ASYMPTOTIC ANALYSIS AND EXPLICIT ESTIMATION OF A CLASS OF STOCHASTIC VOLATILITY MODELS WITH JUMPS USING THE MARTINGALE ESTIMATING FUNCTION APPROACH

Friedrich Hubalek and Petra Posedel

Vienna University of Technology, Austria and Zagreb School of Economics and Management, Croatia

Abstract. We provide and analyze explicit estimators for a class of discretely observed continuous-time stochastic volatility models with jumps. In particular we consider the class of non-Gaussian Ornstein-Uhlenbeck based models, as introduced by Barndorff-Nielsen and Shephard.

We develop in detail the martingale estimating function approach for this kind of processes, which are bivariate Markov processes, that are not diffusions, but admit jumps. We assume that the bivariate process is observed on a discrete grid of fixed width, and the observation horizon tends to infinity.

We 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.

As an illustration we provide a simulation study for daily increments, but the method applies unchanged for any time-scale, including high-frequency observations, without introducing any discretization error.

1. Introduction

Barndorff-Nielsen and Shephard introduced in their seminal paper [BNS01] a class of continuous-time stochastic volatility models that allow flexible modelling, capture many stylized facts of financial time series, and yet are of great analytical tractability. Here the volatility, or rather more precisely,

the instantaneous variance follows an Ornstein-Uhlenbeck type process driven by an increasing Lévy process, and thus exhibits jumps.

BNS-models, as we will call them from now on, are affine models in the sense of [DPS00] and [DFS03], and associated with Riccati type equations that 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], and it is this property that turns out to be extremely useful not only for option pricing, but as we will show below, also for estimation purposes.

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 and in [KMKV11] pricing options on variance is analyzed. Multivariate extensions and option pricing is studied in [MKPS10]. In [BK05, BMB05, BG05, RS06] the minimal entropy martingale measure is investigated. The papers [BKR03, Lin06] address the portfolio optimization problem.[1] Bayesian/MCMC/computer intensive estimation is already in the seminal paper [BNS01], and in the works [RPD04, tH03, GS06, FSS09]. BNS models are also treated in the textbooks [CT04, Sch03].

Unfortunately, it seems that statistical estimation of the model is the most difficult problem, and most of the work in that area is focused on computationally intensive methods. We will show in the present paper that this is not the only possibility and that an explicit estimator for this class of stochastic volatility models with jumps can be obtained. Our contributions are as follows:

- Firstly we develop the closed form estimator using the martingale estimating function approach. Our approach is based on the general theory and framework in [Sø99] and [Sø97]. Our work differs from a large part of the existing literature in so far as we deal with a bivariate process, the process is not a diffusion, and we do not approximate transition probabilities, but can use exact expressions.
- Secondly, we prove consistency and asymptotic normality for the estimator. 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. Since the joint conditional

---

[1]Although the model captures several stylized facts of financial data, there have been some criticisms in that respect as well. The principal objection against the model is that jumps of returns and volatility are exactly the same, up to a multiplicative factor, and jumps in returns are one-sided. The turbo Bates model [Bat00] allows more flexible jump modelling while still being affine, which makes the application of the martingale estimating function approach feasible. Nevertheless, we prefer to present our analysis for the structurally slightly simple BNS models.

Some results on quantitative end empirical research on the relation of jumps of returns and volatility can be found in [JT10] and [JKG12].

moment-generating function of returns X and variance V is known in closed form we obtain closed form expressions for the joint conditional moments up to any desired order which yields a sequence of martingale differences. We employ then the large sample properties and a multivariate martingale central limit theorem.

- Thirdly we provide a detailed application of the theory of martingale estimating functions in a non-diffusion setting, including numerical illustrations.

  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 precision. Yet the model exhibits other peculiarities, see the remarks in Section 2.3.

- The present paper provides the basis and a benchmark for [HP11], where an empirical application, implementation issues, and some modifications of our approach are discussed.

This paper is organized as follows: In Section 2.1 we describe the class of BNS stochastic volatility 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 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. In Section 4 we present numerical illustrations. In conclusion, further and alternative developments are discussed, in particular considering the issue that volatility is typically not observed in discrete time. Explicit moment calculations of any order can be found in Appendix A.

Let us conclude the introduction by providing for the readers' convenience further references on stochastic volatility models, their applications and estimation in a broader context.

For an overview and key references for the role of stochastic volatility in modern time series econometrics, applications for modelling, decision making and prediction on financial markets we refer the reader to [She05] and [AS09].

Derivatives pricing, hedging, and risk management with stochastic volatility can be found in [Lew00]. Prominent stochastic volatility models of diffusion type, thus exhibiting continuous sample paths for return and volatility processes, are for example [HW87, SS91, Hes93]. In the last decade several continuous-time stochastic volatility models that exhibit jumps both in return and volatility processes have been suggested and studied, such as [Bat96]. Further information on BNS model is given in [BNNS02].

For the general theory and powerful methods to analyze the nature and relation of jumps in returns and volatility we refer the reader to the works of Jacod and Ait-Sahalia [ASJ07, ASJ09].

The literature on estimation for discretely observed diffusions is vast, we just mention the books [PR99], [Kut04] and the works [MR03, Jac02, KS99, BS95]. In particular, the martingale estimating function approach is used, developed and studied for example in [Sø99], [Sø00], and [Sø97].

