博士論文

論文題目　Statistical Inference for
Stochastic Differential Equations with Jumps:
Global Filtering Approach

（ジャンプを含む確率微分方程式に対する統計推測：
大域的フィルターによる方法）

氏　　名　稲次　春彦

Contents

Chapter 1

Introduction

In this thesis, we propose a new statistical estimation method for stochastic differential equations (SDEs) with jumps to obtain more stable estimation results by extending several previous studies. We discuss both parametric and non-parametric estimation procedures.

In recent years, high-frequency data have become available in many fields, and one of the most important research topics in mathematical statistics is to establish appropriate statistical inference techniques for effective use of such data. There are various forms of high-frequency data, but, especially in the fields of financial engineering and biology, models described by stochastic differential equations (SDEs) play an important role, and statistical inference theory for SDEs is needed in order to apply these models to real data.

A typical stochastic integral equation is written as

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dw_s + J_t, \quad t \in [0, T], \tag{1.0.1}$$

where $(b_s)_{s\in[0,T]}$ and $(\sigma_s)_{s\in[0,T]}$ are càdlàg adapted processes, $(w_s)_{s\in[0,T]}$ is a standard Brownian motion, and $J_t = \sum_{s\in[0,t]} \Delta Y_s$ the jump part of Y ($\Delta Y_s = Y_s - Y_{s-} = Y_s - \lim_{r\uparrow s} Y_r$). The terminal of the data (observations) is denoted by T. The theme of statistical inference for SDEs is to estimate parameters, given the observation $\{Y_t\}_{t\in\mathcal{T}}$, where $\mathcal{T} \subset [0, T]$ is a set of points of observations, which may be an uncountable set (though unrealistic). Typically, it is a sequence of (possibly random) positive numbers $\mathcal{T} = \{t_i^n;\ i = 1, \dots,\ n\}$. This is the case of discrete sampling scheme, which is a main focus in the recent research. The most typical example is an equidistant sampling scheme: $t_i = iT/n$.

Studies on statistical inference for SDEs without jumps ($J_t \equiv 0$ in (1.0.1)) go back to Prakasa Rao [17], and are then generalized by Yoshida [24] and Kessler [10]. These studies assume a parametric representation of the model, namely, they assume that the drift and diffusion coefficients are given by functions of the form $b_t = b(Y_t, \beta)$ and $\sigma_t = \sigma(Y_t, \theta)$, respectively. Their aim is to construct estimators of β and θ under certain sampling schemes and discuss asymptotic properties of the estimators, such as consistency and asymptotic normality.

For SDEs without jumps, there are also many studies in the context of nonparametric estimation. One of the most active research topic in this field is the estimation of realized volatility. This is an attempt to construct a consistent estimator of the integrated volatility $\Theta = \int_0^T \sigma_t^2 dt$ for models described by SDEs. It is known that, in the absence of jumps, the sum of squared increments (called the realized volatility; RV hereafter) is the consistent estimator of the integrated volatility (Protter [18]):

$$\mathrm{RV}_n = \sum_{i=1}^n |\Delta_i Y|^2 \longrightarrow^p \Theta \qquad (n \to \infty),$$

where $\Delta_i Y = Y_{t_i^n} - Y_{t_{i-1}^n}$, and $\to^p$ denotes convergence in probability. It is then generalized to estimation of the covariation in the case of nonsynchronous observation (Hayashi and Yoshida [6], for example). Volatility estimation is now applied to real data, especially in the field of financial time series analysis.

While there has been such a large number of studies on SDEs without jumps, research on SDEs with jumps ($J_t \not\equiv 0$) remains much less studied. One approach to tackle the existence of jumps is to *detect/eliminate* them by a threshold method, proposed by Mancini [13]. This is based on the idea that when the absolute value of an increment $\Delta_i Y$ within a time interval exceeds a certain threshold given as a function of the length of interval $t_i^n - t_{i-1}^n$, the increment is regarded as a jump. This thresholding is also applied to the parametric context by Shimizu and Yoshida [20]. They consider statistical estimation of parameters in drift, diffusion, and jump terms of SDEs. The jump part is assumed to be driven by a compound Poisson process. Their idea is that, if the absolute value of the increment $|\Delta_i Y|$ is below a threshold given by $n^{-\rho}$, the increment is considered to come from the continuous part (Brownian semimartingale) and is used to estimate the drift and diffusion parameters, while if it exceeds the threshold, the increment is considered to come from the jump part (compound Poisson process) and is used to estimate the jump parameters. Here $\rho \in [0, 1/2)$ and $C > 0$ are parameters that affects the precision of jump detection. Especially, the choice of ρ crucially determines the precision of jump detection. Their estimator of diffusion parameter is defined as a maximizer of the following function:

$$\ell_n(\theta) = -\frac{1}{2}\sum_{i=1}^{n}\left\{h_n^{-1}\bar{Y}_i(\beta)' S_{i-1}(\theta)^{-1}\bar{Y}_i(\beta) - \sum_{i=1}^{n}\log\det S_{i-1}(\theta)\right\} 1_{\{|\Delta_i Y| \le h_n^{\rho}\}},$$

where $\bar{Y}_i(\beta) = \Delta_i Y - h_n b(Y_{t_{i-1}^n}, \beta)$, $S_{i-1}(\theta) = \sigma(Y_{t_{i-1}^n}, \theta)^{\otimes 2}$ and $h_n = t_i^n - t_{i-1}^n$. The indicator $1_{\{|\Delta_i Y| \le h_n^{\rho}\}}$ is called the "filter," which eliminates large increments from the likelihood function. This estimator is shown to be consistent and asymptotically normal under some regularity conditions.

Another approach dealing with jumps is to *mitigate* the effects of jumps with the aid of information about the nearest increment. The conventional RV, the sum of squared increments $|\Delta_i Y|^2$, is known to be extremely vulnerable to jumps, and mitigating the effects of jumps is an important theoretical and practical topic. A seminal work in this direction is the bipower variation (BV) proposed by Barndorff-Nielsen and Shephard [2] in the field of estimation of realized volatility. The idea of the BV is to use an increment within a neighboring interval. It is defined as the sum of the products of two absolute increments within adjacent intervals, $|\Delta_i Y||\Delta_{i+1} Y|$, making use of the fact that the probability of occurrence of successive jumps in adjacent time intervals is small:

$$\mathrm{BV}_n = \mu_1^{-2}\sum_{i=1}^{n-1}|\Delta_i Y||\Delta_{i+1} Y|,$$

where μ_1 is the first-order absolute moment of the standard normal distribution: $\mu_1 = E[|Z|], Z \sim N(0,1)$. Similarly, Andersen et al. [1] propose the minimum realized volatility (minRV), defined as the squares of smaller between two increments within adjacent intervals, $\min\{|\Delta_i Y|^2, |\Delta_{i+1} Y|^2\}$:

$$\mathrm{minRV}_n = \frac{\pi}{\pi - 2}\sum_{i=1}^{n-1}\min\{|\Delta_i Y|^2, |\Delta_{i+1} Y|^2\}.$$

These estimators are shown to be consistent of the integrated volatility Θ even in the presence of jumps.

The ideas of these previous studies can be summarized as "detecting, eliminating or mitigating the effects of jumps by using a single or two neighboring increments." In this sense, the approaches are "local." Since these methods are shown to have theoretically desirable properties, it may be said that problems of dealing with jumps has been solved to some extent.

However, it has been pointed out that these approaches may not always work well in practice, so there are still issues to be solved in terms of application. For example, Shimizu [19] reported that the estimation results by threshold method can vary greatly depending on the setting of a threshold. Also, for BV and minRV, it would not be necessarily sufficient to mitigate the effects of jumps by just using adjacent increments. Hence, a new method has been needed that overcomes such weakness of previous methods.

In this dissertation, we propose a new method for eliminating jumps, called "global filtering," and show its theoretical properties. Moreover, with some numerical simulations, we demonstrate its superiority over previous methods. The global filter compares the absolute increments of the data with all other samples, and excludes increments that exceeds a threshold determined by the data. Our filter is of the form $\{V_i < V_{(s_n)}\}$, where V_i is obtained by dividing $|\Delta_i Y|$ by a normalizing random variable. s_n is a certain positive integer, and $V_{(s_n)}$ denotes the s_n-th order statistic of $\{V_i\}_{i=1}^n$. The integer s_n is determined by a tuning parameter (cut-off ratio) $\alpha \in (0,1)$, which postulates how many observations are trimmed (the larger α is, the more observations are trimmed). For example, the maximum likelihood estimator of the diffusion parameter by global filtering is based on the following function (here we use a slightly different notation from Chapter 2; a more general formulation is presented there):

$$\mathbb{H}_n(\theta;\alpha) = -\frac{1}{2}\sum_{j=1}^{n}\left\{q(\alpha)^{-1}h_n^{-1}\Delta_j Y' S_{j-1}(\theta)^{-1}\Delta_j Y K_{n,j} + p(\alpha)^{-1}\log\det S_{j-1}(\theta)\right\}1_{\{V_j < V_{(s_n(\alpha))}\}},$$

where $p(\alpha)$ and $q(\alpha)$ are coefficients that depend on the cut-off ratio α, and $V_{(s_n(\alpha))}$ is the $s_n(\alpha)$-th order statistic of $\{V_i\}_{i=1}^n$ with $s_n(\alpha) = \lfloor n(1-\alpha)\rfloor$. $K_{n,j}$ is a certain truncation function. The point is that the threshold $V_{(s_n(\alpha))}$ appearing in the filter depends on all observations. On the other hand, the filter proposed by Shimizu and Yoshida [20] is of the form $\{|\Delta_i Y| \le n^{-\rho}\}$ and thus it is independent of the other observations. By comparing with all samples, the judgment of jumps and non-jumps by the global filter becomes more accurate, and it succeeds in overcoming the weakness of the "local" approach in previous studies. Also, the realized volatility by global filtering is defined by

$$\mathbb{V}_n(\alpha) = \sum_{j=1}^{n} q(\alpha)^{-1}|\Delta_j Y|^2 K_{n,j} 1_{\{V_j < V_{(s_n(\alpha))}\}},$$

where $V_{(s_n(\alpha))}$ is the same as above. It is obvious that this is a variant of the RV using global filtering.

The global filter is based on order statistics. The use of order statistic completely destroys the time-series structure of a model (i.e., martingale properties of estimating functions). This makes the proof of the theoretical validity of global filtering highly troublesome. We resolved this difficulty by proving a series of "global filtering lemmas," which help us create statistical inference theory based on global filtering.

In this dissertation, we focus on inference theory regarding the diffusion term σ_t given discrete observations $\{Y_{t_i^n}\}_{i=1}^n$, and discuss asymptotic properties of the estimators when the number of samples tends to infinity: $n \to \infty$ (and T is fixed). We assume equidistant sampling scheme: $t_i^n = iT/n$. Inference on drift and jump terms is beyond the scope of this paper. We assume nothing on its distributional structure (we only assume that the number of jumps is finite almost surely). In this sense, our method is highly versatile for estimation of diffusion parameter.

In Chapter 2, we deal with the parametric estimation of diffusion parameters, based on Inatsugu and Yoshida [8]. In this paper, in (1.0.1), we assume that the diffusion coefficient is of the form

$\sigma_t = \sigma(Y_t, \theta)$, and construct the maximum likelihood estimator and the Bayes estimator based on the global filtering method. As for the choice of cut-off ratio α, we discuss both the case where it is fixed and where it is "moving" (or "shrinking"), i.e., dependent on the number n of samples. We prove their moment convergence, which leads to their consistency and asymptotic mixed normality. To see that the new estimators outperform previous ones, we report results of numerical simulations.

In Chapter 3, we deal with the estimation of the integrated volatility. We do not assume any parametric structure of the diffusion coefficient in (1.0.1) and discuss the estimation of Θ. We introduce the global realized volatility (GRV) and its variant, the winsorized GRV, and give their rate of convergence. Similarly to Chapter 2, we discuss both the fixed-α and shrinking-α cases. Then we conduct numerical simulation to show its accuracy and usefulness.

Acknowledgement

This dissertation is comprised of two joint-works with my supervisor Professor Nakahiro Yoshida. Throughout the master and doctoral courses, I received his enthusiastic and courteous guidance. Without his support, I would not have been able to experience a fulfilling research. Particularly in the doctoral course, my initial simple idea was developed into a much far-reaching theory of global filtering thanks to his outstanding knowledge and techniques. Moreover, in the doctoral course, he patiently and kindly waited for my research even when it did not proceed as planned due to my business schedule. I would like to express my deepest and sincere gratitude to him for his kindness and guidance.

Chapter 2

Parametric Estimation: Global Jump Filters and Quasi-likelihood Analysis for Volatility

2.1 Introduction

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

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma(X_s,\theta)dw_s + J_t, \quad t\in[0,T] \tag{2.1.1}$$

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 (2.1.1) 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:

$$\begin{cases} Y_t = X_t + J_t^Y, \\ X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma(X_s,\theta)dw_s + J_t^X, \end{cases}$$

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 $(X_{t_j}, Y_{t_j})_{j=0,1,\dots,n}$, where $t_j = t_j^n = jT/n$. 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 [17, 16]. The joint asymptotic normality of estimators was given in Yoshida [24] and later generalized in Kessler [10]. The quasi-likelihood analysis (QLA, Yoshida [25]) 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 [21, 23]) and the hybrid multi-step estimators (Kamatani and

Uchida [9]) are of practical importance from computational aspects. Statistics becomes non-ergodic under a finite time horizon $T < \infty$. Dohnal [4] 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 [5]. Uchida and Yoshida [22] 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 non-synchronous case is addressed by Ogihara and Yoshida [15] within QLA. As for inference for jump-diffusion processes, under ergodicity, Ogihara and Yoshida [14] 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 [20].

The filter in the quasi-likelihood functions of Shimizu and Yoshida [20] is based on the magnitude of the absolute value of the increment: $\{|\Delta_i Y| > Ch_n^\rho\}$, 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 data sets 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 [19], Iacus and Yoshida [7]. 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 (2.1.1) 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 Section 2.2.2, we introduce the α-quasi-log likelihood function $\mathbb{H}_n(\theta; \alpha)$, that 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 Section 2.2.4 enable us to recover these properties. Section 2.2.5 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 [25]), the so-called polynomial type large deviation inequality is derived in Theorem 2.2.13 for an annealed version of the quasi-log likelihood $\mathbb{H}_n^\beta(\theta; \alpha)$ of (2.2.10), 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 (2.2.11). 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 2.2.15). 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 Section 2.3, 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 2.3.3 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 2.3.4). Stable convergence of these estimators and moment convergence are validated by Theorem 2.3.13. 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 Section 2.4, 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 Sections 2.2 and 2.3, 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 2.6 that the so-constructed estimators are accurate and quite stable against α, in practice. In Section 2.5, we relax the conditions for stable convergence by a localization argument. Section 2.6 presents some simulation results and shows that the global filter can detect jumps more precisely than the local threshold methods.

2.2 Global filter: α-threshold method

2.2.1 Model structure

We will work with the model (2.1.1). To structure the model suitably, we begin with an example.

Example 2.2.1. 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 component-wise. 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

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

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

2.2.2 Quasi likelihood function by order statistics

In this section, we will give a filter that removes ΔJ. [20] and [14] used certain jump detecting 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 $\{|\Delta_i Y|\}_{i=1,...,n}$.

We prepare statistics $\bar{S}^{(k)}_{n,j-1}$ ($k = 1, ..., \mathsf{k}$; $j = 1, ..., n$; $n \in \mathbb{N}$) 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 2.2.2. 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

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

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

$$\bar{S}^{(k)}_{n,j-1} = \hat{S}^{(k)}_{n,j-1} 1_{\{\lambda_{min}(\hat{S}^{(k)}_{n,j-1}) > 2^{-1}\epsilon_0\}} + 2^{-1}\epsilon_0 I_{\mathsf{m}_k} 1_{\{\lambda_{min}(\hat{S}^{(k)}_{n,j-1}) \leq 2^{-1}\epsilon_0\}}, \tag{2.2.1}$$

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

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

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

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

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

equivalently,

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

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

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

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

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

where

$$K_{n,j}^{(k)} \;=\; 1_{\left\{|\Delta_jY^{(k)}|<C_*^{(k)}n^{-\frac{1}{4}}\right\}} \tag{2.2.2}$$

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

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

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

Remark 2.2.3. (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., $\sup_{t\in[0,T]}|\Delta J_t|\in L^{\infty-}=\cap_{p>1}L^p$. It is also reasonable to use

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

if $\bar{S}_{n,j-1}$ is uniformly $L^{\infty-}$-bounded. In any case, the factor $K_{n,j}^{(k)}$ 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 $K_{n,j}^{(k)}$ is of order $O(n^{-1/4})$, that 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 $\sigma^{(k)}(X_t,\theta)dw_t^{(k)}=S^{(k)}(X_t,\theta)^{1/2}\cdot\big(S^{(k)}(X_t,\theta)^{-1/2}\sigma^{(k)}(X_t,\theta)dw_t^{(k)}\big)$. On the other hand, we could introduce an $\mathsf{m}_k\times\mathsf{m}_k$ random matrix $\bar{\sigma}_{n,j-1}^{(k)}$ approximating $\sigma^{(k)}(X_{t_{j-1}},\theta^*)$ up to scaling, and use $\big(\bar{\sigma}_{n,j-1}^{(k)}\big)^{-1}\Delta_jY^{(k)}$ for $\big(\bar{S}_{n,j-1}^{(k)}\big)^{-1/2}\Delta_jY^{(k)}$, in order to remove the assumption of symmetry.

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

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

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

2.2.3 Assumptions

We assume Sobolev's embedding inequality

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

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 $N^{(k)}_t=\sum_{s\le t}1_{\{\Delta J^{(k)}_s\neq0\}}$ and $N_t=\sum_{s\le t}1_{\{\Delta J_s\neq0\}}$ 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

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

- **(ii)** $\sup_{t\in[0,T]}\|b_t\|_p<\infty$ for every $p>1$.
- **(iii)** $\sigma\in C^{2,\kappa}_{\uparrow}(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R}^{\mathsf{m}}\otimes\mathbb{R}^{\mathsf{r}})$, $S(X_t,\theta)$ 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 $N^X_T\in L^{\infty-}$.

Let $(\kappa_n)_{n\in\mathbb{N}}$ be a sequence of positive integers satisfying $\kappa_n=O(n^{1/2})$ as $n\to\infty$. For $j\in\{1,...,n\}$, let $I_{n,j}=\big\{i\in\{1,...,n\};\ |i-j|\le\kappa_n\big\}$. Let $\mathtt{I}_{n,j}=\cup_{i\in I_{n,j}}[t_{i-1},t_i]$. Define the index set $\mathsf{L}^{(k)}_n$ by

$$\mathsf{L}^{(k)}_n \;=\; \big\{j\in\{1,...,n\};\ N^{(k)}(\mathtt{I}_{n,j})+N^X(\mathtt{I}_{n,j})\neq0\big\}.$$

[F2] **(i)** $\bar{S}^{(k)}_{n,j-1}$ are symmetric, invertible and $\sup_{n\in\mathbb{N}}\max_{j=1,...,n}\big\|(\bar{S}^{(k)}_{n,j-1})^{-1}\big\|_p<\infty$ 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

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

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

Remark 2.2.4. 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 2.2.5. The $\bar{S}^{(k)}_{n,j-1}$ given by (2.2.1) in Example 2.2.2 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.

2.2.4 Global filtering lemmas

The α-quasi-log likelihood function $\mathbb{H}_n(\theta;\alpha)$ involves the summation regarding the index set $\mathcal{J}_n^{(k)}(\alpha^{(k)})$. The global jump filter $\mathcal{J}_n^{(k)}(\alpha^{(k)})$ 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

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

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

$$\sup_{n\in\mathbb{N}}\sup_{j=1,...,n}\left\|R_j^{(k)}1_{\{j\in(\mathsf{L}_n^{(k)})^c\}}\right\|_p \;=\; O(n^{-\gamma_0})$$

for every $p>1$, where

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

Remark that $A^{1/2} = \frac{1}{\pi}\int_0^\infty \lambda^{-1/2}A(\lambda+A)^{-1}d\lambda$ 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 j-th ordered statistic of $\{\overline{W}_1^{(k)},...,\overline{W}_n^{(k)}\}$, and $\overline{U}_{(j)}^{(k)}$ denotes the j-th 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

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

Lemma 2.2.6. *Suppose that* $\alpha^{(k)}\in(0,1)$*. Then* $P\Big[\bigcup_{j=1,..,n}N_{n,j}^{(k)}\Big] \;=\; O(n^{-L})$ *as* $n\to\infty$ *for every* $L>0$.

Proof. We have

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

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 $A_j = \{\overline{W}^{(k)}_j \leq q_{\bar{\alpha}^{(k)}} + n^{-\gamma_1}\}$ and $Z_j = 1_{A_j} - P[A_1]$, by the Burkholder-Davis-Gundy inequality, Jensen's inequality and $|Z_j| \leq 1$, we obtain

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

for every $p > 1$.

Let

$$
B^{(k)}_n = \big\{\big|\overline{W}^{(k)}_{(a^{(k)}_n)} - q_{\bar{\alpha}^{(k)}}\big| > n^{-\gamma_1}\big\}.
$$

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

$$
P\big[B^{(k)}_n\big] = O(n^{-L}) \tag{2.2.3}
$$

for every $L > 0$.

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

$$
\begin{aligned}
&P\Big[\sum_{j'=1}^n 1_{\{\overline{W}^{(k)}_{j'} \in [q_{\bar{\alpha}^{(k)}}-2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}}+2n^{-\gamma_1}]\}} \geq n^{1-\gamma_2}\Big] \\
&= P\Big[n^{-1+\gamma_1}\sum_{j'=1}^n 1_{\{\overline{W}^{(k)}_{j'} \in [q_{\bar{\alpha}^{(k)}}-2n^{-\gamma_1}, q_{\bar{\alpha}^{(k)}}+2n^{-\gamma_1}]\}} \geq n^{\gamma_1-\gamma_2}\Big] \\
&= O(n^{-L})
\end{aligned} \tag{2.2.4}
$$

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

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

is bounded in $L^{\infty-}$ (this can be proved by the same argument as above). Since the estimate (2.2.4) is independent of $j \in \{1, ..., n\}$, combining it with (2.2.3), we obtain

$$
\max_{j=1,..,n} P\big[N^{(k)}_{n,j}\big] = O(n^{-L})
$$

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

Let

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

where

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

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

$$
\Omega_n = \Big\{\sum_k \#\mathsf{L}^{(k)}_n < n^{1-\gamma_2}\Big\}\bigcap\Big(\bigcap_{k=1,...,\mathsf{k}}\ \bigcap_{j=1,...,n}\Big[\big\{|R^{(k)}_j|1_{\{j\in(\mathsf{L}^{(k)}_n)^c\}} < 2^{-1}n^{-\gamma_1}\big\} \cap (N^{(k)}_{n,j})^c\Big]\Big).
$$

Lemma 2.2.7.

$$\hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)}) \cap (\mathsf{L}_n^{(k)})^c \quad \subset \quad \mathcal{J}_n^{(k)}(\alpha^{(k)}) \tag{2.2.5}$$

on Ω_n*. In particular*

$$\#\big[\mathcal{J}_n^{(k)}(\alpha^{(k)}) \ominus \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})\big] \quad \le \quad c_* n^{1-\gamma_2} + \#\,\mathsf{L}_n^{(k)} \tag{2.2.6}$$

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 $(j_1, j_2) \in (\mathsf{L}_n^{(k)})^c \times (\mathsf{L}_n^{(k)})^c$ satisfies $r(\overline{W}_{j_1}^{(k)}) \le \hat{a}_n^{(k)}$ and $r(\overline{W}_{j_2}^{(k)}) \ge a_n^{(k)}$, then $\overline{U}_{j_1}^{(k)} < \overline{W}_{j_1}^{(k)} + 2^{-1}n^{-\gamma_1} \le \overline{W}_{(a_n^{(k)})}^{(k)} - 2^{-1}n^{-\gamma_1} \le \overline{W}_{j_2}^{(k)} - 2^{-1}n^{-\gamma_1} < \overline{U}_{j_2}^{(k)}$. Therefore, if $j \in \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)}) \cap (\mathsf{L}_n^{(k)})^c$, then $j \in \mathcal{J}_n^{(k)}(\alpha^{(k)})$ since one can find at least $\lceil \alpha^{(k)} n \rceil (\le (n - a_n^{(k)} + 1) - n^{1-\gamma_2})$ variables among $\overline{U}_{(a_n^{(k)})}^{(k)}, \ldots\ldots, \overline{U}_{(n)}^{(k)}$ that are larger than $\overline{U}_j^{(k)}$. Therefore (3.4.4) holds, and so does (3.4.5) as follows. From (3.4.4), we have $\#\big[\mathcal{J}_n^{(k)}(\alpha^{(k)}) \ominus \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})\big] \le \mathsf{N} + \#\,\mathsf{L}_n^{(k)}$ for

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

Suppose that $j \in \mathcal{J}_n^{(k)}(\alpha^{(k)}) \cap \hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})^c \cap (\mathsf{L}_n^{(k)})^c$. 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 $r(\overline{W}_j^{(k)}) \ge a_n^{(k)}$, as seen above, $\overline{U}_{j_1}^{(k)} < \overline{U}_j^{(k)}$ on Ω_n for all $j_1 \in (\mathsf{L}_n^{(k)})^c$ satisfying $r(\overline{W}_{j_1}^{(k)}) \le \hat{a}_n^{(k)}$, since $j \in (\mathsf{L}_n^{(k)})^c$ and $r(\overline{W}_j^{(k)}) \ge a_n^{(k)}$. 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 $\#\{j' \in \{1, ..., n\};\ \overline{U}_j^{(k)} < \overline{U}_{j'}^{(k)}\} \ge \lceil \alpha^{(k)} n \rceil$. Therefore

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

on Ω_n. We obtain (3.4.5) on Ω_n with $c_* = 6$ if we use the inequality $4n^{1-\gamma_2} + 2 \le 6n^{1-\gamma_2}$. □

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

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

Lemma 2.2.8. (i) *Let* $p_1 > 1$*. Then*

$$\begin{aligned} \|\mathcal{D}_n^{(k)}\|_p \quad &\le \quad \big(c_* n^{\gamma_3 - \gamma_2} + n^{-1+\gamma_3} \|\#\,\mathsf{L}_n^{(k)}\|_{p_1}\big) \Big\| \max_{j=1,\ldots,n} |V_j| \Big\|_{pp_1(p_1-p)^{-1}} \\ &\quad + n^{\gamma_3} \Big\| \max_{j=1,\ldots,n} |V_j| 1_{\Omega_n^c} \Big\|_p \end{aligned}$$

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

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

$$\begin{aligned} \|\mathcal{D}_n^{(k)}\|_p \quad &\le \quad \big(c_* n^{\gamma_3 - \gamma_2} + n^{-1+\gamma_3} \|\#\,\mathsf{L}_n^{(k)}\|_{p_1}\big) \\ &\quad \times \bigg(n^{\gamma_4} + n \max_{j=1,\ldots,n} \Big\| |V_j| 1_{\{|V_j| > n^{\gamma_4}\}} \Big\|_{pp_1(p_1-p)^{-1}} \bigg) \\ &\quad + n^{\gamma_3} \Big\| \max_{j=1,\ldots,n} |V_j| 1_{\Omega_n^c} \Big\|_p \end{aligned}$$

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

Proof. The estimate in (i) is obvious from (3.4.5). (ii) follows from (i). □

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

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

Lemma 2.2.9. *Let* $\tilde{\Omega}_n = \big\{\big|\overline{W}^{(k)}_{(\hat{a}_n^{(k)})} - q_{\bar{\alpha}^{(k)}}\big| < \check{C}\,n^{-\gamma_2}\big\}$, *where* $\check{C}$ *is a positive constant. Then*

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

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

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

$$\begin{aligned}\|\tilde{\mathcal{D}}_n^{(k)}\|_p \;\le\;& n^{\gamma_3}\bigg\|\max_{j=1,\dots,n}|V_j|\bigg\|_p P\Big[\big|\overline{W}_1^{(k)} - q_{\bar{\alpha}^{(k)}}\big| \le \check{C}\,n^{-\gamma_2}\Big]\\ &+n^{\gamma_3}\bigg\|\max_{j=1,\dots,n}|V_j|\bigg\|_{pp_1(p_1-p)^{-1}}\\ &\quad\times\bigg\|\frac{1}{n}\sum_{j=1}^n\Big(1_{\{|\overline{W}_j^{(k)}-q_{\bar{\alpha}^{(k)}}|\le\check{C}\,n^{-\gamma_2}\}} - P\Big[|\overline{W}_1^{(k)} - q_{\bar{\alpha}^{(k)}}| \le \check{C}\,n^{-\gamma_2}\Big]\Big)\bigg\|_{p_1}\\ &+n^{\gamma_3}P[\tilde{\Omega}_n^c]^{1/p_1}\bigg\|\max_{j=1,\dots,n}|V_j|\bigg\|_{pp_1(p_1-p)^{-1}}.\end{aligned}$$

Proof. (i) follows from

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

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 2.2.9 can be estimated as the proof of Lemma 2.2.6. For example, $P[\tilde{\Omega}_n^c] = O(n^{-L})$ for any $L > 0$.

Lemma 2.2.10. *Let* $k \in \{1, ..., \mathsf{k}\}$ *and let* $f \in C^{1,1}_{\uparrow}(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R})$. *Suppose that* $[F1]_0$ *is fulfilled. Then*

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

for every $p \ge 1$ *and* $\epsilon < \gamma_2$.

Proof. Use Sobolev's inequality and Burkholder's inequality as well as Lemmas 2.2.6, 3.4.5 (ii) and 2.2.9 (ii). More precisely, we have the following decomposition

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

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 3.4.5 (ii) to obtain

$$\begin{aligned}
\left\|\sup_{\theta\in\Theta} n^{\epsilon}\left|I_{1,n}^{(k)}(\theta)\right|\right\|_p &\lesssim \sum_{i=0,1}\sup_{\theta\in\Theta}\left\|n^{\epsilon}\left|\frac{1}{n}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}\partial_\theta^i f(X_{t_{j-1}},\theta) - \frac{1}{n}\sum_{j\in\hat{\mathcal{J}}_n^{(k)}(\alpha^{(k)})}\partial_\theta^i f(X_{t_{j-1}},\theta)\right|\right\|_p \\
&\lesssim \sum_{i=0,1}\sup_{\theta\in\Theta}\Bigg\{\left(c_* n^{\epsilon-\gamma_2} + n^{-1+\epsilon}\|\#\,\mathsf{L}_n^{(k)}\|_{p_1}\right) \\
&\qquad\times\left(n^{\gamma_4} + n\max_j\left\||\partial_\theta^i f(X_{t_{j-1}},\theta)|1_{\{|\partial_\theta^i f(X_{t_{j-1}},\theta)|\ge n^{\gamma_4}\}}\right\|_{\frac{pp_1}{p_1-p}}\right) \\
&\qquad+n^{\epsilon}\left\|\max_j|\partial_\theta^i f(X_{t_{j-1}},\theta)|1_{\Omega_n^c}\right\|_p\Bigg\}.
\end{aligned}$$

By taking $\gamma_4 > 0$ small enough, we can verify that the right-hand side is $o(1)$ since

$$\left\|\#\,\mathsf{L}_n^{(k)}\right\|_p \lesssim \kappa_n\left\|N_T + N_T^X\right\|_p = O(n^{1/2}).$$

Note that we have used the fact $P[\Omega_n^c] = O(n^{-L})$ for any $L > 0$. A similar argument with Lemma 2.2.9 (ii) yields $\left\|\sup_{\theta\in\Theta} n^{\epsilon}\left|I_{2,n}^{(k)}(\theta)\right|\right\|_p = o(1)$.

