UNIVERSITÀ DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
E METODI MATEMATICI°

QUADERNI DEL DIPARTIMENTO

Department of Information Technology and Mathematical Methods

Working Paper

Series "*Mathematics and Statistics*"

n. 7/MS – 2009

Goodness of fit test for ergodic diffusions by tick time sample scheme

by

Ilia Negri, Yoichi Nishiyama

° Viale Marconi. 5, I – 24044 Dalmine (BG), ITALY, Tel. +39-035-2052339; Fax. +39-035-562779

Goodness of fit test for ergodic diffusions by tick time sample scheme

Ilia Negri

University of Bergamo

Viale Marconi, 5, 24044, Dalmine (BG), Italy

ilia.negri@unibg.it

and

Yoichi Nishiyama *

The Institute of Statistical Mathematics

4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan

nisiyama@ism.ac.jp

May 21, 2009

Running title: Goodness of fit test for ergodic diffusions

Abstract

We consider a nonparametric goodness of fit test problem for the drift coefficient of one-dimensional ergodic diffusions, where the diffusion coefficient is a nuisance function which is estimated in some sense. Using a theory for the continuous observation case, we construct a test based on the data observed discretely in space, that is, the so-called tick time sampled data. It is proved that the asymptotic distribution of our test under the null hypothesis is the supremum of the standard Brownian motion, and thus our test is asymptotically distribution free. It is also shown that the test is consistent under any fixed alternative.

Keywords. Ergodic diffusion process, tick time sample, invariance principle, asymptotically distribution free test.

AMS Mathematics Subject Classifications (1991). 62G10, 62G20, 62M02.

*Corresponding author. Tel: +81-3-5421-8726, Fax: +81-3-5421-8796.

1 Introduction

Goodness of fit tests play an important role in theoretical and applied statistics, and the study for them has a long history. Such tests are really useful especially if they are *distribution free*, in the sense that their distributions do not depend on the underlying model. The origin goes back to the Kolmogorov-Smirnov and Crámer-von Mises tests in the i.i.d. case, established early in the 20th century, which are *asymptotically distribution free*. Although their great importance in application, the theory of goodness of fit tests for diffusion processes has not received much attention from researchers until several years ago. Kutoyants [11] discussed some possibilities of the construction of such tests in his Section 5.4, where he considered the Kolmogorov-Smirnov statistics based on the continuous time observation of a diffusion process. The goodness of fit test based on the Kolmogorov-Smirnov statistic is asymptotically consistent and the asymptotic distribution under the null hypothesis follows from the weak convergence of the empirical process to a suitable Gaussian process, but this test is not asymptotically distribution free. Note that the Kolmogorov-Smirnov statistic for ergodic diffusion process was studied in Fournie [4], see also Fournie and Kutoyants [5] for more details, while the weak convergence of the empirical process was proved in Negri [14] (see van der Vaart and van Zanten [20] for further developments). Dachian and Kutoyants [1] and Negri and Nishiyama [15] proposed some asymptotically distribution free tests. Recently Kutoyants [12] proposed some Crámer-von Mises type tests based on the empirical distribution function and the local time estimator of the invariant density; the proposed test is asymptotically distribution free after a suitable transformation of the test statistics. However, all these results are based on *continuous time observation* of the diffusion processes.

One of the interesting points of this paper is that the proposed test is based on *tick time* sample scheme of observations. Tick time sample scheme, roughly speaking, consists in observing the underling process only when the process reaches some fixed values of a suitable grid in the state space. The moments when the process reaches those values are called tick times. Tick time sample arises in many problems in finance, when for example, the prices are sampled with every *continued price changes* in bid or ask quotation data. See for example the work of Fukasawa [6] and reference therein. Usually in finance the most common scheme is one where the prices are sampled at regular interval in calendar time. We should mention that there is a huge literature on discrete time approximations of statistical estimators for diffusion processes; see e.g. the Introduction of Gobet *et al.* [7] for a review including not only high frequency cases but also low frequency cases. On the other hand, with the increasing availability of transaction data alternative sampling scheme, such as tick time sampling and transaction time sampling, has gain popularity. See Griffin and Oomen [8] for an interesting discussion on these different sample schemes. We are interested in this new research direction.

In this work, we extend the approach taken by Negri and Nishiyama [15] who considered the continuous time observation case to the tick time sample case; a similar attempt in the high frequency sampled case can be found in Masuda *et al*

[13] and Nishiyama [18]. The method is essentially based on a certain marked empirical process to construct an asymptotically distribution free test, where "marked empirical process" actually means a random field of innovation martingales, and we will show that with some necessary modifications it can be applied to the tick time sample case.

The approach based on the marked empirical process goes back to and it is motivated by the work of Koul and Stute [10] who considered a non-linear parametric time series model (see also Section 7.3 of Nishiyama [17] which is a reprint of his thesis in 1998). They studied the large sample behavior of the proposed test statistics under the null hypotheses and present a martingale transformation of the underlying process that makes tests based on it asymptotically distribution free. Some considerations on consistency have also been done. The approach is well expounded in Koul [9]. See Delgado and Stute [2] and references therein for more recent information.

Now we turn to the description of the problem treated in this paper. Consider a one-dimensional stochastic differential equation (SDE)

$$X_t = X_0 + \int_0^t S(X_s)ds + \int_0^t \sigma(X_s)dW_s, \tag{1}$$

where the initial value X_0 is finite almost surely, S and σ are functions which satisfy some properties described later, and $t \rightsquigarrow W_t$ is a standard Wiener process defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. We consider a case where a unique strong solution X to this SDE exists, and we shall assume that X is ergodic. We are interested in goodness of fit test for the drift coefficient S, while the diffusion coefficient σ^2 is an unknown nuisance function whose expected value with respect to the invariant law is estimated in our testing procedure. That is, we consider the problem of testing hypothesis $H_0 : S = S_0$ versus $H_1 : S \neq S_0$ for a given S_0. The meaning of the alternatives "$S \neq S_0$" will be precisely stated in Section 2.

