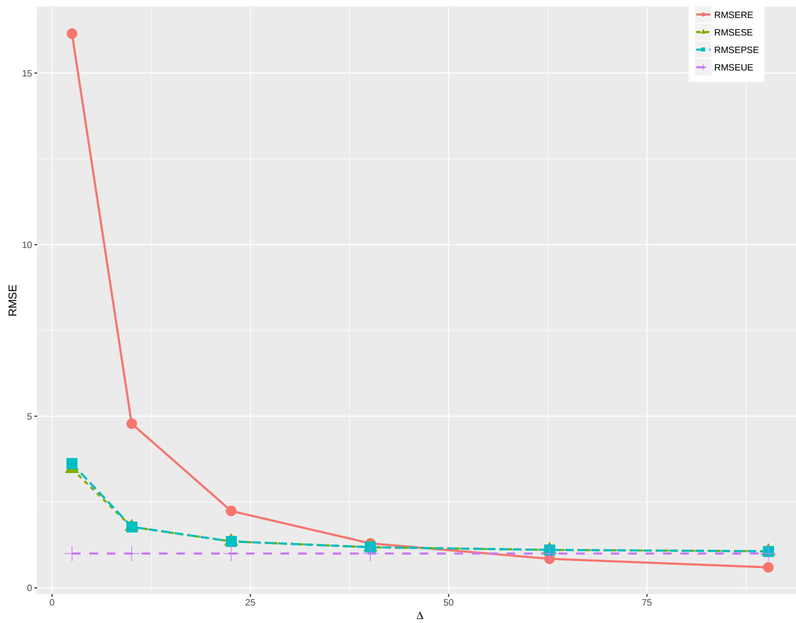


Figure 8.3: RMSE of UE, RE, SE, PSE versus Δ ($m = 2$, $T = 20$)

For the two change point case, from Figures 8.3 and 8.4, near $\Delta = 0$, RMSE of RE is higher than the RMSE of UE, RMSE of SE and RMSE of PSE. It means that, near the restriction, RE is more efficient than other three estimators. These figures also show that the efficiency of RE decreases as one moves far away from the null hypothesis. Further, PSE and SE outperform than UE, and PSE is more efficient than SE.

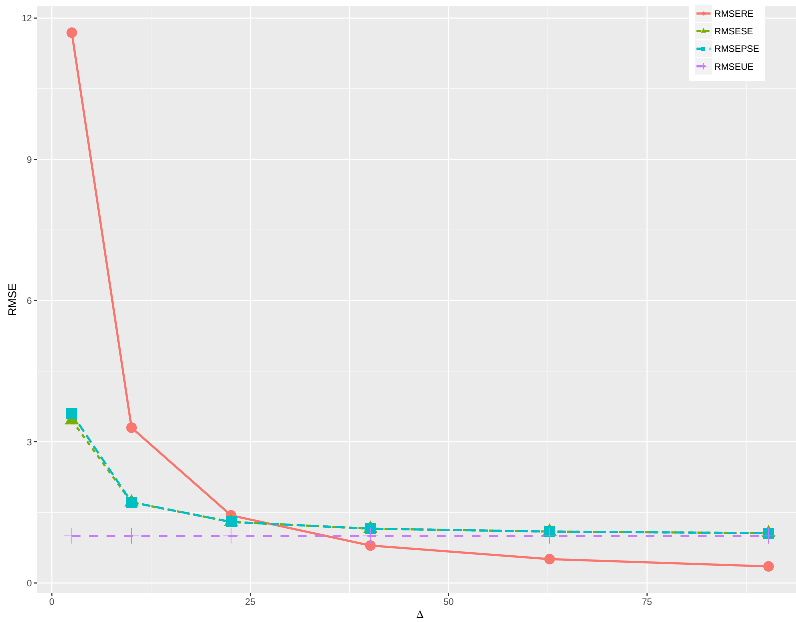


Figure 8.4: RMSE of UE, RE, SE, PSE versus Δ ($m = 2$, $T = 50$)

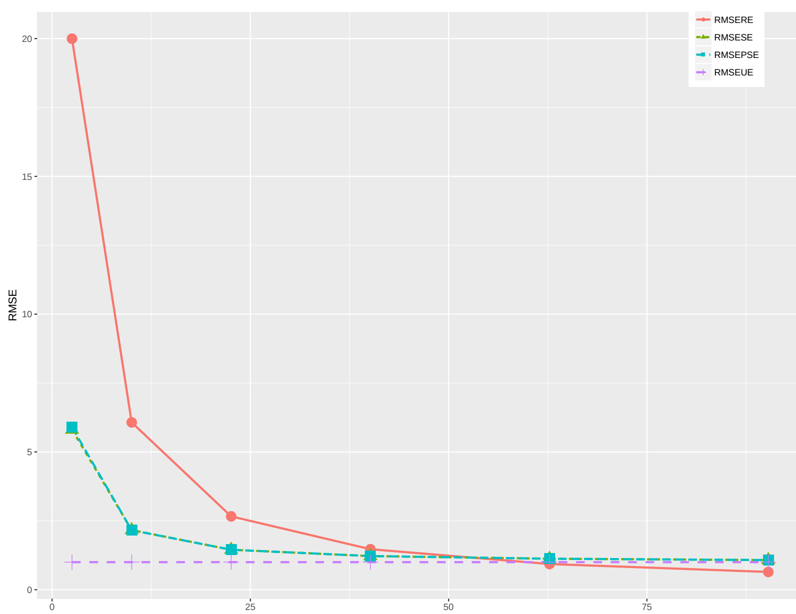


Figure 8.5: RMSE of UMLE, RMLE, SE, PSE versus Δ ($m = 3$, $T = 20$)

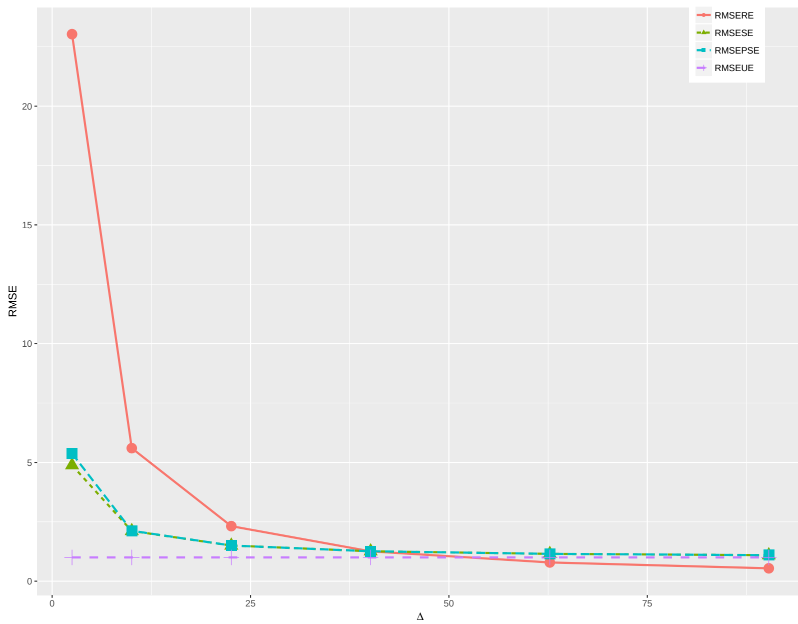


Figure 8.6: RMSE of UMLE, RMLE, SE, PSE versus Δ ($m = 3$, $T = 50$)

For the three change points case, from Figures 8.5 and 8.6, near H_0, the performance of the RE is better than that of UE, SE and PSE. However, as Δ increases, RE performs worse. Both figures show that although the efficiency of SE and PSE decreases as Δ increases, they are more efficient than UE. Also, PSE is more efficient than SE. In conclusion, the numerical results of both cases are in agreement with the theoretical results established in Section 7.2.

From Figures 8.7-8.12, it is obvious that the empirical power tends to 1 as Δ increases to infinity. Also, as T inceases, the empirical power also increases. It means that the numerical results coincide with the theoretical results which show that the test in (6.3) is consistent.

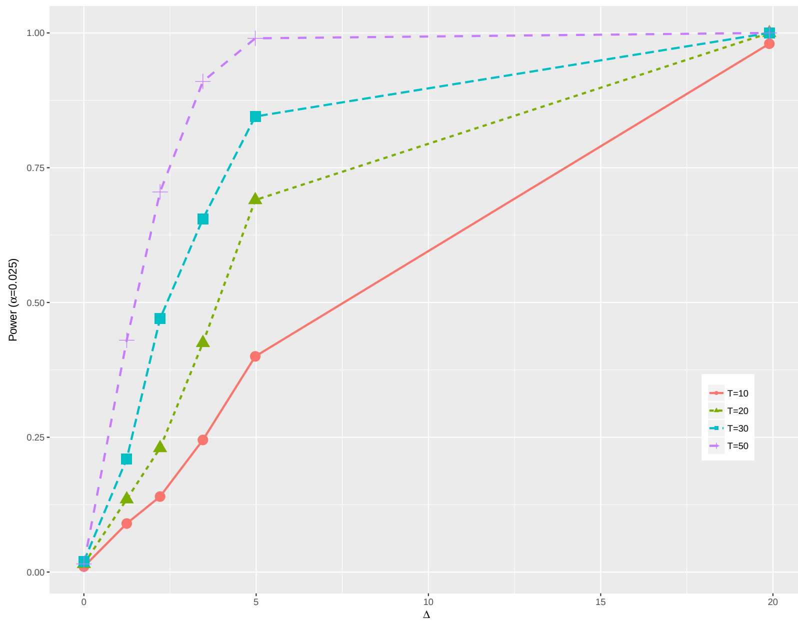


Figure 8.7: The empirical power of the test versus Δ and T ($m = 2$)

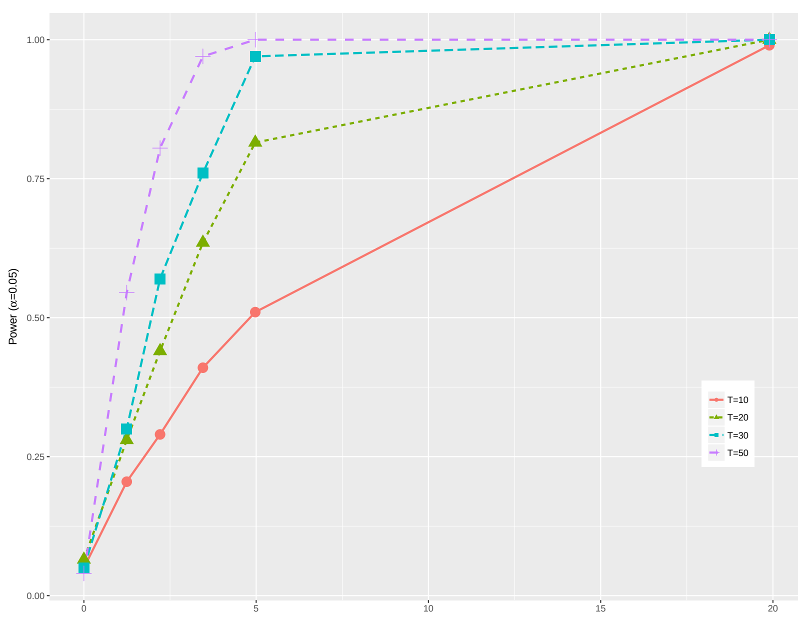


Figure 8.8: The empirical power of the test versus Δ and T ($m = 2$)

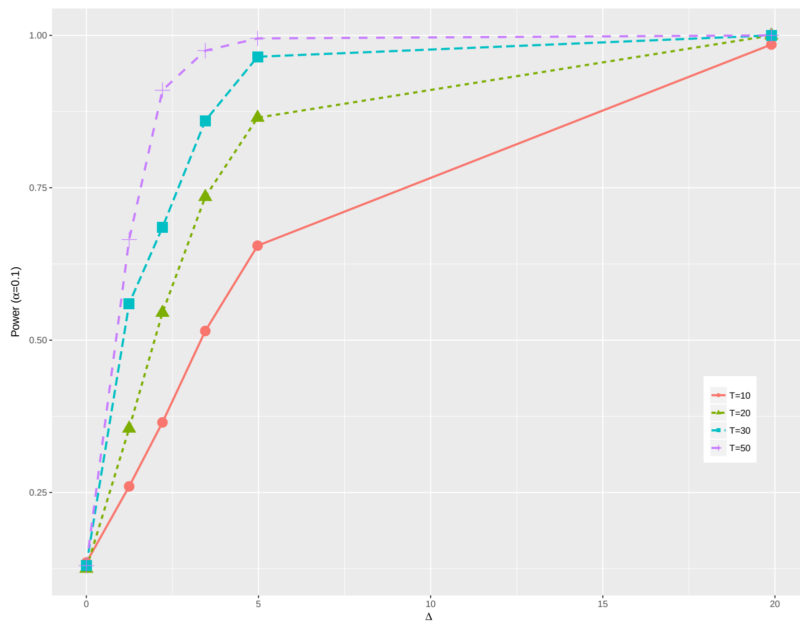


Figure 8.9: The empirical power of the test versus Δ and T ($m = 2$)

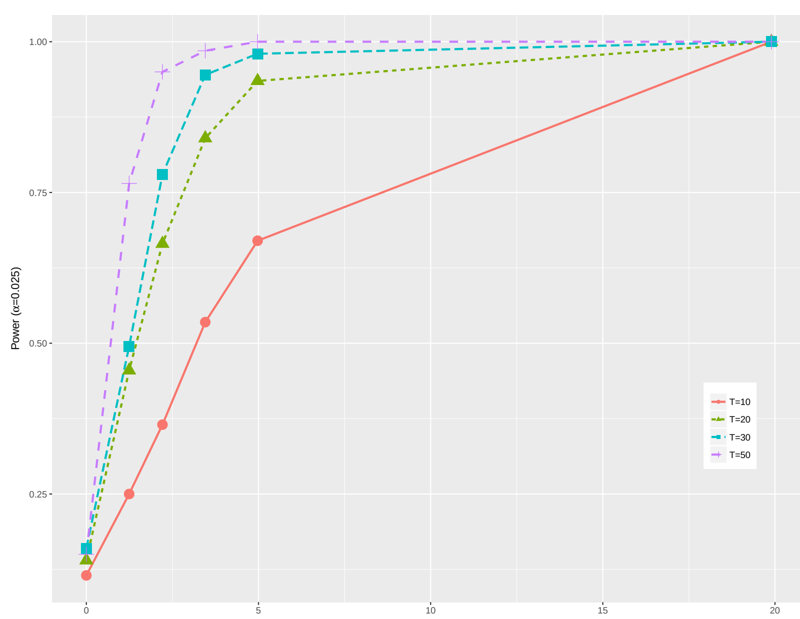


Figure 8.10: The empirical power of the test versus Δ and T ($m = 3$)

