Econometric Theory, **22**, 2006, 677–719. Printed in the United States of America.
DOI: 10.1017/S0266466606060324

LIMIT THEOREMS FOR BIPOWER VARIATION IN FINANCIAL ECONOMETRICS

OLE E. BARNDORFF-NIELSEN AND SVEND ERIK GRAVERSEN
University of Aarhus

JEAN JACOD
Université Pierre et Marie Curie

NEIL SHEPHARD
University of Oxford

In this paper we provide an asymptotic analysis of generalized bipower measures of the variation of price processes in financial economics. These measures encompass the usual quadratic variation, power variation, and bipower variations that 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.

1. INTRODUCTION

In this paper we discuss the limiting theory for a novel, unifying class of nonparametric measures of the variation of financial prices. The theory covers commonly used estimators of variation such as realized volatility, but it also encompasses more recently suggested quantities such as realized power variation and realized 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 (2006). 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 micro-

Ole E. Barndorff-Nielsen's work is supported by the Centre for Analytical Finance (CAF), 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. Address correspondence to Ole E. Barndorff-Nielsen, Department of Mathematical Sciences, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark; e-mail: oebn@imf.au.dk.

© 2006 Cambridge University Press 0266-4666/06 $12.00

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

structure 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 that directly model the impact of market frictions on realized 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, 2001b), whereas 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

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

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., realized 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

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

where n is a positive integer.

We will study the behavior of the realized generalized 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 nonparametric volatility measures used in financial econometrics fall within this class (a measure not included in this setup is the range statistic studied in, e.g., Parkinson, 1980, and its realized 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.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Example 1.

(a) Suppose $g(y) = (y^j)^2$ and $h(y) = 1$; then (2) becomes

$$\sum_{i=1}^{\lfloor nt \rfloor} (\Delta_i^n Y^j)^2, \qquad j = 1,2,\ldots,d,$$

which is called the realized quadratic variation process of Y^j in econometrics (e.g., Jacod, 1994; Jacod and Protter, 1998; Barndorff-Nielsen and Shephard, 2002, 2004a; Mykland and Zhang, 2006). The increments of this quantity, typically calculated over a day or a week, are often called the realized variances in financial economics. The importance of these increments has 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 (2006b). Realized 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,

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

This is the realized 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 $g(y) = |y^j|^r$ for $r > 0$ and $h(y) = 1$; then (2) becomes

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

which is called the realized rth-order power variation. When r is an integer it has been studied from a probabilistic viewpoint by Jacod (1994), whereas 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 that 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), and they have also been informally discussed by Shiryaev (1999, pp. 349–350) and Maheswaran and Sims (1993). Following the work by Barndorff-Nielsen and

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Shephard (2003), Ghysels, Santa-Clara, and Valkanov (2004) and Forsberg and Ghysels (2004) have successfully used realized power variation as an input into volatility forecasting competitions.

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

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

which is called the realized r,sth-order bipower variation process. This measure of variation was introduced by Barndorff-Nielsen and Shephard (2004b), and a more formal discussion of its behavior in the $r = s = 1$ case was developed by Barndorff-Nielsen and Shephard (2006a). These authors' interest in this quantity was motivated by its virtue of being resistant to finite activity jumps as long as $\max(r,s) < 2$. Recently Barndorff-Nielsen, Shephard, and Winkel (2006) and Woerner (2006) have studied how these results on jumps extend to infinite activity processes, and Corradi and Distaso (2006) 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

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

Then (2) becomes

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

Barndorff-Nielsen and Shephard (2006a) used the joint behavior 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 realized quadratic variation (QV), where the limit is the usual 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), thus extending a number of results previously given in the literature on special cases of this framework.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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, and there is an Appendix that 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, et al. (2006).

2. NOTATION AND MODELS

We start with Y on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$. 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, a is a d-dimensional process whose elements are predictable and that has locally bounded sample paths, and the spot covolatility d, d'-dimensional matrix σ has elements that 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,kth element of Σ_t and

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

Remark 1. Because of 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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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, pp. 170–172). Using the Dambis–Dubins–Schwartz theorem, we know that the difference between the entire continuous local martingale class and the stochastic volatility class is the local martingales that have only continuous, not absolutely continuous,[1] QV. The drift $\int_0^t a_u \, \mathrm{d}u$ has elements that 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; 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 functions f, g, h, etc. 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

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

We also need the following notation:

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

and

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

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 $g(y) = |y^j|^r$; then $\rho_\sigma(g) = \mu_r \sigma_j^r$, where σ_j^2 is the j,jth 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, et al. (2006), which is restated here.

THEOREM 1. *Under Assumption (H) and assuming that g and h satisfy Assumption (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.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

The result is quite clean as it 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 that are currently receiving a great deal of attention as measures of variation in financial econometrics. Here we briefly discuss some of the work that has studied the limiting behavior of these objects.