The data is supposed to be sampled as follows. For every $T > 0$, let $\bigcup_p (a_p^T, a_{p+1}^T]$ be a countable partition of the state space of the diffusion process. We assume that $\inf_p |a_{p+1}^T - a_p^T| > 0$ for each T. The process $X = \{X_t : t \in [0, \infty)\}$ is observed at random times $0 = \tau_0^T \leq \tau_1^T < \cdots < \tau_{N(T)}^T < \tau_{N(T)+1}^T = T$, where

$$\tau_1 = \inf\{t > 0 : X_t = a_p^T \text{ for some } p\}$$

and

$$\tau_i = \inf\{t > \tau_{i-1}^T : X_t = a_{p+1}^T \text{ or } a_{p-1}^T \text{ where } X_{\tau_{i-1}^T} = a_p^T\}, \quad \text{for } i = 2, ..., N(T).$$

We suppose that $h_T = o(T^{-1/2})$ as $T \to \infty$, where $h_T = \sup_p |a_{p+1}^T - a_p^T|$. We will prove that $N(T) < \infty$ almost surely for every T; the precise formulation including the case where $\tau_1^T \geq T$ will be described in Section 3. We will propose an asymptotically distribution free test based on this sampling scheme, namely, *tick time sample scheme*.

The organization of the article is as follows. In Section 2, we begin with the continuous time observation case which was considered by Negri and Nishiyama

[15]. Using the result for the continuous time observation case, our main results for the tick time sample case are given in Section 3. Throughout Sections 2 and 3 the proofs for "Theorems" are stated there, and those for "Lemmas' will be given in Section 6. Section 4 contains a proposal of an ad-hoc "estimator" for a value which we have to estimate in our testing procedure. In Section 5 we present some computer simulation results, including the one for the ad-hoc estimator proposed in Section 4. The Appendix contains some known results which are used in the main text.

Let us close this section with making some conventions. We denote by $\ell^\infty(\mathbb{T})$ the class of bounded functions on a set $\mathbb{T}$, and equip the space with the uniform metric. We denote by "$\to^p$" and "$\to^d$" the convergence in probability and in distribution as $T \to \infty$, respectively. The notation "$\to$" means that we take the limit as $T \to \infty$. See e.g. van der Vaart and Wellner [19] for the weak convergence theory in the $\ell^\infty(\mathbb{T})$ space.

2 Continuous observation case

Throughout all this paper we shall assume the following.

A1. The diffusion process X, which is a solution to the SDE (1) for (S, σ), is regular, and the speed measure $m_{S,\sigma}$ is finite. (Thus the process X is ergodic.) The invariant density $f_{S,\sigma}$ satisfies that

$$\Sigma_{S,\sigma}^2 := \int_{-\infty}^{\infty} \sigma(z)^2 f_{S,\sigma}(z) dz \in (0, \infty).$$

$\diamond$

We consider the stochastic process $V^T = \{V^T(x); x \in [-\infty, \infty]\}$ defined by

$$V^T(x) = \frac{1}{\sqrt{T}} \int_0^T 1_{(-\infty,x]}(X_t)[dX_t - S_0(X_t)dt].$$

Negri and Nishiyama [15] called (a slightly different version of) this process the "score marked empirical process", and obtained the following result, which is a fruit of the combination of the weak convergence theory for ℓ^∞-valued continuous martingales based on the *metric entropy* developed by Nishiyama [16], [17] and a theorem for local time of ergodic diffusion processes given by van Zanten [21] (see also van der Vaart and van Zanten [20]). To consider the "metric entropy" we introduce the metric $\rho_{S,\sigma}$ on $[-\infty, \infty]$ given by

$$\rho_{S,\sigma}(x, y) = \sqrt{\int_{x \wedge y}^{x \vee y} (\sigma(z)^2 f_{S,\sigma}(z) + \phi(z)) dz}$$

where ϕ is the density of the standard Gaussian distribution. Without ϕ, the above $\rho_{S,\sigma}$ may not define a *metric* but just define a *semimetric*. Nishiyama's [17] weak convergence theory requires that ρ is a metric, so we have added the Gaussian density ϕ. It is easy to see that the space $[-\infty, \infty]$ is compact and the metric entropy condition is satisfied:

$$\int_0^1 \sqrt{\log N([-\infty, \infty], \rho_{S,\sigma}; \varepsilon)} d\varepsilon < \infty.$$

Here, when a metric space $(\mathbb{T}, \rho)$ is given, $N(\mathbb{T}, \rho; \varepsilon)$ denotes the smallest number of closed balls with ρ-radius ε which cover $\mathbb{T}$.

Lemma 1 *Assume* **A1** *for (S_0, σ), and set $\rho = \rho_{S_0,\sigma}$. Then $V^T \to^d G$ in $\ell^\infty([-\infty, \infty])$, where $G = \{G(x); x \in [-\infty, \infty]\}$ is a zero-mean Gaussian process with co-variance given by*

$$EG(x)G(y) = \int_{-\infty}^{x \wedge y} \sigma(z)^2 f_{S_0,\sigma}(z) dz.$$

Almost all paths of G are uniformly ρ-continuous.

Corollary 2 *Assume* **A1** *for (S_0, σ). Then*

$$\sup_{x \in [-\infty,\infty]} |V^T(x)| \to^d \sup_{t \in [0,\Sigma^2]} |B_t| =^d \Sigma \sup_{t \in [0,1]} |B_t|,$$

where $t \rightsquigarrow B_t$ is a standard Brownian motion, $\Sigma = \Sigma_{S_0,\sigma}$, and where the notation "$=^d$" means that the distributions are the same.

