Jarque-Bera Normality Test for the Driving Levy Process of a Discretely Observed Univariate SDE

Lee, Sangyeol
Department of Statistics, Seoul National University

Masuda, Hiroki
Graduate School of Mathematics, Kyushu University

出版情報：MI Preprint Series. 2008-10, 2008-11-17. 九州大学大学院数理学研究院
バージョン：

MI Preprint Series

Kyushu University
The Grobal COE Program
Math-for-Industry Education & Research Hub

Jarque-Bera Normality Test for the Driving Lévy Process of a Discretely Observed Univariate SDE

S. Lee & H. Masuda

MI 2008-10

(Received November 17, 2008)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

Jarque-Bera Normality Test for the Driving Lévy Process of a Discretely Observed Univariate SDE *

Sangyeol Lee

Department of Statistics, Seoul National University,
Seoul, 151-742, Korea
Email: sylee@stats.snu.ac.kr

Hiroki Masuda†

Graduate School of Mathematics, Kyushu University,
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
Email: hiroki@math.kyushu-u.ac.jp

Version: November 17, 2008

Abstract

We study the validity of the Jarque-Bera test for a class of univariate parametric stochastic differential equations (SDE) $dX_t = b(X_t, \alpha)dt + dZ_t$ observed at discrete time points $t_i^n = ih_n$, $i = 1, 2, \ldots, n$, where Z is a nondegenerate Lévy process with finite moments, and $nh_n \to \infty$ and $nh_n^2 \to 0$ as $n \to \infty$. Under appropriate conditions it is shown that Jarque-Bera type statistics based on the Euler residuals can be used to test the normality of the unobserved Z, and moreover, that the proposed test is consistent against presence of any nontrivial jump component. Our result therefore provides a very easy and asymptotically distribution-free test procedure without any fine-tuning parameter. Some illustrative simulation results are given to reveal good performance of our test statistics.

Running head: Normality Test for SDE.

The AMS Mathematics Subject Classifications (2000): 62M02, 62F05.

Key Words: Jarque-Bera test, Lévy process, stochastic differential equation.

*Partly supported by Grant-in-Aid for Young Scientists (B) 20740061 of Japan, and Cooperative Research Program of the Institute of Statistical Mathematics.

†The corresponding author

1 Introduction

During the past decades, the diffusion process has long been popular among practitioners in various fields such as finance, engineering, physical and medical sciences. Statistical inference for diffusion processes has become very crucial in statistical analysis, especially in stochastic finance, to cope with the demand to resolve statistical problems occurring in actual practice. See, for instance, Karatzas and Shreve (1988), Shiryayev (1999), Prakasa Rao (1999), Lipster and Shiryayev (2001), and Kutoyants (2004). Although the diffusion process is very popular in handling financial time series data, experience suggests that the diffusion process is not well fitted to given data due to high volatilities and discontinuous jumps. To deal with this problem, practitioners often adopt models such as jump diffusion processes and Lévy processes: see Barndorff-Nielsen at al. (2001), Shoutens (2003), and Cont and Tankov (2004). Hence, to employ correct models, there is a need to check whether or not modelling based on diffusion processes is reasonable in handling time series data. In this article, motivated by this viewpoint, we consider the goodness of fit test problem for the parametric univariate stochastic differential equation (SDE for short) given by (1) below, based on high-frequency and long-period data (see **C1** below for the precise meaning). We are interested in testing whether or not the driving Lévy process is a Wiener process (possibly scaled by an unknown constant), against presence of "any" nontrivial jump component.

Among the goodness of fit methods, the Kolmogorov-Smirnov test falls in the category of an empirical process method (cf. D'Agostino and Stephens (1986)) since it is generated from the empirical process. As a reference that addresses the empirical process and the goodness of fit tests for the autoregressive and GARCH models, we employ Lee and Wei (1992) and Lee and Taniguchi (2005). In contrast to the Kolmogorov-Smirnov test, the Bickel-Rosenblatt test (cf. Bickel and Rosenblatt (1973)) belongs to a class of density-based testing methods and is well known to better detect heavy-tailed alternatives: see Lee and Na (2001) and Horváth and Zitikis (2006). Although these tests have their own merit, it is widely accepted that they also have certain shortcomings. For instance, the Kolmogorov-Smirnov test has a tendency to produce low powers in many situations, and the Bickel-Rosenblatt test has difficulty in choosing an optimal bandwidth (see Lee (2006) for the Bickel-Rosenblatt test for diffusion processes). Based on this reasoning, we here employ the Jarque-Bera (JB) test (cf. Jarque and Bera (1980) and Bera and Jarque (1981)), which is asymptotically distribution-free, as an alternative in our study since it is well known that the JB test is easy to implement in actual practice in comparison to other conventional tests.

In the construction of the JB test for diffusion processes, we will use the discrete sampling scheme as seen in Flrores-Zmirou (1989), Yoshida (1992) and Kessler (1997). The key idea to employ the residual-based JB test is that if the data is truly realized from a diffusion process, the residuals obtained from the sampled observations should behave like normal random variables. This idea is actually used in Lee and Wee (2008) who consider the residual empirical process in diffusion processes. In fact, the residual based JB test is widely used for time series models without a theoretical justification since the residual based JB test is believed to behave like the ordinary JB test. However, as seen in Lee and Wei (1999) and Lee and Taniguchi (2005), the residual based test behaves somewhat differently from the test based on true errors, depending upon the characteristic of the structure of the time series models. In particular, the result of Lee and Taniguchi (2005) reveals that the GARCH effects severely affect the limiting null distribution of the residual empirical process. To our knowledge, there exist few articles considering the JB test in financial time series models. We refer to Kulperger and Yu (2005) who study the JB test based on GARCH residuals within the framework of high moment partial sum processes. By considering all these aspects, here we carefully analyze the JB test for diffusion processes. Mainly due to the high-frequency sampling scheme, it turns out that our test based on the statistic $\mathcal{T}_n$ defined in Section 2 is asymptotically distribution-free and consistent.

This article is organized as follows: in Section 2, we introduce our model setup and describe our results; in Section 3, we provide some simulation results to illustrate our findings; finally, Section 4 presents the proofs of our results.

2 Setup and statement of result

Suppose we have a discrete-time data $X_{t_0^n}, X_{t_1^n}, \dots, X_{t_n^n}$ from a solution of the univariate SDE

$$dX_t = b(X_t, \theta)dt + dZ_t \tag{1}$$

defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, P)$, where $t_i^n = ih_n$ are positive constants, $\theta \in \Theta \subset \mathbb{R}^p$ is an unknown vector with Θ being a bounded convex domain, and Z is a nontrivial Lévy process; here the nontriviality means that Z is not a deterministic linear function of t. The initial variable X_0 is supposed to be independent of Z. Denote by $\sigma^2 \geq 0$ and ν the Gaussian variance and Lévy measure of Z. In this article we are interested in testing the normality of the unobserved Z against presence of any nontrivial jump component. Under the nontriviality of Z, this can be formulated as

$$H_0\colon\ \nu(\mathbb{R}) = 0 \ \text{ v.s. } \ H_1\colon\ \nu(\mathbb{R}) \in (0, \infty].$$

Note that $\sigma^2 \geq 0$ may be arbitrary under H_1.

