Variational Sums and Power Variation: a unifying approach to model selection and estimation in semimartingale models

Jeannette H.C. Woerner
University of Oxford

Abstract

In the framework of general semimartingale models we provide limit theorems for variational sums including the p-th power variation, i.e. the sum of p-th absolute powers of increments of a process. This gives new insight in the use of quadratic and realised power variation as an estimate for the integrated volatility in finance. It also provides a criterion to decide from high frequency data, whether a jump component should be included in the model. Furthermore, results on the asymptotic behaviour of integrals with respect to Lévy processes, estimates for integrals with respect to Lévy measures and non-parametric estimation for Lévy processes will be derived and viewed in the framework of variational sums.

key words and phrases: limit theorem, Lévy process, model selection, power variation, non-parametric estimation, quadratic variation, semimartingale

1 Introduction

The concept of power variation, i.e. examining $\sum_i |X_{t_i} - X_{t_{i-1}}|^p$ as $\max_i |t_i - t_{i-1}| \to 0$ and its implications for estimating integrated volatility became increasingly popular in the last years, since on the one hand stochastic volatility models play an important role, and the link between the mathematical concept of quadratic variation and actual (i.e. integrated) volatility was made. Contributions include Barndorff-Nielsen and Shephard (2001a, 2001b, 2002a), 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 seemed to be more attractive to use absolute values of the returns than squares, see e.g. 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 (2001a) provided 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 $\alpha(t) + \int_0^t \sigma(s)dW_s$, where $\sigma > 0$ and α are assumed to be stochastically independent of the Brownian motion W. They also considered the same model when the Brownian motion is replaced by a stable process, cf. Barndorff-Nielsen and Shephard (2002b).

We now examine the general framework of semimartingales, also allowing jumps. Furthermore, we derive results for more general variational sums, where the p-th power function is replaces by a function, decreasing sufficiently fast to zero at the origin. This allows more flexibility in weighting large increments, i.e. possible outliers of the data. Different from most current financial literature, equally spaced data, which is hardly available, is not required for our results. Some theoretical results may already be found in the probability literature back in the sixties and seventies. Berman (1965), Hudson and Tucker (1974) and Hudson and Mason (1976) studied variational sums for additive processes, however only allowing either power functions or bounded functions. Lepingle (1976) derived results for power variation of semimartingales, when $p > 1$. Becker (1998) considered variational sums of random functions. However, this research does not include norming sequences, as introduced in Barndorff-Nielsen and Shephard (2001a) to deal with the continuous part when the power exponent is less than two. We extend the results in this direction, which has deep implications for modelling, and also provide results allowing a more flexible class of functions in the variational sums. Our results not only provide the theoretical background for estimation based on high frequency data, but also to decide which model is appropriate for the underlying data. Namely to decide, if the underlying data is from a continuous semimartingale or if it has jumps, and when jumps are involved even to decide, how much activity the Lévy process should have, e.g. finite activity as a compound Poisson process, bounded variation as e.g. subordinators or infinite activity as e.g. the hyperbolic Lévy motion. Hence we provide a different approach as in Ait-Sahalia (2002) to tackle the question, telling from discrete data whether the underlying process is continuous. He used transitions functions of diffusions and crossing arguments of trajectories.

Furthermore, our results give a different explanation, why empirically stochastic volatility estimates perform better, when using absolute values of returns 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 realised volatility, even when our model assumption was not correct. However, taking absolute values, only works when the jump component has at least slightly less activity than a bounded variation process. If this fails to hold, 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.

Since the calculation of power variation only involves high frequency data in some finite time interval, the concept of checking for jump components and analyzing the activity of jump components may also be used to observe how modelling should change over periods of time and how the correlation is to the economical situation. Hence power variation could give a flexible tool to adjust underlying stochastic models.

Looking at variational sums of the form $\sum E(g(X_{t_{n,i}} - X_{t_{n,i-1}}))$ for Lévy processes, where we sum up expectations, we can both infer the quantities of the Lévy triplet in a non-parametric setting and infer integrals with respect to Lévy processes. These asymptotic relations of integrals are e.g. needed in the context of proving local asymptotic normality for discretely observed Lévy processes, cf. Woerner (2001). We generalize results by Hudson and Tucker (1974) and Rüschendorf and Woerner (2002). Taking a sampling scheme where also the observed time interval tends to infinity, estimation of of integrals with respect to the Lévy measure, including as a special case non-parametric estimation of the jump measure and the drift in Lévy processes can be established in the context of variational sums, even without expectations. For the latter one cf. Rubin and Tucker (1959) and Basawa and Brockwell (1982).

The paper is organized as follows, first we give the basic notation and definitions, then look at variational sums of general semimartingale models, with a special emphasis on power variation, then consider asymptotics of integrals with respect to Lévy processes and finally look at non-parametric inference for Lévy processes.

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. 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 $\max_{1\leq k\leq n}\{t_{n,k} - t_{n,k-1}\} \to 0$ as $n \to \infty$. Now the p-th power variation is defined to be

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

Assume that g is a continuous non-negative function, then the variational sum with respect to g is defined to be

$$\sum_i g(|X_{t_{n,i}} - X_{t_{n,i-1}}|) = V_g(X, S_n).$$

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. An extension of the concept of power variation is to introduce an appropriate norming sequence, as it was done in Barndorff-Nielsen and Shephard (2001a), 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 now introduce our models. We start with a general semimartingale processes X_t, which is widely used in finance. For an overview both under financial and theoretical aspects see Shiryaev (1999). A semimartingale is a process right continuous with left limits of the form $X_t = X_0 + M_t + A_t$, where X_0 is finite-valued and $\mathcal{F}_0$-measurable, M is a local martingale and A some process of finite variation.

