Gaussian and non-Gaussian processes of zero power variation

Francesco Russo *†and Frederi VIENS ‡

November 8, 2018

Abstract

This paper considers the class of stochastic processes X defined on $[0,T]$ by $X(t) = \int_0^T G(t,s)\, dM(s)$ where M is a square-integrable martingale and G is a deterministic kernel. When M is Brownian motion, X is Gaussian, and the class includes fractional Brownian motion and other Gaussian processes with or without homogeneous increments. 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 $\delta(r) = o\left(r^{1/(2m)}\right)$.

In the case of a Gaussian process with homogeneous increments, δ is X's canonical metric, the condition on δ is proved to be necessary, and the zero variation result is extended to non-integer symmetric powers, i.e. using $|X(s+\varepsilon) - X(s)|^m \operatorname{sgn}(X(s+\varepsilon) - X(s))$ for any real value $m \geq 1$. In the non-homogeneous Gaussian case, when $m = 3$, the symmetric (generalized Stratonovich) integral is defined, proved to exist, and its Itô formula is proved to hold for all functions of class C^6.

KEY WORDS AND PHRASES: Power variation, martingale Volterra convolution, covariation, 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 order. Before summarizing our results, we give a brief historical description of the topic of p-variations, as a basis for our motivations.

*ENSTA-ParisTech. Unité de Mathématiques appliquées, 32, Boulevard Victor, F-75739 Paris Cedex 15 (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

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

$$\sum_{i=1}^{N-1} |f(t_{i+1}) - f(t_i)|^p. \tag{1}$$

The analytic monograph [6] contains an interesting study on this concept, showing that a p-variation function is the composition of an increasing function and a Hölder-continuous function. The notion of p-variation of a stochastic process or of a function was rediscovered in stochastic analysis, particularly in the context of pathwise (or quasi-pathwise) stochastic calculus. The fundamental paper [11], due to H. Föllmer, treats the case of 2-variations. More recent dealings with p-variations and their stochastic applications, particularly to rough path and other integration techniques, are described at length for instance in the books [7] and [16] , which also contain excellent bibliographies on the subject.

For $p = 2$, the Itô stochastic calculus for semimartingales has mimicked the notion of 2-variation, with the notion of quadratic variation. Let S be a semimartingale; as in (1), consider the expression

$$\sum_{i=1}^{N-1} |S(t_{i+1}) - S(t_i)|^2. \tag{2}$$

One defines the quadratic variation $[S]$ of S as the limit in probability of the expression in (2) as the partition mesh goes to 0, instead of considering the supremum over all partitions. Moreover, the notion becomes stochastic. In fact for a standard Brownian motion B, its 2-variation $[B]$ is a.s. infinite, but its quadratic variation is equal to T. In order to reconcile 2-variations with the finiteness of $[B]$, many authors have proposed restricting the supremum in (1) to the dyadic partitions. However, in Itô calculus, the idea of quadratic variation is associated with the notion of covariation (also known as joint quadratic variation) $[S^1, S^2]$ of two semimartingales S^1, S^2, something which is not present in analytic treatments of p-variation. This covariation $[S^1, S^2]$ is obtained by polarization of (2), i.e. is the limit in probability of $\sum_{i=1}^{N-1} \left(S^1(t_{i+1}) - S^1(t_i)\right)\left(S^2(t_{i+1}) - S^2(t_i)\right)$ when, again, the partition mesh goes to zero.

In the study of stochastic processes, the p-variation has been analyzed in some specific cases, such as local time processes (see [24]), iterated Brownian motion, whose 4-th variation is finite, and more recently fractional Brownian motion (fBm) and related processes. To work with a general class of processes, the tools of Itô calculus would nonetheless restrict the study of covariation to semimartingales. In [27], the authors enlarged the notion of covariation to general processes X and Y. They achieved this by modifying the definition, considering regularizations instead of discretizations. One starting observation is the following. Let $f : [0,T] \to \mathbf{R}$ be continuous. This f has finite variation (i.e. it admits the 1-variation) if and only if $\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^T |f(s+\varepsilon) - f(s)| ds$ exists. In this case, the previous limit equals the total variation of f. An objective was to produce a more efficient stochastic calculus tool, able to go beyond the case of semimartingales. Given two processes X and Y, their covariation $[X,Y](t)$ is the limit in probability, when ε goes to zero, of

$$[X,Y]_\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t \big(X(s+\varepsilon) - X(s)\big)\big(Y(s+\varepsilon) - Y(s)\big) ds; \quad t \geq 0. \tag{3}$$

The limit is again denoted by $[X,Y](t)$ and this notion coincides with the classical covariation when X,Y are continuous semimartingales. The processes X such that $[X,X]$ exists are called finite quadratic variation processes; their analysis and their applications were performed in [10, 26].

The notion of covariation was also extended in order to involve more processes. In [9] the authors consider the n-covariation $[X^1, X^2, \cdots, X^n]$ of n processes $X^1, \ldots, X^n$, as in formula (3), but with a product of n increments rather than just two. For $n = 4$, for X being an fBm with so-called *Hurst* parameter $H = 1/4$, the paper [13] calculates the 4-covariation $[g(X), X, X, X]$ where g is, for instance, a bounded continuous function. If $X = X^1 = X^2 = X^3$ is a single stochastic process, 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. Note that this variation 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 fractional Brownian motion B^H with Hurst parameter $H \in (0,1)$. Then when $H > 1/6$, in the introduction of [14] is established that the cubic variation $[X,3]$ equals zero, while in [14, Theorem 4.1 part 2(c)] it is shown that $[X,3]$ does not exist if $H < 1/6$. In the limiting case of $H = 1/6$, [14, Theorem 4.1 part 2(b)] shows that the regularization approximation $[X,3]_\varepsilon(t)$ converges in law to a normal for every $t > 0$. This phenomenon is confirmed in the related study of finite-difference approximating sequence of $[X,3]$: for instance one may use the so-called Breuer-Major central limit theorem for stationary Gaussian sequences [4], to prove that the finite-difference approximating sequence of $[X,3](t)$ converges in law to a Gaussian variable; this was noted in [23, Theorem 10], where the authors prove that more is true: considered as a process depending on the upper endpoint of the time interval, the approximation converges in law to κW where W is an independent Brownian motion, and κ is a universal constant given by

$$\kappa^2 = \frac{3}{4} \sum_{r \in \mathbb{Z}} (|r+1|^{\frac{1}{3}} + |r-1|^{\frac{1}{3}} - 2|r|^{\frac{1}{3}}).$$

Beyond a basic interest in the variations of non-semimartingale stochastic processes, the significance of the cubic variation lies in its ability to guarantee the existence of (generalized symmetric) Stratonovich integrals, and their associated Itô formula, for highly irregular processes, notably fBm with $H > 1/6$ (such fBm have Hölder regularity parameter exceeding 1/6; compare with the near 1/2-Hölder-regularity for continuous semimartingales). This existence of a quite general Itô-Stratonovich formula is a relatively well-known phenomenon, established in great generality in [14]. This result is a main motivation for our work, in which we attempt to give specific classes of Gaussian and non-Gaussian processes with the said zero cubic variation property; indeed the conditions in [14] are used therein specifically with the Gaussian class of fBms, which are fractionally self-similar and have stationary increments; the methods in [14] can be extended only to similar cases, i.e. Gaussian processes with canonical metrics that are bounded above and below by multiples of the fBm's, for instance the bi-fractional Brownian motion treated in [25]. Regarding the existence of Itô-Stratonovich formulas, it is instructive to recall a variant on [14] established in [9]: if $f : \mathbf{R} \to \mathbf{R}$ is a function of class C^3 and X has a strong cubic variation, then the following Itô type formula holds:

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) d^\circ X - \frac{1}{12} \int_0^t f'''(X_s) d[X,3](s), \tag{4}$$

and the stochastic integral in the right-hand side is the symmetric-Stratonovich integral introduced for instance in [28], while the other is a Lebesgue-Stieltjes integral. The problem is that until now no examples are known of processes X which have a cubic variation $[X,3]$ which exists but does not vanish. In [21], an analogous formula to (4) is obtained for the case $X = B^H$ with $H = 1/6$, but in the sense of distributions only: the symmetric integral has to be interpreted as existing in law and the integral with respect to the cubic variation makes sense by replacing $[X,3]$ with the term κW, W being the independent Wiener process identified in [23], so that $\int_0^t f'''(X_s)d[X,3](s)$ is merely defined in law as a conditionally Wiener integral.

1.2 Specific motivations

Following the regularization methodology of [27] or [28], the cubic variation of a process X, denoted by $[X,3](t)$, was defined similarly to [9] as the limit in probability or in the mean square, as $\varepsilon \to 0$, of

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

This was already mentioned above. This $[X,3]$ will be null for a deterministic function X as long as it is α-Hölder-continuous with $\alpha > 1/3$. But the main physical reason for being interested in this cubic variation for random processes is that, because the cube function is symmetric, if the process X itself has some probabilistic symmetry as well (such as the Gaussian property and the stationarity of increments), then we can expect $[X,3]$ to be 0 for much more irregular processes than those which are almost-surely α-Hölder-continuous with $\alpha > 1/3$. As mentioned above, [9] proves that fBm has zero cubic variation as soon as $H > 1/6$, in spite of the fact that fBm is only α-Hölder-continuous almost surely for all $\alpha < H$. This doubling improvement over the deterministic situation is due exclusively to the random symmetries of fBm, as they combine with the fact that the cube function is odd. Typically for other, non-symmetric types of variations, H needs to be larger to guarantee existence of the variation, let alone nullity; 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 $H = 1/6$. We have already observed that this threshold represents a critical value in terms of existence of cubic variation, for fBm: we mentioned that whether in the sense of regularization or of finite-difference, the approximating sequences of $[X,3](t)$ converge in law to Gaussian laws. On the other hand, these normal convergences contrast with one further point in the study of variations for fBm: 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, as we said above.