2. THE MODEL

2.1. *The continuous time model.*

2.1.1. *The general setting.* As in Barndorff-Nielsen and Shephard, see [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 returns X and the instantaneous variance process V satisfy

$$dX(t) = (\mu + \beta V(t-))dt + \sqrt{V(t-)}dW + \rho dZ_\lambda(t), \quad X(0) = 0$$

and

$$(2.1) \qquad dV(t) = -\lambda V(t-)dt + dZ_\lambda(t), \quad V(0) = V_0,$$

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 Shephard, 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

$$E[V_0] = \zeta, \qquad \text{Var}[V_0] = \eta.$$

To shorten the notation we introduce the parameter vector

$$\theta = (\lambda, \zeta, \eta, \mu, \beta, \rho)^\top,$$

and the bivariate process

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

If the distribution of V_0 is from a particular class D then $\mathbf{X}$ is called a BNS-DOU(θ) model. For example, in Section 2.1.2 D is the class of Γ distributions and in Section 2.1.3 D is the class of inverse Gaussian distributions. For a general discussion of DOU and OUD processes, see [BNS01, Section 2.3., p.173]. The bivariate process $\mathbf{X}$ 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

$$(2.2) \qquad \zeta = \frac{\nu}{\alpha}, \qquad \eta = \frac{\nu}{\alpha^2}.$$

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 constructing 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 $\text{IG}(\delta/2, \gamma)$ 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

$$(2.3) \qquad \zeta = \frac{\delta}{\gamma}, \qquad \eta = \frac{\delta}{\gamma^3}.$$

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 the variance process on a discrete grid of points in time,

$$0 = t_0 < t_1 < \ldots < t_n.$$

This implies

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

Using

$$(2.4) \qquad V_i := V(t_i), \quad U_i := \int_{t_{i-1}}^{t_i} e^{-\lambda(t_i - s)} dZ_\lambda(s)$$

we have that $(U_i)_{i \geq 1}$ is a sequence of independent 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

$$X(t_i) - X(t_{i-1}) = \mu(t_i - t_{i-1}) + \beta(Y(t_i) - Y(t_{i-1}))$$

$$+ \int_{t_{i-1}}^{t_i} \sqrt{V(s-)}dW(s) + \rho(Z_\lambda(t_i) - Z_\lambda(t_{i-1})),$$

where

$$Y(t) = \int_0^t V(s-)ds$$

is the integrated variance process. This suggests introducing the discrete time quantities

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

and

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

Furthermore, it is also convenient to introduce the discrete quantity

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

It is not difficult to see (conditioning!) that $(W_i)_{i \geq 1}$ is an iid $N(0,1)$ sequence independent from all discrete quantities derived from the BDLP alone, in particular $(Z_i)_{i \geq 1}$ and $(U_i)_{i \geq 1}$. 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 $t_k = k\Delta$, where $\Delta > 0$ is fixed. This implies

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

and

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

where

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

Furthermore,

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

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

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 assumptions are made, and all conditions required for consistency and asymptotic normality of our estimator will be proven rigorously from that assumption.

PROPOSITION 2.1. *We have for all $n \in \mathbb{N}$ that*

$$E[Z_1^n] < \infty, \quad E[U_1^n] < \infty, \quad E[S_1^n] < \infty,$$

and

$$E[Y_1^n] < \infty, \quad E[W_1^n] < \infty, \quad E[X_1^n] < \infty.$$

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-known relation between the existence 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 $E[V_0^n] < \infty$ 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 $E[Z(\lambda\Delta)^n] < \infty$, and as $Z_1 = Z(\lambda\Delta)$ we have shown $E[Z_1^n] < \infty$, for all $n \in \mathbb{N}$.

From (2.4) and (2.6) 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 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 functions is developed for diffusions, see for example [MR03, Jac02, Kes00, Sør97], 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 usually made 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.

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, see [KW56, Joh78], but for the purpose of our analysis that direction is not exploited further. For a related problem 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 a *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 $\left(\mathbb{D}(\mathbb{R})^2, \mathcal{D}(\mathbb{R}^2)\right)$, where $\mathcal{D}(\mathbb{R}^2)$ is the Borel $\sigma-$algebra on $\mathbb{D}(\mathbb{R})^2$ with respect to the Skorohod topology, under each P_θ as a starting point, see [JS03].

3. The explicit estimator

3.1. *The estimating equations and their explicit solution.* For estimation purposes we consider a probability space on which a parameterized family of probability measures is given:

$$\left(\Omega, \mathcal{F}, \left\{P_\theta : \theta \in \Theta\right\}\right),$$

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 and that satisfies Assumption 1 for all $\theta \in \Theta$. We want to find an estimator for θ_0 using observations $X_1, \ldots, X_n, V_1, \ldots, V_n$ and we are interested in asymptotics as $n \to \infty$. To that purpose let us consider the vector estimating function $G_n(\theta) = \left(G_n^1(\theta), \ldots, G_n^6(\theta)\right)$ with the following martingale estimating functions:

$$(3.1)\quad
\begin{aligned}
G_n^1(\theta) &= \sum_{k=1}^n \left[V_k - f^1(V_{k-1}, \theta)\right], & f^1(v, \theta) &= E_\theta[V_1|V_0 = v],\\
G_n^2(\theta) &= \sum_{k=1}^n \left[V_k V_{k-1} - f^2(V_{k-1}, \theta)\right], & f^2(v, \theta) &= E_\theta[V_1 V_0|V_0 = v],\\
G_n^3(\theta) &= \sum_{k=1}^n \left[V_k^2 - f^3(V_{k-1}, \theta)\right], & f^3(v, \theta) &= E_\theta[V_1^2|V_0 = v],\\
G_n^4(\theta) &= \sum_{k=1}^n \left[X_k - f^4(V_{k-1}, \theta)\right], & f^4(v, \theta) &= E_\theta[X_1|V_0 = v],\\
G_n^5(\theta) &= \sum_{k=1}^n \left[X_k V_{k-1} - f^5(V_{k-1}, \theta)\right], & f^5(v, \theta) &= E_\theta[X_1 V_0|V_0 = v],\\
G_n^6(\theta) &= \sum_{k=1}^n \left[X_k V_k - f^6(V_{k-1}, \theta)\right], & f^6(v, \theta) &= E_\theta[X_1 V_1|V_0 = v].
\end{aligned}$$