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 $(B(h), < X^c >, \nu)$, 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 $A \in \mathcal{B}(\mathbb{R} - \{0\})$ 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 $(\int_{(0,t] \times \mathbb{R}} (x^2 \wedge 1) d\nu)_{t \geq 0}$ is locally integrable. Semimartingale models include the well-established continuous diffusions, jump-diffusions, hence stochastic volatility models, as well as Lévy processes and most additive processes.

Additive processes X_t, in most cases a special form of a semimartingale, are processes with independent increments, in general given by their characteristic function

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

or for short by their characteristic triplet $(\alpha(t), \sigma^2(t), \nu_t)$. In contrast to the general semimartingale all quantities are deterministic, α is continuous, σ^2 continuous, non-negative, non-decreasing, and ν_t denotes the compensator of the jump measure, satisfying $\int (x^2 \wedge 1)\nu_t(dx) < \infty$. Here h denotes a truncation function, such that the integrability is insured, common functions are $h(x) = x/(1 + x^2)$ or $x 1_{|x| \leq 1}(x)$. Hence σ determines the Gaussian part and ν the jump part, which are independent. When α is of bounded variation, then the additive process is a semimartingale, which we shall assume in the following.

Lévy processes are a special class of additive processes where we not only have independent but also 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 does not have negative jumps and the process is of bounded variation in addition. For more details see Sato (1999).

A measure for the activity of the jump component of an additive process is the Blumenthal-Getoor index β, defined by

$$\beta = \inf\{\delta > 0 : \int (1 \wedge |x|^{\delta})\nu_t(dx) < \infty\}.$$

This index also determines, that for $p > \beta$ the sum of the p-th power of jumps will be finite. We can extend this index to general semimartingales replacing finiteness by being a locally integrable process,

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

3 Variational Sums and Power Variation

Semimartingales build a large class of models including different directions of more realistic modelling in finance by improving the major problems of the geometric Brownian motion in the Black-Scholes framework. Namely both, stochastic volatility models, which allow the spot rate to be random and serial dependent, but the underlying log-price process still being continuous, or pure jump stochastic differential equations, where the Brownian motion is replaced by some purely discontinuous Lévy process, are semimartingales. Both capture the empirical facts of excess kurtosis, skewness and fat tails. But of course the question occurs if the appropriate model is a continuous or a jump model. Ait-Sahalia (2002) provides a method based on the transition density for diffusion processes to decide on the basis of discrete samples if the underlying process is a continuous diffusion. We provide results based on variational sums and power variation, which allow to decide, whether the underlying process is purely continuous, purely discontinuous or a mixture, when over a fixed time interval the number of high frequency increments tends to infinity.

The existing literature only provides results for power variation, when $p > 1$ and X is a general semimartingale (cf. Lepingle (1976)) or general p, when X is an additive process (cf. Bermann (1965) and Hudson and Mason (1976)). Variational sums, except for theses power functions, were only considered for additive processes X and bounded functions g, which may be dominated by a power function with $p > 1$ (cf. Hudson and Tucker (1974)).

We generalize these results to general semimartingales X and variational sums, where g only has to decay sufficiently fast at the origin. This also provides results for power variation for general semimartingales with $p \leq 1$. Furthermore, we derive results, under which conditions with an appropriate chosen norming sequence, the limit of the continuous semimartingale component dominates the jump component, which is a new approach to variational sums.

Theorem 1 *Let X_t be a semimartingale, S_n be a partition of $[0,t]$, β be the generalized Blumenthal-Getoor index and g a nonnegative continuous function, satisfying the condition, that there exist $\eta > 0$, $C > 0$ and $\gamma > \beta$ such that for $|x| \leq \eta$, $g(x) \leq C|x|^\gamma$, then we obtain for $n \to \infty$*

$$V_g(X, S_n) \to \sum (g(|J(X_s)|) : 0 < s \leq t) \tag{1}$$

a.s. under the conditions:
a) If $\gamma > 2$.
b) If $1 < \gamma \leq 2$, $\beta = 1$ and $< X^c >_t = 0$.
c) If $\gamma \leq 1$, $< X^c >_t = 0$, $B(h) + (x - h) * \nu = 0$ and the jump times of X_t are previsible.
And in probability under the conditions:
a') If $\gamma > 2$.
b') If $1 < \gamma \leq 2$ and $< X^c >_t = 0$.
c') If $\gamma \leq 1$, $< X^c >_t = 0$, $B(h) + (x - h) * \nu = 0$ and the jump times of X_t are previsible.
d) Denote by Y the continuous part of X. Assume $\beta < \gamma < 2$ and $X - Y$ either satisfies b') or c'). Moreover assume that there exist a continuous f, such that f^{-1} exists and is continuous, satisfying for $c > 0$

$$f(cV_g(A + B, S_n)) \leq f(cV_g(A, S_n)) + f(cV_g(B, S_n)) \tag{2}$$

and

$$\Delta_n V_g(Y, S_n) \xrightarrow{p} C < \infty, \tag{3}$$

then also

$$\Delta_n V_g(X, S_n) \xrightarrow{p} C, \tag{4}$$

where Δ_n denotes some norming sequence tending to zero as $n \to \infty$.

Proof. For a) to c) and a') to c') we proceed similarly as in Hudson and Mason (1976) for power variation of additive processes. Denote $X_{nk} = X_{t_{n,k}} - X_{t_{n,k-1}}$, then we wish to show

$$\lim_{n\to\infty} \sum_k g(|X_{nk}|) = \sum (g(|J(X_s)|) : 0 < s \leq t).$$

The idea is to decompose the process in a process with jumps less than or equal to ϵ and one only possessing finitely many jumps of size bigger than ϵ, analyze both components and finally let epsilon tend to zero.