These properties of fBm beg the question of what occurs for other Gaussian processes which may not be self-similar or even have stationary increments, or even for non-Gaussian processes with similar α-Hölder-continuous paths, and to what extent the threshold $\alpha > 1/6$ is sharp. Similarly, can the odd symmetry of the cube function be generalized to any "symmetric" power function, i.e. $x \mapsto |x|^m \mathrm{sgn}(x)$ with arbitrary integer or non-integer $m > 1$? This refers to what we will call the "odd mth variation", defined (when it exists in the mean-square sense) by

$$[X,m](t) := \lim_{\varepsilon \to 0} \varepsilon^{-1} \int_0^t |X(s+\varepsilon) - X(s)|^m \, \mathrm{sgn}\,(X(s+\varepsilon) - X(s)) \, ds. \tag{5}$$

The qualifier "odd" above, when applied to $m = 3$, can easily yield the term "odd cubic variation", which has historically been called simply "cubic variation" as we noted before, as opposed to the "strong cubic variation" which involves absolute values; therefore in this article, we will systematically use the qualifier "odd" for all higher order mth variations based on odd functions, but will typically omit it for the cubic variation.

1.3 Summary of results

This article provides answers to some of the above questions, both in Gaussian and non-Gaussian settings, and we hope that it will stimulate work on resolving some of the remaining open problems. Specifically, we consider the process X defined on $[0,T]$ by

$$X(t) = \int_0^T G(t,s)\, dM(s) \tag{6}$$

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. In other words, X is defined using a Volterra representation with respect to a square-integrable martingale. The quadratic variations of these martingale-based convolutions was studied in [8].

What we call the Gaussian case is that in which M is the standard Wiener process (Brownian motion) W. The itemized list below is a summary of our results. Here, for the reader's convenience, we have not spelled out the technical conditions which are needed for some of our results, indicating instead references to the precise theorem statements in the body of this article. Some conditions become more restrictive as one move from simple Gaussian cases to non-Gaussian cases. Yet we cover much wider classes of processes than has been done in the past. The summary below also provides indications of how wide a scope we reach, and has references to examples in the main body of the paper.

One condition which appears in all cases, is essentially equivalent to requiring that all processes X that we consider are not more regular than standard Brownian motion, i.e. are not 1/2-Hölder-continuous. This typically takes the form of a concavity condition on the process's squared canonical metric $\delta^2(s,t) := \mathbf{E}\left[(X(t) - X(s))^2\right]$. This condition is not a restriction on the range of path regularity, since the main interest of our results occurs around the Hölder exponent 1/6, or more generally the exponent $1/(2m)$ for any $m > 1$: the processes with zero odd mth variation appear as those which are better than $1/(2m)$-Hölder-continuous in the $L^2(\Omega)$-sense. Processes which are better than 1/2-Hölder-continuous are not covered by this paper, but can be treated using classical non-probabilistic tools such as the Young integral.

We now give a summary of all our results. For any number $m \geq 2$, let

$$[X,m]_\varepsilon(t) := \frac{1}{\varepsilon} \int_0^t ds\, |X(s+\varepsilon) - X(s)|^m \operatorname{sgn}(X(s+\varepsilon) - X(s))$$

where $\operatorname{sgn}(x)$ is the sign function $x/|x|$. The limit in probability of $[X,m]_\varepsilon(t)$ as $\varepsilon \to 0$ is the "odd mth variation" of X at time t, denotd by $[X,m](t)$. Except for some results in Section 5, the results in this paper are stated without loss of generality for a fixed value of $t \leq T$, and we typically take $t = T$; we occasionally drop the dependence on T, writing only $[X,m]_\varepsilon$ and $[X,m]$.

- [**Homogeneous Gaussian case, odd powers:** Theorem 6 on page 13]. When X is Gaussian with homogeneous increments (meaning $\delta(s,t)$ depends only on $|t-s|$), for any odd integer $m \geq 3$, X has zero odd mth variation if *and only* if $\delta(r) = o\left(r^{1/(2m)}\right)$ for r near 0.

– This theorem does not require any assumptions beyond δ being increasing and concave.

- [**Homogeneous Gaussian case, arbitrary real powers:** Theorem 10 on page 22]. The sufficient condition of the result above holds for any integer $m > 1$, and for any real non-integer > 1 modulo a mild technical condition.

  – This theorem extends the previous result to even powers without requiring any additional assumptions, and to all real powers under a technical regularity assumption on the covariance which places no regularity restrictions on the paths of X (see Remark 11).

- [**Non-homogeneous Gaussian case:** Theorem 8 on page 19]. When X is Gaussian with non-homogeneous increments, for any odd integer $m \geq 3$, if $\delta^2(s, s+r) = o\left(r^{1/(2m)}\right)$ for r near 0 uniformly in s, and under a technical condition, X has zero odd mth variation.

  – The technical condition is a non-explosion assumption on the mixed partial derivative of δ^2 near the diagonal. It places no regularity restriction on the paths of X. The description on page 20 shows that the condition is satisfied for the so-called Riemann-Liouville version of fBm, and for a wide class of Volterra-convolution-type Gaussian processes with inhomogeneous increments.

- [**Non-Gaussian martingale case:** Theorem 12 on page 26]. Let $m \geq 3$ be an odd integer. When X is non-Gaussian as in (6), based on a martingale M whose quadratic variation process has a derivative with $2m$ moments (the actual condition on M in the theorem is weaker), let $\Gamma(t) = \left(\mathbf{E}\left[(d[M]/dt)^m\right]\right)^{1/(2m)}$ and consider the Gaussian process

$$Z(t) = \int_0^T \Gamma(s) G(t,s) \, dW(s).$$

  Under a technical integrability condition on planar increments of ΓG near the diagonal, if Z satisfies the conditions of Theorem 6 or Theorem 8, then X has zero odd mth variation.

  – Proposition 13 on page 30 provides examples of wide classes of martingales and kernels for which the assumptions of Theorem 12 are satisfied. Details on how to construct these examples, and how to evaluate their regularity properties, are given on page 4.

  – A key consequence of Proposition 13 and Theorem 12 is that this paper's results extend from the Gaussian case to highly non-Gaussian situations, insofar as, for m an odd integer, it is easy to construct a variety of martingales M with no more than m moments, which are comparable to their Gaussian analogues in terms of path regularity, and for which the corresponding X in (7) has null odd mth variation. This is explained on page 4.

  – It is important to note that while the base process M used here is a martingale, the process X in (7) whose variation we study is as far from being a martingale as fBm is.

- [**Itô formula:** Theorem 14 on page 35, and its corollary]. When $m \geq 3$ is an odd integer and X is a Gaussian process with non-homogeneous increments such that $\delta^2(s, s+r) = o\left(r^{1/(2m)}\right)$

uniformly in s, under some additional technical conditions, 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 [14], Theorem 14 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. This formula is in Corollary 15 on page 36.

– The scope of the technical conditions needed for the theorem and its corollary is discussed immediately after the corollary. These conditions include similar monotonicity and concavity conditions as are used in the remainder of the article, plus some coercivity conditions ensuring that the process X is not too far from having homogeneous increments. The discussion after Corollary 15 establishes that the coercivity conditions are satisfied in the homogeneous case.

1.4 Relation with other recent work

We finish this introduction with a description of recent work done by several other authors on problems related to our preoccupations to some extent, in various directions. The authors of the paper [15] consider, as we do, stochastic processes which can be written as Volterra integrals with respect to martingales. In fact, they study the concept of "fractional martingale", which is the generalization of the so-called Riemann-Liouville fractional Brownian motion when the driving noise is a martingale. This 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 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 fractional Brownian motion. The paper [15] does not follow, however, the same motivation as our work: for us, say in the case of $m=3$, we study the threshold $H>1/6$ for vanishing (odd) cubic variations in various Gaussian and non-Gaussian contexts, and its relation to stochastic calculus.

To find a similar motivation to ours, one may look at the recent result of [20], where the authors study the central and non-central behavior of weighted Hermite variations for fBm. Using the Hermite polynomial of order q rather than the power-q function, they show that the threshold value $H=1/(2q)$ poses an interesting open problem, since above this threshold (but below $H=1-1/(2q)$) one obtains Gaussian limits (these limits are conditionally Gaussian when weights are present, and can be represented as stochastic integrals with respect to an independent Brownian motion), while below the threshold, degeneracy occurs. The behavior at the threshold was worked out for $H=1/4, q=2$ in [20], boasting an exotic correction term with an independent Brownian motion, while the general open problem of Hermite variations with $H=1/(2q)$ was settled in [19]. More questions arise, for instance, with a similar result in [18] for $H=1/4$, but this time with bidimensional fBm, in which two independent Brownian motions are needed to characterize the exotic correction term.

The value $H=1/6$ is mentioned again 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 [5] and [23].

Summarizing, when compared to the works described in the above paragraphs, our work situates itself by

- choosing to prove necessary and sufficient conditions for nullity of the cubic variation, around the threshold regularity value $H = 1/6$, for Gaussian processes with homogeneous increments (this is a wider class than previously considered, showing in particular that self-similarity is not related to the question of nullity of the cubic variation);
- studying the nullity threshold for higher order "odd" power functions, with possibly non-integer order, showing that this property relies only on the symmetry of Gaussian processes with homogeneous increments and on the symmetrization of the power functions;
- 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.

The article has the following structure. Section 2 contains some formal definitions and notations. The basic theorems in the Gaussian case are in Section 3, where the homogeneous case, non-homogeneous case, and case of non-integer m are separated in three subsections. The use of non-Gaussian martingales is treated in Section 4. Section 5 presents the Itô formula.

2 Definitions

We recall our process X defined for all $t \in [0,T]$ by

$$X(t) = \int_0^T G(t,s)\, dM(s) \tag{7}$$

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. For any real number $m \geq 2$, let the *odd ε-m-th variation* of X be defined by

$$[X,m]_\varepsilon(T) := \frac{1}{\varepsilon}\int_0^T ds\, |X(s+\varepsilon) - X(s)|^m \operatorname{sgn}(X(s+\varepsilon) - X(s)). \tag{8}$$

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

$$\lim_{\varepsilon\to 0}[X,m]_\varepsilon(T) = 0 \tag{9}$$

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 *homogeneous 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 *homogeneous 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 fractional Brownian motion, while our definition does. In non-Gaussian settings, the usual way to interpret the "homogeneous" 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 the limit in (9) being zero 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 as soon as

$$\delta(s,t) = o\left(|t-s|^{1/(2m)}\right), \tag{10}$$

and that this is a necessary condition in some cases. Although this is a mean-square condition, it can be interpreted as a regularity (local) condition on X; for example, when X is a Gaussian process with homogeneous increments, this condition means precisely that almost surely, the uniform modulus of continuity ω of X on any fixed closed interval, defined by $\omega(r) = \sup\{|X(t) - X(s)| : |t-s| < r\}$, satisfies $\omega(r) = o\left(r^{1/6}\log^{1/2}(1/r)\right)$. The lecture notes [1], as well as the article [29], can be consulted for this type of statement.

