A weak law of large numbers for estimating the correlation in bivariate Brownian semistationary processes

Andrea Granelli[*]   Almut E. D. Veraart[†]

Department of Mathematics, Imperial College London
180 Queen's Gate, London, SW7 2AZ, UK

July 26, 2017

Abstract

This article presents various weak laws of large numbers for the so-called realised covariation of a bivariate stationary stochastic process which is not a semimartingale. More precisely, we consider two cases: Bivariate moving average processes with stochastic correlation and bivariate Brownian semistationary processes with stochastic correlation. In both cases, we can show that the (possibly scaled) realised covariation converges to the integrated (possibly volatility modulated) stochastic correlation process.

Keywords: Weak law of large numbers, moving average process, Brownian semistationary process, stochastic correlation, multivariate setting, high frequency data.
MSC: 60F05, 60F15, 60G15

1   Introduction

The aim of this article is to construct consistent estimators for the (possibly) stochastic correlation between two stochastic processes *outside* the semimartingale framework.

In the semimartingale case, the corresponding results are well-known. For instance, for a bivariate semimartingale $\mathbf{Y} = (Y^{(1)}, Y^{(2)})^\top$ the quadratic covariation denoted by $[Y^{(1)}, Y^{(2)}]$ exists and can be approximated by the so-called realised covariation. More precisely, for $n \in \mathbb{N}$ we write $\Delta_n = n^{-1}$; then we have for $t \geq 0$ that

$$\sum_{i=1}^{\lfloor nt \rfloor} \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \overset{u.c.p.}{\to} [Y^{(1)}, Y^{(2)}]_t, \quad \text{where } \Delta_i^n Y^{(j)} = Y^{(j)}_{i\Delta_n} - Y^{(j)}_{(i-1)\Delta_n}, j \in \{1, 2\},$$

where the convergence is uniform on compacts in probability (u.c.p.) as $n \to \infty$. We know that as soon as we drop the assumption that $\mathbf{Y}$ is a semimartingale, its quadratic covariation does not necessarily exist anymore. Hence we would like to answer the question whether a

---

[*]E-mail: `a.granelli12@imperial.ac.uk`
[†]E-mail: `a.veraart@imperial.ac.uk`

weak law of large numbers can be formulated for a possibly scaled version of the realised covariation

$$\sum_{i=1}^{\lfloor nt \rfloor} \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}$$

when $\mathbf{Y}$ is *not* a semimartingale. We tackle this question in a semi-parametric setting: First of all, we define the class of bivariate moving average processes with stochastic correlation and later extend this class to bivariate Brownian semistationary processes with stochastic correlation. In these cases, we are able to show that, under appropriate assumptions, the (possibly) scaled realised covariation can indeed be used to estimate correlation in a non-semimartingale setting.

While our results are interesting from a purely theoretical point of view, we also believe that they are relevant for various applications. For instance, empirical work in modelling of turbulence suggests that non-semimartingale models, and Brownian semistationary processes in particular, compare favourably to other alternatives, see e.g. Barndorff-Nielsen and Schmiegel (2009); Márquez and Schmiegel (2016); Barndorff-Nielsen, Benth, and Veraart (2017). Also, when it comes to modelling of financial asset prices, we remark that the classical theory has heavily relied on the semimartingale framework, but it has been shown that non-semimartingale models do not necessarily lead to the presence of arbitrage. More precisely, in the presence of small proportional transaction costs, Guasoni, Rásonyi, and Schachermayer (2008) showed that if a non-semimartingale has conditional full support, then it does not lead to so-called free lunches. This property is indeed satisfied by a Brownian semistationary processes, see Pakkanen (2011). Also, in the case of modelling energy prices the semimartingale assumption can be relaxed and (multivariate) Brownian semistationary processes have been used in this context, see e.g. Barndorff-Nielsen, Benth, and Veraart (2013); Veraart and Veraart (2014).

Our work has been motivated by a variety of recent articles which consider weak laws of large numbers and central limit theorems for Gaussian processes, see Guyon and León (1989); Barndorff-Nielsen, Corcuera, and Podolskij (2009); Corcuera (2012); Barndorff-Nielsen, Corcuera, Podolskij, and Woerner (2009), for Brownian semistationary processes, see Barndorff-Nielsen and Schmiegel (2009); Barndorff-Nielsen, Corcuera, and Podolskij (2011), and for Lévy-driven processes, see Basse-O'Connor, Lachieze-Rey, and Podolskij (2017); Basse-O'Connor, Heinrich, and Podolskij (2016). However, we remark that all these articles consider a univariate setting, whereas our article is to the best of our knowledge the first one to consider the multivariate case. This enables us to construct for the first time a consistent estimator for stochastic correlation outside the semimartingale framework.

The remainder of this article is structured as follows. Section 2 introduces the notation and defines the main objects of interest. In particular, it formulates the assumptions which ensure that we are outside the semimartingale framework. Section 3 contains the main contributions of the article by presenting three weak laws of large numbers for the (possibly scaled) realised covariation. Since the proofs of our results are rather technical, we conclude in Section 4 and relegate the proofs to Section 5. Finally, Section A contains some useful background material.

2   The setting

Throughout this article we denote by $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ a filtered, complete probability space and by $\mathcal{B}(\mathbb{R})$ the class of Borel subsets of $\mathbb{R}$. We will consider a finite time horizon $[0, T]$ for some $T > 0$. Let us first recall the definition of a Brownian measure.

Definition 2.1 (Brownian measure). *An $\mathscr{F}_t$-adapted Brownian measure $W : \Omega \times \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ is a Gaussian stochastic measure such that, if $A \in \mathcal{B}(\mathbb{R})$ with $\mathbb{E}[(W(A))^2] < \infty$, then $W(A) \sim N(0, Leb(A))$, where Leb is the Lebesgue measure. Moreover, if $A \subseteq [t, +\infty)$, then $W(A)$ is independent of $\mathscr{F}_t$.*

An introduction to constructing stochastic integrals against such measures can be found in Walsh (1986).

We will assume that $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ supports two independent $\mathscr{F}_t$-Brownian measures $W^{(1)}, \tilde{W}$ on $\mathbb{R}$ and will now define a bivariate moving average process with stochastic correlation.

Definition 2.2 (Bivariate moving average process with stochastic correlation). *Consider two independent Brownian measures $W^{(1)}$ and $\tilde{W}$ adapted to $\mathscr{F}_t$ and two nonnegative deterministic functions $g^{(1)}, g^{(2)} \in L^2((0, \infty))$ which are continuous on $\mathbb{R} \setminus \{0\}$. Let ρ be a càdlàg stochastic process, defined on the whole real line, with paths lying in $[-1, +1]$ a.s. and independent of $\mathscr{F}_t$. Define*

$$Y_t^{(1)} := \int_{-\infty}^t g^{(1)}(t-s)\, dW_s^{(1)},$$

$$Y_t^{(2)} := \int_{-\infty}^t g^{(2)}(t-s)\rho_s\, dW_s^{(1)} + \int_{-\infty}^t g^{(2)}(t-s)\sqrt{1-\rho_s^2}\, d\tilde{W}_s.$$

Then the vector process $(\mathbf{Y}_t)_{t \geq 0} = (Y_t^{(1)}, Y_t^{(2)})_{t \geq 0}^\top$ is called a bivariate moving average process with stochastic correlation.

We call $\mathbf{Y}$ defined as above a bivariate moving average process with stochastic correlation, since we will formally write $dW_t^{(2)} := \rho_s dW_t^{(1)} + \sqrt{1-\rho_t^2}\, d\tilde{W}_t$ and we can then write

$$Y_t^{(2)} = \int_{-\infty}^t g^{(2)}(t-s)\, dW_s^{(2)}.$$

Note that for $t > 0$, $W^{(2)}([0,t]) := \int_0^t \rho_s\, dW_s^{(1)} + \int_0^t \sqrt{1-\rho_s^2}\, d\tilde{W}_s$ is a standard Brownian motion by Lévy's characterisation theorem. Hence $Y^{(1)}$ and $Y^{(2)}$ appear as two univariate moving average processes which feature stochastic dependence.

In what follows we will mostly refer to the Brownian measures $W^{(1)}$ and $\tilde{W}$ as processes, by a slight abuse of notation. An *increment* $W_t - W_s$ will simply be the (Gaussian) random variable $W(\omega, (s, t])$.

2.1 Integrated correlation and realised covariation

Our object of interest is the integrated stochastic correlation coefficient given by $\int_0^t \rho_s ds$. It is well known that in the case when $\mathbf{Y}$ is a semimartingale, then the quadratic covariation of $\mathbf{Y}$ is given by

$$[Y^{(1)}, Y^{(2)}]_t = g^{(1)}(0^+)g^{(2)}(0^+)\int_0^t \rho_s ds. \tag{1}$$

Our aim is to estimate (the right hand side of) equation (1) consistently given high frequency observations. To this end consider the following setting. Suppose that we sample our processes discretely along successive partitions of $[0, T]$. A partition Π_n of $[0, T]$ will be a collection

of times $0 = t_0 < \cdots < t_i < t_{i+1} < \cdots < t_n = T$, where, for simplicity, we assume that the partition is equally spaced. The mesh of the partition will therefore be $\Delta_n = \frac{1}{n}$. Hence when $n \to \infty$, $\Delta_n \to 0$ and we are in the setting of the so-called *infill asymptotics*.

We denote the increments of $Y^{(j)}$ by $\Delta_i^n Y^{(j)} := Y_{i\Delta_n}^{(j)} - Y_{(i-1)\Delta_n}^{(j)}$, for $j = 1, 2$, and find that

$$
\begin{aligned}
\Delta_i^n Y^{(j)} = &\int_{-\infty}^{(i-1)\Delta_n} \left(g^{(j)}\left(i\Delta_n - s \right) - g^{(j)}\left((i-1)\Delta_n - s \right) \right) dW_s^{(j)} \\
&+ \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(j)}(i\Delta_n - s)\, dW_s^{(j)}.
\end{aligned}
\tag{2}
$$

The *realised covariation* is defined as

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

for $n \geq 1, t \in [0, T]$. In the case when $\mathbf{Y}$ is a semimartingale, we know from Protter (2005)(Theorem 23), that

$$
\sum_{i=1}^{\lfloor nt \rfloor} \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \overset{\text{u.c.p.}}{\to} [Y^{(1)}, Y^{(2)}]_t, \qquad \text{as } n \to \infty,
$$

where the convergence is uniform on compacts in probability (u.c.p.), see Section A.1 for details on u.c.p. convergence.

Since the asymptotic theory for realised covariation is well known in the semimartingale framework, we are exclusively interested in the asymptotic behaviour of the realised covariation when $\mathbf{Y}$ is *not* a semimartingale. In that case, we have no guarantee that, as n tends to infinity, the realised covariation tends to a finite limit. In the following we will prove convergence of the (possibly scaled) realised covariation and identify the limiting process.

2.2 (Non-) semimartingale conditions

The question of whether or not $Y^{(j)}$ for $j = 1, 2$ is a semimartingale hinges on the properties of the functions $g^{(j)}$. To simplify the exposition, let us suppress the superscripts in the following and focus on the process $Y_t = \int_{-\infty}^{t} g(t - s) dW_s$ in this subsection. In order to establish whether Y belongs to the semimartingale class, we need to be precise about the filtration we are using. There are three possible filtrations that can be considered: 1) The filtration $\left(\mathscr{F}_t^Y \right)_{t \geq 0}$, the natural filtration of Y, i.e. the smallest filtration to which Y is adapted; 2) the filtration $\left(\mathscr{F}_t^{Y,\infty} \right)_{t \geq 0}$, such that $\mathscr{F}_t^{Y,\infty} := \sigma\{Y_s, s \in (-\infty, t]\}$, i.e. the filtration generated by the history of Y; 3) the filtration $\left(\mathscr{F}_t^{W,\infty} \right)_{t \geq 0}$ which is the smallest filtration with respect to which W is an adapted Brownian measure.

A discussion on the conditions to impose to ensure that Y is a semimartingale in these different filtrations is contained in Basse (2008), Cheridito (2004). We will work with the condition that Y is a semimartingale in the $\mathscr{F}_t^{W,\infty}$ filtration. This is the most restrictive of the three (see Basse (2008)). For this setting, there exists a classical theorem due to Knight (1992) that states the following:

4

Theorem 2.3 (Knight). *The process $(Y_t)_{t \geq 0}$ is an $\mathscr{F}_t^{W,\infty}$-semimartingale if and only if there exist $h \in L^2(\mathbb{R})$ and $\alpha \in \mathbb{R}$ such that:*

$$g(t) = \alpha + \int_0^t h(s)\, ds.$$

Next let us study some sufficient conditions for a more general class of stochastic processes to be a semimartingale: Barndorff-Nielsen and Schmiegel (2009) extended the moving average process to allow for stochastic volatility and defined the so-called Brownian semistationary ($\mathcal{BSS}$) process by

$$X_t = \int_{-\infty}^t g(t - s)\, \sigma_s dW_s, \tag{3}$$

where W is an $\mathscr{F}_t$-adapted Brownian measure, σ is càdlàg and $\mathscr{F}_t$-adapted, $g \colon \mathbb{R} \to \mathbb{R}$ is a deterministic function, continuous in $\mathbb{R} \setminus \{0\}$, with $g(t) = 0$ if $t \leq 0$ and $g \in L^2((0, \infty))$. Also it is assumed that $\int_{-\infty}^t g^2(t - s)\sigma_s^2\, ds < \infty$ a.s. so that a.s. we have $Y_t < \infty$ for all $t \geq 0$. Barndorff-Nielsen and Schmiegel (2009) gave the following sufficient conditions for $\mathcal{BSS}$ process X to be a semimartingale:

Theorem 2.4. *Under the assumptions that*

(i) g is absolutely continuous and $g' \in L^2((0, \infty))$,

(ii) $\lim_{x \to 0^+} g(x) =: g(0^+) < \infty$,

(iii) The process $g'(-\cdot)\sigma_\cdot$ is square integrable,

then X_t defined as in (3) is an $\mathscr{F}_t^{W,\infty}$-semimartingale. In this case X_t admits the decomposition:

$$X_t = g(0^+)W_t + \int_0^t dl \left[\int_{-\infty}^l g'(l - s)\sigma_s\, dW_s \right].$$

We will now reintroduce the superscripts in our notation and formulate conditions which ensure that the bivariate moving average process $\mathbf{Y}$ with stochastic correlation is *not* a semimartingale. To this end, we will relax the first two assumptions in Theorem 2.4 since both assumptions are necessary for $Y^{(j)}$ to belong to the semimartingale class (see Basse (2008)) for $j = 1, 2$.

