THIELE CENTRE
FOR APPLIED MATHEMATICS IN NATURAL SCIENCE

Asymptotic analysis for a simple explicit estimator in Barndorff-Nielsen and Shephard stochastic volatility models

Friedrich Hubalek and Petra Posedel

Research Report

No. 05 | April 2007

Asymptotic analysis for a simple explicit estimator in Barndorff-Nielsen and Shephard stochastic volatility models

This Thiele Research Report is also Research Report number 492 in the Stochastics Series at Department of Mathematical Sciences, University of Aarhus, Denmark.

Asymptotic analysis for a simple explicit estimator in Barndorff-Nielsen and Shephard stochastic volatility models*

Friedrich Hubalek[†] Petra Posedel[‡]

April 12, 2007

Abstract

We provide a simple explicit estimator for discretely observed Barndorff-Nielsen and Shephard models, prove rigorously consistency and asymptotic normality based on the single assumption that all moments of the stationary distribution of the variance process are finite, and give explicit expressions for the asymptotic covariance matrix.

We develop in detail the martingale estimating function approach for a bivariate model, that is not a diffusion, but admits jumps. We do not use ergodicity arguments.

We assume that both, logarithmic returns and instantaneous variance are observed on a discrete grid of fixed width, and the observation horizon tends to infinity. This anaysis is a starting point and benchmark for further developments concerning optimal martingale estimating functions, and for theoretical and empirical investigations, that replace the (actually unobserved) variance process with a substitute, such as number or volume of trades or implied variance from option data.

KEYWORDS:

Martingale estimating functions, stochastic volatility models with jumps, consistency and asymptotic normality

1 Introduction

In [BNS01] Barndorff-Nielsen and Shephard introduced a class of stochastic volatility models in continuous time, where the instantaneous variance follows an Ornstein-Uhlenbeck type process driven by an increasing Lévy process. Those models allow flexible modelling, capture many stylized facts of financial time series, and

*We thank Ole Barndorff-Nielsen, Michael Sørensen, Bent Nielsen and Søren Johansen for helpful comments. We thank Mathieu Kessler for making his PhD-thesis available to us. Financial support from the Austrian Science Fund (FWF) under grant P15889 is gratefully acknowledged.

†Vienna University of Technology, Austria

‡Department of Mathematics, Faculty of Economics & Business Zagreb, Croatia

yet are of great analytical tractability. For further information see also [BNNS02]. BNS-models, as we will call them from now on, are affine models in the sense of [DPS00] and [DFS03], where the associated Riccati type equations can be solved up to quadrature in general. In several concrete cases the integration can be performed explicitly in closed form in terms of elementary functions, see [NV03] and [Ven01].

BNS-models have been studied from various points of view in mathematical finance and related fields. In [NV03] option pricing and structure preserving martingale measures are studied. In [BK05, BMB05, BG05, RS06] the minimal entropy martingale measure is investigated. The papers [BKR03, Lin06] address the portfolio optimization problem. Baysian/MCMC/computer intensive estimation is already in the seminal paper [BNS01], and in the works [RPD04, GS01, FSS01, tH03]. The papers [Jam05, Jam06] exploit the analytical tractability to develop maximum likelihood estimation using the results of [CM00, CR90] for Dirichlet processes. BNS models are also treated in the textbooks [CT04, Sch03].

Strangely though, it seems that statistical estimation of the model is the most difficult problem, and most of the work in that area focused on computationally intensive methods.

The contributions of the present paper are as follows: first we develop a simple and explicit estimator for BNS models. Secondly, we give rigorous proofs of its consistency and asymptotic normality. In doing so we compute explicitly the asymptotic covariance matrix and develop to that purpose formulas for arbitrary bivariate integer moments of returns and variance. Thirdly we provide a detailed application of the theory of martingale estimating functions in a non-diffusion setting, including numerical illustrations.

The literature on estimation for discretely observed diffusions is vast, a few references are [Uch04b, Uch04a, DS04, MR03, KP02, Jac02, Jac01, Sør01, BS01, Kes00, KS99a, Sør97, BS95]. In particular, the martingale estimating function approach is used, developed and studied for example in [Sø99], [Sø00], [Sø97]. In the diffusion setting the major difficulty is that the transition probabilities are not known and are difficult to compute. In contrast to that, the characteristic function of the transition probability is known in closed form for many BNS models and the transition probability can be computed with Fourier methods with high precisions. Yet the model exhibits other peculiarities, see the remarks in section 2.3.

In the present paper we explore the joint distribution of logarithmic returns X and the instantaneous variance V supposing that both processes can be observed in discrete time. Since the joint conditional moment-generating function of (X, V) is known in closed form we obtain close form expressions for the join conditional moments up to any desired order which yields a sequence of martingale differences. We employ then the large sample properties of those, in particular the strong law of large numbers for martingales and martingale central limit theorem. In this way we do not need ergodicity, mixing conditions, etc.[1]

The remainder of the paper is organized as follows: in section 2.1 we describe the class of BNS models in continuous time and present two concrete examples, the $\Gamma-$OU and IG-OU model. In section 2.2 we introduce the quantities observed in

[1]Let us mention though, that the martingale strong law and the ergodic theorem have similar proofs and can be derived from a common source, [Rao73].

discrete time that are used for estimation. Section 2.3 contains some remarks of particular features of the model and its estimation. In section 3 we present the estimating equations, their explicit solution which is our estimator and prove its consistency and asymptotic normality. This estimator is reviewed in section 3.4 in a general framework given in the lecture notes [Sø97]. In section 4 we present numerical illustrations. In section 5 we sketch further and alternative developments, in particular concerning the issue that volatility is typically not observed in discrete time. Explicit moment calculations of any order can be found in the appendix A. Appendix B contains explicit expressions required for the asymptotic covariance, and in Appendix C we provide for the readers convenience a simple multivariate martingale central limit theorem.

2 The model

2.1 The continuous time model

2.1.1 The general setting

As in Barndorff-Nielsen and Shepard [BNS01], we assume that the price process of an asset S is defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and is given by $S_t = S_0 \exp(X_t)$ with $S_0 > 0$ a constant. The process of logarithmic returs X and the instantaneous variance process V satisfy

$$dX(t) = (\mu + \beta V(t-))dt + \sqrt{V(t-)}dW_\theta(t) + \rho dZ_\lambda(t), \quad X(0) = 0. \tag{2.1}$$

and

$$dV(t) = -\lambda V(t-)dt + dZ_\lambda(t), \quad V(0) = V_0, \tag{2.2}$$

where the parameters μ, β, ρ and λ are real constants with $\lambda > 0$. The process W is a standard Brownian motion, the process Z is an increasing Lévy process, and we define $Z_\lambda(t) = Z(\lambda t)$ for notational simplicity. Adopting the terminology introduced by Barndorff-Nielsen and Shepard, we will refer to Z as the *background driving Lévy process* (BDLP). The Brownian motion W and the BDLP Z are independent and $(\mathcal{F}_t)$ is assumed to be the usual augmentation of the filtration generated by the pair (W, Z_λ). The random variable V_0 has a self-decomposable distribution corresponding to the BDLP such that the process V is strictly stationary and

$$E[V_0] = \zeta, \qquad \mathrm{Var}[V_0] = \eta. \tag{2.3}$$

To shorten the notation we introduce the parameter vector

$$\theta = (\lambda, \zeta, \eta, \mu, \beta, \rho)^\top, \tag{2.4}$$

and the bivariate process

$$\mathbf{X} = (X, V). \tag{2.5}$$

If the distribution of V_0 is from a particular class D then $\mathbf{X}$ is called a BNS-DOU(θ) model.

The process $(X_t, V_t)_{t\geq 0}$ is clearly Markovian.

2.1.2 The Γ-OU model

The Γ-OU model is obtained by constructing the BNS-model with stationary gamma distribution, $V_0 \sim \Gamma(\nu, \alpha)$, where the parameters are $\nu > 0$ and $\alpha > 0$. The corresponding background driving Lévy process Z is a compound Poisson processes with intensity ν and jumps from the exponential distribution with parameter α. Consequently both processes Z and V have a finite number of jumps in any finite time interval.

For the Γ-OU model it is more convenient to work with the parameters ν and α. The connection to the generic parameters used in our general development is given by

$$\zeta = \frac{\nu}{\alpha}, \qquad \eta = \frac{\nu}{\alpha^2}. \tag{2.6}$$

As the gamma distribution admits exponential moments we have integer moments of all orders and our Assumption 1 below is satisfied.

2.1.3 The IG-OU model

The IG-OU model is obtained by construction the BNS-model with stationary inverse Gaussian distribution, $V_0 \sim (\delta, \gamma)$, with parameters $\delta > 0$ and $\gamma > 0$.

The corresponding background driving Lévy process is the sum of an IG$(\delta/2, \gamma)$ process and an independent compound Poisson process with intensity $\delta\gamma/2$ and jumps from an $\Gamma(1/2, \gamma^2/2)$ distribution. Consequently both processes Z and V have infinitely many jumps in any finite time interval.

For the IG-OU model it is more convenient to work with the parameters δ and γ. The connection to the generic parameters used in our general development is given by

$$\zeta = \frac{\delta}{\gamma}, \qquad \eta = \frac{\delta}{\gamma^3}. \tag{2.7}$$

As the inverse Gaussian distribution admits exponential moments we have integer moments of all orders and our Assumption 1 below is satisfied.

2.2 Discrete observations

We observe returns and variance process on a discrete grid of points in time,

$$0 = t_0 < t_1 < \cdots < t_n. \tag{2.8}$$

This implies

$$V(t_i) = V(t_{i-1})e^{-\lambda(t_i - t_{i-1})} + \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s). \tag{2.9}$$

Using

$$V_i := V(t_i), \quad U_i := \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s) \tag{2.10}$$

we have that $(U_i)_{i\geq 1}$ is a sequence of idependent random variables, and it is independent of V_0. If the grid is equidistant, then $(U_i)_{i\geq 1}$ are iid. Observing the returns

X on the grid we have

$$X(t_i) - X(t_{i-1}) = \mu(t_i - t_{i-1}) + \beta(Y(t_i) - Y(t_{i-1}) \\ + \int_{t_{i-1}}^{t_i} \sqrt{V(s-)} dW(s) + \rho(Z_\lambda(t_i) - Z_\lambda(t_{i-1})). \quad (2.11)$$

This suggests introducing the discrete time quantities

$$X_i = X(t_i) - X(t_{i-1}), \quad Y_i = Y(t_i) - Y(t_{i-1}), \quad Z_i = Z_\lambda(t_i) - Z_\lambda(t_{i-1}) \quad (2.12)$$

and

$$W_i = \frac{1}{\sqrt{Y_i}} \int_{t_{i-1}}^{t_i} \sqrt{V(s-)} dW(s). \quad (2.13)$$

Furthermore, it is also convenient to introduce the discrete quantity

$$S_i = \frac{1}{\lambda}(Z_i - U_i). \quad (2.14)$$

It is not difficult to see (conditioning!) that $(W_i)_{i \geq 1}$ is an iid $N(0, 1)$ sequence independent from all other discrete quantities. We note also that $(U_i, Z_i)_{i \geq 1}$ is a bivariate iid sequence, but U_i and Z_i are obviously dependent.

From now on, for notational simplicity, we consider the equidistant grid with

$$t_k = k\Delta, \quad (2.15)$$

where $\Delta > 0$ is fixed. This implies

$$V_i = \gamma V_{i-1} + U_i \quad (2.16)$$

and

$$Y_i = \epsilon V_{i-1} + S_i, \quad (2.17)$$

where

$$\gamma = e^{-\lambda\Delta}, \qquad \epsilon = \frac{1 - \gamma}{\lambda}. \quad (2.18)$$

Furthermore,

$$X_i = \mu\Delta + \beta Y_i + \sqrt{Y_i} W_i + \rho Z_i. \quad (2.19)$$

The sequence $(X_i, V_i)_{i \geq 0}$ is clearly Markovian. From now on we assume all moments of the stationary distribution of V_0 exist.

Assumption 1.

$$E[V_0^n] < \infty \qquad \forall n \in \mathbb{N}. \quad (2.20)$$

In the estimating context we assume all moments are finite with respect to all probability measures $P_\theta, \theta \in \Theta$ under consideration, where Θ is the parameter space.

No other assumtions are made, and all conditions required for consistency and asymptotic normality of our estimator will be proven rigorously from that assumption.

Proposition 1. *We have for all $n \in \mathbb{N}$ that*

$$E[Z_1^n] < \infty, \quad E[U_1^n] < \infty, \quad E[S_1^n] < \infty, \tag{2.21}$$

and

$$E[Y_1^n] < \infty, \quad E[W_1^n] < \infty, \quad E[X_1^n] < \infty. \tag{2.22}$$

Consequently the expectation of any (multivariate) polynomial in Z_1, U_1, S_1, $\sqrt{Y_1}$, W_1, X_1 exists under P_θ.

Proof. We will use repeatedly the well-know relation between the existent of moments and the differentiability of the characteristic function of a random variable, see [CT97, Theorem 8.4.1, p.295f], for example.

Let $\phi(t)$ denote the characteristic function of V_0. By assumption $E_\theta[V_0^n] < \infty$ for all $n \in \mathbb{N}$. Thus $\phi(t)$ is arbitrarily many times differentiable. The law of V_0 is self-decomposable, thus infinitely divisible and $\phi(t) \neq 0$ for all $t \in \mathbb{R}$. Thus the Fourier cumulant function $\kappa(t) = \log \phi(t)$ is arbitrarily many times differentiable. It follows from [BNS01, equation (12)], that the characteristic function of $Z(1)$ is $\psi(t) = \exp(t\kappa'(t))$. Thus $\psi(t)$ is arbitrarily many times differentiable and consequently $E[Z(1)^n] < \infty$, for all $n \in \mathbb{N}$. As Z is a Lévy process this implies $E[Z(\lambda)^n] < \infty$, and as $Z_1 = Z(\lambda)$ we have shown $E[Z_1^n] < \infty$, for all $n \in \mathbb{N}$.

From (2.10) and (2.14) we have $U_1 \leq Z_1$ and $S_1 \leq \lambda^{-1} Z_1$ so $E[U_1^n] < \infty$ and $E[S_1^n] < \infty$ for all $n \in \mathbb{N}$. As W_1 has a standard normal distribution it follows trivially $E[W_1^n] < \infty$ for all $n \in \mathbb{N}$. Repeated application of the binomial resp. multinomial theorem, the Hölder and the Cauchy-Schwarz inequalities yields $E[Y_1^n] < \infty$ and $E[X_1^n] < \infty$ for all $n \in \mathbb{N}$, and the final conclusion for polynomials. □

Let us remark that, by the stationarity, the above result holds also for Z_i, U_i, S_i, $\sqrt{Y_i}$, W_i, X_i instead of Z_1, U_1, S_1, $\sqrt{Y_1}$, W_1, X_1, where $i \in \mathbb{N}$ is arbitrary.

2.3 Some remarks

Most work on estimating function is developed for diffusions, see for example [Sø97, Uch04b, Uch04a, DS04, MR03, KP02, Jac02, Jac01, Sør01, BS01, Kes00, KS99a, Sør97, BS95], although it is often remarked that the results extend to Markov chains. Yet the models under consideration here display several peculiarities.