3 Gaussian case

We assume that X is centered Gaussian. Then we can write X as in formula (7) with $M = W$ a standard Brownian motion. More importantly, beginning with the easiest case where m is an odd integer, we can easily show the following.

Lemma 1 *If m is an odd integer ≥ 3, we have*

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

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

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

Proof. The lemma is an easy consequence of the following formula, which can be found as Lemma 5.2 in [14]: for any centered jointly Gaussian pair of r.v.'s (Y,Z), we have

$$\mathbf{E}\left[Y^m Z^m\right] = \sum_{j=0}^{(m-1)/2} c_j \mathbf{E}\left[YZ\right]^{m-2j}\ Var\left[X\right]^j\ Var\left[Y\right]^j.$$

■

We may translate $\Theta^\varepsilon(s,t)$ immediately in terms of Q, and then δ. We have:

$$\begin{aligned}\Theta^\varepsilon(s,t) &= Q(t+\varepsilon,s+\varepsilon) - Q(t,s+\varepsilon) - Q(s,t+\varepsilon) + Q(s,t) \\ &= \frac{1}{2}\left[-\delta^2(t+\varepsilon,s+\varepsilon) + \delta^2(t,s+\varepsilon) + \delta^2(s,t+\varepsilon) - \delta^2(s,t)\right] \qquad (11) \\ &=: -\frac{1}{2}\Delta_{(s,t);(s+\varepsilon,t+\varepsilon)}\delta^2. \qquad (12)\end{aligned}$$

Thus $\Theta^\varepsilon(s,t)$ appears as the opposite of the planar increment of the canonical metric over the rectangle defined by its corners (s,t) and $(s+\varepsilon,t+\varepsilon)$.

3.1 The case of 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 [14, 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 (7) where W is Brownian motion and the kernel G is well-known. Information on G and on the Wiener chaos generated by W can be found respectively in Chapters 5 and 1 of the textbook [22]. The covariance formula for an fBm X is

$$R_H(s,t) := \mathbf{E}\left[X(t)X(s)\right] = 2^{-1}\left(s^{2H} + t^{2H} - |t-s|^{2H}\right). \qquad (13)$$

Lemma 3 *Fix $\varepsilon>0$. Let $\Delta X_s := X(s+\varepsilon) - X(s)$ and $\Delta G_s(u) := G(s+\varepsilon,u) - G(s,u)$. Then*

$$\begin{aligned}[X,3]_\varepsilon(T) &= \mathcal{I}_1 + \mathcal{I}_3 \\ &=: \frac{3}{\varepsilon}\int_0^T ds \int_0^T \Delta G_s(u)\,dW(u)\left(\int_0^T |\Delta G_s(v)|^2\,dv\right) \qquad (14) \\ &\quad + \frac{6}{\varepsilon}\int_0^T dW(s_3)\int_0^{s_3} dW(s_2)\int_0^{s_2} dW(s_1)\int_0^T \left[\prod_{k=1}^3 \Delta G_s(s_k)\right] ds. \qquad (15)\end{aligned}$$

Proof. The proof of this lemma is elementary. It follows from two uses of the multiplication formula for Wiener integrals [22, Proposition 1.1.3], for instance. It can also be obtained directly from Lemma 7 below, or using the Itô formula technique employed further below in finding an expression for $[X,m]_\varepsilon(T)$ in Step 0 of the proof of Theorem 12 on page 26. All details are left to the reader. ■

The above lemma indicates the Wiener chaos decomposition of $[X,3]_\varepsilon(T)$ into the term $\mathcal{I}_1$ of line (14) which is in the first Wiener chaos (i.e. a Gaussian term), and the term $\mathcal{I}_3$ of line (15), 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 *The Gaussian term $\mathcal{I}_1$ converges to 0 in $L^2(\Omega)$ as $\varepsilon \to 0$.*

Lemma 5 *The 3rd chaos term $\mathcal{I}_3$ 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 $\mathbf{E}[|\mathcal{I}_1|^p] \leq c_p\left(\mathbf{E}\left[|\mathcal{I}_1|^2\right]\right)^{p/2}$ and $\mathbf{E}[|\mathcal{I}_3|^p] \leq c_p\left(\mathbf{E}\left[|\mathcal{I}_3|^2\right]\right)^{p/2}$; 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. 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. ■

To conclude this section, we only need to prove the above two lemmas. To improve readability, we write H instead of $1/6$.

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

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

and therefore,

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

We note here that $\mathbf{E}\left[|\mathcal{I}_1|^2\right]$ coincides with what we called J_1 in Lemma 1, but we will not use this fact here. Instead, using the variances of fBm,

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

Bounding the variance of the difference by twice the sum of the variances, and using the fBm covariance formula (13),

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

proving Lemma 4. ■

Proof of Lemma 5. By the proof of Lemma 3, and using the covariance formula (13) for fBm, we first get

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

Again, this expression coincides with the term J_0 from Lemma 1, but this will not be used in this proof. 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

$$\begin{aligned}|\mathcal{ODI}_3| &\leq 384H^3|2H-1|^3\varepsilon^{-2}\int_{2\varepsilon}^T\int_0^{t-2\varepsilon}\left(\varepsilon\cdot 2\varepsilon\cdot(t-s-\varepsilon)^{2H-2}\right)^3 dtds\\ &= 384H^3|2H-1|^3\varepsilon^4\int_{2\varepsilon}^T\int_0^{t-2\varepsilon}(t-s-\varepsilon)^{6H-6}\,dtds\\ &= \frac{384H^3|2H-1|^3\varepsilon^4}{5-6H}\int_{2\varepsilon}^T\left[\varepsilon^{6H-5} - (t-\varepsilon)^{6H-5}\right]dt\\ &\leq \frac{384H^3|2H-1|^3}{5-6H}T\varepsilon^{6H-1} = \frac{384H^3|2H-1|^3}{5-6H}T = \frac{32}{243}T.\end{aligned}$$

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

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

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 [14, 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. This concludes the proof of Lemma 5. ■

3.2 The homogeneous case

We now study the homogeneous case in detail. We are ready to prove a necessary and sufficient condition for having a zero m-th variation when m is an odd integer.

Theorem 6 *Let $m>1$ be an odd integer. Let X be a centered Gaussian process on $[0,T]$ with homogeneous 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.* We denote by $d\delta^2$ the derivative, in the sense of measures, of δ^2; we know that $d\delta^2$ is a positive bounded measure on $[0,T]$. Using homogeneity, we also get

$$Var\left[X(t+\varepsilon)-X(t)\right]=\delta^2(\varepsilon).$$

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. Let us deal first with the diagonal term. We define the ε-diagonal $D_\varepsilon:=\{0\le t-\varepsilon<s<t\le T\}$. Trivially using Cauchy -Schwarz's inequality, we have

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

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:

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

where cst denotes a constant whose value may change in the remainder of the article's proofs (here it depends only on δ and m). 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

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

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

Step 3: off-diagonal. Because of the homogeneity hypothesis, we can calculate from (11) that for any s,t in the ε-off diagonal set $OD_{\varepsilon} := \{0 \leq s < t - \varepsilon < t \leq T\}$

$$\begin{aligned} \Theta^{\varepsilon}(s,t) &= \left(\delta^2(t-s+\varepsilon) - \delta^2(t-s)\right) - \left(\delta^2(t-s) - \delta^2(t-s-\varepsilon)\right) \\ &= \int_{t-s}^{t-s+\varepsilon} d\delta^2(r) - \int_{t-s-\varepsilon}^{t-s} d\delta^2(r). \end{aligned} \tag{16}$$

By the concavity hypothesis, we see that $\Theta^{\varepsilon}(s,t)$ is negative in this off-diagonal set OD_{ε}. Unfortunately, using the notation in Lemma 1, this negativity does not help us because the off-diagonal portion $J_{j,OD}$ of J_j also involves the constant c_j, which could itself be negative. Hence we need to estimate $J_{j,OD}$ more precisely.

The constancy of the sign of Θ^{ε} is still useful, because it enables our first operation in this step, which is to reduce the estimation of $|J_{j,OD}|$ to the case of $j = (m-1)/2$. Indeed, using Cauchy-Schwarz's inequality and the fact that $|\Theta^{\varepsilon}| = -\Theta^{\varepsilon}$, we write

$$\begin{aligned} |J_{j,OD}| &= \varepsilon^{-2}\delta^{4j}(\varepsilon)\, |c_j| \int_{\varepsilon}^{T} dt \int_0^{t-\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right|^{m-2j} \\ &= -\varepsilon^{-2}\delta^{4j}(\varepsilon)\, |c_j| \int_{\varepsilon}^{T} dt \int_0^{t-\varepsilon} ds \Theta^{\varepsilon}(s,t) \left|\Theta^{\varepsilon}(s,t)\right|^{m-2j-1} \\ &\leq \varepsilon^{-2}\delta^{4j}(\varepsilon)\, |c_j| \int_{\varepsilon}^{T} dt \int_0^{t-\varepsilon} ds \left(-\Theta^{\varepsilon}(s,t)\right) \left|\delta^2(\varepsilon)\right|^{m-2j-1} \\ &= \varepsilon^{-2}\delta^{2m-2}(\varepsilon)\, |c_j| \int_{\varepsilon}^{T} dt \int_0^{t-\varepsilon} ds \left(-\Theta^{\varepsilon}(s,t)\right). \end{aligned}$$

It is now sufficient to show that the estimate for the case $j = (m-1)/2$ holds, i.e. that

$$\int_{\varepsilon}^{T} dt \int_0^{t-\varepsilon} ds \left(-\Theta^{\varepsilon}(s,t)\right) \leq cst \cdot \varepsilon \delta^2(2\varepsilon) \tag{17}$$

