Appl. Math. J. Chinese Univ.
2019, 34(3): 309-325

Bias Free Threshold Estimation for Jump Intensity Function

LIN Yi-wei[1] LI Zhen-wei[2] SONG Yu-ping[2,*]

Abstract. In this paper, combining the threshold technique, we reconstruct Nadaraya-Watson estimation using Gamma asymmetric kernels for the unknown jump intensity function of a diffusion process with finite activity jumps. Under mild conditions, we obtain the asymptotic normality for the proposed estimator. Moreover, we have verified the better finite-sampling properties such as bias correction and efficiency gains of the underlying estimator compared with other nonparametric estimators through a Monte Carlo experiment.

§1 Introduction

The diffusion process with jumps can characterize the statistical and economical dynamics of the underlying state variables such as asset prices or interest rates, which is defined as follows:

$$dX_t = \mu(X_{t-})dt + \sigma(X_{t-})dW_t + dJ_t, \qquad t \in [0,T], \tag{1}$$

where J_t is a finite activity (FA) pure jump semimartingale. J_t is usually assumed to be a compound Poisson process and can be written as

$$J_t = \int_0^t \int_{\mathscr{R}} x \cdot m(ds, dx) := \sum_{i=1}^{N_t} \gamma_i,$$

where m is the jump random measure of J_t, $N_t := \int_0^t \int_{\mathscr{R}} 1 \cdot m(ds, dx)$ is a a.s. finite Poisson process with a jump intensity stochastic process $\lambda(\cdot)$, and each γ_i is the size of the jump, more details in Cont and Tankov (2004). Many statisticians and economists focused on the theoretical properties and empirical applications for models (1). Hanif (2012) considered local

Received: 2018-09-16. Revised:2019-05-04.
MR Subject Classification: 62G20, 62M05, 60J75.
Keywords: Jump-diffusion model, Nonparametric estimation, Gamma asymmetric kernel.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-019-3630-4.
LIN Yi-wei is supported by the National Natural Science Foundation of China (No.11701331), Shandong Provincial Natural Science Foundation (No. ZR2017QA007) and Young Scholars Program of Shandong University. SONG Yu-ping is supported by Ministry of Education, Humanities and Social Sciences project (No. 18YJCZH153), National Statistical Science Research Project (No. 2018LZ05), Youth Academic Backbone Cultivation Project of Shanghai Normal University (No. 310-AC7031-19-003021), General Research Fund of Shanghai Normal University (SK201720) and Key Subject of Quantitative Economics (No. 310-AC7031-19-004221) and Academic Innovation Team (No. 310-AC7031-19-004228) of Shanghai Normal University.
*Corresponding author.

linear estimator for the volatility function $\sigma^2(x)+\lambda(x)\sigma_J^2$, where λ and σ_J^2 are the jump intensity and the variance of J_t respectively. Moreover, Lin and Wang (2010) or Hanif et al. (2012) improved the nonparametric estimation for the volatility. The description of jumping behaviors can help explain the sharp fluctuations for the underlying assets, whereas less research is done independently for the characteristic of jumps, especially the parameter $\lambda(x)$ which characterizes the frequency of jumps. In view of an economical point, how to effectively estimate this intensity parameter for a jump-diffusion process has also been a fundamental problem.

Assuming the parametric design for the diffusion model with jumps (1) such as one-factor CKLS model including jumps, Das (2002) and Andersen et al. (2004) considered the characteristic of jump through the parametric estimators. However, due to complex structure of data in the analysis process, people often can not make simple assumptions about the overall distribution pattern. At this time, the method of parameter estimation is no longer applicable. Nonparametric approach does not rely on the assumptions of overall distribution or some of the overall parameter characteristics such as variance, which, to some extent, can reduce the errors caused by errors from misspecified model. Based on the fourth and sixth infinitesimal conditional moments of jump-diffusion process, Johannes (2004) and Bandi and Nguyen (2003) constructed nonparametric estimator for the identify parameter $\lambda(x)$. However, they didn't obtain the central limit theorem for the underlying estimator. Mancini and Renò (2011) improved the nonparametric estimator for intensity function $\lambda(x)$ combining Nadaraya-Watson method with the threshold function and proved the central limit theorem for it.

As is known to all, there exists a "boundary effect" problem for Nadaraya-Watson estimator at the boundary point. Chen (2000) proposed the Gamma asymmetric kernels for nonparametric estimation of positive supported densities and also mentioned that the kernel estimator constructed with gamma probability densities were free of boundary bias. In addition, the variance of nonparametric estimator constructed with gamma asymmetric kernels is inversely proportional to the location of the design point x away from the origin, which could bring about variance reduction. Gospodinov and Hirukawa (2012) proposed nonparametric estimation constructed with gamma asymmetric kernels for scalar diffusion models with application to bond and option pricing using U.S. interest rates. They also proved the asymptotic theorem for this approach and illustrated its advantages such as bias correction and efficiency gains through Monte Carlo experiments. In this paper, we introduce and analyze the bias free nonparametric threshold estimator as Nadaraya-Watson estimator constructed with Gamma asymmetric kernels for the unknown intensity coefficients in the jump-diffusion model. The main contributions of this paper are twofold. Theoretically, different from Song and Wang (2018), the asymptotic normality for the proposed estimator which is a stochastic integral driven by a pure jump Lévy martingale not a Wiener process, is obtained by Jacod's stable convergence theorem. Practically, compared with the estimator constructed with Gaussian kernels considered in Mancini and Renò (2011), the better finite-sample performance of our estimators, especially bias correction at the boundary points, is verified through a Monte Carlo experiment.

An outline of the paper is as follows: Section 2 introduces the bias free threshold estimator constructed with Gamma asymmetric kernels for jump intensity function. Moreover, the regular conditions and main results are shown here. Section 3 presents the finite-sample performance of the underlying estimator through Monte-Carlo simulation study. Section 4 collects the technical lemmas and the detailed proof for the main results.

§2 Bias free threshold estimator and Main results

Chen (2000) introduced nonparametric kernel estimation for positive supported densities constructed with Gamma asymmetric kernels which are defined as

$$K_{G(x/h_n+1,h_n)}(u) = \frac{u^{x/h_n}\exp(-u/h_n)}{{h_n}^{x/h_n+1}\Gamma(x/h_n+1)}, \quad 0 \le u \le \infty, \tag{2}$$

where $\Gamma(m) = \int_0^\infty y^{m-1}\exp(-y)dy,\ m > 0$ is the Gamma function and h_n is the smoothing parameter. The local nonparametric estimator using Gamma kernel functions does not bring about boundary bias for nonnegative variables or nonnegative part of underlying variables due to the fact that the support of the Gamma density function is $[0, \infty)$. Since $K_{G(x/h_n,h_n)}(u)$ is unbounded near at $x = 0$, we use modified Gamma kernel functions $K_{G(x/h_n+1,h_n)}(u)$ instead of $K_{G(x/h_n,h_n)}(u)$. The Gamma function has shapes varying with the smoothing parameter and the design point x, which behaves as the variable bandwidth method. Furthermore, the asymptotic variance of the nonparametric Gamma kernel estimation is inversely proportional to the design point x, which yields the variance reduction of nonnegative kernels.