Let $\epsilon < \eta/4$ and define $I_{nj}(\epsilon)$ to be 1 if there are no jumps of absolute value greater than ϵ in $(t_{n,j-1}, t_{n,j}]$ and 0 otherwise. First we show that

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \sum_k g(|X_{nk}|) I_{nk}(\epsilon) = 0 \tag{5}$$

Let

$$Y_t^\epsilon = X_t - \sum (J(X_s) : |J(X_s)| > \epsilon, 0 < s \leq t).$$

Since g is nonnegative and $X_{nk} = Y_{nk}^\epsilon$ whenever $I_{nk}(\epsilon) = 1$, we have for every n and k

$$g(|X_{nk}|) I_{nk}(\epsilon) \leq g(|Y_{nk}^\epsilon|).$$

Furthermore, since Y_t^ϵ has right and left limits and no jumps of absolute value greater than ϵ, a Heine-Borel argument shows that $\sup_k |Y_{nk}^\epsilon| \leq 2\epsilon$ for sufficiently large n. This implies that for $\delta \in (\beta, \gamma)$ we obtain

$$\begin{aligned}
\limsup_{n\to\infty} \sum_k g(|Y_{nk}^\epsilon|) &\leq C \limsup_{n\to\infty} \sum_k |Y_{nk}^\epsilon|^\gamma \\
&\leq C(2\epsilon)^{\gamma-\delta} \limsup_{n\to\infty} \sum_k |Y_{nk}^\epsilon|^\delta \\
&\leq C(2\epsilon)^{\gamma-\delta} (\sum (|J(X_s)|^\delta : |J(X_s)| < \eta, 0 < s \leq t) \\
&\quad + \limsup_{n\to\infty} \sum_k |Y_{nk}^\eta|^\delta).
\end{aligned}$$

The first sum is finite since $\delta > \beta$. For the second we have to consider the different cases. For a) and b) the conditions in Lepingle (1976, Thm. 1) are satisfied and we obtain finiteness a.s., hence letting $\epsilon \to 0$ yields (5) a.s.. For a') and b') the conditions in Lepingle (1976, Thm. 2) are satisfied and we obtain finiteness in probability. Hence letting $\epsilon \to 0$ yields (5) in probability. Under the conditions of c) and c'), namely $1 > \delta > \beta > 0$, $< X^c >_t = 0$, $B(h) + (x - h) * \nu = 0$ and the jump times are previsible we have finiteness a.s. by Lepingle (1976, Thm. 1), noting that since Y_t^η has bounded jumps it is a special semimartingale and the condition on B ensures that it is a local martingale. Hence letting $\epsilon \to 0$ yields (5) a.s. and also in probability.

The second part of the proof is to look at the component with big jumps. Since g is continuous, we only have a finite number of jumps and the sample paths are right continuous with left limits, we obtain a.s.

$$\lim_{n\to\infty} \sum_k g(|X_{nk}|)(1 - I_{nk}(\epsilon)) = \sum (g(|J(X_s)|) : |J(X_s)| > \epsilon, 0 < s \leq t).$$

Furthermore, we have that

$$\lim_{\epsilon\to 0} \lim_{n\to\infty} \sum_k g(|X_{nk}|)(1 - I_{nk}(\epsilon)) \leq \liminf_{n\to\infty} \sum_k g(|X_{nk}|).$$

Hence under the appropriate conditions for (5), we obtain

$$\begin{aligned} \limsup_{n\to\infty} \sum_k g(|X_{nk}|) &\leq \lim_{\epsilon\to 0} \limsup_{n\to\infty} \sum_k g(|X_{nk}|) I_{nk}(\epsilon) \\ &\quad + \lim_{\epsilon\to 0} \limsup_{n\to\infty} \sum_k g(|X_{nk}|)(1 - I_{nk}(\epsilon)) \\ &\leq \liminf_{n\to\infty} \sum_k g(|X_{nk}|) \end{aligned}$$

This yields under the appropriate conditions for (5)

$$\lim_{n\to\infty} \sum_k g(|X_{nk}|) = \sum (g(|J(X_s)|) : 0 < s \leq t),$$

which is a) to c) and a') to c').

Finally we have to prove d). Let us decompose $X_t = Y_t + Z_t$ where Y_t denotes the continuous part of X_t and Z_t either satisfies b') or c'), then using the subadditivity property of f we obtain

$$\begin{aligned}
&P(|f(\Delta_n V_g(X,S_n)) - f(C)| > \lambda) \\
&\leq P(|f(\Delta_n V_g(X,S_n)) - f(\Delta_n V_g(Y,S_n))| > \lambda/2) \\
&\quad +P(|f(\Delta_n V_g(Y,S_n)) - f(C)| > \lambda/2) \\
&\leq P(|f(\Delta_n V_g(Z,S_n))| > \lambda/2) \\
&\quad +P(|f(\Delta_n V_g(Y,S_n)) - f(C)| > \lambda/2) \\
&\leq P(|f(\Delta_n V_g(Z,S_n)) - f(\Delta_n \sum (g(|J(X_s)|) : 0 < s \leq t))| > \lambda/4) \\
&\quad +P(f(\Delta_n \sum (g(|J(X_s)|) : 0 < s \leq t)) > \lambda/4) \\
&\quad +P(|f(\Delta_n V_g(Y,S_n)) - f(C)| > \lambda/2) < \epsilon,
\end{aligned}$$

since by the assumptions $\Delta_n \sum(g(|J(X_s)|) : 0 < s \leq t) \to 0$, $\Delta_n V_g(Y,S_n) \xrightarrow{p} C$ and by b') and c') $V_g(Z,S_n) \xrightarrow{p} \sum(g(|J(X_s)|) : 0 < s \leq t)$. This implies convergence in probability of $\Delta_n V_p(X,S_n)$ to C, as f and f^{-1} are continuous. $\square$

Remark.
i) Taking $g(x) = x^p$, we obtain the special case of power variation. For $p > 1$ the results are already known from Lepingle (1976). However, $p \leq 1$ and d) are new results. In the framework of power variation the conditions in d) simplify a lot. Namely, when $p > 1$ we have to take $f(x) = x^{1/p}$, then (2) follows by Minkowski's inequality, when $p \leq 1$ (2) is satisfied for f(x)=x.
ii) Note that $p = 2$ without any further restriction on the process is not covered by our Theorem. But it is the quadratic variation result and well-known, namely, if $p = 2$, then in probability

$$\begin{aligned}
V_2(X,S_n) &\xrightarrow{p} [X]_t \\
&\xrightarrow{p} [X^c]_t + \sum(|J(X_s)|^2 : 0 < s \leq t). \qquad (6)
\end{aligned}$$

iii) For the previous Theorem our assumptions are very general. We do not have to impose any further structure on our model and the observations need not be equidistant.
iv) For additive processes the condition of previsible jump times is always satisfies, cf. Sato (1999, Lemma 2.9).
v) For subordinators, i.e. Lévy processes with only positive jumps, 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.