In order to obtain an asymptotically distribution free test, we need a consistent estimator for $\Sigma_{S_0,\sigma}$. In the usual context of the continuous observation the diffusion coefficient is assumed to be known, so we can compute $\Sigma_{S_0,\sigma}$ for the given S_0. Even if the computation is not possible, we will propose a consistent estimator for $\Sigma_{S_0,\sigma}$ in Section 3. In any case, we have the following theorem.

Theorem 3 *Assume* **A1** *for (S_0, σ). Under $H_0 : S = S_0$, suppose that a positive, consistent estimator $\widehat{\Sigma}^T$ for $\Sigma_{S_0,\sigma}$ be given. Then it holds that*

$$D^T := \frac{\sup_{x \in [-\infty,\infty]} |V^T(x)|}{\widehat{\Sigma}^T} \to^d \sup_{t \in [0,1]} |B_t|, \tag{2}$$

where $t \rightsquigarrow B_t$ is a standard Brownian motion.

It is well known that the distribution function of the limit is given by

$$F(x) = P\left(\sup_{t \in [0,1]} |B_t| \le x \right) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left(-\frac{(2n+1)^2 \pi^2}{8x^2} \right); \tag{3}$$

see e.g. page 343 of Feller [3].

Negri and Nishiyama [15] further proved the consistency of (a slightly different version of) D^T under the fixed alternative $S \in \mathcal{S}$ where $\mathcal{S}$ is the class of functions S which satisfies **A1** and

$$\int_{-\infty}^{x_S} (S(z) - S_0(z)) f_{S,\sigma}(z) dz \neq 0 \quad \text{for some } x_S \in (-\infty, \infty]. \tag{4}$$

Theorem 4 *Assume* **A1** *for* (S, σ). *Under* $H_1 : S \in \mathcal{S}$, *if* $\widehat{\Sigma}^T = O_P(1)$, *then it holds for any* $M > 0$ *that* $P(D^T \leq M) \to 0$, *where* D^T *is the statistic defined on (2).*

The method of the proof will appear in that for the extension to the tick time sample case given in the next section, so the proof of the above theorem is omitted here.

3 Tick time sample case

Now we list up some conditions for the pair of functions (S, σ) which are assumed in the tick time sample case.

A2. There exists a constant $K > 0$ such that

$$|S(x) - S(y)| \leq K|x - y|, \quad |\sigma(x) - \sigma(y)| \leq K|x - y|.$$

$\diamond$

Under this condition, the SDE (1) has a unique strong solution X. Notice also that there exists a constant $K' > 0$ such that

$$|S(x)| \leq K^{'}(1 + |x|), \quad |\sigma(x)| \leq K^{'}(1 + |x|).$$

To see this, just put $y = 0$. The constant K' depends on the values $S(0)$ and $\sigma(0)$, however the constant K itself depends on the choice of the functions (S, σ). So it is convenient to introduce the notation

$$K_{S,\sigma} = K \vee K'.$$

This notation will be used throughout this article.

A3. $\sup_{z \in (-\infty, \infty)} (1 + |z|^2) f_{S,\sigma}(z) < \infty$. $\diamond$

A4. $\sup_{t \in [0,\infty)} E|X_t|^2 < \infty$ where X is a solution to the SDE (1) for (S, σ). $\diamond$

In this section we consider the situation where the data is obtained in the following way. For every $T > 0$, let $\bigcup_p (a_p^T, a_{p+1}^T]$ be a countable partition of $(-\infty, \infty)$. We assume that $\inf_p |a_{p+1}^T - a_p^T| > 0$ for each T. Define the stopping times τ_i^T by

$$\begin{aligned}
\tau_0^T &= 0, \\
\tau_1^T &= \inf\{t > 0 : \; X_t = a_p^T \text{ for some } p\}, \\
\tau_i^T &= \inf\{t > \tau_{i-1}^T : \; X_t = a_{p+1}^T \text{ or } a_{p-1}^T \text{ where } X_{\tau_{i-1}^T} = a_p^T\}, \quad i \geq 2.
\end{aligned}$$

Now we define

$$N(T) = \sup\{i : \tau_i^T < T\}.$$

Lemma 5 *Under* **A2** *and* **A4**, *it holds that* $N(T) < \infty$ *almost surely.*

Notice that $N(T)$ may be zero. Including such a case, we re-define $\tau^T_{N(T)+1} := T$ for notational convenience.

Sample Scheme. The process $X = \{X_t; t \in [0,\infty)\}$ is observed at the stopping times $0 = \tau^T_0 \le \tau^T_1 < \cdots \tau^T_{N(T)} < \tau^T_{N(T)+1} = T$. We suppose that $h_T = o(T^{-1/2})$ as $T \to \infty$ where $h_T = \sup_p |a^T_{p+1} - a^T_p|$. $\diamond$

Now, we introduce an array of constants

$$-\infty = x^T_0 < x^T_1 < x^T_2 < \cdots < x^T_{m(T)} < x^T_{m(T)+1} = \infty$$

such that, as $T \to \infty$,

$$\max_{2\le k\le m(T)} |x^T_k - x^T_{k-1}| \to 0, \quad x^T_1 \downarrow -\infty, \quad x^T_{m(T)} \uparrow \infty.$$

For example, one may consider $x^T_k = -[T] + (k/[T])$ with $k = 1, 2, ..., 2[T]^2$. Next we introduce a sequence of functions $z \rightsquigarrow \psi^T_k(z)$ on $(-\infty,\infty)$ which approximates the indicator function $1_{(-\infty,x^T_k]}$.

