Gaussian and non-Gaussian processes of zero power variation, and related stochastic calculus

Francesco Russo *†and Frederi VIENS ‡

July 16, 2014

Abstract

We consider a class of stochastic processes X defined by $X(t) = \int_0^T G(t,s)\,dM(s)$ for $t \in [0,T]$, where M is a square-integrable continuous martingale and G is a deterministic kernel. Let m be an odd integer. Under the assumption that the quadratic variation $[M]$ of M is differentiable with $\mathbf{E}\left[|d[M](t)/dt|^m\right]$ finite, it is shown that the mth power variation

$$\lim_{\varepsilon \to 0} \varepsilon^{-1} \int_0^T ds\,(X(s+\varepsilon) - X(s))^m$$

exists and is zero when a quantity $\delta^2(r)$ related to the variance of an increment of M over a small interval of length r satisfies $\delta(r) = o\left(r^{1/(2m)}\right)$. When M is the Wiener process, X is Gaussian; the class then includes fractional Brownian motion and other Gaussian processes with or without stationary increments. When X is Gaussian and has stationary increments, δ is X's univariate canonical metric, and the condition on δ is proved to be necessary. In the non-stationary Gaussian case, when $m = 3$, the symmetric (generalized Stratonovich) integral is defined, proved to exist, and its Itô formula is established for all functions of class C^6.

KEY WORDS AND PHRASES: Power variation, martingale, calculus via regularization, Gaussian processes, generalized Stratonovich integral, non-Gaussian processes.

MSC Classification 2000: 60G07; 60G15; 60G48; 60H05.

1 Introduction

The purpose of this article is to study wide classes of processes with zero cubic variation, and more generally, zero variation of any odd order. Before summarizing our results, we give a brief historical description of the topic of p-variations, as a basis for our motivations.

1.1 Historical background

The p-variation of a function $f : [0,T] \to \mathbf{R}$ is the supremum over all the possible partitions $\{0 = t_0 < \ldots < t_N = T\}$ of $[0,T]$ of the quantity $\sum_{i=0}^{N-1} |f(t_{i+1}) - f(t_i)|^p$. The analytic monograph [10] contains an interesting study on this concept, showing that a p-variation function is the composition

*ENSTA-ParisTech. Unité de Mathématiques appliquées, 828, bd des Maréchaux, F-91120 Palaiseau (France)
†INRIA Rocquencourt Projet MathFi and Cermics Ecole des Ponts
‡Department of Statistics, Purdue University, 150 N. University St., West Lafayette, IN 47907-2067, USA

of an increasing function and a Hölder-continuous function. The analytic notion of p-variation precedes stochastic calculus and processes (see [10]).

It was rediscovered in stochastic analysis in the context of pathwise stochastic calculus, starting with $p = 2$ as in the fundamental paper [16] of H. Föllmer. Dealings with p-variations and their stochastic applications, particularly to rough path and other recent integration techniques for fractional Brownian motion (fBm) and related processes, are described at length for instance in the books [12], [17], and [24], which also contain excellent bibliographies on the subject. Prior to this, power variations could be seen as related to oscillations of processes in [4], and some specific cases had been treated, such as local time processes (see [32]).

The Itô stochastic calculus for semimartingales defines a *quadratic* variation of a semimartingale S, instead of its 2-variation, by taking the limit in probability of $\sum_{i=0}^{N-1} |S(t_{i+1}) - S(t_i)|^2$ over the smaller set of partitions whose mesh tends to 0, instead of the supremum over all partitions. One defines the quadratic variation $[S]$ of S as the limit *in probability* of the expression above when the partition mesh goes to 0, instead of considering pathwise the supremum over all partitions, in the hopes of making it more likely to have a finite limit; this is indeed the case for standard Brownian motion $M = B$, where its 2-variation $[B]$ is a.s. infinite, but its quadratic variation is equal to T. To reconcile 2-variations with the finiteness of $[B]$, many authors have proposed restricting the supremum over dyadic partitions. But there is a fundamental difference between the deterministic and stochastic versions of "variation", since in Itô calculus, quadratic variation is associated with the notion of covariation (also known as joint quadratic variation), something which is not present in analytic treatments of 2-variation. The co-variation $[S^1, S^2]$ of two semimartingales S^1, S^2 is obtained by polarization, using again a limit in probability when the partition mesh goes to zero.

To work with a general class of processes, the tools of Itô calculus would nonetheless restrict the study of covariation to semimartingales. In [35], the authors enlarged the notion of covariation to general processes, in an effort to create a more efficient stochastic calculus tool to go beyond semimartingales, by considering regularizations instead of discretizations. Drawing some inspiration from the classical fact that a continuous $f : [0, T] \to \mathbf{R}$ has finite variation (1-variation) if and only if $\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^T |f(s + \varepsilon) - f(s)| ds$ exists, for two processes X and Y, their covariation $[X, Y](t)$ is the limit in probability, when ε goes to zero, of

$$[X, Y]_\varepsilon (t) = \varepsilon^{-1} \int_0^t \big(X(s + \varepsilon) - X(s)\big)\big(Y(s + \varepsilon) - Y(s)\big) ds; \quad t \geq 0. \tag{1}$$

$[X, Y]$ coincides with the classical covariation for continuous semimartingales. The processes X such that $[X, X]$ exists are called finite quadratic variation processes, and were analyzed in [15, 34].

The notion of covariation was extended in [14] to more than two processes: the n-covariation $[X^1, X^2, \cdots, X^n]$ of n processes $X^1, \ldots, X^n$ is as in formula (1), but with a product of n increments, with specific analyses for $n = 4$ for fBm with "Hurst" parameter $H = 1/4$ in [19]. If $X = X^1 = X^2 = X^3$ we denote $[X; 3] := [X, X, X]$, which is called the *cubic variation*, and is one of the main topics of investigation in our article. This variation is the limit in probability of

$$[X, 3]_\varepsilon (t) := \varepsilon^{-1} \int_0^t (X(s + \varepsilon) - X(s))^3 ds, \tag{2}$$

when $\varepsilon \to 0$. (2) involves the signed cubes $(X(s + \varepsilon) - X(s))^3$, which has the same sign as the increment $X(s + \varepsilon) - X(s)$, unlike the case of quadratic or 2-variation, or of the so-called *strong* cubic variation, where absolute values are used inside the cube function. Consider the case where X is a fBm B^H with Hurst parameter $H \in (0, 1)$. For fBm, [20] establish that $[X, 3] \equiv 0$ if $H > 1/6$ and $[X, 3]$ does not exist if $H < 1/6$, while for $H = 1/6$, the regularization approximation $[X, 3]_\varepsilon (t)$

converges in law to a normal law for every $t > 0$. This phenomenon was confirmed for the related finite-difference approximating sequence of $[X,3](t)$ which also converges in law to a Gaussian variable; this was proved in [31, Theorem 10] by using the the so-called Breuer-Major central limit theorem for stationary Gaussian sequences [8].

A practical significance of the cubic variation lies in its well-known ability to guarantee the existence of (generalized symmetric) Stratonovich integrals, and their associated Itô-Stratonovich formula, for various highly irregular processes. This was established in [20] in significant generality; technical conditions therein were proved to apply to fBm with $H > 1/6$, and can extend to similar Gaussian cases with canonical metrics that are bounded above and below by multiples of the fBm's, for instance the bi-fractional Brownian motion treated in [33]. A variant on [20]'s Itô formula was established previously in [14] for less irregular processes: if X (not necessarily Gaussian) has a finite strong cubic variation, so that $[X,3]$ exists (but may not be zero), for $f \in C^3(\mathbf{R})$, $f(X_t) = f(X_0) + \int_0^t f'(X_s) d^\circ X - \frac{1}{12}\int_0^t f'''(X_s) d[X,3](s)$, which involves the symmetric-Stratonovich integral of [36], and a Lebesgue-Stieltjes integral. In [29], an analogous formula is obtained for fBm with $H = 1/6$, but in the sense of distribution laws only: $\int_0^t f'(X_s) d^\circ X$ exist only in law, and $\int_0^t f'''(X_s) d[X,3](s)$ is replaced by a conditionally Wiener integral defined in law by replacing $[X,3]$ with a term κW, where W is the independent Wiener process identified in [31].

1.2 Specific motivations

Our work herein is motivated by the properties described in the previous paragraph, particularly as in [20]. We want to avoid situations where Itô formulas can only be established in law, i.e. involving conditionally Wiener integrals defined as limits in a weak sense. Thus we study scales where this term vanishes in a strong sense, while staying as close to the threshold $H = 1/6$ as possible. Other types of stochastic integrals for fBm and related irregular Gaussian processes make use of the Skorohod integral, identified as a divergence operator on Wiener space (see [30] and also [3, 6, 11, 25, 22]), and rough path theory (see [17, 24]). The former method is not restrictive in how small H can be (see [25]), but is known not to represent a pathwise notion of integral; the latter is based in a true pathwise strategy and it is based on giving a Lévy-type area or iterated integrals *a priori*. In principal the objective of the rough path approach is not to link any discretization (or other approximation) scheme. These provide additional motivations for studying the regularlization methodology of [35] or [36], which does not feature these drawbacks for $H > 1/6$.

We come back to the cubic variation approximation $[X,3]$ defined via the limit of (2). The reasons for which $[X,3] = 0$ for fBm with $H > 1/6$, which is considerably less regular than the threshold $H > 1/3$ one has for H-Hölder-continuous deterministic functions, are the odd symmetry of the cube function, and the accompanying probabilistic symmetries of the process X itself (e.g. Gaussian property). This doubling improvement over the deterministic case does not typically hold for non-symmetric variations: H needs to be larger to guarantee existence of the variation; for instance, when X is fBm, its strong cubic variation, defined as the limit in probability of $\varepsilon^{-1}\int_0^t |X(s+\varepsilon) - X(s)|^3 ds$, exists for $H \geq 1/3$ only.

Finally, some brief notes in the case where X is fBm with $H = 1/6$. This threshold is a critical value since, as mentioned above, whether in the sense of regularization or of finite-difference, the approximating sequences of $[X,3](t)$ converge in law to Gaussian laws. In contrast to these normal convergences, in our article, we show as a preliminary result (Proposition 2 herein), that $[X,3]_\varepsilon$ does not converge in probability for $H = 1/6$; the non-convergence of $[X,3]_\varepsilon$ in probability for $H < 1/6$ was known previously.

1.3 Summary of results and structure of article

This article investigates the properties of cubic and other odd power variations for processes which may not be self-similar, or have stationary increments, or be Gaussian, when they have α-Hölder-continuous paths; this helps answer to what extent the threshold $\alpha > 1/6$ is sharp for $[X, 3] = 0$. We consider processes X defined on $[0, T]$ by a Volterra representation

$$X(t) = \int_0^T G(t, s)\, dM(s), \tag{3}$$

where M is a square-integrable martingale on $[0, T]$, and G is a non-random measurable function on $[0, T]^2$, which is square-integrable in s with respect to $d[M]_s$ for every fixed t. The quadratic variations of these martingale-based convolutions was studied in [13]. The "Gaussian" case results when M is the standard Wiener process (Brownian motion) W.

In this paper, we concentrate on processes X which are not more regular than standard Brownian motion; this irregularity is expressed via a concavity condition on the squared canonical metric $\delta^2(s, t) = \mathbf{E}\left[\left(X(t) - X(s)^2\right)\right]$. It is not a restriction since the main interest of our results occurs around the Hölder exponent $1/(2m)$ for odd $m \geq 3$, and processes which are more regular than Brownian motion can be treated using classical non-probabilistic tools such as the Young integral.

After providing some definitions [Section 2], our first main finding is that the processes with zero odd mth variation (same definition as for $[X, 3] = 0$ in (2) but with m replacing 3) are those which are better than $1/(2m)$-Hölder-continuous in the $L^2(\Omega)$-sense, whether for Gaussian processes [Section 3], or non-Gaussian ones [Section 4]. Specifically,

- for X *Gaussian with stationary increments* (i.e. $\delta(s, t) = \delta(t - s)$), for any odd integer $m \geq 3$, $[X, m] = 0$ if and only if $\delta(r) = o\left(r^{1/(2m)}\right)$ for r near 0 [Theorem 6 on page 8];
- for X *Gaussian* but not necessarily with stationary increments, for any odd integer $m \geq 3$, $[X, m] = 0$ if $\delta^2(s, s + r) = o\left(r^{1/(2m)}\right)$ for r near 0 uniformly in s. [Theorem 8 on page 12; this holds under a technical non-explosion condition on the mixed partial derivative of δ^2 near the diagonal; see Examples 9 and 10 on page 12 for a wide class of Volterra-convolution-type Gaussian processes with non-stationary increments which satisfy the condition].
- for X *non-Gaussian* based on a martingale M, for any odd integer $m \geq 3$, with $\Gamma(t) := \left(\mathbf{E}\left[(d[M]/dt)^m\right]\right)^{1/(2m)}$ if it exists, we let $Z(t) := \int_0^T \Gamma(s)\, G(t, s)\, dW(s)$. This Z is a Gaussian process; if it satisfies the conditions of Theorem 6 or Theorem 8, then $[X, m] = 0$. [Section 4, Theorem 11 on page 13; Proposition 12 on page 13 provides examples of wide classes of martingales and kernels for which the assumptions of Theorem 11 are satisfied, with details on how to construct examples and study their regularity properties on page 4.].

Our results shows how broad a class of processes, based on martingale convolutions with only m moments, one can construct which have zero odd mth variation, under conditions which are the same in terms of regularity as in the case of Gaussian processes with stationary increments, where we prove sharpness. Note that X itself can be far from having the martingale property, just as it is generally far from standard Brownian motion in the Gaussian case. Our second main result is an application to weighted variations, generalized Stratonovich integration, and an Itô formula [Section 5.]

- Under the conditions of Theorem 8 (general Gaussian case), and an additional coercivity condition, for every bounded measurable function g on $\mathbf{R}$,

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^2}\mathbf{E}\left[\left(\int_0^T du\,(X_{u+\varepsilon}-X_u)^m\,g\left(\frac{X_{u+\varepsilon}+X_u}{2}\right)\right)^2\right]=0.$$

  If $m=3$, by results in [20], Theorem 15 implies that for any $f\in C^6(\mathbf{R})$ and $t\in[0,T]$, the Itô formula $f(X_t)=f(X_0)+\int_0^t f'(X_u)\,d^\circ X_u$ holds, where the integral is in the symmetric (generalized Stratonovich) sense. [Theorem 15 and its Corollary 16, on page 16.]

Most of the proofs of our theorems are relegated to the Appendix [Section 6].

1.4 Relation with other recent work

The authors of the paper [21] consider, as we do, stochastic processes which can be written as Volterra integrals with respect to martingales. Their "fractional martingale", which generalizes Riemann-Liouville fBm, is a special case of the processes we consider in Section 4, with $K(t,s)=(t-s)^{H-1/2}$. The authors' motivation is to prove an analogue of the famous Lévy characterization of Brownian motion as the only continuous square-integrable martingale with a quadratic variation equal to t. They provide similar necessary and sufficient conditions based on the $1/H$-variation for a process to be fBm. This is a different aspect of the theory than our motivation to study necessary and sufficient conditions for a process to have vanishing (odd) cubic variation, and its relation to stochastic calculus. The value $H=1/6$ is mentioned in the context of the stochastic heat equation driven by space-time white-noise, in which discrete trapezoidal sums converge in distribution (not in probability) to a conditionally independent Brownian motion: see [9] and [31].

To find a similar motivation to ours, one may look at the recent result of [28], where the authors study the central and non-central behavior of weighted Hermite variations for fBm. Using the Hermite polynomial of order m rather than the power-m function, they show that the threshold value $H=1/(2m)$ poses an interesting open problem, since above this threshold (but below $H=1-1/(2m)$) one obtains Gaussian limits (these limits are conditionally Gaussian when weights are present, and can be represented as Wiener integrals with respect to an independent Brownian motion), while below the threshold, degeneracy occurs. The behavior at the threshold was worked out for $H=1/4, m=2$ in [28], boasting an exotic correction term with an independent Brownian motion, while the general open problem of Hermite variations with $H=1/(2m)$ was settled in [27]. More questions arise, for instance, with a similar result in [26] for $H=1/4$, but this time with bidimensional fBm, in which two independent Brownian motions are needed to characterize the exotic correction term. Compared to the above works, our work situates itself by

- establishing necessary and sufficient conditions for nullity of the mth power variation, around the threshold regularity value $H=1/(2m)$, for general Gaussian processes with stationary increments, showing in particular that self-similarity is not related to this nullity, and that the result works for all odd integers, thanks only to the problem's symmetries;

- showing that our method is able to consider processes that are far from Gaussian and still yield sharp sufficient conditions for nullity of odd power variations, since our base noise may be a generic martingale with only a few moments; our ability to prove an Itô formula for such processes attests to our method's power.