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

$$
\begin{aligned}
&\sup_{\theta\in\Theta}\left\|n^{\epsilon}\sum_{j=1}^{n}\frac{1}{n}\partial_\theta^i f(X_{t_{j-1}},\theta)\Big\{1_{\{\overline{W}_j^{(k)}\le q_{\bar\alpha^{(k)}}\}}-p(\alpha^{(k)})\Big\}\right\|_p^p\\
\lesssim\ &\sup_{\theta\in\Theta} n^{-p\left(\frac{1}{2}-\epsilon\right)}E\left[\left|\frac{1}{n}\sum_{j=1}^{n}\big|\partial_\theta^i f(X_{t_{j-1}},\theta)\big|^2\Big\{1_{\{\overline{W}_j^{(k)}\le q_{\bar\alpha^{(k)}}\}}-p(\alpha^{(k)})\Big\}^2\right|^{\frac{p}{2}}\right]\\
=\ &O\Big(n^{-(\frac{1}{2}-\epsilon)p}\Big)
\end{aligned}
$$

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

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

for every $p\ge 1$.

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

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

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

$$
\begin{aligned}
&\left\|n^{\epsilon}\sup_{\theta\in\Theta}\big|I_{4,n}^{(k)}(\theta)\big|\right\|_p\\
\le\ & n^{\epsilon}\times\frac{1}{nh}\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\big\|1_{\{\Delta_jN^X=0\}}C_f(X_{t_{j-1}},X_t)|X_t-X_{t_{j-1}}|\big\|_p\,dt\\
&+n^{\epsilon}\left\|\frac{1}{nh}\sum_{j=1}^{n}1_{\{\Delta_jN^X\neq0\}}\int_{t_{j-1}}^{t_j}C_f(X_{t_{j-1}},X_t)|X_t-X_{t_{j-1}}|\,dt\right\|_p\\
\lesssim\ & n^{-\frac{1}{2}+\epsilon}+n^{-\frac{1}{2}+\epsilon}\left\|(N_T^X)^{\frac{1}{2}}\left\{n^{-1}\sum_{j=1}^{n}\left(h^{-1}\int_{t_{j-1}}^{t_j}C_f(X_{t_{j-1}},X_t)|X_t-X_{t_{j-1}}|\,dt\right)^2\right\}^{\frac{1}{2}}\right\|_p\\
\lesssim\ & n^{-\frac{1}{2}+\epsilon}+n^{-\frac{1}{2}+\epsilon}\|N_T^X\|_p^{\frac{1}{2}}\\
=\ & O(n^{-\frac{1}{2}+\epsilon})
\end{aligned}
$$

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

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

Lemma 2.2.11. *Let $k\in\{1,...,\mathsf{k}\}$ and let $f\in C_\uparrow^{0,1}\big(\mathbb{R}^{\mathsf{d}}\times\Theta;\mathbb{R}^{\mathsf{m}_k}\otimes\mathbb{R}^{\mathsf{m}_k}\big)$. Suppose that $[F1]_0$ is fulfilled. Then*

$$
\sup_{n\in\mathbb{N}}\left\|\sup_{\theta\in\Theta}n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})}f(X_{t_{j-1}},\theta)\Big[\big(\Delta_jY^{(k)}\big)^{\otimes2}K_{n,j}^{(k)}-\big(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_jw^{(k)}\big)^{\otimes2}\Big]\right|\right\|_p\;<\;\infty
$$

for every $p\ge1$ and $\epsilon>0$.

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

$$Q_j \;=\; \big(\sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)}\big)^{\otimes 2}.$$

Then

$$\begin{aligned}
&\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} 1_{\{\Delta_j\check{N}>0\}} f(X_{t_{j-1}},\theta)\Big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}K_{n,j}^{(k)} - Q_j\Big]\right|\right\|_p \\
\le\;& \sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\max_{j=1,\dots,n}\left|f(X_{t_{j-1}},\theta)\Big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}K_{n,j}^{(k)} - Q_j\Big]\right|\right\|_{2p}\|\check{N}_T\|_{2p} \\
=\;& o(1) \qquad (2.2.7)
\end{aligned}$$

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

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

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

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

$$\begin{aligned}
\Delta_j\widetilde{Y}^{(k)} \;=\;& \sigma^{(k)}(X_{t_{j-1}},\theta^*)\Delta_j w^{(k)} + \int_{t_{j-1}}^{t_j}\big\{\sigma^{(k)}(X_t,\theta^*)-\sigma^{(k)}(X_{t_{j-1}},\theta^*)\big\}dw_t^{(k)} + \int_{t_{j-1}}^{t_j} b_t^{(k)}dt \\
=:\;& \xi_{1,j}+\xi_{2,j}+\xi_{3,j}.
\end{aligned}$$

Then we have

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

Thus, we can see

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

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

$$\begin{aligned}
&\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} 1_{\{\Delta_j\check{N}=0\}} f(X_{t_{j-1}},\theta)\Big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}K_{n,j}^{(k)} - Q_j\Big]\right|\right\|_p \\
\le\ &\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} 1_{\{\Delta_j\check{N}=0\}} f(X_{t_{j-1}},\theta)\Big[\big(\Delta_j \widetilde{Y}^{(k)}\big)^{\otimes 2} - Q_j\Big]\right|\right\|_p + O(n^{-L}) \\
=\ &o(1) \qquad (2.2.8)
\end{aligned}$$

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

From (2.2.7) and (2.2.8), we obtain

$$\sup_{\theta\in\Theta}\left\|n^{\frac{1}{2}-\epsilon}\left|\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} f(X_{t_{j-1}},\theta)\Big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}K_{n,j}^{(k)} - Q_j\Big]\right|\right\|_p \;=\; o(1) \qquad (2.2.9)$$

for every $p > 1$. Applying the same estimate as (2.2.9) to $\partial_\theta f$ for f, we conclude the proof by Sobolev's inequality. □

Lemmas 3.4.5, 2.2.9 and 2.2.11 suggest approximation of $n^{-1}\mathbb{H}_n(\theta;\alpha)$ by

$$-\frac{1}{2n}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\widetilde{\mathcal{J}}_n^{(k)}(\alpha^{(k)})}\Big\{q^{(k)}(\alpha^{(k)})^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)^{1/2}S^{(k)}(X_{t_{j-1}},\theta)^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)^{1/2}$$
$$\cdot\big[\big(h^{-1/2}\Delta_j w^{(k)}\big)^{\otimes 2}\big] + p(\alpha^{(k)})^{-1}\log\det S^{(k)}(X_{t_{j-1}},\theta)\Big\},$$

as we will see its validity below.

2.2.5 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 2.2.13 below for a random field associated with $\mathbb{H}_n(\theta;\alpha)$. Proof of Theorem 2.2.13 will be given in Section 2.2.6, based on the QLA theory ([25]) with the aid of the global filtering lemmas in Section 2.2.4. 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 [21] 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

$$\mathbb{H}_n^\beta(\theta;\alpha) \;=\; n^{-1+2\beta}\mathbb{H}_n(\theta;\alpha). \qquad (2.2.10)$$

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

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

Let

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

The key index χ_0 is defined by

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

Non-degeneracy of χ_0 plays an essential role in the QLA.

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

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

Remark 2.2.12. An analytic criterion and a geometric criterion are known to insure Condition $[F3]$ when X is a non-degenerate diffusion process. See Uchida and Yoshida [22] 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 Section 2.5.

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

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

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 Section 2.2.6.

Theorem 2.2.13. *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*

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

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 2.2.13, as usual in the QLA theory.

Proposition 2.2.14. *Suppose that $[F1]_4$, $[F2]$ and $[F3]$ are satisfied. Then $\sup_{n\in\mathbb{N}}\big\|n^\beta\big(\hat\theta_n^{M,\alpha}-\theta^*\big)\big\|_p<\infty$ 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

$$\hat{\theta}_n^{B,\alpha,\beta} = \left[\int_\Theta \exp\big(\mathbb{H}_n^\beta(\theta;\alpha)\big)\varpi(\theta)d\theta\right]^{-1}\int_\Theta \theta\exp\big(\mathbb{H}_n^\beta(\theta;\alpha)\big)\varpi(\theta)d\theta, \tag{2.2.11}$$

where ϖ is a continuous function on Θ satisfying $0 < \inf_{\theta\in\Theta}\varpi(\theta) \le \sup_{\theta\in\Theta}\varpi(\theta) < \infty$. Once again Theorem 2.2.13 ensures $L^{\infty-}$-boundedness of the error of the (α,β)-QBE:

Proposition 2.2.15. *Suppose that $[F1]_4$, $[F2]$ and $[F3]$ are satisfied. Let $\beta \in (0, \gamma_0)$. Then*

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

for every $p > 1$.

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

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

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] \\
\leq\ & E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)\varpi(\theta^*+n^{-\beta}u)du\bigg)^{-1}\int_{\mathbb{U}_n^\beta}|u|^p\mathbb{Z}_n^\beta(u;\alpha)\varpi(\theta^*+n^{-\beta}u)du\bigg] \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(Jensen's inequality, } p\geq 1) \\
\leq\ & C(\varpi)\sum_{r=1}^\infty(r+1)^p\bigg\{E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\bigg)^{-1}\int_{\{u;r<|u|\leq r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du \\
& \times 1_{\left\{\int_{\{u;r<|u|\leq r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du>\frac{C_1'}{r^{D-\mathsf{p}+1}}\right\}}\bigg] \\
& +E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\bigg)^{-1}\int_{\{u;r<|u|\leq r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du \\
& \qquad\qquad\qquad\qquad \times 1_{\left\{\int_{\{u;r<|u|\leq r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\leq\frac{C_1'}{r^{D-\mathsf{p}+1}}\right\}}\bigg]\bigg\} \\
& +C(\varpi)\quad\text{(The last term is for } r=0\text{. The integrand is not greater than one.)} \\
\leq\ & C(\varpi)\sum_{r=1}^\infty(r+1)^p\bigg\{P\bigg[\int_{\{u;r<|u|\leq r+1\}\cap\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du>\frac{C_1'}{r^{D-\mathsf{p}+1}}\bigg] \\
& \qquad\qquad +\frac{C_1'}{r^{D-\mathsf{p}+1}}E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\bigg)^{-1}\bigg]\bigg\}+C(\varpi) \\
\leq\ & C(\varpi)\sum_{r=1}^\infty(r+1)^p\bigg\{P\bigg[\sup_{u\in\mathbb{V}_n^\beta(r)}\mathbb{Z}_n^\beta(u;\alpha))>\frac{C_1}{r^D}\bigg] \\
& \qquad\qquad +\frac{C_1'}{r^{D-\mathsf{p}+1}}E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\bigg)^{-1}\bigg]\bigg\}+C(\varpi) \\
\lesssim\ & \sum_{r=1}^\infty r^{-(L-p)}+\sum_{r=1}^\infty r^{-(D-p-\mathsf{p}+1)}E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\bigg)^{-1}\bigg]+C(\varpi). \\
<\ & \infty
\end{aligned}
$$

by Theorem 2.2.13, suppose that

$$
E\bigg[\bigg(\int_{\mathbb{U}_n^\beta}\mathbb{Z}_n^\beta(u;\alpha)du\bigg)^{-1}\bigg]\ <\ \infty. \tag{2.2.12}
$$

However, one can show (2.2.12) by using Lemma 2 of Yoshida [25]. □

2.2.6 Proof of Theorem 2.2.13

We will prove Theorem 2.2.13 by Theorem 2 of Yoshida [25] with the aid of the global filtering lemmas in Section 2.2.4. Choose parameters η, β_1, ρ_1, ρ_2 and β_2 satisfying the following inequalities:

$$
\begin{aligned}
& 0<\eta<1, \qquad 0<\beta_1<\frac{1}{2}, \qquad 0<\rho_1<\min\{1,\eta(1-\eta)^{-1},2\beta_1(1-\eta)^{-1}\}, \\
& 2\eta<\rho_2, \qquad \beta_2\geq 0, \qquad 1-2\beta_2-\rho_2>0.
\end{aligned} \tag{2.2.13}
$$

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:

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

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 2.2.16. *For every $p\geq 1$,*

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

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

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

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

and

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

Apply Lemma 2.2.11 to $\partial_\theta^i\mathbb{R}_n^\circ(\theta;\alpha)$ $(i=0,...,3)$ to obtain

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

for every $p>1$. Moreover, we apply Sobolev's inequality, Lemma 3.4.5 (ii) and Lemma 2.2.9 (ii). Then it is sufficient to show that

$$\sum_{i=0}^{4}\sup_{\theta\in\Theta}\Big\{\big\|\partial_\theta^i n^{-1}\mathbb{H}_n^\times(\theta;\alpha)\big\|_p+\big\|\partial_\theta^i n^{-1}\mathbb{M}_n^\times(\theta;\alpha)\big\|_p\Big\} \;<\; \infty \tag{2.2.14}$$

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, (2.2.14) is obvious. □

Lemma 2.2.17. *For every $p \geq 1$,*

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

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

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

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

and

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

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

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

by Lemma 2.2.10, and also obtain

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

by Lemma 2.2.11 for every $p > 1$. Moreover, by Lemmas 3.4.5 (ii) and 2.2.9 (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

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

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

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

Lemma 2.2.18. *For every $p \geq 1$, there exists a constant C_p such that*

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

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

Lemma 2.2.19. *For every $p \geq 1$, there exists a constant C_p such that*

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

for all $r > 0$.

Lemma 2.2.20. *For every $p \geq 1$,*

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

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

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

and

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

We see $\sup_{n\in\mathbb{N}}\big\|R_n^\vee(\alpha,\beta)\big\|_p < \infty$ by Lemma 2.2.11. Moreover $\sup_{n\in\mathbb{N}}\big\|M_n^\vee(\alpha,\beta)\big\|_p < \infty$ by Lemmas 3.4.5 (ii) and 2.2.9 (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 2.2.20. The location shift of the random field $\mathbb{Z}_n^\beta(\cdot;\alpha)$ asymptotically vanishes.

Lemma 2.2.21. *For every $p \geq 1$,*

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

Proof. In this situation, we use the decomposition

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

with

$$\mathbb{Y}_n^+(\theta;\alpha) = -\frac{1}{2n}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{J}_n^{(k)}(\alpha^{(k)})} p(\alpha^{(k)})^{-1}\bigg\{\mathrm{Tr}\Big(S^{(k)}(X_{t_{j-1}},\theta)^{-1}S^{(k)}(X_{t_{j-1}},\theta^*)-I_{\mathsf{m}_k}\Big) + \log\frac{\det S^{(k)}(X_{t_{j-1}},\theta)}{\det S^{(k)}(X_{t_{j-1}},\theta^*)}\bigg\},$$

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

and

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

As assumed, $\beta(1-2\beta_2)<\gamma_2\le 1/2$. Lemma 2.2.11 gives

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

for every $p>1$. Furthermore, Lemma 2.2.10 gives

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

On the other hand, Lemmas 3.4.5 (ii) and 2.2.9 (ii) and the Burkholder-Davis-Gundy inequality together with Sobolev's inequality deduce

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

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

Proof of Theorem 2.2.13. Now Theorem 2.2.13 follows from Theorem 2 of Yoshida [25] combined with Lemmas 2.2.16, 2.2.17, 2.2.18, 2.2.19, 2.2.20 and 2.2.21. □

2.3 Global filter with moving threshold

2.3.1 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 $\alpha_n = (\alpha_n^{(1)}, ..., \alpha_n^{(\mathsf{k})})$. Let

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

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 j-th order statistic of $V_1^{(k)}, ..., V_n^{(k)}$.

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

$$\mathbb{H}_n(\theta) = -\frac{1}{2}\sum_{k=1}^{\mathsf{k}}\sum_{j\in\mathcal{K}_n^{(k)}}\bigg\{(q_n^{(k)})^{-1}h^{-1}S^{(k)}(X_{t_{j-1}},\theta)^{-1}\big[\big(\Delta_j Y^{(k)}\big)^{\otimes 2}\big]K_{n,j}^{(k)} + (p_n^{(k)})^{-1}\log\det S^{(k)}(X_{t_{j-1}},\theta)\bigg\}. \tag{2.3.1}$$

Remark 2.3.1. The truncation functional $K_{n,j}^{(k)}$ is given by (2.2.2). It is also reasonable to set it as

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

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

Remark 2.3.2. 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 $[F2']$ below. Thus a modification of $\mathbb{H}_n(\theta)$ is $\mathring{\mathbb{H}}_n(\theta)$ defined by

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

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 (2.3.1). 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

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

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

2.3.2 Polynomial type large deviation inequality

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

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

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

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

for every $p > 1$.

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

A polynomial type large deviation inequality is given by the following theorem, a proof of which is in Section 2.3.3.

Theorem 2.3.3. *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*

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

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

The polynomial type large deviation inequality for $\mathbb{Z}_n$ in Theorem 2.3.3 ensures $L^{\infty-}$-boundedness of the QLA estimators.

Proposition 2.3.4. *Suppose that $[F1]_4$, $[F2']$ and $[F3]$ are satisfied. Then*

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

for every $p > 1$.

2.3.3 Proof of Theorem 2.3.3

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

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

Lemma 2.3.5. *For every $p \geq 1$,*

$$\sum_{i=0}^{4}\sup_{\theta\in\Theta}\left\|n^{-1/2}\partial^i_\theta\mathbb{H}_n(\theta) - n^{-1/2}\partial^i_\theta\widetilde{\mathbb{H}}_n(\theta)\right\|_p \quad \to \quad 0 \tag{2.3.2}$$

as $n \to \infty$.

Proof. Let

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

Let

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

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

$$\begin{aligned} & \Big| (\mathfrak{S}_{n,j'(\omega)-1}^{(k)})^{-1/2} \Delta_{j'(\omega)} J^{(k)}(\omega) \Big| - \Big| (\mathfrak{S}_{n,j'(\omega)-1}^{(k)}(\omega))^{-1/2} \Delta_{j'(\omega)} \widetilde{Y}^{(k)}(\omega) \Big| \\ \leq \; & V_{j'(\omega)}^{(k)}(\omega) \; < \; V_{j(\omega)}^{(k)}(\omega) \; = \; \Big| (\mathfrak{S}_{n,j(\omega)-1}^{(k)}(\omega))^{-1/2} \Delta_{j(\omega)} \widetilde{Y}^{(k)}(\omega) \Big| \end{aligned}$$

and hence

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

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

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

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

$$\big(\mathfrak{A}_n^{(k)} \big)^c \subset \Big\{ \#\{ j \in \{1, ..., n\};\ \Delta_j N^{(k)} \neq 0 \} \geq n - s_n^{(k)} + 1 \Big\} \subset \big\{ N_T^{(k)} \geq B^{(k)} n^{\delta_1^{(k)}} \big\}.$$

Thus

$$P\Big[\bigcap_{k=1}^{\mathsf{k}} \mathfrak{B}_n^{(k)} \Big] = 1 - O(n^{-L}) \tag{2.3.3}$$

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

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

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

where the indicator function controls the moment outside of $\cap_{k=1}^{\mathsf{k}} \mathfrak{B}_n^{(k)}$. Then by (2.3.3), the cap and $N_T \in L^{\infty-}$, we obtain

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

as $n \to \infty$ for every $p \geq 1$. Indeed, we can estimate this difference of the two variables on the event $\mathfrak{C}_n := \cap_{k=1}^{\mathsf{k}} \mathfrak{B}_n^{(k)}$ and on $\mathfrak{C}_n^c$, as follows. On $\mathfrak{C}_n$, $|\Delta_j J^{(k)}| \leq n^{-1/4-\delta_0} 1_{\{\Delta_j J^{(k)} \neq 0\}}$ whenever $j \in \mathcal{K}_n^{(k)}$. The cap $K_{n,j}^{(k)}$ also offers the estimate $|\Delta_j Y^{(k)}| < C_*^{(k)} n^{-1/4}$. On $\mathfrak{C}_n$, after removing the factor $1_{\{|\Delta_j J^{(k)}| \leq 1\}}$ 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

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

for $\mathcal{M}_n^{(k)} = \max_{j=1,\ldots,n} |S^{(k)}(X_{t_{j-1}}, \theta)^{-1}|$, 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 (2.3.3). The cap $1_{\{|\Delta_j J^{(k)}| \leq 1\}}$ 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 (2.3.3), it is easy to show

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

which implies (2.3.2) as $n \to \infty$ for every $p \geq 1$. □

We choose parameters η, β_1, ρ_1, ρ_2 and β_2 satisfying (2.2.13) 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

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

The following two estimates will play a basic role.

Lemma 2.3.6. *Let $f \in C_\uparrow^{0,1}\big(\mathbb{R}^{\mathsf{d}} \times \Theta; \mathbb{R}^{\mathsf{m}_k} \otimes \mathbb{R}^{\mathsf{m}_k}\big)$. Then under $[F1]_0$,*

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

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

Proof. One can validate this lemma in a quite similar way as Lemma 2.2.11. □

Lemma 2.3.7. *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*

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

Proof. Let $p > 1$. By taking an approach similar to the proof of Lemma 2.3.6, we obtain

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

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 2.3.8. *For every $p \geq 1$,*

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

Proof. Applying Lemma 2.3.5 and Sobolev's inequality, one can prove the lemma in a fashion similar to Lemma 2.2.16. □

Lemma 2.3.9. *For every $p \geq 1$,*

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

Proof. Thanks to Lemma 2.3.5, it is sufficient to show that

$$\sup_{n\in\mathbb{N}} E\big[\big(n^{\beta_1}\big|\widetilde{\Gamma}_n-\Gamma\big|\big)^p\big] \ <\ \infty \tag{2.3.4}$$

where

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

Now taking a similar way as Lemma 2.2.17, one can prove the desired inequality by applying Lemmas 2.3.6 and 2.3.7 as well as the Burkholder-Davis-Gundy inequality. □

Lemma 2.3.10. *For every $p \geq 1$, $\sup_{n\in\mathbb{N}} E\big[\big|\Delta_n\big|^p\big] < \infty$.*

Proof. By Lemma 2.3.5, it suffices to show

$$\sup_{n\in\mathbb{N}} E\big[\big|\widetilde{\Delta}_n\big|^p\big] \ <\ \infty \tag{2.3.5}$$

for

$$\widetilde{\Delta}_n \ =\ n^{-1/2}\partial_\theta\widetilde{\mathbb{H}}_n(\theta^*) \ =\ \frac{1}{2\sqrt{n}}\sum_{k=1}^{\mathsf{k}}\sum_{j=1}^{n} f_{t_{j-1}}\big[D_j^{(k)}\big] \tag{2.3.6}$$

where

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

and

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

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

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

for every $p > 1$. Therefore

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

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

$$\Big\| n^{-1/2}\sum_{j=1}^n 1_{\{\Delta_j N^X=0\}} f_{t_{j-1}}\big[D_j^{(k)}\big]\Big\|_p \quad = \quad O(1) \tag{2.3.7}$$

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

Now, we have the equality

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

where

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

Define $C(x,y)$ by

$$C(x,y) \quad = \quad \Big| \int_0^1 \partial_x \sigma^{(k)}(x + r(y-x),\theta^*)dr \Big|.$$

Then, by the same reason as in (2.3.7), and by Itô's formula and the Burkholder-Davis-Gundy inequality,

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

and the last expression is not greater than

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

for $p > 1$ since $\big\|\widetilde{X}_t - \widetilde{X}_{t_{j-1}}\big\|_{2p} \le C_{2p} n^{-1/2}$ 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

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

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

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

by the Burkholder-Davis-Gundy inequality. Therefore we obtained (2.3.7) and hence (2.3.5). □

Lemma 2.3.11. *For every $p \ge 1$,*

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

Proof. We use Lemmas 2.3.5, 2.3.6 and 2.3.7 besides the Burkholder-Davis-Gundy inequality and Sobolev's inequality. Then the proof is similar to Lemma 2.2.21 and also to Lemma 6 of Uchida and Yoshida [22]. □

Proof of Theorem 2.3.3. The result follows from Theorem 2 of Yoshida [25] with the aid of Lemmas 2.2.18, 2.2.19, 2.3.8, 2.3.9, 2.3.10 and 2.3.11. □

2.3.4 Limit theorem and convergence of moments

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

$[\mathbf{F1'}]_\kappa$ 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 $\mathbf{F}$-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,

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

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 (2.3.1). 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 $\hat{u}_n^{\mathsf{A},\alpha_n} = \sqrt{n}\big(\hat{\theta}_n^{\mathsf{A},\alpha_n} - \theta^*\big)$ for $\mathsf{A} \in \{M, B\}$.

Let $B(R) = \{u \in \mathbb{R}^{\mathsf{p}};\ |u| \le R\}$ 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 2.3.12. *Suppose that $[F1']_4$, $[F2']$ and $[F3]$ are fulfilled. Then*

$$\mathbb{Z}_n|_{B(R)} \to^{d_s(\mathcal{F})} \mathbb{Z}|_{B(R)} \qquad \text{in } C(B(R)) \tag{2.3.8}$$

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

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

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

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

$$\Big\| \sum_{j=1}^n n^{1/2} f_{t_{j-1}}[\widetilde{D}_j^{(k)}] \Big\|_p \to 0 \tag{2.3.9}$$

for every $p > 1$. Let

$$\mathsf{B}_j = \int_{t_{j-1}}^{t_j} b_s^{(k)} ds, \quad \mathsf{C}_j = \sigma^{(k)}(X_{t_{j-1}}, \theta^*)\Delta_j w^{(k)},$$
$$\mathsf{D}_j = \int_{t_{j-1}}^{t_j} \big(\sigma^{(k)}(X_s, \theta^*) - \sigma^{(k)}(X_{t_{j-1}}, \theta^*)\big) dw_s, \quad \mathsf{E}_j = \int_{t_{j-1}}^{t_j} \sigma^{(k)}(X_s, \theta^*) dw_s.$$

Then

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

It is easy to see

$$\Big\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}[\mathsf{B}_j^{\otimes 2}] \Big\|_p \quad \to \quad 0. \tag{2.3.10}$$

For $p > 2$, we have

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

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

$$\Big\| \sum_{j=1}^{n} n^{1/2} f_{t_{j-1}}[\mathsf{B}_j \otimes \mathsf{E}_j] \Big\|_p \quad \to \quad 0 \tag{2.3.11}$$

since

$$I_n \;:=\; \Big\| \sum_{j=1}^{n} n^{1/2} \big|f_{t_{j-1}}\big| \big|\mathsf{E}_j\big| \int_{t_{j-1}}^{t_j} \big|b_s - b_{t_{j-1}}\big| ds \Big\|_p \quad \to \quad 0.$$

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

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

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

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

for $n > T/\delta$, where

$$V_j = n^{1/2}|f_{t_{j-1}}||\mathsf{E}_j|\int_{t_{j-1}}^{t_j}\big(|b_s|+|b_{t_{j-1}}|\big)|ds.$$

Thus we obtain $\lim_{n\to\infty} I_n = 0$ and hence (2.3.11).

Itô's formula gives

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

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

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

Obviously

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

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

$$\begin{aligned}\Big\|\sum_{j=1}^{n} n^{1/2}f_{t_{j-1}}\Big[\mathsf{C}_j\otimes\int_{t_{j-1}}^{t_j}\mathsf{d}_j(s)dw_s\Big]\Big\|_p &\le \Big\|\max_{j=1,\dots,n}\hat{V}_jN_T^X\Big\|_p \\ &\le n^{-1/4}\big\|N_T^X\big\|_p+P\Big[\max_{j=1,\dots,n}\hat{V}_j>n^{-1/4}\Big]^{\frac{1}{2p}}\big\|N_T^X\big\|_{2p} \\ &\to 0. \end{aligned} \tag{2.3.12}$$

Therefore

$$\Big\|\sum_{j=1}^{n} n^{1/2}f_{t_{j-1}}[\mathsf{C}_j\otimes\mathsf{D}_j]\Big\|_p \to 0. \tag{2.3.13}$$

Similarly to (2.3.12), we know

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

and also

$$\Big\|\sum_{j=1}^{n} n^{1/2}f_{t_{j-1}}[\mathsf{D}_j\otimes\mathsf{D}_j]\Big\|_p \to 0. \tag{2.3.14}$$

From (2.3.10), (2.3.11), (2.3.13), (2.3.14) and symmetry, we obtain (2.3.9). In particular, (2.3.9) and (2.3.6) give the approximation

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

and so $\widetilde{\Delta}_n \to^{d_s(\mathcal{F})} \Gamma^{\frac{1}{2}}\zeta$ as $n\to\infty$. Furthermore, Lemma 2.3.5 ensures

$$\Delta_n \quad \to^{d_s(\mathcal{F})} \quad \Gamma^{\frac{1}{2}}\zeta \tag{2.3.15}$$

as $n\to\infty$.

