PATHWISE STOCHASTIC INTEGRALS AND ITÖ FORMULA FOR MULTIDIMENSIONAL GAUSSIAN PROCESSES

ZHE CHEN AND LAURI VIITASAARI

ABSTRACT. In this article we study existence of pathwise stochastic integrals with respect to a general class of \(n\)-dimensional Gaussian processes and a wide class of adapted integrands. More precisely, we study integrands which are functions that are of locally bounded variation with respect to all variables. Moreover, multidimensional Itô formula is derived.

1. INTRODUCTION

Let \(X = (X^1, X^2, \ldots, X^n)\) be \(n\)-dimensional Gaussian process on \([0, T]\) such that all components \(X^k\) are independent. In this article we are interested in which generality stochastic integrals of type

\[
\sum_{k=1}^{n} \int_0^T f_k(X^1_u, X^2_u, \ldots, X^n_u) \, dX^k_u
\]

exist in a pathwise sense. In particular, we assume that the processes \(X^k\) are not semimartingales and hence standard integration techniques cannot be applied.

Pathwise generalizations of stochastic integration go back to Young [6] who proved that if the integrand and the integrator are together smooth enough in the sense of \(p\)-variations, then the integral exists as a limit of Riemann–Stieltjes sums. In a particular case of fractional Brownian motion with Hurst index \(H > \frac{1}{2}\) this was considered in Lin [9] and Dai and Heyde [11] who showed the existence of stochastic integral if the integrand has finite \(p\)-variation with \(\frac{1}{p} + H > 1\). The pathwise forward-type Riemann–Stieltjes integration was studied by Föllmer [5] and the pathwise generalised Lebesgue–Stieltjes integrals was introduced by Zähle [12] and later developed by Nualart and Răşcanu [2]. In this case the integral with respect to fractional Brownian motion is well-defined if the integrand has \(\lambda\)-Hölder continuous paths with \(\lambda > 1 - H\). However, all of the mentioned studies
considered too restrictive integrands. In particular, even the simple integral
\[ \int_0^T 1_{X_u > a} dX_u \]
is not covered.

Later on this problem was studied in the case of fractional Brownian motion by Azmoodeh et al. [1] who proved that if \( B^H \) is fractional Brownian motion with \( H > \frac{1}{2} \) and \( F \) is a convex function with left-sided derivative \( f'_{-} \), then
\[ \int_0^T f'_{-}(B^H_u) dB^H_u \]
exists in the sense of generalized Lebesgue–Stieltjes integral. Moreover, it was shown that the integral can be also understood as a Föllmer integral along uniform partition and the following Itô formula
\[ F(B^H_T) = F(0) + \int_0^T f'_{-}(B^H_u) dB^H_u \]
holds almost surely. It was also pointed out in Mishura et al. [8] that the results hold also for functions of locally bounded variation and thus the path-wise integrals with respect to fractional Brownian motion with \( H > \frac{1}{2} \) can be defined for integrands \( f(B^H_u) \) for a wide class of functions \( f \). Moreover, it was shown in Azmoodeh and Viitasaari [4] that under mild integrability condition the limit of the forward-type Riemann–Stieltjes sums can also be understood as \( L^p \)-limit for suitable range of \( p \). The authors also studied rate of convergence of this approximation. Later on these results for fractional Brownian motion was widely generalized by Sottinen and Viitasaari [10] to cover a large class of Gaussian processes. Namely, the authors in [10] proved that, under some mild extra assumptions, these results hold for Gaussian processes which has \( \alpha \)-Hölder continuous trajectories almost surely for some \( \alpha > \frac{1}{2} \).

While the above mentioned works cover large class of Gaussian processes, they only consider one dimensional processes. Motivated by this we study existence of integrals where the integrand is of locally bounded variation separately with respect to all variables. More precisely, we prove that under a natural integrability assumption the integrals
\[ \int_0^T f(X^1_u, X^2_u, \ldots, X^n_u) dX^k_u \]
exist for every \( k = 1, \ldots, n \) in the sense of generalized Lebesgue–Stieltjes integral, provided that processes \( X^k \) are independent Gaussian processes of certain type and the function \( f(x) \) is of locally bounded variation separately with respect to each variable \( x_k \). However, we do not assume that all the processes \( X^k \) are independent copies and hence different stylized facts can be added to the model by adding them to different independent random sources. We also prove the following Itô formula
\[ F(X^1_T, X^2_T, \ldots, X^n_T) = F(X^1_0, X^2_0, \ldots, X^n_0) + \sum_{k=1}^n \int_0^T \frac{\partial}{\partial k} F(X^1_u, X^2_u, \ldots, X^n_u) dX^k_u, \]
where \( \frac{\partial}{\partial k} F \) denotes the one-sided derivative of a continuous function \( F \) and \( \frac{\partial}{\partial k} F \) is assumed to be of locally bounded variation.
The rest of the paper is organized as follows. In Section 2, we recall basic facts on generalized Lebesgue–Stieltjes integrals and Föllmer integrals. Moreover, we introduce our class of integrands and processes together with discussions. Section 3 is devoted to our main results together with the proofs. We end the paper with discussion.

2. Auxiliary Facts

2.1. Pathwise integrals.

**Definition 2.1.** Fix $0 < \beta < 1$.

1. The fractional Besov space $W_{1}^{\beta} = W_{1}^{\beta}(0, T]$ is the space of real-valued measurable functions $f : [0, T] \to \mathbb{R}$ such that
   \[
   \|f\|_{1, \beta} = \sup_{0 \leq s < t \leq T} \left( \frac{|f(t) - f(s)|}{(t - s)^{\beta}} + \int_{s}^{t} \frac{|f(u) - f(s)|}{(u - s)^{1+\beta}} du \right) < \infty.
   \]

2. The fractional Besov space $W_{2}^{\beta} = W_{2}^{\beta}(0, T]$ is the space of real-valued measurable functions $f : [0, T] \to \mathbb{R}$ such that
   \[
   \|f\|_{2, \beta} = \int_{0}^{T} \frac{|f(s)|}{s^{\beta}} ds + \int_{0}^{T} \int_{0}^{s} \frac{|f(u) - f(s)|}{(u - s)^{1+\beta}} duds < \infty.
   \]

**Remark 2.1.** Let $C^\alpha = C^\alpha([0, T])$ denote the space of Hölder continuous functions of order $\alpha$ on $[0, T]$ and let $0 < \epsilon < \beta \wedge (1 - \beta)$. Then $C^{\beta + \epsilon} \subset W_{1}^{\beta} \subset C^{\beta - \epsilon}$ and $C^{\beta + \epsilon} \subset W_{2}^{\beta}$.

**Definition 2.2.** Let $t \in [0, T]$. The Riemann–Liouville fractional integrals $I_{0+}^{\beta}$ and $I_{t-}^{\beta}$ of order $\beta > 0$ on $[0, T]$ are

\[
(I_{0+}^{\beta} f)(s) = \frac{1}{\Gamma(\beta)} \int_{0}^{s} f(u)(s - u)^{\beta - 1} du,
\]

\[
(I_{t-}^{\beta} f)(s) = \frac{(-1)^{-\beta}}{\Gamma(\beta)} \int_{s}^{t} f(u)(u - s)^{\beta - 1} du,
\]

where $\Gamma$ is the Gamma-function. The Riemann–Liouville fractional derivatives $D_{0+}^{\beta}$ and $D_{t-}^{\beta}$ are the left-inverses of the corresponding integrals $I_{0+}^{\beta}$ and $I_{t-}^{\beta}$. They can be also define via the Weyl representation as

\[
(D_{0+}^{\beta} f)(s) = \frac{1}{\Gamma(1 - \beta)} \left( \frac{f(s)}{s^\beta} + \beta \int_{0}^{s} \frac{f(s) - f(u)}{(s - u)^{\beta + 1}} du \right),
\]

\[
(D_{t-}^{\beta} f)(s) = \frac{(-1)^{-\beta}}{\Gamma(1 - \beta)} \left( \frac{f(s)}{(t-s)^\beta} + \beta \int_{s}^{t} \frac{f(s) - f(u)}{(u - s)^{\beta + 1}} du \right)
\]

if $f \in I_{0+}^{\beta}(L^1)$ or $f \in I_{t-}^{\beta}(L^1)$, respectively.

