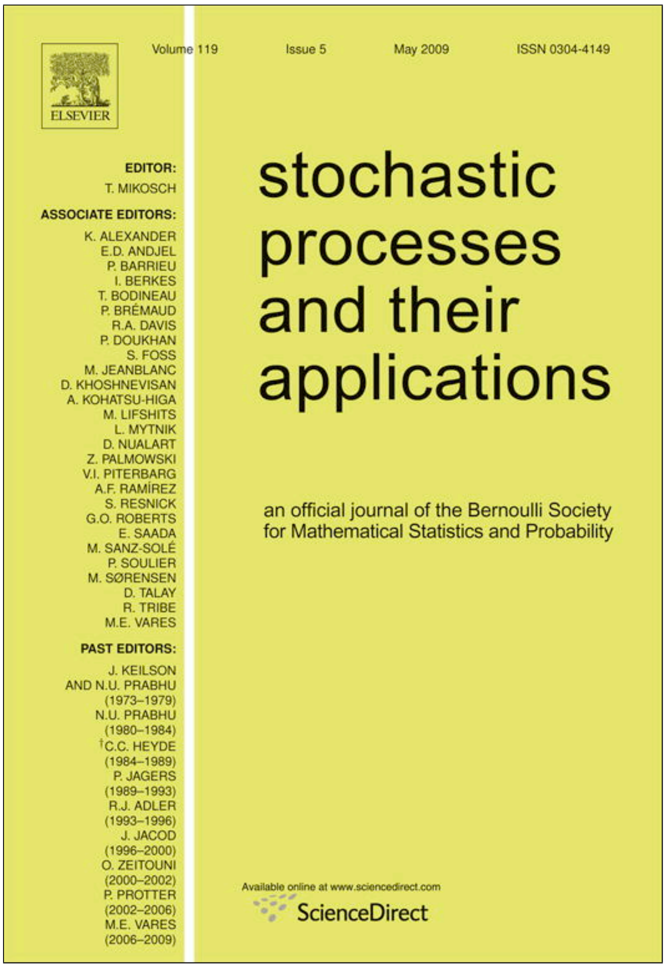
Volume 119
Issue 5
May 2009
ISSN 0304-4149
ELSEVIER
EDITOR:
T. MIKOSCH
ASSOCIATE EDITORS:
K. ALEXANDER
E.D. ANDJEL
P. BARRIEU
I. BERKES
T. BODINEAU
P. BRÉMAUD
R.A. DAVIS
P. DOUKHAN
S. FOSS
M. JEANBLANC
D. KHOSHNEVISAN
A. KOHATSU-HIGA
M. LIFSHITS
L. MYTNIK
D. NUALART
Z. PALMOWSKI
V.I. PITERBARG
A.F. RAMÍREZ
S. RESNICK
G.O. ROBERTS
E. SAADA
M. SANZ-SOLÉ
P. SOULIER
M. SØRENSEN
D. TALAY
R. TRIBE
M.E. VARES
PAST EDITORS:
J. KEILSON
AND N.U. PRABHU
(1973–1979)
N.U. PRABHU
(1980–1984)
†C.C. HEYDE
(1984–1989)
P. JAGERS
(1989–1993)
R.J. ADLER
(1993–1996)
J. JACOD
(1996–2000)
O. ZEITOUNI
(2000–2002)
P. PROTTER
(2002–2006)
M.E. VARES
(2006–2009)
stochastic processes and their applications
an official journal of the Bernoulli Society for Mathematical Statistics and Probability
Available online at www.sciencedirect.com
ScienceDirect

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Stochastic Processes and their Applications 119 (2009) 1580–1600
stochastic processes and their applications
www.elsevier.com/locate/spa

Parametric estimation for partially hidden diffusion processes sampled at discrete times

Stefano Maria Iacus[a], Masayuki Uchida[b,*], Nakahiro Yoshida[c,d]

[a] *Department of Economics, Business and Statistics, University of Milan, Via Conservatorio 7, 20122 Milan, Italy*
[b] *Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*
[c] *Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan*
[d] *Japan Science and Technology Agency (JST), Japan*

Received 1 February 2008; received in revised form 28 July 2008; accepted 12 August 2008
Available online 19 August 2008

Abstract

For a one-dimensional diffusion process $X = \{X(t);\ 0 \le t \le T\}$, we suppose that $X(t)$ is hidden if it is below some fixed and known threshold τ, but otherwise it is visible. This means a partially hidden diffusion process. The problem treated in this paper is the estimation of a finite-dimensional parameter in both drift and diffusion coefficients under a partially hidden diffusion process obtained by a discrete sampling scheme. It is assumed that the sampling occurs at regularly spaced time intervals of length h_n such that $nh_n = T$. The asymptotic is when $h_n \to 0$, $T \to \infty$ and $nh_n^2 \to 0$ as $n \to \infty$. Consistency and asymptotic normality for estimators of parameters in both drift and diffusion coefficients are proved.

MSC: primary 62F12; 62M05; secondary 60J60

Keywords: Discrete observations; Partially observed systems; Diffusion processes

1. Introduction

We consider the estimation of the unknown parameter $\theta = (\theta_1, \theta_2)$ characterizing a one-dimensional diffusion process defined by the stochastic differential equation

$$\mathrm{d}X(t) = b(X(t), \theta_2)\mathrm{d}t + \sigma(X(t), \theta_1)\mathrm{d}W_t, \qquad X(0) = x_0, \quad t \in [0, T],$$

* Corresponding author. Tel.: +81 6 6850 6465; fax: +81 6 6850 6496.
E-mail address: uchida@sigmath.es.osaka-u.ac.jp (M. Uchida).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2008.08.004

where W is a one-dimensional standard Brownian motion, b and σ are supposed to be regular enough to ensure the existence of a (strong) solution to the above stochastic differential equation. In the situation where discrete observations are $\mathbf{X}_n = \{X(t_i);\ i = 0, 1, \ldots, n\}$ with $t_i = ih_n$, $nh_n = T$, the estimation problem for the parameter θ has been considered by several authors, see [2,11,12,16,17,9]. In this paper, however, we generalize it to a different setup. We suppose that $X(t)$ is observable if $X(t) > \tau$ for some threshold τ, and that $X(t)$ cannot be observed if $X(t) \le \tau$. This means that the original process becomes a partially hidden diffusion process based on a threshold τ, and the discretized trajectory $\mathbf{X}_n$ is also influenced by a threshold τ. This type of observation naturally arises in the study of stochastic resonance and has been treated so far in the statistical context for the i.i.d. case in [4], for continuous time ergodic diffusion processes in [6] and for a class of continuous time mixing processes in [7]. In signal theory this corresponds to the problem of signal detection when the signal is so faint that it is not always receivable by some detector. This scheme of observation frequently appears in radio and CCD astronomy in the problem of identification of faint perturbed signals originated by astronomical sources (see e.g. [14]). A partially observed diffusion model also arises in the context of financial markets (see e.g. [18]) and in neuronal activation analysis (see e.g. [10]). In stochastic resonance context the original observation is altered by adding some noise with known structure to the channel in order to have full (but eventually quite noisy) observations, hence the problem is the one of determining the optimal level of noise. In the approach used in this paper, only the available observations are retained and used to estimate θ. In this setup, we need to build a contrast function which is different from the one proposed in the literature of estimation for discretely observed diffusion processes cited above. Other different approaches based on particle filters (see e.g. [1]) and observation augmentation (see e.g. [13]) have been also recently proposed in the literature but our approach and asymptotic scheme adopted are substantially different from these references. Nevertheless, after some refinement it is still possible to prove consistency and asymptotic normality of the proposed estimators along the lines of e.g. [15,16,3, 9]. The organization of the paper is as follows. Section 2 introduces the model, the assumptions and two contrast functions. Section 3 contains the statement of the main result on consistency and asymptotic normality of estimators. Section 4 is devoted to the proofs of the results in Section 3.

2. Model of observation and assumptions

Let $X = \{X(t);\ 0 \le t \le T\}$ denote a diffusion process satisfying

$$\mathrm{d}X(t) = b(X(t), \theta_2)\mathrm{d}t + \sigma(X(t), \theta_1)\mathrm{d}W_t, \qquad X(0) = x_0, \quad t \in [0, T]. \tag{1}$$

The parameter of our interest is $\theta = (\theta_1, \theta_2)$, $\theta \in \Theta$ and Θ is a compact rectangle in $\mathbf{R}^2$. The true value is denoted by $\theta_0 = (\theta_{1,0}, \theta_{2,0})$ and it is assumed that $\theta_0 \in \mathrm{Int}(\Theta)$. Let $X_i = X(t_i)$, $t_i = ih_n$ and $nh_n = T$. For $i = 0, 1, \ldots, n$, we assume that X_i is observable if $X_i > \tau$ for some threshold τ, and that X_i is unobserved if $X_i \le \tau$. The asymptotics will be investigated when $h_n \to 0$, $nh_n \to \infty$ and $nh_n^2 \to 0$ as $n \to \infty$. In order to simplify the description, we use the following notation

$$\sigma_i = \sigma(X_i, \theta_1), \qquad b_i = b(X_i, \theta_2), \qquad \Delta_i X = X_i - X_{i-1}.$$

When the coefficients are evaluated at the true value of the parameter, we will write

$$\sigma_i^* = \sigma(X_i, \theta_{1,0}), \qquad b_i^* = b(X_i, \theta_{2,0}).$$

We further define $\partial_{\theta_i} f = \frac{\partial}{\partial \theta_i} f$. For any real sequence $u_n \in (0, 1]$, $R(u_n, x)$ represents a function such that

$$|R(u_n, x)| \le u_n C(1 + |x|)^C, \tag{2}$$

where C is a positive constant independent of n and x (and eventually θ when x is $X(t)$). In the proof, K and/or C denote generic constants independent of θ, x and n.

Assumptions

A1 There exists $K > 0$ such that for every $x, y \in \mathbf{R}$,

$$|b(x, \theta_{2,0}) - b(y, \theta_{2,0})| + |\sigma(x, \theta_{1,0}) - \sigma(y, \theta_{1,0})| \le K|x - y|,$$

so that (1) has a unique solution for $\theta = \theta_0$.

