Estimation of integrated volatility in continuous time financial models with applications to goodness-of-fit testing

Holger Dette, Mark Podolskij, Mathias Vetter
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany

email: holger.dette@ruhr-uni-bochum.de

FAX: +49 2 34 32 14 559

May 13, 2004

Abstract

Properties of a specification test for the parametric form of the variance function in diffusion processes $dX_t = b\,(t, X_t)\,dt + \sigma\,(t, X_t)\,dW_t$ are discussed. The test is based on the estimation of certain integrals of the volatility function. If the volatility function does not depend on the variable x it is known that the corresponding statistics have an asymptotic normal distribution. However, most models of mathematical finance use a volatility function which depends on the state x. In this paper we prove that in the general case, where σ depends also on x the estimates of integrals of the volatility converge stably in law to random variables with a non-standard limit distribution. The limit distribution depends on the diffusion process X_t itself and we use this result to develop a bootstrap test for the parametric form of the volatility function, which is consistent in the general diffusion model.

Keywords: continuous time financial model, model diagnostics, diffusion process, heteroscedasticity, pseudo residuals, parametric bootstrap, estimation of integrated volatility, delta-method

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. 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. 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 and Ross (1985), Karatzas (1988), Constantinides (1992) or Duffie and Harrison (1993) among many others]. However, economic theory typically does not give much information about the drift and variance and misspecification of such a model may lead to serious errors in the subsequent data analysis. Therefore many authors propose to test the goodness-of-fit of the postulated model [see Ait-Sahalia (1996), Kleinow (2002) among many others].
In a recent paper Dette and von Lieres und Wilkau (2003) suggest a test for the parametric form of the volatility function, which is based on estimates of the integrated volatility function $\int_0^1 \sigma^{2n}(t, X_t)dt$ ($n \in \mathbb{N}$). Under the assumption that the volatility function in the diffusion model (1.1) does not depend on the variable x, i.e. $\sigma(t, x) = \sigma(t)$, these authors show that the corresponding test statistic is asymptotically normal distributed. The test avoids the nonparametric estimation of the volatility function and therefore does not depend on the specification of a smoothing parameter. Because of its simplicity this test is attractive to practitioners, if a hypothesis of a parametric form of the volatility function has to be tested, which is independent of the variable x. However, many of the common models in mathematical finance contain a volatility function, which also depends on the values of X_t, and cannot be checked by this method.
It is the purpose of the present note to give a more refined asymptotic analysis of a generalization of the test statistic proposed by Dette and von Lieres und Wilkau (2003) in the diffusion model (1.1), where the volatility also depends on the variable x. It is shown in Section 2 that in this case the analysis of the asymptotic properties of the estimates for the integrated volatility is substantially more difficult. The estimates of the integrals of the volatility are no longer asymptotically normal distributed, but converge stably in law to a random variable with a nonstandard distribution depending on the process $(X_t)_{t\geq 0}$ itself. Moreover, conditioned on the process $(X_t)_{t\in[0,1]}$ the law of the limit distribution is a normal distribution. These results are used for the construction of a consistent bootstrap test, which is also applicable in the general situation, where the volatility function depends on x. A simulation study is presented which illustrates good finite sample properties of the corresponding test and a data example is analyzed. Finally the technical details are deferred to the appendix.
The work, which is most similar in spirit with the present paper are the papers of Corradi and White (1999), Kleinow (2002), Barndorff-Nielsen and Shephard (2003) and Woerner (2003), who also discuss the situation, where the time span is fixed and the length of the discrete sampling interval converges to zero as the sample size increases. While the latter authors consider the problem of estimating power variation in stochastic volatility models, Corradi and White (1999) and Kleinow (2002) discuss the problem of testing for a parametric form of the volatility function in the diffusion model (1.1). These authors compare a nonparametric estimator of the variance function [see Florens-Zmirou (1993)] with an estimator under the null hypothesis of a specific parametric form 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 parametric form of the 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. To the knowledge of the authors this method provides the first omnibus test of a general parametric hypothesis for the volatility function in a stochastic differential equation based on discretely observed data on a fixed time span. Moreover, our approach 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 volatility function are required.

2 Basic asumptions and main results

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)

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

are locally Lipschitz continuous, i.e. for every integer $M > 0$ there exists a constant K_M such that

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

for all $t \in [0,1], x, y \in [-M, M]$, and there exists a constant K such that

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

for all $t \in [0,1], x \in \mathbb{R}$. 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 condition $X_0 = \xi$ 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]. Throughout this paper we assume additionally that

$$E[|\xi|^8] < \infty. \tag{2.3}$$

The solution of the differential equation can be represented as

$$X_t = \xi + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s \quad \text{a.s.,} \tag{2.4}$$

where X_t is $\mathcal{F}_t$-measurable for all $t \in [0,1]$ and the paths $t \to X_t$ are almost surely continuous. In the literature various parametric functions have been proposed for the different types of options [see e.g. Black and Scholes (1973), Vasicek (1977), Cox, Ingersoll and Ross (1985), Karatzas (1988), Constantinides (1992) or Duffie and Harrison (1993) among many others]. In this paper we are interested in testing the hypothesis that the variance function in the stochastic differential equation (1.1) is of the form

$$H_0 : \sigma^2(t, X_t) = \sigma^2(t, X_t; \theta) \quad \forall\, t \in [0,1] \quad \text{(a.s.)}\ , \tag{2.5}$$

where $\sigma(t,x;\theta)$ is a known volatility function and $\theta = (\theta_1,\dots,\theta_d) \in \Theta \subset \mathbb{R}^d$ an unknown finite dimensional parameter. For the sake of simplicity and transparency we will mainly discuss the problem of a linear hypothesis

$$\tilde{H}_0 : \sigma^2(t,X_t) = \sum_{j=1}^{d} \theta_j \sigma_j^2(t,X_t) \quad \forall\, t \in [0,1] \quad \text{(a.s.)} \tag{2.6}$$

where $\sigma_1,\dots,\sigma_d$ are known linearly independent functions. The general case of a nonlinear hypothesis is obtained by a (standard) linearization [see Remark 2.6]. Note that the hypothesis (2.6) is equivalent to the condition

$$\tilde{H}_0 :\; M^2 = \min_{\theta_1,\dots,\theta_d \in \mathbb{R}} \int_0^1 \Big\{ \sigma^2(t,X_t) - \sum_{j=1}^{d} \theta_j \sigma_j^2(t,X_t) dt \Big\}^2 dt = 0 \quad \text{a.s.} \tag{2.7}$$

Moreover, a standard argument in Hilbert space theory [see Achieser (1956)] shows that

$$M^2 = g(C_1,\dots,C_d,3C_0;D) \;:=\; C_0 - (C_1,\dots,C_d)D^{-1}(C_1,\dots,C_d)^T \;, \tag{2.8}$$

where the random variables $C_0,\dots,C_d$ and $D = (D_{ij})_{i,j=1,\dots,d}$ are defined by

$$\begin{aligned} C_0 &:= \langle \sigma^2, \sigma^2 \rangle_2 = \int_0^1 \sigma^4(t,X_t)dt \\ C_i &:= \langle \sigma^2, \sigma_i^2 \rangle_2 = \int_0^1 \sigma^2(t,X_t)\sigma_i^2(t,X_t)dt \quad 1 \le i \le d \\ D_{ij} &:= \langle \sigma_i^2, \sigma_j^2 \rangle_2 = \int_0^1 \sigma_i^2(t,X_t)\sigma_j^2(t,X_t)dt \quad 1 \le i,j \le d \end{aligned} \tag{2.9}$$

A test for the hypotheses of the form (2.6) has recently been proposed by Dette and von Lieres und Wilkau (2003) under the additional assumption that the functions $\sigma,\sigma_1,\dots,\sigma_d$ do not depend on the the variable x, i.e. $\sigma(t,x) = \sigma(t)$. In the following discussion we generalize this method to the case where the volatility functions also depend on x.
Assume that we observe the diffusion only on the time span $[0,1]$ at discrete points $t_i = i/n\,(i = 1,\dots,n)$. For the construction of an appropriate estimator of M^2 we estimate the quantities in (2.8) separately. We define the statistics

$$T_{1n,j} \;:=\; \sum\nolimits_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}})(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2 \qquad j = 1,\dots,d \tag{2.10}$$

$$T_{2n} \;:=\; \sum\nolimits_{i=1}^{n-1} n(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^4 \tag{2.11}$$

$$\hat{D}_n \;=\; \frac{1}{n}\bar{X}^T\bar{X} \tag{2.12}$$

and obtain an empirical anolog of M^2, i.e.