One assumption that is made usually is that the transition probabilities under P_θ have the same support for each θ. Typically the support of the conditional distribution of V_1 in a BNS model given $V_0 = v$ is $(ve^{-\lambda\Delta}, +\infty)$ under P_θ, thus depends on θ. This does not affect our analysis. The experiment is not homogeneous, cf.[Str85].

If the BDLP is a compound Poisson process, as in the $\Gamma-$OU case, we have the atom of the conditional distribution of V_1 given $V_0 = v$ under P_θ at the parameter dependent position $ve^{-\lambda\Delta}$. Consequently no dominating measure exists and maximum likelihood cannot be *defined* in the usual way. There is an alternative definition covering that case, cf. [KW56, Joh78], but we have not exploited that direction further. See also [NS03]. This problem does not appear with an infinite activity BDLP such as in the IG-OU model and standard maximum likelihood estimation could be studied.

The description given in sections 2.1 and 2.2 provides a BNS model for each θ, but not a statistical experiment as it is taken as a starting point in section 3. The reason is that the processes X and V will depend on θ. This can be avoided by introducing *statistical experiment generated by a BNS model.* In analogy to the statistical experiment generated by a diffusion, see [SS00]. This means we take the distribution of X and V on the Skorohod space $\big(\mathbb{D}^2, \mathcal{B}(\mathbb{D}^2)\big)$ under each P_θ as a starting point.

3 The simple explicit estimator

3.1 The simple estimating equations and their explicit solution

For estimation purposes we consider a probability space on which a parametrized family of probability measures is given:

$$\big(\Omega, \mathcal{F}, \big\{P_\theta : \theta \in \Theta\big\}\big), \tag{3.1}$$

where $\Theta = \{\theta \in \mathbb{R}^6 : \theta^1 > 0, \theta^2 > 0, \theta^3 > 0\}$. The data is generated under the true probability measure P_{θ_0} with some $\theta_0 \in \Theta$. The expectation with respect to P_θ is denoted by $E_\theta[.]$ and with respect to P_{θ_0} simply by $E[.]$.

We assume there is a process $\mathbf{X}$ that is BNS-DOU(θ) under P_θ. We want to find an estimator for θ_0 using observations $X_1, \dots, X_n, V_1, \dots, V_n$. We are interested in asymptotics as $n \to \infty$. To that purpose let us consider the following martingale estimating functions:

$$\begin{aligned}
G_n^1(\theta) &= \textstyle\sum_{k=1}^n \big[V_k - f^1(V_{k-1}, \theta)\big], & f^1(v,\theta) &= E_\theta[V_1 | V_0 = v] \\
G_n^2(\theta) &= \textstyle\sum_{k=1}^n \big[V_k V_{k-1} - f^2(V_{k-1}, \theta)\big], & f^2(v,\theta) &= E_\theta[V_1 V_0 | V_0 = v] \\
G_n^3(\theta) &= \textstyle\sum_{k=1}^n \big[V_k^2 - f^3(V_{k-1}, \theta)\big], & f^3(v,\theta) &= E_\theta[V_1^2 | V_0 = v] \\
G_n^4(\theta) &= \textstyle\sum_{k=1}^n \big[X_k - f^4(V_{k-1}, \theta)\big], & f^4(v,\theta) &= E_\theta[X_1 | V_0 = v] \\
G_n^5(\theta) &= \textstyle\sum_{k=1}^n \big[X_k V_{k-1} - f^5(V_{k-1}, \theta)\big], & f^5(v,\theta) &= E_\theta[X_1 V_0 | V_0 = v] \\
G_n^6(\theta) &= \textstyle\sum_{k=1}^n \big[X_k V_k - f^6(V_{k-1}, \theta)\big], & f^6(v,\theta) &= E_\theta[X_1 V_1 | V_0 = v]
\end{aligned} \tag{3.2}$$

Lemma 1. *We have the explicit expressions*

$$\begin{aligned}
f^1(v,\theta) &= \gamma v + (1-\gamma)\zeta \\
f^2(v,\theta) &= \gamma v^2 + (1-\gamma)\zeta v \\
f^3(v,\theta) &= \gamma^2 v^2 + 2\gamma(1-\gamma)\zeta v + (1-\gamma)^2\zeta^2 + (1-\gamma^2)\eta \\
f^4(v,\theta) &= \beta\epsilon v + \mu\Delta + \beta(1-\epsilon)\zeta + \rho\lambda\zeta \\
f^5(v,\theta) &= \beta\epsilon v^2 + (\mu\Delta + \beta(1-\epsilon)\zeta + \rho\lambda\zeta)v \\
f^6(v,\theta) &= \beta\epsilon\gamma v^2 + ((\mu\Delta + \beta(1-\epsilon)\zeta + \rho\lambda\zeta)\gamma \\
&\qquad + \beta\epsilon(1-\gamma)\zeta)v + (1-\epsilon)(1-\gamma)\zeta^2 + \epsilon^2\lambda\eta
\end{aligned} \tag{3.3}$$

Proof. The formulas are special cases of the general moment calculations given in the appendix. For demonstrating the basic idea we will prove the statements for two special and simple cases here, namely for $f^1(v, \theta)$ and $f^4(v, \theta)$. From (2.16) it follows that

$$E_\theta[V_1 | V_0 = v] = \gamma v + E_\theta[U_1] \tag{3.4}$$

and from the stationarity of V we have

$$E_\theta[U_1] = (1-\gamma)E_\theta(V_0) = (1-\gamma)\zeta. \tag{3.5}$$

Furthermore, from (2.19) and the fact that $E[W_1] = 0$, it follows that

$$E_\theta[X_1 | V_0 = v] = \mu\Delta + \beta E_\theta[Y_1 | V_0 = v] + \rho E_\theta[Z_1 | V_0 = v]. \tag{3.6}$$

But, from (2.17) we have that

$$E_\theta[Y_1 | V_0 = v] = \epsilon v + \tfrac{1}{\lambda} E_\theta[Z_1 - U_1] = \epsilon v + \zeta(1-\epsilon), \tag{3.7}$$

and

$$E_\theta[Z_1] = \lambda\zeta. \tag{3.8}$$

So, from (3.6) it follows that

$$E_\theta[X_1 | V_0 = v] = \mu\Delta + \beta\epsilon v + \beta(1-\epsilon)\zeta + \rho\lambda\zeta. \tag{3.9}$$

□

The estimator $\hat{\theta}_n$ is obtained by solving the estimating equation $G_n(\theta) = 0$ and it turns out that this equation has a simple explicit solution.

Proposition 2. *The estimating equation $G_n(\hat{\theta}_n) = 0$ admits for every $n \geq 2$ on the event*

$$C_n = \left\{\xi_n^2 - \xi_n^1 \upsilon_n^1 > 0, \upsilon_n^2 - (\upsilon_n^1)^2 > 0\right\} \tag{3.10}$$

a unique solution $\hat{\theta}_n = (\lambda_n, \zeta_n, \eta_n, \beta_n, \rho_n, \mu_n)$ that is given by

$$\begin{aligned}
\gamma_n &= (\xi_n^2 - \xi_n^1 \upsilon_n^1)/(\upsilon_n^2 - (\upsilon_n^1)^2); \\
\zeta_n &= (\xi_n^1 - \gamma_n \upsilon_n^1)/(1-\gamma_n); \\
\eta_n &= ((\xi_n^3 - (\xi_n^1)^2) - \gamma_n^2(\upsilon_n^2 - (\upsilon_n^1)^2))/(1-\gamma_n^2); \\
\lambda_n &= -\log(\gamma_n)/\Delta; \\
\epsilon_n &= (1-\gamma_n)/\lambda_n; \\
\beta_n &= (\xi_n^5 - \upsilon_n^1 \xi_n^4)/(\epsilon_n(\upsilon_n^2 - (\upsilon_n^1)^2)); \\
\rho_n &= (\xi_n^6 - \xi_n^4 \xi_n^1 - \beta_n \epsilon_n(\eta_n(1-\gamma_n) + \gamma_n(\upsilon_n^2 - (\upsilon_n^1)^2)))/(2(1-\gamma_n)\eta_n); \\
\mu_n &= (\xi_n^4 - \beta_n \epsilon_n(\upsilon_n^1 - \zeta_n))/\Delta - (\beta_n + \lambda_n \rho_n)\zeta_n;
\end{aligned} \tag{3.11}$$

where

$$\begin{aligned}
\xi_n^1 &= \frac{1}{n}\sum_{i=1}^n V_i, &\quad \xi_n^2 &= \frac{1}{n}\sum_{i=1}^n V_iV_{i-1}, &\quad \xi_n^3 &= \frac{1}{n}\sum_{i=1}^n V_i^2,\\
\xi_n^4 &= \frac{1}{n}\sum_{i=1}^n X_i, &\quad \xi_n^5 &= \frac{1}{n}\sum_{i=1}^n X_iV_{i-1}, &\quad \xi_n^6 &= \frac{1}{n}\sum_{i=1}^n X_iV_i,
\end{aligned} \tag{3.12}$$

and

$$\upsilon_n^1 = \frac{1}{n}\sum_{i=1}^n V_{i-1}, \qquad \upsilon_n^2 = \frac{1}{n}\sum_{i=1}^n V_{i-1}^2. \tag{3.13}$$

Proof. The first three equations $G_n^j(\theta) = 0$, for $j = 1,2,3$ contain only the unknowns ζ, η, λ and are easily solved. In fact we get a familiar estimator for the first two moments and the autocorrelation coefficient of an AR(1) process. The last three equations $G_n^j(\theta) = 0$, for $j = 4,5,6$ can be seen as a linear system for the unknowns μ, β, ρ, once the other parameters have been determined. □

Remark 1. *The exceptional set C_n could be simplified to*

$$C_n' = \left\{\xi_n^2 - \xi_n^1\upsilon_n^1 > 0\right\} \tag{3.14}$$

Since the jump times and the jump size of the BDLP are independent, and the former have an exponential distribution it follows that $V_0, \dots, V_n$ is with probability one not constant, so $P[\upsilon_n^2 - (\upsilon_n^1)^2 > 0] = 1$. But although it can be shown that the probability of C_n tends to zero, for finite n we have $P[\xi_n^2 - \xi_n^1\upsilon_n^1 \le 0] > 0$. This is the common phenomenon that sample moments do not share all properties of their theoretical counterparts. For definiteness we put $\hat{\theta}_n = 0$ outside C_n.

3.2 Consistency

Let us investigate the consistency of the estimator from the previous section. First, we will need the following lemma.

Lemma 2. *For every $k \ge 1$ and $p > 0$*

$$\frac{V_n^k}{n^p} \xrightarrow{a.s.} 0 \qquad as \quad n \to \infty. \tag{3.15}$$

Proof. The random variables $\{V_n,\ n \ge 1\}$ are identically distributed and $m_k = E[|V_1^k|] < \infty$ for all $k \ge 1$. Thus we are in the situation of [Sto74, Exercise 2.1.2(i), p.14].

Let $k \ge 1$ and $\epsilon > 0$ be arbitrarily chosen. Taking any integer $\alpha > 1/p$ and using the Chebyshev inequality we obtain

$$\sum_{n=1}^\infty P\left(\left|\frac{V_n^k}{n^p}\right| > \epsilon\right) \le \sum_{n=1}^\infty \frac{E\big|V_n^k|^\alpha}{n^{\alpha p}\epsilon^\alpha} \le \sum_{n=1}^\infty \frac{m_{k\alpha}}{n^{\alpha p}\epsilon^\alpha} < \infty. \tag{3.16}$$

Therefore from the Borel-Cantelli lemma it follows that $P\big(\limsup_n n^{-p}|V_n^k| > \epsilon\big) = 0$. □

Lemma 3. *We have for all $k \in \mathbb{N}$ that*

$$\frac{1}{n}\sum_{i=1}^{n} V_i^k \xrightarrow{a.s.} E[V_1^k], \tag{3.17}$$

as $n \to \infty$.

Proof. We will prove this statement by induction.

(1) $k = 1$. Let define

$$X_i = V_i - E(V_i|V_{i-1}), \qquad i \geq 1.$$

Obviously, $(X_i,\ i \geq 1)$ is a sequence of martingale differences and is therefore uncorrelated. Using expressions (2.10) and (2.16) we obtain

$$\begin{aligned} E\big[X_i^2\big] &= E(V_i^2) - E\big[E(V_i|V_{i-1})^2\big] \\ &= (1-\gamma^2)E(V_0^2) - 2\gamma E(U_1)E(V_0) + E(U_1^2), \end{aligned}$$

so $E\big[X_i^2\big]$ have a common bound for every $i \geq 1$. Since the assumptions of the theorem 5.1.2 from Chung are satisfied, it follows that

$$\frac{1}{n}\sum_{i=1}^{n} V_i - \frac{1}{n}\sum_{i=1}^{n} E(V_i|V_{i-1}) \xrightarrow{a.s.} 0, \qquad as \quad n \to \infty.$$

But using again the definition (2.16), the last expression is equivalent to

$$\frac{1-\gamma}{n}\sum_{i=1}^{n} V_i + \frac{V_0 - V_n}{n} + E(U_1) \xrightarrow{a.s.} 0, \qquad as \quad n \to \infty.$$

Finally, using the result of the previous lemma, it follows that

$$\frac{1}{n}\sum_{i=1}^{n} V_i \xrightarrow{a.s.} E(V_0), \qquad as \quad n \to \infty.$$

This completes the proof for $k = 1$.

(2) Suppose now that the statement of the theorem holds for $l \leq k-1$, i.e. $E(V_1^{k-1}) < \infty$ and

$$\frac{1}{n}\sum_{i=1}^{n} V_i^l \xrightarrow{a.s.} E(V_0^l), \quad l \leq k-1 \tag{3.18}$$

when $n \to \infty$. For $k > 1$, and for $i \geq 1$, let

$$X_i^k := V_i^k - E\big[V_i^k|V_{i-1}\big] \quad \text{and} \quad S_n^k := \sum_{i=1}^{n} X_i^k. \tag{3.19}$$

Obviously, $(X_i^k,\ i \geq 1)$ is a sequence of martingale differences so it is specially uncorrelated. Moreover, due to the strong stationarity of the volatility sequence and relations (3.19) and (2.16) we obtain

$$\begin{aligned} E\big[X_i^k\big]^2 &= E\big[V_i^{2k} - 2V_i^k E[V_i^k|V_{i-1}] + E[V_i^k|V_{i-1}]^2\big] \\ &= E[V_i^{2k}] - E\big[(E[V_i^k|V_{i-1}])^2\big] \\ &\leq E[V_1^{2k}] =: c_k, \end{aligned} \tag{3.20}$$

c_k denoting some constant that does not depend on i. Hence, by [Chu01, Theorem 5.1.2, p.108], it follows that $S_n^k/n \xrightarrow{a.s.} 0$ when $n \to \infty$, that in our case, due to the definition of S_n^k, is equivalent to

$$\frac{1}{n}\sum_{i=1}^{n} V_i^k - \frac{1}{n}\sum_{i=1}^{n} E\big[V_i^k|V_{i-1}\big] \xrightarrow{a.s.} 0. \tag{3.21}$$

Using again the definition (2.16) and the independency of U_i from V_{i-1}, for $i \geq 1$, we obtain

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} V_i^k - \frac{1}{n}\sum_{i=1}^{n} E\big[V_i^k|V_{i-1}\big] \\
&\quad = \frac{1}{n}\sum_{i=1}^{n} V_i^k - \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{k}\binom{k}{j}\gamma^j V_{i-1}^j E\big[U_1^{k-j}\big] \\
&\quad = \frac{1-\gamma^k}{n}\sum_{i=1}^{n} V_i^k - \frac{\gamma^k}{n}(V_0^k - V_n^k) - \sum_{j=0}^{k-1}\binom{k}{j} E\big[U_1^{k-j}\big]\frac{\gamma^j}{n}\sum_{i=1}^{n} V_{i-1}^j.
\end{aligned}$$

Finally, applying the assumption of the induction, the lemma 2 and the statement (3.21), we obtain

$$\frac{1}{n}\sum_{i=1}^{n} V_i^k \xrightarrow{a.s.} \frac{1}{1-\gamma^k}\sum_{j=0}^{k-1}\binom{k}{j}\gamma^j E\big[U_1^{k-j}\big]E(V_0^j) = E(V_1^k),$$

where the last equality follows calculating $E(\gamma V_0 + U_1)^k$ using (2.16). □

In the next lemma we extend the strong law of large numbers for $(V_i^p, i \geq 1)$ to more general sequences.

