Electronic Journal of Statistics
Vol. 13 (2019) 1400–1442
ISSN: 1935-7524
https://doi.org/10.1214/19-EJS1548

Improved inference in generalized mean-reverting processes with multiple change-points*

Sévérien Nkurunziza†

University of Windsor, Mathematics and Statistics department, 401 Sunset Avenue, Windsor, ON, N9B3P4, Canada e-mail: severien@uwindsor.ca

and

Kang Fu

University of Windsor, Mathematics and Statistics department, 401 Sunset Avenue, Windsor, ON, N9B3P4, Canada e-mail: fu11s@uwindsor.ca

Abstract: In this paper, we consider inference problem 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 restriction. As compared to the results in literature, we generalize some findings in five ways. First, we consider the model which incorporates the uncertain prior knowledge. Second, we derive the unrestricted estimator (UE) and the restricted estimator (RE) and we study their asymptotic properties. 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 encloses the UE, the 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. Finally, we present some simulation results which corroborate the established theoretical findings.

MSC 2010 subject classifications: Primary 62F30; secondary 62M02.
Keywords and phrases: ADR, change-point, drift-parameter, mean-reverting process, Ornstein-Uhleneck process, testing, SDE, shrinkage estimators, unrestricted estimator.

Received May 2018.

Contents

*The first author would like to acknowledge the financial support received from the Natural Sciences and Engineering Research Council of Canada (NSERC).

†Corresponding author: Sévérien Nkurunziza Email: severien@uwindsor.ca

1. Introduction

Nowadays, the Ornstein-Uhlenbeck (O-U) process is applied in different fields, such as physical sciences (Lansky and Sacerdote (2001)) [11] and biology (Rohlfs *et al.* (2010)) [17]. Such a process is also called mean-reverting process since the mean reverting level is the component which has large effect on it. For the classical O-U process, the mean reverting level is constant. However, the classical O-U process does not fit well the data whose mean reverting level may change with the time. This is particularly the case for some phenomena which heavily depend on the factors which change with the time. For instance, government policies are examples of the factors which affect 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 solve such a problem, Dehling *et al.* (2010) [6] proposed generalized Ornstein-Uhlenbeck processes which have a time-dependent periodic mean reverting function. Further, Dehling *et al.* (2014) [7], Nkurunziza and Zhang (2018) [15] considered inference problems in generalized O-U processes with change-point. The problem studied here was mainly inspired by the work in Chen *et al.* (2017) [4]. Namely, Chen *et al.* (2017) [4] proposed a method for detecting multiple change points in generalized O-U processes.

In this paper, we study the inference problem in generalized O-U processes with multiple unknown change points in the context where the drift parameter is suspected to satisfy some restrictions. The proposed method generalizes the work of Chen *et al.* (2017) [4] in five ways. First, we consider the model which incorporates the uncertain prior information. Second, we derive the unrestricted estimator (UE) and the restricted estimator (RE), and we study their asymptotic properties. Third, 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. Fourth, we construct a class of shrinkage estimators (SEs) which are expected to be robust with respect to the restriction. Fifth, we study the relative risk dominance of the proposed estimators. With respect to the established asymptotic power and risk dominance, the novelty of the proposed methods consist in the fact that the dimensions of the UE and RE are random variables. To overcome the difficulty due to the randomness of the dimensions of the UE and RE, we establish an asymptotic result which is of interest in its own. Further, we weaken some conditions underlying the main results in Chen *et al.* (2017) [4]. Specifically, we establish that the findings in Chen *et al.* (2017) [4] hold without their Assumption 2. We also provide a condition, about the initial value of the generalized O-U, which was omitted in Chen *et al.* (2017) [4] although required for their main results to hold.

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

2. Statistical model and preliminary results

In this section, we introduce the statistical model and set up some notations and assumptions. Inspired by the statistical model in Chen *et al.* (2017) [4], we study the inference problem about the drift parameter in generalized O-U processes. To introduce some notations, let $(\Omega, \mathfrak{F}, \mathrm{P})$ be a probability space where $\mathfrak{F}$ is $\mathfrak{S}$-field on the sample space Ω, and P is a probability measure. Further, let L^p denote the space of measurable p-integrable functions, for some $p \geqslant 1$. Let $\boldsymbol{A}'$ denote the transpose of a given matrix $\boldsymbol{A}$, let $\theta = (\theta_1', \ldots, \theta_{m+1}')'$ with $\theta_j = (\mu_{1,j}, \mu_{2,j}, \ldots, \mu_{p,j}, a_j)'$, for $j = 1, \ldots, m+1$, $k = 1, \ldots, p$, $\mu_{k,j}$ is real valued and $a_j > 0$. Let $\varphi_k(t)$ be real-valued function on $(0, T)$, $k = 1, \ldots, p$, and let $\boldsymbol{\varphi}(t) = (\varphi_1(t), \ldots, \varphi_p(t))'$, $t \geqslant 0$. Let $\{W_t, t \geqslant 0\}$ be a one-dimensional standard Brownian motion defined on (Ω, F, P) and let $\sigma > 0$. Let $\mathbb{I}_A$ be the indicator function of the event A. We also use the notations $\xrightarrow[T\to\infty]{d}$, $\xrightarrow[T\to\infty]{P}$, $\xrightarrow[T\to\infty]{a.s}$, and $\xrightarrow[T\to\infty]{L^p}$ to denote, respectively, the convergence in distribution, in probability, almost surely and in L^p-space, as T tends to infinity. Further, let $O_P(a(T))$ stand for a random quantity such that $O_P(a(T))a^{-1}(T)$ is bounded in probability and let $o_P(a(T))$ stand for a random quantity such that $o_P(a(T))a^{-1}(T)$ converges in probability to 0 as T tends to infinity. We say that a stochastic process $\{Y_t,\, t \geqslant 0\}$ is L^p-bounded if there exists $K > 0$ such that $\mathrm{E}\left(|Y_t|^p\right) < K$, for all

$t \geqslant 0$, for some $p \geqslant 1$.

As in Chen *et al.* (2017) [4], we consider that $\{X_t : t \geqslant 0\}$ is a solution of the stochastic differential equation (SDE)

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

where the drift coefficient, $S(\theta, t, X_t)$, is given by

$$S(\theta, t, X_t) = \sum_{j=1}^{m+1} \left(\boldsymbol{\varphi}'(t), -X_t\right)' \theta_j \mathbb{I}_{\{\tau_{j-1} < t \leqslant \tau_j\}}, \ 0 \leqslant t \leqslant T. \tag{2.2}$$

We assume that m ($m \geqslant 1$) is unknown as well as the change-points $\tau_1 < \tau_2 < \cdots < \tau_m$, $\tau_1 > 0$, and $\tau_m < T$. For the sake of simplicity, we suppose that $\tau_j = \phi_j T$, where $j = 1, \ldots, m$ and $0 < \phi_1 < \cdots < \phi_m < 1$. For mathematical convenience, let $\tau_0 = 0$, $\tau_{m+1} = T$, $\phi_0 = 0$, and let $\phi_{m+1} = 1$. The parameter of interest is θ while $\tau_1, \tau_2, \ldots, \tau_m$ and m are unknown nuisance parameters. In the sequel, let $\{\mathfrak{F}_t, t \geqslant 0\}$ denote the natural filtration of the Brownian motion given in the SDE (2.1).

In this paper, we also suppose that there exists a vague prior knowledge about the target parameter, θ. In particular, we consider the scenario where θ may satisfy the linear restriction $\boldsymbol{B}\theta = r$, where $\boldsymbol{B}$ 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. This restriction leads to the testing problem

$$H_0 : \boldsymbol{B}\theta = r \ \text{ vs } \ H_1 : \boldsymbol{B}\theta \neq r. \tag{2.3}$$

Particularly, let $\boldsymbol{I}_p$ be p-dimensional identity matrix, if we choose $r = 0$ and

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{I}_{p+1} & -\boldsymbol{I}_{p+1} & 0 & \ldots & 0 & 0 \\ 0 & \boldsymbol{I}_{p+1} & -\boldsymbol{I}_{p+1} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \boldsymbol{I}_{p+1} & -\boldsymbol{I}_{p+1} \end{pmatrix} = \boldsymbol{B}_0,$$

the restriction in (2.3) corresponds to the case where there are no change points. Thus, the testing problem in (2.3) includes as a special case testing the absence of change points. The optimality of the proposed method requires the following assumptions.

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 $\{W_t : t \geqslant 0\}$ and $\mathrm{E}[|X_0|^d] < \infty$, for some $d \geqslant 2$.*

Assumption 2. *The p-valued function $\boldsymbol{\varphi}(t)$ is square Riemann-integrable on $[0, T]$, for any $T > 0$, and satisfies*
(1) **Periodicity***: there exists $v > 0$ such that $\varphi(t + v) = \varphi(t)$, $\forall t > 0$ (v is the period);*
(2) **Orthonormality** *in $L^2([0, v], \frac{1}{v} d\lambda)$: $\int_0^v \varphi(t)\varphi'(t)dt = v\boldsymbol{I}_p$.*

Since, $\varphi(t)$ is bounded on $[0, T]$ and it is periodic, it is bounded on $\mathbb{R}_+$. As in Chen *et al.* (2017) [4], without loss of generality, we assume that $v = 1$.

Remark 2.1. It should be noticed that Assumption 1 is not explicitly mentioned in Chen et al. (2017) [4]. However, their main results require this assumption to hold. For example, if the distribution of X_0 depends on θ, from Theorem 1.12 of Kutoyants (2004, p 34) [10], it is clear that the likelihood function given in Chen et al. (2017, see p. 2204) [4] does not hold. Moreover, if $E[|X_0|^2] = +\infty$, the relation (3.8) in Chen et al. (2017, see p.2208) [4] does not hold.

The following lemma shows that the SDE in (2.1) and (2.2) admits a strong and unique solution.

Lemma 2.1. *Suppose that (2.1) and (2.2) hold along with Assumptions 1. Then, the SDE admits a strong and unique solution given by*

$$X_t = \sum_{j=1}^{m+1} X_j(t)\mathbb{I}_{(\tau_{j-1},\tau_j]}(t), \tag{2.4}$$

with, for $j = 1, 2, \dots, m+1$,

$$\begin{aligned} X_j(t) &= (e^{-a_j(t-\tau_{j-1})}X_{\tau_{j-1}} + h_j(t-\tau_{j-1}) + z_j(t-\tau_{j-1})), \quad t \geqslant \tau_{j-1}, \\ h_j(t) &= e^{-a_j t}\sum_{k=1}^{p}\mu_{k,j}\int_0^t e^{a_j s}\varphi_k(s+\tau_{j-1})ds, \quad t \geqslant 0, \\ z_j(t) &= \sigma e^{-a_j t}\int_0^t e^{a_j s}dW_s^{(\tau_{j-1})}, \quad W_t^{(u)} = W_{t+u} - W_u, \quad t \geqslant 0, \quad u \geqslant 0. \end{aligned} \tag{2.5}$$

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

The proof of this lemma is given in the Appendix. By using the last statement of Lemma 2.1, we establish the following corollary which is useful in deriving the likelihood function of the SDE in 2.1.

Corollary 2.1. *If Assumptions 1-2 hold, then,*

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

Proof. By some algebraic computations, we have

$$S^2(\theta, t, X_t) \leqslant (m+1)\|\theta\|^2\left(\|\boldsymbol{\varphi}(t)\|^2 + |X_t|^2\right),$$

for all $t \geqslant 0$. Then, from Lemma 2.1 and Assumption 2, we have
$\mathrm{E}\left(\int_0^T S^2(\theta, t, X_t)dt\right) < \infty$ then $\mathrm{P}\left(\int_0^T S^2(\theta, t, X_t)dt < \infty\right) = 1.$ $\square$

Remark 2.2. To make a connection with the work in Chen et al. (2017) [4] which inspired the most this paper, it should be noticed that Corollary 2.1 is given as Assumption 2 in Chen et al. (2017) [4]. Thus, Corollary 2.1 shows that the results established in Chen et al. (2017) [4] work without requiring their Assumption 2.

The process $\{X_t; t \geqslant 0\}$ is not stationary. Thus, it is convenient to introduce the following auxiliary stochastic processes. For $j = 1, ..., m+1$, let

$$\tilde{X}_t = \sum_{j=1}^{m+1} \tilde{X}_j(t)\mathbb{I}_{(\tau_{j-1},\tau_j]}(t) \quad \text{with} \quad \tilde{X}_j(t) = \tilde{h}_j(t) + \tilde{z}_j(t) \tag{2.6}$$

where the function $\tilde{h}_j(t) \ : \ [0,\infty] \to \mathbb{R}$ and the process $\{\tilde{z}_j(t) : t \geqslant 0\}$ are given by

$$\tilde{h}_j(t) = e^{-a_j t} \sum_{k=1}^{p} \mu_{k,j} \int_{-\infty}^{t} e^{a_j s}\varphi_k(s)ds, \ \tilde{z}_j(t) = \sigma e^{-a_j t} \int_{-\infty}^{t} e^{a_j s} d\tilde{B}_s, \tag{2.7}$$

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 $\{B_s\}_{s\geqslant 0}$ and $\{\bar{B}_{-s}\}_{s\geqslant 0}$ are two independent standard Brownian motion. The following theorem gives some relationship between $X(t)$ and $\tilde{X}(t)$.

Theorem 2.1. *Let* $a_{(1)} = \min\limits_{1\leqslant j\leqslant m+1} a_j$*. If Assumptions 1-2 hold, then*

(i) $|X(t) - \tilde{X}(t)| \leqslant C_0 e^{-a_{(1)}t}$ *where* C_0 *is a random variable such that* $\mathrm{E}(|C_0|^2) < \infty$*;*

(ii) $|X(t) - \tilde{X}(t)| \xrightarrow[t\to\infty]{a.s. \ and \ L^2} 0,$ *and* $|X^2(t) - \tilde{X}^2(t)| \xrightarrow[t\to\infty]{a.s. \ and \ L^1} 0;$

(iii) $\sup\limits_{0\leqslant a<b\leqslant 1} \left| \frac{1}{T}\int_{aT}^{bT} \tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{aT}^{bT} X_t\varphi(t)dt \right| \xrightarrow[T\to\infty]{a.s. \ and \ L^2} 0;$

(iv) $\sup\limits_{0\leqslant a<b\leqslant 1} \left| \frac{1}{T}\int_{aT}^{bT} \tilde{X}_t^2 dt - \frac{1}{T}\int_{aT}^{bT} X_t^2 dt \right| \xrightarrow[T\to\infty]{a.s. \ and \ L^1} 0.$

The proof of this theorem is given in the Appendix. To introduce some notations, let $\mathrm{diag}(A_1, A_2, \ldots, A_n)$ denote a diagonal matrix (or a block diagonal matrix) with components $A_1, A_2, \ldots, A_n$, let $\phi = (\phi_1, \phi_2, \ldots, \phi_m)'$, and let

$$\tilde{r}_{(a,b)} = \begin{pmatrix} \int_a^b \varphi(t)dX_t \\ -\int_a^b X_t dX_t \end{pmatrix}, \ Q_{(a,b)} = \begin{bmatrix} \int_a^b \varphi(t)\varphi'(t)dt & -\int_a^b X_t\varphi(t)dt \\ -\int_a^b \varphi'(t)X_t dt & \int_a^b X_t^2 dt \end{bmatrix}, a < b, \tag{2.8}$$

$$\tilde{R}(\phi, m) = (\tilde{r}_{(0,\tau_1)}, ..., (\tilde{r}_{(\tau_m,T)})', \text{ and}$$
$$Q(\phi, m) = \operatorname{diag}\left(Q_{(0,\tau_1)}, Q_{(\tau_1,\tau_2)}, \ldots, Q_{(\tau_m,T)}\right), \tag{2.9}$$
$$M_{(a,b)} = \left(\int_a^b \boldsymbol{\varphi}'(t)dW_t, -\int_a^b X_t dW_t\right)', \quad \boldsymbol{M}(\phi, m) = (M_{(0,\tau_1)}, ..., M_{(\tau_m,T)})'. \tag{2.10}$$

for $0 \leqslant a < b \leqslant T$. In Proposition A.3, we show that $Q(\phi, m)$ is a positive definite matrix. This is useful in deriving the existence of the unrestricted maximum likelihood estimator (UMLE) and the restricted maximum likelihood estimator (RMLE), in case $\boldsymbol{\phi}$ is known.

3. Estimation in case of known change points

In this section, we assume that the change point $\tau_j = \phi_j T$ is known, $j = 1, ..., m$ with m known. By using Proposition A.3, we derive the following lemma which gives the UMLE and the RMLE. We also derive in this section, the joint asymptotic normality of the UMLE and the RMLE. Let $\hat{\theta}(\phi, m)$ and $\tilde{\theta}(\phi, m)$ be the UMLE and the RMLE respectively, let $G = Q^{-1}(\phi, m)B'(BQ^{-1}(\phi, m)B')^{-1}$.

