Likelihood Estimation of Stable Levy Processes from Discrete Data

Masuda, Hiroki
Faculty of Mathematics, Kyushu University

出版情報：MHF Preprint Series. 2006-18, 2006-04-10. 九州大学大学院数理学研究院
バージョン：
権利関係：

MHF Preprint Series

Kyushu University
21st Century COE Program
Development of Dynamic Mathematics with High Functionality

Likelihood estimation of stable Lévy processes from discrete data

H. Masuda

MHF 2006-18

(Received April 10, 2006)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

Likelihood Estimation of Stable Lévy Processes from Discrete Data

Hiroki Masuda

Graduate School of Mathematics, Kyushu University,
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
Email: hiroki@math.kyushu-u.ac.jp

Abstract

We study the likelihood inference for real-valued non-Gaussian stable Lévy processes $X = (X_t)_{t\in\mathbb{R}_+}$ based on sampled data $(X_{ih_n})_{i=0}^n$, where $h_n \downarrow 0$, focusing on cases of either symmetric or completely skewed (one-sided) Lévy density. First, the local asymptotic normality with always degenerate Fisher information matrix is obtained, so that the maximum likelihood estimation is inappropriate for joint estimation of all parameters involved. Second, supposing that either index or scale parameter is known, we obtain the uniform asymptotic normality of the maximum likelihood estimates and their asymptotic efficiency, where the resulting optimal convergence rates reveal that, as opposed to the Gaussian case, that $nh_n \to \infty$ is not necessary for consistent estimation for all parameters.

Keywords: discrete sampling; efficiency; maximum likelihood estimation; stable Lévy process.

1 Introduction

The purpose of this study is to investigate likelihood-based parametric estimation for discretely observed non-Gaussian stable Lévy processes whose Lévy measures are either symmetric or completely skewed (one-sided). Our main results are presented in Section 3: first, we prove the local asymptotic normality (LAN) with an always degenerate Fisher information matrix; second, the uniform asymptotic normalities and the asymptotic efficiencies of the maximum likelihood estimators (MLE) are obtained when either index or scale parameter is supposed to be known. Recall that uniform asymptotic normality is theoretically important for constructing asymptotic confidence intervals. The resulting phenomena turn out to differ considerably from the Wiener case (see below). The proofs of the results are given in Section 4.

Our precise model setup will be described in Section 2, however, prior to it let us recall some well known facts and then give some remarks on our problem.

Consider the real-valued process $X = (X_t)_{t\in\mathbb{R}_+}$ given by

$$X_t = \gamma t + \sigma w_t, \tag{1}$$

where $\gamma \in \mathbb{R}$, $\sigma > 0$, and w is a standard Wiener process. If we can get a continuous record $(X_t)_{0\le t\le T}$, then on account of the quadratic-variation character we can identify σ without error for each $T > 0$: any two statistical experiments corresponding to different σ are mutually singular (entirely separated). So, in this case we may suppose $\sigma > 0$ is known and only estimating γ makes sense. The MLE of γ possesses the asymptotic normality and efficiency with optimal rate $\sqrt{T}$ and asymptotic variance σ^2. On the other hand, if we have only discrete observation $(X_{ih_n})_{i=1}^n$ for some $h_n > 0$ possibly depending on the sample size n, then estimation of σ makes sense too and actually we have the following.

(i) If $\lim_{n\to\infty} nh_n \in [0,\infty)$, then the MLE of σ is asymptotically normal and efficient with optimal rate $\sqrt{n}$ and asymptotic variance $\sigma^2/2$, regardless of the behavior of $h_n \downarrow 0$, while there exists no consistent estimate of γ; more precisely, the sequence of observed information associated with γ is stochastically bounded without norming.

(ii) If $nh_n \to \infty$ and $h_n \to 0$ as $n \to \infty$, then the MLE of (γ, σ) is asymptotically normal and efficient with optimal rate $\mathrm{diag}(\sqrt{nh_n}, \sqrt{n})$ and asymptotic covariance matrix $\mathrm{diag}(\sigma^2, \sigma^2/2)$.

Importantly, much more general diffusions give rise to analogous phenomena; see Gobet [3, 4] for efficiency issues, and Yoshida [14] and Kessler [9] for optimal estimating procedures.

Now consider the question "what about cases where X is instead a non-Gaussian stable Lévy process?", to which we shall give a reply in this article. We know that continuously observed cases makes no sense for all parameters involved: indeed, applying Kabanov et al. [8, Theorem 15] we can easily check that, for every $T > 0$, P_θ^T and $P_{\theta'}^T$ are absolutely continuous if and only if $\theta = \theta'$, where P_θ^T denotes the image measure of $(X_t)_{0\le t\le T}$ associated with θ whose components consist of the index, scale, location. Historically, (possibly skewed) non-Gaussian stable distributions have been rather typical for independent and identically distributed (iid) data whose distribution seem to possess regularly varying tails, and have been deeply investigated by many statisticians as well as probabilists: standard references for basic facts of stable distributions can be found in the nice bibliography of Nolan [10]. Despite the lack of explicit expressions of the densities except for a few particular cases, one can build on numerical procedures when attempting parametric inference: especially, for the implementation of maximum likelihood estimation based on iid samples, one should consult the above Nolan's paper.

When X is sampled at equally spaced time points, the continuous-time background reduces to the usual iid-sample case since a Lévy process has independent and stationary increments. However, it may be often useful and reasonable when we try to accommodate "asymptotic high frequency" of data such as intraday minute-by-minute one into the model, in which cases we may obtain precise asymptotic results by considering sampling length decreasing as sample size increase, such as the case where we observe $(X_{i/n})_{i=1}^n$: actually, this is the scope of our present study. When a distribution of observed data seems to possess the Paretian tail, stable-Lévy processes then may serve as a building block of modelling them.

We end this section with mentioning a few existing results. Woerner [13] proved the LAN property with rate $\sqrt{n}$ for the scale parameter of symmetric stable Lévy processes X based on discrete data $(X_{ih_n})_{i=1}^n$, where either $h_n = h > 0$ or $h_n \to 0$. Aït-Sahalia and Jacod [1] studied asymptotic behaviors of the Fisher information of Lévy processes with symmetric-stable convolution factor sampled at i/n, $i \le n$, and then Aït-Sahalia and Jacod [2] exhibited an explicit construction of an explicit rate-efficient estimator for the scale parameter.

2 Objective

For a random variable ξ we denote its law by $\mathcal{L}(\xi)$. Let $\gamma \in \mathbb{R}$ and $\sigma > 0$ be constants. We shall deal with the following two cases.

Case A. (*Stable Lévy process with drift and symmetric Lévy density*)
For $\alpha \in (0,2)$, let $SS_\alpha(\sigma)$ denote the α-stable distribution on the real line with the characteristic function

$$u \mapsto \exp(-\sigma^\alpha |u|^\alpha), \quad u \in \mathbb{R}.$$

Then let $X = (X_t)_{t\in\mathbb{R}_+}$ with $X_0 = 0$ a.s. be a Lévy process such that $\mathcal{L}(X_1 - \gamma) = SS_\alpha(\sigma)$.

Case B. (*Stable subordinator with drift*)

For $\alpha \in (0,1)$, let $S_\alpha^+(\sigma)$ denote the α-stable distribution on the positive half line with the characteristic function

$$u \mapsto \exp\left\{ -\sigma^\alpha |u|^\alpha \left(1 - i \tan\frac{\pi\alpha}{2} \cdot \mathrm{sign} u \right) \right\}, \quad u \in \mathbb{R}.$$

Then let $X = (X_t)_{t\in\mathbb{R}_+}$ with $X_0 = 0$ a.s. be a Lévy process such that $\mathcal{L}(X_1 - \gamma) = S_\alpha^+(\sigma)$.

