Estimation of integrated volatility in stochastic volatility models

Jeannette H.C. Woerner
Institut für Mathematische Stochastik
Universität Göttingen
Maschmühlenweg 8-10, D-37073 Göttingen, Germany
email: woerner@math.uni-goettingen.de

Abstract

In the framework of stochastic volatility models we examine estimators for the integrated volatility based on the p-th power variation (i.e. the sum of p-th absolute powers of the log-returns). We derive consistency and distributional results for the estimators given high frequency data, especially taking into account what kind of process we may add to our model without affecting the estimate of the integrated volatility. This may on the one hand be interpreted as a possible flexibility in modelling, for example adding jumps or even leaving the framework of semimartingales by adding a fractional Brownian motion, or on the other hand as robustness against model misspecification. We will discuss possible choices of p under different model assumptions and irregularly spaced data.

key words and phrases: stochastic volatility, limit theorem, power variation, quadratic variation, semimartingale, jump process, fractional Brownian motion, high frequency data

1 Introduction

In the last years the concept of power variation (i.e. taking $\sum_i |X_{t_i} - X_{t_{i-1}}|^p$ as $\max_i |t_i - t_{i-1}| \to 0$, where X_t denotes the log-price process), as an estimate for the integrated volatility, became popular as a measure for the change in the volatility, because stochastic volatility models play an important role in overcoming the problems of the Black-Scholes world, especially being able to fit skews and smiles. Moreover, volatility derivatives, such as volatility and variance swaps and swaptions, became increasingly attractive to investors. Namely these financial instruments avoid direct exposure to underlying assets, but make it possible to hedge volatility risk. For pricing these derivatives reliable estimators of the integrated volatility based on the discretely observed log-price process are important, cf. Howison et.al. (2003).

The starting point for the use of power variation was made when the link between the mathematical concept of quadratic variation and integrated volatility was established. Contributions include Barndorff-Nielsen and Shephard (2001, 2002a,b, 2003), Corsi, Zumbach, Muller, and Dacorogna (2001), Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001), Andreou and Ghysels (2001), Bai, Russell, and Tiao (2000), Maheu and McCurdy (2001), Areal and Taylor (2001), Galbraith and Zinde-Walsh (2000), Bollerslev and Zhou (2001) and Bollerslev and Forsberg (2001). However, empirically it seems to be more attractive to use absolute values of the returns than squared returns, see for example Andersen and Bollerslev (1997, 1998), Taylor (1986, Ch.2), Cao and Tsay (1992), Ding, Granger, and Engle (1993), West and Cho (1995), Granger and Ding (1995), Jorion (1995), Shiryaev (1999, Ch. IV) and Granger and Sin (1999). Barndorff-Nielsen and Shephard (2002b, 2003) provide the theoretical background to this work in terms of limit theorems for power variations when the underlying data is obtained from a continuous semimartingale of the form $\alpha_t + \int_0^t \sigma_s dB_s$, where $\sigma_t > 0$ and α_t are assumed to be stochastically independent of the Brownian motion B_t and to satisfy some regularity conditions, especially being satisfied for their Ornstein-Uhlenbeck type stochastic volatility model. They also consider the same model when the Brownian motion is replaced by a stable process, cf. Barndorff-Nielsen and Shephard (2002b). From the practical point of view the concept of power variation is very attractive since it is simple and easy to implement, as it only involves certain powers of the log-returns.

We now leave the framework of continuous semimartingale models and examine how we can choose the process $\alpha_t = (\alpha_t, t \geq 0)$ such that it does not affect the estimate for the integrated volatility. This extends results on consistency of these estimators in a semimartingale framework with jumps (cf. Woerner (2003)). For consistency of the estimates we can relax the conditions on $\sigma_t = (\sigma_t, t \geq 0)$ and α_t, as to include most of the well known stochastic volatility models, such as Hull and White (1987), Scott (1987) and Stein and Stein (1991), as well as the Ornstein-Uhlenbeck type model by Barndorff-Nielsen and Shephard (2001). For α_t it turns out that we can add jump processes or even leave the framework of semimartingales by adding a fractional Brownian motion. Furthermore, we may have a correlation between the mean process and the underlying Brownian motion. The possibility of adding different types of processes α_t can be seen as a greater flexibility in modelling. Our framework allows us for example to include the leverage effect in the Ornstein-Uhlenbeck type model by adding a jump component built of the adjusted subordinator which drives σ_t, as suggested by Barndorff-Nielsen and Shephard (2001). Furthermore, our model allows us to include fractional Brownian motion with Hurst parameter $H \in (1/2, 1)$. Since the covariance function between increments of this process is positive, this is especially interesting for modelling clustering effects of the data, cf. Shiryaev (1999, pp. 232).

In addition, examining what kind of processes we can add to the stochastic volatility model can also be interpreted as examining the robustness of the estimator of the integrated volatility, namely what kind of noise we can add without affecting the estimate.

For the distributional theory our conditions are a bit more restrictive than for consistency, hence in some cases we obtain that adding a certain process does not influence the consistency of the estimator of the integrated volatility but the spread around the true integrated volatility. This means for practical purposes, when we have a model misspecification our error bounds for the estimate will no longer be valid.

We will examine what the best choice for p is, both in terms of the process α_t and in terms of the irregularity of the sampling scheme. We do not require equally spaced data, which is hardy available in practice, but only some mild regularity in terms of the ratio of the decay of the maximal and minimal distance of observations.

Furthermore, our results give some insight, why empirically stochastic volatility estimates perform better, when using absolute values of returns rather than quadratic variation. Namely, when we assume a continuous semimartingale model, but the data involves some jump component, then the quadratic variation possesses an additional unexpected term coming from the jumps. Hence we not only get an estimate for the integrated volatility, but for the integrated volatility plus some extra term. Whereas when using absolute values and the correct norming as for a continuous semimartingale model, the continuous part is dominating the jumps and we get an estimate of the integrated volatility, even when our model assumption was not correct. However, taking absolute values, only works when the jump component has at most as much activity as a bounded variation process. However, for the commonly used financial models this is satisfied since they mostly add jumps derived from a compound Poisson process (i.e. a process with finite activity), or from a subordinator (i.e. a process with bounded variation). If the jump component possesses more activity, an alternative is to choose some exponent of the variation lying between one and two, since for values strictly less than two, the continuous part is still dominating and the jump part is negligible. However, this only describes the setting for consistency. For the distributional theory to hold we need an exponent less than one, hence possibly a mean process of less activity than a bounded variation process.

The paper is organized as follows. First we will introduce our notation and the stochastic volatility models. Then we prove consistency for the estimate of the integrated volatility, discussing the conditions on the mean process α_t and the choice of p. Finally we discuss the distributional theory for the estimate of the integrated volatility and suitable choice of p.

