Limit theorems for bipower variation in financial econometrics

Ole E. Barndorff-Nielsen
Department of Mathematical Sciences,
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
oebn@imf.au.dk

Svend Erik Graversen
Department of Mathematical Sciences,
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
matseg@imf.au.dk

Jean Jacod
Laboratoire de Probabilités et Modéles Aléatoires (CNRS UMR 7599),
Université Pierre et Marie Curie,
4 Place Jussieu, 75252 Paris Cedex 05, France
jj@ccr.jussieu.fr

Neil Shephard
Nuffield College, University of Oxford, Oxford OX1 1NF, UK
neil.shephard@nuf.ox.ac.uk

October 27, 2005

Abstract

In this paper we provide an asymptotic analysis of generalised bipower measures of the variation of price processes in financial economics. These measures encompass the usual quadratic variation, power variation and bipower variations which have been highlighted in recent years in financial econometrics. The analysis is carried out under some rather general Brownian semimartingale assumptions, which allow for standard leverage effects.

Keywords: Bipower variation; Power variation; Quadratic variation; Semimartingales; Stochastic volatility.

1 Introduction

In this paper we discuss the limiting theory for a novel, unifying class of non-parametric measures of the variation of financial prices. The theory covers commonly used estimators of variation such as realised volatility, but it also encompasses more recently suggested quantities like realised power variation and realised bipower variation. We considerably strengthen existing results on the latter two quantities, deepening our understanding and unifying their treatment. We will outline the proofs of these theorems, referring for the very technical, detailed formal proofs of the general results to a companion probability theory paper Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005). Our emphasis is on exposition, explaining where the results come from and how they sit within the econometrics literature.

Our theoretical development is motivated by the advent of complete records of quotes or transaction prices for many financial assets. Although market microstructure effects (e.g. discreteness of prices, bid/ask bounce, irregular trading etc.) mean that there is a mismatch between asset pricing theory based on semimartingales and the data at very fine time intervals it does suggest the desirability of establishing an asymptotic distribution theory for estimators as we use more and more highly frequent observations. Papers which directly model the impact of market frictions on realised volatility include Zhou (1996), Bandi and Russell (2003), Hansen and Lunde (2006), Zhang, Mykland, and Aït-Sahalia (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004) and Zhang (2004). Related work in the probability literature on the impact of noise on discretely observed diffusions can be found in Gloter and Jacod (2001a) and Gloter and Jacod (2001b), while Delattre and Jacod (1997) report results on the impact of rounding on sums of functions of discretely observed diffusions. In this paper we ignore these effects.

Let the d-dimensional vector of the log-prices of a set of assets follow the process

$$Y = \left(Y^1, ..., Y^d\right)'.$$

At time $t \geq 0$ we denote the log-prices as Y_t. Our aim is to calculate measures of the variation of the price process (e.g. realised volatility) over discrete time intervals (e.g. a day or a month). Without loss of generality we can study the mathematics of this by simply looking at what happens when we have n high frequency observations on the time interval $t = 0$ to $t = 1$ and study our measures of variation as $n \to \infty$. In this case returns will be measured over intervals of length n^{-1} as

$$\Delta_i^n Y = Y_{i/n} - Y_{(i-1)/n}, \quad i = 1, 2, ..., n, \tag{1}$$

where n is a positive integer.

We will study the behaviour of the realised generalised bipower variation process

$$Y^n(g, h)_t = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} g(\sqrt{n}\ \Delta_i^n Y) h(\sqrt{n}\ \Delta_{i+1}^n Y), \tag{2}$$

as n becomes large and where g and h are two given, matrix functions of dimensions $d_1 \times d_2$ and $d_2 \times d_3$ respectively, whose elements have at most polynomial growth. Here $\lfloor x \rfloor$ denotes the largest integer less than or equal to x.

Although (2) looks initially rather odd, in fact most of the non-parametric volatility measures used in financial econometrics fall within this class (a measure not included in this setup is the range statistic studied in, for example, Parkinson (1980) and its realised version recently introduced by Christensen and Podolskij (2005) and Martens and van Dijk (2005)). Here we give an extensive list of examples and link them to the existing literature. More detailed discussion of the literature on the properties of these special cases will be given later.

Example 1 ***(a)*** *Suppose $g(y) = \left(y^j\right)^2$ and $h(y) = 1$, then (2) becomes*

$$\sum_{i=1}^{\lfloor nt \rfloor} \left(\Delta_i^n Y^j\right)^2, \quad j = 1, 2, ..., d,$$

which is called the realised quadratic variation process of Y^j in econometrics, e.g. Jacod (1994), Jacod and Protter (1998), Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen and Shephard (2004a) and Mykland and Zhang (2006). The increments of this quantity, typically calculated over a day or a week, are often called the realised variances in financial economics. The importance of these increments have been highlighted by Andersen, Bollerslev, Diebold, and Labys (2001) and Andersen, Bollerslev, and Diebold (2006) in the context of volatility measurement and forecasting. See also the survey by Barndorff-Nielsen and Shephard (2005b). Realised variance has a very long history in financial economics. It appears in, for example, Rosenberg (1972), Officer (1973), Merton (1980), French, Schwert, and Stambaugh (1987), Schwert (1989) and Schwert (1998).

(b) *Suppose $g(y) = yy'$ and $h(y) = I$, then (2) becomes, after some simplification,*

$$\sum_{i=1}^{\lfloor nt \rfloor} \left(\Delta_i^n Y\right)\left(\Delta_i^n Y\right)'.$$

This is the realised covariation process. It has been studied by Jacod and Protter (1998), Barndorff-Nielsen and Shephard (2004a) and Mykland and Zhang (2006). Andersen, Bollerslev, Diebold, and Labys (2003) study the increments of this process to produce forecast distributions for vectors of returns.

(c) *Suppose $g(y) = \left|y^j\right|^r$ for $r > 0$ and $h(y) = 1$, then (2) becomes*

$$n^{-1+r/2} \sum_{i=1}^{\lfloor nt \rfloor} \left|\Delta_i^n Y^j\right|^r, \quad j = 1, 2, ..., d,$$

which is called the realised r-th order power variation. When r is an integer it has been studied from a probabilistic viewpoint by Jacod (1994) while Barndorff-Nielsen and Shephard (2003) look at the econometrics of the case where $r > 0$. Barndorff-Nielsen and Shephard (2004b) extend this work to the case where there are jumps in Y, showing the statistic is robust to certain types of jumps when $r < 2$. Aït-Sahalia and Jacod (2005) have additional insights on that topic. The increments of these types of high frequency volatility measures have been informally used *in the financial econometrics literature for some time when $r = 1$, but until recently without a strong understanding of their theoretical asymptotic properties. Examples of their use include Schwert (1990), Andersen and Bollerslev (1998) and Andersen and Bollerslev (1997), while they have also been informally discussed by Shiryaev (1999, pp. 349–350) and Maheswaran and Sims (1993). Following the work by Barndorff-Nielsen and Shephard (2003), Ghysels, Santa-Clara, and Valkanov (2004) and Forsberg and Ghysels (2004) have successfully used realised power variation as an input into volatility forecasting competitions.*

(d) *Suppose $g(y) = \left|y^j\right|^r$ and $h(y) = \left|y^j\right|^s$ for $r, s > 0$, then (2) becomes*

$$n^{-1+(r+s)/2} \sum_{i=1}^{\lfloor nt \rfloor} \left|\Delta_i^n Y^j\right|^r \left|\Delta_{i+1}^n Y^j\right|^s, \quad j = 1, 2, ..., d,$$

which is called the realised r, s-th order bipower variation process. This measure of variation was introduced by Barndorff-Nielsen and Shephard (2004b), while a more formal discussion of its behaviour in the $r = s = 1$ case was developed by Barndorff-Nielsen and Shephard (2006). These authors' interest in this quantity was motivated by its virtue of being resistant to finite activity jumps so long as $\max(r, s) < 2$. Recently Barndorff-Nielsen, Shephard, and Winkel (2004) and Woerner (2006) have studied how these results on jumps extend to infinite activity processes, while Corradi and Distaso (2004) have used these statistics to test the specification of parametric volatility models. Here we study these statistics in the case where there are no jumps.

(e) *Suppose*

$$g(y) = \begin{pmatrix} \left|y^j\right| & 0 \\ 0 & \left(y^j\right)^2 \end{pmatrix}, \quad h(y) = \begin{pmatrix} \left|y^j\right| \\ 1 \end{pmatrix}.$$

Then (2) becomes,

$$\begin{pmatrix} \sum_{i=1}^{\lfloor nt \rfloor} \left|\Delta_i^n Y^j\right| \left|\Delta_{i+1}^n Y^j\right| \\ \sum_{i=1}^{\lfloor nt \rfloor} \left(\Delta_i^n Y^j\right)^2 \end{pmatrix}.$$

Barndorff-Nielsen and Shephard (2006) used the joint behaviour of the increments of these two statistics to test for jumps in price processes. Huang and Tauchen (2005) have empirically studied the finite sample properties of these types of jump tests. Andersen, Bollerslev, and Diebold (2003) and Forsberg and Ghysels (2004) use bipower variation as an input into volatility forecasting.

We will derive the probability limit of (2) under a general Brownian semimartingale, the workhorse process of modern continuous time asset pricing theory. Only the case of realised quadratic variation, where the limit is the usual quadratic variation QV (defined for general semimartingales), has been previously studied under such wide conditions. Further, under some stronger but realistic conditions, we will derive a limiting distribution theory for (2), so extending a number of results previously given in the literature on special cases of this framework.

The outline of this paper is as follows. Section 2 contains the main notation used in our analysis. Section 3 gives a statement of a weak law of large numbers for these statistics and the corresponding central limit theory is presented in Section 4. Extensions of the results to higher order variations are briefly indicated in Sections 5 and 6. Section 7 concludes, while there is an Appendix which provides an outline of the proofs of the results discussed in this paper. For detailed, quite lengthy and highly technical formal proofs we refer to our companion probability theory paper Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005).

2 Notation and models

We start with Y on some filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P\right)$. In most of our analysis we will assume that Y follows a d-dimensional Brownian semimartingale (written $Y \in \mathcal{BSM}$). It is given in the following statement.

Assumption (H): We have

$$Y_t = Y_0 + \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_{u-} \mathrm{d}W_u, \tag{3}$$

where W is a d'-dimensional standard Brownian motion (BM), a is a d-dimensional process whose elements are predictable and has locally bounded sample paths, and the spot covolatility d, d'-dimensional matrix σ has elements which have càdlàg sample paths.

Throughout we will write

$$\Sigma_t = \sigma_t \sigma_t', \tag{4}$$

the spot covariance matrix. Typically Σ_t will be full rank, but we do not assume that here. We will write Σ_t^{jk} to denote the j, k-th element of Σ_t and

$$\sigma_{j,t}^2 = \Sigma_t^{jj}.$$

Remark 1 *Due to the fact that $t \mapsto \sigma_t^{jk}$ is càdlàg all powers of σ_t^{jk} are locally integrable with respect to the Lebesgue measure. In particular then $\int_0^t \Sigma_u^{jj} \mathrm{d}u < \infty$ for all t and j.*

Remark 2 *Both a and σ can have, for example, jumps, intraday seasonality and long-memory.*

Remark 3 *The stochastic volatility (e.g. Shephard (2005)) component of Y,*

$$\int_0^t \sigma_{u-} \mathrm{d}W_u,$$

is always a vector of local martingales each with continuous sample paths, as $\int_0^t \Sigma_u^{jj} \mathrm{d}u < \infty$ *for all* t *and* j*. All continuous local martingales with absolutely continuous quadratic variation can be written in the form of a stochastic volatility process. This result, which is due to Doob (1953), is discussed in, for example, Karatzas and Shreve (1991, p. 170–172). Using the Dambis-Dubins-Schwartz Theorem, we know that the difference between the entire continuous local martingale class and the SV class are the local martingales which have only continuous, not absolutely continuous*[1]*, QV. The drift* $\int_0^t a_u \mathrm{d}u$ *has elements which are absolutely continuous. This assumption looks ad hoc, however if we impose a lack of arbitrage opportunities and model the local martingale component as a SV process then this property must hold (Karatzas and Shreve (1998, p. 3) and Andersen, Bollerslev, Diebold, and Labys (2003, p. 583)). Hence (3) is a rather canonical model in the finance theory of continuous sample path processes.*

3 Law of large numbers

To build a weak law of large numbers for $Y^n(g,h)_t$ we need to make the pair (g,h) satisfy the following assumption.

Assumption (K): All the elements of f on $\mathbf{R}^d$ are continuous with at most polynomial growth.

This amounts to there being suitable constants $C > 0$ and $p \geq 2$ such that

$$x \in \mathbf{R}^d \quad \Rightarrow \quad \|f(x)\| \leq C(1 + \|x\|^p). \tag{5}$$

We also need the following notation.

$$\rho_\sigma(g) = \mathrm{E}\left\{g(X)\right\}, \quad \text{where} \quad X|\sigma \sim N(0, \sigma\sigma'),$$

and

$$\rho_\sigma(gh) = \mathrm{E}\left\{g(X)h(X)\right\},$$

where the expectations are conditional on σ.

Example 2 ***(a)*** *Let* $g(y) = yy'$ *and* $h(y) = I$*, then* $\rho_\sigma(g) = \Sigma$ *and* $\rho_\sigma(h) = I$*.*
(b) *Suppose* $g(y) = \left|y^j\right|^r$ *then* $\rho_\sigma(g) = \mu_r \sigma_j^r$*, where* σ_j^2 *is the* j,j*-th element of* Σ*,* $\mu_r = \mathrm{E}(|u|^r)$ *and* $u \sim N(0,1)$*.*

This setup is sufficient for the proof of Theorem 1.2 of Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005), which is restated here.

[1] An example of a continuous local martingale which has no SV representation is a time-change Brownian motion where the time-change takes the form of the so-called "devil's staircase," which is continuous and non-decreasing but not absolutely continuous (see, for example, Munroe (1953, Section 27)). This relates to the work of, for example, Calvet and Fisher (2002) on multifractals.

Theorem 1 *Under (H) and assuming g and h satisfy (K) we have that*

$$Y^n(g,h)_t \xrightarrow{p} Y(g,h)_t := \int_0^t \rho_{\sigma_u}(g)\rho_{\sigma_u}(h)\mathrm{d}u, \tag{6}$$

where the convergence is also locally uniform in time.

The result is quite clean as it is requires no additional assumptions on Y and so is very close to dealing with the whole class of financially coherent continuous sample path processes.

Theorem 1 covers a number of existing setups which are currently receiving a great deal of attention as measures of variation in financial econometrics. Here we briefly discuss some of the work which has studied the limiting behaviour of these objects.