Here we firstly consider Nadaraya-Watson estimation for the unknown coefficient of the diffusion model with jumps (1) constructed with the Gamma asymmetric kernels. The Nadaraya-Watson estimator $\hat{\lambda}_{AS}(x)$ for $\lambda(x)$ constructed with the Gamma asymmetric kernels is

$$\hat{\lambda}_{AS}(x) = \frac{\sum_{i=1}^{n} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})c_{i,n}I_{\{|X_{i\Delta_n}-X_{(i-1)\Delta_n}|^2>\vartheta(\Delta_n)\}}}{\Delta_n\sum_{i=1}^{n} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})}, \tag{3}$$

where $\{X_{i\Delta_n}; i = 1, 2, \cdots\}$ are discrete sampling observations for the process (1), $c_{i,n}$ is a double array of constants with $i = 1, ..., n$ and $\vartheta(\Delta_n)$ is the threshold function.

We impose the following assumptions throughout the paper. Assume that $\mathscr{D} = (l, u)$ with $-\infty \le l < u \le \infty$ is the range of the process X_t.

Assumption 1. *For model (1), the jump intensity $\lambda_t := \lambda(X_{t-})$ is bounded and nonnegative, and the coefficients μ_t and σ_t are twice continuously differentiable and progressively measurable processes with càdlàg paths satisfying the following conditions:*

(i) For each $n \in \mathbb{N}$, there exists a positive constant L_n such that for any $|x| \le n$, $|y| \le n$,

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \le L_n|x - y|;$$

(ii) There exists a positive constant C, such that for all $x \in \mathbb{R}$,

$$|\mu(x)| + |\sigma(x)| \le C(1 + |x|);$$

(iii) $\sigma^2(x)$ is strictly positive.

Assumption 2. *The solution of model* (1) *is Harris recurrent.*

Assumption 3. *The local time $L_X(T, x)$ for model (1) satisfies*

$$\sup_{|u|\le 1} |L_X(T, x + ua_n) - L_X| = o_p(L_X(T, x)),$$

as $a_n \to 0$.

Remark 2.1. The assumptions (i) and (ii) in Assumption 1 guarantee that the SDE (1) has a unique strong solution which is adapted and right continuous with left limits on $[0,T]$, see Jacod and Shiryaev (2003) for more details. The Assumption 2 guarantees the existence of a unique invariant measure $s(x)$, that is, $s(A)=\int_{\mathscr{D}} P(X_t^{(x)}\in A)s(dx),\quad \forall A\in\mathfrak{B}(\mathscr{D})$. The Assumption 2 with positive Harris recurrent condition implies that the process X_t has a time-invariant probability measure given by $p(dx)=\frac{s(x)}{s(\mathscr{D})}$ at any initial level $x\in\mathscr{D}$. Assumption 3 actually arises from Theorem 1 in Eisenbaum and Kaspi (2007), which was mentioned as assumption 5 in Wang and Zhou (2017).

In the following theorem, we will obtain the corresponding asymptotic normality of Nadaraya-Watson threshold estimator constructed with Gamma asymmetric kernels (3) by letting n, $T\to\infty$ and $\Delta_n = T/n\to 0$.

Theorem 1. ***[recurrent case]*** *In model* (1)*, under Assumptions 1, 2 and 3, and we also assume that*

(i) J_t *is such that* $\forall\varepsilon>0$, $P\{|\gamma_i|<\varepsilon\}\le c\varepsilon$ *and the jump sizes* $\{\gamma_i\}_i$ *are independent of* N_t*;*

(ii) $\vartheta(\Delta_n)=\Delta_n^{\eta}, \eta\in(0,1)$*, with* $n\Delta_n^{1+\eta/2}\to 0$ *and* $\frac{\Delta_n\ln(\frac{1}{\Delta_n})}{\vartheta(\Delta_n)}\to 0$*;*

(iii) *the bandwidth parameter is of the form* $h_n=\Delta_n^{\phi}$ *with* $\phi\in(0,\eta/2)$*;*

(iv) *as* n*,* $T\to\infty$ *and* h_n*,* $\Delta_n\to 0$*, we have for all visited* x

$$\frac{(\Delta_n \ln(\frac{1}{\Delta_n}))^{\frac12}}{h_n^2}\to 0,\qquad \frac{(\Delta_n\ln\frac{1}{\Delta_n})^{\frac12}}{h_n}\hat{L}_T^{\sharp}(x)\xrightarrow{a.s.}0,\quad h_n\hat{L}_T^{\sharp}(x)\xrightarrow{a.s.}\infty,$$

but $h_n^{\frac52}\hat{L}_T^{\sharp}(x)=O_P(1)$*, where* $\hat{L}_T^{\sharp}(x)=\Delta_n\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})$.

Then if $c_{i,n}$ *is a double array of constants with* $i=1,...,n$ *such that* $\forall x$*,* $\sqrt{h_n^{1/2}\hat{L}_T^{\sharp}(x)}\sup_i|1-c_{i,n}|\to 0$ *as* $n\to\infty$*, then, for the interior* x *visited by* X*, we have*

$$\sqrt{h_n^{1/2}\hat{L}_T^{\sharp}(x)}\left(\hat{\lambda}_{AS}(x)-\lambda(x)-\left[\lambda^{'}(x)\left\{1+x\frac{s^{'}(x)}{s(x)}\right\}+\frac{x}{2}\lambda^{''}(x)\right]\cdot h_n\right)\xrightarrow{d}$$
$$\mathscr{N}\left(0,\lambda(x)\cdot\frac{1}{2\sqrt{\pi}x^{1/2}}\right);$$

for the boundary x *visited by* X*, we have*

$$\sqrt{h_n\hat{L}_T^{\sharp}(x)}\left(\hat{\lambda}_{AS}(x)-\lambda(x)\right)\xrightarrow{d}\mathscr{N}\left(0,\lambda(x)\cdot\frac{\Gamma(2\kappa+1)}{2^{1+2\kappa}\Gamma^2(\kappa+1)}\right),$$

where "interior x*" if "*$x/h_n\longrightarrow\infty$*" or "boundary* x*" if "*$x/h_n\longrightarrow\kappa$*" for* $x\in\mathscr{D}$.

Theorem 2. ***[positive recurrent case]*** *Under Assumptions 1, 2 with positive Harris recurrent condition, 3 and* $n\Delta_n h_n^{\frac52}=O_P(1)$*, then, for the interior* x *visited by* X*, we can obtain*

$$\sqrt{n\Delta_n h_n^{1/2}}\left(\hat{\lambda}_{AS}(x)-\lambda(x)-\left[\lambda^{'}(x)\left\{1+x\frac{s^{'}(x)}{s(x)}\right\}+\frac{x}{2}\lambda^{''}(x)\right]\cdot h_n\right)\xrightarrow{d}$$
$$\mathscr{N}\left(0,\lambda(x)\cdot\frac{1}{2\sqrt{\pi}x^{1/2}p(x)}\right);$$

for the boundary x *visited by* X*, we have*

$$\sqrt{n\Delta_n h_n}\left(\hat{\lambda}_{AS}(x)-\lambda(x)\right)\xrightarrow{d}\mathscr{N}\left(0,\lambda(x)\cdot\frac{\Gamma(2\kappa+1)}{2^{1+2\kappa}\Gamma^2(\kappa+1)p(x)}\right).$$

Remark 2.2. For the positive recurrent case of Assumption 2, the local time $\bar{L}_X(T,x)$ increases consistently with T as

$$\frac{\bar{L}_X(T,x)}{T} \xrightarrow{a.s.} p(x),\ \forall x \in \mathscr{D}. \tag{4}$$

So we can deduce Theorem 2 by means of Theorem 1 easily with the property (4) above.