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

$$\log\mathbb{Z}_n(u) \quad = \quad \Delta_n[u] + \frac{1}{2n}\partial_\theta^2\mathbb{H}_n(\theta^*)[u^{\otimes 2}] + r_n(u), \tag{2.3.16}$$

where

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

with $\theta_n^\dagger(u) = \theta^* + n^{-1/2}u$. Combining (2.3.15), Lemmas 2.3.9 and 2.3.8 with the representation (2.3.16), we conclude the finite-dimensional stable convergence

$$\mathbb{Z}_n \quad \to^{d_{s\text{-}f}(\mathcal{F})} \quad \mathbb{Z} \tag{2.3.17}$$

as $n\to\infty$. Since Lemma 2.3.8 validates the tightness of $\{\mathbb{Z}_n|_{B(R)}\}_{n\geq n(R)}$, we obtain the functional stable convergence (2.3.8). □

Theorem 2.3.13. *Suppose that $[F1']_4$, $[F2']$ and $[F3]$ are fulfilled. Then*

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

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 [25] with the help of Lemma 2.3.12 and Proposition 2.3.4. For the case $\mathsf{A}=B$, we obtain the convergence

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

for any continuous function of at most polynomial growth, by applying Theorem 6 of [25]. For that, we use Lemma 2.3.12 and Theorem 2.3.3. Estimate with Lemma 2 of [25] ensures Condition (i) of Theorem 8 of [25], which proves the stable convergence as well as moment convergence. □

2.4 Efficient one-step estimators

In Section 2.3, 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 Section 2.2.5 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 (2.3.1), 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 2.2.14, $\hat{\theta}_n^{M,\alpha}$ 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

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

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

$$F(\theta_1,\theta_0) = \epsilon(\theta_0)+\sum_{i=2}^{\kappa-2}A_{1,i}(\theta_0)\big[(\theta_1-\theta_0)^{\otimes i}\big], \tag{2.4.1}$$

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

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

are defined on $\{\theta_0\in\Theta;\ \partial_\theta^2\mathbb{H}_n(\theta_0)$ 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

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

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

$$\hat\theta_n^{M,\alpha_n}-\hat\theta_n^{M,\alpha} = F(\hat\theta_n^{M,\alpha_n},\hat\theta_n^{M,\alpha})+A_{1,\kappa-1}(\hat\theta_n^{M,\alpha_n},\hat\theta_n^{M,\alpha})\big[(\hat\theta_n^{M,\alpha_n}-\hat\theta_n^{M,\alpha})^{\otimes(\kappa-1)}\big]. \qquad (2.4.2)$$

Let

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

Then the estimate

$$\bigg\|\bigg\{\hat\theta_n^{M,\alpha_n}-\hat\theta_n^{M,\alpha}-A_1(\hat\theta_n^{M,\alpha})-\sum_{i=2}^{\kappa-2}A_{1,i}(\hat\theta_n^{M,\alpha})\big[(\hat\theta_n^{M,\alpha_n}-\hat\theta_n^{M,\alpha})^{\otimes i}\big]\bigg\}1_{\mathfrak{M}_n'}\bigg\|_p = O(n^{-\beta(\kappa-1)}) \qquad (2.4.3)$$

for every $p > 1$ follows from the representation (2.4.2), Propositions 2.2.14 and 2.3.4 and Lemma 2.2.18. Moreover, Lemmas 2.2.18, 2.3.9 and 2.3.8 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 $\mathfrak{M}^0_n$, we have

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

Therefore it follows from (2.4.3) 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,

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

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

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

and this implies

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

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

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

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

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

and

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

for every $p > 1$.

Write $\check{u}_n^{\mathsf{A}} = \sqrt{n}\big(\check{\theta}_n^{\mathsf{A}} - \theta^*\big)$ for $\mathsf{A} =$"M, α" and "B, α, β". Thus, we have obtained the following result from Theorem 2.3.13 for $\hat{\theta}_n^{M,\alpha_n}$.

Theorem 2.4.1. *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) $E\big[f\big(\check{u}_n^{M,\alpha}\big)\Phi\big] \to \mathbb{E}\big[f\big(\Gamma^{-1/2}\zeta\big)\Phi\big]$ *as* $n \to \infty$.

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

2.5 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 non-degenerate diffusion process (Uchida and Yoshida [22]). 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$^\flat$] $\chi_0 > 0$ 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

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

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,

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

for every $L>0$. Now it is not difficult to follow the proof of Propositions 2.2.14 and 2.2.15 to obtain

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

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 2.5.1. *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 2.3.13.

Theorem 2.5.2. *Suppose that* $[F1']_4$, $[F2']$ *and* $[F3^\flat]$ *are fulfilled. Then*

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

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

Moreover, a modification of the argument in Section 2.4 gives the stable convergence of the one-step estimators.

Theorem 2.5.3. *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

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

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

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

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 non-degenerate 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 non-degeneracy of $\chi_{0,\epsilon}$, we may apply a criterion in Uchida and Yoshida [22].

2.6 Simulation Studies

2.6.1 Setting of simulation

In this section, we numerically investigate the performance of the global threshold estimator. We use the following one-dimensional Ornstein-Uhlenbeck process with jumps

$$dX_t = -\eta X_t dt + \sigma dw_t + dJ_t \qquad (t \in [0,1]) \tag{2.6.1}$$

starting from X_0. Here $w = (w_t)_{t\in[0,1]}$ is a one-dimensional Brownian motion and J is a one-dimensional compound Poisson process defined by

$$J_t = \sum_{i=1}^{N_t} \xi_i, \qquad \xi_i \sim \mathcal{N}(0, \varepsilon^2),$$

where $\varepsilon > 0$ and $N = (N_t)_{t\in[0,1]}$ is a Poisson process with intensity $\lambda > 0$. The parameters η, ε, and λ are nuisance parameters, whereas σ is unknown to be estimated from the discretely observed data $(X_{t_i^n})_{i=0,1,\ldots,n}$.

There are already several parametric estimation methods for stochastic differential equations with jumps. Among them, Shimizu and Yoshida [20] proposed a local threshold method for optimal parametric estimation. They used method of jump detection by comparing each increment $|\Delta_i X|$ with h_n^ρ, where $h_n = t_i^n - t_{i-1}^n$ is the time interval and $\rho \in (0, 1/2)$. More precisely, an increment $\Delta_i X$ satisfying $|\Delta_i X| > h_n^\rho$ is regarded as being driven by the compound Poisson jump part, and is removed when constructing the likelihood function of the continuous part. The likelihood function of the continuous part is defined by

$$l_n(\sigma) = \sum_{i=1}^{n} \left[-\frac{1}{2\sigma^2 h_n} |\bar{X}_i^n|^2 - \frac{1}{2} \log \sigma^2 \right] \mathbf{1}_{\{|\Delta X_i| \le h_n^\rho\}},$$

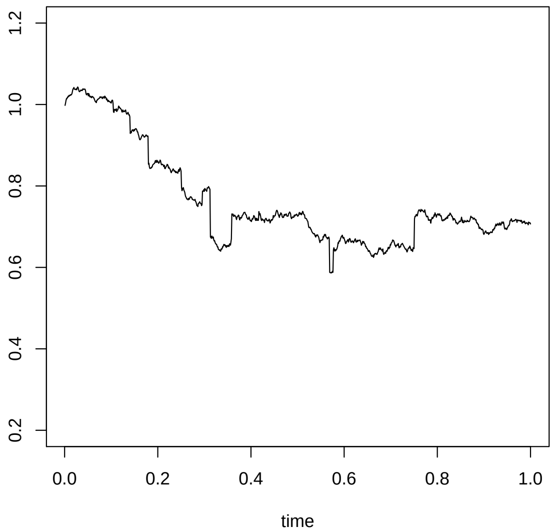
where $\bar{X}_i^n = X_{t_i^n} - X_{t_{i-1}^n} + \eta X_{t_{i-1}^n} h_n$. Obviously, the jump detection scheme is essentially different from our approach in this paper. They do not use any other increments to determine whether an increment has a jump or not. Our approach, however, uses all the increments.

Shimizu and Yoshida [20] proved that this estimator is consistent as the sample size n tends to infinity; that is, asymptotic property of the local and the global threshold approaches are the same from the viewpoint of consistency. However, precision of jump detection may be different in the case of (large but) finite samples. Comparison of two approaches is the main purpose of this section.

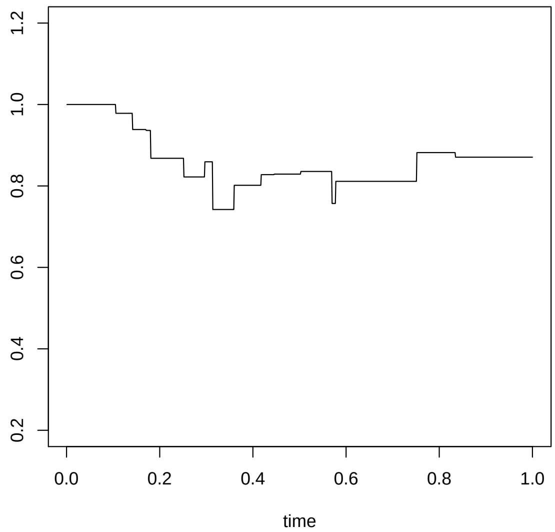
In our setting, however, we assume that the jump size is normally distributed, the case of which is not dealt with in Shimizu and Yoshida [20]. In their original paper, they assume that the jump size must be bounded away from zero. Ogihara and Yoshida [14] accomodated a restrictive assumption on the distribution of jump size. They proved that the local threshold estimator works well under this assumption by using some elaborate arguments. Hence, the local estimator can be used in our setting and thus we can compare its estimates with the global threshold estimator.

Note that, we do not impose too restrictive assumption about the distribution of jump sizes in our paper: we only assume natural moment conditions on the number of jumps. Versatility in this sense can be regarded as the advantage of our approach.

The setting of the simulation is as follows. The initial value is $X_0 = 1$. The true value of the unknown parameter σ is 0.1. Other parameters are all known and given by $\eta = 0.1$, $\varepsilon = 0.05$, and $\lambda = 20$. The sample size is $n = 1,000$ in Section 6.2 to see the accuracy of the jump detection of our filter and $n = 5,000$ in Section 2.6.3 and thereafter to compare the estimates of each estimator. We assume the equidistant case, so that $h_n = 1/n = 0.001$ and $h_n = 0.0002$. Since the time horizon is



(a) Sample path of X



(b) Sample path of the jump part J of X

Figure 2.1: Sample paths of X and its jump part

now finite and η is not consistently estimable, we set η in $l_n(\sigma)$ at the true value 0.1, that is the most preferable value for the estimator in Shimizu and Yoshida [20].

In applying the global estimator, we need to set several tuning parameters. we set $C_*^{(k)} = 1$ for the truncation function $K_{n,j}^{(k)}$ in (2.2.2), that is used for the definition of α-quasi-log likelihood function. For the one-step global estimator, we use the parameter $C_*^{(k)} = 1$ and $\delta_0 = 1/5$ for the truncation function $K_{n,j}^{(k)} = 1_{\left\{V_j^{(k)} < C_*^{(k)} n^{-\frac{1}{4}-\delta_0}\right\}}$. Moreover, we set $\delta_1^{(k)} = 4/9$ so that $p_n^{(k)} = (n - \lfloor n^{4/9} \rfloor)/n$ in the definition of the moving threshold quasi-likelihood function in (2.3.1).

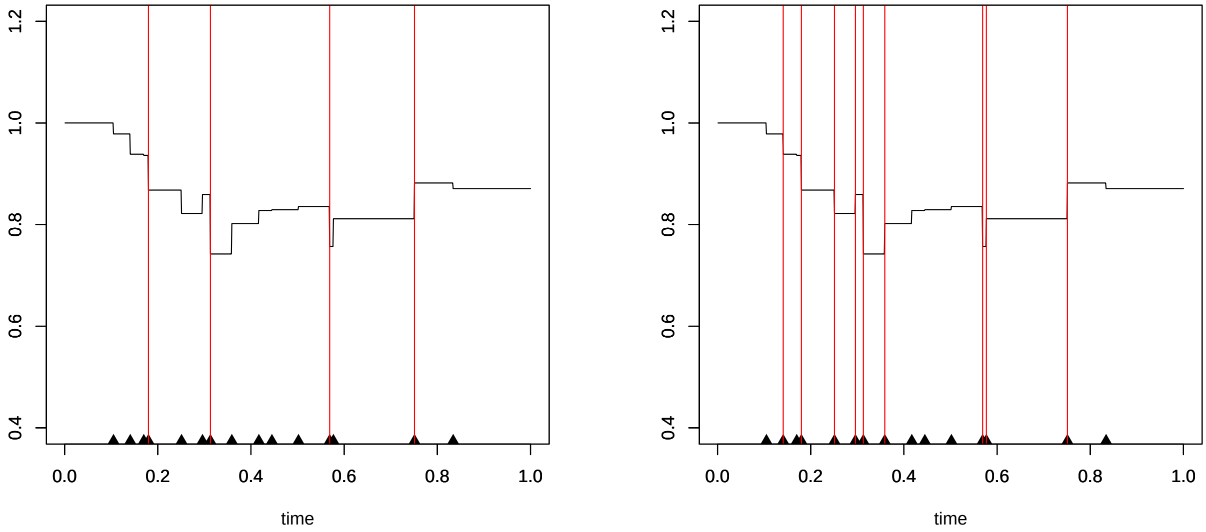
Figure 1 shows a sample path of (X, J). The left panel is the sample path of X and the right panel is its jump part J. Note that the jump part is not observable and thus we need to discriminate the jump from the sample path of X.

2.6.2 Accuracy of jump detection

Before comparing the results of parameter estimation, we check the accuracy of jump detection of each estimation procedure. If there are too many misjudged increments, the estimated value can have a significant bias. Hence it is important how accurately we can eliminate jumps from the observed data X.

Local threshold method

First, we check the accuracy of jump detection of the local threshold method. Figure 2 shows the results of jump detection by the local threshold method of Shimizu and Yoshida [20] for $\rho = 1/3$ in panel (a) and $\rho = 1/2$ in panel (b). The red vertical lines indicate the jump detected by each estimator, whereas the triangles on the horizontal axis indicate the true jumps. As these figures show, the accuracy of the jump detection heavily depend on a choice of the tuning parameter ρ. For relatively



(a) $\rho = \frac{1}{3}$ (b) $\rho = \frac{1}{2}$

Figure 2.2: Results of jump detection by local threshold method

small ρ (say $\rho = 1/3$), we cannot completely detect jumps: the estimator detects only one jump for $\rho = 1/3$. On the other hand, in the case of (theoretically banned) $\rho = 1/2$, the estimator detects the jumps better than the case of $\rho = 1/3$. Note that the case of $\rho = 1/2$ is not dealt with in Shimizu and Yoshida [20], but it is useful for us to compare the local threshold method with the global threshold method later and so we show the result of the exceptional case.

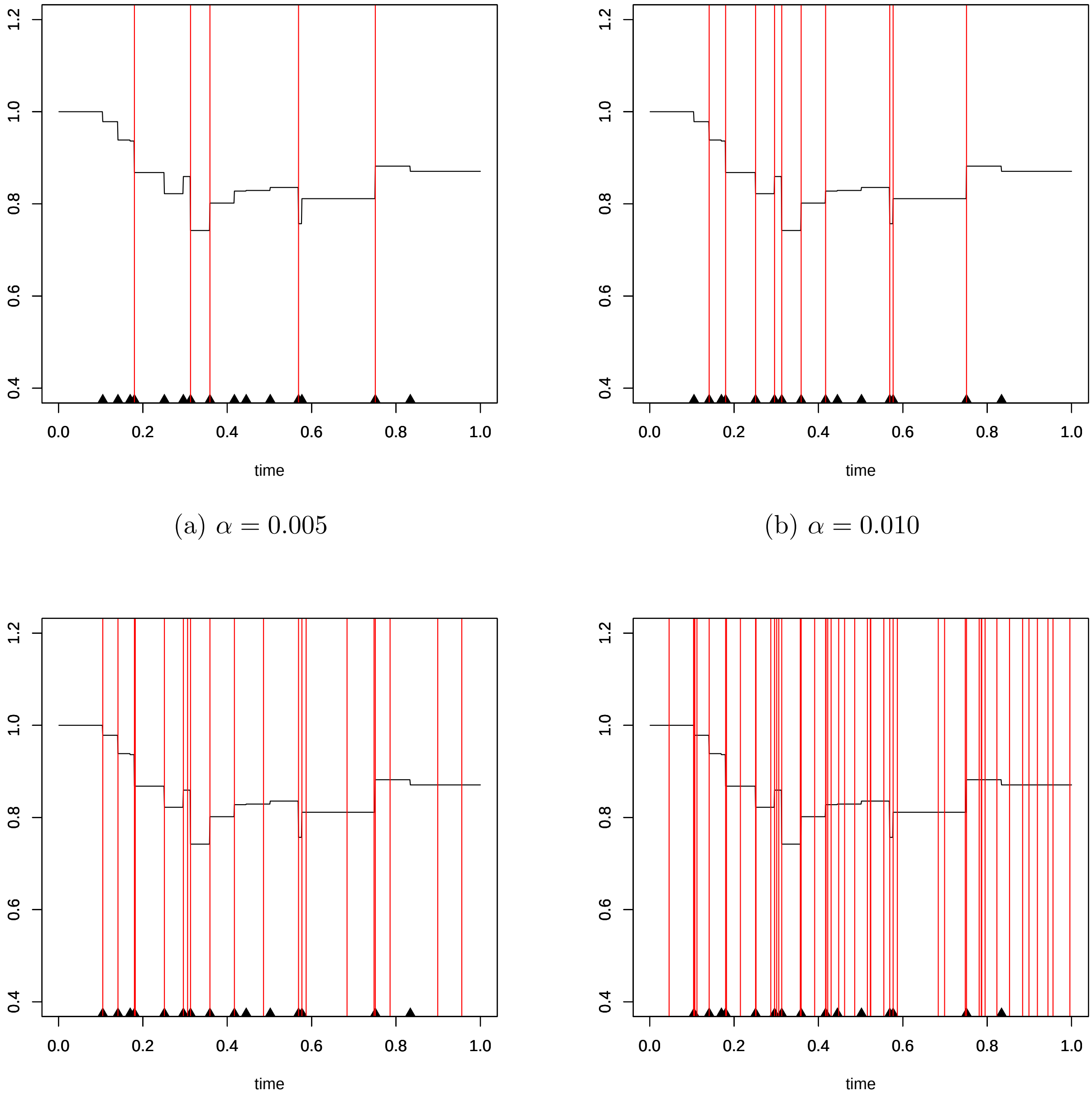
Global threshold method

Next, we discuss the jump detection by global threshold method. The accuracy of jump detection depends on the tuning parameter $\alpha \in (0, 1)$, so we here show results of four cases, namely, the case $\alpha = 0.005, 0.010, 0.020, 0.050$.

From the figures, we see that the too small α cannot detect jumps sufficiently, mistakenly judging some genuine jumps as increments driven by the continuous part, which is similar to the case of small ρ of the Shimizu-Yoshida estimator. By setting α a little larger, the accuracy of jump detection increases, as shown in panels (b) and (c). On the other hand, too large α discriminate too many increments as jumps, as panel (d) shows. In this case, there are many increments that are regarded as jumps but are actually generated by the continuous part of the process only. These figures suggests that one should choose the tuning parameter α carefully to detect jumps appropriately.

We show the false negative / positive ratio of jump detection in Table 2.1. Note that *false negative* means that our method did not judge an increment as a jump, despite it was actually driven by the compound Poisson jump part. The meaning of *false positive* is the opposite; that is, our method judged an increment which was not driven by the jump part as a jump.

The false negative ratio for small α tends to be large because in this case the estimator judges only big increments as jumps, and ignores some jumps of intermediate size. On the other hand, the false positive ratio for large α is high, since the estimator judges small increments as jumps, but almost increments are actually driven by the continous part. From this table as well, we can infer that there



(a) $\alpha = 0.005$

(b) $\alpha = 0.010$

(c) $\alpha = 0.020$

(d) $\alpha = 0.050$

Figure 2.3: Results of jump detection by global threshold method

Table 2.1: False Negative/Positive ratio of jump detection

alpha	0.005	0.01	0.015	0.02	0.025	0.05	0.1	0.25
False Negative	73.333	40.000	26.667	26.667	26.667	26.667	26.667	20.000
False Positive	0.000	0.000	0.305	0.812	1.320	3.858	8.934	24.061

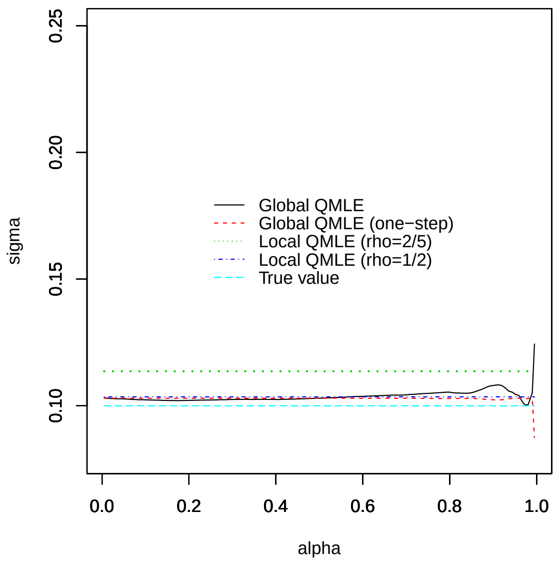


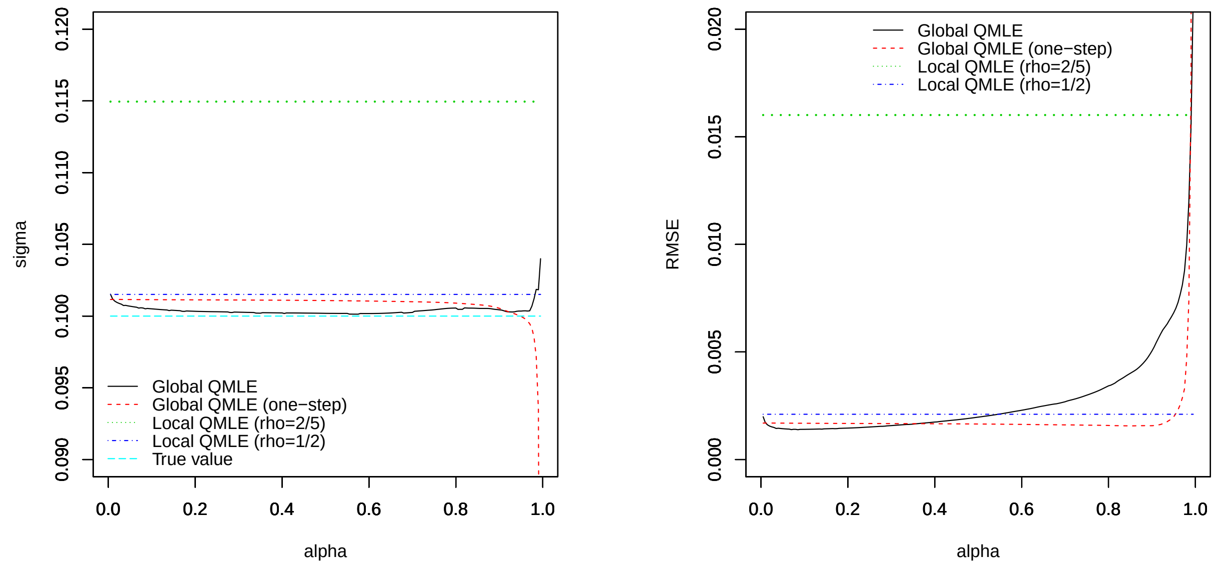
Figure 2.4: Comparison of estimators given a sample path

should be some optimal range of α for jump detection. In any case, a large value of false negative may seriously bias the estimation, while a large value of false positive only decreases efficiency. Sensitivity of the local filter is also essentially observed by this experiment since each value of α of the global filter corresponds to a value of the threshold Lh^{ρ} of the local filter.

2.6.3 Comparison of the estimators

Next, we investigate the estimation results of the global threshold method. In this section, we set the number of samples $n = 5,000$ to let the biases of the estimators as small as possible. Since the estimator depends on the parameter α, we check the stability of the estimator with respect to the parameter α. Remember that too small α is not able to detect jumps effectively, but too large α mistakenly eliminates small increments driven by the Brownian motion which should be used to construct the likelihood function of the continuous part. So there would be a suitable level α.

Figure 3.4 compares the global QMLEs with the local QMLE with $\rho = 2/5$, as $\rho = 1/2$ is theoretically prohibited, and suggests that the global methods are superior to the local methods. Figure 3.4 also compares the performance of the global threshold estimator $\hat{\sigma}_n^{M,\alpha}$ and the one-step estimator $\check{\sigma}_n^{M,\alpha}$ with α ranging in $(0, 1)$, as well as that of the local filters. Here we used $\kappa = 3$ to construct the one-step estimator; that is, the one-step estimator is given by $\check{\sigma}_n^{M,\alpha} = \hat{\sigma}_n^{M,\alpha} + A_1(\hat{\sigma}_n^{M,\alpha})$, where the adjustment term A_1 is defined in Section 2.4. As the figure shows, for suitably small α, both the estimate $\hat{\sigma}_n^{M,\alpha}$ and $\check{\sigma}_n^{M,\alpha}$ are well close to σ. However, as this figure indicates, the global threshold



(a) Average estimates (b) RMSE

Figure 2.5: Results of jump detection by local threshold method: comparison of averaged results

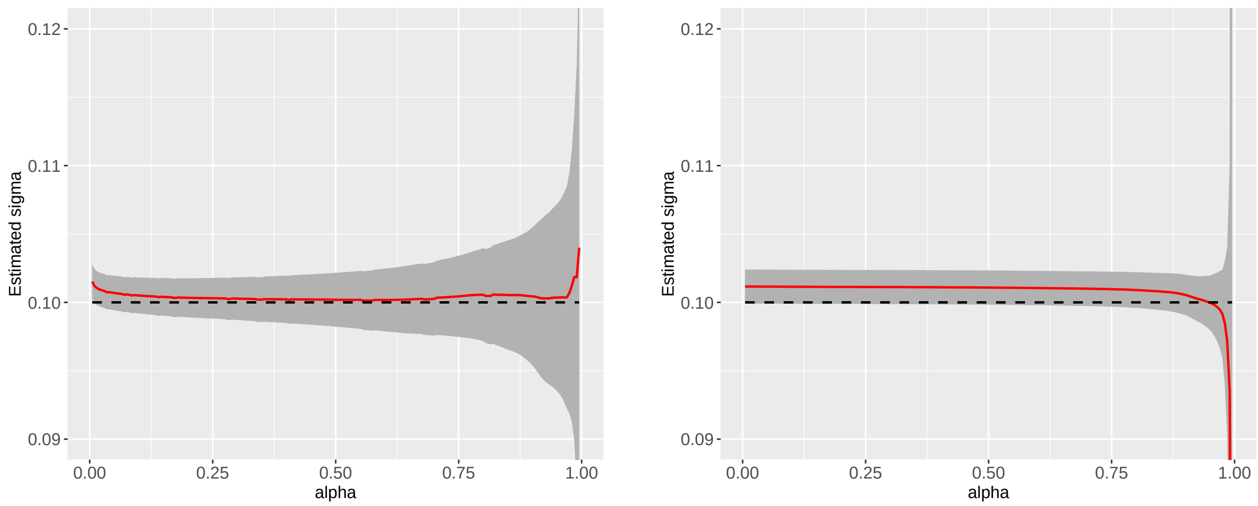
estimator may be somewhat unstable with respect to the choice of α. Although the global estimator with moving α and one-step global estimator are asymptotically equivalent, when we use the original global estimator, it would be recommended to use the one-step estimator as well and to try estimation for several α's in order to check the stability of the estimates.

To compare statistical properties of the estimators, we used the 100 outcomes of Monte Carlo simulation to calculate the average estimates, the root mean square error (RMSE), and the standard deviation of this experiment. Looking at the average values of the estimators shown in the Figure 3.5 (a), we see the global threshold estimators outperform the local threshold estimator. It is concluded that the accuracy of the global estimator is not dependent on a sample path. High average accuracy can also be checked by RMSE. As shown in Figure 3.5 (b), RMSEs of the global estimators are smaller than those of the local estimators, except for the extreme choices of α.

Figure 2.6 indicates the estimates for global QMLE estimator with standard error band. The standard errors are calculated by using 100 Monte carlo trials. It shows that the global QMLE estimator works very well with or without one-step adjustment. We can see, however, the one-step adjusted estimator is robust against the choice of the tuning parameter α. For large α, the global threshold tends to eliminate increments that are not driven by the jump part of the underlying process, and this could result in the large standard deviation of the estimate. The one-step estimator works well for such large α.

A suitably chosen α will yield a good estimate of the unknown parameter, although too small or too large α might tends to bias the estimate. The global threshold estimator seems to generally be robust to the choice of the tuning parameters. The optimal choice of α depends on the situation. Hence, it is desirable to use several values of α and to compare the results to determine the preferable value of α in using the global estimator. Moreover, it is worth considering of using one-step adjustment to get more robust estimates.

The global filter sets a number for the critical value of the threshold though it is determined after observing the data. In this sense, the global filter looks similar to the local filter, that has a



(a) global QMLE (b) global QMLE (one-step)

Figure 2.6: Estimation results of global QMLE estimator with standard error band

predetermined number as its threshold. However, the critical values used by the two methods are fairly different in practice. We consider the situation where, for some n, the local filter with threshold Lh^ρ approximately performs as good as the global filter with α. For simplicity, let us consider a one-dimensional case with $\sigma(x,\theta) = 1$ constantly. Hence the critical value should approximately be near to the upper $\alpha/2$-quantile of $\Delta_j w$. Moreover, let $n = 10^3$, $\rho = 2/5$ and $\alpha = 0.1$. Then the constant L in the threshold of the local filter should satisfy $(10^{-3})^{-1/2} \times 1.64 = (10^{-3})^{-\rho} L$, namely, $L \sim 3.27$ approximately. Since L is a predetermined common constant for different numbers n, the critical value of the threshold of the local filter becomes $10^{-5\rho}L \sim 0.0327$ when $n = 10^5$, while the threshold of the global filter is about $(10^{-5})^{1/2} \times 1.64 \sim 0.00519$. Some of jumps may not be detected by the local filter, since its critical value is not so small, compared with $\epsilon = 0.05$.

2.6.4 Asymmetric jumps

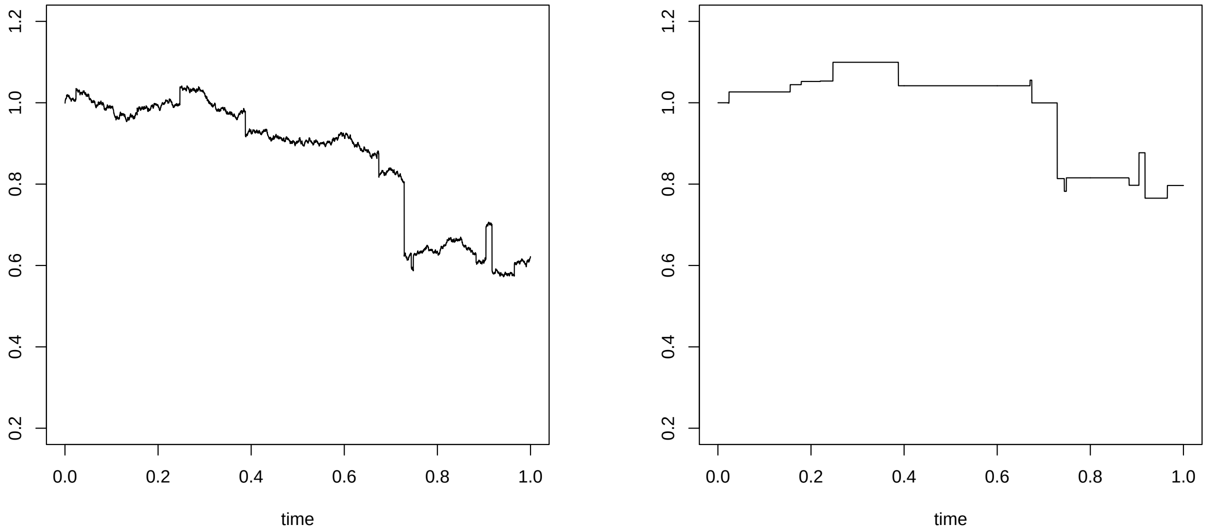
In the previous subsection, we assumed that the distribution of jump size was centered Gaussian and thus symmetric. In a real situations, however, the distribution of the size of jumps might be not symmetric. For example, stock prices have an asymmetric distribution with heavier tail in negative price changes. In this subsection, we show that our global estimator performs well for jumps with asymmetric distribution.