LEMMA 3.1. *We have the explicit expressions*

$$f^1(v, \theta) = \gamma v + (1 - \gamma)\zeta,$$
$$f^2(v, \theta) = \gamma v^2 + (1 - \gamma)\zeta v,$$
$$f^3(v, \theta) = \gamma^2 v^2 + 2\gamma(1 - \gamma)\zeta v + (1 - \gamma)^2\zeta^2 + (1 - \gamma^2)\eta,$$
$$\text{(3.2)} \qquad f^4(v, \theta) = \beta\epsilon v + \mu\Delta + \beta(\Delta - \epsilon)\zeta + \rho\lambda\zeta\Delta,$$
$$f^5(v, \theta) = \beta\epsilon v^2 + (\mu\Delta + \beta(\Delta - \epsilon)\zeta + \rho\lambda\zeta\Delta)v,$$
$$f^6(v, \theta) = 2\epsilon\eta\lambda\rho + \Delta(\gamma v + \zeta\epsilon\lambda)(\mu + \lambda\rho\zeta)$$
$$\qquad\qquad + \beta(\epsilon^2\eta\lambda + (\gamma v + \zeta\epsilon\lambda)(\epsilon(v - \zeta) + \Delta\zeta)).$$

PROOF. The formulas are special cases of the general moment calculations given in Appendix A. Proofs of the statements for two special cases that show the basic calculation idea are given in [HP07]. □

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

PROPOSITION 3.2. *The estimating equation $G_n(\hat{\theta}_n) = 0$ admits for every $n \geq 2$ on the event*

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

a unique solution $\hat{\theta}_n = (\lambda_n, \zeta_n, \eta_n, \mu_n, \beta_n, \rho_n)$ that is given by
(3.3)

$$\gamma_n = (\xi_n^2 - \xi_n^1 v_n^1)/(v_n^2 - (v_n^1)^2),$$
$$\zeta_n = (\xi_n^1 - \gamma_n v_n^1)/(1 - \gamma_n),$$
$$\eta_n = ((\xi_n^3 - (\xi_n^1)^2) - \gamma_n^2(v_n^2 - (v_n^1)^2))/(1 - \gamma_n^2),$$
$$\lambda_n = -\log(\gamma_n)/\Delta,$$
$$\epsilon_n = (1 - \gamma_n)/\lambda_n,$$
$$\beta_n = (\xi_n^5 - v_n^1\xi_n^4)/(\epsilon_n(v_n^2 - (v_n^1)^2)),$$
$$\rho_n = (\xi_n^6 - \xi_n^4\xi_n^1 - \beta_n\epsilon_n(\eta_n(1 - \gamma_n) + \gamma_n(v_n^2 - (v_n^1)^2)))/(2(1 - \gamma_n)\eta_n),$$
$$\mu_n = (\xi_n^4 - \beta_n\epsilon_n(v_n^1 - \zeta_n))/\Delta - (\beta_n + \lambda_n\rho_n)\zeta_n,$$

where

$$\text{(3.4)} \qquad \begin{aligned} &\xi_n^1 = \tfrac{1}{n}\sum_{i=1}^n V_i, \quad \xi_n^2 = \tfrac{1}{n}\sum_{i=1}^n V_i V_{i-1}, \quad \xi_n^3 = \tfrac{1}{n}\sum_{i=1}^n V_i^2, \\ &\xi_n^4 = \tfrac{1}{n}\sum_{i=1}^n X_i, \quad \xi_n^5 = \tfrac{1}{n}\sum_{i=1}^n X_i V_{i-1}, \quad \xi_n^6 = \tfrac{1}{n}\sum_{i=1}^n X_i V_i \end{aligned}$$

and

(3.5)
$$v_n^1 = \frac{1}{n}\sum_{i=1}^{n} V_{i-1}, \quad v_n^2 = \frac{1}{n}\sum_{i=1}^{n} V_{i-1}^2.$$

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 3.3. The exceptional set C_n could be simplified to
$$C_n' = \left\{\xi_n^2 - \xi_n^1 v_n^1 > 0\right\}.$$

Since the jump times and the jump sizes of the BDLP are independent, and the former have an exponential distribution it follows that $V_0, \ldots, V_n$ is with probability one not constant, so $P[v_n^2 - (v_n^1)^2 > 0] = 1$. But although it can be shown that the probability of C_n tends to one, for finite n we have $P[\xi_n^2 - \xi_n^1 v_n^1 \le 0] > 0$. 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.* In this section we investigate the consistency of the estimator from the previous section. First, we will need the following lemma.

LEMMA 3.4. *For every $k \ge 1$ and $p > 0$*
$$\frac{V_n^k}{n^p} \xrightarrow{a.s.} 0 \qquad as \quad n \to \infty.$$

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

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

In Lemma 3.5 and Lemma 3.6 we provide the strong law of large numbers for $(V_i^p, i \ge 1)$ and for more general sequences $(X_i^p V_i^q V_{i-1}^r, i \ge 1)$. The proofs of Lemma 3.5 and Lemma 3.6 are omitted since the results follow directly from the Birkhoff's ergodic theorem. Indeed, the sequence of $\mathbb{R}^3$-valued variables $T_i = (V_{i-1}, V_i, X_i)$ is strictly stationary, so $\frac{1}{n}\sum_{i=1}^n g(T_i) \xrightarrow{a.s.} \mathbb{E}(g(T_1))$ for any

function g such that $g(T_1) \in \mathbf{L}^1$. More elementary proofs, which do not use any ergodicity arguments, but employ a law of large numbers by Rajchman ([Raj32]) can be found in [HP07].

