Bernoulli **23**(1), 2017, 159–190
DOI: 10.3150/15-BEJ740

Mixed domain asymptotics for a stochastic process model with time trend and measurement error

CHIH-HAO CHANG[1], HSIN-CHENG HUANG[2,*] and CHING-KANG ING[2,3,**]

[1]*Institute of Statistics, National University of Kaohsiung, Kaohsiung 811, Taiwan. E-mail: jhow@nuk.edu.tw*
[2]*Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan. E-mail: *hchuang@stat.sinica.edu.tw; **cking@stat.sinica.edu.tw*
[3]*Department of Economics, National Taiwan University, Taipei 106, Taiwan. E-mail: cking1@ntu.edu.tw*

We consider a stochastic process model with time trend and measurement error. We establish consistency and derive the limiting distributions of the maximum likelihood (ML) estimators of the covariance function parameters under a general asymptotic framework, including both the fixed domain and the increasing domain frameworks, even when the time trend model is misspecified or its complexity increases with the sample size. In particular, the convergence rates of the ML estimators are thoroughly characterized in terms of the growing rate of the domain and the degree of model misspecification/complexity.

Keywords: asymptotic normality; consistency; exponential covariance function; fixed domain asymptotics; increasing domain asymptotics

1. Introduction

Learning the covariance structure of a stochastic process from data is a fundamental prerequisite for problems such as prediction, classification and control. For example, to do prediction for an Ornstein–Uhlenbeck (OU) (Uhlenbeck and Ornstein [15]) process $\eta(s)$, $s \in [0, 1]$ with mean 0 and covariance function

$$\operatorname{cov}\bigl(\eta(s_1), \eta(s_2)\bigr) = \sigma_{0,\eta}^2 \exp\bigl(-\kappa_0 |s_1 - s_2|\bigr), \tag{1.1}$$

where $\sigma_{0,\eta}^2, \kappa_0 > 0$ are unknown, Ying [17] proposed the maximum likelihood (ML) estimators for $\sigma_{0,\eta}^2$ and κ_0 based on discrete observations $\eta(s_1), \dots, \eta(s_n)$ with $0 \le s_1 < \cdots < s_n \le 1$, and established the root-n consistency of the corresponding ML estimator for $\sigma_{0,\eta}^2\kappa_0$. Note that since the probability measures induced by two OU processes are absolutely continuous with respect to each other if and only if their $\sigma_{0,\eta}^2\kappa_0$ values are equal (Ibragimov and Rozanov [9]), the parameters in (1.1) are asymptotically identifiable up to $\sigma_{0,\eta}^2\kappa_0$. However, when the OU process is subject to measurement error, the so-called "nugget" effect (see, for example, Cressie [7]) may deteriorate the performance of the ML estimators. In particular, Chen, Simpson and Ying [6] showed that the ML estimator for $\sigma_{0,\eta}^2\kappa_0$ becomes fourth-root-n consistent, depicting the effect of measurement error in estimating the exponential covariance parameters in (1.1). On the other

1350-7265

hand, they also proved that the ML estimator of the measurement-error variance has the usual root-n consistency. In fact, a similar phenomenon can also be found in a driftless Brownian motion (BM) process with measurement error. Let this error-contaminated process be denoted by $y(t), t \in [0, 1]$. Having observed $y(0), y(1/n), \ldots, y(1)$, Stein [14] showed that a modified ML (MML) estimator of the ratio of the variance of the increments of the BM process to that of measurement error is only fourth-root-n consistent, whereas the corresponding MML estimator of the measurement-error variance still remains root-n consistent. Similar asymptotic results for the ML estimators of the two variances have also been established by Aït-Sahalia, Mykland and Zhang [1].

In this article, we shall superimpose a time trend (regression) term on an OU process with measurement error in order to accommodate a broader range of applications. Specifically, we propose the following model for a real-valued stochastic process $\{Z(s); s \in D \subset \mathbb{R}\}$:

$$Z(s) = \beta_0 + \sum_{j=1}^{p} \beta_j x_j(s) + \eta(s) + \epsilon(s), \tag{1.2}$$

where $\mathbf{x}(s) = (x_1(s), \ldots, x_p(s))'$ is a p-dimensional time trend vector, $\eta(s)$ is a zero-mean OU process with covariance function defined in (1.1), $\epsilon(s)$ is a zero-mean Gaussian measurement error with $\mathrm{E}(\epsilon(s)\epsilon(t)) = \theta_{0,1} I_{\{s=t\}}$ for some unknown $\theta_{0,1} > 0$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)'$ is a $(p+1)$-dimensional constant vector, and $\{\mathbf{x}(s)\}$, $\{\eta(s)\}$ and $\{\epsilon(s)\}$ are independent. In a computer experiment, $\eta(s)$ in (1.2) can be used to describe the systematic departure of the response $Z(s)$ from the linear model $\beta_0 + \sum_{j=1}^{p} \beta_j x_j(s)$ and $\epsilon(s)$ denotes the measurement error. For more details, we refer the reader to Sacks, Schiller and Welch [13] and Ying [17]. Model (1.2) can also be applied to one-dimensional geostatistical modeling and $\eta(\cdot)$ therein corresponds to a commonly used exponential covariance model; see Ripley [12] and Cressie [7] for numerous examples. Denote the true time trend by

$$\mu_0(s) = Z(s) - \eta(s) - \epsilon(s), \tag{1.3}$$

where $\{\mu_0(s)\}$ is independent of $\{\eta(s)\}$ and $\{\epsilon(s)\}$, and define $\mathbf{x}_0(s) = (1, \mathbf{x}'(s))'$. The time trend model $\boldsymbol{\beta}'\mathbf{x}_0(s)$ in (1.2) is said to be correctly specified if

$$\mu_0(s) = \boldsymbol{\beta}'\mathbf{x}_0(s) \qquad \text{for some } \boldsymbol{\beta} \in \mathbb{R}^{p+1}, \tag{1.4}$$

and misspecified otherwise. In this article, we shall allow $\boldsymbol{\beta}'\mathbf{x}_0(s)$ to be misspecified, which further increases the flexibility of model (1.2). However, a misspecified time trend will usually create extra challenges in estimating covariance parameters. This motivates us to ask how the ML estimators of the covariance parameters in model (1.2) perform when the corresponding time trend model is subject to misspecification.

To facilitate exposition, we assume in the sequel that $D = [0, n^{\delta}]$ for some $\delta \in [0, 1)$, and the data are observed regularly at $s_i = in^{-(1-\delta)}$, $i = 1, \ldots, n$. In addition, we also allow that the number of regressors (model complexity) $p = p_n$ grows to infinity in order to reduce the model bias. When $\delta = 0$, the domain $D = [0, 1]$ has been considered by the aforementioned authors, and the setup is called fixed domain asymptotics. On the other hand, when $\delta > 0$, the domain D grows to infinity as $n \to \infty$ with a faster growing rate for a larger δ value, and the setup is referred to as the increasing domain asymptotics, even though the minimum inter-data distance $n^{-(1-\delta)}$ goes to zero. This is different from the increasing domain setup considered by Zhang and Zimmerman [18], in which the minimum distance between sampling points is bounded away from zero. By incorporating both fixed and increasing domains, our *mixed* domain asymptotic framework enables us to explore the interplay between the model misspecification/complexity and the growing rate of D on the asymptotic behaviors of the ML estimators, thereby leading to an intriguing answer to the above question.

Re-parameterizing (1.1) by $\theta_{0,2} = \sigma^2_{0,\eta}\kappa_0$ and $\theta_{0,3} = \kappa_0$, the covariance parameter vector in model (1.2) can be written as $\boldsymbol{\theta}_0 = (\theta_{0,1}, \theta_{0,2}, \theta_{0,3})'$. Let Θ, the parameter space, be a compact set in $(0, \infty)^3$ and suppose $\boldsymbol{\theta}_0 \in \Theta$. Based on model (1.2) and observations $(\mathbf{x}'(s_i), Z(s_i)), i = 1, \ldots, n$, we estimate $\boldsymbol{\theta}_0$ using the ML estimator $\hat{\boldsymbol{\theta}}$, which satisfies

$$\ell(\hat{\boldsymbol{\theta}}) = \sup_{\boldsymbol{\theta}=(\theta_1,\theta_2,\theta_3)'\in\Theta} \ell(\boldsymbol{\theta}),$$

where

$$\begin{aligned} \ell(\boldsymbol{\theta}) = &-\tfrac{1}{2}n\log(2\pi) - \tfrac{1}{2}\log\det\bigl(\boldsymbol{\Sigma}(\boldsymbol{\theta})\bigr) \\ &- \tfrac{1}{2}\mathbf{Z}'\bigl(\mathbf{I} - \mathbf{M}(\boldsymbol{\theta})\bigr)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\bigl(\mathbf{I} - \mathbf{M}(\boldsymbol{\theta})\bigr)\mathbf{Z}, \end{aligned} \tag{1.5}$$

is known as the profile log-likelihood function, in which $\mathbf{Z} = (Z(s_1), \ldots, Z(s_n))'$,

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}) + \theta_1\mathbf{I}, \tag{1.6}$$

with

$$\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}) = \left(\frac{\theta_2}{\theta_3}\exp\bigl(-\theta_3|s_i - s_j|\bigr)\right)_{1\le i,j\le n},$$

and

$$\mathbf{M}(\boldsymbol{\theta}) = \mathbf{X}\bigl(\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{X}\bigr)^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}), \tag{1.7}$$

with $\mathbf{X} = (\mathbf{x}_0(s_1), \ldots, \mathbf{x}_0(s_n))'$ being full rank almost surely (a.s.). It is not difficult to show that the ML estimator of $\boldsymbol{\beta}$ is given by $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})$, where

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) = \bigl(\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{X}\bigr)^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{Z}.$$

However, since model (1.2) can be misspecified, investigating the asymptotic properties of $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})$ is beyond the scope of this paper.

Let $\boldsymbol{\mu}_0 = (\mu_0(s_1), \ldots, \mu_0(s_n))'$ and $\boldsymbol{\epsilon} = (\epsilon(s_1), \ldots, \epsilon(s_n))'$. By $\mathbf{M}(\boldsymbol{\theta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})$, (1.3) and (1.5), we have

$$\begin{aligned} -2\ell(\boldsymbol{\theta}) = &-2\ell_0(\boldsymbol{\theta}) + \boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\bigl(\mathbf{I} - \mathbf{M}(\boldsymbol{\theta})\bigr)\boldsymbol{\mu}_0 \\ &+ 2\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\bigl(\mathbf{I} - \mathbf{M}(\boldsymbol{\theta})\bigr)(\boldsymbol{\eta} + \boldsymbol{\epsilon}) \\ &- (\boldsymbol{\eta} + \boldsymbol{\epsilon})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})(\boldsymbol{\eta} + \boldsymbol{\epsilon}), \end{aligned} \tag{1.8}$$

where with $h(\boldsymbol{\theta}) = (\boldsymbol{\eta}+\boldsymbol{\epsilon})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})(\boldsymbol{\eta}+\boldsymbol{\epsilon}) - \operatorname{tr}(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$,

$$\ell_0(\boldsymbol{\theta}) \equiv -\tfrac{1}{2}\bigl\{n\log(2\pi) + \log\det\bigl(\boldsymbol{\Sigma}(\boldsymbol{\theta})\bigr) + \operatorname{tr}\bigl(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\bigr) + h(\boldsymbol{\theta})\bigr\}, \tag{1.9}$$

is the log-density function for $\boldsymbol{\eta}+\boldsymbol{\epsilon}$. As will be seen in Section 2, the contribution of the time trend to $-2\ell(\boldsymbol{\theta})$ is mainly made by

$$\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\bigl(\mathbf{I}-\mathbf{M}(\boldsymbol{\theta})\bigr)\boldsymbol{\mu}_0 - (\boldsymbol{\eta}+\boldsymbol{\epsilon})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})(\boldsymbol{\eta}+\boldsymbol{\epsilon}). \tag{1.10}$$

The first term above, vanishing when (1.4) holds true, is due to model misspecification, and the second term, having an order of magnitude $O_p(p_n)$ uniformly over Θ (see Lemma 4.7), is related to model complexity. We therefore introduce

$$R(\Theta) = \max\Bigl\{\sup_{\boldsymbol{\theta}\in\Theta}\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\bigl(\mathbf{I}-\mathbf{M}(\boldsymbol{\theta})\bigr)\boldsymbol{\mu}_0,\, p_n\Bigr\}, \tag{1.11}$$

as a uniform bound for (1.10) over Θ. Let $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)' = \hat{\boldsymbol{\theta}}$. The growing rates of D needed for $\hat{\theta}_i$, $i = 1, 2, 3$, to achieve consistency are given in the next theorem in terms of the order of magnitude of $R(\Theta)$. It provides a preliminary answer to the question of whether the covariance structures of $\boldsymbol{\eta}$ and $\boldsymbol{\epsilon}$ can be learnt from data under possible model misspecification.