2 Models and Notation

The concept of variational sums and power variation was introduced in the context of studying the path behaviour of stochastic processes in the 1960ties, cf. Berman (1965), Hudson and Tucker (1974), Hudson and Mason (1976) for additive processes or Lepingle (1976) for semimartingales. Assume that we are given a stochastic process X on some finite time interval $[0, t]$. Let n be a positive integer and denote by $S_n = \{0 = t_{n,0}, t_{n,1}, \cdots, t_{n,n} = t\}$ a partition of $[0,t]$, such that $0 < t_{n,1} < t_{n,2<\cdots} < t_{n,n}$ and $\max_{1\leq k\leq n}\{t_{n,k} - t_{n,k-1}\} \to 0$ as $n \to \infty$. Now the p-th power variation is defined to be

$$\sum_i |X_{t_{n,i}} - X_{t_{n,i-1}}|^p = V_p(X, S_n).$$

We are interested in the limit as $n \to \infty$. Well established are for convergence in probability the cases for $p = 1$, where finiteness of the limit means that the processes has bounded variation, and $p = 2$, called quadratic variation, which is finite for all semimartingale processes. However, in the stochastic volatility setting, for the moment assuming that our process has the form $\int_0^t \sigma_s dB_s$, only the case $p = 2$ leads to a non-trivial limit. Obviously, for $p > 2$ the limit is zero and for $p < 2$ the limit is infinity.

An extension of the concept of power variation is to introduce an appropriate norming sequence, as it was done in Barndorff-Nielsen and Shephard (2003), which allows to find non-trivial limits even in the cases where the non-normed power variation limit would be zero or infinity. Let us introduce the following notation for the normed p-th power variation

$$\sum_i \Delta_{n,i}^{\gamma} |X_{t_{n,i}} - X_{t_{n,i-1}}|^p = V_p(X, S_n, \Delta_n^{\gamma}),$$

where $\gamma \in I\!R$ and $t_{n,i} - t_{n,i-1} = \Delta_{n,i}$ denotes the distance between the i-th and the $i-1$-th time-point. When we have equally spaced observations, $\Delta_{n,i}$ is independent of i and the normed power variation reduces to $\Delta_n^{\gamma} V(X, S_n)$.

In the following we need a measure of the regularity for the sequence of partitions. We use the term of ϵ-balanced partitions, $\epsilon \in (0,1)$, which was introduced by Barndorff-Nielsen and Shephard (2002b) and is defined by

$$\frac{\max_i \Delta_{n,i}}{(\min_i \Delta_{n,i})^{\epsilon}} \to 0,$$

as $n \to \infty$. This means we compare how fast the minimum distance in the partition converges to zero compared with the sequence of maxima, for example if $\max_i \Delta_{n,i} = O(1/n)$ and $\min_i \Delta_{n,i} = O(1/n^2)$, then the partition is ϵ-balanced for $\epsilon \in (0, 1/2)$. Obviously, an equally spaced partition is ϵ-balanced for all $\epsilon \in (0,1)$. If a partition is ϵ-balanced for some $\epsilon \in (0,1)$, then it is also δ-balanced for $\delta \in (0, \epsilon]$. Hence the larger ϵ the closer the partition is to an equally-spaced one.

Let us now briefly review the stochastic processes which we will need in the following. We start with a general semimartingale process X_t, which is widely used in finance. For an overview both under financial and theoretical aspects see Shiryaev (1999). In its canonical representation a semimartingale may be written as

$$X_t = X_0 + B(h) + X^c + h * (\mu - \nu) + (x - h(x)) * \mu,$$

or for short with the predictable characteristic triplet $(B(h), \langle X^c \rangle, \nu)$, where X^c denotes the continuous local martingale component, $B(h)$ is predictable of bounded variation and h is a truncation function, behaving like x around the origin. Furthermore, $\mu((0,t] \times A; \omega) = \sum (I_A(J(X_s)), 0 < s \leq t)$, where $J(X_s) = X_s - X_{s-}$ and $A \in \mathcal{B}(I\!R - \{0\})$, is a random measure, the jump measure, and ν denotes its compensator, satisfying $(x^2 \wedge 1) * \nu \in \mathcal{A}_{loc}$ (i.e. the process $(\int_{(0,t] \times I\!R} (x^2 \wedge 1) d\nu)_{t \geq 0}$ is locally integrable). Semimartingale models include the well-established continuous diffusions, jump-diffusions, hence stochastic volatility models, as well as Lévy processes.

Lévy processes are a special class of semimartingales where we have independent and stationary increments. They are given by the characteristic function via the Lévy-Khichin formula

$$E[e^{iuX_t}] = \exp\{t(i\alpha u - \frac{\sigma^2 u^2}{2} + \int (e^{iux} - 1 - iuh(x))\nu(dx))\},$$

where α denotes the drift, σ^2 the Gaussian part and ν the Lévy measure. Hence σ^2 determines the continuous part and the Lévy measure the frequency and size of jumps. If $\int (1 \wedge |x|)\nu(dx) < \infty$ the process has bounded variation. If $\int \nu(dx) < \infty$ the process jumps only finitely many times in any finite time-interval, called finite activity, it is a compound Poisson process. Furthermore, the support of ν determines the size and direction of jumps. A popular example in finance are subordinators, where the the support of the Lévy measure is restricted to the positive half line, hence the process is of bounded variation. For more details see Sato (1999).

A measure for the activity of the jump component of a semimartingale is the generalized Blumenthal-Getoor index,

$$\beta = \inf\{\delta > 0 : (|x|^\delta \wedge 1) * \nu \in \mathcal{A}_{loc}\},$$

where $\mathcal{A}_{loc}$ is the class of locally integrable processes. This index β also determines, that for $p > \beta$ the sum of the p-th power of jumps will be finite. Note that if we are in the framework of Lévy processes, being an element of a locally integrable process reduces to finiteness of the integral with respect to the Lévy measure, since the jump measure is deterministic.

A fractional Brownian motion with Hurst exponent $H \in (0,1)$ is defined by

$$\frac{1}{\Gamma(H+1/2)}(\int_{-\infty}^{0} ((t-s)^{H-1/2} - (-s)^{H-1/2})dW_s + \int_0^t (t-s)^{H-1/2} dW_s),$$

where $(W_s, -\infty < s < \infty)$ denotes a Wiener process extended to the real line. A fractional Brownian motion does not belong to the class of semimartingales, however it is used in finance to model clustering and long range dependence, cf. Shiryaev (1999).