Example 3 ***(Example 1(a) continued)****. Then $g(y) = \left(y^j\right)^2$ and $h(y) = 1$, so (6) becomes*

$$\sum_{i=1}^{\lfloor nt \rfloor} \left(\Delta_i^n Y^j\right)^2 \xrightarrow{p} \int_0^t \sigma_{j,u}^2 \mathrm{d}u = [Y^j]_t,$$

the quadratic variation (QV) of Y^j. This well known result in probability theory is behind much of the modern work on realised volatility, which is compactly reviewed in Barndorff-Nielsen and Shephard (2005b).

*(**Example 1(b) continued**). As $g(y) = yy'$ and $h(y) = I$, then*

$$\sum_{i=1}^{\lfloor nt \rfloor} \left(\Delta_i^n Y\right)\left(\Delta_i^n Y\right)' \xrightarrow{p} \int_0^t \Sigma_u \mathrm{d}u = [Y]_t,$$

the well known multivariate version of QV.

(Example 1(c) continued). *As $g(y) = \left|y^j\right|^r$ and $h(y) = 1$ so*

$$n^{-1+r/2} \sum_{i=1}^{\lfloor nt \rfloor} \left|\Delta_i^n Y^j\right|^r \xrightarrow{p} \mu_r \int_0^t \sigma_{j,u}^r \mathrm{d}u.$$

This result is due to Jacod (1994) and Barndorff-Nielsen and Shephard (2003).

(Example 1(d) continued). *As $g(y) = \left|y^j\right|^r$ and $h(y) = \left|y^j\right|^s$ for $r, s > 0$, so*

$$n^{-1+(r+s)/2} \sum_{i=1}^{\lfloor nt \rfloor} \left|\Delta_i^n Y^j\right|^r \left|\Delta_{i+1}^n Y^j\right|^s \xrightarrow{p} \mu_r \mu_s \int_0^t \sigma_{j,u}^{r+s} \mathrm{d}u,$$

a result due to Barndorff-Nielsen and Shephard (2004b), who derived it under stronger conditions than those used here.

(Example 1(e) continued). *As*

$$g(y) = \begin{pmatrix} \left|y^j\right| & 0 \\ 0 & \left(y^j\right)^2 \end{pmatrix}, \quad h(y) = \begin{pmatrix} \left|y^j\right| \\ 1 \end{pmatrix},$$

so

$$\begin{pmatrix} \sum_{i=1}^{\lfloor nt \rfloor} \left|\Delta_i^n Y^j\right| \left|\Delta_{i+1}^n Y^j\right| \\ \sum_{i=1}^{\lfloor nt \rfloor} \left(\Delta_i^n Y^j\right)^2 \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \mu_1^2 \\ 1 \end{pmatrix} \int_0^t \sigma_{j,u}^2 \mathrm{d}u.$$

Barndorff-Nielsen and Shephard (2006) used this type of result to test for jumps as this particular bipower variation is robust to jumps.

4 Central limit theorem

4.1 Further assumptions on the process

It is important to be able to quantify the difference between the estimator $Y^n(g,h)$ and $Y(g,h)$. In this subsection we do this by giving a central limit theorem for $\sqrt{n}(Y^n(g,h) - Y(g,h))$. We have to make some stronger assumptions both on the process Y and on the pair (g,h) in order to derive this result.

We start with a variety of assumptions which strengthen (H) and (K) given in the previous subsection.

Assumption (H1): We have (H) with

$$\begin{aligned} \sigma_t &= \sigma_0 + \int_0^t a_u^* \mathrm{d}u + \int_0^t \sigma_{u-}^* \mathrm{d}W_u + \int_0^t v_{u-}^* \mathrm{d}V_u \\ &\quad + \int_0^t \int_E \varphi \circ w(u-,x) \left(\mu - \nu\right)(\mathrm{d}u, \mathrm{d}x) + \int_0^t \int_E \left(w - \varphi \circ w\right)(u-,x) \, \mu\left(\mathrm{d}u, \mathrm{d}x\right). \end{aligned} \tag{7}$$

Here a^*, σ^*, v^* are adapted càdlàg arrays, with a^* also being predictable and locally bounded. V is a d''-dimensional Brownian motion independent of W. μ is a Poisson measure on $(0,\infty) \times E$ independent of W and V, with intensity measure $\nu(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}t \otimes F(\mathrm{d}x)$ and F is a σ-finite measure on the Polish space $(E, \mathcal{E})$. φ is a continuous truncation function on $R^{dd'}$ (a function with compact support, which coincide with the identity map on the neighbourhood of 0). Finally $w(\omega, u, x)$ is a map $\Omega \times [0,\infty) \times E$ into the space of $d \times d'$arrays which is $\mathcal{F}_u \otimes \mathcal{E}-$measurable in (ω, x) for all u and càdlàg in u, and such that for some sequences (S_k) of stopping times increasing to $+\infty$ we have

$$\sup_{\omega \in \Omega, u < S_k(\omega)} \|w(\omega,u,x)\| \leq \psi_k(x) \quad \text{where} \quad \int_E \left(1 \wedge \psi_k(x)^2\right) F(\mathrm{d}x) < \infty.$$

Assumption (H2): $\Sigma = \sigma\sigma'$ is everywhere invertible.

Remark 4 *If there were no jumps in the volatility then it would be sufficient to employ*

$$\sigma_t = \sigma_0 + \int_0^t a_u^* \mathrm{d}u + \int_0^t \sigma_{u-}^* \mathrm{d}W_u + \int_0^t v_{u-}^* \mathrm{d}V_u, \tag{8}$$

which is covered by (H1). The assumption (H1) is rather general from an econometric viewpoint as it allows for flexible leverage effects, multifactor volatility effects, jumps, non-stationarities, intraday effects, etc. Indeed we do not know of a continuous time volatility model used in financial economics which is outside this class.

Assumption (H1) looks quite complicated and one might wonder if a simpler assumption could have been used whose jumps enter through a stochastic integral with a Lévy integrator. However, such a condition is somewhat unsatisfactory for that alternative class of processes is not closed under squaring (further comment on this is given in Section 9.2). Hence we have chosen to use Assumption (H1) for it can be applied equally to σ and $\Sigma = \sigma\sigma'$.

4.2 Further assumptions on g and h

In order to derive a central limit theorem we need to impose some regularity on g and h.

Assumption (K1): f is even (that is $f(x) = f(-x)$ for $x \in R^d$) and continuously differentiable, with derivatives having at most polynomial growth.

In order to handle some of the most interesting cases of bipower variation, where we are mostly interested in taking low powers of absolute values of returns which may not be differentiable at zero, we sometimes need to relax (K1). The resulting condition is quite technical and is called (K2). It is discussed in the Appendix.

Assumption (K2): f is even and continuously differentiable on the complement B^c of a closed subset $B \subset \mathbb{R}^d$ and satisfies

$$||y|| \leq 1 \Longrightarrow |f(x+y) - f(x)| \leq C(1 + ||x||^p)||y||^r$$

for some constants C, $p \geq 0$ and $r \in (0, 1]$. Moreover

a) If $r = 1$ then B has Lebesgue measure 0.

b) If $r < 1$ then B satisfies

$$\left.\begin{array}{l} \text{for any positive definite } d \times d \text{ matrix } C \text{ and} \\ \text{any } N(0, C)\text{-random vector } U \text{ the distance } d(U, B) \\ \text{from } U \text{ to } B \text{ has a density } \psi_C \text{ on } R_+, \text{ such that} \\ sup_{x \in R_+, |C| + |C^{-1}| \leq A} \psi_C(x) < \infty \text{ for all } A < \infty, \end{array}\right\} \tag{9}$$

and we have

$$x \in B^c,\ \|y\| \leq 1 \bigwedge \frac{d(x,B)}{2} \quad \Rightarrow \quad \left\{ \begin{array}{l} \|\nabla f(x)\| \leq \frac{C(1+\|x\|^p)}{d(x,B)^{1-r}}, \\ \|\nabla f(x+y) - \nabla f(x)\| \leq \frac{C(1+\|x\|^p)\|y\|}{d(x,B)^{2-r}}. \end{array}\right. \tag{10}$$

Remark 5 *These conditions accommodate the case where f equals $\left|x^j\right|^r$: this function satisfies (K1) when $r > 1$, and (K2) when $r \in (0, 1]$ (with the same r of course). When B is a finite* *union of hyperplanes it satisfies (9). Also, observe that (K1) implies (K2) with $r = 1$ and $B = \emptyset$. Assumption K1 will often be enough to cover many cases of functions with regularly varying properties, as long as they are even, for regularly varying functions are bounded by members of the polynomial at infinity class.*

4.3 Main asymptotic result

Each of the following assumptions (J1) and (J2) are sufficient for the statement of Theorem 1.3 of Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) to hold.
Assumption (J1): We have (H1) and g and h satisfy (K1).
Assumption (J2): We have (H1), (H2) and g and h satisfy (K2).

Clearly J2 makes stronger assumptions about the volatility process and weaker assumptions about the functions g and h, than assumption J1. It is J2 which will be used when analysing the interesting low power versions of bipower variation.

The result of the Theorem is restated in the following.

Theorem 2 *Assume at least one of (J1) and (J2) holds, then the process*

$$\sqrt{n}\ \left(Y^n(g,h)_t - Y(g,h)_t\right)$$

converges in law stably towards a limiting process $U(g,h)$ having the form

$$U(g,h)_t^{jk} = \sum_{j'=1}^{d_1}\sum_{k'=1}^{d_3}\int_0^t \alpha(\sigma_u,g,h)^{jk,j'k'}\ \mathrm{d}B_u^{j',k'}, \tag{11}$$

where

$$\sum_{l=1}^{d_1}\sum_{m=1}^{d_3}\alpha(\sigma,g,h)^{jk,lm}\alpha(\sigma,g,h)^{j'k',lm} = A(\sigma,g,h)^{jk,j'k'},$$

and

$$\begin{aligned} A(\sigma,g,h)^{jk,j'k'} &= \sum_{l=1}^{d_2}\sum_{l'=1}^{d_2}\Big\{\rho_\sigma\left(g^{jl}g^{j'l'}\right)\rho_\sigma\left(h^{lk}h^{l'k'}\right) + \rho_\sigma\left(g^{jl}\right)\rho_\sigma\left(h^{l'k'}\right)\rho_\sigma\left(g^{j'l'}h^{lk}\right) \\ &\quad +\rho_\sigma\left(g^{j'l'}\right)\rho_\sigma\left(h^{lk}\right)\rho_\sigma\left(g^{jl}h^{l'k'}\right) \\ &\quad -3\rho_\sigma\left(g^{jl}\right)\rho_\sigma\left(g^{j'l'}\right)\rho_\sigma\left(h^{lk}\right)\rho_\sigma\left(h^{l'k'}\right)\Big\}. \end{aligned}$$

Furthermore, B is a standard Wiener process which is defined on an extension of $\left(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},P\right)$ and is independent of the σ–field $\mathcal{F}$.

Remark 6 *The concept and role of stable convergence may be unfamiliar to some readers and we therefore add some words of explanation. In the simplest case of stable convergence of sequences* *of random variables, rather than processes, the concise mathematical definition is as follows. Let X_n denote a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Then we say that X_n* converges stably in law *if there exists a probability measure μ on $(\Omega \times \mathbb{R}, \mathcal{F} \times \mathcal{B})$ (where $\mathcal{B}$ denotes the Borel σ-algebra on $\mathbb{R}$) such that for every bounded random variable Z on $(\Omega, \mathcal{F}, P)$ and every bounded and continuous function g on $\mathbb{R}$ we have that, for $n \to \infty$,*

$$\mathrm{E}\left(Zg\left(X_n\right)\right) \to \int Z\left(\omega\right) g\left(x\right) \mu\left(\mathrm{d}\omega, \mathrm{d}x\right).$$

If X_n converges stably in law then, in particular, it converges in distribution (or in law or weak convergence), the limiting law being $\mu\left(\Omega, \cdot\right)$. Accordingly, one says that X_n converges stably to some random variable X if there exists a probability measure μ as above such that X has law $\mu\left(\Omega, \cdot\right)$. This concept and its extension to stable convergence of processes is discussed in Jacod and Shiryaev (2003, pp. 512-518). For earlier expositions, see Hall and Heyde (1980, pp. 56-58) and Jacod (1997). An early use of this concept in econometrics was Phillips and Ouliaris (1990) in their work on the limit distribution of cointegration tests.

However, this formalisation does not reveal the key nature of stable convergence which is that $X_n \to X$ stably implies that for any random variable Z, the pair (Z, X_n) converges in law to (Z, X). In the context the present paper consider the following simple example of the above result. Let

$$X_n = \sqrt{n}\left(\sum_{i=1}^{\lfloor nt \rfloor}\left(\Delta_i^n Y^j\right)^2 - \int_0^t \sigma_{j,u}^2 \mathrm{d}u\right)$$

and

$$Z = \sqrt{\int_0^t \sigma_{j,u}^4 \mathrm{d}u}.$$

Our focus is on $X_n/\sqrt{Z}$ and our convergence in law stably implies that

$$\sqrt{n}\left(\sum_{i=1}^{\lfloor nt \rfloor}\left(\Delta_i^n Y^j\right)^2 - \int_0^t \sigma_{j,u}^2 \mathrm{d}u\right) / \sqrt{\int_0^t \sigma_{j,u}^4 \mathrm{d}u} \overset{law}{\to} N(0,2). \tag{12}$$

Without the convergence in law stably, (12) could not be deduced.

Corollary 1 *Suppose $d_3 = 1$, which is the situation looked at in Example 1(e). Then $Y^n(g,h)_t$ is a vector and so the limiting law of $\sqrt{n}(Y^n(g,h) - Y(g,h))$ simplifies. It takes on the form of*

$$U(g,h)_t^j = \sum_{j'=1}^{d_1} \int_0^t \alpha(\sigma_u, g, h)^{j,j'} \, \mathrm{d}B_u^{j'}, \tag{13}$$

where

$$\sum_{l=1}^{d_1} \alpha(\sigma, g, h)^{j,l} \alpha(\sigma, g, h)^{j',l} = A(\sigma, g, h)^{j,j'}.$$

Here

$$A(\sigma, g, h)^{j,j'} = \sum_{l=1}^{d_2}\sum_{l'=1}^{d_2}\left\{\rho_\sigma(g^{jl}g^{j'l'})\rho_\sigma(h^l h^{l'}) + \rho_\sigma(g^{jl})\rho_\sigma(h^{l'})\rho_\sigma(g^{j'l'}h^l)\right.$$
$$\left. + \rho_\sigma(g^{j'l'})\rho_\sigma(h^l)\rho_\sigma(g^{jl}h^{l'}) - 3\rho_\sigma(g^{jl})\rho_\sigma(g^{j'l'})\rho_\sigma(h^l)\rho_\sigma(h^{l'})\right\}.$$

In particular, for a single point in time t,

$$\sqrt{n}\ (Y^n(g,h)_t - Y(g,h)_t) \xrightarrow{L} MN\left(0, \int_0^t A(\sigma_u, g, h)\mathrm{d}u\right),$$

where MN denotes a mixed Gaussian distribution and $A(\sigma, g, h)$ denotes a matrix whose j, j'-th element is $A(\sigma, g, h)^{j,j'}$.