Denote by $\theta_0 \in \Theta$ the true value of θ, and by P_0 the true law of X associated with θ_0. Throughout this article, the symbol $\to^p$ (resp. $\to^d$) indicates the convergence in P_0-probability (resp. weak convergence along P_0) for $n \to \infty$, and also stochastic-order symbols are taken under P_0. We will denote by ∂_θ the gradient operator with respect to θ, and write $T_n = nh_n$ and

$$\Delta_i^n \zeta = \zeta_{t_i^n} - \zeta_{t_{i-1}^n}$$

for a process ζ. For conciseness we will here focus on $\sqrt{T_n}$-consistent estimators of θ and Z with finite moment of any order.

Our basic regularity conditions are summarized as follows.

C1. $h_n \to 0$, $T_n \to \infty$, and $nh_n^2 \to 0$.

C2. $E[Z_1] = 0$, and $E[|Z_t|^q] < \infty$ for every $q > 0$.

C3. $x \mapsto b(x, \theta_0)$ is globally Lipschitz, and $\theta \mapsto b(x, \theta)$ is of class $\mathcal{C}^2$ for every x.

C4. For every $q > 0$ we have

$$\sup_{t\in\mathbb{R}_+} \left\{ E_0[|b(X_t, \theta_0)|^q] + E_0\left[\sup_{\theta\in\Theta} |\partial_\theta b(X_t, \theta)|^q\right] + E_0\left[\sup_{\theta\in\Theta} |\partial_\theta^2 b(X_t, \theta)|\}^q\right]\right\} < \infty.$$

C5. There exist estimators $\hat{\theta}_n$ of θ_0 such that $\sqrt{T_n}(\hat{\theta}_n - \theta_0) = O_p(1)$.

For convenience we give some remarks on our conditions.

Remark 2.1. *Under* **C2**, *Z admits a Lévy-Itô decomposition*

$$Z_t = \sigma w_t + \int_0^t \int z\tilde{\mu}(ds, dz), \tag{2}$$

where w is a standard Wiener process and $\tilde{\mu}(ds, dz) := \mu(ds, dz) - \nu(dz)ds$ with Poisson random measure $\mu(ds, dz)$ and Lévy measure $\nu(dz)$. Recall that $E[|Z_t|^q] < \infty$ if and only if $\int_{|z|>1} |z|^q \nu(dz) < \infty$. Then we may set $E[Z_t] = 0$ from the beginning without loss of generality, since, if not, we may replace Z_t with the centered version $\tilde{Z}_t := Z_t - E[Z_1]t$ by incorporating $E[Z_1]$ into the drift parameter θ: however, we explicitly stated that $E[Z_t] = 0$ for clarity.

Remark 2.2. *Clearly,* **C4** *is automatic if $x \mapsto b(x, \theta_0)$ and $x \mapsto \sup_{\theta\in\Theta} \partial_\theta^j b(x, \theta)$, $j = 1, 2$, are bounded. Otherwise,* **C4** *is implied by the condition*

$$\forall q > 0 \quad \sup_{t\in\mathbb{R}_+} E_0[|X_t|^q] < \infty, \tag{3}$$

as soon as $x \mapsto \partial_\theta^j b(x,\theta)$, $j=1,2$, *are dominated by some polynomials uniformly in* θ. *For checking (3) we can apply Masuda (2007, Theorem 2.2(i)): among many possibilities, for example, (3) holds true as soon as there exist positive constants* c *and* R *such that for every* $|x| \geq R$ *we have*

$$xb(x,\theta_0) \leq -c|x|^2 \tag{4}$$

which, of course, sets limits to Θ.

Remark 2.3. *In ergodic cases we may use the simple least-squares type estimator for* **C5***: under appropriate conditions including the ergodicity of* X *and (3) (see Masuda (2005, Sections 2 and 3) for details), we can take* $\hat{\theta}_n$ *as a random root of the estimating equation*

$$\sum_{i=1}^{n}\{\Delta_i^n X - b(X_{t_{i-1}^n},\theta)h_n\}\partial_\theta b(X_{t_{i-1}^n},\theta) = 0.$$

When $\nu(\mathbb{R}) = 0$, *several conditions for the ergodicity of* X *are well known. In case of* $\nu(\mathbb{R}) > 0$ *with arbitrary* $\sigma^2 \geq 0$, *we can apply Kulik (2007, Propositions 0.1 and/or A.2) to conclude that the (exponential) ergodicity holds true as soon as there exists a* $\mathcal{C}^2$*-function* $f : \mathbb{R} \to \mathbb{R}_+$ *satisfying* $f(x) \to \infty$ *as* $|x| \to \infty$, *for which there exist positive constants* c *and* R *such that for every* $|x| \geq R$ *we have*

$$\mathcal{A}f(x) \leq -cf(x), \tag{5}$$

where $\mathcal{A}$ *denotes the generator of* X*: in our framework, (4) is sufficient for (5). Moreover, as in the diffusion cases, it can be expected that for non-ergodic* X *the trajectory-fitting type estimator studied in Masuda (2005) may be consistent for* θ_0 *at much faster rate than* $\sqrt{T_n}$ *(possibly exponentially fast).*

Put $b_{i-1}(\theta) = b(X_{t_{i-1}^n},\theta)$ for notational simplicity. We define an Euler-type residual sequence (without variance scaling; see Remark 2.5 below) by

$$\hat{\epsilon}_{ni} = \frac{1}{\sqrt{h_n}}\{\Delta_i^n X - b_{i-1}(\hat{\theta}_n)h_n\}, \tag{6}$$

which will be used to approximate

$$\epsilon_{ni} := \frac{\Delta_i^n Z}{\sqrt{h_n}}.$$

Writing $\bar{\hat{\epsilon}}_n = \frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_{ni}$, we introduce the kth self-normalized residual sums ($k \in \mathbb{N}$, $k \geq 2$):

$$\hat{\Phi}_n^{(k)} = \frac{\hat{\Psi}_n^{(k)}}{(\hat{\Psi}_n^{(2)})^{k/2}}, \quad \text{where} \quad \hat{\Psi}_n^{(k)} := \frac{1}{n}\sum_{i=1}^{n}(\hat{\epsilon}_{ni} - \bar{\hat{\epsilon}}_n)^k.$$

Our test statistics $\mathcal{T}_n$ are then defined by

$$\mathcal{T}_n = \frac{n}{6}(\hat{\Phi}_n^{(3)})^2 + \frac{n}{24}(\hat{\Phi}_n^{(4)} - 3)^2.$$

Now we are in position to state our main result.