Lemma 3.1. *If Assumptions 1-2 hold, then the log likelihood function is* $\log L(\theta, X_t) = \frac{1}{\sigma^2}\theta'\tilde{R}(\phi, m) - \frac{1}{2\sigma^2}\theta' Q(\phi, m)\theta$. *Further, the UMLE and the RMLE are given by* $\hat{\theta}(\phi, m) = Q^{-1}(\phi, m)\tilde{R}(\phi, m)$ *and* $\tilde{\theta}(\phi, m) = \hat{\theta}(\phi, m) - G(B\hat{\theta}(\phi, m) - r)$, *respectively.*

Proof. The derivation of $\log L(\theta, X_t)$ follows from Theorem 1.12 of Kutoyants (2004) [10] or Theorem 7.6 of Lipster and Shiryaev (2001) [12] along with some algebraic computations. Further, by optimizing the function $\log L(\theta, X_t)$, without and with the constraint in (2.3), we get the UMLE and the RMLE. □

In order to derive the asymptotic normality of the UE, we first establish three preliminary propositions. To simplify some mathematical expressions, let

$$\Sigma_j = \begin{bmatrix} \boldsymbol{I}_p & \Lambda_j \\ \Lambda_j' & \omega_j \end{bmatrix}, \ \Lambda_j = -\int_0^1 \tilde{h}_j(t)\varphi(t)dt, \ \omega_j = \int_0^1 \tilde{h}_j^2(t)dt + \frac{\sigma^2}{2a_j}, \tag{3.1}$$

$j = 1, ..., m+1$, and let

$$\Sigma = \operatorname{diag}\left(\phi_1\Sigma_1, (\phi_2 - \phi_1)\Sigma_2, \ldots, (\phi_m - \phi_{m-1})\Sigma_m, (1-\phi_m)\Sigma_{m+1}\right). \tag{3.2}$$

By Proposition A.4 in the Appendix A, Σ_j, $j = 1, 2, \ldots, m+1$ and Σ are invertible. From (3.2), we have

$$\Sigma^{-1} = \operatorname{diag}\Big(\phi_1^{-1}\Sigma_1^{-1}, (\phi_2 - \phi_1)^{-1}\Sigma_2^{-1}, \ldots, (\phi_m - \phi_{m-1})^{-1}\Sigma_m^{-1}, (1-\phi_m)^{-1}\Sigma_{m+1}^{-1}\Big). \tag{3.3}$$

Below, we present a theorem and three propositions which play a crucial role in deriving the joint asymptotic normality of the UMLE and the RMLE.

Theorem 3.1. *If Assumptions 1-2 hold, for* $0 \leqslant \phi_{j-1} < \phi_j \leqslant 1$, $j = 1, ..., m+1$, *then*

(i) $\lim_{T\to\infty} \frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \varphi(t)\varphi'(t)dt = (\phi_j - \phi_{j-1})\boldsymbol{I}_p$;

(ii) $\frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t \varphi(t)dt \xrightarrow[T\to\infty]{a.s.\ and\ L^1} -(\phi_j - \phi_{j-1})\Lambda_j$;

(iii) $\frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{a.s.\ and\ L^1} (\phi_j - \phi_{j-1})\omega_j$.

The proof is given in the Appendix A. By combining Theorems 2.1 and 3.1, we derive the following proposition.

Proposition 3.1. *If Assumptions 1-2 hold, for* $0 \leqslant \phi_{j-1} < \phi_j \leqslant 1$, $j = 1, ..., m+1$, *then*

(i) $\frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} X_t \varphi(t)dt \xrightarrow[T\to\infty]{a.s.\ and\ L^1} -(\phi_j - \phi_{j-1})\Lambda_j$;

(ii) $\frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} X_t^2 dt \xrightarrow[T\to\infty]{a.s.\ and\ L^1} (\phi_j - \phi_{j-1})\omega_j$.

The proof follows directly from Theorems 2.1 and 3.1. By using Propositions A.4 and 3.1, we establish the following proposition.

Proposition 3.2. *If Assumptions 1-2 hold, then, for* $0 \leqslant \phi_{j-1} < \phi_j \leqslant 1$, $j = 1, ..., m+1$,

$T^{-1}Q_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})\Sigma_j$, $\quad TQ^{-1}_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{a.s.} \frac{1}{\phi_j - \phi_{j-1}}\Sigma_j^{-1}$,

$T^{-1}Q(\phi, m) \xrightarrow[T\to\infty]{a.s.} \Sigma$, *and* $TQ^{-1}(\phi, m) \xrightarrow[T\to\infty]{a.s.} \Sigma^{-1}$.

The proof is given in the Appendix A. By using Proposition 3.2, we establish below the asymptotic normality of the $\hat{\theta}(\phi, m)$. Let $\rho_T(\phi, m) = \sqrt{T}(\hat{\theta}(\phi, m) - \theta)$.

Proposition 3.3. *Suppose that Assumptions 1-2 hold. Then,*

(i) $\hat{\theta}(\phi, m) = \theta + \sigma Q^{-1}(\phi, m)\boldsymbol{M}(\phi, m)$;

(ii) $\frac{1}{\sqrt{T}}\boldsymbol{M}(\phi, m) \xrightarrow[T\to\infty]{d} \boldsymbol{M}_0 \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma)$;

(iii) $\rho_T(\phi, m) \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2\Sigma^{-1})$.

The proof is given in the Appendix A. By using this proposition, we derive the joint asymptotic normality of $\hat{\theta}(\phi, m)$ and $\tilde{\theta}(\phi, m)$. Let $\zeta_T(\phi, m) = \sqrt{T}(\tilde{\theta}(\phi, m) - \theta)$ and let $\xi_T(\phi, m) = \sqrt{T}(\hat{\theta}(\phi, m) - \tilde{\theta}(\phi, m))$. We consider the following set of local alternative restrictions,

$$H_{a,T} : B\theta - r = \frac{r_0}{\sqrt{T}},\ T > 0, \tag{3.4}$$

where r_0 is a fixed q-column vector. First, note that

$$\sqrt{T}(\tilde{\theta}(\phi, m) - \theta) = (I_{(m+1)(p+1)} - GB)\sqrt{T}(\hat{\theta}(\phi, m) - \theta) - \sqrt{T}G(B\theta - r). \tag{3.5}$$

From Proposition 3.2, $G \xrightarrow[T\to\infty]{a.s.} G^* = \Sigma^{-1}B'(B\Sigma^{-1}B')^{-1}$. Then,

$$I_{(m+1)(p+1)} - GB \xrightarrow[T\to\infty]{a.s.} I_{(m+1)(p+1)} - G^*B; \; \sqrt{T}G(B\theta - r) \xrightarrow[T\to\infty]{a.s.} G^*r_0. \quad (3.6)$$

The following proposition presents the asymptotic distribution of $(\rho_T'(\phi,m), \zeta_T'(\phi,m), \xi_T'(\phi,m))'$. Let $\Lambda_{22} = \Sigma^{-1} - G^*B\Sigma^{-1}$.

Proposition 3.4. *If Assumptions 1-2 hold along with the set of local alternatives in (3.4), then, if $r_0 \neq 0$, $(\rho_T'(\phi,m), \zeta_T'(\phi,m), \xi_T'(\phi,m))' \xrightarrow[T\to\infty]{d} (\rho', \zeta', \xi')'$ where*

$$\begin{pmatrix}\rho\\ \zeta\\ \xi\end{pmatrix} \sim \mathcal{N}_{3(m+1)(p+1)}\left(\begin{pmatrix}0\\ -G^*r_0\\ G^*r_0\end{pmatrix}, \sigma^2\begin{pmatrix}\Sigma^{-1} & \Lambda_{22} & \Sigma^{-1}-\Lambda_{22}\\ \Lambda_{22} & \Lambda_{22} & 0\\ \Sigma^{-1}-\Lambda_{22} & 0 & \Sigma^{-1}-\Lambda_{22}\end{pmatrix}\right).$$

Further, if $r_0 = 0$, $(\rho_T'(\phi,m), \zeta_T'(\phi,m), \xi_T'(\phi,m))' \xrightarrow[T\to\infty]{d} (\rho_0', \zeta_0', \xi_0')'$ where

$$\begin{pmatrix}\rho_0\\ \zeta_0\\ \xi_0\end{pmatrix} \sim \mathcal{N}_{3(m+1)(p+1)}\left(0, \sigma^2\begin{pmatrix}\Sigma^{-1} & \Lambda_{22} & \Sigma^{-1}-\Lambda_{22}\\ \Lambda_{22} & \Lambda_{22} & 0\\ \Sigma^{-1}-\Lambda_{22} & 0 & \Sigma^{-1}-\Lambda_{22}\end{pmatrix}\right).$$

The proof is given in Appendix A. Proposition 3.4 generalizes the result given in relation (3.19) of Chen *et al.* (2017) [4].

4. Estimation in case of unknown change points

4.1. The unrestricted and restricted estimators

In the previous section, the locations of change-points and the number of change points are assumed to be known. Nevertheless, in practice, the change points is also unknown. Thus, the change points have to be estimated from the data. In this section, we assume that the number of the change points, m, is known but the positions of change points are unknown. We show that Proposition 3.4 holds when one replaces ϕ by one of its 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 $\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_m)'$.

First, for estimating the positions of change points, one can use the least sum of squared errors (LSSE) method, which is similar to that in Chen *et al.* (2017) [4]. We divide the time period $[0, T]$ into n partitions, i.e., $0 = t_0 < \cdots < t_n = T$. For the sake of simplicity, we consider that the time increments are equal, i.e. $\tilde{\Delta} = t_{i+1} - t_i$, $i = 0, ..., n-1$. Moreover, we define $Y_i = X_{t_i+1} - X_{t_i}$ and $z_i = (\varphi_1(t_i), ..., \varphi_p(t_i), -X_{t_i})\tilde{\Delta}$, and let $\tau = (\tau_1, ..., \tau_m)$. From the Euler-Maruyama discretisation method, we have

$$Y_i = z_i\theta_i + \epsilon_i, \quad i = 1, ..., n \quad (4.1)$$

where ϵ_i is the error term $\sigma\sqrt{\tilde{\Delta}}\omega_i$, 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

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

where

$$\mathrm{SSE}([0,T], \tau, \hat{\theta}(\tau)) = \sum_{t_i \in [0,T]} (Y_i - z_i\hat{\theta}_i)'(Y_i - z_i\hat{\theta}_i) \tag{4.3}$$

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.*

For more details about this assumption, we refer to Chen *et al.* (2017) [4] and references therein.

Remark 4.1. The estimator of ϕ is computed as $\hat{\phi} = \dfrac{\hat{\tau}}{T}$. The consistency of $\hat{\phi}$ can be established by using the similar techniques as in Chen et al. (2017) [4]. The estimator $\hat{\phi}$ is $\mathfrak{F}_T$-measurable $[0,1]$-valued, and there exists $\delta_0 > 0$ such that $\hat{\phi} - \phi = O_P(T^{-\delta_0})$.

For convenience, let $\hat{\phi}_0 = 0$ and $\hat{\phi}_{m+1} = 1$. Let $\boldsymbol{M}(\hat{\phi}, m)$ be as $\boldsymbol{M}(\phi, m)$ in (2.10) by replacing ϕ by $\hat{\phi}$. Let $\tilde{R}(\hat{\phi}, m)$ and $Q(\hat{\phi}, m)$ be as $\tilde{R}(\phi, m)$ and $Q(\phi, m)$ by replacing ϕ by $\hat{\phi}$. The UE is given by $\hat{\theta}(\hat{\phi}, m)$ and the RE is given by

$$\tilde{\theta}(\hat{\phi}, m) = \hat{\theta}(\hat{\phi}, m) - J(B\hat{\theta}(\hat{\phi}, m) - r) \tag{4.4}$$

where $J = Q^{-1}(\hat{\phi}, m)B'(BQ^{-1}(\hat{\phi}, m)B')^{-1}$.

As in Chen *et al.* (2017) [4], another estimation method of the locations of the change points is the one which is based on the Maximum log-likelihood estimation. In particular, by Theorem 7.6 of Lipster and Shiryaev (2001) [12], the log-likelihood function is given by

$$\log L(\tau, \theta) = \frac{1}{\sigma^2}\int_0^T S(\theta, t, X_t)dX_t - \frac{1}{2\sigma^2}\int_0^T S^2(\theta, t, X_t)dt, \quad \tau = (\tau_1, \dots, \tau_m)'.$$

Note that

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

and, one can verify that

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

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

$$\log L(\tau, \theta) = \frac{1}{\sigma^2} \sum_{j=1}^{m+1} \int_{\tau_{j-1}}^{\tau_j} S_j(\theta_j, t, X_t) dX_t - \frac{1}{2\sigma^2} \sum_{j=1}^{m+1} \int_{\tau_j - 1}^{\tau_j} S_j^2(\theta_j, t, X_t) dt \tag{4.5}$$

where $S_j(\theta_j, t, X_t) = \sum_{k=1}^{p} \mu_{k,j} \varphi_k(t) - a_j X_t$. From (4.5), when the number of change points, m, is known, the estimator of $\tau = (\tau_1, \ldots, \tau_m)$, is given by

$$\hat{\tau} = \arg\max_{\tau} \log L(\tau, \hat{\theta}(\tau)) \tag{4.6}$$

where $\hat{\theta}(\tau)$ is the MLE of θ for a given change points τ. Auger and Lawrence (1989) [2] introduced a numerical method to approximate the integrals inside the log-likelihood function. We use this method to calculate $\log L(\tau, \hat{\theta}(\tau))$ in (4.6). Divide the interval $[0, T]$ into n partitions, 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 (4.6) is approximated by

$$\begin{aligned} \log L^*([0, T], \tau, \hat{\theta}(\tau)) = &\frac{1}{\sigma^2} \sum_{j=1}^{m+1} \sum_{t_i^* \in (\tau_{j-1}, \tau_j]} \hat{\theta}^{(j)\top} V(t) (X_{t_{i+1}^*} - X_{t_i^*}) \\ &- \frac{1}{2\sigma^2} \sum_{j=1}^{m+1} \sum_{t_i^* \in (\tau_{j-1}, \tau_j]} \left(\hat{\theta}^{(j)\top} V(t)\right)^2 \Delta_t^* \end{aligned} \tag{4.7}$$

where $V(t) = (\varphi_1(t), \ldots, \varphi_p(t), -X_t)'$. Then, one can estimate τ by

$$\hat{\tau} = \arg\max_{\tau} \log L^*([0, T], \tau, \hat{\theta}(\tau)). \tag{4.8}$$

In the sequel, let $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ be the estimators of ϕ_j and ϕ_{j-1} respectively, $j = 1, ..., m+1$, obtained from (4.2) or (4.8). Below, we present three lemma and two propositions which are useful in deriving the joint asymptotic normality of $\hat{\boldsymbol{\theta}}(\hat{\boldsymbol{\phi}}, m)$ and $\tilde{\boldsymbol{\theta}}(\hat{\boldsymbol{\phi}}, m)$.

Lemma 4.1. *Let $\hat{a}$ and $\hat{b}$ be $\mathfrak{F}_T$-measurable and consistent estimators for a and b respectively, with $0 \leqslant a < b \leqslant 1$, and $0 \leqslant \hat{a} < \hat{b} \leqslant 1$ a.s. Let $\{Y_t, t \geqslant 0\}$ be a stochastic process $\{\mathfrak{F}_t, t \geqslant 0\}$-adapted and L^2-bounded. Then,*

(i) $\displaystyle \frac{1}{T} \int_{\hat{a}T}^{\hat{b}T} Y_t dt - \frac{1}{T} \int_{aT}^{bT} Y_t dt \xrightarrow[T \to \infty]{L^1} 0.$

Further, if $\max(|\hat{a} - a|, |\hat{b} - b|) = O_P(T^{-\delta_0})$ *with* $1/2 < \delta_0 \leqslant 1$, *then*

(ii) $\displaystyle \frac{1}{T} \int_{\hat{a}T}^{\hat{b}T} Y_t dt - \frac{1}{T} \int_{aT}^{bT} Y_t dt \xrightarrow[T \to \infty]{L^{2\delta_0}} 0;$

(iii) $\displaystyle \frac{1}{\sqrt{T}} \int_{\hat{a}T}^{\hat{b}T} Y_t dt - \frac{1}{\sqrt{T}} \int_{aT}^{bT} Y_t dt \xrightarrow[T \to \infty]{\mathrm{P}} 0.$

The proof is outlined in Appendix A.