Remark 7 *Suppose $g(y) = I$, then A becomes*

$$A(\sigma, g, h)^{jk,j'k'} = \rho_\sigma(h^{jk}h^{j'k'}) - \rho_\sigma(h^{jk})\rho_\sigma(h^{j'k'}).$$

Example 4 *Suppose $d_1 = d_2 = d_3 = 1$, then*

$$U(g,h)_t = \int_0^t \sqrt{A(\Sigma_u, g, h)}\ \mathrm{d}B_u, \tag{14}$$

where

$$A(\sigma, g, h) = \rho_\sigma(gg)\rho_\sigma(hh) + 2\rho_\sigma(g)\rho_\sigma(h)\rho_\sigma(gh) - 3\left\{\rho_\sigma(g)\rho_\sigma(h)\right\}^2.$$

We consider two concrete examples of this setup.
(i) *Power variation. Suppose $g(y) = \left|y^j\right|^r$ and $h(y) = 1$ where $r > 0$, then $\rho_\sigma(h) = 1$,*

$$\rho_\sigma(g) = \rho_\sigma(gh) = \mu_r\sigma_j^r, \quad \rho_\sigma(gg) = \mu_{2r}\sigma_j^{2r}.$$

This implies that

$$\begin{aligned} A(\sigma, g, h) &= \mu_{2r}\sigma_j^{2r} + 2\mu_r^2\sigma_j^{2r} - 3\mu_r^2\sigma_j^{2r} \\ &= \left(\mu_{2r} - \mu_r^2\right)\sigma_j^{2r} \\ &= v_r\sigma_j^{2r}, \end{aligned}$$

where $v_r = \mathrm{Var}(|u|^r)$ and $u \sim N(0,1)$. When $r = 2$, this yields a central limit theorem for the realised quadratic variation process, with

$$U(g,h)_t = \int_0^t \sqrt{2\sigma_{j,u}^4}\ \mathrm{d}B_u, \tag{15}$$

a result which appears in Jacod (1994), Mykland and Zhang (2006) and, implicitly, Jacod and Protter (1998), while the case of a single value of t appears in Barndorff-Nielsen and Shephard

(2002). For the more general case of $r > 0$ Barndorff-Nielsen and Shephard (2003) derived, under much stronger conditions, a central limit theorem for $U(g,h)_1$. Their result ruled out leverage effects, which are allowed under Theorem 2. The finite sample behaviour of this type of limit theory is studied in, for example, Barndorff-Nielsen and Shephard (2005a), Goncalves and Meddahi (2004) and Nielsen and Frederiksen (2005) in the absence of market frictions.

(ii) *Bipower variation. Suppose $g(y) = \left|y^j\right|^r$ and $h(y) = \left|y^j\right|^s$ where $r, s > 0$, then*

$$\begin{aligned} \rho_\sigma(g) &= \mu_r \sigma_j^r, \quad \rho_\sigma(h) = \mu_s \sigma_j^s, \quad \rho_\sigma(gg) = \mu_{2r}\sigma_j^{2r}, \\ \rho_\sigma(hh) &= \mu_{2s}\sigma_j^{2s}, \quad \rho_\sigma(gh) = \mu_{r+s}\sigma_j^{r+s}. \end{aligned}$$

This implies that

$$\begin{aligned} A(\sigma, g, h) &= \mu_{2r}\sigma_j^{2r}\mu_{2s}\sigma_j^{2s} + 2\mu_r\sigma_j^r\mu_s\sigma_j^s\mu_{r+s}\sigma_j^{r+s} - 3\mu_r^2\sigma_j^{2r}\mu_s^2\sigma_j^{2s} \\ &= \left(\mu_{2r}\mu_{2s} + 2\mu_{r+s}\mu_r\mu_s - 3\mu_r^2\mu_s^2\right)\sigma_j^{2r+2s}. \end{aligned}$$

In the $r = s = 1$ case Barndorff-Nielsen and Shephard (2006) derived, under much stronger conditions, a central limit theorem for $U(g,h)_1$. Their result ruled out leverage effects, which are allowed under Theorem 2. In that special case, writing

$$\vartheta = \frac{\pi^2}{4} + \pi - 5,$$

we have

$$U(g,h)_t = \mu_1^2 \int_0^t \sqrt{(2+\vartheta)\,\sigma_{j,u}^4}\,\mathrm{d}B_u.$$

In the case where $r = \varepsilon$, $s = 2 - \varepsilon$ where $2 > \varepsilon > 0$ then $Y(g,h)_t = \mu_\varepsilon\mu_{2-\varepsilon}\int_0^t \sigma_u^2 \mathrm{d}u$ and the statistic is asymptotically robust to finite activity jumps (Barndorff-Nielsen and Shephard (2004b)). For arbitrarily small ε the error process $U(g,h)_t$ is close to (15), so this jump robust process is basically as efficient as if there are no jumps in the process.

Example 5 *Suppose $g(y) = yy'$, $h = I$. Then we have to calculate*

$$A(\sigma, g, h)^{jk,j'k'} = \rho_\sigma(g^{jk}g^{j'k'}) - \rho_\sigma(g^{jk})\rho_\sigma(g^{j'k'}).$$

However,

$$\rho_\sigma(g^{jk}) = \Sigma^{jk}, \quad \rho_\sigma(g^{jk}g^{j'k'}) = \Sigma^{jk}\Sigma^{j'k'} + \Sigma^{jj'}\Sigma^{kk'} + \Sigma^{jk'}\Sigma^{kj'},$$

so

$$\begin{aligned} A(\sigma, g, h)^{jk,j'k'} &= \Sigma^{jk}\Sigma^{j'k'} + \Sigma^{jj'}\Sigma^{kk'} + \Sigma^{jk'}\Sigma^{kj'} - \Sigma^{jk}\Sigma^{j'k'} \\ &= \Sigma^{jj'}\Sigma^{kk'} + \Sigma^{jk'}\Sigma^{kj'}. \end{aligned}$$

This is the result found in Barndorff-Nielsen and Shephard (2004a), but proved there under stronger conditions. The result is, in fact, implicit in the work of Jacod and Protter (1998).

Example 6 *Suppose $d_1 = d_2 = 2$, $d_3 = 1$ and g is diagonal. Then*

$$U(g,h)^j_t = \sum_{j'=1}^{2} \int_0^t \alpha(\sigma_u, g, h)^{j,j'} \, \mathrm{d}B^{j'}_u, \tag{16}$$

where

$$\sum_{l=1}^{2} \alpha(\sigma,g,h)^{j,l} \alpha(\sigma,g,h)^{j',l} = A(\sigma,g,h)^{j,j'}.$$

Here

$$\begin{aligned} A(\sigma,g,h)^{j,j'} &= \rho_\sigma(g^{jj}g^{j'j'})\rho_\sigma(h^j h^{j'}) + \rho_\sigma(g^{jj})\rho_\sigma(h^{j'})\rho_\sigma(g^{j'j'}h^j) \\ &\quad + \rho_\sigma(g^{j'j'})\rho_\sigma(h^j)\rho_\sigma(g^{jj}h^{j'}) - 3\rho_\sigma(g^{jj})\rho_\sigma(g^{j'j'})\rho_\sigma(h^j)\rho_\sigma(h^{j'}). \end{aligned}$$

Example 7 *Joint behaviour of realised QV and realised bipower variation. This sets*

$$g(y) = \begin{pmatrix} |y^j| & 0 \\ 0 & 1 \end{pmatrix}, \quad h(y) = \begin{pmatrix} |y^j| \\ (y^j)^2 \end{pmatrix}.$$

The implication is that

$$\rho_\sigma(g^{11}) = \rho_\sigma(g^{22}g^{11}) = \rho_\sigma(g^{11}g^{22}) = \mu_1\sigma_j, \ \rho_\sigma(g^{22}) = 1, \ \rho_\sigma(g^{11}g^{11}) = \sigma_j^2, \ \rho_\sigma(g^{22}g^{22}) = 1,$$

$$\rho_\sigma(h^1) = \mu_1\sigma_j, \ \rho_\sigma(h^2) = \rho_\sigma(h^1h^1) = \sigma_j^2, \ \rho_\sigma(h^1h^2) = \rho_\sigma(h^2h^1) = \mu_3\sigma_j^3, \ \rho_\sigma(h^2h^2) = 3\sigma_j^4,$$

$$\rho_\sigma(g^{11}h^1) = \sigma_j^2, \ \rho_\sigma(g^{11}h^2) = \mu_3\sigma_j^3, \ \rho_\sigma(g^{22}h^1) = \mu_1\sigma_j, \ \rho_\sigma(g^{22}h^2) = \sigma_j^2.$$

Thus

$$\begin{aligned} A(\sigma,g,h)^{1,1} &= \sigma_j^2\sigma_j^2 + 2\mu_1\sigma_j\mu_1\sigma_j\sigma_j^2 - 3\mu_1\sigma_j\mu_1\sigma_j\mu_1\sigma_j\mu_1\sigma_j \\ &= \sigma_j^4\left(1 + 2\mu_1^2 - 3\mu_1^4\right) = \mu_1^4(2+\vartheta)\sigma_j^4, \end{aligned}$$

while

$$A(\sigma,g,h)^{2,2} = 3\sigma_j^4 + 2\sigma_j^4 - 3\sigma_j^4 = 2\sigma_j^4,$$

and

$$\begin{aligned} A(\sigma,g,h)^{1,2} &= \mu_1\sigma_j\mu_3\sigma_j^3 + \mu_1\sigma_j\sigma_j^2\mu_1\sigma_j + \mu_1\sigma_j\mu_3\sigma_j^3 - 3\mu_1\sigma_j\mu_1\sigma_j\sigma_j^2 \\ &= 2\sigma_j^4\left(\mu_1\mu_3 - \mu_1^2\right) = 2\mu_1^2\sigma_j^4. \end{aligned}$$

This generalises the result given in Barndorff-Nielsen and Shephard (2006) to the leverage case. In particular we have that

$$\begin{pmatrix} U(g,h)^1_t \\ U(g,h)^2_t \end{pmatrix} = \begin{pmatrix} \mu_1^2 \int_0^t \sqrt{2\sigma_u^4}\mathrm{d}B^1_u + \mu_1^2 \int_0^t \sqrt{\vartheta\sigma_u^4}\mathrm{d}B^2_u \\ \int_0^t \sqrt{2\sigma_u^4}\mathrm{d}B^1_u. \end{pmatrix}$$

5 Multipower variation

A natural extension of generalised bipower variation is to generalised multipower variation

$$Y^n(g)_t = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor}\left\{\prod_{i'=1}^{I\wedge(i+1)} g_{i'}(\sqrt{n}\ \Delta_{i-i'+1}^n Y)\right\},$$

where $a \wedge b$ denotes the minimum of a and b. This measure of variation, for the $g_{i'}$ being absolute powers, was introduced by Barndorff-Nielsen and Shephard (2006).

We will be interested in studying the properties of $Y^n(g)_t$ for given functions $\{g_i\}$ with the following properties.

Assumption (K*): All the $\{g_i\}$ are continuous with at most polynomial growth.

The previous results suggests that if Y is a Brownian semimartingale and Assumption (K*) holds then

$$Y^n(g)_t \xrightarrow{p} Y(g)_t := \int_0^t \prod_{i=0}^{I} \rho_{\sigma_u}(g_i)\mathrm{d}u.$$

See Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) for more details.

Example 8 ***(a)*** *Suppose $I = 4$ and $g_i(y) = \left|y^j\right|$, then $\rho_\sigma(g_i) = \mu_1\sigma_j$ so*

$$Y(g)_t = \mu_1^4 \int_0^t \sigma_{j,u}^4 \mathrm{d}u,$$

a scaled version of integrated quarticity.
(b) *Suppose $I = 3$ and $g_i(y) = \left|y^j\right|^{4/3}$, then*

$$\rho_\sigma(g_i) = \mu_{4/3}\sigma_j^{4/3}$$

so

$$Y(g)_t = \mu_{4/3}^3 \int_0^t \sigma_{j,u}^4 \mathrm{d}u.$$

Example 9 *Of some importance is the generic case where $g_i(y) = \left|y^j\right|^{2/I}$, which implies*

$$Y(g)_t = \mu_{2/I}^I \int_0^t \sigma_{j,u}^2 \mathrm{d}u.$$

Thus this class provides an interesting alternative to realised variance as an estimator of integrated variance. Of course it is important to know a central limit theory for these types of quantities. Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) show that when (H1) and (H2) hold then

$$\sqrt{n}\left[Y^n(g)_t - Y(g)_t\right] \to \int_0^t \sqrt{\omega_I^2 \sigma_{j,u}^4}\ \mathrm{d}B_u,$$

where

$$\omega_I^2 = \mathrm{Var}\left(\prod_{i=1}^{I} |u_i|^{2/I}\right) + 2\sum_{j=1}^{I-1} \mathrm{Cov}\left(\prod_{i=1}^{I} |u_i|^{2/I}, \prod_{i=1}^{I} |u_{i-j}|^{2/I}\right),$$

with $u_i \sim NID(0,1)$. *Thus the asymptotic variance is again a scaled version of integrated quarticity. Clearly* $\omega_1^2 = 2$, *while recalling that* $\mu_1 = \sqrt{2/\pi}$,

$$\begin{aligned}\omega_2^2 &= \mathrm{Var}(|u_1|\,|u_2|) + 2\mathrm{Cov}(|u_1|\,|u_2|, |u_2|\,|u_3|) \\ &= 1 + 2\mu_1^2 - 3\mu_1^4,\end{aligned}$$

and

$$\begin{aligned}\omega_3^2 &= \mathrm{Var}((|u_1|\,|u_2|\,|u_3|)^{2/3}) + 2\mathrm{Cov}((|u_1|\,|u_2|\,|u_3|)^{2/3}, (|u_2|\,|u_3|\,|u_4|)^{2/3}) \\ &\quad +2\mathrm{Cov}((|u_1|\,|u_2|\,|u_3|)^{2/3}, (|u_3|\,|u_4|\,|u_5|)^{2/3}) \\ &= \left(\mu_{4/3}^3 - \mu_{2/3}^6\right) + 2\left(\mu_{4/3}^2\mu_{2/3}^2 - \mu_{2/3}^6\right) + 2\left(\mu_{4/3}\mu_{2/3}^4 - \mu_{2/3}^6\right).\end{aligned}$$

6 Sums of realised generalised bipower

The law of large numbers and the central limit theorem also hold for linear combinations of processes like $Y(g)$ above.