$$T_n := g\Big(T_{1n,1},\dots,T_{1n,d},T_{2n};\hat{D}_n\Big) \;:= \frac{1}{3}T_{2n} - T_{1n}^T\hat{D}_n^{-1}T_{1n} \;, \tag{2.13}$$

where the vector T_{1n} is defined by

$$T_{1n} \;=\; (T_{1n,1}, \ldots, T_{1n,d})^T \tag{2.14}$$

and the matrix $\bar{X}$ is given by

$$\bar{X} = \Big(\sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}})\Big)_{i=1,\ldots,n-1}^{j=1,\ldots,d} \in \mathbb{R}^{n-1\times d}. \tag{2.15}$$

In the following we will show that T_n is a consistent estimate of M^2 and that $\sqrt{n}(T_n - M^2)$ converges stably in law with a nonstandard limit distribution depending on the diffusion $(X_t)_{t\in[0,1]}$. For this purpose we require the following additional assumptions for the volatility and drift function in the stochastic differential equation (1.1). We assume that the drift and variance function satisfy a further Lipschitz condition of order $\gamma > \frac{1}{2}$ with respect to the first argument, i.e.

$$|b\,(t,x) - b\,(s,x)| + |\sigma(t,x) - \sigma(s,x)| \le L\,|t-s|^\gamma \tag{2.16}$$

for all $s,t \in [0,1]$, $L > 0$ and that $\sigma^2 : [0,1]\times\mathbb{R} \to \mathbb{R}$ is two times continuously differentiable such that

$$\sup_{s,t\in[0,1]} E[(\frac{\partial}{\partial y}\sigma^2(s,X_t))^4] < F, \ \sup_{s,t\in[0,1]} E[(\frac{\partial^2}{\partial y^2}\sigma^2(s,X_t))^4] < F \tag{2.17}$$

for some constant $F > 0$, where the differentiation is performed with respect to the second argument. We finally assume that the functions $\sigma_1^2, \ldots, \sigma_d^2$ in the linear hypothesis (2.6) satisfy the same assumptions (2.1), (2.2), (2.16) and (2.17) as the volatility function σ^2. Our first result shows that the matrix $\hat{D}_n$ defined by (2.12) is an approximation of the matrix $D = (D_{ij})_{i,j=1}^d$ given in (2.9), while the second result gives a stochastic expansion for the random variables T_{2n} and $T_{1n,j}$.

Theorem 2.1. *If the assumptions (2.1), (2.2), (2.3), (2.16) and (2.17) are satisfied, then*

$$\hat{D}_n - D \;=\; \frac{1}{n}\bar{X}^T\bar{X} - D = o_p(n^{-\frac{1}{2}})\ , \tag{2.18}$$

where the random matrix $D = (D_{ij})_{i,j=1}^d$ is defined in (2.9).

Theorem 2.2. *If the assumptions (2.1), (2.2), (2.3), (2.16) and (2.17) are satisfied, then*

$$\sqrt{n}(T_{1n,j} - C_j) \;=\; n^{-\frac{1}{2}}\sum\nolimits_{i=1}^{n-1} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}})\sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}})\,\Big[n(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})^2 - 1\Big] + o_p(1)$$

$$\sqrt{n}(T_{2n} - 3C_0) \;=\; n^{-\frac{1}{2}}\sum\nolimits_{i=1}^{n-1} \sigma^4(\frac{i}{n}, X_{\frac{i}{n}})\,\Big[n^2(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})^4 - 3\Big] + o_p(1)$$

We are now in a position to derive the joint asymptotic distribution of the vector $(T_{1n,1}, \ldots, T_{1n,d}, T_{2n})^T$ by an application of a result on the stable convergence of triangular arrays of martingale differences [see the proofs in the appendix]. Throughout this paper the symbol

$$X_n \xrightarrow{\mathcal{D}_{st}} X$$

means that the sequence of random variables converges stably in law. Recall that a sequence of d-dimensional random variables $(X_n)_{n\in\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ converges stably in law with limit X if and only if for any measurable and bounded random variable Y and any bounded and continuous function g the convergence

$$\lim_{n\to\infty} E[Yg(X_n)] \;=\; E[Yg(X)]$$

holds. This is obviously a slightly stronger mode of convergence than convergence in law [see Renyi (1963), Aldous and Eagleson (1978) for more details on stable convergence].

Theorem 2.3. *If the assumptions (2.1), (2.2), (2.3), (2.16) and (2.17) are satisfied, then*

$$\sqrt{n}\left\{\begin{pmatrix} T_{1n,1} \\ \vdots \\ T_{1n,d} \\ T_{2n} \end{pmatrix} - \begin{pmatrix} C_1 \\ \vdots \\ C_d \\ 3C_0 \end{pmatrix}\right\} \xrightarrow{\mathcal{D}_{st}} \int_0^1 \Sigma(t, X_t)^{\frac{1}{2}}\, dB_t \ ,$$

where the random matrix $\Sigma(t, X_t)$ is defined by

$$\Sigma(t,X_t) := \begin{pmatrix} 2s_{11}(t,X_t) & \cdots & 2s_{1d}(t,X_t) & 12s_1(t,X_t) \\ \vdots & \ddots & \vdots & \vdots \\ 2s_{d1}(t,X_t) & \cdots & 2s_{dd}(t,X_t) & 12s_d(t,X_t) \\ 12s_1(t,X_t) & \cdots & 12s_d(t,X_t) & 96v(t,X_t) \end{pmatrix} \tag{2.19}$$

with

$$\begin{aligned} s_{ij}(t,X_t) &:= \sigma_i^2(t,X_t)\sigma_j^2(t,X_t)\sigma^4(t,X_t) \quad 1\le i,j\le d \\ s_i(t,X_t) &:= \sigma_i^2(t,X_t)\sigma^6(t,X_t) \quad 1\le i\le d \\ v(t,X_t) &:= \sigma^8(t,X_t) \end{aligned} \tag{2.20}$$

and $(B_t)_{t\in[0,1]}$ denotes a $(d+1)$-dimensional Brownian motion independent of the process $(W_t)_{t\in[0,1]}$.

As a by-product Theorem 2.3 clarifies the asymptotic properties of estimates of the integrated volatility $C_1 = \int_0^1 \sigma^2(t,X_t)dt$ and similar quantities in the diffusion model (1.1). For example, we obtain for the quadratic variation $Q_n = \sum_{i=1}^{n-1}(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2$

$$\sqrt{n}\{Q_n - \ C_1\} \xrightarrow{\mathcal{D}_{st}} \sqrt{2}\int_0^1 \sigma^2(t,X_t)\, dB_t \ ,$$

where $(B_t)_{t\in[0,1]}$ is a Brownian motion independent of the process $(W_t)_{t\in[0,1]}$. We also note that the property of stable convergence in law of Theorem 2.3 is essential in the following discussion, because due to the randomness of the quantity M^2 the standard delta-method [see e.g. van der Vaart (1998)] is not applicable to derive weak convergence of the random variable

$$\sqrt{n}(T_n - M^2) = \sqrt{n}\Big\{g\Big(T_{1n,1},\dots,T_{1n,d},T_{2n};\hat{D}_n\Big) - g(C_1,\dots,C_d,3C_0;D)\Big\}$$

from the weak convergence of the vector $\sqrt{n}\{(T_{1n,1},\dots,T_{1n,d},T_{2n})^T - (C_1,\dots,C_d,3C_0)^T\}$. Theorem 2.3 yields a slightly stronger statement than weak convergence, which enables us to prove weak convergence of an appropriately standardized version of the statistic T_n by a generalization of the classical delta-method [see the proof of the following result in the appendix].