Remark 2.3. Under some regular conditions and $h_n^5 \hat{L}_X(T,x) = O_{a.s.}(1)$, Mancini and Renò (2011) obtained asymptotic normality of Nadaraya-Watson estimator constructed with Gaussian symmetric kernels for the unknown quantity $\lambda(x)$,

$$\sqrt{h_n \hat{L}_X(T,x)} \left(\hat{\lambda}_{NW}(x) - \lambda(x) - h_n^2 \cdot \frac{1}{2} \left[\lambda^{''}(x) + \frac{\lambda^{'}(x) s^{'}(x)}{s(x)} \right] \right) \xrightarrow{d} \mathscr{N}\left(0, \lambda(x) \cdot \frac{1}{2\sqrt{\pi}}\right), \tag{5}$$

where

$$\hat{L}_X(T,x) = \frac{1}{h_n} \sum_{i=1}^n K\big(\frac{X_{(i-1)\Delta_n} - x}{h_n}\big)\Delta_n,$$

and

$$\hat{\lambda}_{NW}(x) = \frac{\sum_{i=1}^n K\big(\frac{X_{(i-1)\Delta_n} - x}{h_n}\big) c_{i,n} I_{\{|X_{i\Delta_n} - X_{(i-1)\Delta_n}|^2 > \vartheta(\Delta_n)\}}}{\sum_{i=1}^n K\big(\frac{X_{(i-1)\Delta_n} - x}{h_n}\big)\Delta_n}.$$

There are two main differences between the asymptotic result in Theorem 1 and that in Mancini and Renò (2011) above: on one hand, the convergence rate of Nadaraya-Watson estimator using Gamma asymmetric kernel is various for the location of the design point x such as "interior x" and "boundary x"; on the other hand, the variance of of Nadaraya-Watson estimator using Gamma asymmetric kernel for "interior x" is inversely proportional to the design x, which implies that the estimator (3) is resistant to sparse design point x. More theoretical comparison between the results, one can refer to Song et al. (2019) for more details.

It is very important to consider the choice of the bandwidth h_n for the nonparametric estimation. Based on the mean square error (MSE), the optimal bandwidth of Nadaraya-Watson threshold estimator with Gaussian kernels (5) for both interior and boundary points is given

$$h_{n,opt}^{NW} = \left(\frac{\lambda(x)}{\hat{L}_X(T,x) 2\sqrt{\pi} A^2(x)} \right)^{\frac{1}{5}} = O_P\left(L_X(T,x)^{-\frac{1}{5}} \right),$$

where $A(x) = \frac{1}{2}\left[\lambda^{''}(x) + \frac{\lambda^{'}(x)s^{'}(x)}{s(x)} \right]$. The optimal bandwidth of nonparametric threshold estimator constructed with Gamma asymmetric kernels (3) based on Theorem 1 for "interior x" is given

$$h_{n,opt}^{AS} = \left(\frac{\lambda(x)}{L_T^{\sharp}(x) 2\sqrt{\pi} x^{1/2} B^2(x)} \right)^{\frac{2}{5}} = O_P\left(L_X(T,x)^{-\frac{2}{5}} \right),$$

where $B(x) = \left[\lambda^{'}(x) \left\{ 1 + x\frac{s^{'}(x)}{s(x)} \right\} + \frac{x}{2}\lambda^{''}(x) \right]$.

We can observe that for "interior x", the optimal smoothing parameter

$$h_{n,opt}^{AS} = O_P\left(L_X(T,x)^{-\frac{2}{5}} \right) = O_P(h_{n,opt}^{NW})^2,$$

which shows that the asymptotic variance of the nonparametric threshold estimator constructed with Gamma asymmetric kernels is $O_P\left(L_X(T,x)^{-\frac{4}{5}} \right)$, the same as that constructed with symmetric kernels. For the further study of the theoretical optimal value of the bandwidth, one can refer to Aït-Sahalia and Park (2016) and Wang and Zhou (2017) for more details.

§3 Simulation study

In this section, the finite-sample performance of various nonparametric threshold estimators is constructed through a simple Monte Carlo simulation experiment. For simplicity, nonparametric threshold estimator using Gamma asymmetric kernels studied here is denoted as AS and Nadaraya-Watson estimator mentioned in Mancini and Renò (2011) is denoted as NW.

A jump-diffusion model defined as

$$dX_t = 0.103X_t dt + 0.178X_t dW_t + dJ_t, \tag{6}$$

is considered here, where $J_t = \sum_{i=1}^{N_t} Z_i$ with the arrival intensity $\lambda = 1$ and jump size $Z_n \sim \mathscr{N}(0, 0.036^2)$. One sample path of X_t with $T = 10,\ n = 480,\ X_0 = 0.1,\ \Delta_n = \frac{T}{n} = \frac{1}{48}$ for model (6) is shown in FIG 1. The difference of X_t and its time-varying threshold function $\vartheta(\Delta_n)$ depicted in FIG 2, which is implemented as that in Mancini and Renò (2011) and effectively disentangle the jumps.

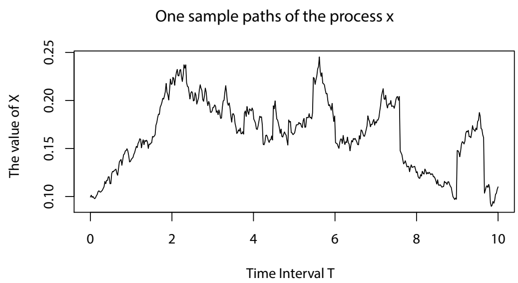

Figure 1: One sample path of X_t for model (6)

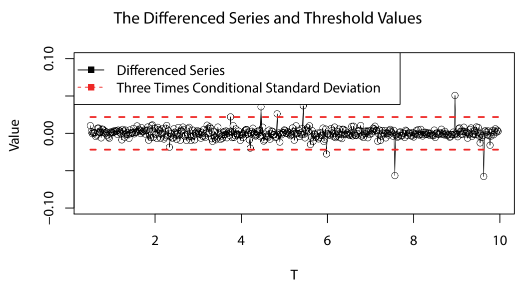

Figure 2: The Difference Sequence and its Threshold Value

Throughout this section, we take Gaussian density $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ and the practical bandwidth $h_n^{NW} = 2.8 \cdot \hat{S} \cdot T^{-\frac{1}{5}}$ for Nadaraya-Watson threshold estimator and $h_n^{AS} = 2.8 \cdot \hat{S} \cdot T^{-\frac{2}{5}}$ for nonparametric threshold estimator constructed with Gamma asymmetric kernels, where $\hat{S}$ denotes the standard deviation of the data. The sequence $c_{i,n}$ is implemented as that in Mancini and Renò (2011). The plot of estimators and their corresponding biases at various quantile points of sample X_t in model (6) are demonstrated in FIG 3 and Table 1. From FIG 3 and Table 1, it is observed that AS threshold estimator performs better than NW threshold estimator in terms of bias, especially at the right sparse design point.