Lemma 4.2. *Let $\hat{a}$ and $\hat{b}$ be $\mathfrak{F}_T$-measurable and consistent estimators for a and b respectively, with $0 \leqslant a < b \leqslant 1$, and $0 \leqslant \hat{a} < \hat{b} \leqslant 1$ a.s. Let $\{Y_t, t \geqslant 0\}$ be a R^p-valued deterministic and bounded function. Then,*

$$\frac{1}{\sqrt{T}} \int_{\hat{a}T}^{\hat{b}T} Y_t dW_t - \frac{1}{\sqrt{T}} \int_{aT}^{bT} Y_t dW_t \xrightarrow[T\to\infty]{\mathrm{L}^2} 0.$$

The proof is outlined in Appendix A. For the case where the process is not purely deterministic, we derive the following lemma.

Lemma 4.3. *Let $\hat{a}$ and $\hat{b}$ be $\mathfrak{F}_T$-measurable and consistent estimators for a and b respectively, with $0 \leqslant a < b \leqslant 1$, $0 \leqslant \hat{a} < \hat{b} \leqslant 1$ a.s, and $\max(|\hat{a} - a|, |\hat{b} - b|) = O_P(T^{-\delta_0})$ with $1/2 < \delta_0 \leqslant 1$. Let $\{Y_t, t \geqslant 0\}$ be a solution of SDE*

$$dY_t = \sum_{k=1}^{m+1} f(\mu_k, Y_t) \mathbb{I}_{\{\tau_{j-1} < t \leqslant \tau_j\}} dt + \sigma dW_t, \;\; 0 \leqslant t \leqslant T \tag{4.9}$$

where $f(\theta, x)$ is a real-valued function such that the processes $\{Y_t, t \geqslant 0\}$ and $\{f(\theta, Y_t), t \geqslant 0\}$ are L^2-bounded. Then,

$$\frac{1}{\sqrt{T}} \int_{\hat{a}T}^{\hat{b}T} Y_t dW_t - \frac{1}{\sqrt{T}} \int_{aT}^{bT} Y_t dW_t \xrightarrow[T\to\infty]{\mathrm{P}} 0.$$

The proof is outlined in the Appendix A. By combining Lemma 4.1 and Proposition A.1, we establish the following proposition.

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

(i) $\displaystyle \frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \varphi(t)\varphi'(t) dt - \frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \varphi(t)\varphi'(t) dt \xrightarrow[T\to\infty]{\mathrm{P}} \mathbf{0}$;

(ii) $\displaystyle \frac{1}{T} \int_{\hat{\phi}_{j-1}T}^{\hat{\phi}_j T} \begin{pmatrix} \varphi(t) \\ -X_t \end{pmatrix} X_t dt - \frac{1}{T} \int_{\phi_{j-1}T}^{\phi_j T} \begin{pmatrix} \varphi(t) \\ -X_t \end{pmatrix} X_t dt \xrightarrow[T\to\infty]{\mathrm{P}} 0.$

The proof follows from Lemma 4.1 and Proposition A.1 in the Appendix A. By combining Propositions 3.2, A.4 and 4.1, we derive the following propositions which give the similar results as Proposition 3.2 in case one replaces the change points ϕ_j and ϕ_{j-1} by their consistent estimators $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ respectively. Let $Q_{(\hat{\tau}_{j-1}, \hat{\tau}_j)}$ be as defined in (2.8) by replacing (τ_{j-1}, τ_j) by $(\hat{\tau}_{j-1}, \hat{\tau}_j)$, where $\hat{\tau}_j = \hat{\phi}_j T$, $\hat{\tau}_{j-1} = \hat{\phi}_{j-1} T$ for $j = 1, ..., m+1$.

Proposition 4.2. *If Assumptions 1-3 hold, then, for $0 \leqslant \phi_{j-1} < \phi_j \leqslant 1$, $j = 1, ..., m+1$,*

$$T^{-1} Q_{(\hat{\tau}_{j-1}, \hat{\tau}_j)} \xrightarrow[T\to\infty]{\mathrm{P}} (\phi_j - \phi_{j-1}) \Sigma_j, \qquad T Q^{-1}_{(\hat{\tau}_{j-1}, \hat{\tau}_j)} \xrightarrow[T\to\infty]{\mathrm{P}} \tfrac{1}{\phi_j - \phi_{j-1}} \Sigma_j^{-1},$$

$$T^{-1} Q(\hat{\phi}, m) \xrightarrow[T\to\infty]{\mathrm{P}} \Sigma, \qquad \textit{and} \qquad T Q^{-1}(\hat{\phi}, m) \xrightarrow[T\to\infty]{\mathrm{P}} \Sigma^{-1}.$$

Proof. The proof of the first statement follows from Propositions 4.1 and A.1. The proof of the second statement follows from the first statement along with Proposition A.4. The proof of the third statement follows from the first statement. The fourth statement follows from the third statement along with Proposition A.4. □

By using Propositions 3.4, 4.2 and Lemma 4.3, we derive the joint asymptotic distribution of the UE $\hat{\theta}(\hat{\phi}, m)$ and the RE $\tilde{\theta}(\hat{\phi}, m)$. Let $\rho_T(\hat{\phi}, m) = \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta)$.

Proposition 4.3. *If Assumptions 1-3 hold, then,*
$\frac{1}{\sqrt{T}}\left(\boldsymbol{M}(\hat{\phi}, m) - \boldsymbol{M}(\phi, m)\right) \xrightarrow[T\to\infty]{\mathrm{P}} 0,$
$\frac{1}{\sqrt{T}}\boldsymbol{M}(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} \boldsymbol{M}_0 \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma)$, *and*
$\rho_T(\hat{\phi}, m) \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2\Sigma^{-1})$.

The proof of this proposition is given in the Appendix A. By using Proposition 4.3, we derive the joint asymptotic distribution of UE $\hat{\theta}(\hat{\phi}, m)$ and RE $\tilde{\theta}(\hat{\phi}, m)$. Let
$\zeta_T(\hat{\phi}, m) = \sqrt{T}(\tilde{\theta}(\hat{\phi}, m) - \theta)$, and let $\xi_T(\hat{\phi}, m) = \sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \tilde{\theta}(\hat{\phi}, m))$. We have

$$\zeta_T(\hat{\phi}, m) = (I_{(m+1)(p+1)} - JB)\sqrt{T}(\hat{\theta}(\hat{\phi}, m) - \theta) - \sqrt{T}J(B\theta - r), \tag{4.10}$$

with $J = TQ^{-1}(\hat{\phi}, m)B'(BTQ(\hat{\phi}, m)^{-1}B')^{-1}$. Then, under the set local alternative restrictions in (3.4), from Proposition 4.2, we have

$$J = TQ(\hat{\phi}, m)^{-1}B'(BTQ(\hat{\phi}, m)^{-1}B')^{-1} \xrightarrow[T\to\infty]{P} \Sigma^{-1}B'(B\Sigma^{-1}B')^{-1} = G^*.$$

$$I_{(m+1)(p+1)} - JB \xrightarrow[T\to\infty]{P} I_{(m+1)(p+1)} - G^*B,\ \sqrt{T}J(B\theta - r) \xrightarrow[T\to\infty]{P} G^*r_0. \tag{4.11}$$

Proposition 4.4. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.4). Then, if $r_0 \neq 0$, $(\rho_T(\hat{\phi}, m)', \zeta_T(\hat{\phi}, m)', \xi_T(\hat{\phi}, m)')' \xrightarrow[T\to\infty]{d}$ $(\rho', \zeta', \xi')'$, and if $r_0 = 0$, $(\rho_T(\hat{\phi}, m)', \zeta_T(\hat{\phi}, m)', \xi_T(\hat{\phi}, m)')' \xrightarrow[T\to\infty]{d} (\rho_0', \zeta_0', \xi_0')'$ where ρ, ζ, ξ, ρ_0, ζ_0 and ξ_0 are defined as in Proposition 3.4.*

The proof follows from the similar steps as of the proof of Proposition 3.4. Proposition 4.4 generalizes Corollary 4.2 in Chen *et al.* (2017).

5. Inference in the case of unknown number of change points

In Sections 3 and 4, we suppose that the number of change points, m, is known. However, for a given data set, m is also unknown. Thus, in this section, we solve a more general problem where the nuisance parameters m, τ_1, τ_2, $\dots$, τ_m are unknown.

5.1. Estimation of the number of change points and algorithm

In this subsection, we describe the algorithm which is used to estimate m and τ. The algorithm is given here for the completeness as it can be found in Chen *et al.* (2017) [4]. Thus, we consider estimating m as selecting the best fitting model i. e. among models with different 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

$$\mathrm{IC}(m) = -2\log L(\tau, \hat{\theta}(\phi, m)) + (m+1)h(p)\Upsilon(T) + \lambda'(B\hat{\theta}(\phi, m) - r) \quad (5.1)$$

where $\log L(\tau, \hat{\theta}(\phi, m))$ is defined in (4.5); $\hat{\tau}$ is established by (4.6) 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; $\boldsymbol{B}$ and r are defined in (2.3). By Riemann sum approximation of $\log L(\tau, \hat{\theta})$, the information criterion is given by

$$\mathrm{IC}(m) = -2\log L^*([0,T], \tau, \hat{\theta}(\tau)) + (m+1)h(p)\Upsilon(T) + \lambda'(\boldsymbol{B}\hat{\theta}(\tau) - r) \quad (5.2)$$

where $\log L^*([0,T], \tau, \hat{\theta}(\tau))$ is defined in (4.7); and $\hat{\tau}$ is established by (4.8) corresponding to each m. It should be noticed that, if m is known, the term $(m+1)h(p)\Upsilon(T)$ is fixed. The approach involving (5.2) is the same as the maximum log-likelihood method introduced in Section 4. Note that (5.2) leads to the well known information criterion so-called Akaike information criterion (AIC) Akaike (1973) [1] when $\Upsilon(T) = 2$. However, as mentioned in Chen *et al.* (2017) [4], due to the problem of consistency of AIC, it is convenient to use the function which yields the Schwarz information criterion (SIC) as given in Schwarz (1978) [19]. In SIC, $\Upsilon(T)$ is set as the logarithm of the sample size. Thus, in the sequel, we use the SIC. One can verify that, as T is large, $\mathrm{IC}(m)$ 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 same as finding the $\mathrm{IC}(m)$ in (5.2) which reaches its minimum. Then, its corresponding m is the number of change points we would like to estimate. Below, we present an algorithm which is useful in finding τ and m. Let $\hat{m}$ be a consistent estimator of m and let $\hat{\tau}(\hat{m})$ be a consistent estimator for $\tau(m)$. In passing, note that, in the steps of the algorithm for the estimation of $\tau(m)$, we apply the LSSE method or the Maximum log-likelihood method in Section 4 along with the dynamic programming algorithm from Bai and Perron (1998) [3], Perron and Qu (2006) [16].

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 (4.8) 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 (4.8) 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 (4.8) 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 $b-a \geqslant h$.
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

$$H_1(1,T_1) = \begin{cases} \min_{a\in[h,T_1-h]}[H_2(0,a)+H_2(a,T_1)] & \text{for (4.2)} \\ \max_{a\in[h,T_1-h]}[H_2(0,a)+H_2(a,T_1)] & \text{for (4.8).} \end{cases}$$

Step 3: Sequentially compute and save

$$H_1(r,T_r) = \begin{cases} \min_{a\in[rh,T_r-h]}[H_1(r-1,a)+H_2(a,T_r)] & \text{for (4.2)} \\ \max_{a\in[rh,T_r-h]}[H_1(r-1,a)+H_2(a,T_r)] & \text{for (4.8).} \end{cases}$$

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

$$H_1(m,T) = \begin{cases} \min_{a\in[mh,T-h]}[H_1(m-1,a)+H_2(a,T)] & \text{for (4.2)} \\ \max_{a\in[mh,T-h]}[H_1(m-1,a)+H_2(a,T)] & \text{for (4.8),} \end{cases}$$

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 $a-b \geqslant h$ 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 $\mathrm{IC}(m)$.
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, $0 < m \leqslant m_{\max}$ where $0 \leqslant m_{\max} \leqslant \lceil [T/h] \rceil$. The $m_{\max}$ can be determined by observing and analyzing the given process. By Proposition A.2 in Appendix A, $\hat{m}$ is consistent, provided that the exact value of the number of the change-points $m^0 \in [0, m_{\max}]$.

5.2. Asymptotic properties of the UE and the RE

In this subsection, we derive a lemma and a theorem which allows us to overcome the difficulty due to the randomness of the dimensions of the UE and the RE. The established results are of interest on their own in addition to be useful in deriving the asymptotic power of the proposed test as well as in studying the asymptotic risk dominance of the UE, the RE and the SEs. Let $\hat{m}$ be a consistent estimator for m and let $\hat{\tau}(\hat{m})$ be the estimator of $\tau(m)$. 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 $\tilde{\theta}(\hat{\phi},\hat{m})$.

Note that since the dimensions of $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$ are functions of $\hat{m}$, it is challenging or rather impossible to derive the limiting distribution of $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$. Because of that, neither the relative risk dominance of the UE and the RE nor the construction of shrinkage estimators follow from the results in literature, as for example in Saleh (2006) [18], Chen and Nkurunziza (2015) [5], Nkurunziza and Zhang (2018) [15] among others. Below, we derive some results which are useful to overcome the problem related to the randomness fact of the dimensions of $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$. To this end, let $\rho_T(\hat{\phi}, \hat{m}) = \sqrt{T}(\hat{\theta}(\hat{\phi}, \hat{m}) - \theta)$, $\zeta_T(\hat{\phi}, \hat{m}) = \sqrt{T}(\tilde{\theta}(\hat{\phi}, \hat{m}) - \theta)$ and $\xi_T(\hat{\phi}, \hat{m}) = \sqrt{T}(\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m}))$. Below, we derive the limiting distribution of $g\left(\rho_T(\hat{\phi}, \hat{m}), \zeta_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})\right)$ for a given continuous function $g : \mathbb{R}^{(m+1)(p+1)} \times \mathbb{R}^{(m+1)(p+1)} \times \mathbb{R}^{(m+1)(p+1)} \longrightarrow \mathbb{R}^q$ with q no depending on m. As a preliminary result, we prove the following lemma.

Lemma 5.1. *Let q be a positive integer, let $\hat{m}$ be nonnegative 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-random vectors such that $X_T(m) \xrightarrow[T\to\infty]{d} X(m)$. Then, $X_T(\hat{m}) \xrightarrow[T\to\infty]{d} X(m)$.*

The proof of this proposition is given in the Appendix A. By combining Proposition 4.4 and Lemma 5.1, we establish the following theorem.

Theorem 5.1. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.4) and let $g\left(\rho_T(\hat{\phi}, \hat{m}), \zeta_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})\right)$ for a given continuous function $g : \mathbb{R}^{(m+1)(p+1)} \times \mathbb{R}^{(m+1)(p+1)} \times \mathbb{R}^{(m+1)(p+1)} \longrightarrow \mathbb{R}^q$ with q no depending on m. Then, if $r_0 \neq 0$, $g(\rho_T(\hat{\phi}, \hat{m}), \zeta_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho, \zeta, \xi)$, and*
if $r_0 = 0$, $g(\rho_T(\hat{\phi}, \hat{m}), \zeta_T(\hat{\phi}, \hat{m}), \xi_T(\hat{\phi}, \hat{m})) \xrightarrow[T\to\infty]{d} g(\rho_0, \zeta_0, \xi_0)$ where ρ, ζ, ξ, ρ_0, ζ_0 and ξ_0 are defined as in Proposition 3.4.

The proof follows from Proposition 4.4 and Lemma 5.1. Theorem 5.1 plays a crucial role in deriving a test for the testing problem in (2.3) as well as in constructing a class of shrinkage estimators.

6. Testing and shrinkage estimators

In this section, 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).

6.1. Testing the restriction

In this section, we develop a test for the hypothesis testing problem in (2.3). First, note that, in the continuous time observation, 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 $\chi_q^2(\lambda)$ be a chi-square random variable (r.v.) with q-degrees of freedom (df), and non-centrality parameter λ; let χ_q^2 be a chi-square r.v. with q degrees of freedom. Let $\chi_{\alpha;q}^2$ be the αth-quantile of a χ_q^2 where $0 < \alpha \leqslant 1$. Also, define $\Delta = \frac{1}{\sigma^2} r_0'(B\Sigma^{-1}B')^{-1}r_0$ where r_0 is given as in (3.4), $\hat{\Gamma} = \frac{1}{\hat{\sigma}^2} B'(BTQ^{-1}(\hat{\phi}, \hat{m})B')^{-1}B$, $\Gamma = \frac{1}{\sigma^2} B'(B\Sigma^{-1}B')^{-1}B$. From Proposition 5.1, we derive the following corollary which is the foundation for testing $H_0 : B\theta = r$ versus $H_a : B\theta \neq r$.

Let $\psi_T(\hat{m}) = \xi_T(\hat{\phi}, \hat{m})'\hat{\Gamma}\xi_T(\hat{\phi}, \hat{m})$, let $\psi(m) = \xi'\Gamma\xi$, and let $\psi_0(m) = \xi_0'\Gamma\xi_0$.