Assumption 2.1. *For $j \in \{1, 2\}$, we assume that $g^{(j)} \colon \mathbb{R} \to \mathbb{R}^+$ are nonnegative functions and continuous, except possibly at $x = 0$. Also, $g^{(j)}(x) = 0$ for $x < 0$ and $g^{(j)} \in L^2((0, +\infty))$. We further ask that $g^{(j)}$ be differentiable everywhere with derivative $\left(g^{(j)}\right)' \in L^2((b^{(j)}, \infty))$ for some $b^{(j)} > 0$ and $\left((g^{(j)})'\right)^2$ non-increasing in $[b^{(j)}, \infty)$.*

To simplify the notation in some of the proofs we set $b = \max\{b^{(1)}, b^{(2)}\}$, then $\left(g^{(j)}\right)' \in L^2((b, \infty))$ and $\left((g^{(j)})'\right)^2$ is non-increasing in $[b, \infty)$ for $j = 1, 2$.

Note that we are not assuming that $\left(g^{(j)}\right)' \in L^2((0, \infty))$ in order to exclude the semimartingale case. In particular, we must have that, for all $\varepsilon > 0$, $\sup_{x \in (0, \varepsilon)} \left|\left(g^{(j)}\right)'(x)\right| = \infty$.

Remark 2.5. *Under Assumption 2.1 we can deduce from the mean value theorem that there exists an $\eta \in [s, s + \Delta_n]$ such that*

$$g^{(j)}(s + \Delta_n) - g^{(j)}(s) = \Delta_n g^{(j)\,\prime}(\eta).$$

Then Assumption 2.1 implies that, for $s \in [b, \infty)$, we have

$$|g^{(j)}(s + \Delta_n) - g^{(j)}(s)| \leq \Delta_n |g^{(j)\,\prime}(s)|.$$

Hence, we can derive the bound:

$$\int_b^\infty \left(g^{(j)}(s + \Delta_n) - g^{(j)}(s) \right)^2 ds \leq \Delta_n^2 \int_b^\infty \left(g^{(j)\,\prime}(s) \right)^2 ds = K\Delta_n^2,$$

for some constant $K > 0$, since $\left(g^{(j)} \right)' \in L^2((b, \infty))$.

We will impose an additional assumption for which we recall here the definition of a slowly varying function.

Definition 2.6 (Slowly and regularly varying function)**.** *A measurable function $L\colon (0, \infty) \to (0, \infty)$ is called* slowly varying at infinity *if, for all $\lambda > 0$ we have that $\lim_{x \to \infty} \frac{L(\lambda x)}{L(x)} = 1$. A function $g\colon (0, \infty) \to (0, \infty)$ is called* regularly varying at infinity *if, for x large enough, it can be written as: $g(x) = x^\delta L(x)$, for a slowly varying function L. The parameter δ is called the* index of regular variation*. Finally, a measurable function $L\colon (0, \infty) \to (0, \infty)$ is called* slowly varying at zero *(resp. regularly varying at zero) if $x \to L\left(\frac{1}{x}\right)$ is slowly varying (resp. regularly varying) at infinity.*

The theory of regular variation allows us to construct functions which behave asymptotically like power functions. See also the discussion in Bennedsen, Lunde, and Pakkanen (2017), while the standard comprehensive reference on the subject is the book by Bingham, Goldie, and Teugels (1989).

Assumption 2.2. *We require that, for $j \in \{1, 2\}$, the following integral functions are regularly varying at zero:*

$$\int_0^x \left(g^{(j)}(s) \right)^2 ds = x^{2\delta^{(j)}+1} L_1^{(j)}(x), \tag{4}$$

$$\int_0^{b^{(j)}} \left(g^{(j)}(s + x) - g^{(j)}(s) \right)^2 ds = x^{2\delta^{(j)}+1} L_2^{(j)}(x), \tag{5}$$

for $\delta^{(j)} \in (-\frac{1}{2}, 0) \cup (0, \frac{1}{2})$, where $L_1^{(j)}(x)$ and $L_2^{(j)}(x)$ are positive, slowly varying functions at zero which are continuous on $(0, \infty)$.

In this situation, the restriction $\delta^{(j)} \in (-\frac{1}{2}, 0) \cup (0, \frac{1}{2})$ ensures that the process leaves the semimartingale class.

Example 2.7. *Suppose for simplicity $g(x) = x^\delta$ (where we suppress the superscripts again). Then we can show that, if $\delta \in (-\frac{1}{2}, 0) \cup (0, \frac{1}{2})$, then*

$$\int_0^b (g(x + \Delta_n) - g(x))^2 dx = O(\Delta_n^{2\delta+1}).$$

With the change of variable $x = y\Delta_n$ we get:

$$\Delta_n \int_0^{bn} \left(g\left(\Delta_n(y+1) \right) - g\left(\Delta_n y \right) \right)^2 dy = \Delta_n^{2\delta+1} \int_0^{bn} \left((y+1)^\delta - y^\delta \right)^2 dy$$

$$\leq \Delta_n^{2\delta+1} \int_0^\infty \left((y+1)^\delta - y^\delta \right)^2 dy.$$

It is then sufficient to show that $\int_0^\infty \left((y+1)^\delta - y^\delta \right)^2 dy < \infty$. First, suppose $\delta > 0$. Then, in a neighbourhood of 0, $(y+1)^\delta - y^\delta \sim 1$, while by Taylor's theorem, $\left((y+1)^\delta - y^\delta \right)^2 \sim y^{2\delta-2}$ away from 0. Hence, if $2\delta - 2 < -1 \iff \delta < \frac{1}{2}$ the function is integrable at infinity. If instead $\delta < 0$, changing variables $z = \frac{1}{y}$ and writing $-\delta =: \beta > 0$, we obtain:

$$\int_0^\infty \left(\left(\frac{z+1}{z} \right)^\delta - \left(\frac{1}{z} \right)^\delta \right)^2 \frac{1}{z^2} dz = \int_0^\infty \left(\frac{z^{\beta-1} \left(1 - (z+1)^\beta \right)}{(z+1)^\beta} \right)^2 dz.$$

For z close to zero, $(z+1)^\beta \sim 1 + z\beta$, so around 0 the integrand is asymptotically equivalent to $z^\beta \sim 0$. At infinity, $\left(\frac{z^{\beta-1}\left(1-(z+1)^\beta\right)}{(z+1)^\beta} \right)^2 \sim \left(\frac{z^{\beta-1}\left(-(z+1)^\beta\right)}{(z+1)^\beta} \right)^2 = z^{2\beta-2}$, so we must impose $2\beta - 2 < -1 \iff \beta < \frac{1}{2} \iff \delta > -\frac{1}{2}$.

Example 2.8. *A popular example in applications is the so-called Gamma kernel which is given by $g(x) = e^{-\lambda x} x^\delta 1_{\{x>0\}}$, for $\lambda > 0, \delta > -\frac{1}{2}$, where superscripts are suppressed. A review paper on the importance of the Gamma kernel in this context is given in Barndorff-Nielsen (2016). By way of example, the Gamma kernel is very relevant in turbulence modelling, indeed $g(x) \sim x^\delta$, close to zero, and δ is called the scaling parameter. A scaling parameter equal to $-\frac{1}{6}$ fits well with Kolmogorov's scaling law in turbulence (see Corcuera, Hedevang, Pakkanen, and Podolskij (2013)).*

Note that with $\delta \in (-\frac{1}{2}, 0) \cup (0, \frac{1}{2}]$, g satisfies Assumptions 2.1 and 2.2 (a proof of the second statement is contained in Barndorff-Nielsen, Corcuera, and Podolskij (2011)). Indeed, for δ in the stated range, $g' \notin L^2(0, \infty)$, since it is unbounded and not square integrable near zero; in this case the moving average process (and also the Brownian semistationary process) is not a semimartingale. If $\delta > 0$, let b be the inflection point of g and $\bar{y}$ its point of maximum. Then $b > \bar{y}$, and, if $y > b$, we have that $((g')^2)'(y) = 2g'(y)g''(y) \leq 0$, since the Gamma kernel decreases after the maximum and the second derivative tends to $0+$ for $y \to \infty$. If $\delta \leq 0$, the condition that $(g')^2$ is non increasing in $[b, \infty)$ is satisfied for any $b > 0$. Finally, g' is obviously square integrable over $[b, \infty)$ thanks to its exponential decay.

The next assumption we assume throughout concerns the stochastic correlation process ρ.

Assumption 2.3. *The paths of ρ are almost surely Hölder continuous with exponent α.*

Example 2.9. *If ρ can be written as a diffusion: $\rho_t = \int_0^t \kappa_s \, dW_s^*$ with W^* a Brownian measure independent of $W^{(1)}, \tilde{W}$, and κ satisfying some mild conditions, then ρ admits a version which is $\alpha-$Hölder continuous, for all $\alpha < \frac{1}{2}$. We can for example take the solution to the SDE:*

$$d\rho_t = \int_0^t \sqrt{(1-\rho_t)(1+\rho_t)} \, dW_t,$$

which defines a stochastic process bound to stay between -1 and +1 (with the addition of an appropriate drift, such a process takes the name of Jacobi process, see e.g. Veraart and Veraart (2012)).

3 Results

We will now state our main results where we distinguish two different scenarios: The first scenario considers the case when the functions $g^{(j)}$ have a limit at zero, but the resulting process $\mathbf{Y}$ falls outside the semimartingale class. In the second scenario we will instead allow the functions $g^{(j)}$ to blow up near zero, but we ask that they are decreasing.

3.1 Weak law of large numbers for bivariate moving average processes with stochastic correlation

3.1.1 Scenario 1: The non-semimartingale case when $g^{(1)}(0^+)g^{(2)}(0^+) < \infty$

This subsection presents the first of our convergence results for the realised covariation. Here we also impose an assumption on the limit behaviour of the functions $g^{(j)}$ at 0.

Assumption 3.1. *We suppose that $\delta^{(1)} + \delta^{(2)} \geq 0$ and that the limit $g^{(1)}(0^+)g^{(2)}(0^+) := \lim_{x \to 0^+} g^{(1)}(x)g^{(2)}(x)$, exists finite. If $\delta^{(1)} + \delta^{(2)} = 0$, then we require that $\lim_{x \to 0^+} L_2^{(1)}(x)L_2^{(2)}(x) = 0$.*

Remark 3.1. *If we further assumed that:*

$$g^{(j)}(x) = x^{\delta^{(j)}} L_3^{(j)}(x), \tag{6}$$

where L_3 is slowly varying, then Assumption 3.1 would tell us that $\lim_{x \to 0^+} g^{(1)}(x)g^{(2)}(x) = 0$. This is a common practical situation, where both the kernels g can be taken to be $x^{\delta^{(j)}} e^{-\lambda^{(j)} x}$, for some positive, potentially different $\delta^{(1)}, \delta^{(2)}$. Note, however, that (4) is more general. Indeed, taking derivatives of (4), we get:

$$\left(g^{(j)}(x)\right)^2 = (2\delta^{(j)} + 1)x^{2\delta^{(j)}} L_1^{(j)}(x) + x^{2\delta^{(j)}+1} \left(L_1^{(j)}(x)\right)'.$$

The derivative of a slowly varying function need not be slowly varying (see De Haan and Resnick (1979)), hence we are not necessarily in the same context of (6).

We can now formulate our first law of large numbers for the realised covariation of two non-semimartingales.

Theorem 3.2 (First law of large numbers). *Suppose that the Assumptions 2.1, 2.2, 2.3 and 3.1 are satisfied. Then, the following* u.c.p. *convergence holds:*

$$\left(\sum_{i=1}^{\lfloor nt \rfloor} \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right)_{t \in [0,T]} \xrightarrow{u.c.p.} \left(g^{(1)}(0^+)g^{(2)}(0^+) \int_0^t \rho_s \, ds\right)_{t \in [0,T]}, \quad as \ n \to \infty. \tag{7}$$

3.1.2 Scenario 2: The non-semimartingale with decreasing functions $g^{(1)}, g^{(2)}$

In the case when the limit $\lim_{x \to 0^+} g^{(1)}(x)g^{(2)}(x)$ does not exist finite we need to proceed differently in order to derive a weak law of large numbers for the realised covariation. A notable example of this scenario is the case when $g^{(j)}$ are Gamma kernels given by $g^{(j)}(x) = x^{\delta^{(j)}} e^{-\lambda^{(j)} x}$, for $\delta^{(j)} < 0, \lambda^{(j)} > 0$ for $j = 1, 2$.

The method we are going to propose in the following is motivated by the method used in the univariate case by Barndorff-Nielsen and Schmiegel (2009). To this end, let us introduce the following notation. We write:

$$\Delta_i^n Y^{(1)} = \int_{-\infty}^{i\Delta_n} \varphi_{\Delta_n}^{(1)}(i\Delta_n - s)\, dW_s^{(1)},$$

where

$$\varphi_{\Delta_n}^{(j)}(s) = \begin{cases} g^{(j)}(s), & s \leq \Delta_n, \\ g^{(j)}(s) - g^{(j)}(s - \Delta_n), & s > \Delta_n, \end{cases}$$

and hence

$$\varphi_{\Delta_n}^{(j)}(i\Delta_n - s) = \begin{cases} g^{(j)}(i\Delta_n - s), & s \geq (i-1)\Delta_n, \\ g^{(j)}(i\Delta_n - s) - g^{(j)}((i-1)\Delta_n - s), & s < (i-1)\Delta_n. \end{cases}$$

For reasons that will become apparent later, we need to impose monotonicity of our kernel functions. Since we want to be able to cover the case where g goes to infinity close to zero, we introduce the following assumption:

Assumption 3.2. *We assume that $g^{(1)}, g^{(2)}$ are both decreasing on $\mathbb{R}^+$.*

We need to formulate another rather technical assumption which has first been formulated in Granelli and Veraart (2017). To this end let us introduce the necessary notation: We define the bivariate Gaussian process $(\mathbf{G}_t)_{t\geq 0}$ whose components are given by $G_t^{(1)} = Y_t^{(1)} = \int_{-\infty}^t g^{(1)}(t-s)dW_s^{(1)}$ and $G_t^{(2)} = \int_{-\infty}^t g^{(2)}(t-s)dW_s^{(1)}$.

Remark 3.3. *As a consequence of Assumption 2.2, if we denote for $i \in \{1,2\}$:*

$$\bar{R}^{(i)}(t) := \mathbb{E}\left[\left(G_{s+t}^{(i)} - G_s^{(i)}\right)^2\right],$$

then it holds that:

$$\bar{R}^{(i)}(t) = t^{2\delta^{(i)}+1} L_0^{(i)}(t),$$

for $\delta^{(i)} \in \left(-\frac{1}{2}, \frac{1}{2}\right) \setminus \{0\}$, where $L_0^{(i)}(t)$ is a slowly varying function at zero. Let us denote by $\tilde{L}_0^{(i,j)}(t) := \sqrt{L_0^{(i)}(t)L_0^{(j)}(t)}$ another slowly varying function at zero.