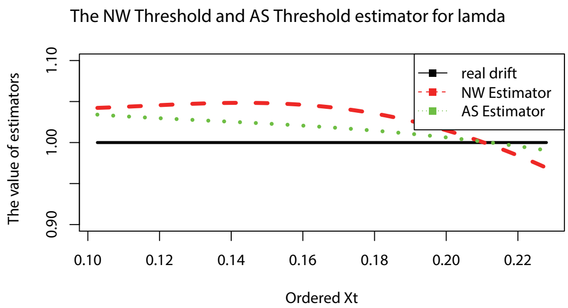

Figure 3: Various Nonparametric Threshold Estimators for $\lambda(x) = 1$

Table 1: The biases of NW and AS at various quantile points of sample X_t in model (6)

Bias	Quantile points of sample X_t								
	10%	20%	30%	40%	50%	60%	70%	80%	90%
NW	0.0446	0.0470	0.0484	0.0477	0.0441	0.0370	0.0261	0.0111	-0.0079
AS	0.0308	0.0280	0.0253	0.0225	0.0192	0.0153	0.0105	0.0049	-0.0018

Furthermore, the overall finite-sample performance between Nadaraya-Watson threshold estimator considered in Mancini and Renò (2011) and bias free nonparametric threshold estimator constructed with Gamma asymmetric kernels for the intensity function λ is calculated via the following four measures used in Fan et al. (2007).

Measure 1: Absolute Mean Error (AME): AME $= \frac{1}{N}\left|\sum_{k=1}^{N}[\hat{\lambda}(x_k) - \lambda(x_k)]\right|$;

Measure 2: Root Mean Square Error (RMSE): RMSE $= \sqrt{\frac{1}{N}\sum_{k=1}^{N}[\hat{\lambda}(x_k) - \lambda(x_k)]^2}$;

Measure 3: Ideal Mean Absolute Deviation Error (IMADE):

$$\text{IMADE} = \frac{1}{N}\sum_{k=1}^{N}\left|\hat{\lambda}(x_k) - \lambda(x_k)\right|;$$

Measure 4: Relative Ideal Mean Absolute Deviation Error (RIMADE):

$$RIMADE = \frac{1}{N}\sum_{k=1}^{N}\frac{\left|\hat{\lambda}(x_k) - \lambda(x_k)\right|}{\lambda(x_k)},$$

where $\hat{\lambda}(x)$ is the various nonparametric threshold estimator for $\lambda(x)$ and $\{x_k\}_1^N$ are chosen uniformly to cover the range of sample path of X_t. From Table 2, we can find that the AS estimator performs better than the NW estimator for various jump intensities and various measures.

Table 2: Comparisons between different nonparametric threshold estimators for intensity function with various measures

Measures	Estimators	$\lambda = 1$	$\lambda = 3$	$\lambda = 5$
AME	NW	0.1943	1.6557	2.8967
	AS	0.1883	1.6311	2.8520
RMSE	NW	0.1949	1.6571	2.8972
	AS	0.1903	1.6322	2.8554
IMADE	NW	0.1943	1.6557	2.8967
	AS	0.1883	1.6311	2.8520
RIMADE	NW	0.1943	0.5519	0.5793
	AS	0.1883	0.5437	0.5704

Finally, the QQ plots of nonparametric threshold estimator using Gamma asymmetric kernels for $\lambda(x) = 1$ at the left or right boundary point and in the interior point are displayed in FIG 4, which confirms the normality of the underlying nonparametric threshold estimator shown in Theorem 1.

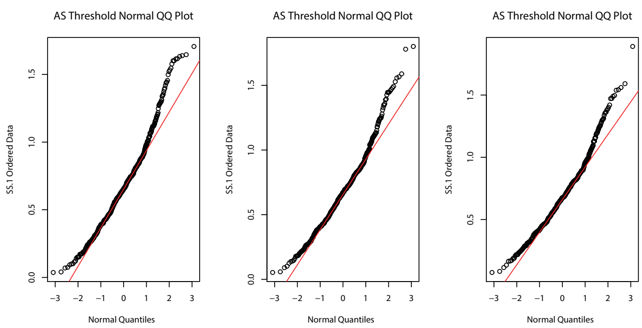

Figure 4: QQ plot of Nonparametric Threshold Estimators constructed with Gamma Asymmetric kernels for $\lambda(x) = 1$

§4 Proof for main result

We recall that $\Delta_n = \frac{T}{n}, t_i = i\Delta_n, \Delta_i X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$. Note that throughout this article, we use C to denote a generic constant, which may vary from line to line. By $\sigma.W$ we denote the stochastic integral of σ with respect to W. We denote by $\left(\tau_j\right)_{j\in\mathbb{N}}$ the jump instants of J_t and by $\tau^{(i)}$ the instant of the first jump in $(t_{i-1}, t_i]$, if $\Delta_i N \geq 1$.

Lemma 4.1. ***[The occupation time formula]*** *Let X_t be a semimartingale with local time $(L_X(\cdot,a))_{a\in\mathscr{D}}$. Let g be a bounded Borel measurable function. Then*

$$\int_{-\infty}^{\infty} L_X(t,a)g(a)da = \int_0^t g(X_{s-})d[X]_s^c, \quad a.s., \tag{7}$$

where $[X]^c$ is the continuous part of the quadratic variation of X and $\mathscr{D}$ denotes the the admissible range of the process of interest.

Lemma 4.2. ***[Jacod's stable convergence theorem]*** *A sequence of $\mathbb{R}-$valued variables $\{\zeta_{n,i} : i \geq 1\}$ defined on the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F})_{t\geq 0}, P)$ is $\mathscr{F}_{i\Delta_n}-$measurable for all n, i. Assume there exists a continuous adapted $\mathbb{R}-$valued process of finite variation B_t and a continuous adapted and increasing process C_t, for any $t > 0$, we have*

$$\sup_{0\leq s\leq t} \Big| \sum_{i=1}^{[s/\Delta_n]} E\big[\zeta_{n,i} \mid \mathscr{F}_{(i-1)\Delta_n}\big] - B_s\Big| \xrightarrow{P} 0, \tag{8}$$

$$\sum_{i=1}^{[t/\Delta_n]} \Big(E\big[\zeta_{n,i}^2 \mid \mathscr{F}_{(i-1)\Delta_n}\big] - \mathbb{E}^2\big[\zeta_{n,i} \mid \mathscr{F}_{(i-1)\Delta_n}\big]\Big) - C_t \xrightarrow{P} 0, \tag{9}$$

$$\sum_{i=1}^{[t/\Delta_n]} E\big[\zeta_{n,i}^4 \mid \mathscr{F}_{(i-1)\Delta_n}\big] \xrightarrow{P} 0. \tag{10}$$

Assume also

$$\sum_{i=1}^{[t/\Delta_n]} E\big[\zeta_{n,i}\Delta_n^i H \mid \mathscr{F}_{(i-1)\Delta_n}\big] \xrightarrow{P} 0, \tag{11}$$

where either H is one of the components of Wiener process W or is any bounded martingale orthogonal (in the martingale sense) to W and $\Delta_n^i H = H_{i\Delta_n} - H_{(i-1)\Delta_n}$.