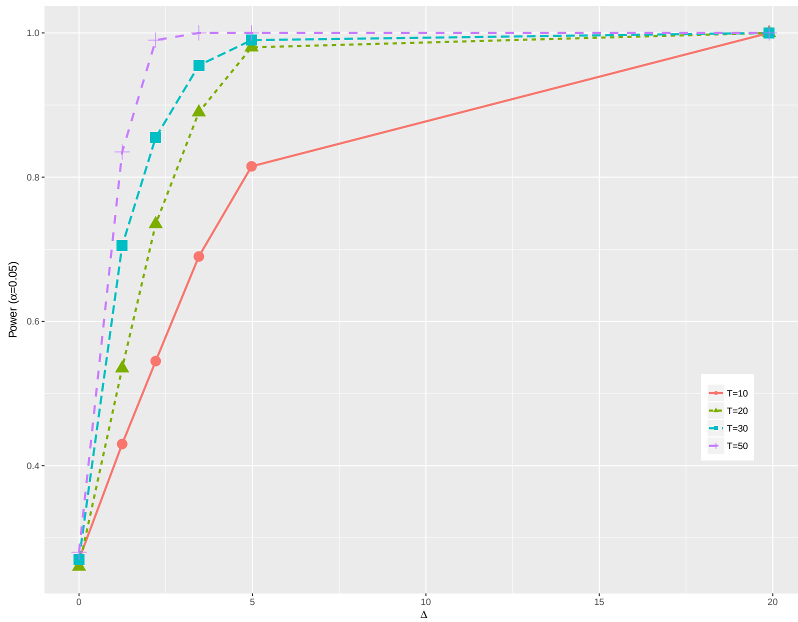


Figure 8.11: The empirical power of the test versus Δ and T $(m = 3)$

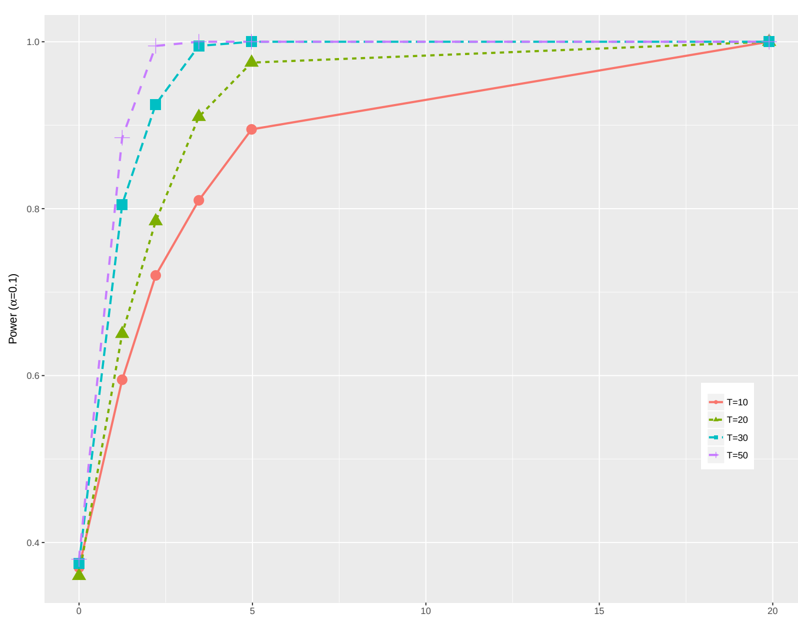


Figure 8.12: The empirical power of the test versus Δ and T $(m = 3)$

Chapter 9

Conclusion

In this thesis, we considered the inference problem in generalized O-U processes with unknown multiple change-points. In particular, the target parameter is the drift parameter whereas the number of change-points and the locations of the change-points are the nuisance parameters. In summary, we generalized the methods in Chen *et al.* (2017) as well as in Nkurunziza and Zhang (2018). More precisely, we generalized the main results in Chen *et al.* (2017) in five ways. First, we considered the statistical model which incorporates uncertain prior information and the uncertain restriction includes as a special case the nonexistence of the change-points. We derived the unrestricted estimators (UE) and the restricted estimators (RE). Second, in context of a known number of change-points, we derived the joint asymptotic normality of the UE and the RE. Third, we developed a hypothesis test for testing the restriction and we derived its asymptotic power. Fourth, we derived a class of shrinkage estimators (SEs) which encloses as special cases the UE, the RE as well as the classical SEs. Fifth, we derived the asymptotic distributional risk (ADR) functions of the UE, the RE, the SEs, and compared their relative risk efficiency. From the simulation study, we found that the simulation results corroborate the derived theoretical results.

On the top of these contributions, we derived two asymptotic properties which are of interest on their own. Further, we waived the conditions for the results in Chen *et al.* (2017) to hold. More precisely, we showed that Assumption 2 in Chen *et al.* (2017) is not required for their results to hold.

APPENDIX A

Appendix A

Theoretical Background

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 $x \in H$, $\sum_{i=1}^{p} |\langle x, \varphi_i \rangle|^2 \leq \|x\|^2$.*

Definition A.1. *(Weakly Stationary) A stochastic process $\{X_k\}_{k \in S}$ is **weakly stationary** if it has finite first and second moments and*

(i) $\mathrm{E}(X_k)$ *is a constant, i.e., it does not depend on k*

(ii) $\mathrm{Cov}(X_k, X_j)$ *is a function of $|j - k|$.*

Definition A.2. *(Strongly Stationary) A stochastic process $\{X_k\}_{k \in S}$ is (strongly) **stationary** if for any finite integer a, the joint distribution of $\{X_k\}_{k \in S}$ is equal to the joint distribution of $\{X_{k+a}\}_{k \in S}$.*

Note: These two definitions of strongly and weakly stationary come with the case of Gaussian process. For more detail, we refer to Koralov and Sinai (2007, p.234). Let $C[0, T]$ be a space of continuous function from $[0, T]$ to $\mathbb{R}$.

Proposition A.1. *(Wiener integral) If $f \in C[0, T]$, then the process defined by*

$$X_t = \int_0^t f(s) dB_s \ , \ t \in [0, T]$$

is a mean zero Gaussian process with independent increment and with covariance function

$$\mathrm{Cov}(X_s, X_t) = \int_0^{\min(s,t)} f^2(u)du.$$

Moreover, if we take the partition of [0,T] given $t_i = \frac{iT}{n}$ *for* $0 \leq i \leq n$ *and choose* t_i^* *to satisfy* $t_{i-1} \leq t_i^* \leq t_i$ *for all* $1 \leq i \leq n$*, then we have*

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(t_i^*)(B_{t_i} - B_{t_{i-1}}) = \int_0^T f(s)dB_s$$

where the limit is understood in the sense of convergence in probability.

For the proof of this result, we refer to Steele (2001 Proposition 7.6, p.101).

Theorem A.1. *Let* $(\Omega, \mathcal{A}, P, \tau)$ *be a measure-preserving dynamical system. Further, suppose that a stationary Gaussian process* $\{X_n\}_{n\in\mathbb{N}_0}$ *with correlation coefficient* R_n *satisfies* $\lim_{n\to\infty} R_n = 0$*. Then,* τ *is weakly-mixing.*

For the proof of this result, we refer to Lemma 5 and Theorem 5 of Chapter II in Gikhman and Skorohod (2004). Also, we refer to Stout (1974, Example 3.5.2, p.185.)

Theorem A.2. $(\Omega, \mathcal{A}, P, \tau)$ *is ergodic if and only if for all* A *,* $B \in \mathcal{A}$*, the measure preserving transformation* τ*is weakly-mixing.*

Definition A.3. *Let* $X = \{X_t\}_{t\in S}$ *be a stochastic process where the index set* S *could be* $\mathbb{R}$*,* $\mathbb{N}$*,* $\mathbb{N}_0, \mathbb{Z}$ *and so on. Then, the stochastic process is* X *is called* ***ergodic*** *if* $(\Omega, \mathcal{A}, P, \tau)$ *is ergodic.*

Theorem A.3. *Let* $\{X_i, i \geq 1\}$ *be stationary ergodic and let* ϕ *be a measurable function* $\phi : \mathbb{R}^\infty \to \mathbb{R}^1$*. Let* $Y_i = \phi(X_i, X_{i+1}, ...)$ *and define* $\{Y_i, i \geq 1\}$*. Then,* $\{Y_i, i \geq 1\}$ *is stationary ergodic.*

The proof of this result is given in Stout (1974 Theorem 3.5.8, p.182).

Theorem A.4. *(Stout, 1974, Theorem 3.5.7, p.181) Let* $\{X_i, i \geq 1\}$ *be a stationary and ergodic process with* $\mathrm{E}[|X_i|] < \infty$*. Then,* $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{a.s.} \mathrm{E}[X_1]$.

Theorem A.5. *(Mathai and Provost, 1992, Theorem 5.1.3) If* $X \sim \mathcal{N}_p(\mu, \Sigma)$ *and* Σ *is positive semidefinite, then a set of necessary and sufficient conditions for* $X^\top AX \sim \chi^2_q(\Delta)$ *is*

(i) $tr(A\Sigma) = q$ *and* $\mu^\top A\mu = \Delta$,

(ii) $\Sigma A\Sigma A\Sigma = \Sigma A\Sigma$,

(iii) $\mu^\top A\Sigma A\mu = \mu^\top A\mu$,

(iv) $\mu^\top (A\Sigma)^2 = \mu^\top A\Sigma$.

The proof is referred to Mathai and Provost (1992 Theorem 5.1.3, p.199).

Proposition A.2. *Let* $X \sim \mathcal{N}_m(\mu, \Sigma)$*. If* A *is* $n \times m$*-matrix and* B *is* n*-column vector, then* $AX + B \sim \mathcal{N}(A\mu + B, A\Sigma A^\top)$.

A.1 Identities in Shrinkage method

Theorem A.6. *(Nkurunziza, 2012b, Theorem 2.2) Let*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}_{2q\times 2k}\left(\begin{pmatrix} M_1 \\ M_2 \end{pmatrix}, \begin{pmatrix} \Upsilon_{11}\otimes\Lambda_{11} & 0 \\ 0 & \Upsilon_{22}\otimes\Lambda_{22} \end{pmatrix}\right)$$

where Λ_{11} *is positive definite matrix, and* Υ_{11}, Υ_{22} *and* Λ_{22} *are non-negative definite matrices with rank* $p \leq k$*. Also, let* Ξ *be a symmetric and positive definite matrix which satisfies the following two conditions:*

(i) $\Upsilon_{11}\Xi$ *is an idempotent matrix ;(ii)* $\Xi\Upsilon_{11}\Xi M_1 = \Xi M_1$.

Then, for any h Borel measurable and integrable function, and any non-negative def-

inite matrix A, we have

$$\mathrm{E}[h(trace(\Lambda_{11}^{-1}X^{\top}\Xi\Upsilon_{11}\Xi X))Y^{\top}AX] = \mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\Lambda_{11}^{-1}M_1^{\top}\Xi\Upsilon_{11}\Xi M_1)))]M_2^{\top}AM_1.$$

Theorem A.7. *(Nkurunziza, 2012b, Theorem 2.3) Let $X \sim \mathcal{N}_{q\times k}(M, \Upsilon \otimes \Lambda)$, where Λ is a positive definite matrix and Υ is a non-negative definite matrix with rank $p \leq k$. Also, let A and Ξ be positive definite symmetric matrices and assume that Ξ satisfies the following two conditions:*
(i) $\Upsilon\Xi$ is an idempotent matrix ;(ii) $\Xi\Upsilon\Xi M = \Xi M$.
Then, for any h Borel measurable and integrable function, we have

$$\begin{aligned}
&\mathrm{E}[h(\mathrm{trace}(\Lambda^{-1}X^{\top}\Xi\Upsilon\Xi X))\mathrm{trace}(X^{\top}AX)] \\
&= \mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\Lambda^{-1}M^{\top}\Xi\Upsilon\Xi M)))]\mathrm{trace}(A\Upsilon)trace(A) \\
&+ \mathrm{E}[h(\chi^2_{pq+4}(\mathrm{trace}(\Lambda^{-1}M^{\top}\Xi\Upsilon\Xi M)))]\mathrm{trace}(M^{\top}AM).
\end{aligned}$$

Theorem A.8. *(Nkurunziza, 2012b, Theorem 3.1) Let $\Sigma = \Lambda^{-1}$, $\Sigma^* = \Lambda L_1^{\top}(L_1\Lambda L_1^{\top})^{-1}L_1\Lambda$, and $\delta = \Lambda L_1^{\top}(L_1\Lambda L_1^{\top})^{-1}(L_1\theta L_2 - d)$. Then, the risk function of the estimator $\hat{\theta}$ is given by*

$$\begin{aligned}
\mathrm{R}(\hat{\theta}, \theta, \Omega) &= \mathrm{trace}(\Omega(\Sigma - \Sigma^*))\mathrm{trace}(L_2^{\top}L_2) + \mathrm{trace}(\delta^{\top}\Omega\delta) \\
&- 2\mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}((L_2^{\top}L_2)^{-1}\delta^{\top}\Xi_1\delta)))]\mathrm{trace}(\delta^{\top}\Omega\delta) \\
&+ \mathrm{E}[h^2(\chi^2_{pq+2}(\mathrm{trace}((L_2^{\top}L_2)^{-1}\delta^{\top}\Xi_1\delta)))]\mathrm{trace}(\Omega\Sigma^*)\mathrm{trace}(L_2^{\top}L_2) \\
&+ \mathrm{E}[h^2(\chi^2_{pq+4}(\mathrm{trace}((L_2^{\top}L_2)^{-1}\delta^{\top}\Xi_1\delta)))]\mathrm{trace}(\delta^{\top}\Omega\delta).
\end{aligned}$$

APPENDIX B

Appendix B

Some Technical Results and Proofs

Proposition B.1. *Suppose that the conditions in Proposition 4.1 hold. Then, $\hat{\phi}$ is a consistent estimator for ϕ. Further, for every $\epsilon > 0$, there exists a $C > 0$ such that for large T,* $\mathrm{P}(T \max_{1 \le j \le m} |\hat{\phi}_j - \phi_j| > C) < \epsilon$.

The proof is similar to that given for Proposition 4.2 of Chen *et al.* (2017).

Proposition B.2. *Under Assumption 1-4. we have that for large T,*

(i) $\mathrm{IC}(m^0) < \mathrm{IC}(m)$ *a.s.* $\forall\ m < m^0$ *and (ii)* $\mathrm{IC}(m^0) < \mathrm{IC}(m)$ *a.s.* $\forall\ m > m^0$.

The proof is similar to that given for Proposition 5.1 of Chen *et al.* (2017).