2 Definitions

We recall our process X defined for all $t \in [0,T]$ by (3). For any integer $m \geq 2$, let the *odd ε-m-th variation* of X be defined by

$$[X,m]_\varepsilon (T) := \frac{1}{\varepsilon} \int_0^T ds \left| X(s+\varepsilon) - X(s) \right|^m \operatorname{sgn}(X(s+\varepsilon) - X(s)). \tag{4}$$

The odd variation is different from the absolute (or strong) variation because of the presence of the sign function, making the function $|x|^m \operatorname{sgn}(x)$ an odd function. In the sequel, in order to lighten the notation, we will write $(x)^m$ for $|x|^m \operatorname{sgn}(x)$. We say that X has *zero odd m-th variation* (in the mean-squared sense) if the limit $\lim_{\varepsilon \to 0} [X,m]_\varepsilon (T) = 0$ holds in $L^2(\Omega)$.

The *canonical metric* δ of a stochastic process X is defined as the pseudo-metric on $[0,T]^2$ given by $\delta^2(s,t) = \mathbf{E}\left[(X(t) - X(s))^2\right]$. The *covariance function* of X is defined by $Q(s,t) = \mathbf{E}[X(t)X(s)]$. The special case of a centered Gaussian process is of primary importance; then the process's entire distribution is characterized by Q, or alternately by δ and the variances $var(X(t)) = Q(t,t)$, since we have $Q(s,t) = \frac{1}{2}\left(Q(s,s) + Q(t,t) - \delta^2(s,t)\right)$. We say that δ has *stationary increments* if there exists a function on $[0,T]$ which we also denote by δ such that $\delta(s,t) = \delta(|t-s|)$. Below, we will refer to this situation as the *stationary case*. This is in contrast to usual usage of this appellation, which is stronger, since for example in the Gaussian case, it refers to the fact that $Q(s,t)$ depends only on the difference $s-t$; this would not apply to, say, standard or fBm, while our definition does. In non-Gaussian settings, the usual way to interpret the "stationary" property is to require that the processes $X(t+\cdot)$ and $X(\cdot)$ have the same law, which is typically much more restrictive than our definition.

The goal of the next two sections is to define various general conditions under which a characterization of $\lim_{\varepsilon \to 0} [X,m]_\varepsilon (T) = 0$ can be established. In particular, we aim to show that X has zero odd m-th variation for well-behaved M's and G's if – and in some cases only if –

$$\delta(s,t) = o(|t-s|^{1/(2m)}). \tag{5}$$

3 Gaussian case

We assume that X is centered Gaussian. Then we can write X as in formula (3) with $M = W$ a standard Brownian motion. We have the following elementary result.

Lemma 1 *If m is an odd integer ≥ 3, we have $\mathbf{E}\left[([X,m]_\varepsilon (T))^2\right] = \sum_{j=0}^{(m-1)/2} J_j$ where*

$$J_j := \frac{1}{\varepsilon^2} \sum_{j=0}^{(m-1)/2} c_j \int_0^T \int_0^t dtds \Theta^\varepsilon (s,t)^{m-2j} \; Var\left[X(t+\varepsilon) - X(t)\right]^j \; Var\left[X(s+\varepsilon) - X(s)\right]^j,$$

the c_j's are positive constants depending only on j, and

$$\Theta^\varepsilon (s,t) := \mathbf{E}\left[(X(t+\varepsilon) - X(t))(X(s+\varepsilon) - X(s))\right].$$

Using Q and δ, $\Theta^\varepsilon (s,t)$ computes as the opposite of the planar increment of the canonical metric over the rectangle defined by its corners (s,t) and $(s+\varepsilon, t+\varepsilon)$:

$$\Theta^\varepsilon (s,t) = \frac{1}{2}\left[-\delta^2(t+\varepsilon, s+\varepsilon) + \delta^2(t, s+\varepsilon) + \delta^2(s, t+\varepsilon) - \delta^2(s,t)\right] =: -\frac{1}{2}\Delta_{(s,t);(s+\varepsilon,t+\varepsilon)}\delta^2. \tag{6}$$

3.1 The case of critical fBm

Before finding sufficient and possibly necessary conditions for various Gaussian processes to have zero cubic (or mth) variation, we discuss the threshold case for the cubic variation of fBm. Recall that when X is fBm with parameter $H = 1/6$, as mentioned in the Introduction, it is known from [20, Theorem 4.1 part (2)] that $[X,3]_\varepsilon(T)$ converges in distribution to a non-degenerate normal law. However, there does not seem to be any place in the literature specifying whether the convergence may be any stronger than in distribution. We address this issue here.

Proposition 2 *Let X be an fBm with Hurst parameter $H = 1/6$. Then X does not have a cubic variation (in the mean-square sense), by which we mean that $[X,3]_\varepsilon(T)$ has no limit in $L^2(\Omega)$ as $\varepsilon \to 0$. In fact more is true: $[X,3]_\varepsilon(T)$ has no limit in probability as $\varepsilon \to 0$.*

In order to prove the proposition, we study the Wiener chaos representation and moments of $[X,3]_\varepsilon(T)$ when X is fBm; X is given by (3) where W is Brownian motion and the kernel G is well-known (see Chapters 1 and 5 of the textbook [30]).

Lemma 3 *Fix $\varepsilon > 0$. Let $\Delta G_s(u) := G(s+\varepsilon, u) - G(s,u)$. Then $[X,3]_\varepsilon(T) = \mathcal{I}_1 + \mathcal{I}_3$ where*

$$\mathcal{I}_1 := \frac{3}{\varepsilon}\int_0^T ds \int_0^T \Delta G_s(u)\, dW(u) \left(\int_0^T |\Delta G_s(v)|^2\, dv\right), \tag{7}$$

$$\mathcal{I}_3 := \frac{6}{\varepsilon}\int_0^T dW(s_3)\int_0^{s_3} dW(s_2)\int_0^{s_2} dW(s_1)\int_0^T \left[\prod_{k=1}^{3}\Delta G_s(s_k)\right] ds. \tag{8}$$

The above lemma indicates the Wiener chaos decomposition of $[X,3]_\varepsilon(T)$ into the term $\mathcal{I}_1$ of line (7) which is in the first Wiener chaos (i.e. a Gaussian term), and the term $\mathcal{I}_3$ of line (8), in the third Wiener chaos. The next two lemmas contain information on the behavior of each of these two terms, as needed to prove Proposition 2.

Lemma 4 *$\mathcal{I}_1$ converges to 0 in $L^2(\Omega)$ as $\varepsilon \to 0$.*

Lemma 5 *$\mathcal{I}_3$ is bounded in $L^2(\Omega)$ for all $\varepsilon > 0$, and does not converge in $L^2(\Omega)$ as $\varepsilon \to 0$.*

Proof of Proposition 2. We prove the proposition by contradiction. Assume $[X,3]_\varepsilon(T)$ converges in probability. For any $p > 2$, there exists c_p depending only on p such that $\mathbf{E}[|\mathcal{I}_1|^p] \le c_p\left(\mathbf{E}\left[|\mathcal{I}_1|^2\right]\right)^{p/2}$ and $\mathbf{E}[|\mathcal{I}_3|^p] \le c_p\left(\mathbf{E}\left[|\mathcal{I}_3|^2\right]\right)^{p/2}$; this is a general fact about random variables in fixed Wiener chaos, and can be proved directly using Lemma 3 and the Burkhölder-Davis-Gundy inequalities. Also see [7]. Therefore, since we have $\sup_{\varepsilon>0}(\mathbf{E}\left[|\mathcal{I}_1|^2\right] + \mathbf{E}\left[|\mathcal{I}_3|^2\right]) < \infty$ by Lemmas 4 and 5, we also get $\sup_{\varepsilon>0}(\mathbf{E}[|\mathcal{I}_1 + \mathcal{I}_3|^p]) < \infty$ for any p. Therefore, by uniform integrability, $[X,3]_\varepsilon(T) = \mathcal{I}_1 + \mathcal{I}_3$ converges in $L^2(\Omega)$. In $L^2(\Omega)$, the terms $\mathcal{I}_1$ and $\mathcal{I}_3$ are orthogonal. Therefore, $\mathcal{I}_1$ and $\mathcal{I}_3$ must converge in $L^2(\Omega)$ separately. This contradicts the non-convergence of $\mathcal{I}_3$ in $L^2(\Omega)$ obtained in Lemma 5. Thus $[X,3]_\varepsilon(T)$ does not converge in probability. ■

3.2 The case of stationary increments

We prove a necessary and sufficient condition for having a zero odd m-th variation for Gaussian processes with stationary increments.

Theorem 6 *Let $m > 1$ be an odd integer. Let X be a centered Gaussian process on $[0,T]$ with stationary increments; its canonical metric is*

$$\delta^2(s,t) := \mathbf{E}\left[(X(t) - X(s))^2\right] = \delta^2(|t-s|)$$

where the univariate function δ^2 is assumed to be increasing and concave on $[0,T]$. Then X has zero mth variation if and only if $\delta(r) = o\left(r^{1/(2m)}\right)$.

Proof. *Step 0: setup.* The derivative $d\delta^2$ of δ^2, in the sense of measures, is positive and bounded on $[0,T]$. By stationarity, $Var[X(t+\varepsilon) - X(t)] = \delta^2(\varepsilon)$. Using the notation in Lemma 1, we get

$$J_j = \varepsilon^{-2}\delta^{4j}(\varepsilon)\, c_j \int_0^T dt \int_0^t ds \Theta^\varepsilon(s,t)^{m-2j}.$$

Step 1: diagonal. We define the ε-diagonal $D_\varepsilon := \{0 \le t-\varepsilon < s < t \le T\}$. Trivially using the Cauchy-Schwarz's inequality,

$$|\Theta^\varepsilon(s,t)| \le \sqrt{Var[X(t+\varepsilon) - X(t)]\, Var[X(s+\varepsilon) - X(s)]} = \delta^2(\varepsilon).$$

Hence, according to Lemma 1, the diagonal portion $\sum_{j=0}^{(m-1)/2} J_{j,D_\varepsilon}$ of $\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right]$ can be bounded above, in absolute value, as:

$$\begin{aligned}\left|\sum_{j=0}^{(m-1)/2} J_{j,D_\varepsilon}\right| &:= \left|\sum_{j=0}^{(m-1)/2} \varepsilon^{-2}\delta^{4j}(\varepsilon)\, c_j \int_\varepsilon^T dt \int_{t-\varepsilon}^t ds \Theta^\varepsilon(s,t)^{m-2j}\right| \\ &\le \frac{1}{\varepsilon^2}\sum_{j=0}^{(m-1)/2} c_j \int_\varepsilon^T dt \int_{t-\varepsilon}^t ds \delta^{2m}(\varepsilon) \le cst \cdot \varepsilon^{-1}\delta^{2m}(\varepsilon)\end{aligned}$$

where *cst* denotes a constant (here depending only on δ and m) whose value may change in the remainder of the article's proofs. The hypothesis on δ^2 implies that the above converges to 0 as ε tends to 0.

Step 2: small t term . The term for $t \in [0,\varepsilon]$ and any $s \in [0,t]$ can be dealt with similarly, and is of a smaller order than the one in Step 1. Specifically we have

$$|J_{j,S}| := \varepsilon^{-2}\delta^{4j}(\varepsilon)\, c_j \left|\int_0^\varepsilon dt \int_0^t ds \Theta^\varepsilon(s,t)^{m-2j}\right| \le \varepsilon^{-2}\delta^{4j}(\varepsilon)\, c_j \delta^{2(m-2j)}(\varepsilon)\, \varepsilon^2 = c_j \delta^{2m}(\varepsilon),$$

which converges to 0 like $o(\varepsilon)$.

Step 3: off-diagonal. By stationarity, from (6), for any s,t in the ε-off diagonal set $OD_\varepsilon := \{0 \le s < t-\varepsilon < t \le T\}$,

$$\begin{aligned}\Theta^\varepsilon(s,t) &= \left(\delta^2(t-s+\varepsilon) - \delta^2(t-s)\right) - \left(\delta^2(t-s) - \delta^2(t-s-\varepsilon)\right) \\ &= \int_{t-s}^{t-s+\varepsilon} d\delta^2(r) - \int_{t-s-\varepsilon}^{t-s} d\delta^2(r). \qquad (9)\end{aligned}$$

By the concavity of δ^2, we see that $\Theta^\varepsilon(s,t)$ is negative in OD_ε. According to Lemma 1, the off-diagonal portion $\sum_{j=0}^{(m-1)/2} J_{j,OD_\varepsilon}$ of $\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right]$ is precisely equal to,

$$\sum_{j=0}^{(m-1)/2} J_{j,OD_\varepsilon} := \sum_{j=0}^{(m-1)/2} \varepsilon^{-2}\delta^{4j}(\varepsilon)\, c_j \int_\varepsilon^T dt \int_0^{t-\varepsilon} ds \Theta^\varepsilon(s,t)^{m-2j}.$$

The negativity of Θ^εon OD_ε, odd power $m-2j$, and positivity of all other factors above implies that the entire off-diagonal contribution to $\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right]$ is negative. Combining this with the results of Steps 1 and 2, we obtain that

$$\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right] \leq cst \cdot \varepsilon^{-1}\delta^{2m}(2\varepsilon)$$

which implies the sufficient condition in the theorem.

Step 4: necessary condition. The proof of this part is more delicate than the above: it requires an excellent control of the off-diagonal term, since it is negative and turns out to be of the same order of magnitude as the diagonal term. We spell out the proof here for $m=3$. The general case is similar, and is left to the reader.

Step 4.1: positive representation. The next elementary lemma (see the product formula in [30, Prop. 1.1.3], or [23, Thm 9.6.9]) uses the following chaos integral notation: for any $n \in \mathbf{N}$, for $g \in L^2([0,T]^n)$, g symmetric in its n variables, then $I_n(g)$ is the multiple Wiener integral of g over $[0,T]^n$ with respect to W.

Lemma 7 *Let $f \in L^2([0,T])$. Then $I_1(f)^3 = 3|f|^2_{L^2([0,T])} I_1(f) + I_3(f \otimes f \otimes f)$*

Using this lemma, as well as definitions (3) and (4), recalling the notation $\Delta G_s(u) := G(s+\varepsilon,u) - G(s,u)$ already used in Lemma 3, and exploiting the fact that the covariance of two multiple Wiener integrals of different orders is 0, we can write

$$\mathbf{E}\left[([X,3]_\varepsilon(T))^2\right] = \frac{9}{\varepsilon^2}\int_0^T ds \int_0^T dt \mathbf{E}\left[I_1(\Delta G_s) I_1(\Delta G_t)\right] |\Delta G_s|^2_{L^2([0,T])} |\Delta G_t|^2_{L^2([0,T])}$$
$$+ \frac{1}{\varepsilon^2}\int_0^T ds \int_0^T dt \mathbf{E}\left[I_3\left((\Delta G_s)^{\otimes 3}\right) I_3\left((\Delta G_t)^{\otimes 3}\right)\right].$$

Now we use the fact that $\mathbf{E}\left[I_3(h) I_3(\ell)\right] = \langle h, \ell\rangle_{L^2([0,T]^3)}$, plus the fact that in our stationary situation $|\Delta G_s|^2_{L^2([0,T])} = \delta^2(\varepsilon)$ for any s. Hence the above equals

$$\frac{9\delta^4(\varepsilon)}{\varepsilon^2}\int_0^T ds \int_0^T dt \langle \Delta G_s, \Delta G_t\rangle_{L^2([0,T])} + \frac{1}{\varepsilon^2}\int_0^T ds \int_0^T dt \left\langle (\Delta G_s)^{\otimes 3}, (\Delta G_t)^{\otimes 3}\right\rangle_{L^2([0,T]^3)}$$
$$= \frac{9\delta^4(\varepsilon)}{\varepsilon^2}\int_0^T ds \int_0^T dt \int_0^T du \Delta G_s(u)\Delta G_t(u) + \frac{1}{\varepsilon^2}\int_0^T ds \int_0^T dt \iiint_{[0,T]^3} \prod_{i=1}^3 (du_i \Delta G_s(u_i)\Delta G_t(u_i))$$
$$= \frac{9\delta^4(\varepsilon)}{\varepsilon^2}\int_0^T du \left|\int_0^T ds \Delta G_s(u)\right|^2 + \frac{1}{\varepsilon^2}\iiint_{[0,T]^3} du_1\, du_2\, du_3 \left|\int_0^T ds \prod_{i=1}^3 (\Delta G_s(u_i))\right|^2.$$