In other words, we shall not consider skewed cases with $\beta \in (-1,0) \cup (0,1)$, where β stands for the skewness parameter, see Nolan [10, Section 1.1]; this is equivalent to excluding cases where $\mathcal{L}(X_t - \gamma t)$ admits an everywhere positive and skewed density for each $t > 0$. By taking the negative of X in Case B, stable Lévy processes with only negative jumps can be treated as well. Note that for every $t > 0$

$$\mathcal{L}(X_t - \gamma t) = \begin{cases} SS_\alpha(\sigma t^{1/\alpha}) & \text{in Case A,} \\ S_\alpha^+(\sigma t^{1/\alpha}) & \text{in Case B.} \end{cases}$$

Now we describe our statistical model. Suppose we have equally spaced discrete data $X_{h_n}, X_{2h_n}, \dots, X_{nh_n}$, where $(h_n)_{n\in\mathbb{N}}$ is a non-random bounded positive sequence fulfilling

$$\lim_{n\to\infty} h_n = 0 \quad \text{and} \quad \liminf_{n\to\infty} nh_n > 0. \tag{2}$$

If $h_n \asymp n^{-a}$ for example, then $a \in (0,1]$ necessarily; the symbol $a_n \asymp b_n$ means that there exists a constant $c > 0$ such that $c^{-1} \le a_n/b_n \le c$ for every n large enough. Let us emphasize that no restriction on decreasing rate of h_n other than (2) will be put in the sequel. Thus we are led to the statistical model $(P_\theta^n)_{\theta\in\Theta}$, the image measure of $(X_{ih_n})_{i=1}^n$, which depends on the unknown parameter

$$\theta = (\alpha, \gamma, \sigma) \in \Theta \subset \mathbb{R}^3.$$

The parameter space Θ is supposed to be a bounded domain whose closure is contained in:

$$\begin{cases} \{(\alpha,\gamma,\sigma) : \alpha \in (0,2), \gamma \in \mathbb{R}, \sigma > 0\} & \text{in Case A,} \\ \{(\alpha,\gamma,\sigma) : \alpha \in (0,1), \gamma \in \mathbb{R}, \sigma > 0\} & \text{in Case B.} \end{cases}$$

We implicitly suppose that there exists a true value $\theta \in \Theta$ generating the observed data.

With the above setup, the first goal of this article is to provide a LAN property for $\theta \in \Theta$ with an always degenerate Fisher information matrix (Section 3.1). This entails a negative conclusion, that is, joint estimation for (α,γ,σ) based on the maximum likelihood estimator is out of place. Nevertheless, this bad job does not arise if we suppose either α or σ is known (see (6) below). As the second goal, we provide uniform asymptotic normalities of the maximum likelihood estimates (MLE) of either (α,γ) or (γ,σ) (Section 3.2). The optimal rates for estimating α, γ, and σ will turn out to be $\sqrt{n}\log(1/h_n)$, $\sqrt{n}h_n^{1-1/\alpha}$, and $\sqrt{n}$, respectively, implying that the answer to the question "which estimate converges with most speed?" changes according as the true value of α; see (10) below. Also seen is that, as opposed to the Wiener case, we need not impose that $nh_n \to \infty$ for consistent estimation of any component of θ; that is to say, the observed information over any (non-empty) bounded time interval is rich enough. It is the $1/\alpha$-selfsimilarity of stable-Lévy processes, which plays a key role in our study as in [13] and [1], that induces these inherent phenomena; recall that a Lévy process is selfsimilar if and only if it is stable.

Remark 2.1. The parameter θ determines the generating triplet $(\gamma, 0, g_{\alpha,\sigma}(z)dz)$ of the process X, where, letting

$$c(\alpha,\sigma) = \sigma^\alpha \left\{ \frac{1}{\alpha}\Gamma(1-\alpha)\cos\frac{\alpha\pi}{2} \right\}^{-1},$$

the Lévy density $g_{\alpha,\sigma}$ is given by

$$g_{\alpha,\sigma}(z) = \begin{cases} 2^{-1}c(\alpha,\sigma)|z|^{-\alpha-1}\mathbf{1}_{\{z\neq 0\}} & \text{in Case A,} \\ c(\alpha,\sigma)z^{-\alpha-1}\mathbf{1}_{\{z>0\}} & \text{in Case B.} \end{cases} \tag{3}$$

(3) is readily seen by invoking Sato [11, Lemma 14.11]. Note that $(\alpha,\sigma)\mapsto c(\alpha,\sigma)$ is continuous on $(0,2)\times(0,\infty)$, especially $\lim_{\alpha\to 1} c(\alpha,\sigma)=\sigma/\pi$ and $\lim_{\alpha\downarrow 0} c(\alpha,\sigma)=\lim_{\alpha\uparrow 2} c(\alpha,\sigma)=0$ for every $\sigma>0$. It is important to remind that any distribution of X possibly infinite-dimensional such as $\mathcal{L}(\inf_{t\le 1} X_t)$ is determined by θ.

3 Statement of results

Under (2), without loss of generality we may suppose that $h_n\in(0,1)$ and hence $\log(1/h_n)>0$, taking the sample size n large enough. Any notation of asymptotics will be used for $n\to\infty$ unless otherwise stated.

Consider Case A and denote by $x\mapsto\phi_\alpha(x;\sigma)$ the density of $SS_\alpha(\sigma)$, especially $\phi_\alpha(x):=\phi_\alpha(x;1)$. By the scaling property we have $\phi_\alpha(x;a\sigma)=a^{-1}\phi_\alpha(a^{-1}x;\sigma)$ for every $a>0$ and $x\in\mathbb{R}$. Noticing that $\mathcal{L}(X_{ih_n}-X_{(i-1)h_n}-\gamma h_n)=SS_\alpha(\sigma h_n^{1/\alpha})$ for each $i\le n$, we see that the log-likelihood function of $(X_{ih_n})_{i=1}^n$, say $\ell_n(\theta)$, is computed as

$$\begin{aligned} \ell_n(\theta) &= \sum_{i=1}^n \log\phi_\alpha(X_{ih_n}-X_{(i-1)h_n}-\gamma h_n;\sigma h_n^{1/\alpha}) \\ &= \sum_{i=1}^n \log\left\{\sigma^{-1}h_n^{-1/\alpha}\phi_\alpha(Y_{ni})\right\} \\ &= \sum_{i=1}^n \left\{-\log\sigma+\alpha^{-1}\log(1/h_n)+\log\phi_\alpha(Y_{ni})\right\}, \end{aligned}$$

where $(Y_{ni})_{i=1}^n$ defined by

$$Y_{ni}=\sigma^{-1}h_n^{-1/\alpha}(X_{ih_n}-X_{(i-1)h_n}-\gamma h_n), \quad i\le n, \tag{4}$$

forms an iid triangular array with common law $SS_\alpha(1)$ independent of n. The exactly same as above remains true for Case B; just replace "$SS_\alpha(\sigma)$" with "$S_\alpha^+(\sigma)$".

3.1 LAN property with degenerate Fisher information

Recall that the experiment $(P_\theta^n)_{n\in\mathbb{N}}$ is said to satisfy LAN at $\theta\in\Theta$ with rate $A_n(\theta)$ if there exist a random vector $\Delta_n(\theta)$ and deterministic matrix $I(\theta)$ such that, for any bounded vector sequence $(u_n)\subset\mathbb{R}^3$ fulfilling $u_n\to u$, we have

$$\log\frac{dP^n_{\theta+\{A_n(\theta)\}^{-1}u_n}}{dP^n_\theta}=u^\top\Delta_n(\theta)-\frac{1}{2}u^\top I(\theta)u+o_{P_\theta^n}(1),$$

where $A_n(\theta)\in\mathbb{R}^{3\otimes 3}$ is invertible for each n, and $\Delta_n(\theta)\in\mathbb{R}^3$ weakly converges along (P_θ^n)-sequence to a centred normal variable with covariance matrix $I(\theta)$, which corresponds to the Fisher information matrix.

Theorem 3.1. *For both of Cases A and B, the experiment (P_θ^n) satisfies LAN at any $\theta=(\alpha,\gamma,\sigma)\in\Theta$ with rate*

$$A_n(\alpha):=\mathrm{diag}\left\{\sqrt{n}\log(1/h),\ \sqrt{n}h^{1-1/\alpha},\ \sqrt{n}\right\} \tag{5}$$

and with the always degenerate Fisher information matrix

$$I(\theta) = \begin{pmatrix} H_\alpha/\alpha^4 & J_\alpha/(\sigma\alpha^2) & H_\alpha/(\sigma\alpha^2) \\ J_\alpha/(\sigma\alpha^2) & M_\alpha/\sigma^2 & J_\alpha/\sigma^2 \\ H_\alpha/(\sigma\alpha^2) & J_\alpha/\sigma^2 & H_\alpha/\sigma^2 \end{pmatrix}. \tag{6}$$

Here

$$\begin{cases} H_\alpha = \int \frac{\{\phi_\alpha(y) + y\partial\phi_\alpha(y)\}^2}{\phi_\alpha(y)} dy, \\ M_\alpha = \int \frac{\{\partial\phi_\alpha(y)\}^2}{\phi_\alpha(y)} dy, \\ J_\alpha = \int \frac{\partial\phi_\alpha(y)\{\phi_\alpha(y) + y\partial\phi_\alpha(y)\}}{\phi_\alpha(y)} dy, \end{cases} \tag{7}$$

all being finite, especially, J_α equals zero in Case A while is positive in Case B.