Let us now discuss the implications of this theorem for our applications. For checking the presence and structure of jumps, considering power variation is sufficient. Whereas the general concept of variational sums gives more flexibility for weighting large values.

- If we are in a situation, where for $p > 2$

  $$V_p(X, S_n) \xrightarrow{p} 0,$$

  we know that our process is purely continuous. Furthermore, this implies that in the framework of stochastic volatility models the quadratic variation may be used as an estimate for the integrated volatility. However, in practice, when the distance between the observations does not tend to zero, one might prefer to construct a function satisfying the assumptions of Theorem 1 in that way that around zero it behaves like x^2 whereas it does not give much weight to large values of x.

- If we are in a situation, where for $p > 2$

  $$V_p(X, S_n) \xrightarrow{p} L > 0,$$

  we know that our process possesses a jump component. Furthermore, this implies that the quadratic variation is not a good estimate for the integrated volatility since it possesses an additional term coming from the jump component, as we can see in (6).

- If we are in the situation that for some $1 < p < 2$ and some appropriate norming sequence Δ_n

  $$\Delta_n V_p(X, S_n) \xrightarrow{p} L,$$

  with $0 < L < \infty$, we know that we have a continuous part and that for this p, for a previously identified jump part with $\sum(|J(X_s)|^p : 0 < s \leq t) < \infty$, the p-th power variation may be taken as an estimate for the integrated volatility, since the jump component has no influence. Of course, in practice when we do not know the structure of the jumps, we do not know β. Hence in general values close to two are more likely to satisfy the condition.

- If we are in the situation that for some $1 < p < 2$, we have

  $$V_p(X, S_n) \xrightarrow{p} L > 0,$$

and for all norming sequences Δ_n

$$\Delta_n V_p(X, S_n) \xrightarrow{p} 0,$$

we have a purely discontinuous process.

- If we are in the framework of Lévy processes without drift and diffusion we can also examine the activity of the process, i.e. how the Lévy measure of our process should behave around the origin. If we have a purely discontinuous Lévy process, we obtain that the process must be a compound Poisson process, when for all $p > 0$

$$V_p(X, S_n) \xrightarrow{p} L,$$

with $0 < L < \infty$. When the smallest number for which this is true, is less than one, our process has bounded variation, hence a subordinator may be appropriate for modelling, as it is e.g. suggested to include leverage effects in Barndorff-Nielsen and Shephard (2001b).

Example (Stochastic volatility model)
Let us start with the assumption that our data is derived from a stochastic volatility model as discussed in Barndorff-Nielsen and Shephard (2001a)

$$X_t = \alpha(t) + \int_0^t \sigma(s) dW_s,$$

where $\sigma > 0$, the spot volatility process, and α, the mean or risk premium, are stochastically independent of the Brownian motion W. For simplicity we assume that our partition of $[0, t]$ is equally spaced with distance Δ_n, $\Delta_n \to 0$ as $n \to \infty$, and $\Delta_n M = t$. Furthermore, we have to assume some regularity conditions. The volatility process $\tau = \sigma^2$ is (pathwise) locally bounded away from zero and has the property as $\delta \to 0$

$${\Delta_n}^{1/2} \sum_{j=1}^{M} |\tau^p(\eta_j) - \tau^p(\chi_j)| \xrightarrow{p} 0$$

for some $p > 0$ and for any χ_j and η_j such that

$$0 \leq \chi_1 \leq \eta_1 \leq \Delta_n \leq \chi_2 \leq \eta_2 \leq 2\Delta_n \leq \cdots \leq \chi_j \leq \eta_j \leq M\Delta_n = t.$$

The mean process α satisfies

$$\limsup_{\Delta_n \to 0} \max_{1 \le j \le M} \Delta_n^{-1} |\alpha(j\Delta_n) - \alpha((j-1)\Delta_n)| < \infty.$$

When we calculate power variations with $p > 2$ without norming sequence and the limit is zero, our data indeed should be from a semimartingale without jump component and we can use the limits and its implication on estimation as provided in Barndorff-Nielsen and Shephard (2001a), namely for $p \ge 1/2$

$$\mu_p^{-1} \Delta_n^{1-p/2} V_p(X, S) \xrightarrow{p} \int_0^t \sigma^p(s) ds, \tag{7}$$

and

$$\frac{\mu_p^{-1} \Delta_n^{1-p/2} V_p(X, S) - \int_0^t \sigma^p(s) ds}{\mu_p^{-1} \Delta_n^{1-p/2} \sqrt{\mu_{2p}^{-1} \upsilon_p \int_0^t \sigma^{2p}(s) ds}} \xrightarrow{\mathcal{D}} N(0, 1), \tag{8}$$

where $\mu_p = E[|u|^p]$ and $\upsilon_p = Var[|u|^p]$ with $u \sim N(0, 1)$. The second limit can be made feasible by applying the first limit and Slutzky's Lemma, which gives.

$$\frac{\mu_p^{-1} \Delta_n^{1-p/2} V_p(X, S) - \int_0^t \sigma^p(s) ds}{\mu_p^{-1} \Delta_n^{1-p/2} \sqrt{\mu_{2p}^{-1} \upsilon_p V_{2p}(X, S)}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{9}$$

If we get some positive limit for $V_p(X, S)$, $p > 2$, this indicates that our data does not match the continuous model and there should be some jump part, e.g. as introduced in the leverage case by Barndorff-Nielsen and Shephard (2001b). Nevertheless, in this case when the jumps are from a subordinator, in the calculation of the power variation with $1 < p < 2$ the jump component does not effect the result of (7) which can be used as an estimate for the integrated volatility. It is not clear if (8) is still valid, but (9) would certainly not be valid since $2p > 2$. If we want to take $p \le 1$, we need a jump component which satisfies the conditions of Lemma 1 d) with $\beta < 1$, e.g. a compound Poisson process.