Let us now introduce the stochastic volatility models. In the Black and Scholes framework the logarithm of an asset price X_t is modelled as a geometric Brownian motion or as the solution of the stochastic differential equation

$$dX_t = (\mu + \beta\sigma^2)dt + \sigma dB_t,$$

where μ, β and σ are constants. One possibility of overcoming the problems of the Black-Scholes framework and capturing the empirical facts of excess kurtosis, skewness, fat tails and volatility smiles, is to introduce a random spot volatility process $\sigma_t = (\sigma_t, t \geq 0)$ leading to the simplest case of a stochastic volatility model. Now the logarithm of an asset price X_t is modelled as the solution to the following diffusion equation

$$dX_t = (\mu + \beta\sigma_t^2)dt + \sigma_t dB_t, \tag{1}$$

where σ_t is assumed to satisfy a second stochastic differential equation. Transforming (1) to an integrated form leads to

$$X_t = \mu t + \beta \int_0^t \sigma_s^2 ds + \int_0^t \sigma_s dB_s,$$

up to a constant, or in a more general formulation

$$X_t = \alpha_t + \int_0^t \sigma_s dB_s,$$

where $\alpha_t = (\alpha_t, t \geq 0)$ is some stochastic process. Traditionally, it is assumed that the mean process α_t is of locally bounded variation, but in our framework we do not need this restriction, we can even allow that the process is not a semimartingale such as a fractional Brownian motion.

The main differences between the various stochastic volatility models lie in the stochastic differential equation the spot volatility process is assumed to satisfy. We will recall different examples, which our estimating results can be applied to. Note that the model by Heston (1993) does not satisfy our assumptions, since we need that σ_t and B_t are independent, whereas in Heston (1993) B_t and the Brownian motion which drives σ_t are correlated.

Example:

Assume that the price process can be described by the following diffusion equation

$$dX_t = \mu X_t dt + \sigma_t X_t dB_t$$

Then Hull and White (1987) model σ_t^2 by a geometric Brownian motion,

$$d\sigma_t^2 = \alpha\sigma_t^2 dt + \chi\sigma_t^2 dW_t,$$

where W_t is a Brownian motion independent of B_t, α and χ are some constants. Scott (1987) and Stein and Stein (1991) model σ_t by an Ornstein-Uhlenbeck process,

$$d\sigma_t = -\delta(\sigma_t - \theta)dt + k dW_t,$$

where again W_t is a Brownian motion independent of B_t and δ, θ and k are some constants.

Barndorff-Nielsen and Shephard (2001) model σ_t^2 by an Ornstein-Uhlenbeck type process of the form

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + dZ_{\lambda t}.$$

Here Z_t is a subordinator without drift, independent of the Brownian motion B_t. The time scale λt is chosen to ensure that the marginal law of σ_t^2 is not affected by the choice of λ. Note that though σ_t^2 exhibits jumps, X_t is still continuous.

3 Estimating integrated volatility

In the context of stochastic volatility models in practice neither the structure of the underlying spot volatility process is known nor the process is observed continuously. This makes it difficult to infer the volatility, which is necessary for option pricing, hedging and risk assessment.

The concept of power variation provides a solution to this problem. It makes it possible to infer the integrated volatility in a simple way, using irregularly spaced high frequency data, only by assuming mild regularity assumptions on the spot volatility and the mean process, which are satisfied by the most popular models .

Under a suitable choice of the power exponent p, a large class of mean processes turns out to be negligible in the estimating procedure of the integrated volatility. This on the one hand provides us with more flexibility in modelling. Namely, we can add jump components as for example desirable to include the leverage effect in the Ornstein-Uhlenbeck type model (cf. Barndorff-Nielsen and Shephard (2001)), or include fractional Brownian motion to model clustering effects. On the other hand allowing to add processes to our original stochastic volatility model can be viewed as adding noise. Hence we can determine how robust our estimator is under a misspecification of the model. As the simplest case take $p = 2$, then we are in the case of quadratic variation. It is obvious that our estimating procedure is not robust against jumps, since in this case we get a quadratic variation part estimating the integrated volatility and one extra part belonging to the jump component. We will discuss in the following how the choice of p affects the robustness of the estimate in different settings.

The following theorem is an extension of Barndorff-Nielsen and Shephard (2002b, 2003) and Woerner (2003) allowing irregularly spaced data as well as a larger class of spot volatility processes and mean processes. Especially we can include jumps and not necessarily have to stay in the framework of semimartingales. Furthermore, the mean process may be correlated with both the spot volatility process and the Brownian motion.

Theorem 1 *Let*

$$X_t = Y_t + \int_0^t \sigma_s dB_s, \tag{2}$$

where Y_t is some stochastic process satisfying for some $p > 0$

$$V_p(Y, S_n, \Delta_n^{1-p/2}) \xrightarrow{p} 0. \tag{3}$$

Furthermore, assume that the volatility process σ_t^2 is independent of the Brownian motion B_t and a.s. locally Riemann-integrable. Then for any $t > 0$ and for any sequence of $1/2$-balanced partitions S_n, we obtain

$$V_p(X, S_n, \Delta_n^{1-p/2}) \xrightarrow{p} \mu_p \int_0^t \sigma_s^p ds, \tag{4}$$

as $n \to \infty$, where $\mu_p = E(|u|^p) = \frac{2^{p/2}\Gamma(\frac{p+1}{2})}{\Gamma(1/2)}$ and $u \sim N(0,1)$.

Proof. The idea of the proof is to use the result for the stochastic volatility component derived by Barndorff-Nielsen and Shephard (2002b) together with a triangular inequality for $p \le 1$ and Minkowski's inequality for $p > 1$. Let us denote the stochastic volatility part by $Z_t = \int_0^t \sigma_s dB_s$.

First we show that we do not need all conditions stated in Theorem 5.3 of Barndorff-Nielsen and Shephard (2002b) to establish the convergence of the stochastic volatility component, but Riemann integrability is sufficient. As $n \to \infty$, we obtain

$$\begin{aligned}
&|\mu_p^{-1} V_p(Z, S_n, \Delta_n^{1-\frac{p}{2}}) - \int_0^t \sigma_s^p ds| \\
\le\ & |\mu_p^{-1} V_p(Z, S_n, \Delta_n^{1-\frac{p}{2}}) - \sum_j \Delta_{n,j}^{1-\frac{p}{2}} |\int_{t_{n,j-1}}^{t_{n,j}} \sigma_s^2 ds|^{p/2}| \\
& + |\sum_j \Delta_{n,j}^{1-\frac{p}{2}} |\int_{t_{n,j-1}}^{t_{n,j}} \sigma_s^2 ds|^{p/2} - \int_0^t \sigma_s^p ds| \\
\to\ & 0.
\end{aligned} \tag{5}$$

The first term tends to zero by equation (5.17) in Barndorff-Nielsen and Shepard (2002c), noting that the denominator in (5.17) tends to zero. The second term tends to zero by (5.21) in Barndorff-Nielsen and Shephard (2002b) using the Riemann-integrability.

