

Pathwise Taylor expansions for random fields on multiple dimensional paths

Rainer Buckdahn^{a,b}, Jin Ma^c, Jianfeng Zhang^{c,*}

^a *Département de Mathématiques, Université de Bretagne-Occidentale, F-29285 Brest Cedex, France*

^b *School of Mathematics, Shandong University, Jinan 250100, PR China*

^c *Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA*

Received 15 October 2013; received in revised form 7 February 2015; accepted 11 February 2015

Available online 20 February 2015

Abstract

In this paper we establish the pathwise Taylor expansions for random fields that are “regular” in terms of Dupire’s path-derivatives [6]. Using the language of pathwise calculus, we carry out the Taylor expansion naturally to any order and for any dimension, which extends the result of Buckdahn et al. (2011). More importantly, the expansion can be both “forward” and “backward”, and the remainder is estimated in a pathwise manner. This result will be the main building block for our new notion of viscosity solution to forward path-dependent PDEs corresponding to (forward) stochastic PDEs in our accompanying paper Buckdahn et al. [4].

© 2015 Elsevier B.V. All rights reserved.

MSC: 60H07, 15, 30; 35R60; 34F05

Keywords: Path derivatives; Pathwise Taylor expansion; Functional Itô formula; Itô–Wentzell formula; Stochastic partial differential equations

1. Introduction

In this paper we are interested in establishing the pathwise Taylor expansions for the Itô-type random field of the form

$$u(t, x) = u_0(x) + \int_0^t \alpha(s, x) ds + \int_0^t \beta(s, x) \circ dB_s, \quad (1.1)$$

* Corresponding author.

E-mail addresses: Rainer.Buckdahn@univ-brest.fr (R. Buckdahn), jinma@usc.edu (J. Ma), jianfenz@usc.edu (J. Zhang).

where B is a d -dimensional standard Brownian motion, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and “ \circ ” means Stratonovich integral. In particular we are interested in such expansions for the solution to the following fully nonlinear stochastic partial differential equations (SPDE):

$$u(t, x) = u_0(x) + \int_0^t f(s, x, \cdot, u, \partial_x u, \partial_{xx} u) ds + \int_0^t g(s, x, \cdot, u, \partial_x u) \circ dB_s, \tag{1.2}$$

where f and g are random fields that have certain regularity in their spatial variables.

In our previous work [3] we studied the so-called *pathwise stochastic Taylor expansion* for a class of Itô-type random fields. The main result can be briefly described as follows. Suppose that u is a random field of the form (1.1), and B is a one dimensional Brownian motion. If we denote $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ to be the natural filtration generated by B and augmented by all \mathbb{P} -null sets in \mathcal{F} , then under reasonable regularity assumptions on the integrands α and β , the following stochastic “Taylor expansion” holds: For any stopping time τ and any \mathcal{F}_τ -measurable, square-integrable random variable ξ , and for any sequence of random variables $\{(\tau_k, \xi_k)\}$ where τ_k ’s are stopping times such that either $\tau_k > \tau, \tau_k \downarrow \tau$; or $\tau_k < \tau, \tau_k \uparrow \tau$, and ξ_k ’s are all $\mathcal{F}_{\tau_k \wedge \tau}$ -measurable, square integrable random variables, converging to ξ in L^2 , it holds that

$$u(\tau_k, \xi_k) = u(\tau, \xi) + a(\tau_k - \tau) + b(B_{\tau_k} - B_\tau) + p(\xi_k - \xi) + \frac{c}{2}(B_{\tau_k} - B_\tau)^2 + q(\xi_k - \xi)(B_{\tau_k} - B_\tau) + \frac{1}{2}X(\xi_k - \xi)^2 + o(|\tau_k - \tau| + |\xi_k - \xi|^2), \tag{1.3}$$

where (a, b, c, p, q, X) are all \mathcal{F}_τ -measurable random variables, and the remainder $o(\zeta_k)$ are such that $o(\zeta_k)/\zeta_k \rightarrow 0$ as $k \rightarrow \infty$, in probability. Furthermore, the six-tuple (a, b, c, p, q, X) can be determined explicitly in terms of α, β and their derivatives (in certain sense).

While the Taylor expansion (1.3) reveals the possibility of estimating the remainder in a stronger form than mean-square (see for example [12]), it is not satisfactory for the study of pathwise property of the random fields which is essential in the study of, for example, stochastic viscosity solution. In a subsequent paper (Buckdahn–Bulla–Ma [2]) the result was extended to the case where the expansion could be made around any random time–space point (τ, ξ) where τ does not have to be a stopping time; and more importantly, the remainder was estimated in a pathwise manner, in the spirit of the Kolmogorov continuity criterion. In other words, modulo a \mathbb{P} -null set, the estimate holds for each ω , locally uniformly in (t, x) . Furthermore, all the coefficients can be calculated explicitly in terms of a certain kind of “derivatives” for Itô-type random field, introduced in [2] (see more detailed description in Section 8 of this paper). It is noted, however, that a main drawback of the result in [2] is that the derivatives involved are not intuitive, and are difficult to verify. A more significant weakness of the result is that the dimension of the Brownian motion is restricted to 1, which, as we shall see in this paper, reduced the complexity of the Taylor expansion drastically.

The main purpose of this paper is to re-investigate the Taylor expansion in a much more general setting, but with a different “language”. In particular, we shall allow both the spatial variable and the Brownian motion to be multi-dimensional, and the random field is “regular” in a very different way. To be more precise, we shall introduce a new notion of “path-derivative” in the spirit of Dupire [6] to impose a different type of regularity, that is, the regularity on the variable $\omega \in \Omega$. Such a language turns out to be very effective, and many originally cumbersome expressions in stochastic analysis becomes intuitive and very easy to understand. For example, the Itô–Wentzell formula reads exactly like the multi-dimensional Itô formula, and both inte-

grands for the Lebesgue integral and the stochastic integral can be memorized simply as “chain rule”, with respect to time and path, respectively. For this reason we shall name it “pathwise Itô–Wentzell formula” (see Section 4 for details). We should note that our path derivative applies to all semi-martingale random fields that are not necessarily “continuous” in ω . In this sense it is weaker than the original one by Dupire (see also [5]). But on the other hand such a generality brings out some intrinsic “rough-path” nature of the Brownian motion. Among other things, for example, the “Hessian” under the current path-derivatives will be *asymmetric* in a general multi-dimensional setting, reflecting the nature of *Lévy area* in the rough-path theory (see, e.g., [8,15]).

We would like to point out that the Taylor expansion for stochastic processes, especially for the solutions of stochastic differential equations, is not new. There is a large amount of literature on the subject, from various perspectives. We refer to the books of Kloeden–Platen [12] from the numerical approximation point of view; and of Friz–Victoir [9] from the rough-path point of view, as well as the numerous references cited therein. In fact, all Taylor expansions resemble each other in their forms, language notwithstanding, and the difference often lies in the error (remainder) estimates. The main feature of our results is the following. First, our Taylor expansion applies to general random fields and stochastic processes, and therefore does not depend on the special structure for being a solution to a differential equation, whence “non-Markovian” in nature. Second, unlike our previous work, we shall provide a unified treatment of the Taylor expansion up to any order, and allowing the temporal increment to be both “forward” and “backward”. Finally, and most importantly, we pursue the pathwise estimate for the remainder, that is, the error of the expansion is estimated uniformly for all paths ω , modulo a common null set. The main difficulty, compared to an L^2 estimate (or in the sense of “in probability”) that we often see in the literature, is that one cannot use the isometry between the L^2 -norms of stochastic integrals and the L^2 -norms of the Lebesgue integrals, thus it requires some novel treatments of multiple integrals. The trade-off for being able to do this, however, is that we require some new regularities of the random field with respect to the “paths”. These requirements, when unified under our new language of “pathwise analysis”, are direct and easy to check. To our best knowledge, the pathwise Taylor expansion in such a generality is new.

It is worth noting that our Taylor expansion is the first step of our study of the viscosity solution to the (forward) “path-dependent PDEs” (PPDEs) corresponding to the forward SPDE (1.2), which will be the main topic of our accompanying paper [4]. We would only like to comment here that a classical solution in the traditional sense does not necessarily permit a pathwise Taylor expansion. Therefore a somewhat convoluted treatment of the solution to the SPDEs will have to be carried out based on the pathwise Taylor expansions, as it was seen in the deterministic viscosity solution theory as well as the existing studies of stochastic viscosity solutions (see, e.g., [3]).

The rest of the paper is organized as follows. In Section 2 we give all necessary notations and introduce the definition of “path-derivatives”. In Section 3 we give a heuristic analysis for a simpler case, the second order expansion for Itô processes, to illustrate the main points of our method. In Section 4 we prove the crucial estimates for the remainders of higher order Taylor expansions. In Section 5 we extend the Taylor expansion to Itô random fields; and in Section 6 we weaken the regularity assumptions of the coefficients to Hölder spaces. In Section 7 we apply the Taylor expansion to the solutions to stochastic PDEs, and finally, in Section 8 we compare the main theorem with our previous result [2].

2. Preliminaries

Throughout this paper we denote $\Omega := \{\omega \in C([0, \infty), \mathbb{R}^d) : \omega_0 = \mathbf{0}\}$ to be the set of continuous paths starting from the origin, B the canonical process on Ω , \mathbb{P}_0 the Wiener measure,

$\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ the \mathbb{P}_0 -augmented filtration generated by B , and $\Lambda := [0, \infty) \times \Omega$. Here and in the sequel, we use $\mathbf{0}$ to denote vectors or matrices with appropriate dimensions whose components are all equal to 0, and for any dimension m , we take the convention that $\mathbb{R}^m := \mathbb{R}^{m \times 1}$ denotes the set of column vectors. Define

$$x \cdot x' := \sum_{i=1}^m x_i x'_i \quad \text{for any } x, x' \in \mathbb{R}^m, \quad \gamma : \gamma' := \text{tr}[\gamma(\gamma')^T] \quad \text{for any } \gamma, \gamma' \in \mathbb{R}^{m \times n},$$

and $|x|^2 := x \cdot x$, $|\gamma|^2 := \gamma : \gamma$. Here T denotes the transpose.