Theorem 2.4. *If the assumptions (2.1), (2.2), (2.3), (2.16) and (2.17) are satisfied, then*

$$\sqrt{n}(T_n - M^2) \xrightarrow{\mathcal{D}_{st}} \Big(-2(C_1,\dots,C_d)D^{-1},\frac{1}{3}\Big)\int_0^1 \Sigma(t,X_t)^{\frac{1}{2}}\,dB_t\ ,$$

where the random matrix $\Sigma(t,X_t)$ is defined by (2.19) and $(B_t)_{t\in[0,1]}$ denotes a $(d+1)$-dimensional Brownian motion independent of the process $(W_t)_{t\in[0,1]}$.

Remark 2.5. Note that conditionally on the diffusion $(X_t)_{t\in[0,1]}$ the limit distribution is normal with conditional variance

$$\beta^2 = \Big(-2(C_1,\dots,C_d)D^{-1},\frac{1}{3}\Big)\int_0^1 \Sigma(t,X_t)\,dt\,\Big(-2(C_1,\dots,C_d)D^{-1},\frac{1}{3}\Big)^T .$$

Now we define an estimator $\hat{\beta}_n^2$ of the asymptotic conditional variance β^2 by the statistic

$$\hat{\beta}_n^2 = \Big(-2(T_{1n,1},\dots,T_{1n,d})\hat{D}_n^{-1},\frac{1}{3}\Big)\,\hat{\Sigma}_n\,\Big(-2(T_{1n,1},\dots,T_{1n,d})\hat{D}_n^{-1},\frac{1}{3}\Big)^T ,$$

where the random matrix $\hat{\Sigma}_n$ is given by

$$\hat{\Sigma}_n := \begin{pmatrix} 2\hat{s}_{11n} & \cdots & 2\hat{s}_{1dn} & 12\hat{s}_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ 2\hat{s}_{d1n} & \cdots & 2\hat{s}_{ddn} & 12\hat{s}_{dn} \\ 12\hat{s}_{1n} & \cdots & 12\hat{s}_{dn} & 96\hat{v}_n \end{pmatrix} \tag{2.21}$$

with

$$\begin{aligned} \hat{s}_{ijn} &:= \frac{1}{3}n\sum_{k=1}^{n-1}\sigma_i^2(\frac{k}{n},X_{\frac{k}{n}})\sigma_j^2(\frac{k}{n},X_{\frac{k}{n}})(X_{\frac{k+1}{n}}-X_{\frac{k}{n}})^4 \quad 1\le i,j\le d \\ \hat{s}_{in} &:= \frac{1}{15}n^2\sum_{k=1}^{n-1}\sigma_i^2(\frac{k}{n},X_{\frac{k}{n}})(X_{\frac{k+1}{n}}-X_{\frac{k}{n}})^6 \quad 1\le i\le d \\ \hat{v}_n &:= \frac{1}{105}n^3\sum_{k=1}^{n-1}(X_{\frac{k+1}{n}}-X_{\frac{k}{n}})^8 \end{aligned} \tag{2.22}$$

One can prove by similar arguments as given in the appendix for the proof of Theorem 2.1 and 2.2 that

$$\hat{\beta}_n^2 \xrightarrow{P} \beta^2$$

[see Podolskij (2003)]. Note that β^2 is a random variable and for this reason Slutzky's lemma is not directly applicable. However, it follows from the stable convergence of $\sqrt{n}(T_n - M^2)$ established in Theorem 2.4 and Remark 2.1 of Jacod (1997) that the standardized version of $\sqrt{n}(T_n - M^2)$ has a standard normal limit distribution, that is

$$\frac{\sqrt{n}(T_n - M^2)}{\hat{\beta}_n} \xrightarrow{\mathcal{D}_{st}} N(0,1)$$

A tedious calculation shows that

$$\beta^2 \;=\; \frac{8}{3}\int_0^1 \sigma^8(t,X_t)dt + 8\int_0^1 \sigma^4(t,X_t)\{\sigma^2 - P_{\mathcal{U}_d}\sigma^2\}^2(t,X_t)dt \;,$$

where $P_{\mathcal{U}_d}$ denotes the orthogonal projection of the function σ^2 onto the space $\mathcal{U}_d$ =span$\{\sigma_1^2,\dots,\sigma_d^2\}$ with respect to the inner product $< f,g >= \int_0^1 f(t,X_t)g(t,X_t)dt$. If the hypothesis (2.6) is satisfied, then $M^2 = 0$ (a.s.),

$$\beta^2 = \frac{8}{3}\int_0^1 \sigma^8(t,X_t)dt,$$

and we obtain the central limit theorem

$$Z_n \;=\; \sqrt{\frac{3n}{8}}\,\frac{T_n}{\hat{\tau}_n} \xrightarrow{\mathcal{D}_{st}} N(0,1) \quad , \tag{2.23}$$

where the estimator $\hat{\tau}_n$ is defined by

$$\hat{\tau}_n^2 = \frac{1}{n}\sum_{i=1}^n \left\{\sum_{j=1}^d \hat{\theta}_j \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}})\right\}^4$$

and converges to the random variable β^2 in probability. This result provides in principle a simple test for the hypothesis of a parametric form by rejecting the hypothesis (2.6) whenever

$$Z_n \;=\; \sqrt{\frac{3n}{8}}\,\frac{T_n}{\hat{\tau}_n} > u_{1-\alpha} \quad ,$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution. In the following section we will investigate the finite sample properties of a corresponding bootstrap version of this test.

Remark 2.6. The results presented in Theorem 2.4 remain true if the general hypothesis (2.5) of a nonlinear volatility function $H_0 : \sigma^2(t,X_t) = \sigma^2(t,X_t;\theta) \quad \forall\, t \in [0,1] \quad$ (a.s.) for some $\theta \in \Theta$ has to be tested, where $\Theta \subset \mathbb{R}^d$, and the minimum of

$$M_\Theta^2 = \min_{\theta\in\Theta}\int_0^1 [\sigma^2(t,X_t) - \sigma^2(t,X_t,\theta)]^2 dt \tag{2.24}$$

is attained at a unique interior point $\theta_0 \in \Theta$ (for any $\omega \in \Omega$). Then under some regularity assumptions [see e.g. Seber and Wild (1989), pp. 572-574 or Gallant (1987), Chapter 4], Theorem 2.3 and 2.4 remain valid. The proof uses a Taylor expansion and the fact that the sum of squared residuals satisfies

$$\hat{M}_{LSE}^2 = \min_{\theta\in\Theta}\frac{1}{n}\sum_{i=1}^{n-1}\Big(n(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2 - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}, \theta)\Big)^2 = \frac{1}{n}\eta^T\Big(I - G(G^TG)^{-1}G^T\Big)\eta + O_p(\frac{1}{n}) \;,$$

where the matrix G is defined by

$$G^T = \left(\frac{\partial \sigma^2}{\partial \theta}(\frac{i}{n}, X_{\frac{i}{n}}, \theta_0) \right)_{i=1}^{n-1} \in \mathbb{R}^{d \times n-1}$$

and $\eta = (n(X_{\frac{i+1}{n}} - X_{\frac{i}{n}})^2 - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}, \theta_0))_{i=1}^{n-1}$. Roughly speaking this means that the nonlinear model can be treated as the linear model with the d regression functions

$$\sigma_1^2(t, X_t) = \frac{\partial}{\partial \theta_1}\sigma^2(t, X_t, \theta_0), \ldots, \sigma_d^2(t, X_t) = \frac{\partial}{\partial \theta_d}\sigma^2(t, X_t, \theta_0) \ .$$

The stable convergence in law now follows along the lines of Section 4 [see the proof of Theorem 2.3 and 2.4].

3 Finite sample properties and a data example

In this section we study the finite sample properties of the new tests proposed in Section 2. In the case, where the variance function depends only on the time parameter t, Dette and von Lieres und Wilkau (2003) pointed out that the approximation by the normal distribution (as proposed in Remark 2.5) is not very accurate. For this reason these authors proposed a resampling method, and the results of Section 2 show that this method can easily be adapted to the more general situation considered in this paper. The basic idea is to consider at first the stochastic differential equation (1.1) with $b(t, x) \equiv 0$ and simulate the level of the corresponding test statistic by appropriate resampling.