Although there are many asymmetric jumps in applications, we use just a normal distribution with a negative average because heavier tails would make jump detection easier. More precisely, we assume that the jump process J is given by

$$J_t = \sum_{i=1}^{N_t} \xi_t, \qquad \xi_i \sim \mathcal{N}(\mu, \varepsilon^2),$$

where $\mu = -0.01$ and $\varepsilon = 0.05$. In this setting, as shown in Figure 2.7, negative jumps appear more frequently than positive ones.

As Figure 2.8 shows, the global estimator performs well even in the case of asymmetric jumps. The estimates are well similarly to those in the case of symmetric jumps in the previous subsection. This



(a) Sample path of X (b) Sample path of the jump part J of X

Figure 2.7: Sample paths of X and its jump part: in the case of asymmetric jump distribution

example implies that out estimator will work very well under realistic circumstances, like financial time series where changes in asset prices have an symmetric distribution with heavy tail in negative price changes.

2.6.5 Location-dependent diffusion coefficient

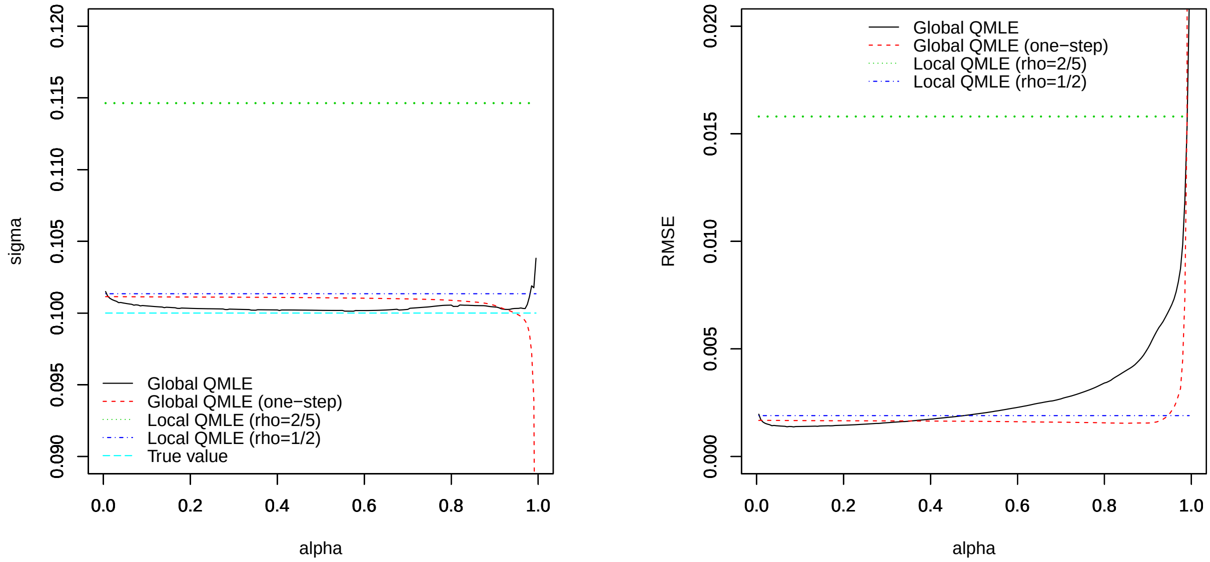
Here we assume that the diffusion coefficient is given by $\sigma\sqrt{1+x^2}$, where σ is an unknown positive parameter to be estimated. Other settings are entirely the same as those given in the Section 2.6.1. In particular, we assume that the distribution of jump size is centered, contrary to the previous subsection.

In this example, we have to set an estimator $\bar{S}_{n,j-1}$ of the volatility matrix, $\left(\sigma\sqrt{1+X^2_{t^n_{j-1}}}\right)^2$, which satisfies the condition [F2](ii). It is obvious that we can choose $\bar{S}_{n,j-1} = 1 + X^2_{t^n_{j-1}}$ to satisfy the condition. The results are shown in Figure 2.9. Like in the case of constant coefficient, the global estimators perform well. Except for too small or large α for which the estimates are unstable and different from those of the case of constant diffusion coefficient, our estimators yield a good estimate even in the case of location-dependent diffusion coefficient.

2.7 Further topics and future work

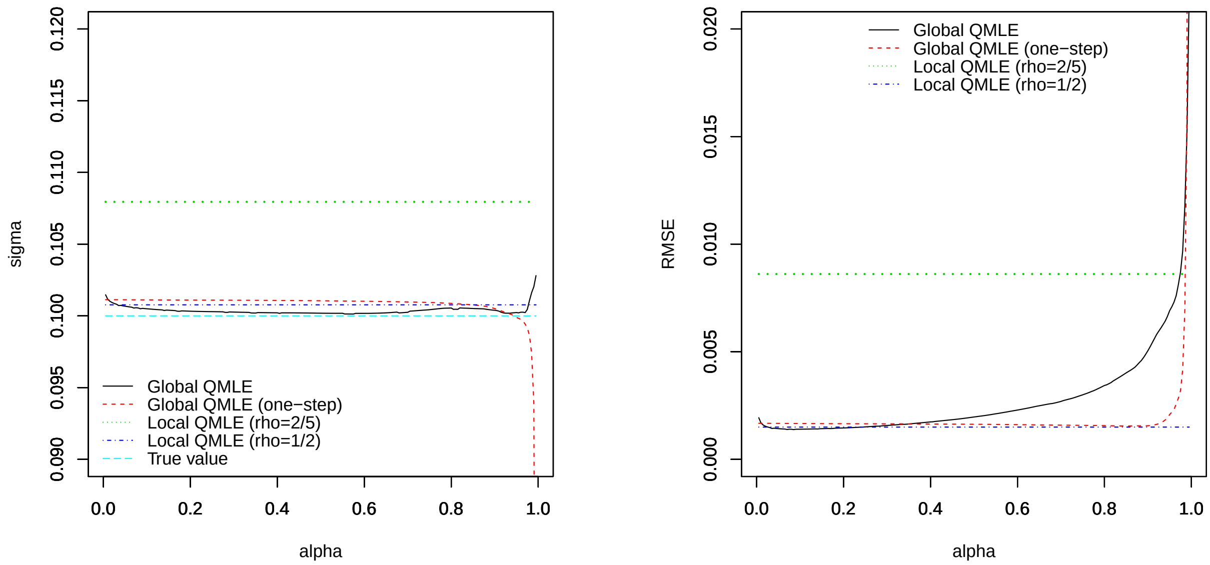
In this paper, we payed main attention to removing jumps and to obtaining stable estimation of the diffusion parameter. The removed data consist of relatively large Brownian increments and the increments having jumps. Then it is possible to apply a suitable testing procedure to the removed data, e.g., the goodness-of-fit test for the cut-off normal distribution, in order to test existence of jumps.

It is also possible to consider asymptotics where the intensity of jumps goes to infinity at a moderate



(a) Average estimates (b) RMSE

Figure 2.8: Results of jump detection: in the case of asymmetric jump disribution



(a) Average estimates (b) RMSE

Figure 2.9: Results of jump detection: a location-dependent diffusion coefficient

rate that does not essentially change the argument of removing jumps. In such a situation, estimation of jumps becomes an issue. Probably, some central limit theorem holds for the error of the estimators of the structure of jumps. Furthermore, a statistical test of the existence of jumps will be possible in this framework. The ergodic case as $T \to \infty$ will be another situation where the parameters of jumps are estimable.

The global jump filter was motivated by data analysis. This scheme is to be implemented on YUIMA, a comprehensive R package for statistical inference and simulation for stochastic processes.

Chapter 3

Application of Global Jump Filters to Estimation of Integrated Volatility

3.1 Model

Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$. We consider a one-dimensional semimartingle $X = (X_t)_{t\in[0,T]}$ having a decomposition

$$X_t \quad = \quad X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dw_s + J_t \quad (t \in [0,T])$$

where X_0 is an $\mathcal{F}_0$-measurable random variable, $b = (b_t)_{t\in[0,T]}$ and $\sigma = (\sigma_t)_{t\in[0,T]}$ are càdlàg $\mathbb{F}$-adapted processes, and $w = (w_t)_{t\in[0,T]}$ is an $\mathbb{F}$-standard Wiener process. $J = (J_t)_{t\in[0,T]}$ is the jump part of X. We will assumed that J is finitely active, that is, $J_t = \sum_{s\in(0,t]} \Delta X_s$ for $\Delta X_s = X_s - X_{s-}$ and $\sum_{t\in[0,T]} 1_{\{\Delta J_t \neq 0\}} < \infty$ a.s. In this paper, we are interested in the estimation of the integrated volatility

$$\Theta \quad = \quad \int_0^T \sigma_t^2 dt \tag{3.1.1}$$

based on the data $(X_{t_j})_{j=0,1,\ldots n}$, where $t_j = t_j{}^n = jT/n$.

The jump part J can be endogenous or exogenous, as well as b and σ, however, J is a nuisance in any case. The simple realized volatility is heavily damaged when jumps exist. To avoid the effects of the jumps, various methods have been proposed so far. For example, the bipower variation (Barndorff-Nielsen and Shephard. [2], Barndorff-Nielsen et al. [3]) and the minimum realized volatility (Andersen et al. [1]) are shown to be consistent estimators of the integrated volatility even in the presence of jumps. The idea of these methods is that, to mitigate the effect of jumps, they employ adjacent increments in constructing the estimator.

Another direction to handle jumps is to introduce a threshold to detect jumps. Threshold method was investigated in Shimizu and Yoshida [20] in the context of the parametric inference for a stochastic differential equation with jumps. The idea of thresholding goes back to the studies of limit theorems for Lévy processes as latest. Mancini [12] introduced a nonparametric threshold that detect jumps by observing the size of increments within each time interval. The threshold is defined as a function of the length of a time inverval. If an increment is so large that exceeds a threshold, it is regarded as a jump. Koike [11] applied the threshold method to covariance estimation for asynchronously observed semimartingales with jumps.

However, examining each individual increment is not always effective in finding jumps. It sometimes overlooks relatively small jumps. To tackle this problem, Inatsugu and Yoshida [8] introduced global filters that examine all increments simultaneously and regard an increment of high rank in order of absolute size as a jump. Using the information about the size of other increments helps us detect jumps more accurately than the previous method that ignores such information.

In this paper, we apply the global filtering method to nonparametric volatility estimation. Specifically, we construct the "global realized volatility (GRV) estimator" of the integrated volatility for the a stochastic differential equation with jumps. We investigate the theoretical properties of GRV and then conduct numerical simulations to study their performance compared with well-known methods, that is, the bipower variation and the minimum realized volatility.

The organization of this chapter is as follows. In Section 3.2, we introduce the GRV and its variant, the winsorized GRV (WGRV). In Section 3.3, we also introduce the –global realized volatility (LGRV) and prove its convergence to spot volatilities. The LGRV is needed to normalize the increments and construct the GRV. In Section 3.4, we prove the rate of convergence of the GRV and WGRV in the situation where the intensity of jumps is high. In this case, need a high and fixed cut-off rate α. In Section 3.5, we allow the cut-off rate to vary according to the sample size. This "moving threshold" method is for the situation where the intensity of jumps is moderate and small cut-off rate is applicable. In Section 3.6, we discuss the situation where true volatility is constant. In this case, we do not need normalizing increments, so the estimator gets a little simpler. In Section 3.7, we show simulation results to compare the performance of the GRV, WGRV, bipower varition, and the mininum realized volatility.

3.2 Realized volatilities with a global jump filter

The global jump filter introduced by Inatsugu and Yoshida [8] uses the order statistics of the transformed increments of the observations. Suppose that an estimator $S_{n,j-1}$ of the spot volatility $\sigma(X_{t_{j-1}})^2$ (up to a common scaling factor) is given for each $j \in I_n = \{1, ..., n\}$. Denote $\Delta_j U = U_{t_j} - U_{t_{j-1}}$ for a process $U = (U_t)_{t\in[0,T]}$. Then the distribution of the scaled increment $S_{n,j-1}^{-1/2}\Delta_j X$ is expected to be well approximated by the standard normal distribution $N(0, 1)$. Therefore, if the value

$$V_j \quad = \quad \left|(S_{n,j-1})^{-1/2}\Delta_j X\right| \tag{3.2.1}$$

is relatively very large among $\mathcal{V}_n = \{V_k\}_{k\in I_n}$, then plausibly we can infer that the V_j involves jumps with high probability. The idea of the global jump filter is to eliminate the increment $\Delta_j X$ from the data if the corresponding V_j is ranked within the top $100\alpha\%$ in $\mathcal{V}_n$. More precisely, let

$$\mathcal{J}_n(\alpha) \quad = \quad \left\{j \in I_n;\ V_j < V_{(s_n(\alpha))}\right\}$$

where

$$s_n(\alpha) \quad = \quad \lfloor n(1-\alpha)\rfloor$$

for $\alpha \in [0, 1)$, and we denote by $r_n(U_j)$ the rank of U_j among the variables $\{U_i\}_{i\in I_n}$. Let

$$q(\alpha) \quad = \quad \int_{\{|z|\le c(\alpha)^{1/2}\}} z^2\phi(z; 0, 1)dz$$

where $\phi(z; 0, 1)$ is the density function of $N(0, 1)$ and $c(\alpha)$ defined by

$$P\left[\zeta^2 \le c(\alpha)\right] \quad = \quad 1-\alpha$$

for $\zeta \sim N(0,1)$ and $\alpha \in [0,1)$. Then the **global realized volatility** (globally truncated realized volatility, GRV) with cut-off ratio α is defined by

$$\mathbb{V}_n(\alpha) \quad = \quad \sum_{j\in\mathcal{J}_n(\alpha)} q(\alpha)^{-1}|\Delta_j X|^2 K_{n,j} \tag{3.2.2}$$

where $K_{n,j} = 1_{\{|\Delta_j X|\leq n^{-1/4}\}}$. As remarked in Inatsugu and Yoshida [8], the indicator function $K_{n,j}$ is set just for relaxing the conditions for validation. Generalization by using like $1_{\{|\Delta_j X|\leq B_1 n^{-\delta_1}\}}$ with constants $B_1 > 0$ and $\delta_1 \in (0, 1/4]$ is straightforward, but we prefer simplicity in presentation of this article. In practice, the probability that $K_{n,j}$ executes the task is exponentially small by the large deviation principle. However, the moments of ΔJ_t are not controllable without assumption, and we can simply avoid it by the cut-off function $K_{n,j}$.

Winsorization is a popular technique in robust statistics. In the present context, the **Winsorized global realized volatility** (WGRV) is given by

$$\mathbb{W}_n(\alpha) \quad = \quad \sum_{j=1}^{n} \mathsf{w}(\alpha)^{-1}\big\{|\Delta_j X| \wedge \big(S_{n,j-1}^{1/2} V_{(s_n(\alpha))}\big)\big\}^2 K_{n,j}$$

where

$$\mathsf{w}(\alpha) \quad = \quad \int_{\mathbb{R}} \big(z^2 \wedge c(\alpha)\big)\phi(z;0,1)dz.$$

The cut-off ratio $\alpha \in [0,1)$ is a tuning parameter in estimation procedures. The bigger α provides the more stable estimates even in high intensity of jumps. On the other hand, the smaller α gives the more precise estimates if the intensity of jumps is low. Making trade-off between stability and precision is necessary in practice. As a matter of fact, these cases require different theoretical treatments. We will consider fixed α in Section 3.4, and shrinking α in Section 3.5.

3.3 Local-global filter

3.3.1 Glocal filtering lemmas

For each $j \in I_n$, let

$$\underline{j}_n \quad = \quad \begin{cases} 1 & (j \leq \kappa_n) \\ j-\kappa_n & (\kappa_n+1 \leq j \leq n-\kappa_n) \\ n-2\kappa_n & (j \geq n-\kappa_n+1) \end{cases}$$

for $\kappa_n \in \mathbb{Z}_+$ satisfying $2\kappa_n + 1 \leq n$. Let $I_{n,j} = \{\underline{j}_n, \underline{j}_n + 1, ..., \underline{j}_n + 2\kappa_n\}$. Let

$$\widehat{U}_{j,k} \;=\; h^{-1/2}\sigma^{-1}_{t_{\underline{j}_n-1}}\Delta_k X \quad \text{and} \quad W_j \;=\; h^{-1/2}\Delta_j w$$

for $j,k \in I_n$. Both variables $\widehat{U}_{j,k}$ and W_j depend on n. Let

$$\widehat{R}_{j,k} \quad = \quad \widehat{U}_{j,k} - W_k - h^{-1/2}\sigma^{-1}_{t_{\underline{j}_n-1}}\Delta_k J$$

for $j,k \in I_n$. Denote $L^{\infty-} = \cap_{p>1}L^p$.

Let $N = \sum_{s\in(0,\cdot]} 1_{\{\Delta J_s\neq 0\}}$. Let $\widetilde{\sigma} = \sigma - J^\sigma$ for $J^\sigma = \sum_{s\in(0,\cdot]}\Delta\sigma_s$, and let $N^\sigma = \sum_{s\in(0,\cdot]} 1_{\{\Delta J^\sigma_s\neq 0\}}$. We assume that $N^\sigma_T < \infty$ a.s. Moreover, let $\overline{N} = N + N^\sigma$. Let $\widetilde{X} = X - J$. A counting process will be identified with a random measure. Let $\mathtt{I}_{n,j} = \big(t_{\underline{j}_n-1}, t_{\underline{j}_n+2\kappa_n}\big]$.

[G1] **(i)** For every $p > 1$, $\sup_{t\in[0,T]} \|\sigma_t\|_p < \infty$ and

$$\left\|\widetilde{\sigma}_t - \widetilde{\sigma}_s\right\|_p \leq C(p)|t-s|^{1/2} \quad (t,s \in [0,T])$$

for some constant $C(p)$ for every $p > 1$.

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

(iii) $\sigma_t \neq 0$ a.s. for every $t \in [0,T]$, an $\sup_{t\in[0,T]} \left\|\sigma_t^{-1}\right\|_p < \infty$ for every $p > 1$.

Lemma 3.3.1. *Under [G1],*

$$\sup_{j\in I_n} \sup_{k\in I_{n,j}} \left\|\widehat{R}_{j,k} 1_{\{N^\sigma(\mathtt{I}_{n,j})=0\}}\right\|_p = O\left(\left(\frac{\kappa_n}{n}\right)^{1/2}\right) \tag{3.3.1}$$

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

Proof. For $j \in I_n$, let $E(j) = \{N^\sigma(\mathtt{I}_{n,j}) = 0\}$. Then, for $k \in I_{n,j}$,

$$\begin{aligned} \widehat{R}_{j,k} 1_{E(j)} &= \left(h^{-1/2}\sigma^{-1}_{t_{\underline{j}_n-1}} \Delta_k \widetilde{X} - h^{-1/2}\Delta_k w\right) 1_{E(j)} \\ &= h^{-1/2} \int_{t_{k-1}}^{t_k} \sigma^{-1}_{t_{\underline{j}_n-1}} (\widetilde{\sigma}_t - \widetilde{\sigma}_{t_{\underline{j}_n-1}}) dw_t 1_{E(j)} \\ &\quad + h^{-1/2}\sigma^{-1}_{t_{\underline{j}_n-1}} \int_{t_{k-1}}^{t_k} b_t dt 1_{E(j)} \end{aligned} \tag{3.3.2}$$

We obtain (3.3.1) by applying the Burkholder-Davis-Gundy inequality to the martingale part of (3.3.2) after the trivial estimate $1_{E(j)} \leq 1$. □

For $j \in I_n$, denote by $r_{n,j}(U_k)$ the rank of the element U_k among a collection of random variables $\{U_\ell\}_{k\in I_{n,j}}$. Let

$$\begin{aligned} &0 < \eta_2 < \eta_1, \qquad \overline{\kappa}_n = 2\kappa_n + 1, \\ &\mathsf{a}_n = \lfloor (1-\alpha_0)\overline{\kappa}_n - \overline{\kappa}_n^{1-\eta_2} \rfloor, \qquad \widehat{\mathsf{a}}_n = \lfloor \mathsf{a}_n - \overline{\kappa}_n^{1-\eta_2} \rfloor \end{aligned}$$

for $\alpha_0 \in [0,1)$. Let

$$L_{n,j,k} = \left\{r_{n,j}(|W_k|) \leq \mathsf{a}_n - \overline{\kappa}_n^{1-\eta_2}\right\} \cap \left\{|W|_{(j,\mathsf{a}_n)} - |W_k| < \overline{\kappa}_n^{-\eta_1}\right\} \tag{3.3.3}$$

where $\left(|W|_{(j,k)}\right)_{k\in I_{n,j}}$ are the ordered statistics made from $\{|W_k|\}_{k\in I_{n,j}}$. In the same way as Lemma 1 of Inatsugu and Yoshida [8], we obtain the following result.

Lemma 3.3.2. *Let $\alpha_0 \in (0,1)$. Suppose that $\eta_1 < 1/2$ and that $n^{-\epsilon}\kappa_n \to \infty$ as $n \to \infty$ for some $\epsilon \in (0,1)$. Then*

$$\sup_{j\in I_n} P\left[\bigcup_{k\in I_{n,j}} L_{n,j,k}\right] = O(n^{-L})$$

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

Define $\mathcal{K}_{n,j}(\alpha_0)$ by

$$\mathcal{K}_{n,j}(\alpha_0) \quad = \quad \big\{k \in I_{n,j};\ r_{n,j}(|\Delta_k X|) \le (1-\alpha_0)\overline{\kappa}_n\big\},$$

where $r_{n,j}(|\Delta_k X|)$ is the rank of $|\Delta_k X|$ among $\{|\Delta_{k'} X|\}_{k' \in I_{n,j}}$. Let

$$\widehat{\mathcal{K}}_{n,j}(\alpha_0) \quad = \quad \big\{k \in I_{n,j};\ r_{n,j}(|W_k|) \ \le \ \hat{\mathsf{a}}_n\big\}.$$

Let

$$\Omega_{n,j} \quad = \quad \bigcap_{k \in I_{n,j}} \left[\left\{|\widehat{R}_{j,k}| 1_{\{N^\sigma(\mathtt{I}_{n,j})=0\}} < 2^{-1}\overline{\kappa}_n^{\ -\eta_1}\right\} \cap L^c_{n,j,k}\right].$$

Let

$$\mathcal{L}_n = \big\{j \in I_n;\ \overline{N}(\mathtt{I}_{n,j}) \ne 0\big\}. \tag{3.3.4}$$

Lemma 3.3.3. (a) *$\widehat{\mathcal{K}}_{n,j}(\alpha_0) \subset \mathcal{K}_{n,j}(\alpha_0)$ on $\Omega_{n,j}$ if $j \in \mathcal{L}^c_n$.*

(b) $1_{\Omega_{n,j}} 1_{\{j \in \mathcal{L}^c_n\}} \#\big(\mathcal{K}_{n,j}(\alpha_0) \setminus \widehat{\mathcal{K}}_{n,j}(\alpha_0)\big) \le 4\,\overline{\kappa}_n^{\,1-\eta_2} \quad (j \in I_{n.j},\ n \in \mathbb{N}).$

Proof. Let $n \in \mathbb{N}$ and suppose that $j \in \mathcal{L}^c_n$. We will work on $\Omega_{n,j}$. For a pair $(k_1, k_2) \in I^2_{n,j}$, suppose that

$$r_{n,j}(|W_{k_1}|) \le \widehat{\mathsf{a}}_n \quad \text{and} \quad r_{n,j}(|W_{k_2}|) \ge \mathsf{a}_n. \tag{3.3.5}$$

Then $|\widehat{U}_{j,k_1}| < |W_{k_1}| + 2^{-1}\overline{\kappa}_n^{\ -\eta_1}$, since $\Delta_{k_1} N = 0$ and $N^\sigma(\mathtt{I}_{n,j}) = 0$ when $j \in \mathcal{L}^c_n$, and then $|\widehat{R}_{j,k_1}| < 2^{-1}\overline{\kappa}_n^{\ -\eta_1}$ on $\Omega_{n,j}$. By the first inequality of (3.3.5), $r_{n,j}(|W_{k_1}|) \le \mathsf{a}_n - \overline{\kappa}_n^{\,1-\eta_2}$, and hence on $\Omega_{n,j} \subset L^c_{n,j,k_1}$, we have $|W|_{(j,\mathsf{a}_n)} - |W_{k_1}| \ge \overline{\kappa}_n^{\ -\eta_1}$ by the definition (3.3.3) of $L_{n,j,k}$. Therefore

$$|\widehat{U}_{j,k_1}| \quad < \quad |W|_{(j,\mathsf{a}_n)} - 2^{-1}\overline{\kappa}_n^{\ -\eta_1}. \tag{3.3.6}$$

The assumption $j \in \mathcal{L}^c_n$ entails $|\widehat{R}_{j,k_2}| < 2^{-1}\overline{\kappa}_n^{\ -\eta_1}$ on $\Omega_{n,j}$, and hence $|W_{k_2}| - 2^{-1}\overline{\kappa}_n^{\ -\eta_1} < |\widehat{U}_{j,k_2}|$ due to $\Delta_{k_2} J = 0$. From (3.3.6), we have got

$$|\widehat{U}_{j,k_1}| < |\widehat{U}_{j,k_2}| \tag{3.3.7}$$

on $\Omega_{n,j}$ if $j \in \mathcal{L}^c_n$ and if a pair $(k_1, k_2) \in I^2_{n,j}$ satisfies (3.3.5).

We are working on $\Omega_{n,j}$ yet. Suppose that $j \in \mathcal{L}^c_n$ and $k_1 \in \widehat{\mathcal{K}}_{n,j}(\alpha_0)$. Then the inequality (3.3.7) holds for any $k_2 \in I_{n,j}$ satisfying $r_{n,j}(|W_{k_2}|) \ge \mathsf{a}_n$. So, there are at least $\lfloor \alpha_0\overline{\kappa}_n + 1\rfloor\big(\le \alpha_0\overline{\kappa}_n + \overline{\kappa}_n^{\,1-\eta_2} + 1 \le \overline{\kappa}_n - \mathsf{a}_n + 1\big)$ variables $\widehat{U}_{j,k_2}$ that satisfy (3.3.7). Then $r_{n,j}(|\widehat{U}_{j,k_1}|) \le (1-\alpha_0)\overline{\kappa}_n$, and hence $k_1 \in \mathcal{K}_{n,j}(\alpha_0)$. Thus, we found

$$\widehat{\mathcal{K}}_{n,j}(\alpha_0) \quad \subset \quad \mathcal{K}_{n,j}(\alpha_0)$$

on $\Omega_{n,j}$ if $j \in \mathcal{L}^c_n$, that is, (a).

We still work on $\Omega_{n,j}$. Suppose that $j\in\mathcal{L}_n^c$ and $k_2\in\mathcal{K}_{n,j}(\alpha_0)\setminus\widehat{\mathcal{K}}_{n,j}(\alpha_0)$. When $r_{n,j}(|W_{k_2}|)<\mathsf{a}_n$, since $r_{n,j}(|W_{k_2}|)>\widehat{\mathsf{a}}_n$ due to $k_2\in\widehat{\mathcal{K}}_{n,j}(\alpha_0)^c$, we see

$$1_{\{j\in\mathcal{L}_n^c\}}\#\big\{k_2\in\mathcal{K}_{n,j}(\alpha_0)\setminus\widehat{\mathcal{K}}_{n,j}(\alpha_0);\ r_{n,j}(|W_{k_2}|)<\mathsf{a}_n\big\}\quad\le\quad\overline{\kappa}_n^{\,1-\eta_2} \tag{3.3.8}$$

on $\Omega_{n,j}$. When $r_{n,j}(|W_{k_2}|)\ge\mathsf{a}_n$, for any k_1 satisfying $r_{n,j}(|W_{k_1}|)\le\widehat{\mathsf{a}}_n$, we have (3.3.7). Therefore

$$\#\big\{k_1\in I_{n,j};\ |\widehat{U}_{j,k_1}|<|\widehat{U}_{j,k_2}|\big\}\quad\ge\quad 1_{\{j\in\mathcal{L}_n^c\}}1_{\{r_{n,j}(|W_{k_2}|)\ge\mathsf{a}_n\}}\,\widehat{\mathsf{a}}_n,$$

in other words,

$$r_{n,j}(|\widehat{U}_{j,k_2}|)\quad>\quad\widehat{\mathsf{a}}_n \tag{3.3.9}$$

on $\Omega_{n,j}$ if $j\in\mathcal{L}_n^c$ and $r_{n,j}(|W_{k_2}|)\ge\mathsf{a}_n$. Moreover, $r_{n,j}(|\widehat{U}_{j,k_2}|)\le\lfloor(1-\alpha_0)\overline{\kappa}_n\rfloor$ since $k_2\in\mathcal{K}_{n,j}(\alpha_0)$. Combining this estimate with (3.3.9), we obtain

$$\begin{aligned}1_{\{j\in\mathcal{L}_n^c\}}\#\big\{k_2\in\mathcal{K}_{n,j}(\alpha_0)\setminus\widehat{\mathcal{K}}_{n,j}(\alpha_0);\ r_{n,j}(|W_{k_2}|)\ge\mathsf{a}_n\big\}\quad&\le\quad(1-\alpha_0)\overline{\kappa}_n-\widehat{\mathsf{a}}_n\\&\le\quad 2\,\overline{\kappa}_n^{\,1-\eta_2}+1\end{aligned} \tag{3.3.10}$$

From (3.3.8) and (3.3.10), we obtain (b). □

For $\eta_3\in\mathbb{R}$, $j\in I_n$ and a sequence of random variables $(V_j)_{j\in I_n}$, let

$$\mathcal{D}_{n,j}\quad=\quad\overline{\kappa}_n^{\eta_3}\bigg|\frac{1}{\overline{\kappa}_n}\sum_{k\in\mathcal{K}_{n,j}(\alpha_0)}V_k-\frac{1}{\overline{\kappa}_n}\sum_{k\in\widehat{\mathcal{K}}_{n,j}(\alpha_0)}V_k\bigg|$$

The following lemma follows from Lemma 3.3.3 immediately.