Example 10 *Let* ζ_i^n *the* $d \times d$ *matrix whose* (k,l) *entry is* $\sum_{j=0}^{d-1} \Delta_{i+j}^n Y^k \Delta_{i+j}^n Y^l$. *Then*

$$Z_t^n = \frac{n^{d-1}}{d!} \sum_{i=1}^{[nt]} \det(\zeta_i^n)$$

is a linear combinations of processes $Y^n(g)$ *for functions* g_l *being of the form* $g_l(y) = y^j y^k$. *It is proved in Jacod, Lejay, and Talay (2005) that under (H)*

$$Z_t^n \xrightarrow{p} Z_t := \int_0^t \det(\sigma_u \sigma_u')\mathrm{d}u,$$

whereas under (H1) and (H2) the associated CLT is the following convergence in law:

$$\sqrt{n}(Z_t^n - Z_t) \to \int_0^t \sqrt{\Gamma(\sigma_u)}\, \mathrm{d}B_u,$$

where $\Gamma(\sigma)$ *denotes the covariance of the variable* $\det(\zeta)/d!$, *and* ζ *is a* $d \times d$ *matrix whose* (k,l) *entry is* $\sum_{j=0}^{d-1} U_j^k U_j^l$ *and the* U_j*'s are i.i.d. centered Gaussian vectors with covariance* $\sigma\sigma'$.

This kind of result may be used for testing whether the rank of the diffusion coefficient is everywhere smaller than d. *In that case one could use a model with a* $d' < d$ *for the dimension of the driving Wiener process* W.

7 Conclusion

This paper provides some rather general limit results for realised generalised bipower variation. In the case of power variation and bipower variation the results are proved under much weaker assumptions than those which have previously appeared in the literature. In particular the no-leverage assumption is removed, which is important in the application of these results to stock data.

There are a number of open questions. It is rather unclear how econometricians might exploit the generality of the g and h functions to learn about interesting features of the variation of price processes. It would be interesting to know what properties g and h must possess in order for these statistics to be robust to finite activity and infinite activity jumps.

It would be attractive to extend the analysis to allow g and h to depend upon the entire path of Y, not just returns, and to depend upon n. This would allow, respectively, the theory to additionally cover the realised range process studied by Christensen and Podolskij (2005) and the truncated estimator studied by Mancini (2004) and more recently by Aït-Sahalia and Jacod (2005).

A challenging extension is to construct a version of realised generalised bipower variation which is robust to market frictions. Following the work on the realised volatility there are two strategies which may be able to help: the kernel based approach, studied in detailed by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004), and the subsampling approach of Zhang, Mykland, and Aït-Sahalia (2005) and Zhang (2004). In the realised volatility case these methods are basically equivalent, however it is perhaps the case that the subsampling method is easier to extend to the non-quadratic case. Further insights into the choice of n may be possible using mean square error based optimal sampling developed by Bandi and Russell (2003) and Hansen and Lunde (2006) for realised variance.

8 Acknowledgments

Ole E. Barndorff-Nielsen's work is supported by CAF (`www.caf.dk`), which is funded by the Danish Social Science Research Council. Neil Shephard's research is supported by the UK's ESRC through the grant "High frequency financial econometrics based upon power variation." We thank the editor, Peter Phillips, and the referees for their stimulating comments on an earlier version.

9 Techniques for the Proof of Theorem 2

9.1 Notational conventions

Below we give a fairly detailed account of the basic techniques in the proof of Theorem 2, in the one-dimensional case and under some relatively minor simplifying assumptions. Throughout we set $h = 1$ for the main difficulty in the proof is being able to deal with the generality in the g function. Once that has been mastered the extension to the bipower measure is not a large obstacle. We refer the reader to Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005) for readers who wish to see the more general case. The outline of this section is as follows. First we introduce our basic notation, while in subsection 9.2 we set out the model and review the assumptions we use. In subsection 9.3 we state the theorem we will prove and outline the steps in the proof. Subsections 9.5, 9.6 and 9.7 give the proofs of the successive steps.

All processes mentioned in the following are defined on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. We shall in general use standard notation and conventions. For instance, given a process (Z_t) we write $\triangle_i^n Z := Z_{\frac{i}{n}} - Z_{\frac{i-1}{n}}$, $i, n \geq 1$.

All results will be proved using convergence 'stably in law' of sequences of càdlàg processes, which is a slightly stronger notion than convergence in law (cf. Remark 6 above). For this we shall use the notation

$$(Z_t^n) \to (Z_t),$$

where (Z_t^n) and (Z_t) are given càdlàg processes. Furthermore we shall write

$$(Z_t^n) \xrightarrow{P} 0 \quad \text{meaning} \quad \sup_{0 \leq s \leq t} |Z_s^n| \to 0 \ \text{ in probability for all } t \geq 0,$$

$$(Z_t^n) \xrightarrow{P} (Z_t) \quad \text{meaning} \quad (Z_t^n - Z_t) \xrightarrow{P} 0.$$

Often

$$Z_t^n = \sum_{i=1}^{[nt]} a_i^n \quad \text{for all } t \geq 0,$$

where the a_i^n's are $\mathcal{F}_{\frac{i-1}{n}}$-measurable. Recall here that given càdlàg processes (Z_t^n), (Y_t^n) and (Z_t) we have

$$(Z_t^n) \to (Z_t) \ \text{ if } \ (Z_t^n - Y_t^n) \xrightarrow{P} 0 \ \text{ and } \ (Y_t^n) \to (Z_t).$$

Moreover, for $h : \mathbf{R} \to \mathbf{R}$ Borel measurable of at most polynomial growth we note that $x \mapsto \rho_x(h)$ is locally bounded and continuous if h is continuous at 0.

In what follows many arguments will consist of a series of estimates of terms indexed by i, n and t. In these estimates we shall denote by C a finite constant which may vary from place to place. Its value will depend on the constants and quantities appearing in the assumptions of the model but it is always independent of i, n and t.

9.2 Model and basic assumptions

Throughout the following (W_t) denotes a $((\mathcal{F}_t), P)$-Wiener process and (σ_t) a given càdlàg $(\mathcal{F}_t)$-adapted process. Define the local martingale

$$Y_t := \int_0^t \sigma_{s-}\,\mathrm{d}W_s \qquad t \geq 0.$$

We have deleted the drift of the (Y_t) process as taking care of it is a simple technical task, while its presence increase the clutter of the notation. Our aim is to study the asymptotic behaviour of the processes

$$\{(Y_t^n(g))\,|\,n \geq 1\}$$

where

$$Y_t^n(g) = \frac{1}{n}\sum_{i=1}^{[nt]} g(\sqrt{n}\,\triangle_i^n Y), \quad t \geq 0,\, n \geq 1.$$

Here $g : \mathbf{R} \to \mathbf{R}$ is a given continuous function of at most polynomial growth. We are especially interested in g's of the form $x \mapsto |x|^r$ $(r > 0)$ but we shall keep the general notation since nothing is gained in simplicity by assuming that g is of power form. Throughout the following we shall assume that g furthermore satisfies the following.

Assumption *(Ka)*: *g is an even function and continuously differentiable in B^c where $B \subseteq \mathbf{R}$ is a closed Lebesgue null-set and $\exists$ M, $p \geq 1$ such that*

$$|g(x+y) - g(x)| \leq M(1 + |x|^p + |y|^p) \cdot |y| \ ,$$

for all $x, y \in \mathbf{R}$.

Remark 8 *The assumption (Ka) implies, in particular, that if $x \in B^c$ then*

$$|g'(x)| \leq M(1 + |x|^p).$$

Observe that only power functions corresponding to $r \geq 1$ do satisfy (Ka). The remaining case $0 < r < 1$ requires special arguments which will be omitted here (for details see Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005)).

In order to prove the CLT-theorem we need some additional structure on the volatility process (σ_t). A natural set of assumptions would be the following.

Assumption *(H0)*: *(σ_t) can be written as*

$$\sigma_t = \sigma_0 + \int_0^t a_s^*\,\mathrm{d}s + \int_0^t \sigma_{s-}^*\,\mathrm{d}W_s + \int_0^t v_{s-}^*\,\mathrm{d}Z_s$$

where (Z_t) is a $((\mathcal{F}_t), P)$-Lévy process independent of (W_t) and (σ_t^*) and (v_t^*) are adapted càdlàg processes and (a_t^*) a predictable locally bounded process.

However, in modelling volatility it is often more natural to define (σ_t^2) as being of the above form, i.e.

$$\sigma_t^2 = \sigma_0^2 + \int_0^t a_s^* \,\mathrm{d}s + \int_0^t \sigma_{s-}^* \,\mathrm{d}W_s + \int_0^t v_{s-}^* \,\mathrm{d}Z_s.$$

Now this does not in general imply that (σ_t) has the same form; therefore we shall replace (H0) by the more general structure given by the following assumption.

Assumption *(H1)*: *(σ_t) can be written, for $t \geq 0$, as*

$$\begin{aligned}\sigma_t &= \sigma_0 + \int_0^t a_s^* \,\mathrm{d}s + \int_0^t \sigma_{s-}^* \,\mathrm{d}W_s + \int_0^t v_{s-}^* \,\mathrm{d}V_s \\ &\quad + \int_0^t \int_E q \circ \phi(s-, x)\,(\mu - \nu)(\mathrm{d}s\,\mathrm{d}x) \\ &\quad + \int_0^t \int_E \{\phi(s-, x) - q \circ \phi(s-, x)\}\ \mu(\mathrm{d}s\,\mathrm{d}x).\end{aligned}$$

Here (a_t^*), (σ_t^*) and (v_t^*) are as in (H0) and (V_t) is another $((\mathcal{F}_t), P)$-Wiener process independent of (W_t) while q is a continuous truncation function on $\mathbf{R}$, i.e. a function with compact support coinciding with the identity on a neighbourhood of 0. Further μ is a Poisson random measure on $(0, \infty) \times E$ independent of (W_t) and (V_t) with intensity measure $\nu(\mathrm{d}s\,\mathrm{d}x) = \mathrm{d}s \otimes F(\mathrm{d}x)$, F being a σ-finite measure on a measurable space $(E, \mathcal{E})$ and

$$(\omega, s, x) \mapsto \phi(\omega, s, x)$$

is a map from $\Omega \times [0, \infty) \times E$ into $\mathbf{R}$ which is $\mathcal{F}_s \otimes \mathcal{E}$ measurable in (ω, x) for all s and càdlàg in s, satisfying furthermore that for some sequence of stopping times (S_k) increasing to $+\infty$ we have for all $k \geq 1$

$$\int_E \left\{1 \wedge \psi_k(x)^2\right\}\ F(\mathrm{d}x) < \infty,$$

where

$$\psi_k(x) = \sup_{\omega \in \Omega,\, s < S_k(\omega)} |\phi(\omega, s, x)|.$$

Remark 9 *(H1) is weaker than (H0), and if (σ_t^2) satisfies (H1) then so does (σ_t).*

Finally we shall also assume a non-degeneracy in the model.

Assumption *(H2)*: *(σ_t) satisfies $0 < \sigma_t^2(\omega)$ for all (t, ω).*

According to general stochastic analysis theory it is known that to prove convergence in law of a sequence (Z_t^n) of càdlàg processes it suffices to prove the convergence of each of the stopped processes $(Z^n_{T_k\wedge t})$ for at least one sequence of stopping times (T_k) increasing to $+\infty$. Applying this together with standard localisation techniques (for details see Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2005)), we may assume that the following more restrictive assumptions are satisfied.

Assumption *(H1a)*: *(σ_t) can be written as*

$$\sigma_t = \sigma_0 + \int_0^t a^*_s \, ds + \int_0^t \sigma^*_{s-} \, \mathrm{d}W_s + \int_0^t v^*_{s-} \, \mathrm{d}V_s + \int_0^t \int_E \phi(s-,x)(\mu-\nu)(\mathrm{d}s\,\mathrm{d}x) \quad t \geq 0.$$

Here (a^*_t), (σ^*_t) and (v^*_t) are real valued uniformly bounded càdlàg $(\mathcal{F}_t)$-adapted processes; (V_t) is another $((\mathcal{F}_t), P)$-Wiener process independent of (W_t). Further μ is a Poisson random measure on $(0,\infty)\times E$ independent of (W_t) and (V_t) with intensity measure $\nu(\mathrm{d}s\,\mathrm{d}x) = \mathrm{d}s \otimes F(\mathrm{d}x)$, F being a σ-finite measure on a measurable space $(E,\mathcal{E})$ and

$$(\omega, s, x) \mapsto \phi(\omega, s, x)$$

is a map from $\Omega \times [0,\infty) \times E$ into $\mathbf{R}$ which is $\mathcal{F}_s \otimes \mathcal{E}$ measurable in (ω, x) for all s and càdlàg in s, satisfying furthermore

$$\psi(x) = \sup_{\omega\in\Omega,\, s\geq 0} |\phi(\omega,s,x)| \leq M < \infty \quad \textit{and} \quad \int \psi(x)^2 \, F(\mathrm{d}x) < \infty.$$

Likewise, by a localisation argument, we may assume

Assumption *(H2a)*: *(σ_t) satisfies $a < \sigma^2_t(\omega) < b$ for all (t,ω) for some $a, b \in (0,\infty)$.*

Observe that under the more restricted assumptions (Y_t) is a continuous martingale having moments of all orders and (σ_t) is represented as a sum of three square integrable martingales plus a continuous process of bounded variation. Furthermore, the increments of the increasing processes corresponding to the three martingales and of the bounded variation process are dominated by a constant times $\triangle t$, implying in particular that

$$\mathrm{E}\left[\,|\sigma_v - \sigma_u|^2\right] \leq C\,(v-u), \qquad \text{for all } 0 \leq u < v. \tag{17}$$

We use $\Upsilon(x)$ as a shorthand for $\rho_x(g)$. Observe that the assumptions on g imply that $x \mapsto \Upsilon(x)$ is differentiable with a bounded derivative on any bounded interval not including 0; in particular (see (H2a))

$$|\,\Upsilon(x) - \Upsilon(y) - \Upsilon'(y)\cdot(x-y)\,| \leq \Psi(|x-y|)\cdot|x-y|, \quad x^2, y^2 \in (a,b), \tag{18}$$

where $\Psi : \mathbf{R}_+ \to \mathbf{R}_+$ is continuous, increasing and $\Psi(0) = 0$.