A2 The process X is stationary and ergodic for $\theta = \theta_0$ with its invariant measure denoted by ν_{θ_0}.

A3 For all $p \ge 0$, $\mathrm{E}\left[|X(0)|^p\right] < \infty$.

A4 $\inf_{x,\theta_1} \sigma^2(x, \theta_1) = K_4 > 0$.

A5 (Polynomial growth) The coefficients b and σ are continuously differentiable with respect to x up to order 2 for all θ_1 and θ_2. These coefficients and their derivatives up to order 2 are of polynomial growth in x, uniformly in θ.

A6 (Polynomial growth) The coefficients b and σ and all their x derivatives up to order 2, are three times continuously differentiable with respect to θ for all x. Moreover, these θ-derivatives are of polynomial growth in x and uniformly on θ.

A7 (Identifiability) $\sigma^2(x, \theta_1) = \sigma^2(x, \theta_{1,0})$ for ν_{θ_0} a.s. all $x \Rightarrow \theta_1 = \theta_{1,0}$, $b(x, \theta_2) = b(x, \theta_{2,0})$ for ν_{θ_0} a.s. all $x \Rightarrow \theta_2 = \theta_{2,0}$.

The contrast function

The main idea of this paper is to fix a new threshold $\tau'\,(>\tau)$ as follows. We fix a number $\alpha \in (0, 1/2)$ and take a sequence $\tau_n\,(>\tau)$ such that $h_n^\alpha/(\tau_n - \tau) = O(1)$; for example, $\tau_n = \tau + h_n^\alpha$. We use τ' instead of τ_n. Notice that $\tau' \to \tau$ slowly. Thus, we introduce the following contrast functions

$$g_n(\theta_1) = \sum_{i=1}^n g(i, i-1; \theta_1) \chi_{\{X_{i-1} > \tau', X_i > \tau\}}, \tag{3}$$

$$\ell_n(\theta) = \sum_{i=1}^n \ell(i, i-1; \theta) \chi_{\{X_{i-1} > \tau', X_i > \tau\}}, \tag{4}$$

where χ is the indicator function and

$$g(i, i-1; \theta_1) = \log \sigma_{i-1}^2 + \frac{(\Delta_i X)^2}{\sigma_{i-1}^2 h_n},$$

$$\ell(i, i-1; \theta) = \log \sigma_{i-1}^2 + \frac{(\Delta_i X - b_{i-1} h_n)^2}{\sigma_{i-1}^2 h_n}.$$

3. Consistent and asymptotically normal estimators

As in [16], we first estimate the parameter belonging to the diffusion coefficient, i.e. θ_1, because, as usual, the estimator of θ_1 has a faster rate of convergence than the one of θ_2. Let $\hat{\theta}_{1,n}$ denote an estimator of θ_1 satisfying

$$g_n(\hat{\theta}_{1,n}) = \inf_{\theta_1} g_n(\theta_1). \tag{5}$$

The measurable selection theorem ensures the existence of such a measurable mapping.

Theorem 3.1. *Under assumptions* A1–A7*, as* $h_n \to 0$ *and* $nh_n \to \infty$*,*

$$\hat{\theta}_{1,n} \xrightarrow{p} \theta_{1,0}.$$

We consider an estimator $\hat{\theta}_{2,n}$ of θ_2 that satisfies

$$\ell_n(\hat{\theta}_{1,n}, \hat{\theta}_{2,n}) = \inf_{\theta_2} \ell_n(\hat{\theta}_{1,n}, \theta_2). \tag{6}$$

Theorem 3.2. *Under assumptions* A1–A7*, as* $h_n \to 0$ *and* $nh_n \to \infty$*,*

$$\hat{\theta}_{2,n} \xrightarrow{p} \theta_{2,0}.$$

Let

$$\Sigma = \begin{pmatrix} 2\int \left(\frac{\partial_{\theta_1}\sigma(x,\theta_{1,0})}{\sigma(x,\theta_{1,0})}\right)^2 \chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x) & 0 \\ 0 & \int \left(\frac{\partial_{\theta_2}b(x,\theta_{2,0})}{\sigma(x,\theta_{1,0})}\right)^2 \chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x) \end{pmatrix}.$$

The next theorem is the main result in this paper.

Theorem 3.3. *Suppose that the assumptions* A1–A7 *are satisfied. If* Σ *is non-singular, then as* $h_n \to 0$*,* $nh_n \to \infty$ *and* $nh_n^2 \to 0$*,*

$$\begin{pmatrix} \sqrt{n}(\hat{\theta}_{1,n} - \theta_{1,0}) \\ \sqrt{nh_n}(\hat{\theta}_{2,n} - \theta_{2,0}) \end{pmatrix} \xrightarrow{d} N(0, \Sigma^{-1}).$$

Remark 1. (i) As seen from the proof of (14), stationarity of the diffusion is used in order to show that $P(\tau' < X_{i-1} \le \tilde{\tau}) = o(1)$ and that $P(\tau'' < X_{i-1} \le \tau') = o(1)$, where $\tilde{\tau} = \tau' + h_n^\alpha$ and $\tau'' = \tau' - h_n^\alpha$ for $\alpha \in (0, 1/2)$. If we suppose A1 and A2 except for stationarity, the moment condition satisfying that $\sup_t \mathrm{E}\left[|X(t)|^p\right] < \infty$ for all $p \ge 0$, and A4–A7 together with regularity conditions for which the above estimates hold, Theorems 3.1–3.3 still hold true. (ii) It seems true that under some regularity conditions, Theorem 3.1 still holds even if $T(= nh_n)$ is fixed. For the case that T is fixed, consistency and asymptotically mixed normality of the estimator will be a future work.

4. Proofs

Proof of Theorem 3.1. First, we will show that

$$\sup_{\theta_1} \left| \frac{1}{n} g_n(\theta_1) - G(\theta_1) \right| \xrightarrow{p} 0, \tag{7}$$

where

$$G(\theta_1) = \int_{\mathbf{R}} \left\{ \log \sigma^2(x, \theta_1) + \frac{\sigma^2(x, \theta_{1,0})}{\sigma^2(x, \theta_1)} \right\} \chi_{\{x>\tau\}} \nu_{\theta_0}(\mathrm{d}x).$$

Noting that

$$\chi_{\{X_{i-1}>\tau', X_i>\tau\}} - \chi_{\{X_{i-1}>\tau'\}} = -\chi_{\{X_{i-1}>\tau', X_i\le\tau\}}, \tag{8}$$

one has

$$\frac{1}{n} g_n(\theta_1) = \frac{1}{n} \sum_{i=1}^n g(i, i-1; \theta_1) \chi_{\{X_{i-1}>\tau'\}} \tag{9}$$

$$- \frac{1}{n} \sum_{i=1}^n g(i, i-1; \theta_1) \chi_{\{X_{i-1}>\tau', X_i\le\tau\}}. \tag{10}$$

In order to show the uniform convergence of (10) to zero, we consider the estimate that

$$\mathrm{E}\left\{ \sup_{\theta_1} \left| \frac{1}{n} \sum_{i=1}^n g(i, i-1; \theta_1) \chi_{\{X_{i-1}>\tau', X_i\le\tau\}} \right| \right\}$$
$$\le \frac{1}{n} \sum_{i=1}^n \left\| \sup_{\theta_1} \left| \log \sigma_{i-1}^2 + \frac{(\Delta_i X)^2}{h_n \sigma_{i-1}^2} \right| \right\|_p P(X_{i-1} > \tau', X_i \le \tau)^{\frac{1}{q}}$$

for $1 < p, q < \infty$ with $1/p + 1/q = 1$. Since A4 and A5 imply that

$$\sup_{\theta_1} \left| \log \sigma_{i-1}^2 \right| \le \max(|\log(K_4)^2|, \sup_{\theta_1} |\sigma_{i-1}^2|) \le K_4' + C(1 + |X_{i-1}|)^C,$$

it follows from A3 that

$$\left\| \sup_{\theta_1} \left| \log \sigma_{i-1}^2 \right| \right\|_p < \infty.$$

By A4 and the estimate that $\mathrm{E}|X_i - X_{i-1}|^{2p} \le C h_n^p$ for $p \ge 1$,

$$\left\| \sup_{\theta_1} \frac{(\Delta_i X)^2}{h_n \sigma_{i-1}^2} \right\|_p^p \le K \left\| \frac{(\Delta_i X)^2}{h_n} \right\|_p^p = O(1).$$

Moreover, for $k > 0$,

$$\begin{aligned} \sup_i P(X_{i-1} > \tau', X_i \le \tau) &\le \sup_i P(|X_{i-1} - X_i| \ge \tau' - \tau) \\ &\le \left(\frac{1}{\tau' - \tau} \right)^k \sup_i \mathrm{E}|X_{i-1} - X_i|^k \\ &\le C \left(\frac{h_n^\alpha}{\tau' - \tau} \right)^k (h_n^{1/2-\alpha})^k \\ &= O\left(h_n^{(1/2-\alpha)k} \right) \to 0 \end{aligned} \tag{11}$$

because $h_n^\alpha/(\tau' - \tau) = O(1)$ for $\alpha \in (0, 1/2)$. Thus, we obtain

$$\sup_{\theta_1} \left| \frac{1}{n} \sum_{i=1}^{n} g(i, i-1; \theta) \chi_{\{X_{i-1} > \tau', X_i \le \tau\}} \right| = o_p(1).$$

In order to prove the uniform convergence of (9) to G, it is enough to show that

$$\frac{1}{n} \sum_{i=1}^{n} g(i, i-1; \theta_1) \chi_{\{X_{i-1} > \tau'\}} \xrightarrow{p} G(\theta_1) \tag{12}$$

for each θ_1, and

$$\sup_n E \left[\sup_{\theta_1} \left| \frac{1}{n} \sum_{i=1}^{n} \partial_{\theta_1} g(i, i-1; \theta_1) \right| \right] < \infty. \tag{13}$$