Definition 6 *Let a sequence of positive constants b_T be given. For every $k = 1, 2, ..., m(T)$, ψ^T_k is the continuous, piecewise linear function on $(-\infty,\infty)$ defined by*

$$\psi^T_k(z) = \begin{cases} 1, & z \in (-\infty, x^T_k], \\ \textit{line}, & z \in [x^T_k, x^T_k + b_T], \\ 0, & z \in [x^T_k + b_T, \infty). \end{cases}$$

Also we define $\psi^T_0 \equiv 0$ and $\psi^T_{m(T)+1} \equiv 1$.

This function satisfies the following properties:

$$|\psi^T_k(z) - \psi^T_k(z')| \le b_T^{-1}|z - z'|;$$

$$|\psi^T_k(z) - 1_{(-\infty,x^T_k]}(z)| \le 1_{[x^T_k, x^T_k+b_T]}(z).$$

Now we make the following condition.

A5. In addition to $h_T = o(T^{-1/2})$, we assume the following:
(i) $b_T^{-1} h_T \cdot \sqrt{\log m(T)} \to 0$;
(ii) $b_T \log m(T) \to 0$. $\diamond$

Typically, $\log m(T) = O(\log T^\alpha)$ for some $\alpha > 0$. In this case, the above (i) and (ii) are satisfied if we take $b_T = T^{-1/4}$.

We approximate V^T by $U^T = \{U^T(x); x \in [-\infty,\infty]\}$ given by

$$U^T(x) := \frac{1}{\sqrt{T}} \sum_{i=1}^{N(T)+1} \psi^T_k(X_{\tau^T_{i-1}})[X_{\tau^T_i} - X_{\tau^T_{i-1}} - S_0(X_{\tau^T_{i-1}})|\tau^T_i - \tau^T_{i-1}|]$$

for $x \in (x^T_{k-1}, x^T_k]$, $1 \le k \le m(T) + 1$. Actually, we have the following lemma.

Lemma 7 *Assume* **A1** – **A5** *for* (S_0, σ). *Then it holds that*

$$\sup_{x\in[-\infty,\infty]} |V^T(x) - U^T(x)| \to^p 0.$$

A consistent estimator for $\Sigma_{S,\sigma}$ is given as follows.

Lemma 8 *Assume* **A1**, **A2** *and* **A4** *for* (S, σ). *The estimator*

$$\widehat{\Sigma}_2^T = \sqrt{\frac{1}{T}\sum_{i=1}^{N(T)+1} |X_{\tau_i^T} - X_{\tau_{i-1}^T}|^2}$$

is consistent for $\Sigma_{S,\sigma}$.

We therefore obtain the following result.

Theorem 9 *Assume* **A1** – **A5** *for* (S_0, σ). *Under* $H_0 : S = S_0$, *it holds that*

$$D_2^T := \frac{\sup_{x\in[-\infty,\infty]} |U^T(x)|}{\widehat{\Sigma}_2^T} \to^d \sup_{t\in[0,1]} |B_t|, \tag{5}$$

where $t \rightsquigarrow B_t$ *is a standard Brownian motion.*

Now let us turn to study the asymptotic behavior of D_2^T under a fixed alternative. We consider the class $\mathcal{S}$ consists of all functions satisfying **A1** – **A5** and (4). The precise description of our problem is testing the null hypothesis $H_0 : S = S_0$ versus the alternatives $H_1 : S \in \mathcal{S}$.

We may write $U^T = \Psi_2^T - \Phi_2^T(S_0)$ where

$$\Psi_2^T(x) = \frac{1}{\sqrt{T}} \sum_{i=1}^{N(T)} \int_{\tau_{i-1}^T}^{\tau_i^T} \psi_k^T(X_{\tau_{i-1}^T}) dX_t$$

and

$$\Phi_2^T(S)(x) = \frac{1}{\sqrt{T}} \sum_{i=1}^{N(T)} \int_{\tau_{i-1}^T}^{\tau_i^T} \psi_k^T(X_{\tau_{i-1}^T}) S(X_{\tau_{i-1}^T}) dt$$

for $x \in (x_{k-1}^T, x_k^T], 1 \le k \le m(T) + 1$. Fix $S \in \mathcal{S}$. Then we have

$$\sup_{x\in[-\infty,\infty]} |U^T(x)| \ge \sup_{x\in[-\infty,\infty]} |\Phi_2^T(S)(x) - \Phi_2^T(S_0)(x)| - \sup_{x\in[-\infty,\infty]} |\Psi_2^T(x) - \Phi_2^T(S)(x)|.$$

Under $H_1 : S \in \mathcal{S}$, the second term on the right hand side is $O_P(1)$ by Lemmas 1 and 7 with S_0 replaced by S. As for the first term, we have the following lemma.

Lemma 10 *Assume* $S \in \mathcal{S}$ *and choose* $x_S \in (-\infty, \infty]$ *as in (4). Assume* **A2** *for* (S_0, σ). *Then we have*

$$\frac{1}{\sqrt{T}}(\Phi_2^T(S)(x_S) - \Phi_2^T(S_0)(x_S)) \to^p \int_{-\infty}^{x_S} (S(z) - S_0(z)) f_{S,\sigma}(z) dz \ne 0.$$

Thus we have that for any $M > 0$

$$P\left(\sup_{x\in[-\infty,\infty]} |\Phi_2^T(S)(x) - \Phi_2^T(S_0)(x)| \le M\right) \to 0.$$

We therefore obtain the consistency of the test.