Lemma 3.3.4. **(i)** *Let $p\ge1$. Then*

$$\big\|\mathcal{D}_{n,j}\big\|_p\quad\le\quad 4\,\overline{\kappa}_n^{\eta_3-\eta_2}\bigg\|\max_{k\in I_{n,j}}|V_k|1_{\Omega_{n,j}\cap\{j\in\mathcal{L}_n^c\}}\bigg\|_p+\overline{\kappa}_n^{\eta_3}\bigg\|\max_{k\in I_{n,j}}|V_k|1_{\Omega_{n,j}^c}\bigg\|_p+\overline{\kappa}_n^{\eta_3}\bigg\|\max_{k\in I_{n,j}}|V_k|1_{\{j\in\mathcal{L}_n\}}\bigg\|_p$$

for $j\in I_n$, $n\in\mathbb{N}$.

(ii) *Let $p\ge1$ and $\eta_4>0$. Then*

$$\begin{aligned}\big\|\mathcal{D}_{n,j}\big\|_p\quad&\le\quad 4\,\overline{\kappa}_n^{\eta_3-\eta_2}\bigg(\kappa_n^{\eta_4}+\overline{\kappa}_n\max_{k\in I_{n,j}}\Big\||V_k|1_{\{|V_k|>\kappa_n^{\eta_4}\}}1_{\Omega_{n,j}\cap\{j\in\mathcal{L}_n^c\}}\Big\|_p\bigg)\\&\qquad+\overline{\kappa}_n^{\eta_3}\bigg\|\max_{k\in I_{n,j}}|V_k|1_{\Omega_{n,j}^c}\bigg\|_p+\overline{\kappa}_n^{\eta_3}\bigg\|\max_{k\in I_{n,j}}|V_k|1_{\{j\in\mathcal{L}_n\}}\bigg\|_p\end{aligned}$$

for $j\in I_n$, $n\in\mathbb{N}$.

Let

$$\widetilde{\mathcal{K}}_{n,j}(\alpha_0)\quad=\quad\big\{k\in I_{n,j};\ |W_k|\le c(\alpha_0)^{1/2}\big\}$$

For $\eta_3>0$, $j\in I_n$ and a sequence of random variables $(V_j)_{j\in I_n}$, let

$$\widetilde{\mathcal{D}}_{n,j}\quad=\quad\overline{\kappa}_n^{\eta_3}\bigg|\frac{1}{\overline{\kappa}_n}\sum_{k\in\widehat{\mathcal{K}}_{n,j}(\alpha_0)}V_k-\frac{1}{\overline{\kappa}_n}\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)}V_k\bigg|$$

Let

$$\widetilde{\Omega}_{n,j}=\big\{\big||W|_{(j,\widehat{\mathsf{a}}_n)}-c(\alpha_0)^{1/2}\big|<\check{C}\kappa_n^{-\eta_2}\big\} \tag{3.3.11}$$

for $j\in I_n$, where $\check{C}$ is a positive constant.

Lemma 3.3.5. *Let $\eta_3 \in \mathbb{R}$. Then*

(i) *For $p \geq 1$ and $j \in I_n$,*

$$\big\|\widetilde{\mathcal{D}}_{n,j}\big\|_p \;\leq\; \overline{\kappa}_n^{\eta_3}\bigg\| \max_{k'\in I_{n,j}} |V_{k'}| \frac{1}{\overline{\kappa}_n}\sum_{k\in I_{n,j}} 1_{\left\{\left||W_k|-c(\alpha_0)^{1/2}\right| < \check{C}\kappa_n^{-\eta_2}\right\}}\bigg\|_p + \overline{\kappa}_n^{\eta_3}\bigg\| 1_{\widetilde{\Omega}^c_{n,j}} \max_{k'\in I_{n,j}} |V_{k'}|\bigg\|_p$$

(ii) *For $p_1 > p \geq 1$ and $j \in I_n$,*

$$\begin{aligned}\big\|\widetilde{\mathcal{D}}_{n,j}\big\|_p \;\leq\; & \overline{\kappa}_n^{\eta_3}\bigg\| \max_{k\in I_{n,j}} |V_k|\bigg\|_p P\Big[\big||W_1|-c(\alpha_0)^{1/2}\big| < \check{C}\kappa_n^{-\eta_2}\Big] \\ & +\overline{\kappa}_n^{\eta_3}\bigg\| \max_{k\in I_{n,j}} |V_k|\bigg\|_{pp_1(p_1-p)^{-1}} \bigg\| \frac{1}{\overline{\kappa}_n}\sum_{k\in I_{n,j}} \bigg(1_{\left\{\left||W_k|-c(\alpha_0)^{1/2}\right| < \check{C}\kappa_n^{-\eta_2}\right\}} \\ & \qquad -P\Big[\big||W_k|-c(\alpha_0)^{1/2}\big| < \check{C}\kappa_n^{-\eta_2}\Big]\bigg)\bigg\|_{p_1} \\ & +\overline{\kappa}_n^{\eta_3} P\big[\widetilde{\Omega}^c_{n,j}\big]^{1/p_1} \bigg\| \max_{k\in I_{n,j}} |V_k|\bigg\|_{pp_1(p_1-p)^{-1}}\end{aligned}$$

Proof. For $k \in I_{n,j}$,

$$\begin{aligned}& \widetilde{\Omega}_{n,j} \cap \big\{\, r_{n,j}(|W_k|) \;\leq\; \widehat{\mathsf{a}}_n\big\}^c \cap \big\{|W_k| \leq c(\alpha_0)^{1/2}\big\} \\ =\; & \big\{\big||W|_{(j,\widehat{\mathsf{a}}_n)} - c(\alpha_0)^{1/2}\big| < \check{C}\kappa_n^{-\eta_2}\big\} \cap \big\{|W_k| > |W|_{(j,\widehat{\mathsf{a}}_n)}\big\} \cap \big\{|W_k| \leq c(\alpha_0)^{1/2}\big\} \\ \subset\; & \big\{\big||W_k| - c(\alpha_0)^{1/2}\big| < \check{C}\kappa_n^{-\eta_2}\big\}\end{aligned}$$

and

$$\begin{aligned}& \widetilde{\Omega}_{n,j} \cap \big\{\, r_{n,j}(|W_k|) \;\leq\; \widehat{\mathsf{a}}_n\big\} \cap \big\{|W_k| \leq c(\alpha_0)^{1/2}\big\}^c \\ =\; & \big\{\big||W|_{(j,\widehat{\mathsf{a}}_n)} - c(\alpha_0)^{1/2}\big| < \check{C}\kappa_n^{-\eta_2}\big\} \cap \big\{|W_k| \leq |W|_{(j,\widehat{\mathsf{a}}_n)}\big\} \cap \big\{|W_k| > c(\alpha_0)^{1/2}\big\} \\ \subset\; & \big\{\big||W_k| - c(\alpha_0)^{1/2}\big| < \check{C}\kappa_n^{-\eta_2}\big\}.\end{aligned}$$

Thus we obtain (i). Property (ii) follows from (i). □

Lemma 3.3.6. *If the constant $\check{C}$ in (3.4.7) is sufficiently large, then*

$$\sup_{j\in I_n} P\big[\widetilde{\Omega}^c_{n,j}\big] \;=\; O(n^{-L})$$

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

Proof. We have

$$\begin{aligned}& P\big[|W|_{(j,\widehat{\mathsf{a}}_n)} - c(\alpha_0)^{1/2} < -\check{C}\kappa_n^{-\eta_2}\big] \\ \leq\; & P\Big[|W|_{\left(j,\lfloor \mathsf{a}_n - \overline{\kappa}_n^{\,1-\eta_2} - 1\rfloor\right)} < c(\alpha_0)^{1/2} - \check{C}\kappa_n^{-\eta_2}\Big] \\ \leq\; & P\Big[\sum_{k\in I_{n,j}} 1_{A_{n,k}} \geq \lfloor \mathsf{a}_n - \overline{\kappa}_n^{1-\eta_2} - 1\rfloor\Big] \\ =\; & P\Big[\overline{\kappa}_n^{-1/2}\sum_{k\in I_{n,j}} \big\{1_{A_{n,k}} - P[A_{n,k}]\big\} \geq C_n\Big] \qquad (3.3.12)\end{aligned}$$

where

$$\begin{aligned} A_{n,k} &= \big\{|W_k| < c(\alpha_0)^{1/2} - \check{C}\kappa_n^{-\eta_2}\big\}, \\ C_n &= \overline{\kappa}_n^{-1/2}\big(\mathsf{a}_n - \overline{\kappa}_n^{1-\eta_2} - 2 - \overline{\kappa}_n P[A_{n,1}]\big). \end{aligned}$$

By using the mean-value theorem, we obtain

$$\begin{aligned} C_n &\sim \overline{\kappa}_n^{-1/2}\Big[(1-\alpha_0)\overline{\kappa}_n - 2\overline{\kappa}_n^{1-\eta_2} - \overline{\kappa}_n\big\{1-\alpha_0 - 2\phi\big(c(\alpha_0)^{1/2};0,1\big)\,\check{C}\kappa_n^{-\eta_2}\big\}\Big] \\ &\gtrsim \overline{\kappa}_n^{\frac{1}{2}-\eta_2} \end{aligned}$$

as $n \to \infty$ if we choose a sufficiently large $\check{C}$. Therefore, the L^p-boundedness of the random variables in (3.3.12) gives

$$\sup_{j\in I_n} P\big[|W|_{(j,\widehat{\mathsf{a}}_n)} - c(\alpha_0)^{1/2} < -\check{C}\kappa_n^{-\eta_2}\big] = O(n^{-L}) \tag{3.3.13}$$

as $n \to \infty$ for any $L > 0$. In a similar way, we know

$$P\big[|W|_{(j,\widehat{\mathsf{a}}_n)} - c(\alpha_0)^{1/2} > \check{C}\kappa_n^{-\eta_2}\big] = O(n^{-L}) \tag{3.3.14}$$

as $n \to \infty$ for any $L > 0$. Then we obtain the result from (3.3.13) and (3.3.14). □

3.3.2 Local-global realized volatility

We introduce the local-global realized volatility (LGRV)

$$\mathbb{L}_{n,j}(\alpha_0) = \frac{n}{\overline{\kappa}_n T}\sum_{k\in\mathcal{K}_{n,j}(\alpha_0)} q(\alpha_0)^{-1}|\Delta_k X|^2 K_{n,k}. \tag{3.3.15}$$

Theorem 3.3.7. *Suppose that [G1] is fulfilled. For $c_0 \in (0,1)$ and $B > 0$, suppose that $\kappa_n \sim Bn^{c_0}$ as $n \to \infty$. Then*

$$\sup_{n\in\mathbb{N}}\sup_{j\in I_n}\sup_{k\in I_{n,j}} n^{\gamma_*}\big\|1_{\{j\in\mathcal{L}_n^c\}}\big(\mathbb{L}_{n,j}(\alpha_0) - \sigma_{t_k}^2\big)\big\|_p < \infty \tag{3.3.16}$$

as $n \to \infty$ for any constant γ_* satisfying

$$\gamma_* < \min\Big\{\frac{1}{2}(1-c_0), \frac{1}{2}c_0\Big\}.$$

Proof. (I) We have $\kappa_n \sim n^{c_0} \sim h^{-c_0}$ and $n/\overline{\kappa}_n \sim n^{1-c_0} \sim h^{c_0-1}$. Let

$$\mathcal{D}^*_{n,j} = \overline{\kappa}_n^{\eta_3}\Big\{\frac{n}{\overline{\kappa}_n}\sum_{k\in\mathcal{K}_{n,j}(\alpha_0)}|\Delta_k X|^2 K_{n,k} - \frac{n}{\overline{\kappa}_n}\sum_{k\in\widehat{\mathcal{K}}_{n,j}(\alpha_0)}|\Delta_k X|^2 K_{n,k}\Big\}.$$

Applied to $V_k = n|\Delta_k X|^2 K_{n,k}1_{\{j\in\mathcal{L}_n^c\}}$, Lemma 3.3.4 (ii) gives

$$\big\|\mathcal{D}^*_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \le \Phi_{n,j}^{(3.3.18)} + \Phi_{n,j}^{(3.3.19)} \tag{3.3.17}$$

for every $p>1$, where

$$\Phi^{(3.3.18)}_{n,j} \;=\; 4\overline{\kappa}_n^{\eta_3-\eta_2}\bigg(\kappa_n^{\eta_4}+\overline{\kappa}_n\max_{k\in I_{n,j}}\Big\|n|\Delta_kX|^2 1_{\{n|\Delta_kX|^2>\kappa_n^{\eta_4}\}}1_{\{j\in\mathcal{L}_n^c\}}\Big\|_p\bigg) \tag{3.3.18}$$

and

$$\Phi^{(3.3.19)}_{n,j} \;=\; \overline{\kappa}_n^{\eta_3}\bigg\|\max_{k\in I_{n,j}} n|\Delta_kX|^2K_{n,k}1_{\Omega^c_{n,j}}\bigg\|_p. \tag{3.3.19}$$

Since there is no jump of J on $\{j\in\mathcal{L}_n^c\}$, we see

$$\sup_{j\in I_n}\sup_{k\in I_{n,j}}\big\|n|\Delta_kX|^2 1_{\{j\in\mathcal{L}_n^c\}}\big\|_p = O(1) \tag{3.3.20}$$

for every $p>1$, as a result, the L^p-norm on the right-hand side of (3.3.18) is of $O(n^{-L})$ for arbitrary $L>0$, and hence

$$\Phi^{(3.3.18)}_{n,j} \;=\; O\big(\kappa_n^{\eta_3-\eta_2+\eta_4}\big) \tag{3.3.21}$$

as $n\to\infty$. Similarly to (3.3.20), we obtain

$$\sup_{j\in I_n} P\big[\Omega^c_{n,j}\big] \;=\; O(n^{-L}) \tag{3.3.22}$$

as $n\to\infty$ for every $L>0$, from Lemma 3.3.2 as well as Lemma 3.3.1 because $(n/\kappa_n)^{1/2}\kappa_n^{-\eta_1}\gg 1$ when $2^{-1}(c_0^{-1}-1)>\eta_1$. Then

$$\Phi^{(3.3.19)}_{n,j} \;\le\; \kappa_n^{\eta_3}n^{1/2}P\big[\Omega^c_{n,j}\big]^{1/p} \;=\; O(n^{-L}) \tag{3.3.23}$$

for every $L>0$ and $p>1$. From (3.3.17), (3.3.21) and (3.3.23),

$$\big\|\mathcal{D}^*_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \;=\; O\big(\kappa_n^{\eta_3-\eta_2+\eta_4}\big) \;=\; O\big(n^{-c_0(\eta_2-\eta_3-\eta_4)}\big) \tag{3.3.24}$$

as $n\to\infty$ for every $p>1$. We recall that the parameters should satisfy

$$0<\eta_2<\eta_1<\min\left\{\frac{1}{2},\frac{1}{2}\left(\frac{1}{c_0}-1\right)\right\},\quad \eta_3+\eta_4<\eta_2.$$

[In particular, if $c_0=1/2$, then $0<\eta_2<\eta_1<1/2$. The positive parameters η_3 and η_4 can be sufficiently small at this stage. Remark that $c_0\eta_2<1/4$ when $c_0\le 1/2$.]

(II) Let

$$\widetilde{\mathcal{D}}^*_{n,j} \;=\; \overline{\kappa}_n^{\eta_3}\bigg\{\frac{n}{\overline{\kappa}_n}\sum_{k\in\widehat{\mathcal{K}}_{n,j}(\alpha_0)}|\Delta_kX|^2K_{n,k}-\frac{n}{\overline{\kappa}_n}\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)}|\Delta_kX|^2K_{n,k}\bigg\}.$$

Applying Lemma 3.3.5 (ii) to $V_k=n|\Delta_kX|^2K_{n,k}1_{\{j\in\mathcal{L}_n^c\}}$, we have

$$\big\|\widetilde{\mathcal{D}}^*_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \;\le\; \Phi^{(3.3.26)}_{n,j}+\Phi^{(3.3.27)}_{n,j}+\Phi^{(3.3.28)}_{n,j}, \tag{3.3.25}$$

where

$$\Phi^{(3.3.26)}_{n,j} \;=\; \overline{\kappa}_n^{\eta_3}\bigg\|\max_{k\in I_{n,j}}n|\Delta_kX|^2K_{n,k}1_{\{j\in\mathcal{L}_n^c\}}\bigg\|_p P\Big[\big||W_1|-c(\alpha_0)^{1/2}\big|<\check{C}\kappa_n^{-\eta_2}\Big], \tag{3.3.26}$$

$$\Phi^{(3.3.27)}_{n,j} = \overline{\kappa}_n^{\eta_3}\Big\|\max_{k\in I_{n,j}} n|\Delta_k X|^2K_{n,k}1_{\{j\in\mathcal{L}_n^c\}}\Big\|_{pp_1(p_1-p)^{-1}} \times\Big\|\frac{1}{\overline{\kappa}_n}\sum_{k\in I_{n,j}}\bigg(1_{\{||W_k|-c(\alpha_0)^{1/2}|<\check{C}\kappa_n^{-\eta_2}\}}-P\Big[\big||W_k|-c(\alpha_0)^{1/2}\big|<\check{C}\kappa_n^{-\eta_2}\Big]\bigg)\Big\|_{p_1} \tag{3.3.27}$$

and

$$\Phi^{(3.3.28)}_{n,j} = \overline{\kappa}_n^{\eta_3}P\big[\widetilde{\Omega}^c_{n,j}\big]^{1/p_1}\Big\|\max_{k\in I_{n,j}} n|\Delta_k X|^2K_{n,k}1_{\{j\in\mathcal{L}_n^c\}}\Big\|_{pp_1(p_1-p)^{-1}} \tag{3.3.28}$$

for $j\in I_n$, $n\in\mathbb{N}$. Then, paying $\kappa_n^{\eta_4}$ for the maximum, we have the following estimates for any $p_1>p\geq 1$:

$$\sup_{j\in I_n}\Phi^{(3.3.26)}_{n,j} = O\big(\kappa_n^{\eta_3+\eta_4-\eta_2}\big)=O\big(n^{-c_0(\eta_2-\eta_3-\eta_4)}\big), \tag{3.3.29}$$

$$\sup_{j\in I_n}\Phi^{(3.3.27)}_{n,j} = O\big(\kappa_n^{\eta_3}\times\kappa_n^{\eta_4}\times\kappa_n^{-(1+\eta_2)/2}\big) = O\big(n^{-c_0\left(\frac{1+\eta_2}{2}-\eta_3-\eta_4\right)}\big), \tag{3.3.30}$$

and

$$\sup_{j\in I_n}\Phi^{(3.3.28)}_{n,j} = O(n^{-L}) \tag{3.3.31}$$

as $n\to\infty$ for any $L>0$ for a sufficiently large $\check{C}$; the estimate (3.3.31) follows from Lemma 3.3.6. In this way,

$$\big\|\widetilde{\mathcal{D}}^*_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p = O\big(n^{-c_0(\eta_2-\eta_3-\eta_4)}\big)+O\big(n^{-c_0\left(\frac{1+\eta_2}{2}-\eta_3-\eta_4\right)}\big) \tag{3.3.32}$$

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

(III) On the event $\{j\in\mathcal{L}_n^c\}$, we have

$$\begin{aligned}\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)}|\Delta_k X|^2K_{n,k} &= \sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)}\bigg(\int_{t_{k-1}}^{t_k}\sigma_t dw_t+\int_{t_{k-1}}^{t_k}b_t dt\bigg)^2K_{n,k}\\ &= \Phi^{(3.3.34)}_{n,j}+\Phi^{(3.3.35)}_{n,j}+\Phi^{(3.3.36)}_{n,j}+\Phi^{(3.3.37)}_{n,j}+\Phi^{(3.3.38)}_{n,j}\end{aligned} \tag{3.3.33}$$

where

$$\Phi^{(3.3.34)}_{n,j} = \sum_{k\in I_{n,j}}(\sigma_{t_{\underline{j}_n}})^2hW_k^21_{\{|W_k|\leq c(\alpha_0)^{1/2}\}}, \tag{3.3.34}$$

$$
\begin{aligned}
\Phi^{(3.3.35)}_{n,j} &= \sum_{k\in I_{n,j}} \big(\sigma_{t_{\underline{j}_n}}\big)^2 h W_k^2 1_{\{|W_k|\le c(\alpha_0)^{1/2}\}}\big(K_{n,k}-1\big) \\
&\quad +2\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t} \big(\widetilde{\sigma}_s-\widetilde{\sigma}_{t_{\underline{j}_n}}\big)dw_s\sigma_t dw_t K_{n,k} \\
&\quad +2\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t} \sigma_{t_{\underline{j}_n}} dw_s\big(\widetilde{\sigma}_t-\widetilde{\sigma}_{t_{\underline{j}_n}}\big)dw_t K_{n,k} \\
&\quad +2\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \int_{t_{k-1}}^{t_k} \widetilde{\sigma}_{t_{\underline{j}_n}}\big(\widetilde{\sigma}_t-\widetilde{\sigma}_{t_{\underline{j}_n}}\big)dt K_{n,k} \\
&\quad +\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \bigg(\int_{t_{k-1}}^{t_k} \big(\widetilde{\sigma}_t-\widetilde{\sigma}_{t_{\underline{j}_n}}\big)dw_t\bigg)^2 K_{n,k},
\end{aligned}
\tag{3.3.35}
$$

$$
\Phi^{(3.3.36)}_{n,j} = 2\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t} b_s ds\sigma_t dw_t K_{n,k}, \tag{3.3.36}
$$

$$
\Phi^{(3.3.37)}_{n,j} = 2\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t} \sigma_s dw_s b_t dt K_{n,k}, \tag{3.3.37}
$$

and

$$
\Phi^{(3.3.38)}_{n,j} = 2\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)} \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t} b_s ds b_t dt K_{n,k}. \tag{3.3.38}
$$

By assumption,

$$
\begin{aligned}
& \sup_{j\in I_n}\sup_{s\in[t_{\underline{j}_n-1},t_{\underline{j}_n+\overline{\kappa}_n}]} \big\|1_{\{j\in\mathcal{L}_n^c\}}\big(\sigma_s^2-\sigma_{t_{\underline{j}_n-1}}^2\big)\big\|_p \\
\le\ & \sup_{j\in I_n}\sup_{s\in[t_{\underline{j}_n-1},t_{\underline{j}_n+\overline{\kappa}_n}]} \big\|\widetilde{\sigma}_s^2-\widetilde{\sigma}_{t_{\underline{j}_n-1}}^2\big\|_p \\
\lesssim\ & (\kappa_n h)^{1/2} \lesssim h^{\frac{1}{2}(1-c_0)}
\end{aligned}
\tag{3.3.39}
$$

for every $p>1$. First, a primitive estimate gives

$$
\sup_{j\in I_n}\frac{n}{\kappa_n}\big\|\Phi^{(3.3.35)}_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \lesssim \frac{n}{\kappa_n}\times\frac{\kappa_n^{3/2}}{n^{3/2}} \lesssim h^{\frac{1}{2}(1-c_0)} \tag{3.3.40}
$$

as $n\to\infty$; we note that the orthogonality cannot apply due to $\widetilde{\mathcal{K}}_{n,j}(\alpha_0)$ even after $K_{n,k}$ is decoupled. We also have

$$
\sup_{j\in I_n}\frac{n}{\kappa_n}\big\|\Phi^{(3.3.36)}_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \lesssim h^{1/2}. \tag{3.3.41}
$$

For $\Phi^{(3.3.37)}_{n,j}$ and $\Phi^{(3.3.38)}_{n,j}$, by the same way, we can get

$$
\sup_{j\in I_n}\frac{n}{\kappa_n}\big\|\Phi^{(3.3.37)}_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \lesssim h^{1/2}, \tag{3.3.42}
$$

and

$$\sup_{j\in I_n}\frac{n}{\kappa_n}\big\|\Phi^{(3.3.38)}_{n,j}1_{\{j\in\mathcal{L}_n^c\}}\big\|_p \quad\lesssim\quad h \tag{3.3.43}$$

as $n\to\infty$. Furthermore, we have

$$\begin{aligned}
&\sup_{j\in I_n}\left\|\left\{\frac{1}{\overline{\kappa}_n h}\Phi^{(3.3.34)}_{n,j}-\big(\sigma_{t_{\underline{j}_n}}\big)^2 q(\alpha_0)\right\}1_{\{j\in\mathcal{L}_n^c\}}\right\|_p\\
\le\;& \sup_{j\in I_n}\left\|\frac{1}{\overline{\kappa}_n h}\left\{\Phi^{(3.3.34)}_{n,j}-\sum_{k\in I_{n,j}}\big(\sigma_{t_{\underline{j}_n}}\big)^2 q(\alpha_0)h\right\}1_{\{j\in\mathcal{L}_n^c\}}\right\|_p\\
\le\;& \sup_{j\in I_n}\left\|\frac{1}{\overline{\kappa}_n}\sum_{k\in I_{n,j}}\big(\sigma_{t_{\underline{j}_n}}\big)^2\big(W_k^2 1_{\{|W_k|\le c(\alpha_0)^{1/2}\}}-q(\alpha_0)\big)\right\|_p\\
=\;& O(\kappa_n^{-1/2})\;=\;O(h^{c_0/2})
\end{aligned}\tag{3.3.44}$$

for every $p>1$. Combining (3.3.33) and (3.3.39)-(3.3.44), we obtain

$$\begin{aligned}
&\sup_{j\in I_n}\sup_{k'\in I_{n,j}}\left\|1_{\{j\in\mathcal{L}_n^c\}}\left(\frac{n}{\overline{\kappa}_n T}\sum_{k\in\widetilde{\mathcal{K}}_{n,j}(\alpha_0)}|\Delta_k X|^2 K_{n,k}-\sigma^2_{t_{k'}}q(\alpha_0)\right)\right\|_p\\
=\;& O\big(n^{-(1-c_0)/2}\big)+O(n^{-c_0/2})
\end{aligned}\tag{3.3.45}$$

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

(IV) From (3.3.24), (3.3.32) and (3.3.45), we obtain the estimate

$$\begin{aligned}
&\sup_{j\in I_n}\sup_{k'\in I_{n,j}}\kappa_n^{\eta_3}\left\|1_{\{j\in\mathcal{L}_n^c\}}\left(\frac{n}{\overline{\kappa}_n T}\sum_{k\in\mathcal{K}_{n,j}(\alpha_0)}|\Delta_k X|^2 K_{n,k}-\sigma^2_{t_{k'}}q(\alpha_0)\right)\right\|_p\\
=\;& O\big(n^{-c_0(\eta_2-\eta_3-\eta_4)}\big)+\left\{O\big(n^{-c_0(\eta_2-\eta_3-\eta_4)}\big)+O\big(n^{-c_0\left(\frac{1+\eta_2}{2}-\eta_3-\eta_4\right)}\big)\right\}\\
&+\kappa_n^{\eta_3}\left\{O\big(n^{-(1-c_0)/2}\big)+O(n^{-c_0/2})\right\}\\
=\;& O\big(n^{-c_0(\eta_2-\eta_3-\eta_4)}\big)+O\big(n^{c_0(\eta_3+\eta_4)-(1-c_0)/2}\big)\;=:\;\mathbb{O}_n
\end{aligned}\tag{3.3.46}$$

as $n\to\infty$ for every $p>1$. Here we are assuming the parameters satisfy

$$\begin{aligned}
&c_0\in(0,1),\quad B>0,\quad \eta_1\in\left(0,\min\left\{\tfrac{1}{2}\big(\tfrac{1}{c_0}-1\big),\tfrac{1}{2}\right\}\right),\\
&\eta_2\in(0,\eta_1),\quad \eta_3>0,\quad \eta_4>0,\quad \eta_3+\eta_4<\eta_2.
\end{aligned}\tag{3.3.47}$$

To obtain the last error bound in (3.3.46), we used the inequalities

$$-c_0\left(\frac{1+\eta_2}{2}-\eta_3-\eta_4\right)<-c_0\big(\eta_2-\eta_3-\eta_4\big)$$

and

$$c_0\eta_3-\frac{c_0}{2}<c_0\eta_3-c_0\eta_2<-c_0(\eta_2-\eta_3-\eta_4).$$

The LGRV $\mathbb{L}_{n,j}(\alpha_0)$ of (3.3.15) does not depend on η_i $(i = 1, 2, 3, 4)$ within the ranges (3.3.47). When $c_0 > 1/2$, we make

$$\frac{1}{2} > \frac{1}{2}\left(\frac{1}{c_0} - 1\right) > \eta_1 > \eta_2 > \eta_3 \uparrow \frac{1}{2}\left(\frac{1}{c_0} - 1\right), \quad \eta_4 \downarrow 0$$

to obtain $\mathbb{O}_n = O(1)$. When $c_0 \leq 1/2$, we make

$$\frac{1}{2} > \eta_1 > \eta_2 > \eta_3 \uparrow \frac{1}{2}, \quad \eta_4 \downarrow 0$$

to obtain $\mathbb{O}_n = O(1)$. Thus, the proof of Theorem 3.3.7 is concluded. □

According to the error bound (3.3.16), we should in general take $c_0 = 1/2$, i.e., $\kappa_n \sim Bn^{1/2}$ to obtain an optimal error estimate. However, this is not always true. If the process σ is (unknown) constant for example, then we do not need any spot volatility estimator to construct a global jump filter, and the convergence of the resulting estimator for Θ becomes much faster than that in the non-constant σ case.

3.3.3 Local minimum RV

Estimation of spot volatilities can be done by the minimum realized volatility method of Andersen et al. [1]. It is defined as follows.

$$\mathbb{M}_{n,j} = \frac{\pi}{\pi - 2}\frac{n}{\bar{\kappa}_n T}\sum_{k \in I_{n,j}} \{|\Delta_k X| \wedge |\Delta_{k+1} X|\}^2.$$

Theorem 3.3.8. *Suppose that [G1] is fulfilled. For $c_0 \in (0, 1)$ and $B > 0$, suppose that $\kappa_n \sim Bn^{c_0}$ as $n \to \infty$. Then*

$$\sup_{n \in \mathbb{N}} \sup_{j \in I_n} \sup_{k \in I_{n,j}} n^{\gamma_*} \left\| 1_{\{j \in \mathcal{L}_n^c\}} \left(\mathbb{M}_{n,j} - \sigma_{t_k}^2 \right) \right\|_p \quad < \quad \infty$$

as $n \to \infty$ for any $p > 1$ and any constant γ_* satisfying

$$\gamma_* \quad < \quad \min\left\{ \frac{1}{2}(1 - c_0), \frac{1}{2}c_0 \right\}.$$

The proof is essentially the same as that of Andersen et al. [1].

