Annals of the Institute of Statistical Mathematics (2021) 73:555–598
https://doi.org/10.1007/s10463-020-00768-x

Global jump filters and quasi-likelihood analysis for volatility

Haruhiko Inatsugu[1] · **Nakahiro Yoshida**[1]

Received: 1 July 2018 / Revised: 21 June 2020 / Accepted: 21 June 2020 / Published online: 16 January 2021

Abstract
We propose a new estimation scheme for estimation of the volatility parameters of a semimartingale with jumps based on a jump detection filter. Our filter uses all of the data to analyze the relative size of increments and to discriminate jumps more precisely. We construct quasi-maximum likelihood estimators and quasi-Bayesian estimators and show limit theorems for them including L^p-estimates of the error and asymptotic mixed normality based on the framework of the quasi-likelihood analysis. The global jump filters do not need a restrictive condition for the distribution of the small jumps. By numerical simulation, we show that our "global" method obtains better estimates of the volatility parameter than the previous "local" methods.

Keywords Volatility · Jump · Global filter · High-frequency data · Quasi-likelihood analysis · Stochastic differential equation · Order statistic · Asymptotic mixed normality · Polynomial-type large deviation · Moment

1 Introduction

We consider an m-dimensional semimartingale $Y = (Y_t)_{t\in[0,T]}$ admitting a decomposition

This work was in part supported by CREST JPMJCR14D7 Japan Science and Technology Agency; Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research No. 17H01702 (Scientific Research); and a Cooperative Research Program of the Institute of Statistical Mathematics.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10463-020-00768-x) contains supplementary material, which is available to authorized users.

✉ Nakahiro Yoshida
nakahiro@ms.u-tokyo.ac.jp

1 Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan

$$Y_t = Y_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma(X_s, \theta)\mathrm{d}w_s + J_t, \quad t \in [0, T] \tag{1}$$

on a stochastic basis $(\Omega, \mathcal{F}, \mathbf{F}, P)$ with a filtration $\mathbf{F} = (\mathcal{F}_t)_{t\in[0,T]}$. Here, $b = (b_t)_{t\in[0,T]}$ is an m-dimensional càdlàg adapted process, $X = (X_t)_{t\in[0,T]}$ is a d-dimensional càdlàg adapted process, $w = (w_t)_{t\in[0,T]}$ is an r-dimensional standard $\mathbf{F}$-Wiener process, θ is a parameter in the closure of an open set Θ in $\mathbb{R}^{\mathsf{p}}$, and $\sigma : \mathbb{R}^{\mathsf{d}} \times \bar{\Theta} \to \mathbb{R}^{\mathsf{m}} \otimes \mathbb{R}^{\mathsf{r}}$ is a continuous function. $J = (J_t)_{t\in[0,T]}$ is the jump part of Y, i.e., $J_t = \sum_{s\in[0,t]} \Delta Y_s$, where $\Delta Y_s = Y_s - Y_{s-}$ and $\Delta Y_0 = 0$. We assume $J_0 = 0$ and $\sum_{t\in[0,T]} 1_{\{\Delta J_t \neq 0\}} < \infty$ a.s. Model (1) is a stochastic regression model, but for example, it can express a diffusion-type process with jumps ΔJ^X contaminated by exogenous jump noise J^Y:

$$\begin{cases} Y_t &= X_t + J^Y_t, \\ X_t &= X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma(X_s, \theta) dw_s + J^X_t, \end{cases}$$

with $J = J^X + J^Y$, and as a special case, a jump-diffusion process. We want to estimate the true value $\theta^* \in \Theta$ of θ based on the data $(X_{t_j}, Y_{t_j})_{j=0,1,\ldots,n}$, where $t_j = t^n_j = jT/n$. Asymptotic properties of estimators will be discussed when $n \to \infty$. That is, the observations are high-frequency data. The data of the processes b and J are not available since they are not directly observed.

Today, a substantial amount of literature is available on parametric estimation of the diffusion parameter θ of diffusion-type processes with/without jumps. In the ergodic diffusion case of $J = 0$ and $T \to \infty$, the drift coefficient is parameterized as well as the diffusion coefficient. Certain asymptotic properties of estimators are found in Prakasa Rao (1983, 1988). The joint asymptotic normality of estimators was given in Yoshida (1992) and later generalized in Kessler (1997). The quasi-likelihood analysis (QLA, Yoshida 2011) ensures not only limit theorems but also moment convergence of the QLA estimators, i.e., the quasi-maximum likelihood estimator (QMLE) and the quasi-Bayesian estimator (QBE). The adaptive estimators (Uchida and Yoshida 2012, 2014) and the hybrid multi-step estimators (Kamatani and Uchida 2014) are of practical importance from computational aspects. Statistics becomes nonergodic under a finite time horizon $T < \infty$. Dohnal (1987) discussed estimation of the diffusion parameter based on high-frequency data. Stable convergence of the quasi-maximum likelihood estimator was given by Genon-Catalot and Jacod (1993). Uchida and Yoshida (2013) showed stable convergence of the quasi-Bayesian estimator and moment convergence of the QLA estimators. The methods of the QLA were essential there and will be applied in this article. The nonsynchronous case is addressed by Ogihara and Yoshida (2014) within QLA. As for inference for jump-diffusion processes, under ergodicity, Ogihara and Yoshida (2011) showed asymptotic normality of the QLA estimators and moment convergence of their error. They used a type of optimal jump-filtered quasi-likelihood function in Shimizu and Yoshida (2006).

The filter in the quasi-likelihood functions of Shimizu and Yoshida (2006) is based on the magnitude of the absolute value of the increment: $\{|\Delta_i Y| > Ch^\rho_n\}$, where $\Delta_i Y = Y_{t_i} - Y_{t_{i-1}}$, $\rho \in [0, 1/2)$ and $C > 0$. If an increment is sufficiently large relative to the threshold, then it is classified as a jump. If, on the other hand, the size of the increment is "moderate," it is regarded as coming from the continuous part. Then, the parameters in the continuous and jump parts can optimally be estimated by respective datasets obtained by classification of increments. This threshold is natural, and in fact, historically, the idea goes back to studies of limit theorems for semimartingales, even further back to Lévy processes.

However, this jump detection filter has a caveat. Though the efficiency of the estimators has been established theoretically, it is known that their real performance strongly depends on a choice of tuning parameters; see, e.g., Shimizu (2009), Iacus and Yoshida (2018). The filter is each time based on only one increment of the data. In this sense, this filter can be regarded as a *local* method. This localism would cause misclassification of increments in practice, even though it should not occur mathematically by the large deviation principle in the limit, and estimated values' instability and strong dependency on the tuning parameters. To overcome these problems, we introduce a *global* filtering method, which we call the *α-threshold method*. It uses all of the data to more accurately detect increments having jumps, based on the order statistics associated with all increments. Another advantage of the global filter is that it does not need any restrictive condition on the distribution of small jumps. This paper provides efficient parametric estimators for the model (1) under a finite time horizon $T < \infty$ by using the α-threshold method, while applications of this method to the realized volatility and other related problems are straightforward. Additionally, it should be remarked that though the α-threshold method involves the tuning parameter α to determine a selection rule for increments, it is robust against the choice of α as we will see later.

The organization of this paper is as follows. In Sect. 2.2, we introduce the α-quasi-log likelihood function $\mathbb{H}_n(\theta;\alpha)$, which is a truncated version of the quasi-log likelihood function made from local Gaussian approximation, based on the global filter for the tuning parameter α. The α-quasi-maximum likelihood estimator (α-QMLE) $\hat{\theta}_n^{M,\alpha}$ is defined with respect to $\mathbb{H}_n(\theta;\alpha)$. Since the truncation is formulated by the order statistics of the increments, this filter destroys adaptivity and martingale structure. However, the global filtering lemmas in Sect. 2.4 enable us to recover these properties. Section 2.5 gives a rate of convergence of the α-QMLE $\hat{\theta}_n^{M,\alpha}$ in L^p sense. In order to prove it, with the help of the QLA theory (Yoshida 2011), the so-called polynomial-type large deviation inequality is derived in Theorem 1 for an annealed version of the quasi-log likelihood $\mathbb{H}_n^{\beta}(\theta;\alpha)$ of (11), where β is the annealing index. Moreover, the (α,β)-quasi-Bayesian estimator ((α,β)-QBE) $\hat{\theta}_n^{B,\alpha,\beta}$ can be defined as the Bayesian estimator with respect to $\mathbb{H}_n^{\beta}(\theta;\alpha)$ as (12). Then, the polynomial-type large deviation inequality makes it possible to prove L^p-boundedness of the error of the (α,β)-QBE $\hat{\theta}_n^{B,\alpha,\beta}$ (Proposition 2). The α-QMLE and (α,β)-QBE do not attain the optimal rate of convergence when the parameter α is fixed though the fixed α-method surely removes jumps as a matter of fact. In Sect. 3, we introduce a quasi-likelihood function $\mathbb{H}_n(\theta)$ depending on a moving level α_n. The random field $\mathbb{H}_n(\theta)$ is more aggressive than $\mathbb{H}_n(\theta;\alpha)$ with a fixed α. Then, a polynomial-type large deviation inequality is obtained in Theorem 2 but the scaling factor is $n^{-1/2}$ in this case so that we can prove $\sqrt{n}$-consistency in L^p sense for both QMLE $\hat{\theta}_n^{M,\alpha_n}$ and QBE $\hat{\theta}_n^{B,\alpha_n}$ associated with the random field $\mathbb{H}_n(\theta)$ (Proposition 3). Stable convergence of these estimators and moment convergence are validated by Theorem 3. The moving threshold method attains the optimal rate of convergence in contrast to the fixed-α method. However, the theory requires the sequence α_n should keep a certain balance: too large α_n causes deficiency and too small α_n may fail to filter out jumps. To balance efficiency of estimation and precision in filtering by taking advantage of the stability of the fixed-α scheme, in Sect. 4, we construct a one-step estimator $\check{\theta}_n^{M,\alpha}$ for a fixed α and the aggressive $\mathbb{H}_n(\theta)$ with the α-QMLE $\hat{\theta}_n^{M,\alpha}$ as the initial estimator. Similarly, the one-step estimator $\check{\theta}_n^{B,\alpha,\beta}$ is constructed for fixed (α, β) and $\mathbb{H}_n(\theta)$ with the (α, β)-quasi-Bayesian estimator $\hat{\theta}_n^{B,\alpha,\beta}$ for the initial estimator. By combining the results in Sects. 2 and 3, we show that these estimators enjoy the same stable convergence and moment convergence as QMLE $\hat{\theta}_n^{M,\alpha_n}$ and QBE $\hat{\theta}_n^{B,\alpha_n}$. It turns out in Section 6 in Supplementary materials that the so-constructed estimators are accurate and quite stable against α, in practice. In Sect. 5, we relax the conditions for stable convergence by a localization argument. Section 6 in Supplementary materials presents some simulation results and shows that the global filter can detect jumps more precisely than the local threshold methods.

2 Global filter: α-threshold method

2.1 Model structure

We will work with model (1). To structure the model suitably, we begin with an example.

Example 1 Consider a two-dimensional stochastic differential equation partly having jumps:

$$\begin{cases} d\xi_t &= b_t^{\xi} dt + \sigma^{\xi}(\xi_t, \eta_t, \zeta_t, \theta) dw_t^{\xi} + dJ_t^{\xi} \\ d\eta_t &= b_t^{\eta} dt + \sigma^{\eta}(\xi_t, \eta_t, \zeta_t, \theta) dw_t^{\eta}. \end{cases}$$

We can set $Y = (\xi, \eta)$, $X = (\xi, \eta, \zeta)$ and $J = (J^{\xi}, 0)$. No jump filter is necessary for the component η.

This example suggests that different treatments should be given componentwise. We assume that

$$\sigma = \mathrm{diag}[\sigma^{(1)}(x, \theta), \dots, \sigma^{(\mathsf{k})}(x, \theta)]$$

for some $\mathsf{m}_k \times \mathsf{m}_k$ nonnegative symmetric matrices $\sigma^{(k)}(x, \theta)$, $k = 1, ..., \mathsf{k}$, and we further assume that $w = (w^{(k)})_{k=1,...,\mathsf{k}}$ with $\mathsf{r} = \sum_{k=1}^{\mathsf{m}} \mathsf{m}_k = \mathsf{m}$. Let $S = \sigma^{\otimes 2} = \sigma\sigma^{\star}$. Then, $S(x, \theta)$ has the form of

$$S(x, \theta) = \mathrm{diag}\left[S^{(1)}(x, \theta), ..., S^{(\mathsf{k})}(x, \theta)\right]$$

for $\mathsf{m}_k \times \mathsf{m}_k$ matrices $S^{(k)}(x, \theta) = \sigma^{(k)}(\sigma^{(k)})^{\star}(x, \theta)$, $k = 1, ..., \mathsf{k}$. According to the blocks of S, we write

$$Y_t = \begin{bmatrix} Y_t^{(1)} \\ \vdots \\ Y_t^{(\mathsf{k})} \end{bmatrix}, \qquad b_t = \begin{bmatrix} b_t^{(1)} \\ \vdots \\ b_t^{(\mathsf{k})} \end{bmatrix}, \qquad w_t = \begin{bmatrix} w_t^{(1)} \\ \vdots \\ w_t^{(\mathsf{k})} \end{bmatrix}, \qquad J_t = \begin{bmatrix} J_t^{(1)} \\ \vdots \\ J_t^{(\mathsf{k})} \end{bmatrix}.$$

Let $N_t^X = \sum_{s\leq t} 1_{\{\Delta X_s \neq 0\}}$. We will pose a condition that $N_T^X < \infty$ a.s. The jump part J^X of X is defined by $J_t^X = \sum_{s\leq t} \Delta X_s$.

2.2 Quasi-likelihood function by order statistics

In this section, we will give a filter that removes ΔJ. Shimizu and Yoshida (2006) and Ogihara and Yoshida (2011) used certain jump detection filters that cut large increments $\Delta_j Y$ by a threshold comparable to diffusion increments. It is a *local* filter because the classification is done for each increment without using other increments. Contrarily, in this paper, we propose a *global* filter that removes increments $\Delta_j Y$ when $|\Delta_j Y|$ is in an upper class among all data $\{|\Delta_i Y|\}_{i=1,\dots,n}$.

We prepare statistics $\bar{S}^{(k)}_{n,j-1}$ ($k = 1,\dots,\mathsf{k}$; $j = 1,\dots,n$; $n \in \mathbb{N}$) such that each $\bar{S}^{(k)}_{n,j-1}$ is an initial estimator of $S^{(k)}(X_{t_{j-1}}, \theta^*)$ up to a scaling constant; that is, there exists a (possibly unknown) positive constant $c^{(k)}$ such that every $S^{(k)}(X_{t_{j-1}}, \theta^*)$ is approximated by $c^{(k)}\bar{S}^{(k)}_{n,j-1}$, as precisely stated later. We do not assume that $\bar{S}^{(k)}_{n,j-1}$ is $\mathcal{F}_{t_{j-1}}$ -measurable.

Example 2 Let K be a positive integer. Let $(\bar{i}_n)$ be a diverging sequence of positive integers, e.g., $\bar{i}_n \sim h^{-1/2}$. Let

$$\hat{S}^{(k)}_{n,j-1} = \frac{\sum_{i=-\bar{i}_n}^{\bar{i}_n} \left(\Delta_{j-i} Y^{(k)}\right)^{\otimes 2} 1_{\left\{|\Delta_{j-i-K+1} Y^{(k)}|\wedge\cdots\wedge|\Delta_{j-i-1} Y^{(k)}| \geq |\Delta_{j-i} Y^{(k)}|\right\}}}{h \max\left\{1, \sum_{i=-\bar{i}_n}^{\bar{i}_n} 1_{\left\{|\Delta_{j-i-K+1} Y^{(k)}|\wedge\cdots\wedge|\Delta_{j-i-1} Y^{(k)}| \geq |\Delta_{j-i} Y^{(k)}|\right\}}\right\}}.$$

Here, $\Delta_j Y^{(k)}$ reads 0 when $j \leq 0$ or $j > n$. An example of $\bar{S}^{(k)}_{n,j-1}$ is

$$\bar{S}^{(k)}_{n,j-1} = \hat{S}^{(k)}_{n,j-1} 1_{\{\lambda_{\min}(\hat{S}^{(k)}_{n,j-1}) > 2^{-1}\epsilon_0\}} + 2^{-1}\epsilon_0 I_{\mathsf{m}_k} 1_{\{\lambda_{\min}(\hat{S}^{(k)}_{n,j-1}) \leq 2^{-1}\epsilon_0\}}, \tag{2}$$

and suppose that $\inf_{x,\theta} \lambda_{\min}(S^{(k)}(x,\theta)) \geq \epsilon_0$ for some positive constant ϵ_0, where $\lambda_{\min}$ is the minimum eigenvalue of the matrix.

Let $\alpha = (\alpha^{(k)})_{k\in\{1,\dots,\mathsf{k}\}} \in [0,1)^{\mathsf{k}}$. Our global jump filter is constructed as follows. Denote by $\mathcal{J}_n^{(k)}(\alpha^{(k)})$ the set of $j \in \{1,\dots,n\}$ such that

$$\#\left\{j' \in \{1,\dots,n\}; |(\bar{S}^{(k)}_{n,j'-1})^{-1/2}\Delta_{j'} Y^{(k)}| > |(\bar{S}^{(k)}_{n,j-1})^{-1/2}\Delta_j Y^{(k)}|\right\} \geq \alpha^{(k)} n$$

for $k = 1,\dots,\mathsf{k}$ and $n \in \mathbb{N}$. If $\alpha^{(k)} = 0$, then $\mathcal{J}_n^{(k)}(\alpha^{(k)}) = \{1,\dots.,n\}$; that is, there is no filter for the kth component. The density function of the multidimensional normal distribution with mean vector μ and covariance matrix C is denoted by $\phi(z;\mu,C)$. Let

$$q^{(k)}(\alpha^{(k)}) = \frac{\mathrm{Tr}\left(\int_{\{|z| \le c(\alpha^{(k)})^{1/2}\}} z^{\otimes 2} \phi(z;0, I_{\mathsf{m}_k}) \mathrm{d}z \right)}{\mathrm{Tr}\left(\int_{\mathbb{R}^{\mathsf{m}_k}} z^{\otimes 2} \phi(z;0, I_{\mathsf{m}_k}) \mathrm{d}z \right)},$$

equivalently,

$$\begin{aligned} q^{(k)}(\alpha^{(k)}) &= (\mathsf{m}_k)^{-1} \mathrm{Tr}\left(\int_{\{|z| \le c(\alpha^{(k)})^{1/2}\}} z^{\otimes 2} \phi(z;0, I_{\mathsf{m}_k}) \mathrm{d}z \right) \\ &= (\mathsf{m}_k)^{-1} E[V 1_{\{V \le c(\alpha^{(k)})\}}], \end{aligned}$$

for a random variable $V \sim \chi^2(\mathsf{m}_k)$, the Chi-squared distribution with m_k degrees of freedom, where $c(\alpha^{(k)})$ is determined by

$$P[V \le c(\alpha^{(k)})] = 1 - \alpha^{(k)}.$$

Let $p(\alpha^{(k)}) = 1 - \alpha^{(k)}$. Now the **$\boldsymbol{\alpha}$ -quasi-log likelihood function** $\mathbb{H}_n(\theta;\alpha)$ is defined by

$$\begin{aligned} \mathbb{H}_n(\theta;\alpha) = -\frac{1}{2} \sum_{k=1}^{\mathsf{k}} \sum_{j \in \mathcal{J}_n^{(k)}(\alpha^{(k)})} &\Big\{ q^{(k)}(\alpha^{(k)})^{-1} h^{-1} S^{(k)}(X_{t_{j-1}}, \theta)^{-1} \big[\big(\Delta_j Y^{(k)} \big)^{\otimes 2} \big] K_{n,j}^{(k)} \\ &+ p(\alpha^{(k)})^{-1} \log \det S^{(k)}(X_{t_{j-1}}, \theta) \Big\} \end{aligned}$$

where

$$K_{n,j}^{(k)} = 1_{\left\{ |\Delta_j Y^{(k)}| < C_*^{(k)} n^{-\frac{1}{4}} \right\}} \tag{3}$$

and $C_*^{(k)}$ are arbitrarily given positive constants. For a tensor $T = (T_{i_1,\dots,i_k})_{i_1,\dots,i_k}$, we write

$$T[x_1, \dots, x_k] = T[x_1 \otimes \cdots \otimes x_k] = \sum_{i_1,\dots,i_k} T_{i_1,\dots,i_k} x_1^{i_1} \cdots x_k^{i_k}$$

for $x_1 = (x_1^{i_1})_{i_1}$, ..., $x_k = (x_k^{i_k})_{i_k}$. We denote $u^{\otimes r} = u \otimes \cdots \otimes u$ (r times). Brackets [ ] stand for a multilinear mapping. This notation also applies to tensor-valued tensors.