Theorem 11 *Assume* **A1** – **A5** *for* (S,σ) *and* **A2** *for* (S_0,σ)*. Under* $H_1 : S \in \mathcal{S}$*, it holds for any* $M > 0$ *that* $P(D_2^T \le M) \to 0$*, where* D_2^T *is the statistic defined on (5).*

4 Bias corrected estimator for $\Sigma_{S_0,\sigma}$

In the preceding section, we proposed the estimator $\widehat{\Sigma}_2^T$ which is consistent for $\Sigma_{S,\sigma}$ for any (S,σ). However, one may think that it is better to use the *bias corrected* estimator $\widehat{\Sigma}_3^T$, which is consistent for $\Sigma_{S_0,\sigma}$ only for the null hypothesis "S_0", given as follows:

$$\widehat{\Sigma}_3^T = \sqrt{\frac{1}{T}\sum_{i=1}^{N(T)+1} |X_{\tau_i^T} - X_{\tau_{i-1}^T} - S_0(X_{\tau_{i-1}^T})|\tau_i^T - \tau_{i-1}^T||^2}.$$

This estimator satisfies the following; the proofs are not difficult, so they are omitted.

(i) Under $H_0 : S = S_0$, it holds that $\widehat{\Sigma}_3^T \to^p \Sigma_{S_0,\sigma}$.

(ii) For general drift S satisfying the Lipschitz condition **A2**, it holds that $\widehat{\Sigma}_3^T = O_P(1)$.

Based on these facts, we have the same conclusions as Theorems 9 and 11 with D_2^T replaced by

$$D_3^T := \frac{\sup_{x\in[-\infty,\infty]} |U^T(x)|}{\widehat{\Sigma}_3^T}.$$

The performance of this test statistic will be reported in the next section.

5 Simulation study

In this section we observe finite sample performance of our test statistics. We consider the following stochastic differential equation as true data generating model

$$X_t = \int_0^t S(X_s)ds + W_t \tag{6}$$

with $S(x) = -2x$, that is an Ornstein-Uhlenbeck process starting from the origin. As the null hypotheses, we consider the following two cases: $H_0 : S_0(x) = -2x$ and $H_0 : S_0(x) = 2 - 2x$.

In our experimental design we consider asymptotic for $h_T \to 0$, considering the values for h_T respectively equal to 0.2, 0.1 and 0.05. The asymptotics for $T \to \infty$ is considered taking T equals to 1, 5 and 10 respectively. We will consider both the test statistics D_2^T and D_3^T. We take the significance level to be $\alpha = 0.05$. We see that $F(x) = 0.95$ when $x = 2.24$, where F is the limit distribution under the null hypothesis given by (3), hence the critical region is $\{x > 2.24\}$. If the asymptotic conditions are realized $P(D_i^T > 2.24)$ should tend to 0.05 under $H_0 : S_0(x) = -2x$ and should tend to 1 under $H_0 : S_0(x) = 2 - 2x$, for $i = 2, 3$. For every configuration in our experimental design we simulate $m = 1000$ trajectories of (6) and we compute the empirical size (e.s.) defined by the sampling proportion of making incorrect rejections of the null $H_0 : S_0(x) = -2x$ and the empirical power (e.p.) defined by the sampling proportion of making successful rejection of the null $H_0 : S_0(x) = 2 - 2x$.

For simplicity we consider the case with $b_T = 0$, where b_T is the sequence introduced in condition **A5**. This case may be more interesting, from a practical point of view, because our theory cannot confirm the asymptotic results when the test statistics is based on the more natural random field $\widetilde{U}^T = \{\widetilde{U}^T(x); x \in [-\infty, \infty]\}$ given by the following

$$\widetilde{U}^T(x) = \frac{1}{\sqrt{T}} \sum_{i=1}^{N(T)+1} 1_{(-\infty,x]}(X_{\tau_{i-1}^T})[X_{\tau_i^T} - X_{\tau_{i-1}^T} - S_0(X_{\tau_{i-1}^T})|\tau_i^T - \tau_{i-1}^T|].$$

We conjecture that the same asymptotic results would also hold for $\widetilde{U}^T$, as the simulation study will show.

Table 1 and Table 2 summarizes the simulation results. We observe that for a fixed h_T empirical power gains along with increasing T. For a fixed T empirical power gains along with decreasing h_T, but gains less. This suggests that the ergodicity assumption is important. Moreover the bias corrected estimator for $\Sigma_{S_0,\sigma}$ that appear in D_3^T does not seem to give better asymptotic results than the simpler one. We report the mean number of observation for each sample scheme.

	$T = 1$		T=5		T=10	
h_T	e.s.	e.p.	e.s.	e.p.	e.s.	e.p.
0.2	0.034	0.376	0.038	0.989	0.041	1.000
	($n = 25$)		($n = 122$)		($n = 242$)	
0.1	0.046	0.397	0.038	0.988	0.047	1.000
	($n = 97$)		($n = 481$)		($n = 962$)	
0.05	0.059	0.418	0.044	0.998	0.053	1.000
	($n = 373$)		($n = 1862$)		($n = 3724$)	

Table 1: Empirical sizes (e.s.) and empirical powers (e.p.) based on 1000 independent statistics of D_2^T. Here the significance level is 0.05, and $b_T = 0$. In the bracket the mean number of observation is reported.

	$T=1$		T=5		T=10	
h_T	e.s.	e.p.	e.s.	e.p.	e.s.	e.p.
0.2	0.019	0.223	0.035	0.987	0.039	1.000
	$(n=25)$		$(n=122)$		$(n=242)$	
0.1	0.041	0.367	0.037	0.987	0.047	1.000
	$(n=97)$		$(n=481)$		$(n=962)$	
0.05	0.058	0.406	0.044	0.998	0.053	1.000
	$(n=373)$		$(n=1862)$		$(n=3724)$	

Table 2: Empirical sizes (e.s.) and empirical powers (e.p.) based on 1000 independent statistics of D_3^T. Here the significance level is 0.05, and $b_T = 0$. In the bracket the mean number of observation is reported.