Then the process

$$\sum_{i=1}^{[t/\Delta_n]} \zeta_{n,i} \xrightarrow{\mathcal{S}-\mathcal{L}} B_t + M_t,$$

where M_t is a continuous process defined on an extension $\big(\widetilde{\Omega}, \widetilde{P}, \widetilde{\mathscr{F}}\big)$ of the filtered probability space $\big(\Omega, P, \mathscr{F}\big)$ and which, conditionally on the the $\sigma-$filter $\mathscr{F}$, is a centered Gaussian $\mathbb{R}-$valued process with $\widetilde{E}\big[M_t^2 \mid \mathscr{F}\big] = C_t$.

Remark 4.1. For lemma 4.2, one can refer to Jacod (2012) (Lemma 4.4) for more details. The stable convergence implies the following crucial property required in the detailed proof of Theorem 1.

If $Z_n \xrightarrow{\mathcal{S}-\mathcal{L}} Z$ and if Y_n and Y are variables defined on $(\Omega, \mathscr{F}, P)$ and with values in the same Polish space F, then

$$Y_n \xrightarrow{P} Y \quad \Rightarrow \quad (Y_n,\ Z_n) \xrightarrow{\mathcal{S}-\mathcal{L}} (Y,\ Z), \tag{12}$$

which implies that $Y_n \times Z_n \xrightarrow{\mathcal{S}-\mathcal{L}} Y \times Z$ through the continuous function $g(x,y) = x \times y$.

Lemma 4.3 (Song et al. (2019))**.** *Under Assumptions 1, 2 and 3, we have*

$$\Delta_n \sum_{i=1}^{n-1} K_{G(x/h_n+1,h_n)}(X_{i\Delta_n})g(X_{i\Delta_n}) = \int_0^T K_{G(x/h_n+1,h_n)}(X_{s-})g(X_{s-})ds + o_p(1). \tag{13}$$

Remark 4.2. Denote $g\bar{L}_X(T,y) := g(y)\cdot\bar{L}_X(T,y)$ and $\xi \sim \Gamma(x/h_n+1, h_n)$, then $E(\xi) = x+h_n$ and $Var(\xi) = xh_n + h_n^2$. For model (1) we have $d[X]_s^c = \sigma^2(X_{s-})ds$, and by the occupation time formula in Lemma 4.1,

$$\begin{aligned}
& \int_0^T K_{G(x/h_n+1,h_n)}(X_{s-})g(X_{s-})ds \\
= & \int_0^T K_{G(x/h_n+1,h_n)}(X_{s-})g(X_{s-})\frac{\sigma^2(X_{s-})ds}{\sigma^2(X_{s-})} \\
= & \int_0^T K_{G(x/h_n+1,h_n)}(X_{s-})g(X_{s-})\frac{d[X]_s^c}{\sigma^2(X_{s-})} \\
= & \int_0^\infty K_{G(x/h_n+1,h_n)}(u)g(u)\bar{L}_X(T,u)du \\
= & E[g\bar{L}_X(T,\xi)] = E[g\bar{L}_X(T,E(\xi)+\xi-E(\xi))] \\
= & E[g\bar{L}_X(T,x+O_p(\sqrt{h_n}))] := E[g\bar{L}_X(T,x+a_n)] \\
= & E[g(T,x+a_n)\bar{L}_X(T,x+a_n)-g(x)\bar{L}_X(T,x+a_n)] \\
& +E[g(x)\bar{L}_X(T,x+a_n)-g(x)\bar{L}_X(T,x)]+E[g(x)\bar{L}_X(T,x)] \\
= & g(x)\bar{L}_X(T,x)+o_p(\bar{L}_X(T,x))
\end{aligned}$$

combining the Assumptions 1 and 3, where $a_n=\sqrt{h_n}$ and $\bar{L}_X(T,x)=\frac{L_X(T,x)}{\sigma^2(x)}$.

Remark 4.3. If $g\equiv 1$, then

$$\Delta_n\sum_{i=1}^{n-1}K_{G(x/h_n+1,h_n)}(X_{i\Delta_n})\xrightarrow{P}\bar{L}_X(T,x). \tag{14}$$

For stationary case,

$$\frac{1}{n}\sum_{i=1}^{n-1}K_{G(x/h_n+1,h_n)}(X_{i\Delta_n})g(X_{i\Delta_n})\xrightarrow{P}g(x)p(x). \tag{15}$$

4.1 The proof of Theorem 1

Proof. Denote

$$\hat{L}_T^\sharp(x):=\Delta_n\sum_{i=1}^{n}K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})$$

and the regularization coefficient $R(h_n)=\begin{cases}\sqrt{h_n^{1/2}}, & if\ x/h_n\to\infty\ (\text{“}interior\ x\text{”});\\ \sqrt{h_n}, & if\ x/h_n\to\kappa\ (\text{“}boundary\ x\text{”}).\end{cases}$

Write

$$\begin{aligned}
&R(h_n)\sqrt{\hat{L}_T^\sharp(x)}(\hat{\lambda}_{AS}(x)-\lambda(x)) \\
=&R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})(c_{i,n}-1)I_{\{(\Delta_iX)^2>\vartheta(\Delta_n)\}}}{\hat{L}_T^\sharp(x)} \\
&+R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\left(\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})I_{\{(\Delta_iX)^2>\vartheta(\Delta_n)\}}}{\hat{L}_T^\sharp(x)}-\lambda(x)\right).
\end{aligned}$$

Firstly, for the second part of $R(h_n)\sqrt{\hat{L}_T^{\sharp}(x)}(\hat{\lambda}_{AS}(x)-\lambda(x))$,

$$\begin{aligned}R_{n,T} :=&R(h_n)\sqrt{\hat{L}_T^{\sharp}(x)}\left(\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})I_{\{(\Delta_i X)^2>\vartheta(\Delta_n)\}}}{\hat{L}_T^{\sharp}(x)}-\lambda(x)\right)\\ =&R(h_n)\sqrt{\hat{L}_T^{\sharp}(x)}\Big(\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})(I_{\{(\Delta_i X)^2>\vartheta(\Delta_n)\}}-\Delta_i N)}{\hat{L}_T^{\sharp}(x)}\\ &+\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\big(\Delta_i N-\int_{t_{i-1}}^{t_i}\lambda_s ds\big)}{\hat{L}_T^{\sharp}(x)}\\ &+\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}\big(\lambda_s-\lambda(x)\big)ds}{\hat{L}_T^{\sharp}(x)}\Big)\\ :=&R_{1_{n,T}}+R_{2_{n,T}}+R_{3_{n,T}}.\end{aligned}$$

As for $R_{1_{n,T}}$, we can show it tends to zero in probability similarly as the detailed proof in Mancini and Renò (2011) replacing $K\left(\frac{X_{t_{i-1}}-x}{h_n}\right)$ with $K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})$.