For $i, j \in \{1, 2\}$, we write $\rho_{i,j} = \rho$ for $i \neq j$ and $\rho_{i,j} = 1$ for $i = j$. Moreover, we introduce the functions mapping $\mathbb{R}^+$ into $\mathbb{R}^+$, with $i, j \in \{1, 2\}$:

$$\bar{R}^{(i,j)}(t) := \mathbb{E}\left[\left(G_t^{(j)} - G_0^{(i)}\right)^2\right] = \left\|g^{(i)}\right\|_{L^2}^2 + \left\|g^{(j)}\right\|_{L^2}^2 - 2\mathbb{E}\left[G_0^{(i)}G_t^{(j)}\right]. \tag{8}$$

We note that we can write

$$\bar{R}^{(i,j)}(t) = C_{i,j} + 2\rho_{i,j} \int_0^\infty (g^{(j)}(x) - g^{(j)}(x+t))g^{(i)}(x)dx,$$

where $C_{i,j} := \left\|g^{(i)}\right\|_{L^2}^2 + \left\|g^{(j)}\right\|_{L^2}^2 - 2\rho_{i,j}\int_0^\infty g^{(i)}(x)g^{(j)}(x)dx$, where in particular $C_{i,i} = 0$. We can now formulate our next assumption.

9

Assumption 3.3. *For all $t \in (0, T)$, there exist slowly varying functions $L_0^{(i,j)}(t)$ and $L_2^{(i,j)}(t)$ which are continuous on $(0, \infty)$ such that*

$$\bar{R}^{(i,j)}(t) = C_{i,j} + \rho_{i,j} t^{\delta^{(i)} + \delta^{(j)} + 1} L_0^{(i,j)}(t), \quad \text{for } i, j \in \{1, 2\}, \tag{9}$$

and

$$\frac{1}{2}(\bar{R}^{(i,j)})''(t) = \rho_{i,j} t^{\delta^{(i)} + \delta^{(j)} - 1} L_2^{(i,j)}(t), \quad \text{for } i, j \in \{1, 2\},$$

where $\delta^{(1)}, \delta^{(2)} \in \left(-\frac{1}{2}, \frac{1}{2}\right) \setminus \{0\}$.

Also, if we denote $\tilde{L}_0^{(i,j)}(t) := \sqrt{L_0^{(i,i)}(t) L_0^{(j,j)}(t)}$, we ask that the functions $L_0^{(i,j)}(t)$ and $L_2^{(i,j)}(t)$ are such that, for all $\lambda > 0$, there exists a $H^{(i,j)} \in \mathbb{R}$ such that:

$$\lim_{t \to 0+} \frac{L_0^{(i,j)}(\lambda t)}{\tilde{L}_0^{(i,j)}(t)} = H^{(i,j)} < \infty, \tag{10}$$

and that there exists $b \in (0, 1)$, such that:

$$\limsup_{x \to 0^+} \sup_{y \in (x, x^b)} \left| \frac{L_2^{(i,j)}(y)}{\tilde{L}_0^{(i,j)}(x)} \right| < \infty. \tag{11}$$

Remark 3.4. *We remark that Assumption 3.3 is important for the setting when $i \neq j$, in the case when $i = j$ the assumption is implied by the previous assumptions as discussed before.*

Remark 3.5. *Assumption 3.3 can be seen as the extension of the classical assumptions on the variogram of a stationary process, see e.g. Corcuera, Hedevang, Pakkanen, and Podolskij (2013); Barndorff-Nielsen, Corcuera, and Podolskij (2013); Guyon and León (1989).*

We need to introduce the following scaling factor for the realised covariation: The scaling factor $c(x)$ is defined as

$$c(x) := \int_0^\infty \varphi_x^{(1)}(s) \varphi_x^{(2)}(s) \, ds$$
$$= \int_0^x g^{(1)}(s) g^{(2)}(s) \, ds + \int_0^\infty \left(g^{(1)}(s + x) - g^{(1)}(s) \right) \left(g^{(2)}(s + x) - g^{(2)}(s) \right) \, ds.$$

Assumption 3.4. *The scaling factor satisfies*

$$c(\Delta_n) = \Delta_n^{\delta^{(1)} + \delta^{(2)} + 1} L_4^{(1,2)}(\Delta_n), \tag{12}$$

where $L_4^{(1,2)}$ is a continuous function on $(0, \infty)$ which is slowly varying at zero and $\delta^{(1)}, \delta^{(2)} \in \left(-\frac{1}{2}, \frac{1}{2}\right) \setminus \{0\}$. Moreover, we assume that there exists a constant $|H| < \infty$ such that

$$\lim_{x \to 0+} \frac{L_4^{(1,2)}(x)}{\tilde{L}_0^{(1,2)}(x)} = H. \tag{13}$$

Note that since the kernel functions are positive, decreasing and differentiable, $c(x)$ is a positive, increasing function of x.

We can now formulate the second weak law of large numbers for the scaled realised covariation of two non-semimartingales.

Theorem 3.6. *Suppose that Assumptions 2.1, 2.2, 2.3, 3.2, 3.3 and 3.4 hold. Then:*

$$\left(\Delta_n \frac{\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}}{c(\Delta_n)} \right)_{t \in [0,T]} \overset{u.c.p.}{\longrightarrow} \left(\int_0^t \rho_l \, dl \right)_{t \in [0,T]}, \quad as \ n \to \infty.$$

We note that this result differs from the result obtained in Theorem 3.2 by the fact that we need to include the additional scaling factor $\frac{\Delta_n}{c(\Delta_n)} \sim \Delta_n^{-\delta^{(1)} - \delta^{(2)}}$. While this might seem innocent at first sight, it does impose restriction on empirical applications, since the scaling factor depends on the typically unknown functions $g^{(1)}$ and $g^{(2)}$. Note further, that for n large enough, the scaling factor $\frac{\Delta_n}{c(\Delta_n)}$ is increasing with n.

3.2 Weak law of large number for bivariate Brownian semistationary processes with stochastic correlation

So far we have considered bivariate moving average processes with stochastic correlation. In a next step, we wish to consider a more general class of stochastic processes which also allows for stochastic volatility. To this end, we give the definition of the bivariate $\mathcal{BSS}$ process with stochastic correlation that generalises Definition 2.2.

Definition 3.7 (Bivariate Brownian semistationary process with stochastic correlation)**.** *Consider two independent Brownian measures $W^{(1)}$ and $\tilde{W}$ adapted to $\mathscr{F}_t$ and two nonnegative deterministic functions $g^{(1)}, g^{(2)} \in L^2((0, \infty))$ which are continuous on $\mathbb{R} \setminus \{0\}$. Let ρ be a càdlàg stochastic process, defined on the whole real line, with paths lying in $[-1, +1]$ a.s. and independent of $\mathscr{F}_t$. Let further $\sigma^{(1)}, \sigma^{(2)}$ be càdlàg , $\mathscr{F}_t$-adapted stochastic processes and assume that for $i \in \{1, 2\}$, and for all $t \in [0, T]$:*

$$\int_{-\infty}^t g^{(i)2}(t - s)\sigma_s^{(i)2} \, ds < \infty.$$

Define

$$X_t^{(1)} := \int_{-\infty}^t g^{(1)}(t - s)\sigma_s^{(1)} \, dW_s^{(1)},$$

$$X_t^{(2)} := \int_{-\infty}^t g^{(2)}(t - s)\sigma_s^{(2)}\rho_s \, dW_s^{(1)} + \int_{-\infty}^t g^{(2)}(t - s)\sigma_s^{(2)}\sqrt{1 - \rho_s^2} \, d\tilde{W}_s.$$

Then the vector process $(\mathbf{X}_t)_{t \geq 0} = (X_t^{(1)}, X_t^{(2)})_{t \geq 0}^\top$ is called a bivariate Brownian semistationary process with stochastic correlation.

We note that we can write

$$\mathbf{X}_t = \int_{-\infty}^t \begin{pmatrix} g^{(1)}(t-s) & 0 \\ 0 & g^{(2)}(t-s) \end{pmatrix} \begin{pmatrix} \sigma_s^{(1)} & 0 \\ 0 & \sigma_s^{(2)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \rho_s & \sqrt{1-\rho_s^2} \end{pmatrix} d \begin{pmatrix} W_s^{(1)} \\ \tilde{W}_s \end{pmatrix},$$

$$= \int_{-\infty}^t \begin{pmatrix} g^{(1)}(t-s) & 0 \\ 0 & g^{(2)}(t-s) \end{pmatrix} \begin{pmatrix} \sigma_s^{(1)} & 0 \\ 0 & \sigma_s^{(2)} \end{pmatrix} d \begin{pmatrix} W_s^{(1)} \\ W_s^{(2)} \end{pmatrix},$$

where the integration is to be understood componentwise. For the increment processes, we obtain for $j \in \{1, 2\}$:

$$\Delta_i^n X^{(j)} = \int_{-\infty}^{(i-1)\Delta_n} \left(g^{(j)}(i\Delta_n - s) - g^{(j)}(i-1)\Delta_n - s) \right) \sigma_s^{(j)} \, dW_s^{(j)}$$
$$+ \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(j)}(i\Delta_n - s)\sigma_s^{(j)} \, dW_s^{(j)}. \quad (14)$$

Lastly, we will need the following technical assumption:

Assumption 3.5. *The volatility processes are assumed to be independent of the Brownian measures, i.e.:* $\sigma\left(\sigma^{(1)}, \sigma^{(2)}\right)$ *is independent of* $\sigma\left(W^{(1)}, \tilde{W}\right)$. *We assume that for* $j \in \{1, 2\}$, *we have that the following property holds almost surely:*

$$\int_1^\infty \left(\frac{d}{ds} g^{(j)}(s) \right)^2 \sigma_{y-s}^{(j)} \, ds < \infty, \qquad \text{for all } y \in \mathbb{R}^+.$$

We can write:

$$\sum_{i=1}^n \Delta_i^n X^{(1)} \Delta_i^n X^{(2)} = \sum_{i=1}^n \int_0^\infty \varphi_{\Delta_n}^{(1)}(s)\sigma_{i\Delta_n-s}^{(1)} \, dW_{i\Delta_n-s}^{(1)} \int_0^\infty \varphi_{\Delta_n}^{(2)}(s)\sigma_{i\Delta_n-s}^{(2)} \, dW_{i\Delta_n-s}^{(2)}.$$

We now consider the sigma algebra $\mathscr{H} := \mathscr{F}^{\rho,\sigma^{(1)},\sigma^{(2)}}$ generated by the processes $\rho, \sigma^{(1)}, \sigma^{(2)}$. We obtain:

$$\mathbb{E}\left[\Delta_i^n X^{(1)} \Delta_i^n X^{(2)} \middle| \mathscr{H} \right] = \int_0^\infty \varphi_{\Delta_n}^{(1)} \varphi_{\Delta_n}^{(2)} \sigma_{i\Delta_n-s}^{(1)} \sigma_{i\Delta_n-s}^{(2)} \rho_{i\Delta_n-s} \, ds$$
$$= \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g^{(1)} \Delta_i^n g^{(2)} \sigma_s^{(1)} \sigma_s^{(2)} \rho_s \, ds + \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n-s)g^{(2)}(i\Delta_n-s)\sigma_s^{(1)}\sigma_s^{(2)}\rho_s \, ds.$$
$$(15)$$

We aim to show the corresponding version of Theorem 3.6, in presence of volatility.

Theorem 3.8 (Law of large numbers). *Suppose that the assumptions of Theorem 3.6 hold. Further assume that the processes* $\sigma^{(1)}, \sigma^{(2)}$ *have Hölder continuous sample paths and satisfy Assumption 3.5. Then*

$$\left(\Delta_n \frac{\sum_{i=1}^{\lfloor nt \rfloor} \Delta_i^n X^{(1)} \Delta_i^n X^{(2)}}{c(\Delta_n)} \right)_{t \in [0,T]} \xrightarrow{u.c.p.} \left(\int_0^t \sigma_l^{(1)} \sigma_l^{(2)} \rho_l \, dl \right)_{t \in [0,T]}, \qquad \text{as } n \to \infty,$$

where $c(\Delta_n) = \int_0^{\Delta_n} g^{(1)}(s)g^{(2)}(s) \, ds + \int_0^\infty \left(g^{(1)}(s+\Delta_n) - g^{(1)}(s) \right) \left(g^{(2)}(s+\Delta_n) - g^{(2)}(s) \right) \, ds.$

4 Conclusion

In this article we have proven two versions of a law of large numbers for the realised covariation of the bivariate moving average process with stochastic correlation. We have also proven a weak law of large numbers for a bivariate Brownian semistationary process with stochastic correlation. These results state the consistency of our estimator for the integrated covariance, the (possibly scaled) realised covariation. The importance of such results can be seen in the more general context of building a fully multivariate theory for general $\mathcal{BSS}$ process,

whose importance in stochastic modelling in e.g. finance and turbulence, see Barndorff-Nielsen, Benth, and Veraart (2013); Corcuera, Hedevang, Pakkanen, and Podolskij (2013), is steadily rising. Performing inference on the dependence of two such processes is of fundamental importance in practical applications. It has to be stressed, though, that our second law of large numbers requires us to scale the estimator by the typically unknown kernel functions $g^{(1)}$ and $g^{(2)}$, and this makes the estimation unfeasible in practice. Further research could be devoted to estimating the kernel functions first and use a plug-in type estimator as the scaling factor of the realised covariation we have considered in this article.

On the stochastic analysis side, our results show that it is possible to obtain convergence of the "realised covariation" between two non-semimartingales, subject to appropriate scaling. Our techniques are very general and one could perhaps try to adapt them to other classes of multivariate fractional process, for example extending the theory to include the integral processes appearing in Corcuera, Nualart, Woerner, et al. (2006).

In future research one could also tackle the problem of deriving a central limit theorem for bivariate Brownian semistationary processes and first results along these lines are available in Granelli and Veraart (2017).

5 Proofs

5.1 Proof of Theorem 3.2

5.1.1 Notation and preliminary remarks

This section contains the full proof of Theorem 3.2, which will be divided into several separate lemmas.

Recall that we are working on a finite, fixed time horizon $[0, T]$, but according to Theorem A.3 and Remark 5.1 below, we do not lose generality if, for simplicity, we restrict ourselves to $T = 1$. Hence we do so in the following.