We rewrite the planar increments of δ^2 as in (16) to show what cancellations occur: with the notation $s' = t - s$,

$$-\Theta^{\varepsilon}(s,t) = -\left(\delta^2(s'+\varepsilon) - \delta^2(s')\right) + \left(\delta^2(s') - \delta^2(s'-\varepsilon)\right) = -\int_{s'}^{s'+\varepsilon} d\delta^2(r) + \int_{s'-\varepsilon}^{s'} d\delta^2(r).$$

Therefore, using the change of variables from s to s', and another to change $[s'-\varepsilon, s']$ to $[s', s'+\varepsilon]$,

$$\begin{aligned}
\int_{\varepsilon}^{T} dt \int_{0}^{t-\varepsilon} ds \left(-\Theta^{\varepsilon}(s,t)\right) &= \int_{\varepsilon}^{T} dt \left[\int_{\varepsilon}^{t} ds' \int_{s'-\varepsilon}^{s'} d\delta^2(r) - \int_{\varepsilon}^{t} ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \\
&= \int_{\varepsilon}^{T} dt \left[\int_{\varepsilon}^{t} ds' \int_{s'-\varepsilon}^{s'} d\delta^2(r) - \int_{\varepsilon}^{t} ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \\
&= \int_{\varepsilon}^{T} dt \left[\int_{0}^{t-\varepsilon} ds'' \int_{s''}^{s''+\varepsilon} d\delta^2(r) - \int_{\varepsilon}^{t} ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \\
&= \int_{\varepsilon}^{T} dt \left[\int_{0}^{\varepsilon} ds'' \int_{s''}^{s''+\varepsilon} d\delta^2(r) - \int_{t-\varepsilon}^{t} ds' \int_{s'}^{s'+\varepsilon} d\delta^2(r)\right] \qquad (18)
\end{aligned}$$

We may now invoke the positivity of $d\delta^2$, to obtain

$$\begin{aligned}
&\int_{\varepsilon}^{T} dt \int_{0}^{t-\varepsilon} ds \left(-\Theta^{\varepsilon}(s,t)\right) \leq \int_{\varepsilon}^{T} dt \int_{0}^{\varepsilon} ds'' \int_{s''}^{s''+\varepsilon} d\delta^2(r) \\
&= \int_{\varepsilon}^{T} dt \int_{0}^{\varepsilon} ds'' \left(\delta^2\left(s''+\varepsilon\right) - \delta^2\left(s''\right)\right) \leq \int_{\varepsilon}^{T} dt \; \varepsilon \; \delta^2(2\varepsilon) \leq T\varepsilon\delta^2(2\varepsilon).
\end{aligned}$$

This is precisely the claim in (17), which finishes the proof that for all j, $|J_{j,OD}| \leq cst \cdot \varepsilon^{-1}\delta^{2m}(2\varepsilon)$ for some constant cst. Combining this with the results of Steps 1 and 2, we obtain that

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

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 lemma 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. This lemma's elementary proof is left to the reader.

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 (7) and (8), 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

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

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

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

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 $|J_{1,OD}|$. Using the calculations performed in Step 3 (note here that $(m-1)/2=1$, in particular line (18), we have

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

We can already see that $K_1\geq 0$ and $K_2\leq 0$, so it's only necessary to find an upper bound on $|K_2|$; but in reality, the reader will easily check that $|K_1|$ is of the order $\delta^6(\varepsilon)$, and we will see that this is much smaller than either $J_{1,D}$ or $|K_2|$. Performing a Fubini on the variables s and r, the integrand in K_2 is calculated as

$$\begin{aligned}
\int_0^{\varepsilon} ds \int_s^{s+\varepsilon} d\delta^2(r) &= \int_{r=0}^{\varepsilon} d\delta^2(r) \int_{s=0}^{r} ds + \int_{r=\varepsilon}^{2\varepsilon} d\delta^2(r) \int_{s=r-\varepsilon}^{\varepsilon} ds \\
&= \int_{r=0}^{\varepsilon} r \; d\delta^2(r) + \int_{r=\varepsilon}^{2\varepsilon} (2\varepsilon - r)\, d\delta^2(r) \\
&= \left[r\delta^2(r)\right]_0^{\varepsilon} - \left[r\delta^2(r)\right]_{\varepsilon}^{2\varepsilon} - \int_0^{\varepsilon} \delta^2(r)\, dr + \int_{\varepsilon}^{2\varepsilon} \delta^2(r)\, dr + 2\varepsilon\left(\delta^2(2\varepsilon) - \delta^2(\varepsilon)\right) \\
&= -\int_0^{\varepsilon} \delta^2(r)\, dr + \int_{\varepsilon}^{2\varepsilon} \delta^2(r)\, dr.
\end{aligned}$$

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

$$\begin{aligned}
|J_{1,OD}| &\leq \frac{9\delta^4(\varepsilon)}{\varepsilon^2} \int_{\varepsilon}^{T} dt \left(\int_{\varepsilon}^{2\varepsilon} \delta^2(r)\, dr - \int_0^{\varepsilon} \delta^2(r)\, dr\right) \\
&= \frac{9(T-\varepsilon)\,\delta^4(\varepsilon)}{\varepsilon^2} \left(\int_{\varepsilon}^{2\varepsilon} \delta^2(r)\, dr - \int_0^{\varepsilon} \delta^2(r)\, dr\right). \qquad (19)
\end{aligned}$$

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

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

Therefore

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

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

$$\begin{aligned}
\int_0^{\varepsilon} ds \int_s^{s+\varepsilon} d\delta^2(r) &= \int_0^{\varepsilon} r \; d\delta^2(r) + \varepsilon \int_{\varepsilon}^{2\varepsilon} d\delta^2(r) \\
&= \left[r\delta^2(r)\right]_0^{\varepsilon} - \int_0^{\varepsilon} \delta^2(r)\, dr + \varepsilon\left(\delta^2(2\varepsilon) - \delta^2(\varepsilon)\right) \\
&= \varepsilon\delta^2(2\varepsilon) - \int_0^{\varepsilon} \delta^2(r)\, dr.
\end{aligned}$$

In other words,

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

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

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

When δ is in the Hölder scale $\delta(r) = r^H$, the above quantity is obviously commensurate with $\delta^6(\varepsilon)/\varepsilon$, 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 $r \in [\varepsilon, 2\varepsilon]$,

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

Thus we can write

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

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

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

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

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

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

3.3 Non-homogeneous case

The concavity and homogeneity assumptions were used heavily above for the proof of the necessary condition in Theorem 6. However, these assumptions can be considerably weakened while still resulting in a sufficient condition. We now show that a weak uniformity condition on the variances, coupled with a natural bound on the second-derivative measure of δ^2, result in zero m-variation processes.

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,

$$\delta(r) = o\left(r^{1/2m}\right). \tag{20}$$

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,

$$|\mu|(OD) \le c\varepsilon^{-(m-1)/m}, \tag{21}$$

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

Proof. *Step 0: setup.* Recall that by Lemma 1,

$$\begin{aligned}\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right] &= \frac{1}{\varepsilon^2}\sum_{j=0}^{(m-1)/2} c_j \int_0^T\int_0^t dtds\Theta^\varepsilon(s,t)^{m-2j}\ \delta^{2j}(s,s+\varepsilon)\ \delta^{2j}(t,t+\varepsilon) \\ &:= \sum_{j=0}^{(m-1)/2} J_j\end{aligned} \tag{22}$$

and now we express

$$\Theta^\varepsilon(s,t) = \mu([s,s+\varepsilon]\times[t,t+\varepsilon)) = \int_s^{s+\varepsilon}\int_t^{t+\varepsilon}\mu(dudv). \tag{23}$$

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

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

The hypothesis on the univariate δ^2 implies that this converges to 0 as ε tends to 0. The case of $t \le 2\varepsilon$ works equally easily.

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

$$J_{j,OD} = \varepsilon^{-2}c_j\int_{2\varepsilon}^T dt\int_0^{t-2\varepsilon} ds\delta^{2j}(s,s+\varepsilon)\,\delta^{2j}(t,t+\varepsilon)\,\Theta^\varepsilon(s,t)^{m-2j} \tag{24}$$

As we will prove below, the dominant term turns out to be $J_{(m-1)/2,OD}$; we deal with it now.

Step 2.1: term $J_{(m-1)/2,OD}$. 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 (23) and Fubini's theorem:

$$\begin{aligned}
\int_{2\varepsilon}^{T} dt \int_0^{t-2\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right| &= \int_{2\varepsilon}^{T} dt \int_0^{t-2\varepsilon} ds \left|\int_s^{s+\varepsilon}\int_t^{t+\varepsilon} \mu(dudv)\right| \\
&\leq \int_{2\varepsilon}^{T} dt \int_0^{t-2\varepsilon} ds \int_s^{s+\varepsilon}\int_t^{t+\varepsilon} |\mu|(dudv) \\
&= \int_{v=2\varepsilon}^{T+\varepsilon}\int_{u=0}^{\min(v,T)-\varepsilon} |\mu|(dudv) \int_{t=\max(2\varepsilon,v-\varepsilon,u+\varepsilon)}^{\min(v,T)} \int_{s=\max(0,u-\varepsilon)}^{\min(u,t-2\varepsilon)} ds\ dt \\
&\leq \int_{v=2\varepsilon}^{T+\varepsilon}\int_{u=0}^{v-\varepsilon} |\mu|(dudv) \int_{t=v-\varepsilon}^{v}\int_{s=u-\varepsilon}^{u} ds\ dt \\
&= \varepsilon^2 \int_{v=2\varepsilon}^{T+\varepsilon}\int_{u=0}^{v-\varepsilon} |\mu|(dudv)
\end{aligned}$$

Hence we have

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

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 (24) 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

$$\begin{aligned}
\left|J_{j,OD}\right| &\leq \frac{\delta^{4j}(\varepsilon)\,c_j}{\varepsilon^2}\int_{2\varepsilon}^{T} dt \int_0^{t-2\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right|^{m-2j-1}\left|\Theta^{\varepsilon}(s,t)\right| \\
&\leq \delta^{2m-2}(\varepsilon)\,c_j\varepsilon^{-2}\int_{2\varepsilon}^{T} dt \int_0^{t-2\varepsilon} ds \left|\Theta^{\varepsilon}(s,t)\right|,
\end{aligned}$$

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

A typical situation covered by the above theorem is that of the Riemann-Liouville fractional Brownian motion. This is the process $B^{H,RL}$ defined by $B^{H,RL}(t) = \int_0^t (t-s)^{H-1/2}\,dW(s)$. Its canonical metric is not homogeneous, but we do have, when $H \in (0,1/2)$,

$$|t-s|^H \leq \delta(s,t) \leq 2\,|t-s|^H, \tag{25}$$

which implies, incidentally, that $B^{H,RL}$ has the same regularity properties as fractional Brownian motion, see [17] for a proof of these inequalities. To apply the theorem, we must choose $H >$

$1/(2m)$ for the condition on the variances. For the other condition, we calculate that $\mu(dsdt) = 2H(1-2H)|t-s|^{2H-2}\,dsdt$, and therefore

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

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

One can generalize this example 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

$$X(t) = \int_0^t \left(\frac{d\gamma^2}{dr}\right)^{1/2} (t-r)\,dW(r). \tag{26}$$

Then one can show (see [17]) 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 inhomogeneous processes 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 $\gamma^2(r) = r^{1/(2m)}/\log(1/r)$.