Step 4.2: J_1 as a lower bound. The above representation is extremely useful because it turns out, as one readily checks, that of the two summands in the last expression above, the first is what we called J_1 and the second is J_0, and we can now see that both these terms are positive, which was not at all obvious before, since, as we recall, the off-diagonal contribution to either term is negative by our concavity assumption. Nevertheless, we may now have a lower bound on the ε-variation by finding a lower bound for the term J_1 alone. Reverting to our method of separating diagonal and off-diagonal terms, and recalling by Step 2 that we can restrict $t \geq \varepsilon$, we have

$$\begin{aligned} J_1 &= \frac{9\delta^4(\varepsilon)}{\varepsilon^2} 2 \int_\varepsilon^T dt \int_0^t ds \int_0^T du \Delta G_s(u) \Delta G_t(u) = \frac{9\delta^4(\varepsilon)}{\varepsilon^2} 2 \int_\varepsilon^T dt \int_0^t ds \Theta_\varepsilon(s,t) \\ &= \frac{9\delta^4(\varepsilon)}{\varepsilon^2} \int_\varepsilon^T dt \int_0^t ds \left(\delta^2(t-s+\varepsilon) - \delta^2(t-s) - \left(\delta^2(t-s) - \delta^2(|t-s-\varepsilon|)\right)\right) \\ &= J_{1,D} + J_{1,OD} \end{aligned}$$

where, performing the change of variables $t - s \mapsto s$

$$\begin{aligned} J_{1,D} &:= \frac{9\delta^4(\varepsilon)}{\varepsilon^2} \int_\varepsilon^T dt \int_0^\varepsilon ds \left(\delta^2(s+\varepsilon) - \delta^2(s) - \left(\delta^2(s) - \delta^2(\varepsilon - s)\right)\right) \\ J_{1,OD} &:= \frac{9\delta^4(\varepsilon)}{\varepsilon^2} \int_\varepsilon^T dt \int_\varepsilon^t ds \left(\delta^2(s+\varepsilon) - \delta^2(s) - \left(\delta^2(s) - \delta^2(s - \varepsilon)\right)\right). \end{aligned}$$

Step 4.3: Upper bound on $|J_{1,OD}|$. We rewrite the planar increments of δ^2 as in (9) to show what cancellations occur: with the change of variable $s' := t - s - \varepsilon$, we get $-\Theta^\varepsilon(s,t) = -\int_{s'}^{s'+\varepsilon} d\delta^2(r) + \int_{s'-\varepsilon}^{s'} d\delta^2(r)$, and

$$\begin{aligned} \int_\varepsilon^T dt \int_0^{t-\varepsilon} ds \left(-\Theta^\varepsilon(s,t)\right) &= \int_\varepsilon^T dt \left[\int_\varepsilon^t ds' \int_{s'-\varepsilon}^{s'} d\delta^2(r) - \int_\varepsilon^t ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \\ &= \int_\varepsilon^T dt \left[\int_0^{t-\varepsilon} ds'' \int_{s''}^{s''+\varepsilon} d\delta^2(r) - \int_\varepsilon^t ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \\ &= \int_\varepsilon^T dt \left[\int_0^{\varepsilon} ds'' \int_{s''}^{s''+\varepsilon} d\delta^2(r) - \int_{t-\varepsilon}^t ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \end{aligned}$$

where we also used the change $s'' := s' - \varepsilon$. Thus we have

$$J_{1,OD} = \frac{9\delta^4(\varepsilon)}{\varepsilon^2} \int_\varepsilon^T dt \left[\int_{t-\varepsilon}^t ds \int_s^{s+\varepsilon} d\delta^2(r) - \int_0^\varepsilon ds \int_s^{s+\varepsilon} d\delta^2(r)\right] =: K_1 + K_2.$$

We can already see that $K_1 \geq 0$ and $K_2 \leq 0$, so it is only necessary to find an upper bound on $|K_2|$; by Fubini on (r,s), the integrand in K_2 is calculated as

$$\int_0^\varepsilon ds \int_s^{s+\varepsilon} d\delta^2(r) = -\int_0^\varepsilon \delta^2(r)\, dr + \int_\varepsilon^{2\varepsilon} \delta^2(r)\, dr.$$

In particular, because $|K_1| \ll |K_2|$ and δ^2 is increasing, we get

$$|J_{1,OD}| \leq \frac{9\left(T-\varepsilon\right)\delta^4\left(\varepsilon\right)}{\varepsilon^2}\left(\int_{\varepsilon}^{2\varepsilon}\delta^2\left(r\right)dr-\int_0^{\varepsilon}\delta^2\left(r\right)dr\right). \tag{10}$$

Step 4.4: Lower bound on $J_{1,D}$. Note first that

$$\int_0^{\varepsilon} ds\left(\delta^2\left(s\right)-\delta^2\left(\varepsilon-s\right)\right)=\int_0^{\varepsilon} ds\ \delta^2\left(s\right)-\int_0^{\varepsilon} ds\ \delta^2\left(\varepsilon-s\right)=0.$$

Therefore

$$J_{1,D}=\frac{9\delta^4\left(\varepsilon\right)}{\varepsilon^2}\int_{\varepsilon}^{T}dt\int_0^{\varepsilon}ds\left(\delta^2\left(s+\varepsilon\right)-\delta^2\left(s\right)\right)=\frac{9\delta^4\left(\varepsilon\right)}{\varepsilon^2}\left(T-\varepsilon\right)\int_0^{\varepsilon}ds\int_s^{s+\varepsilon}d\delta^2\left(r\right).$$

We can also perform a Fubini on the integral in $J_{1,D}$, easily obtaining

$$J_{1,D}=\frac{9\delta^4\left(\varepsilon\right)}{\varepsilon^2}\left(T-\varepsilon\right)\left(\varepsilon\delta^2\left(2\varepsilon\right)-\int_0^{\varepsilon}\delta^2\left(r\right)dr\right).$$

Step 4.5: conclusion. We may now compare $J_{1,D}$ and $|J_{1,OD}|$: by the results of Steps 4.1 and 4.2,

$$\begin{aligned}J_1=J_{1,D}-|J_{1,OD}|&\geq\frac{9\delta^4\left(\varepsilon\right)}{\varepsilon^2}\left(T-\varepsilon\right)\left(\varepsilon\delta^2\left(2\varepsilon\right)-\int_0^{\varepsilon}\delta^2\left(r\right)dr\right)\\&-\frac{9\delta^4\left(\varepsilon\right)}{\varepsilon^2}\left(T-\varepsilon\right)\left(\int_{\varepsilon}^{2\varepsilon}\delta^2\left(r\right)dr-\int_0^{\varepsilon}\delta^2\left(r\right)dr\right)=\frac{9\delta^4\left(\varepsilon\right)}{\varepsilon^2}\left(T-\varepsilon\right)\int_{\varepsilon}^{2\varepsilon}\left(\delta^2\left(2\varepsilon\right)-\delta^2\left(r\right)\right)dr.\end{aligned}$$

When δ is in the Hölder scale $\delta\left(r\right)=r^H$, the above quantity is obviously commensurate with $\delta^6\left(\varepsilon\right)/\varepsilon$, which implies the desired result, but in order to be sure we are treating all cases, we now present a general proof which only relies on the fact that δ^2 is increasing and concave.

Below we use the notation $\left(\delta^2\right)'$ for the density of $d\delta^2$, which exists a.e. since δ^2 is concave. The mean value theorem and the concavity of δ^2 then imply that for any $r\in[\varepsilon,2\varepsilon]$,

$$\delta^2\left(2\varepsilon\right)-\delta^2\left(r\right)\geq\left(2\varepsilon-r\right)\inf_{[\varepsilon,2\varepsilon]}\left(\delta^2\right)'=\left(2\varepsilon-r\right)\left(\delta^2\right)'\left(2\varepsilon\right).$$

Thus we can write

$$\begin{aligned}J_1&\geq 9(T-\varepsilon)\varepsilon^{-1}\delta^4\left(\varepsilon\right)\left(\delta^2\right)'\left(2\varepsilon\right)\int_{\varepsilon}^{2\varepsilon}\left(2\varepsilon-r\right)dr=9(T-\varepsilon)\varepsilon^{-1}\delta^4\left(\varepsilon\right)\left(\delta^2\right)'\left(2\varepsilon\right)\varepsilon^2/2\\&\geq cst\cdot\delta^4\left(\varepsilon\right)\cdot\left(\delta^2\right)'\left(2\varepsilon\right).\end{aligned}$$

Since δ^2 is concave, and $\delta\left(0\right)=0$, we have $\delta^2\left(\varepsilon\right)\geq\delta^2\left(2\varepsilon\right)/2$. Hence, with the notation $f\left(x\right)=\delta^2\left(2x\right)$, we have

$$J_1\geq cst\cdot f^2\left(\varepsilon\right)f'\left(\varepsilon\right)=cst\cdot\left(f^3\right)'\left(\varepsilon\right).$$

Therefore we have that $\lim_{\varepsilon\to0}\left(f^3\right)'\left(\varepsilon\right)=0$. We prove this implies $\lim_{\varepsilon\to0}\varepsilon^{-1}f^3\left(\varepsilon\right)=0$. Indeed, fix $\eta>0$; then there exists $\varepsilon_\eta>0$ such that for all $\varepsilon\in(0,\varepsilon_\eta]$, $0\leq\left(f^3\right)'\left(\varepsilon\right)\leq\eta$ (we used the positivity of $\left(\delta^2\right)'$). Hence, also using $f\left(0\right)=0$, for any $\varepsilon\in(0,\varepsilon_\eta]$,

$$0\leq\frac{f^3\left(\varepsilon\right)}{\varepsilon}=\frac{1}{\varepsilon}\int_0^{\varepsilon}\left(f^3\right)'\left(x\right)dx\leq\frac{1}{\varepsilon}\int_0^{\varepsilon}\eta dx=\eta.$$

This proves that $\lim_{\varepsilon\to0}\varepsilon^{-1}f^3\left(\varepsilon\right)=0$, which is equivalent to the announced necessary condition, and finishes the proof of the theorem. ■

3.3 Non-stationary case

The concavity and stationarity assumptions were used heavily above for the proof of the necessary condition in Theorem 6. We now show they can be considerably weakened while still resulting in a sufficient condition: we only need a weak uniformity condition on the variances, coupled with a natural bound on the second-derivative measure of δ^2.

Theorem 8 *Let $m > 1$ be an odd integer. Let X be a centered Gaussian process on $[0,T]$ with canonical metric*

$$\delta^2(s,t) := \mathbf{E}\left[(X(t) - X(s))^2\right].$$

Define a univariate function on $[0,T]$, also denoted by δ^2, via

$$\delta^2(r) := \sup_{s\in[0,T]} \delta^2(s, s+r),$$

and assume that for r near 0,

$$\delta(r) = o\left(r^{1/2m}\right). \tag{11}$$

Assume that, in the sense of distributions, the derivative $\partial\delta^2/(\partial s\partial t)$ is a finite signed σ finite measure μ on $[0,T]^2 - \Delta$ where Δ is the diagonal $\{(s,s)|s \in [0,T]\}$. Denote the off-diagonal simplex by $OD = \{(s,t) : 0 \le s \le t - \varepsilon \le T\}$; assume μ satisfies, for some constant c and for all ε small enough,

$$|\mu|(OD) \le c\varepsilon^{-(m-1)/m}, \tag{12}$$

where $|\mu|$ is the total variation measure of μ. Then X has zero mth variation.

Example 9 *A typical situation covered by the above theorem is that of the Riemann-Liouville fBm $B^{H,RL}$ and similar non-stationary processes. The process $B^{H,RL}$ is defined by $B^{H,RL}(t) = \int_0^t (t-s)^{H-1/2} dW(s)$; it differs from the standard fBm by a bounded variation process, and as such it has zero mth variation for any $H > 1/(2m)$. This can also be obtained via our theorem, because $B^{H,RL}$ is a member of the class of Gaussian processes whose canonical metric satisfies*

$$|t-s|^H \le \delta(s,t) \le 2|t-s|^H. \tag{13}$$

(see [25]). For any process satisfying (13), our theorem's condition on variances is equivalent to $H > 1/(2m)$, while for the other condition, a direct computation yields $\mu(dsdt)/(dsdt) \asymp |t-s|^{2H-2} dsdt$ off the diagonal, and therefore, for $H < 1/2$,

$$\mu(OD) = |\mu|(OD) \asymp \int_0^T \int_\varepsilon^t s^{2H-2} dsdt \asymp \varepsilon^{2H-1}.$$

This quantity is bounded above by $\varepsilon^{-1+1/m}$ as soon as $H \ge 1/(2m)$, of course, so the strict inequality is sufficient to apply the theorem and conclude that $B^{H,RL}$ all other processes satisfying (13) have zero mth variation.

Example 10 *One can generalize Example 9 to any Gaussian process with a Volterra-convolution kernel: let γ^2 be a univariate increasing concave function, differentiable everywhere except possibly at 0, and define*

$$X(t) = \int_0^t \left(\frac{d\gamma^2}{dr}\right)^{1/2} (t-r)\, dW(r). \tag{14}$$

Then one can show (see [25]) that the canonical metric $\delta^2(s,t)$ of X is bounded above by $2\gamma^2(|t-s|)$, so that we can use the univariate $\delta^2 = 2\gamma^2$, and also $\delta^2(s,t)$ is bounded below by $\gamma^2(|t-s|)$. Similar calculations to the above then easily show that X has zero mth variation as soon as $\delta^2(r) = o\left(r^{1/(2m)}\right)$. Hence there are processes with non stationary increments that are more irregular than fractional Brownian for any $H > 1/(2m)$ which still have zero mth variation: use for instance the X above with $\gamma^2(r) = r^{1/(2m)}/\log(1/r)$.

4 Non-Gaussian case

Now assume that X is given by (3) and M is a square-integrable (non-Gaussian) continuous martingale, m is an odd integer, and define a positive non-random measure μ for $\bar{s} = (s_1, s_2, \cdots, s_m) \in [0,T]^m$ by

$$\mu(d\bar{s}) = \mu(ds_1 ds_2 \cdots ds_m) = \mathbf{E}\left[d[M](s_1)\, d[M](s_2) \cdots d[M](s_m)\right], \tag{15}$$

where $[M]$ is the quadratic variation process of M. We make the following assumption on μ.

(A) The non-negative measure μ is absolutely continuous with respect to the Lebesgue measure $d\bar{s}$ on $[0,T]^m$ and $K(\bar{s}) := d\mu/d\bar{s}$ is bounded by a tensor-power function: $0 \le K(s_1, s_2, \cdots, s_m) \le \Gamma^2(s_1)\Gamma^2(s_2)\cdots\Gamma^2(s_m)$ for some non-negative function Γ on $[0,T]$.

A large class of processes satisfying (A) is the case where $M(t) = \int_0^t H(s)\, dW(s)$ where $H \in L^2([0,T]\times\Omega)$ and W is a standard Wiener process, and we assume $\mathbf{E}\left[H^{2m}(t)\right]$ is finite for all $t \in [0,T]$. Indeed then by Hölder's inequality, since we can take $K(\bar{s}) = \mathbf{E}\left[H^2(s_1)H^2(s_2)\cdots H^2(s_m)\right]$, we see that $\Gamma(t) = \left(\mathbf{E}\left[H^{2m}(t)\right]\right)^{1/(2m)}$ works.

We will show that the sufficient conditions for zero odd variation in the Gaussian cases generalize to the case of condition (A), by associating X with the Gaussian process

$$Z(t) := \int_0^T \tilde{G}(t,s)\, dW(s). \tag{16}$$

where $\tilde{G}(t,s) := \Gamma(s)\, G(t,s)$. We have the following.

Theorem 11 *Let m be an odd integer ≥ 3. Let X and Z be as defined in (3) and (16). Assume M satisfies condition* (A) *and Z is well-defined and satisfies the hypotheses of Theorem 6 or Theorem 8 relative to a univariate function δ. Assume that for some constant $c > 0$, and every small $\varepsilon > 0$,*

$$\int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \int_{u=0}^{T} \left|\Delta\tilde{G}_t(u)\right| \left|\Delta\tilde{G}_s(u)\right| du \le c\varepsilon\delta^2(2\varepsilon), \tag{17}$$

where we use the notation $\Delta\tilde{G}_t(u) = \tilde{G}(t+\varepsilon, u) - \tilde{G}(t,u)$. Then X has zero mth variation.

The next proposition illustrates the range of applicability of Theorem 11. We will use it to construct classes of examples of martingale-based processes X to which the theorem applies.

Proposition 12 *Let X be defined by (3). Assume m is an odd integer ≥ 3 and condition* (A) *holds. Assume that $\tilde{G}(t,s) := \Gamma(s)\, G(t,s)$ can be bounded above as follows: for all s,t,*

$$\tilde{G}(t,s) = \mathbf{1}_{s\le t}\; g(t,s) = \mathbf{1}_{s\le t}\, |t-s|^{1/(2m)-1/2}\, f(t,s)$$

in which the bivariate function $f(t,s)$ is positive and bounded as

$$|f(t,s)| \leq f(|t-s|)$$

where the univariate function $f(r)$ is increasing, and concave on $\mathbf{R}_+$, with $\lim_{r\to 0} f(r) = 0$, and where g has a second mixed derivative such that

$$\left|\frac{\partial g}{\partial t}(t,s)\right| + \left|\frac{\partial g}{\partial s}(t,s)\right| \leq c\,|t-s|^{1/(2m)-3/2};$$
$$\left|\frac{\partial^2 g}{\partial s \partial t}(t,s)\right| \leq c\,|t-s|^{1/(2m)-5/2}.$$

Also assume $t \mapsto g(s,t)$ is decreasing and $t \mapsto f(s,t)$ is increasing. Then X has zero m-variation.

