On a test for constant volatility in continuous time financial models

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: holger.dette@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

Carsten von Lieres und Wilkau
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: carsten.von.lieres@ruhr-uni-bochum.de

March 13, 2001

Abstract

A new specification test for homoscedasticity in diffusion processes is proposed, which does not require specific knowledge of the functional form of the model. The corresponding test statistic has an asymptotic normal distribution under the null hypothesis of constant volatility and diverges at an appropriate rate under the alternative. In contrast to recent work the approach of the present paper does not require the specification of particular time points at which the hypothesis of homoscedasticity is checked. Moreover, the new test does not use nonparametric estimation techniques for estimating the variance function and is therefore independent of the specification of a particular smoothing parameter. The results are illustrated by a small simulation study and a data example is analyzed.

Keywords: model diagnostics, diffusion process, heteroscedasticity, pseudo residuals

1 Introduction

Itô diffusions are commonly used for representing asset prices, because the strong Markov property and the nondifferentiability of the paths capture the idea of no arbitrage opportunities [see e.g. Merton (1990)]. In general the diffusion $(X_t)_t$ is a solution of the stochastic differential equation

$$dX_t = b\,(t, X_t)\,dt + \sigma\,(t, X_t)\,dW_t \tag{1.1}$$

where $(W_t)_t$ is a standard Brownian motion. An appropriate pricing of derivative assets requires a correct specification of the functional form of the drift and variance and different models have been proposed in the literature for the different types of options [see e.g. Black and Scholes (1973), Vasicek (1977), Cox, Ingersoll, Ross (1985), Karatzas (1988), Constantinides (1992) or Duffie and Harrison (1993) among many others]. Parametric models are attractive among practioners, because they often admit a direct interpretation of the observed effects in terms of the parameters and the available information of the observations is increased by applying more efficient methods. However, economic theory typically does not give much information of the drift and variance and misspecification of such a model may lead to serious errors in the subsequent data analysis. For these reasons many authors propose to test the goodness-of-fit of the postulated model [see e.g. Azzalini and Bowman (1993), Ait-Sahalia (1996b), Zheng (1996) or Dette and Munk (1998a) among many others]. If the assumption of a parametric model cannot be justified nonparametric estimates for the drift and variance of the diffusion should be used, which are less efficient from an asymptotic point of view [see e.g. Genon-Catalot, Laredo and Picard (1992), Ait-Sahalia (1996a), Florens-Zmirou (1993) or Jiang and Knight (1997)].

It is the purpose of the present paper to develop a test for homoscedasticity or a specific parametric form of the variance function in a diffusion model of the form (1.1). This problem is of importance in theoretical finance because several continuous-time financial models considered in the literature assume a constant volatility [see e.g. Merton (1973) or Vasicek (1977)] or a specific parametric form of heteroscedasticity [see e.g. Cox, Ingersoll, Ross (1985) or Constantinides (1992)], and it is reasonable to check this assumption by a statistical test. Moreover, specific information about the structure of the variance function (for example a constant volatility) allows the application of more efficient procedures for analyzing the observed data.

We assume discretely observed data on a fixed time span, say $[0, 1]$, with increasing sample size. As pointed out by Corradi and White (1999) this model is appropriate for analyzing the pricing of European, American or Asian options. Following Dette and Munk (1998b) we use an appropriate estimator of the integrated variance function

$$M^2 = \int_0^1 \left\{ \sigma^2(t, X_t) - \int_0^1 \sigma^2(s, X_s) ds \right\}^2 dt \tag{1.2}$$

as a measure of heteroscedasticity in the diffusion model (1.1) and prove its asymptotic normality under the null hypothesis of homoscedasticity. It is also demonstrated that the method can be generalized to the problem of testing for a parametric form of the volatility function and a simulation study is presented which illustrates excellent finite sample properties of "a bootstrap" version of the proposed test.

We conclude this introduction with a brief discussion of the work of Ait-Sahalia (1996b) and Corradi and White (1999), which is most similar in spirit with the present paper. In contrast to the method proposed by Ait-Sahalia (1996b), who compared the density implied by a joint parametric specification for the drift and variance against a nonparametric estimate of the density, our approach does not require such a specification of the parametric model. Moreover, the test of Ait-Sahalia (1996b) requires a time span approaching infinity for an increasing sample size, while for our method the time span has to be fixed and the length of the discrete sampling interval converges to zero as the sample size increases. Thus in this sense the two methods are complementary. Corradi and White (1999) consider a similar model as discussed in this paper and compare a nonparametric estimator of the variance function [see Florens-Zmirou (1993)] with an estimator under the null hypothesis of homoscedasticity at a fixed number of specified points. Consequently, the finite sample size and power of the test of Corradi and White (1999) depend on the evaluation points and the test proposed by these authors is in fact a test for a constant variance of the diffusion at a fixed number of specified points in the time scale [see e.g. Müller (1992) for a similar method in the context of checking the functional form of the mean in a nonparametric regression model]. In contrast to this work the test proposed in the present paper is consistent against any alternative, for which the process $(\sigma^2(t, X_t))_{t\in[0,1]}$ is not constant. Moreover, our method does neither require the specification of particular evaluation points at which the variance function has to be estimated, nor uses a smoothing parameter for a nonparametric estimator of the variance function, because only estimates of the integrated variance function are required.

2 The test statistic and its asymptotic distribution

Let $(W_t, \mathcal{F}_t)_{t\geq 0}$ denote a standard Brownian motion $[\mathcal{F}_t = \sigma(W_s, 0 \leq s \leq t)]$ defined on an appropriate probability space $(\Omega, \mathcal{F}, P)$ and assume that the drift and variance function in the stochastic differential equation (1.1)

$$b \,:\, [0,1] \times \mathbb{R} \to \mathbb{R}$$
$$\sigma \,:\, [0,1] \times \mathbb{R} \to \mathbb{R}$$

are continuous functions satisfying

$$|b\,(t,x) - b\,(t,y)| + |\sigma\,(t,x) - \sigma\,(t,y)| \leq K\ |x-y| \tag{2.1}$$

for all $t \in [0,1], x, y \in \mathbb{R}$, and

$$|b\,(t,x)|^2 + |\sigma\,(t,x)|^2 \leq K^2\,(1+|x|^2) \tag{2.2}$$

for all $t \in [0,1], x \in \mathbb{R}$, where $K > 0$ is a fixed constant. It is well known that for a $\mathcal{F}_0$ measurable square integrable random variable ξ, which is independent of the Brownian motion $(W_t)_{t\in[0,1]}$, the assumptions (2.1) and (2.2) admit a unique strong solution $(X_t)_{t\in[0,1]}$ of the stochastic differential equation (1.1), with initial conditon $X_0 = \xi$, which is adapted to the filtration generated by the Brownian motion $(W_t)_{t\in[0,1]}$, see e.g. Karatzas and Shreve (1991) p. 289. Moreover, the solution can be represented as

$$X_t = \xi + \int_0^t b\,(s, X_s)\,ds + \int_0^t \sigma\,(s, X_s)\,dW_s \quad \text{a.s.,} \tag{2.3}$$