Remark 3.2. Under (2) it is clear that $\{A_n(\alpha)\}^{-1} \to 0$ whenever $\alpha \in (0,2)$; this is not necessarily the case if we drop the second one of (2).

Remark 3.3. Theorem 3.1 says that, in contrast to the Wiener case, the maximum likelihood estimation for θ based on discrete sampling still leads to a degenerate Fisher information matrix. On the other hand, $I(\theta)$ does not depend on γ alike the Wiener case.

3.2 Uniform asymptotic normality when either α or σ is known

The form (6) says that we can proceed to considering efficiency issues if either α or σ is known.

We need to prepare some notation. In the sequel the notation $\Rightarrow_u$ indicates the weak convergence along (P_θ^n)-sequence, which holds uniformly over Θ^- (the closure of Θ): precisely, for any random vectors $\zeta_n(\theta)$ and $\zeta(\theta)$ with distribution P_n and P depending on $\theta \in \Theta$, we write $\zeta_n \Rightarrow_u \zeta(\theta)$ if

$$\sup_{\theta\in\Theta} \left| \int f(y)P_n(dy) - \int f(y)P(dy) \right| \to 0$$

for every continuous bounded function f. Analogous notation will be used for other modes of convergence (including ordinary convergence of a nonrandom sequence). Write $\theta' = (\alpha, \gamma)$ and $\theta'' = (\gamma, \sigma)$, and let Θ' and Θ'' respectively denote the corresponding admissible parameter spaces induced from Θ. Write $\ell_n(\theta')$ (resp. $\ell_n(\theta'')$) instead of $\ell_n(\theta)$ when σ (resp. α) is known, and let

$$D_{1n}(\alpha) = \mathrm{diag}\{\sqrt{n}\log(1/h_n), \sqrt{n}h_n^{1-1/\alpha}\}, \quad D_{2n}(\alpha) = \mathrm{diag}\{\sqrt{n}h_n^{1-1/\alpha}, \sqrt{n}\}.$$

Putting $I(\theta) = [I^{kl}(\theta)]_{k,l=1}^3$ we write

$$\mathcal{I}_1(\theta') = \begin{pmatrix} I^{11}(\theta) & I^{12}(\theta) \\ I^{21}(\theta) & I^{22}(\theta) \end{pmatrix}, \quad \mathcal{I}_2(\theta'') = \begin{pmatrix} I^{22}(\theta) & I^{23}(\theta) \\ I^{32}(\theta) & I^{33}(\theta) \end{pmatrix}.$$

Finally, put $\hat{\theta}'_n = \sup_{\theta'\in\Theta'^-} \ell_n(\theta')$ (resp. $\hat{\theta}''_n = \sup_{\theta''\in\Theta''^-} \ell_n(\theta'')$) when σ (resp. α) is known.

Now we can state our second result.

Theorem 3.4. *Let Z stand for a two-dimensional standard normal variable. For both of Cases A and B, we have the following.*

(*a*) *Suppose σ is known, while θ' is unknown. Then there exists a local maximum $\hat{\theta}'_n$ of $\ell_n(\theta')$ with probability tending to 1, for which*

$$D_{1n}(\alpha)(\hat{\theta}'_n - \theta') \Rightarrow_u \{\mathcal{I}_1(\theta')\}^{-1/2} Z. \tag{8}$$

The estimate is asymptotically efficient.

(*b*) *Suppose α is known, while θ'' is unknown. Then there exists a local maximum $\hat{\theta}''_n$ of $\ell_n(\theta'')$ with probability tending to 1, for which*

$$D_{2n}(\alpha)(\hat{\theta}''_n - \theta'') \Rightarrow_u \{\mathcal{I}_2(\theta'')\}^{-1/2} Z. \tag{9}$$

The estimate is asymptotically efficient.

Remark 3.5. To see that both of $\mathcal{I}_1(\theta')$ and $\mathcal{I}_2(\theta'')$ are positive-definite, it suffices to show that $I^{kk}(\theta) >$ for $k = 1, 2$ and that $I^{kk}(\theta)I^{ll}(\theta) - \{I^{kl}(\theta)\}^2 > 0$ for $(k, l) = (1, 2), (2, 3)$. But the former is obvious and the latter is a direct consequence of the fact $J_\alpha^2 < H_\alpha M_\alpha$ coming from Schwarz's inequality (the equality $J_\alpha^2 = H_\alpha M_\alpha$ cannot be satisfied).

Remark 3.6. Clearly $\{D_{jn}(\alpha)\}^{-1} \to_u 0$, $j = 1, 2$, under our model setup described in Section 2. Since the components of each $D_{jn}(\alpha)$ are different, it may be informative to mention the magnitude relations of the optimal rates. In Case A they are summarized as follows:

$$\begin{cases} \sigma_n \prec \alpha_n \prec \gamma_n & \text{if } \alpha \in (0, 1), \\ \sigma_n \sim \gamma_n \prec \alpha_n & \text{if } \alpha = 1, \\ \gamma_n \prec \sigma_n \prec \alpha_n & \text{if } \alpha \in (1, 2). \end{cases} \tag{10}$$

Here, $\alpha_n \prec \gamma_n$ (resp. $\alpha_n \sim \gamma_n$) means that "any rate-efficient estimate of γ converges faster than (resp. with the same rate as) that of γ". In (10), ignore σ_n (resp. α_n) in case of (a) (resp. (b)). Of course, the order is always $\sigma_n \prec \alpha_n \prec \gamma_n$ in Case B.

Remark 3.7. It follows from (8) (resp. (9)) that $\hat{\alpha}_n$ and $\hat{\gamma}_n$ (resp. $\hat{\gamma}_n$ and $\hat{\sigma}_n$) are asymptotically independent in Case A, especially, it is identical to the Wiener case that the estimations of drift and scale parameters are asymptotically independent (case (b)). In contrast, this is not the case in Case B, where the estimates are asymptotically correlated (namely, $J_\alpha > 0$).

4 Proofs

The proofs for Cases A and B are almost the same, hence for conciseness we shall first complete the proofs of Case A in Sections 4.1 and 4.2, and then turn to Case B in Section 4.3, where only the variation from the proof of Case A will be mentioned.

4.1 Proof of Theorem 3.1 for Case A

It is well known that $(\alpha, y) \mapsto \phi_\alpha(y)$ is everywhere positive and of class C^∞; by means of the Fourier-inversion formula, for any $k, k' \in \mathbb{N} \cup \{0\}$ there exist constants $c_i = c_i(\alpha, k, k') > 0$ such that

$$|\partial^k \partial_\alpha^{k'} \phi_\alpha(y)| \le c_0 \int e^{-|u|^\alpha} |u|^{c_1} \{1 + (\log|u|)^{c_2}\} du,$$

so that $|\partial^k \partial_\alpha^{k'} \phi_\alpha(y)| < \infty$. It also follows from the series expansion of the density (e.g., Sato [11, Remark 14.18]) that for any $k, k' \in \mathbb{N} \cup \{0\}$

$$|\partial^k \partial_\alpha^{k'} \phi_\alpha(y)| \asymp (\log|y|)^{k'} |y|^{-\alpha-1-k}, \tag{11}$$

as $|y| \to \infty$. Therefore:

$$\text{the quantities } H_\alpha, M_\alpha \text{ and } J_\alpha \text{ of (7) and } \int \left(\frac{\partial_\alpha \phi_\alpha(y)}{\phi_\alpha(y)} \right)^2 dy \text{ are finite.} \tag{12}$$

We denote by $\to^{\text{a.s.}}$ the almost sure convergence under P_θ^n, and introduce some notation as follows:

$$\begin{aligned} \psi_n(Y_{ni}; \theta) &:= \sigma^{-1} h_n^{-1/\alpha} \phi_\alpha(Y_{ni}), \\ g_{ni}(\theta) &:= \partial_\theta \log \psi_n(Y_{ni}; \theta), \\ \theta_n &:= \theta + A_n(\alpha) u_n. \end{aligned}$$

To complete the proof it suffices to verify Lemmas 4.1 to 4.3 below (law of large numbers, Lindeberg condition, and $L^2(P)$-differentiability, respectively; see, e.g., Greenwood and Shiryaev [5, Sections 6, 7 and 8] for details).

Lemma 4.1. $C_{1n}(\theta) := A_n(\alpha) \sum_{i=1}^{n} g_{ni}(\theta)\{g_{ni}(\theta)\}^\top A_n(\alpha) \to^{\text{a.s.}} I(\theta)$.