Theorem 2.4. *Suppose the conditions* **C1** *to* **C5**. *Then we have:*

(a) $\mathcal{T}_t \to^d \chi^2(2)$ *under* H_0, *where* $\chi^2(2)$ *denotes the chi-square distribution with* 2 *degrees of freedom; and*

(b) $P_0[\mathcal{T}_n > K] \to 1$ *for every* $K > 0$ *under* H_1.

Given sampling points t_i^n, Theorem 2.4 enables us to perform a consistent Jarque-Bera type test for the normality of the unobserved driving Lévy process Z, without any fine-tuning parameter. The proof of Theorem 2.4 is given in Section 4.

Remark 2.5. *Instead of (6) we may consider the (possibly more natural) residual*

$$\hat{\epsilon}'_{ni} = \frac{1}{\sqrt{\hat{\sigma}_n^2 h_n}}\{\Delta_i^n X - h_n b_{i-1}(\hat{\theta}_n)\},$$

where $\hat{\sigma}_n^2 := \frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_{ni}^2$*, the residual sum of squares (see Yoshida (1992)). This* $\hat{\epsilon}'_{ni}$ *approximates* $\epsilon'_{ni} := \Delta_i^n Z/\sqrt{h_n \mathrm{var}[Z_1]}$*, and actually it turned out that we could deduce the same claim as in Theorem 2.4 with this* $(\hat{\epsilon}'_{ni})_{i=1}^n$*. However, using* $\hat{\epsilon}'_{ni}$ *is clearly redundant in our setup (1), because of the invariance of* $\hat{\Phi}_n^{(k)}$ *under scaling of* $\hat{\epsilon}_{ni}$*.*

3 Simulation experiments

We observe the finite-sample performance of the statistics $\mathcal{T}_n$ when X is a Ornstein-Uhlenbeck type process

$$dX_t = -\theta_0 X_t dt + dZ_t$$

with $X_0 = 0$, targeting at:

$$\begin{aligned} H_0 &: \ Z \text{ is a standard Wiener process;} \\ H_1 &: \ \text{the law of } Z_t \text{ is } NIG(a, 0, \delta t, 0). \end{aligned} \tag{7}$$

See the references cited in Masuda (2005) for the details of the NIG distributions $NIG(\alpha, 0, \delta t, 0)$. Under H_1, Z is a centered and symmetric Lévy process and we know that **C2** is fulfilled for the NIG Lévy process as soon as $\alpha > 0$. **C3** is clearly met. Supposing $\theta_0 > 0$, we can verify **C4** and **C5** with the least squares type estimator

$$\hat{\theta}_n = \frac{1}{h_n}\left\{1 - \left(\sum_{i=1}^n X_{t_{i-1}^n}^2\right)^{-1} \sum_{i=1}^n X_{t_{i-1}^n} X_{t_{i-1}^n}\right\},$$

which fulfils

$$\sqrt{T_n}(\hat{\theta}_n - \theta_0) \to^d \mathcal{N}_1\left(0, \mathrm{var}[Z_1]\left(\int x^2 \pi(dx)\right)^{-1}\right),$$

where π denotes the invariant measure of X (the characteristic function of π is explicit): see the references cited in Remarks 2.2 and 2.3 for details.

For the parameters of the driving NIG Lévy process, we choose $(a, \delta) = (3, 3)$ and $(10, 10)$, for both of which we have $E[Z_1^2] = \delta/\alpha = 1$, comparable with the case of H_0. The tail of $NIG(\alpha, 0, \delta t, 0)$ gets heavier for smaller $\alpha > 0$.

Here we take $\theta_0 = 3$ and $h_n = n^{-0.6}$, so that $T_n = n^{0.4} \to \infty$ and $nh_n^2 = n^{-0.2} \to 0$, making **C1** valid. In order to simulate sample paths of X, we use the Euler scheme with mesh $h_n/50$ in each trial. The Figure 1 shows sample paths of X with $\theta_0 = 3$ under H_0 and H_1. In each panel it seems hard to find distinguished characters of two paths, which exhibit quite similar behaviors. Nonetheless, we will see that under H_0 our test procedure effectively detect the Gaussianity of Z.

We simulate L independent paths of X under H_0 and H_1, yielding L values of $\mathcal{T}_n^{H_0}$ and $\mathcal{T}_n^{H_1}$, where $\mathcal{T}_n^{H_j}$ stands for $\mathcal{T}_n$ under H_j, $j = 0, 1$. Denote these values by

$$(\mathcal{T}_n^{H_0,l}, \mathcal{T}_n^{H_1,l}), \quad l = 1, \dots, L.$$

Based on these values, we compute empirical sizes $\hat{\psi}_{n,L}$ and empirical powers $\hat{\pi}_{n,L}$ corresponding to the significance levels 5 and 1: the upper 5 (resp. 1) percentile of the $\chi^2(2)$ distribution is given by 5.991 (resp. 9.21). Specifically, for the 5% significance level, $\hat{\psi}_{n,L}$ and $\hat{\pi}_{n,L}$ are defined to be

$$\begin{aligned} \hat{\psi}_{n,L} &= \sharp\{l \le L : \ \mathcal{T}_n^{H_0} > 5.991\}/L, \\ \hat{\pi}_{n,L} &= \sharp\{l \le L : \ \mathcal{T}_n^{H_1} > 5.991\}/L, \end{aligned}$$

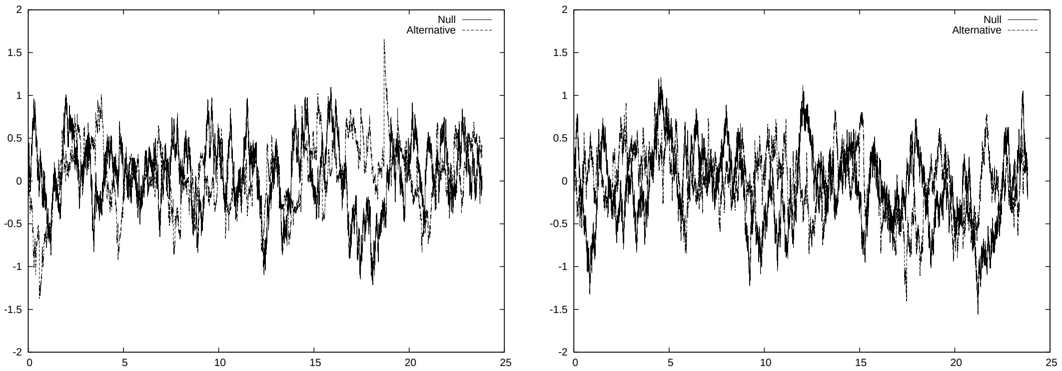

Figure 1: Plots of sample paths of X with $\theta_0 = 3$ under H_0 and H_1: the left panel corresponds to $(\alpha, \delta) = (3, 3)$, and the right panel to $(\alpha, \delta) = (10, 10)$.

and similarly for the 1% case. From Theorem 2.4 we have $\hat{\psi}_n \to 0.05$ (or $\to 0.01$, according as the significance level) and $\hat{\pi}_{n,L} \to 1$ for L and n getting larger.