Summarizing, to choose our exponent p we have to consider a trade of between choosing p close to 2 or small. If p is close to 2 then it is more likely that our conditions on the jumps are satisfied, but on the other hand outliers a weighted quite strongly. Taking $p \le 1$ we are more restrictive to our structure of jumps, we even need a stronger condition as being derived from a subordinator, on the other hand outliers are less strongly weighted.

4 Variational Sums of Expectations for Lévy Processes

In the previous section we examined variational sums for general semimartingales, where the function of the increments considered was a continuous function with appropriate decay at the origin. Now we consider sums of expectations of some functions of the increments of Lévy processes. In this case not only the behaviour around the origin is important, but also the behaviour as $x \to \infty$, namely to ensure the existence of the expectation. In Rüschendorf and Woerner (2002) it was derived that a appropriate class of functions g is the one that can be decomposed into a submultiplicative and subadditive component, which means that g does not increase faster than $x^k \exp\{cx\}$.

Definition 1

$$\begin{aligned} \mathcal{S} &= \{g(x) = h(x)k(x) | \exists H, K s.t. \, \forall x, y \in I\!R : \\ &\quad g(x+y) = h(x+y)k(x+y) \leq HK(h(x)+h(y))k(x)k(y)\}. \end{aligned}$$

Taking expectations instead of deterministic functions g in the variational sums, we can infer all quantities of the Lévy triplet and besides obtain results on the asymptotic behaviour of integrals with respect to Lévy processes when the time or the distance between the observations tends to zero. These integral occur when looking at parametric estimation, e.g. local asymptotic normality and martingale estimating functions, cf. Woerner (2001), for Lévy processes observed under the sampling scheme $n\Delta_n \to \infty$ and $\Delta_n \to 0$, as $n \to \infty$. Here Δ_n denotes the distance of observations and n the number of observations.

The following theorem is a generalization of theorems in Hudson and Tucker (1974, Thm. 1) and Rüschendorf and Woerner (2002, Thm. 5), allowing a wider range of functions and the Lévy process to possess a drift and diffusion component as well.

Theorem 2 *Let X_t be a Lévy process with Blumenthal-Getoor index β, observed at the time points $0, \Delta_n, 2\Delta_n, \cdots, n\Delta_n = t$, let P_{Δ_n} be the distribution of $X_{nk} = X_{k\Delta_n} - X_{(k-1)\Delta_n}$ and g a function satisfying*
1) $\int_{|x|\geq 1} g(x)\nu(dx) < \infty$, $|g(x)| \leq g_1$ for $g_1 \in \mathcal{S}$,
2) There exists $\eta, C > 0$ and $\gamma > \beta$ such that for all $|x| \leq \eta$ $|g(x)| \leq C|x|^\gamma$ and for all $|x| > \eta$ g is continuously differentiable with $|g'| \leq g_2$, $g_2 \in \mathcal{S}$

3) $|\int g(x+y)\nu_\epsilon(dy)| \leq g_3$, $|\frac{\partial}{\partial x}\int g(x+y)\nu_\epsilon(dy)| \leq g_4$ *for* $g_3, g_4 \in \mathcal{S}$ *where* $\nu_\epsilon = 1_{|x|>\epsilon}\nu$, *then we obtain*

$$\lim_{n\to\infty} \frac{1}{t}\sum_k E(g(X_{nk})) = \lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP_{\Delta_n}(x) = \int g(x)\nu(dx),$$

a) for a process with (α, σ^2, ν) *and* $\gamma > 2$,
b) for a process with $(\alpha, 0, \nu)$ *and* $\gamma > 1 \geq \beta$,
c) for a process with $(0, 0, \nu)$ *and* $\gamma > \beta$.
d) For a process with (α, σ^2, ν) *and under the additional condition that there exists* $\eta > 0$ *such that* $g(x) = x^2$ *for* $|x| \leq \eta$, *we obtain*

$$\lim_{n\to\infty} \frac{1}{t}\sum_k E(g(X_{nk})) = \lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP_{\Delta_n}(x) = \sigma^2 + \int g(x)\nu(dx).$$

e) For a process with $(\alpha, 0, \nu)$ *and under the additional condition that there exists* $\eta > 0$ *such that* $g(x) = x$ *for* $|x| \leq \eta$, *we obtain*

$$\lim_{n\to\infty} \frac{1}{t}\sum_k E(g(X_{nk})) = \lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP_{\Delta_n}(x) = \alpha + \int g(x)\nu(dx).$$

Proof. Let us first show that the representation as variational sum and as integral is the same.

$$\frac{1}{t}\sum_k E(g(X_{nk})) = \frac{1}{t}\sum_k \int g(x)dP_{\Delta_n}(x) = \frac{n}{n\Delta_n}\int g(x)dP_{\Delta_n}(x).$$