X_t is $\mathcal{F}_t$ measurable for all $t \in [0,1]$ and the paths $t \to X_t$ are almost surely continuous. Assume that we observe the diffusion only on the time span $[0,1]$ at discrete points $t_i = i/n\,(i = 1, \ldots, n)$. We are interested in testing the hypothesis of homoscedasticity

$$H_0 : \sigma^2(t,x) \;=\; \sigma^2 \qquad \forall\ t \in [0,1]\ ,\ \ \forall\ x \tag{2.4}$$

in the stochastic differential equation (1.1) under assumptions (2.1) and (2.2). For this purpose we note that the hypothesis of constant volatility in (2.4) holds if and only if

$$M^2 = 0 \quad \text{a.s.,} \tag{2.5}$$

where the random variable M^2 is defined in (1.2). Therefore it is reasonable to reject the hypothesis of homoscedasticity for large values of an appropriate estimator of M^2. In order to estimate M^2 from the observed data we define

$$T_{pn} := n^{p-1} \sum_{i=1}^{n-1} (X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^{2p} \quad ; \qquad p = 1, 2, \tag{2.6}$$

and a test statistic by

$$T_n := \frac{1}{3} T_{2n} - T_{1n}^2. \tag{2.7}$$

We assume that the drift and variance function satisfy a Lipschitz condition of order $\gamma > \frac{1}{2}$, i.e.

$$|b\,(t,x) - b\,(s,x)| + |\sigma(t,x) - \sigma(s,x)| \leq L\,|t-s|^\gamma \tag{2.8}$$

for all $s, t \in [0,1]$ and for some fixed constant $L > 0$. Moreover, if the initial condition ξ has an existing eighth moment, i.e.

$$E[|\xi|^8] < \infty, \tag{2.9}$$

then the following theorem shows that the statistic T_n consistently estimates the measure of heteroscedasticity M^2. The proof is deferred to the appendix.

Theorem 2.1. *If the assumptions (2.1), (2.2), (2.8) and (2.9) are satisfied, then the statistic T_n defined in (2.7) is a consistent estimator of M^2. More precisely, if $n \to \infty$, we have*

$$T_n - M^2 = O_p(n^{-1/2} \log n), \tag{2.10}$$

where the random variable M^2 is defined in (1.2).

Recall that the variance function $\sigma(t, X_t)$ in the stochastic differential equation (1.1) is a.s. constant (as a function of t) if and only if (2.5) holds and consequently the hypothesis of homoscedasticity is rejected for large values of the statistic T_n. Our second main result specifies the asymptotic distribution of T_n under the additional assumption that the variance function in (1.1) does not depend on X_t, i.e.

$$\sigma(t,x) = \sigma(t) \quad \forall\, x \in \mathbb{R}. \tag{2.11}$$

Note that this assumption includes the situation of homoscedasticity ($\sigma(t,x) = \sigma > 0 \quad \forall\, t \in [0,1]\ ,\ \forall x$) and that (2.11) implies that M^2 is a nonnegative constant random variable.

Theorem 2.2. *If (2.1), (2.2), (2.8), (2.9) and (2.11) are satisfied, then the statistic T_n defined in (2.7) is asymptotically normal distributed, i.e.*

$$\sqrt{n}(T_n - M^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \frac{32}{3} s_8 - 16 s_2 s_6 + 8 s_2^2 s_4), \tag{2.12}$$

where

$$s_{2j} = \int_0^1 \sigma^{2j}(t)dt, \quad j = 1,2,3,4. \tag{2.13}$$

Especially, if $(X_t)_{t\in[0,1]}$ is a diffusion defined by (1.1) with constant variance $\sigma^2(t,x) = \sigma^2 > 0$, then

$$\sqrt{n}T_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \frac{8}{3}\sigma^8). \tag{2.14}$$

Note that Theorem 2.2 provides a simple test for the hypothesis of homoscedasticity, by rejecting the hypothesis (2.4) whenever

$$\sqrt{\frac{3n}{8}} \frac{T_n}{T_{1n}^2} > u_{1-\alpha}, \tag{2.15}$$

where $u_{1-\alpha}$ is the $(1-\alpha)$ quantile of the standard normal distribution and T_{1n}^4 is used as an estimator of σ^8 in (2.14). By Theroem 2.2 this test has asymptotic level α and its consistency follows directly from Theorem 2.1. The performance of the test will be illustrated in the following section.

Remark 2.3. It is worthwhile to mention that the approach for testing homoscedasticity can easily be extended to the problem of testing for a more general structure of the variance function. For the sake of simple notation we assume (2.11) and only consider the problem of testing the hypothesis

$$H_0^* : \sigma^2(t,x) = \sum_{j=1}^{d} \alpha_j \sigma_j^2(t) \quad \forall\, t \in [0,1]\ \forall\, x \tag{2.16}$$

in the stochastic differential equation (1.1), where $\sigma_1^2, \ldots, \sigma_d^2$ are given nonnegative, linearly independent functions [note that the hypothesis of homoscedasticity corresponds to the case $d = 1, \sigma_1^2(t) \equiv 1$]. Define pseudo residuals

$$\Delta_i = n(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2 \quad i = 1, \ldots, n-1, \tag{2.17}$$

$\Delta = (\Delta_1, \ldots, \Delta_{n-1})^T, \alpha = (\alpha_1, \ldots, \alpha_d)^T$ and a design matrix

$$X = \Big(\sigma_j^2(\frac{i}{n})\Big)_{i=1,\ldots,n-1}^{j=1,\ldots,d} \in \mathbb{R}^{n-1\times d}.$$

Consider the least squares problem

$$\hat{\alpha} = \underset{\alpha\in\mathbb{R}^d}{\operatorname{argmin}}(\Delta - X\alpha)^T(\Delta - X\alpha) = (X^TX)^{-1}X^T\Delta$$

(note that the linear independence of the functions $\sigma_1^2, \ldots, \sigma_d^2$ implies that X has rank d) and define a test statistic for the hypothesis in (2.16) by

$$\hat{T}_n = \frac{1}{n}\{\frac{1}{3}\Delta^T\Delta - \Delta^TX(X^TX)^{-1}X^T\Delta\}.$$

Observing the definition of the pseudo residuals Δ_i in (2.17) it follows from the arguments given in the appendix that

$$E[\frac{1}{3n}\Delta^T\Delta] = \frac{1}{3}\sum_{i=1}^{n-1} E[(X_{\frac{i+1}{n}} - X_{\frac{1}{n}})^4] \approx \int_0^1 \sigma^4(t)dt.$$