Proof of Lemma 2.1. Consider the SDE without change-point,
$dU_t = \left(\sum_{k=1}^{p} \mu_k \varphi_k(t) - aU_t\right) dt + \sigma dW_t,\ 0 \le t \le T$. Let $g(t,x) = e^{at}x$ and $Y_t = g(t, U_t)$
By Itô's lemma,

$$dY_t = \frac{\partial g}{\partial t}(t, U_t)dt + \frac{\partial g}{\partial U_t}(t, U_t)dU_t + \frac{1}{2}\frac{\partial^2 g}{\partial U_t^2}(t, U_t)d\langle U_t, U_t\rangle.$$

Since $dU_t = \left(\sum_{k=1}^{p} \mu_k \varphi_k(t) - aU_t\right) dt + \sigma dW_t$, then,

$$dY_t = ae^{at}U_t dt + e^{at}dU_t = e^{at}\sum_{k=1}^{p} \mu_k \varphi_k(t) dt + e^{at}\sigma dW_t.$$

Integrating both sides from 0 to t, we get

$$U_t = e^{-at}U_0 + e^{-at}\sum_{k=1}^{p}\mu_k\int_0^t e^{as}\varphi_k(s)ds + \sigma e^{-at}\int_0^t e^{as}dW_s.$$

Further, we have

$$\mathrm{E}\left[|X_t|^2\right] = \mathrm{E}\left[\left(\sum_{j=1}^{m+1} X_j(t)\mathbb{I}_{(\tau_{j-1}<t\leq\tau_j)}\right)^2\right].$$

By convexity of the quadratic function,

$$\left(\sum_{j=1}^{m+1}\frac{1}{m+1}X_j(t)\mathbb{I}_{(\tau_{j-1}<t\leq\tau_j)}\right)^2 \leq \sum_{j=1}^{m+1}\frac{1}{m+1}(X_j(t)\mathbb{I}_{(\tau_{j-1}<t\leq\tau_j)})^2.$$

Then,

$$\mathrm{E}\left[\left(\sum_{j=1}^{m+1} X_j(t)\mathbb{I}_{(\tau_{j-1}<t\leq\tau_j)}\right)^2\right] \leq (m+1)\sum_{j=1}^{m+1}\mathrm{E}\left[(X_j(t))^2\mathbb{I}_{(\tau_{j-1}<t\leq\tau_j)}\right],$$

and then, $\sup_{t\geq 0}\mathrm{E}[|X_t|^2] \leq (m+1)\sum_{j=1}^{m+1}\sup_{t\geq 0}\mathrm{E}[|X_j(t)|^2]$. Now, it is sufficient to prove $\sup_{t\geq 0}\mathrm{E}[|X_j(t)|^2] < \infty$, $j = 1, 2, \ldots, m+1$. Then, from the convexity of the quadratic function, $(a+b+c)^2 \leq 3a^2 + 3b^2 + 3c^2$, then,

$$\mathrm{E}[|X_j(t)|^2] = \mathrm{E}[|e^{-a_j t}X_0 + h_j(t) + z_j(t)|^2] \leq 3e^{-2a_j t}\mathrm{E}[|X_0|^2] + 3\mathrm{E}[h_j^2(t)] + 3\mathrm{E}[z_j^2(t)]. \tag{B.1}$$

Let $\sum_{k=1}^{p}|\mu_{k,j}| \leq K_\mu < \infty$ and $|\varphi_k(t)| \leq K_\varphi < \infty$ for all $j = 1, ..., m+1$ and $k = 1, ..., p$, $t \geq 0$. By Triangular Inequality and Jensen's Inequality, we have

$$\mathrm{E}[h_j^2(t)] = \mathrm{E}\left[\left(e^{-a_j t}\sum_{k=1}^{p}\mu_{k,j}\int_0^t e^{a_j s}\varphi_k(s)ds\right)^2\right) \leq e^{-2a_j t}K_\mu^2K_\varphi^2\int_0^t e^{2a_j s}ds.$$

Then, for $t \geq 0$

$$\mathrm{E}[h_j^2(t)] \leq e^{-2a_j t}K_\mu^2K_\varphi^2\frac{1}{2a_j}(e^{2a_j t}-1) = K_\mu^2K_\varphi^2\frac{1}{2a_j}(1-e^{-2a_j t}) \leq \frac{K_\mu^2K_\varphi^2}{2a_j}. \tag{B.2}$$

Also, we have

$$\mathrm{E}[z_j^2(t)] = \mathrm{E}\left[\left(\sigma e^{-a_j t}\int_0^t e^{a_j s}dW_s\right)^2\right] = \sigma^2 e^{-2a_j t}\mathrm{E}\left[\left(\int_0^t e^{a_j s}dW_s\right)^2\right].$$

By using Itô's Isometry, we have

$$\mathrm{E}[z_j^2(t)] = \sigma^2 e^{-2a_j t}\mathrm{E}\left[\int_0^t e^{2a_j s}ds\right] = \sigma^2 e^{-2a_j t}\frac{1}{2a_j}(e^{2a_j t}-1).$$

Then, for $t \geq 0$

$$\mathrm{E}[z_j^2(t)] = \sigma^2\frac{1}{2a_j}(1-e^{-2a_j t}) \leq \frac{\sigma^2}{2a_j}. \tag{B.3}$$

Finally, by combining (B.1), (B.2) and (B.3), we establish

$$\sup_{t\geq 0}\mathrm{E}[|X_j(t)|^2] \leq 3\mathrm{E}[|X_0|^2] + 3\frac{K_\mu^2 K_\varphi^2}{2a_j} + 3\frac{\sigma^2}{2a_j} < \infty,$$

this completes the proof. □

Proof of Proposition 2.2. We have

$\mathrm{E}[\tilde{X}_j(t+k)] = \mathrm{E}[\tilde{h}_j(t+k)] + \mathrm{E}[\tilde{z}_j(t+k)] = \tilde{h}_j(t+k) + \mathrm{E}[\tilde{z}_j(t+k)].$

Let $r = s - k$ and by Assumption 2,

$$\begin{aligned}\tilde{h}_j(t+k) &= e^{-a_j(t+k)}\sum_{i=1}^p \mu_{i,j}\int_{-\infty}^{t+k} e^{a_j s}\varphi_i(s)ds \\ &= e^{-a_j t}\sum_{i=1}^p \mu_{i,j}\int_{-\infty}^{t} e^{a_j r}\varphi_i(r)dr = \tilde{h}_j(t).\end{aligned}$$

Therefore, $\tilde{h}_j(t+k)$ does not depend on k and is a constant for every $t \in [0,1]$.

$$\mathrm{E}[\tilde{z}_j(t+k)] = \sigma e^{-a_j(t+k)}\left[\mathrm{E}\left[\int_{-\infty}^0 e^{a_j s}d\bar{B}_{-s}\right] + \mathrm{E}\left[\int_0^{t+k} e^{a_j s}dB_s\right]\right]$$

Since Itô's integral is martingale, we have $\mathrm{E}\left[\int_0^{t+k} e^{a_j s}dB_s\right] = 0$, and

$$\mathrm{E}\left[\int_{-\infty}^0 e^{a_j s}d\bar{B}_{-s}\right] = \mathrm{E}\left[\int_0^{\infty} e^{-a_j u}d\bar{B}_u\right] = \mathrm{E}\left[\lim_{U\to\infty}\int_0^U e^{-a_j u}d\bar{B}_u\right] = \mathrm{E}\left[\lim_{U\to\infty} I_U\right].$$

By (B.11), $I_U \xrightarrow[U\to\infty]{L^2} I_\infty$. This implies that $I_U \xrightarrow[U\to\infty]{L^1} I_\infty$. Then, by the martingale property, $\mathrm{E}[\lim_{U\to\infty} I_U] = \lim_{U\to\infty} \mathrm{E}[I_U] = 0$, and then,

$$\mathrm{E}[\tilde{z}_j(t+k)] = 0,\ \ k = 0, 1, \ldots. \tag{B.4}$$

Therefore,

$$\mathrm{E}[\tilde{X}_j(t+k)] = \tilde{h}_j(t+k) + \mathrm{E}[\tilde{z}_j(t+k)] = \tilde{h}_j(t),\ \ k = 0, 1, \ldots, \tag{B.5}$$

which is a constant. For $\mathrm{Cov}(\tilde{X}_j(t), \tilde{X}_j(t+k))$, since $\mathrm{E}[\tilde{z}_j(t+k)] = 0$, we have

$$\mathrm{Cov}(\tilde{X}_j(t), \tilde{X}_j(t+k)) = \mathrm{Cov}(\tilde{z}_j(t), \tilde{z}_j(t+k)) = \mathrm{E}[\tilde{z}_j(t)\tilde{z}_j(t+k)].$$

We have

$$\begin{aligned}
\mathrm{E}[\tilde{z}_j(t)\tilde{z}_j(t+k)] &= \mathrm{E}\left[\left(\sigma e^{-a_j t}\int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s\right)\left(\sigma e^{-a_j(t+k)}\int_{-\infty}^{t+k} e^{a_j s} d\tilde{B}_s\right)\right] \\
&= e^{-a_j k}\mathrm{E}\left[\left(\sigma e^{-a_j t}\int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s\right)^2\right] + \sigma^2 e^{-a_j(2t+k)}\mathrm{E}\left[\int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s \int_{t}^{t+k} e^{a_j s} d\tilde{B}_s\right] \\
&= e^{-a_j k}\mathrm{E}[\tilde{z}_j^2(t)]
\end{aligned}$$

since, by the independent increments of Wiener process, we have

$$\mathrm{E}\left[\int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s \int_{t}^{t+k} e^{a_j s} d\tilde{B}_s\right] = \mathrm{E}\left[\int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s\right]\mathrm{E}\left[\int_{t}^{t+k} e^{a_j s} d\tilde{B}_s\right] = 0,$$

then, by (B.13), $\mathrm{E}[\tilde{z}_j^2(t+k)] = \dfrac{\sigma^2}{2a_j}$ for $k = 0, 1, 2, \ldots$, and then, we can establish $\mathrm{E}[\tilde{z}_j^2(t)] = \dfrac{\sigma^2}{2a_j}$. Then, $\mathrm{Cov}(\tilde{X}_j(t), \tilde{X}_j(t+k)) = e^{-a_j k}\dfrac{\sigma^2}{2a_j}$, this completes the proof. □

Proof of Lemma 2.2. First, we prove for all k and t, $\mathrm{E}[|\tilde{X}_j(k+t)|^2] < \infty$. Since $(a+b)^2 \le 2a^2 + 2b^2$, we have

$$\mathrm{E}[|\tilde{X}_j(k+t)|^2] = \mathrm{E}[|\tilde{h}_j(t+k) + \tilde{z}_j(t+k)|^2] \le 2\mathrm{E}[\tilde{h}_j^2(t+k)] + 2\mathrm{E}[\tilde{z}_j^2(t+k)].$$

Let $\sum_{i=1}^{p} |\mu_{i,j}| \leq K_\mu < \infty$ and $|\varphi_i(t+k)| \leq K_\varphi < \infty$ for all $j = 1, ..., m+1$ and $k = 1, ..., p$, $t \geq 0$. By Triangle Inequality and Jensen's Inequality,

$$\begin{aligned}\mathrm{E}[\tilde{h}_j^2(t+k)] &\leq e^{-2a_j(t+k)} \left(\sum_{i=1}^{p} |\mu_{i,j}| \right)^2 \int_{-\infty}^{t+k} (e^{a_j s} K_\varphi)^2 ds \\ &\leq e^{-2a_j(t+k)} K_\mu^2 K_\varphi^2 \int_{-\infty}^{t+k} e^{2a_j s} ds.\end{aligned}$$

Then,

$$\mathrm{E}[\tilde{h}_j^2(t+k)] \leq e^{-2a_j(t+k)} K_\mu^2 K_\varphi^2 \frac{e^{2a_j(t+k)}}{2a_j} = \frac{K_\mu^2 K_\varphi^2}{2a_j} < \infty, \ k = 0, 1, \ldots. \tag{B.6}$$

Since B_s and B_{-s} are independent,

$$\begin{aligned}\mathrm{E}[\tilde{z}_j^2(t+k)] &= \sigma^2 e^{-2a_j(t+k)} \mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s + \int_{0}^{t+k} e^{a_j s} d\tilde{B}_s \right)^2 \right] \\ &= \sigma^2 e^{-2a_j(t+k)} \left[\mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} \right)^2 \right] + 2\mathrm{E}\left[\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} \right] \mathrm{E}\left[\int_{0}^{t+k} e^{a_j s} dB_s \right] \right. \\ &\quad \left. + \mathrm{E}\left[\left(\int_{0}^{t+k} e^{a_j s} dB_s \right)^2 \right] \right].\end{aligned}$$

Then,

$$\mathrm{E}[\tilde{z}_j^2(t+k)] = \sigma^2 e^{-2a_j(t+k)} \left[\mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} \right)^2 \right] + \mathrm{E}\left[\left(\int_{0}^{t+k} e^{a_j s} dB_s \right)^2 \right] \right],$$

since Itô's integral is a martingale, $\mathrm{E}\left[\int_{0}^{t+k} e^{a_j s} dB_s \right] = 0$. Then, we have

$$\mathrm{E}[\tilde{z}_j^2(t+k)] = \sigma^2 e^{-2a_j(t+k)} \left[\mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} \right)^2 \right] + \mathrm{E}\left[\left(\int_{0}^{t+k} e^{a_j s} dB_s \right)^2 \right] \right]. \tag{B.7}$$

From (B.7), by Itô's Isometry,

$$\mathrm{E}\left[\left(\int_{0}^{t+k} e^{a_j s} dB_s \right)^2 \right] = \mathrm{E}\left[\int_{0}^{t+k} e^{2a_j s} ds \right] = \frac{1}{2a_j} (e^{2a_j(t+k)} - 1). \tag{B.8}$$

From (B.7), by using substitution with $s = -u$,

$$\mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} \right)^2 \right] = \mathrm{E}\left[\left(\int_{0}^{\infty} e^{-a_j u} d\bar{B}_u \right)^2 \right]. \tag{B.9}$$

Now, we define $I_U = \int_0^U e^{-a_j u} d\bar{B}_u$. By Itô's Isometry,

$$\mathrm{E}[I_U^2] = \mathrm{E}\left[\left(\int_0^U e^{-a_j u} d\bar{B}_u\right)^2\right] = \mathrm{E}\left[\int_0^U e^{-2a_j u} du\right] = \frac{1}{2a_j}(1 - e^{-2a_j U}) \qquad \text{(B.10)}$$

which is bounded for all $U \geq 0$. Thus, by L^2-Bounded Martingale Convergence Theorem,

$$I_U \xrightarrow[U\to\infty]{L^2} I_\infty = \int_0^\infty e^{-a_j u} d\bar{B}_u \;\; and \;\; \mathrm{E}[I_\infty^2] < \infty. \qquad \text{(B.11)}$$

Then, we have

$$\mathrm{E}\left[\left(\int_{-\infty}^0 e^{a_j s} d\bar{B}_{-s}\right)^2\right] = \mathrm{E}[I_\infty^2] = \lim_{U\to\infty} \mathrm{E}[I_U^2] = \frac{1}{2a_j}. \qquad \text{(B.12)}$$

By (B.7), (B.8) and (B.12), we have

$$\mathrm{E}[\tilde{z}_j^2(t+k)] = \sigma^2 e^{-2a_j(t+k)}\left[\frac{1}{2a_j}(e^{2a_j(t+k)}-1)+\frac{1}{2a_j}\right] = \frac{\sigma^2}{2a_j} < \infty,\ k = 0, 1, \ldots. \qquad \text{(B.13)}$$

Since $\mathrm{E}[\tilde{h}_j^2(t+k)] < \infty$ and $\mathrm{E}[\tilde{z}_j^2(t+k)] < \infty$, it implies

$$\mathrm{E}[|\tilde{X}_j(k+t)|^2] \leq 2\mathrm{E}[\tilde{h}_j^2(t+k)] + 2\mathrm{E}[\tilde{z}_j^2(t+k)] < \infty.$$

Now, we will start to prove $\mathrm{E}[\tilde{X}_j(t+k)]$ is a constant. From Proposition 2.2, $\{\tilde{X}_j(t+k)\}_{k\in\mathbb{N}_0}$ is weakly stationary.