$\sigma^2(t,x) = x^2$	$n = 100$			$n = 200$			$n = 500$		
$b(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
0	20.6	12.2	6.9	20.1	10.1	5.7	21.1	9.6	5.8
2	25.6	14.0	7.2	24.0	12.5	6.2	19.4	9.8	4.0
x	23.0	12.4	5.8	22.5	12.3	6.0	18.9	8.2	4.4
$2-x$	22.8	11.7	6.9	22.9	11.1	6.1	20.5	10.9	6.6
tx	20.3	10.7	4.9	22.5	10.5	5.2	20.9	10.9	5.4

Table 3.1: *Simulated level of the parametric bootstrap test for the hypothesis (3.2) with various drift functions.*

3.1 A bootstrap test for parametric form of the volatility function. In a first step we calculate the least squares estimator

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^d}{\operatorname{argmin}} (\Delta - \bar{X}\theta)^T(\Delta - \bar{X}\theta) = (\bar{X}^T\bar{X})^{-1}\bar{X}^T\Delta \ , \tag{3.1}$$

where the matrix $\bar{X}$ is defined in (2.15) and the vector Δ is given by

$$\Delta = n\Big((X_{\frac{1}{n}} - X_0)^2, (X_{\frac{2}{n}} - X_{\frac{1}{n}})^2, \ldots, (X_1 - X_{\frac{n-1}{n}})^2\Big)^T$$

(in the case of a nonlinear parametric hypothesis we use the corresponding nonlinear least squares estimator). Next the test statistic Z_n defined in (2.23) is calculated from the original data. In a second step data $X^{*(j)}_{i/n}$ $(i = 1,\dots,n;\ j = 1,\dots B)$ from the stochastic differential equation (1.1) with $b(t,x) \equiv 0$ and $\sigma^2(t,x) = \sum_{j=1}^d \hat{\theta}_j \sigma_j^2(t,x)$ are generated and the corresponding bootstrap statistics

$$Z_n^{*(1)}, \dots, Z_n^{*(B)}$$

are calculated. Finally the value of the statistic Z_n is compared with the corresponding quantiles of the bootstrap distribution.

We have investigated the performance of this resampling procedure for the problem of testing various linear hypotheses, where the volatility function depends on the variable x. In Table 3.1 and 3.2 we show for various drift functions $b(t,x)$ the simulated level of the bootstrap test for the hypotheses

$$\tilde{H}_0: \quad \sigma^2(t,X_t) = \theta_1 X_t^2 \quad \text{(a.s.)} \tag{3.2}$$

$$\tilde{H}_0: \quad \sigma^2(t,X_t) = \theta_1 + \theta_2 X_t^2 \quad \text{(a.s.)} \tag{3.3}$$

respectively. The sample size is $n = 100$, 200 and 500, while 1000 simulation runs and $B = 500$ bootstrap replications are used to simulate the rejection probabilities. The null hypothesis (3.3) was chosen as $\theta_1 = \theta_2 = 1$, while we used $\theta_1 = 1$ in (3.2). We observe a rather accurate approximation of the nominal level in all cases under consideration.

$\sigma^2(t,x) = x^2+1$	$n = 100$			$n = 200$			$n = 500$		
$b(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
0	20.5	11.4	6.6	20.6	10.7	5.3	21.1	11.9	5.1
2	20.4	11.0	6.3	20.1	9.6	5.5	17.5	8.8	4.8
x	20.4	10.5	5.8	20.6	11.8	5.5	18.9	8.2	4.4
$2-x$	20.1	9.6	4.2	19.7	9.1	4.6	19.2	9.2	5.0
tx	18.4	8.8	4.0	19.5	10.1	5.5	21.5	10.1	4.3

Table 3.2: *Simulated level of the parametric bootstrap test for the hypothesis (3.3) with various drift functions.*

$b(t,x) = 2-x$	$n = 100$			$n = 200$			$n = 500$		
$\sigma^2(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
$1+x^2$	66.0	62.7	59.5	76.9	74.3	70.8	86.0	84.4	82.3
x^2	71.2	63.6	58.8	80.1	73.6	69.5	92.2	88.0	83.6
$5\|x\|^{\frac{3}{2}}$	94.0	93.2	91.9	99.0	98.5	98.2	99.9	99.9	99.9
$5\|x\|$	83.7	82.9	81.7	92.4	92.4	92.2	97.3	97.3	97.3
$(1+x)^2$	78.5	76.5	74.7	88.3	87.5	86.5	97.5	97.4	97.0

Table 3.3: *Simulated power of the parametric bootstrap test for the hypothesis of a constant volatility function and various alternatives.*

$b(t,x) = 2 - x$	$n = 100$			$n = 200$			$n = 500$		
$\sigma^2(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
$1 + x^2$	46.8	37.4	30.9	50.3	40.4	34.0	50.6	42.0	34.1
1	25.6	13.5	9.8	31.8	20.0	14.2	35.5	22.6	17.9
$5\|x\|^{\frac{3}{2}}$	39.0	37.0	35.2	43.1	40.4	37.2	47.3	44.8	42.2
$5\|x\|$	53.9	48.3	43.2	57.7	51.6	46.0	70.0	63.4	57.3
$(1 + x)^2$	51.4	47.8	43.0	54.3	49.6	45.7	64.9	59.8	55.8

Table 3.4: *Simulated power of the parametric bootstrap test for the hypothesis (3.2) and various alternatives.*

$b(t,x) = 2 - x$	$n = 100$			$n = 200$			$n = 500$		
$\sigma^2(t,x)$	20%	10%	5%	20%	10%	5%	20%	10%	5%
$5\|x\|^{\frac{3}{2}}$	31.7	27.4	22.1	33.2	30.1	25.9	37.3	33.0	29.8
$5\|x\|$	30.3	23.9	19.0	35.2	28.2	23.2	37.0	30.1	24.8
$(1 + x)^2$	27.6	21.4	17.7	29.3	23.6	19.2	35.0	29.5	22.7

Table 3.5: *Simulated power of the parametric bootstrap test for the hypothesis (3.3) and various alternatives.*

In Table 3.3 - 3.5 we illustrate the rejection probabilities of the parametric bootstrap test for different alternatives, where the drift function is given by $b(t,x) = 2 - x$. Table 3.3 corresponds to the problem of testing for a constant volatility function [for simulation results of the nominal level of the test for a constant volatility function, see Dette and von Lieres und Wilkau (2003)], while Table 3.4 and 3.5 show the rejection probabilities of the bootstrap test for the hypothesis (3.2) and (3.3), respectively. We observe that the test is able to detect a deviance from the null hypothesis with reasonable probability, where the size of the power depends sensitively on the distance between the parametric hypothesis and the alternative. For example, deviations from the hypothesis of constant volatility are usually easier detected than deviations from the parametric hypothesis (3.2) and (3.3). It is also worthwhile to mention that the simulated rejection probabilities in Table 3.4 and 3.5 are lower than the corresponding values in Table 3.3.

3.2 A data Example. In this section we illustrate the application of the proposed bootstrap test to tick-by-tick data. As a specific example we consider the exchange rate between the Swiss franc and the US dollar. The data were available at approximately 1530 days between May 20th, 1985 and April 12th, 1991, and were recorded between 20 and 25 times per hour. Obvious errors in the data have been removed. The authors would like to thank V. Spokoiny (WIAS Berlin), who kindly send us this data set for this evaluation.

We illustrate the application of the proposed procedures for the problem of testing the hypothesis of constant volatility and the hypotheses (3.2) and (3.3). For this we applied the proposed tests to the log-returns of the available data at each day and display some of the results in Table 3.6 for two weeks in November 1985. No clear pattern can be observed for the structure of the volatility function. At four days of the first week the hypothesis of homoscedasticity cannot be rejected (at the 5%-level), while there is a clear rejection at most days of the second week. The situation for the hypotesis (3.2) and (3.3) is very similar. There are a few day in the first week, where the null-hypothesis is not rejected, but in general there is not too much evidence that a diffusion with the volatility function of the form (3.2) and (3.3) describes the data adequately. The similarity of the results for the different hypotheses can be partially explained by the observation that there is not too much variation in the log-returns at all.