For details, see the proof of Theorem 4.1 in [15]. As in the proof of the uniform convergence of (10), we can obtain (13). For the proof of (12), we will prove

$$\frac{1}{n} \sum_{i=1}^{n} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \xrightarrow{p} \int_{\mathbf{R}} \log \sigma^2(x, \theta) \chi_{\{x > \tau\}} \nu_{\theta_0}(\mathrm{d}x), \tag{14}$$

$$\frac{1}{nh_n} \sum_{i=1}^{n} \frac{(\Delta_i X)^2}{\sigma_{i-1}^2} \chi_{\{X_{i-1} > \tau'\}} \xrightarrow{p} \int_{\mathbf{R}} \frac{\sigma^2(x, \theta_{1,0})}{\sigma^2(x, \theta_1)} \chi_{\{x > \tau\}} \nu_{\theta_0}(\mathrm{d}x). \tag{15}$$

For the proof of (14), we set $I_i = \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \mathrm{d}s$ for $i = 1, \ldots, n$. Note that

$$\chi_{\{X_{i-1} > \tau'\}} \ge \chi_{\{\inf_{s \in (t_{i-1}, t_i]} X(s) > \tau\}} - \chi_{\{\tau < \inf_{s \in (t_{i-1}, t_i]} X(s) \le \tau'\}}. \tag{16}$$

We first estimate I_i for the case that $\log \sigma_{i-1}^2 \ge 0$. Let $J_i = \chi_{\{\log \sigma_{i-1}^2 \ge 0\}}$ for $i = 1, \ldots, n$.

$$\begin{aligned}
I_i J_i &= J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}}^2 \mathrm{d}s \\
&\ge J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \left[\chi_{\{\inf_{s \in (t_{i-1}, t_i]} X(s) > \tau\}} - \chi_{\{\tau < \inf_{s \in (t_{i-1}, t_i]} X(s) \le \tau'\}} \right] \mathrm{d}s \\
&\ge -J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \chi_{\{\tau < \inf_{s \in (t_{i-1}, t_i]} X(s) \le \tau'\}} \mathrm{d}s \\
&\quad + J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \chi_{\{\inf_{s \in (t_{i-1}, t_i]} X(s) > \tau\}} \chi_{\{X(s) > \tau\}} \mathrm{d}s \\
&= -J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \chi_{\{\tau < \inf_{s \in (t_{i-1}, t_i]} X(s) \le \tau'\}} \mathrm{d}s \\
&\quad + J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1} > \tau'\}} \left[\chi_{\{\inf_{s \in (t_{i-1}, t_i]} X(s) > \tau\}} - 1 \right] \chi_{\{X(s) > \tau\}} \mathrm{d}s \\
&\quad + J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \left[\chi_{\{X_{i-1} > \tau'\}} - 1 \right] \chi_{\{X(s) > \tau\}} \mathrm{d}s + J_i \int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X(s) > \tau\}} \mathrm{d}s.
\end{aligned}$$

Hence,

$$J_i\left(I_i - \int_{t_{i-1}}^{t_i} \log\sigma^2(X(s),\theta_1)\chi_{\{X(s)>\tau\}}\mathrm{d}s\right) \geq J_i\Xi_i^{(1)}, \tag{17}$$

where

$$\Xi_i^{(1)} = -\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau'\}}\mathrm{d}s \tag{18}$$

$$-\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau\}}\chi_{\{X(s)>\tau\}}\mathrm{d}s \tag{19}$$

$$-\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}\leq\tau'\}}\chi_{\{X(s)>\tau\}}\mathrm{d}s \tag{20}$$

$$+\int_{t_{i-1}}^{t_i} \left\{\log\sigma_{i-1}^2 - \log\sigma^2(X(s),\theta_1)\right\}\chi_{\{X(s)>\tau\}}\mathrm{d}s. \tag{21}$$

Next, noting that

$$\begin{aligned} I_iJ_i &= J_i\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)>\tau\}}\mathrm{d}s \\ &\quad + J_i\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau\}}\mathrm{d}s \\ &\leq J_i\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X(s)>\tau\}}\mathrm{d}s + J_i\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau\}}\mathrm{d}s, \end{aligned}$$

we obtain that

$$J_i\left(I_i - \int_{t_{i-1}}^{t_i} \log\sigma^2(X(s),\theta_1)\chi_{\{X(s)>\tau\}}\mathrm{d}s\right) \leq J_i\Xi_i^{(2)}, \tag{22}$$

where

$$\begin{aligned} \Xi_i^{(2)} &= \int_{t_{i-1}}^{t_i} \left\{\log\sigma_{i-1}^2 - \log\sigma^2(X(s),\theta_1)\right\}\chi_{\{X(s)>\tau\}}\mathrm{d}s \\ &\quad + \int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau\}}\mathrm{d}s. \end{aligned}$$

It follows from (17) and (22) that

$$\left|J_i\left(I_i - \int_{t_{i-1}}^{t_i} \log\sigma^2(X(s),\theta_1)\chi_{\{X(s)>\tau\}}\mathrm{d}s\right)\right| \leq \max\{|\Xi_i^{(1)}|, |\Xi_i^{(2)}|\}.$$

For the estimate of (18), we set $\tilde{\tau} = \tau' + h_n^\alpha$, where $\alpha \in (0, 1/2)$.

$$\begin{aligned} &E\left[\left|\int_{t_{i-1}}^{t_i} \log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau'\}}\mathrm{d}s\right|\right] \\ &\quad \leq E\left[\int_{t_{i-1}}^{t_i} \left|\log\sigma_{i-1}^2\right|\left\{\chi_{\{X_{i-1}>\tau'\}} - \chi_{\{X_{i-1}>\tilde{\tau}\}}\right\}\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]}X(s)\leq\tau'\}}\mathrm{d}s\right] \end{aligned}$$

$$
\begin{aligned}
&\quad + E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \chi_{\{X_{i-1}>\tilde{\tau}\}}\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]} X(s)\le\tau'\}} \mathrm{d}s\right]\\
&\le E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \chi_{\{\tau'<X_{i-1}\le\tilde{\tau}\}} \mathrm{d}s\right]\\
&\quad + E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \chi_{\{X_{i-1}>\tilde{\tau}\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]} X(s)\le\tau'\}} \mathrm{d}s\right]\\
&\le h_n C\left[P(\tau'<X_{i-1}\le\tilde{\tau})^{1/2} + P(\sup_{s\in(t_{i-1},t_i]} |X_{i-1}-X(s)|>h_n^{\alpha})^{1/2}\right] = o(h_n).
\end{aligned}
$$

Concerning the estimate of (19),

$$
\begin{aligned}
&E\left[\left|\int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]} X(s)\le\tau\}}\chi_{\{X(s)>\tau\}} \mathrm{d}s\right|\right]\\
&\le h_n C P(\sup_{s\in(t_{i-1},t_i]} |X_{i-1}-X(s)|>h_n^{\alpha})^{1/2} = o(h_n).
\end{aligned}
$$

In order to estimate (20), we set $\tau''=\tau-h_n^{\alpha}$, where $\alpha\in(0,1/2)$.

$$
\begin{aligned}
&E\left[\left|\int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1}\le\tau'\}}\chi_{\{X(s)>\tau\}} \mathrm{d}s\right|\right]\\
&\le E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \left\{\chi_{\{X_{i-1}\le\tau'\}} - \chi_{\{\tau''<X_{i-1}\le\tau'\}}\right\} \chi_{\{X(s)>\tau\}} \mathrm{d}s\right]\\
&\quad + E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \chi_{\{\tau''<X_{i-1}\le\tau'\}}\chi_{\{X(s)>\tau\}} \mathrm{d}s\right]\\
&\le E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \chi_{\{X_{i-1}\le\tau''\}}\chi_{\{X(s)>\tau\}} \mathrm{d}s\right] + E\left[\int_{t_{i-1}}^{t_i} \left|\log \sigma_{i-1}^2\right| \chi_{\{\tau''<X_{i-1}\le\tau'\}} \mathrm{d}s\right]\\
&\le h_n C\left[P(\sup_{s\in(t_{i-1},t_i]} |X(s)-X_{i-1}|>h_n^{\alpha})^{1/2} + P(\tau''<X_{i-1}\le\tau')^{1/2}\right] = o(h_n).
\end{aligned}
$$

As for the estimate of (21),

$$
E\left[\left|\int_{t_{i-1}}^{t_i} \left\{\log \sigma_{i-1}^2 - \log \sigma^2(X(s),\theta_1)\right\} \mathrm{d}s\right|\right] \le C h_n^{3/2} = o(h_n).
$$

Thus, we obtain

$$
E\left[|\Xi_i^{(1)}|\right] = o(h_n). \tag{23}
$$

Moreover,

$$
\begin{aligned}
E\left[|\Xi_i^{(2)}|\right] &\le E\left[\left|\int_{t_{i-1}}^{t_i} \left\{\log \sigma_{i-1}^2 - \log \sigma^2(X(s),\theta_1)\right\} \chi_{\{X(s)>\tau\}} \mathrm{d}s\right|\right]\\
&\quad + E\left[\left|\int_{t_{i-1}}^{t_i} \log \sigma_{i-1}^2 \chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]} X(s)\le\tau\}} \mathrm{d}s\right|\right]\\
&= o(h_n). \qquad (24)
\end{aligned}
$$

It follows from (23) and (24) that

$$E\left[\left|J_i\int_{t_{i-1}}^{t_i}\left\{\log\sigma_{i-1}^2\chi_{\{X_{i-1}>\tau'\}}-\log\sigma^2(X(s),\theta_1)\chi_{\{X(s)>\tau\}}\right\}\mathrm{d}s\right|\right]=o(h_n). \tag{25}$$

For the case that $\log\sigma_{i-1}^2<0$, set $\tilde{J}_i=\chi_{\{\log\sigma_{i-1}^2<0\}}$ for $i=1,\ldots,n$. In a similar way as the upper bound of I_iJ_i together with (16),

$$\begin{aligned}
I_i\tilde{J}_i &= -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\mathrm{d}s=-\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}^2\mathrm{d}s\\
&\le -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\left[\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)>\tau\}}-\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]}X(s)\le\tau'\}}\right]\mathrm{d}s\\
&\le \tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]}X(s)\le\tau'\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)>\tau\}}\chi_{\{X(s)>\tau\}}\mathrm{d}s\\
&= \tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\tau<\inf_{s\in(t_{i-1},t_i]}X(s)\le\tau'\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\left[\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)>\tau\}}-1\right]\chi_{\{X(s)>\tau\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\left[\chi_{\{X_{i-1}>\tau'\}}-1\right]\chi_{\{X(s)>\tau\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2+\log\sigma^2(X(s),\theta_1))\chi_{\{X(s)>\tau\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma^2(X(s),\theta_1))\chi_{\{X(s)>\tau\}}\mathrm{d}s.
\end{aligned}$$