Similary, we have for the j-th component of the vector $X^T\Delta$

$$E[\frac{1}{n}(X^T\Delta)_j] = E[\sum_{i=1}^{n-1}\sigma_j^2(\frac{i}{n})(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2] \approx \int_0^1 \sigma_j^2(t)\sigma^2(t)dt$$

and continuity properties of the variance function imply

$$\frac{1}{n}(X^TX) \approx \Big(\int_0^1 \sigma_j^2(t)\sigma_i^2(t)dt\Big)_{i,j=1}^d =: \Sigma.$$

These approximations motivate (a rigorous proof follows by a straightforward but tedious extension of the arguments given in the appendix)

$$\begin{aligned} E[\hat{T}_n] &\approx \int_0^1 \sigma^4(t)dt - s^T\Sigma^{-1}s \\ &= \min_{\alpha_1,\ldots,\alpha_d\in\mathbb{R}} \int_0^1 \{\sigma^2(t) - \sum_{j=1}^d \alpha_j\sigma_j^2(t)\}^2 dt, \end{aligned} \tag{2.18}$$

where

$$s = \Big(\int_0^1 \sigma_1^2(t)\sigma^2(t)dt, \ldots, \int_0^1 \sigma_d^2(t)\sigma^2(t)dt\Big)^T$$

and the last equality follows by a standard calculation from Hilbert space theory. Note that the right hand side vanishes if the hypothesis H_0^* is valid and consequently this hypothesis should be rejected for large values of the statistic $\hat{T}_n$. It can be shown by similar arguments as given in the proof of Theorem 2.2 that under the null hypothesis (2.16) we have

$$\sqrt{n}\hat{T}_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \frac{8}{3}\int_0^1 \sigma^8(t)dt) \tag{2.19}$$

and an estimator of the asymptotic variance is obtained from the observation that

$$E[(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^8] \approx 105(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(s)ds)^4 \approx \frac{105}{n^4}\sigma^8(\frac{i}{n})$$

[see formula (4.40) in the proof of Theorem 2.2 in the appendix], which gives

$$\hat{\tau}_n^2 = \frac{n^3}{105}\sum_{i=1}^{n-1}(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^8$$

as a consistent estimator for $\int_0^1 \sigma^8(t)dt$. Consequently the hypothesis of the parametric structure is rejected if

$$\sqrt{\frac{3n}{8}}\frac{\hat{T}_n}{\hat{\tau}_n} > u_{1-\alpha},$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution.

3 Finite sample properties

In order to study the finite sample properties of the new test consider at first the stochastic differential equation (1.1) with $b(t, x) \equiv 0$ and assume that the hypothesis of homoscedasticity $\sigma^2(t, x) = \sigma^2 > 0$ is valid. In this case the pseudo residuals $X_{\frac{i+1}{n}} - X_{\frac{i}{n}}$ are independent identically distributed with

$$X_{\frac{i+1}{n}} - X_{\frac{i}{n}} \sim \mathcal{N}(0, \frac{\sigma^2}{n}) \quad i = 1, \ldots, n-1. \tag{3.1}$$

The distribution of the test statistic $\sqrt{3n/8} T_n / T_{1n}^2$ is in this case equal to the distribution of the random variable

$$V_n = \sqrt{\frac{3n}{8}} \frac{\frac{n}{3} \sum_{i=1}^{n-1} Z_i^4 - (\sum_{i=1}^{n-1} Z_i^2)^2}{(\sum_{i=1}^{n-1} Z_i^2)^2}, \tag{3.2}$$

where $Z_1, \ldots, Z_{n-1}$ are independent identically standard normally distributed random variables. Note that the random variable on the right hand side is essentially an estimate of the kurtosis of a random variable with zero mean and it is well know that the normal approximation for this distribution is rather poor. Obviously, because bias is present for a nonvanishing drift, this observation carries over to the normal approximation for the statistic T_n defined in (2.7). For this reason we propose an alternative method for obtaining quantiles for the distribution of T_n. At first note that due to the assumptions (2.1) and (2.8) the pseudo residuals $X_{\frac{i+1}{n}} - X_{\frac{i}{n}}$ are approximately unbiased. Secondly, if this bias is neglected, it follows from the above discussion that under the hypothesis of homoscedasticity the distribution of T_n / T_{1n}^2 is scale invariant and we may assume without loss of generality $\sigma^2 = 1$. Now Theorem 2.1 shows that

$$T_{1n} \xrightarrow{P} 1$$

and for this reason we propose to use the quantiles of the statistic Z_n which is obtained if

$$\sqrt{3n/8} \cdot T_n$$

is evaluated with data generated by the standard Brownian motion (note that we do not estimate the variance). The quantiles of this distribution can easily be obtained via simulation and are listed for various values of n in Table 3.0. These results are based on 100000 simulation runs.

n	80%	90%	95%	97.5%
25	0.2148	0.6823	1.2645	1.9514
50	0.3846	0.9055	1.4810	2.1311
100	0.5098	1.0609	1.6303	2.448
200	0.6237	1.1642	1.6831	2.2119

Table 3.0. *Simulated quantiles of the statistic* $\sqrt{3n/8} T_n$ *for data generated from a standard Brownian motion.*

The normalized statistic $\sqrt{3n/8}T_n/T_{1n}^2$ obtained from the observed data is then compared with the quantiles of the simulated distribution and the hypothesis of homoscedasticity is rejected if

$$\sqrt{\frac{3n}{8}}\frac{T_n}{T_{1n}^2} > z_{1-\alpha} \tag{3.3}$$

where $z_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the simulated distribution (see Table 3.0).

Example 3.1. Our first example investigates the approximation of the level of the test defined in (3.3). Because of the scale invariance we restrict ourselves to the case $\sigma^2 = 1$ and considered the functions

$$b(t,x) = \begin{cases} x \\ \sin(5x) \\ x + x^{-1} \\ tx \\ x\sin(t) \\ xe^t \end{cases}$$

for the drift. The diffusion was usually "normalized" by $X_0 = 0$ a.s., except in the case $b(t,x) = x + x^{-1}$, where $X_0 = 1$ a.s. was used as initial value. Table 3.1 shows the simulated level of the test (3.3) for various sample sizes. The numbers in brackets denote the level obtained by the normal approximation. We observe a nonsatisfactory performance of the test using the quantiles of the standard normal distribution and a reasonable approximation of the level of the test (3.3) for all drift functions under consideration. It is remarkable that the quality of approximation does not change if an additional time parameter is included in the drift function.