For every $t \in [0,1]$ and $k \in \mathbb{N}_0$, we have $\tilde{X}_j(t+k) = \tilde{h}_j(t) + \tilde{z}_j(t+k)$. By some algebraic computations, one can verify that $\{\tilde{X}_j(t+k)\}_{k\in\mathbb{N}}$ is Gaussian process. This implies that the weekly stationary Gaussian process $\{\tilde{X}_j(t+k)\}_{k\in\mathbb{N}_0}$ is also strongly stationary. Now, for $t \in [0,1]$ and $k \in \mathbb{N}_0$, the correlation coefficient function is

$$R_k = \frac{\mathrm{Cov}(\tilde{X}_j(t), \tilde{X}_j(t+k))}{\mathrm{Var}(\tilde{X}_j(t))},$$

where $\mathrm{Var}(\tilde{X}_j(t)) = \mathrm{E}[\tilde{X}_j^2(t)] - \mathrm{E}[\tilde{X}_j(t)]^2 = \tilde{h}_j^2(t) + \mathrm{E}[\tilde{z}_j^2(t)] - \tilde{h}_j^2(t) = \mathrm{E}[\tilde{z}_j^2(t)] = \frac{\sigma^2}{2a_j}$, and $\mathrm{Cov}(\tilde{X}_j(t), \tilde{X}_j(t+k)) = e^{-a_j k}\frac{\sigma^2}{2a_j}$. Then, $R_k = e^{-a_j k}$, and then, $R_k \to 0$ as

$k \to \infty$. By Definition A.3, Theorem A.1 and Theorem A.2, we have $\{\tilde{X}_j(t+k)\}_{k\in\mathbb{N}_0}$ is ergodic, this completes the proof. $\square$

Proof of Proposition 2.3. From Remark 2, we apply Theorem 7.6 of Lipster and Shiryaev (2001). Then, the likelihood function is

$$L(\theta, X_t) = \exp\left(\frac{1}{\sigma^2}\int_0^T S(\theta,t,X_t)dX_t - \frac{1}{2\sigma^2}\int_0^T S^2(\theta,t,X_t)dt\right). \tag{B.14}$$

Then, the log-likelihood function is

$$\log L(\theta, X_t) = \frac{1}{\sigma^2}\int_0^T S(\theta,t,X_t)dX_t - \frac{1}{2\sigma^2}\int_0^T S^2(\theta,t,X_t)dt. \tag{B.15}$$

By (2.2), we have

$$\begin{aligned}\frac{1}{\sigma^2}\int_0^T S(\theta,t,X_t)dX_t &= \frac{1}{\sigma^2}\int_0^T \sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_j X_t\right)\mathbb{I}_{\{\tau_{j-1}<t\leq\tau_j\}}dX_t \\ &= \frac{1}{\sigma^2}\sum_{j=1}^{m+1}\int_0^T \left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_j X_t\right)\mathbb{I}_{\{\tau_{j-1}<t\leq\tau_j\}}dX_t.\end{aligned}$$

Then,

$$\frac{1}{\sigma^2}\int_0^T S(\theta,t,X_t)dX_t = \frac{1}{\sigma^2}\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\int_{\tau_{j-1}}^{\tau_j}\mu_{k,j}\varphi_k(t)dX_t - \int_{\tau_{j-1}}^{\tau_j} a_j X_t dX_t\right). \tag{B.16}$$

This gives

$$\frac{1}{\sigma^2}\int_0^T S(\theta,t,X_t)dX_t = \frac{1}{\sigma^2}\sum_{j=1}^{m+1}\theta_j^\top \tilde{r}_{(\tau_{j-1},\tau_j)}$$

where

$$\tilde{r}_{(\tau_{j-1},\tau_j)} = \left(\int_{\tau_{j-1}}^{\tau_j}\varphi_1(t)dX_t, ..., \int_{\tau_{j-1}}^{\tau_j}\varphi_p(t)dX_t, -\int_{\tau_{j-1}}^{\tau_j}X_t dX_t\right)^\top.$$

Further, from (2.2), we have

$$\frac{1}{2\sigma^2}\int_0^T S^2(\theta,t,X_t)dt = \frac{1}{2\sigma^2}\int_0^T\left[\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_j X_t\right)\mathbb{I}_{\{\tau_{j-1}<t\leq\tau_j\}}\right]^2 dt.$$

Note that

$$\left[\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)-a_jX_t\right)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\right]^2$$

$$=\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)-a_jX_t\right)^2\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}$$

$$+\sum_{i\neq j}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,i}\varphi_k(t)-a_iX_t\right)\mathbb{I}_{\{\tau_{i-1}<t\le\tau_i\}}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)-a_jX_t\right)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}.$$

This gives

$$\left[\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)-a_jX_t\right)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\right]^2=\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)-a_jX_t\right)^2\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}$$

$$=\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}^2\varphi_k(t)^2+\sum_{k\neq q}^{p}\mu_{k,j}\varphi_k(t)\mu_{q,j}\varphi_q(t)-2\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)a_jX_t+a_j^2X_t^2\right)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}.$$

Hence,

$$\begin{aligned}\frac{1}{2\sigma^2}\int_0^T S^2(\theta,t,X_t)dt &= \frac{1}{2\sigma^2}\int_0^T\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}^2\varphi_k(t)^2+\sum_{k\neq q}^{p}\mu_{k,j}\varphi_k(t)\mu_{q,j}\varphi_q(t)\right.\\ &\quad\left.-2\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t)a_jX_t+a_j^2X_t^2\right)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}dt\\ &=\frac{1}{2\sigma^2}\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\int_{\tau_{j-1}}^{\tau_j}\mu_{k,j}^2\varphi_k(t)^2dt+\sum_{k\neq q}^{p}\int_{\tau_{j-1}}^{\tau_j}\mu_{k,j}\varphi_k(t)\mu_{q,j}\varphi_q(t)dt\right.\\ &\quad\left.-2\sum_{k=1}^{p}\int_{\tau_{j-1}}^{\tau_j}\mu_{k,j}\varphi_k(t)a_jX_tdt+\int_{\tau_{j-1}}^{\tau_j}a_j^2X_t^2dt\right).\end{aligned}$$

Then,

$$\frac{1}{2\sigma^2}\int_0^T S^2(\theta,t,X_t)dt=\frac{1}{2\sigma^2}\sum_{j=1}^{m+1}\theta_j^\top Q_{(\tau_{j-1},\tau_j)}\theta_j$$

where

$$Q_{(\tau_{j-1},\tau_j)}=\begin{bmatrix}\int_{\tau_{j-1}}^{\tau_j}\varphi_1^2(t)dt & \dots & \int_{\tau_{j-1}}^{\tau_j}\varphi_1(t)\varphi_p(t)dt & -\int_{\tau_{j-1}}^{\tau_j}\varphi_1(t)X_tdt\\ \vdots & \vdots & \vdots & \vdots\\ -\int_{\tau_{j-1}}^{\tau_j}\varphi_1(t)X_tdt & \dots & -\int_{\tau_{j-1}}^{\tau_j}\varphi_p(t)X_tdt & \int_{\tau_{j-1}}^{\tau_j}X_t^2dt\end{bmatrix}.$$

Finally, we can conclude that

$$\log L(\theta, X_t) = \frac{1}{\sigma^2}\sum_{j=1}^{m+1}\theta_j^\top \tilde{r}_{(\tau_{j-1},\tau_j)} - \frac{1}{2\sigma^2}\sum_{j=1}^{m+1}\theta_j^\top Q_{(\tau_{j-1},\tau_j)}\theta_j.$$

This completes the proof. □

Proof of Proposition 2.4. Let $a = (a_1^\top, a_2)^\top$ where a_1 is p-column vector and a_2 is scalar. Then, we have

$$\begin{aligned}
&a^\top Q_{(\tau_{j-1},\tau_j)}a \\
&= (a_1^\top \int_{\tau_{j-1}}^{\tau_j}\varphi(t)\varphi^\top(t)dt - a_2\int_{\tau_{j-1}}^{\tau_j}\varphi^\top(t)X_t dt, -a_1^\top\int_{\tau_{j-1}}^{\tau_j}X_t\varphi(t)dt + a_2\int_{\tau_{j-1}}^{\tau_j}X_t^2 dt)(a_1^\top, a_2)^\top \\
&= a_1^\top\int_{\tau_{j-1}}^{\tau_j}\varphi(t)\varphi^\top(t)dta_1 - 2a_2a_1^\top\int_{\tau_{j-1}}^{\tau_j}X_t\varphi(t)dt + a_2^2\int_{\tau_{j-1}}^{\tau_j}X_t^2 dt \\
&= \int_{\tau_{j-1}}^{\tau_j}(a_1^\top\varphi(t) - a_2X_t)^2 dt \geq 0 \ for\ all\ t\ on\ (\tau_{j-1},\tau_j),
\end{aligned}$$

and then, $\int_{\tau_{j-1}}^{\tau_j}(a_1^\top\varphi(t) - a_2X_t)^2 dt = 0$ iff $a_1^\top\varphi(t) = a_2X_t$. Taking expected value and variance in both sides, we have $a_1^\top\mathrm{E}[\varphi(t)] = a_2\mathrm{E}[X_t]$ and $a_1^\top\mathrm{Var}(\varphi(t))a_1 = a_2^2Var[X_t]$. Since $\mathrm{Var}(\varphi(t))$=0, we have $a_2^2Var[X_t] = 0$. We know $Var[X_t] \neq 0$, then $a_2 = 0$. Then, $a_1^\top\mathrm{E}[\varphi(t)] = 0$ which implies $a_1^\top\varphi(t) = 0$. From Assumption 2, we have orthogonality of $\varphi(t)$. This implies the linear independence of $\varphi(t)$. Then, we have $a_1 = 0$. Therefore, we can conclude that $a^\top Q_{(\tau_{j-1},\tau_j)}a > 0$ for all t on (τ_{j-1},τ_j) and for all $a \neq 0$, this completes the first part of the proof. Further, we have $Q_{(\tau_{j-1},\tau_j)}$ is positive definite, for $j = 1, ..., m+1$. Since $Q(\phi, m)$ is a Block diagonal matrix, $Q(\phi, m)$ is positive definite, this completes the proof. □

Proof of Lemma 3.1. First, taking first derivative of $\log L(\theta, X_t)$ respect to θ

$$\frac{\partial}{\partial\theta}\log L(\theta, X_t) = \frac{1}{\sigma^2}\tilde{R}(\phi, m) - \frac{1}{\sigma^2}Q(\phi, m)\theta.$$

Then, setting $\frac{\partial}{\partial\theta}\log L(\theta, X_t) = 0$, we get $\dfrac{1}{\sigma^2}\tilde{R}(\phi, m) = \dfrac{1}{\sigma^2}Q(\phi, m)\hat{\theta}$. Then,

$$\hat{\theta} = Q^{-1}(\phi, m)\tilde{R}(\phi, m)$$

since, by Proposition 2.4, $Q(\phi, m)$ is positive definite, which implies that $Q(\phi, m)$ is invertible. Next, taking second derivative of $\log L(\theta, X_t)$ respect to θ

$$\frac{\partial^2}{\partial\theta\partial\theta^\top}\log L(\theta, X_t) = -\frac{1}{\sigma^2}Q(\phi, m) = -\frac{T}{\sigma^2}\frac{1}{T}Q(\phi, m)$$

since, by Proposition 2.4, $Q(\phi, m)$ is positive definite, and $\sigma > 0$, we have $-\dfrac{T}{\sigma^2}\dfrac{1}{T}Q(\phi, m)$ is negative definite, which completes the proof. □

Proof of Proposition 3.1. We have

$$dX_t = \sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_jX_t\right)\mathbb{I}_{\{\tau_{j-1}<t\leq\tau_j\}}dt + \sigma dW_t.$$

For fixed j, $j = 1, ..., m+1$

$$\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)dX_t = \int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)\left[\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_jX_t\right)\mathbb{I}_{\{\tau_{j-1}<t\leq\tau_j\}}\right]dt + \sigma\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)dW_t.$$

Then,

$$\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)dX_t = \int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_jX_t\right)dt + \sigma\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)dW_t,$$

and then,

$$\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)dX_t = \sum_{k=1}^{p}\mu_{k,j}\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)\varphi_k(t)dt - a_j\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)X_tdt + \sigma\int_{\tau_{j-1}}^{\tau_j}\varphi_i(t)dW_t.$$

Further, we have

$$\int_{\tau_{j-1}}^{\tau_j}X_tdX_t = \int_{\tau_{j-1}}^{\tau_j}X_t\left[\sum_{j=1}^{m+1}\left(\sum_{k=1}^{p}\mu_{k,j}\varphi_k(t) - a_jX_t\right)\mathbb{I}_{\{\tau_{j-1}<t\leq\tau_j\}}\right]dt + \sigma\int_{\tau_{j-1}}^{\tau_j}X_tdW_t.$$

Then,

$$\int_{\tau_{j-1}}^{\tau_j} X_t dX_t = \int_{\tau_{j-1}}^{\tau_j} X_t \left(\sum_{k=1}^{p} \mu_{k,j}\varphi_k(t) - a_j X_t \right) dt + \sigma \int_{\tau_{j-1}}^{\tau_j} X_t dW_t$$
$$= \sum_{k=1}^{p} \mu_{k,j} \int_{\tau_{j-1}}^{\tau_j} X_t \varphi_k(t) dt - a_j \int_{\tau_{j-1}}^{\tau_j} X_t^2 dt + \sigma \int_{\tau_{j-1}}^{\tau_j} X_t dW_t.$$