Theorem 1.1. *Suppose*

$$R(\Theta) = O_p\bigl(n^{\xi}\bigr), \tag{1.12}$$

for some $\xi \in [0, 1)$. *Then, for* $\delta \in [0, 1)$,

$$\hat{\theta}_1 = \theta_{0,1} + o_p(1) \qquad \text{if } 0 \le \xi < 1, \tag{1.13}$$

$$\hat{\theta}_2 = \theta_{0,2} + o_p(1) \qquad \text{if } 0 \le \xi < (1+\delta)/2, \tag{1.14}$$

$$\hat{\theta}_3 = \theta_{0,3} + o_p(1) \qquad \text{if } 0 \le \xi < \delta. \tag{1.15}$$

Theorem 1.1 shows that as long as (1.12) holds true, $\hat{\theta}_1$ is a consistent estimator of $\theta_{0,1}$, regardless of the value of δ. In contrast, in order for $\hat{\theta}_2$ and $\hat{\theta}_3$ to achieve consistency, one would require $0 \le \xi < (1+\delta)/2$ and $0 \le \xi < \delta$, respectively. In fact, these two constraints cannot be weakened because we provide counterexamples in Section 3 illustrating that $\hat{\theta}_3$ is no longer consistent when $\xi = \delta$, and both $\hat{\theta}_2$ and $\hat{\theta}_3$ fail to achieve consistency if $\xi = (1+\delta)/2$. It is worth mentioning that $\ell(\boldsymbol{\theta})$ is highly convoluted due to the involvement of regression terms, making it difficult to establish consistency of $\hat{\boldsymbol{\theta}}$. Our strategy is to decompose the nonstochastic part of $-2\ell(\boldsymbol{\theta})$ into several layers whose first three leading orders are $n_1 \equiv n$, $n_2 \equiv n^{(1+\delta)/2}$ and $n_3 \equiv n^{\delta}$, respectively, and express the remainder stochastic part as the sum of $h(\boldsymbol{\theta})$ and two other terms that can be uniformly expressed as $O_p(R(\Theta))$ and $o_p(n^{\delta})$; see (4.17). One distinctive characteristic of these nonstochastic layers is that the coefficient associated with the ith $(1 \le i \le 3)$ leading layer only depends on $\theta_1, \dots, \theta_i$. When (1.12) is assumed, this hierarchical layer structure together with some uniform bounds established for the second moments of

$h(\boldsymbol{\theta})$ enables us to derive the consistency of $\hat{\boldsymbol{\theta}}$ in the order of $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{\theta}_3$ by focusing on one layer and one parameter at a time. Let $\mathrm{tr}(\mathbf{A})$ denote the trace of a matrix $\mathbf{A}$. As shown in the proof of Theorem 1.1, the uniform bounds for $h(\boldsymbol{\theta})$ are first expressed in terms the supremums of $\mathrm{tr}\{(\partial^m \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)/\partial\theta_{j_1}\cdots\theta_{j_m})^2\}$, $1 \le m \le 3$, $j_1 < \cdots < j_m \in \{1,2,3\}$, or other similar trace terms such as those given in (4.31). These expressions are obtained using the idea that the sup-norms of a sufficiently smooth function can be bounded above by suitable integral norms, as suggested in Lai [10], Chan and Ing [3] and Chan, Huang and Ing [2]. We then carefully calculate the orders of magnitude of the aforementioned traces, yielding uniform bounds in terms of $n, n^{(1+\delta)/2}$ or n^{δ}. Note that Dahlhaus [8] has applied the chaining lemma (see Pollard [11]) to obtain uniform probability bounds for some quadratic forms of a discrete time long-memory process. However, since no rates have been reported in his bounds, his approach may not be directly applicable here.

Whereas Theorem 1.1 has demonstrated the performance of $\hat{\boldsymbol{\theta}}$ from the perspective of consistency, the questions of what are the convergence rates of and whether there are central limit theorems (CLTs) for $\hat{\theta}_i, i = 1, \ldots, 3$, still remain unanswered. The next section is devoted to these questions. In particular, it is shown in Theorem 2.2 that for $n_i \to \infty$, $1 \le i \le 3$, $\hat{\theta}_i - \theta_{0,i} = O_p(\max\{n^{\xi} n_i^{-1}, n_i^{-1/2}\})$ if $n^{\xi} = o(n_i)$, and $n_i^{1/2}(\hat{\theta}_i - \theta_{0,i})$ has a limiting normal distribution if $n^{\xi} = o(n_i^{1/2})$. Since the time trend is involved, our proof of Theorem 2.2 is somewhat nonstandard. We first obtain the initial convergence rates of $\hat{\boldsymbol{\theta}}$ using the standard Taylor expansion and an argument similar but subtler than the one used in the proof of Theorem 1.1. Using these initial rates, we can improve the convergence results through the same argument. We then repeat this iterative procedure until the final convergence results are established.

The rest of this article is organized as follows. In Section 2, we begin by establishing the CLT for $\hat{\theta}_i, i = 1, \ldots, 3$ in situations where p_n is fixed and the regression model is correctly specified (namely, (1.4) is true); see Theorem 2.1. We subsequently drop these two restrictions and report in Theorem 2.2 the most general convergence results of this paper. In Section 3, we provide two counterexamples showing that the results obtained in Theorem 1.1 are difficult to improve. The proofs of all theorems and corollaries in the first three sections are given in Section 4. The proofs of the auxiliary lemmas used in Section 4 are provided in the supplementary material (Chang, Huang and Ing [5]) in light of space constraint. Before leaving this section, we remark that although our results are derived under the Gaussianity of $\{\eta(t)\}$ and $\{\epsilon(t)\}$, similar results can be obtained when either $\{\eta(t)\}$ or $\{\epsilon(t)\}$ is not (but pretended to be) Gaussian, provided some fourth moment information is available. On the other hand, while we allow the time trend to be misspecified, we preclude a misspecified covariance model. The interested reader is referred to Xiu [16] for some asymptotic results on the ML estimators when the covariance model considered in Stein [14] or Aït-Sahalia, Mykland and Zhang [1] is misspecified.

2. Central limit theorems and rates of convergence

In this section, we begin with establishing the asymptotic normality of $\hat{\theta}_i$, $1 \le i \le 3$, in situations where the regression model is correctly specified and p_n is fixed.

Theorem 2.1. *Assume that* (1.4) *holds and* p_n *is a fixed nonnegative integer.* (*Note that these assumptions yield* $\xi = 0$ *in* (1.12).) *Then for* $\delta \in [0, 1)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big), \tag{2.1}$$

$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big), \tag{2.2}$$

and for $\delta \in (0, 1)$,

$$n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) \xrightarrow{d} N(0, 2\theta_{0,3}). \tag{2.3}$$

One of the easiest ways to understand Theorem 2.1 is to link the result to the Fisher information matrix. Straightforward calculations show that under the assumption of Theorem 2.1, the diagonal elements of the Fisher information matrix evaluated at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ are given by

$$\begin{aligned}
-\mathrm{E}\bigg(\frac{\partial^2}{\partial\theta_1^2}\ell(\boldsymbol{\theta}_0)\bigg) &= \frac{1}{2}\operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_0)\big) + O(1),\\
-\mathrm{E}\bigg(\frac{\partial^2}{\partial\theta_2^2}\ell(\boldsymbol{\theta}_0)\bigg) &= \frac{1}{2\theta_{0,2}^2}\operatorname{tr}\big\{\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\big)^2\big\} + O(1),\\
-\mathrm{E}\bigg(\frac{\partial^2}{\partial\theta_3^2}\ell(\boldsymbol{\theta}_0)\bigg) &= \frac{1}{2}\operatorname{tr}\bigg\{\bigg(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)}{\partial\theta_3}\bigg)^2\bigg\} + O(1) \qquad \text{if } 0 < \delta < 1,
\end{aligned} \tag{2.4}$$

where the trace terms are solely contributed by the log-density (log-likelihood) function for $\boldsymbol{\eta} + \boldsymbol{\epsilon}$ (defined in (1.9)), and the $O(1)$ terms, which vanish if the time trend is known to be zero, are related to the model complexity. Moreover, by (4.11), (4.13) and (4.22),

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{2n}\operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_0)\big) &= \frac{1}{2\theta_{0,1}^2},\\
\lim_{n\to\infty}\frac{1}{2\theta_{0,2}^2 n^{(1+\delta)/2}}\operatorname{tr}\big\{\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\big)^2\big\} &= \frac{1}{2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}},\\
\lim_{n\to\infty}\frac{1}{2n^\delta}\operatorname{tr}\bigg\{\bigg(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)}{\partial\theta_3}\bigg)^2\bigg\} &= \frac{1}{2\theta_{0,3}} \qquad \text{if } 0 < \delta < 1.
\end{aligned} \tag{2.5}$$

It is interesting pointing out that the denominator on the right-hand side of the first equation of (2.5) coincides exactly with the limiting variance in (2.1). This is reminiscent of a conventional asymptotic theory for the ML estimate which says that the limiting variance of the ML estimate is the reciprocal of the corresponding Fisher information number. On the other hand, while the reciprocals of the right-hand sides of the second and third identities of (2.5) are the same as the limiting variances in (2.2) and (2.3), the divergence rates of the corresponding trace terms $n^{(1+\delta)/2}$ and n^δ are much slower than n. In fact, they are equal to the divergence rates of the second and third leading layers of the nonstochastic part of $-2\ell(\boldsymbol{\theta})$; see (4.17). These findings reveal that the amounts of information related to $\theta_{0,i}$'s have different orders of magnitude, thereby leading to different normalizing constants in the CLTs for $\hat{\theta}_i$'s.

The next theorem improves Theorem 2.1 by deriving rates of convergence of $\hat{\theta}_i$, $1 \le i \le 3$, without requiring $\xi = 0$ in (1.12). It further shows that CLTs for $\hat{\theta}_i$, $1 \le i \le 3$, are still possible if the model misspecification/complexity associated with the time trend has an order of magnitude smaller than $n^{1/2}$, $n^{(1+\delta)/4}$ and $n^{\delta/2}$, respectively.

Theorem 2.2. *Suppose that* (1.12) *is true. Then for* $\delta \in [0, 1)$,

$$\hat{\theta}_1 - \theta_{0,1} = \begin{cases} O_p\big(n^{-1/2}\big); & \textit{if } \xi < 1/2, \\ O_p\big(n^{-(1-\xi)}\big); & \textit{if } 1/2 \le \xi < 1, \end{cases}$$

$$\hat{\theta}_2 - \theta_{0,2} = \begin{cases} O_p\big(n^{-(1+\delta)/4}\big); & \textit{if } \xi < (1+\delta)/4, \\ O_p\big(n^{-\{(1+\delta)/2-\xi\}}\big); & \textit{if } (1+\delta)/4 \le \xi < (1+\delta)/2, \end{cases}$$

and for $\delta \in (0, 1)$,

$$\hat{\theta}_3 - \theta_{0,3} = \begin{cases} O_p\big(n^{-\delta/2}\big); & \textit{if } \xi < \delta/2, \\ O_p\big(n^{-(\delta-\xi)}\big); & \textit{if } \delta/2 \le \xi < \delta. \end{cases}$$

In addition, for $\delta \in [0, 1)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big); \qquad \textit{if } \xi < 1/2,$$

$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big); \qquad \textit{if } \xi < (1+\delta)/4,$$

and for $\delta \in (0, 1)$,

$$n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) \xrightarrow{d} N(0, 2\theta_{0,3}); \qquad \textit{if } \xi < \delta/2.$$

Recall that $n_1 = n$, $n_2 = n^{(1+\delta)/2}$ and $n_3 = n^{\delta}$. It is shown in (2.4) and (2.5) that, ignoring the constant, the amount of information regarding $\theta_{0,i}$ contained in $\boldsymbol{\eta} + \boldsymbol{\epsilon}$ is n_i, $1 \le i \le 3$. On the other hand, as will become clear later, n^{ξ} can be used to measure the amount of information contaminated by model misspecification/complexity (again ignoring the constant). Therefore, the first part of Theorem 2.2 delivers nothing more than the simple idea that

$$\begin{aligned} &\text{Rate of convergence of } \hat{\theta}_i \\ &\quad = \max\Bigg\{\frac{\text{Amount of information contaminated by model misspecification/complexity}}{\text{Amount of information regarding } \theta_{0,i} \text{ contained in } \boldsymbol{\eta} + \boldsymbol{\epsilon}}, \\ &\qquad \frac{1}{(\text{Amount of information regarding } \theta_{0,i} \text{ contained in } \boldsymbol{\eta} + \boldsymbol{\epsilon})^{1/2}}\Bigg\}, \end{aligned} \tag{2.6}$$

provided that

$$\begin{aligned}&\text{Amount of information contaminated by model misspecification/complexity}\\&\quad< \text{Amount of information regarding } \theta_{0,i} \text{ contained in } \boldsymbol{\eta}+\boldsymbol{\epsilon}.\end{aligned} \tag{2.7}$$

Note that the second term on the right-hand side of (2.6) is the best rate one can expect when the time trend is known to be zero. The second part of Theorem 2.2 further indicates that the CLTs for $\hat{\theta}_i$'s in Theorem 2.1 carry over to situations where (2.7) holds with the right-hand side replaced by its square root. To the best of our knowledge, this is one of the most general CLTs established for $\hat{\theta}_i$'s. In the following, we present two specific examples illustrating how the asymptotic behavior of $\hat{\theta}_i$'s is affected by the interaction between ξ and δ. In the first example, the model misspecification yields $R(\Theta) = O(n^{\delta})$, and hence $\xi = \delta$. According to Theorem 2.2, the CLTs for $\hat{\theta}_1$ and $\hat{\theta}_2$ hold for a certain range of δ.

Corollary 2.1. *Consider the intercept-only model of* (1.2) *with $p_n = 0$. Suppose that $\mu_0(s) = \beta_{0,0} + \beta_{0,1} n^{-\delta} s$, where $\beta_{0,0}$ and $\beta_{0,1}$ are nonzero constants. Then for $\delta \in [0, 1)$,*

$$R(\Theta) = O(n^{\delta}), \tag{2.8}$$

$$\begin{aligned} n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) &\xrightarrow{d} N(0, 2\theta_{0,1}^2); \qquad \delta \in [0, 1/2),\\ n^{1-\delta}(\hat{\theta}_1 - \theta_{0,1}) &= O_p(1); \qquad \delta \in [1/2, 1), \end{aligned} \tag{2.9}$$

$$\begin{aligned} n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) &\xrightarrow{d} N(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}); \qquad \delta \in [0, 1/3),\\ n^{(1-\delta)/2}(\hat{\theta}_2 - \theta_{0,2}) &= O_p(1); \qquad \delta \in [1/3, 1). \end{aligned} \tag{2.10}$$

We remark that the scaling factor $n^{-\delta}$ is introduced for the linear term, $x_1(s) = n^{-\delta}s$, so that $\frac{1}{n^{\delta}}\int_0^{n^{\delta}} (x_1(s) - \bar{x}_1)^2\, ds$ does not depend on n, where $\bar{x}_1 = \frac{1}{n^{\delta}}\int_0^{n^{\delta}} x_1(s)\, ds$. The model misspecification in the next example results in $R(\Theta) = O_p(n^{(1+\delta)/2})$, yielding $\xi = (1+\delta)/2$. Therefore, $\hat{\theta}_1$ is guaranteed to be consistent in view of Theorem 2.2.