2.1. Path derivatives for Itô processes

Let $\mathbb{L}^0(\Lambda, \mathbb{R}^{m \times n})$ denote the set of \mathbb{F} -progressively measurable processes $u : \Lambda \rightarrow \mathbb{R}^{m \times n}$, and $\mathbb{L}^0(\Lambda) := \mathbb{L}^0(\Lambda, \mathbb{R})$. Strongly motivated by the functional Itô formula initiated by Dupire [6] (see also Cont and Fournie [5] and a slight variation by Ekren–Touzi–Zhang [7]), in what follows we introduce the notion of “path-derivatives”, which will be the foundation of our pathwise stochastic analysis.

Recall that $u \in \mathbb{L}^0(\Lambda)$ is a semimartingale if there exist $A \in \mathbb{L}^0(\Lambda)$ and $\beta \in \mathbb{L}^0(\Lambda, \mathbb{R}^d)$ such that

$$u_t = u_0 + A_t + \int_0^t \beta_s \cdot dB_s, \quad \text{and} \quad V_0^t(A) + \int_0^t |\beta_s|^2 ds < \infty, \quad t \geq 0, \mathbb{P}_0\text{-a.s.}, \tag{2.1}$$

where $V_0^t(A)$ is the total variation of A on $[0, t]$.

Definition 2.1. Let u be a semimartingale in the form of (2.1). We define:

$$\partial_\omega u := \beta. \tag{2.2}$$

Moreover, if β is a semimartingale and $A_t = \int_0^t \alpha_s ds$ for some $\alpha \in \mathbb{L}^0(\Lambda)$, then we define

$$\partial_{\omega\omega}^2 u := \partial_\omega \beta = \partial_\omega(\partial_\omega u) \quad \text{and} \quad \partial_t u := \alpha - \frac{1}{2} \text{tr}(\partial_{\omega\omega}^2 u). \tag{2.3}$$

We remark that the path derivatives, whenever they exist, are unique in “ \mathbb{P}_0 -a.s.” sense.

Remark 2.2 (Functional Itô Formula). When the path derivatives $\partial_t u$, $\partial_\omega u$, $\partial_{\omega\omega}^2 u$ exist, we have $\alpha = \partial_t u + \frac{1}{2} \text{tr}(\partial_{\omega\omega}^2 u)$ and $\beta = \partial_\omega u$. In other words, the functional Itô formula holds: for $t \geq 0$

$$u_t = u_0 + \int_0^t \left[\partial_t u + \frac{1}{2} \text{tr}(\partial_{\omega\omega}^2 u) \right] (s, \cdot) ds + \int_0^t \partial_\omega u(s, \cdot) \cdot dB_s, \quad \mathbb{P}_0\text{-a.s.} \tag{2.4}$$

In particular, this implies that u is continuous in t . Equivalently, since $\beta = \partial_\omega u$ is a semimartingale, by using the Stratonovich integral, denoted by $\circ dB_s$, one has

$$u_t = u_0 + \int_0^t \partial_t u(s, \cdot) ds + \int_0^t \partial_\omega u(s, \cdot) \circ dB_s, \quad t \geq 0, \mathbb{P}_0\text{-a.s.} \quad \blacksquare \tag{2.5}$$

Remark 2.3. (i) The main result in [6,5] is the functional Itô formula (2.4), and in [7] the functional Itô formula (2.4) is used to define the derivatives. In this sense, our definition is consistent with theirs.

- (ii) In [6,5], one needs to extend the processes from Ω to the space of càdlàg paths. In [7] the definition is restricted to the space Ω only, but it still requires the processes and all the derivatives involved being continuous in ω . Our path-derivatives do not require such regularity, in particular our derivatives are defined only in “ \mathbb{P}_0 -a.s.” sense. In this aspect our definition is weaker, and is convenient for our study of SPDEs in [4], as typically one cannot expect the solution of a SPDE to be continuous in ω .
- (iii) In [6,5,7], the path derivative $\partial_{\omega}u$ is not required to be an Itô process. In this sense our definition is stronger. This is mainly because our pathwise Taylor expansion below requires stronger regularity than the functional Itô formula.
- (iv) When $u(t, \omega) = v(t, \omega_t)$ with $v \in C^{1,2}([0, \infty) \times \mathbb{R}^d)$, by the standard Itô formula we see that $\partial_{\omega}u(t, \omega_t) = \partial_x v(t, \omega_t)$. If we assume further that $\partial_x v \in C^{1,2}([0, \infty) \times \mathbb{R}^d)$, then $\partial_{\omega\omega}^2 u(t, \omega) = \partial_{xx}^2 v(t, \omega_t)$ and $\partial_t u(t, \omega) = \partial_t v(t, \omega_t)$. So our path derivatives are consistent with the standard derivatives in Markovian case. However, as pointed out in (iii), we need slightly stronger regularity requirements. ■

Remark 2.4. (i) In general the differential operators ∂_t and ∂_{ω} cannot commute. Moreover, in the case $d > 1$, the Hessian matrix $\partial_{\omega\omega}^2 u$ may not be symmetric, which implies that in general ∂_{ω^i} and ∂_{ω^j} do not commute either. See Example 2.5.

(ii) When $u(t, \omega) = v(t, \omega_{t_1 \wedge t}, \dots, \omega_{t_n \wedge t})$ for some t_1, \dots, t_n and some deterministic smooth function v , then $\partial_t, \partial_{\omega^i}, \partial_{\omega^j}$ do commute, and in particular $\partial_{\omega\omega}^2 u$ is symmetric.

(iii) The Hessian $\partial_{\omega\omega}^2 u$ in [6] (or [5]) could be symmetric, provided all path derivatives involved are continuous in (t, ω) . In fact, with certain regularity the Hessian $\partial_{\omega\omega}^2 u$ under our definition is also symmetric, see Remark 3.3. However, the non-commutativity of ∂_t and ∂_{ω} seems to be independent of the regularity, as we see in Example 2.5(i).

(iv) In [7] the $\partial_{\omega\omega}^2 u$ is by definition symmetric. Indeed, the $\partial_{\omega\omega}^2 u$ in [7] corresponds to $\frac{1}{2}[\partial_{\omega\omega}^2 u + (\partial_{\omega\omega}^2 u)^T]$ here. However, in this case the relation $\partial_{\omega\omega}^2 u = \partial_{\omega}(\partial_{\omega}u)$ may fail to hold, which not only is somewhat unnatural, but also makes the definition of higher order derivatives much more difficult. Our new definition modified this point. We should also note that the process u in Example 2.5(ii) is not continuous in (t, ω) , and thus is not in the framework of [5,6], or [7]. ■

Example 2.5. (i) Let $d = 1$ and $du = B_t dt$. Then $\partial_{\omega}u = 0, \partial_t u = B_t$. It is clear that $\partial_t \partial_{\omega}u = 0 \neq 1 = \partial_{\omega} \partial_t u$.

(ii) Let $d = 2$ and $du = B_t^2 dB_t^1$. Then $\partial_{\omega}u = [B^2, 0]^T$, and thus $\partial_{\omega\omega}^2 u = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not symmetric. In particular, $\partial_{\omega^1} \partial_{\omega^2} u = 0 \neq 1 = \partial_{\omega^2} \partial_{\omega^1} u$.

We note that the path derivatives can be extended to any order in a natural way. We now introduce some L^p spaces that will be frequently used in the paper. We begin by introducing the following norms on $u \in \mathbb{L}^0(\Lambda)$:

$$\|u\|_{0,p,T} := \|u\|_{p,T} := \sup_{0 \leq t \leq T} \left(\mathbb{E}[|u_t|^p] \right)^{1/p} \tag{2.6}$$

$$\|u\|_{1,p,T} := \|u\|_{p,T} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{0,p,T}; \tag{2.7}$$

$$\|u\|_{n,p,T} := \|u\|_{p,T} + \|\partial_t u\|_{n-2,p,T} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{n-1,p,T}, \quad n \geq 2.$$

We then define the spaces:

$$\mathcal{H}_p^{[n]}(\Lambda) := \left\{ u \in \mathbb{L}^0(\Lambda) : \|u\|_{n,p,T} < \infty, \forall T > 0 \right\}, \quad n \geq 0. \tag{2.8}$$

We shall also define the following Hölder norms: for any $\alpha \in (0, 1)$,

$$\begin{aligned} \|u\|_{0,p,\alpha,T} &:= \|u\|_{p,\alpha,T} := \|u\|_{p,T} + \mathbb{E} \left[\sup_{0 \leq t \leq T, \delta > 0} \frac{|u_{t+\delta} - u_t|^p}{\delta^{\frac{p\alpha}{2}}} \right]^{1/p}, \\ \|u\|_{1,p,\alpha,T} &:= \|u\|_{p,\alpha,T} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{p,\alpha,T}, \\ \|u\|_{n,p,\alpha,T} &:= \|u\|_{p,\alpha,T} + \|\partial_t u\|_{n-2,p,\alpha,T} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{n-1,p,\alpha,T}, \quad n \geq 2. \end{aligned} \tag{2.9}$$

Then we define correspondingly:

$$\mathcal{H}_p^{[n]+\alpha}(\Lambda) := \left\{ u \in \mathcal{H}_p^{[n]}(\Lambda) : \|u\|_{n,p,\alpha,T} < \infty, \forall T > 0 \right\}, \quad n \geq 0. \tag{2.10}$$

It should be noted that $\mathcal{H}_p^{[n+1]}(\Lambda)$ is not a subspace of $\mathcal{H}_p^{[n]+\alpha}(\Lambda)$ (!). However, one can show that if $p > 0$ is large enough (more precisely $p > \frac{2}{1-\alpha}$), then it holds that $\mathcal{H}_p^{[n+2]}(\Lambda) \subset \mathcal{H}_p^{[n]+\alpha}(\Lambda)$ for any $n \geq 0$ (see Lemma 6.1).