Remark 5.1 (From convergence in probability to *u.c.p.* convergence)**.** *We only need to show pointwise convergence in probability of the realised covariation in order to prove Theorem 3.2. Indeed, even though we are not directly in the framework of Theorem A.3, because the realised quadratic covariation process does not have increasing paths, since the product of the increments can be negative, we can still deduce u.c.p. convergence via the polarisation identity:*

$$\sum_{i=1}^{n} \left(\Delta_i^n Y \Delta_i^n X \right) = \frac{1}{2} \sum_{i=1}^{n} \left(\Delta_i^n (X + Y) \right)^2 - \sum_{i=1}^{n} \left(\Delta_i^n X \right)^2 - \sum_{i=1}^{n} \left(\Delta_i^n Y \right)^2.$$

If we can prove that the left hand side converges in probability for all t, it will follow that the realised quadratic variation of the sum $X + Y$ converges for all t, because we already know from the results in Barndorff-Nielsen and Schmiegel (2009) that convergence of the squared increments of X and Y separately holds for all t. But then each term on the right hand side converges u.c.p. because they all have increasing paths. Since u.c.p. convergence is induced by a metric, we can conclude that u.c.p. convergence will hold for the realised quadratic covariation too.

In the following, the notation: $\mathbb{E}_{(i-1)\Delta_n}$ denotes the conditional expectation operator upon the sub-σ-algebra $\mathscr{G}_{(i-1)\Delta_n}$ generated by the whole process ρ and by the increments of the

processes $W^{(1)}$ and $\tilde{W}$ until $(i-1)\Delta_n$, i.e.:

$$\mathscr{G}_{(i-1)\Delta_n} = \sigma\{\rho_s; s \in \mathbb{R}\} \vee \sigma\{W_u^{(1)} - W_t^{(1)}, \tilde{W}_u - \tilde{W}_t; \quad -\infty < t \le u \le (i-1)\Delta_n\}.$$

The main idea to prove (7) is to add and subtract the quantity:

$$\sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right] = \mathbb{E}_{(i-1)\Delta_n}\left[\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right], \tag{16}$$

and split the difference $\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} - g^{(1)}(0^+)g^{(2)}(0^+)\int_0^1 \rho_s\,ds$ into:

$$A_n := \sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} - \sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right], \tag{17}$$

$$B_n := \sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right] - g^{(1)}(0^+)g^{(2)}(0^+)\int_0^1 \rho_s\,ds. \tag{18}$$

The proof that A_n converges to zero will take most of the present section. Proposition 5.11 contains the much shorter proof that B_n converges to zero as well.

Using the notation $\Delta_i^n g_s^{(j)} := g^{(j)}(i\Delta_n - s) - g^{(j)}((i-1)\Delta_n - s)$, the quantity in (16) equals:

$$\mathbb{E}_{(i-1)\Delta_n}\left[\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right]$$

$$= \sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\,dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\,dW_s^{(2)}\right] \tag{19}$$

$$+ \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)}\,dW_s^{(1)} \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(2)}\,dW_s^{(2)}.$$

We have to establish a preliminary lemma:

Lemma 5.2. *The following holds:*

$$\mathbb{E}_{(i-1)\Delta_n}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g(i\Delta_n - s)\,dW_s\right] = 0,$$

$$\mathbb{E}_{(i-1)\Delta_n}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\,dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\,dW_s^{(2)}\right] \tag{20}$$

$$= \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\,\rho_s\,ds.$$

Proof. The first result is immediate. For the second one, using the definition of $W^{(2)}$:

$$\mathbb{E}_{(i-1)\Delta_n}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\,dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\,dW_s^{(2)}\right]$$

14

$$=\mathbb{E}_{(i-1)\Delta_n}\left[\int_{(i-1)\Delta_n}^{i\Delta_n}g^{(1)}(i\Delta_n-s)\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n}g^{(2)}(i\Delta_n-s)\,\rho_s dW_s^{(1)}\right] \tag{21}$$

$$+\mathbb{E}_{(i-1)\Delta_n}\left[\int_{(i-1)\Delta_n}^{i\Delta_n}g^{(1)}(i\Delta_n-s)\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n}g^{(2)}(i\Delta_n-s)\,\sqrt{1-\rho_s^2}d\tilde{W}_s\right]. \tag{22}$$

Line (22) immediately evaluates to zero by the tower property and independence of $W^{(1)},\tilde{W}$.
Observe that

$$\sigma\left\{W_u^{(1)}-W_t^{(1)},\tilde{W}_u-\tilde{W}_t;\quad-\infty<t\leq u\leq(i-1)\Delta_n\right\}$$

is independent of

$$\sigma\left(\int_{(i-1)\Delta_n}^{i\Delta_n}g^{(1)}(i\Delta_n-s)dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n}g^{(2)}(i\Delta_n-s)\rho_s dW_s^{(1)}\right)\vee\sigma\left(\rho_s,s\in\mathbb{R}\right)=:\mathscr{F}^\rho.$$

Hence, conditioning on $\mathscr{G}_{(i-1)\Delta_n}$ is equivalent to conditioning on $\mathscr{F}^\rho$ (Williams (1991)).

By taking the derivative of the conditional characteristic function, and exchanging differentiation and expectation thanks to uniform boundedness, we can write:

$$-\frac{\partial^2}{\partial\vartheta_1\partial\vartheta_2}\bigg|_{\vartheta_1=0,\vartheta_2=0}\mathbb{E}\left[\exp\left\{i\vartheta_1\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(1)}(k\Delta_n-s)\,dW_s^{(1)}\right.\right.$$
$$\left.\left.+i\vartheta_2\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(2)}(k\Delta_n-s)\rho_s\,dW_s^{(1)}\right\}\bigg|\mathscr{F}^\rho\right]$$

$$=-\mathbb{E}\left[\frac{\partial^2}{\partial\vartheta_1\partial\vartheta_2}\bigg|_{\vartheta_1=0,\vartheta_2=0}\exp\left\{i\vartheta_1\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(1)}(k\Delta_n-s)\,dW_s^{(1)}\right.\right.$$
$$\left.\left.+i\vartheta_2\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(2)}(k\Delta_n-s)\rho_s\,dW_s^{(1)}\right\}\bigg|\mathscr{F}^\rho\right]$$

$$=\mathbb{E}\left[\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(1)}(k\Delta_n-s)\,dW_s^{(1)}\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(2)}(k\Delta_n-s)\rho_s\,dW_s^{(1)}\bigg|\mathscr{F}^\rho\right],$$

but on the other hand:

$$-\frac{\partial^2}{\partial\vartheta_1\partial\vartheta_2}\bigg|_{\vartheta_1=0,\vartheta_2=0}\mathbb{E}\left[\exp\left\{i\vartheta_1\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(1)}(k\Delta_n-s)\,dW_s^{(1)}\right.\right.$$
$$\left.\left.+i\vartheta_2\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(2)}(k\Delta_n-s)\rho_s\,dW_s^{(1)}\right\}\bigg|\mathscr{F}^\rho\right]$$

$$=-\frac{\partial^2}{\partial\vartheta_1\partial\vartheta_2}\bigg|_{\vartheta_1=0,\vartheta_2=0}\mathbb{E}\left[\exp\left\{i\int_{(k-1)\Delta_n}^{k\Delta_n}\left[\vartheta_1 g^{(1)}(k\Delta_n-s)+\vartheta_2\rho_s g^{(2)}(k\Delta_n-s)\right]dW_s^{(1)}\right\}\bigg|\mathscr{F}^\rho\right]$$

$$=-\frac{\partial^2}{\partial\vartheta_1\partial\vartheta_2}\bigg|_{\vartheta_1=0,\vartheta_2=0}\exp\left(-\frac{1}{2}\int_{(k-1)\Delta_n}^{k\Delta_n}\left[\vartheta_1 g^{(1)}(k\Delta_n-s)+\vartheta_2\rho_s g^{(2)}(k\Delta_n-s)\right]^2 ds\right)$$

$$=\int_{(k-1)\Delta_n}^{k\Delta_n}g^{(1)}(k\Delta_n-s)g^{(2)}(k\Delta_n-s)\rho_s\,ds,$$

15

proving the claim. □

We are now ready to begin to look at the proof of Theorem 3.2. The next two sections contain the full details on the convergence of A_n as defined in (17).

5.1.2 Convergence of A_n

First we show the convergence of A_n. To this end recall that

$$A_n = \sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} - \sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n} \left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \right].$$

Expanding the conditional expectation using the computations in Lemma 5.2 and simple algebra gives:

$$\sum_{i=1}^n \left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} - \mathbb{E}_{(i-1)\Delta_n} \left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \right] \right) = A_n^{(1)} + A_n^{(2)}, \quad \text{where}$$

$$A_n^{(1)} = \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s) dW_s^{(2)} \tag{23}$$

$$+ \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(2)} dW_s^{(2)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) dW_s^{(1)} \tag{24}$$

$$A_n^{(2)} = \sum_{i=1}^n \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s) dW_s^{(2)} \right. \tag{25}$$

$$\left. - \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) g^{(2)}(i\Delta_n - s) \rho_s \, ds \right). \tag{26}$$

Proposition 5.3. *We have the following L^2-convergence: $A_n^{(1)} \xrightarrow{L^2} 0$ as $n \to \infty$.*

Proof. The two terms (23) and (24) in $A_n^{(1)}$ are symmetric. Hence we only present the proof that the first one converges to zero, and the proof for the second one will follow in an analogous way. We compute the L^2 norm of the first term in $A_n^{(1)}$:

$$\mathbb{E}\left[\left(\sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s) dW_s^{(2)} \right)^2 \right]. \tag{27}$$

It is easy to see that the cross products originating by squaring the sum are equal to:

$$2 \sum_{j<i} \mathbb{E}\left[\int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} dW_s^{(1)} \int_{-\infty}^{(j-1)\Delta_n} \Delta_j^n g_s^{(1)} dW_s^{(1)} \right.$$

$$\left. \times \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s) dW_s^{(2)} \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}(j\Delta_n - s) dW_s^{(2)} \right],$$

which is easily seen to be zero, by the tower property, conditioning on $\mathscr{G}_{(i-1)\Delta_n}$. The other term originating from (27), is the sum of the n terms squared:

$$\sum_{i=1}^{n} \mathbb{E}\left[\left(\int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\, dW_s^{(2)}\right)^2\right]$$
$$= \sum_{i=1}^{n} \int_{-\infty}^{(i-1)\Delta_n} \left(\Delta_i^n g_s^{(1)}\right)^2 ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}(i\Delta_n - s)\right)^2 ds, \tag{28}$$

thanks to an application of Corollary A.5 in the Appendix.

Now (28) becomes, after a change of variable:

$$n \int_0^\infty \left(g^{(1)}(s + \Delta_n) - g^{(1)}(s)\right)^2 ds \int_0^{\Delta_n} \left(g^{(2)}(s)\right)^2 ds. \tag{29}$$

From Remark 2.5 it follows that it is only necessary to consider the bounded interval $(0, b^{(1)})$ in the first integral of (29), since on $[b^{(1)}, +\infty)$ the integral is $O(\Delta_n^2)$.

It then follows from Assumption 2.2 that the quantity in (29) can be written as follows, for some constant C:

$$Cn\Delta_n^{2\delta^{(1)}+2\delta^{(2)}+2} L_1^{(2)}(\Delta_n) L_2^{(1)}(\Delta_n) = C\Delta_n^{2\delta^{(1)}+2\delta^{(2)}+1} L_1^{(2)}(\Delta_n) L_2^{(1)}(\Delta_n) \to 0,$$

since $\delta^{(1)} + \delta^{(2)} \geq 0$, the product of slowly varying function is slowly varying and the Potter bounds apply, which we recall in Remark 5.4 below. $\qquad\square$

Remark 5.4 (Potter's bound for slowly varying functions at zero). *For a slowly varying function L at infinity, we have Potter's bound: there exist constants $\bar{u}, C_1, C_2$ such that, for any $\delta > 0$ and $u > \bar{u}$:*
$$C_1 u^{-\delta} \leq L(u) \leq C_2 u^\delta.$$

Call $M(u) := L(\frac{1}{u})$. M is slowly varying at zero. Potter's bound, for small u, gives us:

$$C_1 \left(\frac{1}{u}\right)^{-\delta} \leq L\left(\frac{1}{u}\right) \leq C_2 \left(\frac{1}{u}\right)^\delta,$$

which we can rewrite as:

$$C_1 u^\delta \leq M(u) \leq C_2 u^{-\delta}.$$

In particular, for any $\alpha > 0$, choose $\delta \in (0, \alpha)$. Then:

$$0 \leq t^\alpha M(t) \leq C_2 t^\alpha t^{-\delta} = C_2 t^{\alpha-\delta} \to 0$$

as $t \to 0$. Henceforth, for all $\alpha > 0$, $\lim_{t\to 0} t^\alpha M(t) = 0$.

For the first part of the proof of Theorem 3.2, we are only left with the term $A_n^{(2)}$, where we recall that

$$A_n^{(2)} = \sum_{i=1}^{n} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\, dW_s^{(2)}$$

17

$$- \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\rho_s \, ds.$$

Using that $dW_t^{(2)} = \rho_t \, dW^{(1)} + \sqrt{(1 - \rho_t^2)} \, d\tilde{W}_t$, we will split it into two terms. I.e. we will write $A_n^{(2)} = A_n^{(2,1)} + A_n^{(2,2)}$, where the first one is given by

$$A_n^{(2,1)} = \sum_{i=1}^{n} \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \right)$$

and is dealt with in Lemma 5.5. The second one, whose convergence is more difficult to prove, is given by:

$$A_n^{(2,2)} = \sum_{i=1}^{n} \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\rho_s \, dW_s^{(1)} \right.$$
$$\left. - \int_{(i-1)\Delta_n}^{i\Delta_n} (g^{(1)}(i\Delta_n - s))(g^{(2)}(i\Delta_n - s)) \, \rho_s \, ds \right).$$

A proof of its convergence can be found in Proposition 5.6. We will start with the easier one:

Lemma 5.5. *We have the following L^2-convergence: $A_n^{(2,1)} \xrightarrow{L^2} 0$ as $n \to \infty$.*

Proof. We get:

$$\mathbb{E}\left[\left(\sum_{i=1}^{n} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \right)^2 \right] \quad (30)$$

$$= \sum_{i=1}^{n} \mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \right)^2 \right] + \quad (31)$$

$$\sum_{i \neq j} \mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \times \right. \quad (32)$$

$$\left. \times \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}(j\Delta_n - s) \, dW_s^{(1)} \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}(j\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \right].$$

Since $\mathscr{F}^{W^{(1)}}$ is independent of $\mathscr{F}^{\rho} \vee \mathscr{F}^{\tilde{W}}$, it is easy to deal with the term in (31):

$$\sum_{i=1}^{n} \mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \right)^2 \right]$$

$$= \sum_{i=1}^{n} \mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \right)^2 \right] \mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1 - \rho_s^2} \, d\tilde{W}_s \right)^2 \right]$$
$$(33)$$

$$= \sum_{i=1}^{n} \int_0^{\Delta_n} \left(g^{(1)}(s) \right)^2 ds \int_0^{\Delta_n} \left(g^{(2)}(s) \right)^2 \left(1 - \rho_{i\Delta_n - s}^2 \right) ds \quad (34)$$

18

$$\leq 2n \int_0^{\Delta_n} \left(g^{(1)}(s)\right)^2 ds \int_0^{\Delta_n} \left(g^{(2)}(s)\right)^2 ds = C\Delta_n^{2(\delta^{(1)}+\delta^{(2)}+1)} L_1^{(1)}(\Delta_n) L_1^{(2)}(\Delta_n) n \to 0,$$

(35)

for a constant $C > 0$, where the asymptotics on line (35) follow from Assumption 2.2 and Remark 5.4. Lastly, the summand in (32) becomes:

$$\sum_{i\neq j} \mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1-\rho_s^2}\, d\tilde{W}_s \times \right.$$
$$\left. \times \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}(j\Delta_n - s)\, dW_s^{(1)} \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}(j\Delta_n - s)\sqrt{1-\rho_s^2}\, d\tilde{W}_s \right]$$
$$= \sum_{i\neq j} \mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\sqrt{1-\rho_s^2}\, d\tilde{W}_s \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}(j\Delta_n - s)\sqrt{1-\rho_s^2}\, d\tilde{W}_s \right] \times$$
$$\mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)} \times \int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}(j\Delta_n - s)\, dW_s^{(1)} \right] = 0,$$

since $((i-1)\Delta_n, i\Delta_n) \cap ((j-1)\Delta_n, j\Delta_n) = \emptyset$. $\qquad\square$

Now we tackle the difficult term:

Proposition 5.6. *We have the following convergence in probability:* $A_n^{(2,2)} \overset{\mathbb{P}}{\to} 0$ *as* $n \to \infty$.