Let $\epsilon < \eta$, now we proceed similarly as in Theorem 1 splitting the process in a component only possessing small jumps, with distribution of the increments $P^\epsilon_{\Delta_n}$, and one possessing big jumps namely a compound Poisson process with distribution of the increments $P^{c,\epsilon}_{\Delta_n}$. Then

$$P_{\Delta_n}(x) = \exp\{-\Delta_n \int d\nu_\epsilon(x)\}(P^\epsilon_{\Delta_n}(x) + \sum_{i=1}^\infty \frac{\Delta_n^i}{i!}\nu_\epsilon^{i*} * P^\epsilon_{\Delta_n}(x)),$$

where $\nu_\epsilon = 1_{|x|>\epsilon}\nu$ and $P^\epsilon_{\Delta_n}$ denotes the distribution belonging to the jumps less than or equal ϵ, i.e Lévy measure $1_{|x|\leq\epsilon}d\nu(x)$. Then

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP_{\Delta_n} \\
&= \lim_{\epsilon\to 0}\lim_{n\to\infty}(\frac{1}{\Delta_n}\int g(x)dP^\epsilon_{\Delta_n}(x) + \frac{1}{\Delta_n}\int g(x)d(\sum_{i=1}^\infty \frac{\Delta_n^i}{i!}\nu_\epsilon^{i*} * P^\epsilon_{\Delta_n})(x)).
\end{aligned}$$

Under condition 3) we know from Rüschendorf and Woerner (2002, Thm.5) that

$$\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)d(\sum_{i=1}^{\infty}\frac{\Delta_n^i}{i!}\nu_\epsilon^{i*}*P_{\Delta_n}^\epsilon)(x)=\int_{|x|>\epsilon}g(x)\nu(dx).$$

Under condition 1) we can take the limit and obtain

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)d(\sum_{i=1}^{\infty}\frac{\Delta_n^i}{i!}\nu_\epsilon^{i*}*P_{\Delta_n}^\epsilon)(x)=\int g(x)\nu(dx). \tag{10}$$

Now we have to look at the small jumps. Let

$$Y_t^\epsilon=X_t-\sum(J(X_s):|J(X_s)|>\epsilon,0<s\le t),$$

which is the process corresponding to P_t^ϵ. First we show that under the conditions of a), b) and c)

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{t}\sum_k E(g(Y_{nk}^\epsilon))=\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP_{\Delta_n}^\epsilon(x)=0 \tag{11}$$

By 2) we obtain

$$\begin{aligned}\frac{1}{t}\sum_k E(g(Y_{nk}^\epsilon)) &= \frac{1}{\Delta_n}\int g(x)dP_{\Delta_n}^\epsilon(x)\\ &\le \frac{C}{\Delta_n}\int|x|^\gamma dP_{\Delta_n}^\epsilon(x)\\ &\quad+\frac{1}{\Delta_n}\int_{|x|>\eta}g(x)dP_{\Delta_n}^\epsilon(x)-\frac{C}{\Delta_n}\int_{|x|>\eta}|x|^\gamma dP_{\Delta_n}^\epsilon(x)\end{aligned}$$

To prove that the second and third term tend to zero as $n\to\infty$ we can use the same technique as in Rüschendorf and Woerner (2002), we give the proof for the second term, but the third can be proved analogous, noting that $|x|^\gamma$ satisfies 1) and 2). Integration by parts yields

$$\int_{|x|>\eta}g(x)dP_{\Delta_n}^\epsilon(x)=\int_{x>\eta}g'(x)\int_x^\infty dP_{\Delta_n}^\epsilon(y)dx-\int_{x<-\eta}g'(x)\int_{-\infty}^x dP_{\Delta_n}^\epsilon(y)dx.$$

We then obtain by Rüschendorf and Woerner (2002, Lemma 2)

$$\left|\frac{1}{\Delta_n}\int_{x>\eta}g'(x)\int_x^\infty dP_{\Delta_n}^\epsilon(y)dx\right|$$

$$
\begin{aligned}
&\le \frac{1}{\Delta_n}\int_{x>\eta}|g'(x)|(1-P^{\epsilon}_{\Delta_n}(x))dx \\
&\le \int_{x>\eta}|g'(x)|\exp\{-\Delta_n a x_0+\Delta_n a x_0\log x_0+ax-ax\log x\}\Delta_n^{ax-1}dx \\
&\le \exp\{-\Delta_n a x_0+\Delta_n a x_0\log x_0\}\int_{x>\eta}|g'(x)|\exp\{ax-ax\log x\}dx<\infty
\end{aligned}
$$

for $\Delta_n \le 1$, $x/\Delta_n \ge x_0$, $a < 1/\epsilon$ and $a\eta > 1$. Hence by dominated convergence we may interchange limit and integration and obtain

$$\lim_{n\to\infty}\frac{1}{\Delta_n}\int_{x>\eta}g'(x)\int_x^\infty dP^{\epsilon}_{\Delta_n}(y)dx=0.$$

The same holds for the other part, which yields

$$\lim_{n\to\infty}\frac{1}{\Delta_n}\int_{|x|>\eta}g(x)dP^{\epsilon}_{\Delta_n}(x)=0. \tag{12}$$

Finally we have to show

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{C}{\Delta_n}\int|x|^{\gamma}dP^{\epsilon}_{\Delta_n}(x)=0.$$

Under b) this term tends to zero as $n\to\infty$ by Hudson and Tucker (1974, Thm.1). Under a) this term tends to zero by Rüschendorf and Woerner (2002, Thm. 5) noting that $|x|^\gamma$ for $\gamma > 2$ satisfies the condition of possessing a derivative belonging to $\mathcal{S}$ except of the origin, which is not important since $\nu(\{0\}) = 0$ anyway. Under c) we can use Lemma 5.2 and 5.3 of Berman (1965). The only point where he uses his condition $\gamma > 1$ is to show by applying Hölder's inequality that

$$\lim_{n\to\infty}\frac{1}{\Delta_n}\int_{|x|>\epsilon}|x|^{\gamma}dP^{\epsilon}_{\Delta_n}(x)=0,$$

However,

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{1}{\Delta_n}\int_{|x|>\epsilon}|x|^{\gamma}dP^{\epsilon}_{\Delta_n}(x)= \\
&\qquad \lim_{n\to\infty}\frac{1}{\Delta_n}\int_{\epsilon<|x|\le\eta}|x|^{\gamma}dP^{\epsilon}_{\Delta_n}(x)+\lim_{n\to\infty}\frac{1}{\Delta_n}\int_{|x|>\eta}|x|^{\gamma}dP^{\epsilon}_{\Delta_n}(x),
\end{aligned}
$$

where the first part is zero by Rüschendorf and Woerner (2002, Lemma 6) and the second is again of the form of (12).