3.4 Rate of convergence of the global realized volatilities in high intensity of jumps

When the frequency of the jumps is high, it is recommend that one should choose a value of α that is not extremely small in order to cover the jumps by the index set $\mathcal{J}_n(\alpha)^c$.

[G2] **(i)** $S_{n,j-1}$ is positive a.s. and

$$\sup_{n \in \mathbb{N}} \sup_{j \in I_n} \left\| S_{n,j-1}^{-1} \right\|_p \quad < \quad \infty$$

for every $p > 1$.

(ii) There exist positive constants γ_0 and c such that

$$\sup_{n\in\mathbb{N}}\sup_{j\in I_n} n^{\gamma_0}\left\|1_{\{j\in\mathcal{L}_n^c\}}\left(\sigma_{t_{j-1}}^2 - c\,S_{n,j-1}\right)\right\|_p \;<\; \infty$$

for every $p > 1$.

In $[G2]$, we do not assume that the value of constant c is known. We note that

$$\sup_{n\in\mathbb{N}}\sup_{j\in I_n}\left\|1_{\{j\in\mathcal{L}_n^c\}}S_{n,j-1}\right\|_p \;<\; \infty$$

for every $p > 1$ under $[G1]$ and $[G2]$. As shown in Theorem 3.3.7, the LGRV in (3.3.15) can serve as $S_{n,j-1}$.

If σ_t is equal to a (possibly unknown) constant, then γ_0 can be arbitrarily large since we can let $S_{n,j-1} = 1$. In other words, we do not need any pre-estimate of $\sigma_{t_{j-1}}^2$. So, the constant volatility case is very special and it will be discussed briefly in Section 3.6 separately. This section logically includes the constant volatility case (hence a less efficient way for it) but we will consider a general non-constant volatility and assume a given local estimator attains a limited rate of convergence.

Remark 3.4.1. When $\mathsf{v} = 2^{-1}\inf_{\omega\in\Omega,t\in[0,T]}\sigma_t^2 > 0$ for a priori known constant v, given a local estimator $\mathbb{L}_{n,j-1}^{loc}$ of $\sigma_{t_{j-1}}^2$, we can use $S_{n,j-1}(\mathsf{v}) = \mathbb{L}_{n,j}^{loc}\vee\mathsf{v}$ for $S_{n,j-1}$. For example, it is the case when X satisfies a stochastic differential equation with jumps and its diffusion coefficient is uniformly elliptic. When $\mathsf{v} = 0$, an appropriate modification of $\mathbb{L}_{n,j}^{loc}$ is necessary and possible. We only give an idea without going into details here. Preset a positive constant v. Using $S_{n,j-1}(\mathsf{v})$ for $S_{n,j-1}$, we obtain an estimator $\widetilde{\mathbb{V}}_n[\mathsf{v}]$ of $\Theta(\mathsf{v}) = \int_0^T\sigma_t^2 1_{\{\sigma_t^2\geq\mathsf{v}\}}dt$, and indeed, the rate of convergence $\widetilde{\mathbb{V}}_n[\mathsf{v}]$ is established in this paper. Then it is natural to use $\widetilde{\mathbb{V}}_n[\mathsf{v}_n]$ to estimate $\Theta = \int_0^T\sigma_t^2dt$ with a sequence of numbers v_n tending to 0 as $n\to\infty$. Consistency does not matter because the mappting $\mathsf{v}\mapsto\Theta(\mathsf{v})$ is continuous and the operation $\mathsf{v}_n\downarrow 0$ is stable. Some work is necessary to give an explicit rate of convergence since the constant of the error bound for each v_n depends on v_n. However, the cause of the error by the truncation at level v_n is the difference $\int_0^T\sigma_t^2 1_{\{\sigma_t^2<\mathsf{v}_n\}}dt$, and it is rather easy to control for small v_n.

3.4.1 Rate of convergence of the GRV with a fixed α

We consider the GRV given by (3.2.2):

$$\mathbb{V}_n(\alpha) \;=\; \sum_{j\in\mathcal{J}_n(\alpha)} q(\alpha)^{-1}|\Delta_j X|^2 K_{n,j}.$$

Denote by $r_n(\mathtt{U}_j)$ the rank of $\mathtt{U}_j$ among the variables $\{\mathtt{U}_i\}_{i\in I_n}$ as before, and $|\mathtt{U}|_{(r)}$ denotes the r-th ordered statistic of $\{|\mathtt{U}_i|\}_{i\in I_n}$. Let $0 < \gamma_2 < \gamma_1 < \gamma_0$, and define numbers a_n and $\widehat{a}_n$ by

$$a_n \;=\; \lfloor(1-\alpha)n - n^{1-\gamma_2}\rfloor \quad\text{and}\quad \widehat{a}_n \;=\; \lfloor a_n - n^{1-\gamma_2}\rfloor,$$

respectively. Define the event $N_{n,j}$ by

$$N_{n,j} \;=\; \left\{r_n(|W_j|)\leq a_n - n^{1-\gamma_2}\right\}\cap\left\{|W|_{(a_n)} - |W_j| < n^{-\gamma_1}\right\}$$

The following lemma is Lemma 2.6 of Inatsugu and Yoshida [8].

Lemma 3.4.2.

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

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

We need some notation:

$$\begin{aligned}
\widehat{\mathcal{J}}_n(\alpha) &= \big\{j \in I_n;\ r_n(|W_j|) \le \widehat{a}_n\big\},\\
U_j &= c^{-1/2}h^{-1/2}(S_{n,j-1})^{-1/2}\Delta_j X\\
R_j &= U_j - W_j - c^{-1/2}h^{-1/2}(S_{n,j-1})^{-1/2}\Delta_j J,
\end{aligned}$$

as well

$$\Omega_n = \Big\{\overline{N}_T < n^{1-\gamma_2}\Big\}\bigcap\bigg(\bigcap_{j=1,...,n}\Big[\big\{|R_j|1_{\{\Delta_j N^\sigma=0\}} < 2^{-1}n^{-\gamma_1}\big\}\cap(N_{n,j})^c\Big]\bigg).$$

We assume that the distribution of the variable $\overline{N}_T$ depends on n. In particular, we consider the case where $\overline{N}_T$ may diverge as $n \to \infty$.

[G3] There exists a constant $\xi \ge 0$ such that $\|\overline{N}_T\|_p = O(n^\xi)$ as $n \to \infty$ for every $p > 1$.

Lemma 3.4.3. *Suppose that [G1] and [G2] are satisfied. Suppose that $0 < \gamma_1 < \gamma_0 < 1/2$. Then*

$$\sup_{j\in I_n} P\big[|R_j|1_{\{\Delta_j N^\sigma=0\}} \ge 2^{-1}n^{-\gamma_1}\big] = O(n^{-L}) \tag{3.4.1}$$

as $n \to \infty$ for every $L > 0$. In particular, if $\gamma_2 < 1 - \xi$ and [G3] is additionally satisfied, then

$$P[\Omega_n^c] = O(n^{-L}) \tag{3.4.2}$$

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

Proof. We have

$$\sup_{j\in I_n}\big\|R_j 1_{\{\Delta_j N^\sigma=0\}}\big\|_p = O(n^{-\gamma_0})$$

for every $p > 1$. The Markov inequality implies (3.4.1). This estimate and Lemma 3.4.2 give (3.4.2). □

Let

$$\mathfrak{L}_n = \big\{j \in I_n;\ \Delta_j\overline{N} \ne 0\big\}. \tag{3.4.3}$$

Lemma 2.7 of Inatsugu and Yoshida [8] is rephrased as follows. We note that the definition of $\mathcal{L}_n^{(k)}$ therein is essentially the same as $\mathfrak{L}_n$, and different from $\mathcal{L}_n$ defined by (3.3.4).

Lemma 3.4.4.

$$\widehat{\mathcal{J}}_n(\alpha)\cap\mathfrak{L}_n^c \subset \mathcal{J}_n(\alpha) \tag{3.4.4}$$

on Ω_n. In particular

$$\#\big[\mathcal{J}_n(\alpha)\ominus\widehat{\mathcal{J}}_n(\alpha)\big] \le c_* n^{1-\gamma_2} + \overline{N}_T \tag{3.4.5}$$

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

For $\gamma_3 > 0$ and random variables $(\mathtt{U}_j)_{j=1,...,n}$, let

$$\mathcal{D}_n = n^{\gamma_3}\left|\frac{1}{n}\sum_{j\in\mathcal{J}_n(\alpha)}\mathtt{U}_j - \frac{1}{n}\sum_{j\in\widehat{\mathcal{J}}_n(\alpha)}\mathtt{U}_j\right|.$$

We refer the reader to Inatsugu and Yoshida [8] (Lemmas 2.8 and 2.9) for proof of the following two lemmas.

Lemma 3.4.5. (i) *Let $p_1 > 1$. Then*

$$\|\mathcal{D}_n\|_p \leq \left(c_* n^{\gamma_3-\gamma_2} + n^{-1+\gamma_3}\|\overline{N}_T\|_{p_1}\right)\left\|\max_{j=1,...,n}\left|\mathtt{U}_j\right|\right\|_{pp_1(p_1-p)^{-1}} + n^{\gamma_3}\left\|\max_{j=1,...,n}\left|\mathtt{U}_j\right|1_{\Omega_n^c}\right\|_p$$

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

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

$$\|\mathcal{D}_n\|_p \leq \left(c_* n^{\gamma_3-\gamma_2} + n^{-1+\gamma_3}\|\overline{N}_T\|_{p_1}\right) \times\left(n^{\gamma_4} + n\max_{j=1,...,n}\left\|\left|\mathtt{U}_j\right|1_{\{|\mathtt{U}_j|>n^{\gamma_4}\}}\right\|_{pp_1(p_1-p)^{-1}}\right) + n^{\gamma_3}\left\|\max_{j=1,...,n}\left|\mathtt{U}_j\right|1_{\Omega_n^c}\right\|_p$$

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

Let

$$\widetilde{\mathcal{D}}_n = n^{\gamma_3}\left|\frac{1}{n}\sum_{j\in\widehat{\mathcal{J}}_n(\alpha)}\mathtt{U}_j - \frac{1}{n}\sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)}\mathtt{U}_j\right|.$$

for a collection of random variables $\{\mathtt{U}_j\}_{j\in I_n}$ and

$$\widetilde{\mathcal{J}}_n(\alpha) = \left\{j \in I_n;\ |W_j| \leq c(\alpha)^{1/2}\right\}. \tag{3.4.6}$$

Let

$$\widetilde{\Omega}_n = \left\{\left||W|_{(\widehat{a}_n)} - c(\alpha)^{1/2}\right| < \check{C}\, n^{-\gamma_2}\right\}, \tag{3.4.7}$$

where $\check{C}$ is a positive constant. See Lemma 4 of Inatsugu and Yoshida [8] for a proof of the following lemma.

Lemma 3.4.6. *Let $\check{C} > 0$ and $\gamma_3 > 0$. Then*

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

$$\|\widetilde{\mathcal{D}}_n\|_p \leq n^{\gamma_3}\left\|\max_{j'=1,...,n}|\mathtt{U}_{j'}|\,\frac{1}{n}\sum_{j=1}^n 1_{\left\{\left||W_j|-c(\alpha)^{1/2}\right|<\check{C}\,n^{-\gamma_2}\right\}}\right\|_p + n^{\gamma_3}\left\|1_{\widetilde{\Omega}_n^c}\max_{j'=1,...,n}|\mathtt{U}_{j'}|\right\|_p.$$

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

$$\begin{aligned}\|\tilde{\mathcal{D}}_n^{(k)}\|_p &\leq n^{\gamma_3}\Big\|\max_{j=1,\dots,n}|\mathtt{U}_j|\Big\|_p P\Big[\big||W_1|-c(\alpha)^{1/2}\big|<\check{C}\,n^{-\gamma_2}\Big]\\ &\quad +n^{\gamma_3}\Big\|\max_{j=1,\dots,n}|\mathtt{U}_j|\Big\|_{pp_1(p_1-p)^{-1}}\\ &\qquad\times\Big\|\frac{1}{n}\sum_{j=1}^{n}\Big(1_{\{||W_j|-c(\alpha)^{1/2}|<\check{C}\,n^{-\gamma_2}\}}-P\Big[\big||W_1|-c(\alpha)^{1/2}\big|<\check{C}\,n^{-\gamma_2}\Big]\Big)\Big\|_{p_1}\\ &\quad +n^{\gamma_3}P[\widetilde{\Omega}_n^c]^{1/p_1}\Big\|\max_{j=1,\dots,n}|\mathtt{U}_j|\Big\|_{pp_1(p_1-p)^{-1}}.\end{aligned}$$

Lemma 3.4.7. *If the constant $\check{C}$ in (3.3.11) is sufficiently large, then*

$$P\big[\widetilde{\Omega}_n^c\big] \;=\; O(n^{-L})$$

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

Now we shall investigate the rate of convergence of $\mathbb{V}_n(\alpha)$ for a constant $\alpha\in(0,1)$. We note that, under $[G1]$ and $[G3]$,

$$\Big\|\sum_{j\in\mathfrak{L}_n}|\Delta_j X|^2K_{n,j}\Big\|_p \;\leq\; n^{-1/2}\big\|\overline{N}_T\big\|_p \;=\; O(n^{-1/2+\xi}). \tag{3.4.8}$$

Let

$$\widehat{\mathbb{V}}_n(\alpha) \;=\; \sum_{j\in\widehat{\mathcal{J}}_n(\alpha)}q(\alpha)^{-1}|\Delta_j X|^2K_{n,j}.$$

Lemma 3.4.8. *Suppose that $[G1]$ $[G2]$ and $[G3]$ are fulfilled. Suppose that $\xi<\frac{1}{2}$. Let $\gamma_5<\min\big\{\gamma_0,\frac{1}{2}-\xi\big\}$. Then*

$$\sup_{n\in\mathbb{N}}n^{\gamma_5}\big\|\mathbb{V}_n(\alpha)-\widehat{\mathbb{V}}_n(\alpha)\big\|_p \;<\; \infty.$$

Proof. By (3.4.8), we obtain

$$\begin{aligned}\big\|\mathbb{V}_n(\alpha)-\widehat{\mathbb{V}}_n(\alpha)\big\|_p &= \Big\|\sum_{j\in\mathcal{J}_n(\alpha)}q(\alpha)^{-1}|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}\\ &\qquad -\sum_{j\in\widehat{\mathcal{J}}_n(\alpha)}q(\alpha)^{-1}|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}\Big\|_p+O(n^{-1/2+\xi}).\end{aligned}$$

By Lemmas 3.4.5 and 3.4.3,

$$\begin{aligned}&n^{\gamma_3}\big\|\mathbb{V}_n(\alpha)-\widehat{\mathbb{V}}_n(\alpha)\big\|_p\\ \lesssim\;&\big(c_*n^{\gamma_3-\gamma_2}+n^{-1+\gamma_3}\|\overline{N}_T\|_{p_1}\big)\\ &\times\Big(n^{\gamma_4}+n\max_{j=1,\dots,n}\Big\|n|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}1_{\{n|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}>n^{\gamma_4}\}}\Big\|_{pp_1(p_1-p)^{-1}}\Big)\\ &+n^{\gamma_3}\Big\|\max_{j=1,\dots,n}\big(n|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}\big)1_{\Omega_n^c}\Big\|_p+O(n^{-1/2+\gamma_3+\xi})\\ \lesssim\;&c_*n^{\gamma_3+\gamma_4-\gamma_2}+n^{-1+\gamma_3+\gamma_4+\xi}+n^{-1/2+\gamma_3+\xi},\end{aligned}$$

where $1 < p < p_1$. The parameters should satisfy

$$0 < \gamma_3 < \gamma_2 < \gamma_1 < \gamma_0 < \frac{1}{2}, \quad \gamma_2 < \frac{1}{2} - \xi, \quad \gamma_4 > 0.$$

We make

$$\gamma_4 \downarrow 0, \qquad \gamma_5 < \gamma_3 <\uparrow \gamma_2 <\uparrow \gamma_1 <\uparrow \min\left\{\gamma_0, \frac{1}{2} - \xi\right\}$$

to obtain the desired exponent. □

For $\widetilde{\mathcal{J}}_n(\alpha)$ defined in (3.4.6), let

$$\widetilde{\mathbb{V}}_n(\alpha) \;=\; \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} q(\alpha)^{-1}|\Delta_j X|^2 K_{n,j}.$$

Lemma 3.4.9. *Suppose that [G1] and [G3] are fulfilled. Suppose that $\xi < \frac{1}{2}$. Let $\gamma_6 < \frac{1}{2} - \xi$. Then*

$$\sup_{n\in\mathbb{N}} n^{\gamma_6}\big\|\widehat{\mathbb{V}}_n(\alpha) - \widetilde{\mathbb{V}}_n(\alpha)\big\|_p \;<\; \infty.$$

Proof. By (3.4.8), we obtain

$$\begin{aligned}\big\|\widehat{\mathbb{V}}_n(\alpha) - \widetilde{\mathbb{V}}_n(\alpha)\big\|_p \;=\; & \Bigg\| \sum_{j\in\widehat{\mathcal{J}}_n(\alpha)} q(\alpha)^{-1}|\Delta_j X|^2 1_{\{\Delta_j \overline{N}=0\}} K_{n,j} \\ & \qquad - \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} q(\alpha)^{-1}|\Delta_j X|^2 1_{\{\Delta_j \overline{N}=0\}} K_{n,j}\Bigg\|_p + O(n^{-1/2+\xi}).\end{aligned}$$

By Lemma 3.4.6, we obtain

$$\begin{aligned} & n^{\gamma_3}\big\|\widehat{\mathbb{V}}_n(\alpha) - \widetilde{\mathbb{V}}_n(\alpha)\big\|_p \\ \lesssim\; & n^{\gamma_3}\Big\|\max_{j=1,\ldots,n}\big(n|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}\big)\Big\|_p P\Big[\big||W_1| - c(\alpha)^{1/2}\big| < \check{C}\, n^{-\gamma_2}\Big] \\ & + n^{\gamma_3}\Big\|\max_{j=1,\ldots,n}\big(n|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}\big)\Big\|_{pp_1(p_1-p)^{-1}} \\ & \quad \times \Bigg\|\frac{1}{n}\sum_{j=1}^{n}\Bigg(1_{\left\{\left||W_j| - c(\alpha)^{1/2}\right| < \check{C}\, n^{-\gamma_2}\right\}} - P\Big[\big||W_1| - c(\alpha)^{1/2}\big| < \check{C}\, n^{-\gamma_2}\Big]\Bigg)\Bigg\|_{p_1} \\ & + n^{\gamma_3} P[\widetilde{\Omega}_n^c]^{1/p_1}\Big\|\max_{j=1,\ldots,n}\big(n|\Delta_j X|^2 1_{\{\Delta_j\overline{N}=0\}}K_{n,j}\big)\Big\|_{pp_1(p_1-p)^{-1}} + O(n^{\gamma_3-1/2+\xi}) \\ \lesssim\; & n^{\gamma_3-\gamma_2} + n^{\gamma_3-\frac{1}{2}-\frac{\gamma_2}{2}} + n^{\gamma_3-1/2+\xi} \;\lesssim\; n^{\gamma_3+\gamma_4-\gamma_2} + n^{\gamma_3+\gamma_4-1/2+\xi},\end{aligned}$$

where $1 \le p < p_1$ and γ_4 is an arbitrary positive number. Lemma 3.4.7 was used in the above derivation. Making

$$\gamma_4 \downarrow 0 \quad \text{and} \quad \gamma_6 < \gamma_3 <\uparrow \gamma_2 <\uparrow \frac{1}{2} - \xi,$$

we conclude the proof. □

Lemma 3.4.10. *Suppose that [G1] and [G3] are satisfied. Suppose that $\xi < 1/2$. Then*

$$\left\|\widetilde{\mathbb{V}}_n(\alpha) - \int_0^T \sigma_t^2 dt\right\|_p \quad = \quad O\big(n^{-\frac{1}{2}+\xi}\big)$$

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

Proof. Recall that $\mathfrak{L}_n$ is defined by (3.4.3). We have

$$\begin{aligned}\sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} |\Delta_j X|^2 K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} &= \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \left(\int_{t_{j-1}}^{t_j} \sigma_t dw_t + \int_{t_{j-1}}^{t_j} b_t dt\right)^2 K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \\ &= \Phi_n^{(3.4.10)} + \Phi_n^{(3.4.11)} + \Phi_n^{(3.4.12)} \qquad (3.4.9)\end{aligned}$$

where

$$\Phi_n^{(3.4.10)} \quad = \quad \sum_{j\in I_n} \sigma_{t_{j-1}}^2 h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}}, \qquad (3.4.10)$$

$$\begin{aligned}\Phi_n^{(3.4.11)} \quad = \quad & \sum_{j\in I_n} \sigma_{t_{j-1}}^2 h W_j^2 1_{\{|W_j|\le c(\alpha_0)^{1/2}\}} \big(K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} - 1\big) \\ & +2 \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} \big(\widetilde{\sigma}_s - \widetilde{\sigma}_{t_{j-1}}\big) dw_s \sigma_t dw_t K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \\ & +2 \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} \sigma_{t_{j-1}} dw_s \big(\widetilde{\sigma}_t - \widetilde{\sigma}_{t_{j-1}}\big) dw_t K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \\ & +2 \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \widetilde{\sigma}_{t_{j-1}} \big(\widetilde{\sigma}_t - \widetilde{\sigma}_{t_{j-1}}\big) dt K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \\ & + \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \left(\int_{t_{j-1}}^{t_j} \big(\widetilde{\sigma}_t - \widetilde{\sigma}_{t_{j-1}}\big) dw_t\right)^2 K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \qquad (3.4.11)\end{aligned}$$

and

$$\begin{aligned}\Phi_n^{(3.4.12)} \quad = \quad & 2 \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} b_s ds \sigma_t dw_t K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \\ & +2 \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} \sigma_s dw_s b_t dt K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}} \\ & +2 \sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} b_s ds b_t dt K_{n,j} 1_{\{j\in\mathfrak{L}_n^c\}}. \qquad (3.4.12)\end{aligned}$$

Let $\epsilon > \xi$. For $p > 1$ and $\epsilon' > 0$,

$$\begin{aligned}
&\Big\| \sum_{j\in I_n} \sigma^2_{t_{j-1}} h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} \big(K_{n,j} 1_{\{j\in \mathfrak{L}_n^c\}} - 1\big) \Big\|_p \\
\le\; & \Big\| \sum_{j\in I_n} \sigma^2_{t_{j-1}} h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} K_{n,j} 1_{\{j\in \mathfrak{L}_n\}} \Big\|_p \\
& + \Big\| \sum_{j\in I_n} \sigma^2_{t_{j-1}} h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} \big(K_{n,j} - 1\big) \Big\|_p \\
\le\; & \Big\| \max_{j\in I_n} \Big(\sigma^2_{t_{j-1}} h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} K_{n,j} \Big) \overline{N}_T \Big\|_p + O(n^{-L}) \\
\le\; & \|\overline{N}_T 1_{\{\overline{N}_T > n^\epsilon\}}\|_{2p} \Big\| \max_{j\in I_n} \Big(\sigma^2_{t_{j-1}} h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} K_{n,j} \Big) \Big\|_{2p} \\
& + n^\epsilon \Big\| \max_{j\in I_n} \Big(\sigma^2_{t_{j-1}} h W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} K_{n,j} \Big) \Big\|_p + O(n^{-L}) \\
\lesssim\; & n^{-\frac{L\epsilon}{2p}} \|\overline{N}_T\|_{2p+L}^{\frac{2p+L}{2p}} \times n^{-1+\epsilon'} + n^{-1+\epsilon+\epsilon'} + O(n^{-L}) \\
\lesssim\; & n^{-\frac{L(\epsilon-\xi)}{2p}+\xi-1+\epsilon'} + n^{-1+\epsilon+\epsilon'} + O(n^{-L}) \\
\lesssim\; & n^{-1+\epsilon+\epsilon'}
\end{aligned} \tag{3.4.13}$$

since $\epsilon > \xi$, where L is a sufficiently large number chosen suitably depending on $(\epsilon, \xi, p, \epsilon')$.

From the estimate (3.4.13), we have

$$\|\Phi_n^{(3.4.11)}\|_p \;\lesssim\; h^{1/2} + h^{1-\epsilon-\epsilon'} \lesssim h^{1/2} \tag{3.4.14}$$

if letting $\epsilon \downarrow \xi < 1/2$ and $\epsilon' \downarrow 0$.

By the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned}
\Big\| \sum_{j\in \widetilde{\mathcal{J}}_n(\alpha)} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} b_s ds \sigma_t dw_t K_{n,j} 1_{\{j\in \mathfrak{L}_n^c\}} \Big\|_p &\le \sum_{j\in I_n} \Big\| \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} b_s ds \sigma_t dw_t \Big\|_p \\
&\lesssim \sum_{j\in I_n} \sqrt{ \Big\| \int_{t_{j-1}}^{t_j} \Big(\int_{t_{j-1}}^{t} b_s ds \sigma_t \Big)^2 dt \Big\|_{p/2} } \\
&\le \sum_{j\in I_n} \sqrt{ \int_{t_{j-1}}^{t_j} \Big\| \int_{t_{j-1}}^{t} b_s ds \sigma_t \Big\|_p^2 dt } \\
&\lesssim h^{1/2}.
\end{aligned}$$

From this and similar estimates, we have

$$\|\Phi_n^{(3.4.12)}\|_p \;\lesssim\; h^{1/2} \tag{3.4.15}$$

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

$$\begin{aligned}
\Big\| \Phi_j^{(3.4.10)} - \sum_{j\in I_n} \sigma^2_{t_{j-1}} q(\alpha) h \Big\|_p &\le \Big\| h \sum_{j\in I_n} \sigma^2_{t_{j-1}} \big(W_j^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}} - q(\alpha) \big) \Big\|_p \\
&= O(h^{1/2})
\end{aligned} \tag{3.4.16}$$

for every $p > 1$.

Obviously,

$$\sup_{j\in I_n}\left\|1_{\{j\in\mathfrak{L}_n^c\}}\left(\sigma^2_{t_k}-\sigma^2_{t_{j-1}}\right)\right\|_p \;\leq\; \sup_{j\in I_n}\left\|\widetilde{\sigma}^2_{t_k}-\widetilde{\sigma}^2_{t_{j-1}}\right\|_p \;\lesssim\; h^{1/2} \tag{3.4.17}$$

for every $p > 1$. In view of (3.4.17), we deduce that

$$\begin{aligned}
&\left\|\sum_{j\in I_n}\sigma^2_{t_{j-1}}h-\int_0^T\sigma^2_t dt\right\|_p \\
\leq\;& \left\|\sum_{j\in I_n}\int_{t_{j-1}}^{t_j}|\widetilde{\sigma}^2_t-\widetilde{\sigma}^2_{t_{j-1}}|dt\right\|_p+\left\|\sum_{j\in I_n}\int_{t_{j-1}}^{t_j}(\sigma^2_t-\sigma^2_{t_{j-1}})dt1_{\{j\in\mathfrak{L}_n\}}\right\|_p \\
\leq\;& O(h^{1/2})+\left\|\max_{j\in I_n}\left\{\int_{t_{j-1}}^{t_j}\left(|\sigma^2_t|+|\sigma^2_{t_{j-1}}|\right)dt\right\}\overline{N}_T\right\|_p \\
=\;& O(h^{1/2}),
\end{aligned} \tag{3.4.18}$$

following the passage from (3.4.13) to (3.4.14).

Easily,

$$\left\|\sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)}|\Delta_j X|^2K_{n,j}1_{\{j\in\mathfrak{L}_n\}}\right\|_p \;\leq\; \left\|n^{-1/2}\overline{N}_T\right\|_p \;\lesssim\; n^{-\frac{1}{2}+\xi}. \tag{3.4.19}$$

Combining (3.4.19), (3.4.9), (3.4.14), (3.4.15) (3.4.16) and (3.4.18), we obtain

$$\left\|\widetilde{\mathbb{V}}_n(\alpha)-\int_0^T\sigma^2_t dt\right\|_p \;=\; O\left(n^{-\frac{1}{2}+\xi}\right)$$

as $n\to\infty$ for every $p > 1$. □

Theorem 3.4.11. *Suppose that [G1] [G2] and [G3] are fulfilled. Suppose that $\xi < \frac{1}{2}$. Let $\alpha \in (0,1)$ and $\beta_0 < \min\left\{\gamma_0, \frac{1}{2}-\xi\right\}$. Then*

$$\left\|\mathbb{V}_n(\alpha)-\Theta\right\|_p \;=\; O(n^{-\beta_0})$$

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

Proof. Use Lemmas 3.4.8, 3.4.9 and 3.4.10. □

3.4.2 Rate of convergence of the WGRV with a fixed α

Next, we discuss the convergence of the WGRV with a fixed α. Recall that the WGRV is defined as

$$\mathbb{W}_n(\alpha)=\sum_{j\in I_n}\mathsf{w}(\alpha)^{-1}\left\{|\Delta_j X|\wedge(S^{1/2}_{n,j-1}V_{(s_n(\alpha))})\right\}^2K_{n,j}.$$

The WGRV has entirely the same rate of convergence as the GRV.

Theorem 3.4.12. *Suppose that [G1], [G2], and [G3] are fulfilled. Suppose that $\xi < \frac{1}{2}$. Let $\alpha \in (0,1)$ and $\beta_0 < \min\left\{\gamma_0, \frac{1}{2} - \xi\right\}$. Moreover, assume that $\kappa_n = O(n^{1/2})$. Then*

$$\|\mathbb{W}_n(\alpha) - \Theta\|_p = O(n^{-\beta_0})$$

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

Proof. Decompose $\mathbb{W}_n(\alpha)$ as

$$\begin{aligned}\mathbb{W}_n(\alpha) &= \sum_{j \in \mathcal{J}_n(\alpha)} \mathsf{w}(\alpha)^{-1} |\Delta_j X|^2 K_{n,j} + \sum_{j \in \mathcal{J}_n(\alpha)^c} \mathsf{w}(\alpha)^{-1} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} \\ &= \frac{q(\alpha)}{\mathsf{w}(\alpha)} \mathbb{V}_n(\alpha) + \sum_{j \in \mathcal{J}_n(\alpha)^c} \mathsf{w}(\alpha)^{-1} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j}.\end{aligned}$$

Note that $\mathsf{w}(\alpha) = q(\alpha) + \alpha c(\alpha)$. Hence, it suffices to show that

$$\left\| \sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} - \alpha c(\alpha) \Theta \right\|_p = O(n^{-\beta_0})$$

as $n \to \infty$ for every $p > 1$. Decompose the left-hand side as

$$\begin{aligned}\sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} - \alpha c(\alpha) \Theta = &\sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} 1_{\{j \in \mathcal{L}_n^c\}} - \alpha c(\alpha) \Theta \\ &+ \sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} 1_{\{j \in \mathfrak{L}_n \cap \mathcal{L}_n\}} \\ &+ \sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} 1_{\{j \in \mathfrak{L}_n^c \cap \mathcal{L}_n\}} \\ =: &A_1 + A_2 + A_3.\end{aligned}$$

Since $S_{n,j-1} V_{(s_n(\alpha))}^2 K_{n,j} \le |\Delta_j X|^2 K_{n,j} \le n^{-1/2}$ for $j \in \mathcal{J}_n(\alpha)^c$, we have $\|A_2\|_p \lesssim n^{-1/2+\xi}$. As for A_3, note that $\# \mathcal{L}_n \lesssim n^{\xi} \times \bar{\kappa}_n = O(n^{\xi+1/2})$ and that $\Delta_j X = \Delta_j \tilde{X}$ for $j \in \mathfrak{L}_n^c$. Hence we have

$$\|A_3\|_p \le \left\| \max_{j \in I_n} |\Delta_j \tilde{X}|^2 \#(\mathfrak{L}_n^c \cap \mathcal{L}_n) \right\|_p \lesssim n^{-1/2+\xi+\epsilon},$$

where ϵ is an arbitrarily small positive number.