Therefore,

$$\tilde{J}_i\left(I_i-\int_{t_{i-1}}^{t_i}\log\sigma^2(X(s),\theta_1)\chi_{\{X(s)>\tau\}}\mathrm{d}s\right)\le\tilde{J}_i\Xi_i^{(1)}. \tag{26}$$

Moreover, since

$$\begin{aligned}
I_i\tilde{J}_i &= -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)>\tau\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)\le\tau\}}\mathrm{d}s\\
&\ge -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2+\log\sigma^2(X(s),\theta_1))\chi_{\{X(s)>\tau\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma^2(X(s),\theta_1))\chi_{\{X(s)>\tau\}}\mathrm{d}s\\
&\quad -\tilde{J}_i\int_{t_{i-1}}^{t_i}(-\log\sigma_{i-1}^2)\chi_{\{X_{i-1}>\tau'\}}\chi_{\{\inf_{s\in(t_{i-1},t_i]}X(s)\le\tau\}}\mathrm{d}s,
\end{aligned}$$

one has that

$$\tilde{J}_i \left(I_i - \int_{t_{i-1}}^{t_i} \log \sigma^2(X(s), \theta_1) \chi_{\{X(s)>\tau\}} \mathrm{d}s \right) \geq \tilde{J}_i \Xi_i^{(2)}. \tag{27}$$

It follows from (26) and (27) that

$$\left| \tilde{J}_i \left(I_i - \int_{t_{i-1}}^{t_i} \log \sigma^2(X(s), \theta_1) \chi_{\{X(s)>\tau\}} \mathrm{d}s \right) \right| \leq \max\{|\Xi_i^{(1)}|, |\Xi_i^{(2)}|\}.$$

By (23) and (24), one has that

$$E\left[\left| (1 - J_i) \int_{t_{i-1}}^{t_i} \left\{ \log \sigma_{i-1}^2 \chi_{\{X_{i-1}>\tau'\}} - \log \sigma^2(X(s), \theta_1) \chi_{\{X(s)>\tau\}} \right\} \mathrm{d}s \right| \right] = o(h_n). \tag{28}$$

Therefore, by (25) and (28),

$$E\left[\left| \int_{t_{i-1}}^{t_i} \left\{ \log \sigma_{i-1}^2 \chi_{\{X_{i-1}>\tau'\}} - \log \sigma^2(X(s), \theta_1) \chi_{\{X(s)>\tau\}} \right\} \mathrm{d}s \right| \right] = o(h_n)$$

and consequently,

$$\left| \frac{1}{nh_n} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \left\{ \log \sigma_{i-1}^2 \chi_{\{X_{i-1}>\tau'\}} - \log \sigma^2(X(s), \theta_1) \chi_{\{X(s)>\tau\}} \right\} \mathrm{d}s \right| = o_p(1). \tag{29}$$

Moreover, by the ergodic theorem,

$$\frac{1}{nh_n} \int_0^{nh_n} \log \sigma^2(X(s), \theta_1) \chi_{\{X(s)>\tau\}} \mathrm{d}s \xrightarrow{p} \int_{\mathbf{R}} \log \sigma^2(x, \theta_1) \chi_{\{x>\tau\}} \nu_{\theta_0}(\mathrm{d}x),$$

which completes the proof of (14). For the proof of (15), we set

$$\Xi_i = \frac{1}{nh_n} \frac{(\Delta_i X)^2}{\sigma_{i-1}^2} \chi_{\{X_{i-1}>\tau'\}}.$$

By Lemma 9 of [3], it is enough to show that

$$\sum_{i=1}^{n} \mathrm{E}_{\theta_0} \{\Xi_i | \mathcal{F}_{i-1}\} \xrightarrow{p} \int_{\mathbf{R}} \frac{\sigma^2(x, \theta_{1,0})}{\sigma^2(x, \theta_1)} \chi_{\{x>\tau\}} \nu_{\theta_0}(\mathrm{d}x), \tag{30}$$

$$\sum_{i=1}^{n} \mathrm{E}_{\theta_0} \left\{ (\Xi_i)^2 | \mathcal{F}_{i-1} \right\} \xrightarrow{p} 0, \tag{31}$$

where $\mathcal{F}_{i-1}$ denotes the *history* up to the time t_{i-1}. In order to evaluate $\mathrm{E}_{\theta_0} \left\{ (\Delta_i X)^2 | \mathcal{F}_{i-1} \right\}$, we can use a well-known Itô–Taylor expansion:

$$\begin{aligned} &\mathrm{E}_{\theta_0}(\phi(X_i, X_{i-1}) | \mathcal{F}_{i-1}) \\ &\quad = \phi(X_{i-1}, X_{i-1}) + h_n L_{\theta_0} \phi(X_{i-1}, X_{i-1}) + \frac{1}{2} h_n^2 L_{\theta_0}^2 \phi(X_{i-1}, X_{i-1}) \\ &\qquad + \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t} E_{\theta_0} \left\{ L_{\theta_0}^2 \phi(X(s), X_{i-1}) - L_{\theta_0}^2 \phi(X_{i-1}, X_{i-1}) | \mathcal{F}_{i-1} \right\} \mathrm{d}s \, \mathrm{d}t \end{aligned}$$

for appropriate functions $\phi(x, y)$, where $L_\theta \phi(x, y) = \frac{1}{2}\sigma^2(x, \theta_1)\frac{\partial^2}{\partial x^2}\phi(x, y) + b(x, \theta_2)\frac{\partial}{\partial x}\phi(x, y)$. Hence

$$\mathrm{E}_{\theta_0}\left\{(\Delta_i X)^2 \,|\mathcal{F}_{i-1}\right\} = h_n \sigma_{i-1}^{*2} + R(h_n^2, X_{i-1}), \tag{32}$$

where $R(\cdot,\cdot)$ is defined in (2). Thus, as in the proof of (14),

$$\begin{aligned}\sum_{i=1}^n E_{\theta_0}\{\Xi_i|\mathcal{F}_{i-1}\} &= \frac{1}{n}\sum_{i=1}^n \frac{(\sigma_{i-1}^*)^2}{\sigma_{i-1}^2}\chi_{\{X_{i-1}>\tau'\}} + \frac{h_n}{n}\sum_{i=1}^n R(1, X_{i-1}) \\ &\xrightarrow{p} \int_{\mathbf{R}} \frac{\sigma^2(x, \theta_{1,0})}{\sigma^2(x, \theta_1)}\chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x)\end{aligned}$$

and in a similar way, we can show (31). This completes the proof of (7).

Next, we see that G attains its minimum only at $\theta_{1,0}$ by noting that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\log x + \frac{a}{x}\right) = \frac{1}{x} - \frac{a}{x^2} = \frac{x-a}{x^2}.$$

Hence, for any $\epsilon > 0$, $\inf_{\theta_1:|\theta_1-\theta_{1,0}|\geq\epsilon} G(\theta_1) > G(\theta_{1,0})$. This implies that if $|\theta_1 - \theta_{1,0}| \geq \epsilon$, then $G(\theta_1) > G(\theta_{1,0}) + \eta$ for some $\eta > 0$. Therefore,

$$\begin{aligned}P\left(|\hat{\theta}_{1,n} - \theta_{1,0}| \geq \epsilon\right) &\leq P\left(G(\hat{\theta}_{1,n}) > G(\theta_{1,0}) + \eta\right) \\ &\leq 2P\left(\sup_{\theta_1}\left|\frac{1}{n}g_n(\theta_1) - G(\theta_1)\right| > \eta/3\right) \\ &\quad + P\left(\frac{1}{n}g_n(\hat{\theta}_{1,n}) - \frac{1}{n}g_n(\theta_{1,0}) > \eta/3\right).\end{aligned} \tag{33}$$

By using (7), the probability of (33) converges to 0. Furthermore, it follows from (5) that

$$P\left(\frac{1}{n}g_n(\hat{\theta}_{1,n}) - \frac{1}{n}g_n(\theta_{1,0}) > \eta/3\right) \leq P\left(\frac{1}{n}g_n(\hat{\theta}_{1,n}) > \frac{1}{n}g_n(\theta_{1,0})\right) \to 0.$$

This competes the proof. □

Proof of Theorem 3.2. We need to prove that

$$\sup_{\theta_2}\left|\frac{1}{nh_n}\left(\ell_n(\hat{\theta}_{1,n}, \theta_2) - \ell_n(\hat{\theta}_{1,n}, \theta_{2,0})\right) - L(\theta_2)\right| \xrightarrow{p} 0, \tag{34}$$

where

$$L(\theta_2) = \int_{\mathbf{R}}\left(\frac{b(x, \theta_2) - b(x, \theta_{2,0})}{\sigma(x, \theta_{1,0})}\right)^2 \chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x).$$