$\sigma \equiv 1$	$n = 50$			$n = 100$			$n = 200$		
$b(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
0	19.94	10.02	4.68	20.14	9.98	4.32	19.93	9.46	4.61
x	20.29	9.59	4.14	19.76	9.47	4.24	20.40	9.69	.476
$\sin(5x)$	20.74	10.02	4.42	20.63	10.45	4.93	20.41	10.42	5.00
$x + \frac{1}{x}$	17.78	8.64	3.76	19.84	9.77	4.34	20.12	9.65	4.71
$t \cdot x$	20.55	9.94	4.61	21.26	10.48	4.94	19.92	9.85	4.94
$x\sin(t)$	20.37	10.31	4.64	20.80	9.81	4.73	20.50	10.10	4.67
xe^t	19.67	9.40	4.34	20.13	9.78	4.60	19.72	9.82	4.94

Table 3.1: *Simulated rejection probabilities of the test (3.3) for various sample sizes and drift functions ($\sigma^2 = 1$). The critical values $z_{1-\alpha}$ are obtained from Table 3.0. The numbers in brackets show the simulated level of the test using a normal approximation.*

Example 3.2. In this example we investigate the power of the proposed test (3.3). To this end we consider the drift functions $b(t,x) = x$ (Table 3.2) and $b(t,x) = xt$ (Table 3.3) as representative examples. For the heteroscedastic alternative we used the functions

$$\sigma(t,x) = \begin{cases} 1+x \\ 1+\sin(5x) \\ 1+xe^t \\ 1+x\sin(5t) \\ 1+tx \end{cases}$$

for the variance in the stochastic differential equation (1.1). The corresponding results are listed in Table 3.2 and 3.3 and show that the test detects heteroscedasticity in all considered cases. It is worthwhile to mention that the size of the power depends on the nonnegative random variable M^2 defined in (1.2), but not directly on the variance function. Note also (comparing the the first with the fifth and third alternative) that the inclusion of an additional time dependence in the variance function can yield a decrease or increase in power. However, comparing Table 3.2 with 3.3 we observe a decrease with respect to power in all cases, if the drift term contains an additional time component.

Acknowledgements. This work of the authors was supported by two grants of the Deutsche Forschungsgmeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen, Teilprojekt A2, Validierung von Hypothesen, De 502/9-1). The authors would also like to thank I. Gottschlich, who typed parts of this paper with considerable technical expertise.

$b(t,x)=x$	$n=50$			$n=100$			$n=200$		
$\sigma(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
$1+x$	75.45	64.79	53.26	86.11	78.41	70.63	92.59	87.97	83.26
$1+\sin(5x)$	99.93	99.75	99.09	100	100	99.99	100	100	100
$1+xe^t$	91.30	83.80	73.69	97.39	94.97	90.99	99.48	98.91	97.84
$1+x\sin(5t)$	71.59	59.11	47.33	82.39	72.76	63.44	88.72	83.13	77.49
$1+tx$	61.97	47.65	35.11	73.84	62.96	52.21	82.28	74.53	67.95

Table 3.2: *Simulated power of the test (3.3) for various sample sizes, drift function $b(t,x)=x$ and different variance functions.*

$b(t,x)=tx$	$n=50$			$n=100$			$n=200$		
$\sigma(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
$1+x$	72.88	60.69	48.93	81.76	73.39	64.30	90.12	84.67	78.69
$1+\sin(5x)$	99.90	99.65	99.29	99.99	99.98	99.92	100	100	100
$1+xe^t$	90.71	83.43	73.85	97.24	94.46	90.46	99.42	98.72	97.66
$1+x\sin(5t)$	67.32	53.87	42.00	78.03	67.71	57.29	86.65	79.78	73.06
$1+xt$	58.52	44.64	31.98	69.56	58.17	47.34	79.06	70.11	62.37

Table 3.3. *Simulated power of the test (3.3) for various sample sizes, drift function $b(t,x)=tx$ and different variance functions.*

4 Appendix: Proofs.

We begin with a decomposition of the diffusion $(X_t)_{t\in[0,1]}$

$$X_t = X_0 + A_t + M_t \quad (0 \le t \le 1)\ , \tag{4.1}$$

where the processes $(A_t)_{t\in[0,1]}$ and $(M_t)_{t\in[0,1]}$ are defined by

$$A_t := \int_0^t b\,(s, X_s)\,ds \quad \text{und} \quad M_t := \int_0^t \sigma\,(s, X_s)\,dW_s \quad (0 \le t \le 1)\,, \tag{4.2}$$

respectively. Our first auxiliary result gives estimates for the the moments of the increments $A_{t+h} - A_t$ and $M_{t+h} - M_t$.

Lemma 4.1. *If assumption (2.1) and (2.2) are satisfied and the solution $(X_t)_{t\in[0,1]}$ of the stochastic differential equation (1.1) is decomposed as in (4.1) with*

$$E\left[|X_0|^{2p}\right] < \infty \tag{4.3}$$

for some $p \in \mathbb{N}$, then the following estimate holds for all $m \ge p$

$$\sup_{0\le t\le 1} E\left[|A_{t+h} - A_t|^p\, |M_{t+h} - M_t|^{m-p}\right] = O\left(h^{(m+p)/2}\right) \quad (h \downarrow 0)\,. \tag{4.4}$$

Proof. Recall from Karatzas and Shreve (1991) p. 306, that there exists a constant $C_{m,K} > 0$ such that the solution of the stochastic differential equation (1.1) satisfies

$$E(\sup_{0\le s\le t} |X_s|^{2m}) \le C_{m,K}(1 + E(|X_0|^{2m}))\, e^{C_{m,K}\,t} \tag{4.5}$$

for all $t \in [0, 1]$, provided that for some $m \in \mathbb{N}$

$$E(|X_0|^{2m}) < \infty\ . \tag{4.6}$$

Under the same assumption it also follows that

$$E(|X_t - X_s|^{2m}) \le C_{m,K}(1 + E(|X_0|^{2m}))(t - s)^m \tag{4.7}$$

holds for all $s, t \in [0, 1]$ with $s \le t$. Now an application of the Cauchy Schwarz inequality yields

$$E\left[|A_{t+h} - A_t|^p\, |M_{t+h} - M_t|^{m-p}\right] \le \{E[|A_{t+h} - A_t|^{2p}]\}^{\frac{1}{2}}\{E[|M_{t+h} - M_t|^{2(m-p)}]\}^{\frac{1}{2}}, \tag{4.8}$$

where the factors of the right hand side can be estimated as follows. Using the definition (4.2) and the estimate (4.5) we obtain by assumption (2.2)

$$\begin{aligned} E[|A_{t+h} - A_t|^{2p}] \;&=\; E\Big[\Big|\int_t^{t+h} b\,(s, X_s)\,ds\Big|^{2p}\Big] \;\le\; h^{2p}K^{2p}E[(1 + \sup_{0\le s\le 1}|X_s|)^{2p}] \\ &\le\; 2^p h^{2p} K^{2p} \sum_{l=0}^{p} \binom{p}{l} E[\sup_{0\le s\le 1} |X_s|^{2l}] \;=\; O(h^{2p}) \end{aligned} \tag{4.9}$$

A further application of (4.7) for the process $(M_t)_{t\in[0,1]}$ defined in (4.2) yields

$$E[|M_{t+h} - M_t|^{2(m-p)}] \le C_{m-p,K}\, h^{m-p} \tag{4.10}$$

(note that $M_0 = 0$) and a combination of (4.8) - (4.10) proves the assertion of Lemma 4.1.

