Global jump filters and quasi-likelihood analysis for volatility

Haruhiko Inatsugu and Nakahiro Yoshida

Graduate School of Mathematical Sciences, University of Tokyo
Japan Science and Technology Agency CREST

Supplementary materials

6 Simulation Studies

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

$$dX_t = -\eta X_t dt + \sigma dw_t + dJ_t \qquad (t \in [0,1]) \tag{36}$$

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 $(X_{t_i^n})_{i=0,1,\dots,n}$.

There are already several parametric estimation methods for stochastic differential equations with jumps. Among them, Shimizu and Yoshida (2006) 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

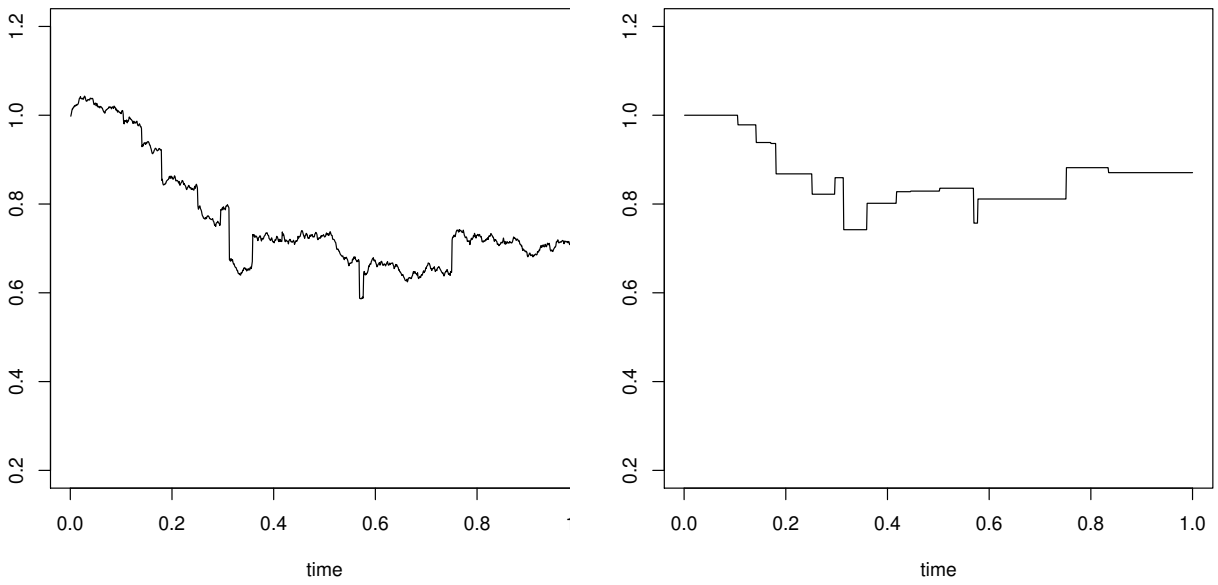
$$l_n(\sigma) = \sum_{i=1}^{n} \left[-\frac{1}{2\sigma^2 h_n} |\bar{X}_i^n|^2 - \frac{1}{2} \log \sigma^2 \right] \mathbf{1}_{\{|\Delta X_i| \le h_n^\rho\}},$$

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 (2006) 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 (2006). In their original paper, they assume that the jump size must be bounded away from zero. Ogihara and Yoshida (2011) 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.



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

Fig. 1 Sample paths of X and its jump part

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 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 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 (2006).

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 (3), 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 (16).