As for $R_{2_{n,T}}$, in order to obtain

$$R(h_n)\sqrt{\hat{L}_T^{\sharp}(x)}\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\big(\Delta_i N-\int_{t_{i-1}}^{t_i}\lambda_s ds\big)}{\hat{L}_T^{\sharp}(x)}\xrightarrow{d} N\left(0,\lambda(x)\cdot R^2(h)B_h(x)\right),$$

we firstly show the numerator of

$$\frac{\frac{R(h_n)}{T\sqrt{T}}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}\bar{\nu}(ds)}{\frac{\hat{L}_T^{\sharp}(x)}{T\sqrt{T}}}:= \frac{\frac{R(h_n)}{T\sqrt{T}}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\big(\Delta_i N-\int_{t_{i-1}}^{t_i}\lambda_s ds\big)}{\frac{\hat{L}_T^{\sharp}(x)}{T\sqrt{T}}} \tag{16}$$

converges stably in law to M_1, where $\bar{\nu}(dt)=N_t dt-\lambda(X_{t-})dt$, M_1 is a Gaussian martingale defined on an extension $\big(\tilde{\Omega},\tilde{P},\tilde{\mathscr{F}}\big)$ of our filtered probability space and having $\tilde{E}[M_1^2|\mathscr{F}]=\lambda(x)\cdot R^2(h_n)\cdot B_{h_n}(x)\cdot\frac{\bar{L}_X(T,x)}{T^3}$ with $B_{h_n}(x)=\begin{cases}\frac{1}{2\sqrt{\pi}}h_n^{-1/2}x^{-1/2}, & if\ x/h_n\to\infty\ (\text{“}interior\ x\text{”});\\ \frac{\Gamma(2\kappa+1)}{2^{1+2\kappa}\Gamma^2(\kappa+1)}h_n^{-1}, & if\ x/h_n\to\kappa\ (\text{“}boundary\ x\text{”}).\end{cases}$

Denote $\sum_{i=1}^n q_i:=\frac{R(h_n)}{T\sqrt{T}}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}\bar{\nu}(ds)$, and Jacod's stable convergence theorem tells us that the following arguments,

$$\begin{aligned}V_1&=\sum_{i=1}^n E_{i-1}[q_i]\xrightarrow{P}0,\\ V_2&=\sum_{i=1}^n\big(E_{i-1}[q_i^2]-E_{i-1}^2[q_i]\big)\xrightarrow{P}\lambda(x)\cdot R^2(h_n)\cdot B_{h_n}(x)\cdot\frac{\bar{L}_X(T,x)}{T^3},\\ V_3&=\sum_{i=1}^n E_{i-1}[q_i^4]\xrightarrow{P}0,\\ V_4&=\sum_{i=1}^n E_{i-1}[q_i\Delta_i H]\xrightarrow{P}0,\end{aligned}$$

imply $\sum_{i=1}^n q_i\xrightarrow{\mathcal{S}-\mathcal{L}}M_1$, where either $H=W$ or H is any bounded martingale orthogonal (in the martingale sense) to W, $E_{i-1}[\ \cdot\ ]=E[\ \cdot\ |X_{(i-1)\Delta_n}]$.

Considering $K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})$ is measurable with respect to the σ-algebra generated by $\{X_u, 0 \le u \le t_{i-1}\}$, q_i is a martingale difference sequence, so $V_1 \equiv 0$.

For V_2,

$$\begin{aligned}
V_2 &= \sum_{i=1}^{n} \left(E_{i-1}[q_i^2] - E_{i-1}^2[q_i]\right) \\
&= \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n}) E_{i-1}\left[\left(\int_{t_{i-1}}^{t_i} \bar{\nu}(ds)\right)^2\right] \\
&= \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n}) E_{i-1}\left[\int_{t_{i-1}}^{t_i} \lambda(X_{s-})ds\right] \\
&= \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n}) \times \\
&\quad E_{i-1}\left[\int_{t_{i-1}}^{t_i} \lambda(X_{(i-1)\Delta_n})ds + \int_{t_{i-1}}^{t_i} (\lambda(X_{s-}) - \lambda(X_{(i-1)\Delta_n}))ds\right] \\
&= \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\lambda(X_{(i-1)\Delta_n})\Delta_n \\
&\quad + \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n}) E_{i-1}\left[\int_{t_{i-1}}^{t_i} (\lambda(X_{s-}) - \lambda(X_{(i-1)\Delta_n}))ds\right] \\
&=: V_{21} + V_{22}.
\end{aligned}$$

For V_{21} :

$$\begin{aligned}
V_{21} &= \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\lambda(X_{(i-1)\Delta_n})\Delta_n \\
&= \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n}) \int_{t_{i-1}}^{t_i} (\lambda(X_{(i-1)\Delta_n}) - \lambda(X_s))ds \\
&\quad + \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} (K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\lambda(X_s) - K^2_{G(x/h_n+1,h_n)}(X_s)\lambda(X_s))ds \\
&\quad + \frac{R^2(h_n)}{T^3} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} K^2_{G(x/h_n+1,h_n)}(X_s)\lambda(X_s)ds \\
&=: V_{211} + V_{212} + V_{213},
\end{aligned}$$

where $V_{211} \overset{a.s.}{\to} 0$ and $V_{212} \overset{a.s.}{\to} 0$ can be dealt with in analogy to (b3) and (b4) in Mancini and Renò (2011) replacing $K\left(\frac{X_{t_{i-1}}-x}{h_n}\right)$ with $K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})$. Using the occupation time formula in Lemma 4.1, Assumption 3 and the result (3.2) in Chen (2000), we can obtain that

$$V_{213} \xrightarrow{a.s.} \lambda(x) \cdot R^2(h_n) \cdot B_{h_n}(x) \cdot \frac{\bar{L}_X(T,x)}{T^3}.$$

For V_{22}, we define the random sets for each n,

$$I_{0,n} = \{i \in \{1, ..., n\} : \Delta_i N = 0\},$$

and

$$I_{1,n} = \{i \in \{1, ..., n\} : \Delta_i N \neq 0\}.$$

Applying the mean-value theorem for $\lambda(\cdot)$, neglecting the terms with $i \in I_{1,n}$ similarly as the detailed proof for Lemma 4.3 and bounding $|X_s - X_{(i-1)\Delta_n}|$ by the property of uniform boundedness of the increments of X_t paths when $i \in I_{0,n}$ (denoted as UBI property), we can reach

$$\begin{aligned}
&\frac{R^2(h_n)}{T^3}\sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\Delta_n \cdot \sup_x |\lambda^{'}(x)| \cdot \sqrt{\Delta_n \ln\frac{1}{\Delta_\mathrm{n}}} \\
= \;& O_{a.s.}\big(\frac{R^2(h_n)}{T^3}\sum_{i\in I_{0,n}} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\Delta_n \cdot \sqrt{\Delta_n \ln\frac{1}{\Delta_\mathrm{n}}}\big) \\
= \;& O_{a.s.}\big(\frac{R^2(h_n)}{T^3}\sum_{i=1}^{n} K^2_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\Delta_n \cdot \sqrt{\Delta_n \ln\frac{1}{\Delta_\mathrm{n}}}\big) \\
= \;& O_{a.s.}\Big(R^2(h_n)\cdot B_{h_n}(x)\cdot \frac{\bar{L}_X(T,x)}{T^3}\Big)\cdot \sqrt{\Delta_n \ln\frac{1}{\Delta_\mathrm{n}}} \xrightarrow{a.s.} 0.
\end{aligned}$$