3.4 Non-odd integer powers

When $m \geq 1$ is not an odd integer, recall that to define the mth odd variation, we use the convention $((x))^m = |x|^m \operatorname{sgn}(x)$, which is an odd function. The idea here is to use the Taylor expansion for this function up to order $[m]$, with a remainder of order $[m]+1$; it can be expressed as the following elementary lemma, whose proof is omitted.

Lemma 9 *Fix $m > 1$ and two reals a and b such that $|a| \geq |b|$. Let $\binom{m}{k}$ denote the formal binomial coefficient $m(m-1)\cdots(m-k+1)/(k(k-1)\cdots 1)$ and let $((x)) := |x|\,sgn(x)$. Then, for all reals a, b,*

1. *if $|b/a| < 1$,*

$$((a))^m ((a+b))^m = \sum_{k=0}^{[m]-1} \binom{m}{k} sgn^k(a)\,|a|^{2m-k} b^k + |a|^{2m} f_{m-1}\left(\frac{b}{a}\right),$$

2. *and if $|a/b| < 1$*

$$((a))^m ((a+b))^m = \sum_{k=0}^{[m]} \binom{m}{k} sgn^{k+1}(a)\, sgn^{k+1}(b)\, |a|^{m+k} |b|^{m-k} + (ab)^m f_m\left(\frac{a}{b}\right),$$

where for all $|x| < 1$, $|f_m(x)| \leq c_m |x|^{[m]+1}$ where c_m depends only on m. When m is an integer, the above formulas have null remainder terms f.

When we apply this lemma to the question of finding Gaussian fields with zero odd mth variation, we are able to prove that the sufficient condition of Theorem 6 still works. The result is the following.

Theorem 10 *Let X be as in Theorem 6 (X with homogeneous increments, with an increasing and concave δ^2). Let m be any real number > 1. Consider the condition*

(S) *δ^2 is twice differentiable, and for some $c < 1$, the function $r \mapsto \left|\left(\delta^2\right)''(r)\right|$ is decreasing and bounded above by $cr^{-2}\delta^2(r)$.*

*If m **is not** an integer, then X has zero odd mth variation as soon as $\delta(r) = o\left(r^{-1/(2m)}\right)$ and condition (S) holds.*

*If m **is** an integer, the same is true without needing condition (S).*

Remark 11 *The technical Condition (S) is not a restriction on the range of regularity of X. Indeed, for all fBm's that are more irregular than Brownian motion, we have $\left|\left(\delta^2\right)''(r)\right| = 2H(1-2H)r^{-2}\delta^2(r)$ which is indeed decreasing, and the constant c can be taken as $1/4$ (this maximum is attained for $H = 1/4$). For perturbations of the fBm scale, where $\delta(r)$ is of the order $r^{2H}\log^{\beta}(1/r)$ for some β, Condition (S) is also typically satisfied. Beyond the Hölder scale, in cases where $\delta(r)$ is of the order $\log^{\beta}(1/r)$ with $\beta < -1/2$, we will actually have a stronger upper-bound condition, of the type $\left|\left(\delta^2\right)''(r)\right| = o\left(r^{-2}\delta^2(r)\right)$. That $\left|\left(\delta^2\right)''(r)\right|$ be decreasing is typical of all Gaussian processes with homogeneous increments, even those which are more regular than Brownian motion.*

Proof of Theorem 10. *Step 0: setup.* Recall that with $Y = X(t+\varepsilon) - X(t)$ and $Z = X(s+\varepsilon) - X(s)$, we have

$$\mathbf{E}\left[\left([X,m]_{\varepsilon}(T)\right)^2\right] = \frac{2}{\varepsilon^2}\int_0^T\int_0^t dtds\mathbf{E}\left[((Y))^m((Z))^m\right].$$

Now introduce the shorthand notation $\sigma^2 = Var[Y]$, $\tau^2 = Var[Z]$, and $\theta = \mathbf{E}[YZ] = \Theta^{\varepsilon}(s,t)$. Thus $Y = \sigma M$ where M is a standard normal r.v.. We can write the "linear regression" expansion of Z w.r.t. Y, using another standard normal r.v. N independent of M:

$$Z = \frac{\theta}{\sigma}M + \rho\tau N$$

where

$$\rho := \left(1 - \left(\frac{\theta}{\sigma\tau}\right)^2\right)^{1/2}.$$

Note that ρ is always well-defined and positive by Cauchy-Schwarz's inequality. Therefore

$$((Y))^m((Z))^m = \sigma^m((M))^m\left(\left(\frac{\theta}{\sigma}M + \rho\tau N\right)\right)^m = \operatorname{sgn}(\theta)\,\sigma^{2m}|\theta|^{-m}((a))^m((a+b))^m$$

where

$$a := \frac{\theta}{\sigma}M \qquad \text{and} \qquad b := \rho\tau N.$$

Applying Lemma 9, we get that $((Y))^m ((Z))^m$ is the sum of the following four expressions:

$$A := \mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1} \sum_{k=0}^{[m]-1} \binom{m}{k} \mathrm{sgn}^{k+1}(\theta) |\theta|^{m-k} \sigma^k \tau^k \rho^k \ |M|^{2m-k} N^k \mathrm{sgn}^k(M) \tag{27}$$

$$A' := \mathrm{sgn}(\theta) \mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1} |\theta|^m |M|^{2m} f_{m-1}\left(\frac{\rho\tau\sigma N}{\theta M}\right) \tag{28}$$

$$B := \mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|>1} \sum_{k=0}^{[m]} \binom{m}{k} \mathrm{sgn}^{k+2}(\theta) |\theta|^k \sigma^{m-k} \tau^{m-k} \rho^{m-k} \mathrm{sgn}^{k+1}(NM) |M|^{m+k} |N|^{m-k} \tag{29}$$

$$B' := \mathrm{sgn}^2(\theta) \mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|>1} (\rho\sigma\tau)^m (MN)^m f_m\left(\frac{\theta M}{\rho\tau\sigma N}\right) \tag{30}$$

Step 1: cancellations in expectation calculation for A and B. In evaluating the ε-mth variation $\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right]$, terms in A and B containing odd powers of M and N will cancel, because of the symmetry of the normal law, of the fact that the indicator functions in the expressions for A and B above are even functions of M and of N, and of their independence. Hence we can perform the following calculations, where $a_{m,k} := \mathbf{E}\left[|M|^{2m-k}|N|^k\right]$ and $b_{m,k} := \mathbf{E}\left[|M|^{m+k}|N|^{m-k}\right]$ are positive constants depending only on m and k.

Step 1.1: expectation of A. In this case, because of the term N^k, the expectation of all the terms in (27) with k odd drop out. We can expand the term ρ^k using the binomial formula, and then perform a change of variables. We then have, with $n = [m]-2$ when $[m]$ is even, or $n = [m]-1$ when $[m]$ is odd,

$$\begin{aligned}
|\mathbf{E}[A]| &\leq \sum_{\substack{k=0 \\ k \text{ even}}}^{[m]-1} \binom{m}{k} |\theta|^{m-k} \sigma^k \tau^k a_{m,k} \sum_{\ell=0}^{k/2} \binom{k}{\ell} (-1)^\ell \left(\frac{\theta}{\sigma\tau}\right)^{2\ell} \\
&= \sum_{j=0}^{n/2} |\theta|^{m-2j} (\sigma\tau)^{2j} \sum_{\substack{k=2j \\ k \text{ even}}}^{n} \binom{m}{k}\binom{k}{k/2-j} (-1)^{k/2-j} a_{m,k} \\
&\leq \sum_{j=0}^{n/2} |\theta|^{m-2j} (\sigma\tau)^{2j} c_{m,j}
\end{aligned}$$

where $c_{m,j}$ are positive constant depending only on m and j. In all cases, the portion of $\mathbf{E}\left[([X,m]_\varepsilon(T))^2\right]$ corresponding to A can be treated using the same method as in the proof of Theorem 6. More precisely, after multiplying by ε^{-2} and integrating over s and t, as we should, each term in the last sum above is of the same form as the term J_j in the proof of Theorem 6, whose upper-estimation is the subject of Steps 1, 2, and 3 in that proof. The lowest power is attained when $j = n/2$, i.e., when $[m]$ is even we have $|\theta|^{2+m-[m]}$, and when $[m]$ is odd, we have $|\theta|^{1+m-[m]}$. In both cases, the power is greater than 1. All other values of j correspond of course to higher powers of $|\theta|$. This means we can use Cauchy-Schwarz's inequality to get the bound, valid for all j,

$$|\theta|^{m-2j} (\sigma\tau)^{2j} = |\theta| |\theta|^{m-2j-1} (\sigma\tau)^{2j} \leq |\theta| (\sigma\tau)^{m-1},$$

and we are back to the situation solved in the proof of Theorem 6, which corresponded therein to the case $j=(m-1)/2$ when m is was odd integer. Thus the portion of $\mathbf{E}\left[\left([X,m]_{\varepsilon}(T)\right)^{2}\right]$ corresponding to A tends to 0 as $\varepsilon\to 0$, as long as $\delta(r)=o\left(r^{1/(2m)}\right)$.