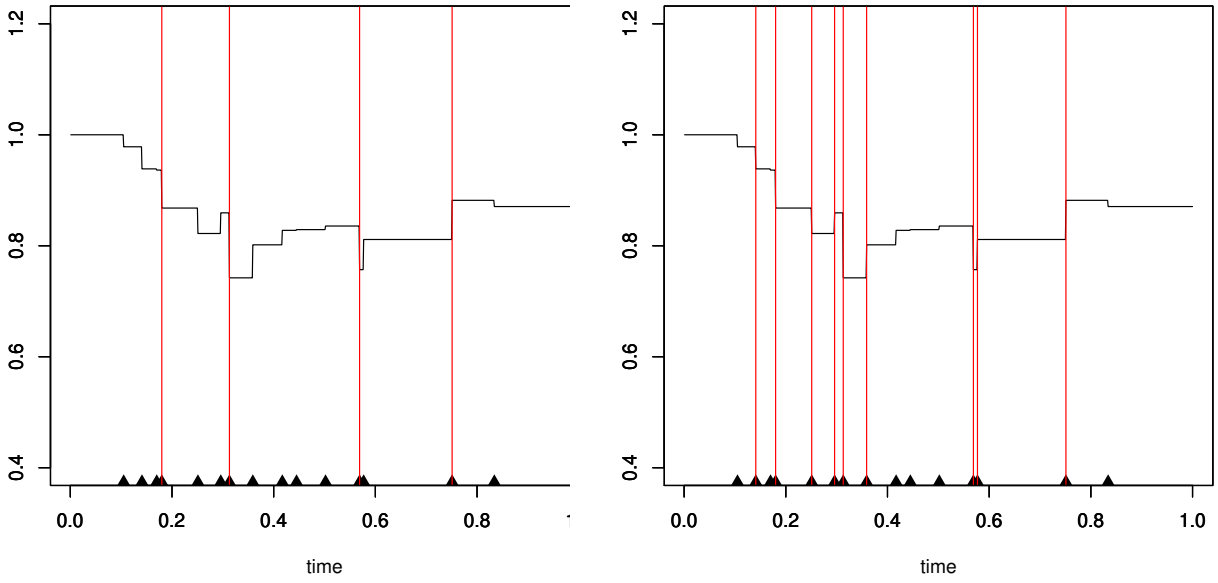
Figure 1 shows a sample path of (X, J). The left panel is the sample path of X and the right panel is its jump part J. Note that the jump part is not observable and thus we need to discriminate the jump from the sample path of X.

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.

6.2.1 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 (2006) 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



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

Fig. 2 Results of jump detection by local threshold method

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 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 (2006), 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.

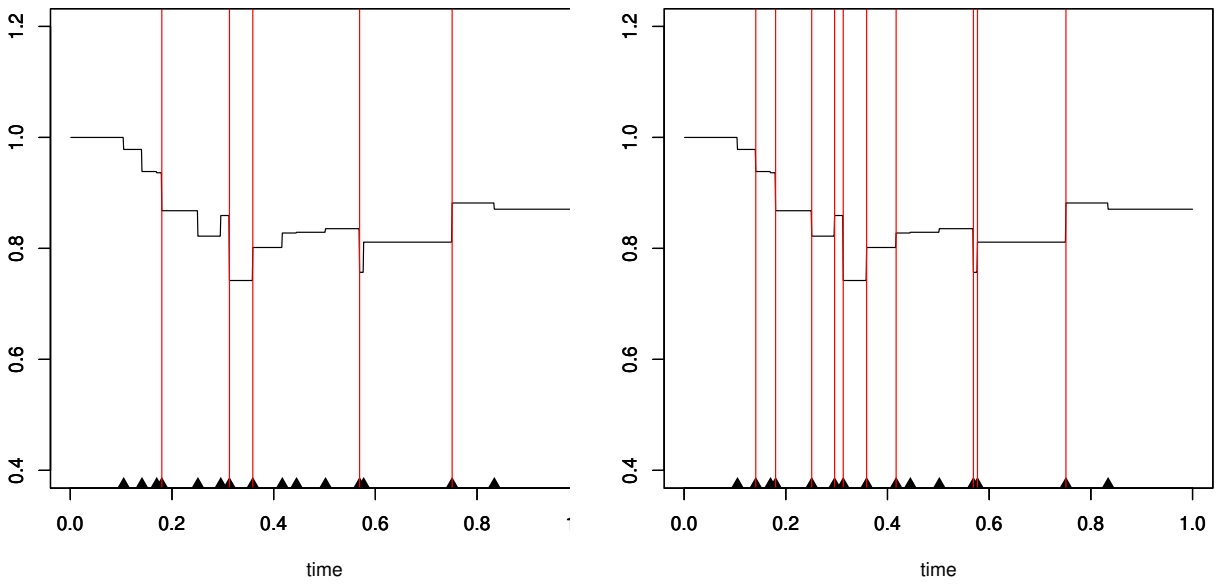
6.2.2 Global threshold method

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 $\alpha = 0.005, 0.010, 0.020, 0.050$.