If $\alpha^{(k)} = 0$, then $\mathcal{J}_n^{(k)}(\alpha^{(k)}) = \{1, \dots, n\}$, $c(\alpha^{(k)}) = +\infty$, $p^{(k)}(\alpha^{(k)}) = 1$ and $q^{(k)}(\alpha^{(k)}) = 1$, so the kth component of $\mathbb{H}_n(\theta;\alpha)$ essentially becomes the ordinary quasi-log likelihood function by local Gaussian approximation.

Remark 1 (i) The cap $K_{n,j}^{(k)}$ can be removed if a suitable condition is assumed for the big jump sizes of J, e.g., $\sup_{t \in [0,T]} |\Delta J_t| \in L^{\infty\text{-}} = \cap_{p>1} L^p$. It is also reasonable to use

$$K_{n,j}^{(k)} = 1_{\left\{ |\bar{S}_{n,j-1}^{-1/2} \Delta_j Y^{(k)}| < C_*^{(k)} n^{-\frac{1}{4}} \right\}}$$

if $\bar{S}_{n,j-1}$ is uniformly $L^{\infty-}$-bounded. In any case, the factor $K^{(k)}_{n,j}$ only serves for removing the effects of too big jumps and the classification is practically never affected by it since the global filter puts a threshold of the order less than $n^{-1/2}\log n$. As a matter of fact, the threshold of $K^{(k)}_{n,j}$ is of order $O(n^{-1/4})$, which is far looser than the ordinary local filters, and the truncation is exercised only with exponentially small probability. On the other hand, the global filter puts no restrictive condition on the distribution of the size of small jumps, like vanishing at the origin or boundedness of the density of the jump sizes, as assumed for the local filters so far. It should be emphasized that the difficulties in jump filtering are focused on the treatments of small jumps that look like the Brownian increments. (ii) The symmetry of $\sigma^{(k)}(x,\theta)$ is not restrictive because $\sigma^{(k)}(X_t,\theta)dw^{(k)}_t = S^{(k)}(X_t,\theta)^{1/2}\cdot\big(S^{(k)}(X_t,\theta)^{-1/2}\sigma^{(k)}(X_t,\theta)dw^{(k)}_t\big)$. On the other hand, we could introduce an $\mathsf{m}_k\times\mathsf{m}_k$ random matrix $\bar{\sigma}^{(k)}_{n,j-1}$ approximating $\sigma^{(k)}(X_{t_{j-1}},\theta^*)$ up to scaling and use $\big(\bar{\sigma}^{(k)}_{n,j-1}\big)^{-1}\Delta_jY^{(k)}$ for $\big(\bar{S}^{(k)}_{n,j-1}\big)^{-1/2}\Delta_jY^{(k)}$, in order to remove the assumption of symmetry.

The **α-quasi-maximum likelihood estimator** of θ (α-QMLE) is any measurable mapping $\hat{\theta}^{M,\alpha}_n$ characterized by

$$\mathbb{H}_n(\hat{\theta}^{M,\alpha}_n;\alpha) = \max_{\theta\in\bar{\Theta}}\mathbb{H}_n(\theta;\alpha).$$

We will identify an estimator of θ that is a measurable mapping of the data, with the pullback of it to Ω since the aim of discussion here is to obtain asymptotic properties of the estimators' distribution.

2.3 Assumptions

We assume Sobolev embedding inequality

$$\sup_{\theta\in\Theta}\big|f(\theta)\big| \le C_{\Theta,p}\left\{\sum_{i=0}^{1}\int_\Theta\big|\partial^i_\theta f(\theta)\big|^p d\theta\right\}^{1/p} \qquad (f\in C^1(\Theta))$$

for a bounded open set Θ in $\mathbb{R}^{\mathsf{p}}$, where $C_{\Theta,p}$ is a constant, $p>\mathsf{p}$. This inequality is valid, e.g., if Θ has a Lipschitz boundary. Denote by $C^{a,b}_{\uparrow}(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R}^{\mathsf{m}}\otimes\mathbb{R}^{\mathsf{r}})$ the set of continuous functions $f:\mathbb{R}^{\mathsf{d}}\times\Theta\to\mathbb{R}^{\mathsf{m}}\otimes\mathbb{R}^{\mathsf{r}}$ that have continuous derivatives $\partial_{s_1}\cdots\partial_{s_\ell}f$ for all $(s_1,...,s_\ell)\in\{\theta,x\}^\ell$ such that $\#\{i\in\{1,...,\ell\};s_i=x\}\le a$ and $\#\{i\in\{1,...,\ell\};s_i=\theta\}\le b$, and each of these derivatives satisfies

$$\sup_{\theta\in\Theta}\big|\partial_{s_1}\cdots\partial_{s_\ell}f(x,\theta)\big| \le C(s_1,...,s_\ell)\big(1+|x|^{C(s_1,...,s_\ell)}\big) \quad (x\in\mathbb{R}^{\mathsf{d}})$$

for some positive constant $C(s_1,...,s_\ell)$. Let $\|V\|_p=\big(E[|V|^p]\big)^{1/p}$ for a vector-valued random variable V and $p>0$. Let $N^{(k)}_t=\sum_{s\le t}1_{\{\Delta J^{(k)}_s\neq 0\}}$ and $N_t=\sum_{s\le t}1_{\{\Delta J_s\neq 0\}}$ We shall consider the following conditions. Let $\widetilde{X}=X-J^X$ for $J^X=\sum_{s\in[0,\cdot]}\Delta X_s$.

[F1]$_\kappa$ (i) For every $p>1$, $\sup_{t\in[0,T]}\|X_t\|_p<\infty$ and there exists a constant $C(p)$ such that

$$\|\widetilde{X}_t - \widetilde{X}_s\|_p \leq C(p)|t-s|^{1/2} \qquad (t,s \in [0,T]).$$

(ii) $\sup_{t\in[0,T]} \|b_t\|_p < \infty$ for every $p > 1$.

(iii) $\sigma \in C_{\uparrow}^{2,\kappa}(\mathbb{R}^{\mathsf{d}} \times \Theta;\mathbb{R}^{\mathsf{m}} \otimes \mathbb{R}^{\mathsf{r}})$, $S(X_t,\theta)$ is invertible a.s. for every $\theta \in \Theta$, and $\sup_{t\in[0,T],\theta\in\Theta} \|S(X_t,\theta)^{-1}\|_p < \infty$ for every $p > 1$.

(iv) $N_T \in L^{\infty-}$ and $N_T^X \in L^{\infty-}$.

[F2] (i) $\bar{S}^{(k)}_{n,j-1}$ are symmetric, invertible and $\sup_{n\in\mathbb{N}} \max_{j=1,...,n} \left\|(\bar{S}^{(k)}_{n,j-1})^{-1}\right\|_p < \infty$ for every $p > 1$ and $k = 1, ..., \mathsf{k}$.

(ii) There exist positive constants γ_0 and $c^{(k)}$ ($k = 1, ..., \mathsf{k}$) such that

$$\sup_{n\in\mathbb{N}} \max_{j=1,...,n} n^{\gamma_0} \left\|S^{(k)}(X_{t_{j-1}},\theta^*) - c^{(k)}\bar{S}^{(k)}_{n,j-1}\right\|_p < \infty$$

for every $p > 1$ and $k = 1, ..., \mathsf{k}$.

Remark 2 In $[F2]$ (ii), we assumed that there exists a positive constant $c^{(k)}$ such that every $S^{(k)}(X_{t_{j-1}},\theta^*)$ is approximated by $c^{(k)}\bar{S}^{(k)}_{n,j-1}$. In estimation of θ, we only assume positivity of $c^{(k)}$ but the values of them can be unknown since the function $\mathbb{H}_n$ does not involve $c^{(k)}$. When $S^{(k)}(X_{t_{j-1}},\theta^*)$ is a scalar matrix, Condition $[F2]$ is satisfied simply by $\bar{S}^{(k)}_{n,j-1} = I_{\mathsf{m}_k}$.

Remark 3 The $\bar{S}^{(k)}_{n,j-1}$ given by (2) in Example 2 satisfies Condition $[F2]$ with $\gamma_0 = 1/4$ if one takes $\bar{i}_n \sim h^{-1/2}$. The constant $c^{(k)}$ depends on the depth K of the threshold. It is possible to give an explicit expression of $c^{(k)}$ but not required by the condition.

2.4 Global filtering lemmas

The α-quasi-log likelihood function $\mathbb{H}_n(\theta;\alpha)$ involves the summation regarding the index set $\mathcal{J}^{(k)}_n(\alpha^{(k)})$. The global jump filter $\mathcal{J}^{(k)}_n(\alpha^{(k)})$ avoids taking jumps but it completely destroys the martingale structure that the ordinary quasi-log likelihood function originally possessed, and without the martingale structure, we cannot follow a standard way to validate desirable asymptotic properties the estimator should have. However, it is possible to recover the martingale structure to some extent by deforming the global jump filter to a suitable deterministic filter. In this section, we will give several lemmas that enable such a deformation.

As before, $\alpha = (\alpha^{(k)})_{k=1,...,\mathsf{k}}$ is a fixed vector in $[0,1)^{\mathsf{k}}$. We may assume that $\gamma_0 \in (0,1/2]$ in $[F2]$. Let

$$U^{(k)}_j = (c^{(k)})^{-1/2} h^{-1/2} (\bar{S}^{(k)}_{n,j-1})^{-1/2} \Delta_j Y^{(k)} \quad \text{and} \quad W^{(k)}_j = h^{-1/2}\Delta_j w^{(k)}.$$

By $[F1]_0$ and $[F2]$, we have

$$\sup_{j=1,...,n} \left\|R^{(k)}_j 1_{\{\Delta_j N^X=0\}}\right\|_p = O(n^{-\gamma_0})$$

for every $p > 1$, where

$$R_j^{(k)} = U_j^{(k)} - W_j^{(k)} - (c^{(k)})^{-1/2}h^{-1/2}(\bar{S}_{n,j-1}^{(k)})^{-1/2}\Delta_j J^{(k)}.$$

Remark that $A^{1/2} = \frac{1}{\pi}\int_0^\infty \lambda^{-1/2}A(\lambda + A)^{-1}\mathrm{d}\lambda$ for a positive-definite matrix A.

Denote $|W_j^{(k)}|$ and $|U_j^{(k)}|$ by $\overline{W}_j^{(k)}$ and $\overline{U}_j^{(k)}$, respectively. $\overline{W}_{(j)}^{(k)}$ denotes the jth-ordered statistic of $\{\overline{W}_1^{(k)}, ..., \overline{W}_n^{(k)}\}$, and $\overline{U}_{(j)}^{(k)}$ denotes the jth-ordered statistic of $\{\overline{U}_1^{(k)}, ..., \overline{U}_n^{(k)}\}$. The rank of $\overline{W}_j^{(k)}$ is denoted by $r(\overline{W}_j^{(k)})$. Denote by $q_{\alpha^{(k)}}$ the $\alpha^{(k)}$-quantile of the distribution of $\overline{W}_1^{(k)}$. The number $q_{\alpha^{(k)}}$ depends on m_k.

Let $0 < \gamma_2 < \gamma_1 < \gamma_0$. Let $a_n^{(k)} = \lfloor \bar{\alpha}^{(k)}n - n^{1-\gamma_2} \rfloor$, where $\bar{\alpha}^{(k)} = 1 - \alpha^{(k)} = p(\alpha^{(k)})$. Define the event $N_{n,j}^{(k)}$ by

$$N_{n,j}^{(k)} = \left\{ r(\overline{W}_j^{(k)}) \le a_n^{(k)} - n^{1-\gamma_2} \right\} \cap \left\{ \overline{W}_{(a_n^{(k)})}^{(k)} - \overline{W}_j^{(k)} < n^{-\gamma_1} \right\}.$$

Lemma 1 *Suppose that* $\alpha^{(k)} \in (0,1)$. *Then* $P\left[\bigcup_{j=1,\dots,n} N_{n,j}^{(k)}\right] = O(n^{-L})$ *as* $n \to \infty$ *for every* $L > 0$.

Proof We have

$$\begin{aligned}
&P\left[\overline{W}_{(a_n^{(k)})}^{(k)} > q_{\bar{\alpha}^{(k)}} + n^{-\gamma_1}\right] \\
=&P\left[\sum_{j=1}^n 1_{\{\overline{W}_j^{(k)} \le q_{\bar{\alpha}^{(k)}} + n^{-\gamma_1}\}} < a_n^{(k)}\right] \\
=&P\left[n^{-1/2}\sum_{j=1}^n \left\{ 1_{\{\overline{W}_j^{(k)} \le q_{\bar{\alpha}^{(k)}} + n^{-\gamma_1}\}} - P\left[\overline{W}_j^{(k)} \le q_{\bar{\alpha}^{(k)}} + n^{-\gamma_1}\right] \right\} < -n^{\frac{1}{2}-\gamma_1}c(n)\right] \\
=&O(n^{-L})
\end{aligned}$$

for every $L > 0$, where $(c(n))_{n\in\mathbb{N}}$ is a sequence of numbers such that $\inf_{n\in\mathbb{N}} c(n) > 0$ (the existence of such $c(n)$ can be proved by the mean value theorem). The last equality in the above estimates is obtained by the following argument. For $A_j = \{\overline{W}_j^{(k)} \le q_{\bar{\alpha}^{(k)}} + n^{-\gamma_1}\}$ and $Z_j = 1_{A_j} - P[A_1]$, by the Burkholder–Davis–Gundy inequality, Jensen's inequality and $|Z_j| \le 1$, we obtain

$$\begin{aligned}
P\left[n^{-1/2}\sum_{j=1}^n Z_j < -n^{\frac{1}{2}-\gamma_1}c(n)\right] &\lesssim n^{-2p(\frac{1}{2}-\gamma_1)}c(n)^{-2p}E\left[n^{-1}\sum_{j=1}^n |Z_j|^{2p}\right] \\
&= O(n^{-p(1-2\gamma_1)})
\end{aligned}$$

for every $p > 1$.

Let

$$B_n^{(k)} = \left\{ \left|\overline{W}_{(a_n^{(k)})}^{(k)} - q_{\bar{\alpha}^{(k)}}\right| > n^{-\gamma_1} \right\}.$$

We can estimate $P\left[\overline{W}_{(a_n^{(k)})}^{(k)} < q_{\bar{\alpha}^{(k)}} - n^{-\gamma_1}\right]$, and so we have

$$P\left[B_n^{(k)}\right] = O(n^{-L}) \tag{4}$$

for every $L > 0$.

By definition, on the event $N_{n,j}^{(k)} \cap (B_n^{(k)})^c$, the number of data $\overline{W}_{j'}^{(k)}$ on the interval $\left[q_{\bar{\alpha}^{(k)}} - 2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}} + 2n^{-\gamma_1}\right]$ is not less than $n^{1-\gamma_2}$. However,

$$\begin{aligned} &P\left[\sum_{j'=1}^{n} 1_{\left\{\overline{W}_{j'}^{(k)} \in \left[q_{\bar{\alpha}^{(k)}} - 2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}} + 2n^{-\gamma_1}\right]\right\}} \geq n^{1-\gamma_2}\right] \\ =&P\left[n^{-1+\gamma_1} \sum_{j'=1}^{n} 1_{\left\{\overline{W}_{j'}^{(k)} \in \left[q_{\bar{\alpha}^{(k)}} - 2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}} + 2n^{-\gamma_1}\right]\right\}} \geq n^{\gamma_1-\gamma_2}\right] \\ =&O(n^{-L}) \end{aligned} \tag{5}$$

for every $L > 0$. Indeed, the family

$$\left\{n^{-1/2} \sum_{j'=1}^{n} \left(1_{\left\{\overline{W}_{j'}^{(k)} \in \left[q_{\bar{\alpha}^{(k)}} - 2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}} + 2n^{-\gamma_1}\right]\right\}} - E\left[1_{\left\{\overline{W}_{j'}^{(k)} \in \left[q_{\bar{\alpha}^{(k)}} - 2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}} + 2n^{-\gamma_1}\right]\right\}}\right]\right)\right\}_{n\in\mathbb{N}}$$

is bounded in $L^{\infty-}$ (this can be proved by the same argument as above). Since the estimate (5) is independent of $j \in \{1, ..., n\}$, combining it with (4), we obtain

$$\max_{j=1,...,n} P\left[N_{n,j}^{(k)}\right] = O(n^{-L})$$

as $n \to \infty$ for every $L > 0$. Now the desired inequality of the lemma is obvious. □

Let

$$\hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)}) = \left\{j \in \{1, ..., n\}; r(\overline{W}_j^{(k)}) \leq \hat{a}_n^{(k)}\right\},$$

where

$$\hat{a}_n^{(k)} = \lfloor a_n^{(k)} - n^{1-\gamma_2} \rfloor.$$

Let $\mathcal{L}_n^{(k)} = \{j; \Delta_j N^{(k)} + \Delta_j N^X \neq 0\}$. Let

$$\begin{aligned} \Omega_n = &\left\{N_T + N_T^X < n^{1-\gamma_2}\right\} \\ &\bigcap\left(\bigcap_{k=1,...,\mathsf{k}} \bigcap_{j=1,...,n} \left[\left\{|R_j^{(k)}| 1_{\{\Delta_j N^X = 0\}} < 2^{-1} n^{-\gamma_1}\right\} \cap (N_{n,j}^{(k)})^c\right]\right). \end{aligned}$$

Lemma 2

$$\hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)}) \cap (\mathcal{L}_n^{(k)})^c \subset \mathcal{J}_n^{(k)}(\alpha^{(k)}) \tag{6}$$

on Ω_n. In particular

$$\#\left[\mathcal{J}_n^{(k)}(\alpha^{(k)}) \ominus \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})\right] \leq c_* n^{1-\gamma_2} + N_T^{(k)} + N_T^X \tag{7}$$

on Ω_n, where c_* is a positive constant. Here, $\ominus$ denotes the symmetric difference operator of sets.