Lemma 4.2. $C_{2n}(\theta) := \sum_{i=1}^{n} E_\theta^n \left[\left\{ u_n^\top A_n(\alpha) g_{ni}(\theta) \right\}^2 \mathbf{1}_{\{|u_n^\top A_n(\alpha) g_{ni}(\theta)| \ge \epsilon\}} \right] \to 0$ *for every* $\epsilon > 0$.

Lemma 4.3. $C_{3n}(\theta) := n \int \left\{ \sqrt{\psi_n(y;\theta_n)} - \sqrt{\psi_n(y;\theta)} - (\theta_n - \theta)^\top \partial_\theta \sqrt{\psi_n(y;\theta)} \right\}^2 dy \to 0.$

4.1.1 Proof of Lemma 4.1

Put $g_{ni}(\theta) = [g_{ni;k}(\theta)]_{k=1}^3$. Direct partial differentiations yield

$$g_{ni;1}(\theta) = -\alpha^{-2} \log(1/h_n) F_1(Y_{ni}) + F_2(Y_{ni}), \tag{13}$$
$$g_{ni;2}(\theta) = -\sigma^{-1} h_n^{1-1/\alpha} F_3(Y_{ni}), \tag{14}$$
$$g_{ni;3}(\theta) = -\sigma^{-1} F_1(Y_{ni}), \tag{15}$$

where

$$F_1(y) = \frac{\phi_\alpha(y) + y\partial\phi_\alpha(y)}{\phi_\alpha(y)}, \quad F_2(y) = \frac{\partial_\alpha \phi_\alpha(y)}{\phi_\alpha(y)}, \quad F_3(y) = \frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)}.$$

Put $A_n(\alpha) = \mathrm{diag}\{A_{1n}(\alpha), A_{2n}(\alpha), A_{3n}(\alpha)\}$ and $C_{1n}(\theta) = [C_{1n;kl}(\theta)]_{k,l=1}^3$. Substituting the expression ($C_{1n}(\theta)$ is symmetric)

$$C_{1n;kl}(\theta) = \sum_{i=1}^{n} \{A_{kn}(\alpha) A_{ln}(\alpha)\}^{-1} g_{ni;k}(\theta) g_{ni;l}(\theta), \quad 1 \le k, l \le 3,$$

with (13) to (15), we get P_θ^n-a.s.

$$C_{1n;11}(\theta) = \alpha^{-4} \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni})^2 + O\big(\{\log(1/h_n)\}^{-1}\big) \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni}) F_2(Y_{ni}),$$
$$C_{1n;22}(\theta) = \sigma^{-2} \frac{1}{n} \sum_{i=1}^{n} F_3(Y_{ni})^2,$$
$$C_{1n;33}(\theta) = \sigma^{-2} \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni})^2,$$
$$C_{1n;12}(\theta) = (\alpha^2\sigma)^{-1} \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni}) F_3(Y_{ni}) + O\big(\{\log(1/h_n)\}^{-1}\big) \frac{1}{n} \sum_{i=1}^{n} F_2(Y_{ni}) F_3(Y_{ni}),$$
$$C_{1n;13}(\theta) = (\alpha^2\sigma)^{-1} \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni})^2 + O\big(\{\log(1/h_n)\}^{-1}\big) \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni}) F_2(Y_{ni}),$$
$$C_{1n;23}(\theta) = \sigma^{-2} \frac{1}{n} \sum_{i=1}^{n} F_1(Y_{ni}) F_3(Y_{ni}).$$

Since $(Y_{ni})_{i=1}^n$ is an iid array with common law $SS_\alpha(1)$ not depending on n (recall (4)), it follows from the strong law of large numbers that

$$\frac{1}{n} \sum_{i=1}^{n} F_k(Y_{ni}) F_l(Y_{ni}) \to^{\text{a.s.}} \int F_k(y) F_l(y) \phi_\alpha(y) dy$$

for $k, l \in \{1, 2, 3\}$, where the finiteness of the limit can be ensured by means of Schwarz's inequality and (12). In particular, we have

$$\int F_1(y)F_3(y)\phi_\alpha(y)dy = 0. \tag{16}$$

since $y \mapsto y\{\partial\phi_\alpha(y)\}^2/\phi_\alpha(y)$ is odd. Thus we get Lemma 4.1.

4.1.2 Proof of Lemma 4.2

Fix any constants $\epsilon, \delta > 0$, and write $u_n = (u_{kn})_{k=1}^3$. In the sequel the notation $a_n \lesssim b_n$ indicate that there exists a constant $c_0 > 0$ such that $a_n \le c_0 b_n$ for every sufficiently large n. Then, using the Lyapunov-type estimate we have

$$\begin{aligned} C_{2n}(\theta) &\lesssim nE_\theta^n\left[\left|\sum_{k=1}^3 u_{kn}A_{kn}(\alpha)^{-1}g_{n1;k}(\theta)\right|^{2+\delta}\right] \\ &\lesssim n\sum_{k=1}^3\{A_{kn}(\alpha)\}^{-(2+\delta)}E_\theta^n\left[\{g_{n1;k}(\theta)\}^{2+\delta}\right]. \end{aligned} \tag{17}$$

At the same time, from the expressions (13) to (15) it is easy to see that

$$\begin{aligned} E_\theta^n[|g_{n1;1}(\theta)|^{2+\delta}] &\lesssim \{\log(1/h_n)\}^{2+\delta} + 1, \\ E_\theta^n[|g_{n1;2}(\theta)|^{2+\delta}] &\lesssim h_n^{(1-1/\alpha)(2+\delta)}, \\ E_\theta^n[|g_{n1;3}(\theta)|^{2+\delta}] &\lesssim 1, \end{aligned}$$

the finiteness being guaranteed by (11), from which combined with (17) we have

$$C_{2n}(\theta) \lesssim O\left(n^{-\delta/2}\left[\{\log(1/h_n)\}^{-(2+\delta)} + 1\right]\right) = o(1),$$

completing the proof of Lemma 4.2.

4.1.3 Proof of Lemma 4.3

For simplicity, let ∂_k stand for the partial differentiation with respect to the kth component of θ. First we estimate $C_{3n}(\theta)$ as, using the standard notation for multi-indices,

$$\begin{aligned} C_{3n}(\theta) &\lesssim n\int\left\{\sum_{|r|=2}\frac{1}{r!}(u_n^\top A_n(\alpha)^{-1})^r\int_0^1(1-v)\partial_\theta^r\sqrt{\psi_n(y;\theta+v(\theta_n-\theta))}dv\right\}^2dy \\ &\lesssim n\sum_{k,l=1}^3\iint_{\mathbb{R}\times[0,1]}\{A_{kn}(\alpha)A_{ln}(\alpha)\}^{-2}\left[\partial_k\partial_l\sqrt{\psi_n(y;\theta+v(\theta_n-\theta))}\right]^2dvdy. \end{aligned} \tag{18}$$

Below we fix a small $\epsilon > 0$ for which we have $B^3(\theta;\epsilon) := \{\rho = (\rho_k)_{k=1}^3 \in \Theta : |\theta - \rho| < \epsilon\} \subset \Theta$. Note that for any $\rho = (\rho_k)_{k=1}^3$

$$\begin{aligned} \left[\partial_l\partial_k\sqrt{\psi_n(y;\rho)}\right]^2 &= \left[\frac{1}{2}\sqrt{\psi_n(y;\rho)}\Big\{\partial_l\partial_k\log\psi_n(y;\rho)\right. \\ &\quad +\Big(\partial_l\log\psi_n(y;\rho)\Big)\Big(\partial_k\log\psi_n(y;\rho)\Big)\Big\} \\ &\quad \left.-\frac{1}{4}\sqrt{\psi_n(y;\rho)}\Big(\partial_k\log\psi_n(y;\rho)\Big)\Big(\partial_l\log\psi_n(y;\rho)\Big)\right]^2 \\ &\lesssim \psi_n(y;\rho)\Big\{\Big(\partial_l\partial_k\log\psi_n(y;\rho)\Big)^2 \\ &\quad +\Big(\partial_l\log\psi_n(y;\rho)\Big)^2\Big(\partial_k\log\psi_n(y;\rho)\Big)^2\Big\}. \end{aligned} \tag{19}$$