Table 1 reports the resulting performances of $\mathcal{T}_n^{H_0}$ with $L = 1000$ and several choices of n. There, also mentioned are empirical means and standard deviations (S.D.) of $\mathcal{T}_n^{H_0}$ computed from $\mathcal{T}_n^{H_0,l}$, $l = 1, \dots, 1000$: both of them are expected to be close to 2, since both of the mean and variance of the exponential distribution $\chi^2(2)$ equal 2. Certainly we see that the asymptotic behavior of $\mathcal{T}_n$ under H_0 are consistent with the first half of Theorem 2.4. Also, Figure 2 shows a standardized histogram based on $\{\mathcal{T}_{1000}^{H_0,l}\}_{l=1}^{L}$, where we now set $L = 5000$ in order to get a more reliable result. The histogram exhibits a good fit to the targeted $\chi^2(2)$-density given by the straight line.

n	T_n	5%-$\hat{\psi}_{n,L}$	1%-$\hat{\psi}_{n,L}$	Mean of $\mathcal{T}_n^{H_0}$	S.D. of $\mathcal{T}_n^{H_0}$
100	6.3096	0.0500	0.0230	1.8697	2.7766
300	9.7915	0.0460	0.0160	1.9669	2.2773
500	12.0112	0.0440	0.0180	1.9533	2.3290
1000	15.8489	0.0530	0.0090	1.9792	2.0438

Table 1: Empirical sizes $\hat{\psi}_{n,L}$ (behaviors of $\mathcal{T}_n^{H_0}$) for $L = 1000$ with different (n, T_n).

Also, Table 2 reports the resulting performances of $\mathcal{T}_n^{H_1}$, again with $L = 1000$ and several choices of n. From the table we can observe very good performances of $\mathcal{T}_n^{H_1}$ for rejecting H_0 in case of $(\alpha, \delta) = (3, 3)$. As for the case of $(\alpha, \delta) = (10, 10)$, the empirical powers badly behave for smaller n, nevertheless, drastically become better with n increases. These numerical results strongly suggest that our test procedure has pretty good power despite of its ease of implementation.

n	T_n	$(\alpha, \delta) = (3, 3)$		$(\alpha, \delta) = (10, 10)$	
		5%-$\hat{\pi}_{n,L}$	1%-$\hat{\pi}_{n,L}$	5%-$\hat{\pi}_{n,L}$	1%-$\hat{\pi}_{n,L}$
100	6.3096	0.8590	0.7990	0.1620	0.0970
300	9.7915	1.0000	1.0000	0.5930	0.4730
500	12.0112	1.0000	1.0000	0.9030	0.8390
1000	15.8489	1.0000	1.0000	1.0000	1.0000

Table 2: Empirical powers $\hat{\pi}_{n,L}$ (behaviors of $\mathcal{T}_n^{H_1}$) for $L = 1000$ with different (n, T_n) and (α, δ).

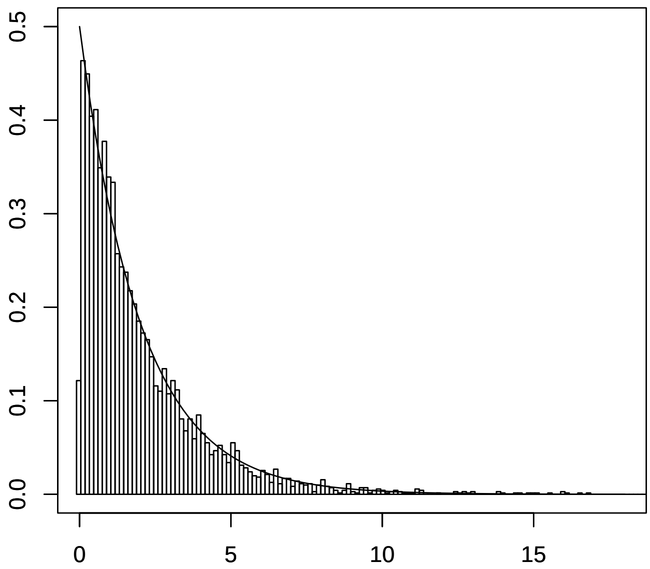

Figure 2: Standardized histogram of $\mathcal{T}_{1000}^{H_0}$ based on $L = 5000$ independent estimates. The straight line indicates the $\chi^2(2)$-density.

4 Proof of Theorem 2.4

We will write $a_n \lesssim b_n$ for a random sequence a_n and b_n if there exists a positive constant C such that $a_n \leq C b_n$, P_0-a.s., for every n large enough. It will be convenient to introduce the following notation:

$$\hat{H}_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n} \hat{\epsilon}_{ni}^k \quad (\text{hence } \hat{H}_n^{(1)} = \bar{\hat{\epsilon}}_n), \quad H_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n} \epsilon_{ni}^k,$$

$$\eta_{ni} = \int_{t_{i-1}^n}^{t_i^n} (b(X_s, \theta_0) - b_{i-1}(\hat{\theta}_n))ds \quad \left(\text{hence } \hat{\epsilon}_{ni} = \frac{1}{\sqrt{h_n}}(\Delta_i^n Z + \eta_{ni}) = \epsilon_{ni} + \frac{\eta_{ni}}{\sqrt{h_n}}\right),$$

$$M_n^{(k,l)} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} (\Delta_i^n Z)^k (\eta_{ni})^l.$$

Then we have

$$\hat{\Psi}_n^{(k)} = \hat{H}_n^{(k)} - k\hat{H}_n^{(1)}\hat{H}_n^{(k-1)} + \sum_{j=2}^{k} \binom{k}{j} (-\hat{H}_n^{(1)})^j \hat{H}_n^{(k-j)}, \tag{8}$$

and

$$\hat{H}_n^{(k)} = H_n^{(k)} + \frac{1}{\sqrt{n}} k h_n^{-k/2} M_n^{(k-1,1)} + \frac{1}{\sqrt{n}} \sum_{j=2}^{k} \binom{k}{j} h_n^{-k/2} M_n^{(k-j,j)}. \tag{9}$$

4.1 Proof of (a): asymptotic behavior under H_0

Our proof is carried out in a similar way to Kulperger and Yu (2005).