6 Proofs of Lemmas

Notations: For $x, y \geq 0$, the inequality $x \lesssim y$ means that there exists a universal constant $C > 0$ such that $x \leq Cy$. We denote $\mathrm{Log}\, m := \log(1+m)$.

Proof of Lemma 1. This is a special case of Theorem 2 of Negri and Nishiyama [15] which is an application of the weak convergence theory of Nishiyama [16], [17] to the family $M^T = \{M^{T,x}; x \in [-\infty, \infty]\}$ of continuous martingales $t \rightsquigarrow M_t^{T,x}$ given by

$$M_t^{T,x} = \frac{1}{\sqrt{T}} \int_0^t 1_{(-\infty,x]}(X_s)\sigma(X_s)dW_s,$$

with help from Theorem 3.1 of van Zanten [21]. □

Proof of Lemma 5. Allowing the possibility $N(T) = \infty$, we consider the random variable

$$\xi^T = \sum_{i:\ \tau_i^T \leq T} |X_{\tau_i^T} - X_{\tau_{i-1}^T}|^2 = \sum_{i:\ \tau_i^T \leq T} \left\{ |X_{\tau_i^T}|^2 - |X_{\tau_{i-1}^T}|^2 - 2X_{\tau_{i-1}^T}(X_{\tau_i^T} - X_{\tau_{i-1}^T}) \right\}.$$

By Itô's formula, we have

$$|X_{\tau_i^T}|^2 - |X_{\tau_{i-1}^T}|^2 = 2\int_{\tau_{i-1}^T}^{\tau_i^T} X_s dX_s + \int_{\tau_{i-1}^T}^{\tau_i^T} \sigma(X_s)^2 ds,$$

hence

$$\xi^T \leq \left| \sum_{i:\ \tau_i^T \leq T} \int_{\tau_{i-1}^T}^{\tau_i^T} (X_s - X_{\tau_{i-1}^T}) dX_s \right| + \int_0^T \sigma(X_s)^2 ds.$$

The first term is bounded by

$$\sum_{i:\ \tau_i^T \leq T} \left| \int_{\tau_{i-1}^T}^{\tau_i^T} (X_s - X_{\tau_{i-1}^T}) S(X_s) ds \right| + \left| \sum_{i:\ \tau_i^T \leq T} \int_{\tau_{i-1}^T}^{\tau_i^T} (X_s - X_{\tau_{i-1}^T}) \sigma(X_s) dW_s \right|$$

The expectation of the first term is bounded by

$$E\int_0^T h_T|S(X_s)|ds \le K_{S,\sigma}\int_0^T h_T E(1+|X_s|)ds < \infty,$$

while the expectation of the square of the second term is bounded by

$$E\int_0^T h_T^2\sigma(X_s)^2 ds \le K_{S,\sigma}\int_0^T h_T^2 E(1+|X_s|)^2 ds < \infty.$$

Thus $\xi^T < \infty$ almost surely. It follows from the assumption $\inf_p |a_{p+1}^T - a_p^T| > 0$ that $N(T) < \infty$ almost surely. Indeed, $\inf_p |a_{p+1}^T - a_p^T|^2(N(T)-1) \le \xi^T$. □

Proof of Lemma 7. Let us introduce the stochastic processes Y_1^T, Y_2^T and Y_3^T given by $Y_1^T(-\infty) = Y_2^T(-\infty) = Y_3^T(-\infty) = 0$ and

$$\begin{aligned}
Y_1^T(x) &= \frac{1}{\sqrt{T}}\sum_{i=1}^{N(T)}\int_{\tau_{i-1}^T}^{\tau_i^T}\psi_k^T(X_{\tau_{i-1}^T})[dX_t - S_0(X_t)dt],\\
Y_2^T(x) &= \frac{1}{\sqrt{T}}\sum_{i=1}^{N(T)}\int_{\tau_{i-1}^T}^{\tau_i^T}\psi_k^T(X_t)[dX_t - S_0(X_t)dt],\\
Y_3^T(x) &= \frac{1}{\sqrt{T}}\sum_{i=1}^{N(T)}\int_{\tau_{i-1}^T}^{\tau_i^T}1_{(-\infty,x_k^T]}(X_t)[dX_t - S_0(X_t)dt],
\end{aligned}$$

for $x \in (x_{k-1}^T, x_k^T]$, $1 \le k \le m(T)+1$. We will prove:

$$\sup_{x\in[-\infty,\infty]}|U^T(x) - Y_1^T(x)| \to^p 0; \tag{7}$$

$$\sup_{x\in[-\infty,\infty]}|Y_1^T(x) - Y_2^T(x)| \to^p 0; \tag{8}$$

$$\sup_{x\in[-\infty,\infty]}|Y_2^T(x) - Y_3^T(x)| \to^p 0; \tag{9}$$

$$\sup_{x\in[-\infty,\infty]}|Y_3^T(x) - V^T(x)| \to^p 0. \tag{10}$$

Proof of (7). We easily have

$$\begin{aligned}
&E\sup_x|U^T(x) - Y_1^T(x)|\\
&\le \frac{1}{\sqrt{T}}E\sum_{i=1}^{N(T)}\int_{\tau_{i-1}^T}^{\tau_i^T}|S_0(X_t) - S_0(X_{\tau_{i-1}^T})|dt\\
&\le \frac{1}{\sqrt{T}}E\sum_{i=1}^{N(T)}\int_{\tau_{i-1}^T}^{\tau_i^T}K_{S_0,\sigma}\sup_{t\in[\tau_{i-1}^T,\tau_i^T]}|X_t - X_{\tau_{i-1}^T}|dt\\
&\le \frac{1}{\sqrt{T}}E\int_0^T K_{S_0,\sigma}h_T dt\\
&\le K_{S_0,\sigma}\sqrt{T}h_T \to 0.
\end{aligned}$$