It is also of interest to compare these findings with the corresponding test for the hypothesis

$$H_0: \ \sigma^2(t,x) = \theta_1 + \theta_2 (t - 1/2)^2 \ . \tag{3.4}$$

The form of this variance function is motivated by the fact that the typical diurnal pattern of market activity results in a U-shaped curve [see e.g. Gerety and Mulherin (1994)]. Note that the test suggested in the present paper reduces to the test proposed by Dette and von Lieres und Wilkau (2003) in the case where the volatility function does not depend on x. These authors observed larger p-values for the hypothesis (3.4). Thus for the data under investigation a quadratic volatility function depending only on the time t seems to be more appropriate than a quadratic volatility function depending only on the state x. Further details of this evaluation are available from the authors.

November	11th	12th	13th	14th	15th	25th	26th	27th	28th	29th
n	213	124	258	263	249	199	171	226	197	190
constant	0.012	1.000	0.212	0.582	0.068	1.000	0.000	0.038	0.014	0.000
(3.2)	0.008	1.000	0.230	0.588	0.006	1.000	0.000	0.038	0.028	0.000
(3.3)	0.008	1.000	0.216	0.650	0.314	1.000	0.050	0.048	0.046	0.010

Table 3.6 *p-values of the bootstrap test for various hypotheses. Ten days between November 11th and November 15th and November 25th and November 29th in 1985 are displayed. The second row shows the number of the available data at each day.*

Acknowledgements. Parts of this paper are based on the second author‘s Diploma thesis [Podolskij (2003)] written at the Ruhr-Universität Bochum. The work of the authors was supported by the Deutsche Forschungsgmeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen). The authors are also grateful to R. Höpfner for some help with the references and to J. Jacod for leading our attention to his 1997 paper, which was crucial for some of the arguments in the proofs of our results.

4 Appendix: Proofs.

We use the decomposition

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

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

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

respectively, and the following auxiliary result, which has been established in Dette and von Lieres und Wilkau (2003).

Lemma 4.1. *If assumptions (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 $m \ge p$

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

4.1 Proof of Theorem 2.1.

We only consider the case $d = 1$. The case $d > 1$ follows by exactly the same arguments with an additional amount of notation. For a function $h : [0,1] \times \mathbb{R} \to \mathbb{R}$ we denote by $\frac{\partial}{\partial y} h$ its partial derivative with respect to the second argument and consider the decomposition

$$D - \frac{1}{n}\bar{X}^T\bar{X} \;=\; H_n^{(1)} + H_n^{(2)} + O_p(n^{-1}), \tag{4.5}$$

where the random variables $H_n^{(1)}$ and $H_n^{(2)}$ are defined by

$$H_n^{(1)} = \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left\{\sigma_1^4(t, X_t) - \sigma_1^4(\frac{i}{n}, X_t)\right\} dt\ , \quad H_n^{(2)} = \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left\{\sigma_1^4(\frac{i}{n}, X_t) - \sigma_1^4(\frac{i}{n}, X_{\frac{i}{n}})\right\} dt\ .$$

Observing (2.16) and the inequality

$$E(\sup_{0\le s\le t} |X_s|^{2m}) \le C_{m,K}(1 + E(|X_0|^{2m}))\, e^{C_{m,K}\, t} \tag{4.6}$$

for some constant $C_{m,K} > 0$ [see Karatzas and Shreve (1991), p. 306], it follows that

$$H_n^{(1)} = o_p(n^{-\frac{1}{2}})\ . \tag{4.7}$$

In order to derive a corresponding estimate for the second term we use a Taylor expansion and obtain

$$H_n^{(2)} \;=\; \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (X_t - X_{\frac{i}{n}})\frac{\partial}{\partial y}\sigma_1^4(\frac{i}{n}, X_{\frac{i}{n}}) + \frac{1}{2}(X_t - X_{\frac{i}{n}})^2 \frac{\partial^2}{\partial y^2}\sigma_1^4(\frac{i}{n}, X^t_{\frac{i}{n}})dt \;=\; H_n^{(2.1)} + H_n^{(2.2)}, \tag{4.8}$$

where $|X^t_{\frac{i}{n}} - X_{\frac{i}{n}}| \le |X_t - X_{\frac{i}{n}}|$ and the last equality defines the random variables $H_n^{(2.1)}$ and $H_n^{(2.2)}$. An application of assumption (2.17), the inequality

$$E(|X_t - X_s|^{2m}) \le C_{m,K}(1 + E(|X_0|^{2m}))(t-s)^m \ , \quad m \in \mathbb{N} \tag{4.9}$$

[see e.g. Karatzas and Shreve (1991)] and the Cauchy-Schwarz inequality yield

$$H_n^{(2.2)} = O_p(n^{-1}) \ . \tag{4.10}$$

The first term on the right hand side in (4.8) can be decomposed as follows

$$\begin{aligned} H_n^{(2.1)} &= \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (A_t - A_{\frac{i}{n}}) \frac{\partial}{\partial y} \sigma_1^4(\frac{i}{n}, X_{\frac{i}{n}}) + \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} (M_t - M_{\frac{i}{n}}) \frac{\partial}{\partial y} \sigma_1^4(\frac{i}{n}, X_{\frac{i}{n}}) dt \\ &= H_n^{(2.1.1)} + H_n^{(2.1.2)} \ , \end{aligned} \tag{4.11}$$

where the last equality defines $H_n^{(2.1.1)}$ and $H_n^{(2.1.2)}$ and the processes $(M_t)_{t\in[0,1]}$ und $(A_t)_{t\in[0,1]}$ are given by (4.2). For the term $H_n^{(2.1.1)}$ we have by a further application of the Cauchy-Scharz inequality , assumption (2.17) and Lemma 4.1

$$H_n^{(2.1.1)} = O_p(n^{-1}) \tag{4.12}$$

For the squared expectation of the remaining term we obtain by similar arguments [see Podolskij (2003)] that

$$E[(H_n^{(2.1.2)})^2] \le n^{-1} \sum_{i=1}^{n-1} E\left[\int_{\frac{i}{n}}^{\frac{i+1}{n}} (M_t - M_{\frac{i}{n}})^2 \left(\frac{\partial}{\partial y} \sigma_1^4(\frac{i}{n}, X_{\frac{i}{n}}) \right)^2 dt \right] = O(n^{-2}) \ , \tag{4.13}$$

which gives $H_n^{(2.1.2)} = O_p(n^{-1})$ and a combination of this result with (4.5), (4.7), (4.8), (4.10), (4.11) and (4.12) yields the assertion of the theorem. □

4.2 Proof of Theorem 2.2.

Note that it follows by a similar reasoning as in Dette and von Lieres und Wilkau (2003) that

$$\begin{aligned} T_{1n,j} &= \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}})(M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^2 + o_p(n^{-\frac{1}{2}}) \ , \qquad j = 1, \ldots, d \ , \\ T_{2n} &= n \sum_{i=1}^{n-1} (M_{\frac{i+1}{n}} - M_{\frac{i}{n}})^4 + o_p(n^{-\frac{1}{2}}) \ , \end{aligned} \tag{4.14}$$

where the process $(M_t)_{t\in[0,1]}$ is defined in (4.2). An application of Ito's formula shows that

$$T_{1n,j} - C_j = U_{1n,j}^{(1)} + U_{1n,j}^{(2)} + o_p(n^{-\frac{1}{2}}) \ , \tag{4.15}$$

where the quantities $U^{(i)}_{1n,j}$ $(i=1,2)$ are defined as

$$U^{(1)}_{1n,j} = \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} 2(M_u - M_{\frac{i}{n}})\sigma(u, X_u) dW_u,$$

$$U^{(2)}_{1n,j} = \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma^2(u, X_u)\{\sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) - \sigma_j^2(u, X_u)\} du \ .$$

Using similar arguments as in the proof of Theorem 2.1 it can be shown that $U^{(2)}_{1n,j} = o_p(n^{-\frac{1}{2}})$, which yields

$$T_{1n,j} - C_j = U^{(1)}_{1n,j} + o_p(n^{-\frac{1}{2}}) \ , \quad j = 1, \ldots, d \ . \tag{4.16}$$