□

Proof of Theorem 2.1. Recalling the definition of the random variables T_n, T_{1n}, T_{2n} in (2.6), (2.7) and the decomposition (4.1) we obtain

$$T_{1n} = T_{1n}^{(1)} + 2T_{1n}^{(2)} + T_{1n}^{(3)} \tag{4.11}$$

where the statistics $T_{1n}^{(i)}$ $(i = 1, 2, 3)$ are defined by

$$\begin{aligned} T_{1n}^{(1)} &= \sum_{i=1}^{n-1} (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2 \\ T_{1n}^{(2)} &= \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})(M_{\frac{i+1}{n}} - M_{\frac{i}{n}}) \\ T_{1n}^{(3)} &= \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})^2. \end{aligned} \tag{4.12}$$

Similary, we have

$$T_{2n} = T_{2n}^{(1)} + 4T_{2n}^{(2)} + 6T_{2n}^{(3)} + 4T_{2n}^{(4)} + T_{2n}^{(5)} \tag{4.13}$$

with statistics $T_{2n}^{(i)}$ $(i = 1, 2, 3, 4, 5)$ defined by

$$\begin{aligned} T_{2n}^{(1)} &= n \sum_{i=1}^{n-1} (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4 \\ T_{2n}^{(2)} &= n \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^3 \\ T_{2n}^{(3)} &= n \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})^2(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2 \\ T_{2n}^{(4)} &= n \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})^3(M_{\frac{i+1}{n}} - M_{\frac{i}{n}}) \\ T_{2n}^{(5)} &= n \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})^4. \end{aligned} \tag{4.14}$$

A straightforward application of Lemma 4.1 gives (using assumption (2.9))

$$
\begin{aligned}
T_{1n}^{(2)} &= O_P(n^{-\frac{1}{2}}) \\
T_{1n}^{(3)} &= O_P\left(n^{-1}\right) \\
T_{2n}^{(2)} &= O_P(n^{-\frac{1}{2}}) \\
T_{2n}^{(3)} &= O_P\left(n^{-1}\right) \\
T_{2n}^{(4)} &= O_P(n^{-\frac{3}{2}}) \\
T_{2n}^{(5)} &= O_P(n^{-2})
\end{aligned}
\tag{4.15}
$$

and observing the definition (2.7) the assertion of Theorem 2.1 can be established by proving the estimates

$$
R_n = T_{1n}^{(1)} - \int_0^1 \sigma^2(s, X_s)ds = O_p(n^{-1/2}) \tag{4.16}
$$

$$
S_n = \frac{1}{3}T_{2n}^{(1)} - \int_0^1 \sigma^4(s, X_s)ds = O_p(n^{-1/2}\log n). \tag{4.17}
$$

In order to prove these estimates we note that Itô's formula [see Karatzas and Shreve (1991)] gives the representation

$$
(M_{t+h} - M_t)^2 = \int_t^{t+h} 2(M_u - M_t)\sigma(u, X_u)dW_u + \int_t^{t+h} \sigma^2(u, X_u)du, \tag{4.18}
$$

which shows

$$
R_n = \sum_{i=1}^{n-1} \Big\{ (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2 - \int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(s, X_s)ds \Big\} + O(n^{-1}) = U_{1n}^{(1)} + O(n^{-1}), \tag{4.19}
$$

where the random variable $U_{1n}^{(1)}$ is defined by

$$
U_{1n}^{(1)} := \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} 2(M_u - M_{\frac{i}{n}})\sigma(u, X_u)dW_u. \tag{4.20}
$$

The martingale properties of the Itô integral show that the terms in the above sum are uncorrelated and Itô's isometry allows an explicit calculation of the L^2-norm of $U_{1n}^{(1)}$, i.e.

$$
E[(U_{1n}^{(1)})^2] = \sum_{i=1}^{n-1} E[\int_{\frac{i}{n}}^{\frac{i+1}{n}} 4(M_u - M_{\frac{i}{n}})^2\sigma^2(u, X_u)du] \tag{4.21}
$$

$$\le \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} 4E[(M_u - M_{\frac{i}{n}})^4]^{\frac{1}{2}} E[\sigma^4(u, X_u)]^{\frac{1}{2}} du$$

$$\le D_1 \cdot \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (u - \frac{i}{n}) \, (1 + E[|X_u|^4])^{\frac{1}{2}} du$$

$$\le D_2 (1 + E[\sup_{0 \le t \le 1} |X_t|^4])^{\frac{1}{2}} \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (u - \frac{i}{n}) \, du \;=\; O(n^{-1}),$$

where we used Cauchy's inequality, the estimates (4.5), (4.7) and the constants D_1, D_2 are independent of n. A combination of this estimate with (4.19) proves (4.16).

For the proof of the remaining estimate (4.17) we note that a twofold application of Itô's formula yields

$$
\begin{aligned}
(M_{t+h} - M_t)^4 \;&=\; \int_t^{t+h} 4(M_u - M_t)^3 \sigma(u, X_u) dW_u \\
&+ \int_t^{t+h} 6(M_u - M_t)^2 \sigma^2(u, X_u) du \\
&= \int_t^{t+h} 4(M_u - M_t)^3 \sigma(u, X_u) dW_u \\
&+ \int_t^{t+h} \int_t^u 12(M_s - M_t) \sigma(s, X_s) dW_s \, \sigma^2(u, X_u) du \\
&+ \int_t^{t+h} \int_t^u 6\sigma^2(s, X_s) ds \, \sigma^2(u, X_u) du,
\end{aligned}
\tag{4.22}
$$

which gives for the left hand side of (4.17) the representation

$$3S_n \;=\; T_{2n}^{(1)} - 3\int_0^1 \sigma^4(s, X_s) ds \;=\; U_{2n}^{(1)} + U_{2n}^{(2)} + U_{2n}^{(3)} + O(n^{-1}) \;, \tag{4.23}$$

where the quantities $U_{2n}^{(i)}$ $(i = 1, 2, 3)$ are defined by

$$
\begin{aligned}
U_{2n}^{(1)} \;&:=\; n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} 4(M_u - M_{\frac{i}{n}})^3 \sigma(u, X_u) dW_u \\
U_{2n}^{(2)} \;&:=\; n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}}) \sigma(s, X_s) dW_s \, \sigma^2(u, X_u) du \\
U_{2n}^{(3)} \;&:=\; \sum_{i=1}^{n-1} \Big\{ n \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{u} 6\sigma^2(s, X_s) ds \, \sigma^2(u, X_u) du - 3\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^4(s, X_s) ds \Big\} \;.
\end{aligned}
\tag{4.24}
$$