Now for $p \le 1$ we obtain

$$\begin{aligned}
& P(|V_p(X, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds| > \lambda) \\
\le\ & P(|V_p(X, S_n, \Delta_n^{1-p/2}) - V_p(Z, S_n, \Delta_n^{1-p/2})| > \lambda/2) \\
& P(|V_p(Z, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds| > \lambda/2) \\
\le\ & P(V_p(Y, S_n, \Delta_n^{1-p/2}) > \lambda/2) \\
& P(|V_p(Z, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds| > \lambda/2) < \epsilon,
\end{aligned}$$

since $V_p(Y, S_n, \Delta_n^{1-p/2}) \xrightarrow{p} 0$ by assumption (3) and $V_p(Z, S_n, \Delta_n^{1-p/2}) \xrightarrow{p} \mu_p \int_0^t \sigma^p(s) ds$ by (5). For the second inequality we used that for $p \le 1$ applying $|a + b|^p \le$

$|a|^p + |b|^p$, we have

$$|\sum_i |a_i + b_i|^p - \sum_i |b_i|^p| \leq \sum_i |a_i|^p.$$

For $p > 1$ we use Minkowski's inequality together with the same technique, which yields

$$\begin{aligned} & P(|(V_p(X, S_n, \Delta_n^{1-p/2}))^{1/p} - (\mu_p \int_0^t \sigma_s^p ds)^{1/p}| > \lambda) \\ \leq \quad & P((V_p(Y, S_n, \Delta_n^{1-p/2}))^{1/p} > \lambda/2) \\ & P(|(V_p(Z, S_n, \Delta_n^{1-p/2}))^{1/p} - (\mu_p \int_0^t \sigma_s^p ds)^{1/p}| > \lambda/2) < \epsilon. \end{aligned}$$

This implies the desired result since the function $f(x) = x^p$ is continuous. □

Now we discuss the condition on the process Y_t. Condition (3) provides lots of flexibility in choosing Y_t. This can be interpreted as for what kind of mean process our estimate is still valid or how robust our estimate is against misspecification of the model, namely what type of noise we can add without affecting the estimate. In the following we look at the conditions on Y_t:

- Condition (3) is satisfied when Y_t is locally Hölder-continuous of the order $1/2+\gamma$, $\gamma > 0$, since this implies that $(Y_{n,j} - Y_{n,j-1})/\Delta_{n,j}^{(1/2)+\gamma} \leq C_j < \infty$ and hence for $p > 0$

$$\begin{aligned} V_p(Y, S_n, \Delta_n^{1-p/2}) &= \sum_{j=1}^n \Delta_{n,j}^{1+\gamma p} |\frac{Y_{n,j} - Y_{n,j-1}}{\Delta_{n,j}^{(1/2)+\gamma}}|^p \\ &\leq (\max_j C_j)(\max_j \Delta_{n,j}^{\gamma p})t \to 0, \end{aligned}$$

  as $n \to \infty$. This is less restrictive than the condition used in Barndorff-Nielsen and Shephard (2003). We can see that we can even leave the framework of semimartingales and include models or noise where Y_t is a fractional Brownian motion with Hurst exponent $H \in (1/2, 1]$. We only need slightly more regularity in the paths than for Brownian motion.

- Condition (3) is also satisfied if we have finite p-th power variation of Y_t and take $p < 2$, which means that the norming sequence tends to zero. Hence for all processes Y_t with bounded variation, which is the standard setting of a stochastic volatility model, we can always take $p = 1$.

  For jump processes this is also very interesting since we can easily determine for which r we have finite r-th power variation and can then take $p \in [r, 2)$. From Woerner (2003) we know that assuming a general semimartingale setting, the r-th power variation is finite if either $1 \leq \beta < r < 2$ and $\langle Y^c \rangle_t = 0$ or $\beta < r \leq 1$, $\langle Y^c \rangle_t = 0$, $B(h) + (x-h) * \nu = 0$ and the jump times of Y_t are previsible. Here β denotes the generalized Blumenthal-Getoor index and $\langle Y^c \rangle_t$ denotes the quadratic variation of the continuous local martingale component of Y_t. Note that for subordinators or Lévy processes of bounded variation in their usual representation with $h(x) = x$, the condition $B(h) + (x - h) * \nu = 0$ reduces to no drift.

- Even less restrictive than the previous condition is to assume that

$$V_p(Y, S_n, \Delta_n^\gamma) \xrightarrow{p} C < \infty,$$

where $\gamma > 0$ and $1 - (p/2) - \gamma > 0$ which implies (3). One example in which this holds is to take

$$Y_t = \int_0^t f_s dZ_s,$$

where Z_t is a symmetric α-stable process, f_t is independent of Z_t and locally Riemann integrable. Then by Barndorff-Nielsen and Shephard (2002b), for $p < \alpha$

$$\Delta_n^{1-p/\alpha} V_p(Y, S_n) \xrightarrow{p} \mu_{\alpha,p} \int_0^t f_s^p ds,$$

where $\mu_{\alpha,p} = E(|Z(1)|^p)$. Here $1 - (p/2) - (1 - p/\alpha) > 0$ is obviously satisfied since $p > 0$ and $\alpha \in (0, 2)$.

Taking $f(x) = 1$ this reduces to an symmetric α-stable process, in which case the Blumenthal-Getoor index is α and hence together with the previous consideration (3) is satisfied for all $p \in (0, 2)$

- Using the same method as in the proof we can see that Y_t obviously also satisfies (3), if it is the sum of processes, each of which satisfies (3).

Now we can discuss the choice of p. Traditionally, $p = 2$, the quadratic variation case, is used and we can see it works quite well, as long as we do not have a jump component. This means the quadratic variation can be used as an estimate for the integrated volatility when the mean process is Hölder continuous of the order $(1/2) + \gamma$. The absence of jumps can be checked by calculating $\lim_{n\to\infty} V_p(X, S_n)$ for $p > 2$, cf. Woerner (2003). If the limit is greater than zero then our data possesses a jump component. However, since we square all differences of data-points in the quadratic variation case, all outliers are weighted quite strongly and from a practical point of view it might be better to choose a smaller exponent p.

As we have seen in the discussion of the jump component, taking $p = 1$ might be a good choice for p as it was also suggested by empirical studies, cf. for example Andersen and Bollerslev (1997, 1998). This makes jumps negligible when they are derived from a process with Blumenthal-Getoor index $\beta \leq 1$. This is satisfied by most financial models since they either use a compound Poisson process or a subordinator. Of course from the point of view of down-weighting outliers it is desirable to choose p as small as possible, whereas choosing p close to 2 capture the widest range of possible jump activities.

The condition on σ_t is quite mild. It is implied by the continuity of most volatility processes, as for example Hull and White (1987), Scott (1987) and Stein and Stein (1991). However, Heston (1993) is not included since in this model σ_t and B_t are correlated. Furthermore, the condition is also satisfied by the jump process model, where σ_t is an Ornstein-Uhlenbeck type process, as it is shown in Barndorff-Nielsen and Shephard (2003).