Corollary 2.2. *Consider the same setup as in Corollary 2.1 except that $\mu_0(s) = \beta_{0,0} + \beta_{0,1}x(s)$, where $x(\cdot)$ is generated from a zero-mean Gaussian spatial process with covariance function*

$$\operatorname{cov}(x(s), x(s')) = \frac{\theta_{1,2}}{\theta_{1,3}} \exp(-\theta_{1,3}|s - s'|); \qquad s, s' \in [0, n^{\delta}],$$

for some constants $\theta_{1,2}, \theta_{1,3} > 0$. Then for $\delta \in [0, 1)$,

$$R(\Theta) = O_p(n^{(1+\delta)/2}), \tag{2.11}$$

$$\hat{\theta}_1 = \theta_{0,1} + O_p(n^{-(1-\delta)/2}). \tag{2.12}$$

It is worth noting that $\hat{\theta}_3$ is inconsistent under the setup of Corollary 2.1. Moreover, both $\hat{\theta}_2$ and $\hat{\theta}_3$ are inconsistent under the setup of Corollary 2.2. These inconsistency results will be reported in detail in the next section. Before closing this section we remark that our theoretical results on $\hat{\boldsymbol{\theta}}$ can be used to make statistical inference about the regression function. For example, when (1.4) holds and $p_n \geq 1$ is a fixed integer, the convergence rate of $\hat{\boldsymbol{\theta}}$ obtained in Theorem 2.1 plays an indispensable role in analyzing the convergence rate of the ML estimator, $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})$, of $\boldsymbol{\beta}$. Recently, by making use of Theorems 2.1 and 2.2, Chang, Huang and Ing [4] established the first model selection consistency result under the mixed domain asymptotic framework. Moreover, some technical results established in the proofs of Theorems 2.1 and 2.2 have been used by Chang, Huang and Ing [4] to develop a model selection consistency result under a misspecified covariance model.

3. Counterexamples

Using the examples constructed in Corollaries 2.1 and 2.2, we show in this section that the constraints $\xi < \delta$ and $\xi < (1+\delta)/2$ imposed in Theorem 1.1 for the consistency of $\hat{\theta}_3$ and $\hat{\theta}_2$, respectively, cannot be relaxed.

Corollary 3.1. *Under the setup of Corollary* 2.1,

$$\hat{\theta}_3 = \frac{12\theta_{0,2}}{12\theta_{0,2} + \beta_{0,1}^2\theta_{0,3}}\theta_{0,3} + o_p(1); \qquad \delta \in (0,1). \tag{3.1}$$

Corollary 3.2. *Under the setup of Corollary* 2.2,

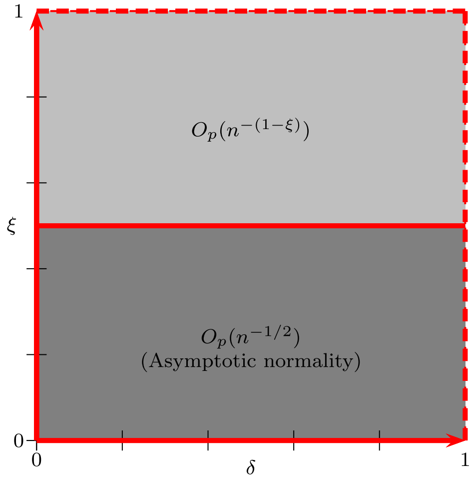
$$\hat{\theta}_2 = \theta_{0,2} + \theta_{1,2}\beta_{0,1}^2 + o_p(1); \qquad \delta \in [0,1), \tag{3.2}$$

$$\hat{\theta}_3 = \frac{\theta_{0,2} + \beta_{0,1}^2\theta_{1,2}}{\beta_{0,1}^2\theta_{1,2}\theta_{1,3}^{-1} + \theta_{0,3}\theta_{0,3}^{-1}} + o_p(1); \qquad \delta \in (0,1). \tag{3.3}$$

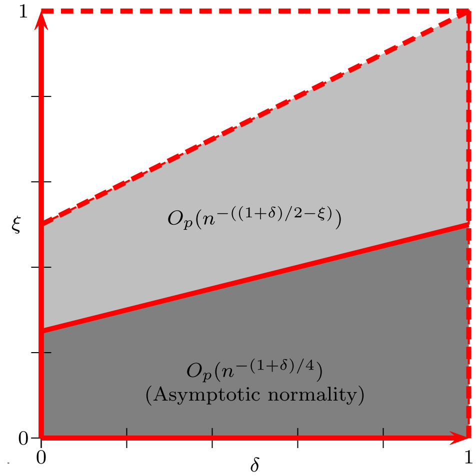
All the above results can be illustrated by Figure 1, in which some change point behavior of $\hat{\theta}_i$'s (in terms of modes of convergence) is exhibited when (δ, ξ) runs through the region $[0,1) \times [0,1)$.

4. Proofs of the theorems and corollaries

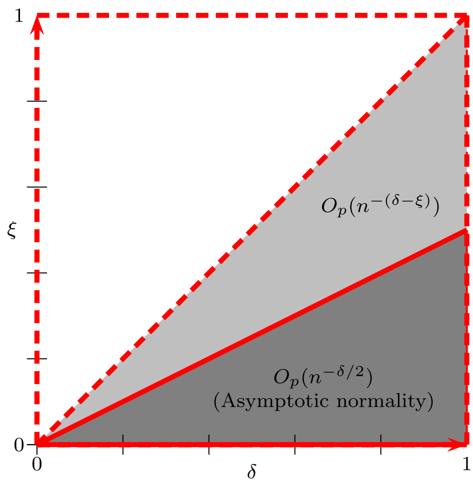
In this section, we first prove the consistency of $\hat{\boldsymbol{\theta}}$ in Section 4.1. The proofs of CLTs for $\hat{\boldsymbol{\theta}}$ with and without the restrictions of correct specification and fixed dimension on the time trend model are given in Sections 4.2 and 4.3, respectively. The proofs of Corollaries 2.1 and 3.1 and those of Corollaries 2.2 and 3.2 are provided in Sections 4.4 and 4.5, respectively.



(a) Convergence rates of $\hat{\theta}_1 - \theta_{0,1}$



(b) Convergence rates of $\hat{\theta}_2 - \theta_{0,2}$



(c) Convergence rates of $\hat{\theta}_3 - \theta_{0,3}$

Figure 1. Convergence rates of $\hat{\theta}_i$ to $\theta_{0,i}$ with respect to (δ, ξ), where $i = 1, \ldots, 3$, δ is the growing rate of the domain and ξ satisfies $R(\Theta) = O_p(n^{\xi})$. Note that $\hat{\theta}_i$ also possesses asymptotic normality when (δ, ξ) falls in the dark gray regions, but may fail to achieve consistency when (δ, ξ) falls in the white regions or on the dash lines. In addition, the points on the lines between the light and dark gray area are referred to as the change points merely in the modes of convergence but not in the convergence rate scenario.

4.1. Proof of Theorem 1.1

To prove Theorem 1.1, we need a series of auxiliary lemmas, Lemmas 4.1–4.9. Lemma 4.1 gives a modified Cholesky decomposition for $\mathbf{\Sigma}^{-1}(\boldsymbol{\theta})$, which can be used to prove Lemma 4.2, asserting that the eigenvalues of $\mathbf{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{\Sigma}(\boldsymbol{\theta}_0)$ are uniformly bounded above and below. Lemmas 4.3 and 4.4 provide the orders of magnitude of the Cholesky factors of $\mathbf{\Sigma}^{-1}(\boldsymbol{\theta})$ and the products of $\mathbf{\Sigma}_\eta(\boldsymbol{\theta})$ and these factors. Based on Lemmas 4.2–4.4, Lemma 4.5 establishes asymptotic expressions for the key components of the nonstochastic part of $-2\ell(\boldsymbol{\theta})$, and Lemma 4.6 provides the orders of magnitude of $\mathbf{\Sigma}^{-1}(\boldsymbol{\theta})\,\partial\mathbf{\Sigma}(\boldsymbol{\theta})/\partial\theta_i$; $i = 1, 2, 3$. Lemmas 4.2 and 4.6 can be used in conjunction with Lemma 4.9, which provides uniform bounds for quadratic forms in i.i.d. random variables, to analyze the asymptotic behavior of $h(\boldsymbol{\theta})$; see (4.19). Lemmas 4.7 and 4.8 explore the effects of the time trend model on $-2\ell(\boldsymbol{\theta})$.

Lemma 4.1. *Let $\mathbf{\Sigma}(\boldsymbol{\theta})$ be given by (1.6) with $\theta_1 \geq 0$, $\theta_2 > 0$ and $\theta_3 > 0$. Then*

$$\mathbf{\Sigma}^{-1}(\boldsymbol{\theta}) = \mathbf{G}_n(\boldsymbol{\theta})'\mathbf{T}_n^{-1}(\boldsymbol{\theta})\mathbf{G}_n(\boldsymbol{\theta}), \tag{4.1}$$

where

$$\mathbf{G}_n(\boldsymbol{\theta}) \equiv \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho_n & 1 & 0 & \ddots & \vdots \\ 0 & -\rho_n & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho_n & 1 \end{pmatrix}_{n\times n},$$

$$\mathbf{T}_n(\boldsymbol{\theta}) = \mathbf{D}_n(\boldsymbol{\theta}) + \theta_1\mathbf{G}_n(\boldsymbol{\theta})\mathbf{G}_n(\boldsymbol{\theta})',$$

$\rho_n = \exp(-\theta_3 n^{-(1-\delta)})$, *and*

$$\mathbf{D}_n(\boldsymbol{\theta}) \equiv \frac{\theta_2}{\theta_3}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1-\rho_n^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1-\rho_n^2 \end{pmatrix}_{n\times n}.$$

Lemma 4.2. *Let $\lambda_{\max}(\mathbf{A})$ and $\lambda_{\min}(\mathbf{A})$ denote the maximum and minimum eigenvalues of the matrix $\mathbf{A}$. For $\mathbf{\Sigma}(\boldsymbol{\theta})$ given by (1.6), suppose that $\Theta \subset (0,\infty)^3$ is compact. Then,*

$$\begin{aligned} 0 &< \liminf_{n\to\infty} \inf_{\boldsymbol{\theta}\in\Theta} \lambda_{\min}\big(\mathbf{\Sigma}^{-1/2}(\boldsymbol{\theta})\mathbf{\Sigma}(\boldsymbol{\theta}_0)\mathbf{\Sigma}^{-1/2}(\boldsymbol{\theta})\big) \\ &\leq \limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Theta} \lambda_{\max}\big(\mathbf{\Sigma}^{-1/2}(\boldsymbol{\theta})\mathbf{\Sigma}(\boldsymbol{\theta}_0)\mathbf{\Sigma}^{-1/2}(\boldsymbol{\theta})\big) < \infty. \end{aligned} \tag{4.2}$$

Lemma 4.3. *Under the setup of Lemma 4.1, for any $\boldsymbol{\theta} \in \Theta \in (0,\infty)^3$, where Θ is compact, and $\delta \in [0, 1)$, the following equation holds uniformly over Θ:*

$$\operatorname{tr}\big(\mathbf{T}_n^{-2}(\boldsymbol{\theta})\big) = \frac{n^{(5-3\delta)/2}}{2^{7/2}\theta_1^{1/2}\theta_2^{3/2}} + o\big(n^{(5-3\delta)/2}\big). \tag{4.3}$$

Lemma 4.4. *Under the setup of Lemma 4.3, for any $\boldsymbol{\theta} \in \Theta$,*

$$\begin{aligned}
&\mathbf{G}_n(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\mathbf{G}_n(\boldsymbol{\theta})' \\
&= \frac{\theta_{0,2}\rho_n}{\theta_{0,3}\rho_{0,n}}\big(1-\rho_{0,n}^2\big)\mathbf{I} + \left(1-\frac{\rho_n}{\rho_{0,n}}\right)(1-\rho_n\rho_{0,n})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0) \\
&\quad + \frac{\theta_{0,2}}{\theta_{0,3}}\left\{\left(1-\frac{\rho_n}{\rho_{0,n}}\right)\big(\mathbf{v}_0\mathbf{e}_1' + \mathbf{e}_1\mathbf{v}_0'\big) + \rho_n^2\mathbf{e}_1\mathbf{e}_1'\right\},
\end{aligned} \tag{4.4}$$

where $\mathbf{e}_1 = (1, 0, \dots, 0)'$, $\mathbf{v}_0 = (1, \rho_{0,n}, \dots, \rho_{0,n}^{n-1})$ and $\rho_{0,n} = \exp(-\theta_{0,3}n^{-(1-\delta)})$. In addition, for any $\delta \in [0, 1)$,

$$\sup_{\boldsymbol{\theta}\in\Theta} \mathbf{v}_0'\mathbf{T}_n^{-1}(\boldsymbol{\theta})\mathbf{v}_0 = O\big(n^{2(1-\delta)}\big), \tag{4.5}$$

$$\sup_{\boldsymbol{\theta}\in\Theta} \mathbf{v}_0'\mathbf{T}_n^{-1}(\boldsymbol{\theta})\mathbf{e}_1 = O\big(n^{1-\delta}\big), \tag{4.6}$$

$$\sup_{\boldsymbol{\theta}\in\Theta} \mathbf{e}_1'\mathbf{T}_n^{-1}(\boldsymbol{\theta})\mathbf{e}_1 = O(1). \tag{4.7}$$

Furthermore, for any $\delta \in (0, 1)$,

$$\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\Big(\big(\mathbf{T}_n^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta})\big)^2\Big) = \frac{1}{4\theta_3^3}n^{4-3\delta} + o\big(n^{4-3\delta}\big). \tag{4.8}$$