We will use a *blocking* technique, which consists in freezing the correlation process in each of the intervals $((i-1)\Delta_n, i\Delta_n)$. Exploiting the Hölder continuity assumption on the paths of ρ, we will show that this "frozen" quantity is a good substitute of the one we have: We need to show that the difference converges to zero (at least) in probability. The following lemma shows that there is essentially no harm in doing so for the Lebesgue integral part:

Lemma 5.7. *We have the following almost sure convergence:*

$$\sum_{i=1}^n \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\left(\rho_s - \rho_{(i-1)\Delta_n}\right) ds \overset{a.s.}{\to} 0, \qquad as\ n \to \infty.$$

Proof. We find that:

$$\left| \sum_{i=1}^n \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\left(\rho_s - \rho_{(i-1)\Delta_n}\right) ds \right|$$
$$\leq \sum_{i=1}^n \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\left|\rho_s - \rho_{(i-1)\Delta_n}\right| ds$$
$$\leq \Delta_n^\alpha n \int_0^{\Delta_n} g^{(1)}(s)g^{(2)}(s)\, ds \leq \Delta_n^\alpha n \sqrt{\int_0^{\Delta_n} \left(g^{(1)}(s)\right)^2 ds} \sqrt{\int_0^{\Delta_n} \left(g^{(2)}(s)\right)^2 ds}$$
$$= C(\Delta_n^{\delta^{(1)}+\delta^{(2)}+\alpha})\sqrt{L_1^{(1)}(\Delta_n) L_1^{(2)}(\Delta_n)} \to 0.$$

$\qquad\square$

19

For the Brownian part, the situation is similar, but we cannot ask for a.s. convergence.

Lemma 5.8. *We have the following L^2-convergence:*

$$\sum_{i=1}^{n} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\left(\rho_s - \rho_{(i-1)\Delta_n}\right) dW_s^{(1)} \xrightarrow{L^2} 0, \quad \text{as } n \to \infty.$$

Proof. Using the inequality $\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \le n \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right]$, we obtain:

$$\mathbb{E}\left[\left(\sum_{i=1}^{n} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\left(\rho_s - \rho_{(i-1)\Delta_n}\right) dW_s^{(1)}\right)^2\right]$$

$$\le n \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)}\right.\right.\right.$$

$$\left.\left.\left. \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\left(\rho_s - \rho_{(i-1)\Delta_n}\right) dW_s^{(1)}\right)^2 \middle| \mathscr{F}^\rho\right]\right]$$

$$= n \sum_{i=1}^{n} \mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(1)}(i\Delta_n - s)\right)^2 ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}(i\Delta_n - s)\right)^2 \left(\rho_s - \rho_{i\Delta_n - s}\right)^2 ds\right.$$

$$\left. + 2\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\left(\rho_s - \rho_{i\Delta_n - s}\right) ds\right)^2\right], \tag{36}$$

where the last equality (36) follows from an application of Corollary A.5 in the appendix. Furthermore, by Hölder's inequality we have

$$\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\left|\rho_s - \rho_{i\Delta_n - s}\right| ds\right)^2$$

$$\le \left(\int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(1)}(i\Delta_n - s)\right)^2 ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}(i\Delta_n - s)\right)^2 \left(\rho_s - \rho_{i\Delta_n - s}\right)^2 ds\right)^2.$$

Finally, using that $|\rho_s - \rho_{(i-1)\Delta_n}| \le C\Delta_n^\alpha$, for some positive constant C:

$$3Cn \sum_{i=1}^{n} \mathbb{E}\left[\int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(1)}(i\Delta_n - s)\right)^2 ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}(i\Delta_n - s)\right)^2 \left(\rho_s - \rho_{i\Delta_n - s}\right)^2 ds\right]$$

$$\le n^2 \Delta_n^{2\alpha} \int_0^{\Delta_n} \left(g^{(1)}(s)\right)^2 ds \int_0^{\Delta_n} \left(g^{(2)}(s)\right)^2 ds = \Delta_n^{2(\delta^{(1)} + \delta^{(2)} + \alpha)} L_1^{(1)}(\Delta_n) L_1^{(2)}(\Delta_n) \to 0,$$

where the asymptotic bound follows again from Assumption 2.2 and Remark 5.4. $\qquad\square$

Proof of Proposition 5.6. We can now prove the statement of Proposition 5.6. In the light of Lemmas 5.7 and 5.8, it will be sufficient to show that:

$$\sum_{i=1}^{n} \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\rho_{(i-1)\Delta_n}\, dW_s^{(1)}\right.$$

$$-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\rho_{(i-1)\Delta_n}\,ds\Bigg) \quad (37)$$

converges to zero. We will show that it does so in L^2. Since we can take out the random variable $\rho_{(i-1)\Delta_n}$ from the stochastic integrals in (37), it remains to prove that:

$$\mathbb{E}\Bigg[\Bigg(\sum_{i=1}^{n}\rho_{(i-1)\Delta_n}\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s)\,dW_s^{(1)}$$

$$-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\,ds\Bigg)\Bigg)^2\Bigg] \quad (38)$$

converges to zero. For ease of notation, we will suppress the arguments of the functions $g^{(1)}, g^{(2)}$ in parts of the present proof. By expanding (38) we obtain two terms: the sum of the squares and the sum of the cross products. We start with the latter:

$$\sum_{\substack{i,j=1\\i\neq j}}^{n}\mathbb{E}\Bigg[\rho_{(i-1)\Delta_n}\rho_{(j-1)\Delta_n}\times\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)$$

$$\Bigg(\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)\Bigg]. \quad (39)$$

The expectation splits thanks to independence between $\mathscr{F}^\rho$ and $\mathscr{F}^{W^{(1)}}$. Now suppose $i < j$; then by the independent increments property of the integrals:

$$\mathbb{E}\Bigg[\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)\times$$

$$\Bigg(\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)\Bigg]$$

$$=\mathbb{E}\Bigg[\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)\Bigg]\times$$

$$\mathbb{E}\Bigg[\Bigg(\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(j-1)\Delta_n}^{j\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)\Bigg]=0.$$

Hence we are only left with considering the sum of squares:

$$\sum_{i=1}^{n}\mathbb{E}\Bigg[\rho_{(i-1)\Delta_n}^2\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)^2\Bigg]$$

$$=\sum_{i=1}^{n}\mathbb{E}\Big[\rho_{(i-1)\Delta_n}^2\Big]$$

$$\cdot\mathbb{E}\Bigg[\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\,dW_s^{(1)}\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\,dW_s^{(1)}-\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\,ds\Bigg)^2\Bigg].$$

$$(40)$$

Now:

$$
\mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\, dW_s^{(1)} - \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\, ds\right)^2\right]
$$

$$
= \mathbb{E}\left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}\, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}\, dW_s^{(1)}\right)^2\right] - \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\, ds\right)^2
$$

$$
= \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(1)}\right)^2 ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}\right)^2 ds + \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}g^{(2)}\, ds\right)^2 \tag{41}
$$

$$
\leq 2 \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(1)}\right)^2 (s)\, ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}\right)^2 (s)\, ds \tag{42}
$$

$$
= C\Delta_n^{2(\delta^{(1)}+\delta^{(2)}+1)} L_1^{(1)}(\Delta_n) L_1^{(2)}(\Delta_n),
$$

where (41) is an application of Lemma 68. Thus, (40) finally goes to zero:

$$
C\Delta_n^{2(\delta^{(1)}+\delta^{(2)}+1)} L_1^{(1)}(\Delta_n) L_1^{(2)}(\Delta_n) \sum_{i=1}^n \mathbb{E}\left[\rho_{(i-1)\Delta_n}^2\right] \leq C\Delta_n^{2(\delta^{(1)}+\delta^{(2)})+1} L_1^{(1)}(\Delta_n) L_1^{(2)}(\Delta_n)
$$

$$
\to 0,
$$

for some positive constant C. The proof of Proposition 5.6 is now complete. $\square$

5.1.3 Convergence of B_n

What remains to be shown is the convergence of the term B_n. Recall that

$$
B_n = \sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right] - g^{(1)}(0^+)g^{(2)}(0^+) \int_0^1 \rho_s\, ds, \tag{43}
$$

where we recall the following expression, which we derived in (19) and (20):

$$
\sum_{i=1}^n \mathbb{E}_{(i-1)\Delta_n}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right] = \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} dW^{(1)} \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(2)} dW^{(2)} +
$$

$$
\sum_{i=1}^n \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\rho_s\, ds,
$$

with $\Delta_i^n g_s^{(j)} := g^{(j)}(i\Delta_n - s) - g^{(j)}((i-1)\Delta_n - s)$. Hence we can write $B_n = B_n^{(1)} + B_n^{(2)}$, where

$$
B_n^{(1)} = \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)}\, dW_s^{(1)} \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(2)}\, dW_s^{(2)},
$$

and

$$
B_n^{(2)} = \sum_{i=1}^n \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s)\rho_s\, ds - g^{(1)}(0^+)g^{(2)}(0^+) \int_0^1 \rho_s\, ds.
$$

Lemma 5.9. *We have the following L^2-convergence: $B_n^{(1)} \xrightarrow{L^2} 0$ as $n \to \infty$.*

Proof. Using the inequality $\mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)^2\right] \leq n \sum_{i=1}^n \mathbb{E}\left[X_i^2\right]$, it suffices to look at:

$$n \sum_{i=1}^n \mathbb{E}\left[\left(\int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} \, dW_s^{(1)} \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(2)} \, dW_s^{(2)}\right)^2\right]$$

$$= n \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} (\Delta_i^n g_s^{(1)})^2 \, ds \int_{-\infty}^{(i-1)\Delta_n} (\Delta_i^n g_s^{(2)})^2 \, ds$$

$$+ 2n \sum_{i=1}^n \left(\int_{-\infty}^{(i-1)\Delta_n} (\Delta_i^n g_s^{(1)})(\Delta_i^n g_s^{(2)}) \rho_s \, ds\right)^2 \tag{44}$$

$$\leq 3n \sum_{i=1}^n \int_{-\infty}^{(i-1)\Delta_n} (\Delta_i^n g_s^{(1)})^2 \, ds \int_{-\infty}^{(i-1)\Delta_n} (\Delta_i^n g_s^{(2)})^2 \, ds, \tag{45}$$

by Hölder's inequality. Each of the two integrals in (45) can be decomposed as:

$$\int_{-\infty}^{(i-1)\Delta_n} \left(\Delta_i^n g_s^{(j)}\right)^2 \, ds$$

$$= \int_0^{b^{(j)}} (g^{(j)}(s + \Delta_n) - g^{(j)}(s))^2 \, ds + \int_{b^{(j)}}^{+\infty} (g^{(j)}(s + \Delta_n) - g^{(j)}(s))^2 \, ds. \tag{46}$$

The integral on the unbounded domain is $O(\Delta_n^2)$, as in Remark 2.5. Hence it suffices to consider the product of the integrals in the bounded domains, i.e.:

$$n \sum_{i=1}^n \int_0^{b^{(1)}} (g^{(1)}(s + \Delta_n) - g^{(1)}(s))^2 \, ds \int_0^{b^{(2)}} (g^{(2)}(s + \Delta_n) - g^{(2)}(s))^2 \, ds. \tag{47}$$

Thanks to Remark 5.4 and Assumption 2.2, the quantity in (47) becomes, for some $C > 0$:

$$Cn^2 \Delta_n^{2+2(\delta^{(1)}+\delta^{(2)})} L_2^{(1)}(\Delta_n) L_2^{(2)}(\Delta_n) \to 0.$$

$\square$

Remark 5.10. *Note that the last bound is sharp, in the sense that we obtained $O(\Delta_n^{2(\delta^{(1)}+\delta^{(2)})})$ instead of $O(\Delta_n^{2(\delta^{(1)}+\delta^{(2)}+\alpha)})$. This is the only term in the proof where we cannot obtain more regularity by the Hölder continuity of the paths of ρ. Hence, Assumption 3.1 which we are imposing in this section is crucial here.*

We conclude with the last part of the convergence theorem.

Proposition 5.11. *We have the following almost sure convergence result: $B_n^{(2)} \xrightarrow{a.s.} 0$ as $n \to \infty$.*

Proof. Using Lemma 5.7 we can simply show that

$$\sum_{i=1}^n \rho_{(i-1)\Delta_n} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) g^{(2)}(i\Delta_n - s) \, ds$$

converges almost surely to the limit $g^{(1)}(0^+)g^{(2)}(0^+) \int_0^1 \rho_s \, ds$. To this end, note that

$$\sum_{i=1}^n \rho_{(i-1)\Delta_n} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)g^{(2)}(i\Delta_n - s) \, ds \tag{48}$$

$$= \sum_{i=1}^n \rho_{(i-1)\Delta_n} \int_0^{\Delta_n} g^{(1)}(y)g^{(2)}(y) \, dy \tag{49}$$

$$= \left(\int_0^{\Delta_n} g^{(1)}(y)g^{(2)}(y) \, dy \right) \sum_{i=1}^n \rho_{(i-1)\Delta_n} \tag{50}$$

$$= \Delta_n g^{(1)}(\zeta)g^{(2)}(\zeta)1_{\{\zeta \in [0,\Delta_n]\}} \sum_{i=1}^n \rho_{(i-1)\Delta_n} \overset{\text{a.s.}}{\to} g^{(1)}(0^+)g^{(2)}(0^+) \int_0^1 \rho_s \, ds. \tag{51}$$

Line (49) follows by setting $y = i\Delta_n - s$, while line (51) follows from the mean value theorem and the fact that the paths of ρ are almost surely Riemann integrable. $\qquad\square$

This concludes the proof of Theorem 3.2, that deals with the case where the indices of regular variation satisfy: $\delta^{(1)} + \delta^{(2)} \geq 0$.