Denote $g_{t-}(s) = g(s) - g(t-)$. The generalized Lebesgue–Stieltjes integral is defined in terms of fractional derivative operators according to the next proposition.
Proposition 2.1. Let \(0 < \beta < 1\) and let \(f \in W^\beta_2[0,T]\) and \(g \in W^{1-\beta}_1[0,T]\). Then for any \(t \in (0,T]\) the generalized Lebesgue–Stieltjes integral exists as the following Lebesgue integral
\[
\int_0^t f(s) \, dg(s) = \int_0^t (D^\beta_0 f)(s)(D^{1-\beta}_s g)(s) \, ds
\]
and is independent of \(\beta\).

Remark 2.2. It is shown in [7] that if \(f \in C^\gamma\) and \(g \in C^{f''}\) with \(\gamma + f'' > 1\), then the generalized Lebesgue–Stieltjes integral \(\int_0^t f(s) \, dg(s)\) exists and coincides with the classical Riemann–Stieltjes integral, i.e., as a limit of Riemann–Stieltjes sums. This is natural, since in this case one can also define the integrals as Young integrals [6].

We will also need the following estimate in order to prove our main theorems.

Theorem 2.1. Let \(f \in W^\beta_2[0,T]\) and \(g \in W^{1-\beta}_1[0,T]\). Then we have the estimation
\[
\left| \int_0^t f \, dg \right| \leq \frac{1}{\Gamma(\beta)} \|f\|_{2,\beta} \|g\|_{1,1-\beta}.
\]

Corollary 2.1. Let \(f, f_n \in W^\beta_2[0,T]\), \(\|f_n - f\|_{2,\beta} \to 0\) as \(n \to \infty\), and \(g \in W^{1-\beta}_1[0,T]\). Then
\[
\int f_n \, dg \to \int f \, dg.
\]

We also recall the definition of a forward-type Riemann–Stieltjes integral due to Föllmer [5] (see also [3]).

Definition 2.3. Let \((\pi_n)_{n=1}^\infty\) be a sequence of partitions \(\pi_n = \{0 = t^n_0 < \ldots < t^n_{k(n)} = T\}\) such that \(|\pi_n| = \max_{j=1,\ldots,k(n)} |t^n_j - t^n_{j-1}| \to 0\) as \(n \to \infty\). Let \(X\) be a continuous process. The Föllmer integral along the sequence \((\pi_n)_{n=1}^\infty\) of \(Y\) with respect to \(X\) is defined as
\[
\int_0^t Y_u \, dX_u = \lim_{n \to \infty} \sum_{t^n_j \in \pi_n \cap [0,t]} Y^n_{t^n_j} \left( X^n_{t^n_j} - X^n_{t^n_{j-1}} \right),
\]
if the limit exists almost surely.

In general it is not clear when the Föllmer integrals exist. In the case of quadratic variation processes the existence is guaranteed by the Itô–Föllmer formula of Lemma 2.1 below, which shows that the Föllmer integral behaves like the Itô integral in the case of integrators with quadratic variation.

Definition 2.4. Let \((\pi_n)_{n=1}^\infty\) be a sequence of partitions \(\pi_n = \{0 = t^n_0 < \ldots < t^n_{k(n)} = T\}\) such that \(|\pi_n| = \max_{j=1,\ldots,k(n)} |t^n_j - t^n_{j-1}| \to 0\) as \(n \to \infty\). Let \(X\) be a continuous process. Then \(X\) is a quadratic variation process along the sequence \((\pi_n)_{n=1}^\infty\) if the limit
\[
\langle X \rangle_t = \lim_{n \to \infty} \sum_{t^n_j \in \pi_n \cap [0,t]} \left( X^n_{t^n_j} - X^n_{t^n_{j-1}} \right)^2
\]
exists almost surely.

**Lemma 2.1.** [5] Let $X$ be a continuous quadratic variation process and let $f \in C^{1,2}([0,T] \times \mathbb{R})$. Let $0 \leq s < t \leq T$. Then

$$f(t, X_t) = f(s, X_s) + \int_s^t \frac{\partial f}{\partial t}(u, X_u) \, du + \int_s^t \frac{\partial f}{\partial x}(u, X_u) \, dX_u + \frac{1}{2} \int_s^t \frac{\partial^2 f}{\partial x^2}(u, X_u) \, d\langle X \rangle_u.$$  

In particular, the Föllmer integral exists and has a continuous modification.

### 2.2. Notation and assumptions.

**Definition 2.5.** Let $X$ be a centered Gaussian process. We denote by $R(t, s)$, $W(t, s)$, and $V(t)$ its covariance, incremental variance and variance, i.e.

$$R(t, s) = \mathbb{E}[X_t X_s],$$
$$W(t, s) = \mathbb{E}[(X_t - X_s)^2],$$
$$V(t) = \mathbb{E}[X_t^2].$$

We denote by $w^*(t)$ the “worst case” incremental variance

$$w^*(t) = \sup_{0 \leq s \leq T-t} W(t + s, s).$$

Recall that a process $X = (X_t)_{t \in [0, T]}$ is Hölder continuous of order $\alpha$ if there exists almost surely finite random variable $C_T$ such that

$$|X_t - X_s| \leq C_T |t - s|^\alpha$$

almost surely for all $s, t \in [0, T]$.

Next we recall the class $\mathcal{X}^\alpha$ of Gaussian processes introduced in [10].

**Definition 2.6.** A centered continuous Gaussian process $X = (X_t)_{t \in [0, T]}$ with covariance $R$ belongs to the class $\mathcal{X}^\alpha$ if

1. There exists a constant $\delta > 0$ such $R(s, t) > 0$ for every $s, t > 0$ provided $|t - s| \leq \delta$,
2. the “worst case” incremental variance satisfies

$$w^*(t) \leq Ct^{2\alpha},$$

where $C > 0$ and $0 < \alpha < 1$,
3. there exists a $c, \delta > 0$ such that

$$V(s) \geq cs^2,$$

when $s \leq \delta$,
4. there exists a $\delta > 0$ such that

$$\sup_{0 < t < 2\delta} \sup_{\frac{1}{2} \leq s \leq t} \frac{R(s, s)}{R(t, s)} < \infty.$$
The definition is rather technical. However, the assumptions are not very restrictive and the following remarks and examples should convince the reader that indeed many processes belong to the given class.

**Remark 2.3.** (1) In the original paper [10] the authors assumed $R(t, s) > 0$ for every $t, s > 0$. However, the assumption is needed only close to diagonal and hence this assumption is rather natural and not very restrictive. Note also that we assume that $X$ is random i.e. for every $s > 0$ we have $R(s, s) = V(s) > 0$. Hence this condition is closely related to the fourth condition i.e. the covariance $R(s, t)$ is not “too far” from the variance when $s$ and $t$ are close to each other.

(2) The second condition on the incremental variance is the most important assumption as it implies that $X$ has a version which is Hölder continuous of order $r$ for any $r < \alpha$ on $[0, T]$.

(3) The third condition is a natural assumption and could be dropped. Indeed, note first that we either have $\inf_{0 \leq s \leq T} V(s) > 0$ or else $X_0 = 0$. Hence if the variance $V(s)$ behaves like $s^\gamma$ for some $\gamma > 2$ near zero, we obtain that the process is Hölder continuous of order $\alpha > 1$ on that interval. Hence it is constant and thus violating first assumption.