An easy computation together with (8) yields that

$$\frac{1}{nh_n}\left(\ell_n(\hat{\theta}_{1,n}, \theta_2) - \ell_n(\hat{\theta}_{1,n}, \theta_{2,0})\right) = \psi_{1,n}(\theta_2) + \psi_{2,n}(\theta_2) + \psi_{3,n}(\theta_2) + R_n(\theta_2),$$

where $\hat{\sigma}_i = \sigma(X_i, \hat{\theta}_{1,n})$,

$$\psi_{1,n}(\theta_2) = -\frac{2}{nh_n}\sum_{i=1}^n \frac{(b_{i-1} - b_{i-1}^*)\int_{t_{i-1}}^{t_i}\sigma(X(s), \theta_{1,0})\mathrm{d}W_s}{\hat{\sigma}_{i-1}^2}\chi_{\{X_{i-1}>\tau'\}},$$

$$\psi_{2,n}(\theta_2) = -\frac{2}{nh_n}\sum_{i=1}^{n}\frac{(b_{i-1}-b^*_{i-1})\int_{t_{i-1}}^{t_i} b(X(s),\theta_{2,0})\mathrm{d}s}{\hat{\sigma}^2_{i-1}}\chi_{\{X_{i-1}>\tau'\}},$$
$$\psi_{3,n}(\theta_2) = \frac{1}{n}\sum_{i=1}^{n}\frac{b^2_{i-1}-b^{*2}_{i-1}}{\hat{\sigma}^2_{i-1}}\chi_{\{X_{i-1}>\tau'\}},$$
$$R_n(\theta_2) = -\frac{1}{nh_n}\sum_{i=1}^{n}\left\{\frac{(\Delta_i X - b_{i-1}h_n)^2}{\hat{\sigma}^2_{i-1}h_n} - \frac{(\Delta_i X - b^*_{i-1}h_n)^2}{\hat{\sigma}^2_{i-1}h_n}\right\}\chi_{\{X_{i-1}>\tau', X_i\le\tau\}}.$$

We first estimate $R_n(\theta_2)$.

$$\begin{aligned}E\left[\sup_{\theta_2}|R_n(\theta_2)|\right] &\le \frac{1}{nh_n}\sum_{i=1}^{n}E\left[\sup_{\theta_2}\left|\frac{(\Delta_i X - b_{i-1}h_n)^2-(\Delta_i X - b^*_{i-1}h_n)^2}{\hat{\sigma}^2_{i-1}h_n}\right|^2\right]^{1/2}\\ &\quad\times P(X_{i-1}>\tau', X_i\le\tau)^{1/2}\\ &\le \frac{1}{h_n^{1/2}}\times C\left(\frac{h_n^\alpha}{\tau'-\tau}\right)^{k/2}(h_n^{1/2-\alpha})^{k/2}\\ &= O\left(h_n^{k/4-\alpha k/2-1/2}\right)\to 0,\end{aligned}$$

where we took $k > 2/(1-2\alpha)$ in (11). This yields that $\sup_{\theta_2}|R_n(\theta_2)| = o_p(1)$. Next, $\psi_{2,n}(\theta_2)$ can be rewritten as

$$\psi_{2,n}(\theta_2) = \psi^{(1)}_{2,n}(\theta_2)+\psi^{(2)}_{2,n}(\theta_2)+\psi^{(3)}_{2,n}(\theta_2),$$

where

$$\psi^{(1)}_{2,n}(\theta_2) = -\frac{2}{n}\sum_{i=1}^{n}\frac{(b_{i-1}-b^*_{i-1})b^*_{i-1}}{\sigma^{*2}_{i-1}}\chi_{\{X_{i-1}>\tau'\}},$$
$$\psi^{(2)}_{2,n}(\theta_2) = -\frac{2}{nh_n}\sum_{i=1}^{n}\frac{(b_{i-1}-b^*_{i-1})\int_{t_{i-1}}^{t_i}\left\{b(X(s),\theta_{2,0})-b^*_{i-1}\right\}\mathrm{d}s}{\sigma^{*2}_{i-1}}\chi_{\{X_{i-1}>\tau'\}},$$
$$\psi^{(3)}_{2,n}(\theta_2) = -\frac{2}{nh_n}\sum_{i=1}^{n}(b_{i-1}-b^*_{i-1})\int_{t_{i-1}}^{t_i}b(X(s),\theta_{2,0})\mathrm{d}s\left(\frac{1}{\hat{\sigma}^2_{i-1}}-\frac{1}{\sigma^{*2}_{i-1}}\right)\chi_{\{X_{i-1}>\tau'\}}.$$

By noting that for $p > 1$ and $K > 0$,

$$\left\|\int_{t_{i-1}}^{t_i}\left\{b(X(s),\theta_{2,0})-b^*_{i-1}\right\}\mathrm{d}s\right\|_p \le Ch_n^{3/2},$$
$$\left|\frac{1}{\hat{\sigma}^2_{i-1}}-\frac{1}{\sigma^{*2}_{i-1}}\right| \le C\left|\sigma^{*2}_{i-1}-\hat{\sigma}^2_{i-1}\right| \le |\hat{\theta}_{1,n}-\theta_{1,0}|K(1+|X_{i-1}|)^K,$$

one has that for $1 < p, q < \infty$ with $1/p+1/q = 1$,

$$\begin{aligned}E\left[\sup_{\theta_2}\left|\psi^{(2)}_{2,n}(\theta_2)\right|\right] &\le \frac{1}{nh_n}\sum_{i=1}^{n}\left\|\sup_{\theta_2}\left|\frac{b_{i-1}-b^*_{i-1}}{\sigma^{*2}_{i-1}}\right|\right\|_p\\ &\quad\times\left\|\int_{t_{i-1}}^{t_i}\left\{b(X(s),\theta_{2,0})-b^*_{i-1}\right\}\mathrm{d}s\right\|_q\\ &\le \frac{C}{nh_n}nh_n^{3/2}\to 0,\end{aligned}$$

and

$$\begin{aligned}\sup_{\theta_2}\left|\psi_{2,n}^{(3)}(\theta_2)\right| &\le \left|\hat{\theta}_{1,n}-\theta_{1,0}\right|\frac{K}{nh_n}\sum_{i=1}^{n}\sup_{\theta_2}\left|b_{i-1}-b_{i-1}^*\right| \\ &\quad\times\left|\int_{t_{i-1}}^{t_i} b(X(s),\theta_{2,0})\mathrm{d}s\right|(1+|X_{t_{i-1}}|)^K \\ &= o_p(1)\times O_p(1)=o_p(1).\end{aligned}$$

As in the proof of the uniform convergence of (9),

$$\sup_{\theta_2}\left|\psi_{2,n}^{(1)}(\theta_2)+2\int\frac{(b(x,\theta_2)-b(x,\theta_{2,0}))b(x,\theta_{2,0})}{\sigma^2(x,\theta_{1,0})}\chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x)\right|=o_p(1).$$

Furthermore, since one estimates

$$\begin{aligned}&\sup_{\theta_2}\left|\psi_{3,n}(\theta_2)-\frac{1}{n}\sum_{i=1}^{n}\frac{b_{i-1}^2-b_{i-1}^{*2}}{\sigma_{i-1}^{*2}}\chi_{\{X_{i-1}>\tau'\}}\right| \\ &\le\left|\hat{\theta}_{1,n}-\theta_{1,0}\right|\frac{K}{n}\sum_{i=1}^{n}\sup_{\theta_2}\left|b_{i-1}^2-b_{i-1}^{*2}\right|(1+|X_{t_{i-1}}|)^K \\ &= o_p(1)\times O_p(1)=o_p(1),\end{aligned}$$

we obtain

$$\sup_{\theta_2}\left|\psi_{3,n}(\theta_2)-\int\frac{b(x,\theta_2)^2-b(x,\theta_{2,0})^2}{\sigma^2(x,\theta_{1,0})}\chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x)\right|=o_p(1).$$

Therefore, we see that

$$\sup_{\theta_2}\left|\psi_{2,n}(\theta_2)+\psi_{3,n}(\theta_2)-L(\theta_2)\right|=o_p(1).$$

To estimate $\psi_1(\theta_1)$, we consider the following process

$$M_n(\theta)=\int_0^{nh_n}\sum_{i=1}^{n}\frac{(b_{i-1}-b_{i-1}^*)\sigma(X(s),\theta_{1,0})}{nh_n\sigma_{i-1}^2}\mathbf{1}_i(s)\mathrm{d}W_s,$$

where $\mathbf{1}_i(s)=\chi_{\{X_{i-1}>\tau'\}}\chi_{\{t_{i-1}\le s\le t_i\}}$. We will prove the following: there exists a constant $\beta>2$ such that for any θ and θ',

$$M_n(\theta)\xrightarrow{p}0, \tag{35}$$

$$E|M_n(\theta)|^\beta\le C, \tag{36}$$

$$\mathrm{E}\left|M_n(\theta)-M_n(\theta')\right|^\beta\le C|\theta-\theta'|^\beta, \tag{37}$$

where C is a constant independent of θ, θ' and n. If (35)–(37) are satisfied, by Theorem 20 in the Appendix of [8] or Lemma 3.1 of [15], we can show that $\sup_\theta|M_n(\theta)|\xrightarrow{p}0$. In fact, (36) and (37) ensure that the family of distributions of $\{M_n(\cdot)\}$ on $C(\Theta)$ with sup-norm is tight. Hence, if (35)–(37) are shown, one can prove that

$$\sup_{\theta_2}\left|\psi_{1,n}(\theta_2)\right|=2\sup_{\theta_2}\left|M_n(\hat{\theta}_{1,n},\theta_2)\right|\le 2\sup_{\theta}|M_n(\theta)|\xrightarrow{p}0. \tag{38}$$

The proof of (37) is as follows. Let us define

$$f_{i-1}(\theta,\theta') = \frac{b_{i-1}(\theta_2)-b^*_{i-1}}{\sigma^2_{i-1}(\theta_1)} - \frac{b_{i-1}(\theta'_2)-b^*_{i-1}}{\sigma^2_{i-1}(\theta'_1)}$$
$$= \frac{b_{i-1}(\theta_2)-b_{i-1}(\theta'_2)}{\sigma^2_{i-1}(\theta'_1)} + (b_{i-1}(\theta_2)-b^*_{i-1})\left(\frac{1}{\sigma^2_{i-1}(\theta_1)} - \frac{1}{\sigma^2_{i-1}(\theta'_1)}\right).$$