LEMMA 3.5. *We have for all $k \in \mathbb{N}$ that*

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

as $n \to \infty$.

LEMMA 3.6. *For all integers $p, q, r \geq 0$ we have*

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

as $n \to \infty$.

THEOREM 3.7. *We have $P(C_n) \to 1$ when $n \to \infty$ and the estimator $\hat{\theta}_n$ is consistent on C_n, namely*

$$\hat{\theta}_n \mathbb{1}_{C_n} \xrightarrow{a.s.} \theta_0$$

as $n \to \infty$, where $\mathbb{1}_{(.)}$ denotes the indicator function.

PROOF. Using the results of Lemma 3.5 it easily follows that

$$\xi_n^2 - \xi_n^1 v_n^1 \to Cov(V_1, V_0) > 0,$$

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

Using again the results of Lemma 3.5 it follows that the empirical moments in (3.4) and (3.5) converge to their theoretical counterparts, $\xi_n^i \xrightarrow{a.s.} \xi^i$ and $v_n^i \xrightarrow{a.s.} v^i$, where

$$\xi^1 = \zeta,$$
$$\xi^2 = \zeta^2 + \gamma\eta,$$
$$\xi^3 = \zeta^2 + \eta, \qquad\qquad\qquad v^1 = \zeta,$$
$$\xi^4 = \mu + (\beta + \lambda\rho)\zeta, \qquad\qquad v^2 = \zeta^2 + \eta.$$
$$\xi^5 = \mu\zeta + (\beta + \lambda\rho)\zeta^2 + \beta\epsilon\eta,$$
$$\xi^6 = \mu\zeta + (\beta + \lambda\rho)\zeta^2 + (\beta + 2\rho\lambda)\epsilon\eta,$$

Plugging the limits into (3.3) shows, after a short mechanical calculation, that the estimator is in fact consistent. $\qquad\square$

3.3. *Asymptotic normality.* For a concise vector notation we introduce

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

and write the estimating equations in the form

$$G_n^i(\theta) = \sum_{k=1}^{n} \left[\Xi_k^i - f^i(V_{k-1}, \theta)\right], \qquad i = 1, \ldots, 6$$

and $f^i(v, \theta)$ given by (3.2). We write

$$(3.6) \qquad f^i(v, \theta) = \sum_{\ell=r_i}^{p_i+r_i+q_i} \phi_\ell^i(\theta) v^\ell.$$

with

$$p = (0, 0, 0, 1, 1, 1), \quad q = (1, 1, 1, 0, 0, 1), \quad r = (0, 1, 0, 0, 1, 0)$$

and

$$\phi_2^1(\theta) = 0, \qquad \phi_1^1(\theta) = \gamma, \qquad \phi_0^1(\theta) = (1 - \gamma)\zeta,$$
$$\phi_2^2(\theta) = \gamma, \qquad \phi_1^2(\theta) = (1 - \gamma)\zeta,$$
$$\phi_2^3(\theta) = \gamma^2, \qquad \phi_1^3(\theta) = 2\gamma(1 - \gamma)\zeta, \qquad \phi_0^3(\theta) = (1 - \gamma)^2\zeta^2 + (1 - \gamma^2)\eta,$$
$$\phi_2^4(\theta) = 0, \qquad \phi_1^4(\theta) = \beta\epsilon, \qquad \phi_0^4(\theta) = \mu + \beta(1 - \epsilon)\zeta + \rho\lambda\zeta,$$
$$\phi_2^5(\theta) = \beta\epsilon, \qquad \phi_1^5(\theta) = \mu + \beta(1 - \epsilon)\zeta + \rho\lambda\zeta,$$
$$\phi_2^6(\theta) = \beta\epsilon\gamma, \qquad \phi_1^6(\theta) = (\mu + \beta(1 - \epsilon)\zeta + \rho\lambda\zeta)\gamma + \beta\epsilon(1 - \gamma)\zeta,$$
$$\phi_0^6(\theta) = (1 - \epsilon)(1 - \gamma)\zeta^2 + \epsilon^2\lambda\eta.$$

We will use, that $f^i(v, \theta)$ is a polynomial in v, and that its coefficients[2] ϕ are smooth functions in θ.

We shall first prove the central limit theorem for the estimating functions. The main tool that we use to that purpose is the multivariate martingale central limit theorem that is given in [HP07], which we recall here. For a vector a, $\|a\|$ denotes the Euclidean norm.

LEMMA 3.8. *Suppose* $(X_{n,k})$ *is a martingale difference array such that for every* $\epsilon > 0$

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

and

$$\sum_{k=1}^n \big[X_{n,k} X_{n,k}^\top|\mathcal{F}_{k-1}\big] \xrightarrow{P} \Upsilon$$

as $n \to \infty$. *Then*

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

PROPOSITION 3.9. *We have*

$$(3.7) \qquad \frac{1}{\sqrt{n}} G_n(\theta_0) \xrightarrow{\mathcal{D}} N(0, \Upsilon),$$

----

[2]ϕ_2^1 and ϕ_2^4 are not used in (3.6), but in (3.10).

as $n \to \infty$, where
$$\Upsilon_{ij} = E\big[Cov(\Xi_1^i, \Xi_1^j | V_0)\big].$$

PROOF. To show the above result, we use Lemma 3.8. To that purpose we introduce the vector martingale difference array