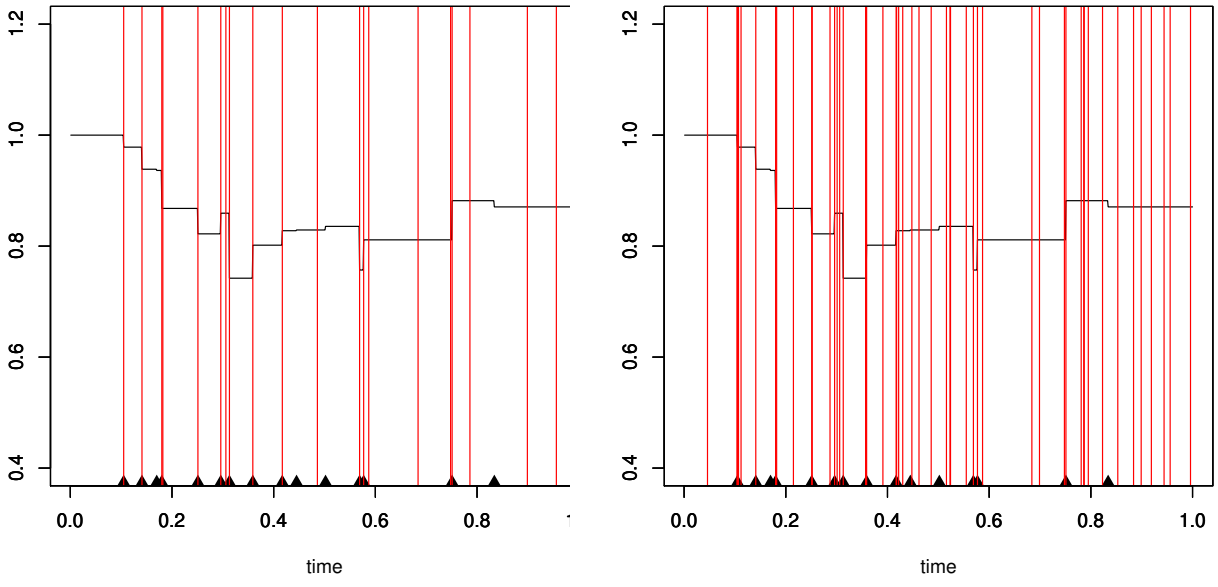
From the figures, we see that the too small α cannot detect jumps sufficiently, mistakenly judging some genuine jumps as increments driven by the continuous part, which is similar to the case of small ρ of the Shimizu-Yoshida estimator. By setting α a little larger, the accuracy of jump detection increases, 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 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 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