By the Burkholder–Davis–Gundy inequality and the Jensen inequality,

$$E|M_n(\theta)-M_n(\theta')|^\beta$$
$$= \frac{1}{(nh_n)^\beta}\mathrm{E}\left|\int_0^{nh_n}\sum_{i=1}^n f_{i-1}(\theta,\theta')\sigma(X(s),\theta_{1,0})\mathbf{1}_i(s)\mathrm{d}W_s\right|^\beta$$
$$\le \frac{C_\beta}{(nh_n)^\beta}\mathrm{E}\left(\sum_{i=1}^n\int_0^{nh_n}\left(f_{i-1}(\theta,\theta')\sigma(X(s),\theta_{1,0})\right)^2\mathbf{1}_i(s)\mathrm{d}s\right)^{\frac{\beta}{2}}$$
$$\le \frac{C_\beta}{(nh_n)^\beta}n^{\beta/2-1}\sum_{i=1}^n\mathrm{E}\left(\int_{t_{i-1}}^{t_i}\left(f_{i-1}(\theta,\theta')\sigma(X(s),\theta_{1,0})\right)^2\mathrm{d}s\right)^{\frac{\beta}{2}}$$
$$\le \frac{C_\beta}{(nh_n)^\beta}(nh_n)^{\beta/2-1}\sum_{i=1}^n\mathrm{E}\left(\int_{t_{i-1}}^{t_i}\left|f_{i-1}(\theta,\theta')\sigma(X(s),\theta_{1,0})\right|^\beta\mathrm{d}s\right).$$

Moreover, it follows from A5 to A6,

$$|f_{i-1}(\theta,\theta')|^\beta \le K(1+|X_{i-1}|)^K|\theta-\theta'|^\beta,$$

which completes the proof of (37). In a similar way, we can show (36). For the proof of (35), we set $g_i = (b_i - b_i^*)/\sigma_i^2$ and

$$\mathrm{E}\,(M_n(\theta))^2 \le \frac{1}{n^2h_n^2}\sum_{i=1}^n\mathrm{E}\left\{\int_{t_{i-1}}^{t_i} g^2_{i-1}\sigma^2(X(s),\theta_{1,0})\mathrm{d}s\right\}\to 0,$$

which completes the proof of (35). Thus, we have (38) and this completes the proof of (34). Finally, note that for any $\epsilon > 0$, $\inf_{\theta_2:|\theta_2-\theta_{2,0}|\ge\epsilon} L(\theta_2) > 0$ because L attains its minimum only at $\theta_{2,0}$. As in the proof of Theorem 3.1, we can show the consistency of $\hat\theta_{2,n}$. This completes the proof. □

Proof of Theorem 3.3. First, we study the asymptotic normality of the score function. Let

$$\mathcal{L}_n = \begin{pmatrix} -\frac{1}{\sqrt{n}}\partial_{\theta_1}g_n(\theta_{1,0}) \\ -\frac{1}{\sqrt{nh_n}}\partial_{\theta_2}\ell_n(\hat\theta_{1,n},\theta_{2,0})\end{pmatrix}, \qquad \bar{\mathcal{L}}_n = \begin{pmatrix} -\frac{1}{\sqrt{n}}\partial_{\theta_1}\bar g_n(\theta_{1,0}) \\ -\frac{1}{\sqrt{nh_n}}\partial_{\theta_2}\bar\ell_n(\theta_0)\end{pmatrix},$$

where

$$\bar g_n(\theta_1) = \sum_{i=1}^n g(i,i-1;\theta_1)\chi_{\{X_{i-1}>\tau'\}},$$
$$\bar\ell_n(\theta) = \sum_{i=1}^n \ell(i,i-1;\theta)\chi_{\{X_{i-1}>\tau'\}}.$$

In order to show that $\mathcal{L}_n - \bar{\mathcal{L}}_n = o_p(1)$, it is sufficient to show that

$$A_n := \frac{1}{\sqrt{n}} \left(\partial_{\theta_1} g_n(\theta_{1,0}) - \partial_{\theta_1} \bar{g}_n(\theta_{1,0})\right) = o_p(1), \tag{39}$$

$$B_n := \frac{1}{\sqrt{nh_n}} \left(\partial_{\theta_2} \ell_n(\hat{\theta}_{1,n}, \theta_{2,0}) - \partial_{\theta_2} \bar{\ell}_n(\theta_0)\right) = o_p(1). \tag{40}$$

For the proof of (39), one estimates

$$\begin{aligned} E|A_n| &\le \frac{1}{\sqrt{n}} \sum_{i=1}^n E\left|\partial_{\theta_1} g(i-1, i; \theta_{1,0}) \chi_{\{X_{i-1} > \tau', X_i \le \tau\}}\right| \\ &\le \frac{C}{\sqrt{n}} \sum_{i=1}^n \left\| \frac{\partial_{\theta_1} \sigma^*_{i-1}}{\sigma^*_{i-1}} \left(1 - \frac{(\Delta_i X)^2}{h_n \sigma^{*2}_{i-1}}\right) \right\|_2 \times O\left(h_n^{(1/4-\alpha/2)k}\right) \\ &\le C\sqrt{n} h_n \times O\left(h_n^{(1/4-\alpha/2)k-1}\right) \to 0, \end{aligned}$$

where we took $k > 4/(1-2\alpha)$ in (11). For the proof of (40), one has that for $\epsilon > 0$,

$$\begin{aligned} |B_n| \chi_{\{|\hat{\theta}_{1,n} - \theta_{1,0}| < \epsilon\}} \le{} & \frac{1}{\sqrt{nh_n}} \sum_{i=1}^n \sup_{\theta_1} \left|\partial_{\theta_1} \partial_{\theta_2} \ell(i, i-1; \theta_1, \theta_{2,0})\right| \left|\hat{\theta}_{1,n} - \theta_{1,0}\right| \\ & + \frac{1}{\sqrt{nh_n}} \sum_{i=1}^n \left|\partial_{\theta_2} \ell(i, i-1; \theta_0) \chi_{\{X_{i-1} > \tau', X_i \le \tau\}}\right|. \end{aligned}$$

As in the proof of (39), $\frac{1}{\sqrt{nh_n}} \sum_{i=1}^n \left|\partial_{\theta_2} \ell(i, i-1; \theta_0) \chi_{\{X_{i-1} > \tau', X_i \le \tau\}}\right| = o_p(1)$. Next, letting $f_{i-1}(\theta_1) = \frac{\partial_{\theta_2} b^*_{i-1} \partial_{\theta_1} \sigma_{i-1}}{\sigma^3_{i-1}}$, we estimate that for $l \ge 1$

$$\begin{aligned} & E\left|\frac{1}{\sqrt{nh_n}} \sum_{i=1}^n \sup_{\theta_1} \left|\partial_{\theta_1} \partial_{\theta_2} \ell(i, i-1; \theta_1, \theta_{2,0})\right|\right|^{2l} \\ &\le \frac{C}{(nh_n)^l} E\left[\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \sup_{\theta_1} |f_{i-1}(\theta_1)| \sigma(X(s), \theta_{1,0}) \mathrm{d}W_s\right]^{2l} \\ &\quad + \frac{C}{(nh_n)^l} E\left[\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \sup_{\theta_1} |f_{i-1}(\theta_1)| (b(X(s), \theta_{2,0}) - b^*_{i-1}) \mathrm{d}s\right]^{2l} \\ &\le \frac{C}{(nh_n)^l} (nh_n)^{l-1} \sum_{i=1}^n E\left[\int_{t_{i-1}}^{t_i} \sup_{\theta_1} |f_{i-1}(\theta_1)|^{2l} \sigma^{2l}(X(s), \theta_{1,0}) \mathrm{d}s\right] \\ &\quad + \frac{C}{(nh_n)^l} (nh_n)^{2l-1} \sum_{i=1}^n E\left[\int_{t_{i-1}}^{t_i} \sup_{\theta_1} |f_{i-1}(\theta_1)|^{2l} (b(X(s), \theta_{2,0}) - b^*_{i-1})^{2l} \mathrm{d}s\right] \\ &= O(1). \end{aligned}$$

Consequently, one has that $|B_n| = o_p(1)$.

Next, we will prove that

$$\bar{\mathcal{L}}_n \xrightarrow{d} N(0, 4\Sigma). \tag{41}$$

Let

$$\xi_i^{(1)} = \frac{1}{\sqrt{n}} \partial_{\theta_1} \ell(i, i-1; \theta_{1,0}) \chi_{\{X_{i-1} > \tau'\}} = \frac{2}{\sqrt{n}} \frac{\partial_{\theta_1} \sigma_{i-1}^*}{\sigma_{i-1}^*} \left(1 - \frac{(\Delta_i X)^2}{h_n \sigma_{i-1}^{*2}}\right) \chi_{\{X_{i-1} > \tau'\}},$$

$$\begin{aligned} \xi_i^{(2)} &= \frac{1}{\sqrt{nh_n}} \partial_{\theta_2} \ell(i, i-1; \theta_0) \chi_{\{X_{i-1} > \tau'\}} \\ &= -\frac{2}{\sqrt{nh_n}} \left\{ \partial_{\theta_2} b_{i-1}^* \frac{\Delta_i X - b_{i-1}^* h_n}{\sigma_{i-1}^{*2}} \right\} \chi_{\{X_{i-1} > \tau'\}}, \end{aligned}$$

$$I(\theta_0) = \begin{pmatrix} I^{(1,1)}(\theta_{1,0}) & 0 \\ 0 & I^{(2,2)}(\theta_0) \end{pmatrix} := 4\Sigma.$$

In order to obtain (41), by the combination of Theorems 3.2 and 3.4 of [5], it is enough to prove the following convergences.

$$\sum_{i=1}^n \mathrm{E}_{\theta_0} \left\{ \xi_i^{(m)} | \mathcal{F}_{i-1} \right\} \xrightarrow{p} 0, \quad m = 1, 2, \tag{42}$$

$$\sum_{i=1}^n \left| \mathrm{E}_{\theta_0} \left\{ \xi_i^{(m)} | \mathcal{F}_{i-1} \right\} \right|^2 \xrightarrow{p} 0, \quad m = 1, 2, \tag{43}$$

$$\sum_{i=1}^n \mathrm{E}_{\theta_0} \left\{ \left(\xi_i^{(m)}\right)^2 | \mathcal{F}_{i-1} \right\} \xrightarrow{p} I^{(m,m)}, \quad m = 1, 2, \tag{44}$$

$$\sum_{i=1}^n \mathrm{E}_{\theta_0} \left\{ \xi_i^{(1)} \xi_i^{(2)} | \mathcal{F}_{i-1} \right\} \xrightarrow{p} 0, \tag{45}$$

$$\sum_{i=1}^n \mathrm{E}_{\theta_0} \left\{ \left(\xi_i^{(m)}\right)^4 | \mathcal{F}_{i-1} \right\} \xrightarrow{p} 0, \quad m = 1, 2. \tag{46}$$