As a special case we obtain the result for equally spaced partitions. Condition (3) reduces to

$$\Delta_n^{1-p/2} V_p(Y, S_n) \xrightarrow{p} 0.$$

For any $t > 0$ and equally spaced partitions S_n we obtain

$$\Delta_n^{1-p/2} V_p(X, S_n) \xrightarrow{p} \mu_p \int_0^t \sigma_s^p ds, \tag{6}$$

as $n \to \infty$, where $\mu_p = E(|u|^p)$ and $u \sim N(0, 1)$.

We can summarize our result in to following table which shows us how we can choose p depending on the components of our mean process. Note that it does not have any influence if the mean process is correlated to or independent of the underlying Brownian motion.

components of mean process			p
Hölder continuous	jumps with index β	α-stable	
yes	no	no	$p > 0$
yes	yes	no	$\beta < p < 2$
yes	no	yes	$0 < p < 2$
no	yes	no	$\beta < p < 2$
no	no	yes	$0 < p < 2$

4 Distributional Theory

In the previous section we derived consistency for estimators of the integrated volatility, however this does not provide the rate of convergence or confidence intervals. As in practice even for high frequency data there is always some distance between the observations, we need a distributional theory to establish error bounds. It turns out that the asymptotic distribution is a normal variance mixture, where the variance is distributed up to a constant as the corresponding $2p$ power variation limit. Hence theoretically the limit theory is feasible. However, we need some restrictions on the mean and the spot volatility process as well as on the sampling scheme. Unfortunately we derive that if we consider the general case where the mean process might be correlated with the underlying Brownian motion B_t only the influence of a purely continuous or a purely discontinuous but not a mixture of both is negligible. However, if we assume that the continuous part of the mean process is independent of B_t, as discussed for the equally spaced sampling scheme in Barndorff-Nielsen and Shephard (2003), the whole mean process including a possible correlated jump component is negligible under a suitable choice of p. An example where these conditions are satisfied is the Ornstein-Uhlenbeck type stochastic volatility model capturing the leverage effect.

The following theorem generalizes the result of Barndorff-Nielsen and Shephard (2002b, 2003) to a large class of mean processes, especially including jumps and fractional Brownian motion, and generalizing the part with the independent continuous mean process to irregularly spaced data.

Theorem 2 *Let*

$$X_t = Y_t + \int_0^t \sigma_s dB_s, \tag{7}$$

where $Y_t = Y_t^{(1)} + Y_t^{(2)}$ *is some stochastic process satisfying for some* $p > 0$

$$V_{2p}(Y, S_n, \Delta_n^{1-p}) \xrightarrow{p} 0 \tag{8}$$

$$\frac{V_p(Y^{(1)}, S_n, \Delta_n^{1-p/2})}{\sqrt{\min_i \Delta_{n,i}}} \xrightarrow{p} 0 \tag{9}$$

$$\frac{V_{2p}(Y^{(1)}, S_n, \Delta_n^{2-p})}{\min_i \Delta_{n,i}} \xrightarrow{p} 0, \tag{10}$$

$$\limsup_{n\to\infty} \max_i \frac{Y_{t_{n,i}}^{(2)} - Y_{t_{n,i-1}}^{(2)}}{\Delta_{n,i}^a} \leq c < \infty, \tag{11}$$

where $a \in (1/2, 1]$ *and* $Y_t^{(2)}$ *is independent of* B_t. *Assume that* $\sigma_t > 0$ *is independent of* B_t*, locally Riemann integrable, pathwise bounded away from zero and has the property that for some* $\gamma > 0$ *and* $n \to \infty$

$$\frac{1}{\sqrt{\min_j \Delta_{n,j}}} \sum_{j=1}^n \Delta_{n,j} |\sigma^\gamma(\eta_{n,j}) - \sigma^\gamma(\chi_{n,j})| \xrightarrow{p} 0 \tag{12}$$

for any $\chi_{n,j}$ *and* $\eta_{n,j}$ *such that*

$$0 \leq \chi_{n,1} \leq \eta_{n,1} \leq t_{n,1} \leq \chi_{n,2} \leq \eta_{n,2} \leq t_{n,2} \cdots \leq \chi_{n,n} \leq \eta_{n,n} \leq t.$$

Then for any $t > 0$ *and for any sequence of* $\max(2/3, 1/(4a-2))$*-balanced partitions* S_n*, we obtain*

$$\frac{V_p(X, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds}{\sqrt{\mu_{2p}^{-1} \nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} \xrightarrow{\mathcal{D}} N(0,1), \tag{13}$$

as $n \to \infty$*, where* $\mu_p = E(|u|^p)$ *and* $\nu_p = Var(|u|^p)$ *with* $u \sim N(0,1)$.

Proof. We use the same notation as in Theorem 1, namely $Z_t = \int_0^t \sigma_s dB_s$. The idea of the proof is to extend the result for Z_t given in Barndorff-Nielsen and Shephard (2002b) to the general X_t by applying Slutzki's Lemma and combine it with an extension of Barndorff-Nielsen and Shephard (2003) to include a wider range of mean processes independent of the underlying Brownian motion. We use the following reformulation

$$\frac{V_p(X, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} = \frac{V_p(X, S_n, \Delta_n^{1-p/2}) - V_p(Z, S_n, \Delta_n^{1-p/2})}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} + \frac{V_p(Z, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(Z, S_n, \Delta_n^{2-p})}} \frac{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(Z, S_n, \Delta_n^{2-p})}}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}}.$$

This leads to the desired result if we can show

$$\frac{V_p(X, S_n, \Delta_n^{1-p/2}) - V_p(Z + Y^{(2)}, S_n, \Delta_n^{1-p/2})}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} + \frac{V_p(Z + Y^{(2)}, S_n, \Delta_n^{1-p/2}) - V_p(Z, S_n, \Delta_n^{1-p/2})}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} \xrightarrow{p} 0 \quad (14)$$

$$\frac{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(Z, S_n, \Delta_n^{2-p})}}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(Z + Y^{(2)}, S_n, \Delta_n^{2-p})}} \frac{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(Z + Y^{(2)}, S_n, \Delta_n^{2-p})}}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} \xrightarrow{p} 1. \quad (15)$$

Recall from Barndorff-Nielsen and Shephard (2002b) Theorem 5.3 that under our assumptions

$$\frac{V_p(Z, S_n, \Delta_n^{1-p/2}) - \mu_p \int_0^t \sigma_s^p ds}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(Z, S_n, \Delta_n^{2-p})}} \xrightarrow{\mathcal{D}} N(0,1). \quad (16)$$

For the proof of the first part of (14) we can use the same technique as in Theorem 1, noting that for $p \leq 1$

$$\frac{|V_p(X, S_n, \Delta_n^{1-p/2}) - V_p(Z + Y^{(2)}, S_n, \Delta_n^{1-p/2})|}{\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})}} \leq \frac{V_p(Y^{(1)}, S_n, \Delta_n^{1-p/2})}{\sqrt{\min_i \Delta_{n,i}}\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{1-p})}},$$

where

$$\frac{V_p(Y^{(1)}, S_n, \Delta_n^{1-p/2})}{\sqrt{\min_i \Delta_{n,i}}} \quad \xrightarrow{p} \quad 0,$$

$$\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{1-p})} \quad \xrightarrow{p} \quad \sqrt{\nu_p \int_0^t \sigma_s^{2p} ds}$$

as $n \to \infty$ by the assumptions together with Theorem 1.