(3.8) $$\chi_{n,k}^{(i)} = \frac{1}{\sqrt{n}}\big[\Xi_k^i - f^i(V_{k-1}, \theta_0)\big].$$

We have to show the two assumptions from Lemma 3.8. First, we prove a multivariate Lyapuonov condition which implies the Lindeberg condition. From (3.8) it follows that $\sqrt{n}\chi_{n,k}^{(i)}$ is of the form $p(V_{k-1}, V_k, X_k)$ 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 Appendix A it follows that

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

where $q(v_0)$ is a polynomial in v_0. From Lemma 3.5 it thus follows

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

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

$$\frac{1}{n}\big(\Xi_k^i - f^i(V_{k-1}, \theta_0)\big)\big(\Xi_k^j - f^j(V_{k-1}, \theta_0)\big).$$

This is again a polynomial in V_{k-1}, V_k and X_k so by Lemma 3.6 it follows that

$$\frac{1}{n}\sum_{k=1}^{n}\big(\Xi_k^i - f^i(V_{k-1}, \theta_0)\big)\big(\Xi_k^j - f^j(V_{k-1}, \theta_0)\big)$$

$$\xrightarrow{a.s.} E\big[\big(\Xi_k^i - f^i(V_{k-1}, \theta_0)\big)\big(\Xi_k^j - f^j(V_{k-1}, \theta_0)\big)\big]$$

as $n \to \infty$. □

REMARK 3.10. A systematic method to evaluate Υ is given in Appendix A.

LEMMA 3.11. *We have*

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

where

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

and

(3.10) $$P_{ij} = \delta_{ij} - \phi_1^i \delta_{1j} - \phi_2^i \delta_{3j}$$

with δ_{ij} denoting the Kronecker delta.

PROOF. We can write

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

with

$$(3.11) \qquad Q_n^i = \frac{1}{\sqrt{n}}\left[\phi_1^i(V_n - V_0) + \phi_2^i(V_n^2 - V_0^2)\right].$$

In view of Lemma 3.4 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

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

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. $\square$

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

THEOREM 3.12. *The estimator*

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

is asymptotically normal, namely

$$(3.12) \qquad \sqrt{n}\left[\hat{\theta}_n - \theta_0\right] \xrightarrow{\mathcal{D}} N(0, T),$$

as $n \to \infty$, where

$$T = D\Sigma D^T$$

and D is explicitly given according to (3.13).

PROOF. We observe from (3.3) 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 in a neighborhood of ξ, and S_n goes to zero in probability in view of Lemma 3.4. 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

$$(3.13) \qquad D_{ij} = \frac{\partial h_i}{\partial x_j}(\xi), \qquad i, j = 1, \ldots, 6.$$

A lengthy elementary calculation shows that the matrix has determinant $\lambda/\left(2(1 - \gamma)^2\gamma\eta^3\right)$, thus it is invertible. $\square$

REMARK 3.13. For comparison it is instructive to study our estimator in the general framework of [Sø99] where the properties of the estimator are studied without exploiting the fact that the estimating equation allows an explicit solution. This is done in [HP07] where the theory is extended in the case of a bivariate Markov process and Condition 2.6 of [Sø99] is proven in order to use his Corollary 2.7 and Theorem 2.8.

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

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

The parameters imply that there are on average 2.6 jumps per day 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 variance is about half a day.

In our example annual log returns have (unconditional) mean 15.6% and 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 $i = 1, \ldots, 250$. Other scenarios can be considered, for example, small jumps arriving every minute, with fast decaying autocorrelation, or few jumps per year, corresponding to exceptional news with heavy impact on the variance process.

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

$$s = \begin{bmatrix} 4.86 \\ 125 \\ 650 \\ 7.36 \\ 253 \\ 0.526 \end{bmatrix}, \qquad r = \begin{bmatrix} 1 & 0.89 & 0.41 & 0.03 & 0.09 & -0.02 \\ 0.89 & 1 & 0.4 & 0.03 & 0.09 & -0.03 \\ 0.41 & 0.4 & 1 & 0.06 & 0.22 & 0 \\ 0.03 & 0.03 & 0.06 & 1 & -0.75 & 0.06 \\ 0.09 & 0.09 & 0.22 & -0.75 & 1 & -0.57 \\ -0.02 & -0.03 & 0 & 0.06 & -0.57 & 1 \end{bmatrix}$$
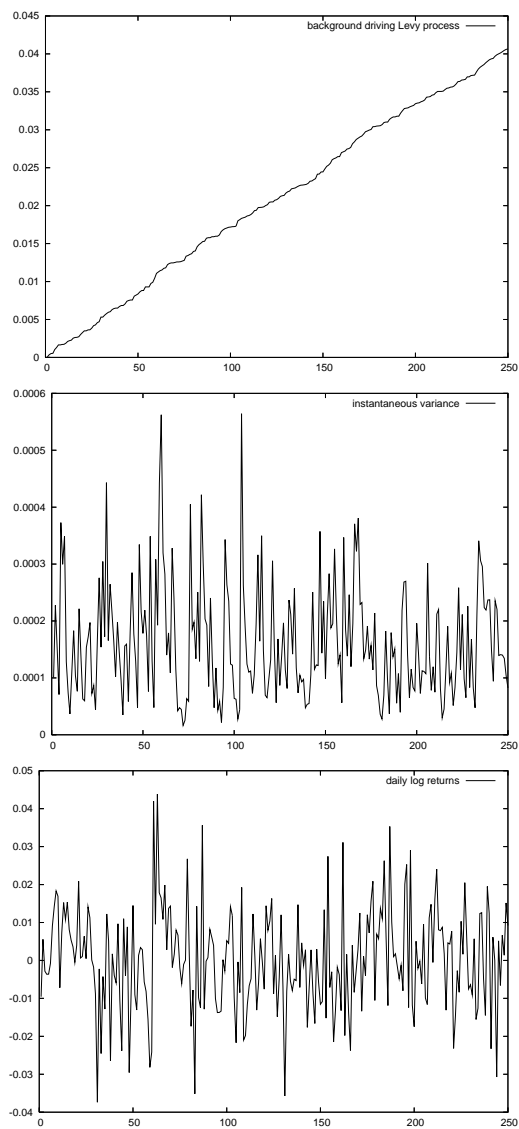
FIGURE 1. Daily observations Z_i, V_i, X_i.