Example 3.

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

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

the QV of Y^j. This well-known result in probability theory is behind much of the modern work on realized volatility, which is compactly reviewed in Barndorff-Nielsen and Shephard (2006b).

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

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

the well-known multivariate version of QV.

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

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

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

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

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

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

(Example 1(e) continued). As

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

so

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

Barndorff-Nielsen and Shephard 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 section we do this by giving a central limit theorem (CLT) 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) to derive this result.

We start with a variety of assumptions that strengthen Assumptions (H) and (K) given in Sections 2 and 3.

Assumption (H1). We have Assumption (H) with

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

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 $(E, \mathcal{E})$. The term ϕ is a continuous truncation function on $R^{dd'}$ (a function with compact support, which coincides with the identity map on a neighborhood of 0). Finally $w(\omega, u, x)$ is a map on $\Omega \times [0,\infty) \times E$ into the space of $d \times d'$arrays that 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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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 Assumption (H1). Assumption (H1) is rather general from an econometric viewpoint as it allows for flexible leverage effects, multifactor volatility effects, jumps, nonstationarities, intraday effects, and so on. Indeed we do not know of a continuous time volatility model used in financial economics that 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 because that alternative class of processes is not closed under squaring (further comment on this is given in Section A.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*

To derive a CLT we need to impose some regularity on g and h.

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

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 that may not be differentiable at zero, we sometimes need to relax (K1). The resulting condition is quite technical and is called (K2).

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

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

and we have

$$x \in B^c,$$

$$\|y\| \leq 1 \bigwedge \frac{d(x,B)}{2} \Rightarrow \begin{cases} \|\nabla f(x)\| \leq \dfrac{C(1+\|x\|^p)}{d(x,B)^{1-r}}, \\ \|\nabla f(x+y) - \nabla f(x)\| \leq \dfrac{C(1+\|x\|^p)\|y\|}{d(x,B)^{2-r}}. \end{cases} \tag{10}$$

Remark 5. These conditions accommodate the case where f equals $|x^j|^r$: this function satisfies Assumption (K1) when $r > 1$, and Assumption (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 Assumption (K1) implies Assumption (K2) with $r = 1$ and $B = \varnothing$. 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), is sufficient for the statement of Theorem 1.3 of Barndorff-Nielsen, Graversen, et al. (2006) to hold.

Assumption (J1). We have Assumption (H1) and g and h satisfy Assumption (K1).

Assumption (J2). We have Assumptions (H1) and (H2) and g and h satisfy Assumption (K2).

Clearly Assumption (J2) makes stronger assumptions about the volatility process and weaker assumptions about the functions g and h, than Assumption (J1). It is Assumption (J2) that will be used when analyzing 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 Assumptions (J1) and (J2) holds; then the process*

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

Furthermore, B is a standard Wiener process that is defined on an extension of $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},P)$ 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 $n\to\infty$,

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

If X_n converges stably in law then, in particular, it converges in distribution (or in law or weak convergence), the limiting law being $\mu(\Omega,\cdot)$. Accordingly, one says that X_n converges stably to some random variable X if there exists a probability measure μ as before such that X has law $\mu(\Omega,\cdot)$. This concept and its extension to stable convergence of processes are 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 the work of Phillips and Ouliaris (1990) on the limit distribution of cointegration tests.

However, this formalization 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 of the present paper consider the following simple example of the preceding result. Let

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

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

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

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

In particular, for a single point in time t,

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

We consider two concrete examples of this setup.

(i) Power variation. Suppose $g(y) = |y^j|^r$ 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

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

where $v_r = \mathrm{Var}(|u|^r)$ and $u \sim N(0,1)$. When $r = 2$, this yields a CLT for the realized 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 that appears in Jacod (1994), Mykland and Zhang (2006), and, implicitly, Jacod and Protter (1998), whereas 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 CLT for $U(g,h)_1$. Their result ruled out leverage effects, which are allowed under Theorem 2. The finite-sample behavior of this type of limit theory is studied in, for example, Barndorff-Nielsen and Shephard (2005), Goncalves and Meddahi (2004), and Nielsen and Frederiksen (2005).

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

$$\rho_\sigma(g) = \mu_r \sigma_j^r, \qquad \rho_\sigma(h) = \mu_s \sigma_j^s, \qquad \rho_\sigma(gg) = \mu_{2r}\sigma_j^{2r},$$

$$\rho_\sigma(hh) = \mu_{2s}\sigma_j^{2s}, \qquad \rho_\sigma(gh) = \mu_{r+s}\sigma_j^{r+s}.$$

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

This implies that

$$A(\sigma,g,h) = \mu_{2r}\sigma_j^{2r}\mu_{2s}\sigma_j^{2s} + 2\mu_r\sigma_j^r\mu_s\sigma_j^s\mu_{r+s}\sigma_j^{r+s} - 3\mu_r^2\sigma_j^{2r}\mu_s^2\sigma_j^{2s}$$

$$= (\mu_{2r}\mu_{2s} + 2\mu_{r+s}\mu_r\mu_s - 3\mu_r^2\mu_s^2)\sigma_j^{2r+2s}.$$

In the $r = s = 1$ case Barndorff-Nielsen and Shephard (2006a) derived, under much stronger conditions, a CLT 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,

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

so

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

$$= \Sigma^{jj'}\Sigma^{kk'} + \Sigma^{jk'}\Sigma^{kj'}.$$

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).