Corollary 6.1. *Suppose that the conditions of Theorem 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. By combining Theorem 5.1 and Proposition 4.2, we have $\psi_T(\hat{m}) = \xi_T(\hat{\phi}, \hat{m})'\hat{\Gamma}\xi_T(\hat{\phi}, \hat{m}) \xrightarrow[T\to\infty]{d} \psi(m) = \xi'\Gamma\xi$. Hence, by Theorem 5.1.3 in Mathai and Provost (1992), we get $\xi'\Gamma\xi \sim \chi_q^2(\Delta)$, this completes the proof. □

From Corollary 6.1, we propose a test for the testing problem in (2.3). We suggest

$$\kappa(\hat{\phi}, T) = \mathbb{I}_{\{\psi_T(\hat{m}) > \chi_{\alpha;q}^2\}}. \tag{6.1}$$

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 (6.1) is given by* $\Pi(\Delta) = \mathrm{P}(\chi_q^2(\Delta) \geqslant \chi_{\alpha;q}^2)$.

The proof follows directly from Corollary 6.1. It is obvious that, under the null hypothesis in (2.3), $r_0 = 0$, then $\Delta = 0$, and by Corollary 6.2, $\Pi(0) = \alpha$. Moreover, the asymptotic power tends to 1 as Δ tends to infinity. The novelty of the proposed test and its asymptotic power consists in the fact that, their derivation is not based on the joint asymptotic normality between $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$, as this is the case in Saleh (2006) [18], Nkurunziza (2012) [14], Nkurunziza and Zhang (2018) [15] and references therein.

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 RE. As an alternative method, we construct shrinkage estimators (SEs) by combining the RE and the UE in the optimal way. Although the construction of SEs has been proposed by several authors, here, 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 a random variable. Because of that, the construction of SEs cannot follow from the existing techniques in statistical literature, as for example in Saleh (2006) [18], Nkurunziza (2012) [14], Nkurunziza and Zhang (2018) [15] among others. We consider the following class of shrinkage type estimators

$$\hat{\theta}^s(\beta) = \tilde{\theta}(\hat{\phi}, \hat{m}) + \beta(\|\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})\|_{\hat{\Gamma}})(\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})), \tag{6.2}$$

where $\|x\|_A = x'Ax$, $\beta(.)$ is continuous real-valued function on $(0, +\infty)$. In particular, if
$\beta(x) = (1 - (q-2)/x),\ x > 0$, we get the shrinkage estimator (SE) given by

$$\hat{\theta}^s = \tilde{\theta}(\hat{\phi}, \hat{m}) + [1 - (q-2)\psi_T^{-1}(\hat{m})](\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})), \tag{6.3}$$

where $2 < q = \text{rank}(B) < (m+1)(p+1)$, and $\psi_T(\hat{m})$ is given as in Corollary 6.1. Let
$a^+ = \max\{0, a\}$. If $\beta(x) = [1 - (q-2)/x]^+,\ x > 0$, we get the positive-part shrinkage estimator (PSE) given by

$$\hat{\theta}^{s+} = \tilde{\theta}(\hat{\phi}, \hat{m}) + [1 - (q-2)\psi_T^{-1}(\hat{m})]^+(\hat{\theta}(\hat{\phi}, \hat{m}) - \tilde{\theta}(\hat{\phi}, \hat{m})). \tag{6.4}$$

Note that the proposed class of estimators includes also the UE and the RE by taking $\beta \equiv 1$ and $\beta \equiv 0$, respectively.

At first glance, the SEs in (6.3) and (6.4) are similar to that in Sen and Saleh (1987) [20], Saleh (2006) [18] among others. However, due to the randomness of the dimensions of the estimators $\hat{\theta}^s$, $\hat{\theta}^{s+}$, $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$, the asymptotic distributional risk analysis of these estimators do not follow from the results in statistical literature (e.g. in Sen and Saleh (1987) [20], Saleh (2006) [18], Nkurunziza and Zhang (2018) [15] and references therein).

7. Comparison between estimators

In this section, we derive 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. The novelty of the proposed methods consists in the fact fact the derivation of the ADR is not based on the joint asymptotic normality between $\hat{\theta}(\hat{\phi}, \hat{m})$ and $\tilde{\theta}(\hat{\phi}, \hat{m})$, as this is the case in Sen and Saleh (1987) [20], Saleh (2006) [18], Nkurunziza (2012) [14], Nkurunziza and Zhang (2018) [15] among others.

7.1. Asymptotic distributional risk (ADR)

Let Ω be $(m+1)(p+1) \times (m+1)(p+1)$ positive symmetric semi-definite weighting matrix. The ADR of an estimator $\hat{\theta}_0$ is defined as

$$\text{ADR}\left(\hat{\theta}_0, \theta, \Omega\right) = \text{E}[\varepsilon'\Omega\varepsilon], \tag{7.1}$$

where ε is the random vector such that $T(\hat{\theta}_0 - \theta)'\Omega(\hat{\theta}_0 - \theta) \xrightarrow[T\to\infty]{d} \varepsilon'\Omega\varepsilon$. Note a slight difference with the ADR in Nkurunziza and Zhang (2018) [15] which is defined as in (7.1) with ε a random vector such that $\sqrt{T}(\hat{\theta}_0 - \theta) \xrightarrow[T\to\infty]{d} \varepsilon$. This last concept implies the one we use in this paper. Let $\Delta = \frac{1}{\sigma^2} r_0'(B\Sigma^{-1}B')^{-1}r_0$.

Theorem 7.1. *If Assumptions 1-3 hold along with the set of local alternatives in (3.4), then,*

$$\begin{aligned}
&\mathrm{ADR}(\hat{\theta}^s(\beta), \theta, \Omega) \\
&\quad = \sigma^2 \mathrm{trace}(\Omega\Lambda_{22}) + r_0' G^{*\top}\Omega G^* r_0 - 2\mathrm{E}[\beta(\chi^2_{q+2}(\Delta))] r_0' G^{*\top}\Omega G^* r_0 \\
&\qquad + \sigma^2 \mathrm{E}[\beta^2(\chi^2_{q+2}(\Delta))]\mathrm{trace}(\Omega(\Sigma^{-1} - \Lambda_{22})) + \mathrm{E}[\beta^2(\chi^2_{q+2}(\Delta))] r_0' G^{*\top}\Omega G^* r_0.
\end{aligned}$$

The proof follows from Theorem 5.1 along with Theorems 2.1-2.3 of Nkurunziza (2012) [14] by taking $L_1 \equiv B$, $L_2 \equiv 1$, $\Xi_1 \equiv \frac{1}{\sigma^2} B'(B\Sigma^{-1}B')^{-1}B$, $\delta \equiv G^* r_0$, $\Sigma^* \equiv \Sigma^{-1} - \Lambda_{22}$, $p \equiv 1$.

Corollary 7.1. *If Assumptions 1-3 and the set of local alternatives in (3.4) hold, then,*

(i) $\mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = \sigma^2 \mathrm{trace}(\Omega\Sigma^{-1})$;

(ii) $\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right)$

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

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

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

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

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

The proof follows directly from Theorem 7.1 by taking $\beta \equiv 1$, $\beta \equiv 0$, $\beta(x) = 1 - (q-2)/x$ and $\beta(x) = [1 - (q-2)/x]^+$, respectively, along with the identity
$\mathrm{E}[\chi^2_{q+2}(\Delta)] - \mathrm{E}[\chi^{-2}_{q+4}(\Delta)] = 2\mathrm{E}[\chi^{-4}_{q+4}(\Delta)]$.

7.2. Risk analysis

In this section, we compare the performance of these estimators. Let λ_1 and λ_n denote, respectively, the smallest and the largest eigenvalues of the matrix $\left[(G^{*\top}\Gamma G^*)^{-1} G^{*\top}\Omega G^*\right]$. The following result compares the UE and the RE.

Proposition 7.1. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.4). If* $\Delta \leqslant \left(\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1})\right)/\lambda_n$*, then the RE dominates the UE, and if*
$\Delta \geqslant \left(\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1})\right)/\lambda_1$*, then the UE dominates the RE.*

The proof of this proposition is outlined in the Appendix. Further, we establish the risk dominance of SEs over the UE.

Proposition 7.2. *Suppose that Assumptions 1-3 hold along with the set of local alternatives in (3.4). If* $\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1})/\lambda_n \geqslant (q+2)/2$ *with* $q \in (2, (m+1)(p+1))$*, then* $\mathrm{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) \leqslant \mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right) \leqslant \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right)$*, for all* $\Delta > 0$.

The proof is given in Appendix A.

8. Simulation study

In this section, 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 two cases, we generate the generalized O-U process with a periodic two-dimensional 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

$$dX_t = \sum_{j=1}^{m}\left(\mu_1^{(j)} + \mu_2^{(j)}\sqrt{2}\cos\left(\frac{2\pi t}{\Delta_t}\right) - \alpha_j X_t\right)\mathbb{I}_{(\tau_{j-1},\tau_j)}(t)dt + \sigma dW_t,\ t \geqslant 0,$$

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

8.1. Performance comparison

We consider first the case of two change points so that we let $m = 2$, with $\phi_1 = 0.35$ and $\phi_2 = 0.7$. Second, we consider the case of three change points with $\phi_1 = 0.25$, $\phi_2 = 0.5$ and $\phi_3 = 0.75$. In order to evaluate the effect of the time period T, we generate the O-U process with $T = 20$ and $T = 50$, with the time increment of $\Delta = 0.001$. Table 1 presents the components of θ which are used to generate the O-U for the case of $m = 2$ and $m = 3$. For the restriction, we take $r = 0$ and the matrix $\boldsymbol{B}$ given by $\boldsymbol{B} = [(\boldsymbol{I}_3, 0)', (-\boldsymbol{I}_3, \boldsymbol{I}_3)', (0, -\boldsymbol{I}_3)']$, for $m = 2$. For, $m = 3$, we choose $\boldsymbol{B} = [(\boldsymbol{I}_3, 0, 0)', (-\boldsymbol{I}_3, \boldsymbol{I}_3, 0)', (0, -\boldsymbol{I}_3, \boldsymbol{I}_3)', (0, 0, -\boldsymbol{I}_3)']$.

From 500 iterations, we estimate the locations of change points based on LSSE method in (4.2). For the case of two change-points, the mean of the estimates of

TABLE 1
Coefficients

Parameter	$m=2$			$m=3$			
	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=4$
$\mu_{1,j}$	10	5	15	10	5	15	20
$\mu_{2,j}$	5	2	8	5	2	7	10
α_j	3	1	4	3	1	3	5

TABLE 2
The mean of estimates of ϕ_1, ϕ_2

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

TABLE 3
The mean of estimates of ϕ_1, ϕ_2, ϕ_3

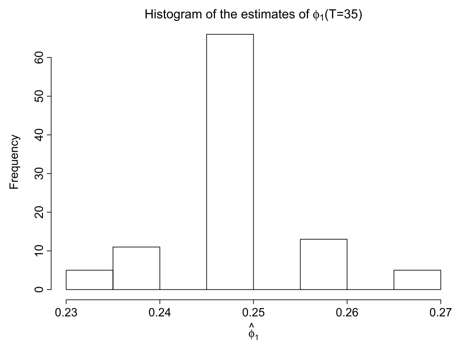
	$T=20$	$T=50$
$\hat{\phi}_1$	0.2519	0.2501
$\hat{\phi}_2$	0.4995	0.5002
$\hat{\phi}_3$	0.7497	0.7502

ϕ_1 and ϕ_2 are recorded in Table 2. Further, for the case of three change-points, the mean of the estimates of ϕ_1, ϕ_2 and ϕ_3 are reported in Table 3. From Tables 2 and 3, it is obvious that, as T becomes large, the estimates of the rate of the change points are closer to the pre-assigned values. In other words, the method is more accurate as T increases.

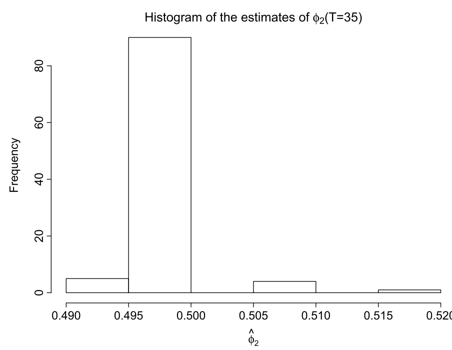
Further, Figure 1, Figure 2 and Figure 3 show that all the histograms are quite symmetric and unimodal with the mode which corresponds to the exact value. As T increases, the estimates become closer to the pre-assigned values.

We also estimate the number of change points based on the algorithm in Section 5. To estimate the number of change points, we take $m_{max} = 6$. From 500 iteration, the cumulative frequency (CF) and the relative frequency (RF) are shown in Table 4. The CF and RF are defined as $\text{CF} = \sum_{i=1}^{500} \mathbb{I}_{(\hat{m}_i = m)}$ and $\text{RF} = \frac{1}{500}\sum_{i=1}^{500} \mathbb{I}_{(\hat{m}_i = m)} \times 100\%$. From Table 4, the cumulative frequency and relative frequency become larger when we change T from 20 to 50. Thus, it seems that the proposed method performs very well when T is relatively large.

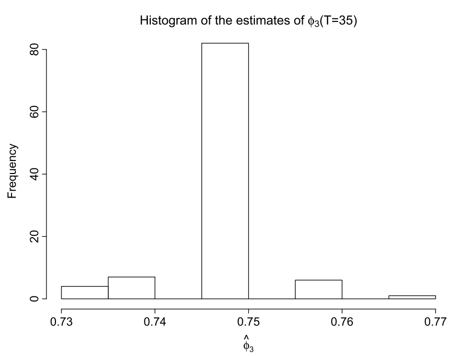
To evaluate the performance of the proposed test in short and medium time period of observation, we present the variation of the empirical power versus the noncentrality parameter $\Delta = \frac{1}{\sigma^2} r_0'(B\Sigma^{-1}B')^{-1} r_0$ with $r_0 = 0.5nr$, $n = 1, 2, 3, 4, 5, 6$. To save the space, we only report here empirical power function for the case of three change-points. Figure 4 corroborates the established theoretical results for which the proposed test is consistent. As in Nkurunziza and



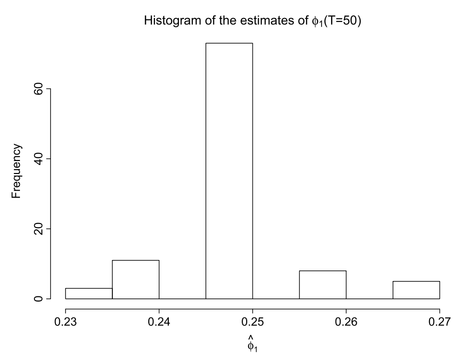
(a) Histogram of $\hat{\phi}_1$ ($T = 35$)



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



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



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

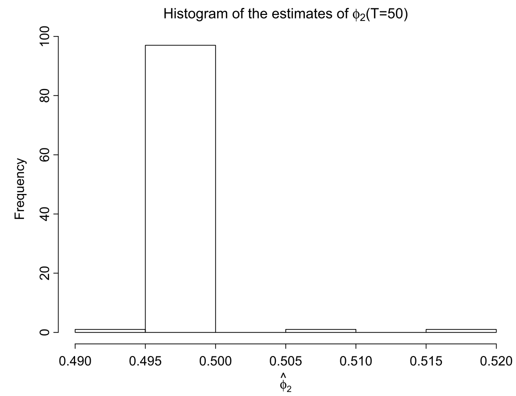
FIG 1. *Histogram of estimates of* $\hat{\phi}$, $m = 3$, $T = (35, 50)$, $\phi = (0.25, 0.5, 0.75)$

Zhang (2018), the relative mean squared efficiency (RMSE) is given as

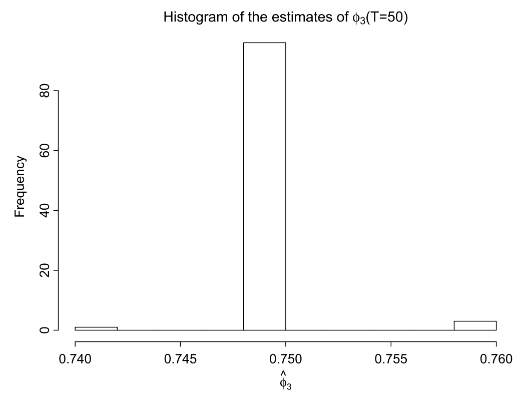
$$\mathrm{RMSE}(\hat{\theta}_0) = \mathrm{ADR}(\hat{\theta}(\hat{\phi}, \hat{m}), \theta; \Omega)/\mathrm{ADR}(\hat{\theta}_0, \theta; \Omega) \tag{8.1}$$