In the following section, we wish to lift such a restriction, and prove another law of large numbers, this time for a scaled version of the realised covariation.

5.2   Proof of Theorem 3.6

As in the last section, thanks to Remark 5.1, we will prove convergence in probability for $t = 1$. In order to prove this result, we denote by $\mathscr{F}^\rho := \sigma\{\rho_s, s \in \mathbb{R}\}$ and we start by considering the following decomposition:

$$\Delta_n \frac{\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}}{c(\Delta_n)} - \int_0^1 \rho_l \, dl = D_n^{(1)} + D_n^{(2)}, \text{ where}$$

$$D_n^{(1)} = \Delta_n \frac{\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}}{c(\Delta_n)} - \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta_n}{c(\Delta_n)} \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \middle| \mathscr{F}^\rho \right], \tag{52}$$

$$D_n^{(2)} = \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta_n}{c(\Delta_n)} \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \middle| \mathscr{F}^\rho \right] - \int_0^1 \rho_l \, dl.$$

5.2.1   Convergence of $D_n^{(2)}$

Note that

$$\mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \middle| \mathscr{F}^\rho \right] = \int_{-\infty}^{i\Delta_n} \varphi_{\Delta_n}^{(1)}(i\Delta_n - s)\varphi_{\Delta_n}^{(2)}(i\Delta_n - s)\rho_s \, ds$$

$$= \int_0^\infty \varphi_{\Delta_n}^{(1)}(s)\varphi_{\Delta_n}^{(2)}(s)\rho_{i\Delta_n - s} \, ds.$$

Hence we can write

$$\Delta_n \frac{\sum_{i=1}^n \mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \middle| \mathscr{F}^\rho\right]}{c(\Delta_n)} = \frac{\int_0^\infty \varphi_{\Delta_n}^{(1)}(s) \varphi_{\Delta_n}^{(2)}(s) \Delta_n \left(\sum_{i=1}^n \rho_{i\Delta_n - s}\right) ds}{c(\Delta_n)} \tag{53}$$

$$= \int_{\mathbb{R}^+} \Delta_n \left(\sum_{i=1}^n \rho_{i\Delta_n - s}\right) d\pi_n(s),$$

where, for each $n \in \mathbb{N}$, π_n is a measure whose density with respect to the Lebesgue measure is given by

$$\frac{\varphi_{\Delta_n}^{(1)}(s) \varphi_{\Delta_n}^{(2)}(s)}{\int_0^\infty \varphi_{\Delta_n}^{(1)}(s) \varphi_{\Delta_n}^{(2)}(s)\, ds}.$$

Note however that $\pi_n(ds)$ is a probability measure if both $g^{(1)}$ and $g^{(2)}$ are positive and for $s > \Delta_n$ one has $\left(g^{(1)}(s) - g^{(1)}(s - \Delta_n)\right) \times \left(g^{(2)}(s) - g^{(2)}(s - \Delta_n)\right) \geq 0$, that is, if both are monotonically increasing or decreasing. This is why we needed to introduce Assumption 3.2.

Let us now formulate a key result needed for establishing the convergence of $D_n^{(2)}$.

Proposition 5.12. *Assume that the assumptions of Theorem 3.6 hold. If the probability measures π_n defined after equation (53) converge weakly to a probability π on $[0, \infty)$ then:*

$$\Delta_n \frac{\sum_{i=1}^n \mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \middle| \mathscr{F}^\rho\right]}{c(\Delta_n)} \overset{a.s.}{\to} \int_{\mathbb{R}^+} \left(\int_{-s}^{1-s} \rho_l\, dl\right) d\pi(s), \quad as\ n \to \infty. \tag{54}$$

Proof. Denote by $f_n(x, \omega) := \Delta_n \sum_{i=1}^n \rho_{i\Delta_n - x}(\omega)$ and by $f(x, \omega) := \int_{-x}^{1-x} \rho_l(\omega)\, dl$. Furthermore, for a measure μ on $[0, +\infty)$ and a measurable function g: $\mu(g) := \int_{\mathbb{R}^+} g(x)\, d\mu(x)$. For a.s. ω, $\lim_{n \to \infty} f_n(x, \omega) = f(x, \omega)$.

Let l_i^* be any point in $[-x + (i-1)\Delta_n, -x + i\Delta_n]$. The error we make at the n-th step with a Riemann sum is:

$$|f_n - f| = \left|\sum_{i=1}^n \int_{-x+(i-1)\Delta_n}^{-x+i\Delta_n} \left(\rho_l - \rho_{l_i^*}\right) dl\right| \leq \sum_{i=1}^n \int_{-x+(i-1)\Delta_n}^{-x+i\Delta_n} K|l - l_i^*|^\alpha\, dl$$

$$\leq \sum_{i=1}^n \int_{-x+(i-1)\Delta_n}^{-x+i\Delta_n} K\Delta_n^\alpha\, dl = Kn\Delta_n^{\alpha+1},$$

for a positive constant K, hence:

$$|\pi_n(f_n) - \pi_n(f)| \leq Kn\Delta_n^{\alpha+1} \pi_n(\mathbb{R}^+) = Kn\Delta_n^{\alpha+1} \to 0,$$

since $\pi_n(\mathbb{R}^+) := \int_{\mathbb{R}^+} d\pi_n(x) = 1$. $\qquad\square$

Proposition 5.13. *Assume that the assumptions of Theorem 3.6 hold. Then $D_n^{(2)} \overset{a.s.}{\to} 0$ as $n \to \infty$.*

Proof. We remark that checking weak convergence of probability measures is equivalent to checking pointwise convergence of the corresponding cumulative distribution functions. Hence we write

$$F_n(x) = \pi_n\left([0,x]\right) = \frac{\int_0^x \varphi_{\Delta_n}^{(1)}(s)\varphi_{\Delta_n}^{(2)}(s)\, ds}{\int_0^\infty \varphi_{\Delta_n}^{(1)}(s)\varphi_{\Delta_n}^{(2)}(s)\, ds}.$$

Now fix $\varepsilon > \Delta_n$. Then, for all $x > \varepsilon$, one has:

$$F_n(x) - F_n(\varepsilon) = \frac{\int_\varepsilon^x \varphi_{\Delta_n}^{(1)}(s)\varphi_{\Delta_n}^{(2)}(s)\, ds}{\int_0^\infty \varphi_{\Delta_n}^{(1)}(s)\varphi_{\Delta_n}^{(2)}(s)\, ds} \le \frac{1}{c(\Delta_n)}\sqrt{\int_\varepsilon^x \left(\varphi_{\Delta_n}^{(1)}(s)\right)^2 ds \int_\varepsilon^x \left(\varphi_{\Delta_n}^{(2)}(s)\right)^2 ds}.$$

Denoting by $b^{(1)}$ a real number such that $\left(g^{(1)'}\right)^2(s)$ is decreasing for $s > b^{(1)}$:

$$\int_\varepsilon^x \left(\varphi_{\Delta_n}^{(1)}(s)\right)^2 ds = \int_\varepsilon^{b^{(1)}} \left(\varphi_{\Delta_n}^{(1)}(s)\right)^2 ds + \int_{b^{(1)}}^x \left(\varphi_{\Delta_n}^{(1)}(s)\right)^2 ds$$

$$\le \Delta_n^2 \left[(b^{(1)} - \varepsilon)\sup_{[\varepsilon, b^{(1)}]} \left(g^{(1)'}\right)^2(s) + \left\|g^{(1)'}\right\|_{L^2[\varepsilon, \infty)} \right] = C\Delta_n^2 \to 0,$$

so

$$\lim_{n\to\infty} F_n(x) - F_n(\varepsilon) = 0, \tag{55}$$

which implies that the limiting cumulative distribution function will be constant on the sets (ε, ∞), for all $\varepsilon > 0$, and hence its mass will intuitively concentrate at zero. We now formalise this observation.

We observe that the family of measures $\{\pi_n\}_{n=1}^\infty$ is easily proven to be tight. Indeed, fix a small $\zeta > 0$, then if $b = \max(b^{(1)}, b^{(2)})$, by the above computations:

$$\int_b^\infty \varphi_{\Delta_n}^{(1)}(s)\varphi_{\Delta_n}^{(2)}(s)\, ds \le C\Delta_n^2,$$

which is smaller than ζ for n big enough, say for $n > n(\zeta)$. Hence the compact set $[0, b]$ is such that $\pi_n[0, b] > 1 - \zeta$ for all $n > n(\zeta)$. The finite collection $\{\pi_i\}_{i=1}^{n(\zeta)}$ is obviously tight, and thus we can conclude that the whole family $\{\pi_n\}$ is tight.

By tightness and Prohorov's theorem, the sequence π_n is relatively compact, that is, every subsequence contains a further subsubsequence converging weakly to some probability measure $\mathbb{Q}$ (that could depend on the chosen subsequence). By (55), the only possible weak limit is the unit mass at zero, and hence every subsequence has a subsubsequence converging to δ_0. But this fact is equivalent to the weak convergence $\pi_n \Rightarrow \delta_0$ (see Billingsley (2009), Theorem 2.6).

Now from (54) and the above remark we obtain that

$$\Delta_n \frac{\sum_{i=1}^n \mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \middle| \mathscr{F}^\rho\right]}{c(\Delta_n)} \xrightarrow{\text{a.s.}} \int_{\mathbb{R}^+} \left(\int_{-s}^{1-s} \rho_l\, dl\right) d\delta_0(s) = \int_0^1 \rho_l\, dl, \quad \text{as } n \to \infty,$$

i.e. $D_n^{(2)} \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. $\qquad\square$

26

5.2.2 Convergence of $D_n^{(1)}$

It remains to prove that $D_n^{(1)}$ converges to 0. To this end note that asking that its L^2-norm goes to zero is equivalent to showing that $\mathrm{Var}(D_n^{(1)}|\mathscr{F}^\rho) \to 0$ as $n \to \infty$.

Proposition 5.14. *Suppose that the assumptions of Theorem 3.6 hold. Then:*

$$\mathrm{Var}\left(\Delta_n \frac{\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}}{c(\Delta_n)}\middle|\mathscr{F}^\rho\right) \to 0, \quad \text{as } n \to \infty.$$

Proof. We can write

$$\mathrm{Var}\left(\Delta_n \frac{\sum_{i=1}^n \Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}}{c(\Delta_n)}\middle|\mathscr{F}^\rho\right) = \frac{\Delta_n^2}{c^2(\Delta_n)}\left(\sum_{i=1}^n \mathrm{Var}\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\middle|\mathscr{F}^\rho\right)\right)$$

$$+2\frac{\Delta_n^2}{c^2(\Delta_n)}\left(\sum_{i=1}^n \sum_{j=i+1}^n \mathrm{Cov}\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}, \Delta_j^n Y^{(1)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right)\right). \quad (56)$$

Now observe that:

$$\mathrm{Var}\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\middle|\mathscr{F}^\rho\right) \le \mathbb{E}\left[\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\right)^2\middle|\mathscr{F}^\rho\right] \sim Kc^2(\Delta_n),$$

for some positive constant $K > 0$ thanks to Assumption 2.2 and the computations in the previous sections. Hence the first term in (56) is asymptotically equivalent to:

$$Kn\Delta_n^2 \to 0.$$

Now we have that

$$\mathrm{Cov}\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}, \Delta_j^n Y^{(1)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right) = \mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)} \Delta_j^n Y^{(1)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right]$$

$$-\mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}\middle|\mathscr{F}^\rho\right]\mathbb{E}\left[\Delta_j^n Y^{(1)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right].$$

$$(57)$$

Thanks to Lemma A.4, we know how to express the first expectation in (57) and write

$$\mathrm{Cov}\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}, \Delta_j^n Y^{(1)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right) =$$

$$\mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_j^n Y^{(1)}\middle|\mathscr{F}^\rho\right]\mathbb{E}\left[\Delta_i^n Y^{(2)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right]$$

$$+\mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_j^n Y^{(2)}\middle|\mathscr{F}^\rho\right]\mathbb{E}\left[\Delta_i^n Y^{(2)} \Delta_j^n Y^{(1)}\middle|\mathscr{F}^\rho\right]. \quad (58)$$

Now, recall the notation $\Delta_i^n g_s^{(k)} := g^{(k)}(i\Delta_n - s) - g^{(k)}((i-1)\Delta_n - s)$, for $k \in \{1, 2\}$. Since $j > i$, we have:

$$\mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_j^n Y^{(1)} \middle| \mathscr{F}^\rho\right] = \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\Delta_j^n g_s^{(1)}\, ds + \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} \Delta_j^n g_s^{(1)}\, ds$$

$$= \int_0^{\Delta_n} g^{(1)}(s)\left(g^{(1)}((j-i)\Delta_n + s) - g^{(1)}((j-i-1)\Delta_n + s)\right)\, ds +$$

$$\int_0^\infty \left(g^{(1)}(\Delta_n + s) - g^{(1)}(s)\right)\left(g^{(1)}((j-i)\Delta_n + s) - g^{(1)}((j-i-1)\Delta_n + s)\right)\, ds, \quad (59)$$

and similarly:

$$\mathbb{E}\left[\Delta_i^n Y^{(1)} \Delta_j^n Y^{(2)} \middle| \mathscr{F}^\rho\right] = \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s)\Delta_i^n g_s^{(2)} \rho_s\, ds + \int_{-\infty}^{(i-1)\Delta_n} \Delta_i^n g_s^{(1)} \Delta_j^n g_s^{(2)} \rho_s\, ds,$$

and analogously for the other terms in (58).