The presence of the indicator function $\mathbf{1}_{s\leq t}$ in the expression for $\tilde{G}$ above is typical of most models, since it coincides with asking that Z be adapted to the filtrations of W, which is equivalent to X being adapted to the filtration of M. The proposition covers a wide variety of martingale-based models, which can be quite far from Gaussian models in the sense that they can have only a few moments. We describe one easily constructed class.

Example 13 *Assume that M is a martingale such that $\mathbf{E}\left[|d[M]/dt|^m\right]$ is bounded above by a constant c^{2m} uniformly in $t \leq T$. For instance we can take $M_t = \int_0^t H_s(s)\,dW(s)$ where H is a W-adapted process with $\mathbf{E}\left[|H_s|^{2m}\right]^{1/2m} \leq c$. This boundedness assumption implies that we can take $\Gamma \equiv c$ in Condition (A), and $\tilde{G} = cG$. Let $G(t,s) = G_{RLfBm}(t,s) := \mathbf{1}_{s\leq t}|t-s|^{1/(2m)-1/2+\alpha}$ for some $\alpha > 0$; in other words, G is the Brownian representation kernel of the Riemann-Liouville fBm with parameter $H = 1/(2m) - \alpha > 1/(2m)$. It is immediate to check that the assumptions of Proposition 12 are satisfied for this class of martingale-based models, which implies that the corresponding X defined by (3) have zero mth variation.*

More generally, assume that G is bounded above by a multiple of G_{RLfBm}, and assume the two partial derivatives of G, and the mixed second order derivative of G, are bounded by the corresponding (multiples of) derivatives of G_{RLfBm}; one can check that the standard fBm's kernel is in this class, and that the martingale-based models of this class also satisfy the assumptions of Proposition 12, resulting again zero mth variations for the corresponding X defined in (3). For the sake of conciseness, we will omit the details, which are tedious and straightforward.

The main quantitative assumption on the univariate function $\delta(\varepsilon)$ corresponding to $\tilde{G}$ in the theorem, i.e. $\delta(r) = o\left(r^{1/(2m)}\right)$, can be reinterpreted as a regularity condition on X.

Example 14 *For example, if X has fractional exponential moments, in the sense that for some constants $a > 0$ and $0 < \beta \leq 2$, $\mathbf{E}\left[\exp\left(a\,|X(t)-X(s)|^\beta\right)\right]$ is finite for all s,t, then an almost-sure uniform modulus of continuity for X is $r \mapsto \delta(r)\log^{\beta/2}(1/r)$. This can be established by using Corollary 4.5 in [38]. By using the Burkholder-Davis-Gundy inequality on the exponential martingale based on M, we can prove that such fractional exponential moments hold, for instance, in the setting of Example 13, if there exists $b > 0$ such that $\mathbf{E}\left[\exp\left(b\,|H_s|^{2\beta}\right)\right]$ is bounded in $s \in [0,T]$. If one only has standard (non-exponential) moments, similar (less sharp) results can be obtained via Kolmogorov's continuity criterion instead of [38]. All details are left to the reader.*

5 Stochastic calculus

This section's goal is to define the so-called symmetric stochastic integral and its associated Itô formula for processes which are not fBm. The reader may refer to the Introduction (Section 1) for motivations on why we study this topic. We concentrate on Gaussian processes under hypotheses similar to those used in Section 3.3 (Theorem 8). The basic strategy is to use the results of [20] which were applied to fBm. Let X be a stochastic process on $[0,1]$. According to Sections 3 and 4 in [20] (specifically, according to the proof of part 1 of Theorem 4.4 therein), if for every bounded measurable function g on $\mathbf{R}$, the limit

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^1 du\,(X_{u+\varepsilon}-X_u)^m\,g\left(\frac{X_{u+\varepsilon}+X_u}{2}\right)=0 \tag{18}$$

holds in probability, for both $m=3$ and $m=5$, then for every $t\in[0,1]$ and every $f\in C^6(\mathbf{R})$, the *symmetric* ("generalized Stratonovich") stochastic integral

$$\int_0^t f'(X_u)\,d^\circ X_u =: \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^t du\,(X_{u+\varepsilon}-X_u)\,\frac{1}{2}\left(f'(X_{u+\varepsilon})+f'(X_u)\right) \tag{19}$$

exists and we have the Itô formula

$$f(X_t)=f(X_0)+\int_0^t f'(X_u)\,d^\circ X_u. \tag{20}$$

Our goal is thus to prove (18) for a wide class of Gaussian processes X, which will in turn imply the existence of (19) and the Itô formula (20).

If X has stationary increments in the sense of Section 3.2, meaning that $\mathbf{E}\left[(X_s-X_t)^2\right]=\delta^2(t-s)$ for some univariate canonical metric function δ, then by using $g\equiv\mathbf{1}$ and our Theorem 6, we see that for (18) to hold, we must have $\delta(r)=o\left(r^{1/6}\right)$. If one wishes to treat non-stationary cases, we notice that (18) for $g\equiv 1$ is the result of our non-stationary Theorem 8, so it is necessary to use that theorem's hypotheses, which include the non-stationary version of $\delta(r)=o\left(r^{1/6}\right)$. But we will also need some non-degeneracy conditions in order to apply the quartic linear regression method of [20]. These are Conditions (i) and (ii) in the next Theorem. Condition (iii) therein is essentially a consequence of the condition that δ^2 be increasing and concave. These conditions are all further discussed after the statement of the next theorem and its corollary.

Theorem 15 *Let $m\geq 3$ be an odd integer. Let X be a Gaussian process on $[0,1]$ satisfying the hypotheses of Theorem 8. This means in particular that we denote as usual its canonical metric by $\delta^2(s,t)$, and that there exists a univariate increasing and concave function δ^2 such that $\delta(r)=o\left(r^{1/(2m)}\right)$ and $\delta^2(s,t)\leq\delta^2(|t-s|)$. Assume that for $u<v$, the functions $u\mapsto Var[X_u]=:Q_u$, $v\mapsto\delta^2(u,v)$, and $u\mapsto-\delta^2(u,v)$ are increasing and concave. Assume there exist positive constants $a>1$, $b<1/2$, $c>1/4$, and $c'>0$ such that for all $\varepsilon<u<v\leq 1$,*

(i) $c\delta^2(u)\leq Q_u$,

(ii) $c'\delta^2(u)\,\delta^2(v-u)\leq Q_uQ_v-Q^2(u,v)$,

(iii)

$$\frac{\delta(au)-\delta(u)}{(a-1)\,u}<b\frac{\delta(u)}{u}. \tag{21}$$

Then for every bounded measurable function g on $\mathbf{R}$,

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^2}\mathbf{E}\left[\left(\int_0^1 du\,(X_{u+\varepsilon}-X_u)^m\,g\left(\frac{X_{u+\varepsilon}+X_u}{2}\right)\right)^2\right]=0.$$

When we apply this theorem to the case $m=3$, the assumption depending on m, namely $\delta(r)=o\left(r^{1/(2m)}\right)$ is satisfied a fortiori for $m=5$ as well, which means that under the assumption $\delta(r)=o\left(r^{1/6}\right)$, the theorem's conclusion holds for $m=3$ and $m=5$. Therefore, as mentioned in the strategy above, we immediately get the following.

Corollary 16 *Assume the hypotheses of Theorem 15 with $m=3$. We have existence of the symmetric integral in (19), and its Itô formula (20), for every $f\in C^6(\mathbf{R})$ and $t\in[0,1]$.*

The end of Section 3.3 contains examples satisfying the hypotheses of Theorem 8; most of these examples also satisfy the monotonicity and convexity conditions in the above theorem. We state this formally, omitting the details of checking the conditions.

Example 17 *The conclusion of Corollary 16 applies to The Riemann-Liouville fBm described in Example 9, which is a Gaussian process with non-stationary increments. It also applies to any member of the wider class of processes in Example 10 for which the function γ defined therein satisfies conditions (i), (ii), and (iii) of Theorem 15. This includes the family of processes such that $\gamma(r)=r^H\log^\beta(1/r)$ for $H\in(1/6,1)$ and $\beta\in\mathbf{R}$, the case $\beta=0$ yielding the Riemann-Liouville fBm processes.*

Before proceeding to the proof of Theorem 15, we provide a broader discussion of its hypotheses.

Condition (i) is a type of coercivity assumption on the non-degeneracy of X's variances in comparison to its increments' variances. The hypotheses of Theorem 8 imply that $Q_u\le\delta^2(u)$, and Condition (i) simply adds that these two quantities should be commensurate, with a lower bound that it not too small. The "Volterra convolution"-type class of processes (14) given at the end of Section 3.3, which includes the Riemann-Liouville fBm's, satisfies Condition (i) with $c=1/2$. In the stationary case, (i) is trivially satisfied since $Q_u\equiv\delta^2(u)$.

Condition (ii) is also a type of coercivity condition. It too is satisfied in the stationary case. We prove this claim, since it is not immediately obvious. In the stationary case, since $\delta^2(u,v)=\delta^2(v-u)=Q_{v-u}$, we calculate

$$Q_uQ_v-Q^2(u,v)=Q_uQ_v-4^{-1}(Q_u+Q_v-Q_{v-u})^2$$

and after rearranging some terms we obtain

$$Q_uQ_v-Q^2(u,v)=2^{-1}Q_{v-u}(Q_u+Q_v)-4^{-1}(Q_v-Q_u)^2-4^{-1}Q_{v-u}^2.$$

We note first that by the concavity of Q, we have $Q_v-Q_u<Q_{v-u}$, and consequently, $(Q_v-Q_u)^2\le(Q_v-Q_u)Q_{v-u}\le Q_vQ_{v-u}$. This implies

$$Q_uQ_v-Q^2(u,v)\ge 2^{-1}Q_{v-u}Q_u+4^{-1}\left(Q_{v-u}Q_v-Q_{v-u}^2\right).$$

Now by monotonicity of Q, we can write $Q_{v-u}Q_v\ge Q_{v-u}^2$. This, together with Condition (i), yield Condition (ii) since we now have

$$Q_uQ_v-Q^2(u,v)\ge 2^{-1}Q_{v-u}Q_u\ge 2^{-1}c^2\delta^2(v-u)\,\delta^2(u).$$

Lastly, Condition (iii) represents a strengthened concavity condition on the univariate function δ. Indeed, the left-hand side in (21) is the slope of the secant of the graph of δ between the points u and au, while the right-hand side is b times the slope of the secant from 0 to u. If b were allowed to be 1, (iii) would simply be a consequence of convexity. Here taking $b \leq 1/2$ means that we are exploiting the concavity of δ^2; the fact that condition (iii) requires slightly more, namely b strictly less than $1/2$, allows us to work similarly to the scale $\delta(r) = r^H$ with $H < 1/2$, as opposed to simply asking $H \leq 1/2$. Since the point of the Theorem is to allow continuity moduli which are arbitrarily close to $r^{1/6}$, Condition (iii) is hardly a restriction.

Proof of Theorem 15.

Step 0: setup. The expectation to be evaluated is written, as usual, as a double integral over $(u,v) \in [0,1]^2$. For $\varepsilon > 0$ fixed, we define the "off-diagonal" set

$$D_\varepsilon = \left\{(u,v) \in [0,1]^2 : \varepsilon^{1-\rho} \leq u \leq v - \varepsilon^{1-\rho} < v \leq 1\right\}$$

where $\rho \in (0,1)$ is fixed. Using the boundedness of g and Cauchy-Schwarz's inequality, thanks to the hypothesis $\delta(r) = o\left(r^{1/(2m)}\right)$, the term corresponding to the diagonal part (integral over D_ε^c) can be treated identically to what was done in [20] in dealing with their term $\mathcal{J}'(\varepsilon)$ following the statement of their Lemma 5.1, by choosing ρ small enough. It is thus sufficient to prove that

$$\mathcal{J}(\varepsilon) := \frac{1}{\varepsilon^2}\mathbf{E}\left[\iint_{D_\varepsilon} dudv\,(X_{u+\varepsilon} - X_u)^m (X_{v+\varepsilon} - X_v)^m\, g\left(\frac{X_{u+\varepsilon} + X_u}{2}\right) g\left(\frac{X_{v+\varepsilon} + X_v}{2}\right)\right]$$

tends to 0 as ε tends to 0. We now use the same method and notation as in Step 3 of the proof of Theorem 4.1 in [20]. It proceeds through the linear regression analysis of the Gaussian vector $(G_1, G_2, G_3, G_4) := (X_{u+\varepsilon} + X_u, X_{v+\varepsilon} + X_v, X_{u+\varepsilon} - X_u, X_{v+\varepsilon} - X_v)$. In order to avoid repeating arguments from that proof, we only state and prove the new lemmas which are required. The new elements come from the analysis of the Gaussian vector $(\Gamma_3, \Gamma_4)^t := A(G_1, G_2)$ where $A := \Lambda_{21}(\Lambda_{11})^{-1}$ where Λ_{11} is the covariance of the vector (G_1, G_2) and Λ_{21} is the matrix $\{Cov(G_{i+2}, G_j) : i,j = 1,2\}$, as well as from the the centered Gaussian vector (Z_3, Z_4) which is the component independent of (G_3, G_4) in its linear regression against (G_1, G_2), i.e. $(G_3, G_4)^t = A(G_1, G_2)^t + (Z_3, Z_4)$.

Step 1: translating Lemma 5.3 from [20]. Using the fact that $\mathbf{E}\left[Z_\ell^2\right] \leq \mathbf{E}\left[G_\ell^2\right] \leq \delta^2(\varepsilon)$, this lemma translates as the following, proved in the Appendix:

Lemma 18 *Let $k \geq 2$ be an integer. Then for $\ell = 3, 4$,*

$$\iint_{D_\varepsilon} \mathbf{E}\left[|\Gamma_\ell|^k\right] dudv \leq cst \cdot \varepsilon\delta^k(\varepsilon).$$

Step 2: translating Lemma 5.4 from [20]. We will prove the following result

Lemma 19 *For all $j \in \{0, 1, \cdots, (m-1)/2\}$,*

$$\iint_{D_\varepsilon} |\mathbf{E}[Z_3 Z_4]|^{m-2j}\, dudv \leq cst \cdot \varepsilon\delta^{2(m-2j)}(\varepsilon).$$

Proof of Lemma 19. As in [20], we have

$$|\mathbf{E}[Z_3Z_4]|^{m-2j} \leq cst \cdot |\mathbf{E}[G_3G_4]|^{m-2j} + cst \cdot |\mathbf{E}[\Gamma_3\Gamma_4]|^{m-2j}.$$

The required estimate for the term corresponding to $|\mathbf{E}[\Gamma_3\Gamma_4]|^{m-2j}$ follows by Cauchy-Schwarz's inequality and Lemma 18. For the term corresponding to $|\mathbf{E}[G_3G_4]|^{m-2j}$, we recognize that $\mathbf{E}[G_3G_4]$ is the negative planar increment $\Theta^{\varepsilon}(u,v)$ defined in (6). Thus the corresponding term was already considered in the proof of Theorem (8). More specifically, up to the factor $\varepsilon^2\delta^{-4j}(\varepsilon)$, we now have to estimated the same integral as in Step 2 of that theorem's proof: see expression (23) for the term we called $J_{j,OD}$. This means that

$$\iint_{D_\varepsilon} |\mathbf{E}[G_3G_4]|^{m-2j}\, dudv \leq \frac{\varepsilon^2}{\delta^{4j}(\varepsilon)} J_{j,OD} \leq \varepsilon^2 |\mu|(OD)\, \delta^{2(m-2j-1)}(\varepsilon).$$

Our hypotheses borrowed from Theorem (8) that $|\mu|(OD) \leq cst \cdot \varepsilon^{1/m-1}$ and that $\delta^2(\varepsilon) = o\left(r^{1/(2m)}\right)$ now imply that the above is $\ll \varepsilon\delta^{2(m-2j)}(\varepsilon)$, concluding the lemma's proof. □

Step 4. Conclusion. The remainder of the proof of the theorem is to check that Lemmas 18 and 19 do imply the claim of the theorem; this is done exactly as in Steps 3 and 4 of the proof of Theorem 4.1 in [20]. Since such a task is only bookkeeping, we omit it, concluding the proof of Theorem 15, modulo the proof of Lemma 18 which is found in the appendix. ■

Acknowledgements

The work of F. Russo was partially supported by the ANR Project MASTERIE 2010 BLAN-0121-01. The work of F. Viens is partially supported by NSF DMS grant 0907321. Constructive comments by referees and editors are gratefully acknowledged and resulted in several improvements.