The correlations among parameters related to the returns distribution, namely μ, β and ρ, are rather small except for those relating to the risk premium parameter β. In contrast to that, correlations among the instantaneous

variance parameters, namely ν, α and λ are rather high. This phenomenon can be related to the particular and common parametrization of the gamma distribution and discussed in more detail in [HP11, Sec.4.3]. Theoretically, the issue of high correlation can be addressed using the optimal martingale estimating function approach, see Section 5 for developments in this direction.

4.3. *Distribution of the estimates.* Figure 2 illustrates the empirical and asymptotic distribution of the estimators for the Γ-OU model. The histograms are produced from $m = 10000$ replications consisting of $n = 8000$ observations each, corresponding to 32 years with 250 daily observations per year. 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 derivatives pricing: 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.

5. Conclusion and further and alternative developments

We provide an explicit estimator for discretely observed Barndorff-Nielsen and Shephard stochastic volatility models, prove its consistency and asymptotic normality based on the single assumption that all moments of the stationary distribution of the variance process are finite. Explicit expressions for the asymptotic covariance matrix are given and we develop in detail the martingale estimating function approach for a bivariate model that is not a diffusion, but admits jumps. In this way we create a framework that opens a new and tractable approach for estimating stochastic volatility models.

Although on one hand large parts of the material are completely standard, on the other hand there is a new feature in the sense that the estimation of the model is not a truly parametric problem, since it is not necessary to specify the law of Z, only the first two moments of Z_1 must be specified. Nevertheless, $\sqrt{n}$ convergence and asymptotic normality are still true, and in that sense the results of the paper do not trivially follow from known results on estimation for ergodic Markov chains.

Our choice of estimating functions is natural, but, mathematically speaking, somewhat arbitrary. In ongoing work [HP09] we show, that the *optimal quadratic estimating function* based on the moments of V_1, X_1, V_1^2, $V_1 X_1$, X_1^2 can be computed explicitly, though the corresponding estimator has to be determined numerically. Our proposed and explicit estimator can be used as a starting point for an iterative root-finding procedure showing in this
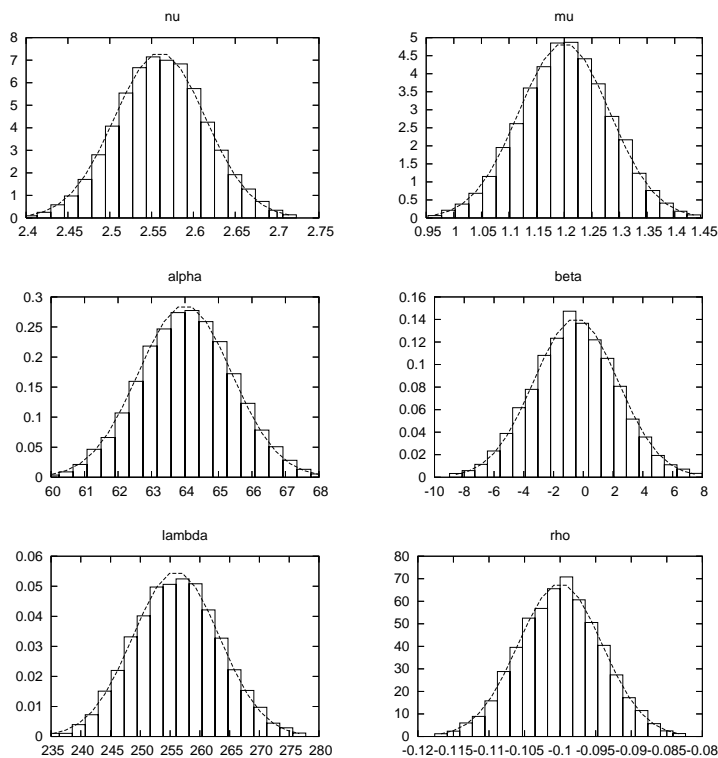
FIGURE 2. Empirical and asymptotic distribution of the estimators for the Γ-OU model. The histograms are produced from $m = 10000$ replications consisting of $n = 8000$ observations each, corresponding to 32 years with 250 daily observations per year. The true values are $\nu = 2.56$, $\alpha = 64$, $\lambda = 256$, $\mu = 1.2$, $\beta = -0.5$, $\rho = -0.1$. The standard deviations used for the normal curves are taken from the explicit asymptotic results, not estimated.

way its potential usefulness also for practical applications. 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.

More efficient estimators than provided by the optimal quadratic estimating function can be obtained by incorporating further moments. As we provided explicit computations for arbitrary integer moments and conditional

moments, our method allows the extension of the procedure 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 Section 2.3. Although 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 for constructing estimating functions. See [AS02] for diffusions, [Sch05] for Lévy type processes, and [Sin01] for affine models.

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, see also [Tod11].