On the other hand, we have the following:

$$
\begin{aligned}
\partial_\alpha^2 \log \psi_n(y;\theta) &= \frac{1}{\alpha^4}[\log(1/h_n)]^2 \left\{ y\frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)} + y^2 \frac{\phi_\alpha(y)\partial^2\phi_\alpha(y) - (\partial\phi_\alpha(y))^2}{\phi_\alpha(y)^2} \right\} \\
&\quad + \frac{2}{\alpha^3}\log(1/h_n) \left\{ 1 + y\frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)} - \alpha y \frac{\phi_\alpha(y)\partial\partial_\alpha\phi_\alpha(y) - (\partial\phi_\alpha(y))(\partial_\alpha(y))}{\phi_\alpha(y)^2} \right\} \\
&\quad + \frac{\phi_\alpha(y)\partial_\alpha^2\phi_\alpha(y) - (\partial_\alpha\phi_\alpha(y))^2}{\phi_\alpha(y)^2}, \qquad (20) \\
\partial_\gamma^2 \log \psi_n(y;\theta) &= \frac{1}{\sigma^2} h_n^{2(1-1/\alpha)} \frac{\phi_\alpha(y)\partial^2\phi_\alpha(y) - (\partial\phi_\alpha(y))^2}{\phi_\alpha(y)^2}, \qquad (21) \\
\partial_\sigma^2 \log \psi_n(y;\theta) &= \frac{1}{\sigma^2}\left\{ 1 + 2y\frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)} + y^2\frac{\phi_\alpha(y)\partial^2\phi_\alpha(y) - (\partial\phi_\alpha(y))^2}{\phi_\alpha(y)^2} \right\}, \qquad (22) \\
\partial_\gamma\partial_\alpha \log \psi_n(y;\theta) &= \frac{1}{\sigma\alpha^2} h_n^{1-1/\alpha}\log(1/h_n) \left\{ \frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)} + y\frac{\phi_\alpha(y)\partial^2\phi_\alpha(y) - (\partial\phi_\alpha(y))^2}{\phi_\alpha(y)^2} \right\} \\
&\quad - \frac{1}{\sigma} h_n^{1-1/\alpha} \frac{\phi_\alpha(y)\partial\partial_\alpha\phi_\alpha(y) - (\partial\phi_\alpha(y))(\partial_\alpha\phi_\alpha(y))}{\phi_\alpha(y)^2}, \qquad (23) \\
\partial_\sigma\partial_\alpha \log \psi_n(y;\theta) &= \frac{1}{\sigma\alpha^2}\log(1/h_n) \left\{ y\frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)} + y^2\frac{\phi_\alpha(y)\partial^2\phi_\alpha(y) - (\partial\phi_\alpha(y))^2}{\phi_\alpha(y)^2} \right\} \\
&\quad - \sigma^{-1} y \left\{ \frac{\partial\partial_\alpha\phi_\alpha(y)}{\phi_\alpha(y)} - \frac{\partial\phi_\alpha(y)\partial_\alpha\phi_\alpha(y)}{\phi_\alpha(y)^2} \right\}, \qquad (24) \\
\partial_\gamma\partial_\sigma \log \psi_n(y;\theta) &= \frac{1}{\sigma^2} h_n^{1-1/\alpha} \left\{ \frac{\partial\phi_\alpha(y)}{\phi_\alpha(y)} + y\frac{\phi_\alpha(y)\partial^2\phi_\alpha(y) - (\partial\phi_\alpha(y))^2}{\phi_\alpha(y)^2} \right\}. \qquad (25)
\end{aligned}
$$

On account of (11) it is easy to see that the above quantities are $\phi_\alpha(y)dy$-integrable. Letting n be so large that $\theta_n \in B^3(\theta;\epsilon)$ and then piecing together the displays (18) to (25), we see that

$$
\begin{aligned}
C_{3n}(\theta) &\lesssim n \sum_{k,l=1}^{3} \int \Big[\sup_{\rho \in B^3(\theta;\epsilon)} \{A_{kn}(\rho_1)A_{ln}(\rho_1)\}^{-2} \psi_n(y;\rho) \\
&\quad \cdot \Big\{ \Big(\partial_l\partial_k \log\psi_n(y;\rho)\Big)^2 + \Big(\partial_l \log\psi_n(y;\rho)\Big)^2 \Big(\partial_k \log\psi_n(y;\rho)\Big)^2 \Big\} \Big] dy \\
&\lesssim O(n^{-1}) = o(1),
\end{aligned}
$$

as desired.

4.2 Proof of Theorem 3.4 for Case A

In both of cases (a) and (b), the asymptotic efficiency directly follows from Theorem 3.1 and the Hajék's minimax theorem [6]. We shall only prove (a) since the proof of (b) is similar.

Denote by $\mathcal{I}_{1n}(\theta')$ the observed information matrix associated with θ':

$$
\mathcal{I}_{1n}(\theta') = [\mathcal{I}_{1n}^{kl}(\theta')]_{k,l=1}^2 := \begin{pmatrix} -\partial_\alpha^2 \ell_n(\theta') & -\partial_\gamma\partial_\alpha \ell_n(\theta') \\ \text{sym.} & -\partial_\gamma^2 \ell_n(\theta') \end{pmatrix}.
$$

For $\theta_1', \theta_2' \in \Theta'$, we also define

$$
\mathcal{I}_{1n}(\theta_1', \theta_2') = [\mathcal{I}_{1n}^{kl}(\theta_1', \theta_2')]_{k,l=1}^2 := \begin{pmatrix} -\partial_\alpha^2 \ell_n(\theta_1') & -\partial_\gamma\partial_\alpha \ell_n(\theta_1') \\ -\partial_\gamma\partial_\alpha \ell_n(\theta_2') & -\partial_\gamma^2 \ell_n(\theta_2') \end{pmatrix}.
$$

Let Θ'_α stand for the second-coordinate space of Θ'. We shall consistently denote by $\to^{\text{a.s.}}$ the $P_{\theta'}^n$-a.s. convergence. According to Sweeting [12, Theorems 1 and 2], the proof is achieved by verifying the following Lemmas 4.4 to 4.6.

Lemma 4.4. $\left|D_{1n}(\alpha)^{-1}\mathcal{I}_{1n}(\theta')D_{1n}(\alpha)^{-1} - \mathcal{I}_1(\theta')\right| \to_u^{\text{a.s.}} 0$.

Lemma 4.5. $\sup{}^*|D_{1n}(\alpha)^{-1}D_{1n}(\alpha') - I_2| \to_u 0$ *for every* $c > 0$*, where* $\sup^*$ *is taken over the set* $\{\alpha' \in \Theta_\alpha'^- : \sqrt{n}\log(1/h_n)|\alpha' - \alpha| \le c\}$.

Lemma 4.6. $\sup{}^{**}|D_{1n}(\alpha)^{-1}\{\mathcal{I}_{1n}(\theta_1', \theta_2') - \mathcal{I}_{1n}(\theta')\}D_{1n}(\alpha)^{-1}| \to_u^{\text{a.s.}} 0$ *for every* $c > 0$*, where* $\sup^{**}$ *is taken over the set* $\{\theta_k' \in \Theta'^-, k = 1, 2 : |D_{1n}(\alpha)(\theta_k' - \theta')| \le c\}$.

Remark 4.7. Concerning Lemmas 4.4 and 4.6, the original Sweeting's conditions require weaker $\Rightarrow_u$ and $\to_u^p$ (convergence in $P_{\theta'}^n$-probability) rather than $\to_u^{\text{a.s.}}$, respectively; see C1 and C2 (ii) of [12]. The primary reason why we prove the stronger uniform $P_{\theta'}^n$-a.s. convergence is that the derivation then becomes much more easy in our framework; when one attempts to prove the uniform convergence in $P_{\theta'}^n$-probability, the techniques of Ibragimov and Has'minskĭ [7, Theorems I.7 and I.20] are often employed, however, one can see that the modulus of continuity of the random field (i.e., the condition (1) of Theorem I.20 of [7]) does not seem to be fulfilled in our model.

4.2.1 Proof of Lemma 4.4

We prepare a simple version of uniform strong laws of large numbers.

Proposition 4.8. *Let* $\mathcal{U} \subset \mathbb{R}^p$ *be compact, and let* $\{(\psi_n(u))_{u\in\mathcal{U}}\}_{n\in\mathbb{N}}$ *be a sequence of real-valued random fields defined on some probability space. Suppose that* $u \mapsto \psi_n(u)$ *is continuous a.s. for every* $n \in \mathbb{N}$*, and that* $\psi_n(u) \to 0$ *a.s. for every* $u \in \mathcal{U}$ *large enough. Then we have* $\sup_{u\in\mathcal{U}} |\psi_n(u)| \to 0$ *a.s.*

Proof. The regularity conditions stated remain true for $\psi_n(u)$ replaced with $-\psi_n(\theta)$, hence it suffices to prove $\overline{\lim}_{n\to\infty} \sup_{u\in\mathcal{U}} \psi_n(u) \le 0$ a.s.