2.2. Path derivatives for random fields

Let $\mathcal{O} \subset \mathbb{R}^{d'}$ be an open domain, $\mathcal{Q} := [0, \infty) \times \mathcal{O}$, and $\hat{\Lambda} := \mathcal{Q} \times \Omega$. We denote by $\mathbb{L}^0(\hat{\Lambda}, \mathbb{R}^{m \times n})$ the set of \mathbb{F} -progressively measurable random fields $u : \hat{\Lambda} \rightarrow \mathbb{R}^{m \times n}$, and $\mathbb{L}^0(\hat{\Lambda}) := \mathbb{L}^0(\hat{\Lambda}, \mathbb{R})$. When there is no confusion, we shall omit the variable ω in u and write it as $u(t, x)$. Given $u \in \mathbb{L}^0(\hat{\Lambda})$, we define its derivatives in x in the standard way, and for any fixed x , its path-derivatives in (t, ω) in the sense of Definition 2.1.

Notice that \mathcal{Q} is not compact. For any $N \geq 1$, denote

$$K_N := \left\{ x \in \mathcal{O} : |x| \leq N, d(x, \mathcal{O}^c) \geq \frac{1}{N} \right\} \quad \text{and} \quad \mathcal{Q}_N := [0, N] \times K_N. \tag{2.11}$$

It is clear that K_N is compact, increasing in N , and

$$\begin{aligned} \bigcup_{N=1}^{\infty} K_N &= \mathcal{O}, \quad \bigcup_{N=1}^{\infty} \mathcal{Q}_N = \mathcal{Q}, \quad \text{and} \\ x + h \in K_{N+1} \quad &\text{for any } x \in K_N, |h| \leq \frac{1}{N(N+1)}. \end{aligned} \tag{2.12}$$

Similar to the process case we can define the norms: for $p \geq 1$ and $n \geq 2$,

$$\begin{aligned} \|u\|_{0,p,N} &:= \|u\|_{p,N} := \sup_{t \in [0, N]} \mathbb{E} \left[\sup_{x \in K_N} |u(t, x)|^p \right]^{1/p}, \\ \|u\|_{1,p,N} &:= \|u\|_{p,N} + \sum_{i=1}^{d'} \|\partial_{x_i} u\|_{0,p,N} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{0,p,N}, \end{aligned}$$

$$\begin{aligned} \|u\|_{n,p,N} &:= \|u\|_{p,N} + \|\partial_t u\|_{n-2,p,N} + \sum_{i=1}^{d'} \|\partial_{x_i} u\|_{n-1,p,N} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{n-1,p,N}, \\ \|u\|_{0,p,\alpha,N} &:= \|u\|_{p,\alpha,N} := \|u\|_{p,N} + \sup_{x \in K_N} \mathbb{E} \left[\sup_{0 \leq t < t' \leq N} \frac{|u(t,x) - u(t',x)|^p}{|t - t'|^{\frac{p\alpha}{2}}} \right]^{1/p} \\ &\quad + \sup_{0 \leq t \leq N} \mathbb{E} \left[\sup_{x,x' \in K_N} \frac{|u(t,x) - u(t,x')|^p}{|x - x'|^{p\alpha}} \right]^{1/p}, \quad \alpha \in (0, 1), \tag{2.13} \\ \|u\|_{1,p,\alpha,N} &:= \|u\|_{p,\alpha,N} + \sum_{i=1}^{d'} \|\partial_{x_i} u\|_{p,\alpha,N} + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{p,\alpha,N}, \\ \|u\|_{n,p,\alpha,N} &:= \|u\|_{p,\alpha,N} + \|\partial_t u\|_{n-2,p,\alpha,N} + \sum_{i=1}^{d'} \|\partial_{x_i} u\|_{n-1,p,\alpha,N} \\ &\quad + \sum_{i=1}^d \|\partial_{\omega^i} u\|_{n-1,p,\alpha,N}. \end{aligned}$$

We now define the following spaces: for $p \geq 1, 0 < \alpha < 1$, and $n \geq 0$,

$$\begin{aligned} \mathcal{H}_p^{[n]}(\hat{\Lambda}) &:= \left\{ u \in \mathbb{L}^0(\hat{\Lambda}) : \|u\|_{n,p,N} < \infty, \forall N > 0 \right\}, \\ \mathcal{H}_p^{[n]+\alpha}(\hat{\Lambda}) &:= \left\{ u \in \mathcal{H}_p^{[n]}(\hat{\Lambda}) : \|u\|_{n,p,\alpha,N} < \infty, \forall N > 0 \right\}. \end{aligned} \tag{2.14}$$

Again, as we shall see in [Lemma 6.1](#), one can show that

$$\mathcal{H}_p^{[n+2]}(\hat{\Lambda}) \subset \mathcal{H}_p^{[n]+\alpha}(\hat{\Lambda}) \quad \text{for any } n \geq 0, \alpha \in (0, 1) \text{ and } p > \frac{2}{1-\alpha}. \tag{2.15}$$

2.3. Itô random fields and Itô–Wentzell formula

We recall that $u \in \mathbb{L}^0(\hat{\Lambda})$ is called an Itô random field if, for any $x \in \mathcal{O}$,

$$u(t, x) = u_0(x) + \int_0^t \alpha(s, x) ds + \int_0^t \beta(s, x) \cdot dB_s, \quad t \geq 0, \mathbb{P}_0\text{-a.s.} \tag{2.16}$$

where $\alpha \in \mathbb{L}^0(\hat{\Lambda}), \beta \in \mathbb{L}^0(\hat{\Lambda}, \mathbb{R}^d)$ satisfy $\int_0^t [|\alpha(s, x)| + |\beta(s, x)|^2] ds < \infty, \mathbb{P}_0\text{-a.s.}$ for all $(t, x) \in \mathcal{Q}$.

It is worth noting that, in contrast to [Remark 2.4](#), the spatial derivative ∂_x commutes with both ∂_t and ∂_ω . In fact, we have the following result. Since the proof is quite straightforward, we omit it.

Lemma 2.6. *Let $u \in \mathbb{L}^0(\hat{\Lambda})$ be an Itô random field in the form of (2.16).*

- (i) *Assume u_0, α, β are differentiable in x , and for any $N > 0$, the processes $|\partial_x \alpha(\cdot, x, \cdot)|$ and $|\partial_x \beta(\cdot, x, \cdot)|^2$ are uniformly integrable on $[0, N] \times \Omega$, uniformly on $x \in K_N$. Then $\partial_x u$ exists and is also an Itô random field: for each $i = 1, \dots, d'$,*

$$\partial_{x_i} u(t, x) = \partial_{x_i} u_0(x) + \int_0^t \partial_{x_i} \alpha(s, x) ds + \int_0^t \partial_{x_i} \beta(s, x) \cdot dB_s, \quad t \geq 0, \mathbb{P}_0\text{-a.s.}$$

In particular, this implies that

$$\partial_\omega \partial_x u = \partial_x \partial_\omega u. \tag{2.17}$$

(ii) Assume further that β is an Itô random field and each of its components satisfies the property of u in (i), then

$$\partial_{\omega\omega}^2 \partial_x u = \partial_x \partial_{\omega\omega}^2 u \quad \text{and} \quad \partial_t \partial_x u = \partial_x \partial_t u. \tag{2.18}$$

As an important application of the path derivatives, we recast the Itô–Wentzell formula, which turns out to be exactly the same as a multidimensional functional Itô formula.

Proposition 2.7 (Itô–Wentzell Formula). *Let $X \in \mathcal{H}_2^{[2]}(\Lambda)$ taking values in \mathcal{O} and $u \in \mathcal{H}_2^{[2]}(\hat{\Lambda})$ such that $u(\cdot, \omega) \in C^{0,2}(Q)$ and $\partial_\omega u(\cdot, \omega) \in C^{0,1}(Q)$, for \mathbb{P}_0 -a.e. ω . Then the following chain rule for our path derivatives holds:*

$$\partial_t [u(t, X_t, \omega)] = \partial_t u + \partial_x u \cdot \partial_t X_t; \quad \partial_\omega [u(t, X_t, \omega)] = \partial_\omega u + [\partial_\omega X_t]^T \partial_x u. \tag{2.19}$$

In particular, if $dX_t = b_t dt + \sigma_t \cdot dB_t$, $du(t, x, \omega) = \alpha(t, x)dt + \beta(t, x) \cdot dB_t$, $t \geq 0$, then the Itô–Wentzell formula holds: for $t \geq 0$, \mathbb{P}_0 -a.s.,

$$\begin{aligned} du(t, X_t, \omega) &= [\partial_t u + \partial_x u \cdot \partial_t X_t]dt + [\partial_\omega u + [\partial_\omega X_t]^T \partial_x u] \circ dB_t \\ &= [\alpha + \partial_x u \cdot b_t + \frac{1}{2} \partial_{xx}^2 u : \sigma_t \sigma_t^T + \partial_x \beta : \sigma_t]dt \\ &\quad + [\beta + \sigma_t^T \partial_x u] \cdot dB_t. \end{aligned} \tag{2.20}$$

Proof. Since $X \in \mathcal{H}_2^{[2]}(\Lambda)$ and $u \in \mathcal{H}_2^{[2]}(\hat{\Lambda})$, one can write

$$dX_t = b_t dt + \sigma_t \cdot dB_t, \quad \text{and} \quad du(t, x, \omega) = \alpha(t, x)dt + \beta(t, x) \cdot dB_t,$$

where $b^i = \partial_t X^i + \frac{1}{2} \text{tr}(\partial_{\omega\omega}^2 X^i)$, $\sigma^i = \partial_\omega X^i$, $i = 1, \dots, d'$, and $\alpha = \partial_t u + \frac{1}{2} \text{tr}(\partial_{\omega\omega}^2 u)$, $\beta = \partial_\omega u$.