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Example 6.

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

$$U(g,h)_t^j = \sum_{j'=1}^{2} \int_0^t \alpha(\sigma_u, g, h)^{j,j'} \, \mathrm{d}B_u^{j'},$$

where

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

Here

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

Example 7.

Joint behavior of realized QV and realized bipower variation. This sets

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

The implication is that

$$\begin{aligned} &\rho_\sigma(g^{11}) = \rho_\sigma(g^{22} g^{11}) = \rho_\sigma(g^{11} g^{22}) = \mu_1 \sigma_j, \qquad \rho_\sigma(g^{22}) = 1, \\ &\rho_\sigma(g^{11} g^{11}) = \sigma_j^2, \qquad \rho_\sigma(g^{22} g^{22}) = 1, \qquad \rho_\sigma(h^1) = \mu_1 \sigma_j, \\ &\rho_\sigma(h^2) = \rho_\sigma(h^1 h^1) = \sigma_j^2, \qquad \rho_\sigma(h^1 h^2) = \rho_\sigma(h^2 h^1) = \mu_3 \sigma_j^3, \\ &\rho_\sigma(h^2 h^2) = 3\sigma_j^4, \qquad \rho_\sigma(g^{11} h^1) = \sigma_j^2, \qquad \rho_\sigma(g^{11} h^2) = \mu_3 \sigma_j^3, \\ &\rho_\sigma(g^{22} h^1) = \mu_1 \sigma_j, \qquad \rho_\sigma(g^{22} h^2) = \sigma_j^2. \end{aligned}$$

Thus

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

whereas

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

and

$$A(\sigma,g,h)^{1,2} = \mu_1\sigma_j\mu_3\sigma_j^3 + \mu_1\sigma_j\sigma_j^2\mu_1\sigma_j + \mu_1\sigma_j\mu_3\sigma_j^3 - 3\mu_1\sigma_j\mu_1\sigma_j\sigma_j^2$$
$$= 2\sigma_j^4(\mu_1\mu_3 - \mu_1^2) = 2\mu_1^2\sigma_j^4.$$

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

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

5. MULTIPOWER VARIATION

A natural extension of generalized bipower variation is to generalized 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 (2006a).

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 suggest 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, et al. (2006) for more details.

Example 8.

(a) Suppose $I = 4$ and $g_i(y) = |y^j|$; 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.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

(b) Suppose $I = 3$ and $g_i(y) = |y^j|^{4/3}$; 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 $g_i(y) = |y^j|^{2/I}$, 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 realized 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, Graverson, et al. (2006) show that when Assumptions (H1) and (H2) hold then

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

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

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

6. SUMS OF REALIZED GENERALIZED BIPOWER

The law of large numbers and the CLT also hold for linear combinations of processes like $Y(g)$ in the preceding discussion.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Example 10.

Let ζ_i^n denote 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 combination 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 Assumption (H)

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

whereas under Assumptions (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 independent and identically distributed (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 realized generalized bipower variation. In the case of power variation and bipower variation the results are proved under much weaker assumptions than those that 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 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 realized 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).

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

A challenging extension is to construct a version of realized generalized bipower variation that is robust to market frictions. Following the work on the realized volatility there are two strategies that may be able to help: the kernel-based approach, studied in detail by Barndorff-Nielsen et al. (2004), and the subsampling approach of Zhang, Mykland, and Aït-Sahalia (2005) and Zhang (2004). In the realized volatility case these methods are basically equivalent; however it is perhaps the case that the subsampling method is easier to extend to the nonquadratic 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 realized variance.

NOTE

1. An example of a continuous local martingale that 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 nondecreasing but not absolutely continuous (see, e.g., Munroe, 1953, Sect. 27). This relates to the work of, for example, Calvet and Fisher (2002) on multifractals.

REFERENCES

Aït-Sahalia, Y. & J. Jacod (2005) Volatility Estimators for Discretely Sampled Lévy Processes. Manuscript, Department of Economics, Princeton University.

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

Andersen, T.G. & 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, & 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. Manuscript, Economics Department, Duke University.