For the dominating term in (4.16) it follows

$$\begin{aligned} U^{(1)}_{1n,j} \ &= \ \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} 2(M_u - M_{\frac{i}{n}})\sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_u \\ &+ \ \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} 2(M_u - M_{\frac{i}{n}})[\sigma(u, X_u) - \sigma(\frac{i}{n}, X_{\frac{i}{n}})] dW_u \\ &= \ \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} 2(M_u - M_{\frac{i}{n}})\sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_u + o_p(n^{-\frac{1}{2}}) \ , \end{aligned} \tag{4.17}$$

where we used the estimate

$$\begin{aligned} & E\left[\left(\sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} 2(M_u - M_{\frac{i}{n}})[\sigma(u, X_u) - \sigma(\frac{i}{n}, X_{\frac{i}{n}})] dW_u\right)^2\right] \\ &= \sum_{i=1}^{n-1} E\left[\sigma_j^4(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} 4(M_u - M_{\frac{i}{n}})^2[\sigma(u, X_u) - \sigma(\frac{i}{n}, X_{\frac{i}{n}})]^2 du\right] \ = \ O(n^{-2}) \end{aligned} \tag{4.18}$$

for the estimation of the second term in (4.17). For the result (4.18) we used (2.1), (2.16) and (4.4). A similar argument then yields

$$\begin{aligned} U^{(1)}_{1n,j} \ &= \ 2\sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left[\int_{\frac{i}{n}}^{u} \sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_s\right] \sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_u + o_p(n^{-\frac{1}{2}}) \\ &= \ 2\sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \int_{\frac{i}{n}}^{\frac{i+1}{n}} (W_u - W_{\frac{i}{n}}) dW_u + o_p(n^{-\frac{1}{2}}) \\ &= \ \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \left[(W^2_{\frac{i+1}{n}} - W^2_{\frac{i}{n}}) - \frac{1}{n} - 2W_{\frac{i}{n}}(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})\right] + o_p(n^{-\frac{1}{2}}) \\ &= \ \sum_{i=1}^{n-1} \sigma_j^2(\frac{i}{n}, X_{\frac{i}{n}}) \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) \left[(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})^2 - \frac{1}{n}\right] + o_p(n^{-\frac{1}{2}}), \end{aligned} \tag{4.19}$$

where the last line follows from the identity

$$\int_a^b W_t^{k-1} dW_t = \frac{1}{k}\left[(W_b^k - W_a^k) - \frac{k(k-1)}{2}\int_a^b W_t^{k-2}dt\right] \quad k \geq 2 \ . \tag{4.20}$$

This establishes the first representation in Theorem 2.2. For the second term we use Ito's formula and obtain

$$T_{2n} - 3C_0 = U_{2n}^{(1)} + U_{2n}^{(2)} + U_{2n}^{(3)} + o_p(n^{-\frac{1}{2}}) \ , \tag{4.21}$$

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

$$U_{2n}^{(1)} = n\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} 4(M_u - M_{\frac{i}{n}})^3\sigma(u, X_u)dW_u \tag{4.22}$$

$$U_{2n}^{(2)} = n\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\left[\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}})\sigma(s, X_s)dW_s\right]\sigma^2(u, X_u)du \tag{4.23}$$

$$U_{2n}^{(3)} = \sum_{i=1}^{n-1}\left[n\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{u} 6\sigma^2(s, X_s)\sigma^2(u, X_u)dsdu - 3\int_{\frac{i}{n}}^{\frac{i+1}{n}}\sigma^4(s, X_s)ds\right] \ . \tag{4.24}$$

For the last term we have

$$\begin{aligned} E|U_{2n}^{(3)}| &= E\Big|\sum_{i=1}^{n-1}\left[3n\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\sigma^2(s, X_s)\{\sigma^2(u, X_u) - \sigma^2(s, X_s)\}dsdu\right]\Big| \\ &= E\Big|\sum_{i=1}^{n-1}\left[3n\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\{\sigma^2(s, X_s) - \sigma^2(\frac{i}{n}, X_{\frac{i}{n}})\}\{\sigma^2(u, X_u) - \sigma^2(s, X_s)\}dsdu\right]\Big| \ = \ O(n^{-1}) \end{aligned} \tag{4.25}$$

The treatment of the two other statistics in (4.21) is more complicated. Note that a Taylor expansion yields

$$U_{2n}^{(2)} = U_{2n}^{(2.1)} + U_{2n}^{(2.2.1)} + U_{2n}^{(2.2.2)} + U_{2n}^{(2.2.3)} \ , \tag{4.26}$$

with

$$U_{2n}^{(2.1)} \ = \ n\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}})\sigma(s, X_s)dW_s \ \sigma^2(\frac{i}{n}, X_{\frac{i}{n}})du \tag{4.27}$$

$$U_{2n}^{(2.2.1)} \ = \ n\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}})\sigma(s, X_s)dW_s\left(\sigma^2(u, X_u) - \sigma^2(\frac{i}{n}, X_u)\right)du \tag{4.28}$$

$$U_{2n}^{(2.2.2)} \ = \ n\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}})\sigma(s, X_s)dW_s(X_u - X_{\frac{i}{n}})\frac{\partial}{\partial y}\sigma^2(\frac{i}{n}, X_{\frac{i}{n}})du \tag{4.29}$$

$$U_{2n}^{(2.2.3)} \ = \ n\sum_{i=1}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}})\sigma(s, X_s)dW_s\frac{1}{2}(X_u - X_{\frac{i}{n}})^2\frac{\partial^2}{\partial y^2}\sigma^2(\frac{i}{n}, X_{\frac{i}{n}}^u)du \tag{4.30}$$

for some $X^u_{\frac{i}{n}}$ with $|X_u - X^u_{\frac{i}{n}}| \leq |X_u - X_{\frac{i}{n}}|$ $(i = 1, \dots, n-1)$. A straightforward but tedious calculation [see Podolskij (2003)] shows

$$E[|U_{2n}^{(2.2.1)}|] = o(n^{-\frac{1}{2}}) \ , \quad E[|U_{2n}^{(2.2.3)})|] = O(n^{-1}) \ . \tag{4.31}$$

For the remaining term $U_{2n}^{(2.2.2)}$ we use a further decomposition

$$U_{2n}^{(2.2.2)} = U_{2n}^{(2.2.2.1)} + U_{2n}^{(2.2.2.2)} + U_{2n}^{(2.2.2.3)} \ , \tag{4.32}$$

where the quantities $U_{2n}^{(2.2.2.i)}$ $(i = 1, 2, 3)$ are defined as follows

$$U_{2n}^{(2.2.2.1)} = n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left[\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}}) \sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_s \right] (M_u - M_{\frac{i}{n}}) \frac{\partial}{\partial y} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) du \tag{4.33}$$

$$\begin{aligned} U_{2n}^{(2.2.2.2)} = {} & n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left[\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}}) [\sigma(s, X_s) - \sigma(\frac{i}{n}, X_{\frac{i}{n}})] dW_s \right] \\ & \times \ (M_u - M_{\frac{i}{n}}) \frac{\partial}{\partial y} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) du \end{aligned} \tag{4.34}$$

$$U_{2n}^{(2.2.2.3)} = n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left[\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}}) \sigma(s, X_s) dW_s \right] (A_u - A_{\frac{i}{n}}) \frac{\partial}{\partial y} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) du \ . \tag{4.35}$$

The second and third term can be treated by similar methods as the terms in (4.28) and (4.31) that is

$$E[|U_{2n}^{(2.2.2.2)}|] = O(n^{-1}) \ , \quad E[|U_{2n}^{(2.2.2.3)}|] = O(n^{-1}) \ . \tag{4.36}$$

Finally we obtain

$$\begin{aligned} U_{2n}^{(2.2.2.1)} = {} & n \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left[\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}}) \sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_s \right] \left[\int_{\frac{i}{n}}^{u} \sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_s \right] \frac{\partial}{\partial y} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) du \\ + n \sum_{i=1}^{n-1} & \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left[\int_{\frac{i}{n}}^{u} 12(M_s - M_{\frac{i}{n}}) \sigma(\frac{i}{n}, X_{\frac{i}{n}}) dW_s \int_{\frac{i}{n}}^{u} (\sigma(s, X_s) - \sigma(\frac{i}{n}, X_{\frac{i}{n}})) dW_s \right] \frac{\partial}{\partial y} \sigma^2(\frac{i}{n}, X_{\frac{i}{n}}) du \end{aligned}$$