References

[1] Adler, R.: *An introduction to continuity, extrema, and related topics for general Gaussian processes.* Inst. Math. Stat., Hayward, CA, (1990).

[2] Alòs, E., Mazet, O., Nualart, D.: *Stochastic calculus with respect to Gaussian processes.* Ann. Probab. **29**, 766–801 (1999)

[3] Alòs, E. and Nualart, D.: *Stochastic integration with respect to fractional Brownian motion.* Stoch. Stoch. Rep., **75**, 3, 129–152 (2003).

[4] Azaïs, J.-M.; Wschebor, M. Almost sure oscillation of certain random processes, *Bernoulli* **2** no. 3, 257–270 (1996).

[5] Bertoin, J.: *Sur une intégrale pour les processus à α-variation bornée.* Ann. Probab., **17**(4):1521–1535, (1989).

[6] Biagini, F., Hu, Y., Øksendal, B. and Zhang, T.: *Stochastic calculus with respect to fractional Brownian motion and applications.* Probability and its Applications. Springer-Verlag, 2008.

[7] Borell, C. On polynomial chaos and integrability, *Probability and Mathematical Statistics*, **3**, no. 2, 191–203 (1984).

[8] Breuer, P., Major, P.: *Central limit theorems for non-linear functionals of Gaussian fields.* J. Multivariate Anal. **13**, 425–441 (1983).

[9] Burdzy, K., Swanson, J.: *A change of variable formula with Itô correction term.* Ann. Probab. **38**, no. 5, 1817–1869 (2010).

[10] Bruneau, M.: *Variation totale d'une fonction.* Lecture Notes in Mathematics, Vol. **413**. Springer-Verlag, Berlin-New York, 1974.

[11] Carmona, Ph, Coutin, L. and Montseny, G.: *Stochastic integration with respect to fractional Brownian motion.* Annales de l'Institut Henry Poincaré. Probabilités et Statistiques, **39**, no 1, 27–68 (2003).

[12] Dudley, R. M., Norvaiša, R.: *Differentiability of six operators on nonsmooth functions and p-variation.* With the collaboration of Jinghua Qian. Lecture Notes in Mathematics, **1703**. Springer-Verlag, Berlin, 1999.

[13] Errami, M., Russo, F.: *Covariation de convolution de martingales.* C. R. Acad. Sci. Paris Sér. I Math. **326**, no. 5, 601–606 (1998).

[14] Errami, M., Russo, F.: *n-covariation,generalized Dirichlet processes and calculus with respect to finite cubic variation processes.* Stochastic Processes Appl. , **104**, 259–299 (2003).

[15] Flandoli, F., Russo, F.: *G*eneralized stochastic integration and stochastic ODE's. Annals of Probability. Vol. **30**, no 1, 270–292 (2002).

[16] Föllmer, H.: Calcul d'Itô sans probabilités. In *Séminaire de Probabilités, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume **850** of *Lecture Notes in Math.*, pages 143–150. Springer, Berlin, 1981.

[17] Friz, P.; Victoir, N. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications.* Cambridge Studies in Advanced Mathematics, Cambridge UP, 2010.

[18] Gradinaru, M., Nourdin, I.: *Approximation at first and second order of m-order integrals of the fractional Brownian motion and of certain semimartingales.* Electron. J. Probab. **8**, no. 18, 26 pp (2003).

[19] Gradinaru, M., Russo F., Vallois, P.: *Generalized covariations, local time and Stratonovich Itô's formula for fractional Brownian motion with Hurst index* $H \geq \frac{1}{4}$. Ann. Pobab., **31** (4), 1772–1820, (2003).

[20] Gradinaru, M., Nourdin, I., Russo, F., Vallois, P.: *m-order integrals and Itô's formula for non-semimartingale processes; the case of a fractional Brownian motion with any Hurst index.* Ann. Inst. H. Poincaré Probab. Statist. **41**, 781–806 (2005).

[21] Hu, Y., Nualart,D., Song, J.: *Fractional martingales and characterization of the fractional Brownian motion.* Ann. Probab. **37**, no. 6, 2404–2430 (2009).

[22] Kruk, I and Russo, F.: *Malliavin-Skorohod calculus and Paley-Wiener integral for covariance singular processes.* Preprint HAL-INRIA 00540914.

[23] Kuo, H.-H. *Introduction to stochastic integration.* Springer, 2006.

[24] Lyons, T., Qian, Z.: *System control and rough paths.* Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002.

[25] Mocioalca, O., Viens, F.: *Skorohod integration and stochastic calculus beyond the fractional Brownian scale.* Journal of Functional Analysis, **222** no. 2, 385–434 (2004).

[26] Nourdin, I.: *A change of variable formula for the 2D fractional Brownian motion of Hurst index bigger or equal to* 1/4. J. Funct. Anal. **256**, 2304–2320 (2009).

[27] Nourdin, I., Nualart, D.: *Central limit theorems for multiple Skorohod integrals.* J. Theoret. Probab. **23**, no. 1, 39–64.

[28] Nourdin, I., Nualart D., Tudor, C.: *Central and non-central limit theorems for weighted power variations of fractional Brownian motion.* Ann. Inst. Henri Poincaré Probab. Stat. **46**, no. 4, 1055–1079 (2010).

[29] Nourdin, I., Réveillac, A., Swanson, J.: *The weak Stratonovich integral with respect to fractional Brownian motion with Hurst parameter 1/6.* Electron. J. Probab. **15**, no. 70, 2117–2162 (2010).

[30] Nualart, D.: *The Malliavin calculus and related topics.* Second edition. Probability and its Applications. Springer-Verlag, 2006.

[31] Nualart, D., Ortiz-Latorre, S.: *Central limit theorems for multiple stochastic integrals and Malliavin calculus.* Stochastic Process. Appl., **118**(4), 614–628 (2008).

[32] Rogers, L. C. G., Walsh, J. B.: *The exact* 4/3*-variation of a process arising from Brownian motion.* Stochastics Stochastics Rep. **51**, no. 3-4, 267–291 (1994).

[33] Russo, F., Tudor, C.: On the bifractional Brownian motion. *Stochastic Processes Appl.* **116**, no. 6, 830–856 (2006).

[34] Russo, F., Vallois, P.: *Stochastic calculus with respect to a finite quadratic variation process.* Stochastics and Stochastic Reports **70**, 1–40 (2000).

[35] Russo, F., Vallois, P.: *The generalized covariation process and Itô formula.* Stochastic Process. Appl. **59**, no. 1, 81–104 (1995).

[36] Russo, F., Vallois, P.: *Elements of stochastic calculus via regularizations.* Séminaire de Probabilités **XL**, Lect. Notes Math. **1899**, Berlin Heidelberg New-York, Springer, pages 147–186 (2007).

[37] Tindel, S., Tudor, C.A., Viens, F.: Sharp Gaussian regularity on the circle and application to the fractional stochastic heat equation. *Journal of Functional Analysis*, **217** (2), 280–313 (2004).

[38] Viens, F., Vizcarra, A.: *Supremum Concentration Inequality and Modulus of Continuity for Sub-nth Chaos Processes.* J. Funct.Analysis **248**, 1–26 (2007).

6 Appendix

Proof of Lemma 1. The formula in the lemma is an easy consequence of the following formula, which can be found as Lemma 5.2 in [20]: for any centered jointly Gaussian pair of r.v.'s (Y, Z), we have $\mathbf{E}\left[Y^m Z^m\right] = \sum_{j=0}^{(m-1)/2} c_j \mathbf{E}\left[YZ\right]^{m-2j} \; Var\left[X\right]^j \; Var\left[Y\right]^j$. To see that the c_j's are positive, note that one can decompose each odd monomial into the basis of odd-order Hermite polynomials: $x^m = \sum_{j=0}^{(m-1)/2} a_{2j+1} H_{2j+1}(x)$, from whence it follows, thanks to the orthogonality of Hermite polynomials of Gaussian rv's, that $c_j = (a_{2j+1})^2$. ■

Proof of Lemma 3. The proof of this lemma is elementary. It follows from two uses of the multiplication formula for Wiener integrals [30, Proposition 1.1.3], for instance. All details are left to the reader. ■

Proof of Lemma 4. Reintroducing the notation X and Θ into the formula in Lemma 3, we get

$$\mathcal{I}_1 = \frac{3}{\varepsilon} \int_0^T ds \left(X(s+\varepsilon) - X(s)\right) Var\left(X(s+\varepsilon) - X(s)\right)$$

and therefore,

$$\mathbf{E}\left[\left|\mathcal{I}_1\right|^2\right] = \frac{9}{\varepsilon^2} \int_0^T \int_0^t dtds \Theta^{\varepsilon}(s,t) \, Var\left(X(t+\varepsilon) - X(t)\right) Var\left(X(s+\varepsilon) - X(s)\right)$$

Using the variances of fBm, writing H instead of $1/6$ to improve readability,

$$\begin{aligned}
\mathbf{E}\left[\left|\mathcal{I}_1\right|^2\right] &= \frac{9}{2}\varepsilon^{-2+4H} \int_0^T \int_0^T dtds \; Cov\left[X(t+\varepsilon) - X(t); X(s+\varepsilon) - X(s)\right] \\
&= \frac{9}{2}\varepsilon^{-2+4H} \; Var\left[\int_0^T \left(X(t+\varepsilon) - X(t)\right) dt\right] \\
&= \frac{9}{2}\varepsilon^{-2+4H} \; Var\left[\int_T^{T+\varepsilon} X(t)\, dt - \int_0^{\varepsilon} X(t)\, dt\right].
\end{aligned}$$

Bounding the variance of the difference by twice the sum of the variances,

$$\mathbf{E}\left[\left|\mathcal{I}_1\right|^2\right] \leq 9\varepsilon^{-2+4H} \left(\int_T^{T+\varepsilon} \int_T^{T+\varepsilon} T^{2H} dsdt + \int_0^{\varepsilon} \int_0^{\varepsilon} \varepsilon^{2H} dsdt\right) = O\left(\varepsilon^{4H}\right),$$

proving Lemma 4. ■

Proof of Lemma 5. By the technique at the start of the proof of Lemma 4, the product formula in [30, Proposition 1.1.3], and the covariance of fBm, we first get

$$\mathcal{I}_3 := \frac{6}{\varepsilon} \int_0^T dW(s_3) \int_0^{s_3} dW(s_2) \int_0^{s_2} dW(s_1) \int_0^T \left[\prod_{k=1}^{3} \Delta G_s(s_k)\right] ds.$$

$$\begin{aligned}
\mathbf{E}\left[\left|\mathcal{I}_3\right|^2\right] &= \frac{12}{\varepsilon^2} \int_0^T \int_0^t dtds \; \left(\Theta^{\varepsilon}(s,t)\right)^3 \\
&= \frac{6}{\varepsilon^2} \int_0^T \int_0^t dtds \; \left(\left|t-s+\varepsilon\right|^{2H} + \left|t-s-\varepsilon\right|^{2H} - 2\left|t-s\right|^{2H}\right)^3.
\end{aligned}$$

We must take care of the absolute values, i.e. of whether ε is greater or less than $t-s$. We define the "off-diagonal" portion of $\mathbf{E}\left[|\mathcal{I}_3|^2\right]$ as

$$\mathcal{ODI}_3 := 6\varepsilon^{-2}\int_{2\varepsilon}^{T}\int_0^{t-2\varepsilon} dtds\left(|t-s+\varepsilon|^{2H}+|t-s-\varepsilon|^{2H}-2\,|t-s|^{2H}\right)^3.$$

For s,t in the integration domain for the above integral, since $\bar{t} := t-s > 2\varepsilon$, by two iterated applications of the Mean Value Theorem for the function x^{2H} on the intervals $[\bar{t}-\varepsilon,\bar{t}]$ and $[\bar{t},\bar{t}+\varepsilon]$,

$$|\bar{t}+\varepsilon|^{2H}+|\bar{t}-\varepsilon|^{2H}-2\bar{t}^{2H} = 2H\,(2H-1)\,\varepsilon\,(\xi_1-\xi_2)\,\xi^{2H-2}$$

for some $\xi_2\in[\bar{t}-\varepsilon,\bar{t}]$, $\xi_1\in[\bar{t},\bar{t}+\varepsilon]$, and $\xi\in[\xi_1,\xi_2]$, and therefore

$$\begin{aligned}|\mathcal{ODI}_3| &\le 384H^3\,|2H-1|^3\,\varepsilon^{-2}\int_{2\varepsilon}^{T}\int_0^{t-2\varepsilon}\left(\varepsilon\cdot 2\varepsilon\cdot(t-s-\varepsilon)^{2H-2}\right)^3 dtds\\ &\le \frac{384H^3\,|2H-1|^3}{5-6H}T\varepsilon^{6H-1} = \frac{32}{243}T.\end{aligned}$$

where in the last line we substituted $H=1/6$. Thus the "off-diagonal" term is bounded. The diagonal part of $\mathcal{I}_3$ is

$$\begin{aligned}\mathcal{DI}_3 &:= 6\varepsilon^{-2}\int_0^{T}\int_{t-2\varepsilon}^{t} dtds\left(|t-s+\varepsilon|^{2H}+|t-s-\varepsilon|^{2H}-2\,|t-s|^{2H}\right)^3\\ &= 6\varepsilon^{-1+6H}T\int_0^2 dr\left(|r+1|^{2H}+|r-1|^{2H}-2\,|r|^{2H}\right)^3 dr = CT\end{aligned}$$

where, having substituted $H=1/6$, yields that C is a universal constant. Thus the diagonal part $\mathcal{DI}_3$ of $\mathbf{E}[|\mathcal{I}_3|^2]$ is constant. This proves that $\mathcal{I}_3$ is bounded in $L^2(\Omega)$, as announced. To conclude that it cannot converge in $L^2(\Omega)$, recall that from [20, Theorem 4.1 part (2)], $[X,3]_\varepsilon(T)=\mathcal{I}_1+\mathcal{I}_3$ converges in distribution to a non-degenerate normal law. By Lemma 4, $\mathcal{I}_1$ converges to 0 in $L^2(\Omega)$. Therefore, $\mathcal{I}_3$ converges in distribution to a non-degenerate normal law; if it also converged in $L^2(\Omega)$, since the 3rd Wiener chaos is closed in $L^2(\Omega)$, the limit would have to be in that same chaos, and thus would not have a non-degenerate normal law. ■

Proof of Theorem 8.
Step 0: setup. Recall the result of Lemma 1, where now we express $Var\left[X(t+\varepsilon)-X(t)\right] = \delta^2(t,t+\varepsilon)$ and

$$\Theta^\varepsilon(s,t) = \mu\left([s,s+\varepsilon]\times[t,t+\varepsilon)\right) = \int_s^{s+\varepsilon}\int_t^{t+\varepsilon}\mu\,(dudv)\,. \tag{22}$$

We again separate the diagonal term from the off-diagonal term, although this time the diagonal is twice as wide: it is defined as $\{(s,t): 0\le t-2\varepsilon\le s\le t\}$.

Step 1: diagonal. Using Cauchy-Schwarz's inequality which implies $|\Theta^\varepsilon(s,t)|\le\delta(s,s+\varepsilon)\,\delta(t,t+\varepsilon)$, and bounding each term $\delta(s,s+\varepsilon)$ by $\delta(\varepsilon)$, the diagonal portion of $\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right]$ can be bounded above, in absolute value, by

$$\frac{1}{\varepsilon^2}\sum_{j=0}^{(m-1)/2} c_j\int_{2\varepsilon}^{T}dt\int_{t-2\varepsilon}^{t}ds\delta^{2m}(\varepsilon) = cst\cdot\varepsilon^{-1}\delta^{2m}(\varepsilon)\,.$$

Hypothesis (11) implies that this converges to 0 with ε. The case of $t \leq 2\varepsilon$ works equally easily.