Proof of (8). We introduce the stopping time

$$\tau_T(H) = \inf\left\{t > 0 : \frac{\sup_{z\in(-\infty,\infty)} l_t(z)}{T} \geq H\right\}, \tag{11}$$

where l_t denotes the local time of X with respect to the speed measure $m_{S_0,\sigma}$. By Theorem 3.1 of van Zanten [21], for every $\varepsilon > 0$ there exists a constant $H > 0$ such that $\limsup_T P(\tau_T(H) < T) < \varepsilon$. So it is enough to see that $\max_{1\leq k\leq m(T)+1} |\xi_k^T| \to^p 0$ where

$$\xi_k^T = \frac{1}{\sqrt{T}} \sum_{i=1}^{n} \int_{\tau_{i-1}^T\wedge\tau_T(H)}^{\tau_i^T\wedge\tau_T(H)} (\psi_k^T(X_{\tau_{i-1}^T}) - \psi_k^T(X_t))\sigma(X_t)dW_t.$$

Clearly $\xi_{m(T)+1}^T = 0$. For every $1 \leq k \leq m(T)$, note that ξ_k^T is a terminal variable of a continuous martingale. To apply the exponential inequality for continuous martingales, let us compute the predictable variation of ξ_k^T:

$$\begin{aligned}
&\frac{1}{T}\sum_{i=1}^{N(T)} \int_{\tau_{i-1}^T\wedge\tau_T(H)}^{\tau_i^T\wedge\tau_T(H)} |\psi_k^T(X_{\tau_{i-1}^T}) - \psi_k^T(X_t)|^2\sigma(X_t)^2 dt \\
&\leq \frac{1}{T}\sum_{i=1}^{N(T)} \int_{\tau_{i-1}^T\wedge\tau_T(H)}^{\tau_i^T\wedge\tau_T(H)} b_T^{-2}|X_{\tau_{i-1}^T} - X_t|^2\sigma(X_t)^2 dt \\
&\leq \frac{1}{T} b_T^{-2}h_T^2 \int_{-\infty}^{\infty} l_{T\wedge\tau_T(H)}(z)\sigma(z)^2 f_{S_0,\sigma}(z)dz \cdot m_{S_0,\sigma}((-\infty,\infty)) \\
&\leq b_T^{-2}h_T^2 H\Sigma_{S_0,\sigma}^2 m_{S_0,\sigma}((-\infty,\infty)).
\end{aligned}$$

Hence, by Lemmas 12 and 13, we have

$$E \max_{1\leq k\leq m(n)} |\xi_k^n| \lesssim \sqrt{b_T^{-2}h_T^2 H\Sigma_{S_0,\sigma}^2 m_{S_0,\sigma}((-\infty,\infty))}\sqrt{\operatorname{Log} m(T)} \to 0.$$

Proof of (9). We introduce the stopping time $\tau_T(H)$ given by (11). Then, it is enough to see that $\max_{1\leq k\leq m(T)+1} |\xi_k^T| \to^p 0$ where

$$\xi_k^T = \frac{1}{\sqrt{T}} \sum_{i=1}^{N(T)} \int_{\tau_{i-1}^T\wedge\tau_T(H)}^{\tau_i^T\wedge\tau_T(H)} (\psi_k^T(X_t) - 1_{(-\infty,x_k^T]}(X_t))\sigma(X_t)dW_t.$$

Clearly $\xi_{m(T)+1}^T = 0$. To apply the exponential inequality for continuous martingales, let us compute the predictable variation of ξ_k^T:

$$\begin{aligned}
&\frac{1}{T}\sum_{i=1}^{N(T)} \int_{\tau_{i-1}^T\wedge\tau_T(H)}^{\tau_i^T\wedge\tau_T(H)} |\psi_k^T(X_t) - 1_{(-\infty,x_k^T]}(X_t)|^2\sigma(X_t)^2 dt \\
&\leq \frac{1}{T}\int_0^{T\wedge\tau_T(H)} 1_{[x_k^T, x_k^T+b_T]}(X_t)\sigma(X_t)^2 dt \\
&\leq \frac{1}{T}\int_{x_k^T}^{x_k^T+b_T} l_{T\wedge\tau_T(H)}(z)\sigma(z)^2 f_{S_0,\sigma}(z)dz \cdot m_{S_0,\sigma}((-\infty,\infty)) \\
&\leq b_T H \sup_z\{\sigma(z)^2 f_{S_0,\sigma}(z)\} \cdot m_{S_0,\sigma}((-\infty,\infty))
\end{aligned}$$

Hence, by Lemmas 12 and 13, we have

$$E \max_{1\leq k\leq m(n)} |\xi_k^T| \lesssim \sqrt{b_T H K_{S_0,\sigma} \sup_z\{(1+|z|)^2 f_{S_0,\sigma}(z)\} \cdot m_{S_0,\sigma}((-\infty,\infty))} \sqrt{\operatorname{Log} m(T)},$$

which tends to zero.

Proof of (10). It is sufficient to show that

$$\max_{1\leq k\leq m(T)+1} \sup_{x\in(x_{k-1}^T, x_k^T]} |V^T(x_k^T) - V^T(x)| \to^p 0.$$

Notice that the weak convergence result for the random fields $x \rightsquigarrow V^T(x)$ implies also the stochastic ρ-equicontinuity, that is, for every $\varepsilon, \eta > 0$ there exists $\delta > 0$ such that

$$\limsup_{T\to\infty} P\left(\sup_{\rho(x,y)<\delta} |V^T(x) - V^T(y)| > \varepsilon \right) < \eta.$$

Since $\rho(x,y) \leq (\sup_z(\sigma(z)^2 f_{S_0,\sigma}(z) + \phi(z)))\sqrt{|x-y|} \leq \text{constant}\sqrt{|x-y|}$, we have $\max_{2\leq k\leq m(T)} \rho(x_{k-1}^T, x_k^T) \to 0$. Also, it is clear that $\rho(-\infty, x_1^T) \to 0$ and $\rho(x_{m(T)}^T, \infty) \to 0$. Hence we have (10).