Step 1.2: expectation of B. This portion is dealt with similarly. Because of the term $\operatorname{sgn}^{k+1}(N)$, the expectation of all the terms in (29) with k even drop out. Contrary to the case of A, we do not need to expand ρ^{m-k} in a binomial series. Since k is now ≥ 1, we simply use Cauchy-Schwarz's inequality to write $|\theta|^{k}\leq|\theta|(\sigma\tau)^{k-1}$. Of course, we also have $\rho<1$. Hence

$$
\begin{aligned}
|\mathbf{E}[B]| &= \left|\mathbf{E}\left[\mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|>1}\sum_{\substack{k=1\\ k\ odd}}^{[m]}\binom{m}{k}\operatorname{sgn}^{k+2}(\theta)|\theta|^{k}\sigma^{m-k}\tau^{m-k}\rho^{m-k}\operatorname{sgn}^{k+1}(MN)|M|^{m+k}|N|^{m-k}\right]\right| \\
&\leq|\theta|(\sigma\tau)^{m-1}\sum_{\substack{k=1\\ k\ odd}}^{[m]}b_{m,k}\binom{m}{k}. \qquad (31)
\end{aligned}
$$

We see here in all cases that we are exactly in the same situation of the proof of Theorem 6 (again, power of $|\theta|$ is $|\theta|^{1}$). Thus the portion of $\mathbf{E}\left[\left([X,m]_{\varepsilon}(T)\right)^{2}\right]$ corresponding to B converges to 0 as soon as $\delta^{2}(r)=o\left(r^{1/(2m)}\right)$.

Step 2. The error term A'. For A' given in (28), we immediately have

$$
\begin{aligned}
\left|\mathbf{E}\left[A'\right]\right| &= \left|\mathbf{E}\left[\mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1}|\theta|^{m}|M|^{2m}f_{m-1}\left(\frac{\rho\tau\sigma N}{\theta M}\right)\right]\right| \\
&\leq c_{m}|\theta|^{m}\,\mathbf{E}\left[\mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1}|M|^{2m}\left|\frac{\rho\tau\sigma N}{\theta M}\right|^{[m]}\right] \\
&= c_{m}|\theta|^{m-[m]}(\rho\tau\sigma)^{[m]}\,\mathbf{E}\left[\mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1}|M|^{2m-[m]}|N|^{[m]}\right].
\end{aligned}
$$

We see here that we cannot ignore the indicator function inside the expectation, because if we did we would be left with $|\theta|$ to the power $m-[m]$, which is less than 1, and therefore does not allow us to use the proof of Theorem 6.

To estimate the expectation, let $x=\frac{\rho\tau\sigma}{|\theta|}$. We can use Hölder's inequality for a conjugate pair p,q with p very large, to write

$$
\mathbf{E}\left[\mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1}|M|^{2m-[m]}|N|^{[m]}\right]\leq\mathbf{P}^{1/q}\left[|xN|<|M|\right]\ \mathbf{E}^{1/p}\left[|M|^{2mp-[m]p}|N|^{[m]p}\right].
$$

The second factor in the right-hand side above is a constant $c_{m,p}$ depending only on m and p. For the first factor, we can use the following standard estimate for all $y>0$: $\int_{y}^{\infty}e^{-z^{2}/2}dz\leq\mathrm{cst}y^{-1}e^{-y^{2}/2}$. Therefore,

$$
\begin{aligned}
\mathbf{P}\left[|xN|<|M|\right] &= 2\int_{0}^{\infty}\frac{du}{\sqrt{2\pi}}e^{-u^{2}/2}\mathbf{P}\left[|M|>xu\right]\leq\sqrt{\frac{2}{\pi}}\int_{0}^{1/x}du\ e^{-u^{2}/2}+\sqrt{\frac{2}{\pi}}\int_{1/x}^{\infty}du\frac{e^{-u^{2}/2}}{ux}e^{-u^{2}x^{2}/2} \\
&\leq\sqrt{2/\pi}\frac{1}{x}+\sqrt{2/\pi}\int_{1/x}^{\infty}du\frac{1}{ux}e^{-u^{2}x^{2}/2}=\sqrt{2/\pi}\left(\frac{1}{x}+\frac{1}{x}\int_{1}^{\infty}dv\frac{1}{v}e^{-v^{2}/2}\right)=\frac{c}{x}
\end{aligned}
$$

where c is a universal constant.

Now choose q so that $m-[m]+1/q=1$, i.e. $q=(1-m+[m])^{-1}$, which exceeds 1 as long as m is not an integer. Then we get

$$\left|\mathbf{E}\left[A'\right]\right| \leq c_m \left|\theta\right|^{m-[m]} \left(\rho\tau\sigma\right)^{[m]} c_{m,p} \left(\frac{c\left|\theta\right|}{\rho\sigma\tau}\right)^{1/q} = c_m c_{m,p} c^{1/q} \left|\theta\right| \left(\rho\tau\sigma\right)^{m-1},$$

and we are again back to the usual computations. The case of m integer is dealt with in Step 4 below.

Step 3. The error term B'. For B' in (30), we have

$$\begin{aligned}\left|\mathbf{E}\left[B'\right]\right| &\leq c_m \sigma^m \tau^m \rho^m \mathbf{E}\left[\left|MN\right|^m \left|\frac{\theta M}{\rho\tau\sigma N}\right|^{[m]+1}\right] \\ &= c_m \left(\rho\sigma\tau\right)^{m-[m]-1} \left|\theta\right|^{1+[m]} \mathbf{E}\left[\left|M\right|^{m+[m]+1} \left|N\right|^{-1+m-[m]}\right].\end{aligned}$$

The expectation above is a constant c'_m depending only on m as long as m is not an integer. The case of m integer is trivial since then we have $B'=0$. Now we can use Cauchy-Schwarz's inequality to say that $\left|\theta\right|^{[m]} \leq (\sigma\tau)^{[m]}$, yielding

$$\left|\mathbf{E}\left[B'\right]\right| \leq c_m c'_m \rho^{m-[m]-1} \left(\sigma\tau\right)^{m-1} \left|\theta\right|^1.$$

This is again identical to the terms we have already dealt with, but for the presence of the negative power on ρ. We will handle this complication by showing that ρ can be bounded below by a universal constant.

First note that integration on the ε-diagonal can be handled by using the same argument as in Steps 1 and 2 of the proof of Theorem 6. Thus we can assume that $t \geq s+2\varepsilon$. Now that we are off the diagonal, note that using the mean value theorem on the expression for θ in (11), we can write that $\theta = \varepsilon^2 \left(\delta^2\right)''(\xi)$ for some ξ in $[t-s-\varepsilon, t-s+\varepsilon]$. At this point, Condition (S) allows us to say first that $\left|\left(\delta^2\right)''(r)\right| \leq cr^{-2}\delta^2(r)$. We now use the expression $\sigma\tau = \delta^2(\varepsilon)$, and the fact that off the diagonal, $\xi > \varepsilon$, combined with the fact that $\left|\left(\delta^2\right)''\right|$ is decreasing according to Condition (S), to write $\left|\theta/(\sigma\tau)\right| \leq \varepsilon^2 \left|\left(\delta^2\right)''(\varepsilon)\right| \delta^{-2}(\varepsilon) \leq c$. Recalling the definition of $\rho := \left(1-\theta^2(\sigma\tau)^{-2}\right)^{1/2}$, we have proved that ρ is bounded below uniformly (off the ε-diagonal) by the positive constant $c'' := \left(1-c^2\right)^{1/2}$. Hence, in the inequality for $\left|\mathbf{E}\left[B'\right]\right|$ above, the term $\rho^{m-[m]-1}$ can be absorbed into the m-dependent constants. In other words, we have proved the upper bound $\left|\mathbf{E}\left[B'\right]\right| \leq c_m c'_m (c'')^{m-[m]-1} (\sigma\tau)^{m-1} \left|\theta\right|^1$, and we are back once again to the situation solved in the proof of Theorem 6, proving the corresponding contribution of B' to $\mathbf{E}\left[\left([X,m]_\varepsilon(T)\right)^2\right]$ converges to 0 as soon as $\delta^2(r) = o\left(r^{1/(2m)}\right)$.

Step 4. The case of m integer. Of course, we already proved the theorem in the case m odd. Now assume m is an even integer. In this special case, we do not need to use a Taylor expansion, since the binomial formula has no remainder. Moreover, on the event $\left|\frac{\rho\tau\sigma N}{\theta M}\right| < 1$, $\operatorname{sgn}(a+b) = \operatorname{sgn}(a)$. Thus $A'=0$ with the understanding that we must replace A by the full sum for $k=0$ to m.

Recalculating this A we get

$$
\begin{aligned}
A &= \sigma^m \operatorname{sgn}(\theta) \operatorname{sgn}(M) |M|^m \mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1} \sum_{k=0}^{m} \binom{m}{k} M^k N^{m-k} \left[\frac{\theta}{\sigma}\right]^k [\rho\tau]^{m-k} \\
&= \operatorname{sgn}(\theta) \mathbf{1}_{\left|\frac{\rho\tau\sigma N}{\theta M}\right|<1} \sum_{k=0}^{m} \binom{m}{k} \operatorname{sgn}^{k+1}(M) |M|^{m+k} N^{m-k} \theta^k \rho^{m-k} \sigma^{m-k} \tau^{m-k}.
\end{aligned}
$$

Here, when we take the expectation $\mathbf{E}$, all terms vanish since we have odd functions of M for k even thanks to the term $\operatorname{sgn}^{k+1}(M)$, and odd functions of N for k odd thanks to the term N^{m-k}. I.e. the term corresponding to A is entirely null when m is even. The term B' is null since we have no error terms in the Taylor expansion. The estimation of the term B in Step 1.2 above applied when m is an integer. The proof of the theorem is finished. ■

4 Non-Gaussian case

Now assume that X is given by (7) and M is a square-integrable (non-Gaussian) 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

$$
\mu(d\bar{s}) = \mu(ds_1 ds_2 \cdots ds_m) = \mathbf{E}\left[d[M](s_1)\, d[M](s_2) \cdots d[M](s_m)\right], \tag{32}
$$

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: $0 \leq K(s_1, s_2, \cdots, s_m) \leq \Gamma^2(s_1)\Gamma^2(s_2)\cdots\Gamma^2(s_m)$ 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 $\Gamma(s) = \left(\mathbf{E}\left[H^{2m}(t)\right]\right)^{1/(2m)}$ 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 a Gaussian process. We let

$$
\tilde{G}(t,s) = \Gamma(s)\, G(t,s)
$$

and

$$
Z(t) := \int_0^T \tilde{G}(t,s)\, dW(s). \tag{33}
$$

We have the following.