The first term is treated similar as the statistic $U_{1n}^{(1)}$ in (4.2) and gives

$$E[(U_{2n}^{(1)})^2] = O(n^{-1}),$$

which implies

$$U_{2n}^{(1)} \;=\; O_P(n^{-\frac{1}{2}}). \tag{4.25}$$

For the estimation of the second term we use integration by parts for the Itô integral [see Karatzes and Shreve (1991), p. 155] and obtain by a straightforward calculation

$$U_{2n}^{(2)} \;=\; 12[U_{2n}^{(2.1)} + U_{2n}^{(2.2)} + U_{2n}^{(2.3)}] \tag{4.26}$$

where the terms $U_{2n}^{(2,i)}$ $(i = 1, 2, 3)$ are defined as follows

$$\begin{aligned} U_{2n}^{(2.1)} &:= n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{u} \sigma^2(s, X_s) ds\, (M_u - M_{\frac{i}{n}}) \sigma(u, X_u) dW_u, \\ U_{2n}^{(2.2)} &:= \sum_{i=1}^{n-1} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} (M_s - M_{\frac{i}{n}}) \sigma(s, X_s) dW_s, \\ U_{2n}^{(2.3)} &:= n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (M_s - M_{\frac{i}{n}}) \sigma(s, X_s) dW_s \int_{\frac{i}{n}}^{\frac{i+1}{n}} (\sigma^2(u, X_u) - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}})) du. \end{aligned} \tag{4.27}$$

For the random variable $U_{2n}^{(2.1)}$ we have

$$\begin{aligned} E[(U_{2n}^{(2.1)})^2] &= n^2 \sum_{i=1}^{n-1} E\Big[\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{u} \sigma^2(s, X_s) ds\, (M_u - M_{\frac{i}{n}}) \sigma(u, X_u) dW_u\Big)^2\Big] \\ &= n^2 \sum_{i=1}^{n-1} E\Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \Big(\int_{\frac{i}{n}}^{u} \sigma^2(s, X_s) ds\Big)^2 (M_u - M_{\frac{i}{n}})^2 \sigma^2(u, X_u) du\Big] \;=\; O(n^{-1}), \end{aligned}$$

where we used similar arguments as in derivation of (4.21). This yields

$$U_{2n}^{(2.1)} = O_p(n^{-1/2}), \tag{4.28}$$

and an analog argument for the second term in (4.26) shows

$$U_{2n}^{(2.2)} = O_p(n^{-1/2}). \tag{4.29}$$

For the remaining term $U_{2n}^{(2.3)}$ we apply Cauchy's inequality

$$\begin{aligned} E[|U_{2n}^{(2.3)}|] \;\le\; n \sum_{i=1}^{n-1} & E\Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} (M_s - M_{\frac{i}{n}})^2 \sigma^2(s, X_s) ds\Big]^{\frac{1}{2}} \\ & \times E\Big[\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(u, X_u) - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}})\, du\Big)^2\Big]^{\frac{1}{2}} \end{aligned} \tag{4.30}$$

and note that it follows from (2.2), (4.5) and (4.7) for the first factor

$$E\Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} (M_s - M_{\frac{i}{n}})^2 \sigma^2(s, X_s) ds\Big]^{\frac{1}{2}} = O(n^{-1}),$$

where the bound is independent of i. For the second factor we have from (2.1) and (2.8) (uniformly in i)

$$\begin{aligned}\int_{\frac{i}{n}}^{\frac{i+1}{n}} |\sigma^2(u,X_u) - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}})|\, du &= \int_{\frac{i}{n}}^{\frac{i+1}{n}} |\sigma(u,X_u) - \sigma(\frac{i}{n}, X_{\frac{i}{n}})||\sigma(u,X_u) + \sigma(\frac{i}{n}, X_{\frac{i}{n}})|\, du \\ &\leq 2 \sup_{0\leq t\leq 1} |\sigma(t,X_t)| \int_{\frac{i}{n}}^{\frac{i+1}{n}} \Big(L(u-\frac{i}{n})^{\gamma} + K\,|X_u - X_{\frac{i}{n}}| \Big)\, du \\ &\leq 2 \sup_{0\leq t\leq 1} |\sigma(t,X_t)| \Big(\frac{L}{n^{3/2}} + \frac{K}{n} \sup_{0\leq s<t\leq 1,\ |t-s|\leq n^{-1}} |X_t - X_s| \Big) \\ &= O(n^{-3/2} + n^{-\frac{3}{2}} (\log n)^{\frac{1}{2}}) \quad a.s.,\end{aligned}$$

where the last line follows from the well known estimate for the modulus of continuity of the diffusion $(X_t)_{t\in[0,1]}$

$$\limsup_{h\downarrow 0} \sup_{0\leq s<t\leq 1,\ |t-s|\leq h} \frac{|X_t - X_s|}{\sqrt{2h\log(h^{-1})}} = O(1) \quad a.s. \tag{4.31}$$

[see McKean (1969) p. 57] and the fact that the (almost surely) continuous function $t \to \sigma(t, X_t)$ is bounded on the compact interval $[0,1]$. A combination of these estimates with (4.30) yields

$$E[|U_{2n}^{(2.3)}|] = O(n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}}),$$

which implies observing the estimates (4.28) and (4.29)

$$U_{2n}^{(2)} = O_p(n^{-1/2}\sqrt{\log n}). \tag{4.32}$$

For the third term in (4.23) we use integration by parts and obtain

$$\begin{aligned}|U_{2n}^{(3)}| &= \Big| \sum_{i=1}^{n-1} 3 \Big\{ n \Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(s,X_s) ds \Big)^2 - \int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^4(s,X_s) ds \Big\} \Big| \\ &\leq \sum_{i=1}^{n-1} 3 \Big\{ n \Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} |\, \sigma^2(s,X_s) - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \,|\, ds \Big)^2 + \int_{\frac{i}{n}}^{\frac{i+1}{n}} |\, \sigma^4(s,X_s) - \sigma^4(\frac{i}{n}, X_{\frac{i}{n}}) \,|\, ds \\ &\qquad + 2\sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} |\, \sigma^2(s,X_s) - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \,|\, ds \Big\} \\ &\leq \sum_{i=1}^{n-1} 3 \Big\{ n \Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} |\, \sigma(s,X_s) - \sigma(\frac{i}{n}, X_{\frac{i}{n}}) \,|\,|\, \sigma(s,X_s) + \sigma(\frac{i}{n}, X_{\frac{i}{n}}) \,|\, ds \Big)^2 \\ &\qquad + \int_{\frac{i}{n}}^{\frac{i+1}{n}} |\, \sigma(s,X_s) - \sigma(\frac{i}{n}, X_{\frac{i}{n}}) \,|\,|\, \sigma(s,X_s) + \sigma(\frac{i}{n}, X_{\frac{i}{n}}) \,|\,|\, \sigma^2(s,X_s) + \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \,| ds \\ &\qquad + 2\sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} |\, \sigma(s,X_s) - \sigma(\frac{i}{n}, X_{\frac{i}{n}}) \,|\,|\, \sigma(s,X_s) + \sigma(\frac{i}{n}, X_{\frac{i}{n}}) \,|\, ds \ \Big\}\end{aligned}$$