For $p > 1$ we have to use the same argument combined with Minkowski's inequality

$$\begin{aligned} & \frac{|(V_p(X, S_n, \Delta_n^{1-p/2}))^{1/p} - (V_p(Z + Y^{(2)}, S_n, \Delta_n^{1-p/2}))^{1/p}|}{(\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{2-p})})^{1/p}} \\ \leq \quad & \frac{(V_p(Y^{(1)}, S_n, \Delta_n^{1-p/2}))^{1/p}}{(\sqrt{\min_i \Delta_{n,i}}\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{1-p})})^{1/p}}. \end{aligned}$$

For the second part of (14) we use the same technique as in Barndorff-Nielsen and Shephard (2003) extended to an irregularly spaced sampling scheme and a relaxed condition (11). The main idea is to use the fact that the mean process $Y_t^{(2)}$ is independent of the underlying Brownian motion B_t. We can express

$$V_p(Z+Y^{(2)}, S_n, \Delta_n^{1-p/2}) - V_p(Z, S_n, \Delta_n^{1-p/2}) = \sum_{i=1}^n \Delta_{n,i}\theta_i^{p/2} h_p(u_i, \gamma_i/\theta_i^{1/2}, \Delta_{n,i}^{a-1/2}),$$

where

$$h_p(u_i, \gamma_i/\theta_i^{1/2}, \Delta_{n,i}^{a-1/2}) = |\frac{\gamma_i}{\theta_i^{1/2}}\Delta_{n,i}^{a-1/2} + u_i|^p - |u_i|^p.$$

The notation is as in Barndorff-Nielsen and Shephard (2003), that is, u_i denotes a sequence of iid (independent identically distributed) standard normal random variables, $\gamma_i = (Y_{t_{n,i}}^{(2)} - Y_{t_{n,i-1}}^{(2)})/\Delta_{n,i}^a$ and $\theta_i = (\sigma_{t_{n,i}}^2 - \sigma_{t_{n,i-1}}^2)/\Delta_{n,i}$. With exactly the same procedure as in Barndorff-Nielsen and Shephard (2003) it can be shown that by the LLN (law of large numbers)

$$\sum_{i=1}^n \Delta_{n,i}^{1-b} h_p(u_i, \gamma_i/\theta_i^{1/2}, \Delta_{n,i}^{a-1/2}) \xrightarrow{p} 0, \tag{17}$$

provided $1 - b + 2a - 1 > 1$, hence $b < 2a - 1$.

We can rewrite the second part of (14) to

$$\frac{(\max_i \Delta_{n,i})^b \sum_{i=1}^n \Delta_{n,i}^{1-b}\theta_i^{p/2} h_p(u_i, \gamma_i/\theta_i^{1/2}, \Delta_{n,i}^{a-1/2})}{\sqrt{\min_i \Delta_{n,i}}\sqrt{\mu_{2p}^{-1}\nu_p V_{2p}(X, S_n, \Delta_n^{1-p})}}.$$

Hence for an ϵ-balanced partition, $\epsilon > 1/(4a-2)$, condition (17) is satisfied which implies the desired result.

Finally, it remains to prove (15). Equivalently, we can show that

$$\begin{aligned} \frac{|\mu_{2p}^{-1}\nu_p V_{2p}(Z+Y^{(2)},S_n,\Delta_n^{2-p}) - \mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{2-p})|}{\mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{2-p})} &\xrightarrow{p} 0 \\ \frac{|\mu_{2p}^{-1}\nu_p V_{2p}(Z,S_n,\Delta_n^{2-p}) - \mu_{2p}^{-1}\nu_p V_{2p}(Z+Y^{(2)},S_n,\Delta_n^{2-p})|}{\mu_{2p}^{-1}\nu_p V_{2p}(Z+Y^{(2)},S_n,\Delta_n^{2-p})} &\xrightarrow{p} 0. \end{aligned}$$

Now, for the first part we can proceed similarly as for the first part of (14) and hence only look at the case $p \le 1$ since $p > 1$ can be derived analogously

$$\begin{aligned} &\frac{|\mu_{2p}^{-1}\nu_p V_{2p}(Z+Y^{(2)},S_n,\Delta_n^{2-p}) - \mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{2-p})|}{\mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{2-p})} \\ &\le \frac{|\mu_{2p}^{-1}\nu_p V_{2p}(Y^{(1)},S_n,\Delta_n^{2-p})|}{\mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{2-p})} \\ &\le \frac{|\mu_{2p}^{-1}\nu_p V_{2p}(Y^{(1)},S_n,\Delta_n^{2-p})|}{\min_i \Delta_{n,i}\mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{1-p})} \end{aligned}$$

The second part can also be shown similarly as the second part of (14), namely it can be reformulated to

$$\frac{(\max_i \Delta_{n,i})^b \sum_{i=1}^n \Delta_{n,i}^{2-b}\theta_i^p h_{2p}(u_i,\gamma_i/\theta_i^{1/2},\Delta_{n,i}^{a-1/2})}{\min_i \Delta_{n,i}\mu_{2p}^{-1}\nu_p V_{2p}(X,S_n,\Delta_n^{1-p})} \xrightarrow{p} 0,$$

which is satisfied by (17) and the condition on the partition.

Piecing together (14), (15) and (16) leads to the desired result. □

Let us discuss the conditions on Y_t first. The conditions are stronger than for Theorem 1, especially since we need that (8) is satisfied for $2p$. We can distinguish two sets of conditions, one which is stronger, when the mean process may be correlated to the underlying Brownian motion and one where the key is to assume that the mean process is independent of the underlying Brownian motion.

In addition, we have two parameters which determine how to choose p, namely the balance coefficient ϵ, describing the regularity of the sampling scheme, and the Hölder coefficient or Blumenthal-Getoor index describing the regularity of the sample paths of $Y_t^{(1)}$. It turns out the more regular the sampling scheme the less regularity we need in the sample paths and vice versa.

- Similarly as for condition (3), condition (8) is satisfied if $Y_t^{(1)}$ is Hölder continuous with exponent $1/2+\gamma$, $\gamma > 0$. Clearly it only makes sense to take $\gamma \in (0, 1/2]$, since if $\gamma > 1/2$ this would imply that $Y_t^{(1)}$ is constant.