Fix any $\epsilon > 0$. Since $u \mapsto \psi_n(u)$ is uniformly continuous, there exists a constant $\delta(\epsilon) > 0$ such that for every large $n \in \mathbb{N}$

$$\sup_{u_i:|u_1-u_2|<\delta(\epsilon)} |\psi_n(u_1) - \psi_n(u_2)| < \epsilon. \tag{26}$$

For this δ_ϵ we can find a finite δ_ϵ-net $(v_j)_{j=1}^{M_\epsilon}$ of $\mathcal{U}$. Next fix any $u \in \mathcal{U}$, and then take a $j(u) \le M_\epsilon$ for which $u \in B^p(v_{j(u)}; \delta_\epsilon)$. Then we have $\psi_n(u) \le \psi_n(v_{j(u)}) + \epsilon$ a.s. for every large $n \in \mathbb{N}$, so that on account of (26) we have

$$\sup_{u\in\mathcal{U}} \psi_n(u) \le \sup_{u\in\mathcal{U}} \psi_n(v_{j(u)}) + \epsilon \le \max_{j\le M_\epsilon} \psi_n(v_j) + \epsilon \quad \text{a.s.} \tag{27}$$

(the net $(u_j)_{j=1}^{M_\epsilon}$ can be taken to be independent of u). Since M_ϵ is finite, we get the claim on taking the limit in (27) together with the assumed u-pointwise convergence. □

Now, recalling the expressions (20), (21) and (23), we can conclude the proof by applying Lemma 4.8 with replacing $\mathcal{U}$ and $\psi_n(u)$ with Θ'^- and the components of

$$G_n(\theta') := D_{1n}(\alpha)^{-1}\mathcal{I}_{1n}(\theta')D_{1n}(\alpha)^{-1} - \mathcal{I}_1(\theta'),$$

respectively; as a matter of fact, by using (11) and the elementary integration by parts (to derive $\int y\partial^2\phi_\alpha(y)dy = 0$, and so on) we can prove that $G_n(\theta') \to^{\text{a.s.}} 0$ for every $\theta' \in \Theta'$. Building on the continuity of $\theta' \mapsto G_n(\theta')$, Proposition 4.8 ends the proof of Lemma 4.4.

4.2.2 Proof of Lemma 4.5

We have $|D_{1n}(\alpha)^{-1}D_{1n}(\alpha') - I_2| = |h_n^{1/\alpha-1/\alpha'} - 1|$. Observe that

$$\sup{}^*\left|\log h_n^{1/\alpha-1/\alpha'}\right| \le \sup{}^*\left|\frac{\alpha'-\alpha}{\alpha\alpha'}\right|\log(1/h_n) \le \frac{c}{\alpha\sqrt{n}}\sup{}^*\frac{1}{\alpha'} \lesssim \frac{1}{\sqrt{n}} \to_u 0,$$

so that $\sup^* h_n^{1/\alpha-1/\alpha'} \to_u 1$. Hence the claim follows.

4.2.3 Proof of Lemma 4.6

First we note that

$$
\begin{aligned}
&\sup_{\theta'\in\Theta'} \sup{}^{**} \left| D_{1n}(\alpha)^{-1}\{\mathcal{I}_{1n}(\theta'_1,\theta'_2) - \mathcal{I}_{1n}(\theta')\} D_{1n}(\alpha)^{-1} \right| \\
&\quad \leq \sup_{\theta'\in\Theta'} \sup{}^{**} \left| n^{-1}\{\log(1/h_n)\}^{-2}\{\mathcal{I}_{1n}^{11}(\theta') - \mathcal{I}_{1n}^{11}(\theta'_1)\} \right| \\
&\qquad + \sup_{\theta'\in\Theta'} \sup{}^{**} \left| n^{-1} h_n^{-2(1-1/\alpha)}\{\mathcal{I}_{1n}^{22}(\theta') - \mathcal{I}_{1n}^{22}(\theta'_2)\} \right| \\
&\qquad + 2\sup_{\theta'\in\Theta'} \sup{}^{**} \left| n^{-1}\{\log(1/h_n)\}^{-1} h_n^{-(1-1/\alpha)}\{\mathcal{I}_{1n}^{12}(\theta') - \mathcal{I}_{1n}^{12}(\theta'_1)\} \right| \\
&\quad = H_{1n} + H_{2n} + H_{3n}, \quad \text{say.}
\end{aligned}
$$

For convenience we denote by P the underlying probability measure (defined on the Skorohod space); the law of the parametric family of X associated with all admissible $\theta' \in \Theta'$. The proof is achieved by proving $H_{kn} \to 0$ P-a.s. for all k, and to this end we shall again utilize Proposition 4.8 partly combined with the fact $\sup{}^{**}|h_n^{1/\alpha - 1/\alpha'}| \to_u 1$, which has seen in Section 4.2.2. We shall only show $H_{1n} \to^{\text{a.s.}} 0$, since the others can be shown in a similar manner.

From Lemma 4.4 we have

$$
\frac{1}{n[\log(1/h_n)]^2}\mathcal{I}_{1n}^{11}(\theta') = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\alpha^4}\frac{\{Y_{ni}\partial\phi_\alpha(Y_{ni})\}^2}{\phi_\alpha(Y_{ni})^2} + o(1) =: \mathcal{J}_n(\theta') + o(1),
$$

$P_{\theta'}^n$-a.s. uniformly in $\theta' \in \Theta'$, where we applied Proposition 4.8 to the $o(1)$ term. Hence, applying Taylor's formula around θ' to the summand of $\mathcal{J}_n(\theta')$ and then taking the definition of $\sup^{**}$ into account, we see that P-a.s.

$$
\begin{aligned}
H_{1n} &\leq \sup_{\theta'\in\Theta'} \sup{}^{**} |\mathcal{J}_n(\theta') - \mathcal{J}_n(\theta'_1)| + o(1) \\
&\lesssim \frac{1}{\sqrt{n}}\Bigg\{ \sup_{\theta'\in\Theta'} \left| [\log(1/h_n)]^{-1}\partial_\alpha \mathcal{J}_n(\theta') \right| \\
&\qquad + \sup_{\theta'\in\Theta'} \left| h_n^{-(1-1/\alpha)}\partial_\gamma \mathcal{J}_n(\theta') \right| \cdot \sup_{\theta'\in\Theta'} \sup{}^{*} \left| h_n^{1/\alpha - 1/\alpha_1} \right| \Bigg\} + o(1), \\
&\lesssim \frac{1}{\sqrt{n}}\left\{ \sup_{\theta'\in\Theta'} \left| [\log(1/h_n)]^{-1}\partial_\alpha \mathcal{J}_n(\theta') \right| + \sup_{\theta'\in\Theta'} \left| h_n^{-(1-1/\alpha)}\partial_\gamma \mathcal{J}_n(\theta') \right| \right\} + o(1). \quad (28)
\end{aligned}
$$

At the same time, the partial differentiations yields that

$$
\begin{aligned}
&[\log(1/h_n)]^{-1}\partial_\alpha \mathcal{J}_n(\theta') \\
&\quad = 2[\log(1/h_n)]^{-1}\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y_{ni}\partial\phi_\alpha(Y_{ni})}{\alpha^2\phi_\alpha(Y_{ni})}\right\}^2\left\{\frac{\partial_\alpha\partial\phi_\alpha(Y_{ni})}{\partial\phi_\alpha(Y_{ni})} - \frac{\partial_\alpha\phi_\alpha(Y_{ni})}{\phi_\alpha(Y_{ni})} - \frac{2}{\alpha}\right\} \\
&\qquad - \frac{2}{\alpha^6}\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y_{ni}\partial\phi_\alpha(Y_{ni})}{\alpha^2\phi_\alpha(Y_{ni})}\right\}^2\left[1 + Y_{ni}\left\{\frac{\partial^2\phi_\alpha(Y_{ni})}{\partial\phi_\alpha(Y_{ni})} - \frac{\partial\phi_\alpha(Y_{ni})}{\phi_\alpha(Y_{ni})}\right\}\right], \\
&h_n^{-(1-1/\alpha)}\partial_\gamma \mathcal{J}_n(\theta') \\
&\quad = \frac{2}{\sigma}\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y_{ni}\partial\phi_\alpha(Y_{ni})}{\alpha^2\phi_\alpha(Y_{ni})}\right\}^2\left\{\frac{1}{Y_{ni}} + \frac{\partial_\alpha\phi_\alpha(Y_{ni})}{\phi_\alpha(Y_{ni})}\frac{\partial^2\phi_\alpha(Y_{ni})}{\partial\phi_\alpha(Y_{ni})}\right\},
\end{aligned}
$$

from which, on account of (11) and Proposition 4.8 once again, it follows that

$$\begin{aligned}
&[\log(1/h_n)]^{-1}\partial_\alpha \mathcal{J}_n(\theta') \\
&\qquad \to_u^{\text{a.s.}} \quad \frac{2}{\alpha^6}\int \left[\frac{y^3\{\partial\phi_\alpha(y)\}^3}{\{\phi_\alpha(y)\}^2} - \frac{y^2\{\partial\phi_\alpha(y)\}^2 + y^3\partial\phi_\alpha(y)\partial^2\phi_\alpha(y)}{\phi_\alpha(y)}\right]dy, \\
&h_n^{-(1-1/\alpha)}\partial_\gamma \mathcal{J}_n(\theta') \\
&\qquad \to_u^{\text{a.s.}} \quad \frac{2}{\sigma\alpha^4}\int \left[\frac{y\{\partial\phi_\alpha(y)\}^2 - y^2\partial\phi_\alpha(y)\partial^2\phi_\alpha(y)}{\phi_\alpha(y)} + \frac{y^2\{\partial\phi_\alpha(y)\}^3}{\{\phi_\alpha(y)\}^2}\right]dy,
\end{aligned}$$

both limits being finite. Therefore we have seen that P-a.s.