For d) and e) we use the same procedure. Under the additional assumption on g we obtain that

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP^{\epsilon}_{\Delta_n}(x) \\
&= \lim_{n\to\infty}\frac{1}{\Delta_n}(\int x^2 dP^{\epsilon}_{\Delta_n}(x)+\int_{|x|>\eta} g(x)dP^{\epsilon}_{\Delta_n}(x)-\int_{|x|>\eta} x^2 dP^{\epsilon}_{\Delta_n}(x) \\
&= \sigma^2+\int_{-\epsilon}^{\epsilon} x^2\nu(dx)
\end{aligned}$$

or respectively

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP^{\epsilon}_{\Delta_n}(x) \\
&= \lim_{n\to\infty}\frac{1}{\Delta_n}(\int x dP^{\epsilon}_{\Delta_n}(x)+\int_{|x|>\eta} g(x)dP^{\epsilon}_{\Delta_n}(x)-\int_{|x|>\eta} x dP^{\epsilon}_{\Delta_n}(x) \\
&= \alpha+\int_{-\epsilon}^{\epsilon} (x-h(x))\nu(dx),
\end{aligned}$$

where for the first part of the integrals we use the moment representation for Lévy processes and for the other two parts the same argument as for (12). Finally we obtain for d) and e)

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP^{\epsilon}_{\Delta_n}(x) = \sigma^2 \tag{13}$$

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{\Delta_n}\int g(x)dP^{\epsilon}_{\Delta_n}(x) = \alpha \tag{14}$$

Piecing together (10), (11), (13) and (14) we obtain the desired result. □

5 Non-parametric Estimation for Lévy Processes

The previous theorem gives us some possibilities to infer the quantities of the Lévy triplet in a non-parametric setting. To estimate σ^2 we can take $g_\epsilon(x) = x^2 1_{|x|\leq\epsilon}(x)$ in b) and let $\epsilon \to 0$. Analogously for α we can take

$g_\epsilon(x) = x1_{|x|\le\epsilon}(x)$ in d) and let $\epsilon \to 0$. To infer the Lévy measure, we can take $g(x) = 1_{x\le y}(x)$, for $y < 0$ or $g(x) = 1_{x\ge y}(x)$, for $y > 0$ and obtain $\nu(-\infty, y]$ or $\nu[y, \infty)$ respectively. However, these estimators all involve expectations of our data. Allowing our sampling scheme to change, namely now looking at an infinite time interval, we can also give estimates for integrals with respect to the Lévy measure, and as a special case for ν itself, not involving expectations, but still lying in the framework of variational sums. The heuristics is easy, simply replacing the expectation by its consistent estimator and using the LLN as the number of replications tends to infinity. The following Theorem makes this idea rigorous.

Theorem 3 *Under the same conditions as in Theorem 2 and furthermore assuming that the same conditions also hold for $|g|$, we obtain*

$$\frac{1}{n\Delta_n}\sum_{k=1}^{n} g(X_{nk}) \to \int g(x)\nu(dx), \tag{15}$$

under the assumptions a) to c),

$$\frac{1}{n\Delta_n}\sum_{k=1}^{n} g(X_{nk}) \to \sigma^2 + \int g(x)\nu(dx),$$

under the assumption d) and

$$\frac{1}{n\Delta_n}\sum_{k=1}^{n} g(X_{nk}) \to \alpha + \int g(x)\nu(dx),$$

under the assumption e). The convergence is in probability as $n \to \infty$, using the sampling scheme $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$.

Proof. The main part of the proof is using a LLN for triangular schemes and checking the conditions by the results from Theorem 2. We restrict ourselves to prove (15) since the other two parts can be shown analogously.

Using the LLN for triangular schemes (cf. Gnedenko and Kolmogorov (1968, p.134)), we obtain

$$\frac{1}{n\Delta_n}\sum_{k=1}^{n} g(X_{nk}) - \frac{1}{\Delta_n}\int_{|g(x)|<n\Delta_n} g(x)dP_{\Delta_n}(x) \xrightarrow{p} 0$$

as $n \to \infty$ under the conditions

$$n \int_{|x| \geq n\Delta_n} dF_{\Delta_n}(x) \quad \to \quad 0 \tag{16}$$

$$\frac{1}{n\Delta_n^2} \int_{|x| < n\Delta_n} x^2 dF_{\Delta_n}(x) \quad \to \quad 0 \tag{17}$$

where $g(X_{nk}) \sim F_{\Delta_n}$. However, as it was shown in Feller (1966, p.232) in the framework for iid random variables (16) already implies (17). Namely integration by parts for (17) leads to

$$\begin{aligned}
&\frac{1}{n\Delta_n^2} \int_{|x| < n\Delta_n} x^2 dF_{\Delta_n}(x) \\
&= \frac{1}{n\Delta_n^2} [x^2 F_{\Delta_n}(x)]_{-n\Delta_n}^{n\Delta_n} - \frac{2}{n\Delta_n^2} \int_{-n\Delta_n}^{n\Delta_n} x F_{\Delta_n}(x) dx \\
&= nF_{\Delta_n}(n\Delta_n) - nF_{\Delta_n}(-n\Delta_n) - \frac{2}{n\Delta_n^2} \int_0^{n\Delta_n} x F_{\Delta_n}(x) dx \\
&\quad - \frac{2}{n\Delta_n^2} \int_{-n\Delta_n}^0 x F_{\Delta_n}(x) dx \\
&= nF_{\Delta_n}(n\Delta_n) - nF_{\Delta_n}(-n\Delta_n) - \frac{2}{n\Delta_n^2} \int_0^{n\Delta_n} x F_{\Delta_n}(x) dx \\
&\quad + \frac{2}{n\Delta_n^2} \int_0^{n\Delta_n} x F_{\Delta_n}(-x) dx \\
&= -n \int_{|x| \geq n\Delta_n} dF_{\Delta_n}(x) + \frac{2}{n\Delta_n^2} \int_0^{n\Delta_n} z \int_{|x| \geq z} dF_{\Delta_n}(x) dz \\
&\leq n \int_{|x| \geq n\Delta_n} dF_{\Delta_n}(x).
\end{aligned}$$