Lemma 4. *For all integers $p, q, r \geq 0$ we have*

$$\frac{1}{n}\sum_{i=1}^{n} X_i^p V_i^q V_{i-1}^r \xrightarrow{a.s.} E\big[X_1^p V_1^q V_0^r\big] \tag{3.22}$$

as $n \to \infty$.

Proof. Let

$$M_i = X_i^p V_i^q V_{i-1}^r - E\big[X_i^p V_i^q V_{i-1}^r|V_{i-1}\big].$$

Obviously, $(M_i, i \geq 1)$ is a sequence of martingale differences, and in particular it is uncorrelated. It is stationary and $E[M_1^2] < \infty$. So we can use again [Chu01, Theorem 5.12] to show

$$\frac{1}{n}\sum_{i=1}^{n} X_i^p V_i^q V_{i-1}^r - \frac{1}{n}\sum_{i=1}^{n} E\big[X_i^p V_i^q V_{i-1}^r|V_{i-1}\big] \xrightarrow{a.s.} 0.$$

The conditional expectation $E\big[X_i^p V_i^q V_{i-1}^r|V_{i-1}\big]$ is a polynomial in V_{i-1}, namely

$$E\big[X_i^p V_i^q V_{i-1}^r|V_{i-1}\big] = \sum_{k=0}^{p+q}\phi_{pqk} V_{i-1}^{k+r}.$$

This is shown in the section A.5 in the appendix where the coefficients ϕ_{pqk} are explicitly calculated. Applying Lemma 3 yields

$$\frac{1}{n}\sum_{i=1}^{n} E\big[X_i^p V_i^q V_{i-1}^r | V_{i-1}\big] \xrightarrow{a.s.} \sum_{k=0}^{p+q} \phi_{pqk} E[V_0^{k+r}].$$

As we have

$$E[X_1^p V_1^q V_0^r] = E[E[X_1^p V_1^q V_0^r | V_0]] = \sum_{k=0}^{p+q} \phi_{pqk} E[V_0^{k+r}], \tag{3.23}$$

the proof is completed. □

Theorem 1. *We have $P(C_n) \to 1$ when $n \to \infty$ and the estimator $\hat{\theta}_n$ is consistent on C_n, namely*

$$\hat{\theta}_n \xrightarrow{a.s.} \theta_0$$

on C_n as $n \to \infty$.

Proof. Using the results of lemma 3 it easily follows that

$$\xi_n^2 - \xi_n^1 \upsilon_n^1 \to Cov(V_1, V_0) > 0, \tag{3.24}$$

so $P(C_n) \to 1$ as $n \to \infty$.

Using again the results of lemma 3 it follows that the empirical moments in (3.12) and (3.13) converge to their theoretical counterparts, $\xi_n^i \xrightarrow{a.s.} \xi^i$ and $\upsilon_n^i \xrightarrow{a.s.} \upsilon^i$, where

$$\begin{array}{ll}
\xi^1 = \zeta, & \\
\xi^2 = \zeta^2 + \gamma\eta, & \\
\xi^3 = \zeta^2 + \eta, & \upsilon^1 = \zeta, \\
\xi^4 = \mu + (\beta + \lambda\rho)\zeta, & \upsilon^2 = \zeta^2 + \eta. \\
\xi^5 = \mu\zeta + (\beta + \lambda\rho)\zeta^2 + \beta\epsilon\eta, & \\
\xi^6 = \mu\zeta + (\beta + \lambda\rho)\zeta^2 + (\beta + 2\rho\lambda)\epsilon\eta, &
\end{array} \tag{3.25}$$

Plugging the limits into (3.11) shows, after a short mechanical calculation, that the estimator is in fact consistent. □

3.3 Asymptotic normality

For a concise vector notation we introduce

$$\Xi_k = (V_k, V_k V_{k-1}, V_k^2, X_k, X_k V_{k-1}, X_k V_k)^\top, \tag{3.26}$$

and write the estimating equations in the form

$$G_n^i(\theta) = \sum_{k=1}^{n} \big[\Xi_k^i - f^i(V_{k-1}, \theta)\big], \qquad i = 1, \dots, 6 \tag{3.27}$$

and $f^i(v, \theta)$ given by (3.3). We write

$$f^i(v, \theta) = \sum_{\ell = r_i}^{p_i + r_i + q_i} \phi_\ell^i(\theta) v^\ell. \tag{3.28}$$

with

$$p = (0,0,0,1,1,1), \quad q = (1,1,1,0,0,1), \quad r = (0,1,0,0,1,0) \tag{3.29}$$

and

$$\begin{aligned}
\phi_1^1(\theta) &= \gamma, & \phi_0^1(\theta) &= (1-\gamma)\zeta, \\
\phi_2^2(\theta) &= \gamma, & \phi_1^2(\theta) &= (1-\gamma)\zeta, \\
\phi_2^3(\theta) &= \gamma^2, & \phi_1^3(\theta) &= 2\gamma(1-\gamma)\zeta, \\
& & \phi_0^3(\theta) &= (1-\gamma)^2\zeta^2 + (1-\gamma^2)\eta, \\
\phi_1^4(\theta) &= \beta\epsilon, & \phi_0^4(\theta) &= \mu + \beta(1-\epsilon)\zeta + \rho\lambda\zeta, \\
\phi_2^5(\theta) &= \beta\epsilon, & \phi_1^5(\theta) &= \mu + \beta(1-\epsilon)\zeta + \rho\lambda\zeta, \\
\phi_2^6(\theta) &= \beta\epsilon\gamma, & \phi_1^6(\theta) &= ((\mu + \beta(1-\epsilon)\zeta + \rho\lambda\zeta)\gamma + \beta\epsilon(1-\gamma)\zeta), \\
& & \phi_0^6(\theta) &= (1-\epsilon)(1-\gamma)\zeta^2 + \epsilon^2\lambda\eta.
\end{aligned} \tag{3.30}$$

We will use, that $f^i(v,\theta)$ is a polynomial in v, and that its coefficients ϕ are smooth functions in θ.

We shall first prove the central limit theorem for the estimating functions.

Proposition 3. *We have*

$$\frac{1}{\sqrt{n}} G_n(\theta_0) \xrightarrow{\mathcal{D}} N(0, \Upsilon), \tag{3.31}$$

as $n \to \infty$, where

$$\Upsilon_{ij} = E\big[\operatorname{Cov}(\Xi_1^i, \Xi_1^j | V_0)\big]. \tag{3.32}$$

Proof. To show the above result, we use the multivariate martingale central limit theorem, that is recapitulated in the appendix. To that purpose we introduce the vector martingale difference array

$$\chi_{n,k} = \frac{1}{\sqrt{n}} \big[\Xi_k^i - f^i(V_{k-1}, \theta)\big]. \tag{3.33}$$

We have to show the two assumptions from the previous theorem. First, we prove a multivariate Lyapuonov condition which implies the Lindeberg condition. From (3.33) it follows that $\sqrt{n}\chi_{n,k}^{(j)}$ is of the form $p(V_0, V_1, X_1)$ where $p(v_0, v_1, x_1)$ is a polynomial in v_0, v_1, x_1 which does not depend on n. Thus, $n^2\|\chi_{n,k}\|^4$ has the same property and from the explicit moment expression from the appendix it follows that

$$E\big[\|\chi_{n,k}\|^4 | \mathcal{F}_{k-1}\big] = \frac{1}{n^2} q(V_{k-1}), \tag{3.34}$$

where $q(v_0)$ is a polynomial in v_0. From Lemma 3 it thus follows

$$\frac{1}{n} \sum_{k=1}^{n} q(V_{k-1}) \xrightarrow{a.s.} E[q(V_0)], \tag{3.35}$$

where the expression on the righthand side exists and is finite. Thus the first condition for the martingale central limit theorem is satisfied. For verifying the second condition from the same theorem we consider the $(i,j)-$th element of the matrix $\chi_{n,k}\chi_{n,k}^\top$ which is given by

$$\tfrac{1}{n}\big(\Xi_k^i - f^i(V_{k-1},\theta)\big)\big(\Xi_k^j - f^j(V_{k-1},\theta)\big). \tag{3.36}$$

This is again a polynomial in V_{k-1}, V_k and X_k so by Lemma 4 it follows that

$$\frac{1}{n}\sum_{k=1}^{n}\big(\Xi_k^i - f^i(V_{k-1},\theta)\big)\big(\Xi_k^j - f^j(V_{k-1},\theta)\big) \\ \xrightarrow{a.s.} E\big[\big(\Xi_k^i - f^i(V_{k-1},\theta)\big)\big(\Xi_k^j - f^j(V_{k-1},\theta)\big)\big] \tag{3.37}$$

as $n \to \infty$. □

Remark 2. *A systematic method to evaluate Υ is given in appendix A and the resulting explicit expressions are listed in appendix B.*

Lemma 5. *We have*

$$\frac{1}{\sqrt{n}}\big[\xi_n - \xi\big] \xrightarrow{\mathcal{D}} N(0,\Sigma), \tag{3.38}$$

where

$$\Sigma = P^{-1}\Upsilon(P^{-1})^\top \tag{3.39}$$

and

$$P_{ij} = \delta_{ij} - \phi_1^i\delta_{1j} - \phi_2^i\delta_{3j} \tag{3.40}$$

with δ_{ij} denoting the Kronecker delta.

Proof. We can write

$$\frac{1}{\sqrt{n}}G_n(\theta_0) = P\sqrt{n}(\xi_n - \xi) + Q_n, \tag{3.41}$$

with

$$Q_n^i = \frac{1}{\sqrt{n}}\big[\phi_1^i(V_n - V_0) + \phi_2^i(V_n^2 - V_0^2)\big]. \tag{3.42}$$

In view of lemma 2 above we see, that the remainder term Q_n goes to zero in probability as $n \to \infty$. As P has determinant $(1-\gamma)^2(1+\gamma) > 0$ it is invertible, and we have

$$\sqrt{n}(\xi_n - \xi) = P^{-1}\left(\frac{1}{\sqrt{n}}G_n(\theta_0)\right) + R_n \tag{3.43}$$

with $R_n = -P^{-1}Q_n$ going to zero in probability as $n \to \infty$. The expression $P^{-1}\left(n^{-1/2}G_n(\theta_0)\right)$ is asymptotically normal with mean 0 and covariance matrix Σ. An application of Slutsky's Theorem proves the lemma. □

Finally, we have all the ingredients for proving the following result.

Theorem 2. *The estimator*

$$\hat{\theta}_n = (\lambda_n, \zeta_n, \eta_n, \beta_n, \rho_n, \mu_n) \tag{3.44}$$

is asymptotically normal, namely

$$\sqrt{n}\big[\hat{\theta}_n - \theta_0\big] \xrightarrow{\mathcal{D}} N(0,T), \tag{3.45}$$

as $n \to \infty$, where

$$T = D\Sigma D^T \tag{3.46}$$

and D is given in appendix B.

Proof. We observe from (3.11) that $\hat{\theta}_n = g(\xi_n, \upsilon_n)$, where g is well defined and continuously differentiable in a neighborhood of (ξ, υ). Using the Taylor expansion in the last two variables we have $\hat{\theta}_n = h(\xi_n) + S_n$, where h is well defined and continuously differentiable for in a neighborhood of ξ, and S_n goes to zero in probability in view of lemma 2. Thus it can be neglected according to Slutsky's Theorem. We apply the delta method, see [Leh99] for example, and compute the Jacobian matrix D with

$$D_{ij} = \frac{\partial h_i}{\partial x_j}(\xi), \qquad i, j = 1, \ldots, 6. \tag{3.47}$$

A lengthy elementary calculation shows that the matrix has determinant

$$\frac{\lambda}{2(1-\gamma)^2\gamma\eta^3},$$

thus it is invertible. □

3.4 The simple estimator in a general framework

For comparison and the preparation to the study of optimal estimating functions we would like to review our simple estimator in the general framework of [Sø99]. There the properties of the estimator are studied without exploiting the fact that the estimating equation allows an explicit solution. We extend the theory in the case of a bivariate Markov process. To do so we want to use Sørensen's corollary 2.7 and below for asymptotic normality we want to use his Theorem 2.8. This requires to show that his Condition 2.6 is satisfied which we will do now: For ease of notation let us write the estimating function in the form

$$G_n^j(\theta) = \sum_{i=1}^{n} \alpha^j(V_{i-1};\theta)\big[X_i^{p_j} \cdot V_i^{q_j} - \phi^j\big(V_{i-1};\theta\big)\big], \quad j = 1, \ldots, d \tag{3.48}$$

where

$$\theta \in \mathbb{R}^d, \quad \phi^j(v;\theta) = \sum_{l=0}^{p_j+q_j} \phi_l^j(\theta) \cdot v^l, \quad \alpha^j(v;\theta) = v^{r_j}. \tag{3.49}$$

Let

$$J_n^{j,k}(\theta^{(1)}, \ldots, \theta^{(d)}) = \frac{\partial G_n^j(\theta^{(j)})}{\partial \theta_k}, \qquad k = 1, \ldots, d \tag{3.50}$$

and

$$M_n^\alpha(\bar{\theta}) = \big\{\theta \in \Theta : \|\theta - \bar{\theta}\| \le \frac{\alpha}{\sqrt{n}}\big\}, \qquad \alpha > 0. \tag{3.51}$$

Proposition 4. *The condition* 2.6 *of [Sø99] is satisfied, namely*

(i) the mapping $\theta \longmapsto G_n(\theta)$ is twice continuously differentiable.