Hence, we observe that

$$\tilde{r}_{(\tau_{j-1},\tau_j)} = \left(\int_{\tau_{j-1}}^{\tau_j} \varphi_1(t) dX_t, ..., \int_{\tau_{j-1}}^{\tau_j} \varphi_p(t) dX_t, -\int_{\tau_{j-1}}^{\tau_j} X_t dX_t \right)^\top$$
$$= Q_{(\tau_{j-1},\tau_j)}\theta_j + \sigma r_{(\tau_{j-1},\tau_j)}.$$

Then,

$$\tilde{R}(\phi, m) = Q(\phi, m)\theta + \sigma R(\phi, m).$$

Finally, we have

$$\hat{\theta} = Q^{-1}(\phi, m)\tilde{R}(\phi, m) = Q^{-1}(\phi, m)Q(\phi, m)\theta + Q^{-1}(\phi, m)\sigma R(\phi, m)$$
$$= \theta + \sigma Q^{-1}(\phi, m)R(\phi, m).$$

This completes the proof. □

Proof of Proposition 3.2. We have, for $\nu \in (0, 1)$,

$$\frac{1}{\nu T} \int_0^{\nu T} \varphi(t)\varphi^\top(t)dt = \frac{1}{\nu T} \sum_{i=1}^{\lfloor \nu T \rfloor} \int_{i-1}^{i} \varphi(t)\varphi^\top(t)dt + \frac{1}{\nu T} \int_{\lfloor \nu T \rfloor}^{\nu T} \varphi(t)\varphi^\top(t)dt.$$

By Assumption 2,

$$\frac{1}{\nu T} \sum_{i=1}^{\lfloor \nu T \rfloor} \int_{i-1}^{i} \varphi(t)\varphi^\top(t)dt = \frac{\lfloor \nu T \rfloor}{\nu T} I_p \xrightarrow{a.s.} I_p \; as \; T \to \infty. \tag{B.17}$$

Also, by Jensen's Inequality, property of periodic function and substitution,

$$\left\| \frac{1}{\nu T} \int_{\lfloor \nu T \rfloor}^{\nu T} \varphi(t)\varphi^\top(t)dt \right\| \leq \frac{1}{\nu T} \int_{\lfloor \nu T \rfloor}^{\nu T} \|\varphi(t)\varphi^\top(t)\|dt \leq \frac{1}{\nu T} \int_{\lfloor \nu T \rfloor}^{\lfloor \nu T \rfloor + 1} \|\varphi(t)\varphi^\top(t)\|dt$$
$$= \frac{1}{\nu T} \int_0^1 \|\varphi(t)\varphi^\top(t)\|dt \xrightarrow{a.s.} 0 \; as \; T \to \infty,$$

which implies

$$\frac{1}{\nu T}\int_{\lfloor \nu T\rfloor}^{\nu T}\varphi(t)\varphi^{\top}(t)dt \xrightarrow[T\to\infty]{a.s.} 0. \tag{B.18}$$

By (B.17) and (B.18), we have

$$\frac{1}{\nu T}\int_{0}^{\nu T}\varphi(t)\varphi^{\top}(t)dt \xrightarrow[T\to\infty]{a.s.} I_p.$$

Therefore, we have

$$\frac{1}{T}\int_{0}^{\nu T}\varphi(t)\varphi^{\top}(t)dt = \nu\frac{1}{\nu T}\int_{0}^{\nu T}\varphi(t)\varphi^{\top}(t)dt \xrightarrow[T\to\infty]{a.s.} \nu I_p.$$

Finally, for $0 \le \phi_{j-1} < \phi_j \le 1$ where $j = 1, ..., m+1$,

$$\begin{aligned}\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\varphi(t)\varphi^{\top}(t)dt &= \frac{1}{T}\int_{0}^{\phi_j T}\varphi(t)\varphi^{\top}(t)dt - \frac{1}{T}\int_{0}^{\phi_{j-1}T}\varphi(t)\varphi^{\top}(t)dt \\ &\xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})I_p.\end{aligned}$$

This completes the proof. □

Proof of Proposition 3.3. By Markov Inequality, we have, for $\xi > 0$ and $\tau_{j-1} < t \le \tau_j$,

$$\mathrm{P}\left(\left\|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}X_t\varphi(t)dt\right\| \ge \xi\right) \le \frac{\mathrm{E}\left[\left\|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)\varphi(t)dt\right\|\right]}{\xi}.$$

By Assumption 2, let $|\varphi_k(t)| \le K_\varphi < \infty$, $t \ge 0$. Then, by Jensen's Inequality, for all $j = 1, ..., m+1$,

$$\begin{aligned}\mathrm{E}\left[\left\|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)\varphi(t)dt\right\|\right] &= \frac{1}{T}\mathrm{E}\left[\left\|\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)\varphi(t)dt\right\|\right] \\ &\le \frac{1}{T}\mathrm{E}\left[\int_{\phi_{j-1}T}^{\phi_j T}\|(\tilde{X}_t - X_t)\varphi(t)\|dt\right] \\ &= \frac{1}{T}\mathrm{E}\left[\int_{\phi_{j-1}T}^{\phi_j T}|\tilde{X}_t - X_t|\|\varphi(t)\|dt\right].\end{aligned}$$

Then,

$$\mathrm{E}\left[\left\|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)\varphi(t)dt\right\|\right] \leq \frac{K_\varphi}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t - X_t|]dt.$$

By Triangle Inequality, for $\tau_{j-1} < t \leq \tau_j$,

$$\begin{aligned}\mathrm{E}[|(\tilde{X}_t - X_t)|] &= \mathrm{E}[|\tilde{h}_j(t) + \tilde{z}_j(t) - e^{-a_j t}X_0 - h_j(t) - z_j(t)|] \\ &\leq \ \mathrm{E}[|\tilde{h}_j(t) - h_j(t)|] + \mathrm{E}[|\tilde{z}_j(t) - z_j(t)|] + \mathrm{E}[|e^{-a_j t}X_0|].\end{aligned}$$

Now, let $\sum_{k=1}^{p}|\mu_{k,j}| \leq K_\mu < \infty$ for all $j = 1,..,m+1$. We have,

$$\begin{aligned}\mathrm{E}[|\tilde{h}_j(t) - h_j(t)|] &= \mathrm{E}\left[\left|e^{-a_j t}\sum_{k=1}^{p}\mu_{k,j}\int_{-\infty}^{0}e^{a_j s}\varphi_k(s)ds\right|\right] \\ &\leq e^{-a_j t}\sum_{k=1}^{p}|\mu_{k,j}|K_\varphi\int_{-\infty}^{0}e^{a_j s}ds,\end{aligned}$$

then,

$$\mathrm{E}[|\tilde{h}_j(t) - h_j(t)|] \leq e^{-a_j t}K_\mu K_\varphi\frac{1}{a_j}, \tag{B.19}$$

and then, by Cauchy Schwartz Inequality,

$$\begin{aligned}\mathrm{E}[|\tilde{z}_j(t) - z_j(t)|] \leq \mathrm{E}[|\tilde{z}_j(t) - z_j(t)|^2]^{\frac{1}{2}} &= \mathrm{E}\left[\left(\sigma e^{-a_j t}\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right]^{\frac{1}{2}} \\ &= \sigma e^{-a_j t}\mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right]^{\frac{1}{2}}.\end{aligned}$$

Since, for $s \in (-\infty, 0)$, $\tilde{B}_s = \bar{B}_{-s}$,

$$\mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right] = \mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\bar{B}_{-s}\right)^2\right] = \mathrm{E}\left[\left(\int_{0}^{\infty}e^{-a_j u}d\bar{B}_u\right)^2\right]. \tag{B.20}$$

Now, we define $I_U = \int_0^U e^{-a_j u}d\bar{B}_u$. By (B.10) and (B.11), we have $\mathrm{E}[I_\infty^2] < \infty$. Let $\mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right]^{\frac{1}{2}} \leq K_1 < \infty$,

$$\mathrm{E}[|\tilde{z}_j(t) - z_j(t)|] \leq \sigma e^{-a_j t}\mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right]^{\frac{1}{2}} \leq \sigma e^{-a_j t}K_1. \tag{B.21}$$

By (B.19) and (B.21), we have

$$\begin{aligned}
\frac{K_\varphi}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|(\tilde{X}_t - X_t)|]dt &\le \frac{K_\varphi}{T}\int_{\phi_{j-1}T}^{\phi_j T}\left(e^{-a_j t}\mathrm{E}[|X_0|] + e^{-a_j t}K_\mu K_\varphi\frac{1}{a_j} + \sigma e^{-a_j t}K_1\right)dt \\
&= \frac{K_\varphi}{T}(\mathrm{E}[|X_0|] + K_\mu K_\varphi\frac{1}{a_j} + \sigma K_1)\int_{\phi_{j-1}T}^{\phi_j T} e^{-a_j t}dt \\
&= \frac{K_\varphi}{T}(\mathrm{E}[|X_0|] + K_\mu K_\varphi\frac{1}{a_j} + \sigma K_1)\frac{1}{a_j}(e^{-a_j\phi_{j-1}T} - e^{-a_j\phi_j T}).
\end{aligned}$$

Then,

$$\frac{K_\varphi}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|(\tilde{X}_t - X_t)|]dt \to 0 \quad as\ T \to \infty.$$

Therefore,

$$0 \le \lim_{T\to\infty}\mathrm{P}\left(\left\|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}X_t\varphi(t)dt\right\| \ge \xi\right) \le 0,$$

which implies

$$\lim_{T\to\infty}\mathrm{P}\left(\left\|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}X_t\varphi(t)dt\right\| \ge \xi\right) = 0.$$

Then, we have

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}X_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \tag{B.22}$$

Further, we have

$$\begin{aligned}
\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\tilde{X}_t\varphi(t)dt &= (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\phi_j T}\tilde{X}_t\varphi(t)dt \\
&= (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\sum_{i=\lfloor\phi_{j-1}T\rfloor+1}^{\lfloor\phi_j T\rfloor}\int_{i-1}^{i}\tilde{X}_t\varphi(t)dt \\
&+ (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor\phi_{j-1}T\rfloor+1}\tilde{X}_t\varphi(t)dt \\
&+ (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\lfloor\phi_j T\rfloor}^{\phi_j T}\tilde{X}_t\varphi(t)dt.
\end{aligned}$$

Now, we define $Y_i = \int_{i-1}^{i} \tilde{X}_t \varphi(t)dt$, and let $u = t - i + 1 \in [0, 1]$. According to Assumption 2, $\varphi(u + i - 1) = \varphi(u)$. Then, we have

$$Y_i = \int_{i-1}^{i} \tilde{X}_t \varphi(t)dt = \int_0^1 \tilde{X}_{u+i-1} \varphi(u + i - 1)du = \int_0^1 \tilde{X}_{u+i-1} \varphi(u)du.$$

According to Lemma 2.2, $\{\tilde{X}_{u+i-1}\}_{i\in\mathbb{N}}$ is a stationary and ergodic process. Since Y_i is measurable function of the stationary and ergodic process $\{\tilde{X}_{u+i-1}\}_{i\in\mathbb{N}}$, by Theorem 3.5.8 in Stout (1974) (see also Theorem A.3), $\{Y_i\}_{i\in\mathbb{N}}$ is stationary and ergodic process. By Birkhoff Ergodic Theorem (see also Theorem A.4), and since $\{Y_i\}_{i\in\mathbb{N}}$ is stationary and ergodic process,

$$\begin{aligned}
&(\phi_j - \phi_{j-1}) \frac{1}{(\phi_j - \phi_{j-1})T} \sum_{i=\lfloor \phi_{j-1}T \rfloor + 1}^{\lfloor \phi_j T \rfloor} \int_{i-1}^{i} \tilde{X}_t \varphi(t)dt \\
&= (\phi_j - \phi_{j-1}) \frac{1}{(\phi_j - \phi_{j-1})T} \sum_{i=\lfloor \phi_{j-1}T \rfloor + 1}^{\lfloor \phi_j T \rfloor} Y_i \\
&= (\phi_j - \phi_{j-1}) \frac{1}{(\phi_j - \phi_{j-1})T} \sum_{i=1}^{\lfloor (\phi_j - \phi_{j-1})T \rfloor} Y_i \\
&\xrightarrow[T\to\infty]{a.s} (\phi_j - \phi_{j-1}) \mathrm{E}\left[\int_0^1 \tilde{X}_t \varphi(t)dt \right],
\end{aligned}$$

which implies

$$(\phi_j - \phi_{j-1}) \frac{1}{(\phi_j - \phi_{j-1})T} \sum_{i=\lfloor \phi_{j-1}T \rfloor + 1}^{\lfloor \phi_j T \rfloor} \int_{i-1}^{i} \tilde{X}_t \varphi(t)dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1}) \mathrm{E}\left[\int_0^1 \tilde{X}_t \varphi(t)dt \right]. \tag{B.23}$$

By Jensen's Inequality,

$$
\begin{aligned}
&\mathrm{E}\left[\left\|(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor + 1} \tilde{X}_t\varphi(t)dt\right\|\right] \\
&\leq \mathrm{E}\left[(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor + 1} |\tilde{X}_t|\|\varphi(t)\|dt\right] \\
&\leq (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\mathrm{E}\left[\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor + 1} |\tilde{X}_t|K_\varphi dt\right] \\
&= (\phi_j - \phi_{j-1})\frac{K_\varphi}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor + 1} \mathrm{E}[|\tilde{X}_t|]dt.
\end{aligned}
$$

By (B.27), $\mathrm{E}[|\tilde{X}_t|^2] \leq K' < \infty$, $t \geq 0$ (in the Proof of Proposition 3.4).

By Cauchy Schwartz Inequality, $\mathrm{E}[|\tilde{X}_t|] \leq \mathrm{E}[|\tilde{X}_t|^2]^{\frac{1}{2}}$, which implies $\mathrm{E}[|\tilde{X}_t|]$ is bounded.