(4) Finally, the fourth assumption is quite mild as for it we simply need that when $s$ and $t$ are both close to each other and at the same time near to zero, the variance $V(s)$ is not “too far” from the covariance $R(s, t)$.

Many processes such as stationary increment processes or stationary processes belong to the class $\mathcal{X}^\alpha$. For examples and more discussion see [10]. We also recall the following technical estimate.

**Lemma 2.2.** [10] Let $X$ be a centered Gaussian process with strictly positive and bounded covariance function $R$, $0 < s < t \leq T$ and $a \in \mathbb{R}$. Then there exists a universal constant $C$ such that

$$
\mathbb{P}(X_s < a < X_t) \leq C e^{-\min_{2, (a-1)^2} \sqrt{V(t, s)}} \left[ 1 + \frac{R(s, s)}{R(t, s)} + \frac{|a|}{\sqrt{V(s)}} \max \left( 1, \frac{R(s, s)}{R(t, s)} \right) \right],
$$

where

$$
V^* = \sup_{s \leq T} V(s).
$$

**Remark 2.4.** In [10] the Lemma was stated in a bit different form. However, by examining the proof it is clear that the above formula is also valid.

Consider now an $n$-dimensional vector $\bar{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ with $\alpha_k > \frac{1}{2}, \forall k$. We consider the following class of processes.

**Definition 2.7.** Let $X = (X^1, X^2, \ldots, X^n)$ be an $n$-dimensional Gaussian process on $[0, T]$. We denote $X \in \mathcal{X}^\bar{\alpha}$ if the processes $X^k$ are independent and for every $k$ we have $X_k \in \mathcal{X}^{\alpha_k}$. 
Remark 2.5. The assumption of independent processes $X^k$ is rather restrictive for many applications. However, our results hold also with obvious changes for dependent processes if all the conditional Gaussian processes $X^k$ given the other $n-1$ variables belong to $X^\alpha$ with some $\alpha > \frac{1}{2}$. Hence the assumption of independent processes is only a simplification.

Consider next a continuous function $F : \mathbb{R}^n \to \mathbb{R}$ such that all one-sided partial derivatives $\frac{\partial}{\partial x_k} F(x_1, \ldots, x_n)$ exist and are of locally bounded variation separately with respect to each variables i.e. for every compact set $K \subset \mathbb{R}$ the function $f : K \to \mathbb{R}$ defined by $f(y) = \frac{\partial}{\partial x_k} F(x_1, \ldots, y, \ldots, x_n)$ is of bounded variation. According to the well-known Jordan decomposition, a function of one variable is of bounded variation if and only if it is a difference of two increasing functions. Moreover, for every increasing function $f$ there exists a convex function $F$ such that

$$\frac{\partial}{\partial x} F(x) = f(x).$$

Moreover, the second derivative $F''$ exist as a distribution $\mu$ and if the Radon measure $\mu$ has compact support, then there exists constants $b$ and $c$ such that

$$F(x) = c + bx + \int_{\text{supp}(\mu)} |x - a| \mu(da)$$

and

$$(2.1) \quad \frac{\partial}{\partial x} F(x) = b + \int_{\text{supp}(\mu)} \text{sgn}(x - a) \mu(da),$$

where $\text{sgn}(x) = 1$ for $x > 0$ and $\text{sgn}(x) = -1$ for $x \leq 0$. Hence for every function of locally bounded variation we can associate a (signed) measure $\mu$ on the given compact set $K$ such that \eqref{2.1} holds. Moreover, for every signed measure $\mu$ we have a decomposition $\mu = \mu^+ - \mu^-$ where $\mu^+$ and $\mu^-$ are positive Radon measures. Note that in our multidimensional setup the measure $\mu$ depends on the fixed variables $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$.

Let now $\sigma = \{\sigma(1), \ldots, \sigma(n)\}$ denote a permutation of indices $\{1, \ldots, n\}$ and for a fixed $k = 1, \ldots, n$, let $P(X^{\sigma(k)} \in d\tilde{X}^k)$ denote the Gaussian measure related to random vector $X^{\sigma(k)} = (X^{\sigma(1)}, X^{\sigma(2)}, \ldots, X^{\sigma(k)})$, where $X^{\sigma(i)} \sim N(0, V_{X^{\sigma(i)}})$, for every $i = 1, \ldots, k$. Now we are ready to define our class of functions.

Definition 2.8. Let $F : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. We denote $F \in Y$ if:

1. for every $k = 1, \ldots, n$ the one-sided partial derivatives $\frac{\partial}{\partial x_k} F(x_1, \ldots, x_n)$ exist and are of locally bounded variation separately with respect to all variables $x_1, \ldots, x_n$,
we have

\[
\sum_{k=1}^{n} \sigma \sum_{\{\mathbb{R} \setminus [-1,1]\}^k} \max_{-1 \leq x_i \leq 1} \int_{\mathbb{R}} \max(|a|, 1) e^{-\min\{|a|^2, (a-1)^2\} / 2x^{(n)}} \left| \mu \right| x_i^{(k+1)} \cdots x_i^{(n-1)} (da) \mathbb{P}(\bar{X}^{(k)} \in d\bar{x}^k) < \infty,
\]

where the second sum is over all possible permutations \(\sigma\), \(|\mu| = \mu^+ + \mu^-\), and the measure \(\mu x^{(1)} \cdots x^{(n-1)}\) is the measure associated to the partial derivative \(\frac{\partial}{\partial x^{(n)}}\).

The definition depends also on the underlying processes \(X^1, \ldots, X^n\) which will be omitted on the notation. Note also that the definition looks rather technical and complicated. However, it is a quite mild and natural assumption. Indeed, take any permutation \(\sigma = (\sigma(1), \ldots, \sigma(n))\) and consider the measure \(\mu\) associated to the partial derivative \(\frac{\partial}{\partial x^{(n)}}\). Then choose any \(k\) variables and take maximum of \(\mu\) in the above sense with respect to the remaining \(n-1-k\) variables over a compact set \([-1,1]^{n-1-k}\). The interpretation of the assumption means that this maximum is “almost” integrable with respect to \(k+1\) dimensional Gaussian measure associated to a random vector \((X^{(1)}, X^{(2)}, \ldots, X^{(k)}, X^{(n)})\). In other words, the term arising from Lemma (2.2)

\[
\max(|a|, 1) e^{-\min\{|a|^2, (a-1)^2\} / 2x^{(n)}}
\]

is close to Gaussian measure related to the random variable \(X^{(n)}\) and \(\mu\) is related to the partial derivative with respect to \(x^{(n)}\). As such the assumption simply means that all the partial derivatives are integrable with respect to a measure that is “close” to Gaussian measure. It is also not obvious how such assumption can be dropped since in order to prove the existence of our integrals we have to show that certain random variables are finite almost surely. In order to do this we prove that expectations are finite and hence we have to assume some kind of integrability. The following remarks and examples give more justification why the assumption is natural and not too restrictive.

**Remark 2.6.** For simplicity we choose to take maximum over a compact set \([-1,1]^k\). However, this could be replaced by any set \([-\epsilon, \epsilon]^k\) with obvious modifications to the assumption. This gives more intuition to the assumption; choose any \(n-1-k\) variables, let them be really close to zero and assume that with those variables close to zero the partial derivatives are “almost” (in the above sense) integrable with respect to a certain Gaussian measure.