9.3 Main result

As already mentioned, our aim is to show the following special version of the general CLT-result given as Theorem 2.

Theorem 3 *Under assumptions (Ka), (H1a) and (H2a), there exists a Wiener process* (B_t) *defined on some extension of* $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ *and independent of* $\mathcal{F}$ *such that*

$$\left(\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{[nt]} g(\sqrt{n}\,\triangle_i^n Y) - \int_0^t \rho_{\sigma_u}(g)\,\mathrm{d}u\right)\right) \to \int_0^t \sqrt{\rho_{\sigma_{u-}}(g^2) - \rho_{\sigma_{u-}}(g)^2}\,\mathrm{d}B_u, \tag{19}$$

where B *is a Brownian motion independent of the process* Y *and the convergence is (stably) in law.*

The first step is to rewrite the left hand side of (19) as follows

$$\begin{aligned}
&\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{[nt]} g(\sqrt{n}\,\triangle_i^n Y) - \int_0^t \rho_{\sigma_u}(g)\mathrm{d}u\right) \\
=\;& \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]}\left\{g(\sqrt{n}\,\triangle_i^n Y) - \mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right]\right\} \\
&+\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{[nt]}\mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right] - \int_0^t \rho_{\sigma_u}(g)\mathrm{d}u\right).
\end{aligned}$$

It is rather straightforward to show that the first term of the right hand side satisfies

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]}\left\{g(\sqrt{n}\,\triangle_i^n Y) - \mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right]\right\} \to \int_0^t \sqrt{\rho_{\sigma_u}(g^2) - \rho_{\sigma_u}(g)^2}\mathrm{d}B_u.$$

Hence what remains is to verify that uniformly

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{[nt]}\mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right] - \int_0^t \rho_{\sigma_u}(g)\mathrm{d}u\right) \xrightarrow{p} 0. \tag{20}$$

We have

$$\begin{aligned}
&\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{[nt]}\mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right] - \int_0^t \rho_{\sigma_u}(g)\mathrm{d}u\right) \\
=\;& \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]}\mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right] - \sqrt{n}\sum_{i=1}^{[nt]}\int_{(i-1)/n}^{i/n}\rho_{\sigma_u}(g)\mathrm{d}u \\
&+\sqrt{n}\left(\sum_{i=1}^{[nt]}\int_{(i-1)/n}^{i/n}\rho_{\sigma_u}(g)\mathrm{d}u - \int_0^t \rho_{\sigma_u}(g)\mathrm{d}u\right)
\end{aligned} \tag{21}$$

where, uniformly

$$\sqrt{n}\left\{\sum_{i=1}^{[nt]}\int_{(i-1)/n}^{i/n}\rho_{\sigma_u}(g)\mathrm{d}u-\int_0^t\rho_{\sigma_u}(g)\mathrm{d}u\right\}\xrightarrow{p}0.$$

The first term on the right hand side of (21) is now split into the difference of

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]}\left\{\mathrm{E}\left[g(\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right]-\rho_{\frac{i-1}{n}}\right\} \tag{22}$$

where

$$\rho_{\frac{i-1}{n}}=\rho_{\sigma_{\frac{i-1}{n}}}(g)=\mathrm{E}\left[g\left(\sigma_{\frac{i-1}{n}}\triangle_i^n W\right)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right]$$

and

$$\sqrt{n}\sum_{i=1}^{[nt]}\int_{(i-1)/n}^{i/n}\left\{\rho_{\sigma_u}(g)\mathrm{d}u-\rho_{\frac{i-1}{n}}\right\}\mathrm{d}u. \tag{23}$$

It is rather easy to show that (22) tends to 0 in probability uniformly in t. The challenge is thus to show the same result holds for (23).

To handle (23) one splits the individual terms in the sum into

$$\sqrt{n}\,\Upsilon'\left(\sigma_{\frac{i-1}{n}}\right)\int_{(i-1)/n}^{i/n}\left(\sigma_u-\sigma_{\frac{i-1}{n}}\right)\,\mathrm{d}u \tag{24}$$

plus

$$\sqrt{n}\int_{(i-1)/n}^{i/n}\left\{\Upsilon(\sigma_u)-\Upsilon\left(\sigma_{\frac{i-1}{n}}\right)-\Upsilon'\left(\sigma_{\frac{i-1}{n}}\right)\cdot\left(\sigma_u-\sigma_{\frac{i-1}{n}}\right)\right\}\,\mathrm{d}u, \tag{25}$$

where $\Upsilon(x)$ is a shorthand for $\rho_x(g)$ and $\Upsilon'(x)$ denotes the derivative with respect to x. That (25) tends to 0 may be shown via splitting it into two terms, each of which tends to 0 as is verified by a sequence of inequalities, using in particular Doob's inequality. To prove that (24) converges to 0, again one splits, this time into three terms, using the differentiability of g in the relevant regions and the mean value theorem for differentiable functions. The first two of these terms can be handled by relatively simple means, the third poses the most difficult part of the whole proof and is treated via splitting it into seven parts. It is at this stage that the assumption that g be even comes into play and is crucial.

9.4 Details of the proof

Introducing the notation

$$U_t(g)=\int_0^t\sqrt{\rho_{\sigma_{u-}}(g^2)-\rho_{\sigma_{u-}}(g)^2}\,\mathrm{d}B_u\quad t\geq 0$$

we may reexpress (19) as

$$\left(\sqrt{n}\left(Y_t^n(g)-\int_0^t\sigma_u(g)\,\mathrm{d}u\right)\right)\to(U_t(g)). \tag{26}$$

To prove this, introduce the set of variables $\{\beta_i^n \mid i,\, n \geq 1\}$ given by

$$\beta_i^n = \sqrt{n} \cdot \sigma_{\frac{i-1}{n}} \cdot \triangle_i^n W, \quad i,\, n \geq 1.$$

The β_i^n's should be seen as approximations to $\sqrt{n}\, \triangle_i^n Y$. In fact, since

$$\sqrt{n}\, \triangle_i^n Y - \beta_i^n = \sqrt{n} \int_{(i-1)/n}^{i/n} (\sigma_s - \sigma_{\frac{i-1}{n}})\, \mathrm{d}W_s$$

and (σ_t) is uniformly bounded, a straightforward application of (17) and the Burkholder-Davis-Gundy-inequalities (e.g. Revuz and Yor (1999, pp. 160-171)) gives for every $p > 0$ the following simple estimates.

$$\mathrm{E}\left[|\sqrt{n}\, \triangle_i^n Y - \beta_i^n|^p \mid \mathcal{F}_{\frac{i-1}{n}} \right] \leq \frac{C_p}{n^{p \wedge 1}} \tag{27}$$

and

$$\mathrm{E}\left[|\sqrt{n}\, \triangle_i^n Y|^p + |\beta_i^n|^p \mid \mathcal{F}_{\frac{i-1}{n}} \right] \leq C_p \tag{28}$$

for all $i, n \geq 1$. Observe furthermore that

$$\mathrm{E}\left[g(\beta_i^n) \mid \mathcal{F}_{\frac{i-1}{n}} \right] = \rho_{\sigma_{\frac{i-1}{n}}}(g), \quad \text{for all } i,\, n \geq 1.$$

Introduce for convenience, for each $t > 0$ and $n \geq 1$, the shorthand notation

$$U_t^n(g) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left\{ g(\sqrt{n}\, \triangle_i^n Y) - \mathrm{E}\left[g(\sqrt{n}\, \triangle_i^n Y) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right\}$$

and

$$\tilde{U}_t^n(g) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left\{ g(\beta_i^n) - \rho_{\sigma_{\frac{i-1}{n}}}(g) \right\} = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \left\{ g(\beta_i^n) - \mathrm{E}\left[g(\beta_i^n) \mid \mathcal{F}_{\frac{i-1}{n}} \right] \right\}.$$

The asymptotic behaviour of $(\tilde{U}_t^n(g))$ is well known. More precisely under the the given assumptions (in fact much less is needed) we have

$$(U_t^n(g)) \to (U_t(g)).$$

This result is a rather straightforward consequence of Jacod and Shiryaev (2003, Theorem IX.7.28). Thus, if $(U_t^n(g) - \tilde{U}_t^n(g)) \xrightarrow{P} 0$ we may deduce the following result.

Theorem 4 *Let (B_t) and $(U_t(g))$ be as above. Then*

$$(\tilde{U}_t^n(g)) \to (U_t(g)).$$

Proof of Theorem 4.

As pointed out just above it is enough to prove that

$$(U_t^n(g) - \tilde{U}_t^n(g)) \xrightarrow{P} 0.$$

But for $t \geq 0$ and $n \geq 1$

$$U_t^n(g) - \tilde{U}_t^n(g) = \sum_{i=1}^{[nt]} \left(\xi_i^n - \mathrm{E}\left[\xi_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]\right)$$

where

$$\xi_i^n = \frac{1}{\sqrt{n}} \left\{g(\sqrt{n}\triangle_i^n Y) - g(\beta_i^n)\right\}, \quad i, n \geq 1.$$

Thus we have to prove

$$\left(\sum_{i=1}^{[nt]} \left\{\xi_i^n - \mathrm{E}\left[\xi_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]\right\}\right) \xrightarrow{P} 0.$$

But, as the left hand side of this relation is a sum of martingale differences, this is implied by Doob's inequality (e.g. Revuz and Yor (1999, pp. 54-55)) if for all $t > 0$

$$\sum_{i=1}^{[nt]} \mathrm{E}[(\xi_i^n)^2] = \mathrm{E}\left[\sum_{i=1}^{[nt]} \mathrm{E}[(\xi_i^n)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}}]\right] \to 0 \quad \text{as } n \to \infty.$$

Fix $t > 0$. Using the Cauchy-Schwarz, Burkholder-Davis-Gundy and Jensen inequalities we have for all $i, n \geq 1$.

$$\begin{aligned}
\mathrm{E}\left[(\xi_i^n)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] &= \frac{1}{n}\mathrm{E}\left[\left\{g(\sqrt{n}\triangle_i^n Y) - \beta_i^n + \beta_i^n - g(\beta_i^n)\right\}^2 \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \\
&\leq \frac{C}{n}\mathrm{E}\left[(1 + |\sqrt{n}\triangle_i^n Y|^p + |\beta_i^n|^p)^2 \cdot (\sqrt{n}\triangle_i^n Y - \beta_i^n)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \\
&\leq \frac{C}{n}\sqrt{\mathrm{E}\left[(1 + |\sqrt{n}\triangle_i^n Y|^{2p} + |\beta_i^n|^{2p}) \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]} \cdot \sqrt{\mathrm{E}\left[(\sqrt{n}\triangle_i^n Y - \beta_i^n)^4 \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]} \\
&\leq C\sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right) \mathrm{d}W_u\right)^4 \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]} \\
&\leq C\sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}u\right)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]}.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{i=1}^{[nt]} \mathrm{E}\left[(\xi_i^n)^2\right] &\leq C \sum_{i=1}^{[nt]} \mathrm{E}\left[\sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}u\right)^2 |\mathcal{F}_{\frac{i-1}{n}}\right]}\right] \\
&\leq C \sum_{i=1}^{[nt]} \sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}u\right)^2\right]} \\
&\leq C\,[nt]\,\frac{1}{[nt]} \sum_{i=1}^{[nt]} \sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}u\right)^2\right]} \\
&\leq C\,[nt] \sqrt{\frac{1}{[nt]} \sum_{i=1}^{[nt]} \mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}u\right)^2\right]} \\
&\leq C\sqrt{[nt]} \sqrt{\sum_{i=1}^{[nt]} \mathrm{E}\left[\left(\frac{1}{n}\int_{i-1}^{i} \left(\sigma_{\frac{v}{n}n-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}v\right)^2\right]} \\
&\leq C\sqrt{\sum_{i=1}^{[nt]} \frac{1}{n}\mathrm{E}\left[\left(\int_{i-1}^{i} \left(\sigma_{\frac{v}{n}n-} - \sigma_{\frac{i-1}{n}}\right)^4 \mathrm{d}v\right)\right]} \\
&\leq C\sqrt{\sum_{i=1}^{[nt]} \mathrm{E}\left[\left(\int_{i-1}^{i} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^4 \mathrm{d}u\right)\right]} \\
&\leq C\sqrt{\sum_{i=1}^{[nt]} \mathrm{E}\left[\left(\int_{i-1}^{i} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}}\right)^2 \mathrm{d}u\right)\right]}.
\end{aligned}$$

as $n \to \infty$, by Lebesgue's Theorem and the boundedness of (σ_t).

□

To prove the convergence (26) it suffices, using Theorem 4 above, to prove that

$$\left(U_t^n(g) - \sqrt{n}\left\{Y_t^n(g) - \int_0^t \rho_{\sigma_u}(g)\,\mathrm{d}u\right\}\right) \xrightarrow{P} 0.$$

But as

$$U_t^n(g) - \sqrt{n}\,Y_t^n(g) = -\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathrm{E}\left[g(\sqrt{n}\,\triangle_i^n Y)\,|\,\mathcal{F}_{\frac{i-1}{n}}\right]$$

and, as is easily seen,

$$\left(\sqrt{n}\int_0^t \rho_{\sigma_u}\ (g)\,\mathrm{d}u - \sum_{i=1}^{[nt]} \sqrt{n}\int_{(i-1)/n}^{i/n} \rho_{\sigma_u}(g)\,\mathrm{d}u\right) \xrightarrow{P} 0,$$

the job is to prove that

$$\sum_{i=1}^{[nt]} \eta_i^n \xrightarrow{P} 0 \quad \text{for all } t > 0,$$

where for $i, n \geq 1$

$$\eta_i^n = \frac{1}{\sqrt{n}} \mathrm{E}\left[g(\sqrt{n} \triangle_i^n Y) \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] - \sqrt{n} \int_{(i-1)/n}^{i/n} \rho_{\sigma_u}(g)\, \mathrm{d}u.$$

Fix $t > 0$ and write, for all $i, n \geq 1$,

$$\eta_i^n = \eta(1)_i^n + \eta(2)_i^n,$$

where

$$\eta(1)_i^n = \frac{1}{\sqrt{n}} \left\{ \mathrm{E}\left[g(\sqrt{n} \triangle_i^n Y) \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] - \rho_{\sigma_{\frac{i-1}{n}}}(g) \right\} \tag{29}$$

and

$$\eta(2)_i^n = \sqrt{n} \int_{(i-1)/n}^{i/n} \left\{ \rho_{\sigma_u}(g) - \rho_{\sigma_{\frac{i-1}{n}}}(g) \right\} \mathrm{d}u. \tag{30}$$