Lemma 4.5. *Under the setup of Lemma 4.3, the following equations hold uniformly over Θ:*

$$\begin{aligned}
\log\big(\det\big(\boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big) &= n\log\theta_1 + \left(\frac{2\theta_2}{\theta_1}\right)^{1/2} n^{(1+\delta)/2} - \left(\frac{\theta_2}{\theta_1}+\theta_3\right)n^\delta \\
&\quad - \frac{1-\delta}{2}\log n + o\big(n^\delta\big) + O(1),
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
\operatorname{tr}\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big) &= \frac{\theta_{0,1}}{\theta_1}n - \frac{\theta_{0,1}}{2\theta_1}\left(\frac{2\theta_2}{\theta_1}\right)^{1/2} n^{(1+\delta)/2} \\
&\quad + \frac{\theta_{0,2}}{(2\theta_1\theta_2)^{1/2}}n^{(1+\delta)/2} + \frac{\theta_{0,2}(\theta_3^2-\theta_{0,3}^2)}{2\theta_2\theta_{0,3}}n^\delta \\
&\quad + o\big(n^\delta\big) + O(1).
\end{aligned} \tag{4.10}$$

Lemma 4.6. *Under the setup of Lemma 4.3, the following equations hold uniformly over Θ:*

$$\operatorname{tr}\Big(\big(\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big)^2\Big) = \left(\frac{\theta_2}{8\theta_1}\right)^{1/2} n^{(1+\delta)/2} + o\big(n^{(1+\delta)/2}\big), \tag{4.11}$$

$$\begin{aligned}
\operatorname{tr}\big(\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big) &= \frac{\theta_{0,2}}{(2\theta_1\theta_2)^{1/2}}n^{(1+\delta)/2} + \frac{\theta_{0,2}(\theta_3^2-\theta_{0,3}^2)}{2\theta_2\theta_{0,3}}n^\delta \\
&\quad + o\big(n^\delta\big) + O(1),
\end{aligned} \tag{4.12}$$

$$\operatorname{tr}\left(\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\frac{\partial}{\partial\theta_3}\boldsymbol{\Sigma}(\boldsymbol{\theta})\right)^2\right)=\frac{1}{\theta_3}n^{\delta}+o\left(n^{\delta}\right). \tag{4.13}$$

Remark 1. As will be shown later, (4.2), (4.11) and (4.12) can be used to derive bounds for $\operatorname{tr}((\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\,\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})/\partial\theta_2)^2)$ and $\operatorname{tr}((\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\,\partial\boldsymbol{\Sigma}(\boldsymbol{\theta})/\partial\theta_1)^2)$. These bounds, together with (4.13), play important roles in establishing the consistency of $\hat{\theta}_1$.

Lemma 4.7. *Let* $\mathbf{X}$ *be full rank a.s. Then under the setup of Lemma 4.3,*

$$\sup_{\boldsymbol{\theta}\in\Theta}\left\{(\boldsymbol{\eta}+\boldsymbol{\epsilon})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})(\boldsymbol{\eta}+\boldsymbol{\epsilon})\right\}=O_p(p_n), \tag{4.14}$$

where $\mathbf{M}(\boldsymbol{\theta})$ *is defined in* (1.7).

Lemma 4.8. *Under the setup up of Lemma 4.3, let* $\mathbf{X}$ *be full rank a.s. Suppose that for some* $\xi\geq 0$,

$$\sup_{\boldsymbol{\theta}\in\Theta}\left\{\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\left(\mathbf{I}-\mathbf{M}(\boldsymbol{\theta})\right)\boldsymbol{\mu}_0\right\}=O_p\left(n^{\xi}\right).$$

Then

$$\sup_{\boldsymbol{\theta}\in\Theta}\left\{\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\left(\mathbf{I}-\mathbf{M}(\boldsymbol{\theta})\right)(\boldsymbol{\eta}+\boldsymbol{\epsilon})\right\}=o_p\left(n^{\xi}\right). \tag{4.15}$$

Before introducing Lemma 4.9, we need some notation. For $1\leq m\leq r<\infty$, define $\mathbf{J}(m,r)=\{(j_1,\ldots,j_m):j_1<\cdots<j_m,\,j_i\in\{1,\ldots,r\},1\leq i\leq m\}$. Let $g(\boldsymbol{\xi})$ be a function of $\boldsymbol{\xi}=(\xi_1,\ldots,\xi_r)'\in\mathbb{R}^r$. For $\mathbf{j}=(j_1,\ldots,j_m)\in\mathbf{J}(m,r)$, define $\mathbf{D_j}g(\boldsymbol{\xi})=\partial^m g(\boldsymbol{\xi})/\partial\xi_{j_1},\ldots,\partial\xi_{j_m}$. Denote by $B_\tau(\boldsymbol{\lambda})$ the r-dimensional closed ball centered at $\boldsymbol{\lambda}=(\lambda_1,\ldots,\lambda_r)'$ with radius $0<\tau<\infty$. For $\mathbf{j}\in\mathbf{J}(m,r)$, define the m-dimensional sphere:

$$\begin{aligned}B_\tau(\boldsymbol{\lambda},\mathbf{j})=\Big\{&(\xi_{j_1},\ldots,\xi_{j_m}):(\lambda_1,\ldots,\lambda_{j_1-1},\xi_{j_1},\lambda_{j_1+1},\ldots,\lambda_{j_2-1},\\&\xi_{j_2},\lambda_{j_2+1},\ldots,\lambda_{j_m-1},\xi_{j_m},\lambda_{j_m+1},\ldots,\lambda_r)\in B_\tau(\boldsymbol{\lambda})\Big\}.\end{aligned}$$

Lemma 4.9. *Assume that* $w_1,\ldots,w_n$ *are i.i.d. random variables with* $\mathrm{E}(w_1)=0,\mathrm{E}(w_1^2)=1$ *and* $\mathrm{E}(w_1^4)<\infty$. *Let* $\mathbf{A}(\boldsymbol{\xi})=[a_{i,j}(\boldsymbol{\xi})]_{1\leq i,j\leq n}$ *be an* $n\times n$ *matrix whose* (i,j)*th component is* $a_{i,j}(\boldsymbol{\xi})$, *a function of* $\boldsymbol{\xi}$ *with a continuous partial derivative* $\mathbf{D_j}a_{i,j}(\boldsymbol{\xi})$ *on* $B_\tau(\boldsymbol{\lambda})$, *for* $\mathbf{j}\in\mathbf{J}(m,r)$. *Define* $q_1(\boldsymbol{\xi})=\mathbf{w}'\mathbf{A}(\boldsymbol{\xi})\mathbf{w}-\operatorname{tr}(\mathbf{A}(\boldsymbol{\xi}))$, *where* $\mathbf{w}=(w_1,\ldots,w_n)'$. *Then for* $\boldsymbol{\xi}\in B_\tau(\boldsymbol{\lambda})$, *there exists a constant* $C>0$ *such that*

$$\begin{aligned}&\mathrm{E}\Big(\sup_{\boldsymbol{\xi}\in B_\tau(\boldsymbol{\lambda})}\left(q_1(\boldsymbol{\xi})-q_1(\boldsymbol{\lambda})\right)^2\Big)\\&\quad\leq C\sum_{m=1}^{r}\sum_{\mathbf{j}\in\mathbf{J}(m,r)}\operatorname{vol}^2\left(B_\tau(\boldsymbol{\lambda},\mathbf{j})\right)\sup_{\boldsymbol{\xi}\in B_\tau(\boldsymbol{\lambda})}\operatorname{var}\left(\mathbf{D_j}q_1(\boldsymbol{\xi})\right),\end{aligned} \tag{4.16}$$

where $\operatorname{vol}(\Theta)$ *denotes the volume of* Θ.

First, we prove (1.13). By (1.12), (4.9), (4.10), (4.14) and (4.15), it follows that

$$
\begin{aligned}
-2\ell(\boldsymbol{\theta}) = {} & n\log(2\pi) - \frac{1-\delta}{2}\log n + \left(\log\theta_1 + \frac{\theta_{0,1}}{\theta_1}\right)n \\
& + \left(\frac{2\theta_2}{\theta_1}\right)^{1/2}\left(1 - \frac{\theta_{0,1}}{2\theta_1} + \frac{\theta_{0,2}}{2\theta_2}\right)n^{(1+\delta)/2} \\
& - \left\{\frac{\theta_2}{\theta_1} + \theta_3 - \frac{\theta_{0,2}(\theta_3^2 - \theta_{0,3}^2)}{2\theta_2\theta_{0,3}}\right\}n^{\delta} \\
& + h(\boldsymbol{\theta}) + O_p\big(n^{\xi}\big) + o_p\big(n^{\delta}\big),
\end{aligned}
\tag{4.17}
$$

uniformly in Θ, where $h(\boldsymbol{\theta}) = (\boldsymbol{\eta}+\boldsymbol{\epsilon})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})(\boldsymbol{\eta}+\boldsymbol{\epsilon}) - \operatorname{tr}(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$. Hence, (1.13) is ensured by for any $\varepsilon > 0$,

$$
P\Big(\inf_{\boldsymbol{\theta}\in\Theta_1(\epsilon)}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell(\boldsymbol{\theta}_0)\big\} > 0\Big) \to 1, \tag{4.18}
$$

as $n\to\infty$, where $\Theta_1(\epsilon) = \{\boldsymbol{\theta}\in\Theta : |\theta_1 - \theta_{0,1}| > \varepsilon\}$. Since by (4.17),

$$
\begin{aligned}
\inf_{\boldsymbol{\theta}\in\Theta_1(\varepsilon)}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell(\boldsymbol{\theta}_0)\big\} \geq {} & \inf_{\boldsymbol{\theta}\in\Theta_1(\varepsilon)}\left\{\log\theta_1 + \frac{\theta_{0,1}}{\theta_1} - \log(\theta_{0,1}) - 1\right\}n \\
& - \sup_{\boldsymbol{\theta}\in\Theta_1(\varepsilon)}\big|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_0)\big| + o_p(n),
\end{aligned}
$$

and since $\inf_{\boldsymbol{\theta}\in\Theta_1(\varepsilon)}\{\log\theta_1 + \frac{\theta_{0,1}}{\theta_1} - \log(\theta_{0,1}) - 1\} > 0$, (4.18) follows immediately from

$$
\mathrm{E}\Big(\sup_{\boldsymbol{\theta}\in\Theta}\big|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_0)\big|^2\Big) = O(n). \tag{4.19}
$$

Since $h(\boldsymbol{\theta})$ is continuous on Θ and Θ is compact, in the rest of the proof, we assume without loss of generality that $\Theta = B_\tau(\boldsymbol{\theta}_0)$, a closed ball centered at $\boldsymbol{\theta}_0$ with radius τ for some $0 < \tau < \infty$. By (4.16) with $\mathbf{w} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\theta}_0)(\boldsymbol{\eta}+\boldsymbol{\epsilon})$ and $\mathbf{A}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)$, we obtain $h(\boldsymbol{\theta}) = \mathbf{w}'\mathbf{A}(\boldsymbol{\theta})\mathbf{w} - \operatorname{tr}(\mathbf{A}(\boldsymbol{\theta}))$ and

$$
\begin{aligned}
& \mathrm{E}\Big(\sup_{\boldsymbol{\theta}\in\Theta}\big|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_0)\big|^2\Big) \\
& \quad \leq C\sup_{\boldsymbol{\theta}\in\Theta}\bigg\{\operatorname{var}\Big(\frac{\partial}{\partial\theta_1}h(\boldsymbol{\theta})\Big) + \operatorname{var}\Big(\frac{\partial}{\partial\theta_2}h(\boldsymbol{\theta})\Big) + \operatorname{var}\Big(\frac{\partial}{\partial\theta_3}h(\boldsymbol{\theta})\Big) \\
& \qquad + \operatorname{var}\Big(\frac{\partial^2}{\partial\theta_1\,\partial\theta_2}h(\boldsymbol{\theta})\Big) + \operatorname{var}\Big(\frac{\partial^2}{\partial\theta_1\,\partial\theta_3}h(\boldsymbol{\theta})\Big) + \operatorname{var}\Big(\frac{\partial^2}{\partial\theta_2\,\partial\theta_3}h(\boldsymbol{\theta})\Big) \\
& \qquad + \operatorname{var}\Big(\frac{\partial^3}{\partial\theta_1\,\partial\theta_2\,\partial\theta_3}h(\boldsymbol{\theta})\Big)\bigg\},
\end{aligned}
\tag{4.20}
$$

for some constant $C > 0$. By (4.2), (4.12),

$$\operatorname{tr}(A)\lambda_{\min}(B) \le \operatorname{tr}(AB) \le \operatorname{tr}(A)\lambda_{\max}(B), \tag{4.21}$$

for the nonnegative definite matrices A and B, and using $\mathbf{I} - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\Sigma_\eta(\boldsymbol{\theta}_0) = \theta_0 \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)$ twice, we obtain

$$\begin{aligned}
\operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_0)\big) &= \frac{1}{\theta_{0,1}}\big\{\operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\big) - \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\big)\big\} \\
&= \frac{1}{\theta_{0,1}}\bigg\{\frac{1}{\theta_{0,1}}\big(n - \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\big)\big) \\
&\quad - \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_0)\big)\bigg\} \\
&= \frac{1}{\theta_{0,1}^2} n + O\big(n^{(1+\delta)/2}\big).
\end{aligned} \tag{4.22}$$

Equations (4.2), (4.21) and (4.22) lead to

$$\begin{aligned}
\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\bigg(\frac{\partial}{\partial\theta_1} h(\boldsymbol{\theta})\bigg) &= \sup_{\boldsymbol{\theta}\in\Theta} 2\operatorname{tr}\bigg(\bigg(\frac{\partial}{\partial\theta_1}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\bigg)^2\bigg) \\
&= \sup_{\boldsymbol{\theta}\in\Theta} 2\operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\big)^2\big) = O(n).
\end{aligned} \tag{4.23}$$

Similarly, (4.2), (4.11) and (4.21) imply

$$\begin{aligned}
\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\bigg(\frac{\partial}{\partial\theta_2} h(\boldsymbol{\theta})\bigg) &= \sup_{\boldsymbol{\theta}\in\Theta} 2\operatorname{tr}\bigg(\bigg(\frac{\partial}{\partial\theta_2}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\bigg)^2\bigg) \\
&= \sup_{\boldsymbol{\theta}\in\Theta} \frac{2}{\theta_2^2}\operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\big)^2\big) \\
&= O\big(n^{(1+\delta)/2}\big).
\end{aligned} \tag{4.24}$$

Moreover, by (4.2), (4.13) and (4.21), one gets

$$\begin{aligned}
\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\bigg(\frac{\partial}{\partial\theta_3} h(\boldsymbol{\theta})\bigg) &= \sup_{\boldsymbol{\theta}\in\Theta} 2\operatorname{tr}\bigg(\bigg(\frac{\partial}{\partial\theta_3}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\bigg)^2\bigg) \\
&= \sup_{\boldsymbol{\theta}\in\Theta} 2\operatorname{tr}\bigg(\bigg(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\bigg(\frac{\partial}{\partial\theta_3}\boldsymbol{\Sigma}(\boldsymbol{\theta})\bigg)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\bigg)^2\bigg) \\
&= O\big(n^{\delta}\big).
\end{aligned} \tag{4.25}$$

In a similar way, it can be shown that

$$\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_2} h(\boldsymbol{\theta})\right) = O\left(n^{(1+\delta)/2}\right), \tag{4.26}$$

and

$$\begin{aligned} &\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_3} h(\boldsymbol{\theta})\right) + \sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial^2}{\partial\theta_2\,\partial\theta_3} h(\boldsymbol{\theta})\right) \\ &\quad + \sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial^3}{\partial\theta_1\,\partial\theta_2\,\partial\theta_3} h(\boldsymbol{\theta})\right) = O\left(n^{\delta}\right). \end{aligned} \tag{4.27}$$

Consequently, (4.19) follows from (4.20)–(4.27), and hence (1.13) holds true.