(a) $\alpha = 0.005$ (b) $\alpha = 0.010$



(c) $\alpha = 0.020$ (d) $\alpha = 0.050$

Fig. 3 Results of jump detection by global threshold method

Table 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

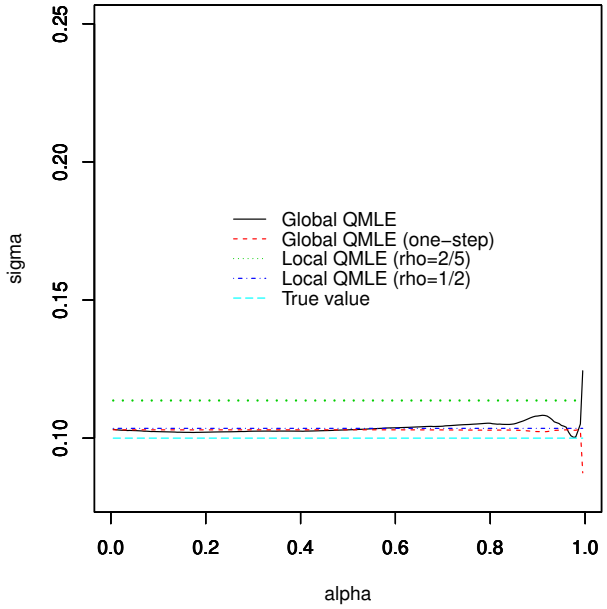


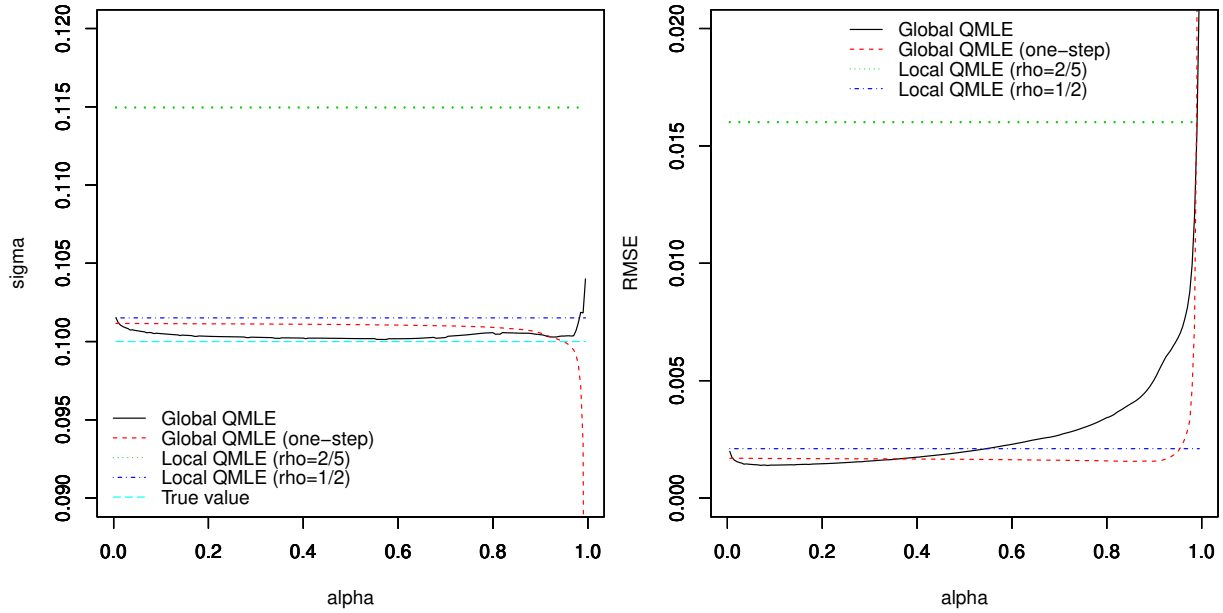
Fig. 4 Comparison of estimators given a sample path

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.

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 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 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 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 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.



(a) Average estimates (b) RMSE

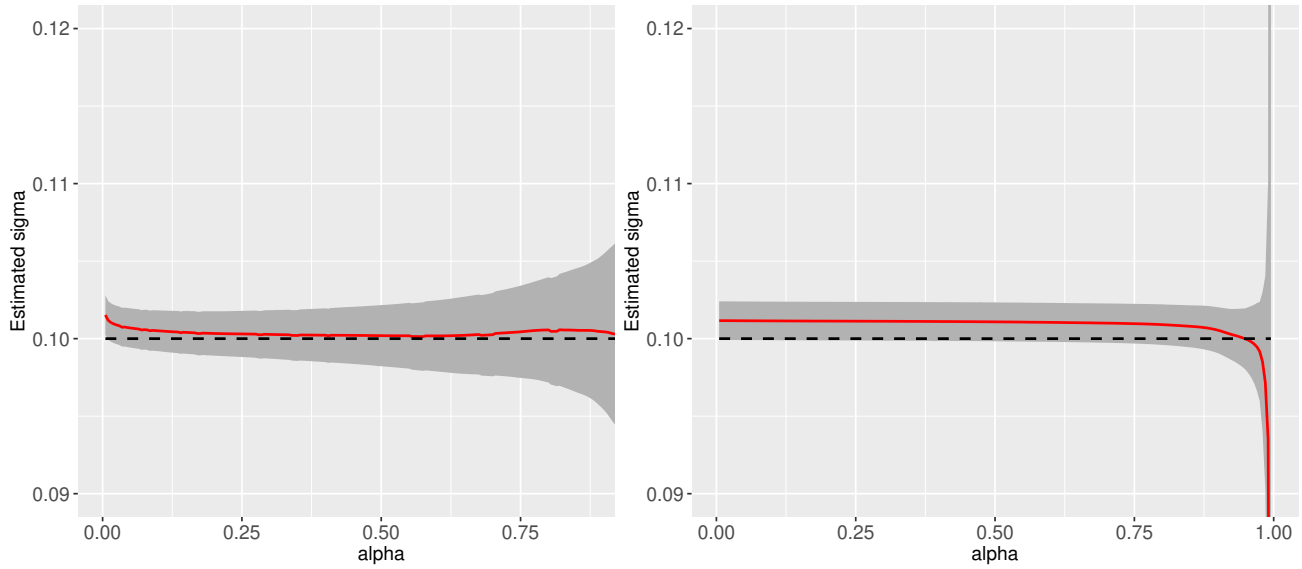
Fig. 5 Results of jump detection by local threshold method: comparison of averaged results

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 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 5 (b), RMSEs of the global estimators are smaller than those of the local estimators, except for the extreme choices of α.