Then, we have

$$
(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor + 1} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{L^1} 0,
$$

which implies

$$
(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor + 1} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \tag{B.24}
$$

Similarly, we also have

$$
(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\lfloor \phi_j T\rfloor}^{\phi_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{L^1} 0,
$$

which implies

$$
(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\lfloor \phi_j T\rfloor}^{\phi_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \tag{B.25}
$$

By (B.23),(B.24) and (B.25),

$$
\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\mathrm{E}\left[\int_0^1 \tilde{X}_t\varphi(t)dt\right]. \tag{B.26}
$$

We know

$$
\mathrm{E}\left[\int_0^1 \tilde{X}_t\varphi_(t)dt\right] = \int_0^1 \mathrm{E}[\tilde{X}_t]\varphi(t)dt,
$$

and, for $\tau_{j-1} < t \le \tau_j$,

$$\mathrm{E}[\tilde{X}_t] = \tilde{h}_j(t) + \mathrm{E}[\tilde{z}_j(t)].$$

We observe that

$$\begin{aligned}\mathrm{E}[\tilde{z}_j(t)] &= \mathrm{E}\left[\sigma e^{-a_j t}\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} + \sigma e^{-a_j t}\int_0^t e^{a_j s} dB_s\right] \\ &= \sigma e^{-a_j t}\mathrm{E}\left[\int_0^{\infty} e^{-a_j u} d\bar{B}_u\right] + \sigma e^{-a_j t}\mathrm{E}\left[\int_0^t e^{a_j s} dB_s\right].\end{aligned}$$

By L^2-Bounded Martingale Convergence Theorem, (B.9) and (B.10), $I_U \xrightarrow[T\to\infty]{L^2} I_\infty = \int_0^\infty e^{-a_j u} d\bar{B}_u$, which implies $I_U \xrightarrow[T\to\infty]{L^1} I_\infty = \int_0^\infty e^{-a_j u} d\bar{B}_u$. Then, by the Martingale of It$\hat{o}$'s integral,

$$\mathrm{E}\left[\int_0^\infty e^{-a_j u} d\bar{B}_u\right] = \lim_{U\to\infty} \mathrm{E}\left[\int_0^U e^{-a_j u} d\bar{B}_u\right] = 0,$$

and

$$\mathrm{E}\left[\int_0^t e^{a_j s} dB_s\right] = 0.$$

Hence, $\mathrm{E}[\tilde{z}_j(t)] = 0$. Then, $\mathrm{E}[\tilde{X}_t] = \tilde{h}_j(t)$. By (B.26),

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t \varphi(t) dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\int_0^1 \tilde{h}_j(t)\varphi(t) dt.$$

Also, by (B.22), we have

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t \varphi(t) dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t \varphi(t) dt \xrightarrow[T\to\infty]{P} 0.$$

Therefore, we establish

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t \varphi(t) dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\int_0^1 \tilde{h}_j(t)\varphi(t) dt.$$

This completes the proof. □

Proof of Proposition 3.4. By Markov Inequality, for $\xi > 0$ and $\tau_{j-1} < t \leq \tau_j$,

$$\mathrm{P}\left(\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt\right| \geq \xi\right) \leq \frac{\mathrm{E}\left[\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)dt\right|\right]}{\xi}.$$

By Jensen's Inequality,

$$\begin{aligned}
&\mathrm{E}\left[\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)dt\right|\right] = \frac{1}{T}\mathrm{E}\left[\left|\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)dt\right|\right] \\
&\leq \frac{1}{T}\mathrm{E}\left[\int_{\phi_{j-1}T}^{\phi_j T}|(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)|dt\right] = \frac{1}{T}\mathrm{E}\left[\int_{\phi_{j-1}T}^{\phi_j T}|\tilde{X}_t - X_t||\tilde{X}_t + X_t|dt\right].
\end{aligned}$$

Then, by Triangle Inequality,

$$\begin{aligned}
&\mathrm{E}\left[\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)dt\right|\right] \leq \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t - X_t||\tilde{X}_t + X_t|]dt \\
&\leq \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t||\tilde{X}_t - X_t|]dt + \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|X_t||\tilde{X}_t - X_t|]dt.
\end{aligned}$$

Then, by Cauchy Swartz Inequality,

$$\begin{aligned}
&\mathrm{E}\left[\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)dt\right|\right] \\
&\leq \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t|^2]^{\frac{1}{2}}\mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}}dt + \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|X_t|^2]^{\frac{1}{2}}\mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}}dt.
\end{aligned}$$

We know the unique strong solution X_t is bounded in L^2, i.e., $\sup \mathrm{E}[|X_t|^2] < \infty$. Since $(x+y)^2 + (x-y)^2 = 2x^2 + 2y^2$ and $(x-y)^2 \geq 0$, $(x+y)^2 \leq 2x^2 + 2y^2$. Then, for $\tau_{j-1} < t \leq \tau_j$,

$$\mathrm{E}[|\tilde{X}_t|^2] = \mathrm{E}[(\tilde{h}_j(t) + \tilde{z}_j(t))^2] \leq 2\mathrm{E}[(\tilde{h}_j(t))^2] + 2\mathrm{E}[(\tilde{z}_j(t))^2].$$

By (B.6) and (B.13), we have

$$\mathrm{E}[|\tilde{X}_t|^2] \leq K' < \infty, \ \forall\ t \geq 0. \tag{B.27}$$

By Lemma 2.1, $\sup_{t\geq 0} \mathrm{E}[X_t^2] < \infty$.

Now, we denote $\sup_{t\geq 0}\{\mathrm{E}[X_t^2]^{\frac{1}{2}}, \mathrm{E}[\tilde{X}_t^2]^{\frac{1}{2}}\} \leq K_2 < \infty$. Then,

$$\begin{aligned}
&\mathrm{E}\left[\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T}(\tilde{X}_t - X_t)(\tilde{X}_t + X_t)dt\right|\right] \\
&\leq \frac{K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}}dt + \frac{K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}}dt \\
&= \frac{2K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}}dt.
\end{aligned}$$

From the convexity of the quadratic function, $(a+b+c)^2 \leq 3a^2 + 3b^2 + 3c^2$. Hence,

$$\frac{2K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}}dt = \frac{2K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T}\mathrm{E}[|\tilde{h}_j(t) + \tilde{z}_j(t) - e^{-a_j t}X_0 - h_j(t) - z_j(t)|^2]^{\frac{1}{2}}dt.$$

We have

$$\begin{aligned}
&\mathrm{E}[|\tilde{h}_j(t) + \tilde{z}_j(t) - e^{-a_j t}X_0 - h_j(t) - z_j(t)|^2] \\
&\leq 3\mathrm{E}[|\tilde{h}_j(t) - h_j(t)|^2] + 3\mathrm{E}[|\tilde{z}_j(t) - z_j(t)|^2] + 3\mathrm{E}[|e^{-a_j t}X_0|^2].
\end{aligned}$$

We observe that

$$\begin{aligned}
\mathrm{E}[|\tilde{h}_j(t) - h_j(t)|^2] &= \mathrm{E}\left[\left|e^{-a_j t}\sum_{k=1}^{p}\mu_{k,j}\int_{-\infty}^{0}e^{a_j s}\varphi_k(s)ds\right|^2\right] \\
&\leq e^{-2a_j t}\left(\sum_{k=1}^{p}\mu_{k,j}\right)^2 K_\varphi^2\left(\int_{-\infty}^{0}e^{a_j s}ds\right)^2.
\end{aligned}$$

Then,

$$\mathrm{E}[|\tilde{h}_j(t) - h_j(t)|^2] \leq e^{-2a_j t}K_\mu^2 K_\varphi^2\frac{1}{a_j^2}. \tag{B.28}$$

Also, we observe that

$$\mathrm{E}[|\tilde{z}_j(t) - z_j(t)|^2] = \mathrm{E}\left[\left(\sigma e^{-a_j t}\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right] = \sigma^2 e^{-2a_j t}\mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right].$$

Then, from (B.9)-(B.11), let $\mathrm{E}\left[\left(\int_{-\infty}^{0}e^{a_j s}d\tilde{B}_s\right)^2\right]^{\frac{1}{2}} \leq K_1$,

$$\mathrm{E}[|\tilde{z}_j(t) - z_j(t)|^2] \leq \sigma^2 e^{-2a_j t}K_1^2. \tag{B.29}$$

Hence, by (B.28) and (B.29),

$$\begin{aligned}
&\frac{2K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T} \mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}} dt \\
&\le \frac{2K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T} (3e^{-2a_j t} K_\mu^2 K_\varphi^2 \frac{1}{a_j^2} + 3\sigma^2 e^{-2a_j t} K_1^2 + 3e^{-2a_j t}\mathrm{E}[X_0^2])^{\frac{1}{2}} dt \\
&= \frac{2K_2}{T}(3K_\mu^2 K_\varphi^2 \frac{1}{a_j^2} + 3\sigma^2 K_1^2 + 3\mathrm{E}[X_0^2])^{\frac{1}{2}} \int_{\phi_{j-1}T}^{\phi_j T} e^{-a_j t} dt \\
&= \frac{2K_2}{T}(3K_\mu^2 K_\varphi^2 \frac{1}{a_j^2} + 3\sigma^2 K_1^2 + 3\mathrm{E}[X_0^2])^{\frac{1}{2}} \frac{1}{a_j}(e^{-a_j\phi_{j-1}T} - e^{-a_j \phi_j T}).
\end{aligned}$$

Then,

$$\frac{2K_2}{T}\int_{\phi_{j-1}T}^{\phi_j T} \mathrm{E}[|\tilde{X}_t - X_t|^2]^{\frac{1}{2}} dt \to 0 \ as \ T \to \infty.$$

Therefore,

$$0 \le \lim_{T\to\infty} \mathrm{P}\left(\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt\right| \ge \xi\right) \le 0.$$

Then, we have

$$\lim_{T\to\infty} \mathrm{P}\left(\left|\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt\right| \ge \xi\right) = 0,$$

which implies that

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{B.30}$$

Further,

$$\begin{aligned}
\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt &= (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \\
&= (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\sum_{i=\lfloor \phi_{j-1}T\rfloor+1}^{\lfloor \phi_j T\rfloor}\int_{i-1}^{i} \tilde{X}_t^2 dt \\
&+ (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor \phi_{j-1}T\rfloor+1} \tilde{X}_t^2 dt \\
&+ (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\lfloor \phi_j T\rfloor}^{\phi_j T} \tilde{X}_t^2 dt.
\end{aligned}$$

According to Lemma 2.2, $\{\tilde{X}_{u+i-1}\}_{i\in\mathbb{N}}$ is stationary and ergodic process. Let $Y_i = \int_{i-1}^{i} \tilde{X}_t^2 dt$ which is a measurable function. Then, by Theorem 3.5.8 in Stout (1974) (see also Theorem A.3), $\{Y_i\}_{i\in\mathbb{N}}$ is stationary and ergodic process. By Birkhoff Ergodic Theorem (see also Theorem A.4),

$$\begin{aligned}
&(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\sum_{i=\lfloor\phi_{j-1}T\rfloor+1}^{\lfloor\phi_j T\rfloor}\int_{i-1}^{i}\tilde{X}_t^2 dt \\
&= (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\sum_{i=1}^{\lfloor(\phi_j-\phi_{j-1})T\rfloor}\int_{i-1}^{i}\tilde{X}_t^2 dt \\
&\xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})\mathrm{E}\left[\int_0^1 \tilde{X}_t^2 dt\right],
\end{aligned}$$

which implies

$$(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\sum_{i=\lfloor\phi_{j-1}T\rfloor+1}^{\lfloor\phi_j T\rfloor}\int_{i-1}^{i}\tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\mathrm{E}\left[\int_0^1 \tilde{X}_t^2 dt\right]. \tag{B.31}$$

By Jensen's Inequality,

$$\begin{aligned}
&\mathrm{E}\left[\left|(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor\phi_{j-1}T\rfloor+1}\tilde{X}_t^2 dt\right|\right] \\
&\le \mathrm{E}\left[(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor\phi_{j-1}T\rfloor+1}|\tilde{X}_t^2| dt\right] \\
&= (\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor\phi_{j-1}T\rfloor+1}\mathrm{E}[\tilde{X}_t^2] dt.
\end{aligned}$$

Since $\mathrm{E}[|\tilde{X}_t|^2] \le K' < \infty,\ t \ge 0$, $\mathrm{E}[\tilde{X}_t^2]$ is bounded.

Hence,

$$(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor\phi_{j-1}T\rfloor+1}\tilde{X}_t^2 dt \xrightarrow[T\to\infty]{L^1} 0,$$

which implies

$$(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\phi_{j-1}T}^{\lfloor\phi_{j-1}T\rfloor+1}\tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{B.32}$$

Similarly, we also have

$$(\phi_j - \phi_{j-1})\frac{1}{(\phi_j - \phi_{j-1})T}\int_{\lfloor \phi_j T \rfloor}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{B.33}$$

By (B.31), (B.32) and (B.33),

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\mathrm{E}\bigg[\int_0^1 \tilde{X}_t^2 dt\bigg]. \tag{B.34}$$

We know that, for $\tau_{j-1} < t \leq \tau_j$,

$$\mathrm{E}\bigg[\int_0^1 \tilde{X}_t^2 dt\bigg] = \mathrm{E}\bigg[\int_0^1 (\tilde{h}_j(t) + \tilde{z}_j(t))^2 dt\bigg].$$

By the Proof of Proposition 3.3, we have $\mathrm{E}[\tilde{z}_j(t)] = 0$. Also, by (B.13), $\mathrm{E}[\tilde{z}_j^2(t)] = \dfrac{\sigma^2}{2a_j}$.

Then,

$$\begin{aligned}\mathrm{E}\bigg[\int_0^1 \tilde{X}_t^2 dt\bigg] &= \mathrm{E}\bigg[\int_0^1 (\tilde{h}_j(t))^2 + 2\tilde{h}_j(t)\tilde{z}_j(t) + (\tilde{z}_j(t))^2 dt\bigg] \\ &= \int_0^1 (\tilde{h}_j(t))^2 dt + \int_0^1 \frac{\sigma^2}{2a_j} dt = \int_0^1 \tilde{h}_j^2(t) dt + \frac{\sigma^2}{2a_j}.\end{aligned}$$

Hence,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\bigg(\int_0^1 \tilde{h}_j^2(t) dt + \frac{\sigma^2}{2a_j}\bigg).$$

Also, by (B.30), we know

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} 0.$$

Combining these two results, we establish that

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\bigg(\int_0^1 \tilde{h}_j^2(t) dt + \frac{\sigma^2}{2a_j}\bigg),$$

this completes the proof. □

Proof of Proposition 3.5. To check the positive definiteness of Σ_j, by Schur Complement Theorem, we need to show that $\omega_j - \Lambda^{(j)T} I_P^{-1} \Lambda_j > 0$. Then, we just need to show

$$\int_0^1 \tilde{h}_j^2(t)dt + \frac{\sigma^2}{2a_j} - \sum_{k=1}^{p} \left(\int_0^1 \tilde{h}_j(t)\varphi_k(t)dt \right)^2 > 0.$$

By Bessel's Inequality (see also Lemma A.1),

$$\int_0^1 \tilde{h}_j^2(t)dt \geq \sum_{k=1}^{p} \left(\int_0^1 \tilde{h}_j(t)\varphi_k(t)dt \right)^2$$

$$\int_0^1 \tilde{h}_j^2(t)dt - \sum_{k=1}^{p} \left(\int_0^1 \tilde{h}_j(t)\varphi_k(t)dt \right)^2 \geq 0,$$

then,

$$\int_0^1 \tilde{h}_j^2(t)dt + \frac{\sigma^2}{2a_j} - \sum_{k=1}^{p} \left(\int_0^1 \tilde{h}_j(t)\varphi_k(t)dt \right)^2 \geq \frac{\sigma^2}{2a_j} > 0.$$

Therefore, Σ_j is positive definite. Further, since Σ is block diagonal matrix whose block components matrices are positive definite, it is also positive definite matrix, this completes the proof. □