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

From Figures 5, 6, 7 and 8, near $\Delta = 0$, RMSE of RE is the highest. This



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



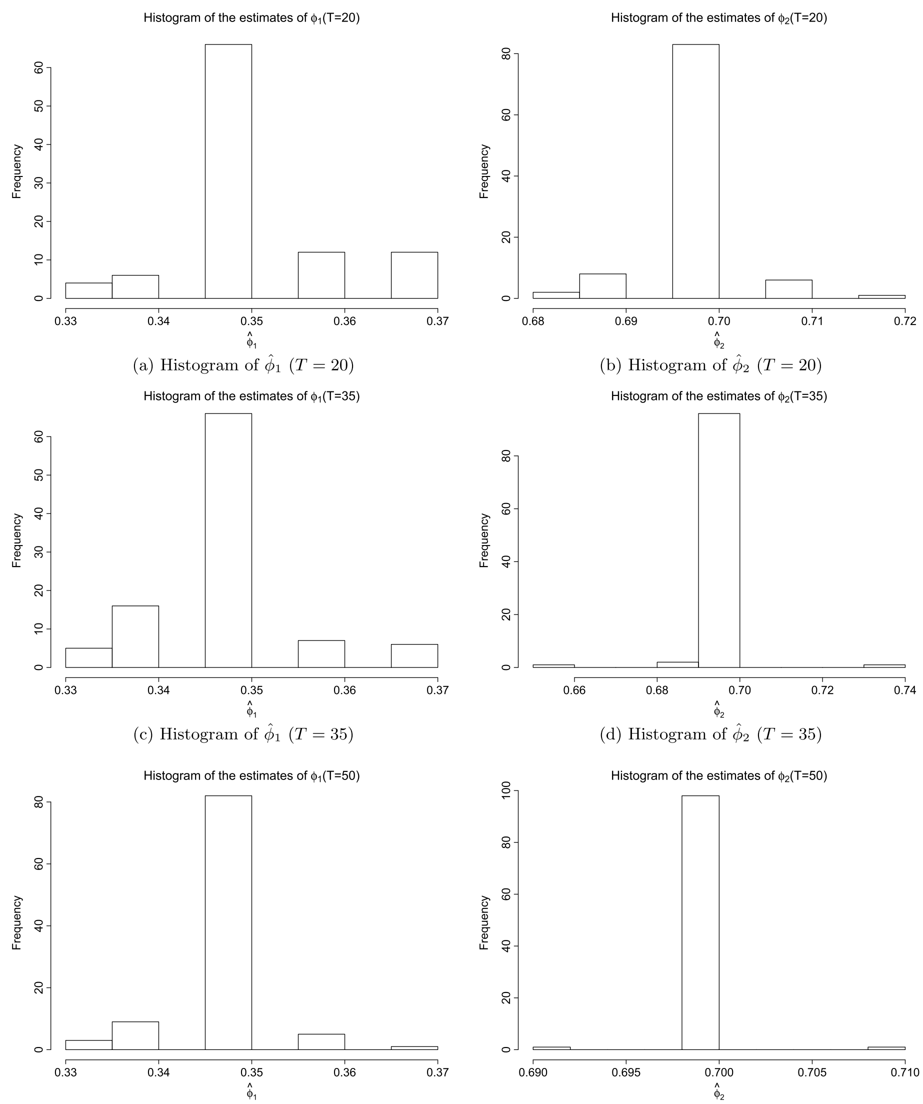
(b) Histogram of $\hat{\phi}_3$ $(T = 50)$

FIG 2. *Histogram of estimates of* $\hat{\phi}$, $m = 3$, $T = (35, 50)$, $\phi = (0.25, 0.5, 0.75)$

shows that, near the restriction, RE is more efficient than the other three estimators. These figures also show that the efficiency of the RE decreases as one moves far away from the null hypothesis. Further, PSE and SE outperform over UE. In conclusion, the numerical results are in agreement with the theoretical results established in Section 7.

9. Conclusion

In this paper, we proposed improved estimation and testing methods in generalized O-U processes with multiple unknown change-points when the drift parameter satisfies uncertain constraint. A Salient feature of this paper consists in the fact that the number of change-points and the locations of the change-points are unknown. The novelty of the established results consists in the fact that the dimensions of the proposed estimators are random. Because of that, the



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

FIG 3. *Histogram of estimates of* $\hat{\phi}$, $m = 2$, $T = (20, 35, 50)$, $\phi = (0.35, 0.7)$

asymptotic power of the proposed test and the asymptotic risk analysis do not follow from the results in statistical literature. In comparison with the results in recent statistical literature, we generalized the findings in Chen *et al.* (2017) [4] as well as that in Nkurunziza and Zhang (2018) [15]. Specifically, we generalized the methods in Chen *et al.* (2017) [4] in five ways. First, we considered the model which incorporates the uncertain prior knowledge. Second, we derived the unrestricted estimator (UE) and the restricted estimator (RE). Further, in

Table 4
Cumulative frequency and relative frequency of 500 iterations (SNS method)

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

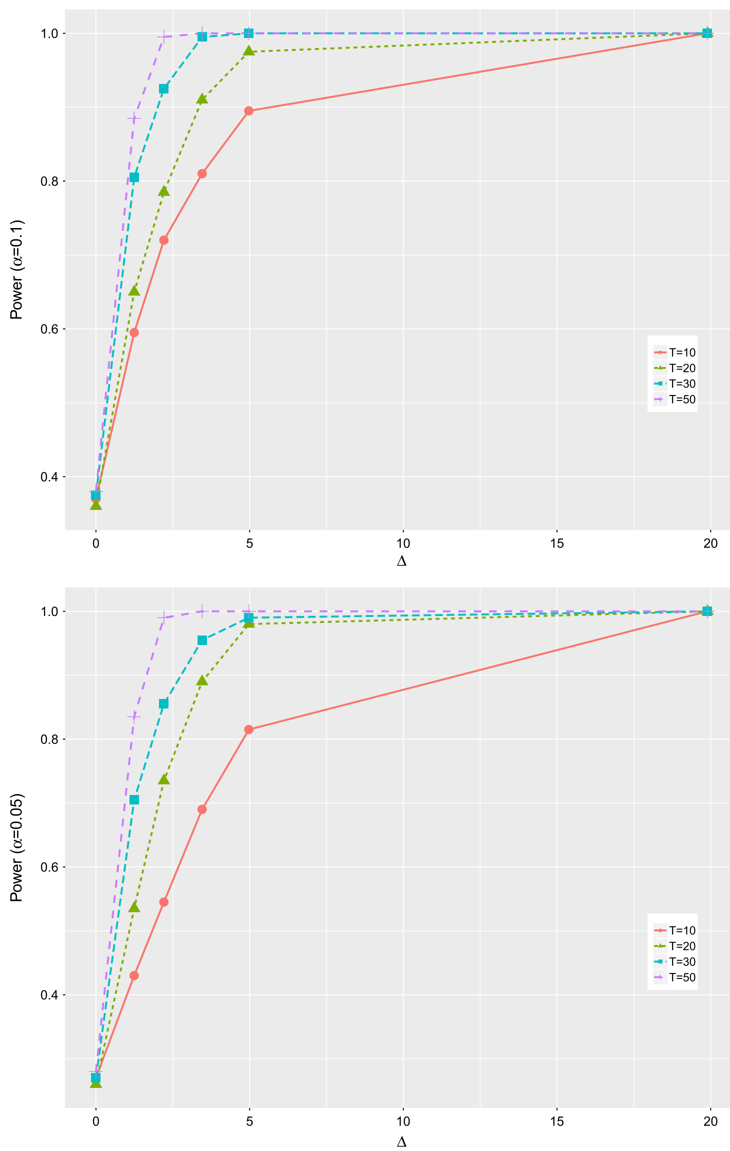


Fig 4. *The power function versus* Δ $(T = 10, T = 20, T = 30, T = 50)$

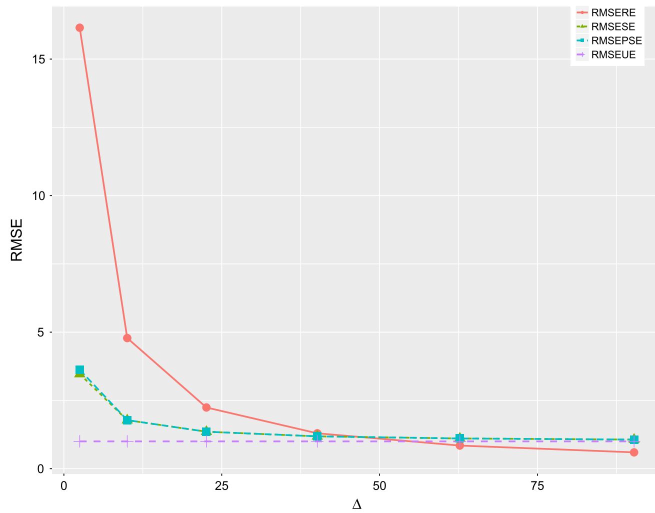


FIG 5. *RMSE of UE, RE, SE, PSE versus* Δ *(*$T = 20$*)*

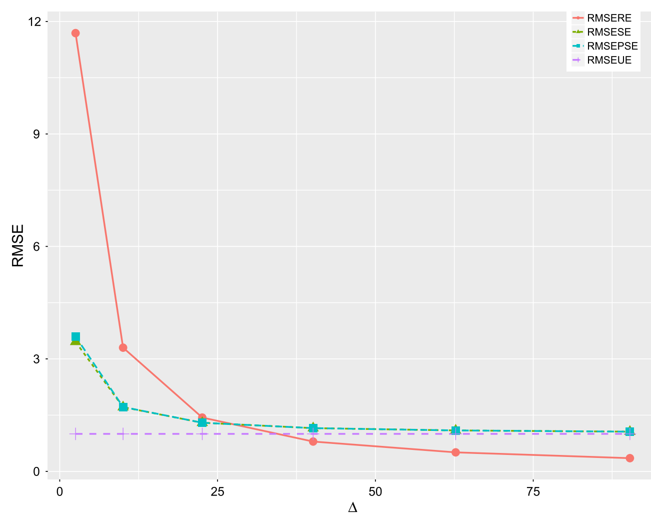


FIG 6. *RMSE of UE, RE, SE, PSE versus* Δ *(*$T = 50$*)*

the known number of the change-points case, we derived the joint asymptotic normality between the UE and the RE, under the set of local alternative restrictions; this generalizes particularly Corollary 4.2 in Chen *et al.* (2017) [4]. Third, we derived a test for testing the hypothesized restriction and we derived its asymptotic power. The proposed test is also useful for testing the absence of change points. Fourth, we constructed a class of shrinkage estimators (SEs) which are expected to be robust with respect to the restriction. Fifth, we studied

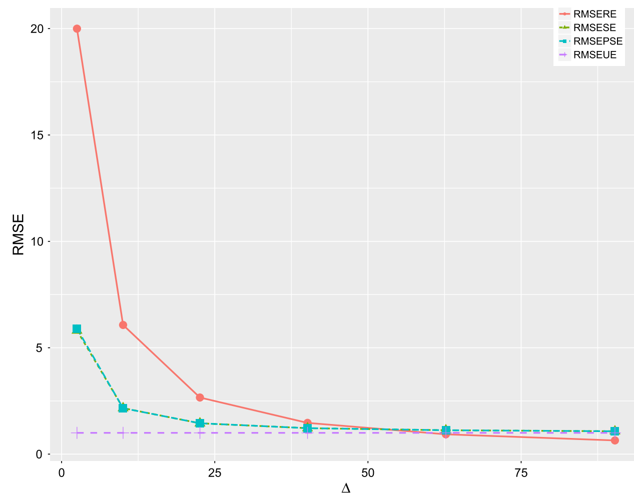


FIG 7. *RMSE of UMLE, RMLE, SE, PSE versus* Δ *(*$T = 20$*)*

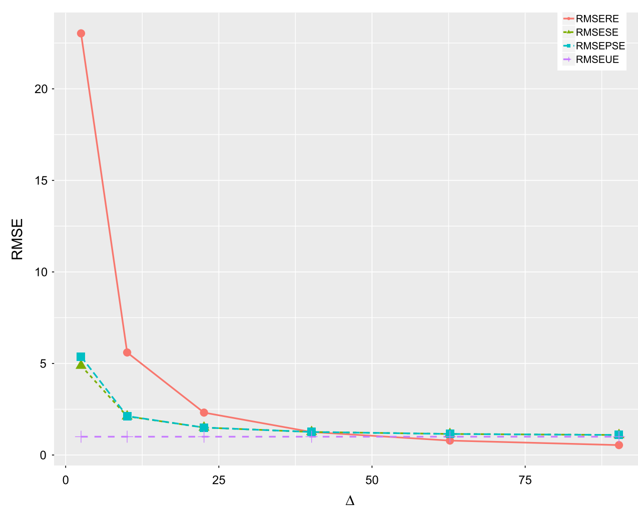


FIG 8. *RMSE of UMLE, RMLE, SE, PSE versus* Δ *(*$T = 50$*)*

the relative risk dominance of the proposed estimators. Specifically, we established that SEs dominate the UE and the RE performs very well near the null hypothesis, but this performs poorly when the restriction is seriously violated. On the top of these contributions, we weakened some conditions underlying the main results in Chen *et al.* (2017) [4]. In particular, it was established that the findings in Chen *et al.* (2017) [4] hold without their Assumption 2. We also added a condition about the initial value, of the SDE, which is necessary for the results in Chen *et al.* (2017) [4] to hold.

Appendix A: Technical results and proofs

Proof of Lemma 2.1. To prove the existence of the unique and strong solution, it suffices to verify that the coefficients of SDE satisfy both space-variable Lipschitz condition and the spatial growth condition. For more details, see the proof of Proposition 3.1 in Chen *et al.* (2017) [4]. The relation (2.4) follows from the relation (3.9) in Chen *et al.* (2017) [4]. Further, by combining Minkowsky's inequality along with Itô's isometry, we get $\sup_{t\geqslant 0} \mathrm{E}[|X_t|^2] < \infty$. □

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

The proof is similar to that given for Lemma 4.3 of Dehling *et al.* (2010) [6].

Proposition A.1. *If Assumptions 1-3 hold, then, $\hat{\phi}$ is a consistent estimator for ϕ, and there exists a $C > 0$ such that for every $\epsilon > 0$, for large T, $\mathrm{P}(T \max_{1\leqslant j\leqslant m} |\hat{\phi}_j - \phi_j| > C) < \epsilon$.*

The proof follows Proposition 4.2 of Chen *et al.* (2017) [4].

Proposition A.2. *Under Assumptions 1-3, we have that for large T,*
(i) $IC(m^0) < IC(m)$ a.s. $\forall\ m < m^0$ and (ii) $IC(m^0) < IC(m)$ a.s. $\forall\ m > m^0$.

The proof is similar to that given for Proposition 5.1 of Chen *et al.* (2017) [4].