These terms do not simply depend on the difference $j - i$ because of the presence of the correlation process ρ. Nevertheless, ρ is bounded by 1 in absolute value, so passing to the absolute values, we can write:

$$\left|\mathrm{Cov}\left(\Delta_i^n Y^{(1)} \Delta_i^n Y^{(2)}, \Delta_j^n Y^{(1)} \Delta_j^n Y^{(2)} \middle| \mathscr{F}^\rho\right)\right|$$

$$\leq \mathbb{E}\left[\left|\Delta_i^n Y^{(1)} \Delta_j^n Y^{(1)}\right| \middle| \rho \equiv 1\right] \mathbb{E}\left[\left|\Delta_i^n Y^{(2)} \Delta_j^n Y^{(2)}\right| \middle| \rho \equiv 1\right]$$

$$+ \mathbb{E}\left[\left|\Delta_i^n Y^{(1)} \Delta_j^n Y^{(2)}\right| \middle| \rho \equiv 1\right] \mathbb{E}\left[\left|\Delta_i^n Y^{(2)} \Delta_j^n Y^{(1)}\right| \middle| \rho \equiv 1\right]. \quad (60)$$

The proof that the quantity goes to zero follows from Theorem 4.1 in Granelli and Veraart (2017). Details on how to apply that theorem in this setting are as follows.

Denote by $r_{a,b}^{(n)}(j - i) = \mathbb{E}\left[\left|\Delta_i^n Y^{(a)} \Delta_j^n Y^{(b)}\right| \middle| \rho \equiv 1\right]$, for $a, b \in \{1, 2\}$. Then, in fact, $r_{a,b}^{(n)}(j - i) = \mathbb{E}\left[\Delta_i^n G^{(a)} \Delta_j^n G^{(b)}\right]$, for $a, b \in \{1, 2\}$. The second term in line (56) is then dominated by:

$$2\frac{\Delta_n^2}{c^2(\Delta_n)} \sum_{i=1}^n \sum_{j=i+1}^n \left[r_{1,1}^{(n)}(j-i)r_{2,2}^{(n)}(j-i) + r_{1,2}^{(n)}(j-i)r_{2,1}^{(n)}(j-i)\right] \quad (61)$$

$$= 2\frac{\Delta_n^2}{c^2(\Delta_n)} \sum_{i=1}^n \sum_{j-i=1}^n \left[r_{1,1}^{(n)}(j-i)r_{2,2}^{(n)}(j-i) + r_{1,2}^{(n)}(j-i)r_{2,1}^{(n)}(j-i)\right] \quad (62)$$

$$= 2\frac{\Delta_n^2}{c^2(\Delta_n)} n \sum_{i=1}^n \left[r_{1,1}^{(n)}(i)r_{2,2}^{(n)}(i) + r_{1,2}^{(n)}(i)r_{2,1}^{(n)}(i)\right] \quad (63)$$

$$= 2\frac{1}{n} \sum_{i=1}^n \left[\frac{r_{1,1}^{(n)}(i)r_{2,2}^{(n)}(i)}{c^2(\Delta_n)} + \frac{r_{1,2}^{(n)}(i)r_{2,1}^{(n)}(i)}{c^2(\Delta_n)}\right]. \quad (64)$$

With our assumptions, Theorem 4.3 in Granelli and Veraart (2017) gives us the uniform bound for a positive constant C:

$$\frac{|\mathbb{E}(\Delta_1^n G^{(a)} \Delta_{1+i}^n G^{(b)})|}{\tau_n^{(a)} \tau_n^{(b)}} \leq C(i-1)^{\delta^{(a)} + \delta^{(b)} + \varepsilon - 1},$$

for all $\varepsilon > 0$, where, for $i \in \{a, b\}$:

$$\tau_n^{(i)} := \sqrt{\mathbb{E}\left[\left(\Delta_1^n G^{(i)}\right)^2\right]} = \sqrt{\int_0^\infty \left(g^{(i)}(s + \Delta_n) - g^{(i)}(s)\right)^2 ds + \int_0^{\Delta_n} \left(g^{(i)}(s)\right)^2 ds}$$

$$= \sqrt{\bar{R}^{(i)}(\Delta_n)} = \Delta_n^{\delta^{(i)}+1/2}\sqrt{L_0^{(i)}(\Delta_n)}.$$

We can write

$$\frac{\left|r_{a,b}^{(n)}(k)\right|}{c(\Delta_n)} = \frac{|\mathbb{E}(\Delta_1^n G^{(a)} \Delta_{1+k}^n G^{(b)})|}{c(\Delta_n)} = \frac{|\mathbb{E}(\Delta_1^n G^{(a)} \Delta_{1+k}^n G^{(b)})|}{\tau_n^{(a)}\tau_n^{(b)}} \frac{\tau_n^{(a)}\tau_n^{(b)}}{c(\Delta_n)}.$$

According to Assumption 3.4, we have

$$\frac{\tau_n^{(a)}\tau_n^{(b)}}{c(\Delta_n)} = \frac{\Delta_n^{\delta^{(a)}+\delta^{(b)}+1}\sqrt{L_0^{(a)}(\Delta_n)L_0^{(b)}(\Delta_n)}}{\Delta_n^{\delta^{(1)}+\delta^{(2)}+1}L_4^{(1,2)}(\Delta_n)} = \frac{\sqrt{L_0^{(a)}(\Delta_n)L_0^{(b)}(\Delta_n)}}{L_4^{(1,2)}(\Delta_n)}.$$

Hence

$$\frac{\tau_n^{(1)}\tau_n^{(1)}}{c(\Delta_n)}\frac{\tau_n^{(2)}\tau_n^{(2)}}{c(\Delta_n)} = \frac{L_0^{(1)}(\Delta_n)L_0^{(2)}(\Delta_n)}{(L_4^{(1,2)}(\Delta_n))^2} = \frac{\tau_n^{(1)}\tau_n^{(2)}}{c(\Delta_n)}\frac{\tau_n^{(2)}\tau_n^{(1)}}{c(\Delta_n)}.$$

According to Assumption 3.4 the term $\frac{L_0^{(1)}(\Delta_n)L_0^{(2)}(\Delta_n)}{(L_4^{(1,2)}(\Delta_n))^2}$ converges, so it can be bounded by a constant K. Hence we can deduce that for all $\varepsilon > 0$ there exists an $n_0(\varepsilon) \in \mathbb{N}$ such that

$$\frac{r_{1,1}^{(n)}(i)r_{2,2}^{(n)}(i)}{c(\Delta_n)^2} \le KC(i-1)^{2\delta^{(1)}+2\delta^{(2)}+2\varepsilon-2}, \quad \frac{r_{1,2}^{(n)}(i)r_{2,1}^{(n)}(i)}{c(\Delta_n)^2} \le KC(i-1)^{2\delta^{(1)}+2\delta^{(2)}+2\varepsilon-2}$$

and for all $n \ge n_0(\varepsilon)$. Convergence to zero then follows from Cesáro's Theorem and from:

$$\lim_{i \to \infty}(i-1)^{2\delta^{(1)}+2\delta^{(2)}+2\varepsilon-2} = 0 \iff 2\delta^{(1)} + 2\delta^{(2)} + 2\varepsilon - 2 < 0 \iff \varepsilon < 1 - \delta^{(1)} - \delta^{(2)},$$

which is always possible, as $\delta^{(1)} + \delta^{(2)} \in (-1, +1)$.

$\square$

5.3 Proof of Theorem 3.8

We first observe that, thanks to Theorem A.3 and Remark 5.1, *u.c.p.* convergence on $[0, T]$ is implied by convergence in probability for all $t \in [0, T]$. Without loss of generality, we will then restrict ourselves to $t = 1$. The proof will be split in several steps. We start with the familiar splitting:

$$\Delta_n \frac{\sum_{i=1}^n \Delta_i^n X^{(1)} \Delta_i^n X^{(2)}}{c(\Delta_n)} - \int_0^1 \sigma_l^{(1)}\sigma_l^{(2)}\rho_l \, dl = G_n^{(1)} + G_n^{(2)}, \text{ where}$$

$$G_n^{(1)} = \Delta_n \frac{\sum_{i=1}^n \Delta_i^n X^{(1)} \Delta_i^n X^{(2)}}{c(\Delta_n)} - \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta_n}{c(\Delta_n)}\Delta_i^n X^{(1)} \Delta_i^n X^{(2)}\bigg|\mathscr{H}\right], \quad (65)$$

$$G_n^{(2)} = \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta_n}{c(\Delta_n)}\Delta_i^n X^{(1)} \Delta_i^n X^{(2)}\bigg|\mathscr{H}\right] - \int_0^1 \sigma_l^{(1)}\sigma_l^{(2)}\rho_l \, dl.$$

Proposition (5.12) easily extends to this situation and we can show that $G_n^{(2)} \overset{\text{a.s.}}{\to} 0$, as $n \to \infty$:

29

Proposition 5.15. *Let $\sigma^{(1)}, \sigma^{(2)}, \rho$ have Hölder continuous sample paths, almost surely. Let π_n be a sequence of probability measures converging weakly to the probability π. Then, the almost sure convergence of the Riemann sums*

$$f_n(\omega, x) := \sum_{i=1}^{n} \rho_{i\Delta_n - x} \sigma^{(1)}_{i\Delta_n - x} \sigma^{(2)}_{i\Delta_n - x} \to \int_{-x}^{1-x} \sigma^{(1)}_l \sigma^{(2)}_l \rho_l \, dl =: f(\omega, x)$$

implies the convergence of the integrals:

$$\pi_n(f_n) \to \pi(f).$$

In particular, we have the almost sure convergence:

$$\sum_{i=1}^{n} \mathbb{E}\left[\frac{\Delta_n}{c(\Delta_n)} \Delta^n_i X^{(1)} \Delta^n_i X^{(2)} \bigg| \mathscr{H} \right] \overset{a.s.}{\to} \int_{\mathbb{R}^+} \left(\int_{-x}^{1-x} \sigma^{(1)}_l \sigma^{(2)}_l \rho_l \, dl \right) d\pi(x), \quad as \ n \to \infty.$$

Proof. Let $\alpha_\rho, \alpha_{\sigma^{(1)}}, \alpha_{\sigma^{(2)}}$ denote the Hölder coefficients of $\rho, \sigma^{(1)}, \sigma^{(2)}$ and set $\alpha := \min\{\alpha_\rho, \alpha_{\sigma^{(1)}}, \alpha_{\sigma^{(2)}}\}$, and $|u - t| < 1$. Then

$$\left| \sigma^{(1)}_u \sigma^{(2)}_u \rho_u - \sigma^{(1)}_t \sigma^{(2)}_t \rho_t \right| \leq \left| \sigma^{(1)}_u \sigma^{(2)}_u (\rho_u - \rho_t) + \rho_t (\sigma^{(1)}_u \sigma^{(2)}_u - \sigma^{(1)}_t \sigma^{(2)}_t) \right|$$

$$\leq \left| \sigma^{(1)}_u \sigma^{(2)}_u (\rho_u - \rho_t) + \rho_t \left(\sigma^{(1)}_u \left(\sigma^{(2)}_u - \sigma^{(2)}_t + \sigma^{(2)}_t \right) - \sigma^{(1)}_t \sigma^{(2)}_t \right) \right|$$

$$\leq |u - t|^{\alpha_\rho} |\sigma^{(1)}_u \sigma^{(2)}_u| + |\sigma^{(1)}_u||u - t|^{\alpha_{\sigma^{(2)}}} + |\sigma^{(2)}_t||u - t|^{\alpha_{\sigma^{(1)}}}$$

$$\leq |u - t|^\alpha \left(|\sigma^{(1)}_u \sigma^{(2)}_u| + |\sigma^{(1)}_u| + |\sigma^{(2)}_t| \right).$$

The error we make with a Riemann sum becomes:

$$|f_n(x) - f(x)| = \left| \Delta_n \sum_{i=1}^{n} \rho_{i\Delta_n - x} \sigma^{(1)}_{i\Delta_n - x} \sigma^{(2)}_{i\Delta_n - x} - \int_{-x}^{1-x} \rho_l \sigma^{(1)}_l \sigma^{(2)}_l \, dl \right|$$

$$\leq \sum_{i=1}^{n} \int_{-x+(i-1)\Delta_n}^{-x+i\Delta_n} |\sigma^{(1)}_l \sigma^{(2)}_l \rho_l - \sigma^{(1)}_{l^*_i} \sigma^{(2)}_{l^*_i} \rho_{l^*_i}| \, dl$$

$$\leq \sum_{i=1}^{n} \int_{-x+(i-1)\Delta_n}^{-x+i\Delta_n} |l - l^*_i|^\alpha \left(|\sigma^{(1)}_l \sigma^{(2)}_l| + |\sigma^{(1)}_l| + |\sigma^{(2)}_{l^*_i}| \right) dl$$

$$\leq \Delta_n^\alpha \left(\int_{-x}^{1-x} \left(|\sigma^{(1)}_l \sigma^{(2)}_l| + |\sigma^{(1)}_l| \right) dl + \Delta_n \sum_{i=1}^{n} |\sigma^{(2)}_{l^*_i}| \right).$$

Now we can finally prove the bound for the integrals:

$$|\pi_n(f_n) - \pi_n(f)| \leq \Delta_n^\alpha \pi_n \left(\int_{-x}^{1-x} \left(|\sigma^{(1)}_l \sigma^{(2)}_l| + |\sigma^{(1)}_l| \right) dl + \Delta_n \sum_{i=1}^{n} |\sigma^{(2)}_{l^*_i}| \right)$$

$$= \Delta_n^\alpha \pi_n \left(\int_{-x}^{1-x} \left(|\sigma^{(1)}_l \sigma^{(2)}_l| + |\sigma^{(1)}_l| \right) dl \right) + \Delta_n^\alpha \pi_n \left(\Delta_n \sum_{i=1}^{n} |\sigma^{(2)}_{l^*_i}| \right).$$

The first term converges to $\pi \left(\int_{-x}^{1-x} |\sigma_l^{(1)} \sigma_l^{(2)}| + |\sigma_l^{(1)}| \, dl \right)$, since the function $x \to \int_{-x}^{1-x} |\sigma_l^{(1)} \sigma_l^{(2)}| + |\sigma_l^{(1)}| \, dl$ is continuous and tends to zero for $x \to \infty$. Since the paths of $\sigma^{(2)}$ are almost surely Hölder continuous, Proposition 5.12 applies, and we can conclude that:

$$\pi_n \left(\Delta_n \sum_{i=1}^{n} |\sigma_{l_i^*}^{(2)}| \right) \to \pi \left(\int_{-x}^{1-x} |\sigma_l^{(2)}| \, dl \right);$$

henceforth:

$$|\pi_n(f_n) - \pi_n(f)| \to 0 \times \pi \left(\int_{-x}^{1-x} |\sigma_l^{(1)} \sigma_l^{(2)}| + |\sigma_l^{(1)}| + |\sigma_l^{(2)}| \, dl \right) = 0.$$

$\square$

Analogously to Proposition 5.14, we need to show that:

$$\mathrm{Var} \left(\frac{\Delta_n}{c(\Delta_n)} \sum_{i=1}^{n} \Delta_i^n X^{(1)} \Delta_i^n X^{(2)} \middle| \mathscr{H} \right) \to 0, \quad \text{as } n \to \infty,$$

which implies that $G_n^{(1)} \xrightarrow{L^2} 0$ as $n \to \infty$.