Next, under our conditions we may apply the standard Itô–Wentzell formula (see, for example, [17]) and obtain

$$\begin{aligned} du(t, X_t) &= \left[\alpha + \partial_x u \cdot b_t + \frac{1}{2} \partial_{xx}^2 u : \sigma_t \sigma_t^T + \partial_x \beta : \sigma_t \right](t, X_t)dt \\ &\quad + [\beta + \sigma_t^T \partial_x u](t, X_t) \cdot dB_t. \end{aligned}$$

Therefore the definition of path derivatives leads to that

$$\begin{cases} \partial_\omega [u(t, X_t)] = \beta + \sigma_t^T \partial_x u = (\partial_\omega u)(t, X_t) + [\partial_\omega X_t]^T [(\partial_x u)(t, X_t)]; \\ \partial_t [u(t, X_t, \omega)] = [\alpha + \partial_x u \cdot b_t + \frac{1}{2} \partial_{xx}^2 u : \sigma_t \sigma_t^T + \partial_x \beta : \sigma_t] \\ \quad - \frac{1}{2} \text{tr}(\partial_{\omega\omega} [u(t, X_t)]). \end{cases} \tag{2.21}$$

Note that $\partial_{\omega\omega}^2 u = \partial_\omega [\partial_\omega u]$, differentiating $\partial_\omega u(t, X_t)$ again we have

$$\partial_{\omega\omega}^2 [u(t, X_t)] = \partial_{\omega\omega} u + \partial_\omega X_t [\partial_{x\omega} u]^T + \sum_{i=1}^{d'} [\partial_{\omega\omega} X_t^i \partial_{x_i} u + \partial_\omega X_t^i [\partial_{x_i \omega} u + \partial_{x_i x} u \partial_\omega X^i]^T].$$

Now plugging this into (2.21) and recalling the definition of $\partial_\omega X$, $\partial_{\omega\omega}^2 X$, $\partial_\omega u$, $\partial_{\omega\omega}^2 u$, with some simple computation we prove (2.19), whence (2.20), immediately. ■

Remark 2.8. (i) If u is deterministic, then $\beta = \partial_\omega u = 0$, and we have the Itô formula.
 (ii) As the “chain rule” (2.19) completely characterizes the expression (2.20), we may refer to it as “pathwise Itô–Wentzell formula”. ■

2.4. Multiple differentiation and integration

Our Taylor expansion will involve multiple differentiation and integration. However, due to the noncommutative property of the path derivatives in Remark 2.4 and Example 2.5, we need to specify the differentiation and integration indices precisely. To simplify presentation, we first introduce some notations. For $i = 0, 1, \dots, d$, define

$$\begin{aligned} \partial_i u &:= \partial_t u, & u_t d_i t &:= u_t dt, & \text{if } i = 0; \\ \partial_i u &:= \partial_{\omega_i} u, & u_t d_i t &:= u_t \circ dB_t^i, & \text{if } 1 \leq i \leq d. \end{aligned} \tag{2.22}$$

Next, for $\theta = (\theta_1, \tilde{\theta}) = (\theta_1, \theta_2, \dots, \theta_n) \in \{0, 1, \dots, d\}^n$ and $s < t$, we define recursively by:

$$\mathcal{D}_\omega^\theta u := \partial_{\theta_1}(\mathcal{D}_\omega^{\tilde{\theta}} u), \quad \mathcal{I}_{s,t}^\theta(u) := \mathcal{I}_{s,t}^{\tilde{\theta}}\left(\int_s^t u_r d_{\theta_1} r\right), \quad \mathcal{I}_{s,t}^\theta := \mathcal{I}_{s,t}^\theta(1). \tag{2.23}$$

Notice that the above definition also implies, for $\theta = (\tilde{\theta}, \theta_n)$,

$$\begin{aligned} \mathcal{D}_\omega^\theta u &= \partial_{\theta_1} \cdots \partial_{\theta_n} u = \mathcal{D}_\omega^{\tilde{\theta}}(\partial_{\theta_n} u), \\ \mathcal{I}_{s,t}^\theta(u) &:= \int_s^t \int_s^{t_n} \cdots \int_s^{t_2} u_{t_1} d_{\theta_1} t_1 \cdots d_{\theta_n} t_n = \int_s^t \mathcal{I}_{s,r}^{\tilde{\theta}}(u) d_{\theta_n} r. \end{aligned} \tag{2.24}$$

Moreover, for the purpose of backward expansion later, we introduce

$$-\theta := (\theta_n, \dots, \theta_1) \quad \text{and} \quad \mathcal{I}_{t,s}^\theta(u) := (-1)^n \mathcal{I}_{s,t}^{-\theta}(u) \quad \text{for } s < t. \tag{2.25}$$

Noting the relation between the horizontal derivative $\partial_t u$ and $\partial_{\omega\omega}^2 u$ (cf. (2.3)), we introduce the following “weighted norm”: for $\theta \in \{0, 1, \dots, d\}^n$,

$$|\theta|_0 := n, \quad |\theta| := n + \sum_{i=1}^n 1_{\{\theta_i=0\}}. \tag{2.26}$$

Moreover, when $|\theta| = 0$, we take the notational convention that

$$\mathcal{D}_\omega^\theta u := u, \quad \mathcal{I}_{s,t}^\theta(u) := u_t. \tag{2.27}$$

Due to the commutative property of Lemma 2.6, the high order differentiation operator in x is simpler. Let \mathbb{N} be the set of nonnegative integers. For $\ell = (\ell_1, \dots, \ell_{d'}) \in \mathbb{N}^{d'}$, denote:

$$\mathcal{D}_x^\ell u := \partial_{x_1}^{\ell_1} \cdots \partial_{x_{d'}}^{\ell_{d'}} u, \quad x^\ell := \prod_{i=1}^{d'} x_i^{\ell_i}, \quad \ell! := \prod_{i=1}^{d'} \ell_i!, \quad |\ell| := \sum_{i=1}^{d'} \ell_i. \tag{2.28}$$

We shall set $\mathcal{D}_x^\ell u := u$, $x^\ell := 1$, and $\ell! := 1$, if $|\ell| = 0$.

Furthermore, together with (2.26), we can introduce a “weighted norm” on the index set $\Theta := \cup_{n=0}^\infty \{0, 1, \dots, d\}^n \times \mathbb{N}^{d'}$:

$$|(\theta, \ell)| := |\theta| + |\ell|, \quad \forall (\theta, \ell) \in \Theta. \tag{2.29}$$

Note that if we denote $\Theta_n := \{(\theta, \ell) \in \Theta : |(\theta, \ell)| \leq n\}$, then by applying Lemma 2.6 one can easily check that: if $u \in \mathcal{H}_2^{[n]}$, then all derivatives of u up to order n can be written as $\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u$ for some $(\theta, \ell) \in \Theta_n$ (counting “ ∂_t ” as a second order derivative!).

3. Taylor expansion for Itô processes (second order case)

In this section we give some heuristic arguments for the simplest second order Taylor expansion for Itô processes. We shall establish both forward and backward temporal expansions.

In what follows we shall always denote, for $s < t$, $\varphi_{s,t} := \varphi_t - \varphi_s$, and we will use the following simple fact frequently: for any semimartingales ξ, η, γ ,

$$\int \xi_t \circ (\eta_t \circ d\gamma_t) = \int (\xi_t \eta_t) \circ d\gamma_t = \int \eta_t \circ (\xi_t \circ d\gamma_t). \tag{3.1}$$

3.1. Forward temporal expansion

Let $t \geq 0, \delta > 0$, and denote $t_\delta := t + \delta$. Repeatedly applying the functional Itô formula (2.5) formally we have

$$\begin{aligned} u_{t_\delta} &= u_t + \int_t^{t_\delta} \partial_t u_s ds + \sum_{i=1}^d \int_t^{t_\delta} \partial_{\omega^i} u_s \circ dB_s^i \\ &= u_t + \sum_{i=1}^d \partial_{\omega^i} u_t B_{t,t_\delta}^i + \int_t^{t_\delta} \partial_t u_s ds + \sum_{i=1}^d \int_t^{t_\delta} [\partial_{\omega^i} u]_{t,s} \circ dB_s^i \\ &= u_t + \sum_{i=1}^d \partial_{\omega^i} u_t B_{t,t_\delta}^i + \int_t^{t_\delta} \partial_t u_s ds \\ &\quad + \sum_{i=1}^d \int_t^{t_\delta} \left[\int_t^s \partial_{t\omega^i} u_r dr + \sum_{j=1}^d \int_t^s \partial_{\omega^j \omega^i} u_r \circ dB_r^j \right] \circ dB_s^i \\ &= u_t + \sum_{i=1}^d \partial_{\omega^i} u_t B_{t,t_\delta}^i + \partial_t u_t \delta + \sum_{i,j=1}^d \partial_{\omega^j \omega^i} u_t \int_t^{t_\delta} B_{t,s}^j \circ dB_s^i \\ &\quad + \int_t^{t_\delta} [\partial_t u]_{t,s} ds + \sum_{i=1}^d \int_t^{t_\delta} \left[\int_t^s \partial_{t\omega^i} u_r dr \circ dB_s^i \right. \\ &\quad \left. + \sum_{i,j=1}^d \int_t^{t_\delta} \int_t^s [\partial_{\omega^j \omega^i} u]_{t,r} \circ dB_r^j \right] \circ dB_s^i. \end{aligned}$$

Here we used the fact that $\partial_{\omega^i} u_t$ and $\partial_{\omega^j \omega^i} u_t$ are \mathcal{F}_t -measurable and can be moved out from the related stochastic integrals. (We note that this will not be the case when we consider backward temporal expansion later.) Then

$$u_{t_\delta} = u_t + \sum_{i=1}^d \partial_{\omega^i} u_t B_{t,t_\delta}^i + \partial_t u_t \delta + \sum_{i,j=1}^d \partial_{\omega^j \omega^i} u_t \int_t^{t_\delta} B_{t,s}^j \circ dB_s^i + R_2(t, \delta), \tag{3.2}$$

where

$$R_2(t, \delta) := \int_t^{t_\delta} \int_t^s \partial_{tt} u_r dr ds + \sum_{i,j=1}^d \int_t^{t_\delta} \int_t^s \int_t^r \partial_{t\omega^j \omega^i} u_\kappa d\kappa \circ dB_r^j \circ dB_s^i$$

$$\begin{aligned}
 & + \sum_{i=1}^d \int_t^{t_\delta} \int_t^s \partial_{\omega^i} u_r \circ dB_r^i ds + \sum_{i=1}^d \int_t^{t_\delta} \int_t^s \partial_{t\omega^i} u_r dr \circ dB_s^i \\
 & + \sum_{i,j,k=1}^d \int_t^{t_\delta} \int_t^s \int_t^r \partial_{\omega^k \omega^j \omega^i} u_\kappa \circ dB_\kappa^k \circ dB_r^j \circ dB_s^i. \tag{3.3}
 \end{aligned}$$