Figure 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 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



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

Fig. 6 Estimation results of global QMLE estimator with standard error band

$(10^{-5})^{1/2} \times 1.64 \sim 0.00519$. Some of jumps may not be detected by the local filter, since its critical value is not so small, compared with $\epsilon = 0.05$.

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

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

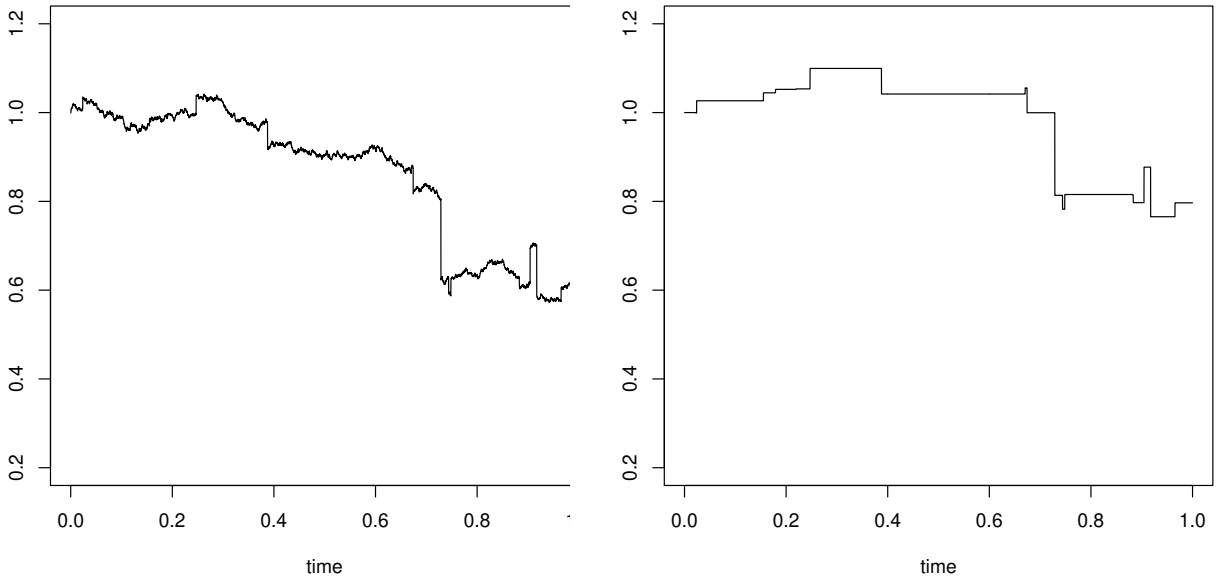
where $\mu = -0.01$ and $\varepsilon = 0.05$. In this setting, as shown in Figure 7, negative jumps appear more frequently than positive ones.

As Figure 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 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.

6.5 Location-dependent diffusion coefficient

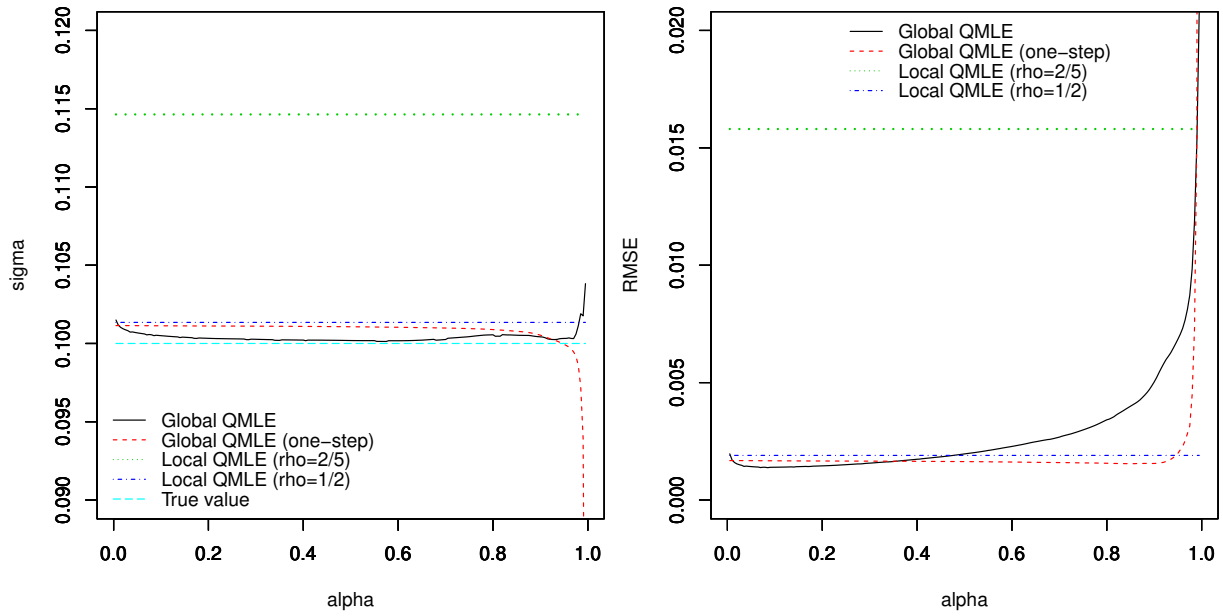
Here we assume that the diffusion coefficient is given by $\sigma\sqrt{1+x^2}$, where σ is an unknown positive parameter to be estimated. Other settings are entirely the same as those given in the Section 6.1. In particular, we assume that the distribution of jump size is centered, contrary to the previous 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 9. Like in the case of constant coefficient, the global estimators perform