$$\begin{aligned}
\sup_{\theta'\in\Theta'} |\{\log(1/h_n)\}^{-1}\partial_\alpha \mathcal{J}_n(\theta')| &= O(1), \\
\sup_{\theta'\in\Theta'} |h_n^{-(1-1/\alpha)}\partial_\gamma \mathcal{J}_n(\theta')| &= O(1),
\end{aligned}$$

from which together with (28) we get $H_{1n} \le O(n^{-1/2}) + o(1) = o(1)$ P-a.s., as desired.

4.3 Proof of Theorems 3.1 and 3.4 for Case B

The proof of Case B can be achieved along with that of Case A, except for the following:

(i) the display (16) changes to

$$\int F_1(y)F_3(y)\phi_\alpha(y)dy = J_\alpha;$$

(ii) the asymptotic behavior (11) remains valid only for $y \uparrow \infty$.

Actually, (i) does not matter; it just changes the expression of $I(\theta)$, and the positivity of J_α is clear from the definition. As for (ii), invoking the series expansion of Sato [11, Remark 14.18 (vi)] together with the scaling property of $\phi_\alpha(y)$, we see that there exist constants c_α, c'_α, and $c''_{\alpha j}$ $(j \ge 1)$ for which, given any $m \in \mathbb{N}$,

$$\begin{aligned}
\phi_\alpha(y) = c_\alpha &\exp\{-c'_\alpha y^{-\alpha/(1-\alpha)}\} y^{-(2-\alpha)/(1-\alpha)} \\
&\cdot \left\{1 + \sum_{j=1}^{m} c''_{\alpha j} y^{\alpha j/(1-\alpha)} + O\left(y^{\alpha(1+m)/(1-\alpha)}\right)\right\} \qquad (29)
\end{aligned}$$

for $y \downarrow 0$. Here the constants c_α, c'_α, and $c''_{\alpha j}$ smoothly depend on α over $(0,1)$. Due to the exponential factor $\exp\{-c'_\alpha y^{-\alpha/(1-\alpha)}\}$ appearing in (29), we see that $\partial^k\partial_\alpha^{k'}\phi_\alpha(y)$ decreases to 0 very fast as $y \downarrow 0$ for any $k, k' \in \mathbb{N}\cup\{0\}$, which enables us to verify, especially, the finiteness of the Fisher information matrix $I(\theta)$, and indeed, to follow the same line as in Case A.

Acknowledgment

This work was partly supported by Grant-in-Aid for Scientific Research from the Ministry of Education, Japan, and Cooperative Research Program of the Institute of Statistical Mathematics, Japan.

References

[1] Aït-Sahalia, Y. and Jacod, J. (2004), Fisher's information for discretely sampled Lévy processes. preprint.

[2] Aït-Sahalia, Y. and Jacod, J. (2004), Volatility estimators for discretely sampled Lévy processes. preprint.

[3] Gobet, E. (2001), Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. *Bernoulli* **7**, 899–912.

[4] Gobet, E. (2002), LAN property for ergodic diffusions with discrete observations. *Ann. Inst. H. Poincare Probab. Statist.* **38**, 711–737.

[5] Greenwood, P. E. and Shiryayev, A. N. (1985), *Contiguity and the Statistical Invariance Principle.* Gordon and Breach Science Publishers, New York.

[6] Hajék, J. (1972), Local asymptotic minimax and admissibility in estimation. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics, pp. 175–194. Univ. California Press, Berkeley, Calif.

[7] Ibragimov, I. A. and Has'minskĭ, R. Z. (1981), *Statistical Estimation. Asymptotic Theory.* Springer-Verlag, New York-Berlin.

[8] Kabanov, Yu., Liptser, R. S. and Shiryaev, A. N. (1980), Absolute continuity and singularity of locally absolutely continuous probability distributions. II. *Math. USSR Sbornik* **36**, 31–58. (English translation)

[9] Kessler, M. (1997), Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24**, 211–229.

[10] Nolan, J. P. (2001), Maximum likelihood estimation and diagnostics for stable distributions. *Levy Processes. Theory and Applications.* Barndorff-Nielsen, O. E., Mikosch, T. and Resnick, S. I. (eds.), 379–400, Birkhäuser Boston, Boston.

[11] Sato, K. (1999), *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

[12] Sweeting, T. J. (1980), Uniform asymptotic normality of the maximum likelihood estimator. *Ann. Statist.* **8**, 1375–1381. [Corrections: (1982) *Ann. Statist.* **10**, 320.]

[13] Woerner, J. H. C. (2001), *Statistical Analysis for Discretely Observed Lévy Processes.* PhD thesis, University of Freiburg.

[14] Yoshida, N. (1992), Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.* **41**, 220–242.

List of MHF Preprint Series, Kyushu University

21st Century COE Program
Development of Dynamic Mathematics with High Functionality

MHF2004-1 Koji YONEMOTO & Takashi YANAGAWA
Estimating the Lyapunov exponent from chaotic time series with dynamic noise

MHF2004-2 Rui YAMAGUCHI, Eiko TSUCHIYA & Tomoyuki HIGUCHI
State space modeling approach to decompose daily sales of a restaurant into time-dependent multi-factors

MHF2004-3 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Cubic pencils and Painlevé Hamiltonians

MHF2004-4 Atsushi KAWAGUCHI, Koji YONEMOTO & Takashi YANAGAWA
Estimating the correlation dimension from a chaotic system with dynamic noise

MHF2004-5 Atsushi KAWAGUCHI, Kentarou KITAMURA, Koji YONEMOTO, Takashi YANAGAWA & Kiyofumi YUMOTO
Detection of auroral breakups using the correlation dimension

MHF2004-6 Ryo IKOTA, Masayasu MIMURA & Tatsuyuki NAKAKI
A methodology for numerical simulations to a singular limit

MHF2004-7 Ryo IKOTA & Eiji YANAGIDA
Stability of stationary interfaces of binary-tree type

MHF2004-8 Yuko ARAKI, Sadanori KONISHI & Seiya IMOTO
Functional discriminant analysis for gene expression data via radial basis expansion

MHF2004-9 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Hypergeometric solutions to the q - Painlevé equations

MHF2004-10 Raimundas VIDŪNAS
Expressions for values of the gamma function

MHF2004-11 Raimundas VIDŪNAS
Transformations of Gauss hypergeometric functions

MHF2004-12 Koji NAKAGAWA & Masakazu SUZUKI
Mathematical knowledge browser

MHF2004-13 Ken-ichi MARUNO, Wen-Xiu MA & Masayuki OIKAWA
Generalized Casorati determinant and Positon-Negaton-Type solutions of the Toda lattice equation

MHF2004-14 Nalini JOSHI, Kenji KAJIWARA & Marta MAZZOCCO
Generating function associated with the determinant formula for the solutions of the Painlevé II equation

MHF2004-15 Kouji HASHIMOTO, Ryohei ABE, Mitsuhiro T. NAKAO & Yoshitaka WATANABE
Numerical verification methods of solutions for nonlinear singularly perturbed problem

MHF2004-16 Ken-ichi MARUNO & Gino BIONDINI
Resonance and web structure in discrete soliton systems: the two-dimensional Toda lattice and its fully discrete and ultra-discrete versions

MHF2004-17 Ryuei NISHII & Shinto EGUCHI
Supervised image classification in Markov random field models with Jeffreys divergence