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

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

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

Bandi, F.M. & J.R. Russell (2003) Microstructure Noise, Realized Volatility, and Optimal Sampling. Manuscript, Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O.E., S.E. Graversen, J. Jacod, M. Podolskij, & N. Shephard (2006) A central limit theorem for realised power and bipower variations of continuous semimartingales. In Y. Kabanov & R. Lipster (eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*, pp. 33–68. Springer.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde, & N. Shephard (2004) Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Unpublished paper, Nuffield College, Oxford.

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

Barndorff-Nielsen, O.E. & N. Shephard (2006a) Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 217–252.

Barndorff-Nielsen, O.E. & N. Shephard (2006b) Variation, jumps, market frictions and high frequency data in financial econometrics. In Richard Blundell, Torsten Persson, and Whitney K. Newey (eds.), *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress*, Econometric Society Monographs. Cambridge University Press, forthcoming.

Barndorff-Nielsen, O.E., N. Shephard, & M. Winkel (2006) Limit theorems for multipower variation in the presence of jumps in financial econometrics. *Stochastic Processes and Their Applications*, forthcoming.

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

Christensen, K. & M. Podolskij (2005) Asymptotic Theory for Range-Based Estimation of Integrated Volatility of a Continuous Semi-Martingale. Manuscript. Aarhus School of Business.

Corradi, V. & W. Distaso (2006) Semiparametric comparison of stochastic volatility models via realized measures. *Review of Economic Studies*, forthcoming.

Delattre, S. & 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*. Wiley.

Forsberg, L. & E. Ghysels (2004) Why Do Absolute Returns Predict Volatility So Well? Unpublished paper, Economics Department, University of North Carolina, Chapel Hill.

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

Ghysels, E., P. Santa-Clara, & R. Valkanov (2004) Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131, 59–95.

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

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

Goncalves, S. & N. Meddahi (2004) Bootstrapping Realized Volatility. Manuscript, CIRANO, Montreal.

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Applications*. Academic Press.

Hansen, P.R. & A. Lunde (2006) Realized variance and market microstructure noise (with discussion) *Journal of Business & Economic Statistics* 24, 127–218.

Huang, X. & 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 120, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.

Jacod, J. (1997) On continuous conditional Gaussian martingales and stable convergence in law. In Jacques Azema, M. Emery, & Marc Yor (eds.), *Séminaire Probability XXXI, Lecture Notes in Mathematics*, Volume 1655, pp. 232–246. Springer-Verlag.

Jacod, J., A. Lejay, & D. Talay (2005) Testing the Multiplicity of a Diffusion. Manuscript, Laboratoire de Probabilitiés, Université Pierre et Marie Curie, Paris.

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Jacod, J. & A.N. Shiryaev (2003) *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag.
Karatzas, I. & S.E. Shreve (1991) *Brownian Motion and Stochastic Calculus*, 2nd ed. Graduate Texts in Mathematics 113. Springer-Verlag.
Karatzas, I. & S.E. Shreve (1998) *Methods of Mathematical Finance*. Springer-Verlag.
Maheswaran, S. & 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. & D. van Dijk (2005) Measuring Volatility with the Realized Range. Manuscript, 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*. Addison-Wesley.
Mykland, P.A. & L. Zhang (2006) ANOVA for diffusions and Ito processes. *Annals of Statistics* 33, forthcoming.
Nielsen, M.O. & P.H. Frederiksen (2005) Finite Sample Accuracy of Integrated Volatility Estimators. Manuscript, 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. & S. Ouliaris (1990) Asymptotic properties of residual based tests for cointegration. *Econometrica* 58, 165–193.
Revuz, D. & M. Yor (1999) *Continuous Martingales and Brownian Motion*, 3rd ed. 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 N. Shephard, *Stochastic Volatility: Selected Readings* (Oxford University Press, 2005, pp. 83–108).
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 University Press.
Shiryaev, A.N. (1999) *Essentials of Stochastic Finance: Facts, Models and Theory*. World Scientific.
Woerner, J. (2006) Power and multipower variation: Inference for high frequency data. In A.N. Shiryaev, M. do Rosario Grossihno, P. Oliviera, & M. Esquivel (eds.), *Proceedings of the International Conference on Stochastic Finance 2004*, pp. 343–364. Springer-Verlag.
Zhang, L. (2004) Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach. Manuscript, Department of Statistics, Carnegie Mellon University.
Zhang, L., P.A. Mykland, & Y. Aït-Sahalia (2005) A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394–1411.
Zhou, B. (1996) High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics* 14, 45–52.

APPENDIX: Techniques for the Proof of Theorem 2

A.1. Notational Conventions. Here we give a fairly detailed account of the basic techniques in the proof of Theorem 2, in the one-dimensional case and under some rel-

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

atively 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 readers who wish to see the more general case to Barndorff-Nielsen, Graversen, et al. (2006). The outline of this section is as follows. First we introduce our basic notation, and in Section A.2 we set out the model and review the assumptions we use. In Section A.3 we state the theorem we will prove and outline the steps in the proof. Sections A.5, A.6, and A.7 give the proofs of the successive steps.