(ii) *There exist a* $\bar\theta \in \operatorname{int}\Theta$ *and an invertible non-random* $d \times d$ *matrix* $A(\bar\theta)$ *such that*

$$\sup_{\theta^{(i)} \in M_n^\alpha(\bar\theta)} \left\| \tfrac{1}{n} J_n(\theta^{(1)}, \dots, \theta^{(d)}) - A(\bar\theta) \right\| \to 0$$

in probability as $n \to \infty$ *for all* $\alpha > 0$.

(iii) *There exist* d *non-random* $d \times d$ *matrices* $B^i(\bar\theta)$, $i = 1, \dots, d$, *such that*

$$\sup_{\theta^{(i)} \in M_n^\alpha(\bar\theta)} \left\| \tfrac{1}{n} Q_n^{(i)}(\theta^{(1)}, \dots, \theta^{(d)}) - B^i(\bar\theta) \right\| \to 0$$

in probability as $n \to \infty$ *for all* $\alpha > 0$ *and all* $i = 1, \dots, d$, *where* $Q_n^{(i)}(\theta) = \partial_\theta^2 G_n^i(\theta)$.

(iv) $\left\{ \dfrac{G_n(\bar\theta)}{n} : n \in \mathbb{N} \right\}$ *is stochastically bounded.*

(v) $\sup_{\theta \in M_n^\alpha(\bar\theta)} \left\| \dfrac{G_n(\theta)}{n} \right\| \to 0$ *in probability as* $n \to \infty$ *for all* $\alpha > 0$.

Proof. In our case the number of parameters is $d = 6$. From the definitions above it immediately follows that the mapping

$$\theta \mapsto G_n^j(\theta)$$

is twice continuously differentiable w.r.t. θ.
Let us consider the matrix $J_n = (J_n^{j,k})$, $j, k = 1, \dots, d$ componentwise. For $k = 1, \dots, d$ let

$$\alpha^{j,k}(v;\theta) = \frac{\partial \alpha^j(v;\theta)}{\partial \theta_k}, \qquad \phi^{j,k}(v;\theta) = \frac{\partial \phi^j(v;\theta)}{\partial \theta_k}, \qquad j = 1, \dots, d, \tag{3.52}$$

$$\phi_l^{j,k}(\theta) = \frac{\partial \phi_l^j(\theta)}{\partial \theta_k}, \qquad \phi_l^{j,k,l}(\theta) = \frac{\partial \phi_l^{j,k}(\theta)}{\partial \theta_l}, \qquad j = 1, \dots, d. \tag{3.53}$$

Using the definitions of ϕ^j and $\phi^{j,k}$ we obtain

$$\begin{aligned}
&J_n^{j,k}(\theta) \\
&= \sum_{i=1}^n \Big[\alpha^{j,k}(V_{i-1};\theta)\big[X_i^{p_j} V_i^{q_j} - \phi^j(V_{i-1};\theta)\big] - \alpha^j(V_{i-1};\theta)\phi^{j,k}(V_{i-1};\theta) \Big] \\
&= \sum_{i=1}^n \alpha^{j,k}(V_{i-1};\theta)\big[X_i^{p_j} V_i^{q_j} - \phi^j(V_{i-1};\theta)\big] - \sum_{i=1}^n \alpha^j(V_{i-1};\theta)\phi^{j,k}(V_{i-1};\theta) \\
&= \sum_{i=1}^n \alpha^{j,k}(V_{i-1};\theta)\big[X_i^{p_j} V_i^{q_j} - \phi^j(V_{i-1};\theta)\big] - \sum_{i=1}^n \alpha^j(V_{i-1};\theta) \sum_{l=0}^{p_j+q_j} \phi_l^{j,k}(\theta) V_{i-1}^l.
\end{aligned} \tag{3.54}$$

Let us define

$$A^{j,k}(\bar\theta) = -\sum_{l=0}^{p_j+q_j} \phi_l^{j,k}(\bar\theta) E\big[\alpha^j(V_0;\bar\theta) V_0^l\big]. \tag{3.55}$$

Using the definitions of ϕ^j the just obtained expression for $J_n^{j,k}$ and (3.55) we obtain

$$\begin{aligned}
&\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}J_n^{j,k}(\theta^{(p)})-A^{j,k}(\bar\theta)\right| \\
&\quad=\sup_{\theta^{(p)}\in M_n(\bar\theta)}\Bigg|\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\theta^{(p)})\big[X_i^{p_j}V_i^{q_j}-\phi^j(V_{i-1};\theta^{(p)})\big] \\
&\qquad\qquad-\frac{1}{n}\sum_{i=1}^{n}\alpha^j(V_{i-1};\theta^{(p)})\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\theta^{(p)})V_{i-1}^l \\
&\qquad\qquad+\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\bar\theta)E\big[\alpha^j(V_0;\bar\theta)V_0^l\big]\Bigg| \\
&\quad\leq\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\theta^{(p)})\big[X_i^{p_j}V_i^{q_j}-\phi^j(V_{i-1};\theta^{(p)})\big]\right| \\
&\qquad+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\Bigg|-\frac{1}{n}\sum_{i=1}^{n}\alpha^j(V_{i-1};\theta^{(p)})\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\theta^{(p)})V_{i-1}^l \\
&\qquad\qquad+\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\bar\theta)E\big[\alpha^j(V_0;\bar\theta)V_0^l\big]\Bigg|.
\end{aligned}\tag{3.56}$$

Adding and subtracting expressions

$$\frac{1}{n}\sum_{i=1}^{n}\alpha^j(V_{i-1};\bar\theta)\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\theta^{(p)})V_{i-1}^l \quad\text{and}\quad \frac{1}{n}\sum_{i=1}^{n}\alpha^j(V_{i-1};\bar\theta)\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\bar\theta)V_{i-1}^l$$

to (3.56) it follows that

$$\begin{aligned}
&\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}J_n^{j,k}(\theta^{(p)})-A^{j,k}(\bar\theta)\right| \\
&\quad\leq\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\theta^{(p)})\big[X_i^{p_j}V_i^{q_j}-\phi^j(V_{i-1};\theta^{(p)})\big]\right| \\
&\qquad+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}\sum_{i=1}^{n}\big[\alpha^j(V_{i-1};\bar\theta)-\alpha^j(V_{i-1};\theta^{(p)})\big]\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\theta^{(p)})V_{i-1}^l\right| \\
&\qquad+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\Bigg|\sum_{l=0}^{p_j+q_j}\big[\phi_l^{j,k}(\bar\theta)-\phi_l^{j,k}(\theta^{(p)})\big]\frac{1}{n}\sum_{i=1}^{n}\alpha^j(V_{i-1};\bar\theta)V_{i-1}^l \\
&\qquad\qquad+\sum_{l=0}^{p_j+q_j}\phi_l^{j,k}(\bar\theta)\left[E[\alpha^j(V_0;\bar\theta)V_0^l]-\frac{1}{n}\sum_{i=1}^{n}\alpha^j(V_{i-1};\bar\theta)V_{i-1}^l\right]\Bigg|.
\end{aligned}\tag{3.57}$$

For ease of notation let

$$K_n^{j,k}(\theta^{(p)})=\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\theta^{(p)})\big[X_i^{p_j}V_i^{q_j}-\phi^j(V_{i-1};\theta^{(p)})\big].\tag{3.58}$$

Since a function belonging to $\mathcal{C}^\infty$ is bounded on a compact set, using a definition of the set M_n it follows that

$$\begin{aligned}
&\sup_{\theta^{(p)}\in M_n(\bar\theta)} \left| \sum_{l=0}^{p_j+q_j} \left[\phi_l^{j,k}(\bar\theta) - \phi_l^{j,k}(\theta^{(p)})\right] \right| \\
&\quad \le \sum_{l=0}^{p_j+q_j} \sup_{\theta^{(p)}\in M_n(\bar\theta)} \sum_{m=1}^{d} \sup_{\theta^*\in M_n^\alpha(\bar\theta)} \left| \frac{\partial \phi_l^{j,k}}{\partial \theta_m^{(p)}}(\theta^*) \right| \left| \bar\theta_m - \theta_m^{(p)} \right| \\
&\quad \le \sum_{l=0}^{p_j+q_j} \sum_{m=1}^{d} M \frac{\alpha}{\sqrt{n}} =: \frac{C_{1,j}}{\sqrt{n}}, \qquad (3.59)
\end{aligned}$$

where $\sup_{\theta^*\in M_n^\alpha(\bar\theta)} \left| \frac{\partial \phi_l^{j,k}}{\partial \theta_m^{(p)}}(\theta^*) \right| \le M$ and $C_{1,j} = M\alpha(p_j+q_j)d$.

Since $\alpha(v,\theta)=1$ or $\alpha(v,\theta)=v$, using definition (3.58), from the relation (3.57) it follows that

$$\begin{aligned}
&\sup_{\theta^{(p)}\in M_n(\bar\theta)} \left| \frac{1}{n} J_n^{j,k}(\theta^{(p)}) - A^{j,k}(\bar\theta) \right| \\
&\quad \le \sup_{\theta^{(p)}\in M_n(\bar\theta)} \left| K_n^{j,k}(\theta^{(p)}) \right| + \frac{\alpha}{\sqrt{n}} \left[\sup_{\theta^{(p)}\in M_n(\bar\theta)} \left| \sum_{l=0}^{p_j+q_j} \phi_l^{j,k}(\theta^{(p)}) \right| \right] \left| \frac{1}{n} \sum_{i=1}^{n} V_{i-1}^l \right| \\
&\quad + \frac{C_{1,j}}{\sqrt{n}} \left| \frac{1}{n} \sum_{i=1}^{n} \alpha^j(V_{i-1};\bar\theta) V_{i-1}^l \right| \\
&\quad + \sum_{l=0}^{p_j+q_j} \left| \phi_l^{j,k}(\bar\theta) \left[\frac{1}{n} \sum_{i=1}^{n} \alpha^j(V_{i-1};\bar\theta) V_{i-1}^l - E[\alpha^j(V_0;\bar\theta)V_0^l] \right] \right|. \qquad (3.60)
\end{aligned}$$

Adding and subtracting expressions

$$\frac{1}{n} \sum_{i=1}^{n} \alpha^{j,k}(V_{i-1};\bar\theta) \left[X_i^{p_j} V_i^{q_j} - \phi^j(V_{i-1};\theta^{(p)}) \right]$$

and

$$\frac{1}{n} \sum_{i=1}^{n} \left[\alpha^{j,k}(V_{i-1};\theta^{(p)}) - \alpha^{j,k}(V_{i-1};\bar\theta) \right] \phi^j(V_{i-1};\bar\theta)$$

to $K_n^{j,k}(\theta^{(p)})$ in relation (3.60) and since functions $\alpha^{j,k}(v,\theta)=0$ or 1, we obtain

$$
\begin{aligned}
&\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|K_n^{j,k}(\theta^{(p)})\right| \\
&\quad\le \sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}\sum_{i=1}^{n}\left[\alpha^{j,k}(V_{i-1};\theta^{(p)})-\alpha^{j,k}(V_{i-1};\bar\theta)\right]\left[X_i^{p_j}V_i^{q_j}-\phi^j(V_{i-1};\bar\theta)\right]\right| \\
&\qquad+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}\sum_{i=1}^{n}\left[\alpha^{j,k}(V_{i-1};\theta^{(p)})-\alpha^{j,k}(V_{i-1};\bar\theta)\right]\left[\phi^j(V_{i-1};\bar\theta)-\phi^j(V_{i-1};\theta^{(p)})\right]\right| \\
&\qquad+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\bar\theta)\left[\phi^j(V_{i-1};\bar\theta)-\phi^j(V_{i-1};\theta^{(p)})\right]\right| \\
&\qquad+\left|\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\bar\theta)\left[X_i^{p_j}V_i^{q_j}-\phi^j(\bar\theta)\right]\right| \\
&\quad\le \frac{\alpha}{\sqrt{n}}\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\left[X_i^{p_j}V_i^{q_j}-\phi^j(V_{i-1};\bar\theta)\right]\right|\right. \\
&\qquad\qquad\left.+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\frac{1}{n}\sum_{i=1}^{n}\left[\phi^j(V_{i-1};\bar\theta)-\phi^j(V_{i-1};\theta^{(p)})\right]\right|\right\} \\
&\qquad+\sup_{\theta^{(p)}\in M_n(\bar\theta)}\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\bar\theta)\left[\phi^j(V_{i-1};\bar\theta)-\phi^j(V_{i-1};\theta^{(p)})\right] \\
&\qquad+\left|\frac{1}{n}\sum_{i=1}^{n}\alpha^{j,k}(V_{i-1};\bar\theta)\left[X_i^{p_j}V_i^{q_j}-\phi^j(\bar\theta)\right]\right|. \qquad (3.61)
\end{aligned}
$$

But again, for every $i=1,\dots n$, we have

$$
\begin{aligned}
\left|\phi^j(V_{i-1};\bar\theta)-\phi^j(V_{i-1};\theta^{(p)})\right| &\le \sum_{l=0}^{p_j+q_j}\left|\phi_l^j(\bar\theta)-\phi_l^j(\theta^{(p)})\right|V_{i-i}^l \\
&\le \sum_{l=0}^{p_j+q_j}\sum_{m=1}^{d}\sup_{\theta^*\in M_n^\alpha(\bar\theta)}\left|\frac{\partial\phi_l^{j,k}}{\partial\theta_m^{(p)}}(\theta^*)\right|\cdot\left|\theta_m^{(p)}-\bar\theta_m\right|V_{i-1}^l \\
&\le \frac{C_{1,j}}{\sqrt{n}}V_{i-1}^l, \qquad (3.62)
\end{aligned}
$$

so from (3.61) using lemma 4 it follows that

$$\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|K_n^{j,k}(\theta^{(p)})\right| \xrightarrow{a.s.} 0 \qquad (3.63)$$

when $n\to\infty$. Applying the just obtained result into the expression (3.60) and using again lemma 4, it follows that

$$\sup_{\theta^{(p)}\in M_n(\bar\theta)}\left|\tfrac{1}{n}J_n^{j,k}(\theta^{(p)})-A^{j,k}(\bar\theta)\right| \xrightarrow{a.s.} 0 \qquad (3.64)$$

as $n\to\infty$. This proves the part (ii). Let

$$Q_n^{j,k,l}(\theta)=\frac{\partial J_n^{j,k}(\theta)}{\partial\theta_l}, \quad l=1,\dots,d. \qquad (3.65)$$

From (3.65) it follows that we consider one higher order derivative of the function G_n than in the matrix J_n. Let us denote

$$\alpha^{j,k,l}(v;\theta) = \frac{\partial \alpha^{j,k}(v;\theta)}{\partial \theta_l}, \qquad l = 1, \dots, d. \tag{3.66}$$

From (3.54) we have

$$\begin{aligned} Q_n^{j,k,l}(\theta) = & \sum_{i=1}^{n} \alpha^{j,k,l}(V_{i-1};\theta)\big[X_i^{p_j} V_i^{q_j} - \phi^j(V_{i-1};\theta)\big] - \sum_{i=1}^{n} \alpha^{j,k}(V_{i-1};\theta)\phi^{j,l}(V_{i-1};\theta) \\ & - \sum_{i=1}^{n} \alpha^{j,l}(V_{i-1};\theta) \sum_{m=0}^{p_j+q_j} \phi_m^{j,k}(\theta) V_{i-1}^m - \sum_{i=1}^{n} \alpha^{j}(V_{i-1};\theta) \sum_{m=0}^{p_j+q_j} \phi_m^{j,k,l}(\theta) V_{i-1}^m. \end{aligned} \tag{3.67}$$

Let us define

$$\begin{aligned} B^{j,k,l}(\theta) = - \sum_{m=0}^{p_j+q_j} & \big\{ \phi_m^{j,l}(\theta) E\big[\alpha^{j,k}(V_0;\theta) V_0^m\big] \\ & + \phi_m^{j,k} E\big[\alpha^{j,l}(V_0;\theta) V_0^m + \phi_m^{j,k,l}(\theta) E\big[\alpha^{j}(V_0;\theta) V_0^m\big]\big\}. \end{aligned} \tag{3.68}$$

Thus, the part (iii) can be proved following the lines of the proof of part (ii) considering $\phi_m^{j,k}, \phi_m^{j,k,l}$ instead of ϕ_m^j and $\phi_m^{j,k}$, respectively and taking in consideration that $\alpha^{j,k}(v,\theta) = 0$ and $\alpha^{j,k,l}(v,\theta) = 0$ since $\alpha^j(v,\theta) = 1$ or v.