Proof of Theorem 2.1. **(i)** Note that $\left(\sum_{j=1}^{m+1} b_j\right)^2 \leqslant (m+1)\sum_{j=1}^{m+1} b_j^2$, for all $b_1, b_2, \ldots, b_{m+1}$ real numbers. Then, by triangle inequality, we have

$$|\tilde{X}_t - X_t|^2 \leqslant 3(m+1)\sum_{j=1}^{m+1}|\tilde{h}_j(t) - h_j(t)|^2 + 3(m+1)\sum_{j=1}^{m+1}|\tilde{z}_j(t) - z_j(t)|^2 + 3(m+1)\sum_{j=1}^{m+1} e^{-2a_j t}|X_0|^2. \quad \text{(A.1)}$$

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

$$\begin{aligned}
[|\tilde{h}_j(t) - h_j(t)|] &\leqslant e^{-a_j t}\sum_{k=1}^{p}|\mu_{k,j}|K_\varphi \int_{-\infty}^{0} e^{a_j s} ds; \\
|\tilde{z}_j(t) - z_j(t)|^2 &= \sigma^2 e^{-2a_j t}\left|\int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s\right|^2.
\end{aligned}$$

Let $K^* = K_\mu^2 K_\varphi^2$. Then,

$$\begin{aligned}
\sum_{j=1}^{m+1} |\tilde{h}_j(t) - h_j(t)|^2 &\leqslant \sum_{j=1}^{m+1} \frac{e^{-2a_j t} K^*}{a_j^2}, \\
\sum_{j=1}^{m+1} |\tilde{z}_j(t) - z_j(t)|^2 &= \sigma^2 \sum_{j=1}^{m+1} e^{-2a_j t} \left| \int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s \right|^2 .
\end{aligned} \tag{A.2}$$

Then, by (A.1) and (A.2), we have

$$|\tilde{X}_t - X_t|^2 \leqslant C_0^2 e^{-2a_{(1)} t} \tag{A.3}$$

where

$$C_0^2 = 3(m+1) K_\mu^2 K_\varphi^2 \sum_{j=1}^{m+1} \frac{1}{a_j^2} + 3\sigma^2 (m+1) \sum_{j=1}^{m+1} \left| \int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s \right|^2 + 3(m+1)|X_0|^2,$$

the proof of Part (i) is completed if we prove that $\mathrm{E}\left(C_0^2\right) < \infty$. This holds iff

$$c_1 = \mathrm{E}\left(\sum_{j=1}^{m+1} \left| \int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s \right|^2 \right) = \sum_{j=1}^{m+1} \mathrm{E}\left(\left| \int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s \right|^2 \right) < \infty. \tag{A.4}$$

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

$$\mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s \right)^2 \right] = \mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\bar{B}_{-s} \right)^2 \right] = \mathrm{E}\left[\left(\int_{0}^{\infty} e^{-a_j u} d\bar{B}_u \right)^2 \right]. \tag{A.5}$$

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

$$\mathrm{E}[\mathcal{I}_U^2] = \mathrm{E}\left[\left(\int_0^U e^{-a_j u} d\bar{B}_u \right)^2 \right] = \mathrm{E}\left[\int_0^U e^{-2a_j u} du \right] = \frac{1}{2a_j}(1 - e^{-2a_j U}) \tag{A.6}$$

which is bounded for all $U \geqslant 0$. Thus, by L^2-bounded martingale convergence theorem, $\mathcal{I}_U \xrightarrow[U\to\infty]{a.s.} \mathcal{I}_\infty = \int_0^\infty e^{-a_j u} d\bar{B}_u$ *and* $\mathrm{E}[\mathcal{I}_\infty^2] < \infty$.

$$c_1 = \sum_{j=1}^{m+1} \mathrm{E}\left[\left(\int_{-\infty}^{0} e^{a_j s} d\tilde{B}_s \right)^2 \right] = \sum_{j=1}^{m+1} (1/(2a_j)) < \infty. \tag{A.7}$$

(ii) From Part (i), we get $|\tilde{X}_t - X_t| \xrightarrow[t\to\infty]{a.s.} 0$ and $|\tilde{X}_t - X_t| \xrightarrow[t\to\infty]{L^2} 0$. We also have

$$\left|\tilde{X}_t^2 - X_t^2\right| \leqslant \left[|\tilde{X}_t - X_t||\tilde{X}_t + X_t| \right] \leqslant |\tilde{X}_t - X_t||\tilde{X}_t| + |\tilde{X}_t - X_t||X_t|,$$

then

$$\left|\tilde{X}_t^2 - X_t^2\right| \leqslant C_0 e^{-a_{(1)} t}\left(|\tilde{X}_t| + |X_t|\right).$$

This gives

$$\sup_{2^n \leqslant t \leqslant 2^{n+1}} \left|\tilde{X}_t^2 - X_t^2\right| \leqslant C_0 e^{-a_{(1)} 2^n}\left(\sup_{2^n \leqslant t \leqslant 2^{n+1}} |\tilde{X}_t| + \sup_{2^n \leqslant t \leqslant 2^{n+1}} |X_t|\right). \quad \text{(A.8)}$$

Then, since the processes $\{X_t : t \geqslant 0\}$ and $\{\tilde{X}_t : t \geqslant 0\}$ have continuous trajectories, we get

$$\sup_{2^n \leqslant t \leqslant 2^{n+1}} \left|\tilde{X}_t^2 - X_t^2\right| \leqslant C_0 e^{-a_{(1)} 2^n}\left(|\tilde{X}_{t_n}| + |X_{t_n^*}|\right), \quad 2^n \leqslant t_n, t_n^* \leqslant 2^{n+1}. \text{(A.9)}$$

From Lemma 2.1, $\sup_{t \geqslant 0} \mathrm{E}[|X_t|^2] < \infty$. We have

$$\mathrm{E}[|\tilde{X}(t)|^2] \leqslant (m+1)\sum_{j=1}^{m+1} \mathrm{E}[(\tilde{h}_j(t) + \tilde{z}_j(t))^2].$$

Then,

$$\mathrm{E}[|\tilde{X}(t)|^2] \leqslant 2(m+1)\sum_{j=1}^{m+1}(\tilde{h}_j(t))^2 + 2(m+1)\sum_{j=1}^{m+1}\mathrm{E}[(\tilde{z}_j(t))^2]. \quad \text{(A.10)}$$

Note that, $(\tilde{h}_j(t))^2 = \left(e^{-a_j t}\sum_{k=1}^{p}\mu_{k,j}\int_{-\infty}^{t} e^{a_j s}\varphi_k(s)ds\right)^2$. Then,

$$\sum_{j=1}^{m+1}(\tilde{h}_j(t))^2 \leqslant \sum_{j=1}^{m+1}\left(e^{-2a_j t}K_\mu^2 K_\varphi^2 \frac{e^{2a_j t}}{a_j^2}\right) = \sum_{j=1}^{m+1}\frac{K_\mu^2 K_\varphi^2}{a_j^2} < \infty, \ \forall\, t \geqslant 0. \quad \text{(A.11)}$$

Since B_s and $\bar{B}_s$ are independent, by martingale property of Itô's integral, we have

$$\mathrm{E}[(\tilde{z}_j(t))^2] = \sigma^2 e^{-2a_j t}\left(\mathrm{E}\left[\left(\int_0^t e^{a_j s}dB_s\right)^2\right] + \mathrm{E}\left[\left(\int_{-\infty}^0 e^{a_j s}d\bar{B}_{-s}\right)^2\right]\right).$$

Also, by Itô's isometry, $\mathrm{E}\left[\left(\int_0^t e^{a_j s}dB_s\right)^2\right] = \mathrm{E}\left[\int_0^t e^{2a_j s}ds\right] = \frac{1}{2a_j}(e^{2a_j t}-1)$. By L^2-bounded martingale convergence theorem, (A.5), (A.6) and Itô's isometry,

$$\sum_{j=1}^{m+1}(\tilde{z}_j(t))^2 \leqslant \sigma^2\sum_{j=1}^{m+1}\left(e^{-2a_j t}\left(\frac{1}{2a_j}(e^{2a_j t}-1) + \frac{1}{2a_j}\right)\right) = \sum_{j=1}^{m+1}\frac{\sigma^2}{2a_j} < \infty. \quad \text{(A.12)}$$

Then, by combining (A.10), (A.11) and (A.12), we get

$$\mathrm{E}[|\tilde{X}(t)|^2] \leqslant 2(m+1)\sum_{j=1}^{m+1}\frac{K_\mu^2 K_\varphi^2}{a_j^2} + 2(m+1)\sum_{j=1}^{m+1}\frac{\sigma^2}{2a_j} < \infty,$$

and then, together with Lemma 2.1, we conclude that, there exists $K_2 > 0$ such that

$$\mathrm{E}(|\tilde{X}_{t_n}| + |X_{t_n^*}|) \leqslant \sup_{t\geqslant 0}\{\mathrm{E}[X_t^2]^{\frac{1}{2}} + \mathrm{E}[\tilde{X}_t^2]^{\frac{1}{2}}\} \leqslant K_2 < \infty.$$

Then
$\mathrm{E}\left\{\sup_{2^n\leqslant t\leqslant 2^{n+1}}\left|\tilde{X}_t^2 - X_t^2\right|\right\} \leqslant K_0^* K_2^* e^{-a_{(1)}2^n}$, for some $K_0^* > 0$, $K_2^* > 0$, and then

$$\sum_{n=1}^{\infty}\mathrm{E}\left\{\sup_{2^n\leqslant t\leqslant 2^{n+1}}\left|\tilde{X}_t^2 - X_t^2\right|\right\} < \infty.$$

Then, together with Markov's inequality, we get

$$\sum_{n=1}^{\infty}\mathrm{P}\left\{\sup_{2^n\leqslant t\leqslant 2^{n+1}}\left|\tilde{X}_t^2 - X_t^2\right| \geqslant \epsilon\right\} < \infty, \text{ for all } \epsilon > 0.$$

Hence, by Borel-Cantelli's lemma along with the fact that the process $\left\{\tilde{X}_t^2 - X_t^2 : t \geqslant 0\right\}$ has continuous paths, we get the second statement of Part (ii).

(iii) For $0 \leqslant a < b \leqslant 1$, we have

$$\begin{aligned}\left\|\int_{aT}^{bT}\tilde{X}_t\varphi(t)dt - \int_{aT}^{bT}X_t\varphi(t)dt\right\| &\leqslant \int_{aT}^{bT}|\tilde{X}_t - X_t|\,\|\varphi(t)\|\,dt \\ &\leqslant K_\varphi\int_{aT}^{bT}|\tilde{X}_t - X_t|dt. \end{aligned}\tag{A.13}$$

Then,

$$\sup_{0\leqslant a<b\leqslant 1}\left\|\frac{1}{T}\int_{aT}^{bT}\tilde{X}_t\varphi(t)dt - \frac{1}{T}\int_{aT}^{bT}X_t\varphi(t)dt\right\| \leqslant \frac{K_\varphi}{T}\int_0^T|\tilde{X}_t - X_t|dt. \tag{A.14}$$

By using part (i) along with the continuous version of Kronecker's lemma, we have

$$\frac{K_\varphi}{T}\int_0^T|\tilde{X}_t - X_t|dt \xrightarrow[T\to\infty]{a.s.} 0, \quad \text{and} \quad \frac{K_\varphi}{T}\int_0^T|\tilde{X}_t - X_t|dt \xrightarrow[T\to\infty]{L^2} 0,$$

then, by combining this last relation with (A.14), we get the statement in (iii).

(iv) For $0 \leqslant a < b \leqslant 1$, we have

$$\left|\frac{1}{T}\int_{aT}^{bT}\tilde{X}_t^2 dt - \frac{1}{T}\int_{aT}^{bT}X_t^2 dt\right| \leqslant \frac{1}{T}\int_{aT}^{bT}|\tilde{X}_t^2 - X_t^2|dt.$$

Then,

$$\sup_{0\leqslant a<b\leqslant 1}\left(\left|\frac{1}{T}\int_{aT}^{bT}\tilde{X}_t^2dt-\frac{1}{T}\int_{aT}^{bT}X_t^2dt\right|\right)\leqslant\frac{1}{T}\int_0^T|\tilde{X}_t^2-X_t^2|dt. \tag{A.15}$$

The proof follows directly from Part (ii) along with the continuous version of Kronecker's lemma, this completes the proof. □

Proposition A.3. *Suppose that Assumptions 1-2 hold, and suppose that* $T\geqslant 1\big/\min\limits_{1\leqslant j\leqslant m+1}(\phi_j-\phi_{j-1})$*, then* $Q(\phi,m)$ *is positive definite.*

Proof. Let $\boldsymbol{b}=(b_1,b_2,\ldots,b_{p+1})'$, with $b_i\in\mathbb{R}$, $i=1,2,\ldots,p+1$, let $\boldsymbol{b}_{(1)}=(b_1,b_2,\ldots,b_p)'$, i.e. $\boldsymbol{b}=\left(\boldsymbol{b}_{(1)}',b_{p+1}\right)'$. We have

$$\boldsymbol{b}'Q_{(\tau_{j-1},\tau_j)}\boldsymbol{b}=\int_{\phi_{j-1}T}^{\phi_jT}\|\boldsymbol{b}'\left(\boldsymbol{\varphi}'(t),-X(t)\right)'\|^2dt. \tag{A.16}$$

Then, since $\|\boldsymbol{b}'\left(\boldsymbol{\varphi}'(t),-X(t)\right)'\|^2\geqslant 0$, for all $\phi_{j-1}T\leqslant t\leqslant\phi_jT$, from (A.16), we have $\boldsymbol{b}'Q_{(\tau_{j-1},\tau_j)}\boldsymbol{b}\geqslant 0$ for all $\boldsymbol{b}\in\mathbb{R}^{p+1}$. Further, if $\boldsymbol{b}'Q_{(\tau_{j-1},\tau_j)}\boldsymbol{b}=0$, we must have

$$\|\boldsymbol{b}'\left(\boldsymbol{\varphi}'(t),-X(t)\right)'\|^2=0,$$

almost everywhere (a.e.) on $[\tau_{j-1},\,\tau_j]$, this implies that

$$\boldsymbol{b}'\left(\boldsymbol{\varphi}'(t),-X(t)\right)'=0,\qquad\text{a.e. on }[\tau_{j-1},\,\tau_j], \tag{A.17}$$

and then

$$\mathrm{P}\left(\left\{\omega:\boldsymbol{b}_{(1)}'\varphi(t)-b_{p+1}X(t,\omega)=0,\forall t\in[\tau_{j-1},\,\tau_j]\right\}\right)=1. \tag{A.18}$$

First, one can verify that whenever $b_{p+1}=0$ then $\boldsymbol{b}_{(1)}=0$. Thus, we first prove that $b_{p+1}=0$. To this end, suppose that $b_{p+1}\neq 0$. From Lemma 2.1, if $t\in[\tau_{j-1},\,\tau_j]$, we have

$$X(t)\big|X_{\tau_{j-1}}\sim\mathcal{N}\left(\mu(t,X_{\tau_{j-1}}),\,\Sigma_0(t)\right)$$

where $\mu(t,X_{\tau_{j-1}})=\mathrm{E}\left(X_j(t)\big|X_{\tau_{j-1}}\right)$, $\sigma_0(t)=\mathrm{Var}\left(X(t)\big|X_{\tau_{j-1}}\right)$, $t\in[\tau_{j-1},\,\tau_j]$. Then,

$$\left(b_{(1)}'\varphi(t)-b_{p+1}X(t)\right)\Big|X_{\tau_{j-1}}\sim\mathcal{N}\left(b_{(1)}'\varphi(t)-b_{p+1}\boldsymbol{\mu}(t,X_{\tau_{j-1}}),\,b_{p+1}^2\sigma_0^2(t)\right), \tag{A.19}$$

$t\in[\tau_{j-1},\,\tau_j]$. Further, since $a_j>0$, for $j=1,2,\ldots,m+1$, from Lemma 2.1, we have

$$\sigma_0^2(t)=\frac{\sigma^2}{2a_j}\left(1-e^{-2a_j(t-\tau_{j-1})}\right)>0, \tag{A.20}$$

for all $t \in]\tau_{j-1},\, \tau_j]$. Then, if $b_{p+1} \neq 0$, $b_{p+1}^2\sigma_0(t) > 0$, for all $t \in]\tau_{j-1},\, \tau_j]$, and then, by using (A.19), we get

$$\mathrm{P}\left(\left\{\omega : \boldsymbol{b}'_{(1)}\varphi(t) - b_{p+1}X(t,\omega) = 0, \forall t \in]\tau_{j-1},\, \tau_j]\right\} \Big| X_{\tau_{j-1}}\right) = 0,$$

this gives

$$\mathrm{P}\left(\left\{\omega : \boldsymbol{b}'_{(1)}\varphi(t) - b_{p+1}X(t,\omega) = 0, \forall t \in]\tau_{j-1},\, \tau_j]\right\}\right) = 0, \tag{A.21}$$

this is a contradiction with the relation (A.18). Therefore, $b_{p+1} = 0$. Hence, together with the relation (A.17), we get $\boldsymbol{b}'_{(1)}\boldsymbol{\varphi}(t) = 0$, for all $\tau_{j-1} = \phi_{j-1}T \leqslant t \leqslant \phi_j T = \tau_j$. Further, provided that $T \geqslant 1/(\phi_j - \phi_{j-1})$, we have $[0,\, 1] \subset [0,\, \phi_j T - \lfloor \phi_{j-1}T \rfloor]$, then, since from Assumption 2, the family $\{\varphi_1(t), \varphi_2(t), \ldots, \varphi_p(t)\}$ is linearly independent on $[0,\, 1]$, this implies that $\boldsymbol{b}_{(1)} = \mathbf{0}$, this completes the proof. □

Proof of Theorem 3.1. **(i)** We have, for $a \in (0,1)$, by the property of periodic function we have

$$\lim_{T\to\infty} \frac{1}{aT}\int_0^{aT} \varphi(t)\varphi'(t)dt = \boldsymbol{I}_p.$$

Therefore, for $0 \leqslant \phi_{j-1} < \phi_j \leqslant 1$, $j = 1, ..., m+1$,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \varphi(t)\varphi'(t)dt \xrightarrow[T\to\infty]{} (\phi_j - \phi_{j-1})\boldsymbol{I}_p.$$

This proves part (i).