Proof On Ω_n, if a pair $(j_1,j_2) \in (\mathcal{L}_n^{(k)})^c \times (\mathcal{L}_n^{(k)})^c$ satisfies $r(\overline{W}_{j_1}^{(k)}) \leq \hat{a}_n^{(k)}$ and $r(\overline{W}_{j_2}^{(k)}) \geq a_n^{(k)}$, then $\overline{U}_{j_1}^{(k)} < \overline{W}_{j_1}^{(k)} + 2^{-1}n^{-\gamma_1} \leq \overline{W}_{(a_n^{(k)})}^{(k)} - 2^{-1}n^{-\gamma_1} \leq \overline{W}_{j_2}^{(k)} - 2^{-1}n^{-\gamma_1} < \overline{U}_{j_2}^{(k)}$. Therefore, if $j \in \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)}) \cap (\mathcal{L}_n^{(k)})^c$, then $j \in \mathcal{J}_n^{(k)}(\alpha^{(k)})$ since one can find at least $\lceil \alpha^{(k)} n \rceil (\leq (n - a_n^{(k)} + 1) - n^{1-\gamma_2})$ variables among $\overline{U}_{(a_n^{(k)})}^{(k)}, \ldots.., \overline{U}_{(n)}^{(k)}$ that are larger than $\overline{U}_j^{(k)}$. Therefore, (6) holds, and so does (7) as follows. From (6), we have $\#\big[\mathcal{J}_n^{(k)}(\alpha^{(k)}) \ominus \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})\big] \leq \mathsf{N} + \#\mathcal{L}_n^{(k)}$ for

$$\mathsf{N} = \#\big[\mathcal{J}_n^{(k)}(\alpha^{(k)}) \cap \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})^c \cap (\mathcal{L}^{(k)})^c\big].$$

Suppose that $j \in \mathcal{J}_n^{(k)}(\alpha^{(k)}) \cap \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})^c \cap (\mathcal{L}^{(k)})^c$. In case $r(\overline{W}_j^{(k)}) < a_n^{(k)}$, since $\hat{a}_n^{(k)} < r(\overline{W}_j^{(k)}) < a_n^{(k)}$, we know the number of such j is less than or equal to $n^{1-\gamma_2}$. In Case $r(\overline{W}_j^{(k)}) \geq a_n^{(k)}$, as seen above, $\overline{U}_{j_1}^{(k)} < \overline{U}_j^{(k)}$ on Ω_n for all $j_1 \in (\mathcal{L}^{(k)})^c$ satisfying $r(\overline{W}_{j_1}^{(k)}) \leq \hat{a}_n^{(k)}$, since $j \in (\mathcal{L}^{(k)})^c$ and $r(\overline{W}_j^{(k)}) \geq a_n^{(k)}$. The number of such j_1s is at least $\hat{a}_n^{(k)} - \lfloor n^{1-\gamma_2} \rfloor$. On the other hand, $j \in \mathcal{J}_n^{(k)}(\alpha^{(k)})$ gives $\#\{j' \in \{1, \ldots, n\}; \overline{U}_j^{(k)} < \overline{U}_{j'}^{(k)}\} \geq \lceil \alpha^{(k)} n \rceil$. Therefore

$$\mathsf{N} \leq n^{1-\gamma_2} + n - \big(\hat{a}_n^{(k)} - \lfloor n^{1-\gamma_2} \rfloor\big) - \lceil \alpha^{(k)} n \rceil \ \leq\ 4n^{1-\gamma_2} + 2$$

on Ω_n. Since $\#\mathcal{L}_n^{(k)} \leq N_T^{(k)} + N_T^X$, we obtain (7) on Ω_n with $c_* = 6$ if we use the inequality $4n^{1-\gamma_2} + 2 \leq 6n^{1-\gamma_2}$. □

Let $\gamma_3 > 0$. For random variables $(V_j)_{j=1,\ldots,n}$, let

$$\mathcal{D}_n^{(k)} = n^{\gamma_3} \left| \frac{1}{n} \sum_{j \in \mathcal{J}_n^{(k)}(\alpha^{(k)})} V_j - \frac{1}{n} \sum_{j \in \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})} V_j \right|.$$

Lemma 3

(i) *Let $p_1 > 1$. Then,*

$$\|\mathcal{D}_n^{(k)}\|_p \leq \big(c_* n^{\gamma_3 - \gamma_2} + n^{-1+\gamma_3} \|N_T^{(k)} + N_T^X\|_{p_1}\big) \left\| \max_{j=1,\ldots,n} |V_j| \right\|_{pp_1(p_1-p)^{-1}} + n^{\gamma_3} \left\| \max_{j=1,\ldots,n} |V_j| 1_{\Omega_n^c} \right\|_p$$

for $p \in (1, p_1)$.

(ii) Let $\gamma_4 > 0$ *and* $p_1 > 1$. *Then,*

$$\|\mathcal{D}_n^{(k)}\|_p \leq \left(c_* n^{\gamma_3-\gamma_2} + n^{-1+\gamma_3}\|N_T^{(k)} + N_T^X\|_{p_1}\right) \times \left(n^{\gamma_4} + n \max_{j=1,\ldots,n} \left\| |V_j| 1_{\{|V_j|>n^{\gamma_4}\}} \right\|_{pp_1(p_1-p)^{-1}}\right) + n^{\gamma_3} \left\| \max_{j=1,\ldots,n} |V_j| 1_{\Omega_n^c} \right\|_p$$

for $p \in (1, p_1)$.

Proof The estimate in (i) is obvious from (7). (ii) follows from (i). □

Let $\widetilde{\mathcal{J}}_n^{(k)}(\alpha^{(k)}) = \{j; |h^{-1/2}\Delta_j w^{(k)}| \leq q_{\bar{\alpha}^{(k)}}\} = \{j; \overline{W}_j^{(k)} \leq q_{\bar{\alpha}^{(k)}}\}$. Let

$$\tilde{\mathcal{D}}_n^{(k)} = n^{\gamma_3} \left| \frac{1}{n} \sum_{j \in \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})} V_j - \frac{1}{n} \sum_{j \in \widetilde{\mathcal{J}}_n^{(k)}(\alpha^{(k)})} V_j \right|.$$

Lemma 4 *Let* $\tilde{\Omega}_n = \left\{ \left| \overline{W}_{(\hat{a}_n^{(k)})}^{(k)} - q_{\bar{\alpha}^{(k)}} \right| < \check{C}\, n^{-\gamma_2} \right\}$, *where* $\check{C}$ *is a positive constant. Then,*

(i) *For* $p \geq 1$,

$$\|\tilde{\mathcal{D}}_n^{(k)}\|_p \leq n^{\gamma_3} \left\| \max_{j'=1,\ldots,n} |V_{j'}| \frac{1}{n} \sum_{j=1}^n 1_{\left\{ \left|\overline{W}_j^{(k)} - q_{\bar{\alpha}^{(k)}}\right| \leq \check{C}\, n^{-\gamma_2} \right\}} \right\|_p + n^{\gamma_3} \left\| 1_{\tilde{\Omega}_n^c} \max_{j'=1,\ldots,n} |V_{j'}| \right\|_p.$$

(ii) *For* $p_1 > p \geq 1$,

$$\|\tilde{\mathcal{D}}_n^{(k)}\|_p \leq n^{\gamma_3} \left\| \max_{j=1,\ldots,n} |V_j| \right\|_p P\left[\left|\overline{W}_1^{(k)} - q_{\bar{\alpha}^{(k)}}\right| \leq \check{C}\, n^{-\gamma_2} \right] + n^{\gamma_3} \left\| \max_{j=1,\ldots,n} |V_j| \right\|_{pp_1(p_1-p)^{-1}} \left\| \frac{1}{n} \sum_{j=1}^n \left(1_{\left\{ \left|\overline{W}_j^{(k)} - q_{\bar{\alpha}^{(k)}}\right| \leq \check{C}\, n^{-\gamma_2} \right\}} - P\left[\left|\overline{W}_1^{(k)} - q_{\bar{\alpha}^{(k)}}\right| \leq \check{C}\, n^{-\gamma_2} \right] \right) \right\|_{p_1} + n^{\gamma_3} P[\tilde{\Omega}_n^c]^{1/p_1} \left\| \max_{j=1,\ldots,n} |V_j| \right\|_{pp_1(p_1-p)^{-1}}.$$

Proof (i) follows from

$$1_{\tilde{\Omega}_n} \left| 1_{\left\{ \overline{W}_j^{(k)} \leq \overline{W}_{(\hat{a}_n^{(k)})}^{(k)} \right\}} - 1_{\left\{ \overline{W}_j^{(k)} \leq q_{\bar{\alpha}^{(k)}} \right\}} \right| \leq 1_{\left\{ \left|\overline{W}_j^{(k)} - q_{\bar{\alpha}^{(k)}}\right| \leq \check{C}\, n^{-\gamma_2} \right\}},$$

and (ii) follows from (i). □

We take a sufficiently large $\check{C}$. Then, the term involving $\tilde{\Omega}_n^c$ on the right-hand side of each inequality in Lemma 4 can be estimated as the proof of Lemma 1. For example, $P[\tilde{\Omega}_n^c] = O(n^{-L})$ for any $L > 0$.

Lemma 5 *Let $k \in \{1, ..., \mathsf{k}\}$ and let $f \in C_{\uparrow}^{1,1}\left(\mathbb{R}^{\mathsf{d}} \times \Theta;\mathbb{R}\right)$. Suppose that $[F1]_0$ is fulfilled. Then,*

$$\sup_{n\in\mathbb{N}} \left\| \sup_{\theta\in\Theta} n^{\epsilon} \left| \frac{1}{n} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} p(\alpha^{(k)})^{-1} f(X_{t_{j-1}}, \theta) - \frac{1}{T}\int_0^T f(X_t, \theta)\mathrm{d}t \right| \right\|_p < \infty$$

for every $p \geq 1$ and $\epsilon < \gamma_2$.

Proof Use Sobolev inequality and Burkholder inequality as well as Lemmas 1, 3 (ii) and 4 (ii). More precisely, we have the following decomposition

$$\begin{aligned}
&\frac{1}{n} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} p(\alpha^{(k)})^{-1} f(X_{t_{j-1}}, \theta) - \frac{1}{T}\int_0^T f(X_t, \theta)\mathrm{d}t \\
=& p(\alpha^{(k)})^{-1} \left\{ \frac{1}{n} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} f(X_{t_{j-1}}, \theta) - \frac{1}{n} \sum_{j\in\hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})} f(X_{t_{j-1}}, \theta) \right\} \\
&+ p(\alpha^{(k)})^{-1} \left\{ \frac{1}{n} \sum_{j\in\hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})} f(X_{t_{j-1}}, \theta) - \frac{1}{n} \sum_{j\in\check{\mathcal{J}}_n^{(k)}(\alpha^{(k)})} f(X_{t_{j-1}}, \theta) \right\} \\
&+ \frac{1}{np(\alpha^{(k)})} \sum_{j=1}^n f(X_{t_{j-1}}, \theta) \left\{ 1_{\left\{\overline{W}_j^{(k)} \leq q_{\tilde{\alpha}^{(k)}}\right\}} - p(\alpha^{(k)}) \right\} \\
&+ \frac{1}{nh} \sum_{j=1}^n \int_{t_{j-1}}^{t_j} [f(X_{t_{j-1}}, \theta) - f(X_t, \theta)]\mathrm{d}t \\
=:& I_{1,n}^{(k)}(\theta) + I_{2,n}^{(k)}(\theta) + I_{3,n}^{(k)}(\theta) + I_{4,n}^{(k)}(\theta).
\end{aligned}$$

We may assume $\alpha^{(k)} > 0$ since only $I_{4,n}^{(k)}(\theta)$ remains when $\alpha^{(k)} = 0$, and it will be estimated below.

As for $I_{1,n}^{(k)}(\theta)$, we apply Lemma 3 (ii) to obtain

$$\left\| \sup_{\theta\in\Theta} n^{\epsilon} \left| I^{(k)}_{1,n}(\theta) \right| \right\|_p \lesssim \sum_{i=0,1} \sup_{\theta\in\Theta} \left\| n^{\epsilon} \left| \frac{1}{n} \sum_{j\in\mathcal{J}^{(k)}_n(\alpha^{(k)})} \partial^i_\theta f(X_{t_{j-1}},\theta) - \frac{1}{n} \sum_{j\in\hat{\mathcal{J}}^{(k)}_n(\alpha^{(k)})} \partial^i_\theta f(X_{t_{j-1}},\theta) \right| \right\|_p$$

$$\lesssim \sum_{i=0,1} \sup_{\theta\in\Theta} \left\{ \left(c_* n^{\epsilon-\gamma_2} + n^{-1+\epsilon} \| N^{(k)}_T + N^X_T \|_{p_1} \right) \right.$$

$$\times \left(n^{\gamma_4} + n \max_j \left\| |\partial^i_\theta f(X_{t_{j-1}},\theta)| 1_{\{|\partial^i_\theta f(X_{t_{j-1}},\theta)| \ge n^{\gamma_4}\}} \right\|_{\frac{pp_1}{p_1-p}} \right)$$

$$\left. + n^{\epsilon} \left\| \max_j |\partial^i_\theta f(X_{t_{j-1}},\theta)| 1_{\Omega^c_n} \right\|_p \right\}.$$

By taking $\gamma_4 > 0$ small enough, we can verify that the right-hand side is $o(1)$. Note that we have used the fact $P[\Omega^c_n] = O(n^{-L})$ for any $L > 0$. A similar argument with Lemma 4 (ii) yields $\| \sup_{\theta\in\Theta} n^{\epsilon} |I^{(k)}_{2,n}(\theta)| \|_p = o(1)$.

As for $I^{(k)}_{3,n}(\theta)$, applying the Burkholder–Davis–Gundy inequality for the discrete-time martingales as well as Jensen's inequality, we have

$$\sup_{\theta\in\Theta} \left\| n^{\epsilon} \sum_{j=1}^{n} \frac{1}{n} \partial^i_\theta f(X_{t_{j-1}},\theta) \left\{ 1_{\left\{\overline{W}^{(k)}_j \le q_{\bar{\alpha}(k)}\right\}} - p(\alpha^{(k)}) \right\} \right\|_p^p$$

$$\lesssim \sup_{\theta\in\Theta} n^{-p\left(\frac{1}{2}-\epsilon\right)} E\left[\left| \frac{1}{n} \sum_{j=1}^{n} \left| \partial^i_\theta f(X_{t_{j-1}},\theta) \right|^2 \left\{ 1_{\left\{\overline{W}^{(k)}_j \le q_{\bar{\alpha}(k)}\right\}} - p(\alpha^{(k)}) \right\}^2 \right|^{\frac{p}{2}} \right]$$

$$= O\left(n^{-(\frac{1}{2}-\epsilon)p} \right)$$

for every $p \ge 2$ and $i = 0, 1$. Hence, by Sobolev inequality, we conclude

$$\left\| \sup_{\theta\in\Theta} n^{\epsilon} \left| I^{(k)}_{3,n}(\theta) \right| \right\|_p = O\left(n^{-\frac{1}{2}+\epsilon} \right)$$

for every $p \ge 1$.

Finally, we will estimate $I^{(k)}_{4,n}(\theta)$. Since $f \in C^{1,1}_{\uparrow}(\mathbb{R}^{\mathsf{d}} \times \Theta; \mathbb{R})$, there exists a positive constant C such that

$$C_f(x,y) \le C(1 + |x|^C + |y|^C)$$

where $C_f(x,y) = \int_0^1 \sup_{\theta\in\Theta} \left| \partial_x f(x + \xi(y-x),\theta) \right| d\xi$ for $x, y \in \mathbb{R}^{\mathsf{d}}$. Then, by $[F1]_0$ (i) and (ii), we obtain

$$\begin{aligned}\left\|n^{\epsilon}\sup_{\theta\in\Theta}|I^{(k)}_{4,n}(\theta)|\right\|_p \leq & n^{\epsilon}\times\frac{1}{nh}\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\|1_{\{\Delta_jN^X=0\}}C_f(X_{t_{j-1}},X_t)|X_t-X_{t_{j-1}}|\|_p\,\mathrm{d}t\\ &+n^{\epsilon}\left\|\frac{1}{nh}\sum_{j=1}^{n}1_{\{\Delta_jN^X\neq 0\}}\int_{t_{j-1}}^{t_j}C_f(X_{t_{j-1}},X_t)|X_t-X_{t_{j-1}}|\,\mathrm{d}t\right\|_p\\ \lesssim & n^{-\frac{1}{2}+\epsilon}+n^{-\frac{1}{2}+\epsilon}\left\|(N^X_T)^{\frac{1}{2}}\left\{n^{-1}\sum_{j=1}^{n}\left(h^{-1}\int_{t_{j-1}}^{t_j}C_f(X_{t_{j-1}},X_t)|X_t-X_{t_{j-1}}|\,\mathrm{d}t\right)^2\right\}^{\frac{1}{2}}\right\|_p\\ \lesssim & n^{-\frac{1}{2}+\epsilon}+n^{-\frac{1}{2}+\epsilon}\|N^X_T\|_p^{\frac{1}{2}}\\ = & O(n^{-\frac{1}{2}+\epsilon})\end{aligned}$$

for every $p\geq 1$. This completes the proof. □

By L^p-estimate, we obtain the following lemma.

Lemma 6 *Let* $k\in\{1,...,\mathsf{k}\}$ *and let* $f\in C^{0,1}_{\uparrow}\left(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R}^{\mathsf{m}_k}\otimes\mathbb{R}^{\mathsf{m}_k}\right)$. *Suppose that* $[F1]_0$ *is fulfilled. Then,*

$$\sup_{n\in\mathbb{N}}\left\|\sup_{\theta\in\Theta}n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}^{(k)}_n(\alpha^{(k)})}f(X_{t_{j-1}},\theta)\left[\left(\Delta_jY^{(k)}\right)^{\otimes 2}K^{(k)}_{n,j}-\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_jw^{(k)}\right)^{\otimes 2}\right]\right|\right\|_p<\infty$$

for every $p\geq 1$ *and* $\epsilon>0$.

Proof Let $\widetilde{Y}^{(k)}=Y^{(k)}-J^{(k)}$. Let $\check{N}=N+N^X$. Let

$$Q_j=\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_jw^{(k)}\right)^{\otimes 2}.$$

Then

$$\begin{aligned}&\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}^{(k)}_n(\alpha^{(k)})}1_{\{\Delta_j\check{N}>0\}}f(X_{t_{j-1}},\theta)\left[\left(\Delta_jY^{(k)}\right)^{\otimes 2}K^{(k)}_{n,j}-Q_j\right]\right|\right\|_p\\ \leq&\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\max_{j=1,...,n}\left|f(X_{t_{j-1}},\theta)\left[\left(\Delta_jY^{(k)}\right)^{\otimes 2}K^{(k)}_{n,j}-Q_j\right]\right|\right\|_{2p}\|\check{N}_T\|_{2p}\\ =&o(1)\end{aligned}\tag{8}$$

as $n\to\infty$ thanks to $K^{(k)}_{n,j}$.