To simplify the presentations let us make use of the notations for multiple derivatives and integrations defined in (2.22)–(2.24). Then it is straightforward to check that (3.2) and (3.3) can be rewritten as a more compact form:

$$u_{t_\delta} = u_t + \sum_{i=0}^d \mathcal{D}_\omega^{(i)} u_t \mathcal{I}_{t,t_\delta}^{(i)} + \sum_{i,j=1}^d \mathcal{D}_\omega^{(j,i)} u_t \mathcal{I}_{t,t_\delta}^{(j,i)} + R_2(t, \delta) \tag{3.4}$$

$$\begin{aligned}
 R_2(t, \delta) := & \mathcal{I}_{t,t_\delta}^{(0,0)} (\mathcal{D}_\omega^{(0,0)} u) + \sum_{i,j=1}^d \mathcal{I}_{t,t_\delta}^{(0,j,i)} (\mathcal{D}_\omega^{(0,j,i)} u) \\
 & + \sum_{i=1}^d \mathcal{I}_{t,t_\delta}^{(i,0)} (\mathcal{D}_\omega^{(i,0)} u) + \sum_{i=1}^d \mathcal{I}_{t,t_\delta}^{(0,i)} (\mathcal{D}_\omega^{(0,i)} u) + \sum_{i,j,k=1}^d \mathcal{I}_{t,t_\delta}^{(k,j,i)} (\mathcal{D}_\omega^{(k,j,i)} u). \tag{3.5}
 \end{aligned}$$

3.2. Backward temporal expansion

Let $0 < \delta \leq t$, and denote $t_\delta^- := t - \delta$. Then similar to the forward expansion we can obtain

$$\begin{aligned}
 u_{t_\delta^-} & = u_t - \int_{t_\delta^-}^t \partial_t u_s ds - \sum_{i=1}^d \int_{t_\delta^-}^t \partial_{\omega^i} u_s^i \circ dB_s^i \\
 & = u_t - \partial_t u_{t_\delta^-} \delta - \int_{t_\delta^-}^t [\partial_t u]_{t_\delta^-,s} ds - \sum_{i=1}^d \left[\partial_{\omega^i} u_{t_\delta^-} B_{t_\delta^-,t}^i + \int_{t_\delta^-}^t [\partial_{\omega^i} u]_{t_\delta^-,s} \circ dB_s^i \right]. \tag{3.6}
 \end{aligned}$$

We should note that the above expansion would be around t_δ^- instead of t , we therefore modify it as follows. First, we write

$$\partial_t u_{t_\delta^-} \delta = \partial_t u_t \delta - [\partial_t u]_{t_\delta^-,t} \delta, \quad \partial_{\omega^i} u_{t_\delta^-} B_{t_\delta^-,t}^i = \partial_{\omega^i} u_t B_{t_\delta^-,t}^i - [\partial_{\omega^i} u]_{t_\delta^-,t} B_{t_\delta^-,t}^i. \tag{3.7}$$

Next, we apply integration by parts formula and/or (standard) Itô formula to get

$$\begin{aligned}
 [\partial_t u]_{t_\delta^-,t} \delta - \int_{t_\delta^-}^t [\partial_t u]_{t_\delta^-,s} ds & = \int_{t_\delta^-}^t (s - t_\delta^-) d(\partial_t u_s) \\
 & = \int_{t_\delta^-}^t \partial_{tt} u_s (s - t_\delta^-) ds + \sum_{i=1}^d \int_{t_\delta^-}^t \partial_{\omega^i} u_s (s - t_\delta^-) \circ dB_s^i; \\
 [\partial_{\omega^i} u]_{t_\delta^-,t} B_{t_\delta^-,t}^i - \int_{t_\delta^-}^t [\partial_{\omega^i} u]_{t_\delta^-,s} \circ dB_s^i & \\
 = \int_{t_\delta^-}^t B_{t_\delta^-,s}^i \circ d(\partial_{\omega^i} u_s) & = \int_{t_\delta^-}^t B_{t_\delta^-,s}^i \circ \left(\partial_{t\omega^i} u_s ds + \sum_{j=1}^d \partial_{\omega^j \omega^i} u_s \circ dB_s^j \right)
 \end{aligned}$$

$$\begin{aligned}
 &= \int_{t_\delta^-}^t \partial_{t\omega^i} u_s B_{t_\delta^-,s}^i ds + \sum_{j=1}^d \int_{t_\delta^-}^t \partial_{\omega^j \omega^i} u_s B_{t_\delta^-,s}^i \circ dB_s^j \\
 &= \int_{t_\delta^-}^t \partial_{t\omega^i} u_s B_{t_\delta^-,s}^i ds + \sum_{j=1}^d \partial_{\omega^j \omega^i} u_t \int_{t_\delta^-}^t B_{t_\delta^-,s}^i \circ dB_s^j \\
 &\quad - \sum_{j=1}^d [\partial_{\omega^j \omega^i} u]_{t_\delta^-,t} \int_{t_\delta^-}^t B_{t_\delta^-,s}^i \circ dB_s^j + \sum_{j=1}^d \int_{t_\delta^-}^t [\partial_{\omega^j \omega^i} u]_{t_\delta^-,s} B_{t_\delta^-,s}^i \circ dB_s^j, \tag{3.8}
 \end{aligned}$$

and

$$\begin{aligned}
 &[\partial_{\omega^j \omega^i} u]_{t_\delta^-,t} \int_{t_\delta^-}^t B_{t_\delta^-,s}^i \circ dB_s^j - \int_{t_\delta^-}^t [\partial_{\omega^j \omega^i} u]_{t_\delta^-,s} B_{t_\delta^-,s}^i \circ dB_s^j \\
 &= \int_{t_\delta^-}^t \left(\int_{t_\delta^-}^s B_{t_\delta^-,r}^i \circ dB_r^j \right) \circ d(\partial_{\omega^j \omega^i} u_s) \\
 &= \int_{t_\delta^-}^t \partial_{t\omega^j \omega^i} u_s \left(\int_{t_\delta^-}^s B_{t_\delta^-,r}^i \circ dB_r^j \right) ds + \sum_{k=1}^d \int_{t_\delta^-}^t \partial_{\omega^k \omega^j \omega^i} u_s \left(\int_{t_\delta^-}^s B_{t_\delta^-,r}^i \circ dB_r^j \right) \circ dB_s^k. \tag{3.9}
 \end{aligned}$$

Plugging (3.9) into (3.8) and then plugging (3.7)–(3.8) into (3.6) we obtain

$$u_{t_\delta^-} = u_t - \sum_{i=1}^d \partial_{\omega^i} u_t B_{t_\delta^-,t}^i - \partial_t u_t \delta + \sum_{i,j=1}^d \partial_{\omega^j \omega^i} u_t \int_{t_\delta^-}^t B_{t_\delta^-,s}^i \circ dB_s^j + R_2(t, -\delta), \tag{3.10}$$

where

$$\begin{aligned}
 R_2(t, -\delta) &= \int_{t_\delta^-}^t \partial_{tt} u_s (s - t_\delta^-) ds - \sum_{i,j=1}^d \int_{t_\delta^-}^t \partial_{t\omega^j \omega^i} u_s \left(\int_{t_\delta^-}^s B_{t_\delta^-,r}^i \circ dB_r^j \right) ds \\
 &\quad + \sum_{i=1}^d \int_{t_\delta^-}^t \partial_{t\omega^i} u_s B_{t_\delta^-,s}^i ds + \sum_{i=1}^d \int_{t_\delta^-}^t \partial_{\omega^i} u_s (s - t_\delta^-) \circ dB_s^i \\
 &\quad - \sum_{i,j,k=1}^d \int_{t_\delta^-}^t \partial_{\omega^k \omega^j \omega^i} u_s \left(\int_{t_\delta^-}^s B_{t_\delta^-,r}^i \circ dB_r^j \right) \circ dB_s^k. \tag{3.11}
 \end{aligned}$$

Using the notations for multiple derivatives and integrations again, we see that (3.10) and (3.11) can again be written as the compact form:

$$u_{t_\delta^-} = u_t - \sum_{i=0}^d \mathcal{D}_\omega^{(i)} u_t \mathcal{I}_{t_\delta^-,t}^{(i)} + \sum_{i,j=1}^d \mathcal{D}_\omega^{(j,i)} u_t \mathcal{I}_{t_\delta^-,t}^{(i,j)} + R_2(t, -\delta); \tag{3.12}$$

where

$$\begin{aligned}
 R_2(t, -\delta) &:= \mathcal{I}_{t_\delta^-,t}^{(0)} (\mathcal{D}_\omega^{(0,0)} u, \mathcal{I}_{t_\delta^-,\cdot}^{(0)}) - \sum_{i,j=1}^d \mathcal{I}_{t_\delta^-,t}^{(0)} (\mathcal{D}_\omega^{(0,j,i)} u, \mathcal{I}_{t_\delta^-,\cdot}^{(i,j)}) \\
 &\quad + \sum_{i=1}^d \mathcal{I}_{t_\delta^-,t}^{(0)} (\mathcal{D}_\omega^{(0,i)} u, \mathcal{I}_{t_\delta^-,\cdot}^{(i)}) \\
 &\quad + \sum_{i=1}^d \mathcal{I}_{t_\delta^-,t}^{(i)} (\mathcal{D}_\omega^{(i,0)} u, \mathcal{I}_{t_\delta^-,\cdot}^{(0)}) - \sum_{i,j,k=1}^d \mathcal{I}_{t_\delta^-,t}^{(k)} (\mathcal{D}_\omega^{(k,j,i)} u, \mathcal{I}_{t_\delta^-,\cdot}^{(i,j)}). \tag{3.13}
 \end{aligned}$$

We should point out here that (3.12) is slightly different from (3.4). But by applying the relation (2.25) we can rewrite (3.12) as

$$u_{t_\delta^-} = u_t + \sum_{i=0}^d \mathcal{D}_\omega^{(i)} u_t \mathcal{T}_{t,t_\delta^-}^{(i)} + \sum_{i,j=1}^d \mathcal{D}_\omega^{(j,i)} u_t \mathcal{T}_{t,t_\delta^-}^{(j,i)} + R_2(t, -\delta). \tag{3.14}$$

We see that (3.14) is indeed consistent with the forward expansion (3.4)(!).