Under H_0 we have $Z = \sigma w$ (recall (2)), hence $(\epsilon_{ni})_{i=1}^n$ forms an i.i.d. array with common normal law $\mathcal{N}_1(0, \sigma^2)$ with $\sigma^2 > 0$ unknown, and $\frac{1}{n}\sum_{i=1}^n |\epsilon_{ni}|^l = O_p(1)$ for every $l \in \mathbb{N}$; here and in the sequel, $\mathcal{N}_1(\gamma, \Sigma)$ denotes the univariate normal distribution with mean γ and variance Σ. We will derive stochastic expansions of $\hat{\Psi}_n^{(k)}$ in (8) up to order $O_p(1/\sqrt{n})$. Write $\bar{\theta}_n = \sqrt{T_n}(\hat{\theta}_n - \theta_0) = O_p(1)$. We have

$$\begin{aligned}
\eta_{ni} &= -h_n\{b_{i-1}(\hat{\theta}_n) - b_{i-1}(\theta_0)\} + \int_{t_{i-1}^n}^{t_i^n} \{b(X_s, \theta_0) - b_{i-1}(\hat{\theta}_0)\} ds \\
&= -\frac{h_n}{\sqrt{T_n}} \partial_\theta^\top b_{i-1}(\theta_0)\bar{\theta}_n + \bigg\{ \int_{t_{i-1}^n}^{t_i^n} \{b(X_s, \theta_0) - b_{i-1}(\theta_0)\} ds \\
&\quad - \frac{1}{n}\bar{\theta}_n^\top \left(\int_0^1 u \int_0^1 \partial_\theta^2 b_{i-1}(\theta_0 + uv(\hat{\theta}_n - \theta_0)) dv du \right) \bar{\theta}_n \bigg\} \\
&= -\frac{h_n}{\sqrt{T_n}} \partial_\theta^\top b_{i-1}(\theta_0)\bar{\theta}_n + \eta'_{ni}, \quad \text{say.}
\end{aligned} \tag{10}$$

First let us look at the second term on the right-hand side of (9). From (10),

$$\begin{aligned}
h_n^{-k/2} M_n^{(k-1,1)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \epsilon_{ni}^{k-1} \frac{1}{\sqrt{h_n}} \eta_{ni} \\
&= -\frac{1}{n} \sum_{i=1}^n \epsilon_{ni}^{k-1} \partial_\theta^\top b_{i-1}(\theta_0)\bar{\theta}_n + \frac{1}{\sqrt{T_n}} \sum_{i=1}^n \epsilon_{ni}^{k-1} \eta'_{ni}.
\end{aligned} \tag{11}$$

Hölder's inequality yields that for every $q \geq 2$

$$\begin{aligned}
\sum_{i=1}^n |\eta'_{ni}|^q &\lesssim h_n^q \sum_{i=1}^n \frac{1}{h_n} \int_{t_{i-1}^n}^{t_i^n} |X_s - X_{t_{i-1}^n}|^q ds + n^{1-q} \left(\frac{1}{n} \sum_{i=1}^n \sup_{\theta \in \Theta} |\partial_\theta^2 b_{i-1}(\theta)|^q \right) |\bar{\theta}_n|^{2q} \\
&= O_p(n h_n^{3p/2}) + O_p(n^{1-q}) = O_p(n h_n^{3q/2} \vee n^{1-q}),
\end{aligned} \tag{12}$$

since under H_0 it hold that $\sup_{|t-s| \leq h} E_0[|X_t - X_s|^q] \lesssim h^{q/2}$ for every $h \leq 1$. Therefore Schwarz's inequality gives

$$\begin{aligned}
\left| \frac{1}{\sqrt{T_n}} \sum_{i=1}^n \epsilon_{ni}^{k-1} \eta'_{ni} \right| &\leq \frac{1}{\sqrt{h_n}} \left(\frac{1}{n} \sum_{i=1}^n |\epsilon_{ni}|^{2(k-1)} \right)^{1/2} \left(\sum_{i=1}^n |\eta'_{ni}|^2 \right)^{1/2} \\
&= O_p\left(\sqrt{n h_n^2} \vee \frac{1}{\sqrt{T_n}} \right) = o_p(1).
\end{aligned}$$

Accordingly, it follows from (11) that

$$h_n^{-k/2} M_n^{(k-1,1)} = -\frac{1}{n} \sum_{i=1}^n \epsilon_{ni}^{k-1} \partial_\theta^\top b_{i-1}(\theta_0)\bar{\theta}_n + o_p(1). \tag{13}$$

Next, for each $j \in \{2, \dots, k\}$, by using (12) as before we get

$$\begin{aligned}
|h_n^{-k/2} M_n^{(k-j,j)}| &\leq \frac{1}{\sqrt{n}} h_n^{-j/2} \sum_{i=1}^n |\epsilon_{ni}|^{k-j} |\eta'_{ni}|^j \\
&\leq \left(\frac{1}{n} \sum_{i=1}^n |\epsilon_{ni}|^{2(k-j)} \right)^{1/2} \left(h_n^{-j} \sum_{i=1}^n |\eta'_{ni}|^{2j} \right)^{1/2} \\
&= O_p\left(\sqrt{n h_n^{2j}} \vee \sqrt{T_n^{-j} n^{1-j}} \right) = o_p(1).
\end{aligned} \tag{14}$$

This implies that the third term on the right-hand side of (9) is $o_p(1/\sqrt{n})$. Substituting (13) and (14) in (9), we get for each $k \in \mathbb{N}$

$$\begin{aligned}\hat{H}_n^{(k)} &= H_n^{(k)} - \frac{k}{\sqrt{n}}\bigg\{\frac{1}{n}\sum_{i=1}^n \epsilon_{ni}^{k-1}\partial_\theta^\top b_{i-1}(\theta_0)\bigg\}\bar{\theta}_n + o_p\bigg(\frac{1}{\sqrt{n}}\bigg)\\ &=: H_n^{(k)} - \frac{k}{\sqrt{n}}F_n^{(k-1)}\bar{\theta}_n + o_p\bigg(\frac{1}{\sqrt{n}}\bigg), \quad \text{say.}\end{aligned} \tag{15}$$

Clearly $F_n^{(l)} = O_p(1)$ for every $l \geq 0$ under the assumptions, so that the second term of the right-hand side of (15) is $O_p(1/\sqrt{n})$. In particular, we have $\hat{H}_n^{(k)} = O_p(1)$ for every $k \in \mathbb{N}$ since $H_n^{(k)} = O_p(1)$ under H_0. Hence we arrive at

$$\hat{\Psi}_n^{(k)} = \hat{H}_n^{(k)} - k\hat{H}_n^{(1)}\hat{H}_n^{(k-1)} + O_p\bigg(\frac{1}{n}\bigg) \tag{16}$$

in view of (8).