For arbitrary $\epsilon > 0$ and $K_\epsilon > 0$ using the Chebyshev inequality and the stationarity of the volatility process we have

$$\begin{aligned} P\left[\left|\frac{G_n^j(\bar\theta)}{\sqrt{n}}\right| > K_\epsilon\right] & \le \frac{E\big[|G_n^j(\bar\theta)|^2\big]}{nK_\epsilon^2} \\ & = \frac{\sum_{i=1}^{n} E\big[\alpha^j(V_{i-1};\theta)^2 \cdot (X_i^{p_j} V_i^{q_j} - \phi^j(V_{i-1};\theta))^2\big]}{nK_\epsilon^2} \\ & = \frac{E\big[\alpha^j(V_0;\theta)^2 \big\{E[X_1^{2p_j} V_1^{2q_j} | V_0] - \phi^j(V_0;\theta)^2\big\}\big]}{K_\epsilon^2} \\ & = \frac{E\big[\alpha^j(V_0;\theta)^2 \big(X_1^{2p_j} V_1^{2q_j} - \phi^j(V_0;\theta)^2\big)\big]}{K_\epsilon^2} \\ & \le \frac{E\big[\alpha^j(V_0;\theta)^2 X_1^{2p_j} V_1^{2q_j}\big]}{K_\epsilon^2}. \end{aligned} \tag{3.69}$$

Taking

$$K_\epsilon^2 := \frac{E\big[\alpha^j(V_0;\theta)^2 X_1^{2p_j} V_1^{2q_j}\big]}{\epsilon}$$

it follows that for every $\epsilon > 0$ exists a $K_\epsilon > 0$ such that

$$\sup_{n\in\mathbb{N}} P\left[\left|\frac{G_n^j(\bar\theta)}{\sqrt{n}}\right| > K_\epsilon\right] < \epsilon,$$

that proves the part (iv). Finally, using the just proven result, the definition of $M_n^\alpha(\bar\theta)$ and (3.64), componentwise we have

$$\begin{aligned}\sup_{\theta\in\ M_n^\alpha(\bar\theta)}\left|\tfrac{1}{n}G_n^j(\theta)\right| &= \sup_{\theta\in\ M_n^\alpha(\bar\theta)}\left|\tfrac{1}{n}G_n^j(\bar\theta)+\tfrac{1}{n}J_n^{j,k}(\theta^*)(\theta-\bar\theta)\right|\\ &\le \left|\tfrac{1}{n}G_n^j(\bar\theta)\right| + \sup_{\theta^*\in M_n^\alpha(\bar\theta)}\left|\tfrac{1}{n}J_n^{j,k}(\theta^*)\right| \sup_{\theta\in M_n^\alpha(\bar\theta)}|\theta-\bar\theta|\\ &\le \frac{1}{\sqrt{n}}K_\epsilon + \sup_{\theta^*\in M_n^\alpha(\bar\theta)}\left|\tfrac{1}{n}J_n^{j,k}(\theta^*)\right|\frac{\alpha}{\sqrt{n}} \longrightarrow 0\end{aligned}$$

in probability as $n\to\infty$. □

From our proposition 4 and corollary 2.7 from [Sø99] consistency of our estimator follows.

We have shown the asymptotic normality of the estimating function in proposition 3. By Theorem 2.8 from [Sø99] and our Proposition 4 it follows that our estimator is asymptotically normal. The covariance matrix equals $T = A(\bar\theta)^{-1}\Upsilon(A(\bar\theta)^{-1})^\top$. In particular, $A(\bar\theta) = -PD^{-1}$ which is a consequence of the implicit function theorem.

4 Numerical illustrations

4.1 Description of the model and its parameter values

To illustrate the results from the previous sections numerically, we consider the Γ-OU model from Section 2.1.2, where the variance V has a stationary gamma distribution. We use as time unit one year consisting of 250 trading days. The true parameters are

$$\nu = 2.56,\quad \alpha = 64,\quad \lambda = 256,\quad \beta = -0.5,\quad \rho = -0.1,\quad \mu = 1.2. \tag{4.1}$$

The parameters imply that there are 2.6 jumps per day on the average, and the jumps in the BDLP and in the volatility are exponentially distributed with mean 0.0156. The interpretation is, that typically every day two or three new pieces of information arrive and make the variance process jump. The stationary mean of the variance is 0.04, hence if we define instantaneous volatility to be the square root of the variance, it will fluctuate around 20% in our example. The half-life of the autocorrelation of returns is about half a day.

In our example annual log returns have (unconditional) mean 25.6% and a annual volatility 20%. Figure 1 displays a simulation of one year of daily observations from the background driving Lévy process, from the instantaneous variance process, and log returns, or more precisely, simulated realizations of Z_i, V_i, and X_i for $i = 1,\dots,250$. In [Pos06] other scenarios are considered, for example, small jumps arriving every minute, with fast decaying autocorrelation, or few jumps per year, corresponding to exceptional new, with heavy impact on the variance process.

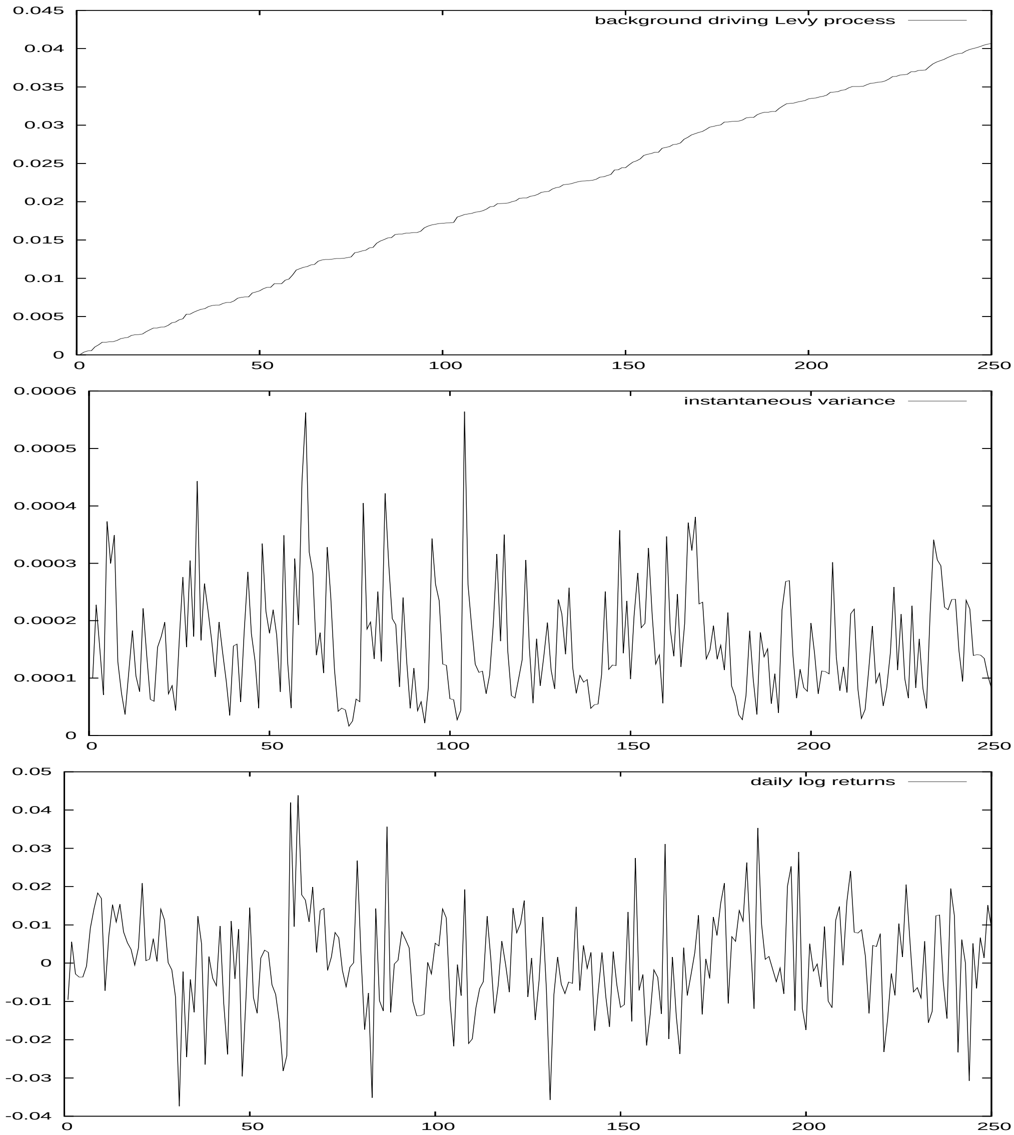

Figure 1: Daily observations Z_i, V_i, X_i.

4.2 The asymptotic covariance matrix of the estimator

As our goal is an analysis of the estimator, and not an empirical study, we do not estimate the asymptotic covariance, but evaluate the explicit expression using the true parameters. Denoting the vector of asymptotic standard deviations of the estimates and the correlation matrix by $s/\sqrt{n}$ resp. r we have

$$s = \begin{bmatrix} 4.86 \\ 125 \\ 650 \\ 7.36 \\ 253 \\ 0.526 \end{bmatrix}, \qquad r = \begin{bmatrix} 1 & 0.89 & 0.41 & 0.03 & 0.09 & -0.02 \\ 0.89 & 1 & 0.4 & 0.03 & 0.09 & -0.03 \\ 0.41 & 0.4 & 1 & 0.06 & 0.22 & 0 \\ 0.03 & 0.03 & 0.06 & 1 & -0.75 & 0.06 \\ 0.09 & 0.09 & 0.22 & -0.75 & 1 & -0.57 \\ -0.02 & -0.03 & 0 & 0.06 & -0.57 & 1 \end{bmatrix} \tag{4.2}$$

4.3 Distribution of the estimates

Figure 4.3 illustrates the empirical and asymptotic distribution of the simple estimators for the Γ-OU model. The histograms are produced from $m = 10000$ replications consiting of $n = 8000$ observations each, corresponding to 32 years with 250 daily obervations per year.

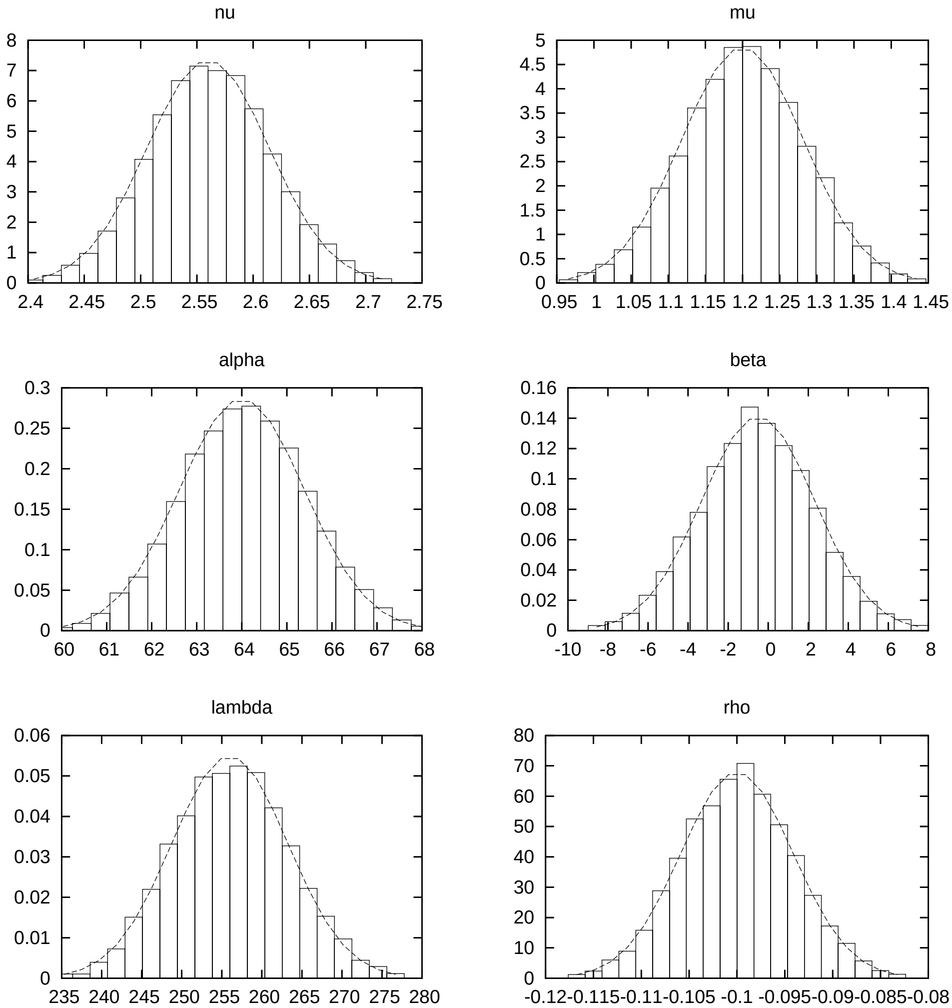

Figure 2: Empirical and asymptotic distribution of the simple estimators for the Γ-OU model. The histograms are produced from $m = 10000$ replications consiting of $n = 8000$ observations each, corresponding to 32 years with 250 daily obervations per year. The true values are $\nu = 2.56$, $\alpha = 64$, $\lambda = 256$, $\mu = 1.2$, $\beta = -0.5$, $\rho = -0.1$. The standard deviations used for the normal curves are taken from the explicit asymptotic results, not estimated.

We see from the graphs that in our illustration the parameters ν, α, λ, and μ can be estimated quite accurately, in the sense that the usual confidence intervals yield one or two significant digits at least. The estimate for ρ is not as accurate and the accuracy for the estimate for β is unsatisfactory.