Remark 3.1. (i) If we define, for $s < t$,

$$\underline{B}_{s,t} := \left[\int_s^t B_{s,r}^i \circ dB_r^j \right]_{1 \leq i,j \leq d}; \quad A_{s,t} := \underline{B}_{s,t} - (\underline{B}_{s,t})^T, \tag{3.15}$$

then we can write

$$\begin{cases} u_{t_\delta} = u_t + \partial_t u_t \delta + \partial_\omega u_t \cdot B_{t,t_\delta} + \partial_{\omega\omega}^2 u_t : \underline{B}_{t,t_\delta} + R_2(t, \delta); \\ u_{t_\delta^-} = u_t - \partial_t u_t \delta - \partial_\omega u_t \cdot B_{t_\delta^-,t} + \partial_{\omega\omega}^2 u_t : \underline{B}_{t_\delta^-,t} + R_2(t, -\delta). \end{cases} \tag{3.16}$$

It is worth noting that $\underline{B}_{s,t}$ and $A_{s,t}$ are essentially the “Step-2 signature” and the “Lévy area”, respectively, in Rough Path theory (see, for example, [8]).

(ii) Note that

$$\begin{aligned} \int_s^t B_{s,r}^i \circ dB_r^j + \int_s^t B_{s,r}^j \circ dB_r^i &= B_{s,t}^i B_{s,t}^j, \\ \text{or equivalently, } \underline{B}_{s,t} + (\underline{B}_{s,t})^T &= B_{s,t} B_{s,t}^T. \end{aligned}$$

Then (3.16) becomes

$$\begin{aligned} u_{t_\delta} &= u_t + \partial_t u_t \delta + \partial_\omega u_t \cdot B_{t,t_\delta} + \frac{1}{2} \partial_{\omega\omega}^2 u_t : B_{t,t_\delta} B_{t,t_\delta}^T \\ &\quad + \frac{1}{2} \partial_{\omega\omega}^2 u_t : A_{t,t_\delta} + R_2(t, \delta); \\ u_{t_\delta^-} &= u_t - \partial_t u_t \delta - \partial_\omega u_t \cdot B_{t_\delta^-,t} + \frac{1}{2} \partial_{\omega\omega}^2 u_t : B_{t_\delta^-,t} B_{t_\delta^-,t}^T \\ &\quad + \frac{1}{2} \partial_{\omega\omega}^2 u_t : A_{t_\delta^-,t} + R_2(t, -\delta). \end{aligned} \tag{3.17}$$

Clearly, if $\partial_{\omega\omega}^2 u$ is symmetric, in particular when $u(t, \omega) = v(t, \omega_t)$ for some deterministic smooth function v , we have $\partial_{\omega\omega}^2 u_t : A_{t,t_\delta} = \partial_{\omega\omega}^2 u_t : A_{t_\delta^-,t} = 0$, and thus

$$\begin{aligned} u_{t_\delta} &= u_t + \partial_t u_t \delta + \partial_\omega u_t \cdot B_{t,t_\delta} + \frac{1}{2} \partial_{\omega\omega}^2 u_t : B_{t,t_\delta} B_{t,t_\delta}^T + R_2(t, \delta); \\ u_{t_\delta^-} &= u_t - \partial_t u_t \delta - \partial_\omega u_t \cdot B_{t_\delta^-,t} + \frac{1}{2} \partial_{\omega\omega}^2 u_t : B_{t_\delta^-,t} B_{t_\delta^-,t}^T + R_2(t, -\delta). \end{aligned} \tag{3.18}$$

This is exactly the standard Taylor expansion. We shall emphasize that in general $\partial_{\omega\omega}^2 u$ is not symmetric (see Example 2.5(ii)), which is reflected by the extra correction term $\frac{1}{2} \partial_{\omega\omega}^2 u : A$ in the Taylor expansion (3.17). ■

Remark 3.2. By Bichteler [1] or Karandikar [11], one may interpret $u_t dB_t^i$ in a pathwise manner, whenever u is continuous in t . In particular, $\underline{B}_{s,t}$ and $A_{s,t}$ can be understood pathwisely. ■

Remark 3.3. Assume that u has enough regularity and the path derivatives are continuous in (t, ω) , then $\partial_{\omega\omega}^2 u$ is symmetric as we pointed out in Remark 2.4(iii), and thus (3.18) holds true. Indeed, consider $d = 2$ and $t = 0$ for simplicity, and denote by $a_\delta := \int_0^\delta B_s^1 \circ dB_s^2 - \int_0^\delta B_s^2 \circ dB_s^1$ the $(1, 2)$ -component of the 2×2 -matrix $A_{0,\delta}$. By Theorem 4.1 we see that

$$\lim_{\delta \downarrow 0} \mathbb{P}_0(E_\delta) = 1, \quad \text{where } E_\delta := \{\omega : |R_2(0, \omega, \delta)| \leq \delta\}. \tag{3.19}$$

Note that for any $\omega, \bar{\omega} \in E_\delta$, we can easily deduce from the first line of (3.17) that

$$\begin{aligned} u_\delta(\bar{\omega}) - u_\delta(\omega) &= \partial_\omega u_0 \cdot [\bar{\omega}_\delta - \omega_\delta] + \frac{1}{2} \partial_{\omega\omega}^2 u_0 : [\bar{\omega}_\delta(\bar{\omega}_\delta)^T - \omega_\delta \omega_\delta^T] \\ &\quad + \frac{1}{2} [\partial_{\omega^1} \partial_{\omega^2} u_0 - \partial_{\omega^2} \partial_{\omega^1} u_0] [a_\delta(\bar{\omega}) - a_\delta(\omega)] \\ &\quad + [R_2(0, \bar{\omega}, \delta) - R_2(0, \omega, \delta)]. \end{aligned} \tag{3.20}$$

Here we emphasize that $\partial_\omega u_0$ and $\partial_{\omega\omega}^2 u_0$ are \mathcal{F}_0 -measurable and thus do not depend on ω . Now, let $\delta > 0$ be small enough so that $\mathbb{P}_0(E_\delta) > 0$. In light of [8, Exercise 2.17], one can construct $\omega^n, \omega \in E_\delta$, such that $\omega^n \rightarrow \omega$ but $|a_\delta(\omega^n) - a_\delta(\omega)| \geq 1$. Then, replacing $\bar{\omega}$ by ω^n in (3.20) we have

$$\begin{aligned} \frac{1}{2} \left| \partial_{\omega^1} \partial_{\omega^2} u_0 - \partial_{\omega^2} \partial_{\omega^1} u_0 \right| &\leq |u_\delta(\omega^n) - u_\delta(\omega)| + |\partial_\omega u_0 \cdot [\omega_\delta^n - \omega_\delta]| \\ &\quad + \frac{1}{2} |\partial_{\omega\omega}^2 u_0 : [\omega_\delta^n (\omega_\delta^n)^T - \omega_\delta \omega_\delta^T]| + 2\delta. \end{aligned}$$

Sending $n \rightarrow \infty$ we have $|\partial_{\omega^1} \partial_{\omega^2} u_0 - \partial_{\omega^2} \partial_{\omega^1} u_0| \leq 4\delta$. Since $\delta > 0$ is arbitrary, we see that $\partial_{\omega^1} \partial_{\omega^2} u_0 = \partial_{\omega^2} \partial_{\omega^1} u_0$. That is, $\partial_{\omega\omega}^2 u_0$ is symmetric. ■

As we pointed out in the Introduction, the main results of this paper are the (pathwise) remainder estimates. Since the proof of the second order estimate is similar to that of the inductional argument for the m -th order estimate, we shall prove a general result directly.

4. Taylor expansion for Itô processes (general case)

We now consider the general form of pathwise Taylor expansion up to any order m . Denote, for $0 \leq t_1 < t_2$ and $\varepsilon > 0$,

$$\begin{aligned} D &:= \left\{ (t, \delta) \in [0, \infty) \times \mathbb{R} \setminus \{0\} : t + \delta \geq 0 \right\}, \\ D_{[t_1, t_2]}^\varepsilon &:= \{(t, \delta) \in D : t_1 \leq t \leq t_2, |\delta| \leq \varepsilon\}. \end{aligned} \tag{4.1}$$

For any $m \geq 0$ and $u \in \mathcal{H}_2^{[m]}(\Lambda)$, in light of (3.4) and (3.14) we shall define the m -th order remainder by: for any $(t, \delta) \in D$ and $\omega \in \Omega$,

$$u(t + \delta, \omega) = \sum_{|\theta| \leq m} \mathcal{D}_\omega^\theta u(t, \omega) \mathcal{I}_{t, t+\delta}^\theta + R_m(u, t, \omega, \delta). \tag{4.2}$$

We emphasize that δ can be negative here, and the right side of (4.2) is pathwise, in light of Remark 3.2. Moreover, when there is no confusion, we shall always omit the variable ω .

The main result of this section is the following pathwise estimate for the remainder R_m .

Theorem 4.1. Assume that $u \in \mathcal{H}_{p_0}^{[m+2]}(\Lambda)$ for some $m \geq 0$ and $p_0 > 2$. Then for any $0 < \alpha < 1 - \frac{2}{p_0}$ and $p < p_0$, it holds that, for any $T > 0$,

$$\mathbb{E} \left\{ \sup_{(t, \delta) \in D_{[0, T]}^1} \left| \frac{R_m(u, t; \delta)}{|\delta|^{\frac{m+\alpha}{2}}} \right|^p \right\} < \infty. \tag{4.3}$$

To prove [Theorem 4.1](#) we need the following crucial estimate. Since its proof is quite lengthy, we shall complete its proof after we prove [Theorem 4.1](#).