For V_3, by Burkholder-Davis-Gundy inequality, Lemma 4.3 and Lemma 4.1 in Rosa and Nogueira (2016), we get

$$\begin{aligned}
V_3 \;&=\; \sum_{i=1}^{n} E_{i-1}[q_i^4] \\
&=\; \frac{R^4(h_n)}{T^6}\sum_{i=1}^{n} K^4_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})E_{i-1}\left[\left(\int_{t_{i-1}}^{t_i}\bar{\nu}(ds)\right)^4\right] \\
&\leq\; O_p(1)\cdot\frac{R^4(h_n)}{T^6}\sum_{i=1}^{n} K^4_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})(\Delta_n^2+\Delta_n) \\
&=\; O_p(1)\cdot\left(R^4(h_n)\cdot\frac{1+\Delta_n}{T^5}\cdot\frac{L_X(T,x)}{T}\cdot B(4,h_n,x)\right) \\
&=\; O_p(1)\cdot\left(R^4(h_n)\cdot\frac{1+\Delta_n}{T^5}\cdot\frac{L_X(T,x)}{T}\cdot\begin{cases} h_n^{-3/2}, & if\ x/h_n\to\infty\ (\text{“}interior\ x\text{”}); \\ h_n^{-3}, & if\ x/h_n\to\kappa\ (\text{“}boundary\ x\text{”}).\end{cases}\right) \\
&\xrightarrow{P}\; 0.
\end{aligned}$$

For V_4, in addition, when $H = W$, according to the model assumption that J_t is independent of W_t, we easily get $V_4 \equiv 0$. Moreover, when H is any bounded martingale orthogonal (in the martingale sense) to W,

$$\begin{aligned}
\sum_{i=1}^{n} E_{i-1}[q_i\Delta_i H] \;&=\; \frac{R(h_n)}{T\sqrt{T}}\sum_{i=1}^{n} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})E_{i-1}\left[\int_{t_{i-1}}^{t_i}\bar{\nu}(ds)\Delta_i H\right] \\
&=\; O_P\left(\frac{R(h_n)}{T\sqrt{T}}\sum_{i=1}^{n} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})E_{i-1}\left[\int_{t_{i-1}}^{t_i}\bar{\nu}(ds)\right]\right) \\
&=\; O_P\left(\frac{R(h_n)}{T\sqrt{T}}\sum_{i=1}^{n} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\Delta_n^{\frac{1}{2}}\right) \\
&=\; O_P\left(\frac{\sqrt{nR^2(h_n)}}{T^2}\sum_{i=1}^{n} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\Delta_n\right) \\
&=\; O_P\left(\frac{\sqrt{nR^2(h_n)}}{T}\cdot\frac{\bar{L}_X(T,x)}{T}\right)\xrightarrow{P} 0,
\end{aligned}$$

provided the bounded of H such that $\Delta_i H \le C$ for the second equality, Hölder inequality and Burkerholder-Davis-Gundy inequality for the third equality, the equation (14) by Lemma 4.3 for the fifth equality and

$$\frac{\sqrt{nR^2(h_n)}}{T} = o_P(1)$$

as $T \longrightarrow \infty$.

As for $R_{3_{n,T}}$,

$$\begin{aligned}
&\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}(\lambda_s-\lambda(x))ds}{\hat{L}_T^{\sharp}(x)}\\
=&\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}(\lambda_s-\lambda_{t_{i-1}})ds}{\hat{L}_T^{\sharp}(x)}\\
&+\frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}(\lambda_{t_{i-1}}-\lambda(x))ds}{\hat{L}_T^{\sharp}(x)}\\
:=&D_{1_{n,T}}+D_{2_{n,T}},
\end{aligned} \tag{17}$$

where $\hat{L}_T^{\sharp}(x) \xrightarrow{P} \bar{L}_X(T,x)$ as the equation (14) by Lemma 4.3.

We now obtain the asymptotic bias for the expressions $D_{1_{n,T}}$ and $D_{2_{n,T}}$ above. That is,

$$D_{1_{n,T}} = o_P(D_{2_{n,T}}).$$

By the Taylor expansion for $\lambda_{t_{i-1}} - \lambda(x)$ in $D_{2_{n,T}}$ up to order 2,

$$\lambda_{t_{i-1}}-\lambda(x)=\lambda^{'}(x)(X_{(i-1)\Delta_n}-x)+\frac{1}{2}\lambda^{''}(x)(x+\theta(X_{(i-1)\Delta_n}-x))(X_{(i-1)\Delta_n}-x)^2,$$

where θ is a random variable satisfying $\theta \in [0,1]$.

For $D_{2_{n,T}}$, by Lemma 4.3 and the results in Chen (2002), we get

$$\begin{aligned}
D_{2_{n,T}} &= \frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}(\lambda_{t_{i-1}}-\lambda(x))\,ds}{\hat{L}_T^{\sharp}(x)}\\
&= \frac{\sum_{i=1}^{n-1} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})(X_{(i-1)\Delta_n}-x)\Delta_n\cdot\lambda^{'}(x)}{\hat{L}_T^{\sharp}(x)}\\
&+ \frac{\sum_{i=1}^{n-1} K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})(X_{(i-1)\Delta_n}-x)^2\Delta_n\cdot\frac{1}{2}\lambda^{''}(x+\theta(X_{(i-1)\Delta_n}-x))}{\hat{L}_T^{\sharp}(x)}\\
&\xrightarrow{P} h_n\left[\lambda^{'}(x)\left\{1+x\frac{s^{'}(x)}{s(x)}\right\}+\frac{x}{2}\lambda^{''}(x)\right],
\end{aligned}$$

using

$$E(\xi-x)=h_n,\quad E(\xi-x)^2=h_n(x+2h_n)$$

and

$$E(\xi-x)^l=O(h_n^2)$$

for $3 \le l$ and $\xi \sim \Gamma(x/h_n+1,h_n)$.

Furthermore, we use the mean-value theorem to $\lambda_s - \lambda_{t_{i-1}}$ for $D_{1_{n,T}}$, then

$$D_{1_{n,T}} = \frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}(\lambda_s - \lambda_{t_{i-1}})ds}{\hat{L}_T^\sharp(x)}$$

$$= \frac{\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\int_{t_{i-1}}^{t_i}\lambda^{'}(\xi_i)(X_s - X_{(i-1)\Delta_n})\,ds}{\hat{L}_T^\sharp(x)}$$

$$\overset{P}{\le} \frac{1}{\bar{L}_X(T,x)} \times \left[\left(\Delta_n \ln\frac{1}{\Delta_n}\right)^{\frac{1}{2}} \cdot \sup_x|\lambda^{'}(x)| \cdot \sum_{i\in I_{0,n}} K_{G(x/h_n+1,h_n)}(X_{i\Delta_n})\Delta_n + 2CN_1\Delta_n \ (for\ i \in I_{1,n})\right]$$

$$\to O\left[\left(\Delta_n \ln\frac{1}{\Delta_n}\right)^{\frac{1}{2}}\right] = o(h_n)$$

by the UBI property of $i \in I_{0,n}$.