The bad quality of the estimator for β is neither surprising nor very troublesome. It has little impact on the model. The main reason for including the parameter β in the specification of BNS models is, for derviatives pring: A risk-neutral BNS-model must have $\beta = -1/2$. In most applications working under a physical probability measure $\beta = 0$ can be assumed without much loss of generality or flexibility.

In ongoing work [HP06] we compare this asymptotic covariance with the covariance of the optimal quadratic estimating function.

5 Further and alternative developments

5.1 Optimal quadratic estimating functions

Our choice of estimating functions is natural, but, mathematically speaking, somewhat arbitrary. In ongoing work [HP06] we show, that the *optimal quadratic estimating function* based on the moments of $V_1, X_1, V_1^2, V_1X_1, X_1^2$ can be computed explicitly, though the corresponding estimator has to be determined numerically. Our simple estimator can be used as a starting point for an iterative root-finding procedure. Consistency and asymptotic normality can be shown using the general theory as presented in [Sø99] along the lines of the present paper, although the expressions involved are slightly more complicated.

5.2 Using more integer or trigonometric moments for better efficiency

More efficient estimators than provided by the optimal quadratic estimating function can be obtained by incorporating further moments. As we have provided explicit computations for arbitrary integer moments and conditional moments, our methods can be extended to that situation. We might even have the number of moments tend to infinity with the number of observations, and obtain an estimator that is asymptotically equivalent to the maximum likelihood estimator, when the latter exists resp. can be defined, see 2.3. The reader might object, that very high moments are not reliable for empirical investigations. BNS-models allow also explicit computation of the characteristic function and thus of conditional and unconditional *trigonometric moments* $E[e^{i(\xi_k V_1 + \psi_k X_1)}]$ and $E[e^{i(\xi_k V_1 + \psi_k X_1)}|V_0]$ for arbitrary constants ξ_k and ψ_k, that could be used instead to construct estimating functions. See [AS02] for diffusions, [Sch05] for Lévy type processes, and [Sin01] for affine models.

5.3 Intra-day observations

Our approach is based on the explicit calculation of conditional and unconditional moments. Those calculations can be done for BNS-models on arbitrary time intervals. Hence our analysis is not restricted to a fixed time grid with the number of observation intervals tending to infinity, but could be performed also on a fixed horizon, with the number of intra-day observations increasing to infinity. The resulting estimators should then be compared to power-variation methods, cf. [Tod06].

5.4 Comparison to the generalized method of moments

We would be interested in a comparison of our results to the related *generalized methods of moments*. For a rigorous treatment of the latter, a precise specification of the weighting matrix is required, see [HHY96] and the references therein.

5.5 Unobserved volatility and substitutes for volatility

Finally, perhaps the biggest issue is, that the instantaneous variance is not observed in discrete time. In [Lin05] it is reported, that the number of trades is an excellent substitute for statistical purposes. This is certainly a promising starting point for an empirical analysis. For a theoretical analysis a joint model for the number prices and number of trades has to be specified.

Another direction would be, to adapt the implied state method (IS-GMM) as introduced in [Pan02] to our martingale estimating function approach: We replace the unobserved V_i in the estimating equations by the model-implied variance $V_i(\theta)$ that is obtained from option prices, assuming that the dynamics are governed by BNS-models both under the physical probability measure P_{θ_0} and a risk-neutral measure $P_{\tilde{\theta}_0}$. The resulting estimating function will not be a martingale estimating function any more, and the bias has to be accounted for in a rigorous analysis. Nevertheless, in view of the results of [Pan02], we are optimistic, that consistency and asymptotic normality will hold also here.

A Explicit moment calculations

This section is about computing explicitly $E[X_1^n V_1^m | V_0 = 0]$ and $E[X_1^n V_1^m]$. All moments below will be given in terms of the cumulants of the stationary distribution, denoted by K_n. We set

$$\zeta = K_1, \qquad \eta = K_2. \tag{A.1}$$

If the stationary distribution is determined by the two parameters ζ and η the higher cumulants are obviously functions of ζ and η, but the formulae hold in more general cases.

The calculations exploit the analytical tractability of the BNS-model, namely conditional Gaussianity of the logarithmic returns X and the linear structure of the OU-type process V. From that it follows, and it is well-known, that univariate and multivariate cumulants can be computed easily. It remains to transform multivariate cumulants to multivariate moments, again a topic that is well-understood, and explicit expressions involve the multivariate Faa di Bruno formula, multivariate Bell polynomials and integer partitions, see for example [McC87].

We have chosen to use simple recursions, that are easy to implement on a computer algebra system, in particular, since the expression, though completely explicit and elementary, are rather lengthy when it comes to evaluating moments of order four for the asymptotic covariance matrix. For the readers convenience, we give the details in this appendix.

A.1 Preliminaries

Let us recaptiulate the variables and notation from section 2.2, that is required in the following calculations. We use

$$\gamma = e^{-\lambda\Delta}, \qquad \epsilon = \frac{1 - e^{-\lambda\Delta}}{\lambda}. \tag{A.2}$$

We have

$$V_1 = \gamma V_0 + U_1, \qquad Y_1 = \epsilon V_0 + S_1 \tag{A.3}$$

where

$$U_1 = \int_0^\Delta e^{-\lambda(\Delta - s)} dZ_{\lambda s}, \qquad S_1 = \int_0^\Delta \lambda^{-1}(1 - e^{-\lambda(\Delta - s)}) dZ_{\lambda s}. \tag{A.4}$$

Note, that we have the simpler formula $S_1 = (Z_1 - U_1)/\lambda$, but the integral above is sometimes notationally more convenient. We have

$$X_1 = A_1 + \sqrt{Y_1} W_1, \qquad A_1 = \mu\Delta + \beta Y_1 + \rho Z_1. \tag{A.5}$$

A.2 Stationary moments

We use the well-known recursion to compute moments from cumulants

$$E[V_0^n] = \delta_{n0} + \sum_{i=0}^{n-1} \binom{n-1}{i} K_{i+1} E[V_0^{n-1-i}]. \tag{A.6}$$

Alternatively we have $E[V_0^n] = Y_n(K_1, \ldots, K_n)$, where $Y_n(x_1, \ldots, x_n)$ denotes the complete Bell polynomials. Explicit non-recursive expressions can be given, but we do not use them.

A.3 Trivariate cumulants

From the key formula for Wiener-type integrals with Lévy process integrator, it follows that the joint cumulants of (S_1, U_1, Z_1) are given by

$$K_{nm\ell} = \lambda \epsilon_{nm} (n + m + \ell) K_{n+m+\ell}, \tag{A.7}$$

with

$$\epsilon_{ij} = \begin{cases} \lambda^{-i}\left(1 + \sum_{k=1}^{i} \binom{i}{k} (-1)^k \frac{1-\gamma^k}{k\lambda}\right) & j = 0 \\ \lambda^{-i}\left(\frac{1-\gamma^j}{j\lambda} + \sum_{k=1}^{i} \binom{i}{k} (-1)^k \frac{1-\gamma^k}{k\lambda}\right) & j > 0 \end{cases} \tag{A.8}$$

A.4 Trivariate Moments

Trivariate moments can be computed recursively from trivariate cumulants

$$E[S_1^n U_1^m Z_1^\ell] = \sum_{i=0}^{n-1} \sum_{j=0}^{m} \sum_{k=0}^{\ell} \binom{n-1}{i} \binom{m}{j} \binom{\ell}{k} K_{i+1,j,k} E[S_1^{n-1-i} U_1^{m-j} Z_1^{\ell-k}] \quad (A.9)$$

$$E[S_1^n U_1^m Z_1^\ell] = \sum_{i=0}^{n} \sum_{j=0}^{m-1} \sum_{k=0}^{\ell} \binom{n}{i} \binom{m-1}{j} \binom{\ell}{k} K_{i,j+1,k} E[S_1^{n-i} U_1^{m-1-j} Z_1^{\ell-k}] \quad (A.10)$$

$$E[S_1^n U_1^m Z_1^\ell] = \sum_{i=0}^{n} \sum_{j=0}^{m} \sum_{k=0}^{\ell-1} \binom{n}{i} \binom{m}{j} \binom{\ell-1}{k} K_{i,j,k+1} E[S_1^{n-i} U_1^{m-j} Z_1^{\ell-1-k}] \quad (A.11)$$

Alternatively, we can express $E[S_1^n U_1^m Z_1^\ell]$ as trivariate complete Bell polynomials $Y_{nm\ell}$ evaluated at the trivariate cumulants of S_1, U_1, Z_1, and explicit non-recursive expressions are available, but not very useful for us.

A.5 Some conditional expectations

Using (A.3) gives

$$E[Y_1^n V_1^m Z_1^\ell | V_0 = v] = \sum_{i=0}^{n} \sum_{j=0}^{m} \binom{n}{i} \binom{m}{j} \epsilon^i \gamma^j E[S_1^{n-i} U_1^{m-j} Z_1^\ell] \cdot v^{i+j} \quad (A.12)$$

Collection powers of v gives

$$E[Y_1^n V_1^m Z_1^\ell | V_0 = v] = \sum_{k=0}^{n+m} \xi_{nm\ell k} v^k \quad (A.13)$$

with

$$\xi_{nm\ell k} = \sum_{j=0}^{m \wedge k} \binom{n}{k-j} \binom{m}{j} \epsilon^{k-j} \gamma^j E[S_1^{n-k+j} U_1^{m-j} Z_1^\ell] \quad (A.14)$$

Then using (A.5) and conditioning gives

$$E[A_1^n Y_1^m V_1^\ell | V_0 = v] = \sum_{i=0}^{n} \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} \beta^i \rho^j \mu^{n-i-j} E[Y_1^{m+i} V_1^\ell Z_1^j | V_0 = v] \quad (A.15)$$

Collecting powers of v gives

$$E[A_1^n Y_1^m V_1^\ell | V_0 = v] = \sum_{k=0}^{n+m+\ell} \psi_{nm\ell k} v^k \quad (A.16)$$

with

$$\psi_{nm\ell k} = \sum_{i=(k-m-l)_+}^{n} \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} \beta^i \rho^j \mu^{n-i-j} \xi_{m+i,\ell,j,k} \quad (A.17)$$

Finally using (A.5) and the Gaussian moments gives

$$E[X_1^n V_1^m | V_0 = v] = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} \frac{(2i)!}{2^i i!} E[A_1^{n-2i} Y_1^i V_1^m | V_0 = v] \tag{A.18}$$

Collecting powers of v gives

$$E[X_1^n V_1^m | V_0 = v] = \sum_{k=0}^{n+m} \phi_{nmk} v^k \tag{A.19}$$

with

$$\phi_{nmk} = \sum_{i=0}^{(n+m-k)\wedge \lfloor \frac{n}{2} \rfloor} \binom{n}{2i} \frac{(2i)!}{2^i i!} \psi_{n-2i,i,m,k} \tag{A.20}$$

It follows from the calculations above that ϕ_{nmk} are polynomials in $\gamma, \epsilon, \mu, \beta, \rho$.

A.6 Some unconditional expectations

The same structure pertains for the unconditional expectations,

$$E[Y_1^n V_1^m Z_1^\ell] = \sum_{i=0}^{n} \sum_{j=0}^{m} \binom{n}{i} \binom{m}{j} \epsilon^i \gamma^j E[S_1^{n-i} U_1^{m-j} Z_1^\ell] E[V_0^{i+j}] \tag{A.21}$$

then

$$E[A_1^n Y_1^m V_1^\ell] = \sum_{i=0}^{n} \sum_{j=0}^{n-i} \binom{n}{i} \binom{n-i}{j} \beta^i \rho^j \mu^{n-i-j} E[Y_1^{m+i} V_1^\ell Z_1^j] \tag{A.22}$$

and finally

$$E[X_1^n V_1^m] = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} \frac{(2i)!}{2^i i!} E[A_1^{n-2i} Y_1^i V_1^m]. \tag{A.23}$$

B Tables with explicit expressions

B.1 The matrix Υ

Let κ_3 and κ_4 denote the third and the fourth cumulant of the stationary distribution of V_0. We recall that the first and second cumulants are the parameters ζ and η. If the distribution is determined by those parameters then κ_3 and κ_4 will be functions of ζ and η. In particular, for the Γ-OU model we have

$$\kappa_3 = \frac{2\eta^2}{\zeta}, \qquad \kappa_4 = \frac{6\eta^3}{\zeta^2}. \tag{B.1}$$

Using those cumulants we obtain for the matrix Υ the entries given below.

$$\Upsilon_{11} = \eta - \gamma^2\eta$$

$$\Upsilon_{12} = -\left(\gamma^2 - 1\right)\zeta\eta$$

$$\Upsilon_{13} = -2\left(\gamma^2 - 1\right)\zeta\eta - \left(\gamma^3 - 1\right)\kappa_3$$

$$\Upsilon_{14} = \frac{\gamma - 1}{\lambda}\ \eta\big(\beta(\gamma - 1) - 2\lambda\rho\big)$$

$$\Upsilon_{15} = \frac{\gamma - 1}{\lambda}\ \zeta\eta\big(\beta(\gamma - 1) - 2\lambda\rho\big)$$

$$\begin{aligned}\Upsilon_{16} = -\frac{\gamma - 1}{2\lambda}\Big[&2\eta\Big(\beta\zeta\left(\gamma^2 + (\epsilon\lambda + \lambda - 1)\gamma + (\epsilon + 1)\lambda\right)\\ &\qquad + \lambda\left((\gamma + 1)\mu + \zeta(\gamma\lambda + \lambda + 2)\rho\right)\Big)\\ &+ \left(\beta(-2\gamma^2 + \gamma + 1) + 3(\gamma + 1)\lambda\rho\right)\kappa_3\Big]\end{aligned}$$

$$\Upsilon_{22} = -\left(\gamma^2 - 1\right)\eta\left(\zeta^2 + \eta\right)$$

$$\Upsilon_{23} = \left(\zeta - \gamma^3\zeta\right)\kappa_3 - 2\left(\gamma^2 - 1\right)\eta\left(\zeta^2 + \gamma\eta\right)$$

$$\Upsilon_{24} = \frac{\gamma - 1}{\lambda}\ \zeta\eta\big(\beta(\gamma - 1) - 2\lambda\rho\big)$$

$$\Upsilon_{25} = \frac{\gamma - 1}{\lambda}\ \eta\left(\zeta^2 + \eta\right)\left(\beta(\gamma - 1) - 2\lambda\rho\right)$$

$$\begin{aligned}\Upsilon_{26} = -\frac{\gamma - 1}{2\lambda}\Big[&2\eta\Big\{\beta\Big(\left(\zeta^2 - \eta\right)\gamma^2 + \left((\epsilon\lambda + \lambda - 1)\zeta^2 + \eta + \epsilon\eta\lambda\right)\gamma\\ &\qquad\quad + \left((\epsilon + 1)\zeta^2 + \epsilon\eta\right)\lambda\Big)\\ &\qquad + \lambda\left((\gamma\lambda + \lambda + 2)\rho\zeta^2 + (\gamma + 1)\mu\zeta + 2\gamma\eta\rho\right)\Big\}\\ &+ \zeta\left(\beta\left(-2\gamma^2 + \gamma + 1\right) + 3(\gamma + 1)\lambda\rho\right)\kappa_3\Big]\end{aligned}$$