Step 2: off diagonal. The off-diagonal contribution is the sum for $j = 0, \cdots, (m-1)/2$ of the terms

$$J_{j,OD} = \varepsilon^{-2} c_j \int_{2\varepsilon}^{T} dt \int_{0}^{t-2\varepsilon} ds \delta^{2j}(s, s+\varepsilon) \delta^{2j}(t, t+\varepsilon) \Theta^{\varepsilon}(s,t)^{m-2j} \tag{23}$$

Step 2.1: term $J_{(m-1)/2,OD}$. This is the dominant term. Denoting $c = \left|c_{(m-1)/2}\right|$, we have

$$\left|J_{(m-1)/2,OD}\right| \leq \frac{c\delta^{2m-2}(\varepsilon)}{\varepsilon^2} \int_{2\varepsilon}^{T} dt \int_{0}^{t-2\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right|.$$

We estimate the integral, using the formula (22) and Fubini's theorem:

$$\begin{aligned}
&\int_{2\varepsilon}^{T} dt \int_{0}^{t-2\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right| = \int_{2\varepsilon}^{T} dt \int_{0}^{t-2\varepsilon} ds \left| \int_{s}^{s+\varepsilon} \int_{t}^{t+\varepsilon} \mu(dudv) \right| \\
&\leq \int_{2\varepsilon}^{T} dt \int_{0}^{t-2\varepsilon} ds \int_{s}^{s+\varepsilon} \int_{t}^{t+\varepsilon} |\mu|(dudv) = \int_{2\varepsilon}^{T+\varepsilon} \int_{0}^{v\wedge(T-\varepsilon)} |\mu|(dudv) \int_{2\varepsilon\vee(v-\varepsilon)\vee(u+\varepsilon)}^{v\wedge T} \int_{0\vee(u-\varepsilon)}^{u\wedge(t-2\varepsilon)} ds\ dt \\
&\leq \int_{2\varepsilon}^{T+\varepsilon} \int_{0}^{v-\varepsilon} |\mu|(dudv) \int_{v-\varepsilon}^{v} \int_{u-\varepsilon}^{u} ds\ dt = \varepsilon^2 \int_{2\varepsilon}^{T+\varepsilon} \int_{0}^{v-\varepsilon} |\mu|(dudv).
\end{aligned}$$

Hence we have

$$J_{(m-1)/2,OD} \leq c\delta^{2m-2}(\varepsilon) \int_{v=2\varepsilon}^{T+\varepsilon} \int_{u=0}^{v-\varepsilon} |\mu|(dudv) \leq c\delta^{2m-2}(\varepsilon) |\mu|(OD),$$

which again converges to 0 by hypothesis as ε goes to 0.

Step 2.2: other $J_{j,OD}$ terms. Let now $j < (m-1)/2$. Using Cauchy-Schwarz's inequality for all but one of the $m-2j$ factors Θ in the expression (23) for $J_{j,OD}$, which is allowed because $m-2j \geq 1$ here, exploiting the bounds on the variance terms via the univariate function δ, we have

$$\left|J_{j,OD}\right| \leq \delta^{2m-2}(\varepsilon) c_j \varepsilon^{-2} \int_{2\varepsilon}^{T} dt \int_{0}^{t-2\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right|,$$

which is the same term we estimated in Step 2.1. This finishes the proof of the theorem. ■

Proof of Theorem 11. *Step 0: setup.* We use an expansion for powers of martingales written explicitly at Corollary 2.18 of [14]. For any integer $k \in [0, [m/2]]$, let Σ_m^k be the set of permutations σ of $m-k$ defined as those for which the first k terms $\sigma^{-1}(1), \sigma^{-1}(2), \cdots, \sigma^{-1}(k)$ are chosen arbitrarily and the next $m-2k$ terms are chosen arbitrarily among the remaining integers $\{1, 2, \cdots, m-k\} \setminus \left\{\sigma^{-1}(1), \sigma^{-1}(2), \cdots, \sigma^{-1}(k)\right\}$. Let Y be a fixed square-integrable martingale. We define the process $Y_{\sigma,\ell}$ (denoted in the above reference by σ_Y^{ℓ}) by setting, for each $\sigma \in \Sigma_m^k$ and each $\ell = 1, 2, \cdots, m-k$,

$$Y_{\sigma,\ell}(t) = \begin{cases} [Y](t) & \text{if } \sigma(\ell) \in \{1, 2, \cdots, k\} \\ Y(t) & \text{if } \sigma(\ell) \in \{k+1, \cdots, m-k\}. \end{cases}$$

From Corollary 2.18 of [14], we then have for all $t \in [0, T]$

$$(Y_t)^m = \sum_{k=0}^{[m/2]} \frac{m!}{2^k} \sum_{\sigma \in \Sigma_m^k} \int_0^t \int_0^{u_{m-k}} \cdots \int_0^{u_2} dY_{\sigma,1}(u_1)\ dY_{\sigma,2}(u_2) \cdots dY_{\sigma,m-k}(u_{m-k}).$$

We use this formula to evaluate

$$[X,m]_\varepsilon(T) = \frac{1}{\varepsilon}\int_0^T ds\,(X(s+\varepsilon)-X(s))^m$$

by noting that the increment $X(s+\varepsilon)-X(s)$ is the value at time T of the martingale $Y_t := \int_0^t \Delta G_s(u)\,dM(u)$ where we set

$$\Delta G_s(u) := G(s+\varepsilon,u) - G(s,u).$$

Hence

$$\begin{aligned}
&(X(s+\varepsilon)-X(s))^m \\
&= \sum_{k=0}^{[m/2]} \frac{m!}{2^k} \sum_{\sigma\in\Sigma_m^k} \int_0^T \int_0^{u_{m-k}} \cdots \int_0^{u_2} d[M]\left(u_{\sigma(1)}\right)\left|\Delta G_s\left(u_{\sigma(1)}\right)\right|^2 \cdots d[M]\left(u_{\sigma(k)}\right)\left|\Delta G_s\left(u_{\sigma(k)}\right)\right|^2 \\
&dM\left(u_{\sigma(k+1)}\right)\Delta G_s\left(u_{\sigma(k+1)}\right)\cdots dM\left(u_{\sigma(m-k)}\right)\Delta G_s\left(u_{\sigma(m-k)}\right).
\end{aligned}$$

Therefore we can write

$$\begin{aligned}
&[X,m]_\varepsilon(T) \\
&= \frac{1}{\varepsilon}\sum_{k=0}^{[m/2]} \frac{m!}{2^k} \sum_{\sigma\in\Sigma_m^k} \int_0^T \int_0^{u_{m-k}} \cdots \int_0^{u_2} d[M]\left(u_{\sigma(1)}\right)\cdots d[M]\left(u_{\sigma(k)}\right) dM\left(u_{\sigma(k+1)}\right)\cdots dM\left(u_{\sigma(m-k)}\right) \\
&\left[\Delta G.\left(u_{\sigma(k+1)}\right);\cdots;\Delta G.\left(u_{\sigma(m-k)}\right);\Delta G.\left(u_{\sigma(1)}\right);\Delta G.\left(u_{\sigma(1)}\right);\cdots;\Delta G.\left(u_{\sigma(k)}\right);\Delta G.\left(u_{\sigma(k)}\right)\right],
\end{aligned}$$

where we have used the notation

$$[f_1,f_2,\cdots,f_m] := \int_0^T f_1(s)\,f_2(s)\cdots f_m(s)\,ds.$$

To calculate the expected square of the above, we will bound it above by the sum over k and σ of the expected square of each term. Writing squares of Lebesgue integrals as double integrals, and using Itô's formula, each term's expected square is thus, up to (m,k)-dependent multiplicative constants, equal to the expression

$$\begin{aligned}
K = &\frac{1}{\varepsilon^2}\int_{u_{m-k}=0}^T \int_{u'_{m-k}=0}^T \int_{u_{m-k-1}=0}^{u_{m-k}} \int_{u'_{m-k-1}=0}^{u_{m-k}} \cdots \int_{u_1=0}^{u_2}\int_{u'_1=0}^{u_2} \\
&\mathbf{E}\left[d[M]^{\otimes k}\left(u_{\sigma(1)},\cdots,u_{\sigma(k)}\right) d[M]^{\otimes k}\left(u'_{\sigma(1)},\cdots,u'_{\sigma(k)}\right) d[M]^{\otimes(m-2k)}\left(u_{\sigma(k+1)},\cdots,u_{\sigma(m-k)}\right)\right] \\
&\cdot\left[\Delta G.\left(u_{\sigma(k+1)}\right);\cdots;\Delta G.\left(u_{\sigma(m-k)}\right);\Delta G.\left(u_{\sigma(1)}\right);\Delta G.\left(u_{\sigma(1)}\right);\cdots;\Delta G.\left(u_{\sigma(k)}\right);\Delta G.\left(u_{\sigma(k)}\right)\right] \\
&\cdot\left[\Delta G.\left(u_{\sigma(k+1)}\right);\cdots;\Delta G.\left(u_{\sigma(m-k)}\right);\Delta G.\left(u'_{\sigma(1)}\right);\Delta G.\left(u'_{\sigma(1)}\right);\cdots;\Delta G.\left(u'_{\sigma(k)}\right);\Delta G.\left(u'_{\sigma(k)}\right)\right],
\end{aligned} \tag{24}$$

modulo the fact that one may remove the integrals with respect to those u'_j's not represented among $\{u'_{\sigma(1)},\cdots,u'_{\sigma(k)}\}$. If we can show that for all $k\in\{0,1,2,\cdots,[m/2]\}$ and all $\sigma\in\Sigma_m^k$, the above expression $K = K_{m,k,\sigma}$ tends to 0 as ε tends to 0, the theorem will be proved.

A final note about notation. The bracket notation in the last two lines of the expression (24) above means that we have the product of two separate Riemann integrals over $s \in [0,T]$. Below we will denote these integrals as being with respect to $s \in [0,T]$ and $t \in [0,T]$.

Step 1: diagonal. As in Step 1 of the proofs of Theorems 6 and 8, we can Cauchy-Schwarz to deal with the portion of $K_{m,k,\sigma}$ in (24) where $|s-t| \leq 2\varepsilon$. The details are omitted.

Step 2: term for $k=0$. When $k=0$, there is only one permutation $\sigma = Id$, and we have, using hypothesis (A)

$$\begin{aligned}
K_{m,0,Id} &= \frac{1}{\varepsilon^2}\int_{u_m=0}^{T}\int_{u_{m-1}=0}^{u_m}\cdots\int_{u_1=0}^{u_2}\mathbf{E}\left[d\left[M\right]^{\otimes m}\left(u_1,\cdots,u_m\right)\right]\cdot\left[\Delta G.\left(u_1\right);\cdots;\Delta G.\left(u_m\right)\right]^2 \\
&\leq \frac{1}{\varepsilon^2}\int_{u_{m-k}=0}^{T}\int_{u_{m-k-1}=0}^{u_{m-k}}\cdots\int_{u_1=0}^{u_2}\Gamma^2\left(u_1\right)\Gamma^2\left(u_2\right)\cdots\Gamma^2\left(u_m\right)\left[\Delta G.\left(u_1\right);\cdots;\Delta G.\left(u_m\right)\right]^2 du_1du_2\cdots du_m \\
&= \frac{1}{\varepsilon^2}\int_{u_{m-k}=0}^{T}\int_{u_{m-k-1}=0}^{u_{m-k}}\cdots\int_{u_1=0}^{u_2}\left[\Delta\tilde{G}.\left(u_1\right);\cdots;\Delta\tilde{G}.\left(u_m\right)\right]^2 du_1du_2\cdots du_m.
\end{aligned}$$

This is precisely the expression one gets for the term corresponding to $k=0$ when $M=W$, i.e. when X is the Gaussian process Z with kernel $\tilde{G}$. Hence our hypotheses from the previous two theorems guarantee that this expression tends to 0.

Step 3: term for $k=1$. Again, in this case, $\sigma = Id$, and we thus have, using hypothesis (A),

$$\begin{aligned}
K_{m,1,Id} &= \frac{1}{\varepsilon^2}\int_{u_{m-1}=0}^{T}\int_{u_{m-2}=0}^{u_{m-1}}\cdots\int_{u_1=0}^{u_2}\int_{u_1'=0}^{u_2}\mathbf{E}\left[d\left[M\right]\left(u_1\right)d\left[M\right]\left(u_1'\right)d\left[M\right]^{\otimes(m-2)}\left(u_2,\cdots,u_{m-1}\right)\right] \\
&\cdot\left[\Delta G.\left(u_2\right);\cdots;\Delta G.\left(u_{m-1}\right);\Delta G.\left(u_1\right);\Delta G.\left(u_1\right)\right]\cdot\left[\Delta G.\left(u_2\right);\cdots;\Delta G.\left(u_{m-1}\right);\Delta G.\left(u_1'\right);\Delta G.\left(u_1'\right)\right] \\
&\leq \frac{1}{\varepsilon^2}\int_{u_{m-1}=0}^{T}\int_{u_{m-2}=0}^{u_{m-1}}\cdots\int_{u_1=0}^{u_2}\int_{u_1'=0}^{u_2}du_1du_1'du_2\cdots du_{m-1}\Gamma^2\left(u_1\right)\Gamma^2\left(u_1'\right)\Gamma^2\left(u_2\right)\cdots\Gamma^2\left(u_m\right) \\
&\cdot\left[\left|\Delta G\right|.\left(u_2\right);\cdots;\left|\Delta G\right|.\left(u_{m-1}\right);\left|\Delta G\right|.\left(u_1\right);\left|\Delta G\right|.\left(u_1\right)\right]\cdot\left[\left|\Delta G\right|.\left(u_2\right);\cdots;\left|\Delta G\right|.\left(u_{m-1}\right);\left|\Delta G\right|.\left(u_1'\right);\left|\Delta G\right|.\left(u_1'\right)\right. \\
&= \frac{1}{\varepsilon^2}\int_{u_{m-1}=0}^{T}\int_{u_{m-2}=0}^{u_{m-1}}\cdots\int_{u_1=0}^{u_2}\int_{u_1'=0}^{u_2}du_1du_1'du_2\cdots du_{m-1} \\
&\left[\left|\Delta\tilde{G}\right|_{.}\left(u_2\right);\cdots;\left|\Delta\tilde{G}\right|_{.}\left(u_{m-1}\right);\left|\Delta\tilde{G}\right|_{.}\left(u_1\right);\left|\Delta\tilde{G}\right|_{.}\left(u_1\right)\right] \\
&\cdot\left[\left|\Delta\tilde{G}\right|_{.}\left(u_2\right);\cdots;\left|\Delta\tilde{G}\right|_{.}\left(u_{m-1}\right);\left|\Delta\tilde{G}\right|_{.}\left(u_1'\right);\left|\Delta\tilde{G}\right|_{.}\left(u_1'\right)\right].
\end{aligned}$$

Note now that the product of two bracket operators $[\cdots][\cdots]$ means we integrate over $0 \leq s \leq t-2\varepsilon$ and $2\varepsilon \leq t \leq T$, and get an additional factor of 2, since the diagonal term was dealt with in Step 1.

In order to exploit the additional hypothesis (17) in our theorem, our first move is to use Fubini by bringing the integrals over u_1 all the way inside. We get

$$\begin{aligned}
K_{m,1,Id} \leq \frac{2}{\varepsilon^2}&\int_{u_{m-1}=0}^{T}\int_{u_{m-2}=0}^{u_{m-1}}\cdots\int_{u_2=0}^{u_3}du_2\cdots du_{m-1} \\
&\int_{t=2\varepsilon}^{T}\int_{s=0}^{t-2\varepsilon}ds\ dt\left|\Delta\tilde{G}_s\left(u_2\right)\right|\cdots\left|\Delta\tilde{G}_s\left(u_{m-1}\right)\right|\left|\Delta\tilde{G}_t\left(u_2\right)\right|\cdots\left|\Delta\tilde{G}_t\left(u_{m-1}\right)\right| \\
&\int_{u_1=0}^{u_2}\int_{u_1'=0}^{u_2}du_1du_1'\left(\Delta\tilde{G}_s\left(u_1\right)\right)^2\left(\Delta\tilde{G}_t\left(u_1'\right)\right)^2.
\end{aligned}$$

The term in the last line above is trivially bounded above by

$$\int_{u_1=0}^{T}\int_{u_1'=0}^{T} du_1 du_1' \left(\Delta\tilde{G}_s\left(u_1\right)\right)^2 \left(\Delta\tilde{G}_t\left(u_1'\right)\right)^2$$

precisely equal to $Var\left[Z\left(s+\varepsilon\right)-Z\left(s\right)\right]\ Var\left[Z\left(t+\varepsilon\right)-Z\left(t\right)\right]$, which by hypothesis is bounded above by $\delta^4\left(\varepsilon\right)$. Consequently, we get

$$\begin{aligned}K_{m,1,Id} \leq 2\frac{\delta^4\left(\varepsilon\right)}{\varepsilon^2}&\int_{u_{m-1}=0}^{T}\int_{u_{m-2}=0}^{u_{m-1}}\cdots\int_{u_2=0}^{u_3} du_2\cdots du_{m-1}\\ &\int_{t=2\varepsilon}^{T}\int_{s=0}^{t-2\varepsilon} ds\ dt\left|\Delta\tilde{G}_s\left(u_2\right)\right|\cdots\left|\Delta\tilde{G}_s\left(u_{m-1}\right)\right|\left|\Delta\tilde{G}_t\left(u_2\right)\right|\cdots\left|\Delta\tilde{G}_t\left(u_{m-1}\right)\right|.\end{aligned}$$