**Example 2.1.** If \(|\mu|\) is absolutely continuous with respect to Lebesgue measure, then the inner integral of the assumption becomes

\[
\int_{\mathbb{R}} \max(|a|, 1) e^{-\min\{|a|^2, (a-1)^2\} / 2x^{(n)}} f_X^{(n)} x^{(k+1)} \cdots x^{(n-1)} (a) da.
\]
In a region \([K, \infty)\) we get by integration by parts formula that

\[
\sum_{k=1}^{n} \sum_{\sigma} \int_{\{R \setminus [-1,1]\}^k} \max_{-1 \leq x_i \leq 1} \int_{K} \alpha^3 e^{-\frac{(a-1)^2}{2}} \frac{X_{\sigma(n)}}{V^{1/2}} f(a) da \mathbb{P}(\bar{X}^{(k)} \in d\bar{x}^k) < \infty
\]

needs to be satisfied, where \(K\) is some real number \(K > 1\) and here \(f(a)\) depends on \(x^{\sigma(1)}, \ldots, x^{\sigma(n)}\). The interval \((-\infty, -K]\) can be treated similarly and the compact interval \([-K, K]\) is obvious.

**Example 2.2.** Another motivating example is a function \(f(x_1, x_2, \ldots, x_n) = |\sum_{k=1}^{n} x_k|\). In this case Radon measure is just the Dirac delta function \(\delta(da)\). Then our assumption turns out to be

\[
\sum_{k=1}^{n} \sum_{\sigma} \int_{\{R \setminus [-1,1]\}^k} \mathbb{P}(\bar{X}^{(k)} \in d\bar{x}^k) < \infty,
\]

which is obviously true.

### 3. Main Results

**3.1. Existence of integrals.** First we will show the existence of multidimensional stochastic integrals with respect to some class of Gaussian processes in the generalized Lebesgue-Stieltjes sense.

**Theorem 3.1.** Let \(X \in X^{\bar{\alpha}}\) and \(f \in \mathcal{Y}\). Then for every \(k = 1, \ldots, n\), the integral

\[
\int_{0}^{T} \frac{\partial^{-}}{\partial x^k} f(X^1_s, \ldots, X^n_s) dX^k_s
\]

exists almost surely in generalized Lebesgue-Stieltjes sense.

**Proof.** For simplicity and without loss of generality we assume that

\[
\sup_{0 \leq s \leq T} V_k(s) \leq 1
\]

for every \(k\). For one-sided partial derivatives we use short notation \(\partial^{-}_k\). Moreover, we assume that all the measures \(\mu\) associated to partial derivatives are positive Radon measures. In a general case \(\mu\) is signed measure with decomposition \(\mu = \mu^+ - \mu^-\) and in this case we consider positive measure \(|\mu| = \mu^+ + \mu^-\). Let \(\beta \in (1 - \inf_{k=1,\ldots,n} \alpha_k, \frac{1}{2})\). According to Proposition 2.1 the generalized Lebesgue-Stieltjes integral exists if we have

\[
\|\partial^{-}_k f(X^1_s, \ldots, X^n_s)\|_{2, \beta} < \infty, \text{ a.s.}
\]

First it is trivial that

\[
\int_{0}^{T} \frac{\|\partial^{-}_k f(X^1_s, \ldots, X^n_s)\|_{s^{-\beta}}}{s^{-\beta}} ds \leq \sup_{0 \leq s \leq T} |\partial^{-}_k f(X^1_s, \ldots, X^n_s)| \int_{0}^{T} \frac{1}{s^{-\beta}} ds < \infty.
\]
For the second part in \( \| \partial_k^- f(X^n_1, \ldots, X^n_k) \|_{2,\beta} \) we split the integral with respect to \( s \) and \( u \) as
\[
\int_0^T \int_0^s \left| \partial_k^- f(X^n_1, \ldots, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_k) \right| ds du = \int_0^\infty \int_0^s \left| \partial_k^- f(X^n_1, \ldots, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_k) \right| ds du,
\]
\[
= \left( \int_0^{2\delta} \int_0^s + \int_0^T \int_0^s + \int_0^T \int_{s-\delta}^T + \int_0^{2\delta} \int_0^s \right) \left| \partial_k^- f(X^n_1, \ldots, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_k) \right| ds du =: I_1 + I_2 + I_3 + I_4.
\]
For terms \( I_3 \) and \( I_4 \) we notice that
\[
|\partial_k^- f(X_u) - \partial_k^- f(X_s)| \leq 2 \sup_{0<s<T} |\partial_k^- f(X^n_1, \ldots, X^n_k)|.
\]
Consequently, \( I_3 < \infty \) and \( I_4 < \infty \). For \( I_1 \) and \( I_2 \) it is sufficient to show that \( \mathbb{E}I_1 < \infty \) and \( \mathbb{E}I_2 < \infty \). First we write
\[
\begin{align*}
|\partial_k^- f(X^n_1, \ldots, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_k)| \\
= |\partial_k^- f(X^n_1, X^n_2, \ldots, X^n_k) - \partial_k^- f(X^n_1, X^n_2, \ldots, X^n_k) + X^n_k - X^n_k)| \\
\leq |\partial_k^- f(X^n_1, X^n_2, \ldots, X^n_k) - \partial_k^- f(X^n_1, X^n_2, X^n_k)| + \ldots \\
+ |\partial_k^- f(X^n_1, \ldots, X^n_{n-1}, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_{n-1}, X^n_k)|.
\end{align*}
\]
We consider only the term \( |\partial_k^- f(X^n_1, X^n_2, \ldots, X^n_k) - \partial_k^- f(X^n_1, X^n_2, X^n_k)| \), and the rest can be treated similarly.

Next we argue why we can apply local representation (2.1) globally even in the case when \( \text{supp}(\mu) \) is not necessarily compact. We define a set
\[
\Omega_N = \{ \omega \in \Omega : \sup_{k=1,\ldots,n} \sup_{t \in [0,T]} |X^n_k| \in [0,n] \}, \quad n \in \mathbb{N}.
\]
By monotone convergence theorem we have
\[
\mathbb{E}\left[ \left( \int_0^{2\delta} \int_0^s + \int_0^T \int_0^s \right) \left| \partial_k^- f(X^n_1, \ldots, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_k) \right| ds du \right] = \lim_{n \to \infty} \mathbb{E}\left[ \mathbf{1}_{\Omega_n} \left( \int_0^{2\delta} \int_0^s + \int_0^T \int_0^s \right) \left| \partial_k^- f(X^n_1, \ldots, X^n_k) - \partial_k^- f(X^n_1, \ldots, X^n_k) \right| ds du \right].
\]
Moreover, we define auxiliary function \( f_n \) by
\[
f_n(y) = \begin{cases} 
\partial_k^- f(-n, x^2, \ldots, x^n)(x^1 + n) + f(-n, x^2, \ldots, x^n), & x^1 < -n, \\
f(x^1, \ldots, x^n), & -n \leq x^1 \leq n, \\
\partial_k^+ f(n, x^2, \ldots, x^n)(x^1 - n) + f(n, x^2, \ldots, x^n), & x^1 > n.
\end{cases}
\]
We take partial derivative with respect to \( x^1, \ldots, x^k \) separately, and for different \( k, f_n \) is different. Now the measure \( \mu \) associated to \( f_n \) has compact support and \( \partial_k^- f_n = \partial_k^- f \) on \( \Omega_N \). Hence we may apply representation (2.1).
to obtain
\[
\mathbb{E} \left[ 1_{\Omega_N} \left( \int_0^{2\delta} \int_s^T \int_{s-\delta}^T |\partial_k^- f(X^1_u, \ldots, X^n_u) - \partial_k^- f(X^1_{u'}, \ldots, X^n_{u'})| \frac{du}{(s-u)^{\beta+1}} \right) \right] \\
= \mathbb{E} \left[ 1_{\Omega_N} \int_0^T \int_0^s \left| \int_{-N}^N \text{sgn}(X^1_s - a) \mu_{X^1}(da) - \int_{-N}^N \text{sgn}(X^1_u - a) \mu_{X^1_u}(da) \right| \frac{du}{(s-u)^{\beta+1}} \right] \\
\leq \mathbb{E} \left[ \int_0^T \int_0^s \left| \int_{-N}^N \text{sgn}(X^1_s - a) \mu_{X^1}(da) - \int_{-N}^N \text{sgn}(X^1_u - a) \mu_{X^1_u}(da) \right| \frac{du}{(s-u)^{\beta+1}} \right].
\]