Next, we prove (1.14), which in turn is implied by the property that for any $\varepsilon_2 > 0$, there exists an $\varepsilon_1 > 0$ such that

$$\mathrm{P}\left(\inf_{\boldsymbol{\theta}\in\Theta_2(\boldsymbol{\varepsilon})} \left\{-2\ell(\boldsymbol{\theta}) + 2\ell\left((\theta_1, \theta_{0,2}, \theta_{0,3})'\right)\right\} > 0\right) \to 1, \tag{4.28}$$

as $n \to \infty$, where $\Theta_2(\boldsymbol{\varepsilon}) = \{\boldsymbol{\theta} \in \Theta : |\theta_1 - \theta_{0,1}| \le \varepsilon_1, |\theta_2 - \theta_{0,2}| > \varepsilon_2\}$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)'$. Let $\boldsymbol{\theta}_b = (\theta_1, \theta_{0,2}, \theta_{0,3})'$. Since $\xi < (1+\delta)/2$, by (4.17), we have

$$\begin{aligned} &\inf_{\boldsymbol{\theta}\in\Theta_2(\boldsymbol{\varepsilon})} \left\{-2\ell(\boldsymbol{\theta}) + 2\ell(\boldsymbol{\theta}_b)\right\} \\ &\quad \ge \inf_{\boldsymbol{\theta}\in\Theta_2(\boldsymbol{\varepsilon})} \frac{1}{(2\theta_1\theta_2)^{1/2}} \left\{\left(\theta_2^{1/2} - \theta_{0,2}^{1/2}\right)^2 + \theta_2^{1/2}\left(\theta_2^{1/2} - \theta_{0,2}^{1/2}\right)\left(1 - \frac{\theta_{0,1}}{\theta_1}\right)\right\} n^{(1+\delta)/2} \\ &\qquad - \sup_{\boldsymbol{\theta}\in\Theta_2(\boldsymbol{\varepsilon})} \left|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_b)\right| + o_p\left(n^{(1+\delta)/2}\right). \end{aligned}$$

Therefore (4.28) is given by

$$\mathrm{E}\left(\sup_{\boldsymbol{\theta}\in\Theta} \left|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_b)\right|^2\right) = O_p\left(n^{(1+\delta)/2}\right). \tag{4.29}$$

By (4.16) with $\mathbf{w} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\theta}_0)(\boldsymbol{\eta} + \boldsymbol{\epsilon})$ and $\mathbf{A}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\}\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)$, we obtain $h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_b) = \mathbf{w}'\mathbf{A}(\boldsymbol{\theta})\mathbf{w} - \operatorname{tr}(\mathbf{A}(\boldsymbol{\theta}))$ and

$$\begin{aligned} &\mathrm{E}\left(\sup_{\boldsymbol{\theta}\in\Theta} \left|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_b)\right|^2\right) \\ &\quad \le C \sup_{\boldsymbol{\theta}\in\Theta} \left\{\operatorname{var}\left(\frac{\partial}{\partial\theta_1}\left(h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_b)\right)\right) + \operatorname{var}\left(\frac{\partial}{\partial\theta_2} h(\boldsymbol{\theta})\right) + \operatorname{var}\left(\frac{\partial}{\partial\theta_3} h(\boldsymbol{\theta})\right) \right. \\ &\qquad + \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_2} h(\boldsymbol{\theta})\right) + \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_3} h(\boldsymbol{\theta})\right) + \operatorname{var}\left(\frac{\partial^2}{\partial\theta_2\,\partial\theta_3} h(\boldsymbol{\theta})\right) \\ &\qquad \left. + \operatorname{var}\left(\frac{\partial^3}{\partial\theta_1\,\partial\theta_2\,\partial\theta_3} h(\boldsymbol{\theta})\right)\right\}, \end{aligned} \tag{4.30}$$

for some constant $C > 0$. In addition, it follows from (4.2), (4.11) and (4.21) that

$$
\begin{aligned}
&\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\Big(\frac{\partial}{\partial\theta_1}\big(h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_b)\big)\Big) \\
&= 2\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\Big(\Big(\frac{\partial}{\partial\theta_1}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\Big)^2\Big) \\
&= 2\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\big)^2\big) \\
&= 2\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0) \\
&\qquad \times \boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\big) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big) \\
&\qquad\quad \times \boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\big)\Big) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\big)\Big) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\big)\Big) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big)\big(\boldsymbol{\Sigma}^{2}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}^{2}(\boldsymbol{\theta})\big)\big)\Big) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big) + \big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big) \\
&\qquad \times \big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big) + \big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big)\Big) \qquad (4.31) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big)\boldsymbol{\Sigma}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big)\Big) \\
&\qquad + O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\Big) \\
&\qquad + O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\boldsymbol{\Sigma}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big)\Big) \\
&\qquad + O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\Big) \\
&= O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)^2\big)\Big) \\
&\qquad + O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+ O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big)\Big) \\
&+ O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta}_b)\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}(\boldsymbol{\theta})\big)^2\big)\Big) \\
=\;& O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\boldsymbol{\Sigma}^{-2}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_b) - \boldsymbol{\Sigma}_\eta(\boldsymbol{\theta})\big)^2\big)\Big) \\
=\;& O\big(n^{(1+\delta)/2}\big).
\end{aligned}
$$

Combining (4.30) and (4.31), with (4.24)–(4.27), yields (4.29), and hence (1.14) is established.

Finally, we prove (1.15). It suffices to show that for any $\varepsilon_3 > 0$, there exist $\varepsilon_1, \varepsilon_2 > 0$ such that

$$
\mathrm{P}\Big(\inf_{\boldsymbol{\theta}\in\Theta_3(\boldsymbol{\varepsilon})}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell\big((\theta_1, \theta_2, \theta_{0,3})'\big)\big\} > 0\Big) \to 1, \tag{4.32}
$$

as $n \to \infty$, where $\Theta_3(\boldsymbol{\varepsilon}) = \{\boldsymbol{\theta} \in \Theta : |\theta_1 - \theta_{0,1}| \le \varepsilon_1, |\theta_2 - \theta_{0,2}| \le \varepsilon_2, |\theta_3 - \theta_{0,3}| > \varepsilon_3\}$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)'$. Let $\boldsymbol{\theta}_c = (\theta_1, \theta_2, \theta_{0,3})'$. Since $\xi < \delta$, by (4.17), we have

$$
\begin{aligned}
\inf_{\boldsymbol{\theta}\in\Theta_3(\boldsymbol{\varepsilon})}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell(\boldsymbol{\theta}_c)\big\} \ge{}& \inf_{\boldsymbol{\theta}\in\Theta_3(\boldsymbol{\varepsilon})}\left\{\frac{\theta_{0,2}(\theta_3 - \theta_{0,3})^2}{2\theta_{0,3}\theta_2} - (\theta_3 - \theta_{0,3})\left(1 - \frac{\theta_{0,2}}{\theta_2}\right)\right\} n^\delta \\
&- \sup_{\boldsymbol{\theta}\in\Theta_3(\boldsymbol{\varepsilon})}\big|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big| + o_p\big(n^\delta\big).
\end{aligned}
$$

Therefore, it suffices for (4.32) to show that

$$
\mathrm{E}\Big(\sup_{\boldsymbol{\theta}\in\Theta}\big|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big|^2\Big) = O\big(n^\delta\big). \tag{4.33}
$$

By (4.16) with $\mathbf{w} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\theta}_0)(\boldsymbol{\eta} + \boldsymbol{\epsilon})$ and $\mathbf{A}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_c)\}\boldsymbol{\Sigma}^{1/2}(\boldsymbol{\theta}_0)$, we obtain $h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c) = \mathbf{w}'\mathbf{A}(\boldsymbol{\theta})\mathbf{w} - \operatorname{tr}(\mathbf{A}(\boldsymbol{\theta}))$ and

$$
\begin{aligned}
&\mathrm{E}\Big(\sup_{\boldsymbol{\theta}\in\Theta}\big|h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big|^2\Big) \\
&\quad\le C\sup_{\boldsymbol{\theta}\in\Theta}\left\{\operatorname{var}\left(\frac{\partial}{\partial\theta_1}\big(h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big)\right) + \operatorname{var}\left(\frac{\partial}{\partial\theta_2}\big(h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big)\right) + \operatorname{var}\left(\frac{\partial}{\partial\theta_3}h(\boldsymbol{\theta})\right)\right. \\
&\qquad + \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_2}\big(h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big)\right) + \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_3}h(\boldsymbol{\theta})\right) + \operatorname{var}\left(\frac{\partial^2}{\partial\theta_2\,\partial\theta_3}h(\boldsymbol{\theta})\right) \\
&\qquad \left. + \operatorname{var}\left(\frac{\partial^3}{\partial\theta_1\,\partial\theta_2\,\partial\theta_3}h(\boldsymbol{\theta})\right)\right\},
\end{aligned}
\tag{4.34}
$$

for some constant $C > 0$. In view of (4.34), (4.25) and (4.27), (4.33) is guaranteed by

$$
\sup_{\boldsymbol{\theta}\in\Theta}\operatorname{var}\left(\frac{\partial}{\partial\theta_1}\big(h(\boldsymbol{\theta}) - h(\boldsymbol{\theta}_c)\big)\right) = O\big(n^\delta\big), \tag{4.35}
$$

$$\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial}{\partial\theta_2}\big(h(\boldsymbol{\theta})-h(\boldsymbol{\theta}_c)\big)\right) = O\big(n^{\delta}\big), \tag{4.36}$$

$$\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial^2}{\partial\theta_1\,\partial\theta_2}\big(h(\boldsymbol{\theta})-h(\boldsymbol{\theta}_c)\big)\right) = O\big(n^{\delta}\big). \tag{4.37}$$

In what follows, we only focus on the proof of (4.35) since the proofs of (4.36) and (4.37) are similar. Note first that by an argument similar to that used to prove (4.31), one obtains

$$\begin{aligned}
&\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{var}\left(\frac{\partial}{\partial\theta_1}\big(h(\boldsymbol{\theta})-h(\boldsymbol{\theta}_c)\big)\right) \\
&\quad = 2\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\left(\left(\frac{\partial}{\partial\theta_1}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})-\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_c)\big)\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\right)^2\right) \\
&\quad = O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}(\boldsymbol{\theta}_c)-\boldsymbol{\Sigma}(\boldsymbol{\theta})\big)\big)^2\big)\Big) \\
&\quad = O\Big(\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\mathbf{T}_n^{-1}(\boldsymbol{\theta})\mathbf{G}_n(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_c)-\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta})\big)\mathbf{G}_n(\boldsymbol{\theta})'\big)^2\big)\Big).
\end{aligned} \tag{4.38}$$

In addition, (4.4) and some algebraic manipulations yield

$$\begin{aligned}
&\mathbf{G}_n(\boldsymbol{\theta})\big(\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_c)-\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta})\big)\mathbf{G}_n(\boldsymbol{\theta})' \\
&\quad = \left(\frac{\theta_2\rho_n}{\theta_{0,3}\rho_{0,n}}\big(1-\rho_{0,n}^2\big)-\frac{\theta_2}{\theta_3}\big(1-\rho_n^2\big)\right)\mathbf{I} + \left(1-\frac{\rho_n}{\rho_{0,n}}\right)(1-\rho_n\rho_{0,n})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_c) \\
&\qquad + \frac{\theta_2}{\theta_{0,3}}\left(1-\frac{\rho_n}{\rho_{0,n}}\right)\big(\mathbf{v}_0\mathbf{e}_1'+\mathbf{e}_1\mathbf{v}_0'\big) + \theta_2\left(\frac{1}{\theta_{0,3}}-\frac{1}{\theta_3}\right)\rho_n^2\mathbf{e}_1\mathbf{e}_1',
\end{aligned} \tag{4.39}$$

where $\rho_{0,n} = \exp(-\theta_{0,3}n^{-(1-\delta)})$, and

$$1-\rho_n^k\rho_{0,n}^\ell = (k\theta_3+\ell\theta_{0,3})n^{-(1-\delta)} + O\big(n^{-2(1-\delta)}\big); \qquad k,\ell\in\mathbb{Z}, \tag{4.40}$$

uniformly in Θ. Moreover, by (4.3), (4.5)–(4.8) and

$$\limsup_{n\to\infty}\sup_{\boldsymbol{\theta}\in\Theta} \lambda_{\max}\big(\boldsymbol{\Sigma}_\eta^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_c)\big) < \infty,$$

which can be shown using an argument similar to that used to prove (B.2) in the supplementary document (Chang, Huang and Ing [5]), we have

$$\begin{aligned}
\sup_{\boldsymbol{\theta}\in\Theta} n^{-4(1-\delta)} \operatorname{tr}\big(\mathbf{T}_n^{-2}(\boldsymbol{\theta})\big) &= O\big(n^{\delta}\big), \\
\sup_{\boldsymbol{\theta}\in\Theta} n^{-4(1-\delta)} \operatorname{tr}\big(\big(\mathbf{T}_n^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(\boldsymbol{\theta}_c)\big)^2\big) &= O\big(n^{\delta}\big), \\
\sup_{\boldsymbol{\theta}\in\Theta} n^{-2(1-\delta)} \operatorname{tr}\big(\big(\mathbf{T}_n^{-1}(\boldsymbol{\theta})\big(\mathbf{v}_0\mathbf{e}_1'+\mathbf{e}_1\mathbf{v}_0'\big)\big)^2\big) &= O(1), \\
\sup_{\boldsymbol{\theta}\in\Theta} \operatorname{tr}\big(\big(\mathbf{T}_n^{-1}(\boldsymbol{\theta})\mathbf{e}_1\mathbf{e}_1'\big)^2\big) &= O(1).
\end{aligned} \tag{4.41}$$

Combining (4.38)–(4.41) leads to (4.35) and hence (4.33). This completes the proof of (1.15).