All processes mentioned in the following discussion 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 $\triangle_i^n Z := Z_{i/n} - Z_{(i-1)/n}$, $i, n \geq 1$.

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). 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

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

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

Often

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

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

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

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 that 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.

A.2. Model and Basic Assumptions. Throughout the following discussion (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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

where

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

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 because 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 assumptions.

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. 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 Assumption (Ka). The remaining case $0 < r < 1$ requires special arguments that will be omitted here (for details, see Barndorff-Nielsen, Graversen, et al., 2006).

To prove the central limit theorem we need some additional structure on the volatility process (σ_t). A natural set of assumptions follows.

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 modeling volatility it is often more natural to define (σ_t^2) as being of the preceding form, that is,

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

$$+ \int_0^t \int_E q \circ \phi(s-, x)(\mu - \nu)(\mathrm{d}s \, \mathrm{d}x)$$

$$+ \int_0^t \int_E \{\phi(s-, x) - q \circ \phi(s-, x)\}\mu(\mathrm{d}s \, \mathrm{d}x).$$

Here $(a_t^*), (\sigma_t^*)$, and (v_t^*), are as in Assumption (H0) and (V_t) is another $((\mathcal{F}_t), P)$-Wiener process independent of (W_t), whereas q is a continuous truncation function on $\mathbf{R}$, that is, a function with compact support coinciding with the identity on a neighborhood of 0. Further, μ is a Poisson random measure on $(0,\infty) \times E$ independent of (W_t) and (V_t) and 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 $\mathbf{R}$ that 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 $k \geq 1$

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

where

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

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

Finally we shall also assume nondegeneracy 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 $(Z_{T_k \wedge t}^n)$ for at least one sequence of stopping times (T_k) increasing to $+\infty$. Applying this together with standard localization techniques (for details, see Barndorff-Nielsen, Graversen, et al., 2006), we may assume that the following more restrictive assumptions are satisfied.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Assumption (H1a). (σ_t) can be written as

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

$$+ \int_0^t \int_E \phi(s-, x)(\mu - \nu)(\mathrm{d}s \, \mathrm{d}x) \qquad t \geq 0.$$

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 $\mathbf{R}$ that is $\mathcal{F}_s \otimes \mathcal{E}$ measurable in (ω, x) for all s and càdlàg in s, satisfying furthermore

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

Likewise, by a localization argument, we may make the following assumption.

Assumption (H2a). (σ_t) satisfies $a < \sigma_t^2(\omega) < b$ 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

$$\mathrm{E}[|\sigma_v - \sigma_u|^2] \leq C(v - u), \quad \text{for all } 0 \leq u < v. \tag{A.1}$$

We use $\Upsilon(x)$ as 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 Assumption (H2a))

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

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

A.3. Main Result. As already mentioned, our aim is to show the following special version of the general CLT result given as Theorem 2.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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*

$$\left(\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} g(\sqrt{n}\triangle_i^n Y) - \int_0^t \rho_{\sigma_u}(g)\,\mathrm{d}u\right)\right) \to \int_0^t \sqrt{\rho_{\sigma_{u-}}(g^2) - \rho_{\sigma_{u-}}(g)^2}\,\mathrm{d}B_u, \tag{A.3}$$

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 (A.3) as follows:

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

It is relatively simple to show that the first term of the right-hand side satisfies

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

Hence what remains is to verify that uniformly

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor}\mathrm{E}[g(\triangle_i^n Y)|\mathcal{F}_{(i-1)/n}] - \int_0^t \rho_{\sigma_u}(g)\,\mathrm{d}u\right) \xrightarrow{p} 0. \tag{A.4}$$

We have

$$\begin{aligned}
&\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor}\mathrm{E}[g(\triangle_i^n Y)|\mathcal{F}_{(i-1)/n}] - \int_0^t \rho_{\sigma_u}(g)\,\mathrm{d}u\right)\\
&\quad= \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\mathrm{E}[g(\triangle_i^n Y)|\mathcal{F}_{(i-1)/n}] - \sqrt{n}\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}\rho_{\sigma_u}(g)\,\mathrm{d}u\\
&\qquad+ \sqrt{n}\left(\sum_{i=1}^{\lfloor nt\rfloor}\int_{(i-1)/n}^{i/n}\rho_{\sigma_u}(g)\,\mathrm{d}u - \int_0^t \rho_{\sigma_u}(g)\,\mathrm{d}u\right),
\end{aligned} \tag{A.5}$$

where, uniformly

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

The first term on the right-hand side of (A.5) is now split into the difference of

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \{\mathrm{E}[g(\triangle_i^n Y)|\mathcal{F}_{(i-1)/n}] - \rho_{(i-1)/n}\}, \tag{A.6}$$

where

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

and

$$\sqrt{n} \sum_{i=1}^{\lfloor nt \rfloor} \int_{(i-1)/n}^{i/n} \{\rho_{\sigma_u}(g)\,\mathrm{d}u - \rho_{(i-1)/n}\}\,\mathrm{d}u. \tag{A.7}$$

It is rather easy to show that (A.6) tends to 0 in probability uniformly in t. The challenge is thus to show that the same result holds for (A.7).