Proposition 4.2. Assume that $u \in \mathcal{H}_{p_0}^{[m+2]}(\Lambda)$ for some $m \geq 0$ and $p_0 > 2$. Then for any $p < p_0, t_0 \geq 0$, and $\varepsilon > 0$, it holds that

$$\mathbb{E} \left[\sup_{(t, \delta) \in D_{[t_0, t_0 + \varepsilon]}^\varepsilon} |R_m(u, t; \delta)|^p \right] \leq C \varepsilon^{\frac{p(m+1)}{2}}, \tag{4.4}$$

where C may depend on $\|u\|_{m+2, p_0, T}$ for some $T \geq t_0 + 2\varepsilon$. ■

Proof of Theorem 4.1. In what follows we shall fix T , and allow the generic constant $C > 0$ to depend on $\|u\|_{m+2, p_0, T+1}$. Clearly it suffices to prove [\(4.3\)](#) for large p , and we thus assume without loss of generality that $\frac{2}{1-\alpha} < p < p_0$.

For any $0 < \varepsilon < 1$, set $t_i := i\varepsilon, i = 0, \dots, \lfloor \frac{T}{\varepsilon} \rfloor + 1$. Then, by [Proposition 4.2](#) we have

$$\mathbb{E} \left[\sup_{(t, \delta) \in D_{[0, T]}^\varepsilon} |R_m(u, t, \delta)|^p \right] \leq \sum_{i=0}^{\lfloor \frac{T}{\varepsilon} \rfloor} \mathbb{E} \left[\sup_{(t, \delta) \in D_{[t_i, t_i + \varepsilon]}^\varepsilon} |R_m(u, t, \delta)|^p \right] \leq C \varepsilon^{\frac{p(m+1)}{2} - 1}.$$

Consequently, since $0 < \alpha < 1 - \frac{2}{p_0}, \frac{2}{1-\alpha} < p < p_0$, it holds that

$$\begin{aligned} \mathbb{E} \left[\sup_{(t, \delta) \in D_{[0, T]}^1} \left| \frac{R_m(u, t, \delta)}{\delta^{\frac{m+\alpha}{2}}} \right|^p \right] &\leq \sum_{n=0}^{\infty} \mathbb{E} \left[\sup_{(t, \delta) \in D_{[0, T]}^1 : 2^{-(n+1)} \leq |\delta| \leq 2^{-n}} \left| \frac{R_m(u, t, \delta)}{\delta^{\frac{m+\alpha}{2}}} \right|^p \right] \\ &\leq \sum_{n=0}^{\infty} 2^{\frac{p(m+\alpha)(n+1)}{2}} \mathbb{E} \left[\sup_{(t, \delta) \in D_{[0, T]}^{2^{-n}}} |R_m(u, t, \delta)|^p \right] \\ &\leq C \sum_{n=0}^{\infty} 2^{\frac{p(m+\alpha)(n+1)}{2}} 2^{-n \left(\frac{p(m+1)}{2} - 1 \right)} \\ &= C 2^{\frac{p(m+\alpha)}{2}} \sum_{n=0}^{\infty} 2^{-\frac{n}{2} [p(1-\alpha) - 2]} < \infty, \end{aligned}$$

completing the proof. ■

Remark 4.3. (i) The estimate [\(4.3\)](#) amounts to saying that, (for each $m \geq 0, T > 0$, there exist a set $\Omega_m \subseteq \Omega$ with $\mathbb{P}(\Omega_m) = 1$ and a nonnegative random variable $C_{m, T}$, such that

$$|R_m(u, t, \omega; \delta)| \leq C_{m, T}(\omega) |\delta|^{\frac{m+\alpha}{2}}, \quad \forall \omega \in \Omega_m, \forall (t, \delta) \in D_{[0, T]}^1.$$

In [Theorem 5.2](#), this pathwise estimate also holds true locally uniformly in the spatial variable x .

- (ii) We should point out that in (4.2) $\delta < 0$ is allowed. That is, the temporal expansion can be “backward”. Such an expansion, along with the pathwise estimates, is crucial for the study of viscosity solutions of SPDEs in [4]. In our previous works [3,2] these results were also obtained in the case $m = 2, d = 1$, but the present treatment is more direct and the conditions are easier to verify. (See Section 8 for a detailed comparison with the result in [2].)
- (iii) There have been many works on stochastic Taylor expansions (see, e.g., the books [9,12], and the references cited therein). We also note that in a recent work [14], the Dupire-type path-derivatives were also used. The main difference between the existing results and ours, however, lies in that in these works the supremum is taken outside of the expectation (and hence they do not yield pathwise estimates), which is not desirable for our study of viscosity solutions. Moreover, no backward expansion was considered in these works. ■

In the rest of this section we prove the key estimate (4.4) in Proposition 4.2. To simplify the presentation, we split the proof into several lemmas that are interesting in their own rights. We begin by establishing a representation formula for R_m , extending (3.5) and (3.13). In what follows we denote $t_\delta := t + \delta$ and $t_\delta^- := t - \delta$, for $\delta > 0$.

Lemma 4.4. *Let $u \in \mathcal{H}_2^{[m+2]}(\Lambda)$ for some $m \geq 0$. Then for any $\delta > 0$, it holds that:*

$$R_m(u, t, \delta) = \sum_{|\theta|=m+1} \mathcal{I}_{t,t_\delta}^\theta(\mathcal{D}_\omega^\theta u) + \sum_{|\theta|=m} \mathcal{I}_{t,t_\delta}^\theta \left(\int_t^{\cdot} \partial_t \mathcal{D}_\omega^\theta u_s ds \right). \tag{4.5}$$

Furthermore, denoting $\theta = (\theta_1, \tilde{\theta})$, then for $t \geq \delta$ one has

$$\begin{aligned} R_m(u, t, -\delta) &= \sum_{|\theta|=m+1} (-1)^{|\theta|_0} \int_{t_\delta^-}^t \left(\mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-,s}^{-\tilde{\theta}} \right) d\theta_1 s \\ &\quad - \sum_{|\theta|=m} (-1)^{|\theta|_0} \int_{t_\delta^-}^t \left[\partial_t \mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-,s}^{-\theta} \right] ds. \end{aligned} \tag{4.6}$$

Proof. (i) We first verify (4.5) by induction. For $m = 0$ we recall the notational convention (2.27). Then it is readily seen that the right side of (4.5) reads

$$\sum_{|\theta|=1} \int_t^{t_\delta} \mathcal{D}_\omega^\theta u_s d\theta s + \int_t^{t_\delta} \partial_t u_s ds = \sum_{i=1}^d \int_t^{t_\delta} \partial_{\omega^i} u_s \circ dB_s^i + \int_t^{t_\delta} \partial_t u_s ds.$$

Thus the equality follows immediately from the functional Itô formula (2.5).

Now assume (4.5) holds for m . Then

$$\begin{aligned} R_{m+1}(u, t, \delta) &= R_m(u, t, \delta) - \sum_{|\theta|=m+1} \mathcal{D}_\omega^\theta u_t \mathcal{I}_{t,t_\delta}^\theta \\ &= \sum_{|\theta|=m+1} \mathcal{I}_{t,t_\delta}^\theta(\mathcal{D}_\omega^\theta u) + \sum_{|\theta|=m} \mathcal{I}_{t,t_\delta}^\theta \left(\int_t^{\cdot} \partial_t \mathcal{D}_\omega^\theta u_s ds \right) - \sum_{|\theta|=m+1} \mathcal{D}_\omega^\theta u_t \mathcal{I}_{t,t_\delta}^\theta \\ &= \sum_{|\theta|=m+1} \mathcal{I}_{t,t_\delta}^\theta([\mathcal{D}_\omega^\theta u]_{t,\cdot}) + \sum_{|\theta|=m} \mathcal{I}_{t,t_\delta}^\theta \left(\int_t^{\cdot} \partial_t \mathcal{D}_\omega^\theta u_s ds \right). \end{aligned}$$

Applying the functional Itô formula (2.5) on $\mathcal{D}_\omega^\theta u_s$ we obtain

$$\begin{aligned} R_{m+1}(u, t, \delta) &= \sum_{|\theta|=m+1} \mathcal{I}_{t, t_\delta}^\theta \left(\int_t^\cdot \partial_t \mathcal{D}_\omega^\theta u_s ds + \sum_{i=1}^d \int_t^\cdot \partial_{\omega^i} \mathcal{D}_\omega^\theta u_s \circ dB_s^i \right) \\ &\quad + \sum_{|\theta|=m} \mathcal{I}_{t, t_\delta}^\theta \left(\int_t^\cdot \partial_t \mathcal{D}_\omega^\theta u_s ds \right) \\ &= \sum_{|\theta|=m+1} \mathcal{I}_{t, t_\delta}^\theta \left(\int_t^\cdot \partial_t \mathcal{D}_\omega^\theta u_s ds \right) \\ &\quad + \sum_{|\theta|=m+1} \sum_{i=1}^d \mathcal{I}_{t, t_\delta}^\theta \left(\int_t^\cdot \partial_{\omega^i} \mathcal{D}_\omega^\theta u_s \circ dB_s^i \right) + \sum_{|\theta|=m} \mathcal{I}_{t, t_\delta}^\theta \left(\int_t^\cdot \partial_t \mathcal{D}_\omega^\theta u_s ds \right). \end{aligned}$$

One may check directly that the last line above is exactly equal to $\sum_{|\theta|=m+2} \mathcal{I}_{t, t_\delta}^\theta (\mathcal{D}_\omega^\theta u)$. Namely (4.5) holds for $m + 1$. Thus (4.5) holds for all m .