Now, letting $\tilde{H}_n^{(k)} := H_n^{(k)} - E[\epsilon_{n1}^k]$ (in particular, $\tilde{H}_n^{(1)} = H_n^{(1)}$), we have $\tilde{H}_n^{(k)} = O_p(1/\sqrt{n})$ by the classical central limit theorem for i.i.d. arrays. Hence (15) and (16) lead to

$$\begin{aligned}\hat{\Psi}_n^{(k)} &= H_n^{(k)} - \frac{k}{\sqrt{n}}F_n^{(k-1)}\bar{\theta}_n + o_p\bigg(\frac{1}{\sqrt{n}}\bigg)\\ &\quad - k\bigg\{H_n^{(1)} - \frac{1}{\sqrt{n}}F_n^{(0)}\bar{\theta}_n + o_p\bigg(\frac{1}{\sqrt{n}}\bigg)\bigg\}\\ &\quad \cdot\bigg\{H_n^{(k-1)} - \frac{k-1}{\sqrt{n}}F_n^{(k-2)}\bar{\theta}_n + o_p\bigg(\frac{1}{\sqrt{n}}\bigg)\bigg\} + O_p\bigg(\frac{1}{n}\bigg)\\ &= E[\epsilon_{n1}^k] + \frac{1}{\sqrt{n}}\{\sqrt{n}\tilde{H}^{(k)} - kE[\epsilon_{n1}^{k-1}]\sqrt{n}\tilde{H}_n^{(1)}\}\\ &\quad + \frac{k}{\sqrt{n}}\{E[\epsilon_{n1}^{k-1}]F_n^{(0)} - F_n^{(k-1)}\}\bar{\theta}_n + o_p\bigg(\frac{1}{\sqrt{n}}\bigg).\end{aligned}$$

However, we see that $Q_n^{(k-1)} := E[\epsilon_{n1}^{k-1}]F_n^{(0)} - F_n^{(k-1)} \to^p 0$ since both of its mean and variance converge to 0: indeed, writing $E_0^{i-1}[\cdot] = E_0[\cdot|\mathcal{F}_{t_{i-1}^n}]$ and supposing $p = 1$ (the dimension of θ) without loss of generality, we have

$$E_0[Q_n^{(k-1)}] = \frac{1}{n}\sum_{i=1}^n E_0\big[(E_0^{i-1}[E[\epsilon_{n1}^{k-1}] - \epsilon_{ni}^{k-1}])\partial_\theta b_{i-1}(\theta_0)\big] = 0,$$

and

$$\begin{aligned}&E_0[\{Q_n^{(k-1)}\}^2]\\ &= \frac{1}{n^2}\sum_{i=1}^n E_0[\{\partial_\theta b_{i-1}(\theta_0)\}^2]\mathrm{var}[\epsilon_{n1}^{k-1}]\\ &\quad + \frac{2}{n^2}\sum_{i<j} E_0\Big[\{\partial_\theta b_{i-1}(\theta_0)\partial_\theta b_{j-1}(\theta_0)(E_0[\epsilon_{n1}^{k-1}] - \epsilon_{ni}^{k-1})\}E_0^{j-1}[E_0[\epsilon_{n1}^{k-1}] - \epsilon_{nj}^{k-1}]\Big]\\ &\lesssim \frac{1}{n^2}\sum_{i=1}^n E_0[|\partial_\theta b_{i-1}(\theta_0)|^2] = O\bigg(\frac{1}{n}\bigg) = o(1).\end{aligned}$$

Thus we get

$$\hat{\Psi}_n^{(k)} = E[\epsilon_{n1}^k] + \frac{1}{\sqrt{n}}\{\sqrt{n}\tilde{H}^{(k)} - kE[\epsilon_{n1}^{k-1}]\sqrt{n}\tilde{H}_n^{(1)}\} + o_p\bigg(\frac{1}{\sqrt{n}}\bigg), \tag{17}$$

rendering that $\hat{\theta}_n$ does not appear in $\hat{\Psi}_n^{(k)}$ up to order $O_p(1/\sqrt{n})$.

Since the kth self-normalized partial sum $\hat{\Phi}_n^{(k)}$ is invariant under scale change of $\hat{\epsilon}_{ni}$, in order to investigate the asymptotic behavior of $\hat{\Phi}_n^{(k)}$ we may consider $\sigma^{-k}\hat{\Psi}_n^{(k)}$ in place of $\hat{\Psi}_n^{(k)}$, so that we may set $\sigma = 1$ without loss of generality. Let ρ_k denote the kth moment of the standard normal distribution. Then (17) is rewritten as

$$\hat{\Psi}_n^{(k)} = \rho_k + \frac{1}{\sqrt{n}}\{\sqrt{n}\tilde{H}^{(k)} - k\rho_{k-1}\sqrt{n}\tilde{H}_n^{(1)}\} + o_p\left(\frac{1}{\sqrt{n}}\right), \tag{18}$$

and in particular,

$$\{\hat{\Psi}_n^{(2)}\}^{k/2} = 1 + \frac{k}{2}\frac{1}{\sqrt{n}}\sqrt{n}\tilde{H}_n^{(2)} + o_p\left(\frac{1}{\sqrt{n}}\right) \tag{19}$$

in view of the Taylor expansion $f(x+y) = f(x)+f'(x)y+\int_0^1\{f'(x+uy)-f'(x)\}du\cdot y$ for $f(x) = x^{k/2}$ with $k \geq 2$. From (18) and (19), expanding the fraction we get

$$\hat{\Phi}_n^{(k)} = \rho_k + \frac{1}{\sqrt{n}}\left[\sqrt{n}\tilde{H}_n^{(k)} - k\rho_{k-1}\sqrt{n}\tilde{H}_n^{(1)} - \frac{k}{2}\rho_k\sqrt{n}\tilde{H}_n^{(2)}\right] + o_p\left(\frac{1}{\sqrt{n}}\right),$$

hence arriving at

$$\sqrt{n}(\hat{\Phi}_n^{(k)} - \rho_k) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{(\epsilon_{ni}^k - \rho_k) - k\rho_{k-1}\epsilon_{ni} - \frac{k}{2}\rho_k(\epsilon_{ni}^2 - 1)\right\} + o_p(1). \tag{20}$$

By means of the expression (20), it is straightforward to deduce

$$\sqrt{n}(\hat{\Phi}_n^{(3)}, \hat{\Phi}_n^{(4)} - 3)^\top \to^d \mathcal{N}_2(0, \mathrm{diag}(6, 24)).$$

Theorem 2.4 (a) now follows on applying the continuous mapping theorem.