$$\Upsilon_{33} = -4\left(\gamma^3 - 1\right)\zeta\kappa_3 - \left(\gamma^2 - 1\right)\left(2\eta\left(\eta\gamma^2 + 2\zeta^2 + \eta\right) + \left(\gamma^2 + 1\right)\kappa_4\right)$$

$$\Upsilon_{34} = \frac{\gamma - 1}{2\lambda}\Big[4\zeta\eta(\beta(\gamma - 1) - 2\lambda\rho) + \left(\beta\left(2\gamma^2 - \gamma - 1\right) - 3(\gamma + 1)\lambda\rho\right)\kappa_3\Big]$$

$$\begin{aligned}\Upsilon_{35} = \frac{\gamma - 1}{2\lambda}\Big[&4\eta\left(\zeta^2 + \gamma\eta\right)\left(\beta(\gamma - 1) - 2\lambda\rho\right)\\ &+ \zeta\left(\beta\left(2\gamma^2 - \gamma - 1\right) - 3(\gamma + 1)\lambda\rho\right)\kappa_3\Big]\end{aligned}$$

$$
\begin{aligned}
\Upsilon_{36} = & -\frac{\gamma-1}{6\lambda}\Big[3\Big(\beta\zeta\Big\{2\gamma^3+2(\epsilon\lambda+\lambda-3)\gamma^2\\
&\qquad\qquad +(2(\epsilon+1)\lambda+3)\gamma+2(\epsilon+1)\lambda+1\Big\}\\
&\qquad +\lambda\Big\{2\left(\gamma^2+\gamma+1\right)\mu\\
&\qquad\qquad +\zeta\left(2\lambda\gamma^2+(2\lambda+9)\gamma+2\lambda+9\right)\rho\Big\}\Big)\kappa_3\\
&\quad +2\Big(6\eta\Big\{\beta\Big(\left(\zeta^2+\epsilon\eta\lambda\right)\gamma^2\\
&\qquad\qquad +\big((\epsilon\lambda+\lambda-1)\zeta^2+\eta(\epsilon\lambda-1)\big)\gamma\\
&\qquad\qquad +\eta+(\epsilon+1)\zeta^2\lambda\Big)\\
&\qquad +\lambda\Big((\gamma\lambda+\lambda+2)\rho\zeta^2+(\gamma+1)\mu\zeta+2\eta\rho\Big)\Big\}\\
&\qquad +\Big(\beta\left(-3\gamma^3+\gamma^2+\gamma+1\right)+4\left(\gamma^2+\gamma+1\right)\lambda\rho\Big)\kappa_4\Big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon_{44} = & \frac{1}{\lambda^2}\Big[-\eta\left(\gamma^2-4\gamma-2\lambda+3\right)\beta^2+4\eta\lambda(\gamma+\lambda-1)\rho\beta\\
&+\lambda\big(2\eta\lambda^2\rho^2+\zeta(\gamma+\epsilon\lambda+\lambda-1)\big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon_{45} = & \frac{1}{\lambda^2}\Big[-\zeta\eta\left(\gamma^2-4\gamma-2\lambda+3\right)\beta^2+4\zeta\eta\lambda(\gamma+\lambda-1)\rho\beta\\
&+\lambda\big((\gamma+\epsilon\lambda+\lambda-1)\zeta^2+2\eta\lambda^2\rho^2\zeta+\epsilon\eta\lambda\big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon_{46} = & \frac{1}{\lambda^2}\Big[\zeta\eta\Big(\gamma^3+(\epsilon\lambda+\lambda-4)\gamma^2+(7-2(\epsilon+1)\lambda)\gamma+(\epsilon+3)\lambda-4\Big)\beta^2\\
&+\eta\lambda\Big\{\mu(\gamma-1)^2+\zeta\big((\lambda-2)\gamma^2-2((\epsilon+2)\lambda-4)\gamma+2\epsilon\lambda+7\lambda-6\big)\rho\Big\}\beta\\
&+\lambda\Big\{(\gamma+\epsilon\lambda+\lambda-1)\zeta^2-2(\gamma-2)\eta\lambda^2\rho^2\zeta\\
&\qquad +\eta\Big(\gamma^2+(\epsilon\lambda-2\mu\rho\lambda-2)\gamma+2\lambda\mu\rho+1\Big)\Big\}\\
&-(\gamma-1)\Big(\beta^2(\gamma-1)^2-3\beta\lambda\rho(\gamma-1)+3\lambda^2\rho^2\Big)\kappa_3\Big]
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon_{55} = & \frac{1}{\lambda^2}\Big[\big(\zeta^2+\eta\big)\Big\{-\eta\big(\gamma^2-4\gamma-2\lambda+3\big)\beta^2+4\eta\lambda(\gamma+\lambda-1)\rho\beta\\
&\qquad\qquad +\lambda\big(2\eta\lambda^2\rho^2+\zeta(\gamma+\lambda-1)\big)\Big\}\Big]+\epsilon\big(\zeta^3+3\eta\zeta+\kappa_3\big)
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon_{56} = & \frac{1}{\lambda^2}\Big[\eta\Big\{\Big(\gamma^3+(\epsilon\lambda+\lambda-4)\gamma^2+(7-2(\epsilon+1)\lambda)\gamma+(\epsilon+3)\lambda-4\Big)\zeta^2\\
&\quad +\eta\Big(-\gamma^3+(\epsilon\lambda+4)\gamma^2+(-2(\epsilon-1)\lambda-3)\gamma+\epsilon\lambda\Big)\Big\}\beta^2\\
&+\eta\lambda\Big\{\zeta\mu(\gamma-1)^2+2\eta\Big(2\gamma^2-((\epsilon-2)\lambda+2)\gamma+\epsilon\lambda\Big)\rho
\end{aligned}
$$

Υ_{56} *continues*

$$
\begin{aligned}
&+\zeta^2\Big((\lambda-2)\gamma^2-2((\epsilon+2)\lambda-4)\gamma+2\epsilon\lambda+7\lambda-6\Big)\rho\Big\}\beta\\
&+\lambda\Big\{(\gamma+\epsilon\lambda+\lambda-1)\zeta^3-2(\gamma-2)\eta\lambda^2\rho^2\zeta^2\\
&\quad+\eta\Big(2\gamma^2+(2\epsilon\lambda-2\mu\rho\lambda+\lambda-3)\gamma+\epsilon\lambda+2\lambda\mu\rho+1\Big)\zeta+2\gamma\eta^2\lambda^2\rho^2\Big\}\\
&+\Big\{-\beta^2\zeta(\gamma-1)^3+3\beta\zeta\lambda\rho(\gamma-1)^2+\lambda^2\Big(3\zeta\rho^2+\gamma\big(\epsilon-3\zeta\rho^2\big)\Big)\Big\}\kappa_3\Big]
\end{aligned}
$$

$$
\begin{aligned}
\Upsilon_{66}=\frac{1}{6\lambda^2}\Big[3\Big\{\quad &2\beta^2\zeta\big(2\gamma^2+(2(\epsilon+1)\lambda-3)\gamma+\epsilon\lambda+\lambda+1\big)(\gamma-1)^2\\
&+2\beta\lambda\Big[\big(2\gamma^2-\gamma-1\big)\mu+\zeta\Big((2\lambda-3)\gamma^2+(-3\epsilon\lambda-4\lambda+6)\gamma\\
&\qquad-3\epsilon\lambda-4\lambda-3\Big)\rho\Big](\gamma-1)\\
&+\lambda\Big(2\gamma^3+\big(-6\zeta\lambda^2\rho^2-6\lambda\mu\rho+2\epsilon\lambda-3\big)\gamma^2\\
&\qquad-12\zeta\lambda\rho^2\gamma+6\zeta\lambda^2\rho^2+6\lambda\rho(\mu+2\zeta\rho)+1\Big)\quad\Big\}\kappa_3\\
-2\Big\{3\Big[&\eta\Big(\zeta^2\Big\{\gamma^4+2(\epsilon\lambda+\lambda-2)\gamma^3+\big((\epsilon+1)^2\lambda^2-4(\epsilon+1)\lambda+7\big)\gamma^2\\
&\qquad+2(\epsilon\lambda+\lambda-4)\gamma-(\epsilon+1)^2\lambda^2-2\lambda+4\Big\}\\
&\quad-\eta\Big\{\gamma^4+2(\epsilon\lambda-2)\gamma^3-\big(\epsilon^2\lambda^2+4\epsilon\lambda-5\big)\gamma^2\\
&\qquad+2\epsilon\lambda\gamma+\epsilon^2\lambda^2+2\lambda-2\Big\}\Big)\beta^2\\
&+2\eta\lambda\Big\{\big(\lambda\rho\zeta^2+\mu\zeta+2\eta\rho\big)\gamma^3\\
&\qquad+\Big(\big((\epsilon+1)\lambda^2-2\lambda+2\big)\rho\zeta^2+(\epsilon\lambda+\lambda-2)\mu\zeta\\
&\qquad\quad+2\eta(\epsilon\lambda-3)\rho\Big)\gamma^2\\
&\qquad+\Big((2\epsilon\lambda+3\lambda-6)\rho\zeta^2+\mu\zeta-2\eta(\epsilon\lambda-2)\rho\Big)\gamma\\
&\qquad-(\epsilon+1)\zeta\lambda\mu-2\eta\lambda\rho-\zeta^2\big((\epsilon+1)\lambda^2+2(\epsilon+2)\lambda-4\big)\rho\Big\}\beta\\
&+\lambda\Big\{-(\gamma+\epsilon\lambda+\lambda-1)\zeta^3+\eta\lambda^2\big(\lambda\gamma^2+4\gamma-\lambda-6\big)\rho^2\zeta^2\\
&\quad-\eta\Big(\big(2-2\lambda^2\mu\rho\big)\gamma^2+(2\epsilon\lambda-4\mu\rho\lambda-3)\gamma+2\lambda^2\mu\rho\\
&\qquad+\lambda(\epsilon+4\mu\rho+1)+1\Big)\zeta\\
&\quad-\eta\lambda\Big(2\eta\big(2\gamma^2-4\gamma+\lambda+2\big)\rho^2-\big(\gamma^2-1\big)\mu^2\Big)\Big\}\qquad\Big]\\
&+(\gamma-1)\Big(\beta^2(3\gamma+1)(\gamma-1)^2+6(\gamma+1)\lambda^2\rho^2\\
&\qquad+4\beta\big(-2\gamma^2+\gamma+1\big)\lambda\rho\Big)\kappa_4\\
&\Big\}\qquad\Big]
\end{aligned}
$$

B.2 The Jacobian D

$$D_{1,1} = \frac{2\left(-1+e^{\lambda}\right)\zeta}{\eta}$$
$$D_{1,2} = -\frac{e^{\lambda}}{\eta}$$
$$D_{1,3} = \frac{1}{\eta}$$
$$D_{1,4} = 0$$
$$D_{1,5} = 0$$
$$D_{1,6} = 0$$
$$D_{2,1} = 1$$
$$D_{2,2} = 0$$
$$D_{2,3} = 0$$
$$D_{2,4} = 0$$
$$D_{2,5} = 0$$
$$D_{2,6} = 0$$
$$D_{3,1} = -2\zeta$$
$$D_{3,2} = 0$$
$$D_{3,3} = 1$$
$$D_{3,4} = 0$$
$$D_{3,5} = 0$$
$$D_{3,6} = 0$$
$$D_{4,1} = \frac{\beta\zeta\left(-e^{\lambda}\left(\lambda^2+4\right)+2e^{2\lambda}+2\right)-e^{\lambda}\lambda^2(\mu+\zeta\lambda\rho)}{\left(-1+e^{\lambda}\right)\eta\lambda}$$
$$D_{4,2} = -\frac{e^{\lambda}\beta\left(-\lambda+e^{\lambda}-1\right)}{\left(-1+e^{\lambda}\right)\eta\lambda}$$
$$D_{4,3} = \frac{-e^{\lambda}\beta+e^{\lambda}\lambda\beta+\beta}{\eta\lambda-e^{\lambda}\eta\lambda}$$
$$D_{4,4} = \frac{e^{\lambda}\zeta\lambda}{\eta-e^{\lambda}\eta}$$
$$D_{4,5} = \frac{e^{\lambda}\lambda}{\left(-1+e^{\lambda}\right)\eta}$$
$$D_{4,6} = 0$$
$$D_{5,1} = 0$$
$$D_{5,2} = \frac{e^{\lambda}\rho}{\left(-1+e^{\lambda}\right)\eta}$$
$$D_{5,3} = \frac{e^{\lambda}\rho}{\eta-e^{\lambda}\eta}$$

$$D_{5,4} = 0$$

$$D_{5,5} = \frac{e^{\lambda}}{2\eta - 2e^{\lambda}\eta}$$

$$D_{5,6} = -\frac{e^{\lambda}}{2\eta - 2e^{\lambda}\eta}$$

$$\begin{aligned} D_{6,1} = \frac{1}{\left(-1+e^{\lambda}\right)\eta\lambda} \Big[& \beta\Big(\zeta^2\left(e^{\lambda}\left(\lambda^2+4\right) - 2e^{2\lambda} - 2\right) - \left(-1+e^{\lambda}\right)\eta\lambda\Big) \\ & + \lambda\Big(-2e^{2\lambda}\rho\zeta^2 - 2\rho\zeta^2 + \eta\lambda\rho \\ & \qquad + e^{\lambda}\left(\left(\lambda^2+4\right)\rho\zeta^2 + \lambda\mu\zeta - \eta\lambda\rho\right)\Big)\Big] \end{aligned}$$

$$D_{6,2} = \frac{e^{\lambda}\zeta\left(-\lambda + e^{\lambda} - 1\right)(\beta + \lambda\rho)}{\left(-1+e^{\lambda}\right)\eta\lambda}$$

$$D_{6,3} = \frac{\zeta\left(e^{\lambda}(\lambda - 1) + 1\right)(\beta + \lambda\rho)}{\left(-1+e^{\lambda}\right)\eta\lambda}$$

$$D_{6,4} = \frac{e^{\lambda}\lambda\zeta^2 + \left(-1+e^{\lambda}\right)\eta}{\left(-1+e^{\lambda}\right)\eta}$$

$$D_{6,5} = \frac{e^{\lambda}\zeta\lambda}{2\eta - 2e^{\lambda}\eta}$$

$$D_{6,6} = \frac{e^{\lambda}\zeta\lambda}{2\eta - 2e^{\lambda}\eta}$$

C The simple multivariate martingale central limit theorem

The following simple version of a multivariate martingale central limit theorem is certainly well-known or obvious for experts, some references are [CP05, KS99b, vZ00].

However, when looking for references, we found statments that do not exactly apply, or that are much more general (continuous time, random normalizations, . . .). It turned out that the elementary proof below is shorter, than an attempt to verify the assumtions and deduce the result from a more 'advanced' theorem. Yet, any *concrete and precise* hint for an appropriate reference would be most welcome to the authors.