4.2. Proof of Theorem 2.1

To prove Theorem 2.1, we need two additional lemmas, Lemmas 4.10–4.11, which provide the orders of magnitude of $\partial\ell(\hat{\boldsymbol{\theta}})/\partial\theta_i$ and $\partial^2\ell(\hat{\boldsymbol{\theta}})/\partial\theta_i^2$; $i = 1, 2, 3$, when the convergence rate of $\hat{\boldsymbol{\theta}}$ is given. On the contrary, using the orders of the magnitude of $\partial\ell(\hat{\boldsymbol{\theta}})/\partial\theta_i$ and $\partial^2\ell(\hat{\boldsymbol{\theta}})/\partial\theta_i^2$, $i = 1, 2, 3$, one can also derive the convergence rate of $\hat{\boldsymbol{\theta}}$; see (4.55)–(4.57). As a result, the convergence rate of $\hat{\boldsymbol{\theta}}$ can be sequentially improved via an initial convergence rate and applying this argument repeatedly.

Lemma 4.10. *Under the setup of Lemma 4.8, define for $k = 1, 2, 3$,*

$$g_k(\boldsymbol{\theta}) = -\frac{\partial}{\partial\theta_k} 2\ell(\boldsymbol{\theta}),$$

where $\ell(\boldsymbol{\theta})$ is given by (1.5). Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)'$ be an estimate of $\boldsymbol{\theta}$ with $\hat{\theta}_1 = \theta_{0,1} + O_p(n^{-r_1})$, $\hat{\theta}_2 = \theta_{0,2} + O_p(n^{-r_2})$ and $\hat{\theta}_3 = \theta_{0,3} + O_p(n^{-r_3})$ for some constants $r_1 \in [0, 1/2]$, $r_2 \in [0, (1+\delta)/4]$ and $r_3 \in [0, \delta/2]$; $\delta \in [0, 1)$. Then for any $\delta \in [0, 1)$,

$$\begin{aligned} g_1\big((\theta_{0,1}, \hat{\theta}_2, \hat{\theta}_3)'\big) &= O_p\big(n^{1/2}\big) + O_p\big(n^{(1+\delta)/2-r_2}\big) + O_p\big(n^{\delta-r_3}\big) \\ &\quad + O_p\big(n^{\xi}\big) + O(1), \end{aligned} \tag{4.42}$$

$$\begin{aligned} g_2\big((\hat{\theta}_1, \theta_{0,2}, \hat{\theta}_3)'\big) &= O_p\big(n^{(1+\delta)/4}\big) + O_p\big(n^{(1+\delta)/2-r_1}\big) + O_p\big(n^{\delta-r_3}\big) \\ &\quad + O_p\big(n^{\xi}\big) + O(1), \end{aligned} \tag{4.43}$$

and for $\delta \in (0, 1)$,

$$\begin{aligned} g_3\big((\hat{\theta}_1, \hat{\theta}_2, \theta_{0,3})'\big) &= O_p\big(n^{\delta/2}\big) + O_p\big(n^{\delta-r_1}\big) + O_p\big(n^{\delta-r_2}\big) \\ &\quad + O_p\big(n^{\xi}\big) + O(1). \end{aligned} \tag{4.44}$$

In addition, for any $\delta \in [0, 1)$, if $\xi < 1/2$ and $r_2 \geq \delta/2$,

$$n^{-1/2} g_1\big((\theta_{0,1}, \hat{\theta}_2, \hat{\theta}_3)'\big) \xrightarrow{d} N\big(0, 2\theta_{0,1}^{-2}\big); \tag{4.45}$$

if $\xi < (1+\delta)/4$, $r_1 > (1+\delta)/4$ and $r_3 > -(1-3\delta)/4$,

$$n^{-(1+\delta)/4} g_2\big((\hat{\theta}_1, \theta_{0,2}, \hat{\theta}_3)'\big) \xrightarrow{d} N\big(0, 2^{-1/2}\theta_{0,1}^{-1/2}\theta_{0,2}^{-3/2}\big). \tag{4.46}$$

Furthermore, for any $\delta \in (0, 1)$, if $\xi < \delta/2$, $r_1 > \delta/2$ and $r_2 > \delta/2$,

$$n^{-\delta/2} g_3\big((\hat{\theta}_1, \hat{\theta}_2, \theta_{0,3})'\big) \xrightarrow{d} N\big(0, 2\theta_{0,3}^{-1}\big). \tag{4.47}$$

Lemma 4.11. *Under the setup of Lemma 4.8, let*

$$g_{kk}(\boldsymbol{\theta}) = -\frac{\partial^2}{\partial\theta_k^2} 2\ell(\boldsymbol{\theta}); \qquad k = 1, 2, 3. \tag{4.48}$$

Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)'$ *be an estimate of* $\boldsymbol{\theta}$. *Suppose that* $\hat{\theta}_1 = \theta_{0,1} + o_p(1)$. *Then for* $\delta \in [0, 1)$, *there exists a constant* $\theta^*_{0,1} > 0$ *satisfying* $|\theta^*_{0,1} - \hat{\theta}_1| \le |\theta_{0,1} - \hat{\theta}_1|$ *such that*

$$g_{11}\big((\theta^*_{0,1}, \hat{\theta}_2, \hat{\theta}_3)'\big) = \frac{n}{\theta^2_{0,1}} + o_p(n). \tag{4.49}$$

In addition, suppose that $\hat{\theta}_1 = \theta_{0,1} + o_p(1)$ *and* $\hat{\theta}_2 = \theta_{0,2} + o_p(1)$, *then for* $\delta \in [0, 1)$, *there exists a constant* $\theta^*_{0,2} > 0$ *satisfying* $|\theta^*_{0,2} - \hat{\theta}_2| \le |\theta_{0,2} - \hat{\theta}_2|$ *such that*

$$g_{22}\big((\hat{\theta}_1, \theta^*_{0,2}, \hat{\theta}_3)'\big) = \frac{n^{(1+\delta)/2}}{2^{3/2}\theta^{1/2}_{0,1}\theta^{3/2}_{0,2}} + O_p\big(n^{\xi}\big) + o_p\big(n^{(1+\delta)/2}\big). \tag{4.50}$$

Furthermore, suppose that $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + o_p(1)$, *then for* $\delta \in (0, 1)$, *there exists a constant* $\theta^*_{0,3} > 0$ *satisfying* $|\theta^*_{0,3} - \hat{\theta}_3| \le |\theta_{0,3} - \hat{\theta}_3|$ *such that*

$$g_{33}\big((\hat{\theta}_1, \hat{\theta}_2, \theta^*_{0,3})'\big) = \frac{n^{\delta}}{\theta_{0,3}} + O_p\big(n^{\xi}\big) + o_p\big(n^{\delta}\big). \tag{4.51}$$

We shall prove (2.1)–(2.3) by iteratively applying (4.42)–(4.51). For the first iteration, we show that

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{-(1-\delta)/2}\big) \qquad \text{if } \delta \in [0, 1), \tag{4.52}$$

$$\begin{aligned} n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) &\xrightarrow{d} N\big(0, 2^{5/2}\theta^{1/2}_{0,1}\theta^{3/2}_{0,2}\big) \qquad \text{if } \delta \in [0, 1/3), \\ \hat{\theta}_2 - \theta_{0,2} &= O_p\big(n^{-(1-\delta)/2}\big) \qquad \text{if } \delta \in [1/3, 1), \end{aligned} \tag{4.53}$$

$$\begin{aligned} n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) &\xrightarrow{d} N(0, 2\theta_{0,3}) \qquad \text{if } \delta \in (0, 1/2), \\ \hat{\theta}_3 - \theta_{0,3} &= O_p\big(n^{-(1-\delta)/2}\big) \qquad \text{if } \delta \in [1/2, 1). \end{aligned} \tag{4.54}$$

Proof of (4.52). Taking the Taylor expansion of $g_1(\hat{\boldsymbol{\theta}})$ at $\hat{\boldsymbol{\theta}}_a = (\theta_{0,1}, \hat{\theta}_2, \hat{\theta}_3)'$ yields

$$0 = g_1(\hat{\boldsymbol{\theta}}) = g_1(\hat{\boldsymbol{\theta}}_a) + g_{11}\big(\hat{\boldsymbol{\theta}}^*_a\big)(\hat{\theta}_1 - \theta_{0,1}), \tag{4.55}$$

where $\hat{\boldsymbol{\theta}}^*_a = (\theta^*_{0,1}, \hat{\theta}_2, \hat{\theta}_3)'$ satisfies $|\theta^*_{0,1} - \hat{\theta}_1| \le |\theta_{0,1} - \hat{\theta}_1|$. Therefore, for (4.52) to hold, it suffices to show that

$$\begin{aligned} g_1(\hat{\boldsymbol{\theta}}_a) &= O_p\big(n^{(1+\delta)/2}\big), \\ g_{11}\big(\hat{\boldsymbol{\theta}}^*_a\big) &= \frac{n}{\theta^2_{0,1}} + o_p(n), \end{aligned}$$

where the first equation follows from (1.14) and (4.42) with $r_2 = 0$, and the second one is given by (1.13) and (4.49). □

Proof of (4.53). Let $\hat{\boldsymbol{\theta}}_b = (\hat{\theta}_1, \theta_{0,2}, \hat{\theta}_3)'$. Taking the Taylor expansion of $g_2(\hat{\boldsymbol{\theta}})$ at $\hat{\boldsymbol{\theta}}_b = (\hat{\theta}_1, \theta_{0,2}, \hat{\theta}_3)'$ yields

$$0 = g_2(\hat{\boldsymbol{\theta}}) = g_2(\hat{\boldsymbol{\theta}}_b) + g_{22}(\hat{\boldsymbol{\theta}}_b^*)(\hat{\theta}_2 - \theta_{0,2}), \tag{4.56}$$

where $\hat{\boldsymbol{\theta}}_b^* = (\hat{\theta}_1, \theta_{0,2}^*, \hat{\theta}_3)'$ satisfies $|\theta_{0,2}^* - \hat{\theta}_2| \le |\theta_{0,2} - \hat{\theta}_2|$. Therefore, for (4.53) to hold, it suffices to show that

$$\begin{aligned} n^{-(1+\delta)/4} g_2(\hat{\boldsymbol{\theta}}_b) \xrightarrow{d} N\big(0, 2^{-1/2}\theta_{0,1}^{-1/2}\theta_{0,2}^{-3/2}\big) \qquad &\text{if } \delta \in [0, 1/3), \\ g_2(\hat{\boldsymbol{\theta}}_b) = O_p\big(n^{\delta}\big) \qquad &\text{if } \delta \in [1/3, 1), \\ g_{22}(\hat{\boldsymbol{\theta}}_b^*) = \frac{n^{(1+\delta)/2}}{2^{3/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}} + o_p\big(n^{(1+\delta)/2}\big), & \end{aligned}$$

where the first two equations follow from (4.43) with $r_1 = (1-\delta)/2$, (4.46) and (4.52), and the last one is ensured by (1.14), (4.50) and (4.52). □

Proof of (4.54). Taking the Taylor expansion of $g_3(\hat{\boldsymbol{\theta}})$ at $\hat{\boldsymbol{\theta}}_c$ yields

$$0 = g_3(\hat{\boldsymbol{\theta}}) = g_3(\hat{\boldsymbol{\theta}}_c) + g_{33}(\hat{\boldsymbol{\theta}}_c^*)(\hat{\theta}_3 - \theta_{0,3}), \tag{4.57}$$

where $\hat{\boldsymbol{\theta}}_c^* = (\hat{\theta}_1, \hat{\theta}_2, \theta_{0,3}^*)'$ satisfies $|\theta_{0,3}^* - \hat{\theta}_3| \le |\theta_{0,3} - \hat{\theta}_3|$. Therefore, for (4.54) to hold, it suffices to show that

$$\begin{aligned} n^{-\delta/2} g_3(\hat{\boldsymbol{\theta}}_c) \xrightarrow{d} N\big(0, 2\theta_{0,3}^{-1}\big) \qquad &\text{if } \delta \in (0, 1/2), \\ g_3(\hat{\boldsymbol{\theta}}_c) = O_p\big(n^{-(1-3\delta)/2}\big) \qquad &\text{if } \delta \in [1/2, 1), \\ g_{33}(\hat{\boldsymbol{\theta}}_c^*) = \frac{n^{\delta}}{\theta_{0,3}} + o_p\big(n^{\delta}\big), & \end{aligned}$$

where the first two equations follow from (4.44) with $r_1 = r_2 = (1-\delta)/2$, (4.47), (4.52) and (4.53), and the last one is ensured by (4.51). Thus, (4.54) is established. □

For the second iteration, we show that

$$\begin{aligned} n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big) \qquad &\text{if } \delta \in [0, 1/2), \\ \hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{-(1-\delta)}\big) \qquad &\text{if } \delta \in [1/2, 1), \end{aligned} \tag{4.58}$$

$$\begin{aligned} n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big) \qquad &\text{if } \delta \in [0, 3/5), \\ \hat{\theta}_2 - \theta_{0,2} = O_p\big(n^{-(1-\delta)}\big) \qquad &\text{if } \delta \in [3/5, 1), \end{aligned} \tag{4.59}$$

$$
\begin{aligned}
n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) &\xrightarrow{d} N(0, 2\theta_{0,3}) \qquad \text{if } \delta \in (0, 2/3), \\
\hat{\theta}_3 - \theta_{0,3} &= O_p\big(n^{-(1-\delta)}\big) \qquad \text{if } \delta \in [2/3, 1).
\end{aligned} \tag{4.60}
$$