Theorem 12 *Let m be an odd integer ≥ 3. Let X and Z be as defined in (7) and (33). 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 $\varepsilon > 0$,*

$$
\int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \int_{u=0}^{T} \left|\Delta\tilde{G}_t(u)\right| \left|\Delta\tilde{G}_s(u)\right| du \leq c\varepsilon\delta^2(2\varepsilon), \tag{34}
$$

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.

Proof.

Step 0: setup. We use an expansion for powers of martingales written explicitly at Corollary 2.18 of [9]. 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 [9], 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

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

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

$$K = \frac{1}{\varepsilon^2} \int_{u_{m-k}=0}^{T} \int_{u'_{m-k}=0}^{T} \int_{u_{m-k-1}=0}^{u_{m-k}} \int_{u'_{m-k-1}=0}^{u_{m-k}} \cdots \int_{u_1=0}^{u_2} \int_{u'_1=0}^{u_2}$$
$$\mathbf{E}\left[d\left[M\right]^{\otimes k} \left(u_{\sigma(1)}, \cdots, u_{\sigma(k)}\right) d\left[M\right]^{\otimes k} \left(u'_{\sigma(1)}, \cdots, u'_{\sigma(k)}\right) d\left[M\right]^{\otimes(m-2k)} \left(u_{\sigma(k+1)}, \cdots, u_{\sigma(m-k)}\right)\right]$$
$$\cdot \left[\Delta G_{\cdot}\left(u_{\sigma(k+1)}\right); \cdots; \Delta G_{\cdot}\left(u_{\sigma(m-k)}\right); \Delta G_{\cdot}\left(u_{\sigma(1)}\right); \Delta G_{\cdot}\left(u_{\sigma(1)}\right); \cdots; \Delta G_{\cdot}\left(u_{\sigma(k)}\right); \Delta G_{\cdot}\left(u_{\sigma(k)}\right)\right]$$
$$\cdot \left[\Delta G_{\cdot}\left(u_{\sigma(k+1)}\right); \cdots; \Delta G_{\cdot}\left(u_{\sigma(m-k)}\right); \Delta G_{\cdot}\left(u'_{\sigma(1)}\right); \Delta G_{\cdot}\left(u'_{\sigma(1)}\right); \cdots; \Delta G_{\cdot}\left(u'_{\sigma(k)}\right); \Delta G_{\cdot}\left(u'_{\sigma(k)}\right)\right], \tag{35}$$

modulo the fact that one may remove the integrals with respect to those u'_j's that are not represented among $\{u'_{\sigma(1)}, \cdots, u'_{\sigma(k)}\}$. The theorem will now be proved 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.

A final note about notation. The bracket notation in the last two lines of the expression (35) 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 Steps 1 of the proofs of Theorems 6 and 8, we can use brutal applications of Cauchy-Schwarz's inequality to deal with the portion of $K_{m,k,\sigma}$ in (35) 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)

$$\begin{aligned} K_{m,0,Id} &= \frac{1}{\varepsilon^2} \int_{u_m=0}^{T} \int_{u_{m-1}=0}^{u_m} \cdots \int_{u_1=0}^{u_2} \mathbf{E}\left[d\left[M\right]^{\otimes m} \left(u_1, \cdots, u_m\right)\right] \cdot \left[\Delta G_{\cdot}\left(u_1\right); \cdots; \Delta G_{\cdot}\left(u_m\right)\right]^2 \\ &\leq \frac{1}{\varepsilon^2} \int_{u_{m-k}=0}^{T} \int_{u_{m-k-1}=0}^{u_{m-k}} \cdots \int_{u_1=0}^{u_2} \Gamma^2\left(u_1\right) \Gamma^2\left(u_2\right) \cdots \Gamma^2\left(u_m\right) \left[\Delta G_{\cdot}\left(u_1\right); \cdots; \Delta G_{\cdot}\left(u_m\right)\right]^2 du_1 du_2 \cdots du_m \\ &= \frac{1}{\varepsilon^2} \int_{u_{m-k}=0}^{T} \int_{u_{m-k-1}=0}^{u_{m-k}} \cdots \int_{u_1=0}^{u_2} \left[\Delta \tilde{G}_{\cdot}\left(u_1\right); \cdots; \Delta \tilde{G}_{\cdot}\left(u_m\right)\right]^2 du_1 du_2 \cdots du_m. \end{aligned}$$

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, there is only one possible permutation, $\sigma = Id$, and we thus have, using hypothesis (A),

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

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

In order to exploit the additional hypothesis (34) in our theorem, our first move is to use Fubini by bringing the integrals over u_1 all the way inside. We get

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

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

$$
\int_{u_1=0}^{T}\int_{u_1'=0}^{T}du_1du_1'\left(\Delta\tilde G_s(u_1)\right)^2\left(\Delta\tilde G_t\left(u_1'\right)\right)^2
$$

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

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

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

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

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

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

Condition (34) implies immediately $K_{m,1,Id} \leq \delta^{2m}(2\varepsilon)\varepsilon^{-1}$ 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 $\delta^4(\varepsilon)$, as in the previous step, i.e. $\delta^{4k}(\varepsilon)$ altogether. In other words, for every $\sigma \in \Sigma_m^k$, and deleting the diagonal term, we have

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

Since $k \leq (m-1)/2$, 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 $\delta^{2(m-2k-1)}(\varepsilon)$. The remaining integral over u_{k+1} yields, by Condition (34), 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. ■

We state and prove the next proposition, in order to illustrate the range of applicability of Theorem 12. It provides a class of martingale-based processes X which can be associated to non-homogeneous Gaussian processes Z satisfying the assumptions of Theorem 8 and the additional assumption (34).

Proposition 13 *Let X be defined by (7) via the kernel G and the martingale M. 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,*

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

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

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

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

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

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

Also assume that g is decreasing in t and the bivariate f is increasing in t. Then X has zero m-variation.

The presence of the indicator function $\mathbf{1}_{s\le t}$ 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. In the case of irregular processes, which is the focus of this paper, the presence of the indicator function makes $\tilde{G}$ non-monotone in both s and t, which creates technical difficulties. Examples of non-adapted irregular processes are easier to treat, since it is possible to require that $\tilde{G}$ be monotone. We do not consider such non-adapted processes further. Specific examples of adapted processes which fall in the class defined in the above proposition are given below, after the proposition's proof.

Proof of Proposition 13. 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 (33), which, by Theorem 12, is sufficient, together with Condition (34), 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 (20) holds in Step 1; Step 2 will show Condition (21) holds; Condition (34) 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 $t = s+\varepsilon$

$$\begin{aligned} \tilde{\delta}^2(s,s+\varepsilon) &= \int_0^t \left|\tilde{G}(t,r) - \tilde{G}(s,r)\right|^2 dr \\ &= \int_0^s \left|(s+\varepsilon-r)^\alpha f(s+\varepsilon,r) - (s-r)^\alpha f(s,r)\right|^2 dr \\ &+ \int_s^{s+\varepsilon} |s+\varepsilon-r|^{2\alpha} f^2(s+\varepsilon,r)\,dr \\ &=: A+B. \end{aligned} \tag{36}$$

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. By the fact that f is increasing and g is decreasing in t,

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

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

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

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

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

This finishes the proof of Condition (20).

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 $|\mu|(OD)\leq\varepsilon^{2\alpha}$, which is condition (21), and $\mu(dsdt)=ds\;dt\;\partial^2\tilde{\delta}^2/\partial s\partial t$. We have, for $0\leq s\leq t-\varepsilon$,

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

We calculate

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

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,

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

which is of the correct order for Condition (21). For the term corresponding to A_1, using our hypotheses, we have

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

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

$$\int_{\varepsilon}^{T} dt \int_{0}^{t-\varepsilon} ds\, |A_1| \leq 2T \int_{\varepsilon}^{T} dt \int_{0}^{t-\varepsilon} ds\ s^{\alpha-1} t^{\alpha-1} = 2T\alpha^{-1} \int_{\varepsilon}^{T} dt\ t^{\alpha-1} (t-\varepsilon)^{\alpha} \leq \alpha^{-2} T^{1+2\alpha}.$$

This is much smaller than the right-hand side $\varepsilon^{2\alpha}$ of Condition (21), 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 (34). In fact, we modify the proof of Theorem 12, in particular Steps 3 and 4, so that we only need to prove

$$\int_{t=2\varepsilon}^{T} dt \int_{s=0}^{t-2\varepsilon} ds \int_{u=0}^{T} \left| \Delta \tilde{G}_t (u) \right| \left| \Delta \tilde{G}_s (u) \right| du \leq c\varepsilon^{2+2\alpha} = c\varepsilon^{1/m+1}, \tag{37}$$

instead of Condition (34). 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 12 holds if we assume (37) instead of Condition (34).

We now prove (37). 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. ■

The above 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. Assume that M is a martingale such that $\mathbf{E}\left[|d[M]/dt|^{2m}\right]$ is bounded above by a constant c^{2m} uniformly in $t \leq T$. This uniform boundedness assumption implies that we can take $\Gamma \equiv c$ in Condition (A). In particular, $\tilde{G}$ can be chosen to be proportional to G. 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 fractional Brownian motion with parameter $H = 1/(2m) - \alpha > 1/(2m)$. It is immediate to check that the assumptions of Proposition 13 are satisfied for this class of martingale-based models, which implies that the corresponding X defined by (7) 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 13, resulting again zero mth variations for the corresponding X defined in (7). For the sake of conciseness, we will omit the details, which are tedious and straightforward.