$$\leq C_1 \sum_{i=1}^{n-1} \Big\{ n \Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} (L(s-\frac{i}{n})^\gamma + K\,|X_s - X_{\frac{i}{n}}|)\, ds \cdot \sup_{0\leq t\leq 1} |\sigma(t,X_t)| \Big)^2$$

$$+ \int_{\frac{i}{n}}^{\frac{i+1}{n}} (L(s-\frac{i}{n})^\gamma + K\,|X_s - X_{\frac{i}{n}}|)\, ds \cdot \sup_{0\leq t\leq 1} |\sigma^3(t,X_t)| \Big\}$$

$$\leq C_2 \sum_{i=1}^{n-1} \Big\{ n \sup_{0\leq t\leq 1} |\sigma^2(t,X_t)| \Big(\frac{L}{n^{3/2}} + \frac{K}{n} \sup_{0\leq s<t\leq 1,\ |t-s|\leq n^{-1}} |X_t - X_s| \Big)^2$$

$$+ \sup_{0\leq t\leq 1} |\sigma^3(t,X_t)| \Big(\frac{L}{n^{3/2}} + \frac{K}{n} \sup_{0\leq s<t\leq 1,\ |t-s|\leq n^{-1}} |X_t - X_s| \Big) \Big\}$$

with constants C_1, C_2 independent of n. Using again the estimate (4.31) for the modulus of continuity it follows that

$$U_{2n}^{(3)} = O_P(n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}}). \qquad (4.33)$$

and a combination with (4.32) and (4.25) yields the estimate (4.17), i.e.

$$S_n = O_p(n^{-1/2} \log n).$$

The assertion of the theorem now follows from the estimates (4.16) and (4.17).

□

Proof of Theorem 2.2. Recalling the definition of s_{2j} in (2.13) and of T_{2n}, T_{1n} in (2.6) the assertion of Theorem 2.2 follows from Cramér's rule if the weak convergence

$$\sqrt{n} \begin{pmatrix} T_{1n} - s_2 \\ T_{2n} - 3s_4 \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, V) \quad (n \to \infty) \qquad (4.34)$$

with

$$V = \begin{pmatrix} 2s_4 & 12s_6 \\ 12s_6 & 96s_8 \end{pmatrix} \qquad (4.35)$$

can be established. To this end we use the decomposition (4.11) and (4.13) introduced in the proof of Theorem 2.1. From the estimates (4.15) it is clear that (4.34) follows from the weak convergence of

$$\sqrt{n} \begin{pmatrix} T_{1n}^{(1)} + 2T_{1n}^{(2)} - s_2 \\ T_{2n}^{(1)} + 4T_{2n}^{(2)} - 3s_4 \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, V). \qquad (4.36)$$

The proof of the lastnamed statement is performed in two steps. At first we show that $T_{1n}^{(2)}$ and $T_{2n}^{(2)}$ are of order $o_p(n^{-1/2})$ and therefore neglectible in (4.36) and secondly we prove asymptotic convergence of the "remaining" random vector. For the first part we use the decomposition

$$T_{kn}^{(2)} = n^{k-1} \sum_{i=1}^{n-1} (A_{\frac{i+1}{n}} - A_{\frac{i}{n}})(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^{2k-1} = T_{kn}^{(2.1)} + T_{kn}^{(2.2)} + T_{kn}^{(2.3)}\ ;\ k = 1, 2, \qquad (4.37)$$

where [recall the definition of the process $(A_t)_{t\in[0,1]}$ in (4.2)]

$$
\begin{aligned}
T_{kn}^{(2.1)} &= n^{k-2}\sum_{i=1}^{n-1} b(\frac{i}{n}, X_{\frac{i}{n}})(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^{2k-1}, \\
T_{kn}^{(2.2)} &= n^{k-1}\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} (b(s, X_s) - b(s, X_{\frac{i}{n}}))\, ds\, (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^{2k-1}, \\
T_{kn}^{(2.3)} &= n^{k-1}\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} (b(s, X_{\frac{i}{n}}) - b(\frac{i}{n}, X_{\frac{i}{n}}))\, ds\, (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^{2k-1}.
\end{aligned} \tag{4.38}
$$

A straightforward application of Itô's formula shows for $k \geq 1$

$$E\left[(M_{t+h} - M_t)^{2k-1} \mid \mathcal{F}_t\right] = 0 \,, \tag{4.39}$$

$$E\left[(M_{t+h} - M_t)^{2k} \mid \mathcal{F}_t\right] = d_k\Big[\int_t^{t+h} \sigma^2(s)\, ds\Big]^k \tag{4.40}$$

where $d_k = 1\cdot 3\cdot \ldots \cdot(2k-1)$. This gives for the expectation and variance of the first term on the right hand side of (4.37)

$$E\left[T_{kn}^{(2.1)}\right] = 0 \;;\; k = 1, 2$$

$$Var\left[T_{kn}^{(2.1)}\right] = d_{2k-1}n^{2k-4}\sum_{i=1}^{n-1} E\left[b^2(\frac{i}{n}, X_{\frac{i}{n}})\right]\left[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(s)\, ds\right]^{2k-1} = O\left(n^{-2}\right) \;;\; k = 1, 2,$$

where we used (2.2) and (4.5) for the last estimate. This implies

$$T_{kn}^{(2.1)} = O_P(n^{-1}) = o_P(n^{-\frac{1}{2}}) \;;\; k = 1, 2. \tag{4.41}$$

For the second term $T_{kn}^{(2.2)}$ in (4.37) we use Cauchy's inequality, (2.1), (4.7) [applied to the process $(M_t)_{t\in[0,1]}$] and obtain with constants F_1, F_2, F_3 independent of n

$$
\begin{aligned}
E\left[\left|T_{kn}^{(2.2)}\right|\right] &\leq n^{k-1}\sum_{i=1}^{n-1}\Big\{E\Big[\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} (b(s, X_s) - b(s, X_{\frac{i}{n}}))\, ds\Big)^2\Big]\Big\}^{\frac{1}{2}}\Big\{E[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^{4k-2}]\Big\}^{\frac{1}{2}} \\
&\leq F_1 n^{k-1}\sum_{i=1}^{n-1}\Big\{E\Big[\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}} |X_s - X_{\frac{i}{n}}|\, ds\Big)^2\Big]\Big\}^{\frac{1}{2}}\; n^{-\frac{2k-1}{2}} \\
&\leq F_2 n^{-\frac{1}{2}}\sum_{i=1}^{n-1}\Big\{n^{-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} E[|X_s - X_{\frac{i}{n}}|^2]\, ds\Big\}^{\frac{1}{2}} \\
&\leq F_2 n^{-1}\sum_{i=1}^{n-1}\Big\{\int_{\frac{i}{n}}^{\frac{i+1}{n}} (1 + E[|X_0|^2])(s - \frac{i}{n})\, ds\Big\}^{\frac{1}{2}} \\
&= F_3\,(1 + E[|X_0|^2])^{\frac{1}{2}}\, n^{-2}(n-1) = O\left(n^{-1}\right),
\end{aligned}
$$

which proves

$$T_{nk}^{(2.2)} = O_P\left(n^{-1}\right) = o_P(n^{-\frac{1}{2}})\ ;\ k = 1,2. \tag{4.42}$$