For the proof of (42), by using the Itô–Taylor expansion and (32), one has

$$\sum_{i=1}^n \mathrm{E}_{\theta_0} \left\{ \xi_i^{(1)} | \mathcal{F}_{i-1} \right\} = \sqrt{nh_n^2} \cdot \frac{1}{n} \sum_{i=1}^n R(1, X_{i-1}) \xrightarrow{p} 0.$$

Moreover, since

$$\mathrm{E}_{\theta_0}(X_i - X_{i-1} | \mathcal{F}_{i-1}) = h_n b_{i-1}^* + R(h_n^2, X_{i-1}),$$

we have

$$\sum_{i=1}^n \mathrm{E}_{\theta_0}(\xi_i^{(2)} | \mathcal{F}_{i-1}) = \frac{-2\sqrt{nh_n^3}}{n} \sum_{i=1}^n R(1, X_{i-1}) \xrightarrow{p} 0.$$

This completes the proof of (42). In a similar way,

$$\sum_{i=1}^n \left| \mathrm{E}_{\theta_0} \left\{ \xi_i^{(1)} | \mathcal{F}_{i-1} \right\} \right|^2 = \frac{h_n^2}{n} \sum_{i=1}^n R(1, X_{i-1}) \xrightarrow{p} 0,$$

$$\sum_{i=1}^n \left| \mathrm{E}_{\theta_0}(\xi_i^{(2)} | \mathcal{F}_{i-1}) \right|^2 = \frac{h_n^3}{n} \sum_{i=1}^n R(1, X_{i-1}) \xrightarrow{p} 0,$$

which complete the proof of (43). For the proof of (44), noting that

$$E\left\{\left(1-\frac{(\Delta_i X)^2}{\sigma_{i-1}^{*2}h_n}\right)^2 |\mathcal{F}_{i-1}\right\} = 1+\frac{3h_n^2\sigma_{i-1}^{*4}+R(h_n^{5/2},X_{i-1})}{\sigma_{i-1}^{*4}h_n^2} - 2\frac{h_n\sigma_{i-1}^{*2}+R(h_n^2,X_{i-1})}{\sigma_{i-1}^{*2}h_n} = 2+\sqrt{h_n}R(1,X_{i-1}),$$

one has

$$\sum_{i=1}^n \mathrm{E}_{\theta_0}\left\{\left(\xi_i^{(1)}\right)^2 |\mathcal{F}_{i-1}\right\} = \sum_{i=1}^n \frac{4}{n}\left(\frac{\partial_{\theta_1}\sigma_{i-1}^*}{\sigma_{i-1}^*}\right)^2 (2+\sqrt{h_n}R(1,X_{i-1}))\chi_{\{X_{i-1}>\tau'\}} \xrightarrow{p} I^{(1,1)}(\theta_{1,0}),$$

which proves (44) for $m=1$. It follows from the Itô–Taylor expansion of $\mathrm{E}_{\theta_0}\{(X_i-X_{i-1}-h_nb_{i-1}^*)^2|\mathcal{F}_{i-1}\}$ that

$$\sum_{i=1}^n \mathrm{E}_{\theta_0}\left\{\left(\xi_i^{(2)}\right)^2 |\mathcal{F}_{i-1}\right\} = \frac{4}{nh_n}\sum_{i=1}^n \frac{(\partial_{\theta_2}b_{i-1}^*)^2}{\sigma_{i-1}^{*4}}(h_n\sigma_{i-1}^{*2}+R(h_n^2,X_{i-1}))\chi_{\{X_{i-1}>\tau'\}} \xrightarrow{p} I^{(2,2)}(\theta_0)$$

and (44) is proved. For the proof of (45), we consider

$$\xi_i^{(1)}\xi_i^{(2)} = -\frac{4}{n\sqrt{h_n}}\frac{\partial_{\theta_1}\sigma_{i-1}^*\partial_{\theta_2}b_{i-1}^*}{\sigma_{i-1}^{*3}}\left(1-\frac{(\Delta_iX)^2}{\sigma_{i-1}^{*2}h_n}\right)\left\{\Delta_iX-b_{i-1}^*h_n\right\}\chi_{\{X_{i-1}>\tau'\}}.$$

Since

$$\mathrm{E}_{\theta_0}\left\{(\Delta_iX)^2\left(\Delta_iX-b_{i-1}^*h_n\right)|\mathcal{F}_{i-1}\right\} = R(h_n^2,X_{i-1})$$

and

$$\mathrm{E}_{\theta_0}\left\{\Delta_iX-b_{i-1}^*h_n|\mathcal{F}_{i-1}\right\} = R(h_n^2,X_{i-1}),$$

one has

$$\sum_{i=1}^n \mathrm{E}_{\theta_0}\left\{\xi_i^{(1)}\xi_i^{(2)}|\mathcal{F}_{i-1}\right\} = -\frac{4}{n}\sum_{i=1}^n\frac{\partial_{\theta_1}\sigma_{i-1}^*\partial_{\theta_2}b_{i-1}^*}{\sigma_{i-1}^{*3}}\chi_{\{X_{i-1}>\tau'\}} \times\frac{1}{\sqrt{h_n}}\left(h_n^2R(1,X_{i-1})-\frac{h_nR(1,X_{i-1})}{\sigma_{i-1}^{*2}}\right) \xrightarrow{p} 0.$$

Hence (45) is proved. For the proof of (46), using the estimate that for $p\geq 1$,

$$\mathrm{E}_{\theta_0}\{(\Delta_iX)^{2p}|\mathcal{F}_{i-1}\} = h_n^pR(1,X_{i-1}),$$

one has

$$\mathrm{E}_{\theta_0}\left\{\sum_{i=1}^{n}\left(\xi_i^{(1)}\right)^4 |\mathcal{F}_{i-1}\right\} \le \frac{C'}{n}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial_{\theta_1}\sigma_{i-1}^*}{\sigma_{i-1}^*}\right)^4 \chi_{\{X_{i-1}>\tau'\}}\{1+R(1,X_{i-1})\} \xrightarrow{p} 0,$$

which completes the proof of (46) for $m=1$. For the case $m=2$, by using the following estimate

$$\mathrm{E}_{\theta_0}\left\{\left(\Delta_i X - b_{i-1}^* h_n\right)^4 |\mathcal{F}_{i-1}\right\} = h_n^2 R(1, X_{i-1}),$$

we have that

$$\sum_{i=1}^{n}\mathrm{E}_{\theta_0}\left\{\left(\xi_i^{(2)}\right)^4 |\mathcal{F}_{i-1}\right\} \le \frac{C'}{n}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial_{\theta_2} b_{i-1}^*}{\sigma_{i-1}^{*2}}\right)^4 \chi_{\{X_{i-1}>\tau'\}} R(1,X_{i-1}) \xrightarrow{p} 0.$$

Thus (46) is proved. This completes the proof of (41). It follows from (39)–(41) that

$$\mathcal{L}_n \xrightarrow{d} N(0, 4\Sigma). \tag{47}$$

Next we consider asymptotic properties of the observed information. Let

$$D_n(\theta) = \begin{pmatrix} \frac{1}{n}\partial_{\theta_1}^2 g_n(\theta_1) & 0 \\ 0 & \frac{1}{nh_n}\partial_{\theta_2}^2 \ell_n(\hat{\theta}_{1,n},\theta_2) \end{pmatrix}, \qquad D(\theta) = \begin{pmatrix} \bar{\mathcal{G}}(\theta_1) & 0 \\ 0 & \bar{\mathcal{L}}(\theta_2) \end{pmatrix},$$

where

$$\begin{aligned}\bar{\mathcal{G}}(\theta_1) &= 2\int_{\mathbf{R}} \frac{\partial_{\theta_1}^2\sigma(X,\theta_1)}{\sigma^3(x,\theta_1)}\left(\sigma^2(x,\theta_1)-\sigma^2(x,\theta_{1,0})\right)\chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x) \\ &\quad + 2\int_{\mathbf{R}} \frac{\left(3\sigma^2(x,\theta_{1,0})-\sigma^2(x,\theta_1)\right)\left(\partial_{\theta_1}\sigma(x,\theta_1)\right)^2}{\sigma^4(x,\theta_1)}\chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x), \\ \bar{\mathcal{L}}(\theta_2) &= 2\int_{\mathbf{R}}\left(\frac{\partial_{\theta_2} b(x,\theta_2)}{\sigma(x,\theta_{1,0})}\right)^2 \chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x) \\ &\quad - 2\int_{\mathbf{R}} \frac{(b(x,\theta_{2,0})-b(x,\theta_2))\partial_{\theta_2}^2 b(x,\theta_2)}{\sigma^2(x,\theta_{1,0})}\chi_{\{x>\tau\}}\nu_{\theta_0}(\mathrm{d}x).\end{aligned}$$

In order to prove that

$$\sup_{\theta} |D_n(\theta) - D(\theta)| = o_p(1), \tag{48}$$

it is sufficient to show that

$$\sup_{\theta_1}\left|\frac{1}{n}\partial_{\theta_1}^2 g_n(\theta_1) - \frac{1}{n}\partial_{\theta_1}^2 \bar{g}_n(\theta_1)\right| = o_p(1), \tag{49}$$

$$\sup_{\theta}\left|\frac{1}{nh_n}\partial_{\theta_2}^2 \ell_n(\theta) - \frac{1}{nh_n}\partial_{\theta_2}^2 \bar{\ell}_n(\theta)\right| = o_p(1), \tag{50}$$

$$\sup_{\theta_1}\left|\frac{1}{n}\partial_{\theta_1}^2 \bar{g}_n(\theta_1) - \bar{\mathcal{G}}(\theta_1)\right| = o_p(1), \tag{51}$$

$$\sup_{\theta_2}\left|\frac{1}{nh_n}\partial_{\theta_2}^2 \bar{\ell}_n(\hat{\theta}_{1,n},\theta_2) - \bar{\mathcal{L}}(\theta_2)\right| = o_p(1). \tag{52}$$