We will now separately prove

$$\eta(1)^n = \sum_{i=1}^{[nt]} \eta(1)_i^n \xrightarrow{P} 0 \tag{31}$$

and

$$\eta(2)^n = \sum_{i=1}^{[nt]} \eta(2)_i^n \xrightarrow{P} 0. \tag{32}$$

9.5 Some auxiliary estimates

In order to show (31) and (32) we need some refinements of the estimate (17) above. To state these we split up $(\sqrt{n} \triangle_i^n Y - \beta_i^n)$ into several terms. By definition

$$\sqrt{n} \triangle_i^n Y - \beta_i^n = \sqrt{n} \int_{(i-1)/n}^{i/n} \left(\sigma_{u-} - \sigma_{\frac{i-1}{n}} \right) \mathrm{d}W_u$$

for all $i, n \geq 1$. Writing

$$E_n = \{x \in \mathrm{E} \,|\, |\Psi(x)| > 1/\sqrt{n}\}$$

the difference $\sigma_u - \sigma_{\frac{i-1}{n}}$ equals

$$\begin{aligned} & \int_{(i-1)/n}^{u} a_s^* \, ds + \int_{(i-1)/n}^{u} \sigma_{s-}^* \,\mathrm{d}W_s + \int_{(i-1)/n}^{u} v_{s-}^* \,\mathrm{d}V_s + \int_{(i-1)/n}^{u} \int_E \phi(s-, x)\, (\mu - \nu)(\mathrm{d}s\, \mathrm{d}x) \\ = \; & \sum_{j=1}^{5} \xi(j)_i^n(u), \end{aligned}$$

for $i, n \geq 1$ and $u \geq (i-1)/n$ where

$$\begin{aligned}
\xi(1)_i^n(u) &= \int_{(i-1)/n}^{u} a_s^* \,\mathrm{d}s + \int_{(i-1)/n}^{u} \left(\sigma_{s-}^* - \sigma_{\frac{i-1}{n}}^*\right) \mathrm{d}W_s + \int_{(i-1)/n}^{u} \left(v_{s-}^* - v_{\frac{i-1}{n}}^*\right) \mathrm{d}V_s \\
\xi(2)_i^n(u) &= \sigma_{\frac{i-1}{n}}^* \left(W_u - W_{\frac{i-1}{n}}\right) + v_{\frac{i-1}{n}}^* \left(V_u - V_{\frac{i-1}{n}}\right) \\
\xi(3)_i^n(u) &= \int_{(i-1)/n}^{u} \int_{E_n^c} \phi(s-, x)\,(\mu - \nu)(\mathrm{d}s\,\mathrm{d}x) \\
\xi(4)_i^n(u) &= \int_{(i-1)/n}^{u} \int_{E_n} \left\{\phi(s-, x) - \phi\left(\frac{i-1}{n}, x\right)\right\} (\mu - \nu)(\mathrm{d}s\,\mathrm{d}x) \\
\xi(5)_i^n(u) &= \int_{(i-1)/n}^{u} \int_{E_n} \phi\left(\frac{i-1}{n}, x\right) (\mu - \nu)(\mathrm{d}s\,\mathrm{d}x).
\end{aligned}$$

That is, for $i, n \geq 1$,

$$\sqrt{n}\,\triangle_i^n Y - \beta_i^n = \sum_{j=1}^{5} \xi(j)_i^n, \tag{33}$$

where

$$\xi(j)_i^n = \sqrt{n} \int_{(i-1)/n}^{i/n} \xi(j)_i^n(u-)\,\mathrm{d}W_u, \quad \text{for } j = 1, 2, 3, 4, 5.$$

The specific form of the variables implies, using Burkholder-Davis-Gundy inequalities, that for every $q \geq 2$ we have

$$\begin{aligned}
\mathrm{E}[\,|\xi(j)_i^n|^q\,] &\leq C_q\, n^{q/2}\, \mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} \xi(j)_i^n(u)^2\,\mathrm{d}u\right)^{q/2}\right] \\
&\leq n \int_{(i-1)/n}^{i/n} \mathrm{E}[\,|\xi(j)_i^n(u)|^q\,]\,\mathrm{d}u \\
&\leq \sup_{(i-1)/n \leq u \leq i/n} \mathrm{E}[\,|\xi(j)_i^n(u)|^q\,],
\end{aligned}$$

for all $i, n \geq 1$ and all j. These terms will now be estimated. This is done in the following series of lemmas where i and n are arbitrary and we use the notation

$$d_i^n = \int_{(i-1)/n}^{i/n} \mathrm{E}\left[\left(\sigma_{s-}^* - \sigma_{\frac{i-1}{n}}^*\right)^2 + \left(v_{s-}^* - v_{\frac{i-1}{n}}^*\right)^2 + \int_E \left\{\phi(s-, x) - \phi\left(\frac{i-1}{n}, x\right)\right\}^2 F(\mathrm{d}x)\right] \mathrm{d}s.$$

Lemma 1

$$\mathrm{E}[\,(\xi(1)_i^n)^2] \leq C_1 \cdot (1/n^2 + d_i^n).$$

Lemma 2

$$\mathrm{E}[\,(\xi(2)_i^n)^2] \leq C_2/n.$$

Lemma 3

$$\mathrm{E}[\,(\xi(3)_i^n)^2] \leq C_3\, \varphi(1/\sqrt{n})/n,$$

where

$$\varphi(\epsilon) = \int_{\{\,|\Psi| \leq \epsilon\}} \Psi(x)^2\, F(\mathrm{d}x).$$

Lemma 4

$$\mathrm{E}[\,(\xi(4)_i^n)^2] \le C_4\, d_i^n.$$

Lemma 5

$$\mathrm{E}[\,(\xi(5)_i^n)^2] \le C_5/n.$$

The proofs of these five Lemmas rely on straightforward martingale inequalities.

Observe that Lebesgue's Theorem ensures, since the processes involved are assumed càdlàg and uniformly bounded, that as $n \to \infty$

$$\sum_{i=1}^{[nt]} d_i^n \; \to 0 \qquad \text{for all } t > 0.$$

Taken together these statements imply the following result.

Corollary 2 *For all $t > 0$ as $n \to \infty$*

$$\sum_{i=1}^{[nt]} \left\{ \mathrm{E}[\,(\xi(1)_i^n)^2] + \mathrm{E}[\,(\xi(3)_i^n)^2] + \mathrm{E}[\,(\xi(4)_i^n)^2] \right\}) \; \to 0.$$

Below we shall invoke this Corollary as well as Lemmas 2 and 5.

9.6 Proof of $\eta(2)^n \xrightarrow{P} 0$

Recall we wish to show that

$$\eta(2)^n = \sum_{i=1}^{[nt]} \eta(2)_i^n \; \xrightarrow{P} 0. \tag{34}$$

From now on let $t > 0$ be fixed. We split the $\eta(2)_i^n$'s according to

$$\eta(2)_i^n = \eta'(2)_i^n + \eta''(2)_i^n \;, \quad i, n \ge 1,$$

where, writing $\Upsilon(x)$ for $\rho_x(g)$,

$$\eta'(2)_i^n = \sqrt{n}\; \Upsilon'\left(\sigma_{\frac{i-1}{n}}\right) \int_{(i-1)/n}^{i/n} \left(\sigma_u - \sigma_{\frac{i-1}{n}}\right) \mathrm{d}u$$

and

$$\eta''(2)_i^n = \sqrt{n} \int_{(i-1)/n}^{i/n} \left\{ \Upsilon(\sigma_u) - \Upsilon\left(\sigma_{\frac{i-1}{n}}\right) - \Upsilon'\left(\sigma_{\frac{i-1}{n}}\right) \cdot \left(\sigma_u - \sigma_{\frac{i-1}{n}}\right) \right\} \mathrm{d}u.$$

With this notation we shall prove (34) by showing

$$\sum_{i=1}^{[nt]} \eta'(2)_i^n \; \xrightarrow{P} 0$$

and

$$\sum_{i=1}^{[nt]} \eta''(2)_i^n \xrightarrow{P} 0.$$

Inserting the description of (σ_t) (see (H1a)) we may write

$$\eta'(2)_i^n = \eta'(2,1)_i^n + \eta'(2,2)_i^n,$$

where for all $i, n \geq 1$

$$\eta'(2,1)_i^n = \sqrt{n}\ \Upsilon'\left(\sigma_{\frac{i-1}{n}}\right) \int_{(i-1)/n}^{i/n} \left(\int_{(i-1)/n}^{u} a_s^* \,\mathrm{d}s\right) \,\mathrm{d}u$$

and

$$\begin{aligned}\eta'(2,2)_i^n \quad &= \quad \sqrt{n}\ \Upsilon'\left(\sigma_{\frac{i-1}{n}}\right) \int_{(i-1)/n}^{i/n} \left[\int_{(i-1)/n}^{u} \sigma_{s-}^* \,\mathrm{d}W_s + \int_{(i-1)/n}^{u} v_{s-}^* \,\mathrm{d}V_s \right. \\ &\quad \left. + \int_E \phi(s-,x)\,(\mu-\nu)(\mathrm{d}s\,\mathrm{d}x)\right] \mathrm{d}u.\end{aligned}$$

By (H2a) and (18) and the uniform boundedness of (a_t^*) we have

$$|\eta'(2,1)_i^n| \leq C\,\sqrt{n} \int_{(i-1)/n}^{i/n} \{u-(i-1)/n\} \,\mathrm{d}u \leq C/n^{3/2}$$

for all $i, n \geq 1$ and thus

$$\sum_{i=1}^{[nt]} \eta'(2,1)_i^n \xrightarrow{P} 0.$$

Since

$$(W_t),\ (V_t) \text{ and } \left(\int_0^t \int_E \phi(s-,x)(\mu-\nu)(\mathrm{d}s\,\mathrm{d}x)\right)$$

are all martingales we have

$$\mathrm{E}\left[\eta'(2,2)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] = 0 \quad \text{for all} \quad i, n \geq 1.$$

By Doob's inequality it is therefore feasible to estimate

$$\sum_{i=1}^{[nt]} \mathrm{E}[\,(\eta'(2,2)_i^n)^2].$$

Inserting again the description of (σ_t) we find, applying simple inequalities, in particular Jensen's, that

$$
\begin{aligned}
&(\eta'(2,2)_i^n)^2 \\
\leq\ & C\,n \left(\int_{(i-1)/n}^{i/n} \left\{ \int_{(i-1)/n}^{u} \sigma^*_{s-}\,\mathrm{d}W_s \right\} \mathrm{d}u \right)^2 + C\,n \left(\int_{(i-1)/n}^{i/n} \left\{ \int_{(i-1)/n}^{u} v^*_{s-}\,\mathrm{d}V_s \right\} \mathrm{d}u \right)^2 \\
& + C\,n \left(\int_{(i-1)/n}^{i/n} \int_{(i-1)/n}^{u} \left\{ \int_E \phi(s-,x)\,(\mu-\nu)(\mathrm{d}s\,\mathrm{d}x) \right\} \mathrm{d}u \right)^2 \\
\leq\ & C \int_{(i-1)/n}^{i/n} \left(\int_{(i-1)/n}^{u} \sigma^*_{s-}\,\mathrm{d}W_s \right)^2 \mathrm{d}u + C \int_{(i-1)/n}^{i/n} \left(\int_{(i-1)/n}^{u} v^*_{s-}\,\mathrm{d}V_s \right)^2 \mathrm{d}u \\
& + C \int_{(i-1)/n}^{i/n} \left(\int_{(i-1)/n}^{u} \int_E \phi(s-,x)\,(\mu-\nu)(\mathrm{d}s\,\mathrm{d}x) \right)^2 \mathrm{d}u.
\end{aligned}
$$

The properties of the Wiener integrals and the uniform boundedness of (σ^*_t) and (v^*_t) ensure that

$$
\mathrm{E}\left[\left(\int_{(i-1)/n}^{u} \sigma^*_{s-}\,\mathrm{d}W_s \right)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}} \right] \leq C \cdot \left(u - \frac{i-1}{n} \right)
$$

and likewise

$$
\mathrm{E}\left[\left(\int_{(i-1)/n}^{u} v^*_{s-}\,\mathrm{d}V_s \right)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}} \right] \leq C \cdot \left(u - \frac{i-1}{n} \right)
$$

for all $i, n \geq 1$. Likewise for the Poisson part we have

$$
\begin{aligned}
& \mathrm{E}\left[\left(\int_{(i-1)/n}^{u} \int_E \phi(s-,x)\,(\mu-\nu)(\mathrm{d}s\,\mathrm{d}x) \right)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}} \right] \\
\leq\ & C \int_{(i-1)/n}^{u} \int_E \mathrm{E}[\phi^2(s,x) \,|\, \mathcal{F}_{\frac{i-1}{n}}]\, F(\mathrm{d}x)\,\mathrm{d}s
\end{aligned}
$$

yielding a similar bound. Putting all this together we have for all $i, n \geq 1$

$$
\begin{aligned}
\mathrm{E}[\,(\eta'(2,2)_i^n)^2 \,|\, \mathcal{F}_{\frac{i-1}{n}}] &\leq C \int_{(i-1)/n}^{i/n} (u-(i-1)/n)\,\mathrm{d}u \\
&\leq C/n^2.
\end{aligned}
$$

Thus as $n \to \infty$ so

$$
\sum_{i=1}^{[nt]} \mathrm{E}[\,(\eta'(2,2)_i^n)^2] \to 0.
$$

and since

$$
\mathrm{E}\left[\eta'(2,2)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}} \right] = 0 \quad \text{for all} \quad i, n \geq 1
$$

we deduce from Doob's inequality that

$$
\sum_{i=1}^{[nt]} \eta'(2,2)_i^n \xrightarrow{P} 0
$$

proving altogether

$$\sum_{i=1}^{[nt]} \eta'(2)_i^n \xrightarrow{P} 0.$$

Applying once more (H2a) and (18) we have for every $\epsilon > 0$ and every i, n that

$$\begin{aligned} |\eta''(2)_i^n| &\leq \sqrt{n} \int_{(i-1)/n}^{i/n} \Psi\left(\left|\sigma_u - \sigma_{\frac{i-1}{n}}\right|\right) \cdot \left|\sigma_u - \sigma_{\frac{i-1}{n}}\right| \, \mathrm{d}u \\ &\leq \sqrt{n}\, \Psi(\epsilon) \int_{(i-1)/n}^{i/n} \left|\sigma_u - \sigma_{\frac{i-1}{n}}\right| \, \mathrm{d}u + \sqrt{n}\, \Psi(2\sqrt{b})/\epsilon \int_{(i-1)/n}^{i/n} \left|\sigma_u - \sigma_{\frac{i-1}{n}}\right|^2 \, \mathrm{d}u. \end{aligned}$$

Thus from (17) and its consequence

$$\mathrm{E}\left[\left|\sigma_u - \sigma_{\frac{i-1}{n}}\right|\right] \leq C/\sqrt{n}$$

we get

$$\sum_{i=1}^{[nt]} \mathrm{E}[\,|\eta''(2)_i^n|\,] \leq Ct\, \Psi(\epsilon) + \frac{C\, \Psi(b)}{\sqrt{n}\, \epsilon}$$

for all n and all ϵ. Letting here first $n \to \infty$ and then $\epsilon \to 0$ we may conclude that as $n \to \infty$

$$\sum_{i=1}^{[nt]} \mathrm{E}[\,|\eta''(2)_i^n|\,] \to 0$$

implying the convergence

$$\sum_{i=1}^{[nt]} \eta(2)_i^n \xrightarrow{P} 0.$$

Thus ending the proof of (32).