Here \( X^1_k = (X^1_s, \ldots, X^n_s) \) is an \( n-1 \) dimensional random variable without the first term. Hence by monotone convergence theorem again, we have
\[
\mathbb{E} \left( \int_0^{2\delta} \int_0^s \int_{2\delta}^T \frac{|\partial_k^- f(X^1_u, \ldots, X^n_u) - \partial_k^- f(X^1_{u'}, \ldots, X^n_{u'})|}{(s-u)^{\beta+1}} \right) \right) \\
\leq \mathbb{E} \left( \int_0^{2\delta} \int_0^s \int_{2\delta}^T \left| \int_{-N}^N \text{sgn}(X^1_s - a) \mu_{X^1}(da) - \int_{-N}^N \text{sgn}(X^1_u - a) \mu_{X^1_u}(da) \right| \frac{du}{(s-u)^{\beta+1}} \right).
\]

It follows that we only need to prove the following
\[
\mathbb{E} \left( \int_0^{2\delta} \int_0^s \int_{2\delta}^T \frac{\int_{-N}^N \text{sgn}(X^1_s - a) \mu_{X^1}(da) \mu(X^1_s - a < X^1_s, \bar{X}^1_n \in d\bar{x}^{n-1})}{|u-s|^{\beta+1}} \right) \right) \\
= \mathbb{E} I_1 + \mathbb{E} I_2 < \infty.
\]

Next we prove that \( \mathbb{E} I_1 < \infty \), the other term \( I_2 \) is easier and can be treated similarly. For \( I_1 \), by applying Tonelli’s theorem, we get
\[
\mathbb{E} I_1 = \\
\int_0^{2\delta} \int_0^s \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \frac{P(X^1_s < a < X^1_u) + P(X^1_s < a < X^1_u)}{|u-s|^{\beta+1}} \mu_{X^1_u}(da) \mu(X^1_s - a < X^1_s, \bar{X}^1_n \in d\bar{x}^{n-1}) duds.
\]

We only consider the case \( P(X^1_u < a < X^1_s) \), and by symmetry the result also holds for the other one. Now introduce time points
\[
s_0 = 2\delta, \quad s_k = 2\delta - 2\delta \sum_{j=1}^{k} \left( \frac{1}{2} \right)^j = 2\delta \left( \frac{1}{2} \right)^k, \quad k \geq 1.
\]
and split the integral as
\[
\int_0^{2\delta} \int_0^s \ldots duds = \sum_{k=0}^{\infty} \int_{s_k}^{s_{k+1}} \int_{\frac{s}{2}}^s \ldots duds.
\]
Next we note that by applying Lemma 2.2 it is sufficient to show

$$\sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{R^{n-1}} \int_{R} C(a) \frac{\sqrt{W_1(u,s)}}{\sqrt{V_1(u)(s-u)^{1+\beta}}} \left( 1 + \frac{|a|}{\sqrt{V_1(u)}} \right) \mu_{\bar{X}_u^1}(da) P(\bar{X}_u^1 \in d\bar{x}^{n-1}) duds$$

$$= \sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{R^{n-1}} \int_{R} C(a) \frac{\sqrt{W_1(u,s)}}{\sqrt{V_1(u)(s-u)^{1+\beta}}} \mu_{\bar{X}_u^1}(da) P(\bar{X}_u^1 \in d\bar{x}^{n-1}) duds$$

$$+ \sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{R^{n-1}} \int_{R} C(a) \frac{|a| \sqrt{W_1(u,s)}}{\sqrt{V_1(u)(s-u)^{1+\beta}}} \mu_{\bar{X}_u^1}(da) P(\bar{X}_u^1 \in d\bar{x}^{n-1}) duds$$

$$= J_1 + J_2$$

$$< \infty,$$

where $W_1(u,s)$ and $V_1(u)$ denote incremental variance and variance functions of the process $X^1$. Next we prove $J_2 < \infty$, the finiteness of $J_1$ can be treated similarly and is more easy to handle. Now by assumption we have $W_1(u,s) \leq (s-u)^{2\alpha_1}$. Moreover,

$$\sup_{s_{k+1} \leq u \leq s_k} V_1(u) \leq C s_k^{2\alpha_1},$$

or

$$\inf_{s_{k+1} \leq u \leq s_k} V_1(u) > 0$$

in which case the proof is trivial. Next we note that, by independency of $X^1, \ldots, X^n$, we have

$$P(\bar{X}_u^1 \in d\bar{x}) = f_{u,2}(x^2) \ldots f_{u,n}(x^n)d\bar{x}^{n-1},$$

where $f_{u,i}(x^i)$ are the density functions of the $i$-th Gaussian process at time $u$. Split the integral with respect to $\bar{X}_u^1$ into

$$\int_{R^{n-1}} \ldots d\bar{x}^{n-1} = \sum_{i=0, \ldots, n-1}^{i+j=n-1} \int_{|x^i| > 1} \ldots \int_{|x^i| \leq 1} d\bar{x}^{n-1}.$$ 

When $|x^i| > 1$, we have

$$f_{u,i}(x^i) = \frac{e^{-\frac{(x^i)^2}{2V_1(u)}}}{\sqrt{V_1(u)2\pi}} \leq C e^{-\frac{|x^i|^2}{2\pi}}$$

since now $u < 2\delta$ and $\delta$ can be chosen small enough. The idea of the proof is that we take maximum with respect to those variables that lie on interval
$[-1, 1]$. We obtain

$$
\sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} |a| e^{-\frac{a^2}{2C_k s_k^2}} \frac{(s-u)^\alpha_1}{u^2(s-u)^{\beta+1}} \mu_{X_1}^a(da) \mathbb{P}(\bar{X}_1^i \in d\bar{x}^{n-1}) duds
$$

$$
= \sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \left( \sum_{\sigma \in \{0, \sigma(1), \ldots, \sigma(n-1)\}} \int_{[-1,1]^{n-1-i}} \right)
$$

$$
\int_{\mathbb{R}\setminus[-1,1]^i} \max_{-1 \leq x^i \leq 1} \int_{\mathbb{R}} |a| e^{-\frac{a^2}{2C_k s_k^2}} \frac{(s-u)^\alpha_1}{u^2(s-u)^{\beta+1}} \mu_{X_1^i}^a(da) \mathbb{P}(\bar{X}_1^i \in d\bar{x}^{n-1}) duds
$$

$$
\leq C \sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \left( \sum_{\sigma \in \{0, \sigma(1), \ldots, \sigma(n-1)\}} \int_{\mathbb{R}\setminus[-1,1]^i} \right)
$$

$$
\int_{\mathbb{R}\setminus[-1,1]^i} \max_{-1 \leq x^i \leq 1} \int_{\mathbb{R}} |a| e^{-\frac{a^2}{2C_k s_k^2}} \frac{(s-u)^\alpha_1}{u^2(s-u)^{\beta+1}} \mu_{X_1^i}^a(da) \mathbb{P}(\bar{X}_1^{\sigma(i)} \in dx^i) ds
$$

$$
\leq C \sum_{k=0}^{\infty} \int_{s_{k+1}}^{s_k} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \left( \sum_{\sigma \in \{0, \sigma(1), \ldots, \sigma(n-1)\}} \int_{\mathbb{R}\setminus[-1,1]^i} \right)
$$

$$
\int_{\mathbb{R}\setminus[-1,1]^i} \max_{-1 \leq x^i \leq 1} \int_{\mathbb{R}} |a| e^{-\frac{a^2}{2C_k s_k^2}} s^{-2s_1^{\alpha_1-\beta-1}} \mu_{X_1^{\sigma(i)}, \bar{x}^i}(da) \mathbb{P}(\bar{X}_1^{\sigma(i)} \in dx^i)
$$

where $\mathbb{P}(\bar{X}_1^{\sigma(i)} \in dx^i)$ is the Gaussian measure related to $(X^{\sigma(1)}, \ldots, X^{\sigma(i)})$ and on the second inequality we have used (5.1) to get rid of dependence of time $u$ of the Gaussian measure.