As for A_1, we can set $c = 1$ in the condition $[G2]$(ii) without loss of generality.

$$
\begin{aligned}
\|A_1 - \alpha c(\alpha)\Theta\|_p \leq & \left\| \Big(h^{-1}V^2_{(s_n(\alpha))} - c(\alpha)\Big) h \sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1}K_{n,j}1_{\{j \in \mathcal{L}_n^c\}} \right\|_p \\
& + c(\alpha) \left\| h \sum_{j \in \mathcal{J}_n(\alpha)^c} \big(S_{n,j-1} - \sigma^2_{t_{j-1}}\big) 1_{\{j \in \mathcal{L}_n^c\}} \right\|_p \\
& + c(\alpha) \left\| h \sum_{j \in \mathcal{J}_n(\alpha)^c} S_{n,j-1}\big(1 - K_{n,j}\big) 1_{\{j \in \mathcal{L}_n^c\}} \right\|_p \\
& + c(\alpha) \left\| h \sum_{j \in \mathcal{J}_n(\alpha)^c} \sigma^2_{t_{j-1}} 1_{\{j \in \mathcal{L}_n^c\}} - \alpha\Theta \right\|_p \\
=: & B_1 + B_2 + B_3 + B_4.
\end{aligned}
$$

By condition $[G2]$, $B_2 = O(n^{-\gamma_0})$. As for B_3, with the estimate $\left\|1_{\{j \in \mathcal{L}_n^c\}}(1 - K_{n,j})\right\|_p \leq P[|\Delta_j \tilde{X}| > n^{-1/4}]^{1/p} = O(n^{-L})$ (for all $p > 1$ and $L > 0$) and the Cauchy-Schwarz inequality, we have

$$
\|B_3\|_p \leq h \sum_{j \in I_n} \left\|1_{\{j \in \mathcal{L}_n^c\}} S_{n,j-1}\right\|_{2p} \left\|1_{\{j \in \mathcal{L}_n^c\}}(1 - K_{n,j})\right\|_{2p} = O(n^{-L}).
$$

For B_4, we use the following decomposition:

$$
\begin{aligned}
& h \sum_{j \in \mathcal{J}_n(\alpha)^c} \sigma^2_{t_{j-1}} 1_{\{j \in \mathcal{L}_n^c\}} - \alpha\Theta \\
= & \left(h \sum_{j \in I_n} \sigma^2_{t_{j-1}} - \Theta \right) - h \sum_{j \in I_n} \sigma^2_{t_{j-1}} 1_{\{j \in \mathcal{L}_n\}} + h \sum_{j \in \mathcal{J}_n(\alpha)} \sigma^2_{t_{j-1}} 1_{\{j \in \mathcal{L}_n\}} \\
& + \left((1-\alpha)\Theta - h \sum_{j \in \widetilde{\mathcal{J}}_n(\alpha)} \sigma^2_{t_{j-1}} \right) + \left(h \sum_{j \in \widetilde{\mathcal{J}}_n(\alpha)} \sigma^2_{t_{j-1}} - h \sum_{j \in \mathcal{J}_n(\alpha)} \sigma^2_{t_{j-1}} \right).
\end{aligned}
$$

Hence, with the aid of Lemmas 3.4.5, 3.4.6 and the estimate $\|\mathcal{L}_n\|_p \lesssim n^{\xi + 1/2}$, we have

$$
\begin{aligned}
& \left\| h \sum_{j \in \mathcal{J}_n(\alpha)^c} \sigma^2_{t_{j-1}} 1_{\{j \in \mathcal{L}_n^c\}} - \alpha\Theta \right\|_p \\
\lesssim & \left\| h \sum_{j \in I_n} \sigma^2_{t_{j-1}} - \Theta \right\|_p + \left\| (1-\alpha)\Theta - h \sum_{j \in \widetilde{\mathcal{J}}_n(\alpha)} \sigma^2_{t_{j-1}} \right\|_p + O(n^{-\beta_0}) \qquad (3.4.20)
\end{aligned}
$$

since $\beta_0 < \frac{1}{2} - \xi$. The first term of the right-hand side of the above inequality is $O(n^{-1/2})$ by (3.4.18).

As for the second term on the right-hand side of (3.4.20),

$$\begin{aligned}\left\|(1-\alpha)\Theta - h\sum_{j\in\widetilde{\mathcal{J}}_n(\alpha)}\sigma_{t_{j-1}}^2\right\|_p &= \left\|(1-\alpha)\Theta - h\sum_{j\in I_n}\sigma_{t_{j-1}}^2 1_{\{|W_j|\le c(\alpha)^{1/2}\}}\right\|_p \\ &\le \left\|h\sum_{j\in I_n}\sigma_{t_{j-1}}^2\Big(1_{\{|W_j|\le c(\alpha)^{1/2}\}} - P\Big[|W_j|\le c(\alpha)^{1/2}\Big]\Big)\right\|_p \\ &\quad + (1-\alpha)\left\|h\sum_{j\in I_n}\sigma_{t_{j-1}}^2 - \Theta\right\|_p \\ &= O(n^{-1/2}).\end{aligned}$$

since Hence we have $B_4 = O(n^{-\frac{1}{2}+\xi})$.

Finally, for B_1, it suffices to show that

$$P\left[\left|h^{-1/2}V_{(s_n(\alpha))} - c(\alpha)^{1/2}\right| > n^{-\beta_0}\right] \quad = \quad O(n^{-L}) \tag{3.4.21}$$

as $n\to\infty$ for every $L>0$ and for every $\beta_0 < \min\{\gamma_0, \frac{1}{2}-\xi\}$. Let

$$\begin{aligned} A_{n,j} &= \left\{h^{-1/2}V_j < c(\alpha)^{1/2} - n^{-\beta_0}\right\} \\ \mathcal{V}_{n,j} &= 1_{\left\{|W_j|\le c(\alpha)^{1/2} - n^{-\beta_0} + 2^{-1}n^{-\gamma_1}\right\}}\end{aligned}$$

and

$$\mu_n \quad = \quad (1-\alpha)n - 1 - n^{\frac{1}{2}+\xi+\epsilon} - nE[\mathcal{V}_{n,j}]$$

for $\epsilon > 0$. Then

$$\begin{aligned} & P\left[h^{-1/2}V_{(s_n(\alpha))} - c(\alpha)^{1/2} < -n^{-\beta_0}\right] \\ \le\ & P\left[\sum_{j\in I_n} 1_{A_{n,j}} \ge (1-\alpha)n - 1\right] \\ \le\ & P\left[\sum_{j\in I_n} 1_{A_{n,j}\cap\{j\in\mathcal{L}_n^c\}} + \sum_{j\in I_n} 1_{A_{n,j}\cap\{j\in\mathcal{L}_n\}} \ge (1-\alpha)n - 1\right] \\ \le\ & P\left[\sum_{j\in I_n} 1_{A_{n,j}\cap\{j\in\mathcal{L}_n^c\}} + \#\mathcal{L}_n \ge (1-\alpha)n - 1\right] \\ \le\ & P\left[\sum_{j\in I_n} \mathcal{V}_{n,j} \ge (1-\alpha)n - 1 - n^{\frac{1}{2}+\xi+\epsilon}\right] + P[\#\mathcal{L}_n > n^{\frac{1}{2}+\xi+\epsilon}] + P[\Omega_n^c] \\ \le\ & P\left[\sum_{j\in I_n} \big(\mathcal{V}_{n,j} - E[\mathcal{V}_{n,j}]\big) \ge \mu_n\right] + P[\#\mathcal{L}_n > n^{\frac{1}{2}+\xi+\epsilon}] + P[\Omega_n^c]\end{aligned}$$

We see

$$\mu_n \ \sim \ (1-\alpha)n - 1 - n^{\frac{1}{2}+\xi+\epsilon} - n\big\{(1-\alpha) - c^*\big(n^{-\beta_0} - 2^{-1}n^{-\gamma_1}\big)\big\} \ \ge \ \frac{1}{2}c^* n^{1-\beta_0}$$

for large n, where c^* is some positive constant, if we take a sufficiently small ϵ and $\gamma_1 \in (\beta_0, \gamma_0)$ thanks to $\beta_0 < \frac{1}{2} - \xi$. Since $n^{-1/2}\mu_n \geq 2^{-1}c^* n^{\frac{1}{2}-\beta_0}$ from $\beta_0 < 1/2$, we obtain

$$P\left[n^{-1/2}\sum_{j\in I_n}\left(\mathcal{V}_{n,j} - E[\mathcal{V}_{n,j}]\right) \geq n^{-1/2}\mu_n\right] = O(n^{-L})$$

for every $L > 0$. Therefore,

$$P\left[h^{-1/2}V_{(s_n(\alpha))} - c(\alpha)^{1/2} < -n^{-\beta_0}\right] = O(n^{-L})$$

as $n \to \infty$ for every $L > 0$. Similarly, we can obtain the estimate $P\left[h^{-1/2}V_{(s_n(\alpha))} - c(\alpha)^{1/2} > n^{-\beta_0}\right] = O(n^{-L})$ to show (3.4.21), which concludes the proof. □

3.5 Asymptotic mixed normality of the global realized volatilities with a moving threshold

3.5.1 The GRV with a moving threshold

In this section, we will consider a situation where the intensity of jumps is moderate. Then it is possible to keep the cut-off ratio of the data small, and to get a precise estimate for the integrated volatility. Let

$$\delta_0 \in \left(0, \frac{1}{4}\right) \quad \text{and} \quad \delta_1 \in \left(0, \frac{1}{2}\right). \tag{3.5.1}$$

In what follows, we will only consider sufficiently large n. In the context of the global jump filtering, given a collection $(\mathfrak{S}_{n,j-1})_{j\in I_n}$ ($n \in \mathbb{N}$) of positive random variables, we may use the index set $\mathcal{M}_n$ given by

$$\mathcal{M}_n = \left\{j \in I_n;\ V_j < V_{(s_n)}\right\} \tag{3.5.2}$$

where

$$V_j = \left|(\mathfrak{S}_{n,j-1})^{-1/2}\,\Delta_j X\right| \tag{3.5.3}$$

and

$$s_n = n - \lfloor Bn^{\delta_1}\rfloor$$

for a positive constant B. We note that the definition of V_j is different from that in (3.2.1). In the terminology of the previous sections, the cut-off ratio is $\alpha_n = \lfloor Bn^{\delta_1}\rfloor/n$, $\mathcal{M}_n = \mathcal{J}_n(\alpha_n)$ and α_n goes to 0 as n tends to ∞.

For estimation of Θ of (3.1.1), we consider the global realized volatility with a moving threshold

$$\mathbf{V}_n = \sum_{j\in\mathcal{M}_n} q_n^{-1}\left|\Delta_j X\right|^2 H_{n,j} \tag{3.5.4}$$

where $(q_n)_{n\in\mathbb{N}}$ is a sequence of positive numbers, and

$$H_{n,j} = 1_{\left\{|\Delta_j X| < B_0 n^{-\frac{1}{4}-\delta_0}\right\}} \tag{3.5.5}$$

for a positive constant B_0.

Remark 3.5.1. $\mathfrak{S}_{n,j-1} = 1$ and $q_n = 1$ satisfy Condition $[G2']$. Asymptotically this choice is sufficient and valid. However, in practice, a natural choice is the pair $S_{n,j-1}$ in $[G2]$ and $q_n = q(\alpha_n)$.

We are about establishing asymptotic mixed normality of the integrated volatility estimator having a moving threshold. We will solve this problem by showing a stability of estimation under elimination of a certain portion of the data. In what follows, we will consider the variable $\mathbf{V}_n$ defined by (3.5.4) with (3.5.5) for $\mathcal{M}_n$, that is just a random index set in I_n. It is not necessary to specify it by (3.5.2) and (3.5.3). We will assume (3.5.1) and

[G2′] **(i)** For every $n \in \mathbb{N}$, $\mathcal{M}_n$ is a random set in I_n such that $\#\big(I_n \setminus \mathcal{M}_n\big) \le B_1 n^{\delta_1}$ $(n \in \mathbb{N})$ for some positive constant B_1.

(ii) $q_n > 0$ $(n \in \mathbb{N})$ and $q_n - 1 = o(n^{-1/2})$ as $n \to \infty$.

Let

$$\mathbf{V}_n^\dagger \;=\; \sum_{j \in \mathcal{M}_n} q_n^{-1} \big|\Delta_j \widetilde{X}\big|^2 H_{n,j}$$

for $\widetilde{X} = X - J$.

Lemma 3.5.2. *Suppose that* $[G1]$, $[G2']$ *and* $[G3]$ *are satisfied. Suppose that* $\xi < 2\delta_0$. *Then*

$$n^{1/2}\big\|\mathbf{V}_n - \mathbf{V}_n^\dagger\big\|_p \;\to\; 0$$

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

Proof. We have the estimate

$$n^{1/2}\big\|\mathbf{V}_n - \mathbf{V}_n^\dagger\big\|_p \;\le\; 2q_n^{-1}\Phi_n^{(3.5.7)} + q_n^{-1}\Phi_n^{(3.5.8)}, \tag{3.5.6}$$

where

$$\Phi_n^{(3.5.7)} \;=\; n^{1/2}\bigg\| \sum_{j \in \mathcal{M}_n} \big|\Delta_j \widetilde{X} \Delta_j J\big| H_{n,j} \bigg\|_p, \tag{3.5.7}$$

and

$$\Phi_n^{(3.5.8)} \;=\; n^{1/2}\bigg\| \sum_{j \in \mathcal{M}_n} \big|\Delta_j J\big|^2 H_{n,j} \bigg\|_p \tag{3.5.8}$$

for $p > 1$. By using the inequality

$$|\Delta_j J| H_{n,j} \;\le\; \big(|\Delta_j \widetilde{X}| + B_0\, n^{-\frac{1}{4}-\delta_0}\big) 1_{\{\Delta_j J \neq 0\}},$$

we obtain

$$\begin{aligned}\Phi_n^{(3.5.7)} \;&\le\; n^{1/2}\bigg\| \max_{j \in I_n} \big\{ \big|\Delta_j \widetilde{X}\big| \big(|\Delta_j \widetilde{X}| + B_0\, n^{-\frac{1}{4}-\delta_0}\big) \big\} \bigg\|_{2p} \big\|N_T\big\|_{2p} \\ &\lesssim\; n^{-\frac{1}{4}-\delta_0+\xi+\epsilon}\end{aligned}$$

as $n \to \infty$ for any $\epsilon > 0$ and $p > 1$. Therefore,

$$\Phi_n^{(3.5.7)} \;\to\; 0 \tag{3.5.9}$$

for every $p > 1$ since $\xi < 2\delta_0 < \frac{1}{4} + \delta_0$. Similarly,

$$\begin{aligned}\Phi_n^{(3.5.8)} &\lesssim n^{1/2}\Big\|\max_{j\in I_n}(|\Delta_j\widetilde{X}|^2 + n^{-\frac{1}{2}-2\delta_0})\Big\|_{2p}\|N_T\|_{2p} \\ &\lesssim n^{-2\delta_0+\xi+\epsilon}\end{aligned}$$

as $n \to \infty$ for any $\epsilon > 0$ and $p > 1$ since $2\delta_0 > \xi$. In particular,

$$\Phi_n^{(3.5.8)} \to 0 \tag{3.5.10}$$

as $n \to \infty$ since $\xi < 2\delta_0$. Now the proof is completed with (3.5.6), (3.5.9) and (3.5.10). □

Define $\widetilde{\mathbf{V}}_n$ by

$$\widetilde{\mathbf{V}}_n = \sum_{j\in I_n}|\Delta_j\widetilde{X}|^2.$$

Lemma 3.5.3. *Suppose that $\xi < 1/2$. Then*

$$n^{1/2}\|\mathbf{V}_n^\dagger - \widetilde{\mathbf{V}}_n\|_p \to 0$$

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

Proof. Recall that $\delta_0 < 1/4$ and $\delta_1 < 1/2$. Define $\mathbf{V}_n^\ddagger$ by

$$\mathbf{V}_n^\ddagger = \sum_{j\in\mathcal{M}_n} q_n^{-1}|\Delta_j\widetilde{X}|^2.$$

Then

$$\begin{aligned}n^{1/2}\|\mathbf{V}_n^\dagger - \mathbf{V}_n^\ddagger\|_p &\lesssim n^{1/2}\Big\|\sum_{j\in\mathcal{M}_n}|\Delta_j\widetilde{X}|^2|H_{n,j}-1|\Big\|_p \\ &\le n^{1/2}\Big\|\sum_{j\in\mathcal{M}_n}|\Delta_j\widetilde{X}|^2|H_{n,j}-1|1_{\{\Delta_j N>0\}}\Big\|_p \\ &\quad +n^{1/2}\Big\|\sum_{j\in\mathcal{M}_n}|\Delta_j\widetilde{X}|^2 1_{\{|\Delta_j\widetilde{X}|>n^{-\frac{1}{4}-\delta_0}\}}1_{\{\Delta_j N=0\}}\Big\|_p \\ &\le n^{1/2}\Big\|\max_{j\in I_n}|\Delta_j\widetilde{X}|^2 N_T\Big\|_p + O(n^{-L}) \\ &\lesssim n^{-\frac{1}{2}+\epsilon+\xi}\end{aligned}$$

for any positive number $\epsilon > 0$. Here L is an arbitrary positive number greater than 1/2, and we used the inequality $\delta_0 < 1/4$ to get $O(n^{-L})$. Since $\xi < 1/2$, we obtain

$$n^{1/2}\|\mathbf{V}_n^\dagger - \mathbf{V}_n^\ddagger\|_p = o(1) \tag{3.5.11}$$

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

From the condition $q_n - 1 = o(n^{-1/2})$ of [$G2'$] (ii), obviously,

$$\begin{aligned}n^{1/2}\|\mathbf{V}_n^\ddagger - \widetilde{\mathbf{V}}_n\|_p &\le n^{1/2}\Big\|\sum_{j\in I_n\setminus\mathcal{M}_n}|\Delta_j\widetilde{X}|^2\Big\|_p + o(1) \\ &\lesssim n^{-\frac{1}{2}+\epsilon+\delta_1} + o(1) = o(1)\end{aligned} \tag{3.5.12}$$

as $n \to \infty$ for every $p > 1$ since $\#(I_n \setminus \mathcal{M}_n) \lesssim n^{\delta_1}$ with $\delta_1 < 1/2$ thanks to $[G2']$ (i) and (3.5.1), and

$$\left\| \max_{j \in I_n} \left| \Delta_j \widetilde{X} \right|^2 \right\|_p = O(n^{-1+\epsilon})$$

for any $p > 1$ and any positive number ϵ. Proof ends with (3.5.11) and (3.5.12). □

Define Γ by

$$\Gamma = 2T \int_0^T \sigma_t^4 dt.$$

Extend $(\Omega, \mathcal{F}, P)$ so that there is a standard normal random variable ζ independent of $\mathcal{F}$ on the extension. The $\mathcal{F}$-stable convergence is denoted by $\to^{d_s}$.

Lemma 3.5.4. *Suppose that $[G1]$ is satisfied. Then*

$$n^{1/2}\big(\widetilde{\mathbf{V}}_n - \Theta\big) \to^{d_s} \Gamma^{1/2}\zeta$$

as $n \to \infty$.

Proof. We have

$$\begin{aligned} \widetilde{\mathbf{V}}_n &= \sum_{j \in I_n} \left(\int_{t_{k-1}}^{t_k} \sigma_t dw_t + \int_{t_{k-1}}^{t_k} b_t dt \right)^2 \\ &= \Phi_n^{(3.5.14)} + \Phi_n^{(3.5.15)} + 2\Phi_n^{(3.5.16)} + \Phi_n^{(3.5.17)} \end{aligned} \tag{3.5.13}$$

where

$$\Phi_n^{(3.5.14)} = \sum_{j \in I_n} 2 \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^{t} \sigma_s dw_s \sigma_t dw_t, \tag{3.5.14}$$

$$\Phi_n^{(3.5.15)} = \sum_{j \in I_n} \int_{t_{j-1}}^{t_j} \sigma_t^2 dt, \tag{3.5.15}$$

$$\Phi_n^{(3.5.16)} = \sum_{j \in I_n} \int_{t_{j-1}}^{t_j} \sigma_t dw_t \int_{t_{j-1}}^{t_j} b_t dt, \tag{3.5.16}$$

and

$$\Phi_n^{(3.5.17)} = \sum_{j \in I_n} \left(\int_{t_{j-1}}^{t_j} b_t dt \right)^2. \tag{3.5.17}$$

Since b is a càdlàg process, for any $\epsilon > 0$, there exists a number $\delta > 0$ such that $P\big[w'(b,\delta) \geq \epsilon\big] < \epsilon$. Here $w'(b,\delta)$ is a modulus of continuity defined by

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

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

$$\dot{\Phi}_n^{(3.5.16)} = \sum_{j\in I_n} \int_{t_{k-1}}^{t_k} \sigma_t dw_t \int_{t_{j-1}}^{t_j} (b_t - b_{t_{j-1}})dt. \tag{3.5.18}$$

Write

$$\begin{aligned} \mathsf{E}_j &= \int_{t_{j-1}}^{t_j} \sigma_t dw_t, \\ \mathsf{V}_j &= n^{1/2}\left|\int_{t_{j-1}}^{t_j} \sigma_t dw_t\right| \int_{t_{j-1}}^{t_j} \big(|b_t| + |b_{t_{j-1}}|\big)dt. \end{aligned}$$

For $\omega \in \Omega$ such that $w'(b(\omega), \delta) < \epsilon$, there exists a (s_i) (depending on ω) such that

$$\begin{aligned} &\max_i \sup_{r_1,r_2\in[s_{i-1},s_i)} \big|b_{r_1}(\omega) - b_{r_2}(\omega)\big| < \epsilon, \\ &\min_{i=1,\ldots,v-1} (s_i - s_{i-1}) > \delta. \end{aligned}$$

For $n > T/\delta$, all intervals $[t_{j-1}, t_j)$ $(j \in I_n)$ includes at most one point among (s_i), therefore the number of intervals $[t_{j-1}, t_j)$ that include some one s_i is at most T/δ. The increment of $b(\omega)$ in $[t_{j-1}, t_j)$ is less than ϵ if $[t_{j-1}, t_j) \cap \{s_i\}$. Thus, we have the inequality

$$\big\|n^{1/2}\dot{\Phi}_n^{(3.5.16)}\big\|_p \le \left\|\sum_{j\in I_n} n^{1/2}|\mathsf{E}_j|\right\|_p \epsilon h + \left\|\max_{j\in I_n} \mathsf{V}_j\right\|_p \frac{T}{\delta} + \left\|\sum_{j\in I_n} \mathsf{V}_j\right\|_{2p} P\big[w'(b,\delta) \le \epsilon\big]^{\frac{1}{2p}}$$

for every $p > 1$. Therefore,

$$\begin{aligned} \big\|n^{1/2}\dot{\Phi}_n^{(3.5.16)}\big\|_p &\le C\left[\epsilon + \left(n^{-1/2} + \sum_{j\in I_n} \big\|\mathsf{V}_j 1_{\{\mathsf{V}_j > n^{-1/2}\}}\big\|_p\right)\frac{T}{\delta} + \epsilon^{\frac{1}{2p}}\right] \\ &\le C'\big(\epsilon + n^{-1/2} + \epsilon^{\frac{1}{2p}}\big) \end{aligned}$$

for all $n > T/\delta$, where C and C' are some constants independent of n. Consequently,

$$\lim_{n\to\infty} \big\|n^{1/2}\dot{\Phi}_n^{(3.5.16)}\big\|_p = 0 \tag{3.5.19}$$

for every $p > 1$. Moreover, for

$$\ddot{\Phi}_n^{(3.5.16)} = \sum_{j\in I_n} \int_{t_{j-1}}^{t_j} \sigma_t dw_t \int_{t_{k-1}}^{t_k} b_{t_{j-1}} dt = \sum_{j\in I_n} h b_{t_{j-1}} \int_{t_{j-1}}^{t_j} \sigma_t dw_t,$$

we have

$$\lim_{n\to\infty} \big\|n^{1/2}\ddot{\Phi}_n^{(3.5.16)}\big\|_p = 0 \tag{3.5.20}$$

for every $p > 1$, by orthogonality. From (3.5.19) and (3.5.20),

$$\lim_{n\to\infty} \big\|n^{1/2}\Phi_n^{(3.5.16)}\big\|_p = 0 \tag{3.5.21}$$

for every $p > 1$.

Obviously,

$$\left\| n^{1/2} \Phi_n^{(3.5.17)} \right\|_p = 0 \tag{3.5.22}$$

for every $p > 1$. Now, we can show the claim of the lemma by using (3.5.13), (3.5.21) and (3.5.22) together with the mixture type of martingale central limit theorem applied to $\Phi_n^{(3.5.14)}$. □

Theorem 3.5.5. *Suppose that [G1], [G2′] and [G3] are satisfied. Suppose that $\xi < 2\delta_0$. Then*

$$n^{1/2}\big(\mathbf{V}_n - \Theta\big) \quad \to^{d_s} \quad \Gamma^{1/2}\zeta$$

as $n \to \infty$.

Proof. Just combine Lemmas 3.5.2, 3.5.3 and 3.5.4. □

3.5.2 The WGRV with a moving threshold

We consider the winsorized global realized volatility with a moving threshold. Differently from the GRV without winsorization, we need an explicit description of $\mathcal{M}_n$ here. Let $\mathcal{M}_n$ and V_j be given by (3.5.2) and (3.5.3), respectively. Set $s_n = n - \lfloor Bn^{\delta_1} \rfloor$ with $\delta_1 \in (0, 1/2)$. Define

$$\mathbf{W}_n = \sum_{j \in I_n} q_n^{-1} \big\{ |\Delta_j X| \wedge \mathfrak{S}_{n,j-1}^{1/2} V_{(s_n)} \big) \big\}^2 H_{n,j},$$

where $(q_n)_{n \in \mathbb{N}}$ is a sequence of positive numbers, We prove the rate of convergence of WGRV is the same as that of $\mathbf{V}_n$.

Theorem 3.5.6. *Suppose that [G1], [G2′](ii) and [G3] are satisfied. Suppose that $\xi < 2\delta_0$. Then*

$$n^{1/2}\big(\mathbf{W}_n - \Theta\big) \quad \to^{d_s} \quad \Gamma^{1/2}\zeta$$

as $n \to \infty$, where Γ and ζ are the same random variables in Theorem 3.5.5.

Proof. It suffices to show that $n^{1/2}\|\mathbf{W}_n - \mathbf{V}_n\|_p \to 0$ as $n \to \infty$ for every $p > 1$. Rewrite the definition of $\mathbf{W}_n$ as

$$\mathbf{W}_n - \mathbf{V}_n = \sum_{j \in \mathcal{M}_n^c} q_n^{-1} \mathfrak{S}_{n,j-1} V_{(s_n)}^2 H_{n,j}.$$

The first term obviously converges to zero in L^p with rate $\sqrt{n}$. As for the second term, we have

$$\begin{aligned} \left\| \sum_{j \in \mathcal{M}_n^c} \mathfrak{S}_{n,j-1} V_{(s_n)}^2 H_{n,j} \right\|_p &\leq \left\| \sum_{j \in \mathcal{M}_n^c} |\Delta_j X|^2 1_{\{j \in \mathfrak{L}_n\}} H_{n,j} \right\|_p + \left\| \sum_{j \in \mathcal{M}_n^c} |\Delta_j \tilde{X}|^2 1_{\{j \in \mathfrak{L}_n^c\}} H_{n,j} \right\|_p \\ &\lesssim n^{-1/2 - 2\delta_0 + \xi} + n^{-1 + \epsilon + \delta_1}. \end{aligned}$$

Since $\delta_1 < 1/2$ and $\xi < 2\delta_0$, we obtain the desired rate of convergence. □

3.6 Constant volatility

The case of constant σ is specific and theoretical treatments can be slightly different from those of the previous sections. In this situation, we do not need to pre-estimate the local spot volatility, and hence, we can take $S_{n,j} = 1$ constantly and no approximation error is caused. $\sigma_t = \theta\tilde{\sigma}_t$ is also the case if $\tilde{\sigma}_{t_{j-1}}$ are observable. For example, the GRV with a fixed cut-off rate α is redefined as

$$\mathbb{V}_n^0(\alpha) = \sum_{j \in \mathcal{J}_n^0(\alpha)} q(\alpha)^{-1} |\Delta_j X|^2 K_{n,j},$$

where

$$\mathcal{J}_n^0(\alpha) = \big\{ j \in I_n;\ |\Delta_j X| < |\Delta X|_{(s_n(\alpha))} \big\}.$$

Then we have the following theorem. Note that we do not need the condition $[G2]$, and γ_0 in $[G2]$(ii) can be arbitrarily close to $1/2$.

Theorem 3.6.1. *Suppose that $[G1]$ and $[G3]$ are fulfilled. Suppose that $\xi < \frac{1}{2}$. Let $\alpha \in (0,1)$ and $\beta_0 < \frac{1}{2} - \xi$. Then*

$$\| \mathbb{V}_n^0(\alpha) - \Theta \|_p = O(n^{-\beta_0})$$

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

The other global-threshold estimators are discussed similarly.

3.7 Simulation studies

In this section, we conduct several numerical simulations to see that our global realized volatility estimators outperform those proposed in previous studies.

3.7.1 The case of compound Poisson jumps

Here we consider the following process:

$$dX_t = \theta X_t dt + (\sigma + \eta X_t^2)^{\frac{1}{4}} dw_t + dJ_t, \quad t \in [0,1], \tag{3.7.1}$$

where J_t is the jump part of X.