To handle (A.7) one splits the individual terms in the sum into

$$\sqrt{n}\,\Upsilon'(\sigma_{(i-1)/n}) \int_{(i-1)/n}^{i/n} (\sigma_u - \sigma_{(i-1)/n})\,\mathrm{d}u \tag{A.8}$$

plus

$$\sqrt{n} \int_{(i-1)/n}^{i/n} \{\Upsilon(\sigma_u) - \Upsilon(\sigma_{(i-1)/n}) - \Upsilon'(\sigma_{(i-1)/n})\cdot(\sigma_u - \sigma_{(i-1)/n})\}\,\mathrm{d}u, \tag{A.9}$$

where $\Upsilon(x)$ is shorthand for $\rho_x(g)$ and $\Upsilon'(x)$ denotes the derivative with respect to x. That (A.9) 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 (A.8) 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 is even comes into play and is crucial.

A.4. Details of the Proof. Introducing the notation

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

we may reexpress (A.3) as

$$\left(\sqrt{n}\left(Y_t^n(g) - \int_0^t \sigma_u(g)\,\mathrm{d}u\right)\right) \to (U_t(g)). \tag{A.10}$$

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

$$\beta_i^n = \sqrt{n} \cdot \sigma_{(i-1)/n} \cdot \triangle_i^n W, \qquad i, n \geq 1.$$

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

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

and (σ_t) is uniformly bounded, a straightforward application of (A.1) and the Burkholder–Davis–Gundy inequalities (e.g., Revuz and Yor, 1999, pp. 160–171) gives for every $p > 0$ the following simple estimates:

$$\mathrm{E}[|\sqrt{n}\triangle_i^n Y - \beta_i^n|^p | \mathcal{F}_{(i-1)/n}] \leq \frac{C_p}{n^{p\wedge 1}} \tag{A.11}$$

and

$$\mathrm{E}[|\sqrt{n}\triangle_i^n Y|^p + |\beta_i^n|^p | \mathcal{F}_{(i-1)/n}] \leq C_p \tag{A.12}$$

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

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

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

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

and

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

The asymptotic behavior of $(U_t^n(g))$ is well known. More precisely under 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, Thm. 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 before. Then*

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

Proof of Theorem 4. As pointed out just before it is enough to prove that

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

where

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

Thus we have to prove

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

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 $t > 0$

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Thus by means of Jensen inequality and the boundedness of (σ_t)

$$\sum_{i=1}^{\lfloor nt\rfloor} \mathrm{E}[(\xi_i^n)^2] \le C \sum_{i=1}^{\lfloor nt\rfloor} \mathrm{E}\left[\sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} (\sigma_{u-} - \sigma_{(i-1)/n})^2 \, \mathrm{d}u \right)^2 \middle| \mathcal{F}_{(i-1)/n} \right]} \right]$$

$$\le C \sum_{i=1}^{\lfloor nt\rfloor} \sqrt{\mathrm{E}\left[\left(\int_{(i-1)/n}^{i/n} (\sigma_{u-} - \sigma_{(i-1)/n})^2 \, \mathrm{d}u \right)^2 \right]}$$

$$\le C[nt] \frac{1}{[nt]} \sum_{i=1}^{\lfloor nt\rfloor} \sqrt{\mathrm{E}\left[\frac{1}{n} \int_{(i-1)/n}^{i/n} (\sigma_{u-} - \sigma_{(i-1)/n})^2 \, \mathrm{d}u \right]}$$

$$\le C[nt] \sqrt{\frac{1}{[nt]} \sum_{i=1}^{\lfloor nt\rfloor} \mathrm{E}\left[\frac{1}{n} \int_{(i-1)/n}^{i/n} (\sigma_{u-} - \sigma_{(i-1)/n})^2 \, \mathrm{d}u \right]}$$

$$\le C \sqrt{\sum_{i=1}^{\lfloor nt\rfloor} \mathrm{E}\left[\int_{(i-1)/n}^{i/n} (\sigma_{u-} - \sigma_{(i-1)/n})^2 \, \mathrm{d}u \right]},$$

which tends to 0 as $n \to \infty$, by Lebesgue's theorem and the boundedness of (σ_t). ■

To prove the convergence (A.10) it suffices, using Theorem 4, 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