Let $\eta=1-\epsilon/2$. Then, by the Burkholder–Davis–Gundy inequality, for any $L\geq 2$,

$$\begin{aligned}
P_n :=& P\left[\max_{j=1,\ldots,n}\left|1_{\{\Delta_j\check{N}=0\}}\int_{t_{j-1}}^{t_j}\left\{\sigma(X_t,\theta^*)-\sigma(X_{t_{j-1}},\theta^*)\right\}\mathrm{d}w_t\right|>n^{-\eta}\right]\\
\le& P\left[\max_{j=1,\ldots,n}\left|\int_{t_{j-1}}^{t_j}\left\{\sigma(\widetilde{X}_t+J^X_{t_{j-1}},\theta^*)-\sigma(X_{t_{j-1}},\theta^*)\right\}\mathrm{d}w_t\right|>n^{-\eta}\right]\\
\lesssim& \sum_{j=1}^n n^{L\eta}E\left[\left(\int_{t_{j-1}}^{t_j}\left|\sigma(\widetilde{X}_t+J^X_{t_{j-1}},\theta^*)-\sigma(X_{t_{j-1}},\theta^*)\right|^2\mathrm{d}t\right)^{L/2}\right]\\
\le& \sum_{j=1}^n n^{L\eta}h^{L/2-1}\int_{t_{j-1}}^{t_j}E\left[\left|\sigma(\widetilde{X}_t+J^X_{t_{j-1}},\theta^*)-\sigma(\widetilde{X}_{t_{j-1}}+J^X_{t_{j-1}},\theta^*)\right|^L\right]\mathrm{d}t\\
=& O\big(n\times n^{L\eta}\times n^{-L/2+1}\times n^{-1}\times n^{-L(1/2-\epsilon/4)}\big)\\
=& O(n^{1-L\epsilon/4}).
\end{aligned}$$

In the last part, we used Taylor's formula and Hölder's inequality. Therefore, $P_n = O(n^{-L})$ for any $L > 0$.

Expand $\Delta_j\widetilde{Y}^{(k)}$ with the formula

$$\begin{aligned}
\Delta_j\widetilde{Y}^{(k)} =& \sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}+\int_{t_{j-1}}^{t_j}\left\{\sigma^{(k)}(X_t,\theta^*)-\sigma^{(k)}(X_{t_{j-1}},\theta^*)\right\}\mathrm{d}w^{(k)}_t+\int_{t_{j-1}}^{t_j}b^{(k)}_t\mathrm{d}t\\
=:& \xi_{1,j}+\xi_{2,j}+\xi_{3,j}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}^{(k)}_n(\alpha^{(k)})}1_{\{\Delta_j\check{N}=0\}}f(X_{t_{j-1}},\theta)\left[\xi_{1,j}\otimes\xi_{2,j}\right]\right|\right\|_p\\
\lesssim& n^{\frac{1}{2}-\frac{\epsilon}{2}}\sup_{\substack{j=1,\ldots,n\\ \theta\in\Theta}}\left\||f(X_{t_{j-1}},\theta)||\xi_{1,j}|\right\|_p+n^{1-\epsilon}P_n^{\frac{1}{2p}}\\
=& o(1).
\end{aligned}$$

Thus, we can see

$$\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}^{(k)}_n(\alpha^{(k)})}1_{\{\Delta_j\check{N}=0\}}f(X_{t_{j-1}},\theta)\left[\xi_{i_1,j}\otimes\xi_{i_2,j}\right]\right|\right\|_p = o(1)$$

for $(i_1,i_2)\in\{1,2,3\}^2\setminus\{(1,1)\}$. Consequently,

$$
\begin{aligned}
&\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}1_{\{\Delta_j\check{N}=0\}}f(X_{t_{j-1}},\theta)\left[\left(\Delta_jY^{(k)}\right)^{\otimes 2}K_{n,j}^{(k)}-Q_j\right]\right|\right\|_p \\
&\leq\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}1_{\{\Delta_j\check{N}=0\}}f(X_{t_{j-1}},\theta)\left[\left(\Delta_j\widetilde{Y}^{(k)}\right)^{\otimes 2}-Q_j\right]\right|\right\|_p+O(n^{-L}) \qquad (9)\\
&=o(1)
\end{aligned}
$$

for every $p>1$ and $L>0$.

From (8) and (9), we obtain

$$
\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}f(X_{t_{j-1}},\theta)\left[\left(\Delta_jY^{(k)}\right)^{\otimes 2}K_{n,j}^{(k)}-Q_j\right]\right|\right\|_p=o(1) \qquad (10)
$$

for every $p>1$. Applying the same estimate as (10) to $\partial_\theta f$ for f, we conclude the proof by Sobolev inequality. □

Lemmas 3, 4 and 6 suggest approximation of $n^{-1}\mathbb{H}_n(\theta;\alpha)$ by

$$
\begin{aligned}
-\frac{1}{2n}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\widetilde{\mathcal{J}}_n^{(k)}(\alpha^{(k)})}&\Big\{q^{(k)}(\alpha^{(k)})^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)^{1/2}S^{(k)}(X_{t_{j-1}},\theta)^{-1}\\
&\times S^{(k)}(X_{t_{j-1}},\theta^*)^{1/2}\left[\left(h^{-1/2}\Delta_jw^{(k)}\right)^{\otimes 2}\right]\\
&+p(\alpha^{(k)})^{-1}\log\det S^{(k)}(X_{t_{j-1}},\theta)\Big\},
\end{aligned}
$$

as we will see its validity below.

2.5 Polynomial-type large deviation inequality and the rate of convergence of the α-QMLE and the (α,β)-QBE

We will show convergence of the α-QMLE. To this end, we will use a polynomial-type large deviation inequality given in Theorem 1 for a random field associated with $\mathbb{H}_n(\theta;\alpha)$. Proof of Theorem 1 will be given in Sect. 2.6, based on the QLA theory (Yoshida 2011) with the aid of the global filtering lemmas in Sect. 2.4. Though the rate of convergence is less optimal, the global filter has the advantage of eliminating jumps with high precision, and we can use it as a stable initial estimator to obtain an efficient estimator later. We do not assume any restrictive condition of the distribution of small jumps though the previous jump filters required such a condition for optimal estimation.

We introduce a middle-resolution (or annealed) random field. A similar method was used in Uchida and Yoshida (2012) to relax the so-called balance condition between the number of observations and the discretization step for an ergodic diffusion model. For $\beta \in (0, \gamma_0)$, let

$$\mathbb{H}_n^\beta(\theta;\alpha) = n^{-1+2\beta}\mathbb{H}_n(\theta;\alpha). \tag{11}$$

The random field $\mathbb{H}_n^\beta(\theta;\alpha)$ mitigates the sharpness of the contrast $\mathbb{H}_n(\theta;\alpha)$. Let

$$\mathbb{Y}_n(\theta;\alpha) = n^{-2\beta}\left\{\mathbb{H}_n^\beta(\theta;\alpha) - \mathbb{H}_n^\beta(\theta^*;\alpha)\right\} = n^{-1}\left\{\mathbb{H}_n(\theta;\alpha) - \mathbb{H}_n(\theta^*;\alpha)\right\}.$$

Let

$$\mathbb{Y}(\theta) = -\frac{1}{2T}\sum_{k=1}^{\mathsf{k}}\int_0^T \left\{\mathrm{Tr}\left(S^{(k)}(X_t,\theta)^{-1}S^{(k)}(X_t,\theta^*) - I_{\mathsf{m}_k}\right)\right. \\ \left. + \log\frac{\det S^{(k)}(X_t,\theta)}{\det S^{(k)}(X_t,\theta^*)}\right\}\mathrm{d}t.$$

The key index χ_0 is defined by

$$\chi_0 = \inf_{\theta\neq\theta^*}\frac{-\mathbb{Y}(\theta)}{|\theta-\theta^*|^2}.$$

Nondegeneracy of χ_0 plays an essential role in the QLA.

[F3] For every positive number L, there exists a constant C_L such that

$$P\left[\chi_0 < r^{-1}\right] \leq C_L\, r^{-L} \qquad (r > 0).$$

Remark 4 An analytic criterion and a geometric criterion are known to ensure Condition [$F3$] when X is a nondegenerate diffusion process. See Uchida and Yoshida (2013) for details. Since the proof of this fact depends on short-time asymptotic properties, we can modify it by taking the same approach before the first jump even when X has finitely active jumps. Details will be provided elsewhere. On the other hand, those criteria can apply to the jump diffusion X without remaking them if we work under localization. See Sect. 5.

Let $\mathbb{U}_n^\beta = \{u \in \mathbb{R}^{\mathsf{p}}; \theta^* + n^{-\beta}u \in \Theta\}$. Let $\mathbb{V}_n^\beta(r) = \{u \in \mathbb{U}_n^\beta; |u| \geq r\}$. The quasi-likelihood ratio random field $\mathbb{Z}_n^\beta(\cdot;\alpha)$ of order β is defined by

$$\mathbb{Z}_n^\beta(u;\alpha) = \exp\left\{\mathbb{H}_n^\beta\left(\theta^* + n^{-\beta}u;\alpha\right) - \mathbb{H}_n^\beta\left(\theta^*;\alpha\right)\right\} \qquad (u \in \mathbb{U}_n^\beta).$$

The random field $\mathbb{Z}_n^\beta(u;\alpha)$ is "annealed" since the contrast function $-\mathbb{H}_n^\beta(\theta;\alpha)$ becomes a milder penalty than $-\mathbb{H}_n(\theta;\alpha)$ because $\beta < 1/2$.

The following theorem will be proved in Sect. 2.6.

Theorem 1 *Suppose that* $[F1]_4$, $[F2]$ *and* $[F3]$ *are fulfilled. Let* $c_0 \in (1, 2)$. *Then, for every positive number L, there exists a constant* $C(\alpha, \beta, c_0, L)$ *such that*

$$P\left[\sup_{u\in\mathbb{V}_n(r)} \mathbb{Z}_n^{\beta}(u;\alpha) \geq e^{-r^{c_0}}\right] \leq \frac{C(\alpha,\beta,c_0,L)}{r^L}$$

for all $r > 0$ *and* $n \in \mathbb{N}$.

Obviously, an α-QMLE $\hat{\theta}_n^{M,\alpha}$of θ with respect to $\mathbb{H}_n(\cdot;\alpha)$ is a QMLE with respect to $\mathbb{H}_n^{\beta}(\cdot;\alpha)$. The following rate of convergence is a consequence of Theorem 1, as usual in the QLA theory.

Proposition 1 *Suppose that* $[F1]_4$, [F2] *and* [F3] *are satisfied. Then,* $\sup_{n\in\mathbb{N}} \left\|n^{\beta}\left(\hat{\theta}_n^{M,\alpha} - \theta^*\right)\right\|_p < \infty$ *for every* $p > 1$ *and every* $\beta < \gamma_0$.

The **(α,β)-quasi-Bayesian estimator** ((α,β)-QBE) $\hat{\theta}_n^{B,\alpha,\beta}$ of θ is defined by

$$\hat{\theta}_n^{B,\alpha,\beta} = \left[\int_{\Theta} \exp\left(\mathbb{H}_n^{\beta}(\theta;\alpha)\right)\varpi(\theta)\mathrm{d}\theta\right]^{-1} \int_{\Theta} \theta \exp\left(\mathbb{H}_n^{\beta}(\theta;\alpha)\right)\varpi(\theta)\mathrm{d}\theta, \tag{12}$$

where ϖ is a continuous function on Θ satisfying $0 < \inf_{\theta\in\Theta} \varpi(\theta) \leq \sup_{\theta\in\Theta} \varpi(\theta) < \infty$. Once again, Theorem 1 ensures $L^{\infty-}$-boundedness of the error of the (α,β)-QBE:

Proposition 2 *Suppose that* $[F1]_4$, [F2] *and* [F3] *are satisfied. Let* $\beta \in (0,\gamma_0)$. *Then,*

$$\sup_{n\in\mathbb{N}} \left\|n^{\beta}\left(\hat{\theta}_n^{B,\alpha,\beta} - \theta^*\right)\right\|_p < \infty$$

for every $p > 1$.

Proof Let $\hat{u}_n^{B,\alpha,\beta} = n^{\beta}\left(\hat{\theta}_n^{B,\alpha,\beta} - \theta^*\right)$. Then,

$$\hat{u}_n^{B,\alpha,\beta} = \left(\int_{\mathbb{U}_n^{\beta}} \mathbb{Z}_n^{\beta}(u;\alpha)\varpi(\theta^* + n^{-\beta}u)\mathrm{d}u\right)^{-1} \int_{\mathbb{U}_n^{\beta}} u\, \mathbb{Z}_n^{\beta}(u;\alpha)\varpi(\theta^* + n^{-\beta}u)\mathrm{d}u;$$

recall $\mathbb{U}_n^{\beta} = \{u \in \mathbb{R}^{\mathsf{p}}; \theta^* + n^{-\beta}u \in \Theta\}$.

Let $C_1 > 0$, $p > 1$, $L > p + 1$ and $D > \mathsf{p} + p$. In what follows, we take a sufficiently large positive constant C_1'. We have

$$
\begin{aligned}
E\big[|\hat{u}_n^{B,\alpha,\beta}|^p\big] &\le E\left[\left(\int_{\mathbb{U}_n^\beta} \mathbb{Z}_n^\beta(u;\alpha)\varpi(\theta^* + n^{-\beta}u)\mathrm{d}u\right)^{-1} \int_{\mathbb{U}_n^\beta} |u|^p \mathbb{Z}_n^\beta(u;\alpha)\varpi(\theta^* + n^{-\beta}u)\mathrm{d}u\right] \quad \text{(Jensen's inequality, } p \ge 1\text{)} \\
&\le C(\varpi)\sum_{r=1}^{\infty}(r+1)^p\Bigg\{E\Bigg[\left(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\right)^{-1} \\
&\qquad \times \int_{\{u;r<|u|\le r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\, 1_{\left\{\int_{\{u;r<|u|\le r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u > \frac{C_1'}{r^{D-\mathsf{p}+1}}\right\}}\Bigg] \\
&\quad + E\Bigg[\left(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\right)^{-1}\int_{\{u;r<|u|\le r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u \\
&\qquad\qquad \times 1_{\left\{\int_{\{u;r<|u|\le r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u \le \frac{C_1'}{r^{D-\mathsf{p}+1}}\right\}}\Bigg]\Bigg\} + C(\varpi)
\end{aligned}
$$

(The last term is for $r = 0$. The integrand is not greater than one.)

$$
\begin{aligned}
&\le C(\varpi)\sum_{r=1}^{\infty}(r+1)^p\Bigg\{P\left[\int_{\{u;r<|u|\le r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u > \frac{C_1'}{r^{D-\mathsf{p}+1}}\right] \\
&\qquad\qquad + \frac{C_1'}{r^{D-\mathsf{p}+1}}E\left[\left(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\right)^{-1}\right]\Bigg\} + C(\varpi) \\
&\le C(\varpi)\sum_{r=1}^{\infty}(r+1)^p\Bigg\{P\left[\sup_{u\in\mathbb{V}_n^\beta(r)}\mathbb{Z}_n^\beta(u;\alpha)) > \frac{C_1}{r^D}\right] \\
&\qquad\qquad + \frac{C_1'}{r^{D-\mathsf{p}+1}}E\left[\left(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\right)^{-1}\right]\Bigg\} + C(\varpi) \\
&\lesssim \sum_{r=1}^{\infty} r^{-(L-p)} + \sum_{r=1}^{\infty} r^{-(D-p-\mathsf{p}+1)}E\left[\left(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\right)^{-1}\right] + C(\varpi). \\
&< \infty
\end{aligned}
$$

by Theorem 1, suppose that

$$
E\left[\left(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\mathrm{d}u\right)^{-1}\right] < \infty. \tag{13}
$$

However, one can show (13) by using Lemma 2 of Yoshida (2011). □

2.6 Proof of Theorem 1

We will prove Theorem 1 by Theorem 2 of Yoshida (2011) with the aid of the global filtering lemmas in Sect. 2.4. Choose parameters η, β_1, ρ_1, ρ_2 and β_2 satisfying the following inequalities:

$$
\begin{aligned}
&0 < \eta < 1, \qquad 0 < \beta_1 < \frac{1}{2}, \qquad 0 < \rho_1 < \min\{1, \eta(1-\eta)^{-1}, 2\beta_1(1-\eta)^{-1}\}, \\
&2\eta < \rho_2, \qquad \beta_2 \ge 0, \qquad 1 - 2\beta_2 - \rho_2 > 0.
\end{aligned} \tag{14}
$$

Let

$$
\Delta_n(\alpha,\beta) = n^{-\beta}\partial_\theta\mathbb{H}_n^\beta(\theta^*;\alpha) \;=\; n^{-1+\beta}\partial_\theta\mathbb{H}_n(\theta^*;\alpha).
$$

Let

$$\Gamma_n(\alpha) = -\,n^{-2\beta}\partial_\theta^2\mathbb{H}_n^\beta(\theta^*;\alpha) \;=\; -n^{-1}\partial_\theta^2\mathbb{H}_n(\theta^*;\alpha).$$

The $\mathsf{p}\times\mathsf{p}$ symmetric matrix $\Gamma^{(k)}$ is defined by the following formula:

$$\Gamma^{(k)}[u^{\otimes 2}] = \frac{1}{2T}\int_0^T \mathrm{Tr}\Big((\partial_\theta S^{(k)}[u])(S^{(k)})^{-1}(\partial_\theta S^{(k)}[u])(S^{(k)})^{-1}(X_t,\theta^*)\Big)\mathrm{d}t,$$

where $u\in\mathbb{R}^{\mathsf{p}}$, and Γ by $\Gamma=\sum_{k=1}^{\mathsf{k}}\Gamma^{(k)}$. We will need several lemmas. We choose positive constants γ_i $(i=1,2)$ so that $\beta<\gamma_2<\gamma_1<\gamma_0$. Then, we can choose parameters $\beta_1(\downarrow 0)$, $\beta_2(\uparrow 1/2)$, $\rho_2(\downarrow 0)$, $\eta(\downarrow 0)$ and $\rho_1(\downarrow 0)$ so that $\max\{2\beta\beta_1,\beta(1-2\beta_2)\}<\gamma_2$. Then, there is an $\epsilon\in(\max\{2\beta\beta_1,\beta(1-2\beta_2)\},\gamma_2)$.