  Condition (9) is satisfied, if $\max_j \Delta_{n,j}^{p\gamma}/\min_j \Delta_{n,j}^{1/2} \to 0$, hence if S_n is $1/(2p\gamma)$-balanced. Hence we must satisfy at least $1 > 1/(2p\gamma)$ or $p\gamma > 1/2$, which is however only for the equally spaced case. Taking into account the range of γ, we see that $p > 1$. If we have an ϵ-balanced partition $\epsilon \in [2/3, 1)$, we need $\epsilon \geq 1/(2p\gamma)$, hence

$$p\gamma \geq \frac{1}{2\epsilon} \tag{18}$$

The relation (18) allows us to compute the possible range p, when we know the regularity of the paths and the regularity of the partition. The more regularity we have in the sample paths the closer we can choose p to $1/\epsilon$.

Finally (10) holds if (9) is satisfied since this yields $\max_j \Delta_{n,j}^{1+p\gamma} / \min_j \Delta_{n,j} \to 0$ as well.

- Similarly to the discussion for Theorem 1, for (8) to hold, we need $2p < 2$ hence $p < 1$ which means more regularity than bounded variation to get finite $2p$-th power variation. Furthermore, (8) is satisfied if we have finite p-th power variation and $\max_i \Delta_{n,i}^{1-p/2} / \min_i \Delta_{n,i}^{1/2} \to 0$, hence if the partition is $1/(2-p)$-balanced. This leads to the range $1 > p$ in an equally spaced setting, or $2 - 1/\epsilon \geq p$ for an ϵ-balanced partition. Hence knowing the regularity of our partition we can calculate the possible range of p, for example in the most irregular spaces case we have to take $1/2 \geq p$. Obviously conditions (8) and (9) imply (10).

- Condition (11) is satisfied if $Y_t^{(2)}$ is Hölder continuous of the order a. Hence this condition does not impose any restrictions on the choice of p. However, a and ϵ are related through $a > (2\epsilon + 1)/(4\epsilon)$ or $\epsilon > 1/(4a - 2)$. Hence the more regularity we have in the sampling scheme the less we need in $Y_t^{(2)}$ and vice versa. For the equally spaces case we need $a > 3/4$ whereas for the most irregular spaces case, $\epsilon = 2/3$ we need $a > 7/8$.

Summarizing, these considerations show that, compared to the considerations of consistency, it is much harder to find a good choice for p to obtain a reliable distributional theory and hence good error bounds for the estimate of the integrated volatility. They even show that it is impossible to obtain a distributional theory when the mean process possesses a continuous part which is correlated to the underlying Brownian motion and a jump part at the same time. Namely for the continuous part to be negligible we need at least $p > 1$ and for the jump part $p < 1$. Hence $p = 1$ cannot be used in neither case by this technique.

The situation improves if we know that the continuous part of the mean process is independent of the underlying Brownian motion. Then the continuous part imposes no restriction on the choice of p, hence we can choose some $p < 1$ to satisfy the condition for the jump part.

Let us discuss the choice of p in more detail. For the continuous, possibly correlated setting, we have two parameters which determine the choice of p, it is the balance parameter which is known and the Hölder coefficient of the process Y_t which might be unknown. The balance parameter ϵ determines the range of $p\gamma$, namely $p\gamma \geq 1/(2\epsilon)$. Hence if we do not know the regularity of Y_t it is safest to choose p large. However, if we choose p large outliers are weighted strongly. Note that the quadratic variation satisfies the conditions for equally spaces data if $\gamma \in (1/4, 1/2]$ or $\gamma \in (3/8, 1/2]$ for the most irregular 2/3-balanced partition.

If the continuous part is independent of the underlying Brownian motion, we have no restrictions on p.

If Y_t is a jump process, we need low activity, that is a Blumenthal-Getoor index $\beta < 1$ even for equally spaced data. For the most irregularly spaced data we need $\beta \leq 1/2$. This means that compound Poisson processes satisfy our assumptions, whereas not all bounded variation processes do.

Hence for the setting where the mean process possesses a jump component $Y_t^{(1)}$ and a continuous component $Y_t^{(2)}$, independent of the underlying Brownian motion, as for example in the Ornstein-Uhlenbeck stochastic volatility model with leverage, the best choice is to take p close to the Blumenthal-Getoor index of the driving Lévy process.

Condition (12) ensures that $\sum_i \Delta_{n,i}^{1-p/2} |\int_{t_{n,i-1}}^{t_{n,i}} \sigma_s ds|^p$ converges fast enough to $\int \sigma_s^p ds$ as required for (16) to hold. In Barndorff-Nielsen and Shephard (2003) it is shown that (12) is satisfied for Ornstein-Uhlenbeck type processes. Unfortunately, for continuous processes σ_t condition (12) only holds if we have Hölder continuity with exponent $> 1/2$, hence more regularity than for a Brownian motion or a diffusion process, normally used as volatility process. Recently, there has been made some progress towards relaxing this condition, namely to include the Cox-Ingersoll-Ross model as volatility process, cf. Barndorff-Nielsen et.al. (2003). However, so far this extension only holds for $p \geq 2$ which rules out the possibility showing that a jump component in the mean process is negligible.

As a special case we obtain the result for equally spaced data:

The assumption on Y_t reduces to

$$\begin{aligned} \Delta_n^{1-p} V_{2p}(Y, S_n) &\xrightarrow{p} 0 \\ \Delta_n^{(1-p)/2} V_p(Y^{(1)}, S_n) &\xrightarrow{p} 0, \end{aligned}$$

$Y_t^{(2)}$ satisfies (11) with $a \in (3/4, 1)$ and the stochastic volatility component satisfies the same conditions as in Theorem 2. Then for any $t > 0$ and for equally spaced partitions S_n, we obtain

$$\frac{\Delta_n^{1-p/2} V_p(X, S_n) - \mu_p \int_0^t \sigma^p(s) ds}{\sqrt{\mu_{2p}^{-1} \nu_p \Delta_n^{2-p} V_{2p}(X, S_n)}} \xrightarrow{\mathcal{D}} N(0,1),$$

as $n \to \infty$, where $\mu_p = E(|u|^p)$ and $\nu_p = Var(|u|^p)$ with $u \sim N(0,1)$.

Example:(Ornstein-Uhlenbeck-type model including leverage)
The Ornstein-Uhlenbeck-type stochastic volatility model including leverage, as it was suggested by Barndorff-Nielsen and Shephard (2001) is given by

$$\begin{aligned} dX_t &= \{\mu + \beta \sigma_t^2\} dt + \sigma_t dB_t + \rho d\bar{Z}_{\lambda t} \\ d\sigma_t^2 &= -\lambda \sigma_t^2 dt + dZ_{\lambda t}, \end{aligned}$$

where $\bar{Z}_t = Z_t - E(Z_t)$ and it is assumed that the subordinator Z_t is independent of the Brownian motion B_t. We can see that our conditions are satisfied, namely the continuous part of the mean process is independent of B_t and Lipschitz continuous, and condition (12) is satisfied by Bardorff-Nielsen and Shephard (2003). Hence for the equally spaced setting we obtain the distributional theory for $\beta < p < 1$, or $\beta < p \leq 2 - 1/\epsilon$ for the ϵ-balanced setting. This implies that we have to choose a Lévy process with sufficiently small Blumenthal-Getoor index, for example a compound Poisson process or a Gamma process.