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

and, as is easily seen,

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

the job is to prove that

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

where for $i, n \ge 1$

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

where

$$\eta(1)_i^n = \frac{1}{\sqrt{n}} \{ \mathrm{E}[g(\sqrt{n} \triangle_i^n Y) | \mathcal{F}_{(i-1)/n}] - \rho_{\sigma_{(i-1)/n}}(g) \} \tag{A.13}$$

and

$$\eta(2)_i^n = \sqrt{n} \int_{(i-1)/n}^{i/n} \{ \rho_{\sigma_u}(g) - \rho_{\sigma_{(i-1)/n}}(g) \} \, \mathrm{d}u. \tag{A.14}$$

We will now separately prove

$$\eta(1)^n = \sum_{i=1}^{\lfloor nt \rfloor} \eta(1)_i^n \xrightarrow{P} 0 \tag{A.15}$$

and

$$\eta(2)^n = \sum_{i=1}^{\lfloor nt \rfloor} \eta(2)_i^n \xrightarrow{P} 0. \tag{A.16}$$

A.5. Some Auxiliary Estimates. To show (A.15) and (A.16) we need some refinements of the estimate (A.1). To state these we split up $(\sqrt{n} \triangle_i^n Y - \beta_i^n)$ into several terms. By definition

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

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

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

the difference $\sigma_u - \sigma_{(i-1)/n}$ equals

$$\int_{(i-1)/n}^{u} a_s^* \, ds + \int_{(i-1)/n}^{u} \sigma_{s-}^* \, \mathrm{d}W_s + \int_{(i-1)/n}^{u} v_{s-}^* \, \mathrm{d}V_s + \int_{(i-1)/n}^{u} \int_E \phi(s-, x)(\mu - \nu)(\mathrm{d}s \, \mathrm{d}x)$$

$$= \sum_{j=1}^{5} \xi(j)_i^n(u),$$

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

$$\xi(1)_i^n(u) = \int_{(i-1)/n}^{u} a_s^* \, \mathrm{d}s + \int_{(i-1)/n}^{u} (\sigma_{s-}^* - \sigma_{(i-1)/n}^*) \, \mathrm{d}W_s + \int_{(i-1)/n}^{u} (v_{s-}^* - v_{(i-1)/n}^*) \, \mathrm{d}V_s,$$

$$\xi(2)_i^n(u) = \sigma_{(i-1)/n}^*(W_u - W_{(i-1)/n}) + v_{(i-1)/n}^*(V_u - V_{(i-1)/n}),$$

$$\xi(3)_i^n(u) = \int_{(i-1)/n}^{u} \int_{E_n^c} \phi(s-, x)(\mu - \nu)(\mathrm{d}s \, \mathrm{d}x),$$

$$\xi(4)_i^n(u) = \int_{(i-1)/n}^{u} \int_{E_n} \left\{ \phi(s-, x) - \phi\left(\frac{i-1}{n}, x \right) \right\} (\mu - \nu)(\mathrm{d}s \, \mathrm{d}x),$$

$$\xi(5)_i^n(u) = \int_{(i-1)/n}^{u} \int_{E_n} \phi\left(\frac{i-1}{n}, x \right) (\mu - \nu)(\mathrm{d}s \, \mathrm{d}x).$$

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

$$\sqrt{n} \triangle_i^n Y - \beta_i^n = \sum_{j=1}^{5} \xi(j)_i^n, \tag{A.17}$$

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

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

LEMMA 1.

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

LEMMA 2.

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

LEMMA 3.

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

where

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

LEMMA 4.

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

LEMMA 5.

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

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

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

$$\sum_{i=1}^{\lfloor nt \rfloor} d_i^n \to 0 \quad \text{for all } t > 0.$$

Taken together these statements imply the following result.

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

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

Subsequently we shall invoke this corollary in addition to Lemmas 2 and 5.

A.6. Proof of $\eta(2)^n \xrightarrow{P} 0$. Recall that we wish to show that

$$\eta(2)^n = \sum_{i=1}^{\lfloor nt \rfloor} \eta(2)_i^n \xrightarrow{P} 0. \tag{A.18}$$

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

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

and

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

With this notation we shall prove (A.18) by showing

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

and

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

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

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

where for all $i, n \geq 1$

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

and

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

By Assumption (H2a) and (A.2) 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

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

Because

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

are all martingales we have

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

By Doob's inequality it is therefore feasible to estimate

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

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

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

The properties of the Wiener integrals and the uniform boundedness of (σ_t^*) and (v_t^*) ensure that

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

and likewise

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

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

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Thus as $n \to \infty$ so

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

and because

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

we deduce from Doob's inequality that

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

proving altogether

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

Applying once more Assumption (H2a) and (A.2) we have for every $\epsilon > 0$ and every i, n that

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

Thus from (A.1) and its consequence

$$\mathrm{E}[|\sigma_u - \sigma_{(i-1)/n}|] \leq C/\sqrt{n}$$

we get

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

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

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

implying the convergence

$$\sum_{i=1}^{\lfloor nt \rfloor} \eta(2)_i^n \xrightarrow{P} 0$$

and thus ending the proof of (A.16). ■

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

A.7. Proof of $\eta(1)^n \xrightarrow{P} 0$. Recall that we are to show that

$$\eta(1)^n = \sum_{i=1}^{\lfloor nt \rfloor} \eta(1)_i^n \xrightarrow{P} 0. \tag{A.19}$$