Lemma 7 *For every $p\ge 1$,*

$$\sup_{n\in\mathbb{N}} E\left[\left(n^{-2\beta}\sup_{\theta\in\Theta}\left|\partial_\theta^3\mathbb{H}_n^\beta(\theta;\alpha)\right|\right)^p\right]<\infty.$$

Proof We have $\mathbb{H}_n(\theta;\alpha)=\mathbb{H}_n^\circ(\theta;\alpha)+\mathbb{M}^\circ(\theta;\alpha)+\mathbb{R}^\circ(\theta;\alpha)$, where

$$\mathbb{H}_n^\circ(\theta;\alpha) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} p(\alpha^{(k)})^{-1}\Big\{S^{(k)}(X_{t_{j-1}},\theta)^{-1}\big[S^{(k)}(X_{t_{j-1}},\theta^*)\big] + \log\det S^{(k)}(X_{t_{j-1}},\theta)\Big\},$$

$$\mathbb{M}_n^\circ(\theta;\alpha) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1} \cdot\Big[q^{(k)}(\alpha^{(k)})^{-1}\big(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\big)^{\otimes 2} - hp(\alpha^{(k)})^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)\Big]$$

and

$$\mathbb{R}_n^\circ(\theta;\alpha) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} q^{(k)}(\alpha^{(k)})^{-1}h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1}\Big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}K_{n,j}^{(k)} - \big(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\big)^{\otimes 2}\Big].$$

Apply Lemma 6 to $\partial_\theta^i\mathbb{R}_n^\circ(\theta;\alpha)$ $(i=0,\dots,3)$ to obtain

$$\sum_{i=0}^{3}\left\|\sup_{\theta\in\Theta}\left|\partial_\theta^i n^{-1}\mathbb{R}_n^\circ(\theta;\alpha)\right|\right\|_p<\infty$$

for every $p>1$. Moreover, we apply Sobolev inequality, Lemma 3 (ii) and Lemma 4 (ii). Then, it is sufficient to show that

$$\sum_{i=0}^{4} \sup_{\theta\in\Theta} \left\{ \left\|\partial_\theta^i n^{-1}\mathbb{H}_n^\times(\theta;\alpha)\right\|_p + \left\|\partial_\theta^i n^{-1}\mathbb{M}_n^\times(\theta;\alpha)\right\|_p \right\} < \infty \tag{15}$$

for proving the lemma, where $\mathbb{H}_n^\times(\theta;\alpha)$ and $\mathbb{M}_n^\times(\theta;\alpha)$ are defined by the same formula as $\mathbb{H}_n^\circ(\theta;\alpha)$ and $\mathbb{M}_n^\circ(\theta;\alpha)$, respectively, but with $\widetilde{\mathcal{J}}_n^{(k)}(\alpha^{(k)})$ in place of $\mathcal{J}_n^{(k)}(\alpha^{(k)})$. However, (15) is obvious. □

Lemma 8 *For every* $p \geq 1$,

$$\sup_{n\in\mathbb{N}} E\left[\left(n^{2\beta\beta_1}\left|\Gamma_n(\alpha) - \Gamma\right|\right)^p\right] < \infty.$$

Proof Consider the decomposition $\Gamma_n(\alpha) = \Gamma_n^* + M_n^* + R_n^*$ with

$$\Gamma_n^* = \frac{1}{2n}\sum_{k=1}^{\mathsf{k}} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} p(\alpha^{(k)})^{-1}\left\{\partial_\theta^2 \log\det S^{(k)}(X_{t_{j-1}},\theta^*) + \left(\partial_\theta^2(S^{(k)\,-1})\right)(X_{t_{j-1}},\theta^*)\left[S(X_{t_{j-1}},\theta^*)\right]\right\},$$

$$M_n^* = \frac{1}{2n}\sum_{k=1}^{\mathsf{k}} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} \left(\partial_\theta^2(S^{(k)\,-1})\right)(X_{t_{j-1}},\theta^*) \cdot \left[q^{(k)}(\alpha^{(k)})^{-1}h^{-1}\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\right)^{\otimes 2} - p(\alpha^{(k)})^{-1}S(X_{t_{j-1}},\theta^*)\right]$$

and

$$R_n^* = \frac{1}{2n}\sum_{k=1}^{\mathsf{k}} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} q^{(k)}(\alpha^{(k)})^{-1}h^{-1}\left(\partial_\theta^2(S^{(k)\,-1})\right)(X_{t_{j-1}},\theta^*) \cdot \left[\left(\Delta_j Y^{(k)}\right)^{\otimes 2}K_{n,j}^{(k)} - \left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\right)^{\otimes 2}\right].$$

Since $2\beta\beta_1 < \gamma_2$, we obtain

$$\sup_{n\in\mathbb{N}} \left\|n^{2\beta\beta_1}\left|\Gamma_n^* - \Gamma\right|\right\|_p < \infty$$

by Lemma 5, and also obtain

$$\sup_{n\in\mathbb{N}} \left\|n^{2\beta\beta_1}\left|R_n^*\right|\right\|_p < \infty$$

by Lemma 6 for every $p > 1$. Moreover, by Lemmas 3 (ii) and 4 (ii) applied to $2\beta\beta_1(<\gamma_2)$ for "γ_3," we replace $\mathcal{J}_n^{(k)}(\alpha^{(k)})$ in the expression of M_n^* by $\widetilde{\mathcal{J}}_n^{(k)}(\alpha^{(k)})$ and then apply the Burkholder–Davis–Gundy inequality to show

$$\sup_{n\in\mathbb{N}} \left\|n^{2\beta\beta_1}\left|M_n^*\right|\right\|_p < \infty$$

for every $p > 1$. This completes the proof. □

The following two lemmas are obvious under [$F3$].

Lemma 9 *For every* $p \geq 1$, *there exists a constant* C_p *such that*

$$P\left[\lambda_{\min}(\Gamma) < r^{-\rho_1}\right] \leq \frac{C_p}{r^p}$$

for all $r > 0$, *where* $\lambda_{\min}(\Gamma)$ *denotes the minimum eigenvalue of* Γ.

Lemma 10 *For every* $p \geq 1$, *there exists a constant* C_p *such that*

$$P\left[\chi_0 < r^{-(\rho_2 - 2\eta)}\right] \leq \frac{C_p}{r^p}$$

for all $r > 0$.

Lemma 11 *For every* $p \geq 1$,

$$\sup_{n\in\mathbb{N}} E\left[\left|\Delta_n(\alpha,\beta)\right|^p\right] < \infty.$$

Proof We consider the decomposition $\Delta_n(\alpha,\beta) = n^{-1+\beta}\partial_\theta \mathbb{H}_n(\theta^*;\alpha) = M_n^\vee + R_n^\vee$ with

$$M_n^\vee = -\frac{n^\beta}{2n}\sum_{k=1}^{\mathsf{k}} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} \left(\partial_\theta(S^{(k)\,-1})\right)(X_{t_{j-1}},\theta^*) \cdot \left[q^{(k)}(\alpha^{(k)})^{-1}h^{-1}\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\right)^{\otimes 2} - p(\alpha^{(k)})^{-1}S(X_{t_{j-1}},\theta^*)\right]$$

and

$$R_n^\vee = -\frac{n^\beta}{2n}\sum_{k=1}^{\mathsf{k}} \sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} q^{(k)}(\alpha^{(k)})^{-1}h^{-1} \times \left(\partial_\theta(S^{(k)\,-1})\right)(X_{t_{j-1}},\theta^*)\left[\left(\Delta_j Y^{(k)}\right)^{\otimes 2}K_{n,j}^{(k)} - \left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\right)^{\otimes 2}\right].$$

We see $\sup_{n\in\mathbb{N}}\left\|R_n^\vee(\alpha,\beta)\right\|_p < \infty$ by Lemma 6. Moreover, $\sup_{n\in\mathbb{N}}\left\|M_n^\vee(\alpha,\beta)\right\|_p < \infty$ by Lemmas 3 (ii) and 4 (ii) and the Burkholder–Davis–Gundy inequality. We note that symmetry between the components of $W_j^{(k)}$ is available. □

As a matter of fact, $\Delta_n(\alpha,\beta)$ converges to 0, as seen in the proof of Lemma 11. The location shift of the random field $\mathbb{Z}_n^\beta(\cdot;\alpha)$ asymptotically vanishes.

Lemma 12 *For every $p \geq 1$,*

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta} n^{\beta(1-2\beta_2)}\left|\mathbb{Y}_n(\theta;\alpha)-\mathbb{Y}(\theta)\right|\right)^p\right]<\infty.$$

Proof In this situation, we use the decomposition

$$\mathbb{Y}_n(\theta;\alpha)=\mathbb{Y}_n^+(\theta;\alpha)+\mathbb{M}_n^+(\theta;\alpha)+\mathbb{R}_n^+(\theta;\alpha)$$

with

$$\begin{aligned}\mathbb{Y}_n^+(\theta;\alpha)=&-\frac{1}{2n}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}p(\alpha^{(k)})^{-1}\Bigg\{\mathrm{Tr}\left(S^{(k)}(X_{t_{j-1}},\theta)^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)-I_{\mathsf{m}_k}\right)\\&+\log\frac{\det S^{(k)}(X_{t_{j-1}},\theta)}{\det S^{(k)}(X_{t_{j-1}},\theta^*)}\Bigg\},\\ \mathbb{M}_n^+(\theta;\alpha)=&-\frac{1}{2n}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}\left(S^{(k)}(X_{t_{j-1}},\theta)^{-1}-S^{(k)}(X_{t_{j-1}},\theta^*)^{-1}\right)\\&\cdot\left[q^{(k)}(\alpha^{(k)})^{-1}h^{-1}\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_jw^{(k)}\right)^{\otimes2}-p(\alpha^{(k)})^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)\right]\end{aligned}$$

and

$$\begin{aligned}\mathbb{R}_n^+(\theta;\alpha)=&-\frac{1}{2n}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}q^{(k)}(\alpha^{(k)})^{-1}h^{-1}\left(S^{(k)}(X_{t_{j-1}},\theta)^{-1}-S^{(k)}(X_{t_{j-1}},\theta^*)^{-1}\right)\\&\cdot\left[\left(\Delta_jY^{(k)}\right)^{\otimes2}K_{n,j}^{(k)}-\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_jw^{(k)}\right)^{\otimes2}\right].\end{aligned}$$

As assumed, $\beta(1-2\beta_2)<\gamma_2\leq1/2$. Lemma 6 gives

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta} n^{\beta(1-2\beta_2)}\left|\mathbb{R}_n^+(\theta;\alpha)\right|\right)^p\right]<\infty$$

for every $p>1$. Furthermore, Lemma 5 gives

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta} n^{\beta(1-2\beta_2)}\left|\mathbb{Y}_n^+(\theta;\alpha)-\mathbb{Y}(\theta)\right|\right)^p\right]<\infty.$$

On the other hand, Lemmas 3 (ii) and 4 (ii) and the Burkholder–Davis–Gundy inequality together with Sobolev inequality deduce

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta} n^{\beta(1-2\beta_2)}\left|\mathbb{M}_n^+(\theta;\alpha)\right|\right)^p\right]<\infty$$

for every $p>1$, which completes the proof. □

Proof of Theorem 1 Now Theorem 1 follows from Theorem 2 of Yoshida (2011) combined with Lemmas 7, 8, 9, 10, 11 and 12. □

3 Global filter with moving threshold

3.1 Quasi-likelihood function with moving quantiles

Though the threshold method presented in the previous section removes jumps surely, it is conservative and does not attain the optimal rate of convergence that is attained by the QLA estimators (i.e., QMLE and QBE) in the case without jumps. On the other hand, it is possible to give more efficient estimator by aggressively taking bigger increments while it may cause miss detection of certain portion of jumps.

Let $\delta_0 \in (0, 1/4)$ and $\delta_1^{(k)} \in (0, 1/2)$. For simplicity, let $s_n^{(k)} = n - B^{(k)}\lfloor n^{\delta_1^{(k)}} \rfloor$ with positive constants $B^{(k)}$. Let $\alpha_n^{(k)} = 1 - s_n^{(k)}/n$ and $\alpha_n = (\alpha_n^{(1)}, \ldots, \alpha_n^{(\mathsf{k})})$. Let

$$\mathcal{K}_n^{(k)} = \left\{ j \in \{1, \ldots, n\}; V_j^{(k)} < V_{(s_n^{(k)})}^{(k)} \right\}$$

where

$$V_j^{(k)} = |(\mathfrak{S}_{n,j-1}^{(k)})^{-1/2} \Delta_j Y^{(k)}|$$

with some positive definite random matrix $\mathfrak{S}_{n,j-1}^{(k)}$, and $V_{(j)}^{(k)}$ is the jth-order statistic of $V_1^{(k)}, \ldots, V_n^{(k)}$.

We consider a random field by removing increments of Y including jumps from the full quasi-likelihood function. Define $\mathbb{H}_n(\theta)$ by

$$\begin{aligned}\mathbb{H}_n(\theta) = -\frac{1}{2} \sum_{k=1}^{\mathsf{k}} \sum_{j \in \mathcal{K}_n^{(k)}} \Big\{ &(q_n^{(k)})^{-1} h^{-1} S^{(k)}(X_{t_{j-1}}, \theta)^{-1} \left[\left(\Delta_j Y^{(k)}\right)^{\otimes 2} \right] K_{n,j}^{(k)} \\ &+ (p_n^{(k)})^{-1} \log \det S^{(k)}(X_{t_{j-1}}, \theta) \Big\}.\end{aligned} \tag{16}$$

Remark 5 The truncation functional $K_{n,j}^{(k)}$ is given by (3). It is also reasonable to set it as

$$K_{n,j}^{(k)} = 1_{\left\{ V_j^{(k)} < C_*^{(k)} n^{-\frac{1}{4} - \delta_0} \right\}},$$

where $C_*^{(k)}$ is an arbitrarily given positive constant.

Remark 6 The threshold is larger than $n^{-\frac{1}{2}+0}$. The truncation $K_{n,j}^{(k)}$ is for stabilizing the increments of Y, not for filtering. The factors $\mathfrak{S}_{n,j-1}^{(k)}$, $q_n^{(k)}$ and $p_n^{(k)}$ can freely be chosen if $\mathfrak{S}_{n,j-1}^{(k)}$ and its inverse are uniformly bounded in $L^{\infty-}$ and if $q_n^{(k)}$ and $p_n^{(k)}$ are sufficiently close to 1. $\bar{S}_{n,j-1}^{(k)}$, $q^{(k)}(\alpha_n^{(k)})$ and $p(\alpha_n^{(k)})$ are natural choices for $\mathfrak{S}_{n,j-1}^{(k)}$, $q_n^{(k)}$

and $p_n^{(k)}$, respectively. Asymptotic theoretically, the factors $(q_n^{(k)})^{-1}$ and $(p_n^{(k)})^{-1}$ can be replaced by 1, and one can take $\mathfrak{S}_{n,j-1}^{(k)} = I_{\mathsf{m}_\mathsf{k}}$; see Condition [$F2'$]. Thus, a modification of $\mathbb{H}_n(\theta)$ is $\mathring{\mathbb{H}}_n(\theta)$ defined by

$$\mathring{\mathbb{H}}_n(\theta) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{K}_n^{(k)}}\Big\{h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1}\big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}\big]K_{n,j}^{(k)} + \log\det S^{(k)}(X_{t_{j-1}},\theta)\Big\}$$

with $\mathcal{K}_n^{(k)}$ for $V_j^{(k)} = |\Delta_j Y^{(k)}|$. The quasi-log likelihood function $\mathring{\mathbb{H}}_n$ gives the same asymptotic results as $\mathbb{H}_n$.

We denote by $\hat{\theta}_n^{M,\alpha_n}$ a QMLE of θ with respect to $\mathbb{H}_n$ given by (16). We should remark that $\hat{\theta}_n^{M,\alpha_n}$ defined by $\mathbb{H}_n(\theta)$ can differ from $\hat{\theta}_n^{M,\alpha}$ previously defined by $\mathbb{H}_n(\theta;\alpha)$. The **quasi-Bayesian estimator** (QBE) $\hat{\theta}_n^{B,\alpha_n}$ of θ is defined by

$$\hat{\theta}_n^{B,\alpha_n} = \left[\int_\Theta \exp\big(\mathbb{H}_n(\theta)\big)\varpi(\theta)\mathrm{d}\theta\right]^{-1}\int_\Theta \theta\exp\big(\mathbb{H}_n(\theta)\big)\varpi(\theta)d\theta,$$

where ϖ is a continuous function on Θ satisfying $0 < \inf_{\theta\in\Theta}\varpi(\theta) \le \sup_{\theta\in\Theta}\varpi(\theta) < \infty$.

3.2 Polynomial-type large deviation inequality

Let $\mathbb{U}_n = \{u \in \mathbb{R}^{\mathsf{p}}; \theta^* + n^{-1/2}u \in \Theta\}$. Let $\mathbb{V}_n(r) = \{u \in \mathbb{U}_n; |u| \ge r\}$. We define the quasi-likelihood ratio random field $\mathbb{Z}_n$ by

$$\mathbb{Z}_n(u) = \exp\Big\{\mathbb{H}_n(\theta^* + n^{-1/2}u) - \mathbb{H}_n(\theta^*)\Big\} \qquad (u \in \mathbb{U}_n).$$

[F2$'$] (i) The positive-definite measurable random matrices $\mathfrak{S}_{n,j-1}^{(k)}$ $(k \in \{1,...,\mathsf{k}\},\ n \in \mathbb{N},\ j \in \{1,...,n\})$ satisfy

$$\sup_{\substack{k\in\{1,...,\mathsf{k}\}\\ n\in\mathbb{N},\, j\in\{1,...,n\}}}\Big(\|\mathfrak{S}_{n,j-1}^{(k)}\|_p + \|(\mathfrak{S}_{n,j-1}^{(k)})^{-1}\|_p\Big) < \infty$$

for every $p > 1$.

(ii) Positive numbers $q_n^{(k)}$ and $p_n^{(k)}$ satisfy $|q_n^{(k)} - 1| = o(n^{-1/2})$ and $|1 - p_n^{(k)}| = o(n^{-1/2})$.

A polynomial-type large deviation inequality is given by the following theorem, a proof of which is in Sect. 3.3.

Theorem 2 *Suppose that* $[F1]_4$, $[F2']$ *and* $[F3]$ *are fulfilled. Let* $c_0 \in (1,2)$. *Then, for every positive number L, there exists a constant* $C(c_0, L)$ *such that*

$$P\left[\sup_{u\in\mathbb{V}_n(r)} \mathbb{Z}_n(u) \geq e^{-r^{c_0}}\right] \leq \frac{C(c_0,L)}{r^L}$$

for all $r > 0$ *and* $n \in \mathbb{N}$.

The polynomial-type large deviation inequality for $\mathbb{Z}_n$ in Theorem 2 ensures $L^{\infty-}$-boundedness of the QLA estimators.

Proposition 3 *Suppose that* $[F1]_4, [F2']$ *and* $[F3]$ *are satisfied. Then,*

$$\sup_{n\in\mathbb{N}} \left\| \sqrt{n}\left(\hat{\theta}_n^{\mathsf{A},\alpha_n} - \theta^*\right)\right\|_p < \infty \qquad (\mathsf{A} = M, B)$$

for every $p > 1$.

3.3 Proof of Theorem 2

Recall $\widetilde{Y}^{(k)} = Y^{(k)} - J^{(k)}$. Let

$$\widetilde{\mathbb{H}}_n(\theta) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j=1}^{n}\left\{h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1}\left[\left(\Delta_j\widetilde{Y}^{(k)}\right)^{\otimes 2}\right] + \log\det S^{(k)}(X_{t_{j-1}},\theta)\right\}.$$

Lemma 13 *For every* $p \geq 1$,

$$\sum_{i=0}^{4}\sup_{\theta\in\Theta}\left\|n^{-1/2}\partial_\theta^i\mathbb{H}_n(\theta) - n^{-1/2}\partial_\theta^i\widetilde{\mathbb{H}}_n(\theta)\right\|_p \to 0 \tag{17}$$

as $n \to \infty$.