It can be shown by a tedious calculation [see Podolskij (2003)] that both terms are of order $O_p(n^{-1})$ (note that the first term is a sum of uncorrelated and centered random variables) which yields observing (4.26), (4.30), (4.32) and (4.36)

$$U_{2n}^{(2)} = U_{2n}^{(2.1)} + o(n^{-\frac{1}{2}}) \ , \tag{4.37}$$

where the statistic $U_{2n}^{(2.1)}$ is defined in (4.27). From (4.21) and (4.25) we thus obtain

$$T_{2n} - 3C_0 \ = \ U_{2n}^{(1)} + U_{2n}^{(2.1)} + o_p(n^{-\frac{1}{2}}) \ , \tag{4.38}$$

where the random variable $U_{2n}^{(1)}$ is defined in (4.22). Observing formula (4.20) it follows that

$$
\begin{aligned}
\sqrt{n}(T_{2n} - 3C_0) &= 4n^{3/2}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n}, X_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}(W_u - W_{\frac{i}{n}})^3 dW_u \\
&+ 12n^{3/2}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n}, X_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\left[\int_{\frac{i}{n}}^{u}(W_s - W_{\frac{i}{n}})dW_s\right]du + o_p(1) \\
&= n^{-\frac{1}{2}}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n}, X_{\frac{i}{n}})\left[n^2(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})^4 - 6n^2\int_{\frac{i}{n}}^{\frac{i+1}{n}}(W_u - W_{\frac{i}{n}})^2 du\right] \\
&+ 6n^{\frac{3}{2}}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n}, X_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\left[(W_u - W_{\frac{i}{n}})^2 - (u - \frac{i}{n})\right]du + o_p(1) \\
&= n^{-\frac{1}{2}}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n}, X_{\frac{i}{n}})\left[n^2(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})^4 - 3\right] + o_p(1)\ , \qquad (4.39)
\end{aligned}
$$

which is the second assertion of the theorem. □

4.3 Proof of Theorem 2.3.

We only consider the case $d = 1$, the general case is treated exactly in the same way. We will make extensive use of a limit theorem, which has recently been established by Jacod (1997). To be precise let $M = (M_s)_{s\in[0,t]}$ denote a fixed d-dimensional martingale defined on a probability space $(\Omega, \mathcal{F}, I\!\!F, P)$, where $I\!\!F = (\mathcal{F}_s)_{s\in[0,t]}$ is the canonical filtration of M and $\mathcal{F} = \mathcal{F}_1$. In the following we discuss further q-dimensional martingales $N = (N_s)_{s\in[0,t]}$ defined on $(\Omega, \mathcal{F}, I\!\!F, P)$ such that $N_0 = 0$ a.s.. Let $\mathcal{M}_b$ denote the set of all bounded martingales on $(\Omega, \mathcal{F}, I\!\!F, P)$, $\mathcal{M}_2$ be the set of all quadratic integrable martingales on $(\Omega, \mathcal{F}, I\!\!F, P)$ and define $\mathcal{M}_b(M^\perp) := \{N \in \mathcal{M}_b \mid \langle M, N\rangle = 0\}$ as the set of all bounded martingales, which are orthogonal with respect M (here $\langle M, N\rangle$ denote the quadratic variation between M and N). We are interested in the asymptotic properties of a sequence of q-dimensional martingales $(Z^n)_{n\geq 1}$ with respect to the filtrations $(\mathcal{F}_{\frac{[ns]}{n}})_{s\in[0,t]}$, which are representable in the form

$$
Z_t^n = \sum_{i=1}^{[nt]} Y_{in}\ , \qquad (4.40)
$$

where Y_{in} are q-dimensional $\mathcal{F}_{\frac{i}{n}}$-measurable random variables. In the following we denote with $(\bar\Omega, \bar{\mathcal{F}}, \bar P)$ an extension of $(\Omega, \mathcal{F}, P)$, which contains a d-dimensional Brownian motion $B = (B_s)_{s\in[0,t]}$, which is independent of the basic martingale M. The following result has been established by Jacod (1997).

Theorem 4.2. *Assume that the basic martingale $M \in \mathcal{M}_2$ is continuous and $Y_{in} \in \mathcal{M}_2$ for all $i, n \geq 1$. Assume further that there exist two continuous stochastic processes F and G on $(\Omega, \mathcal{F}, P)$ with values onto $\mathbb{R}^{q\times q}$ and $\mathbb{R}^{q\times d}$, respectively, which satisfy the conditions*

$$
\sum\nolimits_{i=1}^{[nt]} E[Y_{in}Y_{in}^T|\mathcal{F}_{\frac{i-1}{n}}] \xrightarrow{P} F_t \quad \forall t \in [0, 1] \qquad (4.41)
$$

(4.42) $$\sum_{i=1}^{[nt]} E[Y_{in}(M_{\frac{i}{n}} - M_{\frac{i-1}{n}})^T|\mathcal{F}_{\frac{i-1}{n}}] \xrightarrow{P} G_t \quad \forall t \in [0,1]$$

(4.43) $$\sum_{i=1}^{[nt]} E[|Y_{in}|^2 I_{\{|Y_{in}|>\varepsilon\}}|\mathcal{F}_{\frac{i-1}{n}}] \xrightarrow{P} 0 \quad \forall \varepsilon > 0$$

(4.44) $$\sum_{i=1}^{[nt]} E[Y_{in}(N_{\frac{i}{n}} - N_{\frac{i-1}{n}})^T|\mathcal{F}_{\frac{i-1}{n}}] \xrightarrow{P} 0 \quad \forall t \in [0,1], \forall N \in \mathcal{M}_b(M^\perp)$$

(4.45) $$d\langle M^i, M^i\rangle_s \ll ds \quad and \quad dF_s^{ii} \ll ds \quad \forall i$$

where M^i denotes the i-th component of M and F^{ii} the element in the position (i,i) of F. If additionally there exist predictable processes u,v,w onto $\mathbb{R}^{q\times d}$, $\mathbb{R}^{d\times d}$ and $\mathbb{R}^{q\times q}$, respectively, such that the representations

(4.46) $$\langle M, M^T\rangle_s = \int_0^s v(x)v(x)^T dx \quad \forall s \in [0,t]$$

(4.47) $$G_s = \int_0^s u(x)v(x)v(x)^T dx \quad \forall s \in [0,t]$$

(4.48) $$F_s = \int_0^s u(x)v(x)v(x)^T u(x)^T + w(x)w(x)^T dx \quad \forall s \in [0,t]$$

are valid, then the stochastic process defined in (4.40) converges stably in law, that is

(4.49) $$Z_t^n \xrightarrow{\mathcal{D}_{st}} Z_t = \int_0^t v(s)dM_s + \int_0^t w(s)dB_s \ .$$

For an application of Theorem 4.2 in the proof of Theorem 2.3 we define the random variables

$$Y_{in}^{(1)} := n^{-\frac{1}{2}}\sigma^2(\frac{i-1}{n}, X_{\frac{i-1}{n}})\sigma_1^2(\frac{i-1}{n}, X_{\frac{i-1}{n}}) \left[n(W_{\frac{i}{n}} - W_{\frac{i-1}{n}})^2 - 1\right] ,$$

$$Y_{in}^{(2)} := n^{-\frac{1}{2}}\sigma^4(\frac{i-1}{n}, X_{\frac{i-1}{n}}) \left[n^2(W_{\frac{i}{n}} - W_{\frac{i-1}{n}})^4 - 3\right]$$

and $Y_{in} = (Y_{in}^{(1)}, Y_{in}^{(2)})^T$. Theorem 2.2 shows that the assertion of Theorem 2.3 follows if the stable convergence in law

(4.50) $$\sum_{i=2}^{n} Y_{in} \xrightarrow{\mathcal{D}_{st}} \int_0^1 \Sigma(s, X_s)^{\frac{1}{2}} \, dB_s$$

can be established. Obviously $(Y_{in})_{2\le i\le n, n\in\mathbb{N}}$ is a triangular array of martingale differences with respect to the filtration $(\mathcal{F}_{\frac{i}{n}})_{2\le i\le n, n\in\mathbb{N}}$ and Theorem 4.1 is applicable. For this we have to check the assumptions (4.41) - (4.48), where the basic martingale is given by the Brownian motion W. Note that for $M \equiv W$ condition (4.44) is obvioulsy satisfied, because the orthogonality with respect to the Brownian motion implies $N_t = E[N_0] = 0$ for $t \in [0,1]$ a.s. We prove the remaining properties for the first component