□

9.7 Proof of $\eta(1)^n \xrightarrow{P} 0$

Recall we are to show that

$$\eta(1)^n = \sum_{i=1}^{[nt]} \eta(1)_i^n \xrightarrow{P} 0. \tag{35}$$

Let still $t > 0$ be fixed. Recall that

$$\begin{aligned} \eta(1)_i^n &= \frac{1}{\sqrt{n}} \left\{ \mathrm{E}\left[g(\sqrt{n}\, \triangle_i^n Y) \,|\, \mathcal{F}_{\frac{i-1}{n}} \right] - \rho_{\sigma_{\frac{i-1}{n}}}(g) \right\} \\ &= \frac{1}{\sqrt{n}} \, \mathrm{E}\left[g(\sqrt{n}\, \triangle_i^n Y) - g(\beta_i^n) \,|\, \mathcal{F}_{\frac{i-1}{n}} \right]. \end{aligned}$$

Introduce the notation (recall the assumption (K2))

$$A_i^n = \{\, |\sqrt{n}\, \triangle_i^n Y - \beta_i^n| > \; d(\beta_i^n, B)/2 \,\}.$$

Since B is a Lebesgue null set and β_i^n is absolutely continuous, $g'(\beta_i^n)$ is defined *a.s.* and, by assumption, g is differentiable on the interval joining $\triangle_i^n Y(\omega)$ and $\beta_i^n(\omega)$ for all $\omega \in A_i^{n\,c}$. Thus, using the Mean Value Theorem, we may for all $i, n \geq 1$ write

$$
\begin{aligned}
& g(\sqrt{n}\,\triangle_i^n Y) - g(\beta_i^n) \\
= \; & \left\{ g(\sqrt{n}\,\triangle_i^n Y) - g(\beta_i^n) \right\} \cdot \mathbf{1}_{A_i^n} \\
& + g'(\beta_i^n) \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) \cdot \mathbf{1}_{A_i^{n\,c}} \\
& + \left\{ g'(\alpha_i^n) - g'(\beta_i^n) \right\} \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) \cdot \mathbf{1}_{A_i^{n\,c}} \\
= \; & \sqrt{n}\, \left\{ \delta(1)_i^n + \delta(2)_i^n + \delta(3)_i^n \right\},
\end{aligned}
$$

where α_i^n are random points lying in between $\sqrt{n}\,\triangle_i^n Y$ and β_i^n, i.e.

$$
\sqrt{n}\,\triangle_i^n Y \wedge \beta_i^n \leq \alpha_i^n \leq \sqrt{n}\,\triangle_i^n Y \vee \beta_i^n,
$$

and

$$
\begin{aligned}
\delta(1)_i^n &= \left[\left\{ g(\sqrt{n}\,\triangle_i^n Y) - g(\beta_i^n) \right\} - g'(\beta_i^n) \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) \right] \cdot \mathbf{1}_{A_i^n} / \sqrt{n} \\
\delta(2)_i^n &= \left\{ g'(\alpha_i^n) - g'(\beta_i^n) \right\} \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) \cdot \mathbf{1}_{A_i^{n\,c}} / \sqrt{n} \\
\delta(3)_i^n &= g'(\beta_i^n) \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) / \sqrt{n}.
\end{aligned}
$$

Thus it suffices to prove

$$
\sum_{i=1}^{[nt]} \mathrm{E}\left[\delta(k)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}} \right] \xrightarrow{P} 0, \quad k = 1, 2, 3.
$$

Consider the case $k = 1$. Using (Ka) and the fact that β_i^n is absolutely continuous we have a.s.

$$
\begin{aligned}
& |g(\sqrt{n}\,\triangle_i^n Y) - g(\beta_i^n)| \\
\leq \; & M(1 + |\sqrt{n}\,\triangle_i^n Y - \beta_i^n|^p + |\beta_i^n|^p) \cdot |\sqrt{n}\,\triangle_i^n Y - \beta_i^n| \\
\leq \; & (2^p + 1) M (1 + |\sqrt{n}\,\triangle_i^n Y|^p + |\beta_i^n|^p) \cdot |\sqrt{n}\,\triangle_i^n Y - \beta_i^n|,
\end{aligned}
$$

and

$$
|\, g'(\beta_i^n) \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) \,| \leq M(1 + |\beta_i^n|^p) \cdot |\sqrt{n}\,\triangle_i^n Y - \beta_i^n|.
$$

By Cauchy-Schwarz's inequality $\mathrm{E}[\,|\delta(1)_i^n|\,]$ is therefore for all $i, n \geq 1$ less than

$$
C \cdot \mathrm{E}[\,1 + |\sqrt{n}\,\triangle_i^n Y|^{3p} + |\beta_i^n|^{3p}]^{1/3} \cdot \mathrm{E}[\,(\sqrt{n}\,\triangle_i^n Y - \beta_i^n)^2 / n\,]^{1/2} \cdot P(A_i^n)^{1/6}
$$

implying for fixed t, by means of (20), that

$$
\begin{aligned}
\mathrm{E}\left[\left[\sum_{i=1}^{[nt]} |\,\delta(1)_i^n| \right] \right. & \leq C \cdot \sup_{i \geq 1} P(A_i^n)^{1/6} \sum_{i=1}^{[nt]} \mathrm{E}[\,(\triangle_i^n Y - \beta_i^n)^2 / n\,]^{1/2} \\
& \leq C \cdot \sup_{i \geq 1} P(A_i^n)^{1/6} \sum_{i=1}^{[nt]} 1/n \\
& \leq Ct \cdot \sup_{i \geq 1} P(A_i^n)^{1/6}.
\end{aligned}
$$

For all $i, n \geq 1$ we have for every $\epsilon > 0$

$$\begin{aligned} P(A_i^n) &\leq P(A_i^n \cap \{d(\beta_i^n, B) \leq \epsilon\}) + P(A_i^n \cap \{d(\beta_i^n, B) > \epsilon\}) \\ &\leq P(d(\beta_i^n, B) \leq \epsilon) + P(|\sqrt{n}\,\triangle_i^n Y - \beta_i^n| > \epsilon/2) \\ &\leq P(d(\beta_i^n, B) \leq \epsilon) + \frac{4}{\epsilon^2} \cdot \mathrm{E}[\,(\sqrt{n}\,\triangle_i^n Y - \beta_i^n)^2] \\ &\leq P(d(\beta_i^n, B) \leq \epsilon) + \frac{C}{n\,\epsilon^2}. \end{aligned}$$

But (H2a) implies that the densities of β_i^n are pointwise dominated by a Lebesgue integrable function $h_{a,b}$ providing, for all $i, n \geq 1$, the estimate

$$\begin{aligned} P(A_i^n) &\leq \int_{\{x \,|\, d(x,B) \leq \epsilon\}} h_{a,b}\,\mathrm{d}\lambda_1 + \frac{C}{n\,\epsilon^2} \\ &= \alpha_\epsilon + \frac{C}{n\,\epsilon^2}. \end{aligned} \tag{36}$$

Observe $\lim_{\epsilon \to 0} \alpha_\epsilon = 0$. Taking now in (36) sup over i and then letting first $n \to \infty$ and then $\epsilon \downarrow 0$ we get

$$\lim_n \sup_{i \geq 1} P(A_i^n) = 0$$

proving that

$$\mathrm{E}\left[\sum_{i=1}^{[nt]} |\,\delta(1)_i^n|\right] \to 0$$

and thus

$$\sum_{i=1}^{[nt]} \mathrm{E}\left[\,\delta(1)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \xrightarrow{P} 0.$$

Consider next the case $k = 2$. As assumed in (Ka), g is continuously differentiable outside of B. Thus for each $A > 1$ and $\epsilon > 0$ there exists a function $G_{A,\epsilon} : (0,1) \to \mathbf{R}_+$ such that for given $0 < \epsilon' < \epsilon/2$

$$\big|g'(x+y) - g'(x)\big| \leq G_{A,\epsilon}(\epsilon') \text{ for all } |x| \leq A,\ |y| \leq \epsilon' < \epsilon < d(x,B).$$

Observe that $\lim_{\epsilon' \downarrow 0} G_{A,\epsilon}(\epsilon') = 0$ for all A and ϵ. Fix $A > 1$ and $\epsilon \in (0,1)$. For all $i, n \geq 1$ we have

$$\begin{aligned} & |g'(\alpha_i^n) - g'(\beta_i^n)| \cdot \mathbf{1}_{A_i^{n\,c}} \\ = \;& |g'(\alpha_i^n) - g'(\beta_i^n)| \cdot \mathbf{1}_{A_i^{n\,c}} \left(\mathbf{1}_{\{|\alpha_i^n| + |\beta_i^n| > A\}} + \mathbf{1}_{\{|\alpha_i^n| + |\beta_i^n| \leq A\}}\right) \\ \leq \;& |g'(\alpha_i^n) - g'(\beta_i^n)| \cdot \frac{|\alpha_i^n| + |\beta_i^n|}{A} + |g'(\alpha_i^n) - g'(\beta_i^n)| \cdot \mathbf{1}_{A_i^{n\,c} \cap \{|\alpha_i^n| + |\beta_i^n| \leq A\}} \\ \leq \;& \frac{C}{A} \cdot (1 + |\alpha_i^n|^p + |\beta_i^n|^p)^2 + |g'(\alpha_i^n) - g'(\beta_i^n)| \cdot \mathbf{1}_{A_i^{n\,c} \cap \{|\alpha_i^n| + |\beta_i^n| \leq A\}} \\ \leq \;& \frac{C}{A} \cdot (1 + |\sqrt{n}\,\triangle_i^n Y|^{2p} + |\beta_i^n|^{2p}) + |g'(\alpha_i^n) - g'(\beta_i^n)| \cdot \mathbf{1}_{A_i^{n\,c} \cap \{|\alpha_i^n| + |\beta_i^n| \leq A\}}. \end{aligned}$$

Now writing

$$\begin{aligned}
1 &= \mathbf{1}_{\{d(\beta_i^n,B)\le\epsilon\}} + \mathbf{1}_{\{d(\beta_i^n,B)>\epsilon\}}\\
&= \mathbf{1}_{\{d(\beta_i^n,B)\le\epsilon\}}\\
&\quad + \mathbf{1}_{\{d(\beta_i^n,B)>\epsilon\}\cap\{|\alpha_i^n-\beta_i^n|\le\epsilon'\}}\\
&\quad + \mathbf{1}_{\{d(\beta_i^n,B)>\epsilon\}\cap\{|\alpha_i^n-\beta_i^n|>\epsilon'\}}
\end{aligned}$$

for all $0<\epsilon'<\epsilon/2$ we have

$$\begin{aligned}
\mathbf{1}_{A_i^{n\,c}\cap\{|\alpha_i^n|+|\beta_i^n|\le A\}} &\le \mathbf{1}_{\{d(\beta_i^n,B)\le\epsilon\}\cap A_i^{n\,c}\cap\{|\alpha_i^n|+|\beta_i^n|\le A\}}\\
&\quad + \mathbf{1}_{A_i^{n\,c}\cap\{|\alpha_i^n|+|\beta_i^n|\le A\}\cap\{d(\beta_i^n,B)>\epsilon\}\cap\{|\alpha_i^n-\beta_i^n|\le\epsilon'\}}\\
&\quad + \mathbf{1}_{A_i^{n\,c}\cap\{|\alpha_i^n|+|\beta_i^n|\le A\}\cap\{d(\beta_i^n,B)>\epsilon\}}\cdot\frac{|\alpha_i^n-\beta_i^n|}{\epsilon'}.
\end{aligned}$$

Combining this with the fact that

$$\begin{aligned}
|g'(\alpha_i^n)-g'(\beta_i^n)| &\le C(1+|\alpha_i^n|^p+|\beta_i^n|^p)\\
&\le CA^p
\end{aligned}$$

on $A_i^{n\,c}\cap\{|\alpha_i^n|+|\beta_i^n|\le A\}$ we obtain that

$$\begin{aligned}
&|g'(\alpha_i^n)-g'(\beta_i^n)|\cdot\mathbf{1}_{A_i^{n\,c}\cap\{|\alpha_i^n|+|\beta_i^n|\le A\}}\\
\le\;& CA^p\cdot\left(\mathbf{1}_{\{d(\beta_i^n,B)\le\epsilon\}}+\frac{|\alpha_i^n-\beta_i^n|}{\epsilon'}\right)+G_{A,\epsilon}(\epsilon')\\
\le\;& CA^p\cdot\Big(\mathbf{1}_{\{d(\beta_i^n,B)\le\epsilon\}}+\frac{|\sqrt{n}\,\triangle_i^nY-\beta_i^n|}{\epsilon'}\Big)+G_{A,\epsilon}(\epsilon').
\end{aligned}$$

Putting this together means that

$$\begin{aligned}
\sqrt{n}\,|\delta(2)_i^n| &= |g'(\alpha_i^n)-g'(\beta_i^n)|\cdot|\sqrt{n}\,\triangle_i^nY-\beta_i^n|\cdot\mathbf{1}_{A_i^{n\,c}}\\
&\le \left\{\frac{C}{A}\cdot(1+|\sqrt{n}\,\triangle_i^nY|^{2p}+|\beta_i^n|^{2p})+G_{A,\epsilon}(\epsilon')\right\}\cdot|\sqrt{n}\,\triangle_i^nY-\beta_i^n|\\
&\quad + CA^p\cdot\left(\mathbf{1}_{\{d(\beta_i^n,B)\le\epsilon\}}\cdot|\sqrt{n}\,\triangle_i^nY-\beta_i^n|+\frac{|\sqrt{n}\,\triangle_i^nY-\beta_i^n|^2}{\epsilon'}\right).
\end{aligned}$$