Proof Let

$$\mathfrak{A}_n^{(k)} = \bigcup_{j=1}^{n}\left[\left\{j \in (\mathcal{K}_n^{(k)})^c\right\} \cap \left\{\Delta_j N^{(k)} = 0\right\}\right].$$

Let

$$\mathfrak{B}_n^{(k)} = \bigcap_{j=1}^{n}\left[\left\{V_j^{(k)} \geq V_{(s_n)}^{(k)}\right\} \cup \left\{|\Delta_j J^{(k)}| \leq n^{-\frac{1}{4}-\delta_0}\right\}\right].$$

For $\omega \in \mathfrak{A}_n^{(k)} \cap (\mathfrak{B}_n^{(k)})^c$, there exists $j(\omega) \in (\mathcal{K}_n^{(k)})^c$ such that $\Delta_{j(\omega)}N^{(k)}(\omega) = 0$, and also there exists $j'(\omega) \in \{1,...,n\}$ such that $V_{j'(\omega)}^{(k)}(\omega) < V_{(s_n)}^{(k)}(\omega)$ and $|\Delta_{j'(\omega)}J^{(k)}(\omega)| > n^{-\frac{1}{4}-\delta_0}$. Then,

$$\left|(\mathfrak{S}^{(k)}_{n,j'(\omega)-1})^{-1/2}\Delta_{j'(\omega)}J^{(k)}(\omega)\right| - \left|(\mathfrak{S}^{(k)}_{n,j'(\omega)-1}(\omega))^{-1/2}\Delta_{j'(\omega)}\widetilde{Y}^{(k)}(\omega)\right|$$
$$\leq V^{(k)}_{j'(\omega)}(\omega) \ < \ V^{(k)}_{j(\omega)}(\omega) \ = \ \left|(\mathfrak{S}^{(k)}_{n,j(\omega)-1}(\omega))^{-1/2}\Delta_{j(\omega)}\widetilde{Y}^{(k)}(\omega)\right|$$

and hence

$$n^{-\frac{1}{4}-\delta_0} \leq 2\left|\mathfrak{S}^{(k)}_{n,j'(\omega)-1}\right|^{1/2}\max_{j=1,..,n}\left|(\mathfrak{S}^{(k)}_{n,j-1}(\omega))^{-1/2}\Delta_j\widetilde{Y}^{(k)}(\omega)\right|$$

where $\left|M\right| = \{\mathrm{Tr}(MM^{\star})\}^{1/2}$ for a matrix M. Since $\left\{h^{-1/2}\left|\Delta_j\widetilde{Y}^{(k)}\right|; j=1,...,n,\ n\in\mathbb{N}\right\}$ is bounded in $L^{\infty-}$, we obtain

$$P\left[\mathfrak{A}^{(k)}_n \cap (\mathfrak{B}^{(k)}_n)^c\right] = O(n^{-L})$$

as $n\to\infty$ for every $L>0$. Moreover, $P[(\mathfrak{A}^{(k)}_n)^c] = O(n^{-L})$ from the assumption for $N^{(k)}$ since

$$\left(\mathfrak{A}^{(k)}_n\right)^c \subset \left\{\#\left\{j\in\{1,...,n\};\ \Delta_j N^{(k)}\neq 0\right\}\ \geq\ n - s^{(k)}_n + 1\right\}$$
$$\subset \left\{N^{(k)}_T \geq B^{(k)} n^{\delta^{(k)}_1}\right\}.$$

Thus,

$$P\left[\bigcap_{k=1}^{\mathsf{k}}\mathfrak{B}^{(k)}_n\right] = 1 - O(n^{-L}) \tag{18}$$

as $n\to\infty$ for every $L>0$.

Define $\mathbb{H}^{\dagger}_n(\theta)$ by

$$\mathbb{H}^{\dagger}_n(\theta) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{K}^{(k)}_n}\left\{(q^{(k)}_n)^{-1}h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1}[\left(\Delta_j Y^{(k)} - \Delta_j J^{(k)}\right)^{\otimes 2}]\right.$$
$$\left.\times K^{(k)}_{n,j}1_{\left\{|\Delta_j J^{(k)}|\leq 1\right\}} + (p^{(k)}_n)^{-1}\log\det S^{(k)}(X_{t_{j-1}},\theta)\right\},$$

where the indicator function controls the moment outside of $\cap_{k=1}^{\mathsf{k}}\mathfrak{B}^{(k)}_n$. Then, by (18), the cap and $N_T\in L^{\infty-}$, we obtain

$$\sum_{i=0}^{4}\sup_{\theta\in\Theta}\left\|n^{-1/2}\partial^i_\theta\mathbb{H}_n(\theta) - n^{-1/2}\partial^i_\theta\mathbb{H}^{\dagger}_n(\theta)\right\|_p \to 0$$

as $n\to\infty$ for every $p\geq 1$. Indeed, we can estimate this difference of the two variables on the event $\mathfrak{C}_n := \cap_{k=1}^{\mathsf{k}}\mathfrak{B}^{(k)}_n$ and on $\mathfrak{C}^c_n$ as follows. On $\mathfrak{C}_n$, $|\Delta_j J^{(k)}| \leq n^{-1/4-\delta_0}1_{\{\Delta_j J^{(k)}\neq 0\}}$ whenever $j\in\mathcal{K}^{(k)}_n$. The cap $K^{(k)}_{n,j}$ also offers the estimate $|\Delta_j Y^{(k)}| < C^{(k)}_* n^{-1/4}$. On $\mathfrak{C}_n$, after removing the factor $1_{\left\{|\Delta_j J^{(k)}|\leq 1\right\}}$ from the expression of $n^{-1/2}\partial_\theta^i \mathbb{H}_n^\dagger(\theta)$ with the help of $N_T \in L^{\infty-}$ and the L^p-estimate of $h^{-1}|\Delta_j \widetilde{Y}|^2$, we can estimate the cross term in the difference with

$$n^{-1/2}\sum_{j\in\mathcal{K}_n^{(k)}}\left|h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1}[\Delta_j Y^{(k)}\otimes\Delta_j J^{(k)}]K_{n,j}^{(k)}\right|$$
$$\leq \mathcal{M}_n^{(k)} n^{-\delta_0}\sum_{j=1}^n 1_{\{\Delta_j J^{(k)}\neq 0\}} \leq \left(n^{\delta_0/2}+\mathcal{M}_n^{(k)}1_{\{\mathcal{M}_n>n^{\delta_0/2}\}}\right)n^{-\delta_0}N_T$$

for $\mathcal{M}_n^{(k)}=\max_{j=1,\dots,n}|S^{(k)}(X_{t_{j-1}},\theta)^{-1}|$, as well as the term involving $\left(\Delta_j J^{(k)}\right)^{\otimes 2}$ and admitting a similar estimate. Estimation is much simpler on $\mathfrak{C}_n^c$ thanks to (18). The cap $1_{\left\{|\Delta_j J^{(k)}|\leq 1\right\}}$ helps.

We know that $\#(\mathcal{K}_n^{(k)})^c \sim B^{(k)}n^{\delta_1^{(k)}}$ and have assumed that $|q_n^{(k)}-1|=o(n^{-1/2})$ and that $|1-p_n^{(k)}|=o(n^{-1/2})$. Then, with (18), it is easy to show

$$\sum_{i=0}^4 \sup_{\theta\in\Theta}\left\|n^{-1/2}\partial_\theta^i\mathbb{H}_n^\dagger(\theta)-n^{-1/2}\partial_\theta^i\widetilde{\mathbb{H}}_n(\theta)\right\|_p \to 0,$$

which implies (17) as $n\to\infty$ for every $p\geq 1$. □

We choose parameters η, β_1, ρ_1, ρ_2 and β_2 satisfying (14) with $\beta_2>0$. Let

$$\Delta_n = n^{-1/2}\partial_\theta\mathbb{H}_n(\theta^*) \quad\text{and}\quad \Gamma_n = -n^{-1}\partial_\theta^2\mathbb{H}_n(\theta^*).$$

Let

$$\mathbb{Y}_n(\theta) = n^{-1}\left\{\mathbb{H}_n(\theta)-\mathbb{H}_n(\theta^*)\right\}.$$

The following two estimates will play a basic role.

Lemma 14 *Let $f\in C_\uparrow^{0,1}\left(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R}^{\mathsf{m}_k}\otimes\mathbb{R}^{\mathsf{m}_k}\right)$. Then, under $[F1]_0$,*

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta}\left|n^{\frac{1}{2}-\epsilon}\sum_{j=1}^n f(X_{t_{j-1}},\theta)\left[\left(\Delta_j\widetilde{Y}^{(k)}\right)^{\otimes 2}-\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\right)^{\otimes 2}\right]\right|\right)^p\right]<\infty$$

for every $p>1$ and $\epsilon>0$.

Proof One can validate this lemma in a quite similar way as Lemma 6. □

Lemma 15 *Let $p>1$ and $\epsilon>0$. Let $f\in C_\uparrow^{1,1}(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R})$. Suppose that $[F1]_0$ is satisfied. Then,*

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta} n^{\frac{1}{2}-\epsilon}\left|\frac{1}{n}\sum_{j=1}^n f(X_{t_{j-1}},\theta)-\frac{1}{T}\int_0^T f(X_t,\theta)\mathrm{d}t\right|\right)^p\right]<\infty.$$

Proof Let $p>1$. By taking an approach similar to the proof of Lemma 14, we obtain

$$
\begin{aligned}
&\sup_{\theta\in\Theta} n^{\frac{1}{2}-\epsilon}\left\| h\sum_{j=1}^{n} f(X_{t_{j-1}},\theta) - \int_0^T f(X_t,\theta)\mathrm{d}t \right\|_p \\
&\leq \sup_{\theta\in\Theta} n^{\frac{1}{2}-\epsilon} \sum_{j=1}^{n} \left\| \left| \int_{t_{j-1}}^{t_j} \left\{ f(X_t,\theta) - f(X_{t_{j-1}},\theta) \right\} \mathrm{d}t \right| 1_{\{\Delta_j N^X = 0\}} \right\|_p \\
&\quad + \sup_{\theta\in\Theta} n^{\frac{1}{2}-\epsilon} \left\| \max_{j=1,\dots,n} \left| \int_{t_{j-1}}^{t_j} \left\{ f(X_t,\theta) - f(X_{t_{j-1}},\theta) \right\} \mathrm{d}t \right| \right\|_{2p} \left\| E\left[N_T^X\right] \right\|_{2p} \\
&\leq O(n^{\frac{1}{2}-\epsilon} \times n \times n^{-1.5}) + o(n^{1/2-\epsilon} \times n^{-1/2+\epsilon} \times 1) \\
&= o(1)
\end{aligned}
$$

as $n \to \infty$. We also have the same estimate for $\partial_\theta f$ in place of f. Then, the Sobolev inequality implies the result. □

We have the following estimates.

Lemma 16 *For every* $p \geq 1$,

$$
\sup_{n\in\mathbb{N}} E\left[\left(n^{-1}\sup_{\theta\in\Theta}\left|\partial_\theta^3 \mathbb{H}_n(\theta)\right|\right)^p\right] < \infty.
$$

Proof Applying Lemma 13 and Sobolev inequality, one can prove the lemma in a fashion similar to Lemma 7. □

Lemma 17 *For every* $p \geq 1$,

$$
\sup_{n\in\mathbb{N}} E\left[\left(n^{\beta_1}\left|\Gamma_n - \Gamma\right|\right)^p\right] < \infty.
$$

Proof Thanks to Lemma 13, it is sufficient to show that

$$
\sup_{n\in\mathbb{N}} E\left[\left(n^{\beta_1}\left|\widetilde{\Gamma}_n - \Gamma\right|\right)^p\right] < \infty \tag{19}
$$

where

$$
\widetilde{\Gamma}_n = -n^{-1}\partial_\theta^2 \widetilde{\mathbb{H}}_n(\theta^*)
$$

Now taking a similar way as Lemma 8, one can prove the desired inequality by applying Lemmas 14 and 15 as well as the Burkholder–Davis–Gundy inequality. □

Lemma 18 *For every* $p \geq 1$, $\sup_{n\in\mathbb{N}} E\left[\left|\Delta_n\right|^p\right] < \infty$.

Proof By Lemma 13, it suffices to show

$$
\sup_{n\in\mathbb{N}} E\left[\left|\widetilde{\Delta}_n\right|^p\right] < \infty \tag{20}
$$

for

$$\widetilde{\Delta}_n = n^{-1/2}\partial_\theta \widetilde{\mathbb{H}}_n(\theta^*) = \frac{1}{2\sqrt{n}} \sum_{k=1}^{\mathsf{k}} \sum_{j=1}^{n} f_{t_{j-1}}\left[D_j^{(k)}\right] \tag{21}$$

where

$$f_{t_{j-1}} = \left((S^{(k)})^{-1}(\partial_\theta S^{(k)})(S^{(k)})^{-1}\right)(X_{t_{j-1}},\theta^*)$$

and

$$D_j^{(k)} = h^{-1}\left(\Delta_j \widetilde{Y}^{(k)}\right)^{\otimes 2} - S^{(k)}(X_{t_{j-1}},\theta^*).$$

We have $N_T^X \in L^{\infty\text{-}}$ and

$$\left\| \max_{j=1,\dots,n} \left| f_{t_{j-1}}\left[D_j^{(k)}\right]\right| \right\|_p = O(n^{1/4})$$

for every $p > 1$. Therefore,

$$\left\| n^{-1/2} \sum_{j=1}^{n} f_{t_{j-1}}\left[D_j^{(k)}\right]\right\|_p = \left\| n^{-1/2} \sum_{j=1}^{n} 1_{\{\Delta_j N^X = 0\}} f_{t_{j-1}}\left[D_j^{(k)}\right]\right\|_p + o(1)$$

for every $p > 1$. In this situation, it suffices to show that

$$\left\| n^{-1/2} \sum_{j=1}^{n} 1_{\{\Delta_j N^X = 0\}} f_{t_{j-1}}\left[D_j^{(k)}\right]\right\|_p = O(1) \tag{22}$$

as $n \to \infty$ for every $p > 1$.

Now, we have the equality

$$1_{\{\Delta_j N^X = 0\}} \Delta_j \widetilde{Y}^{(k)} = 1_{\{\Delta_j N^X = 0\}} \left(\Xi_{1,j} + \Xi_{2,j} + \Xi_{3,j}\right),$$

where

$$\begin{aligned}
\Xi_{1,j} &= \sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}, \\
\Xi_{2,j} &= \int_{t_{j-1}}^{t_j} \left\{\sigma^{(k)}(X_{t_{j-1}} + \widetilde{X}_t - \widetilde{X}_{t_{j-1}},\theta^*) - \sigma^{(k)}(X_{t_{j-1}},\theta^*)\right\} \mathrm{d}w_t^{(k)}, \\
\Xi_{3,j} &= \int_{t_{j-1}}^{t_j} b_t^{(k)} \mathrm{d}t.
\end{aligned}$$

Define $C(x, y)$ by

$$C(x,y) = \left| \int_0^1 \partial_x \sigma^{(k)}(x + r(y-x),\theta^*)\mathrm{d}r \right|.$$

Then, by the same reason as in (22), and by Itô's formula and the Burkholder–Davis–Gundy inequality,

$$\begin{aligned}
&\left\| n^{-1/2}\sum_{j=1}^{n} 1_{\{\Delta_j N^X=0\}} h^{-1} f_{t_{j-1}}\left[\Xi_{1,j}\otimes \Xi_{2,j}\right]\right\|_p\\
=&\left\| n^{-1/2}\sum_{j=1}^{n} h^{-1} f_{t_{j-1}}\left[\Xi_{1,j}\otimes \Xi_{2,j}\right]\right\|_p + o(1)\\
\lesssim&\left\| n^{-1/2}\sum_{j=1}^{n} h^{-1} |f_{t_{j-1}}|\left|\sigma^{(k)}(X_{t_{j-1}},\theta^*)\right|\right.\\
&\left.\times\int_{t_{j-1}}^{t_j}\left|\sigma^{(k)}(X_{t_{j-1}}+\widetilde{X}_t-\widetilde{X}_{t_{j-1}},\theta^*)-\sigma^{(k)}(X_{t_{j-1}},\theta^*)\right|\mathrm{d}t\right\|_p + O(1)
\end{aligned}$$

and the last expression is not greater than

$$\begin{aligned}
&\left\| n^{-1/2}\sum_{j=1}^{n} h^{-1} |f_{t_{j-1}}|\left|\sigma^{(k)}(X_{t_{j-1}},\theta^*)\right|\right.\\
&\left.\quad\times\int_{t_{j-1}}^{t_j} C(X_{t_{j-1}},\widetilde{X}_t-\widetilde{X}_{t_{j-1}})\left|\widetilde{X}_t-\widetilde{X}_{t_{j-1}}\right|\mathrm{d}t\right\|_p + O(1)\\
\lesssim&\; n^{-1/2}\sum_{j=1}^{n} h^{-1}\int_{t_{j-1}}^{t_j}\Big\| |f_{t_{j-1}}|\left|\sigma^{(k)}(X_{t_{j-1}},\theta^*)\right| C(X_{t_{j-1}},\widetilde{X}_t-\widetilde{X}_{t_{j-1}})\\
&\qquad\qquad\times\left|\widetilde{X}_t-\widetilde{X}_{t_{j-1}}\right|\Big\|_p\,\mathrm{d}t + O(1)\\
\lesssim&\; n^{-1/2}\sum_{j=1}^{n}\Bigg\{\sup_{t\in[t_{j-1},t_j]}\|\widetilde{X}_t-\widetilde{X}_{t_{j-1}}\|_{2p}\\
&\quad\times\sup_{\substack{t\in[t_{j-1},t_j]\\ j=1,\dots,n}}\Big\| |f_{t_{j-1}}|\left|\sigma^{(k)}(X_{t_{j-1}},\theta^*)\right| C(X_{t_{j-1}},\widetilde{X}_t-\widetilde{X}_{t_{j-1}})\Big\|_{2p}\Bigg\} + O(1)\\
=&\; O(1)
\end{aligned}$$

for $p>1$ since $\|\widetilde{X}_t-\widetilde{X}_{t_{j-1}}\|_{2p}\le C_{2p}n^{-1/2}$ and $\sup_{t\in[0,T]}\|X_t\|_p+\sup_{t\in[0,T]}\|\widetilde{X}_t\|_p<\infty$ by the continuity of the mapping $t\mapsto \widetilde{X}_t\in L^p$ for every $p>1$. In a similar manner, we obtain

$$\left\| n^{-1/2}\sum_{j=1}^{n} 1_{\{\Delta_j N^X=0\}} h^{-1} f_{t_{j-1}}\left[\Xi_{i_1,j}\otimes \Xi_{i_2,j}\right]\right\|_p = O(1)$$

for every $p>1$ and $(i_1,i_2)\in\{1,2,3\}^2\setminus\{(1,1)\}$. Finally, for $(i_1,i_2)=(1,1)$,

$$\begin{aligned}&\left\| n^{-1/2}\sum_{j=1}^{n} 1_{\{\Delta_j N^X=0\}} f_{t_{j-1}}\left[h^{-1}\Xi_{1,j}\otimes\Xi_{1,j}-S^{(k)}(X_{t_{j-1}},\theta^*)\right]\right\|_p\\ =&\left\| n^{-1/2}\sum_{j=1}^{n} f_{t_{j-1}}\left[h^{-1}\Xi_{1,j}\otimes\Xi_{1,j}-S^{(k)}(X_{t_{j-1}},\theta^*)\right]\right\|_p+o(1)\\ =&O(1)\end{aligned}$$

by the Burkholder–Davis–Gundy inequality. Therefore, we obtained (22) and hence (20). □

Lemma 19 *For every* $p \geq 1$,

$$\sup_{n\in\mathbb{N}} E\left[\left(\sup_{\theta\in\Theta} n^{\frac{1}{2}-\beta_2}\left|\mathbb{Y}_n(\theta)-\mathbb{Y}(\theta)\right|\right)^p\right]<\infty.$$

Proof We use Lemmas 13, 14 and 15 besides the Burkholder–Davis–Gundy inequality and Sobolev inequality. Then, the proof is similar to Lemma 12 and also to Lemma 6 of Uchida and Yoshida (2013). □

Proof of Theorem 2 The result follows from Theorem 2 of Yoshida (2011) with the aid of Lemmas 9, 10, 16, 17, 18 and 19. □

3.4 Limit theorem and convergence of moments

In this section, asymptotic mixed normality of the QMLE and QBE will be established.