By (4.42) with $r_2 = r_3 = (1-\delta)/2$, (4.45) and (4.53), we have

$$
\begin{aligned}
n^{-1/2} g_1(\hat{\boldsymbol{\theta}}_a) &\xrightarrow{d} N\big(0, 2\theta_{0,1}^{-2}\big) \qquad \text{if } \delta \in [0, 1/2), \\
g_1(\hat{\boldsymbol{\theta}}_a) &= O_p\big(n^{\delta}\big) \qquad \text{if } \delta \in [1/2, 1).
\end{aligned}
$$

The above two equations, (4.49) and (4.55) give (4.58). By (4.43) with $r_1 = 1-\delta$ and $r_3 = (1-\delta)/2$, (4.46), (4.54) and (4.58), we have

$$
\begin{aligned}
n^{-(1+\delta)/4} g_2(\hat{\boldsymbol{\theta}}_b) &\xrightarrow{d} N\big(0, 2^{-1/2}\theta_{0,1}^{-1/2}\theta_{0,2}^{-3/2}\big) \qquad \text{if } \delta \in [0, 3/5), \\
g_2(\hat{\boldsymbol{\theta}}_b) &= O_p\big(n^{-(1-3\delta)/2}\big) \qquad \text{if } \delta \in [3/5, 1).
\end{aligned}
$$

Combining these two equations together with (4.50) and (4.56) yields (4.59). By (4.44) with $r_1 = r_2 = 1-\delta$, (4.47), (4.58) and (4.59), we have

$$
\begin{aligned}
n^{-\delta/2} g_3(\hat{\boldsymbol{\theta}}_c) &\xrightarrow{d} N\big(0, 2\theta_{0,3}^{-1}\big) \qquad \text{if } \delta \in (0, 2/3), \\
g_3(\hat{\boldsymbol{\theta}}_c) &= O_p\big(n^{-(1-3\delta)/2}\big) \qquad \text{if } \delta \in [2/3, 1),
\end{aligned}
$$

which, together with (4.51) and (4.57), lead immediately to (4.60).

Following the same argument as in the second iteration, we can recursively show that for each $i = 3, 4, \ldots$

$$
\begin{aligned}
n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) &\xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big) \qquad \text{if } \delta \in \big[0, (i-1)/i\big), \\
\hat{\theta}_1 - \theta_{0,1} &= O_p\big(n^{-i(1-\delta)/2}\big) \qquad \text{if } \delta \in \big[(i-1)/i, 1\big), \\
n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) &\xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big) \qquad \text{if } \delta \in \big[0, (2i-1)/(2i+1)\big), \\
\hat{\theta}_2 - \theta_{0,2} &= O_p\big(n^{-i(1-\delta)/2}\big) \qquad \text{if } \delta \in \big[(2i-1)/(2i+1), 1\big], \\
n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) &\xrightarrow{d} N(0, 2\theta_{0,3}) \qquad \text{if } \delta \in \big(0, i/(i+1)\big), \\
\hat{\theta}_3 - \theta_{0,3} &= O_p\big(n^{-i(1-\delta)/2}\big) \qquad \text{if } \delta \in \big[i/(i+1), 1\big).
\end{aligned}
$$

Thus (2.1)–(2.3) are proved.

4.3. Proof of Theorem 2.2

We divide the proof into three parts corresponding to $\delta \in [0, 1/3)$, $\delta \in [1/3, 1/2)$ and $\delta \in [1/2, 1)$.

First, we consider $\delta \in [0, 1/3)$. We further divide the proof into six subparts with respect to ξ in terms of a partition of $[0, 1)$, corresponding to $\xi \in [0, \delta/2)$, $\xi \in [\delta/2, \delta)$, $\xi \in [\delta, (1+\delta)/4)$, $\xi \in [(1+\delta)/4, 1/2)$, $\xi \in [1/2, (1+\delta)/2)$ and $\xi \in [(1+\delta)/2, 1)$. We shall prove each of the following six subparts separately:

(a1) For $\xi \in [(1+\delta)/2, 1)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\bigl(n^{\xi-1}\bigr).$$

(a2) For $\xi \in [1/2, (1+\delta)/2)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\bigl(n^{\xi-1}\bigr), \tag{4.61}$$

$$\hat{\theta}_2 - \theta_{0,2} = O_p\bigl(n^{\xi-(1+\delta)/2}\bigr). \tag{4.62}$$

(a3) For $\xi \in [(1+\delta)/4, 1/2)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\bigl(0, 2\theta_{0,1}^2\bigr), \tag{4.63}$$

$$\hat{\theta}_2 - \theta_{0,2} = O_p\bigl(n^{\xi-(1+\delta)/2}\bigr). \tag{4.64}$$

(a4) For $\xi \in [\delta, (1+\delta)/4)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\bigl(0, 2\theta_{0,1}^2\bigr), \tag{4.65}$$

$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\bigl(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\bigr). \tag{4.66}$$

(a5) For $\xi \in [\delta/2, \delta)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\bigl(0, 2\theta_{0,1}^2\bigr), \tag{4.67}$$

$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\bigl(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\bigr), \tag{4.68}$$

$$\hat{\theta}_3 - \theta_{0,3} = O_p\bigl(n^{\xi-\delta}\bigr). \tag{4.69}$$

(a6) For $\xi \in [0, \delta/2)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\bigl(0, 2\theta_{0,1}^2\bigr), \tag{4.70}$$

$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\bigl(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\bigr), \tag{4.71}$$

and if in addition $\delta \neq 0$, then

$$n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) \xrightarrow{d} N(0, 2\theta_{0,3}). \tag{4.72}$$

Proof of (a1). Applying (4.42) with $r_1 = r_2 = r_3 = 0$ and $\xi \in [(1+\delta)/2, 1)$, we have

$$g_1(\hat{\boldsymbol{\theta}}_a) = O_p\big(n^{\xi}\big). \tag{4.73}$$

According to (1.13) and (4.49), we have

$$g_{11}(\hat{\boldsymbol{\theta}}_a) = \frac{n}{\theta_{0,1}^2} + o_p(n). \tag{4.74}$$

The desired conclusion (a1) now follows from plugging (4.73) and (4.74) into (4.55). □

Proof of (a2). Applying (4.42) with $r_1 = r_2 = r_3 = 0$ and $\xi \in [1/2, (1+\delta)/2)$, we have

$$g_1(\hat{\boldsymbol{\theta}}_a) = O_p\big(n^{(1+\delta)/2}\big).$$

Combining this with (4.55) and (4.74) gives

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{-(1-\delta)/2}\big).$$

Applying (4.43) with $r_1 = (1-\delta)/2$, $r_2 = r_3 = 0$ and $\xi \in [1/2, (1+\delta)/2)$, we obtain

$$g_2(\hat{\boldsymbol{\theta}}_b) = O_p\big(n^{\xi}\big). \tag{4.75}$$

From (1.14) and (4.50), we have

$$g_{22}\big(\hat{\boldsymbol{\theta}}_b^*\big) = \frac{n^{(1+\delta)/2}}{2^{3/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}} + o_p\big(n^{(1+\delta)/2}\big). \tag{4.76}$$

Combining this with (4.56) and (4.75) leads to (4.62). In addition, applying (4.42) with $r_2 = (1+\delta)/2 - \xi$, $r_3 = 0$ and $\xi \in [1/2, (1+\delta)/2)$, we have

$$g_1(\hat{\boldsymbol{\theta}}_a) = O_p\big(n^{1/2}\big) + O_p\big(n^{\xi}\big) = O_p\big(n^{\xi}\big).$$

This together with (4.55) and (4.74) gives (4.61). □

Proof of (a3). Following the same arguments as the one used in the proof of (4.62) leads to (4.64). Applying (4.45) with $r_1 = (1-\delta)/2$, $r_2 = (1+\delta)/2 - \xi$, $r_3 = 0$ and $\xi \in [(1+\delta)/4, 1/2)$, we have

$$n^{-1/2} g_1(\hat{\boldsymbol{\theta}}_a) \xrightarrow{d} N\big(0, 2\theta_{0,1}^{-2}\big).$$

This together with (4.55) and (4.74) gives (4.63). □

Proof of (a4). Applying (4.46) with $r_1 = (1-\delta)/2$, $r_2 = r_3 = 0$ and $\xi \in [\delta, (1+\delta)/4)$, we have

$$n^{-(1+\delta)/4} g_2(\hat{\boldsymbol{\theta}}_b) \xrightarrow{d} N\big(0, 2^{-1/2}\theta_{0,1}^{-1/2}\theta_{0,2}^{-3/2}\big).$$

This, (4.56) and (4.76) imply (4.66). Moreover, (4.65) can be shown by an argument similar to that used to prove (4.63). □

Proof of (a5). The proofs of (4.67) and (4.68) are similar to those of (4.65) and (4.66), respectively. Applying (4.44) with $r_1 = r_2 = (1-\delta)/2$, $r_3 = 0$ and $\xi \in [(1+\delta)/4, 1/2)$, we have

$$g_3(\hat{\boldsymbol{\theta}}_c) = O_p\big(n^{\xi}\big). \tag{4.77}$$

From (1.13)–(1.15) and (4.51), we obtain

$$g_{33}\big(\hat{\boldsymbol{\theta}}_c^*\big) = \frac{n^{\delta}}{\theta_{0,3}} + o\big(n^{\delta}\big). \tag{4.78}$$

Combining this with (4.57) and (4.77) leads to (4.69). □

Proof of (a6). Equations (4.70) and (4.71) can be proved in a way similar to the proofs of (4.65) and (4.66). Applying (4.47) with $r_1 = r_2 = (1-\delta)/2$, $r_3 = 0$ and $\xi \in (0, \delta/2)$, we have

$$n^{-\delta/2} g_3(\hat{\boldsymbol{\theta}}_c) \xrightarrow{d} N\big(0, 2\theta_{0,3}^{-1}\big).$$

This together with (4.57) and (4.78) gives (4.72). □

Second, we consider $\delta \in [1/3, 1/2)$. Following an argument similar to that used in the first part, we obtain

(b1) For $\xi \in [(1+\delta)/2, 1)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{\xi-1}\big).$$

(b2) For $\xi \in [1/2, (1+\delta)/2)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{\xi-1}\big),$$
$$\hat{\theta}_2 - \theta_{0,2} = O_p\big(n^{\xi-(1+\delta)/2}\big).$$

(b3) For $\xi \in [\delta, 1/2)$,

$$\begin{aligned} n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) &\xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big), \\ \hat{\theta}_2 - \theta_{0,2} &= O_p\big(n^{\xi-(1+\delta)/2}\big). \end{aligned}$$

(b4) For $\xi \in [(1+\delta)/4, \delta)$,

$$\begin{aligned} n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) &\xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big), \\ \hat{\theta}_2 - \theta_{0,2} &= O_p\big(n^{\xi-(1+\delta)/2}\big), \\ \hat{\theta}_3 - \theta_{0,3} &= O_p\big(n^{\xi-\delta}\big). \end{aligned}$$

(b5) For $\xi \in [\delta/2, (1+\delta)/4)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big),$$
$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big),$$
$$\hat{\theta}_3 - \theta_{0,3} = O_p\big(n^{\xi-\delta}\big).$$

(b6) For $\xi \in [0, \delta/2)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big),$$
$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big),$$
$$n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) \xrightarrow{d} N(0, 2\theta_{0,3}).$$

Third, for $\delta \in [1/2, 1)$, one can similarly show that

(c1) For $\xi \in [(1+\delta)/2, 1)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{\xi-1}\big).$$

(c2) For $\xi \in [\delta, (1+\delta)/2)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{\xi-1}\big),$$
$$\hat{\theta}_2 - \theta_{0,2} = O_p\big(n^{\xi-(1+\delta)/2}\big).$$

(c3) For $\xi \in [1/2, \delta)$,

$$\hat{\theta}_1 - \theta_{0,1} = O_p\big(n^{\xi-1}\big),$$
$$\hat{\theta}_2 - \theta_{0,2} = O_p\big(n^{\xi-(1+\delta)/2}\big),$$
$$\hat{\theta}_3 - \theta_{0,3} = O_p\big(n^{\xi-\delta}\big).$$

(c4) For $\xi \in [(1+\delta)/4, 1/2)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big),$$
$$\hat{\theta}_2 - \theta_{0,2} = O_p\big(n^{\xi-(1+\delta)/2}\big),$$
$$\hat{\theta}_3 - \theta_{0,3} = O_p\big(n^{\delta/2}\big).$$

(c5) For $\xi \in [\delta/2, (1+\delta)/4)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big),$$
$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big),$$
$$\hat{\theta}_3 - \theta_{0,3} = O_p\big(n^{\xi-\delta}\big).$$

(c6) For $\xi \in [0, \delta/2)$,

$$n^{1/2}(\hat{\theta}_1 - \theta_{0,1}) \xrightarrow{d} N\big(0, 2\theta_{0,1}^2\big),$$

$$n^{(1+\delta)/4}(\hat{\theta}_2 - \theta_{0,2}) \xrightarrow{d} N\big(0, 2^{5/2}\theta_{0,1}^{1/2}\theta_{0,2}^{3/2}\big),$$

$$n^{\delta/2}(\hat{\theta}_3 - \theta_{0,3}) \xrightarrow{d} N(0, 2\theta_{0,3}).$$

Thus the proof of the theorem is complete.

4.4. Proofs of Corollaries 2.1 and 3.1

To prove Corollaries 2.1 and 3.1, the following lemma, which provides the order of magnitude of $R(\Theta)$ defined in (1.11), is needed.