MHF2004-18 Kouji HASHIMOTO, Kenta KOBAYASHI & Mitsuhiro T. NAKAO
Numerical verification methods of solutions for the free boundary problem

MHF2004-19 Hiroki MASUDA
Ergodicity and exponential β-mixing bounds for a strong solution of Lévy-driven stochastic differential equations

MHF2004-20 Setsuo TANIGUCHI
The Brownian sheet and the reflectionless potentials

MHF2004-21 Ryuei NISHII & Shinto EGUCHI
Supervised image classification based on AdaBoost with contextual weak classifiers

MHF2004-22 Hideki KOSAKI
On intersections of domains of unbounded positive operators

MHF2004-23 Masahisa TABATA & Shoichi FUJIMA
Robustness of a characteristic finite element scheme of second order in time increment

MHF2004-24 Ken-ichi MARUNO, Adrian ANKIEWICZ & Nail AKHMEDIEV
Dissipative solitons of the discrete complex cubic-quintic Ginzburg-Landau equation

MHF2004-25 Raimundas VIDŪNAS
Degenerate Gauss hypergeometric functions

MHF2004-26 Ryo IKOTA
The boundedness of propagation speeds of disturbances for reaction-diffusion systems

MHF2004-27 Ryusuke KON
Convex dominates concave: an exclusion principle in discrete-time Kolmogorov systems

MHF2004-28 Ryusuke KON
Multiple attractors in host-parasitoid interactions: coexistence and extinction

MHF2004-29 Kentaro IHARA, Masanobu KANEKO & Don ZAGIER
Derivation and double shuffle relations for multiple zeta values

MHF2004-30 Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Generalized partitioned quantum cellular automata and quantization of classical CA

MHF2005-1 Hideki KOSAKI
Matrix trace inequalities related to uncertainty principle

MHF2005-2 Masahisa TABATA
Discrepancy between theory and real computation on the stability of some finite element schemes

MHF2005-3 Yuko ARAKI & Sadanori KONISHI
Functional regression modeling via regularized basis expansions and model selection

MHF2005-4 Yuko ARAKI & Sadanori KONISHI
Functional discriminant analysis via regularized basis expansions

MHF2005-5 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Point configurations, Cremona transformations and the elliptic difference Painlevé equations

MHF2005-6 Kenji KAJIWARA, Tetsu MASUDA, Masatoshi NOUMI, Yasuhiro OHTA & Yasuhiko YAMADA
Construction of hypergeometric solutions to the q - Painlevé equations

MHF2005-7 Hiroki MASUDA
Simple estimators for non-linear Markovian trend from sampled data:
I. ergodic cases

MHF2005-8 Hiroki MASUDA & Nakahiro YOSHIDA
Edgeworth expansion for a class of Ornstein-Uhlenbeck-based models

MHF2005-9 Masayuki UCHIDA
Approximate martingale estimating functions under small perturbations of dynamical systems

MHF2005-10 Ryo MATSUZAKI & Masayuki UCHIDA
One-step estimators for diffusion processes with small dispersion parameters from discrete observations

MHF2005-11 Junichi MATSUKUBO, Ryo MATSUZAKI & Masayuki UCHIDA
Estimation for a discretely observed small diffusion process with a linear drift

MHF2005-12 Masayuki UCHIDA & Nakahiro YOSHIDA
AIC for ergodic diffusion processes from discrete observations

MHF2005-13 Hiromichi GOTO & Kenji KAJIWARA
Generating function related to the Okamoto polynomials for the Painlevé IV equation

MHF2005-14 Masato KIMURA & Shin-ichi NAGATA
Precise asymptotic behaviour of the first eigenvalue of Sturm-Liouville problems with large drift

MHF2005-15 Daisuke TAGAMI & Masahisa TABATA
Numerical computations of a melting glass convection in the furnace

MHF2005-16 Raimundas VIDŪNAS
Normalized Leonard pairs and Askey-Wilson relations

MHF2005-17 Raimundas VIDŪNAS
Askey-Wilson relations and Leonard pairs

MHF2005-18 Kenji KAJIWARA & Atsushi MUKAIHIRA
Soliton solutions for the non-autonomous discrete-time Toda lattice equation

MHF2005-19 Yuu HARIYA
Construction of Gibbs measures for 1-dimensional continuum fields

MHF2005-20 Yuu HARIYA
Integration by parts formulae for the Wiener measure restricted to subsets in $\mathbb{R}^d$

MHF2005-21 Yuu HARIYA
A time-change approach to Kotani's extension of Yor's formula

MHF2005-22 Tadahisa FUNAKI, Yuu HARIYA & Mark YOR
Wiener integrals for centered powers of Bessel processes, I

MHF2005-23 Masahisa TABATA & Satoshi KAIZU
Finite element schemes for two-fluids flow problems

MHF2005-24 Ken-ichi MARUNO & Yasuhiro OHTA
Determinant form of dark soliton solutions of the discrete nonlinear Schrödinger equation

MHF2005-25 Alexander V. KITAEV & Raimundas VIDŪNAS
Quadratic transformations of the sixth Painlevé equation

MHF2005-26 Toru FUJII & Sadanori KONISHI
Nonlinear regression modeling via regularized wavelets and smoothing parameter selection

MHF2005-27 Shuichi INOKUCHI, Kazumasa HONDA, Hyen Yeal LEE, Tatsuro SATO, Yoshihiro MIZOGUCHI & Yasuo KAWAHARA
On reversible cellular automata with finite cell array

MHF2005-28 Toru KOMATSU
Cyclic cubic field with explicit Artin symbols

MHF2005-29 Mitsuhiro T. NAKAO, Kouji HASHIMOTO & Kaori NAGATOU
A computational approach to constructive a priori and a posteriori error estimates for finite element approximations of bi-harmonic problems

MHF2005-30 Kaori NAGATOU, Kouji HASHIMOTO & Mitsuhiro T. NAKAO
Numerical verification of stationary solutions for Navier-Stokes problems

MHF2005-31 Hidefumi KAWASAKI
A duality theorem for a three-phase partition problem

MHF2005-32 Hidefumi KAWASAKI
A duality theorem based on triangles separating three convex sets

MHF2005-33 Takeaki FUCHIKAMI & Hidefumi KAWASAKI
An explicit formula of the Shapley value for a cooperative game induced from the conjugate point

MHF2005-34 Hideki MURAKAWA
A regularization of a reaction-diffusion system approximation to the two-phase Stefan problem

MHF2006-1 Masahisa TABATA
Numerical simulation of Rayleigh-Taylor problems by an energy-stable finite element scheme

MHF2006-2 Ken-ichi MARUNO & G R W QUISPEL
Construction of integrals of higher-order mappings

MHF2006-3 Setsuo TANIGUCHI
On the Jacobi field approach to stochastic oscillatory integrals with quadratic phase function

MHF2006-4 Kouji HASHIMOTO, Kaori NAGATOU & Mitsuhiro T. NAKAO
A computational approach to constructive a priori error estimate for finite element approximations of bi-harmonic problems in nonconvex polygonal domains

MHF2006-5 Hidefumi KAWASAKI
A duality theory based on triangular cylinders separating three convex sets in R^n

MHF2006-6 Raimundas VIDŪNAS
Uniform convergence of hypergeometric series

MHF2006-7 Yuji KODAMA & Ken-ichi MARUNO
N-Soliton solutions to the DKP equation and Weyl group actions

MHF2006-8 Toru KOMATSU
Potentially generic polynomial

MHF2006-9 Toru KOMATSU
Generic sextic polynomial related to the subfield problem of a cubic polynomial

MHF2006-10 Shu TEZUKA & Anargyros PAPAGEORGIOU
Exact cubature for a class of functions of maximum effective dimension

MHF2006-11 Shu TEZUKA
On high-discrepancy sequences

MHF2006-12 Raimundas VIDŪNAS
Detecting persistent regimes in the North Atlantic Oscillation time series

MHF2006-13 Toru KOMATSU
Tamely Eisenstein field with prime power discriminant

MHF2006-14 Nalini JOSHI, Kenji KAJIWARA & Marta MAZZOCCO
Generating function associated with the Hankel determinant formula for the solutions of the Painlevé IV equation

MHF2006-15 Raimundas VIDŪNAS
Darboux evaluations of algebraic Gauss hypergeometric functions

MHF2006-16 Masato KIMURA & Isao WAKANO
New mathematical approach to the energy release rate in crack extension

MHF2006-17 Toru KOMATSU
Arithmetic of the splitting field of Alexander polynomial

MHF2006-18 Hiroki MASUDA
Likelihood estimation of stable Lévy processes from discrete data