$[F1']_\kappa$ Conditions (*ii*), (*iii*) and (*iv*) of $[F1]_\kappa$ are satisfied in addition to

(i) the process X has a representation

$$X_t = X_0+\int_0^t \tilde{b}_s ds+\int_0^t \tilde{a}_s d\tilde{w}_s + J_t^X \qquad (t\in[0,T])$$

where $J^X=(J_t^X)_{t\in[0,T]}$ is a càdlàg adapted pure jump process, $\tilde{w}=(\tilde{w}_t)_{t\in[0,T]}$ is an r_1-dimensional **F**-Wiener process, $\tilde{b}=(\tilde{b}_t)_{t\in[0,T]}$ is a d-dimensional càdlàg adapted process and $\tilde{a}=(\tilde{a}_t)_{t\in[0,T]}$ is a progressively measurable processes taking values in $\mathbb{R}^{\mathsf{d}}\otimes\mathbb{R}^{\mathsf{r}_1}$. Moreover,

$$\|X_0\|_p+\sup_{t\in[0,T]}\left(\|\tilde{b}_t\|_p+\|\tilde{a}_t\|_p+\|J_t^X\|_p\right)<\infty$$

for every $p>1$.

The Wiener process $\tilde{w}$ is possibly correlated with w.

Recall that $\hat{\theta}_n^{B,\alpha_n}$ denotes the quasi-Bayesian estimator (QBE) of θ with respect to $\mathbb{H}_n$ defined by (16). We extend the probability space $(\Omega,\mathcal{F},P)$ so that a p-dimensional standard Gaussian random vector ζ independent of $\mathcal{F}$ is defined on the extension $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$. Define a random field $\mathbb{Z}$ on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$ by

$$\mathbb{Z}(u) = \exp\left(\Delta[u] - \frac{1}{2}\Gamma[u^{\otimes 2}]\right) \qquad (u \in \mathbb{R}^{\mathsf{p}})$$

where $\Delta[u] = \Gamma^{1/2}[\zeta, u]$. We write $\hat{u}_n^{\mathsf{A},\alpha_n} = \sqrt{n}\left(\hat{\theta}_n^{\mathsf{A},\alpha_n} - \theta^*\right)$ for $\mathsf{A} \in \{M, B\}$.

Let $B(R) = \{u \in \mathbb{R}^{\mathsf{p}}; |u| \leq R\}$ for $R > 0$. Equip the space $C(B(R))$ of continuous functions on $B(R)$ with the sup-norm. Denote by $d_s(\mathcal{F})$ the $\mathcal{F}$-stable convergence.

Lemma 20 *Suppose that* $[F1']_4, [F2']$ *and* $[F3]$ *are fulfilled. Then,*

$$\mathbb{Z}_n|_{B(R)} \to^{d_s(\mathcal{F})} \mathbb{Z}|_{B(R)} \qquad \text{in } C(B(R)) \tag{23}$$

as $n \to \infty$ *for every* $R > 0$.

Proof Fix $k \in \{1, ..., \mathsf{k}\}$. Let

$$\widetilde{D}_j^{(k)} = \left(\Delta_j \widetilde{Y}^{(k)}\right)^{\otimes 2} - \left(\sigma^{(k)}(X_{t_{j-1}}, \theta^*)\Delta_j w^{(k)}\right)^{\otimes 2},$$

and let $f_{t_{j-1}} = \left((S^{(k)})^{-1}(\partial_\theta S^{(k)})(S^{(k)})^{-1}\right)(X_{t_{j-1}}, \theta^*)$. We will show

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}[\widetilde{D}_j^{(k)}] \right\|_p \to 0 \tag{24}$$

for every $p > 1$. Let

$$\begin{aligned} \mathsf{B}_j &= \int_{t_{j-1}}^{t_j} b_s^{(k)} \mathrm{d}s, \quad \mathsf{C}_j = \sigma^{(k)}(X_{t_{j-1}}, \theta^*)\Delta_j w^{(k)}, \\ \mathsf{D}_j &= \int_{t_{j-1}}^{t_j} \left(\sigma^{(k)}(X_s, \theta^*) - \sigma^{(k)}(X_{t_{j-1}}, \theta^*)\right)\mathrm{d}w_s, \quad \mathsf{E}_j = \int_{t_{j-1}}^{t_j} \sigma^{(k)}(X_s, \theta^*)\mathrm{d}w_s. \end{aligned}$$

Then,

$$\widetilde{D}_j^{(k)} = (\mathsf{B}_j)^{\otimes 2} + \left\{\mathsf{B}_j \otimes \mathsf{E}_j + \mathsf{E}_j \otimes \mathsf{B}_j\right\} + \left\{\mathsf{C}_j \otimes \mathsf{D}_j + \mathsf{D}_j \otimes \mathsf{C}_j + \mathsf{D}_j \otimes \mathsf{D}_j\right\}.$$

It is easy to see

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}[\mathsf{B}_j^{\otimes 2}] \right\|_p \to 0. \tag{25}$$

For $p > 2$, we have

$$\begin{aligned}\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}} [\mathsf{B}_j \otimes \mathsf{E}_j] \right\|_p \leq & \left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}} [h b_{t_{j-1}} \otimes \mathsf{E}_j] \right\|_p \\ & + \left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}} \left[\int_{t_{j-1}}^{t_j} \left(b_s - b_{t_{j-1}} \right) \mathrm{d}s \otimes \mathsf{E}_j \right] \right\|_p \\ \lesssim & \left\| \sum_{j=1}^{n} n^{-1} |f_{t_{j-1}}|^2 |b_{t_{j-1}}|^2 |\mathsf{E}_j|^2 \right\|_{p/2}^{1/2} \\ & + \left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}} \left[\int_{t_{j-1}}^{t_j} \left(b_s - b_{t_{j-1}} \right) \mathrm{d}s \otimes \mathsf{E}_j \right] \right\|_p \\ \leq & \left\{ \sum_{j=1}^{n} n^{-1} \| |f_{t_{j-1}}| \|_{3p}^2 \| |b_{t_{j-1}}| \|_{3p}^2 \| |\mathsf{E}_j| \|_{3p}^2 \right\}^{1/2} \\ & + \left\| \sum_{j=1}^{n} n^{1/2} |f_{t_{j-1}}| |\mathsf{E}_j| \int_{t_{j-1}}^{t_j} \left| b_s - b_{t_{j-1}} \right| \mathrm{d}s \right\|_p \end{aligned}$$

by the Burkholder–Davis–Gundy inequality and Hölder's inequality. Therefore,

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}} [\mathsf{B}_j \otimes \mathsf{E}_j] \right\|_p \to 0 \tag{26}$$

since

$$I_n := \left\| \sum_{j=1}^{n} n^{1/2} |f_{t_{j-1}}| |\mathsf{E}_j| \int_{t_{j-1}}^{t_j} \left| b_s - b_{t_{j-1}} \right| \mathrm{d}s \right\|_p \to 0.$$

Indeed, for any $\epsilon > 0$, there exists a number $\delta > 0$ such that $P\left[w'(b, \delta) > \epsilon\right] < \epsilon$, where $w'(x, \delta)$ is the modulus of continuity defined by

$$w'(x, \delta) = \inf_{(s_i) \in \mathcal{S}_\delta} \max_i \sup_{r_1, r_2 \in [s_{i-1}, s_i)} |x(r_1) - x(r_2)|,$$

where $\mathcal{S}_\delta$ is the set of sequences (s_i) such that $0 = s_0 < s_1 < \cdots < s_v = T$ and $\min_{i=1,\dots,v-1}(s_i - s_{i-1}) > \delta$. Then,

$$\begin{aligned} I_n \leq & \left\| \sum_{j=1}^{n} n^{1/2} |f_{t_{j-1}}| |\mathsf{E}_j| \right\|_p \epsilon h + \left\| \max_{j=1,\dots,n} V_j \right\|_p \frac{T}{\delta} + \left\| \sum_{j=1}^{n} V_j \right\|_{2p} P\left[w'(b, \delta) > \epsilon\right]^{\frac{1}{2p}} \\ \lesssim & \epsilon + \left(n^{-1/2} + \sum_{j=1}^{n} \left\| V_j 1_{\{V_j > n^{-1/2}\}} \right\|_p \right) \frac{T}{\delta} + \epsilon^{\frac{1}{2p}} \end{aligned}$$

for $n > T/\delta$, where

$$V_j = n^{1/2} |f_{t_{j-1}}| |\mathsf{E}_j| \int_{t_{j-1}}^{t_j} \left(|b_s| + |b_{t_{j-1}}| \right) |\mathrm{d}s.$$

Thus, we obtain $\lim_{n\to\infty} I_n = 0$ and hence (26).

Itô's formula gives

$$\begin{aligned}\sigma^{(k)}(X_t,\theta^*) - \sigma^{(k)}(X_{t_{j-1}},\theta^*) &= \int_{t_{j-1}}^{t}\left(\partial_x\sigma^{(k)}(X_s,\theta^*)[\tilde{b}_s] + \frac{1}{2}\partial_x^2\sigma^{(k)}(X_s,\theta^*)\big[\tilde{a}_s\tilde{a}_s^{\star}\big]\right)\mathrm{d}s \\ &\quad + \int_{t_{j-1}}^{t}\partial_x\sigma^{(k)}(X_{s-},\theta^*)[\tilde{a}_s\mathrm{d}\tilde{w}_s] \\ &\quad + \int_{(t_{j-1},t]}\big(\sigma^{(k)}(X_s,\theta^*) - \sigma^{(k)}(X_{s-},\theta^*)\big)\mathrm{d}N_s^X \\ &=: \mathsf{b}_j(t) + \mathsf{a}_j(t) + \mathsf{d}_j(t)\end{aligned}$$

for $t \in [t_{j-1}, t_j]$. With Itô's formula, one can show

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}\left[\mathsf{C}_j \otimes \int_{t_{j-1}}^{t_j}\mathsf{a}_j(s)\mathrm{d}w_s\right]\right\|_p \to 0.$$

Obviously,

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}\left[\mathsf{C}_j \otimes \int_{t_{j-1}}^{t_j}\mathsf{b}_j(s)\mathrm{d}w_s\right]\right\|_p \to 0.$$

Moreover, for $\hat{V}_j = n^{1/2}|f_{t_{j-1}}||\mathsf{C}_j||\int_{t_{j-1}}^{t_j}\mathsf{d}_j(s)\mathrm{d}w_s|$, we have

$$\begin{aligned}\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}\left[\mathsf{C}_j \otimes \int_{t_{j-1}}^{t_j}\mathsf{d}_j(s)\mathrm{d}w_s\right]\right\|_p &\leq \left\| \max_{j=1,\dots,n}\hat{V}_j N_T^X\right\|_p \\ &\leq n^{-1/4}\|N_T^X\|_p + P\left[\max_{j=1,\dots,n}\hat{V}_j > n^{-1/4}\right]^{\frac{1}{2p}}\|N_T^X\|_{2p} \\ &\to 0.\end{aligned} \tag{27}$$

Therefore,

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}[\mathsf{C}_j \otimes \mathsf{D}_j]\right\|_p \to 0. \tag{28}$$

Similarly to (27), we know

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}\left[\left(\int_{t_{j-1}}^{t_j}\mathsf{d}_j(s)\mathrm{d}w_s\right)^{\otimes 2}\right]\right\|_p \to 0$$

and also

$$\left\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}[\mathsf{D}_j \otimes \mathsf{D}_j]\right\|_p \to 0. \tag{29}$$

From (25), (26), (28), (29) and symmetry, we obtain (24). In particular, (24) and (21) give the approximation

$$\begin{aligned}\widetilde{\Delta}_n \equiv& n^{-1/2}\partial_\theta \widetilde{\mathbb{H}}_n(\theta^*)\\ =&\frac{1}{2\sqrt{n}}\sum_{k=1}^{\mathsf{k}}\sum_{j=1}^{n} f_{t_{j-1}}\left[h^{-1}\left(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\right)^{\otimes 2} - S^{(k)}(X_{t_{j-1}},\theta^*)\right] + o_p(1),\end{aligned}$$

and so $\widetilde{\Delta}_n \to^{d_s(\mathcal{F})} \Gamma^{\frac{1}{2}}\zeta$ as $n \to \infty$. Furthermore, Lemma 13 ensures

$$\Delta_n \to^{d_s(\mathcal{F})} \Gamma^{\frac{1}{2}}\zeta \tag{30}$$

as $n \to \infty$.

Let $R > 0$. Then, there exists $n(R)$ such that for all $n \geq n(R)$ and all $u \in B(R)$,

$$\log \mathbb{Z}_n(u) = \Delta_n[u] + \frac{1}{2n}\partial_\theta^2 \mathbb{H}_n(\theta^*)[u^{\otimes 2}] + r_n(u), \tag{31}$$

where

$$r_n(u) = \int_0^1 (1-s)\left\{n^{-1}\partial_\theta^2 \mathbb{H}_n(\theta_n^\dagger(su))[u^{\otimes 2}] - n^{-1}\partial_\theta^2 \mathbb{H}_n(\theta^*)[u^{\otimes 2}]\right\}\mathrm{d}s$$

with $\theta_n^\dagger(u) = \theta^* + n^{-1/2}u$. Combining (30), Lemmas 17 and 16 with representation (31), we conclude the finite-dimensional stable convergence

$$\mathbb{Z}_n \to^{d_{s\text{-}f}(\mathcal{F})} \mathbb{Z} \tag{32}$$

as $n \to \infty$. Since Lemma 16 validates the tightness of $\{\mathbb{Z}_n|_{B(R)}\}_{n \geq n(R)}$, we obtain the functional stable convergence (23). □

Theorem 3 *Suppose that* $[F1']_4$, $[F2']$ *and* $[F3]$ *are fulfilled. Then,*

$$E\left[f\left(\hat{u}_n^{\mathsf{A},\alpha_n}\right)\Phi\right] \to \mathbb{E}\left[f\left(\Gamma^{-1/2}\zeta\right)\Phi\right]$$

as $n \to \infty$ *for* $\mathsf{A} \in \{M, B\}$, *any continuous function* f *of at most polynomial growth, and any* $\mathcal{F}$*-measurable random variable* $\Phi \in \cup_{p>1} L^p$.

Proof To prove the result for $\mathsf{A} = M$, we apply Theorem 5 of Yoshida (2011) with the help of Lemma 20 and Proposition 3. For the case $\mathsf{A} = B$, we obtain the convergence

$$\int_{\mathbb{U}_n} f(u)\mathbb{Z}_n(u)\varpi(\theta^* + n^{-1/2}u)\mathrm{d}u \to^{d_s(\mathcal{F})} \int_{\mathbb{R}^\mathsf{p}} f(u)\mathbb{Z}(u)\varpi(\theta^*)\mathrm{d}u$$

for any continuous function of at most polynomial growth, by applying Theorem 6 of Yoshida (2011). For that, we use Lemma 20 and Theorem 2. Estimate with Lemma 2 of Yoshida (2011) ensures Condition (i) of Theorem 8 of Yoshida (2011), which proves the stable convergence as well as moment convergence. □

4 Efficient one-step estimators

In Sect. 3, the asymptotic optimality was established for the QMLE $\hat{\theta}_n^{M,\alpha_n}$ and the QBE $\hat{\theta}_n^{B,\alpha_n}$ having a moving threshold specified by α_n converging to 0. However, in practice for fixed n, these estimators are essentially the same as the α-QMLE and α-QBE for a fixed α though they gained some freedom of choice of $\mathfrak{S}_{n,j-1}^{(k)}$, $p_n^{(k)}$ and $q_n^{(k)}$ in the asymptotic theoretical context.

It was found in Sect. 2.5 that the α-QMLE $\hat{\theta}_n^{M,\alpha}$ and the (α,β)-QBE $\hat{\theta}_n^{B,\alpha,\beta}$ based on a fixed α-threshold are consistent. However, they have pros and cons. They are expected to remove jumps completely but they are conservative and the rate of convergence is not optimal. In this section, as the second approach to optimal estimation, we try to recover efficiency by combining these less optimal estimators with the aggressive random field $\mathbb{H}_n$ given by (16), expecting to keep high precision of jump detection by the fixed α filters.

Suppose that $\kappa \in \mathbb{N}$ satisfies $\kappa > 1 + (2\gamma_0)^{-1}$. We assume $[F1']_{\kappa\vee 4}$, $[F2]$, $[F2']$ and $[F3]$. According to Proposition 1, $\hat{\theta}_n^{M,\alpha}$ attains $n^{-\beta}$-consistency for any $\beta \in \big(2^{-1}(\kappa-1)^{-1}, \gamma_0\big)$, and then $\beta(\kappa-1) > 1/2$. For $\theta^* \in \Theta$, there exists an open ball $B(\theta^*) \subset \Theta$ around θ^*. If $\partial_\theta^2 \mathbb{H}_n(\theta_0)$ is invertible, then Taylor's formula gives

$$\begin{aligned}\theta_1 - \theta_0 =& \big(\partial_\theta^2 \mathbb{H}_n(\theta_0)\big)^{-1}\big[\partial_\theta \mathbb{H}_n(\theta_1) - \partial_\theta \mathbb{H}_n(\theta_0)\big] + \sum_{i=2}^{\kappa-2} A_{1,i}(\theta_0)\big[(\theta_1-\theta_0)^{\otimes i}\big]\\ &+ A_{1,\kappa-1}(\theta_1,\theta_0)\big[(\theta_1-\theta_0)^{\otimes(\kappa-1)}\big]\end{aligned}$$

for $\theta_1, \theta_0 \in B(\theta^*)$. The second term on the right-hand side reads 0 when $\kappa = 3$. Here, $A_{1,i}$ $(i = 2, ..., \kappa-2)$ are written by $\big(\partial_\theta^2 \mathbb{H}_n(\theta_0)\big)^{-1}$ and $\partial_\theta^i \mathbb{H}_n(\theta_0)$ $(i = 3, ..., \kappa-1)$, respectively, and $A_{1,\kappa-1}(\theta_0,\theta_1)$ is by $\big(\partial_\theta^2 \mathbb{H}_n(\theta_0)\big)^{-1}$ and $\partial_\theta^\kappa \mathbb{H}_n(\theta)$ $(\theta \in B(\theta^*))$. Let

$$F(\theta_1,\theta_0) = \epsilon(\theta_0) + \sum_{i=2}^{\kappa-2} A_{1,i}(\theta_0)\big[(\theta_1-\theta_0)^{\otimes i}\big], \tag{33}$$

where

$$\epsilon(\theta_0) = -\big(\partial_\theta^2 \mathbb{H}_n(\theta_0)\big)^{-1}[\partial_\theta \mathbb{H}_n(\theta_0)],$$

i.e., $\epsilon(\theta_0)[u] = -\big(\partial_\theta^2 \mathbb{H}_n(\theta_0)\big)^{-1}[\partial_\theta \mathbb{H}_n(\theta_0), u]$ for $u \in \mathbb{R}^{\mathsf{p}}$. We write $\sum_{i=2}^{\kappa-2} A_{1,i}(\theta_0)\big[F(\theta_1,\theta_0)^{\otimes i}\big]$ in the form

$$\sum_{i=2}^{\kappa-2} A_{1,i}(\theta_0)\big[F(\theta_1,\theta_0)^{\otimes i}\big] = A_2(\theta_0) + \sum_{i_1+i_2\geq 3} A_{2,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes i_1}, (\theta_1-\theta_0)^{\otimes i_2}\big]$$

with

$$A_2(\theta_0) = \sum_{i=2}^{\kappa-2} A_{1,i}(\theta_0)\big[\epsilon(\theta_0)^{\otimes i}\big].$$

Next, we write

$$\sum_{i_1+i_2\geq 3} A_{2,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes i_1}, F(\theta_1,\theta_0)^{\otimes i_2}\big] = A_3(\theta_0) + \sum_{i_1+i_2\geq 4} A_{3,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes i_1}, (\theta_1-\theta_0)^{\otimes i_2}\big]$$

with

$$A_3(\theta_0) = \sum_{i_1+i_2\geq 3} A_{2,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes(i_1+i_2)}\big].$$

Repeat this procedure up to

$$\begin{aligned}&\sum_{i_1+i_2\geq \kappa-2} A_{\kappa-3,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes i_1}, F(\theta_1,\theta_0)^{\otimes i_2}\big]\\ =&A_{\kappa-2}(\theta_0) + \sum_{i_1+i_2\geq \kappa-1} A_{\kappa-2,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes i_1}, (\theta_1-\theta_0)^{\otimes i_2}\big]\end{aligned}$$

with

$$A_{\kappa-2}(\theta_0) = \sum_{i_1+i_2\geq \kappa-2} A_{\kappa-3,i_1,i_2}(\theta_0)\big[\epsilon(\theta_0)^{\otimes(i_1+i_2)}\big].$$

Let $A_1(\theta_0) = \epsilon(\theta_0)$. Thus, the sequence of $\mathbb{R}^{\mathsf{p}}$-valued random functions

$$A_i(\theta_0) \qquad (i = 1, ..., \kappa-2)$$

is defined on $\{\theta_0 \in \Theta;\ \partial_\theta^2 \mathbb{H}_n(\theta_0) \text{ is invertible}\}$. For example, when $\kappa = 4$,

$$\begin{aligned}A_1(\theta_0) =& -\big(\partial_\theta^2\mathbb{H}_n(\theta_0)\big)^{-1}[\partial_\theta\mathbb{H}_n(\theta_0)],\\ A_2(\theta_0) =& -\frac{1}{2}\big(\partial_\theta^2\mathbb{H}_n(\theta_0)\big)^{-1}\big[\partial_\theta^3\mathbb{H}_n(\theta_0)[A_1(\theta_0)^{\otimes 2}]\big].\end{aligned}$$

Let

$$\mathfrak{M}_n = \left\{\hat\theta_n^{M,\alpha}\in\Theta,\ \det\partial_\theta^2\mathbb{H}_n(\hat\theta_n^{M,\alpha})\neq 0,\ \hat\theta_n^{M,\alpha} + \sum_{i=1}^{\kappa-2} A_i(\hat\theta_n^{M,\alpha})\in\Theta\right\}.$$

Define $\check\theta_n^{M,\alpha}$ by

$$\check\theta_n^{M,\alpha} = \begin{cases}\hat\theta_n^{M,\alpha} + \sum_{i=1}^{\kappa-2} A_i(\hat\theta_n^{M,\alpha}) & \text{on } \mathfrak{M}_n\\ \theta_* & \text{on } \mathfrak{M}_n^c\end{cases}$$

where θ_* is an arbitrary value in Θ.