The most quantitatively significant condition in Theorem 12, that the univariate function $\delta(\varepsilon)$ corresponding to $\tilde{G}$ be equal to $o\left(\varepsilon^{1/(2m)}\right)$, can be interpreted as a regularity condition. In the Gaussian case, it means that there is a function $f(\varepsilon) = o\left(\varepsilon^{1/(2m)} \log^{1/2}(1/\varepsilon)\right)$ such that f is an almost-sure uniform modulus of continuity for X. In non-Gaussian cases, similar interpretations can be given for the regularity of X itself, provided enough moments of X exist. If X has fractional exponential moments, in the sense that for some constants $c > 0, 0 < \beta \leq 2$, $\mathbf{E}\left[\exp\left(c|X(t)-X(s)|^{\beta}\right)\right]$ is finite for all s,t, then the function f above will also serve as an almost-sure uniform modulus of continuity for X, provided the logarithmic correction term in f is raised to the power $1/\beta$ rather than $1/2$. Details of how this can be established are in the non-Gaussian regularity theory in [30]. If X has standard moments of all orders, then one can replace $f(\varepsilon)$ by $\varepsilon^{1/(2m)-\alpha}$ for any $\alpha > 0$. This is easily achieved using Kolmogorov's continuity criterion. If X only has finitely many moments, Kolmogorov's continuity criterion can only guarantee that one may take α greater than some $\alpha_0 > 0$. We do not delve into the details of these regularity issues in the non-Gaussian martingale case.

5 Stochastic calculus

In this section, we investigate the possibility of defining the so-called symmetric stochastic integral and its associated Itô formula for processes which are not fractional Brownian motion; fBm was treated in [14]. 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 [14]. Let X be a stochastic process on $[0,1]$. According to Sections 3 and 4 in [14] (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

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^1 du \left(X_{u+\varepsilon} - X_u\right)^m g\left(\frac{X_{u+\varepsilon} + X_u}{2}\right) = 0 \tag{38}$$

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

$$\int_0^t f'(X_u)\, d^\circ X_u =: \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t du \left(X_{u+\varepsilon} - X_u\right) \frac{1}{2} \left(f'(X_{u+\varepsilon}) + f'(X_u)\right) \tag{39}$$

exists and we have the Itô formula

$$f(X_t) = f(X_0) + \int_0^t f'(X_u)\, d^\circ X_u. \tag{40}$$

Our goal is thus to prove (38) for a wide class of Gaussian processes X, which will in turn imply the existence of (39) and the Itô formula (40).

If X has homogeneous 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 (38) to hold, we must have $\delta(r) = o\left(r^{1/6}\right)$. If one wishes to treat non-homogeneous cases, we notice that (38) for $g \equiv 1$ is the result of our non-homogeneous Theorem 8, so it is necessary to use that theorem's hypotheses, which include the non-homogeneous 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 [14]. 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 14 *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) $c'\delta^2(u)\, \delta^2(v-u) \leq Q_u Q_v - Q^2(u,v)$,

(iii)

$$\frac{\delta(au)-\delta(u)}{(a-1)u} < b\frac{\delta(u)}{u}. \tag{41}$$

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 15 *Assume the hypotheses of Theorem 14 with $m=3$. We have existence of the symmetric integral in (39), and its Itô formula (40), for every $f\in C^6(\mathbf{R})$ and $t\in[0,1]$.*

Before proceeding to the proof of this theorem, we discuss its hypotheses. We refer to the description at the end of Section 3.3 for examples satisfying the hypotheses of Theorem 8; these examples also satisfy the monotonicity and convexity conditions in the above theorem.

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 (26) given at the end of Section 3.3, which includes the Riemann-Liouville fBm's, satisfies Condition (i) with $c=1/2$. In the homogeneous 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 homogeneous case. We prove this claim, since it is not immediately obvious. In the homogeneous 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 (41) 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 14.

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 [14] 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 [14]. In order to avoid repeating arguments from that proof, we only state and prove the new lemmas which are required.

Step 1: translating Lemma 5.3 from [14]. 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:

Lemma 16 *Let $k \geq 2$ be an integer. Then*

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

This step and the next 4 steps are devoted to the **Proof of lemma 16.** We only need to show that for all $i,j \in \{1,2\}$,

$$\iint_{D_\varepsilon} |r_{ij}|^k\, dudv \leq cst \cdot \varepsilon\delta^k(\varepsilon). \tag{42}$$

Recall the function K defined in [14]

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

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 [14], and used below.

To follow the proof in [14], we need to prove the following items for some constants c_1 and c_2:

1. $c_1\delta^2(u) \leq K(u,u) \leq c_2\delta^2(u)$;

2. $K(u,v) \leq c_2 \delta(u) \delta(v)$;

3. $\Delta(u,v) := K(u,u) K(v,v) - K(u,v)^2 \geq c_1 \delta^2(u) \delta^2(v-u)$.

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

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

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 $\delta^2(s,t) \leq \delta^2(|t-s|)$, now implies, for $s=0$, that

$$\delta^2(0,u) = Q_u \leq \delta^2(u). \tag{43}$$

We write, via Cauchy-Schwarz's inequality and the fact that δ^2 is increasing, and thanks to (43),

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

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

$$\begin{aligned} K(u,v) &\leq 4\delta(2u)\delta(2v) \\ &\leq 8\delta(u)\delta(v), \end{aligned}$$

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 16 restricted to those values where $v \leq \varepsilon^{\rho/2}$ 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,

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

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

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

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

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

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 (42) for $i=j=1$ now follows by the next lemma.

Lemma 17 *For every $k \geq 2$, there exists $c_k > 0$ such that for every $\varepsilon \in (0,1)$,*

$$\int_{\varepsilon}^{1}\left|\frac{\delta(u)}{u}\right|^{k}du \leq c_k\varepsilon\left|\frac{\delta(\varepsilon)}{\varepsilon}\right|^{k}.$$

Proof of lemma 17. Our hypothesis (iii) can be rewritten as

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

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

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

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 17. $\square$

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

$$\begin{aligned}
|\Lambda_{21}[12]| &= \left|2(Q_{u+\varepsilon} - Q_u) + \delta^2(u+\varepsilon, v+\varepsilon) - \delta^2(u, v+\varepsilon) + \delta^2(u+\varepsilon, v) - \delta^2(u,v)\right| \\
&\leq 2|\Lambda_{21}[11]| + \varepsilon\delta^2(u+\varepsilon, v+\varepsilon)/(v-u) + \varepsilon\delta^2(u+\varepsilon, v)/(v-u-\varepsilon) \\
&\leq 2|\Lambda_{21}[11]| + \varepsilon\delta^2(v-u)/(v-u) + \varepsilon\delta^2(v-u-\varepsilon)/(v-u-\varepsilon) \\
&\leq 2cst \cdot \varepsilon\delta^2(u)/u + 2\varepsilon\delta^2(v-u-\varepsilon)/(v-u-\varepsilon). \qquad (44)
\end{aligned}$$

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 17 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 17 for the second term in the previous line. This finishes the proof of (42) 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 (44). 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 (42) 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

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

Moreover, we have, as in Step 2,

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

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

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

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 (42) is proved for r_{22}, which finishes the entire proof of Lemma 16. □

Step 6: translating Lemma 5.4 from [14]. We will prove the following result

Lemma 18 *For all $j\in\{0,1,\cdots,(m-1)/2\}$,*

$$\iint_{D_{\varepsilon}}\left|\mathbf{E}\left[Z_{3}Z_{4}\right]\right|^{m-2j}dudv \leq cst\cdot\varepsilon\delta^{2(m-2j)}\left(\varepsilon\right).$$

Proof of Lemma 18. As in [14], we have

$$\left|\mathbf{E}\left[Z_{3}Z_{4}\right]\right|^{m-2j} \leq cst\cdot\left|\mathbf{E}\left[G_{3}G_{4}\right]\right|^{m-2j}+cst\cdot\left|\mathbf{E}\left[\Gamma_{3}\Gamma_{4}\right]\right|^{m-2j}.$$

The required estimate for the term corresponding to $\left|\mathbf{E}\left[\Gamma_{3}\Gamma_{4}\right]\right|^{m-2j}$ follows by Cauchy-Schwarz's inequality and Lemma 16. For the term corresponding to $\left|\mathbf{E}\left[G_{3}G_{4}\right]\right|^{m-2j}$, we recognize that $\mathbf{E}\left[G_{3}G_{4}\right]$ is the negative planar increment $\Theta^{\varepsilon}\left(u,v\right)$ defined in (12). Thus the corresponding term was already considered in the proof of Theorem (8). More specifically, up to the factor $\varepsilon^{2}\delta^{-4j}\left(\varepsilon\right)$, we now have to estimated the same integral as in Step 2 of that theorem's proof: see expression (24) for the term we called $J_{j,OD}$. This means that

$$\begin{aligned}\iint_{D_{\varepsilon}}\left|\mathbf{E}\left[G_{3}G_{4}\right]\right|^{m-2j}dudv &\leq \frac{\varepsilon^{2}}{\delta^{4j}\left(\varepsilon\right)}J_{j,OD}\\ &\leq \varepsilon^{2}\left|\mu\right|\left(OD\right)\delta^{2(m-2j-1)}\left(\varepsilon\right).\end{aligned}$$

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

Step 7. Conclusion. The remainder of the proof of the theorem is to check that Lemmas 16 and 18 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 [14]. Since such a task is only bookkeeping, we omit it, concluding the proof of Theorem 14. ■

Acknowledgements

The work of F. Russo was partially supported by the ANR Project MASTERIE 2010 BLAN-0121-01. Part of it was done during a stay of this author at the Bernoulli Center of the EPF Lausanne. The work of F. Viens is partially supported by NSF DMS grant 0907321.

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

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

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

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

[7] 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.

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

[9] 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).

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

[11] 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.

[12] 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).

[13] 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).

[14] 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).

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

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

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

[18] 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).

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

[20] 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).

[21] 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).

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

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

[24] 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).

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

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

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

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

[29] 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).

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