For the proof of (49), as in the proof of the uniform convergence of (10), one has that

$$\begin{aligned}
&\mathrm{E}\left\{\sup_{\theta_1}\left|\frac{1}{n}\sum_{i=1}^{n}\partial^2_{\theta_1} g(i,i-1;\theta_1)\chi_{\{X_{i-1}>\tau',X_i\le\tau\}}\right|\right\}\\
&\quad\le \frac{1}{n}\sum_{i=1}^{n}\left\|\sup_{\theta_1}\left|\partial^2_{\theta_1} g(i,i-1;\theta_1)\right|\right\|_2 P(X_{i-1}>\tau',X_i\le\tau)^{\frac{1}{2}}\to 0.
\end{aligned}$$

For the proof of (50), in a quite similar way as in the proof of (49), one has that

$$\begin{aligned}
&\mathrm{E}\left\{\sup_{\theta}\left|\frac{1}{nh_n}\sum_{i=1}^{n}\partial^2_{\theta_1}\ell(i,i-1;\theta)\chi_{\{X_{i-1}>\tau',X_i\le\tau\}}\right|\right\}\\
&\quad\le \frac{1}{nh_n}\sum_{i=1}^{n}\left\|\sup_{\theta}\left|\partial^2_{\theta_1}\ell(i,i-1;\theta)\right|\right\|_2 P(X_{i-1}>\tau',X_i\le\tau)^{\frac{1}{2}}\\
&\quad\le C\frac{1}{h_n^{1/2}}\times\left(\frac{h_n^{\alpha}}{\tau'-\tau}\right)^{k/2}(h_n^{1/2-\alpha})^{k/2}\\
&\quad= O\left(h_n^{k/4-\alpha k/2-1/2}\right)\to 0,
\end{aligned}$$

where we took $k>2/(1-2\alpha)$ in (11). For the proof of (51), we set

$$\begin{aligned}
\eta_i(\theta_1) &= \frac{1}{n}\partial^2_{\theta_1} g(i,i-1;\theta_1)\chi_{\{X_{i-1}>\tau'\}}\\
&= \frac{2}{nh_n\sigma^4_{i-1}}\\
&\quad\times\left\{(3(\Delta_i X)^2-h_n\sigma^2_{i-1})(\partial_{\theta_1}\sigma_{i-1})^2+\sigma_{i-1}(h_n\sigma^2_{i-1}-(\Delta_i X)^2)\partial^2_{\theta_1}\sigma_{i-1}\right\}\chi_{\{X_{i-1}>\tau'\}}.
\end{aligned}$$

It follows from standard arguments that

$$\sum_{i=1}^{n}\mathrm{E}_{\theta_0}\{\eta_i(\theta_1)|\mathcal{F}_{i-1}\}\xrightarrow{p}\bar{\mathcal{G}}(\theta_1),\qquad \sum_{i=1}^{n}\mathrm{E}_{\theta_0}\left\{(\eta_i(\theta_1))^2|\mathcal{F}_{i-1}\right\}\xrightarrow{p}0.$$

Therefore one has that for each θ_1,

$$\frac{1}{n}\partial^2_{\theta_1}\bar{g}_n(\theta_1)\xrightarrow{p}\bar{\mathcal{G}}(\theta_1).$$

It is easy to show that $\sup_n E[\sup_{\theta_1}|\frac{1}{n}\partial^3_{\theta_1}\bar{g}_n(\theta_1)|]<\infty$, which completes the proof of (51). For the proof of (52), we set

$$\frac{1}{nh_n}\partial^2_{\theta_2}\bar{\ell}_n(i,i-1;\hat{\theta}_{1,n},\theta_2)=\Xi_1(\theta_2)+\Xi_2(\theta_2)+\Xi_3(\theta_2),$$

where

$$\begin{aligned}
\Xi_1(\theta_2) &= \frac{2}{n}\sum_{i=1}^{n}\left\{\left(\frac{\partial_{\theta_2}b_{i-1}}{\hat{\sigma}_{i-1}}\right)^2-\frac{(b^*_{i-1}-b_{i-1})\partial^2_{\theta_2}b_{i-1}}{\hat{\sigma}^2_{i-1}}\right\}\chi_{\{X_{i-1}>\tau'\}},\\
\Xi_2(\theta_2) &= -\frac{2}{nh_n}\sum_{i=1}^{n}\frac{\partial^2_{\theta_2}b_{i-1}\int_{t_{i-1}}^{t_i}\{b(X(s),\theta_{2,0})-b^*_{i-1}\}\mathrm{d}s}{\hat{\sigma}^2_{i-1}}\chi_{\{X_{i-1}>\tau'\}},
\end{aligned}$$

$$\Xi_3(\theta_2) = -\frac{2}{nh_n}\sum_{i=1}^{n}\frac{\partial_{\theta_2}^2 b_{i-1}\int_{t_{i-1}}^{t_i}\sigma(X(s),\theta_{1,0})\mathrm{d}W_s}{\hat{\sigma}_{i-1}^2}\chi_{\{X_{i-1}>\tau'\}}.$$

In a quite similar way as in the proof of (34), one has that

$$\sup_{\theta_2}\left|\Xi_1(\theta_2)-\bar{\mathcal{L}}(\theta_2)\right| = o_p(1), \qquad \sup_{\theta_2}|\Xi_2(\theta_2)| = o_p(1), \qquad \sup_{\theta_2}|\Xi_3(\theta_2)| = o_p(1).$$

This completes the proof of (52). Thus, (48) is proved.

By the Taylor expansion, $\int_0^1 D_n(\theta_0 + u(\hat{\theta}_n - \theta_0))\mathrm{d}u S_n = \mathcal{L}_n$ on an event with probability tending to one, where $S_n = \begin{pmatrix}\sqrt{n}(\hat{\theta}_{1,n}-\theta_{1,0})\\ \sqrt{nh_n}(\hat{\theta}_{2,n}-\theta_{2,0})\end{pmatrix}$. It follows from (47) that

$$\mathcal{L}_n \xrightarrow{d} N(0, 4\Sigma). \tag{53}$$

By (48) and the continuity of $D(\theta)$ with respect to θ, one has

$$D_n(\theta_0) \xrightarrow{p} 2\Sigma, \tag{54}$$

$$\sup_{|\theta|\le\epsilon_n} |D_n(\theta_0+\theta) - D_n(\theta_0)| = o_p(1) \tag{55}$$

for any sequence ϵ_n of positive numbers tending to zero. By using (53)–(55), it is easy to obtain the desired result. This completes the proof. □

Acknowledgments

The authors wish to thank the referee, the associate editor and the editor for their valuable comments. This work was done during the period of stay of the first author at the ISM (Institute of Statistical Mathematics) and the University of Tokyo. The ISM is acknowledged with thanks. The work has been supported by the JSPS (Japan Society for the Promotion of Science) Program FY2006, grant ID No. S06174. Masayuki Uchida's research was partially supported by Grant-in-Aid for Scientific Research (C) No. 19540137 from the JSPS, and by Cooperative Research Program of the ISM. Nakahiro Yoshida's research was in part supported by Grant-in-Aid for Scientific Research (B) No. 19340021 from the JSPS, and by Cooperative Research Program of the ISM.

References

[1] P. Fearnhead, O. Papaspiliopoulos, G.O. Roberts, Particle filters for partially observed diffusions, J. Roy. Statist. Soc. Ser. B 70 (2008) 755–777.
[2] D. Florens-Zmirou, Approximate discrete time schemes for statistics of diffusion processes, Statistics 20 (1989) 547–557.
[3] V. Genon-Catalot, J. Jacod, On the estimation of the diffusion coefficient for multidimensional diffusion processes, Ann. Inst. Henri Poincaré Probab. Statist. 29 (1993) 119–151.
[4] P.E. Greenwood, L.M. Ward, W. Wefelmeyer, Statistical analysis of stochastic resonance in a simple setting, Phys. Rev. E 60 (2000) 4687–4695.
[5] P. Hall, C. Heyde, Martingale Limit Theory and its Applications, Academic Press, New York, 1980.
[6] S.M. Iacus, Statistical analysis of stochastic resonance with ergodic diffusion noise, Stochast. Stochast. Rep. 73 (2002) 271–285.
[7] S.M. Iacus, I. Negri, Estimating unobservable signal with Markovian noise induction, Statist. Methods Appl. 12 (2003) 153–167.
[8] I.A. Ibragimov, R.Z. Has'minskii, Statistical Estimation, Springer-Verlag, New York, 1981.
[9] M. Kessler, Estimation of an ergodic diffusion from discrete observations, Scand. J. Statist. 24 (1997) 211–229.

[10] J.R. Movellan, P. Mineiro, A Monte Carlo EM approach for partially observable diffusion processes: Theory and applications to neural networks, Neural Computation 14 (2002) 1507–1544.
[11] B.L.S. Prakasa Rao, Asymptotic theory for nonlinear least squares estimator for diffusion processes, Math. Operationsforsch. Statist. Ser. Statist. 14 (1983) 195–209.
[12] B.L.S. Prakasa Rao, Statistical Inference from Sampled Data for Stochastic Processes, in: Contemp. Math., vol. 80, Amer. Math. Soc., Providence, RI, 1988, pp. 249–284.
[13] G.O. Roberts, O. Stramer, Likelihood based inference for observed and partially observed diffusions, Biometrika 88 (2001) 603–621.
[14] J.L. Starck, H. Aussel, D. Elbaz, D. Fadda, C. Cesarsky, Faint source detection in ISOCAM images, Astron. Astrophys. Suppl. Ser. 138 (1999) 365–379.
[15] N. Yoshida, Asymptotic behavior of M-estimator and related random field for diffusion process, Ann. Inst. Statist. Math. 42 (1990) 221–251.
[16] N. Yoshida, Estimation for diffusion processes from discrete observations, J. Multivar. Anal. 41 (1992) 220–242.
[17] N. Yoshida, Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations, 2005. Ann. Inst. Statist. Math. (in press).
[18] Y. Zeng, A partially observed model for micromovement of asset prices with Bayes estimation via filtering, Math. Finance 13 (2003) 411–444.