On the event $\mathfrak{M}_n^0 := \{\hat\theta_n^{M,\alpha_n}, \hat\theta_n^{M,\alpha}\in B(\theta^*)\}\cap\mathfrak{M}_n$, the QMLE $\hat\theta_n^{M,\alpha_n}$ for $\mathbb{H}_n$ satisfies

$$\hat\theta_n^{M,\alpha_n} - \hat\theta_n^{M,\alpha} = F(\hat\theta_n^{M,\alpha_n}, \hat\theta_n^{M,\alpha}) + A_{1,\kappa-1}(\hat\theta_n^{M,\alpha_n}, \hat\theta_n^{M,\alpha})\big[(\hat\theta_n^{M,\alpha_n} - \hat\theta_n^{M,\alpha})^{\otimes(\kappa-1)}\big]. \quad (34)$$

Let

$$\mathfrak{M}'_n = \left\{ \hat{\theta}_n^{M,\alpha_n},\ \hat{\theta}_n^{M,\alpha} \in B(\theta^*),\ |\det n^{-1}\partial_\theta^2 \mathbb{H}_n(\hat{\theta}_n^{M,\alpha})| \geq 2^{-1}\det \Gamma, \right.$$
$$\left. \hat{\theta}_n^{M,\alpha} + \sum_{i=1}^{\kappa-2} A_i(\hat{\theta}_n^{M,\alpha}) \in \Theta \right\}.$$

Then, the estimate

$$\left\| \left\{ \hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha} - A_1(\hat{\theta}_n^{M,\alpha}) - \sum_{i=2}^{\kappa-2} A_{1,i}(\hat{\theta}_n^{M,\alpha})\big[(\hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha})^{\otimes i}\big] \right\} 1_{\mathfrak{M}'_n} \right\|_p = O(n^{-\beta(\kappa-1)}) \tag{35}$$

for every $p > 1$ follows from representation (34), Propositions 1 and 3 and Lemma 9. Moreover, Lemmas 9, 17 and 16 together with L^p-boundedness of the estimation errors yield $P[(\mathfrak{M}'_n)^c] = O(n^{-L})$ for every $L > 0$.

Now on the event $\mathfrak{M}_n^0$, we have

$$\begin{aligned}
&\sum_{i=2}^{\kappa-2} A_{1,i}(\hat{\theta}_n^{M,\alpha})\big[\big(\hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha}\big)^{\otimes i}\big] \\
&= \sum_{i=2}^{\kappa-2} A_{1,i}(\hat{\theta}_n^{M,\alpha})\left[\left(F(\hat{\theta}_n^{M,\alpha_n}, \hat{\theta}_n^{M,\alpha}) + A_{1,\kappa-1}(\hat{\theta}_n^{M,\alpha_n}, \hat{\theta}_n^{M,\alpha})\big[(\hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha})^{\otimes(\kappa-1)}\big]\right)^{\otimes i}\right].
\end{aligned}$$

Therefore, it follows from (35) that

$$\begin{aligned}
&\Bigg\| \Bigg\{ \hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha} - A_1(\hat{\theta}_n^{M,\alpha}) - A_2(\hat{\theta}_n^{M,\alpha}) \\
&\quad - \sum_{i_1+i_2\geq 3} A_{2,i_1,i_2}(\hat{\theta}_n^{M,\alpha})\big[\epsilon(\hat{\theta}_n^{M,\alpha})^{\otimes i_1}, (\hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha})^{\otimes i_2}\big] \Bigg\} 1_{\mathfrak{M}'_n} \Bigg\|_p \\
&= O(n^{-\beta(\kappa-1)})
\end{aligned}$$

for every $p > 1$. Inductively,

$$\left\| \left\{ \hat{\theta}_n^{M,\alpha_n} - \hat{\theta}_n^{M,\alpha} - \sum_{i=1}^{\kappa-2} A_i(\hat{\theta}_n^{M,\alpha}) \right\} 1_{\mathfrak{M}'_n} \right\|_p = O(n^{-\beta(\kappa-1)}).$$

Consequently, using boundedness of Θ on $(\mathfrak{M}'_n)^c$, we obtain

$$\|\hat{\theta}_n^{M,\alpha_n} - \check{\theta}_n^{M,\alpha}\|_p = O(n^{-\beta(\kappa-1)}) = o(n^{-1/2})$$

and this implies

$$\|\check{\theta}_n^{M,\alpha} - \theta^*\|_p = O(n^{-1/2})$$

for every $p > 1$. We note that β in the above argument is a working parameter chosen so that $\beta > 2^{-1}(\kappa - 1)^{-1}$.

Next, we will consider a Bayesian estimator as the initial estimator. We are supposing that $\kappa > 1 + (2\gamma_0)^{-1}$, and furthermore we suppose β satisfies $\beta \in (2^{-1}(\kappa-1)^{-1}, \gamma_0)$. Remark that this β is the parameter involved in the estimator $\hat{\theta}_n^{B,\alpha,\beta}$, not a working parameter. Let

$$\mathfrak{B}_n = \left\{ \hat{\theta}_n^{B,\alpha,\beta} \in \Theta,\ \det \partial_\theta^2 \mathbb{H}_n(\hat{\theta}_n^{B,\alpha,\beta}) \neq 0,\ \hat{\theta}_n^{B,\alpha,\beta} + \sum_{i=1}^{\kappa-2} A_i(\hat{\theta}_n^{B,\alpha,\beta}) \in \Theta \right\}.$$

Define $\check{\theta}_n^{B,\alpha,\beta}$ by

$$\check{\theta}_n^{B,\alpha,\beta} = \begin{cases} \hat{\theta}_n^{B,\alpha,\beta} + \sum_{i=1}^{\kappa-2} A_i(\hat{\theta}_n^{B,\alpha,\beta}) & \text{on } \mathfrak{B}_n \\ \theta_* & \text{on } \mathfrak{B}_n^c. \end{cases}$$

Then, we obtain

$$\left\| \hat{\theta}_n^{M,\alpha_n} - \check{\theta}_n^{B,\alpha,\beta} \right\|_p = O(n^{-\beta(\kappa-1)}) = o(n^{-1/2})$$

and

$$\left\| \check{\theta}_n^{B,\alpha,\beta} - \theta^* \right\|_p = O(n^{-1/2})$$

for every $p > 1$.

Write $\check{u}_n^{\mathsf{A}} = \sqrt{n}\left(\check{\theta}_n^{\mathsf{A}} - \theta^*\right)$ for $\mathsf{A} =$"M, α" and "B, α, β." Thus, we have obtained the following result from Theorem 3 for $\hat{\theta}_n^{M,\alpha_n}$.

Theorem 4 *Suppose that* $[F1']_{\kappa\vee 4}$, [F2], $[F2']$ *and* $[F3]$ *are fulfilled. Let f be any continuous function of at most polynomial growth, and let* Φ *be any* $\mathcal{F}$*-measurable random variable in* $\cup_{p>1} L^p$. *Suppose that an integer* κ *satisfies* $\kappa > 1 + (2\gamma_0)^{-1}$. *Then,*

(a) $E\left[f\left(\check{u}_n^{M,\alpha}\right)\Phi\right] \to \mathbb{E}\left[f\left(\Gamma^{-1/2}\zeta\right)\Phi\right]$ *as* $n \to \infty$.

(b) $E\left[f\left(\check{u}_n^{B,\alpha,\beta}\right)\Phi\right] \to \mathbb{E}\left[f\left(\Gamma^{-1/2}\zeta\right)\Phi\right]$ *as* $n \to \infty$, *suppose that* $\beta \in (2^{-1}(\kappa-1)^{-1}, \gamma_0)$.

5 Localization

In the preceding sections, we established asymptotic properties of the estimators, in particular, L^p-estimates for them. Though it was thanks to $[F3]$, verifying it is not straightforward. An analytic criterion and a geometric criterion are known to insure Condition $[F3]$ when X is a nondegenerate diffusion process (Uchida and Yoshida 2013). It is possible to give similar criteria even for jump-diffusion processes but we do not pursue this problem here. Instead, it is also possible to relax $[F3]$ in order to only obtain stable convergences.

We will work with

$[F3^\flat]$ $\quad \chi_0 > 0$ a.s.

in place of [$F3$].

Let $\epsilon > 0$. Then, there exists a $\delta > 0$ such that $P[A_\delta] \geq 1 - \epsilon$ for $A_\delta = \{\chi_0 > \delta\}$. Define ${}^\delta\mathbb{H}_n(\theta;\alpha)$ by

$$ {}^\delta\mathbb{H}_n(\theta;\alpha)_\omega = \begin{cases} \mathbb{H}_n(\theta;\alpha)_\omega & (\omega \in A_\delta) \\ -n|\theta - \theta^*|^2 & (\omega \in A_\delta^c). \end{cases} $$

The way of modification of $\mathbb{H}_n$ on A_δ^c is not essential in the following argument. Let

$$ {}^\delta\mathbb{Z}_n^\beta(u;\alpha) = \exp\left\{ {}^\delta\mathbb{H}_n^\beta\big(\theta^* + n^{-\beta}u;\alpha\big) - {}^\delta\mathbb{H}_n^\beta\big(\theta^*;\alpha\big) \right\} \qquad (u \in \mathbb{U}_n^\beta) $$

for ${}^\delta\mathbb{H}_n^\beta(\theta;\alpha) = n^{-1+2\beta}\,{}^\delta\mathbb{H}_n(\theta;\alpha)$. The random field ${}^\delta\mathbb{Y}_n(\theta;\alpha)$ is defined by

$$ {}^\delta\mathbb{Y}_n(\theta;\alpha) = n^{-2\beta}\left\{ {}^\delta\mathbb{H}_n^\beta(\theta;\alpha) - {}^\delta\mathbb{H}_n^\beta(\theta^*;\alpha)\right\} = n^{-1}\left\{ {}^\delta\mathbb{H}_n(\theta;\alpha) - {}^\delta\mathbb{H}_n(\theta^*;\alpha)\right\}. $$

The limit of ${}^\delta\mathbb{Y}_n(\theta;\alpha)$ is now

$$ {}^\delta\mathbb{Y}(\theta) = \mathbb{Y}(\theta)1_{A_\delta} - |\theta - \theta^*|^2 1_{A_\delta^c}. $$

The corresponding key index is

$$ {}^\delta\chi_0 = \inf_{\theta \neq \theta^*} \frac{-\,{}^\delta\mathbb{Y}(\theta)}{|\theta - \theta^*|^2}. $$

Then, Condition [$F3$] holds for ${}^\delta\chi_0$ under the conditional probability given A_δ, that is,

$$ P\left[\,{}^\delta\chi_0 < r^{-1} \middle| A_\delta\right] \leq C_{L,\delta}\, r^{-L} \quad (r > 0) $$

for every $L > 0$. Now it is not difficult to follow the proof of Propositions 1 and 2 to obtain

$$ \sup_{n\in\mathbb{N}} \left\{ E\left[\left|n^\beta\big(\hat{\theta}_n^{M,\alpha} - \theta^*\big)\right|^p 1_{A_\delta}\right] + E\left[\left|n^\beta\big(\hat{\theta}_n^{B,\alpha,\beta} - \theta^*\big)\right|^p 1_{A_\delta}\right] \right\} < \infty $$

for every $p > 1$ and every $\beta < \gamma_0$, under $[F1]_4$ and [$F2$] in addition to [$F3^\flat$]. Thus, we obtained the following results.

Proposition 4 *Suppose that $[F1]_4$, [$F2$] and [$F3^\flat$] are satisfied. Then, $n^\beta\big(\hat{\theta}_n^{M,\alpha} - \theta^*\big) = O_p(1)$ and $n^\beta\big(\hat{\theta}_n^{B,\alpha,\beta} - \theta^*\big) = O_p(1)$ as $n \to \infty$ for every $\beta < \gamma_0$.*

In a similar way, we can obtain the stable convergence of the estimators with moving α, as a counterpart to Theorem 3.

Theorem 5 *Suppose that $[F1']_4$, [$F2'$] and [$F3^\flat$] are fulfilled. Then,*

$$ \hat{u}_n^{\mathsf{A},\alpha_n} \to^{d_s} \Gamma^{-1/2}\zeta $$

as $n \to \infty$ *for* $\mathsf{A} \in \{M, B\}$.

Moreover, a modification of the argument in Sect. 4 gives the stable convergence of the one-step estimators.

Theorem 6 *Suppose that* $[F1']_{\kappa\vee 4}$, $[F2]$, $[F2']$ *and* $[F3^\flat]$ *are fulfilled. Suppose that an integer* κ *satisfies* $\kappa > 1 + (2\gamma_0)^{-1}$. *Then,*

(a) $\check{u}_n^{M,\alpha} \to^{d_s} \Gamma^{-1/2}\zeta$ *as* $n \to \infty$.
(b) $\check{u}_n^{B,\alpha,\beta} \to^{d_s} \Gamma^{-1/2}\zeta$ *as* $n \to \infty$, *suppose that* $\beta \in (2^{-1}(\kappa-1)^{-1}, \gamma_0)$.

Suppose that the process X satisfies the stochastic integral equation

$$X_t = X_0 + \int_0^t \tilde{b}(X_s)\mathrm{d}s + \int_0^t \tilde{a}(X_s)\mathrm{d}\tilde{w}_s + J_t^X \qquad (t \in [0,T])$$

with a finitely active jump part J^X with $\Delta J_0^X = 0$. The first jump time T_1 of J^X satisfies $T_1 > 0$ a.s. Suppose that X' is a solution to

$$X'_t = X_0 + \int_0^t \tilde{b}(X'_s)\mathrm{d}s + \int_0^t \tilde{a}(X'_s)d\tilde{w}_s \qquad (t \in [0,T])$$

and that $X' = X^{T_1}$ on $[0, T_1)$ for the stopped process X^{T_1} of X at T_1. This is the case where the stochastic differential equation has a unique strong solution. Furthermore, suppose that the key index $\chi_{0,\epsilon}$ defined for $(X'_t)_{t\in[0,\epsilon]}$ is nondegenerate for every $\epsilon > 0$ in that $\sup_{r>0} r^L P[\chi_{0,\epsilon} < r^{-1}] < \infty$ for every $L > 0$. Then, on the event $\{T_1 > \epsilon\}$, we have positivity of χ_0. This implies Condition $[F3^\flat]$. To verify nondegeneracy of $\chi_{0,\epsilon}$, we may apply a criterion in Uchida and Yoshida (2013).

Acknowledgements The authors thank the reviewers for their valuable comments to improve this article.

References

Dohnal, G. (1987). On estimating the diffusion coefficient. *Journal of Applied Probability*, 105–114.
Genon-Catalot, V., Jacod, J. (1993). On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l'Institut Henri Poincaré Probability and Statistics*, *29*(1), 119–151.
Iacus, S. M., Yoshida, N. (2018). *Simulation and inference for stochastic processes with YUIMA*. Basel: Springer.
Kamatani, K., Uchida, M. (2014). Hybrid multi-step estimators for stochastic differential equations based on sampled data. *Statistical Inference for Stochastic Processes*, *18*(2), 177–204.
Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, *24*(2), 211–229.
Ogihara, T., Yoshida, N. (2011). Quasi-likelihood analysis for the stochastic differential equation with jumps. *Statistical Inference for Stochastic Processes*, *14*(3), 189.
Ogihara, T., Yoshida, N. (2014). Quasi-likelihood analysis for nonsynchronously observed diffusion processes. *Stochastic Processes and Their Applications*, *124*(9), 2954–3008.
Prakasa Rao, B. (1988). Statistical inference from sampled data for stochastic processes. *Statistical Inference from Stochastic Processes (Ithaca, NY, 1987)*, *80*, 249–284.

Prakasa Rao, B. L. S. (1983). Asymptotic theory for nonlinear least squares estimator for diffusion processes. *Mathematische Operationsforschung und Statistik Series Statistics*, *14*(2), 195–209.
Shimizu, Y. (2009). A practical inference for discretely observed jump-diffusions from finite samples. *Journal of the Japan Statistical Society*, *38*(3), 391–413.
Shimizu, Y., Yoshida, N. (2006). Estimation of parameters for diffusion processes with jumps from discrete observations. *Statistical Inference for Stochastic Processes*, *9*(3), 227–277.
Uchida, M., Yoshida, N. (2012). Adaptive estimation of an ergodic diffusion process based on sampled data. *Stochastic Processes and Their Applications*, *122*(8), 2885–2924.
Uchida, M., Yoshida, N. (2013). Quasi likelihood analysis of volatility and nondegeneracy of statistical random field. *Stochastic Processes and Their Applications*, *123*(7), 2851–2876.
Uchida, M., Yoshida, N. (2014). Adaptive Bayes type estimators of ergodic diffusion processes from discrete observations. *Statistical Inference for Stochastic Processes*, *17*(2), 181–219.
Yoshida, N. (1992). Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*, *41*(2), 220–242.
Yoshida, N. (2011). Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Annals of the Institute of Statistical Mathematics*, *63*(3), 431–479.