(ii) Define $Y_i = \displaystyle\int_{i-1}^{i} \tilde{X}_j(t)\varphi(t)dt$, and let $u = t - i + 1 \in [0,1]$. By Lemma A.1, $\{\tilde{X}_j(u+i-1)\}_{i\in\mathbb{N}}$ is stationary and ergodic sequence, then, $\{Y_i\}_{i\in\mathbb{N}}$ is stationary and ergodic sequence. Then, by some algebraic computations along with Birkhoff's ergodic theorem, we prove that

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t\varphi(t)dt \xrightarrow[T\to\infty]{a.s. \quad \text{and } L^1} (\phi_j - \phi_{j-1})\mathrm{E}\left[\int_0^1 \tilde{X}_j(t)\varphi(t)dt\right]. \tag{A.22}$$

Since $\{\tilde{X}_j(t), t \geqslant 0\}$ and $\{\varphi(t), t \geqslant 0\}$ are L^2-bounded, with $\mathrm{E}[\tilde{X}_j(t)] = \tilde{h}_j(t) + \mathrm{E}[\tilde{z}_j(t)]$, we have

$$\mathrm{E}\left[\int_0^1 \tilde{X}_j(t)\varphi(t)dt\right] = \int_0^1 \mathrm{E}[\tilde{X}_j(t)]\varphi(t)dt = \int_0^1 \left(\tilde{h}_j(t) + \mathrm{E}[\tilde{z}_j(t)]\right)\varphi(t)dt, \tag{A.23}$$

with $\mathrm{E}[\tilde{z}_j(t)] = \sigma e^{-a_j t}\mathrm{E}\left[\displaystyle\int_0^\infty e^{-a_j u}d\bar{B}_u\right] + \sigma e^{-a_j t}\mathrm{E}\left[\displaystyle\int_0^t e^{a_j s}dB_s\right]$. As in proof of Theorem 2.1, let $\mathcal{I}_U = \displaystyle\int_0^U e^{-a_j u}d\bar{B}_u$, $U \geqslant 0$. By L^2-bounded martingale convergence theorem, (A.5) and (A.6), $\mathcal{I}_U \xrightarrow[T\to\infty]{L^2} \mathcal{I}_\infty = \int_0^\infty e^{-a_j u} d\bar{B}_u$, which implies that

$$\mathcal{I}_U \xrightarrow[T\to\infty]{L^1} \mathcal{I}_\infty = \int_0^\infty e^{-a_j u} d\bar{B}_u.$$

Then,

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

Hence, $\mathrm{E}[\tilde{z}_j(t)] = 0$. Then, $\mathrm{E}[\tilde{X}_t] = \tilde{h}_j(t)$, and then, by (A.22) and (A.23), we get

$$\begin{aligned}
&\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t \varphi(t) dt \xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1}) \int_0^1 \tilde{h}_j(t)\varphi(t) dt,\\
&\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t \varphi(t) dt \xrightarrow[T\to\infty]{L^1} (\phi_j - \phi_{j-1}) \int_0^1 \tilde{h}_j(t)\varphi(t) dt,
\end{aligned}$$

this proves Part (ii).

(iii) As in Part (ii), by Birkhoff's ergodic theorem, we get

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_t^2 dt \xrightarrow[T\to\infty]{a.s \ and \ L^1} (\phi_j - \phi_{j-1}) \mathrm{E}\left[\int_0^1 \tilde{X}_j^2(t) dt\right]. \tag{A.24}$$

One can verify that $\mathrm{E}[\tilde{z}_j(t)] = 0$, and $\mathrm{E}[\tilde{z}_j^2(t)] = \dfrac{\sigma^2}{2a_j}$. Then,

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_j^2(t) dt \xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})\omega_j,$$

and

$$\frac{1}{T}\int_{\phi_{j-1}T}^{\phi_j T} \tilde{X}_j^2(t) dt \xrightarrow[T\to\infty]{L^1} (\phi_j - \phi_{j-1})\omega_j,$$

this completes the proof. □

Proposition A.4. *If Assumption 2 holds. Then, Σ_j, $j = 1, ..., m+1$, and Σ are a positive definite matrices.*

Proof. By Schur complement theorem, Σ_j is positive definite iff $\omega_j - \Lambda_j^T I_P^{-1} \Lambda_j = \omega_j - \sum_{k=1}^{p} \left(\int_0^1 \tilde{h}_j(t)\varphi_k(t) dt\right)^2 > 0$. Indeed, by Bessel's inequality,

$$\int_0^1 \tilde{h}_j^2(t) dt \geqslant \sum_{k=1}^{p} \left(\int_0^1 \tilde{h}_j(t)\varphi_k(t) dt\right)^2,$$

and then

$$\omega_j - \sum_{k=1}^{p}\left(\int_0^1 \tilde{h}_j(t)\varphi_k(t)dt\right)^2 \geqslant \frac{\sigma^2}{2a_j} > 0,$$

this proves the first statement. The second statement follows from the fact that Σ is a block diagonal matrix whose diagonal block-components are positive definite, the proof is completed. □

Proof of Proposition 3.2. By Proposition 3.1, we get

$$\frac{1}{T}Q_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{a.s.} (\phi_j - \phi_{j-1})\Sigma_j.$$

Therefore, $T^{-1}Q(\phi, m) \xrightarrow[T\to\infty]{a.s.} \Sigma$. Further, since Σ_j and Σ are positive definite matrices,
$TQ^{-1}_{(\tau_{j-1},\tau_j)} \xrightarrow[T\to\infty]{a.s.} \frac{1}{\phi_j-\phi_{j-1}}(\Sigma_j)^{-1}$ and $TQ^{-1}(\phi, m) \xrightarrow[T\to\infty]{a.s.} \Sigma^{-1}$, this completes the proof. □

Proof of Proposition 3.3. **(i)** By combining (2.1), (2.2) and (2.10) along with some computations, we get

$$\tilde{r}_{(\tau_{j-1},\tau_j)} = \left(\int_{\tau_{j-1}}^{\tau_j} \varphi'(t)dX_t, -\int_{\tau_{j-1}}^{\tau_j} X_t dX_t\right)' = Q_{(\tau_{j-1},\tau_j)}\theta_j + \sigma \boldsymbol{M}_{(\tau_{j-1},\tau_j)},$$

this proves the Part (i).
(ii) $\boldsymbol{M}(\phi, m)$ is a martingale with quadratic variation $Q(\phi, m)$. Then, the proof follows by combining martingale central limit theorem along with Proposition 3.2 and Slutsky's theorem.
(iii) We have $\rho_T(\phi, m) = \sigma TQ^{-1}(\phi, m)\frac{1}{\sqrt{T}}\boldsymbol{M}(\phi, m)$. By combining Part (ii) and Proposition 3.2 along with Slutsky's theorem, we get

$$\begin{aligned}\rho_T(\phi, m) &= \sigma TQ^{-1}(\phi, m)\frac{1}{\sqrt{T}}\boldsymbol{M}(\phi, m) \xrightarrow[T\to\infty]{d} \sigma\Sigma^{-1}\boldsymbol{M}_0\\ &= \rho \sim \sigma\Sigma^{-1}\mathcal{N}_{(m+1)(p+1)}(0, \Sigma),\end{aligned}$$

this gives $\rho_T(\phi, m) \xrightarrow[T\to\infty]{d} \rho \sim \mathcal{N}_{(m+1)(p+1)}(0, \sigma^2\Sigma^{-1})$, this completes the proof. □

Proof of Proposition 3.4. One can verify that

$$(\rho'_T(\phi, m), \zeta'_T(\phi, m), \xi'_T(\phi, m))' = \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - GB \\ GB \end{pmatrix}\rho_T(\phi, m) + \begin{pmatrix} 0 \\ -Gr_0 \\ Gr_0 \end{pmatrix}.$$

By (3.6), we have

$$(I_{(m+1)(p+1)}, I_{(m+1)(p+1)} - GB, GB)'$$

$$\xrightarrow[T\to\infty]{P} (I_{(m+1)(p+1)}, I_{(m+1)(p+1)} - G^*B, G^*B)';$$
$$(0, -r_0'G', r_0'G')' \xrightarrow[T\to\infty]{P} (0, -r_0'G^{*\top}, r_0'G^{*\top})'. \tag{A.25}$$

By combining Proposition 3.3 and (A.25), and by Slutsky's theorem,

$$(\rho_T'(\phi, m), \zeta_T'(\phi, m), \xi_T'(\phi, m))'$$
$$\xrightarrow[T\to\infty]{d} \begin{pmatrix} I_{(m+1)(p+1)} \\ I_{(m+1)(p+1)} - G^*B \\ G^*B \end{pmatrix} \rho + \begin{pmatrix} 0 \\ -G^*r_0 \\ G^*r_0 \end{pmatrix} = \begin{pmatrix} \rho \\ \zeta \\ \xi \end{pmatrix}.$$

Further, one can verify that

$$G^*B\Sigma^{-1}B'G^{*\top} = \Sigma^{-1}B'(B\Sigma^{-1}B')^{-1}B\Sigma^{-1} = \Sigma^{-1}B'G^{*\top} = G^*B\Sigma^{-1}. \tag{A.26}$$

The rest of the proof follows directly from algebraic computation. □

Proof of Lemma 4.1. We have

$$\int_{\hat{a}T}^{\hat{b}T} Y_t dt - \int_{aT}^{bT} Y_t dt = \left(\int_0^{\hat{b}T} Y_t dt - \int_0^{bT} Y_t dt\right) - \left(\int_0^{\hat{a}T} Y_t dt - \int_0^{aT} Y_t dt\right).$$

Then, the result in Part (i) follows directly from Lemma 3.1 in Nkurunziza and Zhang (2018) [15], and the result in Parts (ii) and (iii) follow from Lemma 3.2 in Nkurunziza and Zhang (2018) [15]. □

Proof of Lemma 4.2. Since $\{Y_t : t \geqslant 0\}$ is bounded, there exits $K_3 > 0$ such that $\|Y_t\|^2 \leqslant K_3$ for all $t \geqslant 0$. Let $\epsilon > 0$, we have

$$\lim_{T\to\infty} \mathrm{P}\left(|\hat{b} - b| \geqslant \frac{\epsilon}{4K_3}\right) = 0. \tag{A.27}$$

Further, let

$$\begin{aligned}
\mathrm{I}_{11}(T) &= \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{b}T} Y_t dW_t - \int_0^{bT} Y_t dW_t\right\|^2 \mathbb{I}_{\{\hat{b}>b\}}\mathbb{I}_{\{|\hat{b}-b|<\frac{\epsilon}{4K_3}\}}\right], \\
\mathrm{I}_{12}(T) &= \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{b}T} Y_t dW_t - \int_0^{bT} Y_t dW_t\right\|^2 \mathbb{I}_{\{\hat{b}\leqslant b\}}\mathbb{I}_{\{|\hat{b}-b|\geqslant\frac{\epsilon}{4K_3}\}}\right], \\
\mathrm{I}_{21}(T) &= \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{b}T} Y_t dW_t - \int_0^{bT} Y_t dW_t\right\|^2 \mathbb{I}_{\{\hat{b}>b\}}\mathbb{I}_{\{|\hat{b}-b|<\frac{\epsilon}{4K_3}\}}\right], \\
\mathrm{I}_{22}(T) &= \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{b}T} Y_t dW_t - \int_0^{bT} Y_t dW_t\right\|^2 \mathbb{I}_{\{\hat{b}\leqslant b\}}\mathbb{I}_{\{|\hat{b}-b|\geqslant\frac{\epsilon}{4K_3}\}}\right].
\end{aligned} \tag{A.28}$$

One can verify that

$$\frac{1}{T}\mathrm{E}\left[\left\|\int_0^{\hat{b}T} Y_t dW_t - \int_0^{bT} Y_t dW_t\right\|^2\right] = \mathrm{I}_{11}(T)+\mathrm{I}_{12}(T)+\mathrm{I}_{21}(T)+\mathrm{I}_{22}(T). \quad \text{(A.29)}$$

Hence, the proof is completed if we prove that

$$\lim_{T\to\infty} \mathrm{I}_{11}(T) = \lim_{T\to\infty} \mathrm{I}_{12}(T) = \lim_{T\to\infty} \mathrm{I}_{21}(T) = \lim_{T\to\infty} \mathrm{I}_{22}(T) = 0.$$

To this end, we have

$$\begin{aligned}\mathrm{I}_{11}(T) &= \frac{1}{T}\mathrm{E}\left[\left\|\int_{bT}^{\hat{b}T} Y_t dW_t\right\|^2 \mathbb{I}_{\left\{\hat{b}>b\right\}}\mathbb{I}_{\left\{|\hat{b}-b|<\frac{\epsilon}{4K_3}\right\}}\right] \\ &= \frac{1}{T}\mathrm{E}\left[\left\|\int_{bT}^{\hat{b}T} Y_t dW_t\right\|^2 \mathbb{I}_{\left\{0<\hat{b}-b\leqslant\frac{\epsilon}{4K_3}\right\}}\right],\end{aligned}$$

then

$$\mathrm{I}_{11}(T) = \frac{1}{T}\mathrm{E}\left[\left\|\int_{bT}^{\hat{b}T} Y_t dW_t\right\|^2 \mathbb{I}_{\left\{0<T(\hat{b}-b)\leqslant\frac{\epsilon T}{4K_3}\right\}}\right]. \quad \text{(A.30)}$$

This gives

$$\mathrm{I}_{11}(T) = \frac{1}{T}\mathrm{E}\left[\left\|\int_0^{|\hat{b}-b|T} Y_{u+Tb} dW_{u,Tb}\right\|^2 \mathbb{I}_{\left\{0<T(\hat{b}-b)\leqslant\frac{\epsilon T}{4K_3}\right\}}\right], \quad \text{(A.31)}$$

where $W_{t,Tb} = W_{t+Tb} - W_{Tb}$, $t \geqslant 0$. From (A.31), we get

$$\mathrm{I}_{11}(T) \leqslant \frac{1}{T}\mathrm{E}\left[\sup_{0\leqslant t\leqslant\frac{\epsilon T}{4K_3}}\left\|\int_0^{t} Y_{u+Tb} dW_{u,Tb}\right\|^2 \mathbb{I}_{\left\{0<T(\hat{b}-b)\leqslant\frac{\epsilon T}{4K_3}\right\}}\right],$$

then,

$$\mathrm{I}_{11}(T) \leqslant \frac{1}{T}\mathrm{E}\left[\sup_{0\leqslant t\leqslant\frac{\epsilon T}{4K_3}}\left\|\int_0^{t} Y_{u+Tb} dW_{u,Tb}\right\|^2\right]. \quad \text{(A.32)}$$

Note that, since $\{Y_u : u \geqslant 0\}$ is a deterministic process, the stochastic process $\left\{\int_0^t Y_{u+Tb} dW_{u,Tb} : t \geqslant 0\right\}$ is martingale with respect to the natural filtration generated by $\{W_{t,Tb} : t \geqslant 0\}$. Then, by combining (A.32) with Doob's maximal inequality, we have

$$\mathrm{I}_{11}(T) \leqslant \frac{4}{T}\mathrm{E}\left[\int_0^{\frac{\epsilon T}{4K_3}} \|Y_{u+Tb}\|^2 du\right] \leqslant \frac{4}{T}\mathrm{E}\left[\int_0^{\frac{\epsilon T}{4K_3}} K_3 du\right] = \frac{4}{T}\frac{\epsilon T}{4K_3}K_3 = \epsilon. \quad \text{(A.33)}$$

This proves that $\lim_{T\to\infty} \mathrm{I}_{11}(T) = 0$. Further, since $0 < b < 1$ and $0 < \hat{b} \leqslant 1$ a.s., we have

$$\begin{aligned} \mathrm{I}_{12}(T) &= \frac{1}{T}\mathrm{E}\left[\left\|\int_{bT}^{\hat{b}T} Y_t dW_t\right\|^2 \mathbb{I}_{\left\{1\geqslant\hat{b}>b>0\right\}}\mathbb{I}_{\left\{|\hat{b}-b|\geqslant\frac{\epsilon}{4K_3}\right\}}\right] \\ &= \frac{1}{T}\mathrm{E}\left[\left\|\int_{bT}^{\hat{b}T} Y_t dW_t\right\|^2 \mathbb{I}_{\left\{(\hat{b}-b)>\frac{\epsilon}{4K_3}\right\}}\mathbb{I}_{\left\{T\geqslant T(\hat{b}-b)>0\right\}}\right]. \end{aligned}$$

This gives

$$\mathrm{I}_{12}(T) = \frac{1}{T}\mathrm{E}\left[\left\|\int_{0}^{|\hat{b}-b|T} Y_{u+Tb} dW_{u,Tb}\right\|^2 \mathbb{I}_{\left\{(\hat{b}-b)>\frac{\epsilon}{4K_3}\right\}}\mathbb{I}_{\left\{T\geqslant T(\hat{b}-b)>0\right\}}\right], \quad \text{(A.34)}$$

where $W_{t,Tb} = W_{t+Tb} - W_{Tb}$, $t \geqslant 0$. Then, by combining (A.34) with Cauchy-Schwarz's inequality, we have

$$\begin{aligned} \mathrm{I}_{12}(T) \leqslant \frac{1}{T}&\left\{\mathrm{E}\left[\left\|\int_{0}^{|\hat{b}-b|T} Y_{u+Tb} dW_{u,Tb}\right\|^4 \mathbb{I}_{\left\{T\geqslant T(\hat{b}-b)>0\right\}}\right]\right\}^{1/2} \\ &\times\left\{\mathrm{E}\left[\mathbb{I}_{\left\{(\hat{b}-b)>\frac{\epsilon}{4K_3}\right\}}\right]\right\}^{1/2}. \end{aligned}$$