We get an upper bound by integrating all the u_j's over their entire range $[0,T]$. I.e. we have,

$$\begin{aligned}&K_{m,1,Id} \leq \frac{\delta^4\left(\varepsilon\right)}{\varepsilon^2}\int_{t=2\varepsilon}^{T} dt\int_{s=0}^{t-2\varepsilon} ds\\ &\int_0^T\int_0^T\cdots\int_0^T du_3\cdots du_{m-1}\left|\Delta\tilde{G}_s\left(u_3\right)\right|\cdots\left|\Delta\tilde{G}_s\left(u_{m-1}\right)\right|\left|\Delta\tilde{G}_t\left(u_3\right)\right|\cdots\left|\Delta\tilde{G}_t\left(u_{m-1}\right)\right|\\ &\cdot\int_{u_2=0}^{T}\left|\Delta\tilde{G}_t\left(u_2\right)\right|\left|\Delta\tilde{G}_s\left(u_2\right)\right| du_2\\ &=2\frac{\delta^4\left(\varepsilon\right)}{\varepsilon^2}\int_{t=2\varepsilon}^{T} dt\int_{s=0}^{t-2\varepsilon} ds\left(\int_0^T du\left|\Delta\tilde{G}_s\left(u\right)\right|\left|\Delta\tilde{G}_t\left(u\right)\right|\right)^{m-3}\cdot\int_{u_2=0}^{u_3}\left|\Delta\tilde{G}_t\left(u_2\right)\right|\left|\Delta\tilde{G}_s\left(u_2\right)\right| du_2..\end{aligned}$$

Now we use a simple Cauchy-Schwarz inequality for the integral over u, but not for u_2. Recognizing that $\int_0^T\left|\Delta\tilde{G}_s\left(u\right)\right|^2 du$ is the variance $Var\left[Z\left(s+\varepsilon\right)-Z\left(s\right)\right]\leq\delta^2\left(\varepsilon\right)$, we have

$$\begin{aligned}K_{m,1,Id} &\leq 2\frac{\delta^4\left(\varepsilon\right)}{\varepsilon^2}\int_{t=2\varepsilon}^{T} dt\int_{s=0}^{t-2\varepsilon} ds\left(\int_0^T du\left|\Delta\tilde{G}_s\left(u\right)\right|^2\right)^{m-3}\cdot\int_{u_2=0}^{u_3}\left|\Delta\tilde{G}_t\left(u_2\right)\right|\left|\Delta\tilde{G}_s\left(u_2\right)\right| du_2.\\ &\leq 2\frac{\delta^{4+2m-6}\left(\varepsilon\right)}{\varepsilon^2}\int_{t=2\varepsilon}^{T} dt\int_{s=0}^{t-2\varepsilon} ds\int_{u_2=0}^{T}\left|\Delta\tilde{G}_t\left(u_2\right)\right|\left|\Delta\tilde{G}_s\left(u_2\right)\right| du_2.\end{aligned}$$

Condition (17) implies immediately $K_{m,1,Id}\leq\delta^{2m}\left(2\varepsilon\right)\varepsilon^{-1}$ which tends to 0 with ε by hypothesis.

Step 4: $k\geq 2$. This step proceeds using the same technique as Step 3. Fix $k\geq 2$. Now for each given permutation σ, there are k pairs of parameters of the type (u,u'). Each of these contributes precisely a term $\delta^4\left(\varepsilon\right)$, as in the previous step, i.e. $\delta^{4k}\left(\varepsilon\right)$ altogether. In other words, for every $\sigma\in\Sigma_m^k$, and deleting the diagonal term, we have

$$\begin{aligned}&K_{m,k,\sigma}\\ &\leq 2\frac{\delta^{4k}\left(\varepsilon\right)}{\varepsilon^2}\int_{t=2\varepsilon}^{T} dt\int_{s=0}^{t-2\varepsilon} ds\int_0^T\int_0^{u_{m-k}}\cdots\int_0^{u_{k+2}} du_{k+1}\cdots du_{m-k}\left[\int_0^T ds\left|\Delta\tilde{G}_s\left(u_{k+1}\right)\right|\cdots\left|\Delta\tilde{G}_s\left(u_{m-k}\right)\right|\right]^2.\end{aligned}$$

Since $k\leq\left(m-1\right)/2$, there is at least one integral, the one in u_{k+1}, above. We treat all the remaining integrals, if any, over $u_{k+2},\cdots,u_{m-k}$ with Cauchy-Schwarz's inequality as in Step 3, yielding a contribution $\delta^{2(m-2k-1)}\left(\varepsilon\right)$. The remaining integral over u_{k+1} yields, by Condition (17),

a contribution of $\delta^2(2\varepsilon)\varepsilon$. Hence the contribution of $K_{m,k,\sigma}$ is again $\delta^{2m}(2\varepsilon)\varepsilon^{-1}$, which tends to 0 with ε by hypothesis, concluding the proof of the Theorem. ■

Proof of Proposition 12. Below the value $1/(2m)-1/2$ is denoted by α. We now show that we can apply Theorem 8 directly to the Gaussian process Z given in (16), which, by Theorem 11, is sufficient, together with Condition (17), to obtain our desired conclusion. Note the assumption about $\tilde{G}$ implies that $s \mapsto \tilde{G}(t,s)$ is square-integrable, and therefore Z is well-defined. We will prove Condition (11) holds in Step 1; Step 2 will show Condition (12) holds; Condition (17) will be established in Step 3.

Step 1. Variance calculation. We need only to show $\tilde{\delta}^2(s,s+\varepsilon) = o\left(\varepsilon^{1/m}\right)$ uniformly in s. We have, for given s and $t = s+\varepsilon$

$$\begin{aligned}\tilde{\delta}^2(s,s+\varepsilon) &= \int_0^s \left|(s+\varepsilon-r)^\alpha f(s+\varepsilon,r) - (s-r)^\alpha f(s,r)\right|^2 dr \\ &\quad + \int_s^{s+\varepsilon} |s+\varepsilon-r|^{2\alpha} f^2(s+\varepsilon,r)\,dr =: A+B \end{aligned} \tag{25}$$

Since $f^2(s+\varepsilon,r) \le f(s+\varepsilon-r)$ and the univariate f increases, in B we can bound this last quantity by $f(\varepsilon)$, and we get

$$B \le f^2(\varepsilon)\int_0^\varepsilon r^{2\alpha}dr = 3f^2(\varepsilon)\varepsilon^{2\alpha+1} = o\left(\varepsilon^{1/m}\right).$$

The term A is slightly more delicate to estimate. Since f is increasing and g is decreasing in t,

$$\begin{aligned} A &\le \int_0^s f^2(s+\varepsilon,r)\left|(s+\varepsilon-r)^\alpha - (s-r)^\alpha\right|^2 dr = \int_0^s f^2(\varepsilon+r)\left|r^\alpha - (r+\varepsilon)^\alpha\right|^2 dr \\ &= \int_0^\varepsilon f^2(\varepsilon+r)\left|r^\alpha-(r+\varepsilon)^\alpha\right|^2 dr + \int_\varepsilon^s f^2(\varepsilon+r)\left|r^\alpha-(r+\varepsilon)^\alpha\right|^2 dr \\ &=: A_1 + A_2. \end{aligned}$$

We have, again from the univariate f's increasingness, and the limit $\lim_{r\to 0} f(r) = 0$,

$$A_1 \le f^2(2\varepsilon)\int_0^\varepsilon \left|r^\alpha-(r+\varepsilon)^\alpha\right|^2 dr = cst\cdot f^2(2\varepsilon)\varepsilon^{2\alpha+1} = o\left(\varepsilon^{1/m}\right).$$

For the other part of A, we need to use f's concavity at the point 2ε in the interval $[0,\varepsilon+r]$ (since $\varepsilon+r>2\varepsilon$ in this case), which implies $f(\varepsilon+r) < f(2\varepsilon)(\varepsilon+r)/(2\varepsilon)$. Also using the mean-value theorem for the difference of negative cubes, we get

$$\begin{aligned} A_2 &\le cst\cdot\varepsilon^2\int_\varepsilon^s f^2(\varepsilon+r)r^{2\alpha-2}dr \le cst\cdot\varepsilon f(2\varepsilon)\int_\varepsilon^s(\varepsilon+r)r^{2\alpha-2}dr \\ &\le cst\cdot\varepsilon f(2\varepsilon)\int_\varepsilon^s r^{2\alpha-1} = cst\cdot\varepsilon^{2\alpha+1}f(2\varepsilon) = o\left(\varepsilon^{1/3}\right). \end{aligned}$$

This finishes the proof of Condition (11).

Step 2. Covariance calculation. We first calculate the second mixed derivative $\partial^2\tilde{\delta}^2/\partial s\partial t$, where $\tilde{\delta}$ is the canonical metric of Z, because we must show $|\mu|(OD) \le \varepsilon^{2\alpha}$, which is condition (12), and $\mu(dsdt) = ds\ dt\ \partial^2\tilde{\delta}^2/\partial s\partial t$. We have, for $0 \le s \le t-\varepsilon$,

$$\tilde{\delta}^2(s,t) = \int_0^s (g(t,s-r)-g(s,s-r))^2 dr + \int_s^t g^2(t,r)\,dr =: A+B.$$

We calculate

$$\begin{aligned}\frac{\partial^2 A}{\partial s \partial t}(t,s) &= 2\frac{\partial g}{\partial t}(t,0)\left(g(t,0)-g(s,0)\right)\\ &+\int_0^s 2\frac{\partial g}{\partial t}(t,s-r)\left(\frac{\partial g}{\partial s}(t,s-r)-\frac{\partial g}{\partial t}(s,s-r)-\frac{\partial g}{\partial s}(s,s-r)\right)\\ &+\int_0^s 2\left(g(t,s-r)-g(s,s-r)\right)\frac{\partial^2 g}{\partial s\partial t}(t,s-r)\,dr.\\ &= A_1+A_2+A_3,\end{aligned}$$

and

$$\frac{\partial^2 B}{\partial s\partial t}(t,s) = -2g(t,s)\frac{\partial g}{\partial t}(t,s).$$

Next, we immediately get, for the portion of $|\mu|(OD)$ corresponding to B, using the hypotheses of our proposition,

$$\begin{aligned}\int_\varepsilon^T dt\int_0^{t-\varepsilon} ds\left|\frac{\partial^2 B}{\partial s\partial t}(t,s)\right| &\le 2c\int_\varepsilon^T dt\int_0^{t-\varepsilon} ds f(|t-s|)\,|t-s|^\alpha\,|t-s|^{\alpha-1}\\ &\le 2c\,\|f\|_\infty\int_\varepsilon^T dt\;\varepsilon^{2\alpha} = cst\cdot\varepsilon^{2\alpha},\end{aligned}$$

which is of the correct order for Condition (12). For the term corresponding to A_1, using our hypotheses, we have

$$\int_\varepsilon^T dt\int_0^{t-\varepsilon} ds\,|A_1| \le 2\int_\varepsilon^T dt\int_0^{t-\varepsilon} ds\;t^\alpha\left|\frac{\partial g}{\partial t}(\xi_{t,s},0)\right||t-s|$$

where $\xi_{t,s}$ is in the interval (s,t). Our hypothesis thus implies $\left|\frac{\partial g}{\partial t}(\xi_{t,s},0)\right| \le s^\alpha$, and hence

$$\int_\varepsilon^T dt\int_0^{t-\varepsilon} ds\,|A_1| \le 2T\int_\varepsilon^T dt\int_0^{t-\varepsilon} ds\;s^{\alpha-1}t^{\alpha-1} = 2T\alpha^{-1}\int_\varepsilon^T dt\;t^{\alpha-1}(t-\varepsilon)^\alpha \le \alpha^{-2}T^{1+2\alpha}.$$

This is much smaller than the right-hand side $\varepsilon^{2\alpha}$ of Condition (12), since $2\alpha = 1/m-1<0$. The terms A_2 and A_3 are treated similarly, thanks to our hypotheses.

Step 3: proving Condition (17). We modify the proof of Theorem 11, in particular Steps 3 and 4, so that we only need to prove

$$\int_{t=2\varepsilon}^T dt\int_{s=0}^{t-2\varepsilon} ds\int_{u=0}^T\left|\Delta\tilde{G}_t(u)\right|\left|\Delta\tilde{G}_s(u)\right|du \le c\varepsilon^{2+2\alpha} = c\varepsilon^{1/m+1}, \tag{26}$$

instead of Condition (17). Indeed, for instance in Step 3, this new condition yields a final contribution of order $\delta^{2m-2}(\varepsilon)\,\varepsilon^{-2}\varepsilon^{1/m+1}$. With the assumption on δ that we have, $\delta(\varepsilon) = o\left(\varepsilon^{1/(2m)}\right)$, and hence the final contribution is of order $o\left(\varepsilon^{(2m-2)/(2m)-1+1/m}\right) = o(1)$. This proves that the conclusion of Theorem 11 holds if we assume (26) instead of Condition (17).

We now prove (26). We can write

$$\begin{aligned}
&\int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \int_{u=0}^{T} \left|\Delta\tilde{G}_t(u)\right| \left|\Delta\tilde{G}_s(u)\right| du \\
&= \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \int_{0}^{s} |g(t+\varepsilon,u) - g(t,u)| \, |g(s+\varepsilon,u) - g(s,u)| \, du \\
&+ \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \int_{s}^{s+\varepsilon} |g(t+\varepsilon,u) - g(t,u)| \, |g(s+\varepsilon,u)| \, du =: A + B.
\end{aligned}$$

For A, we use the hypotheses of this proposition: for the last factor in A, we exploit the fact that g is decreasing in t while f is increasing in t; for the other factor in A, use the bound on $\partial g/\partial t$; thus we have

$$A \leq \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} \varepsilon \, |t-s|^{\alpha-1} \, ds \int_{0}^{s} f(s+\varepsilon,u) \left((s-u)^{\alpha} - (s+\varepsilon-u)^{\alpha}\right) du.$$

We separate the integral in u into two pieces, for $u \in [0, s-\varepsilon]$ and $u \in [s-\varepsilon, s]$. For the first integral in u, since f is bounded, we have

$$\int_{0}^{s-\varepsilon} f(s+\varepsilon,u) \left((s-u)^{\alpha} - (s+\varepsilon-u)^{\alpha}\right) du \leq \|f\|_{\infty} \, \varepsilon \int_{0}^{s-\varepsilon} (s-u)^{\alpha-1} \, du \leq \|f\|_{\infty} \, c_{\alpha} \varepsilon^{1+\alpha}.$$

For the second integral in u, we use the fact that $s-u+\varepsilon > \varepsilon$ and $s-u < \varepsilon$ implies $s-u+\varepsilon > 2(s-u)$, so that the negative part of the integral can be ignored, and thus

$$\int_{s-\varepsilon}^{s} f(s+\varepsilon,u) \left((s-u)^{\alpha} - (s+\varepsilon-u)^{\alpha}\right) du \leq \|f\|_{\infty} \int_{s-\varepsilon}^{s} (s-u)^{\alpha} \, du = \|f\|_{\infty} \, c_{\alpha} \varepsilon^{1+\alpha},$$

which is the same upper bound as for the other part of the integral in u. Thus

$$A \leq cst \cdot \varepsilon^{2+\alpha} \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} |t-s|^{\alpha-1} \, ds \leq cst \cdot \varepsilon^{2+\alpha} \int_{t=2\varepsilon}^{T} dt \; \varepsilon^{\alpha} \leq cst \cdot \varepsilon^{2+2\alpha} = cst \cdot \varepsilon^{1/m+1},$$

which is the conclusion we needed at least for A.

Lastly, we estimate B. We use the fact that f is bounded, and thus $|g(s+\varepsilon,u)| \leq \|f\|_{\infty} |s+\varepsilon-u|^{\alpha}$, as well as the estimate on the derivative of g as we did in the calculation of A, yielding

$$\begin{aligned}
B &\leq \|f\|_{\infty} \, \varepsilon \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \, |t-s-\varepsilon|^{\alpha-1} \int_{s}^{s+\varepsilon} |s+\varepsilon-u|^{\alpha} \, du \\
&= cst \cdot \varepsilon^{\alpha+2} \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \, |t-s-\varepsilon|^{\alpha-1} \\
&\leq 2^{1+|\alpha|} cst \cdot \varepsilon^{\alpha+2} \int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \, |t-s|^{\alpha-1} \leq cst \cdot \varepsilon^{2\alpha+2} = cst \cdot \varepsilon^{1/m+1}.
\end{aligned}$$

This is the conclusion we needed for B,which finishes the proof of the proposition. ■

Proof of Lemma 18.

Step 1: Setup. We only need to show that for all $i, j \in \{1, 2\}$,

$$\iint_{D_{\varepsilon}} |r_{ij}|^{k} \, dudv \leq cst \cdot \varepsilon \delta^{k}(\varepsilon). \tag{27}$$

Recall the function K defined in [20]

$$\begin{aligned} K(u,v) &:= \mathbf{E}\left[\left(X_{u+\varepsilon}+X_u\right)\left(X_{v+\varepsilon}+X_v\right)\right] \\ &= Q(u+\varepsilon,v+\varepsilon)+Q(u,v+\varepsilon)+Q(u+\varepsilon,v)+Q(u,v). \end{aligned}$$

This is not to be confused with the usage of the letter K in previous sections, to which there will be made no reference in this proof; the same remark hold for the notation Δ borrowed again from [20], and used below.