Now (7) – (10) have been proved, and the proof of Lemma 7 is finished. □

Proof of Lemma 8. By the same reason as that in the beginning of the proof of Lemma 5, it is enough to show that

$$\frac{1}{T}\sum_{i=1}^{N(T)} \int_{\tau_{i-1}^n}^{\tau_i^n} (X_s - X_{t_{i-1}^n}) dX_s \to^p 0$$

and

$$\frac{1}{T}\int_0^T \sigma(X_s)^2 ds \to^p \Sigma_{S,\sigma}^2.$$

The latter is nothing else than the ergodicity. The same computation as the proof of Lemma 5 yields the former. The proof is finished. □

Proof of Lemma 10. The proof is similar to and easier than that for Lemma 7, hence it is omitted. □

Appendix

We state the *exponential inequality* for continuous martingales.

Lemma 12 *Let M be a continuous martingale, and let τ be a bounded stopping time. For every $x, v > 0$ it holds that*

$$P\left(\sup_{t\in[0,\tau]} |M_t| > x,\ \langle M\rangle_\tau \leq v \right) \leq 2\exp\left(-\frac{x^2}{2v}\right).$$

The following lemma, the *maximal inequality* for general random variables, is used in connection with Lemma 12 and plays a key role in our approach.

Lemma 13 *Let $X_1, ..., X_m$ be arbitrary random variables which satisfy*

$$P(|X_i| > x) \leq 2\exp\left(-\frac{x^2}{b}\right)$$

for all x and i and a fixed constant $b > 0$. Then there exists a universal constant $C > 0$ such that

$$E\left(\max_{1\leq i\leq m} |X_i|\right) \leq C\sqrt{b}\sqrt{\log(1+m)}.$$

Proof. Use Lemmas 2.2.1 and 2.2.2 of van der Vaart and Wellner [19]. □

Acknowledgements. Part of this work was performed during the first author was a Visiting Associate Professor at the Institute of Statistical Mathematics in December 2008.

References

[1] Dachian, S. and Kutoyants, Yu.A. (2008). On the goodness-of-fit tests for some continuous time processes. In *Statistical Models and Methods for Biomedical and Technical Systems,* 395-413. Vonta, F., Nikulin, M., Limnios, N. and Huber-Carol C. (Eds), Birkhuser, Boston.

[2] Delgado, M. and Stute, W. (2008). Distribution-free specification tests of conditional models. *J. Econometrics* **143**, 37-55.

[3] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Volume II.* (2nd ed.) John Wiley & Sons, Inc., New York.

[4] Fournie, E. (1992). Un test de type Kolmogorov-Smirnov pour processus de diffusion ergodiques, *Rapporte de recherche*, **1696**, INRIA, Sophia-Antipolis.

[5] Fournie, E. and Kutoyants, Yu.A. (1993). Estimateur de la distance maximale pour le processus de diffusion ergodiques, *Rapporte de recherche*, **1952**, INRIA, Sophia-Antipolis.

[6] Fukasawa, M. (2008). Central limit theorem for the realized volatility based on tick time sampling. *To appear in Finance and Stochastics.*

[7] Gobet, E., Hoffmann, M. and Reiss, M. (2004). Nonparametric estimation of scalar diffusions based on low frequency data. *Ann. Statist.* **32**, 2223-2253.

[8] Griffin, J.E. and Oomen, R.C.A. (2008). Sampling returns for realized variance calculations: tick time or transaction time? *Econometric Review* **27**, 230-253.

[9] Koul, H. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models.* (2nd ed.) Springer, New York.

[10] Koul, H. and Stute. W. (1999). Nonparametric model checks in time series. *Ann. Statist.* **27**, 204-237.

[11] Kutoyants, Yu.A. (2004). *Statistical Inference for Ergodic Diffusion Processes.* Springer, London.

[12] Kutoyants, Yu.A. (2009). On the goodness-of-fit testing for ergodic diffusion processes. *Preprint.*

[13] Masuda, H., Negri, I. and Nishiyama, Y. (2008). Goodness of fit test for ergodic diffusions by discrete time observations: an innovation martingale approach. *Research Memorandum* **1069**, Inst. Statist. Math., Tokyo.

[14] Negri, I. (1998). Stationary distribution function estimation for ergodic diffusion process. *Statistical Inference for Stochastic Processes*, **1**, 61–84.

[15] Negri, I. and Nishiyama, Y. (2008). Goodness of fit test for ergodic diffusion processes. *To appear in Ann. Inst. Statist. Math. (Already published online.)*

[16] Nishiyama, Y. (1999). A maximal inequality for continuous martingales and M-estimation in a Gaussian white noise model. *Ann. Statist.* **27**, 675-696.

[17] Nishiyama, Y. (2000). *Entropy Methods for Martingales.* CWI Tract **128**, Centrum voor Wiskunde en Informatica, Amsterdam.

[18] Nishiyama, Y. (2009). Nonparametric goodness of fit tests for ergodic diffusion processes by discrete observations. *Research Memorandum* **1094**, Inst. Statist. Math., Tokyo.

[19] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer-Verlag, New York.

[20] van der Vaart, A.W. and van Zanten, H. (2005). Donsker theorems for diffusions: necessary and sufficient conditions. *Ann. Probab.* **33**, 1422-1451.

[21] van Zanten, H. (2003). On uniform laws of large numbers for ergodic diffusions and consistency of estimators. *Statist. Inference Stoch. Process.* **6**, 199-213.