Next we split the integral with respect to $a$ into

$$
\int_{\mathbb{R}} \ldots da = \int_{-1}^{1} \ldots da + \int_{\mathbb{R}\setminus[-1,1]} \ldots da,
$$

and observe that when $|a| > 1$,

$$
\sum_{k=0}^{\infty} |a| e^{-\frac{a^2}{2C_k s_k^2}} s_1^{\alpha_1-\beta-1} \leq C |a| e^{-\frac{a^2}{2}}.
$$
Hence we obtain

\[ C \sum_{k=0}^{\infty} \sum_{\sigma \in \{0, \sigma(1), \ldots, \sigma(n-1)\}} \int_{[0,1]} \max_{j=0, \ldots, n-1-\iota} \int_{\mathbb{R}} |a| e^{-\frac{s_k^2}{2} \mu_{\bar{X}^{(i)}, \bar{x}}(da)} \mathbb{P}(\bar{X}^{(i)} \in d\bar{x}) \]

\[ \leq C \sum_{\sigma \in \{0, \sigma(1), \ldots, \sigma(n-1)\}} \int_{[0,1]} \max_{j=0, \ldots, n-1-\iota} \int_{\mathbb{R}} |a| e^{-\frac{s_k^2}{2} \mu_{\bar{X}^{(i)}, \bar{x}}(da)} \mathbb{P}(\bar{X}^{(i)} \in d\bar{x}), \]

which is finite by Definition 2.3.

Now take to the case \( a \in [0, 1] \), and the case \( a \in [-1, 0] \) can be treated similarly. We need to show that

(3.2)

\[ C \sum_{k=0}^{\infty} \sum_{\sigma \in \{0, \sigma(1), \ldots, \sigma(n-1)\}} \int_{[0,1]} \max_{j=0, \ldots, n-1-\iota} \int_{0}^{1} |a| e^{-\frac{s_k^2}{2} \mu_{\bar{X}^{(i)}, \bar{x}}(da)} \mathbb{P}(\bar{X}^{(i)} \in d\bar{x}) < \infty. \]

Now take a smooth function \( f_n(x) \) and take \( \psi_{\epsilon} \in C^\infty \) for \( \epsilon > 0 \) with compact support which approximates in uniform norm Dirac delta function \( \delta_\alpha \), i.e. (see [1])

\[ \lim_{\epsilon \to 0} \psi_{\epsilon}(a) = \delta_\alpha. \]

We get

\[ \int_{\mathbb{R}} \psi_{\epsilon} f_n'(x) dx \to \int_{\mathbb{R}} \psi_{\epsilon} \mu(dx) \]

\[ \to \int_{\mathbb{R}} \delta_\alpha \mu(dx) \]

\[ = \mu(a) \]

\[ < \infty. \]

On the other hand, by dominated convergence theorem we have

\[ \int_{\mathbb{R}} \psi_{\epsilon} f_n''(x) dx \to \int_{\mathbb{R}} \delta_\alpha f_n''(x) dx = f_n''(a), \quad \epsilon \to 0. \]

And we know that

\[ \int_{0}^{1} ae^{-\frac{s_k^2}{2} \mu_{\bar{X}^{(i)}, \bar{x}}(da)} \leq C \max_{0 \leq a \leq 1} f_n''(a) s_{2\alpha_1}, \]

hence

\[ \sum_{k=0}^{\infty} \sum_{i=\sigma} \int_{[0,1]} \max_{j=0, \ldots, n-1-\iota} \int_{0}^{1} ae^{-\frac{s_k^2}{2} \mu_{\bar{X}^{(i)}, \bar{x}}(da)} \mathbb{P}(\bar{X}^{(i)} \in d\bar{x}) \]

\[ \leq C \sum_{i=\sigma} \int_{[0,1]} \sum_{k=0}^{\infty} \max_{0 \leq a \leq 1, 0 \leq x \leq 1} f_n''(a) s_{k}^{2\alpha_1 - 1} \mathbb{P}(\bar{X}^{(i)} \in d\bar{x}). \]

Now \( \sum_{k=0}^{\infty} s_{k}^{2\alpha_1 - 1} < \infty \) since \( 3\alpha_1 - \beta - 1 > 0 \) and therefore (3.2) holds.
The message of next theorem is that the above stochastic integral can also be understood as a Föllmer integral which is more natural for many applications.

**Theorem 3.2.** Let $X \in X^g$ and $f \in \mathcal{Y}$. Then for any partition $\pi_n = \{0 = t^n_0 < \ldots < t^n_{k(n)} = T\}$ we have

$$
\sum_{i=1}^{k(n)} \partial^- f(X^n_{t^n_{i-1}}, \ldots, X^n_{t^n_{i-1}}) (X^n_{t^n_{i-1}} - X^n_{t^n_{i-1}}) \overset{a.s.}{\longrightarrow} \int_0^T \partial^- f(X^1_t, \ldots, X^n_t) dX^k_t.
$$

**Proof.** The proof is similar to the proof in [1]. Moreover, as in Theorem 3.1, we can use the representation (2.1) globally. First we have

$$
I_n = \sum_{i=1}^{k(n)} \partial^- f(X^n_{t^n_{i-1}}, \ldots, X^n_{t^n_{i-1}}) (X^n_{t^n_{i-1}} - X^n_{t^n_{i-1}}) - \int_0^T \partial^- f(X^1_t, \ldots, X^n_t) dX^k_t
$$

$$
= \int_0^T \left( \sum_{i=1}^{k(n)} \partial^- f(X^n_{t^n_{i-1}}, \ldots, X^n_{t^n_{i-1}}) \mathbf{1}_{(t^n_{i-1}, t^n_i]}(t) - \partial^- f(X^1_t, \ldots, X^n_t) \right) dX^k_t.
$$

Denote

$$
h_n(t) = \left( \sum_{i=1}^{k(n)} \partial^- f(X^n_{t^n_{i-1}}, \ldots, X^n_{t^n_{i-1}}) \mathbf{1}_{(t^n_{i-1}, t^n_i]}(t) - \partial^- f(X^1_t, \ldots, X^n_t) \right),
$$

and since $\partial^- f$ is continuous except on a countable set, thus $|h_n| \to 0$ pointwise. According to Theorem 2.1 we need to show that

$$
\|h_n(t)\|_{2,\beta} \to 0.
$$

Now

$$
|h_n(t)| \leq 2 \sup_{t \in [0,T]} |\partial^- f(X^1_t, \ldots, X^n_t)|, \quad a.s.
$$