To follow the proof in [20], we need to prove the following items for some constants c_1 and c_2:

1. $c_1\delta^2(u) \le K(u,u) \le c_2\delta^2(u)$;
2. $K(u,v) \le c_2\delta(u)\delta(v)$;
3. $\Delta(u,v) := K(u,u)K(v,v) - K(u,v)^2 \ge c_1\delta^2(u)\delta^2(v-u)$.

By the Theorem's upper bound assumption on the bivariate δ^2 (borrowed from Theorem 8), its assumptions on the monotonicity of Q and the univariate δ, and finally using the coercivity assumption (i), we have

$$\begin{aligned} K(u,u) &= Q_u + Q_{u+\varepsilon} + 2Q(u,u+\varepsilon) = 2\left(Q_u+Q_{u+\varepsilon}\right) - \delta^2(u,u+\varepsilon) \\ &\ge 2\left(Q_u+Q_{u+\varepsilon}\right) - \delta^2(\varepsilon) \ge \left(4-c^{-1}\right)Q_u. \end{aligned}$$

This proves the lower bound in Item 1 above. The upper bound in Item 1 is a special case of Item 2, which we now prove. Again, the assumption borrowed from Theorem 8, which says that $\delta^2(s,t) \le \delta^2(|t-s|)$, now implies, for $s=0$, that

$$\delta^2(0,u) = Q_u \le \delta^2(u). \tag{28}$$

We write, via Cauchy-Schwarz's inequality and the fact that δ^2 is increasing, and thanks to (28),

$$K(u,v) \le 4\delta(u+\varepsilon)\delta(v+\varepsilon).$$

However, since δ^2 is concave with $\delta(0)=0$, we have $\delta^2(2u)/2u \le \delta^2(u)/u$. Also, since we are in the set D_ε, $u+\varepsilon \le 2u$ and $v+\varepsilon \le 2v$. Hence

$$K(u,v) \le 4\delta(2u)\delta(2v) \le 8\delta(u)\delta(v),$$

which is Item 2.

We now verify Item 3 for all $u,v \in D_\varepsilon$, assuming in addition that v is not too small, specifically $v > \varepsilon^{\rho/2}$. One can estimate the integral in Lemma 18 restricted to those values where $v \le \varepsilon^{\rho/2}$ using coarser tools than we use below; we omit the corresponding calculations. From the definition of K above, using the fact that, by our concavity assumptions, Q is, in both variables, a sum of Lipschitz functions, we have, for small ε,

$$K(u,v) = 4Q(u,v) + O(\varepsilon).$$

Therefore,

$$\Delta = 16\left(Q_uQ_v - Q^2(u,v)\right) + O(\varepsilon).$$

Assumption (ii) in the Theorem now implies

$$\Delta \geq 16c'\delta^2(u)\,\delta^2(v-u) + O(\varepsilon).$$

The concavity of Q and Assumption (i) imply $\delta^2(r) \geq Q_r \geq cst \cdot r$. Moreover, because of the restriction on v, either $v-u > cst\cdot\varepsilon^{\rho/2}$ or $u > cst\cdot\varepsilon^{\rho/2}$. Therefore $\delta^2(u)\,\delta^2(v-u) \geq cst\cdot\varepsilon^{1-\rho}\varepsilon^{\rho/2} \gg \varepsilon$. Therefore, for ε small enough, $\Delta \geq 8c'\delta^2(u)\,\delta^2(v-u)$, proving Item 3.

It will now be necessary to reestimate the components of the matrix Λ_{21} where we recall

$$\begin{aligned}
\Lambda_{21}[11] &:= \mathbf{E}\left[(X_{u+\varepsilon}+X_u)(X_{u+\varepsilon}-X_u)\right],\\
\Lambda_{21}[12] &:= \mathbf{E}\left[(X_{v+\varepsilon}+X_v)(X_{u+\varepsilon}-X_u)\right],\\
\Lambda_{21}[21] &:= \mathbf{E}\left[(X_{u+\varepsilon}+X_u)(X_{v+\varepsilon}-X_v)\right],\\
\Lambda_{21}[22] &:= \mathbf{E}\left[(X_{v+\varepsilon}+X_v)(X_{v+\varepsilon}-X_v)\right].
\end{aligned}$$

Step 2: the term r_{11}. We have by the lower bound of item 1 above on $K(u,u)$,

$$|r_{11}| = \left|\frac{1}{\sqrt{K(u,u)}}\Lambda_{21}[11]\right| \leq \frac{cst}{\delta(u)}\left|\Lambda_{21}[11]\right|.$$

To bound $|\Lambda_{21}[11]|$ above, we write

$$\begin{aligned}
|\Lambda_{21}[11]| &= \left|\mathbf{E}\left[(X_{u+\varepsilon}+X_u)(X_{u+\varepsilon}-X_u)\right]\right|\\
&= Q_{u+\varepsilon} - Q_u \leq \varepsilon\delta^2(u)/u
\end{aligned}$$

where we used the facts that Q_u is increasing and concave, and that $Q_u \leq \delta^2(u)$. Thus we have

$$|r_{11}| \leq \varepsilon\; cst\frac{\delta(u)}{u}.$$

The result (27) for $i=j=1$ now follows by the next lemma.

Lemma 20 *For every $k \geq 2$, there exists $c_k > 0$ such that for every $\varepsilon \in (0,1)$, $\int_\varepsilon^1 |\delta(u)/u|^k\,du \leq c_k\varepsilon\,|\delta(\varepsilon)/\varepsilon|^k$.*

Proof of lemma 20. Our hypothesis (iii) can be rewritten as

$$\frac{\delta(au)}{au} < \left(\frac{1+(a-1)b}{a}\right)\frac{\delta(u)}{u} =: K_{a,b}\frac{\delta(u)}{u}.$$

The concavity of δ also implies that $\delta(u)/u$ is increasing. Thus we can write

$$\begin{aligned}
\int_\varepsilon^1 \left|\frac{\delta(u)}{u}\right|^k du &\leq \sum_{j=0}^{\infty}\int_{\varepsilon a^j}^{\varepsilon a^{j+1}} \left|\frac{\delta(u)}{u}\right|^k du \leq \sum_{j=0}^{\infty}\left(\varepsilon a^{j+1}-\varepsilon a^j\right)|K_{a,b}|^{jk}\left|\frac{\delta(\varepsilon)}{\varepsilon}\right|^k\\
&= \varepsilon(a-1)\left|\frac{\delta(\varepsilon)}{\varepsilon}\right|^k \sum_{j=0}^{\infty}\left(|K_{a,b}|^k a\right)^j.
\end{aligned}$$

The lemma will be proved if we can show that $f(a) := |K_{a,b}|^k a < 1$ for some $a > 1$. We have $f(1) = 0$ and $f'(1) = k(1-b) - 1$. This last quantity is strictly positive for all $k \geq 2$ as soon as $b < 1/2$. This finishes the proof of the lemma 20. $\square$

Step 3: the term r_{12}. We have

$$r_{12} = \Lambda_{21}[11] \frac{-K(u,v)}{\sqrt{K(u,u)\Delta(u,v)}} + \Lambda_{21}[12] \frac{\sqrt{K(u,u)}}{\sqrt{\Delta(u,v)}}.$$

We saw in the previous step that $|\Lambda_{21}[11]| = |Q_{u+\varepsilon} - Q_u| \leq cst \cdot \varepsilon\delta^2(u)/u$. For $\Lambda_{21}[12]$, using the hypotheses on our increasing and concave functions, we calculate

$$\begin{aligned}
|\Lambda_{21}[12]| &= \left|2(Q_{u+\varepsilon} - Q_u) + \delta^2(u+\varepsilon, v+\varepsilon) - \delta^2(u, v+\varepsilon) + \delta^2(u+\varepsilon, v) - \delta^2(u,v)\right| \\
&\leq 2|\Lambda_{21}[11]| + \varepsilon\delta^2(u+\varepsilon, v+\varepsilon)/(v-u) + \varepsilon\delta^2(u+\varepsilon, v)/(v-u-\varepsilon) \\
&\leq 2|\Lambda_{21}[11]| + \varepsilon\delta^2(v-u)/(v-u) + \varepsilon\delta^2(v-u-\varepsilon)/(v-u-\varepsilon) \\
&\leq 2cst \cdot \varepsilon\delta^2(u)/u + 2\varepsilon\delta^2(v-u-\varepsilon)/(v-u-\varepsilon). \qquad (29)
\end{aligned}$$

The presence of the term $-\varepsilon$ in the last expression above is slightly aggravating, and one would like to dispose of it. However, since $(u,v) \in D_\varepsilon$, we have $v - u > \varepsilon^\rho$ for some $\rho \in (0,1)$. Therefore $v - u - \varepsilon > \varepsilon^\rho - \varepsilon > \varepsilon^{\rho/2}$ for ε small enough. Hence by using $\rho/2$ instead of ρ in the definition of D_ε in the current calculation, we can ignore the term $-\varepsilon$ in the last displayed line above. Together with items 1, 2, and 3 above which enable us to control the terms K and Δ in r_{12}, we now have

$$\begin{aligned}
|r_{12}| &\leq cst \cdot \varepsilon \frac{\delta^2(u)}{u} \left(\frac{\delta(u)\delta(v)}{\delta(u)\delta(u)\delta(v-u)} + \frac{\delta(u)}{\delta(u)\delta(v-u)} \right) \\
&\quad + cst \cdot \varepsilon \frac{\delta^2(v-u)}{v-u} \frac{\delta(u)}{\delta(u)\delta(v-u)} \\
&= cst \cdot \varepsilon \left(\frac{\delta(u)\delta(v)}{u\delta(v-u)} + \frac{\delta^2(u)}{u\delta(v-u)} + \frac{\delta(v-u)}{v-u} \right).
\end{aligned}$$

We may thus write

$$\iint_{D_\varepsilon} |r_{12}|^k \, dudv \leq cst \cdot \varepsilon^k \iint_{D_\varepsilon} \left(\left| \frac{\delta(u)\delta(v)}{u\delta(v-u)} \right|^k + \left| \frac{\delta^2(u)}{u\delta(v-u)} \right|^k + \left| \frac{\delta(v-u)}{v-u} \right|^k \right) dudv.$$

The last term $\iint_{D_\varepsilon} \left| \frac{\delta(v-u)}{v-u} \right|^k dudv$ is identical, after a trivial change of variables, to the one dealt with in Step 2. Since δ is increasing, second the term $\iint_{D_\varepsilon} \left| \frac{\delta^2(u)}{u\delta(v-u)} \right|^k dudv$ is smaller than the first term $\iint_{D_\varepsilon} \left| \frac{\delta(u)\delta(v)}{u\delta(v-u)} \right|^k dudv$. Thus we only need to deal with that first term; it is more delicate than what we estimated in Step 2.

We separate the integral over u at the intermediate point $v/2$. When $u \in [v/2, v-\varepsilon]$, we use the estimate

$$\frac{\delta(u)}{u} \leq \frac{\delta(v/2)}{v/2} \leq 2\frac{\delta(v)}{v}.$$

On the other hand when $u \in [\varepsilon, v/2]$ we simply bound $1/\delta(v-u)$ by $1/\delta(v/2)$. Thus

$$\begin{aligned}
&\iint_{D_\varepsilon} \left| \frac{\delta(u)\,\delta(v)}{u\delta(v-u)} \right|^k dudv = \int_{v=2\varepsilon}^{1} dv \int_{u=\varepsilon}^{v/2} \left| \frac{\delta(u)\,\delta(v)}{u\delta(v-u)} \right|^k du + \int_{v=\varepsilon}^{1} dv \int_{u=v/2}^{v-\varepsilon} \left| \frac{\delta(u)\,\delta(v)}{u\delta(v-u)} \right|^k du \\
&\leq \int_{v=2\varepsilon}^{1} dv \left| \frac{\delta(v)}{\delta(v/2)} \right|^k \int_{u=\varepsilon}^{v/2} \left| \frac{\delta(u)}{u} \right|^k du + 2\int_{v=\varepsilon}^{1} \left| \frac{\delta^2(v)}{v} \right|^k dv \int_{u=v/2}^{v-\varepsilon} \left| \frac{1}{\delta(v-u)} \right|^k du \\
&\leq 2^k \int_{u=\varepsilon}^{1} \left| \frac{\delta(u)}{u} \right|^k du + 2\frac{1}{\delta^k(\varepsilon)} \int_{v=\varepsilon}^{1} v^k \left| \frac{\delta(v)}{v} \right|^{2k} dv \\
&\leq cst \cdot \varepsilon \left(\frac{\delta(\varepsilon)}{\varepsilon} \right)^k;
\end{aligned}$$

here we used the concavity of δ to imply that $\delta(v)/\delta(v/2) \leq 2$, and to obtain the last line, we used Lemma 20 for the first term in the previous line, and we used the fact that δ is increasing and that $v \leq 1$, together again with Lemma 20 for the second term in the previous line. This finishes the proof of (27) for r_{12}.

Step 4: the term r_{21}. We have

$$r_{21} = \Lambda_{21}[21] \frac{1}{\sqrt{K(u,u)}}$$

and similarly to the previous step,

$$\begin{aligned}
|\Lambda_{21}[21]| &= |Q(u+\varepsilon, v+\varepsilon) - Q(u+\varepsilon, v) + Q(u, v+\varepsilon) - Q(u,v)| \\
&= \left| 2(Q_{v+\varepsilon} - Q_v) + \delta^2(u+\varepsilon, v) - \delta^2(u+\varepsilon, v+\varepsilon) + \delta^2(u,v) - \delta^2(u, v+\varepsilon) \right| \\
&\leq 2|\Lambda_{21}[11]| + \varepsilon \frac{\delta^2(u+\varepsilon, v)}{v-u-\varepsilon} + \varepsilon \frac{\delta^2(u,v)}{v-u} \\
&\leq 2cst \cdot \varepsilon \delta^2(u)/u + 4\varepsilon\delta^2(v-u)/(v-u),
\end{aligned}$$

which is the same expression as in (29). Hence with the lower bound of Item 1 on $K(u,u)$ we have

$$\begin{aligned}
\iint_{D_\varepsilon} |r_{21}|^k dudv &\leq cst \cdot \varepsilon^k \iint_{D_\varepsilon} \left(\left| \frac{\delta(u)}{u} \right|^k + \left| \frac{\delta^2(v-u)}{(v-u)\,\delta(u)} \right|^k \right) dudv \\
&= cst \cdot \varepsilon^k \iint_{D_\varepsilon} \left(\left| \frac{\delta(u)}{u} \right|^k + \left| \frac{\delta^2(u)}{u\delta(v-u)} \right|^k \right) dudv.
\end{aligned}$$

This is bounded above by the expression obtained as an upper bound in Step 3 for $\iint_{D_\varepsilon} |r_{12}|^k dudv$, which finishes the proof of (27) for r_{21}.

Step 5: the term r_{22}. Here we have

$$r_{22} = \Lambda_{21}[21] \frac{-K(u,v)}{\sqrt{K(u,u)\,\Delta(u,v)}} + \Lambda_{21}[22] \frac{\sqrt{K(u,u)}}{\sqrt{\Delta(u,v)}}.$$

We have already seen in the previous step that

$$|\Lambda_{21}[21]| \leq cst \cdot \varepsilon \left(\frac{\delta^2(u)}{u} + \frac{\delta^2(v-u)}{v-u} \right).$$

Moreover, we have, as in Step 2,

$$|\Lambda_{21}[22]| = |Q_{v+\varepsilon} - Q_v| \leq cst \cdot \varepsilon \frac{\delta^2(v)}{v}.$$

Thus using the bounds in items 1, 2, and 3,

$$\begin{aligned} |r_{22}| &\leq cst \cdot \varepsilon \left[\left(\frac{\delta^2(u)}{u} + \frac{\delta^2(v-u)}{v-u} \right) \frac{\delta(u)\delta(v)}{\delta^2(u)\delta(v-u)} + \frac{\delta^2(v)}{v} \frac{\delta(u)}{\delta(u)\delta(v-u)} \right] \\ &= cst \cdot \varepsilon \left[\frac{\delta(u)\delta(v)}{u\delta(v-u)} + \frac{\delta(v)\delta(v-u)}{\delta(u)(v-u)} + \frac{\delta^2(v)}{v\delta(v-u)} \right]. \end{aligned}$$

Of the last three terms, the first term was already treated in Step 3, the second is, up to a change of variable, identical to the first, and the third is smaller than $\frac{\delta^2(u)}{u\delta(v-u)}$ which was also treated in Step 3. Thus (27) is proved for r_{22}, which finishes the entire proof of Lemma 18.

■