Exploiting here the inequalities (20) and (21) we obtain, for all $A>1$ and $0<2\epsilon'<\epsilon<1$ and all $i,n\ge 1$, using Hölder's inequality, the following estimate

$$\mathrm{E}[\,|\delta(2)_i^n|\,]\le C\left(\frac{1}{A\,n}+\frac{G_{A,\epsilon}(\epsilon')}{n}+\frac{A^p\sqrt{\alpha_\epsilon}}{n}+\frac{A^p}{\epsilon'\,n^{3/2}}\right)$$

implying for all $n\ge 1$ and $t\ge 0$ that

$$\sum_{i=1}^{[nt]}\mathrm{E}[\,|\delta(2)_i^n|\,]\le Ct\left(\frac{1}{A}+G_{A,\epsilon}(\epsilon')+A^p\sqrt{\alpha_\epsilon}+\frac{A^p}{\epsilon'\,n^{1/2}}\right).$$

Choosing in this estimate first A sufficiently big, then ϵ small (recall that $\lim_{\epsilon\to 0}\alpha_\epsilon = 0$) and finally ϵ' small, exploiting that $\lim_{\epsilon'\downarrow 0} G_{A,\epsilon}(\epsilon') = 0$ for all A and ϵ, we may conclude that

$$\lim_n \sum_{i=1}^{[nt]} \mathrm{E}\,[\,|\delta(2)_i^n|\,] = 0$$

and thus

$$\sum_{i=1}^{[nt]} \mathrm{E}\left[\,\delta(2)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \xrightarrow{P} 0.$$

So what remains to be proved is the convergence

$$\sum_{i=1}^{[nt]} \mathrm{E}\left[\delta(3)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \xrightarrow{P} 0.$$

As introduced in (33)

$$\sqrt{n}\,\triangle_i^n Y - \beta_i^n = \sum_{j=1}^{5} \xi(j)_i^n = \psi(1)_i^n + \psi(2)_i^n$$

for all $i, n \geq 1$ where

$$\psi(1)_i^n = \xi(1)_i^n + \xi(3)_i^n + \xi(4)_i^n,$$

$$\psi(2)_i^n = \xi(2)_i^n + \xi(5)_i^n,$$

and as

$$\delta(3)_i^n = g'(\beta_i^n)\cdot(\psi(1)_i^n + \psi(2)_i^n)/\sqrt{n}$$

it suffices to prove

$$\left(\sum_{i=1}^{[nt]} \mathrm{E}\left[\,g'(\beta_i^n)\cdot\psi(k)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \,/\sqrt{n}\right) \xrightarrow{P} 0, \quad k = 1, 2.$$

The case $k = 1$ is handled by proving

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathrm{E}[\,|g'(\beta_i^n)\cdot\xi(j)_i^n|\,] \to 0, \quad j = 1, 3, 4. \tag{37}$$

Using Jensen's inequality it is easily seen that for $j = 1, 3, 4$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathrm{E}[\,|g'(\beta_i^n)\cdot\xi(j)_i^n|\,] \leq C\,t\cdot\sqrt{\frac{1}{n}\sum_{i=1}^{[nt]} \mathrm{E}[\,g'(\beta_i^n)^2]}\cdot\sqrt{\sum_{i=1}^{[nt]} \mathrm{E}[\,(\xi(j)_i^n)^2]}$$

and so using (28)

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \mathrm{E}[\,|g'(\beta_i^n)\cdot\xi(j)_i^n|\,] \leq C\,t\cdot\sqrt{\sum_{i=1}^{[nt]} \mathrm{E}[\,(\xi(j)_i^n)^2]}$$

since almost surely

$$|g'(\beta_i^n)| \le C\,(1 + |\beta_i^n|^p)$$

for all $i, n \ge 1$. From here, (37) is an immediate consequence of Lemmas 1, 3 and 4.

The remaining case $k = 2$ is different. The definition of $\psi(2)_i^n$ implies, using basic stochastic calculus, that $\psi(2)_i^n/\sqrt{n}$, for all $i, n \ge 1$, may be written as

$$\begin{aligned}
&\int_{(i-1)/n}^{i/n} \left\{\sigma'_{\frac{i-1}{n}} \left(W_u - W_{\frac{i-1}{n}}\right) + M(n,i)_u\right\} \mathrm{d}W_u \\
= \;& \sigma'_{\frac{i-1}{n}} \int_{(i-1)/n}^{i/n} \left(W_u - W_{\frac{i-1}{n}}\right) \mathrm{d}W_u \\
& + \triangle_i^n M(n,i) \cdot \triangle_i^n W \\
& + \int_{(i-1)/n}^{i/n} \left(W_u - W_{\frac{i-1}{n}}\right) \mathrm{d}M(n,i)_u,
\end{aligned}$$

where $(M(n,i)_t)$ is the martingale defined by $M(n,i)_t \equiv 0$ for $t \le (i-1)/n$ and

$$M(n,i)_t = v^*_{\frac{i-1}{n}} \left(V_t - V_{\frac{i-1}{n}}\right) + \int_{(i-1)/n}^{t} \int_{E_n} \phi\left(\frac{i-1}{n}, x\right) (\mu - \nu)(\mathrm{d}s\,\mathrm{d}x)$$

otherwise. Thus for fixed $i, n \ge 1$

$$\mathrm{E}\left[g'(\beta_i^n) \cdot \psi(2)_i^n \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] \,/\sqrt{n}$$

is a linear combination of the following three terms

$$\mathrm{E}\left[g'(\beta_i^n) \cdot \sigma'_{\frac{i-1}{n}} \int_{(i-1)/n}^{i/n} \left(W_u - W_{\frac{i-1}{n}}\right) \mathrm{d}W_u \,|\, \mathcal{F}_{\frac{i-1}{n}}\right],$$

$$\mathrm{E}\left[g'(\beta_i^n) \cdot \triangle_i^n M(n,i) \cdot \triangle_i^n W \,|\, \mathcal{F}_{\frac{i-1}{n}}\right]$$

and

$$\mathrm{E}[\, g'(\beta_i^n) \cdot \int_{(i-1)/n}^{i/n} W_u \,\mathrm{d}M(n,i)_u \,|\, \mathcal{F}_{\frac{i-1}{n}}\,].$$

But these three terms are all equal to 0 as seen by the following arguments.

The conditional distribution of

$$\left(W_t - W_{\frac{i-1}{n}}\right)_{t \ge \frac{i-1}{n}} |\mathcal{F}_{\frac{i-1}{n}}$$

is clearly not affected by a change of sign. Thus since g being assumed even and g' therefore odd we have

$$\mathrm{E}\left[g'(\beta_i^n) \int_{(i-1)/n}^{i/n} \left(W_u - W_{\frac{i-1}{n}}\right) \mathrm{d}W_u \,|\, \mathcal{F}_{\frac{i-1}{n}}\right] = 0$$

implying the vanishing of the first term.

Secondly, by assumption, $\left(W_t - W_{\frac{i-1}{n}}\right)_{t\geq\frac{i-1}{n}}$ and $(M(n,i)_t)_{t\geq\frac{i-1}{n}}$ are independent given $\mathcal{F}_{\frac{i-1}{n}}$. Therefore, denoting by $\mathcal{F}^0_{i,n}$ the σ-field generated by

$$\left(W_t - W_{\frac{i-1}{n}}\right)_{\frac{i-1}{n}\leq t\leq i/n} \quad \text{and} \quad \mathcal{F}_{\frac{i-1}{n}},$$

the martingale property of $(M(n,i)_t)$ ensures that

$$\mathrm{E}[\, g'(\beta_i^n) \cdot \triangle_i^n M(n,i) \cdot \triangle_i^n W \,|\, \mathcal{F}^0_{i,n}\,] = 0$$

and

$$\mathrm{E}[\left[g'(\beta_i^n) \cdot \int_{(i-1)/n}^{i/n} W_u \,\mathrm{d}M(n,i)_u \,|\, \mathcal{F}^0_{i,n} \right] = 0.$$

Using this the vanishing of

$$\mathrm{E}\left[g'(\beta_i^n) \cdot \triangle_i^n M(n,i) \cdot \triangle_i^n W \,|\, \mathcal{F}_{\frac{i-1}{n}} \right]$$

and

$$\mathrm{E}\left[g'(\beta_i^n) \cdot \int_{(i-1)/n}^{i/n} W_u \,\mathrm{d}M(n,i)_u \,|\, \mathcal{F}_{\frac{i-1}{n}} \right]$$

is easily obtained by successive conditioning.

The proof of (31) is hereby completed.

□

References

Aït-Sahalia, Y. and J. Jacod (2005). Volatility estimators for discretely sampled Lévy processes. Unpublished paper: Department of Economics, Princeton University.

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

Andersen, T. G. and T. Bollerslev (1998). Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance 53*, 219–265.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2003). Some like it smooth, and some like it rough: untangling continuous and jump components in measuring, modeling and forecasting asset return volatility. Unpublished paper: Economics Dept, Duke University.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2006). Parametric and nonparametric measurement of volatility. In Y. Aït-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland. Forthcoming.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55. Correction published in 2003, volume 98, page 501.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Bandi, F. M. and J. R. Russell (2003). Microstructure noise, realized volatility, and optimal sampling. Unpublished paper, Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2005). A central limit theorem for realised power and bipower variations of continuous semimartingales. In Y. Kabanov and R. Lipster (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*. Springer. Forthcoming.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2004). Regular and modified kernel-based estimators of integrated variance: the case with independent noise. Unpublished paper: Nuffield College, Oxford.

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

Barndorff-Nielsen, O. E. and N. Shephard (2003). Realised power variation and stochastic volatility. *Bernoulli 9*, 243–265. Correction published in pages 1109–1111.

Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica 72*, 885–925.

Barndorff-Nielsen, O. E. and N. Shephard (2004b). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics 2*, 1–48.

Barndorff-Nielsen, O. E. and N. Shephard (2005a). How accurate is the asymptotic approximation to the distribution of realised volatility? In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg*, pp. 306–331. Cambridge: Cambridge University Press.

Barndorff-Nielsen, O. E. and N. Shephard (2005b). Variation, jumps, market frictions and high frequency data in financial econometrics. Unpublished discussion paper: Nuffield College, Oxford.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*. Forthcoming.

Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel (2004). Limit theorems for multipower variation in the presence of jumps in financial econometrics. Unpublished paper: Nuffield College, Oxford.

Calvet, L. and A. Fisher (2002). Multifractality in asset returns: theory and evidence. *Review of Economics and Statistics 84*, 381–406.

Christensen, K. and M. Podolskij (2005). Asymptotic theory for range-based estimation of integrated volatility of a continuous semi-martingale. Unpublished paper: Aarhus School of Business.

Corradi, V. and W. Distaso (2004). Specification tests for daily integrated volatility, in the presence of possible jumps. Unpublished paper: Queen Mary College, London.

Delattre, S. and J. Jacod (1997). A central limit theorem for normalized functions of the increments of a diffusion process in the presence of round off errors. *Bernoulli 3*, 1–28.

Doob, J. L. (1953). *Stochastic Processes*. New York: John Wiley and Sons.

Forsberg, L. and E. Ghysels (2004). Why do absolute returns predict volatility so well. Unpublished paper: Economics Department, UNC, Chapel Hill.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*, 3–29.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*. Forthcoming.

Gloter, A. and J. Jacod (2001a). Diffusions with measurement errors. I — local asymptotic normality. *ESAIM: Probability and Statistics 5*, 225–242.

Gloter, A. and J. Jacod (2001b). Diffusions with measurement errors. II — measurement errors. *ESAIM: Probability and Statistics 5*, 243–260.

Goncalves, S. and N. Meddahi (2004). Bootstrapping realized volatility. Unpublished paper, CIRANO, Montreal.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and its Applications*. San Diego: Academic Press.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise (with discussion). *Journal of Business and Economic Statistics 24*. Forthcoming.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variation. *Journal of Financial Econometrics 3*, 456–499.

Jacod, J. (1994). Limit of random measures associated with the increments of a Brownian semimartingale. Preprint number 120, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.

Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire Probability XXXI, Lecture Notes in Mathematics*, Volume 1655, pp. 232–246. Berlin: Springer Verlag.

Jacod, J., A. Lejay, and D. Talay (2005). Testing the multiplicity of a diffusion. In preparation, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.

Jacod, J. and P. Protter (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability 26*, 267–307.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2 ed.). Springer-Verlag: Berlin.

Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus* (2 ed.), Volume 113 of *Graduate Texts in Mathematics*. Berlin: Springer–Verlag.

Karatzas, I. and S. E. Shreve (1998). *Methods of Mathematical Finance*. New York: Springer–Verlag.

Maheswaran, S. and C. A. Sims (1993). Empirical implications of arbitrage-free asset markets. In P. C. B. Phillips (Ed.), *Models, Methods and Applications of Econometrics*, pp. 301–316. Basil Blackwell.

Mancini, C. (2004). Estimation of the characteristics of jump of a general Poisson-diffusion process. *Scandinavian Actuarial Journal 1*, 42–52.

Martens, M. and D. van Dijk (2005). Measuring volatility with the realized range. Unpublished paper: Econometric Institute, Erasmus University, Rotterdam.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*, 323–361.

Munroe, M. E. (1953). *Introduction to Measure and Integration*. Cambridge, MA: Addison-Wesley Publishing Company, Inc.

Mykland, P. A. and L. Zhang (2006). ANOVA for diffusions. *Annals of Statistics 33*. Forthcoming.

Nielsen, M. O. and P. H. Frederiksen (2005). Finite sample accuracy of integrated volatility estimators. Unpublished paper, Department of Economics, Cornell University.

Officer, R. R. (1973). The variability of the market factor of the New York stock exchange. *Journal of Business 46*, 434–453.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*, 61–66.

Phillips, P. C. B. and S. Ouliaris (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica 58*, 165–193.

Revuz, D. and M. Yor (1999). *Continuous Martingales and Brownian motion* (3 ed.). Heidelberg: Springer-Verlag.

Rosenberg, B. (1972). The behaviour of random variables with nonstationary variance and the distribution of security prices. Working paper 11, Graduate School of Business Administration, University of California, Berkeley. Reprinted in Shephard (2005).

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance 44*, 1115–1153.

Schwert, G. W. (1990). Indexes of U.S. stock prices from 1802 to 1987. *Journal of Business 63*, 399–426.

Schwert, G. W. (1998). Stock market volatility: Ten years after the crash. *Brookings-Wharton Papers on Financial Services 1*, 65–114.

Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.

Shiryaev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models and Theory*. Singapore: World Scientific.

Woerner, J. (2006). Power and multipower variation: inference for high frequency data. In A. N. Shiryaev, M. do Rosario Grossihno, P. Oliviera, and M. Esquivel (Eds.), *Proceedings of the International Conference on Stochastic Finance 2004*. Berlin: Springer Verlag. Forthcoming.

Zhang, L. (2004). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. Unpublished paper: Department of Statistics, Carnegie Mellon University.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*. Forthcoming.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics 14*, 45–52.