Theorem 3. *Suppose* $(X_{n,k})$ *is a martingale difference array such that for every* $\epsilon > 0$

$$\sum_{k=1}^{n} E\left[\|X_{n,k}\|^2 \mathbb{1}_{\{\|X_{n,k}\|>\epsilon\}} | \mathcal{F}_{k-1}\right] \xrightarrow{P} 0 \tag{C.1}$$

and

$$\sum_{k=1}^{n} \left[X_{n,k} X_{n,k}^{\top} | \mathcal{F}_{k-1}\right] \xrightarrow{P} \Upsilon \tag{C.2}$$

as $n \to \infty$. Then

$$\sum_{k=1}^{n} X_{n,k} \xrightarrow{\mathcal{D}} N(0, \Upsilon). \tag{C.3}$$

Proof. We will use the Cramer-Wald device. For $\beta \in \mathbb{R}^d$, $\beta \neq 0$, let us define a random variable

$$Y_{n,k} = \beta^\top X_{n,k}. \tag{C.4}$$

Then we have

$$\begin{aligned} \sum_{k=1}^{n} E\big[Y_{n,k}^2 | \mathcal{F}_{k-1}\big] &= \sum_{k=1}^{n} E\big[(\beta^\top X_{n,k})(X_{n,k}^\top \beta) | \mathcal{F}_{k-1}|\big] \\ &= \beta^\top \sum_{k=1}^{n} E\big[X_{n,k} X_{n,k}^\top | \mathcal{F}_{k-1}\big]\beta. \end{aligned} \tag{C.5}$$

From assumption (C.2) it follows that the expression (C.5) converges to $\beta^\top \Upsilon \beta$ and thus

$$\sum_{k=1}^{n} E\big[Y_{n,k}^2 | \mathcal{F}_{k-1}\big] \to \beta^\top \Upsilon \beta \tag{C.6}$$

as $n \to \infty$. Furthermore, it holds

$$\big|\beta^\top X_{n,k}\big| \leq \|\beta\| \cdot \|X_{n,k}\|. \tag{C.7}$$

Thus, for an arbitrarily $\epsilon > 0$ we have

$$\begin{aligned} 0 &\leq \sum_{k=1}^{n} E\big[Y_{n,k}^2 \mathbb{1}_{\{|Y_{n,k}| > \epsilon\}} | \mathcal{F}_{k-1}\big] \\ &= \sum_{k=1}^{n} E\big[(\beta^\top X_{n,k})^2 \mathbb{1}_{\{|\beta^\top X_{n,k}| > \epsilon\}} | \mathcal{F}_{k-1}\big] \\ &\leq \|\beta\|^2 \sum_{k=1}^{n} E\big[\|X_{n,k}\|^2 \mathbb{1}_{\{|\beta^\top X_{n,k}| > \epsilon\}} | \mathcal{F}_{k-1}\big]. \end{aligned} \tag{C.8}$$

Since for $\beta \neq 0$ the condition $|\beta^\top X_{n,k}| > \epsilon$ implies

$$\|X_{n,k}\| \geq \|\beta\|^{-1} |\beta^\top X_{n,k}| > \|\beta\|^{-1}\epsilon, \tag{C.9}$$

it follows that

$$\mathbb{1}_{\{|\beta^\top X_{n,k}| > \epsilon\}} \leq \mathbb{1}_{\{\|X_{n,k}\| > \|\beta\|^{-1}\epsilon\}} \tag{C.10}$$

and thus using the assumption (C.1), from (C.8) it follows that

$$\sum_{k=1}^{n} E\big[Y_{n,k}^2 \mathbb{1}_{\{|Y_{n,k}| > \epsilon\}} | \mathcal{F}_{k-1}\big] \xrightarrow{P} 0 \tag{C.11}$$

as $n \to \infty$. Now, the statement follows from the univariate martingale central limit theorem from [HH80]. □

Lemma 6. *The conditional Lyapounov condition implies the conditional Lindeberg condition, namely, if*

$$\sum_{k=1}^{n} E\big[\|X_{n,k}\|^4 | \mathcal{F}_{k-1}\big] \longrightarrow 0 \tag{C.12}$$

then

$$\sum_{k=1}^{n} E\big[\|X_{n,k}\|^2 \mathbb{1}_{\{\|X_{n,k}\|>\epsilon\}} | \mathcal{F}_{k-1}\big] \longrightarrow 0 \tag{C.13}$$

as $n \to \infty$.

Proof. For every $\epsilon > 0$ we have

$$\begin{aligned}
&\sum_{k=1}^{n} E\big[\|X_{n,k}\|^4 | \mathcal{F}_{k-1}\big] \\
&\quad = \sum_{k=1}^{n} E\big[\|X_{n,k}\|^4 \mathbb{1}_{\{\|X_{n,k}\|>\epsilon\}} | \mathcal{F}_{k-1}\big] + \sum_{k=1}^{n} E\big[\|X_{n,k}\|^4 \mathbb{1}_{\{\|X_{n,k}\|\le\epsilon\}} | \mathcal{F}_{k-1}\big] \\
&\quad \geq \epsilon^2 \sum_{k=1}^{n} E\big[\|X_{n,k}\|^2 \mathbb{1}_{\{\|X_{n,k}\|>\epsilon\}} | \mathcal{F}_{k-1}\big],
\end{aligned} \tag{C.14}$$

since

$$\sum_{k=1}^{n} E\big[\|X_{n,k}\|^4 \mathbb{1}_{\{\|X_{n,k}\|\le\epsilon\}} | \mathcal{F}_{k-1}\big] \geq 0.$$

From assumption (C.12) the statement follows. □

References

[AS02] Yacine Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002.

[BG05] F.E. Benth and M. Groth. The minimal entropy martingale measure and numerical option pricing for the Barndorff-Nielsen–Shephard stochastic volatility model. E-print, Department of Mathematics, University of Oslo, 2005.

[BK05] Fred Espen Benth and Kenneth Hvistendahl Karlsen. A PDE representation of the density of the minimal entropy martingale measure in stochastic volatility markets. *Stochastics*, 77(2):109–137, 2005.

[BKR03] Fred Espen Benth, Kenneth Hvistendahl Karlsen, and Kristin Reikvam. Merton's portfolio optimization problem in a Black and Scholes market with non-Gaussian stochastic volatility of Ornstein-Uhlenbeck type. *Math. Finance*, 13(2):215–244, 2003.

[BMB05] Fred Espen Benth and Thilo Meyer-Brandis. The density process of the minimal entropy martingale measure in a stochastic volatility model with jumps. *Finance Stoch.*, 9(4):563–575, 2005.

[BNNS02] Ole E. Barndorff-Nielsen, Elisa Nicolato, and Neil Shephard. Some recent developments in stochastic volatility modelling. *Quant. Finance*, 2(1):11–23, 2002. Special issue on volatility modelling.

[BNS01] Ole E. Barndorff-Nielsen and Neil Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63(2):167–241, 2001.

[BS95] Bo Martin Bibby and Michael Sørensen. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, 1(1-2):17–39, 1995.

[BS01] Bo Martin Bibby and Michael Sørensen. Simplified estimating functions for diffusion models with a high-dimensional parameter. *Scand. J. Statist.*, 28(1):99–112, 2001.

[Chu01] Kai Lai Chung. *A course in probability theory.* Academic Press Inc., San Diego, CA, third edition, 2001.

[CM00] Donato Michele Cifarelli and Eugenio Melilli. Some new results for Dirichlet priors. *Ann. Statist.*, 28(5):1390–1413, 2000.

[CP05] Irene Crimaldi and Luca Pratelli. Convergence results for multivariate martingales. *Stochastic Processes and their Applications*, 115(4):571–577, 2005.

[CR90] Donato Michele Cifarelli and Eugenio Regazzini. Distribution functions of means of a Dirichlet process. *Ann. Statist.*, 18(1):429–442, 1990.

[CT97] Yuan Shih Chow and Henry Teicher. *Probability theory.* Springer-Verlag, New York, third edition, 1997.

[CT04] Rama Cont and Peter Tankov. *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[DFS03] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 2003.

[DPS00] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[DS04] Susanne Ditlevsen and Michael Sørensen. Inference for observations of integrated diffusion processes. *Scand. J. Statist.*, 31(3):417–429, 2004.

[FSS01] Sylvia Frühwirt-Schnatter and Leopold Sögner. MCMC estimation of the Barndorff-Nielsen-Shephard stochastic volatility model. Unpublished talks, 2001.

[GS01] James E. Griffin and Mark F.J. Steel. Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. Journal of Econometrics, forthcoming, 2001.

[HH80] P. Hall and C. C. Heyde. *Martingale limit theory and its application.* Academic Press Inc., New York, 1980.

[HHY96] L.P. Hansen, J. Heaton, and A. Yaron. Finite-sample properties of some alternative GMM estimators. *Journal of Business Economic Statistics*, 14:262–280, 1996.

[HP06] Friedrich Hubalek and Petra Posedel. Asymptotic analysis for an optimal estimating function for Barndorff-Nielsen-Shephard stochastic volatility models. Work in progress, 2006.

[Jac01] Martin Jacobsen. Discretely observed diffusions: classes of estimating functions and small Δ-optimality. *Scand. J. Statist.*, 28(1):123–149, 2001.

[Jac02] Martin Jacobsen. Optimality and small Δ-optimality of martingale estimating functions. *Bernoulli*, 8(5):643–668, 2002.

[Jam05] Lancelot F. James. Analysis of a class of likelihood based continuous time stochastic volatility models including Ornstein-Uhlenbeck models in financial economics. `arXiv:math/0503055`, 2005.

[Jam06] Lancelot F. James. Laws and likelihoods for Ornstein Uhlenbeck-gamma and other BNS OU stochastic volatilty models with extensions. `arXiv:math/0604086`, 2006.

[Joh78] Søren Johansen. The product limit estimator as maximum likelihood estimator. *Scand. J. Statist.*, 5(4):195–199, 1978.

[Kes00] Mathieu Kessler. Simple and explicit estimating functions for a discretely observed diffusion process. *Scand. J. Statist.*, 27(1):65–82, 2000.

[KP02] Mathieu Kessler and Silvestre Paredes. Computational aspects related to martingale estimating functions for a discretely observed diffusion. *Scand. J. Statist.*, 29(3):425–440, 2002.

[KS99a] Mathieu Kessler and Michael Sørensen. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5(2):299–314, 1999.

[KS99b] Uwe Küchler and Michael Sørensen. A note on limit theorems for multivariate martingales. *Bernoulli*, 5(3):483–493, 1999.

[KW56] J. Kiefer and J. Wolfowitz. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Ann. Math. Statist.*, 27:887–906, 1956.

[Leh99] E. L. Lehmann. *Elements of large-sample theory.* Springer-Verlag, New York, 1999.

[Lin05] Carl Lindberg. *Portfolio Optimization and Statistics in Stochastic Volatility Markets.* PhD dissertation, Chalmers University of Technology, 2005.

[Lin06] Carl Lindberg. News-generated dependence and optimal portfolios for n stocks in a market of Barndorff-Nielsen and Shephard type. *Math. Finance*, 16(3):549–568, 2006.

[McC87] Peter McCullagh. *Tensor methods in statistics.* Chapman & Hall, London, 1987.

[MR03] Rosa Maria Mininni and Silvia Romanelli. Martingale estimating functions for Feller diffusion processes generated by degenerate elliptic operators. *J. Concr. Appl. Math.*, 1(3):191–216, 2003.

[NS03] B. Nielsen and N. Shephard. Likelihood analysis of a first-order autoregressive model with exponential innovations. *J. Time Ser. Anal.*, 24(3):337–344, 2003.

[NV03] Elisa Nicolato and Emmanouil Venardos. Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type. *Math. Finance*, 13(4):445–466, 2003.

[Pan02] Jun Pan. The jump-risk premia implicit in options: evidence from an integrated time-series study. *J. Financial Economics*, 63:3–50, 2002.

[Pos06] Petra Posedel. *Inference for a class of stochastic volatility models in presence of jumps with a special emphasis on martingale estimating functions.* PhD dissertation, Bocconi University, 2006. submitted.

[Rao73] M. M. Rao. Abstract martingales and ergodic theory. In *Multivariate analysis, III (Proc. Third Internat. Sympos., Wright State Univ., Dayton, Ohio, 1972)*, pages 45–60. Academic Press, New York, 1973.

[RPD04] Gareth O. Roberts, Omiros Papaspiliopoulos, and Petros Dellaportas. Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 66(2):369–393, 2004.

[RS06] Thorsten Rheinländer and Gallus Steiger. The minimal entropy martingale measure for general Barndorff-Nielsen/Shephard models. *The Annals of Applied Probability*, 16(3), 2006.

[Sch03] Wim Schoutens. *Lévy processes in finance.* Wiley, 2003.

[Sch05] Ernst Schaumburg. Estimation of Markov processes with Lévy type generators. Working paper, 2005.

[Sin01] Kenneth J. Singleton. Estimation of affine asset pricing models using the empirical characteristic function. *J. Econometrics*, 102(1):111–141, 2001.

[Sø97] Michael Sørensen. Statistical inference for discretely observed diffusion. Lectures at Berliner Graduiertenkolleg 'Stochastische Prozesse und Probabilistische Analysis', 1997.

[Sø99] Michael Sørensen. On asymptotics of estimating functions. *Brazilian Journal of Probability and Statistics*, (13):111–136, 1999.

[Sø00] Helle Sørensen. *Inference for Diffusion Processes and Stochastic Volatility Models.* PhD thesis, 2000.

[Sør97] Michael Sørensen. Estimating functions for discretely observed diffusions: a review. In *Selected Proceedings of the Symposium on Estimating Functions (Athens, GA, 1996)*, volume 32 of *IMS Lecture Notes Monogr. Ser.*, pages 305–325, Hayward, CA, 1997. Inst. Math. Statist.

[Sør01] Helle Sørensen. Discretely observed diffusions: approximation of the continuous-time score function. *Scand. J. Statist.*, 28(1):113–121, 2001.

[SS00] A. N. Shiryaev and V. G. Spokoiny. *Statistical experiments and decisions.* World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[Sto74] William F. Stout. *Almost sure convergence.* Academic Press, New York-London, 1974.

[Str85] Helmut Strasser. *Mathematical theory of statistics.* Walter de Gruyter & Co., Berlin, 1985.

[tH03] Enrique ter Horst. *A Lévy generalization of compound Poisson processes in finance: Theory and application (Chapter 3).* PhD thesis, Duke University, 2003.

[Tod06] Viktor Todorov. Econometric analysis of jump-driven stochastic volatility models. preprint, 2006.

[Uch04a] Masayuki Uchida. Estimation for discretely observed small diffusions based on approximate martingale estimating functions. *Scand. J. Statist.*, 31(4):553–566, 2004.

[Uch04b] Masayuki Uchida. Minimum contrast estimation for discretely observed diffusion processes with small dispersion parameter. *Bull. Inform. Cybernet.*, 36:35–49, 2004.

[Ven01] Emmanouil Venardos. *Derivatives Pricing and Ornstein-Uhlenbeck type Stochastic Volatility.* Dissertation, University of Oxford, 2001.

[vZ00] Harry van Zanten. A multivariate central limit theorem for continuous local martingales. *Statistics & Probability Letters*, 50(3):229–235, 2000.