and by dominated convergence theorem we have

$$
\int_0^T \frac{|h_n(t)|}{t^\beta} dt \to 0 \quad a.s. \quad as \quad n \to \infty.
$$
Moreover, following the same argument in [1], we obtain for $0 \leq s \leq t \leq T$
\[
\int_{k(n)} \sum_{i=1}^{k(n)} \partial_k f(X_{t_{i-1}}^{1}, \ldots, X_{t_{i}}^{n}) (t_{i-1}, t_{i}) - \sum_{j=1}^{k(n)} \partial_k f(X_{s_{j-1}}^{1}, \ldots, X_{s_{j}}^{n}) (t_{j-1}, t_{j}) \int_{(t_{j-1}, t_{j})}^{(t_{n}, t_{n})} (s, t) \leq \int_{k(n), j<i} \sum_{k=1}^{n} (\partial_k f(X_{t_{j-1}}^{1}, \ldots, X_{t_{j}}^{n}) - \partial_k f(X_{s_{j-1}}^{1}, \ldots, X_{s_{j}}^{n})) (t_{j-1}, t_{j}) \int_{(t_{j-1}, t_{j})}^{(t_{n}, t_{n})} (s, t) \leq \int_{k(n), j<i} \sum_{k=1}^{n} \int_{R} \left( X_{t_{j-1}}^{1} < a < X_{t_{j}}^{1} + X_{t_{j-1}}^{1} < a < X_{t_{j}}^{1} \right) \mu_{X_{t_{j-1}}^{1}} (da) \int_{(t_{j-1}, t_{j})}^{(t_{n}, t_{n})} (s, t) \mu_{X_{t_{j-1}}^{1}} (da) + \ldots
\]
where $\check{X}^{k} = (X^{1}, \ldots, X^{k-1}, X^{k+1}, \ldots, X^{n})$ which is $n - 1$ dimensional random variable without the $k$-th item and notice that different $\mu$ depends on different random variables now. Hence according to Theorem 3.1 we obtained integrable dominants and by dominated convergence theorem we have
\[
\int_{0}^{T} \int_{0}^{t} \frac{|h_{n}(t) - h_{n}(s)|}{(t-s)^{\beta+1}} ds dt \to 0 \quad a.s. \text{ as } n \to \infty.
\]

\[\square\]

3.2. **Itô formula.** In this section we will prove Itô formula which is the main theorem of this paper. First we begin with the following smooth version.

**Theorem 3.3.** Let $X \in X^{0}$ and $f \in C^{2}(\mathbb{R}^{n})$. Then
\[
f(X_{T}^{1}, \ldots, X_{T}^{n}) = f(X_{0}^{1}, \ldots, X_{0}^{n}) + \sum_{k=1}^{n} \int_{0}^{T} \partial_{k} f(X_{t}^{1}, \ldots, X_{t}^{n}) dX_{t}^{k}.
\]

**Proof.** By Taylor expansion we have
\[
f(X_{T}^{1}, \ldots, X_{T}^{n}) = f(X_{0}^{1}, \ldots, X_{0}^{n}) + \sum_{k=1}^{n} \int_{0}^{T} \partial_{k} f(X_{t}^{1}, \ldots, X_{t}^{n}) dX_{t}^{k}
\]
\[+ \frac{1}{2} \sum_{k=1, j=1}^{n} \int_{0}^{T} \partial_{kj} f(X_{t}^{1}, \ldots, X_{t}^{n}) d[X_{t}^{k}, X_{t}^{j}].\]
Then because $X^k_t$ has zero quadratic variation, we have

$$f(X^1_T, \ldots, X^n_T) = f(X^1_0, \ldots, X^n_0) + \sum_{k=1}^{n} \int_0^T \partial_k f(X^1_t, \ldots, X^n_t) dX^k_t.$$ 

\[\square\]

The next theorem is our main result in this section. It turns out that, as in the one dimensional case, the existence of the integral is the crucial fact in order to obtain Itô formula.

**Theorem 3.4.** Let $X \in \mathcal{X}^\alpha$ and $f \in \mathcal{Y}$. Then

$$f(X^1_T, \ldots, X^n_T) = f(X^1_0, \ldots, X^n_0) + \sum_{k=1}^{n} \int_0^T \partial_k f(X^1_t, \ldots, X^n_t) dX^k_t.$$ 

**Proof.** The idea of the proof is to use Theorem 3.3. In one-dimensional case the proof for fractional Brownian motion can be found in [1] and for general Gaussian processes in [10]. In our multidimensional case, for simplicity we present the steps only in two dimensional case as the general case follows similar arguments. Assume now that $f(x) \in \mathcal{Y}$ with $x = (x^1, x^2)$. Moreover, we again assume that $\mu$ is positive measure. Let $\eta^\epsilon(x) = \frac{1}{\epsilon^2} \eta(\frac{x}{\epsilon})$, where for $x \in \mathbb{R}^2$,

$$\eta(x) = \begin{cases} C \exp\left(\frac{-1}{1-|x|^2}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

We choose $C$ such that $\int_{\mathbb{R}^2} \eta(x) dx = 1$, then $\eta^\epsilon(x) \in C^\infty$ and has finite support. Now define

$$f_\epsilon(x) = \int_{\mathbb{R}^2} \eta^\epsilon(x-y)f(y)dy = \int_{\mathbb{R}^2} \eta^\epsilon(z)f(x-z)dz.$$ 

Now $f_\epsilon(x) \in C^\infty$ and $f_\epsilon(x)$ converges to $f(x)$ pointwise. For $k = 1, 2$ we have

$$\partial_{x^k} f_\epsilon(x) = \partial_{x^k} \int_{\mathbb{R}^2} \eta^\epsilon(x-y)f(y)dy = (-1) \int_{\mathbb{R}^2} \partial_{y^k} \eta^\epsilon(x-y)f(y)dy$$

by chain rule. Then by weak derivative property we have

$$(-1) \int_{\mathbb{R}^2} \partial_{y^k} \eta^\epsilon(x-y)f(y)dy = \int_{\mathbb{R}^2} \eta^\epsilon(x-y) \partial_{y^k} f(y)dy.$$ 

This is indeed true since the partial derivative exists almost everywhere except on a countable set and thus $\partial_k f_\epsilon(x)$ converge to $\partial_k f(x)$ almost everywhere.

By Theorem 3.3 we have

$$f_\epsilon(X^1_T, Y_T) = f_\epsilon(X^1_0, Y_0) + \int_0^T \partial_{x^1} f_\epsilon(X^1_t, Y_t) dX_t + \int_0^T \partial_{y^1} f_\epsilon(X^1_t, Y_t) dY_t.$$
Now \( f_\varepsilon \) converges to \( f \) pointwise and \( \partial_k f_\varepsilon \) converges to \( \partial_k^- f \) almost everywhere for \( k = x, y \). Hence \( f_\varepsilon(X_t, Y_t) \rightarrow f(X_T, Y_T) \) and \( f_\varepsilon(x_0, y_0) \rightarrow f(x_0, y_0) \). The only thing left to show whether

\[
\int_0^T \partial_y f_\varepsilon(X_t, Y_t) \, dx_t \rightarrow \int_0^T \partial_y^- f(X_t, Y_t) \, dx_t
\]

and

\[
\int_0^T \partial_x f_\varepsilon(X_t, Y_t) \, dy_t \rightarrow \int_0^T \partial_x^- f(X_t, Y_t) \, dy_t.
\]

Here we only consider the first one and the second one is similar. Now by Theorem 2.1 and almost everywhere convergence of \( \partial_k f_\varepsilon \) we only need to prove

\[
\|\partial_y f_\varepsilon(X, Y) - \partial_y^- f(X, Y)\|_{2, \beta} \rightarrow 0, \text{ a.s.}
\]

as \( n \rightarrow \infty \).