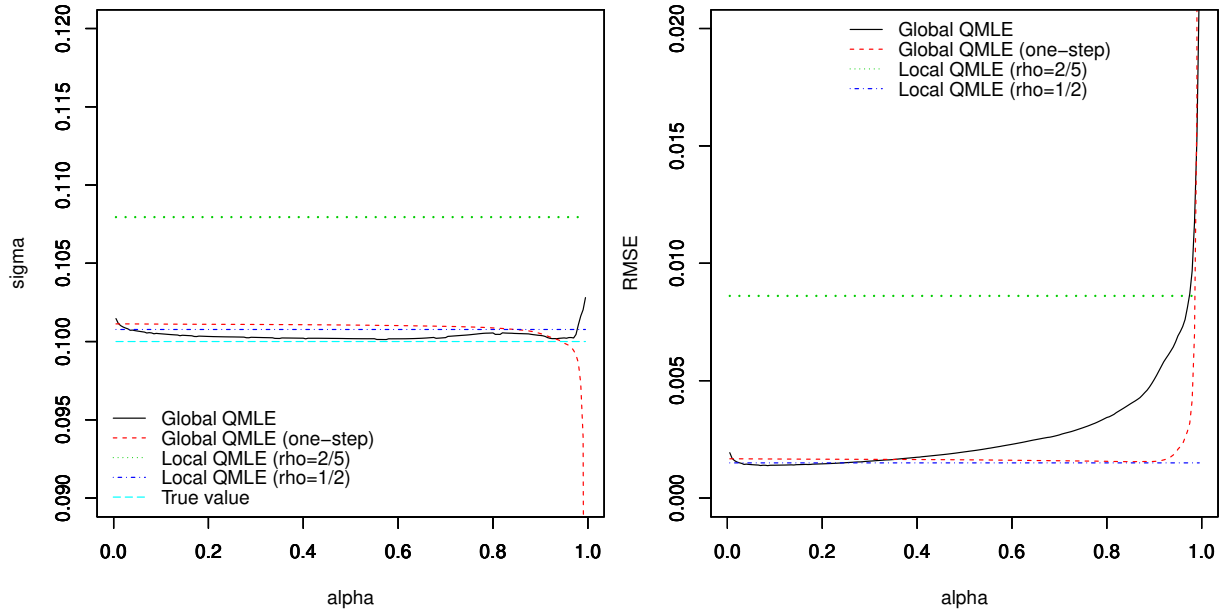
(a) Sample path of X (b) Sample path of the jump part J of X

Fig. 7 Sample paths of X and its jump part: in the case of asymmetric jump distribution



(a) Average estimates (b) RMSE

Fig. 8 Results of jump detection: in the case of asymmetric jump disribution



(a) Average estimates (b) RMSE

Fig. 9 Results of jump detection: a location-dependent diffusion coefficient

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.

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 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 filtering methods can apply to the realized volatility to estimate the integrated volatility. The superiority of the global filter to the several existing filtering methods used in this context is numerically observed as well as a mathematical proof. For details, see the forthcoming paper by the authors.

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.