We can summarize our result in to following table which shows us how we can choose p depending on the components of our mean process. Note that it does not have any influence if the jump component of the mean process is correlated to or independent of the underlying Brownian motion. First we look at the equally spaces setting. The last column shows the maximal range of p which is obtained for $\gamma = 1/2$.

components of mean process			p	p_{max}
Hölder cts. $1/2+\gamma$, $\gamma > 0$	Hölder cts. $a > 3/4$, indep. of B	jumps with index β		
yes	no	no	$p > 1/(2\gamma)$	$p > 1$
no	yes	no	$p > 0$	$p > 0$
no	no	yes	$\beta < p < 1$	$\beta < p < 1$
yes	yes	no	$p > 1/(2\gamma)$	$p > 1$
yes	yes	yes	–	–
yes	no	yes	–	–
no	yes	yes	$\beta < p < 1$	$\beta < p < 1$

For the most irregularly spaced setting, i.e. 2/3-balanced partitions we obtain:

components of mean process			p	p_{max}
Hölder cts. $1/2+\gamma$, $\gamma > 0$	Hölder cts. $a > 7/8$, indep. of B	jumps with index β		
yes	no	no	$p > 3/(4\gamma)$	$p > 3/2$
no	yes	no	$p > 0$	$p > 0$
no	no	yes	$\beta < p \leq 1/2$	$\beta < p \leq 1/2$
yes	yes	no	$p > 3/(4\gamma)$	$p > 3/2$
yes	yes	yes	–	–
yes	no	yes	–	–
no	yes	yes	$\beta < p \leq 1/2$	$\beta < p \leq 1/2$

Acknowledgements. I would like to thank Prof. Dr. N. Shephard for drawing my attention to the concept of power variation, Prof. Dr. O.E. Barndorff-Nielsen for helpful discussions and both for making their latest research available to me. Furthermore, I would like to thank the referee for helpful comments.

References

[1] T. G. Andersen and T. Bollerslev. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4:115–158, 1997.

[2] T. G. Andersen and T. Bollerslev. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economics Review*, 39:885–905, 1998.

[3] T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realised stock return volatility. *Journal of Financial Economics*, 61:43–76, 2001.

[4] T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. The distribution of exchange rate volatility. *Journal of the American Statistical Association*, 96:42–55, 2001.

[5] E. Andreou and E. Ghysels. Rolling-sampling volatility estimators: some new theoretical, simulation and empirical results. *Journal of Business and Economic Statistics*, 19, forthcoming.

[6] N. M. P. C. Areal and S. J. Taylor. The realised volatility of FTSE-100 futures prices. Technical report, Department of Accounting and Finance, Lancaster University, 2001.

[7] X. Bai, J. R. Russell, and G. C. Tiao. Beyond Merton's utopia: effects of non-normality and dependence on the precision of variance estimates using high-frequency data. Technical report, Graduate School of Business, University of Chicago, 2000.

[8] O. E. Barndorff-Nielsen, S. E. Graversen, and N. Shephard. Power variation & stochastic volatility: a review and some new results. Technical report, www.levyprocess.org, 2003.

[9] O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B*, 63:167–241, 2001.

[10] O. E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, 64:253–280, 2002a.

[11] O. E. Barndorff-Nielsen and N. Shephard. Power variation and time change. Technical report, www.levyprocess.org, 2002b.

[12] O. E. Barndorff-Nielsen and N. Shephard. Realised power variation and stochastic volatility models. *Bernoulli*, 9:243–265, 2003.

[13] S. M. Berman. Sign-invariant random variables and stochastic processes with sign invariant increments. *Trans. Amer. Math. Soc*, 119:216–243, 1965.

[14] T. Bollerslev and L. Forsberg. The distribution of realised volatility and the normal inverse gaussian GARCH model: an application to the ECU and EURO exchange rates. Technical report, Department of Economics, Duke University, 2001.

[15] T. Bollerslev and H. Zhou. Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, forthcoming, 2001.

[16] C. Q. Cao and R. S. Tsay. Nonlinear time series of stock volatility. *Journal of Applied Econometrics*, 7:S165–S185, 2001.

[17] F. Corsi, G. Zumbach, U. Muller, and M. Dacorogna. Consistent high-precision volatility from high-frequency data. Technical report, Olsen and Associates, Zurich, 2001.

[18] Z. Ding, C. W. J. Granger, and R.F. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

[19] J. W. Galbraith and V. Zinde-Walsh. Properties of estimates of daily GARCH parameters based on intra-day observations. Technical report, Economics Department, Mac Gill University, 2000.

[20] C. J. W. Granger and C.-T. Sin. Modelling the absolute returns of different stock indices: exploring the forecastability of an alternative measure of risk. Working paper, Department of Economics, University of California at San Diego, 1999.

[21] C. W. J. Granger and Z. Ding. Some properties of absolute returns, an alternative measure of risk. *Annals d'Economie et de Statistique*, 40:67–91, 1995.

[22] S. Heston. A closed form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[23] S. Howison, A. Rafailidis, and H. Rasmussen. On the pricing and hedging of volatility derivatives. Working paper, OCIAM, University of Oxford, 2003.

[24] W. N. Hudson and J. D. Mason. Variational sums for additive processes. *Proc. Amer. Math. Soc*, 55:395–399, 1976.

[25] W. N. Hudson and H. G. Tucker. Limit theorem for variational sums. *Trans. Amer. Math. Soc*, 191:405–426, 1974.

[26] D. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

[27] P. Jorion. Predicting volatility in the foreign exchange market. *Journal of Finance*, 50:507–528, 1995.

[28] D. Lepingle. La variation d'ordre p des semi-martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36:295–316, 1976.

[29] J. M. Maheu and T. H. McCudry. Nonlinear features of realised FX volatility. Technical report, Rotman School of Management, University of Toronto, 2001.

[30] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.

[31] L. Scott. Option pricing when the variance changes randomly: theory, estimation and an application. *Journal of Financial and Quantitative Analysis*, 22:419–438, 1987.

[32] A. N. Shiryaev. *Essentials of Stochastic Finance: Facts, Models and Theory*. World Scientific, Singapore, 1999.

[33] E. Stein and C. Stein. Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4:727–752, 1991.

[34] S. J. Taylor. *Modelling Financial Time Series*. John Wiley, Chichester, 1986.

[35] K. D. West and D. Cho. The predictive ability of several models for exchange rate volatility. *Journal of Econometrics*, 69:367–391, 1995.

[36] J. H. C. Woerner. Variational sums and power variation: a unifying approach to model selection and estimating in semimartingale models. *Statistics & Decisions*, 21:47–68, 2003.