Finally, we address the issue that volatility, or instantaneous variance, is not directly observable in discrete time. As a practical solution to the problem we suggest to use our estimation framework with the unobservable instantaneous variance replaced by a suitable observable substitute. An attractive candidate could be realized variance, multipower variation, or related measures for non-parametric volatility estimation, see [BNSW06, BNGJ$^+$06, Jac08] and the references therein. From a practical and empirical point of view this approach was implemented and analyzed in [CP10], where the incorporation of information from option prices is also analyzed. They find that the results can be sensitive to the choice of the substitute, in particular when considering different time increments Δ. Power variation of the returns was also used by [Tod09] to perform indirect inference on the volatility. However, the method as it is is a benchmark, implementation in practice will require the use of proxies, for example the popular realized volatility, and although consistency is preserved the rates will change. A detailed analysis of the modified scheme is left open for future research.

As a second approach, we suggest to replace the unobservable instantaneous variance by some measure of trading intensity. In [Lin05] it is reported, that the number of trades is an excellent choice for statistical purposes. For a theoretical analysis, a joint model for the prices and number of trades has to be specified, possibly leading to more parameters describing the relation of number of trades and volatility in a concrete way. In [HP11] implementation issues and some modifications of the framework from the present paper are used, and it is shown that the development of an explicit estimator and its rigorous asymptotic analysis is still feasible. This demonstrates that our

method allows an explicit and tractable estimation procedure for a class of stochastic volatility models with jumps, making the approach suitable for practical applications.

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 anymore, and the bias has to be accounted for in a rigorous analysis.

Appendix A. Explicit moment calculations

This section is about computing explicitly $E[X_1^n V_1^m | V_0 = v]$ 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

$$\zeta = K_1, \qquad \eta = K_2.$$

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. It remains to transform multivariate cumulants to multivariate moments 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 in order to implement them on a computer algebra system, in particular, since the expressions, though completely explicit, are rather lengthy when it comes to evaluating moments of order four for the asymptotic covariance matrix. For the reader's convenience, we give the details in this appendix.

A.1. *Preliminaries.* Let us recapitulate the variables and notation from Section 2.2, that are required in the following calculations. We use

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

We have

(A.1) $$V_1 = \gamma V_0 + U_1, \qquad Y_1 = \epsilon V_0 + S_1$$

where

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

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

(A.2) $$X_1 = A_1 + \sqrt{Y_1} W_1, \qquad A_1 = \mu\Delta + \beta Y_1 + \rho Z_1.$$

A.2. *Stationary moments.* We use the well-known recursion to compute moments from cumulants

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

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

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

with

$$\epsilon_{ij} = \begin{cases} \lambda^{-i}\left(1 + \sum_{k=1}^i \binom{i}{k}(-1)^k \dfrac{1-\gamma^k}{k\lambda}\right), & j = 0, \\[4mm] \lambda^{-i}\left(\dfrac{1-\gamma^j}{j\lambda} + \sum_{k=1}^i \binom{i}{k}(-1)^k \dfrac{1-\gamma^k}{k\lambda}\right), & j > 0. \end{cases}$$

A.4. *Trivariate Moments.* Trivariate moments can be computed recursively from trivariate cumulants

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

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

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

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.1) gives

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

Collecting powers of v gives

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

with

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

Then using (A.2) and conditioning gives

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

Collecting powers of v gives

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

with

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

Finally using (A.2) and the Gaussian moments gives

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

Collecting powers of v gives

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

with

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

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,

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

then

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

and finally

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

REFERENCES

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

[AS09]   T. G. Andersen and N. Shephard, *Stochastic volatility: origins and overview,* in: Handbook of Financial Time Series, Springer-Verlag, 2009, 233–254.

[ASJ07] Y. Aït-Sahalia and J. Jacob, *Volatility estimators for discretely sampled Lévy processes,* Ann. Statist. **35** (2007), 355–392.

[ASJ09] Y. Aït-Sahalia and J. Jacob, *Estimating the degree of activity of jumps in high frequency data,* Ann. Statist. **37** (2009), 2202–2244.

[Bat96] D. Bates, *Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options,* Rev. Financ. Stud. **9** (1996), 69–107.

[Bat00] D. S. Bates, *Post-'87 crash fears in the s&p 500 futures option market,* J. Econometrics **94** (2000), 181–238.

[BG05]   F. E. Benth and M. Groth. *The minimal entropy martingale measure and numerical option pricing for the Barndorff–Nielsen–Shephard stochastic volatility model,* Stoch. Anal. Appl. **27** (2009), 875–896.

[BK05]   F. E. Benth and K. H. Karlsen, *A PDE representation of the density of the minimal entropy martingale measure in stochastic volatility markets,* Stochastics **77** (2005), 109–137.

[BKR03] F. E. Benth, K. H. Karlsen and K. Reikvam, *Merton's portfolio optimization problem in a Black and Scholes market with non-Gaussian stochastic volatility of Ornstein-Uhlenbeck type,* Math. Finance **13** (2003), 215–244.

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

[BNGJ$^+$06]   O. E. Barndorff-Nielsen, S. E. Graversen, J. Jacod, M. Podolskij and N. Shephard, *A central limit theorem for realised power and bipower variations of continuous semimartingales*, in: From stochastic calculus to mathematical finance, Springer, Berlin, 2006, 33–68.

[BNNS02]   O. E. Barndorff-Nielsen, E. Nicolato and N. Shephard, *Some recent developments in stochastic volatility modelling*, Quant. Finance **2** (2002), 11–23.

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

[BNSW06]   O. E. Barndorff-Nielsen, N. Shephard and M. Winkel, *Limit theorems for multipower variation in the presence of jumps*, Stochastic Process. Appl. **116** (2006), 796–806.

[BS95]   B. M. Bibby and M. Sørensen, *Martingale estimation functions for discretely observed diffusion processes*, Bernoulli **1** (1995), 17–39.