To establish (16) we can use

$$\begin{aligned}
n \int_{|x| \geq n\Delta_n} dF_{\Delta_n}(x) &\leq \frac{1}{\Delta_n} \int_{|x| \geq n\Delta_n} |x| dF_{\Delta_n}(x) \\
&= \frac{1}{\Delta_n} \int_{|g(x)| \geq n\Delta_n} |g(x)| dP_{\Delta_n}(x). \qquad (18)
\end{aligned}$$

Hence it remains to show that

$$\lim_{n \to \infty} \frac{1}{\Delta_n} \int_{|g(x)| < n\Delta_n} g(x) dP_{\Delta_n}(x) \quad = \quad \int g(x) \nu(dx) \tag{19}$$

$$\lim_{n\to\infty} \frac{1}{\Delta_n} \int_{|g(x)|\geq n\Delta_n} |g(x)| dP_{\Delta_n}(x) \quad = \quad 0. \tag{20}$$

(19) can be show as in Theorem 2, noting that $g(x)1_{|g(x)|<n\Delta_n} \leq g(x)$. For (20) we can use Theorem 2 and (19) for $|g|$, namely

$$\begin{aligned}
&\int |g(x)|\nu(dx) \\
&= \lim_{n\to\infty} \frac{1}{\Delta_n} \int |g(x)| dP_{\Delta_n}(x) \\
&= \lim_{n\to\infty} \frac{1}{\Delta_n} \int_{|g(x)|\geq n\Delta_n} |g(x)| dP_{\Delta_n}(x) + \lim_{n\to\infty} \frac{1}{\Delta_n} \int_{|g(x)|<n\Delta_n} |g(x)| dP_{\Delta_n}(x) \\
&= \int |g(x)|\nu(dx) + \lim_{n\to\infty} \frac{1}{\Delta_n} \int_{|g(x)|<n\Delta_n} |g(x)| dP_{\Delta_n}(x) \\
&\geq \int |g(x)|\nu(dx),
\end{aligned}$$

which yields (20). Hence the assumptions for the LLN are satisfied and we obtain the desired result. □

For the special case, where g is the indicator function, a similar result was proved by Rubin and Tucker (1959) for general Lévy processes and for subordinators by Basawa and Brockwell (1982) under slightly different conditions.

To avoid the problem with the singularity of ν at the origin, Rubin and Tucker (1959) used a different notation of their Lévy process. They considered the process given by the characteristic function

$$\exp\{it\alpha u + t \int (e^{iux} - 1 - \frac{iux}{1+x^2}) \frac{1+x^2}{x^2} dG(x)\},$$

where G is bounded and non-decreasing with $G(-\infty) = 0$ and $G(+0) - G(-0) = \sigma^2$, determining the diffusion component. Then they obtained that

$$G^*_{\Delta_n,n}(u) = \frac{1}{n\Delta_n} \sum_{k=1}^{n} \frac{X_{nk}^2}{1+X_{nk}^2} 1_{X_{nk}\leq u}$$

is strongly consistent as $n \to \infty$ for all continuity points of G.

Acknowledgements. I would like to thank Prof. Dr. N. Shephard for drawing my attention to the concept of power variation and Prof. Dr. L. Rüschendorf for helpful discussions.

References

[1] Y. Ait-Sahalia. Telling from discrete data whether the underlying continuous-time model is a diffusion. Technical report, forthcoming in the Journal of Finance, 2002.

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

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

[4] 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.

[5] 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.

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

[7] 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.

[8] 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.

[9] O. E. Barndorff-Nielsen and N. Shephard. Realised power variation and stochastic volatility models. MPS-RR 2002-6, MaPhySto, Aarhus, 2001a.

[10] 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, 2001b.

[11] 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, 2002a.

[12] O. E. Barndorff-Nielsen and N. Shephard. Jumps and power variation. Technical report, 2002b.

[13] I. V. Basawa and P. J. Brockwell. Non-parametric estimation for non-decreasing lévy processes. *Journal of the Royal Statistical Society, Series B*, 44:262–269, 1982.

[14] E. Becker. *Theoremes limites pour des processus discretises*. PhD thesis, University Paris VI, 1998.

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

[16] 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.

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

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

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

[20] 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.

[21] 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.

[22] 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.

[23] 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.

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

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

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

[29] H. Rubin and H. G. Tucker. Estimating the parameters of a differntial process. *Ann. Math. Statist.*, 30:641–658, 1959.

[30] L. Rüschendorf and J. H. C. Woerner. Expansions of transition distributions of Lévy processes in small time. *Bernoulli*, 8(1):81–96, 2002.

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

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

[33] S. J. Taylor. *Modelling Financial Time Series.* John Wiley, Chichester, 186.

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

[35] J. H. C. Woerner. *Statistical analysis for discretely observed Lévy processes.* PhD thesis, University of Freiburg, 2001.

OCIAM, Mathematical Institute, University of Oxford, 24-29 St. Giles, Oxford OX1 3LB, UK; email: woerner@maths.ox.ac.uk