For the first term in norm \( \| \cdot \|_{2, \beta} \) we have

\[
\frac{|\partial_y f_\varepsilon(X_t, Y_t) - \partial_y^- f(X_t, Y_t)|}{t^\beta} \leq \frac{C \sup_{t \in [0, T]} |\partial_y^- f(X_t, Y_t)|}{t^\beta} \in L^1([0, T], dt).
\]

Hence, thanks to Lebesgue dominated convergence theorem, we have

\[
\int_0^T \frac{|\partial_y f_\varepsilon(X_t, Y_t) - \partial_y^- f(X_t, Y_t)|}{t^\beta} \, dt \rightarrow 0 \quad \text{a.s.}
\]

Next consider

\[
\frac{|\partial_y f_\varepsilon(X_t, Y_t) - \partial_y^- f(X_t, Y_t) - \partial_y f_\varepsilon(X_s, Y_s) + \partial_y^- f(X_s, Y_s)|}{|t - s|^{1+\beta}}
\]

\[
\leq \frac{|\partial_y f_\varepsilon(X_t, Y_t) - \partial_y f_\varepsilon(X_s, Y_s)|}{|t - s|^{1+\beta}} + \frac{|\partial_y^- f(X_t, Y_t) - \partial_y^- f(X_s, Y_s)|}{|t - s|^{1+\beta}}
\]

From Theorem 3.1 we know that the second term is integrable. Consequently, we only have to consider the first term. We write

\[
\frac{|\partial_y f_\varepsilon(X_t, Y_t) - \partial_y f_\varepsilon(X_s, Y_s)|}{|t - s|^{1+\beta}}
\]

\[
\leq \frac{|\partial_y f_\varepsilon(X_s, Y_s)|}{|t - s|^{1+\beta}} + \frac{|\partial_y f_\varepsilon(X_s, Y_s)|}{|t - s|^{1+\beta}}
\]

By mean value theorem and Hölder continuity of \( Y_t \) we have

\[
\frac{|\partial_y f_\varepsilon(X_s, Y_s)|}{|t - s|^{1+\beta}}
\]

\[
\leq \frac{|\partial_{yy} f_\varepsilon(X_s, \theta_\omega)|}{|t - s|^{1+\beta}} |Y_t - Y_s|
\]

\[
\leq \sup_{s \in [0, T]} |\partial_{yy} f_\varepsilon(X_s, \theta_\omega)| (t - s)^{\alpha_y - \delta},
\]
where \( \theta_\omega \) is between \( Y_s(\omega) \) and \( Y_t(\omega) \). According to the proof of Theorem 3.1, we know that \( \partial_yf_\varepsilon(\theta_\omega) \to \mu(\theta_\omega) \). Therefore \( \partial_yf_\varepsilon(\theta_\omega) \) is uniformly bounded in \( n \). Consequently,

\[
\int_0^T \int_0^t \frac{|\partial_yf_\varepsilon(X_s, Y_t) - \partial_yf_\varepsilon(X_s, Y_s)|}{|t-s|^{1+\beta}} ds dt \to 0
\]

by dominated convergence theorem.

For the remaining term by mean value theorem and Hölder continuity of \( X_t \), we have

\[
|\partial_yf_\varepsilon(X_t, Y_t) - \partial_yf_\varepsilon(X_s, Y_t)| \leq |\partial_yf_\varepsilon(\hat{\theta}_\omega, Y_t)||X_t - X_s| |t-s|^{1+\beta} \leq \sup_{t \in [0,T]} |\partial_yf_\varepsilon(\hat{\theta}_\omega, Y_t)|(t-s)^{\alpha_x - \delta}.
\]

Again, according to the proof of Theorem 3.1 and assumption we know that \( \partial_yf_\varepsilon(\hat{\theta}_\omega) \to \hat{\mu}(\hat{\theta}_\omega) \) where now \( \hat{\mu}(\hat{\theta}_\omega) \) is different from \( \mu(\theta_\omega) \). Therefore \( \partial_yf_\varepsilon(\hat{\theta}_\omega) \) is uniformly bounded in \( n \). Consequently,

\[
\int_0^T \int_0^t \frac{|\partial_yf_\varepsilon(X_t, Y_t) - \partial_yf_\varepsilon(X_s, Y_t)|}{|t-s|^{1+\beta}} ds dt \to 0
\]

by dominated convergence theorem.

To conclude we obtain

\[
\left\| \partial_yf_\varepsilon(X,Y) - \partial_yf(X,Y) \right\|_{2,\beta} \to 0, \quad n \to \infty
\]

by dominated convergence theorem. Following similar arguments in \( n \)-dimensional case we can show that for each \( k = 1, \ldots, n \)

\[
\left\| \partial_k f_\varepsilon(X_t^1, \ldots, X_t^n) - \partial_k f(X_t^1, \ldots, X_t^n) \right\|_{2,\beta} \to 0, \quad n \to \infty.
\]

\( \square \)

**Example 3.1.** If \( f(\bar{x}) = |\sum_{k=1}^n x_k| \), then the Itô formula becomes

\[
\left| \sum_{k=1}^n X_t^k \right| = \sum_{k=1}^n \int_0^T sgn(X_t^k) dX_t^k.
\]

This also comes directly from one dimensional case as in [10].

**Example 3.2.** In the case of \( f(x,y) = |xy| \), we obtain the following integration by parts formula

\[
|X_T Y_T| = |X_0 Y_0| + \int_0^T \partial_x |X_t Y_t| dX_t + \int_0^T \partial_y |X_t Y_t| dY_t
= |X_0 Y_0| + \int_0^T |Y_t| sgn(X_t) dX_t + \int_0^T |X_t| sgn(Y_t) dY_t
\]
4. Discussions

In this article we have studied existence of multidimensional stochastic integrals with respect to Gaussian processes and our results have several significant benefits. Firstly, multidimensional integrals are not widely studied in the literature. In particular, usually the considered integrands have more regularity than in our case such as Hölder continuous trajectories. Secondly, our results cover wide class of Gaussian processes. Moreover, we do not assume that our processes $X^k$ are independent copies of each others. Hence we can cover different kind of mixed models. For particular example, our results cover multidimensional fractional Brownian motion where all the processes $B^H_k$ may have different Hurst index $H_k > \frac{1}{2}$. Similarly, some processes in the model can have stationary increments, some can be stationary and some processes can be neither. This is particularly interest for applications where there exist several random sources, and observations suggest different stylized facts for different random sources. Moreover, our results hold also for dependent processes with obvious modifications provided that conditional Gaussian processes belong to the class considered.

References

[1] Mishura Y. Azmoodeh E. and Valkeila E. On hedging european options in geometric fractional brownian motion market model. Statistics and Decisions, 27:129–143, 2010.
[2] Nualart D. and Răşcanu A. Differential equations driven by fractional brownian motion. Collect. Math., 53:55–81, 2002.
[3] Sondermann D. Introduction to Stochastic Calculus for Finance: A New Didactic Approach. Springer, 2006.
[4] Azmoodeh E. and Viitasaari L. Rate of convergence for discretization of integrals with respect to fractional Brownian motion. J. Theor. Probab., DOI:10.1007/s10959–013–0495–y, 2013.
[5] Föllmer H. Calcul d’ito sans probabilités. Séminaire de probabilités, 15:143–150, 1981.
[6] Young L.C. An inequality of the hölder type, connected with stieltjes integration. Acta. Math., 67:251–282, 1936.
[7] Zähle M. Integration with respect to fractal functions and stochastic calculus, part i. Probab. Theory Relat. Fields, 111:333–372, 1998.
[8] Shevchenko G. Mishura Y. and Valkeila E. Random variables as pathwise integrals with respect to fractional Brownian motion. Stochastic processes and their Applications, 123(6):2353–2369, 2013.
[9] Lin S.J. Stochastic analysis of fractional Brownian motions. Stochastics and Stoch. Reports, 55:121–140, 1995.
[10] Sottinen T. and Viitasaari L. Pathwise integrals and itô-tanaka formula for gaussian processes. submitted, arxiv:1307.3578, 2013.
[11] Dai W. and Heyde C.C. Itô’s formula with respect to fractional Brownian motion and its application. Journal of Appl. Math. and Stoch. An., 9:439–448, 1996.