We assume that J is a compound Poisson process of the form $J_t = \sum_{i=1}^{N_t} \xi_i$, where $(N_t)_t$ is a Poisson process with intensity $\lambda > 0$ and $(\xi_i)_i$ are independently and normally distributed random variable with mean μ and variance ν^2. For the intensity parameter, we consider both cases where λ is high and low. Our aim is to estimate the integrated volatility $\Theta = \int_0^1 (\sigma + \eta Y_t^2)^{\frac{1}{2}} dt$.

We compare the estimation results of the bipower variation (BV), minimum realized volatility (minRV), the GRV, and the WGRV.

The set-up of simulation is as follows. The number of samples is $n = 2000$. We repeat calculating the estimators 500 times to obtain their average and quantile. The true parameters are $\theta = 0.2,\ \sigma = 1,\ \eta = 3,\ \mu = 0.3,\ \nu = 0.2$. Throughout this subsection, we set the cut-off ratio $\alpha = 0.2$ for GRV and WGRV with local volatility. That is, we trim the upper 20% of absolute increments. While it may seem that we eliminate too many observations and the estimator suffers from downside bias, GRV and WGRV estimate the integrated volatility well thanks to the adjustment coefficient by $q(\alpha)$ and $\mathsf{w}(\alpha)$.

Note that σ_t is not directly observable and depends on the Y_t. Hence, we have to calculate the local GRV first to normalize the increment $\Delta_i Y$ when constructing the GRV. In this simulation, we use the LGRV (3.3.15) and local minRV (3.3.48) as estimators of spot volatilities. Moreover, we calculate

GRV without normalization (defined in Section 3.6) for comparison. For the length of a subinterval to calculate these local volatilities, we set $\kappa_n = 10 \times n^{0.45}$.

We use the following labels as in Table 3.1 to describe the estimators.

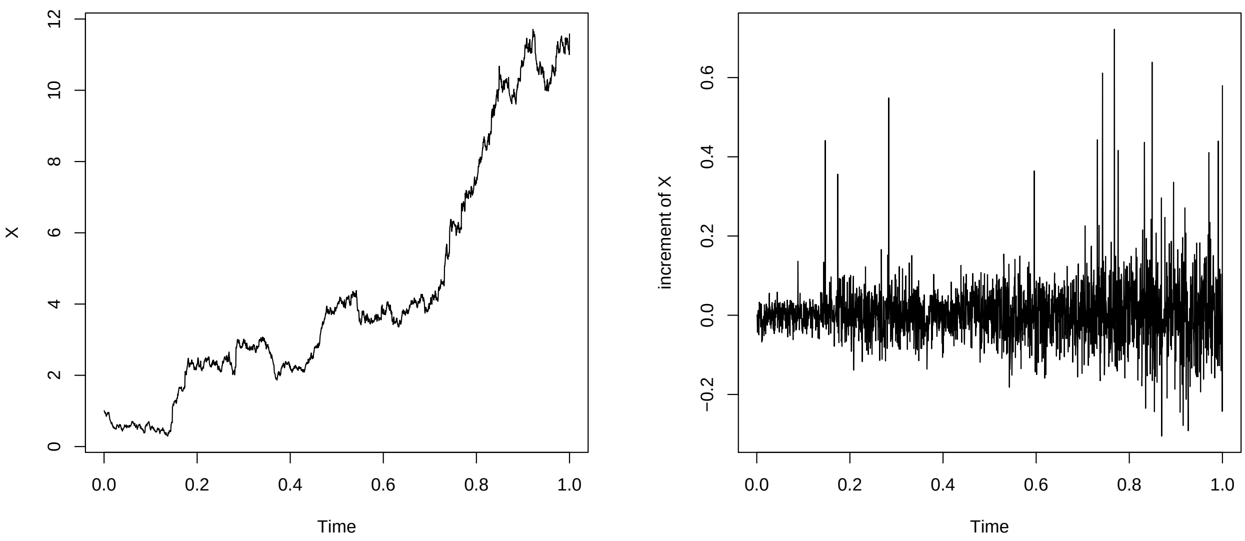
Table 3.1: Definitions of estimators

Label	Method	Cut-off ratio α	Local volatility
`bv`	BV	0.2	-
`mrv`	minRV	0.2	-
`grv.lgrv`	GRV	0.2	GRV
`grv.mrv`	GRV	0.2	minRV
`wgrv.lgrv`	WGR	0.2	GRV
`wgrv.mrv`	WGRV	0.2	minRV
`grv`[α]	GRV	0.2, 0.1, 0.05	-
`grv.lgrv.mov`	GRV with moving threshold	(depends on n)	GRV
`wgrv.lgrv.mov`	WGRV with moving threshold	(depends on n)	GRV

The case of high intensity: GRV with fixed cut-off ratio

First, we deal with the case of high intensity. Here we set $\lambda = 30$ so that the data includes many jumps. The example of a sample path and its increments are shown in Figure 3.1. Obviously, there are many large spikes in the data, suggesting the existence of jumps.

Note that the volatility is non-constant here. In fact, in Panel (b) of Figure 3.1, the size of increments tend to increase as time passes. Hence, to estimate the volatility, we have to use estimated spot volatilities to normalize the increments.



(a) Sample path of X (b) Increment of X

Figure 3.1: Sample path of X and its increments ($\lambda = 30$)

Figure 3.2 is the estimation results. In this case, both BV (`bv`) and minRV (`mrv`) seem to suffer from upward bias due to jumps. In particular, the BV deviates from the true value considerably. On the other hand, GRV with normalization perform well with errors concentrating around zero (`grv.lgrv, grv.mrv`). Note that, although WGRV performs relatively well, it seems to have a small upward bias (`wgrv.lgrv, wgrv.mrv`). This suggests that, if there are many large jumps, using an upper quantile ($V_{(s_n(\alpha))}$) may sometimes lead to biases rather than obtaining a robust estimate.

The three right barplots in this figure (`grv[0.20], grv[0.10], grv[0.05]`) are the results of GRV without normalizing increments by local-global filters, with the cut-off ratio $\alpha = 0.2,\ 0.1,\ 0.05$, respectively. We see that they seem to be less precise (especially when α is large) than GRV or WGRV with local volatility. This result suggests that, if we do not normalize increments by spot volatilities in the case of non-constant volatility, we end up obtaining inappropriate estimates.

Intuitively, when we ignore normalization, we tend to eliminate increments where volatility is high (because they are typically large), even if they come from the Brownian motion, while keeping relatively small jumps which we should actually remove. In addition, theoretically, the adjusting constant $q(\alpha)$ in the definition of GRV (3.2.2) comes from the standard normal distribution. Therefore, when the volatility is non-constant, we should normalize the increments $|\Delta_i Y|$ by local volatility to make them approximately normally distributed.

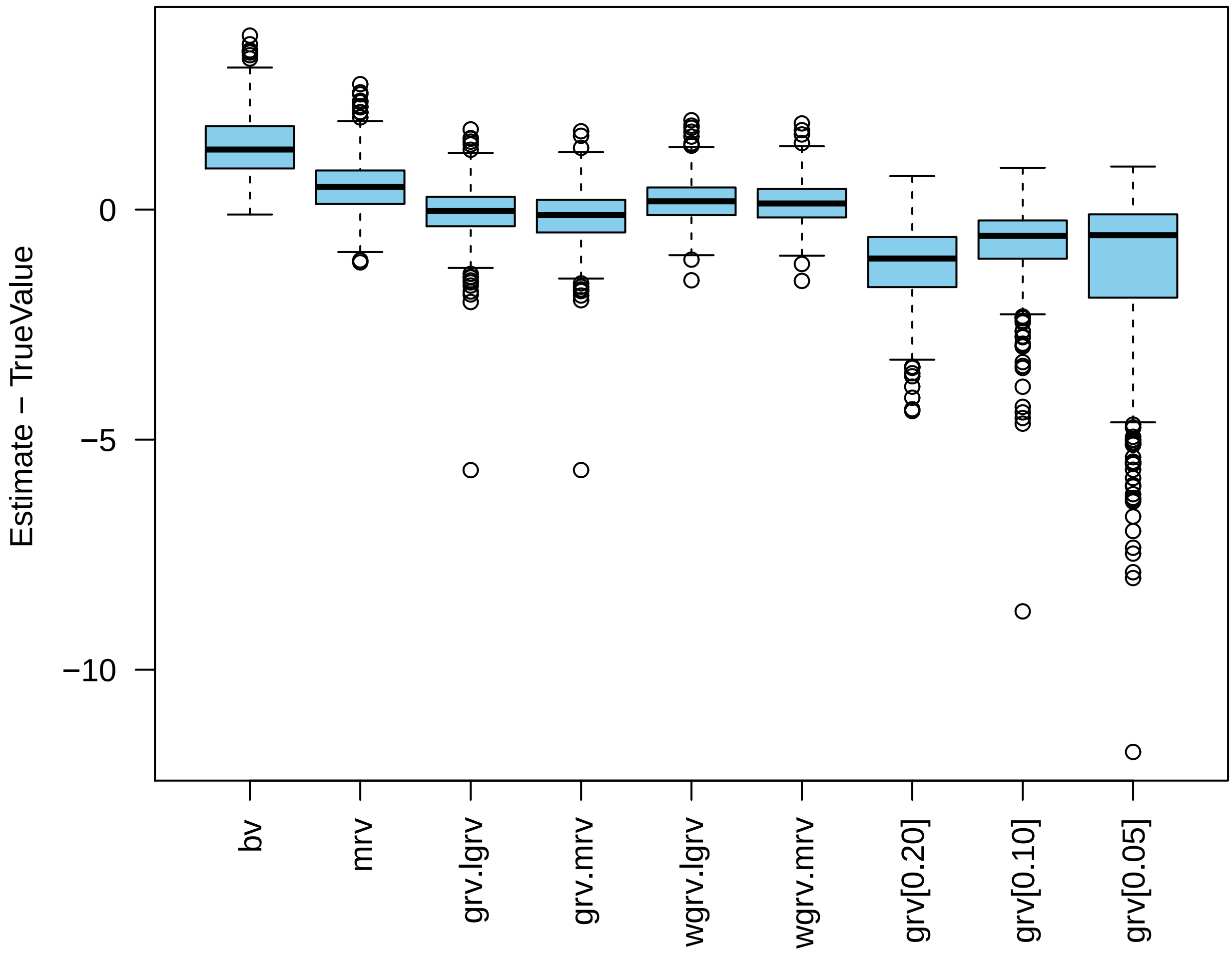


Figure 3.2: Estimation results for the case of high intensity: $\lambda = 30$

The case of moderate intensity: GRV with a shrinking cut-off ratio

Next, we consider the case of low intensity. In this case, we can use shrinking cut-off rate. Recall that the shrinking cut-off rate is defined by $\alpha_n = \lfloor Bn^{\delta_1} \rfloor / n$. In this simulation, we set $B = 10$ and $\delta_1 = 0.45$, so the cut-off rate is then $\alpha_n = 0.1525$.

The estimation results are shown in Figure 3.3. All global-filtering estimators perform well (for GRVs with fixed cut-off ratio, we set $\alpha = 0.2$ as before). These results suggest that if there are not so many jumps in the data, it would be advisable to use as many data as possible by making the cut-off ratio small.

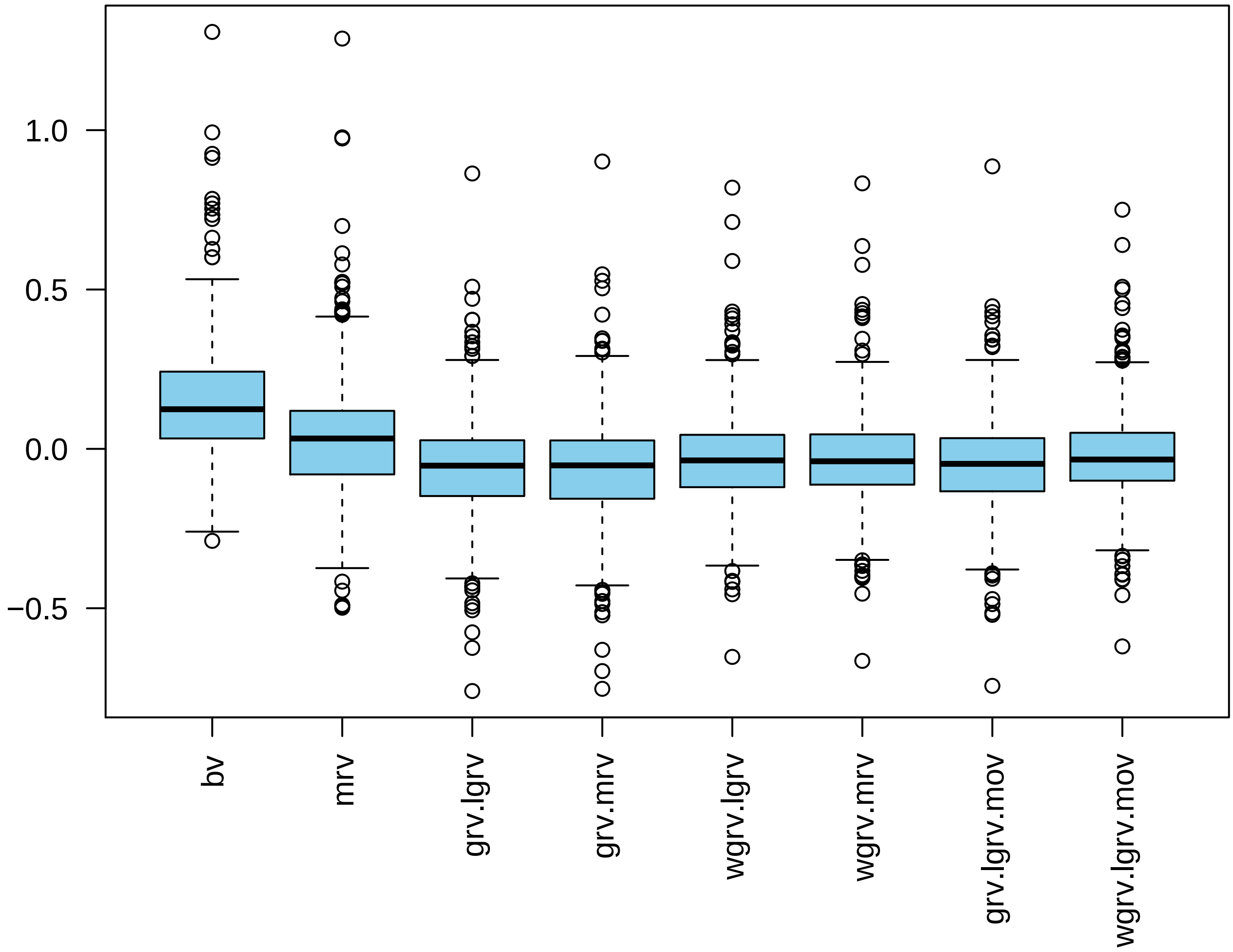


Figure 3.3: Estimation results for the case of low intensity: $\lambda = 5$

The case of constant volatility

Since we assumed that the volatility is location-dependent in the previous sections, the normalization by estimated spot volatilities is needed to obtain an accurate estimator. However, if the true volatility of data is constant, we may ignore normalization.

Here we set $\eta = 0$ so that the data is driven by a constant-volatility diffusion process. The intensity is $\lambda = 30$. Figure 3.4 shows the estimation results of this case. The GRVs without normalization (`grv[0.20]`, `grv[0.10]` and `grv[0.05]`) perform as well as those with normalization. This suggests that, if the true process can be thought as constant-volatility, we may skip normalization (calculation of spot volatilities) procedure.

However, it would be more typical that the volatility is non-constant. Thus, basically, it would be advisable to use normalization.

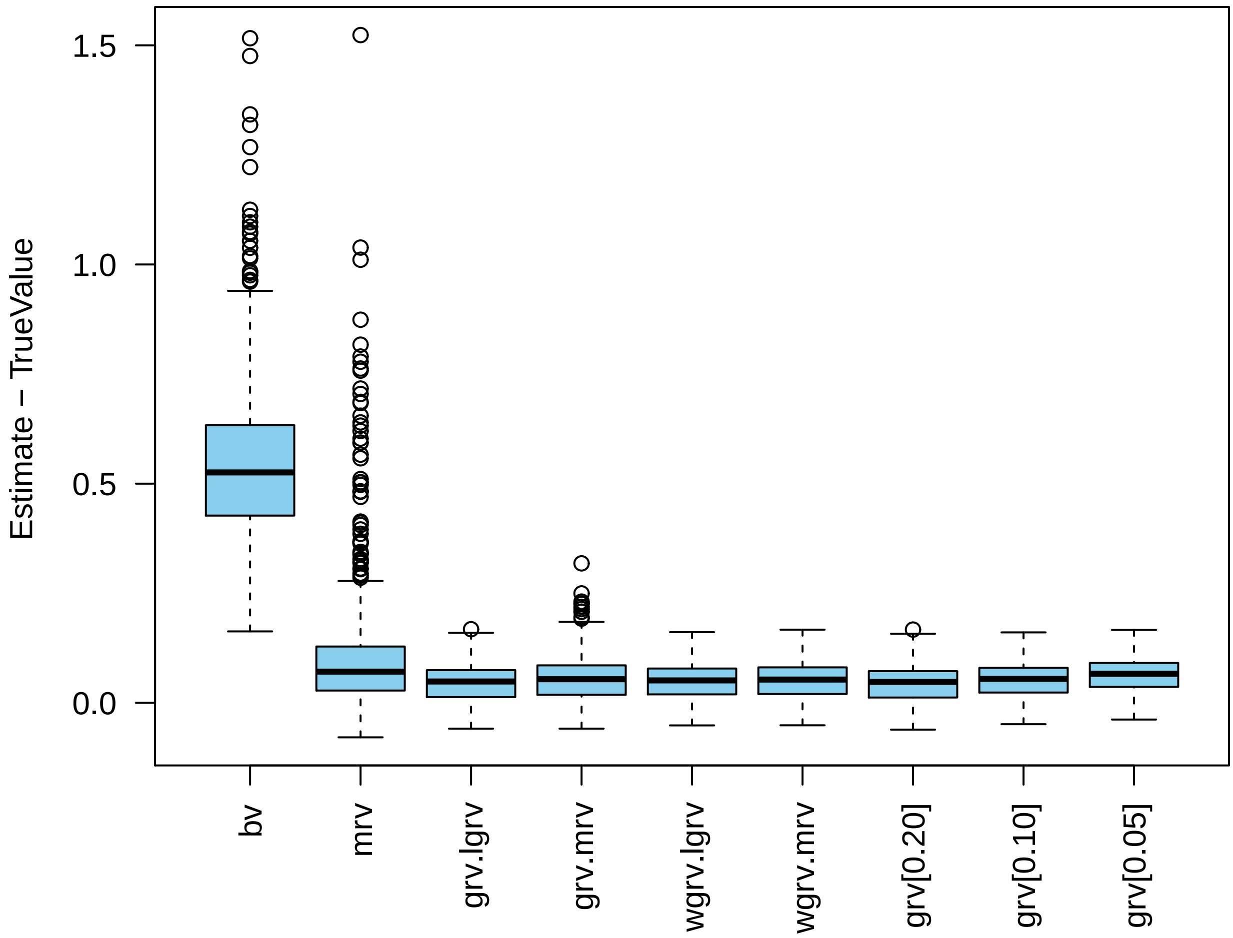


Figure 3.4: Estimation results for the constant volatility: $\lambda = 30$

3.7.2 The case of Neyman-Scott type clustering jumps

As the previous examples show, the minRV performs relatively well in the case of compound Poisson type jumps. However, even if the intensity of jumps is small, the minRV may suffer from an upward bias depending on the structure of jumps. In particular, if there are consecutive jumps (which is quite rare for compound Poisson processes), the minRV loses it advantage. Here we show an example of such a situation.

We consider the case that the data-generating process is given by $Y = X + J$, where X is the continuous part and J is the jump part. Here we assume that J follows a *Neyman-Scott clustering process* (NS hereafter), instead of a compound Poisson process.

NS process is a typical point process representing consecutive jumps. That is, there may be jumps within some consecutive intervals. This leads to upward bias of BV and minRV because the both of two adjacent increments can consist of large jumps. The NS process is constructed as follows.

(1) Set "centers" on the time interval $[0, T]$ by a Poisson process (N_t^0) with intensity λ_0. A center is defined as the point $t \in [0, T]$ which satisfies $\Delta N_t = 1$.

(2) For each center $c \in [0, T]$, choose the number N_c of "children," assuming N_c is Poisson-distributed with mean λ_c.

(3) For each center $c \in [0, T]$, generate independently and exponentially distributed random variables $\left(v_i^{(c)}\right)_{1 \le i \le N_c}$ with mean h. Then the location of child i derived from center c is defined as $c - v_i^{(c)}$. This defines the location of a jump.

(4) For each child i, generate an independently and normally distributed random variable $\xi_i \sim N(0, \nu_J^2)$. This determines the size and direction of a jump ΔJ_s.

(5) The NS process is defined as $J_t = \sum_{s \in [0,t]} \Delta J_s$.

We set the data-generating $Y = X + J$, where X is the Brownian semimartingale satisfying the following SDE: Here we consider the following process:

$$dX_t = \theta X_t dt + (\sigma + \eta X_t^2)^{\frac{1}{4}} dw_t, \tag{3.7.2}$$

We set $\lambda_0 = \lambda_c = 5$ and $\nu_J = 0.5$. For the continuous part X, we use $\theta = 0.2,\ \sigma = 1,\ \eta = 3$.

Figure 3.5 show the estimation results in the case of NS jumps. Because of the possible consecutive jumps, both bipower variation and minRV have upward bias, whereas GRV and WRGV are all robust to such clustering jumps. This suggests that the GRV and WGRV perform very well for various structures of jumps.

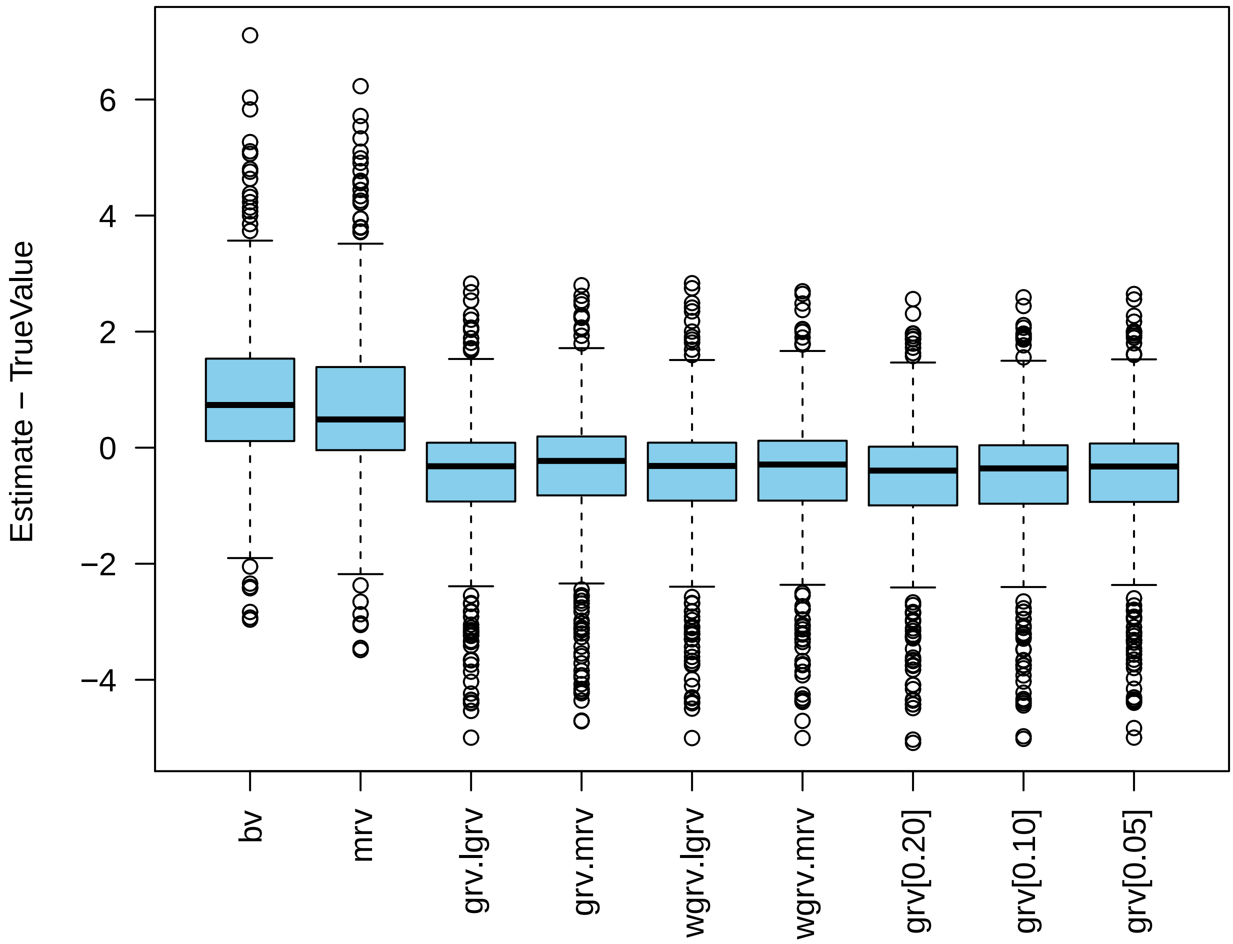


Figure 3.5: Estimation results for the case of Neyman-Scott clustering jumps

3.8 Concluding remarks

In this paper, we construct the global realized volatility estimator in the nonparametric context. We proved the consistency and the asymptotic normality of GRV and WGRV, and, by numerical simulations, we show that these new approaches outperform previous studies which use increments within a single or two intervals.

Our new approach for eliminating jumps is highly versatile. For example, by normalization, it works well when the volatility of data is driven by a nonconstant-volatility process. Moreover, both GRV and WGRV are accurate enough in the case of not only compound-Poisson sporadic jumps but also Neyman-Scott consecutive jumps.

The global-filtering method could be extended to the covariance estimation even under the non-synchronous sampling scheme. Furthermore, this approach could also be applied to construct a test statistic for jump. Also, it is valuable to apply our approach to empirical research of high-frequency time series data. These are important topics for future research.

Bibliography

[1] Andersen, T.G., Dobrev, D., Schaumburg, E.: Jump-robust volatility estimation using nearest neighbor truncation. J. Econometrics **169**(1), 75–93 (2012). DOI 10.1016/j.jeconom.2012.01.011

[2] Barndorff-Nielsen, O.E., Shephard, N.: Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics **2**, 1–48 (2004)

[3] Barndorff-Nielsen, O.E., Shephard, N., Winkel, M.: Limit theorems for multipower variation in the presence of jumps **116**(5), 796–806 (2006). DOI 10.1016/j.spa.2006.01.007

[4] Dohnal, G.: On estimating the diffusion coefficient. J. Appl. Probab. **24**(1), 105–114 (1987)

[5] Genon-Catalot, V., Jacod, J.: On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Ann. Inst. H. Poincaré Probab. Statist. **29**(1), 119–151 (1993)

[6] Hayashi, T., Yoshida, N.: Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. Ann. Inst. Statist. Math. **60**(2), 367–406 (2008). DOI 10.1007/s10463-007-0138-0. URL http://dx.doi.org/10.1007/s10463-007-0138-0

[7] Iacus, S.M., Yoshida, N.: Simulation and inference for stochastic processes with YUIMA. Springer (2018)

[8] Inatsugu, H., Yoshida, N.: Global jump filters and quasi-likelihood analysis for volatility. Forthcoming in Annals of the Institute of Statistical Mathematics.

[9] Kamatani, K., Uchida, M.: Hybrid multi-step estimators for stochastic differential equations based on sampled data. Statistical Inference for Stochastic Processes **18**(2), 177–204 (2014)

[10] Kessler, M.: Estimation of an ergodic diffusion from discrete observations. Scand. J. Statist. **24**(2), 211–229 (1997)

[11] Koike, Y.: An estimator for the cumulative co-volatility of asynchronously observed semimartingales with jumps. Scandinavian Journal of Statistics **41**(2), 460–481 (2014)

[12] Mancini, C.: Disentangling the jumps of the diffusion in a geometric jumping brownian motion **64**(1), 19–47 (2001)

[13] Mancini, C.: Non- parametric threshold estimation for models with stochastic diffusion coefficient and jumps **36**(2), 270–296 (2006). DOI 10.1111/j.1467-9469.2008.00622.x

[14] Ogihara, T., Yoshida, N.: Quasi-likelihood analysis for the stochastic differential equation with jumps. Stat. Inference Stoch. Process. **14**(3), 189–229 (2011). DOI 10.1007/s11203-011-9057-z. URL http://dx.doi.org/10.1007/s11203-011-9057-z

[15] Ogihara, T., Yoshida, N.: Quasi-likelihood analysis for nonsynchronously observed diffusion processes. Stochastic Processes and their Applications **124**(9), 2954–3008 (2014)

[16] Prakasa Rao, B.: Statistical inference from sampled data for stochastic processes. Statistical inference from stochastic processes (Ithaca, NY, 1987) **80**, 249–284 (1988)

[17] Prakasa Rao, B.L.S.: Asymptotic theory for nonlinear least squares estimator for diffusion processes. Math. Operationsforsch. Statist. Ser. Statist. **14**(2), 195–209 (1983)

[18] Protter, P.: Stochastic integration and differential equations. Springer-Verlag (1990)

[19] Shimizu, Y.: A practical inference for discretely observed jump-diffusions from finite samples. J. Japan Statist. Soc. **38**(3), 391–413 (2008)

[20] Shimizu, Y., Yoshida, N.: Estimation of parameters for diffusion processes with jumps from discrete observations. Stat. Inference Stoch. Process. **9**(3), 227–277 (2006). DOI 10.1007/s11203-005-8114-x. URL http://dx.doi.org/10.1007/s11203-005-8114-x

[21] Uchida, M., Yoshida, N.: Adaptive estimation of an ergodic diffusion process based on sampled data. Stochastic Process. Appl. **122**(8), 2885–2924 (2012). DOI 10.1016/j.spa.2012.04.001. URL http://dx.doi.org/10.1016/j.spa.2012.04.001

[22] Uchida, M., Yoshida, N.: Quasi likelihood analysis of volatility and nondegeneracy of statistical random field. Stochastic Processes and their Applications **123**(7), 2851–2876 (2013)

[23] Uchida, M., Yoshida, N.: Adaptive bayes type estimators of ergodic diffusion processes from discrete observations. Statistical Inference for Stochastic Processes **17**(2), 181–219 (2014)

[24] Yoshida, N.: Estimation for diffusion processes from discrete observation. J. Multivariate Anal. **41**(2), 220–242 (1992)

[25] Yoshida, N.: Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. Ann. Inst. Statist. Math. **63**(3), 431–479 (2011). DOI 10.1007/s10463-009-0263-z. URL http://dx.doi.org/10.1007/s10463-009-0263-z