Hence,

$$\begin{aligned} \mathrm{I}_{12}(T) \leqslant \frac{1}{T}&\left\{\mathrm{E}\left[\sup_{0\leqslant t\leqslant T}\left\|\int_{0}^{t} Y_{u+Tb} dW_{u,Tb}\right\|^4 \mathbb{I}_{\left\{T\geqslant T(\hat{b}-b)>0\right\}}\right]\right\}^{1/2} \\ &\times\left\{\mathrm{P}\left[\hat{b}-b>\frac{\epsilon}{4K_3}\right]\right\}^{1/2}, \end{aligned}$$

and this gives,

$$\mathrm{I}_{12}(T) \leqslant \frac{1}{T}\left\{\mathrm{E}\left[\sup_{0\leqslant t\leqslant T}\left\|\int_{0}^{t} Y_{u+Tb} dW_{u,Tb}\right\|^4\right]\right\}^{1/2}\left\{\mathrm{P}\left[\hat{b}-b>\frac{\epsilon}{4K_3}\right]\right\}^{1/2}.$$

Then, by Burkholder-Davis-Gundy's inequality, we have

$$\begin{aligned} \mathrm{I}_{12}(T) &\leqslant \frac{C_4}{T}\left\{\mathrm{E}\left[\left[\int_{0}^{T} \|Y_{u+Tb}\|^2 du\right]^2\right]\right\}^{1/2}\left\{\mathrm{P}\left[\hat{b}-b>\frac{\epsilon}{4K_3}\right]\right\}^{1/2} \\ &\leqslant \frac{C_4}{T}\left\{\mathrm{E}\left[\left[\int_{0}^{T} K_3 du\right]^2\right]\right\}^{1/2}\left\{\mathrm{P}\left[\hat{b}-b>\frac{\epsilon}{4K_3}\right]\right\}^{1/2}, \end{aligned}$$

where C_4 is a strictly positive real number. Therefore,

$$\mathrm{I}_{12}(T) \leqslant C_4 K_3 \left\{ \mathrm{P}\left[\hat{b} - b > \frac{\epsilon}{4K_3}\right]\right\}^{1/2}, \tag{A.35}$$

then, together with (A.27), we prove that $\lim_{T\to\infty} \mathrm{I}_{12}(T) = 0$. By using the similar techniques as for $\mathrm{I}_{11}(T)$ and $\mathrm{I}_{12}(T)$, one proves that

$$\lim_{T\to\infty} \mathrm{I}_{21}(T) = \lim_{T\to\infty} \mathrm{I}_{22}(T) = 0,$$

this completes the proof. □

Proof of Lemma 4.3. Let $f_0(\mu, s)$ represent the drift term of the SDE in (4.9). Let

$$\mathcal{L}_2(T; b, a) = \frac{1}{\sqrt{T}}\left(\int_0^{bT} Y_s f_0(\mu, s)ds - \int_0^{aT} Y_s f_0(\mu, s)ds\right).$$

By Itô's lemma,we have,

$$Y_t^2 = Y_0^2 + 2\int_0^t Y_s f_0(\mu, s)ds + \sigma^2 t + 2\sigma\int_0^t Y_s dW_s,$$

$t \geqslant 0$. Then, by triangle inequality,

$$\frac{1}{\sqrt{T}}\left|\int_{\hat{a}T}^{\hat{b}T} Y_s dW_s - \int_{aT}^{bT} Y_s dW_s\right| \leqslant \frac{|Y_{\hat{b}T}^2 - Y_{bT}^2|}{2\sigma\sqrt{T}} + \frac{\sigma}{2}\sqrt{T}|\hat{b} - b| + |\mathcal{L}_2(T; \hat{b}, b)|/\sigma + \frac{|Y_{\hat{a}T}^2 - Y_{aT}^2|}{2\sigma\sqrt{T}} + \frac{\sigma}{2}\sqrt{T}|\hat{a} - a| + |\mathcal{L}_2(T; \hat{a}, a)|/\sigma. \tag{A.36}$$

Since $\{Y_t, t \geqslant 0\}$ is L^2-bounded, we have

$$|Y_{\hat{b}T}^2 - Y_{bT}^2|/(2\sigma\sqrt{T}) \xrightarrow[T\to\infty]{P} 0, \text{ and } |Y_{\hat{a}T}^2 - Y_{aT}^2|/(2\sigma\sqrt{T}) \xrightarrow[T\to\infty]{P} 0. \tag{A.37}$$

Since there exists $\delta_0 > \frac{1}{2}$ such that $\max\left(|\hat{a} - a|, |\hat{b} - b|\right) = O_P(T^{-\delta_0})$, we have

$$\frac{\sigma}{2}\sqrt{T}|\hat{b} - b| \xrightarrow[T\to\infty]{P} 0, \text{ and } \frac{\sigma}{2}\sqrt{T}|\hat{a} - a| \xrightarrow[T\to\infty]{P} 0. \tag{A.38}$$

By Lemma 4.1, we have $|\mathcal{L}_2(T; \hat{a}, a)|/\sigma \xrightarrow[T\to\infty]{P} 0$ and $|\mathcal{L}_2(T; \hat{b}, b)|/\sigma \xrightarrow[T\to\infty]{P} 0$. Then, by combining this last relation with (A.36)-(A.38), we complete the proof. □

Proof of Proposition 4.3. **(i)** Since $\hat{\phi}_j$ and $\hat{\phi}_{j-1}$ are consistent estimators for ϕ_j and ϕ_{j-1}, $0 \leqslant \phi_{j-1} < \phi_j \leqslant 1$, $j = 1, ..., m+1$, by Lemma 4.3, we have, for $0 \leqslant$

$\phi_{j-1} < \phi_j \leqslant 1, j = 1, ..., m+1, \frac{1}{\sqrt{T}} \int_{\hat{\phi}_{j-1} T}^{\hat{\phi}_j T} X_t dW_t - \frac{1}{\sqrt{T}} \int_{\phi_{j-1} T}^{\phi_j T} X_t dW_t \xrightarrow[T\to\infty]{P} 0.$
Then, by Lemma 4.2, we also have

$$\frac{1}{\sqrt{T}} \int_{\hat{\phi}_{j-1} T}^{\hat{\phi}_j T} \varphi(t) dW_t - \frac{1}{\sqrt{T}} \int_{\phi_{j-1} T}^{\phi_j T} \varphi(t) dW_t \xrightarrow[T\to\infty]{P} 0,$$

these two conditions complete the proof of Part (i).
(ii) By Propositions 3.3, we have $\frac{1}{\sqrt{T}} \boldsymbol{M}(\phi, m) \xrightarrow[T\to\infty]{d} \boldsymbol{M}_0 \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma)$ and by Lemma 4.2 and 4.3,

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

Hence, by Slutsky's theorem,

$$\frac{1}{\sqrt{T}} M(\hat{\phi}, m) = \frac{1}{\sqrt{T}} \Big(M(\hat{\phi}, m) - \boldsymbol{M}(\phi, m) \Big) + \frac{1}{\sqrt{T}} \boldsymbol{M}(\phi, m) \xrightarrow[T\to\infty]{d} \boldsymbol{M}_0 \\ \sim \mathcal{N}_{(m+1)(p+1)}(0, \Sigma),$$

this proves Part (ii).
(iii) It suffices to note that $\rho_T(\hat{\phi}, m) = \sigma T Q(\hat{\phi}, m)^{-1} \frac{1}{\sqrt{T}} M(\hat{\phi}, m)$. Then, the proof follows from the second statement and Proposition 4.2 along with Slutsky's theorem. □

Proof of Lemma 5.1. For the sake of simplicity, for two q-column vectors a and b, we denote $a \leqslant b$ to stand for $a_i \leqslant b_i$, $i = 1, 2, \ldots, q$. Let x be a point of continuity of the cdf of $X(m)$. We have

$$\begin{aligned} \lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leqslant x) = & \lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leqslant x, \hat{m} = m) \\ & + \lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leqslant x, \hat{m} \neq m); \\ \lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x) = & \lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x, \hat{m} = m) \\ & + \lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x, \hat{m} \neq m). \end{aligned}$$

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

$$\lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leqslant x) = \lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x, \hat{m} = m), \tag{A.39}$$

$$\lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x) = \lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x, \hat{m} = m). \tag{A.40}$$

By combining (A.39), (A.40) and $\lim_{T\to\infty} \mathrm{P}(X_T(m) \leqslant x) = \mathrm{P}(X(m) \leqslant x)$, we get $\lim_{T\to\infty} \mathrm{P}(X_T(\hat{m}) \leqslant x) = \mathrm{P}(X(m) \leqslant x)$, this completes the proof. □

Proof of Proposition 7.1. By Corollary 7.1, we have

$$\varrho_0(\Delta) = \mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) - \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) = r_0' G^{*\top} \Omega G^* r_0 \\ -\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}).$$

We observe that, since $G^* = \Sigma^{-1} B' (B \Sigma^{-1} B')^{-1}$ and $\Gamma = \frac{1}{\sigma^2} B' (B \Sigma^{-1} B')^{-1} B$, $G^{*\top} \Gamma G^* = \frac{1}{\sigma^2} (B \Sigma^{-1} B')^{-1}$, which is positive definite for $\sigma > 0$.

Then, if $\Delta > 0$, by Theorem 2.4.7 in Mathai and Provost (1992) [13], we have

$$\lambda_1 \Delta - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \leqslant \varrho_0(\Delta) \leqslant \lambda_n \Delta - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}), \qquad (\mathrm{A.41})$$

From (A.41), provided that $\lambda_1 \Delta - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \geqslant 0$, $\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) \geqslant \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right)$. Further, by (A.41), provided that $\lambda_n \Delta - \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \leqslant 0$, $\mathrm{ADR}\left(\tilde{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) \leqslant \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right)$, this completes the proof. $\square$

Proof of Proposition 7.2. By Theorem 7.1, we have

$$\begin{aligned} \varrho_1(\Delta) &= \mathrm{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) - \mathrm{ADR}\left(\hat{\theta}(\hat{\phi}, \hat{m}), \theta, \Omega\right) \\ &= (q+2)(q-2) r_0' G^{*\top} \Omega G^* r_0 \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] \\ &\quad - (q-2)\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1})(2\mathrm{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]). \end{aligned}$$

Then, by the identity in Saleh (2006, p. 32) [18], we have $\Delta \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] = \mathrm{E}[\chi_{q+2}^{-2}(\Delta)] - (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]$, this gives

$$\begin{aligned} \varrho_1(\Delta) &= (q+2)(q-2) r_0' G^{*\top} \Omega G^* r_0 \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] \\ &\quad - (q-2)\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1})(2\Delta \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] + (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]). \end{aligned}$$

Note that, $\Delta \geqslant 0$ and that, $\Delta = 0$ if and only if $r_0 = 0$. If $\Delta = 0$, we have

$$\varrho_1(\Delta) = -(q-2)^2 \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \mathrm{E}[\chi_{q+2}^{-4}] \leqslant 0.$$

Let $H = \left[1 - ((q+2) r_0' G^{*\top} \Omega G^* r_0) \Big/ (2\Delta \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}))\right]$. If $\Delta > 0$, we have

$$\varrho_1(\Delta) = -(q-2)\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \left[2\Delta \mathrm{E}[\chi_{q+4}^{-4}(\Delta)] H + (q-2)\mathrm{E}[\chi_{q+2}^{-4}(\Delta)]\right].$$

Thus, $\varrho_1(\Delta) \leqslant 0$ for all $\Delta > 0$ provided that $q > 2$ and $H \geqslant 0$. This last inequality holds if and only if $(\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}))/\lambda_n \geqslant (q+2)/2$, this completes the first part of the proof. Further, let $\varrho_2(\Delta) = \mathrm{ADR}\left(\hat{\theta}^s, \theta, \Omega\right) - \mathrm{ADR}\left(\hat{\theta}^{s+}, \theta, \Omega\right)$, from Theorem 7.1, have

$$\begin{aligned} \varrho_2(\Delta) &= -2 r_0' G^{*\top} \Omega G^* r_0 \mathrm{E}[(1 - (q-2)\chi_{q+2}^{-2}(\Delta)) \mathbb{I}_{\{\chi_{q+2}^2(\Delta) < q-2\}}] \\ &\quad + \sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \mathrm{E}[(1 - (q-2)\chi_{q+2}^{-2}(\Delta))^2 \mathbb{I}_{\{\chi_{q+2}^2(\Delta) < q-2\}}] \end{aligned}$$

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

The proof follows from the inequalities

$$-2r_0' G^{*\top} \Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))\mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}] \geqslant 0, \tag{A.42}$$

$$\sigma^2 \mathrm{trace}(\Omega G^* B \Sigma^{-1}) \mathrm{E}[(1-(q-2)\chi_{q+2}^{-2}(\Delta))^2 \mathbb{I}_{\{\chi_{q+2}^2(\Delta)<q-2\}}] \geqslant 0, \tag{A.43}$$

$$r_0' G^{*\top} \Omega G^* r_0 \mathrm{E}[(1-(q-2)\chi_{q+4}^{-2}(\Delta))^2 \mathbb{I}_{\{\chi_{q+4}^2(\Delta)<q-2\}}] \geqslant 0. \tag{A.44}$$

□

Acknowledgments

The first author would like to acknowledge the financial support received from the Natural Sciences and Engineering Research Council of Canada (NSERC). Further, the authors would like to thank the referee for helpful comments.

References

[1] Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*, in B. Petrov, F. Csaki, 2nd *International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó*, 267–281. MR0483125

[2] Auger, I. & Lawrence, C. (1989). Algorithms for the optimal identification of segment neighborhoods. *Bulletin of Mathematical Biology*, **51**(1), 39–54. MR0978902

[3] Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**(1), 47–78. MR1616121

[4] Chen, F., Mamon, R., & Davison, M. (2017). Inference for a mean-reverting stochastic process with multiple change points. *Electronic Journal of Statisitics*, **11**, 2199–2257. MR3654824

[5] Chen, F. & Nkurunziza, S. (2015). Optimal method in multiple regression with structural changes. *Bernoulli*, **21** (4), 2217–2241. MR3378465

[6] Dehling, H., Franke, B., & Kott, T. (2010). Drift estimation for a periodic mean reversion process. *Statistical Inference for Stochastic Process (SISP)*, **13**, 175–192. MR2729648

[7] Dehling, H., Franke, B., Kott, T., & Kulperger, R. (2014). Change point testing for the drift parameters of a periodic mean reversion process. *SISP*, **17**(1), 1–18. MR3178011

[8] Jackson, B., Scargle, J., Barnes, D., Arabhi, S., Alt, A., Gioumousis, P., Gwin, E., Sangtrakulcharoen, P., Tan, L. & Tsai, T. (2002). An algorithm for optimal partitioning of data on an interval. *IEEE Signal Processing Letters*, **12**, 105–108.

[9] Killick, R., Fearnhead, P.& Eckley, I. (2012). Optimal detection of change points with a linear computational cost. *Journal of the American Statistical Association*, **107**(500), 1590–1598. MR3036418

[10] Kutoyants, Y. A. (2004). *Statistical inference for ergodic diffusion processes*. Springer-Verlag, London. MR2144185

[11] Lansky, P., & Sacerdote, L. (2001). The Ornstein-Uhlenbeck neuronal model with signal-dependent noise. *Physics Letters A*, **285**(3-4), 132–140. MR1844801

[12] Liptser, R. S., & Shiryaev, A. N. (2001). *Statistics of Random Processes I*: I. General Theory (Vol. 1). Springer-Verlag, Berlin Heidelberg. MR1800857

[13] Mathai, A. M. & Provost, S. B. (1992). *Quadratic forms in random variables: theory and applications*. CRC Press, New York. MR1192786

[14] Nkurunziza, S. (2012). The risk of pretest and shrinkage estimators. *Statistics*, **46**(3), 305–312. MR2929155

[15] Nkurunziza, S. & Zhang P. (2018). Estimation and testing in generalized mean-reverting processes with change-point. *SISP*, **21**, 191–215. MR3769837

[16] Perron, P. & Qu, Z. (2006). Estimating restricted structural change models. *Journal of Econometrics*, **134**, 373–399. MR2328414

[17] Rohlfs, R., Harrigan, P., & Nielsen, R. (2010). Modeling gene expression evolution with an extended Ornstein-Uhlenbeck process accounting for within-species variation. *Scandinavian Journal of Statistics*, **37**(2), 200–220. MR2682296

[18] Saleh, AME. (2006). *Theory of preliminary test and Stein-type estimation with applications*, **517**. Wiley, New Jersey. MR2218139

[19] Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**(2), 461–464. MR0468014

[20] Sen, PK. & Saleh AKME. (1987). On preliminary test and shrinkage M-estimation in linear models. *The Annals of Statistics*, **15**(4), 1580–1592. MR0913575