Proof of Proposition 3.6. By Proposition 3.2,

$$\frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \varphi(t)\varphi^\top(t)dt \xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})I_p,$$

which implies

$$\frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \varphi(t)\varphi^\top(t)dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})I_p,$$

and combining Proposition 3.3 and Proposition 3.4, we have

$$\frac{1}{T} Q_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\Sigma_j.$$

By Proposition 2.4, we have, for $T > 0$ and $j = 1, ..., m+1$, $\frac{1}{T} Q_{(\tau_{j-1},\tau_j)}$ is positive definite and, by Proposition 3.5, Σ_j is also positive definite. Then, by continuous mapping theorem,

$$TQ^{-1}_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{P} \frac{1}{\phi_j - \phi_{j-1}}(\Sigma_j)^{-1},$$

this completes the first part of proof. Further, by Proposition 2.4, we have, for $T > 0$ and $j = 1, ..., m+1$, $Q_{(\tau_{j-1},\tau_j)}$ is positive definite and, by proposition 3.5, Σ_j is also positive definite. Then, the block diagonal matrix $Q(\phi, m)$ and Σ are positive definite so they are invertible. Hence, since $TQ^{-1}_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{P} \frac{1}{\phi_j-\phi_{j-1}}(\Sigma_j)^{-1}$ for $j = 1, 2, \ldots, m+1$, we have

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

This completes the proof. □

Proof of Proposition 3.7. By Proposition 1.21 in Kutoyants (2004), it is a special case of Proposition 1.21 for which $d_1 = (m+1)(p+1)$ and $d_2 = 1$. Then, we have,

$$\int_0^T \frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\frac{1}{\sqrt{T}}\varphi^\top(t)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}dt = \frac{1}{T}\int_{\tau_{j-1}}^{\tau_j} \varphi(t)\varphi^\top(t)dt, \qquad \text{(B.35)}$$

$$\int_0^T \Big(\frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)\Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)dt = -\frac{1}{T}\int_{\tau_{j-1}}^{\tau_j} \varphi(t)X_t dt, \qquad \text{(B.36)}$$

and

$$\int_0^T \Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)\Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)dt = \frac{1}{T}\int_{\tau_{j-1}}^{\tau_j} X_t^2 dt. \qquad \text{(B.37)}$$

Then, we can see that (B.35),(B.36) and (B.37) are the elements of $\frac{1}{T}Q_{(\tau_{j-1},\tau_j)}$.

By Proposition 3.6,

$$\frac{1}{T}Q_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})\Sigma_j.$$

Also, for $i \neq j$,

$$\int_0^T \frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{\tau_{i-1}<t\le\tau_i\}}\frac{1}{\sqrt{T}}\varphi^\top(t)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}dt = 0,$$

$$\int_0^T \Big(\frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{\tau_{i-1}<t\le\tau_i\}}\Big)\Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)dt = 0,$$

and

$$\int_0^T \Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{i-1}<t\le\tau_i\}}\Big)\Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)dt = 0$$

since $\mathbb{I}_{\{\tau_{i-1}<t\le\tau_i\}}\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}} = 0$ for $i \neq j$, $0 \le \phi_{i-1} < \phi_i \le 1$ and $0 \le \phi_{j-1} < \phi_j \le 1$.

Therefore,

$$\Sigma = \begin{bmatrix} \phi_1\Sigma_1 & 0 & \dots & 0 \\ 0 & (\phi_2-\phi_1)\Sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (1-\phi_m)\Sigma_{m+1} \end{bmatrix}.$$

Further, one can prove that, for $T > 0$, $j = 1, ...m+1$,

$$\mathrm{P}\Bigg(\int_0^T \Big(\frac{1}{\sqrt{T}}\varphi(t)\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)^2 dt < \infty\Bigg) = 1,$$

and

$$\mathrm{P}\Bigg(\int_0^T \Big(-\frac{1}{\sqrt{T}}X_t\mathbb{I}_{\{\tau_{j-1}<t\le\tau_j\}}\Big)^2 dt < \infty\Bigg) = 1.$$

This completes the proof. □

Proof of Proposition 3.8. From Proposition 3.1, we have

$\hat{\theta}(\phi, m) = \theta + \sigma Q^{-1}(\phi, m)R(\phi, m)$. Then,

$$\sqrt{T}(\hat{\theta}(\phi, m) - \theta) = \sigma\sqrt{T}Q^{-1}(\phi, m)R(\phi, m) = \sigma T Q^{-1}(\phi, m)\frac{1}{\sqrt{T}}R(\phi, m).$$

By Proposition 3.6,

$$\sigma T Q^{-1}(\phi, m) \xrightarrow[T\to\infty]{P} \sigma\Sigma^{-1}.$$

By Proposition 3.7,

$$\frac{1}{\sqrt{T}}R(\phi, m) \xrightarrow[T\to\infty]{d} r^* \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma).$$

Then, by Slutsky's Theorem,

$$\sqrt{T}(\hat{\theta}(\phi,m)-\theta)=\sigma T Q^{-1}(\phi,m)\frac{1}{\sqrt{T}}R(\phi,m)\xrightarrow[T\to\infty]{d}\sigma\Sigma^{-1}r^{*}=\rho.$$

We see that Σ is non-random and symmetric. Hence, by Proposition A.2 in Appendix A, we have $\rho\sim\mathcal{N}_{(m+1)(p+1)}(0,\sigma^2\Sigma^{-1})$, and then,

$$\sqrt{T}(\hat{\theta}(\phi,m)-\theta)\xrightarrow[T\to\infty]{d}\rho\sim\mathcal{N}_{(m+1)(p+1)}(0,\sigma^2\Sigma^{-1}).$$

This completes the proof. □

Proof of Proposition 3.9. Let $\log L(\theta, X_T)$ be the log-likelihood function of the stochastic process $X_T=\{X_t, 0\le t\le T\}$, which satisfies the SDE (2.5).

By introducing the Lagrange Multiplier λ into the log-likelihood function

$$\log L(\theta,X_t)=\frac{1}{\sigma^2}\theta^\top\tilde{R}(\phi,m)-\frac{1}{2\sigma^2}\theta^\top Q(\phi,m)\theta.$$

We have

$$\log L(\theta,\lambda,X_t)=\frac{1}{\sigma^2}\theta^\top\tilde{R}(\phi,m)-\frac{1}{2\sigma^2}\theta^\top Q(\phi,m)\theta+\lambda^\top(B\theta-r).$$

First, taking the first derivative with respect to θ and λ respectively, we have

$$\frac{\partial}{\partial\theta}\log L(\theta,\lambda,X_t)=\frac{1}{\sigma^2}\tilde{R}(\phi,m)-\frac{1}{\sigma^2}Q(\phi,m)\theta+B^\top\lambda$$

and

$$\frac{\partial}{\partial\lambda}\log L(\theta,\lambda,X_t)=B\theta-r.$$

Then, setting $\frac{\partial}{\partial\theta}\log L(\theta,\lambda,X_t)$ and $\frac{\partial}{\partial\lambda}\log L(\theta,\lambda,X_t)$ equal to 0,

$$\frac{1}{\sigma^2}\tilde{R}(\phi,m)-\frac{1}{\sigma^2}Q(\phi,m)\tilde{\theta}(\phi,m)+B^\top\hat{\lambda}=0 \tag{B.38}$$

$$B\tilde{\theta}(\phi,m)-r=0. \tag{B.39}$$

By (B.38),

$$\frac{1}{\sigma^2}Q(\phi,m)\tilde{\theta}(\phi,m) = \frac{1}{\sigma^2}\tilde{R}(\phi,m) + B^\top\hat{\lambda},$$

which shows that

$$\tilde{\theta}(\phi,m) = Q^{-1}(\phi,m)\tilde{R}(\phi,m) + \sigma^2 Q^{-1}(\phi,m)B^\top\hat{\lambda}. \tag{B.40}$$

Taking (B.40) into (B.39),

$$BQ^{-1}(\phi,m)\tilde{R}(\phi,m) + B\sigma^2 Q^{-1}(\phi,m)B^\top\hat{\lambda} = r$$

$$\sigma^2 BQ^{-1}(\phi,m)B^\top\hat{\lambda} = r - BQ^{-1}(\phi,m)\tilde{R}(\phi,m)$$

$$\hat{\lambda} = \frac{1}{\sigma^2}(BQ^{-1}(\phi,m)B^\top)^{-1}r - \frac{1}{\sigma^2}(BQ^{-1}(\phi,m)B^\top)^{-1}BQ^{-1}(\phi,m)\tilde{R}(\phi,m).$$

Taking $\hat{\lambda}$ into (B.40),

$$\begin{aligned}\tilde{\theta}(\phi,m) &= Q^{-1}(\phi,m)\tilde{R}(\phi,m) + \sigma^2 Q^{-1}(\phi,m)B^\top\frac{1}{\sigma^2}(BQ^{-1}(\phi,m)B^\top)^{-1}r \\ &\quad - \sigma^2 Q^{-1}(\phi,m)B^\top\frac{1}{\sigma^2}(BQ^{-1}(\phi,m)B^\top)^{-1}BQ^{-1}(\phi,m)\tilde{R}(\phi,m).\end{aligned}$$

Then,

$$\begin{aligned}\tilde{\theta}(\phi,m) &= Q^{-1}(\phi,m)\tilde{R}(\phi,m) + Q^{-1}(\phi,m)B^\top(BQ^{-1}(\phi,m)B^\top)^{-1}r \\ &\quad - Q^{-1}(\phi,m)B^\top(BQ^{-1}(\phi,m)B^\top)^{-1}BQ^{-1}(\phi,m)\tilde{R}(\phi,m),\end{aligned}$$

this gives

$$\tilde{\theta}(\phi,m) = \hat{\theta}(\phi,m) + Gr - GB\hat{\theta}(\phi,m)$$

where, $\hat{\theta}(\phi,m) = Q^{-1}(\phi,m)\tilde{R}(\phi,m)$ and $G = Q^{-1}(\phi,m)B^\top(BQ^{-1}(\phi,m)B^\top)^{-1}$.

Finally,

$$\tilde{\theta}(\phi,m) = \hat{\theta}(\phi,m) + Gr - GB\hat{\theta}(\phi,m) = \hat{\theta}(\phi,m) - G(B\hat{\theta}(\phi,m) - r),$$

this completes the proof. □

Proof of Proposition 3.10. By (3.6),

$$\sqrt{T}(\tilde{\theta}(\phi,m)-\theta)=(I_{(m+1)(p+1)}-GB)\sqrt{T}(\hat{\theta}(\phi,m)-\theta)-\sqrt{T}G(B\theta-r).$$

By Proposition 3.8, we have

$$\sqrt{T}(\hat{\theta}(\phi,m)-\theta)\xrightarrow[T\to\infty]{d}\rho\sim\mathcal{N}_{(m+1)(p+1)}(0,\sigma^2\Sigma^{-1}),$$

and by combining (3.7), (3.9), and Slutsky's Theorem,

$$\sqrt{T}(\tilde{\theta}(\phi,m)-\theta)\xrightarrow[T\to\infty]{d}(I_{(m+1)(p+1)}-G^*B)\rho-G^*r_0=\zeta.$$

Then, by Proposition A.2 in Appendix A,

$$\zeta_T(\phi,m)\xrightarrow[T\to\infty]{d}\zeta\sim\mathcal{N}_{(m+1)(p+1)}(-G^*r_0,\sigma^2(I_{(m+1)(p+1)}-G^*B)\Sigma^{-1}(I_{(m+1)(p+1)}-G^*B)^\top).$$

Note that

$$\begin{aligned}&\sigma^2(I_{(m+1)(p+1)}-G^*B)\Sigma^{-1}(I_{(m+1)(p+1)}-G^*B)^\top\\&=\sigma^2(\Sigma^{-1}-\Sigma^{-1}B^\top G^{*\top}-G^*B\Sigma^{-1}+G^*B\Sigma^{-1}B^\top G^{*\top}).\end{aligned}$$

And, since $G^*=\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}$, we get

$$G^*B\Sigma^{-1}B^\top G^{*\top}=\Sigma^{-1}B^\top(B\Sigma^{-1}B^\top)^{-1}B\Sigma^{-1}=\Sigma^{-1}B^\top G^{*\top}.$$

Hence,

$$\begin{aligned}&\sigma^2(I_{(m+1)(p+1)}-G^*B)\Sigma^{-1}(I_{(m+1)(p+1)}-G^*B)^\top\\&=\sigma^2(\Sigma^{-1}-\Sigma^{-1}B^\top G^{*\top}-G^*B\Sigma^{-1}+\Sigma^{-1}B^\top G^{*\top})\\&=\sigma^2(\Sigma^{-1}-G^*B\Sigma^{-1}).\end{aligned}$$

Finally, we get

$$\zeta_T(\phi,m)\xrightarrow[T\to\infty]{d}\zeta\sim\mathcal{N}_{(m+1)(p+1)}(-G^*r_0,\sigma^2(\Sigma^{-1}-G^*B\Sigma^{-1})).$$

This completes the proof. □

Proof of Lemma 4.1. We have

$$\begin{aligned}
&\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} Y_t dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} Y_t dt \\
&= \left(\frac{1}{T}\int_0^{\hat{\phi}_jT} Y_t dt - \frac{1}{T}\int_0^{\phi_jT} Y_t dt\right) - \left(\frac{1}{T}\int_0^{\hat{\phi}_{j-1}T} Y_t dt - \frac{1}{T}\int_0^{\phi_{j-1}T} Y_t dt\right)
\end{aligned}$$

By Lemma 3.1 in Nkurunziza and Zhang (2018), and since $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ are consistent estimators for ϕ_j and ϕ_{j-1},

$$\frac{1}{T}\int_0^{\hat{\phi}_jT} Y_t dt - \frac{1}{T}\int_0^{\phi_jT} Y_t dt \xrightarrow[T\to\infty]{L^1} 0, \tag{B.41}$$

and

$$\frac{1}{T}\int_0^{\hat{\phi}_{j-1}T} Y_t dt - \frac{1}{T}\int_0^{\phi_{j-1}T} Y_t dt \xrightarrow[T\to\infty]{L^1} 0. \tag{B.42}$$

Therefore, by (B.41) and (B.42), we have

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} Y_t dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} Y_t dt \xrightarrow[T\to\infty]{L^1} 0.$$

This completes the proof. □

Proof of Lemma 4.2. Let $\mathrm{I}_1 = \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{\phi}_jT} Y_t dW_t - \int_0^{\phi_jT} Y_t dW_t\right\|^2 \mathbb{I}_{(\hat{\phi}_j>\phi_j)}\right]$ and $\mathrm{I}_2 = \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{\phi}_jT} Y_t dW_t - \int_0^{\phi_jT} Y_t dW_t\right\|^2 \mathbb{I}_{(\hat{\phi}_j\leq\phi_j)}\right]$. Then, we have

$$\frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{\phi}_jT} Y_t dW_t - \int_0^{\phi_jT} Y_t dW_t\right\|^2\right] = \mathrm{I}_1 + \mathrm{I}_2. \tag{B.43}$$