Lemma 4.12. *Under the setup of Lemma 4.3, let* $\mathbf{x} = n^{-1}(1, 2, \dots, n)'$ *and* $\mathbf{1} = (1, \dots, 1)'$. *Then for any* $\delta \in [0, 1)$, *the following equations hold uniformly in* Θ:

$$\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1} = \frac{\theta_3^2}{2\theta_2}n^\delta + o\big(n^\delta\big) + O(1), \tag{4.79}$$

$$\mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1} = \frac{\theta_3^2}{4\theta_2}n^\delta + o\big(n^\delta\big) + O(1), \tag{4.80}$$

$$\mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{x} = \frac{\theta_3^2}{6\theta_2}n^\delta + o\big(n^\delta\big) + O(1). \tag{4.81}$$

We first prove Corollary 2.1. Note that

$$\begin{aligned}
&\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\mathbf{I} - \mathbf{M}(\boldsymbol{\theta})\big)\boldsymbol{\mu}_0 \\
&\quad = \boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\mu}_0 - \boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})\boldsymbol{\mu}_0 \\
&\quad = \beta_{0,1}^2\mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{x} - \beta_{0,1}^2\mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1}\big(\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1}\big)^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{x} \\
&\quad = \frac{\beta_{0,1}^2\theta_3^2}{24\theta_2}n^\delta + o\big(n^\delta\big),
\end{aligned} \tag{4.82}$$

uniformly in Θ, where $\mathbf{x} = n^{-1}(1, \dots, n)'$ and the last equality is obtained from (4.79)–(4.81). Therefore, (2.8) holds. With the help of (2.8), (2.9) and (2.10) follow directly from Theorem 2.2.

Second, we prove Corollary 3.1. By (1.8), (4.15) and (4.82), we have

$$\begin{aligned}
-2\ell(\boldsymbol{\theta}) &= n\log(2\pi) + \log\det\big(\boldsymbol{\Sigma}(\boldsymbol{\theta})\big) + \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)\big) \\
&\quad + \frac{\beta_{0,1}^2\theta_3^2}{24\theta_2}n^\delta + h(\boldsymbol{\theta}) + o_p\big(n^\delta\big) + O_p(1),
\end{aligned}$$

uniformly in Θ, noting that $h(\boldsymbol{\theta}) = (\boldsymbol{\eta} + \boldsymbol{\epsilon})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})(\boldsymbol{\eta} + \boldsymbol{\epsilon}) - \mathrm{tr}(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$. Therefore, by (4.9) and (4.10),

$$\begin{aligned}
-2\ell(\boldsymbol{\theta}) &= n\log(2\pi) - \frac{1-\delta}{2}\log n + \left(\log\theta_1 + \frac{\theta_{0,1}}{\theta_1}\right)n \\
&\quad + \left(\frac{2\theta_2}{\theta_1}\right)^{1/2}\left(1 - \frac{\theta_{0,1}}{2\theta_1} + \frac{\theta_{0,2}}{2\theta_2}\right)n^{(1+\delta)/2} \\
&\quad - \left(\frac{\theta_2}{\theta_1} + \theta_3 - \frac{\theta_{0,2}(\theta_3^2 - \theta_{0,3}^2)}{2\theta_2\theta_{0,3}} - \frac{\beta_{0,1}^2\theta_3^2}{24\theta_2}\right)n^{\delta} \\
&\quad + h(\boldsymbol{\theta}) + o_p\big(n^{\delta}\big) + O_p(1) \\
&= n\log(2\pi) - \frac{1-\delta}{2}\log n + \left(\log\theta_1 + \frac{\theta_{0,1}}{\theta_1}\right)n \\
&\quad + \left(\frac{2\theta_2}{\theta_1}\right)^{1/2}\left(1 - \frac{\theta_{0,1}}{2\theta_1} + \frac{\theta_{0,2}}{2\theta_2}\right)n^{(1+\delta)/2} \\
&\quad - \left\{\frac{\theta_2}{\theta_1} + \theta_3\left(1 - \frac{\theta_{0,2}}{\theta_2}\right) + \frac{\theta_{0,2}\theta_{0,3} + \theta_{0,2}\theta_{0,3}^*}{2\theta_2}\right. \\
&\quad \left. - \frac{\theta_{0,2}}{2\theta_2\theta_{0,3}^*}\big(\theta_3 - \theta_{0,3}^*\big)^2\right\}n^{\delta} + h(\boldsymbol{\theta}) + o_p\big(n^{\delta}\big) + O_p(1),
\end{aligned} \tag{4.83}$$

uniformly in Θ, where $\theta_{0,3}^* = \frac{12\theta_{0,2}}{12\theta_{0,2}+\beta_{0,1}^2\theta_{0,3}}\theta_{0,3}$. It follows from (4.83) and the same argument as in the proof of (4.32) that for any $\varepsilon_3 > 0$, there exist $\varepsilon_1, \varepsilon_2 > 0$ such that

$$P\Big(\inf_{\boldsymbol{\theta}\in\Theta_2(\varepsilon_1,\varepsilon_2,\varepsilon_3)}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell\big((\theta_1, \theta_2, \theta_{0,3}^*)'\big)\big\} > 0\Big) \to 1,$$

as $n \to \infty$, where $\Theta_3(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \{\boldsymbol{\theta} \in \Theta : |\theta_1 - \theta_{0,1}| \le \varepsilon_1, |\theta_2 - \theta_{0,2}| \le \varepsilon_2, |\theta_3 - \theta_{0,3}^*| > \varepsilon_3\}$. Thus (3.1) is established, and hence the proof is complete.

4.5. Proofs of Corollaries 2.2 and 3.2

We first prove (2.11). Let $\mathbf{x} = (x(s_1), \dots, x(s_n))'$. By an argument similar to that used to prove (4.15), it can be shown that

$$\sup_{\boldsymbol{\theta}\in\Theta} n^{-\delta/2}\mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1} = O_p(1).$$

This, together with (4.81) and (4.12), gives

$$\begin{aligned}
&\boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\big(\mathbf{I} - \mathbf{M}(\boldsymbol{\theta})\big)\boldsymbol{\mu}_0 \\
&\quad = \boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\mu}_0 - \boldsymbol{\mu}_0'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{M}(\boldsymbol{\theta})\boldsymbol{\mu}_0
\end{aligned}$$

$$
\begin{aligned}
&= \beta_{0,1}^2 \mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{x} - \beta_{0,1}^2 \mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1}\big(\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{1}\big)^{-1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{x} \\
&= \beta_{0,1}^2 \operatorname{tr}\big(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(0,\theta_{1,2},\theta_{1,3})'\big) + h_x(\boldsymbol{\theta}) + O_p(1) \\
&= \frac{\beta_{0,1}^2\theta_{1,2}}{(2\theta_1\theta_2)^{1/2}} n^{(1+\delta)/2} + \frac{\beta_{0,1}^2\theta_{1,2}(\theta_3^2-\theta_{1,3}^2)}{2\theta_2\theta_{1,3}} n^\delta \\
&\quad + h_x(\boldsymbol{\theta}) + o\big(n^\delta\big) + O_p(1),
\end{aligned}
$$

uniformly in Θ, where $h_x(\boldsymbol{\theta}) = \beta_{0,1}^2(\mathbf{x}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{x} - \operatorname{tr}(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\eta(0,\theta_{1,2},\theta_{1,3})'))$. In addition, an argument similar to that used to prove (4.29) yields

$$
\sup_{\boldsymbol{\theta}\in\Theta} h_x(\boldsymbol{\theta}) = o_p\big(n^{(1+\delta)/2}\big).
$$

Hence (2.11) follows. In view of (2.11) and Theorem 2.2, we obtain (2.12). Thus, the proof of Corollary 2.2 is complete.

To prove (3.2), note first that by the same line of reasoning as in (4.83), one gets

$$
\begin{aligned}
-2\ell(\boldsymbol{\theta}) &= n\log(2\pi) - \frac{1-\delta}{2}\log n + \left(\log\theta_1 + \frac{\theta_{0,1}}{\theta_1}\right) n \\
&\quad + \left(\frac{2\theta_2}{\theta_1}\right)^{1/2}\left(1 - \frac{\theta_{0,1}}{2\theta_1} + \frac{\theta_{0,2}^*}{2\theta_2}\right) n^{(1+\delta)/2} \\
&\quad - \left\{\frac{\theta_2}{\theta_1} + \theta_3\left(1 - \frac{\theta_{0,2}^*}{\theta_2}\right) + \frac{\theta_{0,2}\theta_{0,3} + \beta_{0,1}^2\theta_{1,2}\theta_{1,3} + \theta_{0,2}^*\theta_{0,3}^*}{2\theta_2}\right. \\
&\quad \left. - \frac{\theta_{0,2}^*}{2\theta_2\theta_{0,3}^*}\big(\theta_3 - \theta_{0,3}^*\big)^2\right\} n^\delta \\
&\quad + h_x(\boldsymbol{\theta}) + h(\boldsymbol{\theta}) + o_p\big(n^\delta\big) + O_p(1),
\end{aligned}
\tag{4.84}
$$

uniformly in Θ, where $\theta_{0,2}^* = \theta_{0,2} + \beta_{0,1}^2\theta_{1,2}$ and $\theta_{0,3}^* = \frac{\theta_{0,2}+\beta_{0,1}^2\theta_{1,2}}{\beta_{0,1}^2\theta_{1,2}\theta_{1,3}^{-1}+\theta_{0,3}\theta_{0,3}^{-1}}$. Moreover, using arguments similar to those used in the proofs of (4.28) and (4.32), respectively, one can show that for any $\varepsilon_2 > 0$, there exists an $\varepsilon_1 > 0$ such that

$$
\lim_{n\to\infty} P\Big(\inf_{\boldsymbol{\theta}\in\Theta_2(\varepsilon_1,\varepsilon_2)}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell\big((\theta_1,\theta_{0,2}^*,\theta_{0,3}^*)'\big)\big\} > 0\Big) = 1,
\tag{4.85}
$$

and for any $\varepsilon_3 > 0$, there exist $\varepsilon_1, \varepsilon_2 > 0$ such that

$$
\lim_{n\to\infty} P\Big(\inf_{\boldsymbol{\theta}\in\Theta_3(\varepsilon_1,\varepsilon_2,\varepsilon_3)}\big\{-2\ell(\boldsymbol{\theta}) + 2\ell\big((\theta_1,\theta_2,\theta_{0,3}^*)'\big)\big\} > 0\Big) = 1,
\tag{4.86}
$$

where $\Theta_2(\varepsilon_1,\varepsilon_2) = \{\boldsymbol{\theta}\in\Theta : |\theta_1-\theta_{0,1}| \le \varepsilon_1, |\theta_2-\theta_{0,2}^*| > \varepsilon_2\}$ and $\Theta_3(\varepsilon_1,\varepsilon_2,\varepsilon_3) = \{\boldsymbol{\theta}\in\Theta : |\theta_1-\theta_{0,1}| \le \varepsilon_1, |\theta_2-\theta_{0,2}^*| \le \varepsilon_2, |\theta_3-\theta_{0,3}^*| > \varepsilon_3\}$. Combining (4.84)–(4.86) yields (3.2) and (3.3). This completes the proof of Corollary 3.2.

Acknowledgements

The authors would like to thank the Associate Editor and an anonymous reviewer for their insightful and constructive comments, which greatly improve the presentation of this paper. The research of Chih-Hao Chang and Hsin-Cheng Huang was supported by Ministry of Science and Technology of Taiwan under grants MOST 103-2118-M-390-005-MY2 and MOST 100-2628-M-001-004-MY3, respectively. The research of Ching-Kang Ing was supported by Academia Sinica Investigator Award.

Supplementary Material

Supplement to "Mixed domain asymptotics for a stochastic process model with time trend and measurement error" (DOI: 10.3150/15-BEJ740SUPP; .pdf). The supplementary material contains the proofs of lemmas in Section 4, following some technical lemmas needed in the proofs.

References

[1] Aït-Sahalia, Y., Mykland, P. and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Rev. Financ. Stud.* **18** 351–416.

[2] Chan, N.H., Huang, S.-F. and Ing, C.-K. (2013). Moment bounds and mean squared prediction errors of long-memory time series. *Ann. Statist.* **41** 1268–1298. MR3113811

[3] Chan, N.H. and Ing, C.-K. (2011). Uniform moment bounds of Fisher's information with applications to time series. *Ann. Statist.* **39** 1526–1550. MR2850211

[4] Chang, C.-H., Huang, H.-C. and Ing, C.-K. (2014). Asymptotic theory of generalized information criterion for geostatistical regression model selection. *Ann. Statist.* **42** 2441–2468. MR3269985

[5] Chang, C.-H., Huang, H.-C. and Ing, C.-K. (2015). Supplement to "Mixed domain asymptotics for a stochastic process model with time trend and measurement error." DOI:10.3150/15-BEJ740SUPP.

[6] Chen, H.-S., Simpson, D.G. and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statist. Sinica* **10** 141–156. MR1742105

[7] Cressie, N.A.C. (1993). *Statistics for Spatial Data*. New York: Wiley. MR1239641

[8] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766. MR1026311

[9] Ibragimov, I.A. and Rozanov, Y.A. (1978). *Gaussian Random Processes*. *Applications of Mathematics* **9**. New York: Springer. MR0543837

[10] Lai, T.L. (1994). Asymptotic properties of nonlinear least squares estimates in stochastic regression models. *Ann. Statist.* **22** 1917–1930. MR1329175

[11] Pollard, D. (1984). *Convergence of Stochastic Processes*. New York: Springer. MR0762984

[12] Ripley, B.D. (1981). *Spatial Statistics*. New York: Wiley. MR0624436

[13] Sacks, J., Schiller, S.B. and Welch, W.J. (1989). Designs for computer experiments. *Technometrics* **31** 41–47. MR0997669

[14] Stein, M.L. (1990). A comparison of generalized cross validation and modified maximum likelihood for estimating the parameters of a stochastic process. *Ann. Statist.* **18** 1139–1157. MR1062702

[15] Uhlenbeck, G.E. and Ornstein, L.S. (1930). On the theory of Brownian motion. *Phys. Rev.* **36** 823–841.

[16] Xiu, D. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *J. Econometrics* **159** 235–250. MR2720855
[17] Ying, Z. (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *J. Multivariate Anal.* **36** 280–296. MR1096671
[18] Zhang, H. and Zimmerman, D.L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika* **92** 921–936. MR2234195

Received May 2014 and revised May 2015