4.2 Proof of (b): asymptotic behavior under H_1

In view of the definition of $\mathcal{T}_n$, it suffices to prove that $|\sqrt{n}\hat{\Phi}_n^{(3)}| = |(\sqrt{n}\hat{\Psi}_n^{(3)})/(\hat{\Psi}_n^{(2)})^{3/2}| \to^p \infty$ under H_1 (here, $|\sqrt{n}\hat{\Phi}_n^{(3)}| \to^p \infty$ means that $P_0[|\sqrt{n}\hat{\Phi}_n^{(3)}| > a] \to 1$ for every $a > 0$). To this end we are going to look at the expressions

$$\hat{\Psi}_n^{(2)} = \hat{H}_n^{(2)} - (\hat{H}_n^{(1)})^2, \tag{21}$$

$$\sqrt{n}\hat{\Psi}_n^{(3)} = \sqrt{n}\{\hat{H}_n^{(3)} - 3\hat{H}_n^{(1)}\hat{H}_n^{(2)} + 2(\hat{H}_n^{(1)})^3\}. \tag{22}$$

We can write $Z_t = \sigma w_t + \int_0^t z\tilde{\mu}(ds, dz) =: \sigma w_t + J_t$ with $E[J_t] = 0$ (recall (2)). It follows from the independence between w and J that $\frac{1}{h_n}E[Z_{h_n}^2] = \sigma^2 + E[J_1^2] = \sigma^2 + \int |z|^2\nu(dz) = \mathrm{var}[Z_1]$, and that for each integer $m \geq 3$

$$\begin{aligned}\frac{1}{h_n}E[Z_{h_n}^m] &= \frac{1}{h_n}E[J_{h_n}^m] + \binom{m}{2}\sigma^2 E[J_{h_n}^{m-2}] \\ &\quad + \sum_{\substack{3\leq j\leq m \\ j:\ \mathrm{even}}}\binom{m}{j}\{\sigma^j(j-1)!!\}h_n^{j/2-1}E[J_{h_n}^{m-j}] \\ &\sim \frac{1}{h_n}E[J_{h_n}^m],\end{aligned} \tag{23}$$

where $a_n' \sim a_n''$ means that $a_n'/a_n'' \to 1$ and $a_n''/a_n' \to 1$. Combining (23) and Asmussen and Rosiński (2001, Lemma 3.1), we see that $\frac{1}{h_n}E[|\Delta_1^n Z|^{m'}] \to \int |z|^{m'}\nu(dz)$ for each even $m' \geq 3$. Based on these observations, we see from the central limit theorem that

$$\sqrt{T_n}\left\{\frac{1}{T_n}\sum_{i=1}^{n}(\Delta_i^n Z)^q - \frac{1}{h_n}E[(\Delta_1^n Z)^q]\right\} \to^d \begin{cases}\mathcal{N}_1\left(0, \sigma^2 + \int |z|^2\nu(dz)\right), & q = 1, \\ \mathcal{N}_1\left(0, \int |z|^{2q}\nu(dz)\right), & q \geq 2,\end{cases} \tag{24}$$

for $q \in \mathbb{N}$: in particular, $\frac{1}{T_n}\sum_{i=1}^n (\Delta_i^n Z)^q = O_p(1)$ for every $q \in \mathbb{N}$.

Let us note that, differently from the case of H_0, we have $\sup_{|t-s|\le h} E_0[|X_t - X_s|^q] \lesssim h$ for $q \ge 2$ and $h \le 1$ under H_1. Observe that for every $q \ge 2$

$$
\begin{aligned}
\sum_{i=1}^n |\eta_{ni}|^q &\lesssim \sum_{i=1}^n h_n^q \left\{ \left| \frac{1}{h_n}\int_{t_{i-1}^n}^{t_i^n} (b(X_s,\theta_0) - b_{i-1}(\theta_0))ds \right|^q + |b_{i-1}(\hat{\theta}_n) - b_{i-1}(\theta_0)|^q \right\} \\
&\lesssim h_n^{q-1}\sum_{i=1}^n \int_{t_{i-1}^n}^{t_i^n} |X_s - X_{t_{i-1}^n}|^q ds + |\bar{\theta}_n|^q n h_n^q T_n^{-q/2}\left(\frac{1}{n}\sum_{i=1}^n \sup_{\theta\in\Theta}|\partial_\theta b_{i-1}(\theta)|^q\right), \\
&= O_p(nh_n^{q+1} \vee nh_n^q T_n^{-q/2}) = O_p(nh_n^{q+1} \vee n^{1-q/2}h_n^{q/2}).
\end{aligned} \tag{25}
$$

Hence, for any integer $l \ge 2$ we get

$$
|M_n^{(0,l)}| \le \frac{1}{\sqrt{n}}\sum_{i=1}^n |\eta_{ni}|^l = O_p(\sqrt{n}h_n^{l+1} \vee n^{(1-l)/2}h_n^{l/2}). \tag{26}
$$

Moreover, the stochastic orders of $M_n^{(k,l)}$ for $k, l \in \mathbb{N}$ are estimated as follows:

$$
\begin{aligned}
|M_n^{(k,l)}| &\le \left(\frac{1}{T_n}\sum_{i=1}^n |\Delta_i^n Z|^{2k}\right)^{1/2}\left(h_n \sum_{i=1}^n |\eta_{ni}|^{2l}\right)^{1/2} \\
&= O_p(\sqrt{n}h_n^{l+1} \vee n^{(1-l)/2}h_n^{(l+1)/2}).
\end{aligned} \tag{27}
$$

Also, from (24) we have $H^{(1)} = O_p(1/\sqrt{n})$ and $\sqrt{T_n}(H_n^{(2)} - \mathrm{var}[Z_1]) \to^d \mathcal{N}_1(0, \int |z|^4\nu(dz))$. Thus

$$
\hat{H}_n^{(1)} = H_n^{(1)} + \frac{1}{\sqrt{n}}kh_n^{-1/2}M_n^{(0,1)} = O_p\left(\frac{1}{\sqrt{n}}\right) + \frac{1}{\sqrt{n}}O_p(\sqrt{nh_n^3} \vee 1) = O_p\left(\frac{1}{\sqrt{n}}\right), \tag{28}
$$

and moreover, $\hat{H}_n^{(2)} \to^p \mathrm{var}[Z_1]$ because

$$
\begin{aligned}
\hat{H}_n^{(2)} &= H_n^{(2)} + \frac{1}{\sqrt{n}}\left(\frac{k}{h_n}M_n^{(1,1)} + \frac{1}{h_n}M_n^{(0,2)}\right) \\
&= H_n^{(2)} + \frac{1}{\sqrt{n}}\left\{O_p(\sqrt{nh_n^2} \vee 1) + O_p\left(\sqrt{nh_n^4} \vee \frac{1}{\sqrt{n}}\right)\right\} = H_n^{(2)} + O_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned} \tag{29}
$$

(in particular, we have $\sqrt{T_n}(\hat{H}_n^{(2)} - \mathrm{var}[Z_1]) \to^p \mathcal{N}_1(0, \int |z|^4\nu(dz))$.) Hence it remains to show $|\sqrt{n}\hat{\Psi}_n^{(3)}| \to^p \infty$, and this in turn amounts to showing $|\sqrt{n}\hat{H}_n^{(3)}| \to^p \infty$ since we have

$$
\sqrt{n}\hat{\Psi}_n^{(3)} = \sqrt{n}\hat{H}_n^{(3)} + O_p(1)
$$

in view of (22), (28) and (29).