Proposition 5.16. *Suppose that the assumption of Theorem 3.8 hold. Then*

$$Var \left(\frac{\Delta_n}{c(\Delta_n)} \sum_{i=1}^{n} \Delta_i^n X^{(1)} \Delta_i^n X^{(2)} \middle| \mathscr{H} \right) \to 0, \quad as \ n \to \infty,$$

Sketch of Proof. We will only sketch the proof, which follows the same lines as the proof of Proposition 5.14. The basic idea is the following: Since the volatility process is càdlàg, it is bounded on compact intervals, while unbounded intervals are controlled via Assumption 3.5.

All the terms concerning intervals on the sets of the form $[(i-1)\Delta_n, i\Delta_n]$ of the proof of Proposition 5.14 are treated exactly in the same way, since we can uniformly bound σ by its maximum on $[0,1]$. When σ appears integrated on an interval of infinite length, bounded away from 0, we can write for $j \in \{1,2\}$:

$$\int_{b^{(j)}}^{\infty} \left(g^{(j)}(s+\Delta_n) - g^{(j)}(s) \right)^2 \left(\sigma_{(i-1)\Delta_n - s}^{(j)} \right)^2 ds \leq \tag{66}$$

$$\Delta_n^2 \int_{b^{(j)}}^{\infty} \left(\frac{d}{ds} g^{(j)}(s) \right)^2 \left(\sigma_{(i-1)\Delta_n - s}^{(j)} \right)^2 ds = O(\Delta_n^2), \tag{67}$$

since the integral in (67) is finite, thanks to Assumption 3.5.

$\square$

A   Background results

A.1   *U.c.p.* convergence of processes

We briefly recall the mode of convergence that we work with throughout this article: *uniform convergence on compacts in probability*, or *u.c.p.* for short.

31

Definition A.1 (*u.c.p.* convergence)**.** *The sequence of càdlàg processes $X^{(n)}$ is said to converge* uniformly on compacts in probability *to the càdlàg process X if, for all $t > 0$ and all $\varepsilon > 0$:*

$$\lim_{n \to \infty} \mathbb{P} \left(\sup_{s < t} \left| X_s^{(n)} - X_s \right| > \varepsilon \right) = 0$$

Remark A.2. *Note that the assumption that the processes are càdlàg is sufficient to ensure that the supremum is $\mathbb{P}$-measurable.*

In some cases, pointwise convergence in probability is enough to prove u.c.p. convergence of processes as can be seen in the following theorem.

Theorem A.3 (Sufficient condition for u.c.p. convergence)**.** *Let the assumption and notation of Definition A.1 prevail. Suppose that, for all t in a dense subset $D \subset \mathbb{R}^+$ we have $X_t^{(n)} \xrightarrow{\mathbb{P}} X_t$. Assume further, that the paths of $X^{(n)}$ are increasing with time and the paths of X are continuous, almost surely. Then, the (stronger) convergence $X^{(n)} \overset{u.c.p.}{\to} X$. holds.*

A.2 Expectation of the product of four stochastic integrals

Lemma A.4. *Let W denote a standard Brownian motion and let $H^{(i)}(t)$, for $i \in \{1, 2, 3, 4\}$ be deterministic functions such that $\int_0^T \left(H^{(i)}(t) \right)^2 dt < \infty$. Then, one has:*

$$\mathbb{E} \left[\int_0^T H_s^{(1)} \, dW_s \int_0^T H_s^{(2)} \, dW_s \int_0^T H_s^{(3)} \, dW_s \int_0^T H_s^{(4)} \, dW_s \right]$$

$$= \int_0^T H_s^{(1)} H_s^{(3)} \, ds \int_0^T H_s^{(2)} H_s^{(4)} \, ds + \int_0^T H_s^{(1)} H_s^{(2)} \, ds \int_0^T H_s^{(3)} H_s^{(4)} \, ds$$

$$+ \int_0^T H_s^{(1)} H_s^{(4)} \, ds \int_0^T H_s^{(2)} H_s^{(3)} \, ds. \quad (68)$$

Proof. We will prove the claim by differentiating the characteristic function. In details, we compute:

$$\left. \frac{\partial^4}{\partial \vartheta_4 \partial \vartheta_3 \partial \vartheta_2 \partial \vartheta_1} \right|_{\boldsymbol{\vartheta} = 0} \mathbb{E} \left[e^{i\vartheta_1 \int_0^T H_s^{(1)} \, dW_s + i\vartheta_2 \int_0^T H_s^{(2)} \, dW_s + \vartheta_3 \int_0^T H_s^{(3)} \, dW_s + i\vartheta_4 \int_0^T H_s^{(4)} \, dW} \right], \quad (69)$$

for $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)^\top$. The random variable

$$\vartheta_1 \int_0^T H_s^{(1)} \, dW_s + \vartheta_2 \int_0^T H_s^{(2)} \, dW_s + \vartheta_3 \int_0^T H_s^{(3)} \, dW_s + \vartheta_4 \int_0^T H_s^{(4)} \, dW_s$$

follows a Gaussian distribution with mean zero and variance:

$$\int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right)^2 \, ds.$$

Note that the variable has moments of all orders, hence our approach is justified.

The expectation in (69) equals:

$$\tilde{H}(\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4) := \exp \left(-\frac{1}{2} \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right)^2 \right) \, ds.$$

We proceed with differentiating the function $\tilde{H}$.

$$\frac{\partial \tilde{H}}{\partial \vartheta_1} = -\tilde{H} \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(1)} \, ds$$

$$\frac{\partial^2 \tilde{H}}{\partial \vartheta_2 \partial \vartheta_1} = -\tilde{H} \int_0^T H_s^{(1)} H_s^{(2)} \, ds + \tilde{H} \left[\int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(1)} \, ds \times \right.$$
$$\left. \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(2)} \, ds \right]$$

$$\frac{\partial^3 \tilde{H}}{\vartheta_3 \vartheta_2 \vartheta_1} = \tilde{H} \int_0^T H_s^{(1)} H_s^{(2)} \, ds \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(3)} \, ds$$
$$+ \tilde{H} \left[\int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(1)} \, ds \right.$$
$$\times \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(2)} \, ds$$
$$\left. \times \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(3)} \, ds \right]$$
$$+ \tilde{H} \left[\int_0^T H_s^{(1)} H_s^{(3)} \, ds \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(2)} \, ds \right.$$
$$\left. + \int_0^T H_s^{(2)} H_s^{(3)} \, ds \int_0^T \left(\vartheta_1 H_s^{(1)} + \vartheta_2 H_s^{(2)} + \vartheta_3 H_s^{(3)} + \vartheta_4 H_s^{(4)} \right) H_s^{(1)} \, ds \right].$$

And finally, since $\tilde{H}(0,0,0,0) = 1$ and $\frac{\partial \tilde{H}}{\partial \vartheta_i} \big|_{\boldsymbol{\vartheta}=0} = 0$,

$$\frac{\partial^4 \tilde{H}}{\partial \vartheta_4 \partial \vartheta_3 \partial \vartheta_2 \partial \vartheta_1} \bigg|_{\boldsymbol{\vartheta}=0} = \int_0^T H_s^{(1)} H_s^{(2)} \, ds \int_0^T H_s^{(3)} H_s^{(4)} \, ds +$$
$$\int_0^T H_s^{(1)} H_s^{(3)} \, ds \int_0^T H_s^{(2)} H_s^{(4)} \, ds + \int_0^T H_s^{(2)} H_s^{(3)} \, ds \int_0^T H_s^{(1)} H_s^{(4)} \, ds, \quad (70)$$

proving the claim. $\qquad \square$

A particularly useful consequence of Lemma A.4 is the following result:

Corollary A.5. *Let $(K_t)_{t \geq 0}$ be an a.s. bounded stochastic process, then:*

$$\mathbb{E} \left[\left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) \, dW_s^{(1)} \int_{(i-1)\Delta_n}^{i\Delta_n} g^{(2)}(i\Delta_n - s) K_s \, dW_s^{(1)} \right)^2 \bigg| \mathscr{F}^K \right]$$
$$= \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(1)}(i\Delta_n - s) \right)^2 \, ds \int_{(i-1)\Delta_n}^{i\Delta_n} \left(g^{(2)}(i\Delta_n - s) \right)^2 K_s^2 \, ds$$
$$+ 2 \left(\int_{(i-1)\Delta_n}^{i\Delta_n} g^{(1)}(i\Delta_n - s) g^{(2)}(i\Delta_n - s) K_s \, ds \right)^2.$$

Acknowledgement

References

Barndorff-Nielsen, O. E., J. M. Corcuera, and M. Podolskij (2009). "Power variation for Gaussian processes with stationary increments". In: *Stochastic Processes and Their Applications* 119.6, pp. 1845–1865 (cit. on p. 2).

— (2011). "Multipower variation for Brownian semistationary processes". In: *Bernoulli* 17.4, pp. 1159–1194 (cit. on pp. 2, 7).

— (2013). "Limit theorems for functionals of higher order differences of Brownian semistationary processes". In: *Prokhorov and contemporary probability theory*. Springer, pp. 69–96 (cit. on p. 10).

Barndorff-Nielsen, O. E., J. M. Corcuera, M. Podolskij, and J. Woerner (2009). "Bipower variation for Gaussian processes with stationary increments". In: *Journal of Applied Probability* 46.1, pp. 132–150 (cit. on p. 2).

Barndorff-Nielsen, O. E. (2016). "Gamma kernels and BSS/LSS processes". In: *Advanced Modelling in Mathematical Finance: In Honour of Ernst Eberlein*. Springer Proceedings in Mathematics & Statistics Volume 189. Springer, pp. 41–61 (cit. on p. 7).

Barndorff-Nielsen, O. E., F. E. Benth, and A. E. D. Veraart (2013). "Modelling energy spot prices by Lévy-driven Volterra processes". In: *Bernoulli* 19.3, pp. 803–845 (cit. on pp. 2, 13).

— (2017). *Ambit stochastics*. Draft of research monograph (cit. on p. 2).

Barndorff-Nielsen, O. E. and J. Schmiegel (2009). "Brownian Semistationary Processes and Volatility/Intermittency". In: *Advanced Financial Modelling*. Ed. by H. Albrecher, W. Rungaldier, and W. Schachermeyer. Radon Series on Computational and Applied Mathematics 8. Berlin: W. de Gruyter, pp. 1–26 (cit. on pp. 2, 5, 9, 13).

Basse, A. (2008). "Gaussian moving averages and semimartingales". In: *Electronic Journal of Probability* 13.39, pp. 1140–1165 (cit. on pp. 4, 5).

Basse-O'Connor, A., C. Heinrich, and M. Podolskij (2016). "On limit theory for Lévy semistationary processes". In: *ArXiv e-prints*. E-print 1604.02307 (cit. on p. 2).

Basse-O'Connor, A., R. Lachieze-Rey, and M. Podolskij (2017). "Power variation for a class of stationary increments Lévy driven moving averages". In: *Annals of Probability*. To appear (cit. on p. 2).

Bennedsen, M., A. Lunde, and M. S. Pakkanen (2017). "Hybrid scheme for Brownian semistationary processes". In: *Finance and Stochastics*. To appear (cit. on p. 6).

Billingsley, P. (2009). *Convergence of probability measures*. Vol. 493. John Wiley & Sons (cit. on p. 26).

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1989). *Regular variation*. Vol. 27. Cambridge university press (cit. on p. 6).

Cheridito, P. (2004). "Gaussian moving averages, semimartingales and option pricing". In: *Stochastic processes and their applications* 109.1, pp. 47–68 (cit. on p. 4).

Corcuera, J. M. (2012). "New Central Limit Theorems for Functionals of Gaussian Processes and their Applications". English. In: *Methodology and Computing in Applied Probability* 14.3, pp. 477–500 (cit. on p. 2).

Corcuera, J. M., E. Hedevang, M. S. Pakkanen, and M. Podolskij (2013). "Asymptotic theory for Brownian semi-stationary processes with application to turbulence". In: *Stochastic Processes and their Applications* 123.7, pp. 2552–2574 (cit. on pp. 7, 10, 13).

Corcuera, J. M., D. Nualart, J. H. Woerner, et al. (2006). "Power variation of some integral fractional processes". In: *Bernoulli* 12.4, pp. 713–735 (cit. on p. 13).

De Haan, L and S. I. Resnick (1979). "Derivatives of regularly varying functions in R d and domains of attraction of stable distributions". In: *Stochastic Processes and their Applications* 8.3, pp. 349–355 (cit. on p. 8).

Granelli, A. and A. E. D. Veraart (2017). "A central limit theorem for the realised covariation of a bivariate Brownian semistationary process". ArXiv (cit. on pp. 9, 13, 28).

Guasoni, P., M. Rásonyi, and W. Schachermayer (2008). "Consistent price systems and face-lifting pricing under transaction costs". In: *The Annals of Applied Probability* 18.2, pp. 491–520 (cit. on p. 2).

Guyon, X. and J. León (1989). "Convergence en loi des H-variations d'un processus gaussien stationnaire sur **R**". In: *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques* 25.3, pp. 265–282 (cit. on pp. 2, 10).

Knight, F. (1992). *Foundations of the predictions process*. Oxford: Clarendon (cit. on p. 4).

Márquez, J. and J. Schmiegel (2016). "Modelling turbulent time series by BSS-processes". In: *The Fascination of Probability, Statistics and their Applications: In Honour of Ole E. Barndorff-Nielsen*. Ed. by M. Podolskij, R. Stelzer, S. Thorbjørnsen, and A. E. D. Veraart. Springer, pp. 29–52 (cit. on p. 2).

Pakkanen, M. S. (2011). "Brownian semistationary processes and conditional full support". In: *International Journal of Theoretical and Applied Finance* 14.04, pp. 579–586 (cit. on p. 2).

Protter, P. (2005). *Stochastic Integration and Differential Equations*. Vol. 21. Springer (cit. on p. 4).

Veraart, A. E. and L. A. Veraart (2012). "Stochastic volatility and stochastic leverage". In: *Annals of finance* 8.2-3, pp. 205–233 (cit. on p. 7).

— (2014). "Modelling electricity day-ahead prices by multivariate Lévy semistationary processes". In: *Quantitative Energy Finance*. Springer, pp. 157–188 (cit. on p. 2).

Walsh, J. (1986). "An introduction to stochastic partial differential equations". In: *École d'Été de Probabilités de Saint Flour XIV-1984*, pp. 265–439 (cit. on p. 3).

Williams, D. (1991). *Probability with martingales*. Cambridge University Press (cit. on p. 15).