[CP10]   B. Jesper Christensen and P. Posedel, *The risk-return tradeoff and leverage effect in a stochastic volatility-in-mean model*, CREATES research paper 2010–50, 2010.

[CT97]   Y. S. Chow and H. Teicher, Probability theory, Springer-Verlag, New York, 1997.

[CT04]   R. Cont and P. 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** (2003), 984–1053.

[DPS00]   D. Duffie, J. Pan and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica **68** (2000), 1343–1376.

[FSS09]   S. Frühwirth-Schnatter and L. Sögner, *Bayesian estimation of stochastic volatility models based on OU processes with marginal Gamma law*, Ann. Inst. Statist. Math. **61** (2009), 159–179.

[GS06]   J. E. Griffin and M. F. J. Steel, *Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility*, J. Econometrics **134** (2006), 605–644.

[Hes93]   S. L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Fin. Studies **6** (1993), 327–343.

[HP07]   F. Hubalek and P. Posedel, Asymptotic analysis for a simple explicit estimator in Barndorff-Nielsen and Shephard stochastic volatility models, Thiele Research Report No.5, 2007.

[HP09]   F. Hubalek and P. Posedel, *Asymptotic analysis for an optimal estimating function for Barndorff-Nielsen-Shephard stochastic volatility models*, work in progress, 2009.

[HP11]   F. Hubalek and P. Posedel, *Joint analysis and estimation of stock prices and trading volume in Barndorff-Nielsen and Shephard stochastic volatility models*, Quant. Finance **11** (2011), 917–932.

[HW87]   J. Hull and A. White, *The pricing of options on assets with stochastic volatilities*, J. Finance **42** (1987), 281–300.

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

[Jac08]   J. Jacod, *Asymptotic properties of realized power variations and related functionals of semimartingales*, Stochastic Process. Appl. **118** (2008), 517–559.

[JKG12]   J. Jacod, C. Klüppelberg and M. Gernot, *Functional relationships between price and volatility jumps and its consequences for discretely observed data*, preprint, TU Münhen, 2012.

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

[JS03]   J. Jacod and A. N. Shiryaev, Limit theorems for stochastic processes, Springer-Verlag, Berlin, second edition, 2003.

[JT10]   J. Jacod and V. Todorov, *Do price and volatility jump together?* Ann. Appl. Probab. **20** (2010), 1425–1469.

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

[KMKV11]   J. Kallsen, J. Muhle-Karbe and M. Voß, *Pricing options on variance in affine stochastic volatility models*, Math. Finance **21** (2011), 627–641.

[KS99]   M. Kessler and M. Sørensen, *Estimating equations based on eigenfunctions for a discretely observed diffusion process*, Bernoulli **5** (1999), 299–314.

[Kut04]   Y. A. Kutoyants, Statistical inference for ergodic diffusion processes, Springer-Verlag, London, 2004.

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

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

[Lew00]   A. L. Lewis, Option valuation under stochastic volatility, Finance Press, Newport Beach, 2000.

[Lin05]   C. Lindberg, Portfolio optimization and statistics in stochastic volatility markets, PhD dissertation, Chalmers University of Technology, 2005.

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

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

[MKPS10]   J. Muhle-Karbe, O. Pfaffel and R. Stelzer, *Option pricing in multivariate stochastic volatility models of OU type*, to appear in SIAM J. Financial Math., `arXiv:1001.3223v1`, 2010.

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

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

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

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

[PR99]   B. L. S. Prakasa Rao, Statistical inference for diffusion type processes, Arnold, London, 1999.

[Raj32]   A. Rajchman, *Zaostrzone prawo wielkich liczb*, Mathesis Polska **6** (1932), 145–161.

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

[RS06]   T. Rheinländer and G. Steiger, *The minimal entropy martingale measure for general Barndorff-Nielsen/Shephard models*, Ann. Appl. Probab. **16**, 2006, 1319–1351.

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

[Sch05]   E. Schaumburg, *Estimation of Markov processes with Lévy type generators*, working paper, 2005.

[She05]   N. Shephard, editor, Stochastic volatility: selected readings, Oxford University Press, 2005.

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

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

[Sø99]   M. Sørensen, *On asymptotics of estimating functions*, Braz. J. Probab. Stat. **13** (1999), 111–136.

[Sø00]   H. Sørensen, Inference for diffusion processes and stochastic volatility models, PhD thesis, University of Copenhagen, 2000.

[Sør97]   M. Sørensen, *Estimating functions for discretely observed diffusions: a review*, in: Selected Proceedings of the Symposium on estimating functions (Athens, GA, 1996), Inst. Math. Statist., Hayward, 1997, 305–325.

[SS91]   E. M. Stein and J. C. Stein, *Stock price distributions with stochastic volatility: an analytic approach*, Rev. Financ. Stud. **4** (1991), 727–752.

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

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

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

[Tod09]   V. Todorov, *Estimation of continuous-time stochastic volatility models with jumps using high-frequency data*, J. Econometrics **148** (2009), 131–148.

[Tod11]   V. Todorov, *Econometric analysis of jump-driven stochastic volatility models*, J. Econometrics **160** (2011), 12–21.

[Ven01]   E. Venardos, Derivatives pricing and Ornstein-Uhlenbeck type stochastic volatility, Dissertation, University of Oxford, 2001.

F. Hubalek
Vienna University of Technology
Financial and Actuarial Mathematics
Wiedner Hauptstraße 8 / 105-1
A-1040 Vienna
Austria
E-mail: `fhubalek@fam.tuwien.ac.at`

P. Posedel
Zagreb School of Economics and Management
Jordanovac 100
10000 Zagreb
Croatia
E-mail: `pposedel@zsem.hr`