$Y_{in}^{(1)}$ of the vector Y_{in}. It is easy to see that the argument for the vector Y_{in} is exactly the same. From the independence of the increments of the Brownian motion we have

$$\begin{aligned}\sum_{i=2}^{n} E[|Y_{in}^{(1)}|^2|\mathcal{F}_{\frac{i-1}{n}}] &= n^{-1}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n},X_{\frac{i}{n}})\sigma_1^4(\frac{i}{n},X_{\frac{i}{n}})E\left[n(W_{\frac{i+1}{n}}-W_{\frac{i}{n}})^2-1\right]^2 \\ &= 2n^{-1}\sum_{i=1}^{n-1}\sigma^4(\frac{i}{n},X_{\frac{i}{n}})\sigma_1^4(\frac{i}{n},X_{\frac{i}{n}}) \xrightarrow{P} 2\int_0^1 \sigma^4(s,X_s)\sigma_1^4(s,X_s)ds ,\end{aligned}$$

which gives (4.41). Obviously the identity

$$E[Y_{in}^{(1)}(W_{\frac{i}{n}}-W_{\frac{i-1}{n}})|\mathcal{F}_{\frac{i-1}{n}}] = 0$$

yields (4.42). For a proof of the property (4.43) we note that

$$\begin{aligned}E[|Y_{in}^{(1)}|^2 I_{\{Y_{in}^{(1)}|>\varepsilon\}}|\mathcal{F}_{\frac{i-1}{n}}] &\leq \left\{E[|Y_{in}^{(1)}|^4|\mathcal{F}_{\frac{i-1}{n}}]\right\}^{\frac{1}{2}}\left\{E[I_{\{|Y_{in}^{(1)}|>\varepsilon\}}|\mathcal{F}_{\frac{i-1}{n}}]\right\}^{\frac{1}{2}} \\ &= \frac{\sqrt{96}}{n}\sigma^4(\frac{i-1}{n},X_{\frac{i-1}{n}})\sigma_1^4(\frac{i-1}{n},X_{\frac{i-1}{n}})\left\{E[I_{\{|Y_{in}^{(1)}|>\varepsilon\}}|\mathcal{F}_{\frac{i-1}{n}}]\right\}^{\frac{1}{2}} = o_p(n^{-1})\end{aligned}$$

uniformly with respect to i, where the last estimate follows from the fact

$$E|E[I_{\{|Y_{in}^{(1)}|>\varepsilon\}}|\mathcal{F}_{\frac{i-1}{n}}]| \leq P(|Y_{in}^{(1)}|>\varepsilon) \leq \varepsilon^{-1}E[|Y_{in}^{(1)}|] = O(n^{-\frac{1}{2}})$$

(uniformly with respect to i) and Markov's inequality. It is easy to see that the functions

$$v(s)\equiv 1 \qquad u(s)\equiv 0 \qquad w(s)=\sqrt{2}\sigma^2(s,X_s)\sigma_1^2(s,X_s)$$

satisfy the conditions (4.46) - (4.48) of Theorem 4.2. The assertion (4.50) now follows from Theorem 4.2, which completes the proof of Theorem 2.3. □

4.4 Proof of Theorem 2.4.

Applying the mean value theorem we obtain for some random variable ζ_n

$$\begin{aligned}\sqrt{n}(T_n-M^2) &= \sqrt{n}\Big\{g\Big(T_{1n,1},\ldots,T_{1n,d},T_{2n};\hat{D}_n\Big)-g\Big(C_1,\ldots,C_d,3C_0;D\Big)\Big\} \\ &= \sqrt{n}\Big\{g\Big(T_{1n,1},\ldots,T_{1n,d},T_{2n};D\Big)-g\Big(C_1,\ldots,C_d,3C_0;D\Big)\Big\} \\ &+ \sqrt{n}\Big\{g\Big(T_{1n,1},\ldots,T_{1n,d},T_{2n};\hat{D}_n\Big)-g\Big(T_{1n,1},\ldots,T_{1n,d},T_{2n};D\Big)\Big\} \\ &= \sqrt{n}\ \mathrm{grad}\ g(\zeta_n;D)\Big(T_{1n,1}-C_1,\ldots,T_{1n,d}-C_d,T_{2n}-3C_0\Big)^T \\ &+ \sqrt{n}\Big\{g\Big(T_{1n,1},\ldots,T_{1n,d},T_{2n};\hat{D}_n\Big)-g\Big(T_{1n,1},\ldots,T_{1n,d},T_{2n};D\Big)\Big\} ,\end{aligned}$$

where

$$||\zeta_n - (C_1, \ldots, C_d, 3C_0)^T|| \ \leq \ ||(T_{1n,1}, \ldots, T_{1n,d}, T_{2n})^T - (C_1, \ldots, C_d, 3C_0)^T|| \ ,$$

the function g is defined by (2.8) and grad g denotes the gradient of the function g. An application of Theorem 2.1 yields the following estimation for the second term in the above sum

$$\sqrt{n}\Big\{g\Big(T_{1n,1}, \ldots, T_{1n,d}, T_{2n}; \hat{D}_n\Big) - g\Big(T_{1n,1}, \ldots, T_{1n,d}, T_{2n}; D\Big)\Big\} = o_p(1) \ .$$

Observing the stochastic convergence

$$\zeta_n \xrightarrow{P} \Big(C_1, \ldots, C_d, 3C_0\Big)^T$$

and the result of Theorem 2.3, we obtain (see Aldous and Eagleson (1978) and Jacod (1997) for details) that

$$\Big(D, \zeta_n, \sqrt{n}\Big(T_{1n,1} - C_1, \ldots, T_{1n,d} - C_d, T_{2n} - 3C_0\Big)^T\Big) \xrightarrow{\mathcal{D}_{st}} \Big(D, \Big(C_1, \ldots, C_d, 3C_0\Big)^T, Z\Big) \ ,$$

where Z is given by

$$Z := \int_0^1 \Sigma(t, X_t)^{\frac{1}{2}} \ dB_t \ .$$

The assertion of Theorem 2.4 now follows from the continuous mapping theorem for stable convergence (see Aldous and Eagleson (1978)). $\square$

References

Achieser, N.J. (1956). Theory of approximation. Dover Publications Inc., New York

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

Aldous, D.J., Eagleson, G.K. (1978). On mixing and stability of limit theorems. Ann. of Prob. 6(2), 325-331.

O.E. Barndorff-Nielsen, N. Shephard (2003). Realised power variation and stochastic volatility models. Bernoulli 9, 243-265.

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., von Lieres und Wilkau, C. (2003). On a test for a parametric form of volatility in continuous time financial models. Financ. Stoch. 7, 363-384

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

Engle, R.F., Russell, J.R. (1998). Autoregressive conditional duration: a new model for irregularly spaced times series data, Econometric 66, 1127-1162.

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

Gallant, A.R. (1987). Nonlinear statistical models. Wiley, New York.

Gerety, M.S., Mulherin, J.H. (1994). Price formation on stock exchanges: The evolution of trading within the day, Review of Financial Studies 7, 609-629.

Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. Seminaire de Probabilites XXXI, 232-246

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.

Kleinow, T. (2002). Testing the diffusion coefficient. Discussion paper, Sonderforschungsbereich 373 Humboldt-Universität zu Berlin

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.

Podolskij, M. (2003). Tests auf parametrische Struktur der Volatilität in stochastischen Differentialgleichungen. Diplomarbeit, Fakultät für Mathematik, Ruhr-Universität Bochum, Germany.

Renyi, A. (1963). On stable sequences of events. Sankya Ser. A 25, 293-302

Seber, G.A.F., Wild, C.J. (1989). Nonlinear regression. Wiley, New York.

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

Van der Vaart, A. (1998). Asymptotic Statistics. Cambridge University Press.

Woerner, J.H.C. (2003). Estimation of integrated volatility : a unifying approach to model selection and estimation in semimartingale models. Statistics and Decisions 21, 47-68.