Just like (28) and (29), it follows from (9), (26), and (27) that

$$
\begin{aligned}
\sqrt{n}\hat{H}_n^{(3)} &= \frac{1}{h_n}\left\{\sqrt{n}h_n H_n^{(3)} + O_p\left(h_n^{3/2} \vee \sqrt{\frac{h_n}{n}}\right)\right\} \\
&= \frac{1}{h_n}\left[\sqrt{T_n}\left\{\frac{1}{T_n}\sum_{i=1}^n (\Delta_i^n Z)^3 - \frac{1}{h_n}E[(\Delta_1^n Z)^3]\right\} + \left(\frac{1}{h_n}E[(\Delta_1^n Z)^3]\right)\sqrt{T_n} + o_p(1)\right].
\end{aligned} \tag{30}
$$

Differentiating the characteristic function of $\mathcal{L}(Z_h)$ three times, we get $E[(\Delta_1^n Z)^3] = h_n \int z^3\nu(dz)$, from which combined with (24) and (30) we deduce that:

- $\sqrt{n}h_n\hat{H}_n^{(3)} \to^d \mathcal{N}_1(0, \int |z|^6\nu(dz))$ if $\int z^3\nu(dz) = 0$; while
- $|\sqrt{n}h_n\hat{H}_n^{(3)}| \to^p \infty$ if $\int z^3\nu(dz) \ne 0$.

In either case we get $|\sqrt{n}\hat{H}_n^{(3)}| \to^p \infty$ as was to be shown. The proof of Theorem 2.4 (b) is thus complete.

References

[1] Asmussen, S. and Rosiński, J. (2001), Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.* **38**, 482–493.

[2] Barndorff-Nielsen, O. E., Mikosch, T. and Resnick, S. I. (2001), *Lévy Processes. Theory and Applications.* Birkhäuser, Boston.

[3] Bera, A. and Jarque, C. (1981), Efficient tests for normality, heteroskedasticity, and serial independence of regression residuals: monte carlo evidence. *Econometric Lett.* **7**, 313–318.

[4] Bickel, P. and Rosenblatt, M. (1973), On some global measures of the deviations of density function estimates. *Ann. Statist.* **1**, 1071–1095.

[5] Cont, P. and Tankov, P. (2004), *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, London.

[6] D'Agostino, R. B. and Stephens, M. A. (1986), *Goodness-of-Fit Techniques.* Marcel Dekker, New York.

[7] Florens-Zmirou, D. (1989), Approximate discrete time schemes for statistics of diffusions. *Statistics* **4**, 547–557.

[8] Horváth, L. and Zitikis, R. (2006), Testing goodness of fit based on densities of GARCH innovations. *Econometric theory* **22**, 457–482.

[9] Jarque, C. M. and Bera, A. K. (1980), Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econom. Lett.* **6**, 255–259.

[10] Karatzas, I. and Shreve, S. E. (1991), *Brownian Motion and Stochastic Calculus. Second Edition.* Springer-Verlag, New York.

[11] Kessler, M. (1997), Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24**, 211–229.

[12] Kulik, A. M., Exponential ergodicity of the solutions to SDE's with a jump noise. To appear in *Stochastic Process. Appl.*

[13] Kulperger, R. and Yu, H. (2005), High moment partial sum processes of residuals in GARCH models and their applications. *Ann. Statist.* **33**, 2395–2422.

[14] Kutoyants, Y. (2004), *Statistical Inference for Ergodic Diffusion Processes.* Springer, New York.

[15] Lee, S. (2006), The Bickel-Rosenblatt test for diffusion processes. *Statist. Probab. Lett.* **76**, 1494–1502.

[16] Lee, S. and Na, S. (2001), On the Bickel-Rosenblatt test for first-order autoregressive models. *Statist. Probab. Lett.* **56**, 23–35.

[17] Lee, S. and Taniguchi, M. (2005), Asymptotic theory for ARCH-SM Models: LAN and residual empirical processes. *Statistica Sinica* **15**, 215–234.

[18] Lee, S. and Wei, C. Z. (1999), On residual empirical processes of stochastic regression models with application to time series. *Ann. Statist.* **27**, 237–261.

[19] Lee, S. and Wee, I.-S. (2008), Residual empirical process for diffusion processes. *J. Korean Math. Soc.* **45**, 683–693.

[20] Liptser, R. S. and Shiryaev, A. N. (2001), *Statistics of Random Processes. I and II. Second, Revised and Expanded Editions.* Springer-Verlag, Berlin.

[21] Masuda, H. (2005), Simple estimators for parametric Markovian trend of ergodic processes based on sampled data. *J. Japan Statist. Soc.* **35**, 147–170.

[22] Masuda, H. (2007), Ergodicity and exponential β-mixing bound for multidimensional diffusions with jumps. *Stochastic Processes Appl.* **117**, 35–56. [Erratum to this paper: to appear from *Stochastic Processes Appl.*]

[23] Prakasa Rao, B. L. S. (1999), *Statistical Inference for Diffusion Type Processes.* Arnold, London.

[24] Schoutens, W. (2003), *Lévy Processes in Finance.* Wiley, New York.

[25] Shiryayev, A. N. (1999), *Essentials of Stochastic Finance. Facts, Models, Theory.* World Science, Singapore.

[26] Yoshida, N. (1992), Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.* **41**, 220–242.

List of MI Preprint Series, Kyushu University

The Grobal COE Program
Math-for-Industry Education & Research Hub

MI

2008-1 Takahiro ITO, Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Abstract collision systems simulated by cellular automata

2008-2 Eiji ONODERA
The intial value problem for a third-order dispersive flow into compact almost Hermitian manifolds

2008-3 Hiroaki KIDO
On isosceles sets in the 4-dimensional Euclidean space

2008-4 Hirofumi NOTSU
Numerical computations of cavity flow problems by a pressure stabilized characteristic-curve finite element scheme

2008-5 Yoshiyasu OZEKI
Torsion points of abelian varieties with values in nfinite extensions over a p-adic field

2008-6 Yoshiyuki TOMIYAMA
Lifting Galois representations over arbitrary number fields

2008-7 Takehiro HIROTSU & Setsuo TANIGUCHI
The random walk model revisited

2008-8 Silvia GANDY, Masaaki KANNO, Hirokazu ANAI & Kazuhiro YOKOYAMA
Optimizing a particular real root of a polynomial by a special cylindrical algebraic decomposition

2008-9 Kazufumi KIMOTO, Sho MATSUMOTO & Masato WAKAYAMA
Alpha-determinant cyclic modules and Jacobi polynomials

2008-10 Sangyeol LEE & Hiroki MASUDA
Jarque-Bera Normality Test for the Driving Lévy Process of a Discretely Observed Univariate SDE