We prove that $D_{1_{n,T}} = o_P(D_{2_{n,T}})$, so the dominant bias arises from $D_{2_{n,T}}$, which is

$$h_n\left[\lambda^{'}(x)\left\{1 + x\frac{s^{'}(x)}{s(x)}\right\} + \frac{x}{2}\lambda^{''}(x)\right].$$

Finally, we can show the first part of $R(h_n)\sqrt{\hat{L}_T^\sharp(x)}(\hat{\lambda}_{AS}(x) - \lambda(x))$,

$$R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\frac{\frac{1}{h_n}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})(c_{i,n}-1)I_{\{(\Delta_i X)^2 > \vartheta(\Delta_n)\}}}{\hat{L}_T^\sharp(x)}$$

$$\le \sup_i |1 - c_{i,n}|R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\frac{\frac{1}{h_n}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})I_{\{(\Delta_i X)^2 > \vartheta(\Delta_n)\}}}{\hat{L}_T^\sharp(x)}$$

$$= \sup_i |1 - c_{i,n}|R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\frac{\frac{1}{h_n}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\big(I_{\{(\Delta_i X)^2 > \vartheta(\Delta_n)\}} - \lambda(x)\Delta_n\big)}{\hat{L}_T^\sharp(x)}$$

$$+ \sup_i |1 - c_{i,n}|R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\frac{\frac{1}{h_n}\sum_{i=1}^n K_{G(x/h_n+1,h_n)}(X_{(i-1)\Delta_n})\lambda(x)\Delta_n}{\hat{L}_T^\sharp(x)}$$

$$= \sup_i |1 - c_{i,n}|\big(O_p(1) + \lambda(x)R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\big)$$

$$= O_p\big(\sup_i |1 - c_{i,n}|R(h_n)\sqrt{\hat{L}_T^\sharp(x)}\big) \longrightarrow 0,$$

with the help of asymptotic normality for $R_{n,T}$.

Combining the results (12), (14) and the above conclusions, we can obtain that for the interior x visited by X,

$$\sqrt{h_n^{1/2}\hat{L}_T^\sharp(x)}\left(\hat{\lambda}_{AS}(x) - \lambda(x) - \left[\lambda^{'}(x)\left\{1 + x\frac{s^{'}(x)}{s(x)}\right\} + \frac{x}{2}\lambda^{''}(x)\right]\cdot h_n\right) \xrightarrow{d} \mathscr{N}\left(0, \lambda(x)\cdot\frac{1}{2\sqrt{\pi}x^{1/2}}\right),$$

for the boundary x visited by X,

$$\sqrt{h_n\hat{L}_T^\sharp(x)}\left(\hat{\lambda}_{AS}(x) - \lambda(x)\right) \xrightarrow{d} \mathscr{N}\left(0, \lambda(x)\cdot\frac{\Gamma(2\kappa+1)}{2^{1+2\kappa}\Gamma^2(\kappa+1)}\right).$$

□

References

[1] Y Aït-Sahalia, J Park. *Bandwidth selection and asymptotic properties of local nonparametric estimators in possibly nonstationary continuous-time models*, J Econometrics, 2016, 192(1): 119-138.

[2] T Andersen, L Benzoni, J Lund. *Stochastic volatility, mean drift, and jumps in the short-term interest rate*, Working paper, https://www.researchgate.net/publication/228793214, 2004.

[3] F Bandi, T Nguyen. *On the functional estimation of jump-diffusion models*, J Econometrics, 2003, 116(1): 293-328.

[4] S Chen. *Probability density function estimation using Gamma kernels*, Ann Inst Statist Math, 2000, 52(3): 471-480.

[5] S Chen. *Local linear smoothers using asymmetric kernels*, Ann Inst Statist Math, 2002, 54(2): 312-323.

[6] R Cont, P Tankov. *Financial Modelling with Jump Processes*, Chapman and Hall/CRC, Boca Raton, 2004.

[7] S Das. *The surprise element: jumps in interest rates*, J Econometrics, 2002, 106(1): 27-65.

[8] N Eisenbaum, H Kaspi. *On the continuity of local times of borel right Markov processes*, Ann Probab, 2007, 35(3): 915-934.

[9] J Fan, Y Fan, J Jiang. *Dynamic integration of time- and state-domain methods for volatility estimation*, J Am Stat Assoc, 2007, 102(478): 618-631.

[10] N Gospodinovy, M Hirukawa. *Nonparametric estimation of scalar diffusion processes of interest rates using asymmetric kernels*, J Empir Financ, 2012, 19(4): 595-609.

[11] M Hanif. *Local linear estimation of recurrent jump-diffusion models*, Commun Stat-Theor M, 2012, 41(22): 4142-4163.

[12] M Hanif. *A nonparametric approach to the estimation of jump-diffusion models with asymmetric kernels*, Cogent Mathematics, 2016, 3(1): 1179247.

[13] M Hanif, H Wang, Z Lin. *Reweighted Nadaraya-Watson estimation of jump-diffusion models*, Sci China Math, 2012, 55(5): 1005-1016.

[14] J Jacod, A N Shiryaev. *Limit Theorems for Stochastic Processes*, Springer, New York, 2003.

[15] J Jacod. *Statistics and high-frequency data*, In: M Kessler, A Lindner, M Sørensen, *Statistical Methods for Stochastic Differential Equations*, CRC Press, Boca Raton, 2012, 191-310.

[16] M Johannes. *The statistical and economic role of jumps in continuous-time interest rate models*, J Finance, 2004, 59(1): 227-260.

[17] Z Lin, H Wang. *Empirical likelihood inference for diffusion processes with jumps*, Sci China Math, 2010, 53(7): 1805-1816.

[18] C Mancini, R Renò. *Threshold estimation of Markov models with jumps and interest rate modeling*, J Econometrics, 2011, 160(1): 77-92.

[19] A Rosa, M Nogueira. *Nonparametric estimation of a regression function using the gamma kernel method in ergodic processes*, arXiv:1605.07520v2, 2016.

[20] Y Song, W Hou, S Zhou. *Variance reduction estimation for return models with jumps using Gamma asymmetric kernels*, accepted by Studies in Nonlinear Dynamics & Econometrics, 2019.

[20] Y Song, H Li, Y Feng, Y Fu. *Nonparametric threshold estimation for expected return of financial assets using asymmetric kernels*, Working paper, Shanghai Normal University, 2019.

[21] Y Song, H Wang. *Central limit theorems of local polynomial threshold estimators for diffusion processes with jumps*, Scand J Statist, 2018, 45(3): 644-681.

[22] H Wang, L Zhou. *Bandwidth selection of nonparametric threshold estimator in jump-diffusion models*, Comput Math Appl, 2017, 73(2): 211-219.

[1] School of Mathematics, Shandong University, Jinan 250100, China.
Email: lin_yiwei@126.com

[2] School of Finance and Business, Shanghai Normal University, Shanghai 200234, China.
Email: songyuping@shnu.edu.cn
Email: 15253172256@163.com