For I_1, by Jensen's Inequality and It$\hat{o}$'s Isometry,

$$\begin{aligned}
\mathrm{I}_1 &= \frac{1}{T}\mathrm{E}\left[\left\|\int_{\phi_jT}^{\hat{\phi}_jT} Y_t dW_t\right\|^2 \mathbb{I}_{(\hat{\phi}_j>\phi_j)}\right] = \frac{1}{T}\mathrm{E}\left[\mathrm{E}\left[\left\|\int_{\phi_jT}^{\hat{\phi}_jT} Y_t dW_t\right\|^2 \mathbb{I}_{(\hat{\phi}_j>\phi_j)}\middle|\hat{\phi}_j\right]\right] \\
&= \frac{1}{T}\mathrm{E}\left[\left\|\int_{\phi_jT}^{\hat{\phi}_jT} Y_t Y_t^\top dt\right\| \mathbb{I}_{(\hat{\phi}_j>\phi_j)}\right] \leq \frac{1}{T}\mathrm{E}\left[\int_{\phi_jT}^{\hat{\phi}_jT} \left\|Y_t Y_t^\top\right\| dt\, \mathbb{I}_{(\hat{\phi}_j>\phi_j)}\right].
\end{aligned} \tag{B.44}$$

Let $\left\|Y_tY_t^\top\right\| \leq K_3$ for all $t \geq 0$. From (B.44),

$$\mathrm{I}_1 \leq \frac{1}{T}\mathrm{E}\left[(\hat{\phi}_j - \phi_j)TK_3\mathbb{I}_{(\hat{\phi}_j > \phi_j)}\right]. \tag{B.45}$$

Similarly, we have

$$\mathrm{I}_2 \leq \frac{1}{T}\mathrm{E}\left[(\phi_j - \hat{\phi}_j)TK_3\mathbb{I}_{(\hat{\phi}_j \leq \phi_j)}\right]. \tag{B.46}$$

By (B.43), (B.45) and (B.46), we establish that $\frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{\phi}_jT} Y_t dW_t - \int_0^{\phi_jT} Y_t dW_t\right\|^2\right] \leq K_3\mathrm{E}\left[|\hat{\phi}_j - \phi_j|\right]$. Then, by Lebesgue's dominated convergence theorem, $\frac{1}{\sqrt{T}}\int_0^{\hat{\phi}_jT} Y_t dW_t - \frac{1}{\sqrt{T}}\int_0^{\phi_jT} Y_t dW_t \xrightarrow[T\to\infty]{L^2} 0$. Similarly, we have $\frac{1}{\sqrt{T}}\int_0^{\hat{\phi}_{j-1}T} Y_t dW_t - \frac{1}{\sqrt{T}}\int_0^{\phi_{j-1}T} Y_t dW_t \xrightarrow[T\to\infty]{L^2} 0$. By combining these two conditions, this completes the proof. □

Proof of Proposition 4.1. We have, for $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$,

$$\begin{aligned}\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \varphi(t)\varphi^\top(t)dt = &\left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \varphi(t)\varphi^\top(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \varphi(t)\varphi^\top(t)dt\right)\\ &+ \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \varphi(t)\varphi^\top(t)dt.\end{aligned}$$

By Lemma 4.1, $\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \varphi(t)\varphi^\top(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \varphi(t)\varphi^\top(t)dt \xrightarrow[T\to\infty]{P} 0$. Further, by Proposition 3.2, $\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \varphi(t)\varphi^\top(t)dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})I_p$. Then, combining these two conditions, for $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$,

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \varphi(t)\varphi^\top(t)dt \xrightarrow[T\to\infty]{P} (\phi_j - \phi_{j-1})I_p.$$

This completes the proof. □

Proof of Proposition 4.2. For $T > 0$, $0 \leq \phi_{j-1} < \phi_j \leq 1$ where $j = 1, ..., m+1$, we have

$$\begin{aligned}
&\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t\varphi(t)dt \\
&= \left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t\varphi(t)dt\right) \\
&+ \left(\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt\right) \\
&+ \left(\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t\varphi(t)dt\right). \qquad \text{(B.47)}
\end{aligned}$$

By Lemma 4.1, we have

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t\varphi(t)dt \xrightarrow[T\to\infty]{L^1} 0,$$

and

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{L^1} 0,$$

which implies that

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0, \qquad \text{(B.48)}$$

and

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \qquad \text{(B.49)}$$

Also, by the Proof of Proposition 3.3,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0. \qquad \text{(B.50)}$$

Hence, combining (B.47), (B.48), (B.49) and (B.50),

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t\varphi(t)dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{P} 0.$$

Similarly, we observe that

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t^2 dt = \left(\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t^2 dt\right) + \left(\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t^2 dt\right) + \left(\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t^2 dt\right). \tag{B.51}$$

By Lemma 4.1, we have

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t^2 dt \xrightarrow[T\to\infty]{L^1} 0,$$

and

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{L^1} 0,$$

which implies that

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t^2 dt \xrightarrow[T\to\infty]{P} 0, \tag{B.52}$$

and

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{B.53}$$

Also, by the Proof of Proposition 3.4,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} X_t^2 dt - \frac{1}{T}\int_{\phi_{j-1}T}^{\phi_jT} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0. \tag{B.54}$$

Hence, combining (B.51), (B.52), (B.53) and (B.54),

$$\frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} X_t^2 dt - \frac{1}{T}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_jT} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{P} 0.$$

This completes the proof. □

Proof of Proposition 4.6. By Proposition 2.4, $Q_{(\hat{\tau}_{j-1},\hat{\tau}_j)}$ is positive definite and, by Proposition 3.5, Σ_j is also positive definite. Then, the block matrix $\frac{1}{T}Q(\hat{\phi}, m)$ and Σ are positive definite so they are invertible. We have

$$(\frac{1}{T}Q(\hat{\phi}, m))^{-1} = TQ^{-1}(\hat{\phi}, m).$$

Hence, by Proposition 4.5,

$$TQ^{-1}(\hat{\phi}, m) \xrightarrow[T\to\infty]{P} \Sigma^{-1}$$

which completes the proof. □

Proof of Lemma 4.3. We have

$$\frac{1}{\sqrt{T}}\left(\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} Y_s dW_s - \int_{\phi_{j-1}T}^{\phi_j T} Y_s dW_s\right)$$
$$= \frac{1}{\sqrt{T}}\left(\int_0^{\hat{\phi}_j T} Y_s dW_s - \int_0^{\phi_j T} Y_s dW_s\right) - \frac{1}{\sqrt{T}}\left(\int_0^{\hat{\phi}_{j-1} T} Y_s dW_s - \int_0^{\phi_{j-1} T} Y_s dW_s\right).$$

Then, by Lemma 3.3 in Nkurunziza and Zhang (2018), we get

$$\frac{1}{\sqrt{T}}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} Y_s dW_s - \frac{1}{\sqrt{T}}\int_{\phi_{j-1}T}^{\phi_j T} Y_s dW_s \xrightarrow[T\to\infty]{P} 0,$$

this completes the proof. □

Proof of Proposition 4.7. We know that $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ are consistent estimators for ϕ_j and ϕ_{j-1} where $j = 1, ..., m+1$. By Lemma 4.3, we have, for $0 \leq \phi_{j-1} < \phi_j \leq 1$, $j = 1, ..., m+1$, $\frac{1}{\sqrt{T}}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} X_t dW_t - \frac{1}{\sqrt{T}}\int_{\phi_{j-1}T}^{\phi_j T} X_t dW_t \xrightarrow[T\to\infty]{P} 0$. Then, by Lemma 4.2, we also have $\frac{1}{\sqrt{T}}\int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \varphi(t) dW_t - \frac{1}{\sqrt{T}}\int_{\phi_{j-1}T}^{\phi_j T} \varphi(t) dW_t \xrightarrow[T\to\infty]{P} 0$. Combining these two conditions completes the proof. □

Proof of Corollary 4.1. From Proposition 3.1, we know

$$\hat{\theta}(\phi, m) = \theta + \sigma Q^{-1}(\phi, m) R(\phi, m).$$

Similarly, we have

$$\hat{\theta}(\hat{\phi}, m) = \theta + \sigma Q^{-1}(\hat{\phi}, m) R(\hat{\phi}, m).$$

Then,

$$\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) = \sigma\sqrt{T} Q^{-1}(\hat{\phi}, m) R(\hat{\phi}, m) = \sigma T Q^{-1}(\hat{\phi}, m) \frac{1}{\sqrt{T}} R(\hat{\phi}, m).$$

By Proposition 4.6,

$$\sigma T Q^{-1}(\hat{\phi}, m) \xrightarrow[T\to\infty]{P} \sigma \Sigma^{-1}.$$

By Proposition 4.8,

$$\frac{1}{\sqrt{T}} R(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} r^* \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma).$$

Then, by Slutsky's Theorem,

$$\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) = \sigma T Q^{-1}(\hat{\phi}, m) \frac{1}{\sqrt{T}} R(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} \sigma \Sigma^{-1} r^* = \rho.$$

We see that Σ is non-random and symmetric. Hence, by Proposition A.2, we have $\rho \sim \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2 \Sigma^{-1})$, then

$$\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) \xrightarrow[T\to\infty]{d} \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2 \Sigma^{-1}),$$

this completes the proof. □

Alternative Proof of Theorem 7.2. From (7.2) and Proposition 5.1, we have

$$\begin{aligned}\mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) &= \mathrm{E}[\rho^\top \Omega \rho] = \mathrm{E}[\mathrm{trace}(\rho^\top \Omega \rho)] = \mathrm{E}[\mathrm{trace}(\Omega \rho \rho^\top)] = \mathrm{trace}(\Omega \mathrm{E}[\rho \rho^\top]) \\ &= \mathrm{trace}(\Omega(\mathrm{var}(\rho) + \mathrm{E}(\rho)\mathrm{E}(\rho)^\top)) = \mathrm{trace}(\Omega \mathrm{var}(\rho)) + \mathrm{E}(\rho)^\top \Omega \mathrm{E}(\rho).\end{aligned}$$

From Corollary 4.1, we have $\mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \sigma^2 \mathrm{trace}(\Omega \Sigma^{-1})$, this completes the proof. □

Alternative Proof of Theorem 7.3. By Proposition 5.1, we have

$$\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \mathrm{E}[\zeta^\top \Omega \zeta],$$

where $\zeta \sim \mathcal{N}_{(m+1)(p+1)}\left(-G^* r_0, \sigma^2(\Sigma^{-1} - G^* B \Sigma^{-1})\right)$. Following similar steps of the proof of Proposition 7.2, we have

$$\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \mathrm{E}[\zeta^\top \Omega \zeta] = \mathrm{trace}(\Omega \mathrm{var}(\zeta)) + \mathrm{E}(\zeta)^\top \Omega \mathrm{E}(\zeta).$$

This gives

$$\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \mathrm{trace}(\Omega \sigma^2(\Sigma^{-1} - G^* B \Sigma^{-1})) + (-G^* r_0)^\top \Omega (-G^* r_0).$$

Then, we have

$$\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \sigma^2 \mathrm{trace}(\Omega \Sigma^{-1}) - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) + r_0^\top G^{*\top} \Omega G^* r_0.$$

Therefore, by Theorem 7.2, we establish that

$$\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) + r_0^\top G^{*\top} \Omega G^* r_0.$$

This completes the proof. □

BIBLIOGRAPHY

[1] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, in B. Petrov, F. Csáki, 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, 267-281.

[2] Auger, I. & Lawrence, C. (1989). Algorithms for the optimal identification of segment neighborhoods, Bulletin of Mathematical Biology. 51(1), 39-54.

[3] Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes, Econometrica. 66(1), 47-78.

[4] Chen, F. Mamon, R., & Davison, M. (2017). Inference for a mean-reverting stochastic process with multiple change points. Electronic Journal of Statisitics, (Vol. 11), 2199-2257.

[5] Chen, F. & Nkurunziza, S. (2015). Optimal method in multiple regression with structural changes. Bernoulli 21 (4), 2217-2241.

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

[7] 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). 118.

[8] Gikhman, I. I. & Skorohod, A. V. (2004). The Theory of Stochastic Processes: I (Vol. 210). Springer-Verlag, Berlin Heidelberg.

[9] Koralov, L. & Sinai, Y. G. (2007). Theory of probability and random processes. Springer-Verlag, Berlin Heidelberg.

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

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

[12] Nkurunziza, S. (2012a). Shrinkage strategies in some multiple multi-factor dynamical systems. ESAIM: Probab Stat 16, 139-150.

[13] Nkurunziza, S. (2012b). The risk of pretest and shrinkage estimators. Statistics, 46 (3), 305-312.

[14] Nkurunziza, S. & Ahmed, SE. (2010). Shrinkage drift parameter estimation for multi-factor Ornstein-Uhlenbeck processes. ASMBI 26 (2), 103-124.

[15] Nkurunziza, S. & Zhang, P. (2018). Estimation and testing in generalized mean-reverting processes with change-point. Stat Inference Stoch Process, 21, 191-215.

[16] Nkurunziza, S. & Fu, K. (2018). Improved inference in generalized mean-reverting processes with multiple change-points, under review.

[17] Nkurunziza, S. & Shen, L. (2018). Inference in a multivariate generalized mean-reverting process with a change-point, under review.

[18] Perron, P. & Qu, Z. (2006). Estimating restricted structural change models. Journal of Econometrics 134, 373-399.

[18] Saleh, AME. (2006). Theory of preliminary test and Stein-type estimation with applications, vol 517. Wiley, New Jersey.

[20] Schwarz, G. (1978). Estimating the dimension of a model, Annals of Statistics.

6(2), 461-464.

[21] Sen, PK. & Saleh AKME. (1987). On preliminary test and shrinkage M-estimation in linear models. Ann Stat, 15(4), 1580-1592.

[22] Shen, L. (2018). Inference in multivariate generalized Ornstein-Uhlenbeck processes with a change-point. MSc, thesis, Department of Mathematics and Statistics, The University of Windsor, Windsor, Ontario, Canada.

[23] Steele, J. M. (2001). Stochastic calculus and financial applications (Vol. 45). Springer Science + Business Media, New York.

[24] Stout, W. F. (1974). Almost sure convergence (Vol. 154). Academic press, New York.

[25] Zhang, F. (2005). The Schur complement and its applications (Vol. 4). Springer-Verlag, U.S.

VITA AUCTORIS

NAME:	Kang Fu
PLACE OF BIRTH:	Beijing, China
YEAR OF BIRTH:	1992
EDUCATION:	High School Affiliated to Renmin University of China, Beijing, China, 2007-2011
	University of Waterloo, B.Sc., Waterloo, ON, 2011-2016
	University of Windsor, M.Sc., Windsor, ON, 2016-2018