Finally, a similar argument shows

$$T_{nk}^{(2.3)} = O_P\left(n^{-1}\right) = o_P(n^{-\frac{1}{2}}) \quad ,\ k = 1,2 \tag{4.43}$$

and combining (4.41) – (4.43) with (4.36) shows that the weak convergence of (4.36) can be established by proving

$$\sqrt{n}\left(a^T\begin{pmatrix} T_{1n}^{(1)} \\ T_{2n}^{(1)}\end{pmatrix} - a^T\begin{pmatrix} s_2 \\ 3s_4\end{pmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, a^TVa\right) \quad (n\to\infty)\,. \tag{4.44}$$

for all vectors $a = (a_1, a_2)^T \in \mathbb{R}^2\backslash\{0\}$. To this and let

$$T_{0n}^{(1)} = a_1 T_{1n}^{(1)} + a_2 T_{2n}^{(1)} = \sum_{i=1}^{n-1} a_1 (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2 + a_2 n (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4, \tag{4.45}$$

then it follows from (4.40) by a straightforward calculation

$$\begin{aligned} E\left[T_{0n}^{(1)}\right] &= \sum_{i=1}^{n-1} a_1 \Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(s)\,ds\Big] + a_2\,3n\Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(s)\,ds\Big]^2 \\ &= \sum_{i=1}^{n-1} a_1 n^{-1}\sigma^2(\frac{i}{n}) + a_2\,3n^{-1}\sigma^4(\frac{i}{n}) + O\left(n^{-\gamma}\right) \\ &= a_1 \int_0^1 \sigma^2(t)\,dt + a_2\,3\int_0^1 \sigma^4(t)\,dt + O\left(n^{-\gamma}\right) \\ &= a_1 s_2 + 3a_2 s_4 + o\left(n^{-1/2}\right) \end{aligned} \tag{4.46}$$

For the variance we obtain by similar arguments (observing that the terms in the sum (4.45) are uncorrelated)

$$\begin{aligned} S_n^2 &= Var\left[T_{0n}^{(1)}\right] \\ &= \sum_{i=1}^{n-1} a_1^2 E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4\right] + 2a_1a_2 n\, E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^6\right] \\ &\quad + a_2^2 n^2\, E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^8\right] - a_1^2 E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2\right]^2 \\ &\quad - 2a_1a_2 n\, E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2\right] E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4\right] \\ &\quad - a_2^2 n^2\, E\left[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4\right]^2 \\ &= \sum_{i=1}^{n-1} 2a_1^2\Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(t)\,dt\Big]^2 + 24 a_1a_2 n\Big[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(t)\,dt\Big]^3 \end{aligned} \tag{4.47}$$

$$+ \ 96a_2^2 n^2 \left[\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(t)\,dt\right]^4$$

$$= 2a_1^2 n^{-1} \int_0^1 \sigma^4(t)\,dt + 24a_1a_2n^{-1} \int_0^1 \sigma^6(t)\,dt$$

$$+ \ 96a_2^2 n^{-1} \int_0^1 \sigma^8(t)\,dt \ + \ O\left(n^{-1-\gamma}\right) \ + \ O\left(n^{-2\gamma}\right)$$

$$= n^{-1} a^T V a + o\left(n^{-1}\right).$$

Moreover, (4.39) and (4.40) also imply

$$S_n^{-4} \sum_{i=1}^{n-1} E\Big[\Big(a_1(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2 - a_1 E[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2]$$

$$+a_2 n(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4 - a_2 n E[(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4]\Big)^4\Big] \ = \ O\left(n^{-1}\right).$$

and the weak convergence in (4.44) follows from the central limit theorem which completes the proof of Theorem 2.2.

□

References

Ait-Sahalia, Y. (1996a). Nonparametric pricing of interest rate derivative securities. Econometrica 64, 527-560.

Ait-Sahalia, Y. (1996b). Testing continuous time models of the spot interest rate. Rev. Financ. Stud. 9, 385-426.

Azzalini, A., Bowman, A. (1993). On the use of nonparametric regression for checking linear relationships. J. Roy. Statist. Soc., Ser. B, 55, 549-559.

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. J. Polit. Econ. 81, 637-654.

Constantinides, G.M. (1992). A theory of the nominal term structure of interest rates. Rev. Finan. Studies 5, 531-552.

Corradi, V., White, H. (1999). Specification tests for the variance of a diffusion. J. Time Series 20, 253-270.

Cox, J.C., Ingersoll, J.E., Ross, S.A. (1985). A theory of the term structure of interest rates. Econometrica 53, 385-407.

Dette, H., Munk, A. (1998a). Validation of linear regression models. Ann. Statist. 26, 778-800.

Dette, H., Munk, A. (1998b). Testing heteroscedasticity in nonparametric regression. J. Roy. Stat. Soc., Ser. B, 60, 693-708.

Duffie, J.D., Harrison, J.M. (1993). Arbitrage pricing of Russian options and perpetual lookback options. Ann. Appl. Probab. 3, 641-651.

Florens-Zmirou, D. (1993). Estimation de la variance à parti d'une observation discrètisée. C.R. Acad. Sci. 309, Ser. I, 195-200.

Genon-Catalod, V., Laredo, C., Picard, D. (1992). Non-parametric estimation of the diffusion coefficient by wavelet methods. Scand. J. Statist. 19, 317-335.

Jiang, J.G., Knight, J.L. (1997). A nonparametric approach to the estimation of diffusion processes with an application to a short-term interest rate model. Economet. Theory 13, 647-667.

Karatzas, I. (1988). On pricing of American options. Appl. Math. Optimization 17, 37-60.

Karatzas, I., Shreve, S.E. (1991). Brownian Motion and Stochastic Calculus. Springer, N.Y.

McKean, H. (1969). Stochastic Integrals. Academic Press, N.Y.

Merton, R.C. (1973). Theory of rational option pricing. Bell Journal of Economics and Management Science 4, 141-183.

Merton, R.C. (1990). Continuous-Time Finance. Blackwell, Cambridge.

Müller, H.G. (1992). Goodness of fit diagnostics of regression models. Scand. J. Statist. 19, 157-172.

Vasicek, O. (1977). An equilibrium characterization of the term structure. J. Finan. Economics 5, 177-188.

Zheng, J.X. (1996). A consistent test of a functional form via nonparametric estimation techniques. J. Econometrics 75, 263-289.