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

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

$$= \frac{1}{\sqrt{n}} \mathrm{E}[g(\sqrt{n}\,\triangle_i^n Y) - g(\beta_i^n)|\mathcal{F}_{(i-1)/n}].$$

Introduce the notation (recall Assumption (K2))

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

Because 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 $\omega \in A_i^{nc}$. Thus, using the mean value theorem, we may for all $i, n \geq 1$ write

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

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

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

and

$$\delta(1)_i^n = [\{g(\sqrt{n}\,\triangle_i^n Y) - g(\beta_i^n)\} - g'(\beta_i^n) \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n)] \cdot \mathbf{1}_{A_i^n} / \sqrt{n}$$

$$\delta(2)_i^n = \{g'(\alpha_i^n) - g'(\beta_i^n)\} \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) \cdot \mathbf{1}_{A_i^{nc}} / \sqrt{n}$$

$$\delta(3)_i^n = g'(\beta_i^n) \cdot (\sqrt{n}\,\triangle_i^n Y - \beta_i^n) / \sqrt{n}.$$

Thus it suffices to prove

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

$$|g(\sqrt{n}\triangle_i^n Y) - g(\beta_i^n)| \le M(1 + |\sqrt{n}\triangle_i^n Y - \beta_i^n|^p + |\beta_i^n|^p)\cdot|\sqrt{n}\triangle_i^n Y - \beta_i^n|$$
$$\le (2^p + 1)M(1 + |\sqrt{n}\triangle_i^n Y|^p + |\beta_i^n|^p)\cdot|\sqrt{n}\triangle_i^n Y - \beta_i^n|$$

and

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

By Cauchy–Schwarz's inequality $\mathrm{E}[|\delta(1)_i^n|]$ is therefore for all $i, n \ge 1$ 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 (A.4), that

$$\mathrm{E}\left[\sum_{i=1}^{\lfloor nt\rfloor}|\delta(1)_i^n|\right] \le C\cdot\sup_{i\ge1}P(A_i^n)^{1/6}\sum_{i=1}^{\lfloor nt\rfloor}\mathrm{E}[(\triangle_i^n Y - \beta_i^n)^2/n]^{1/2}$$
$$\le C\cdot\sup_{i\ge1}P(A_i^n)^{1/6}\sum_{i=1}^{\lfloor nt\rfloor}1/n$$
$$\le Ct\cdot\sup_{i\ge1}P(A_i^n)^{1/6}.$$

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

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

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

$$P(A_i^n) \le \int_{\{x|d(x,B)\le\epsilon\}} h_{a,b}\,\mathrm{d}\lambda_1 + \frac{C}{n\epsilon^2} \tag{A.20}$$
$$= \alpha_\epsilon + \frac{C}{n\epsilon^2}.$$

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

Observe $\lim_{\epsilon\to 0}\alpha_\epsilon = 0$. Taking now in (A.20) sup over i and then first letting $n\to\infty$ and then $\epsilon\downarrow 0$ we get

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

proving that

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

and thus

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

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

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

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

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

Now writing

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

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

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

Putting this together means that

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

Exploiting here the inequalities (A.11) and (A.12) 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:

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

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

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

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

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

and thus

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

So what remains to be proved is the convergence

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

As introduced in (A.17)

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

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

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

The case $k = 1$ is handled by proving

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \mathrm{E}[|g'(\beta_i^n) \cdot \xi(j)_i^n|] \to 0, \qquad j = 1,3,4. \tag{A.21}$$

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

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \mathrm{E}[|g'(\beta_i^n) \cdot \xi(j)_i^n|] \leq Ct \cdot \sqrt{\frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathrm{E}[g'(\beta_i^n)^2]} \cdot \sqrt{\sum_{i=1}^{\lfloor nt \rfloor} \mathrm{E}[(\xi(j)_i^n)^2]}$$

and so using (A.12)

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \mathrm{E}[|g'(\beta_i^n) \cdot \xi(j)_i^n|] \leq Ct \cdot \sqrt{\sum_{i=1}^{\lfloor nt \rfloor} \mathrm{E}[(\xi(j)_i^n)^2]}$$

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

because almost surely

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

for all $i, n \ge 1$. From here, (A.21) 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

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

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

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

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

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

is a linear combination of the following three terms:

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

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

and

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

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

The conditional distribution of

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

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

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

implying the vanishing of the first term.

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press

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

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

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

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

and

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

is easily obtained by successive conditioning.

The proof of (A.15) is hereby completed. ■

https://doi.org/10.1017/S0266466606060324 Published online by Cambridge University Press