(ii) We now prove (4.6), again by induction. For $m = 0$ the argument is similar to (i). Assume now (4.6) holds for m . Then

$$\begin{aligned} R_{m+1}(u, t, -\delta) &= R_m(u, t, -\delta) - \sum_{|\theta|=m+1} (-1)^{|\theta|_0} \mathcal{D}_\omega^\theta u_t \mathcal{I}_{t_\delta^-, t}^{-\theta} \\ &= \sum_{|\theta|=m+1} (-1)^{|\theta|_0} \int_{t_\delta^-}^t [\mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}}] d\theta_1 s \\ &\quad - \sum_{|\theta|=m} (-1)^{|\theta|_0} \int_{t_\delta^-}^t [\partial_t \mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-, s}^{-\theta}] ds \\ &\quad - \sum_{|\theta|=m+1} (-1)^{|\theta|_0} \mathcal{D}_\omega^\theta u_t \int_{t_\delta^-}^t \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}} d\theta_1 s. \end{aligned}$$

Applying integration by parts formula we have

$$\begin{aligned} \mathcal{D}_\omega^\theta u_t \int_{t_\delta^-}^t \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}} d\theta_1 s - \int_{t_\delta^-}^t [\mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}}] d\theta_1 s &= \int_{t_\delta^-}^t \mathcal{I}_{t_\delta^-, s}^{-\theta} \circ d(\mathcal{D}_\omega^\theta u_s) \\ &= \int_{t_\delta^-}^t \mathcal{I}_{t_\delta^-, s}^{-\theta} \left[\partial_t \mathcal{D}_\omega^\theta u_s ds + \sum_{i=1}^d \partial_{\omega^i} \mathcal{D}_\omega^\theta u_s \circ dB_s^i \right]. \end{aligned}$$

Consequently we obtain:

$$\begin{aligned} R_{m+1}(u, t, -\delta) &= - \sum_{|\theta|=m+1} (-1)^{|\theta|_0} \int_{t_\delta^-}^t [\partial_t \mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-, s}^{-\theta}] ds \\ &\quad - \sum_{|\theta|=m+1} (-1)^{|\theta|_0} \int_{t_\delta^-}^t \mathcal{I}_{t_\delta^-, s}^{-\theta} \partial_\omega \mathcal{D}_\omega^\theta u_s \circ dB_s \\ &\quad - \sum_{|\theta|=m} (-1)^{|\theta|_0} \int_{t_\delta^-}^t [\partial_t \mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-, s}^{-\theta}] ds. \end{aligned}$$

One may now check directly that the last line above is exactly equal to, denoting $\theta = (\theta_1, \tilde{\theta})$, $\sum_{|\theta|=m+2} (-1)^{|\theta|_0} \int_{t_\delta}^t \left[\mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta, s}^{-\tilde{\theta}} \right] d\theta_1 s$. Thus (4.6) holds for $m + 1$, proving (ii), whence the lemma. ■

To simplify notation, in what follows we denote, for any semi-martingale φ , and any $p \geq 1, 0 \leq t_1 < t_2$,

$$I(\varphi, p, t_1, t_2) := \left(\mathbb{E} \left[\left(\int_{t_1}^{t_2} |\varphi_s|^2 ds \right)^{\frac{p}{2}} + \left(\int_{t_1}^{t_2} |\partial_\omega \varphi_s| ds \right)^p \right] \right)^{\frac{1}{p}}. \tag{4.7}$$

It is clear that I is increasing in p , and

$$I^p(\varphi, p, t_1, t_2) \leq \sup_{t_1 \leq t \leq t_2} \mathbb{E}[|\varphi_t|^p](t_2 - t_1)^{\frac{p}{2}} + \sup_{t_1 \leq t \leq t_2} \mathbb{E}[|\partial_\omega \varphi_t|^p](t_2 - t_1)^p. \tag{4.8}$$

In light of the above representations, the following estimate is crucial.

Lemma 4.5. *Let $t_0 \geq 0, 0 < \varepsilon < 1, q > p \geq 1$, and φ be a semimartingale. Then, for any $|\theta| \geq 1$, there exists constant $C = C_{p,q,|\theta|_0}$ such that*

$$\mathbb{E} \left\{ \sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \leq t \leq t_0 + \varepsilon} |\mathcal{I}_{t, t_\delta}^\theta(\varphi)|^p \right\} \leq C \varepsilon^{\frac{p(|\theta|-1)}{2}} I^p(\varphi, q, t_0, t_0 + 2\varepsilon). \tag{4.9}$$

Proof. Let $a(\varphi, \theta, p)$ denote the left side of (4.9), and $I(\varphi, p) := I(\varphi, p, t_0, t_0 + 2\varepsilon)$. Without loss of generality, we may assume $I(\varphi, q) < \infty$. We proceed by induction on $n := |\theta|_0$.

(i) First assume $n = 1$, namely $\theta = (\theta_1)$. We estimate $a(\varphi, \theta, p)$ in two cases.

Case 1. $|\theta| = 2$, namely $\theta_1 = 0$. Then $\mathcal{I}_{t, t_\delta}^\theta(\varphi) = \int_t^{t_\delta} \varphi_s ds$, and thus

$$a(\varphi, \theta, p) \leq \mathbb{E} \left[\left(\int_{t_0}^{t_0 + 2\varepsilon} |\varphi_s| ds \right)^p \right] \leq C \varepsilon^{\frac{p}{2}} I(\varphi, p)^p \leq C \varepsilon^{\frac{p}{2}} I(\varphi, q)^p.$$

Case 2. $|\theta| = 1$, namely $\theta_1 = i$ for some $i = 1, \dots, d$. Then

$$\mathcal{I}_{t, t_\delta}^\theta(\varphi) = \left[\int_{t_0}^{t_\delta} - \int_{t_0}^t \right] \varphi(s) dB_s^i + \frac{1}{2} \int_t^{t_\delta} \partial_{\omega^i} \varphi_s ds,$$

and thus

$$a(\varphi, \theta, p) \leq C \mathbb{E} \left[\sup_{t_0 \leq t \leq t_0 + 2\varepsilon} \left| \int_{t_0}^t \varphi_s dB_s^i \right|^p \right] + C \mathbb{E} \left[\left(\int_{t_0}^{t_0 + 2\varepsilon} |\partial_{\omega^i} \varphi_s| ds \right)^p \right],$$

which, together with the Burkholder–Davis–Gundy inequality, implies (4.9) immediately.

(ii) We next prove (4.9) by induction. Assume it holds true for n and we now assume $|\theta|_0 = n + 1$. Denote $\theta = (\theta_1, \tilde{\theta})$, $\psi_s := \int_{t_0}^s \varphi_r d\theta_1 r$, and $\tilde{q} := \frac{p+q}{2}$. Notice that $|\tilde{\theta}|_0 = n$ and $I(1, p) = \sqrt{2\varepsilon}$ for any p , then we may use the induction assumption and obtain

$$\begin{aligned} a(\varphi, \theta, p) &= \mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \leq t \leq t_0 + \varepsilon} |\mathcal{I}_{t, t_\delta}^{\tilde{\theta}}(\psi) - \psi_t \mathcal{I}_{t, t_\delta}^{\tilde{\theta}}(1)|^p \right] \\ &\leq C a(\psi, \tilde{\theta}, p) + C \left(a(\varphi, (\theta_1), \tilde{q}) \right)^{\frac{p}{\tilde{q}}} \left(a \left(1, \tilde{\theta}, \frac{2p\tilde{q}}{q-p} \right) \right)^{\frac{q-p}{2\tilde{q}}} \end{aligned}$$

$$\begin{aligned} &\leq C\varepsilon^{\frac{p(|\tilde{\theta}|-1)}{2}} (I(\psi, \tilde{q}))^p + C\varepsilon^{\frac{p(|\theta_1|-1)}{2}} (I(\varphi, q))^p \varepsilon^{\frac{p(|\tilde{\theta}|-1)}{2}} \left(I\left(1, \frac{3p\tilde{q}}{q-p}\right) \right)^p \\ &= C\varepsilon^{\frac{p(|\tilde{\theta}|-1)}{2}} (I(\psi, \tilde{q}))^p + C\varepsilon^{\frac{p(|\theta_1|-1)}{2}} (I(\varphi, q))^p. \end{aligned}$$

Then clearly (4.9) with $|\theta|_0 = n + 1$ follows from the following claim:

$$I(\psi, \tilde{q}) \leq C\varepsilon^{\frac{|\theta_1|}{2}} I(\varphi, q). \tag{4.10}$$

We again proceed in two cases.

Case 1. $|\theta_1| = 2$, namely $\theta_1 = 0$. Then $\psi_s = \int_{t_0}^s \varphi_r dr$ and $\partial_{\omega_j} \psi_s = 0$. Thus,

$$I(\psi, \tilde{q})^{\tilde{q}} \leq \mathbb{E} \left[\left(\int_{t_0}^{t_0+2\varepsilon} \left| \int_{t_0}^s \varphi_r dr \right|^2 ds \right)^{\frac{\tilde{q}}{2}} \right] \leq \varepsilon^{\frac{\tilde{q}}{2}} \mathbb{E} \left[\left(\int_{t_0}^{t_0+2\varepsilon} |\varphi_r| dr \right)^{\tilde{q}} \right] \leq \varepsilon^{\tilde{q}} I(\varphi, q)^{\tilde{q}}.$$

Case 2. $|\theta_1| = 1$, namely $\theta_1 = i$ for some $i = 1, \dots, d$. Then

$$\psi_s = \int_{t_0}^s \varphi_r \circ dB_r^i = \int_{t_0}^s \varphi_r dB_r^i + \frac{1}{2} \int_{t_0}^s \partial_{\omega^i} \varphi_r dr, \quad \partial_{\omega^j} \psi_s = \varphi_s 1_{\{j=i\}},$$

and thus

$$\begin{aligned} I(\psi, \tilde{q})^{\tilde{q}} &\leq C\mathbb{E} \left[\left(\int_{t_0}^{t_0+2\varepsilon} \left[\left| \int_{t_0}^s \varphi_r dB_r^i \right|^2 + \left| \int_{t_0}^s \partial_{\omega^i} \varphi_r dr \right|^2 \right] ds \right)^{\frac{\tilde{q}}{2}} + \left(\int_{t_0}^{t_0+2\varepsilon} |\varphi_s| ds \right)^{\tilde{q}} \right] \\ &\leq C\varepsilon^{\frac{\tilde{q}}{2}} \mathbb{E} \left[\sup_{t_0 \leq s \leq t_0+2\varepsilon} \left[\left| \int_{t_0}^s \varphi_r dB_r^i \right|^{\tilde{q}} \right] + \left[\int_{t_0}^{t_0+2\varepsilon} |\partial_{\omega^i} \varphi_r| dr \right]^{\tilde{q}} + \left(\int_{t_0}^{t_0+2\varepsilon} |\varphi_s|^2 ds \right)^{\frac{\tilde{q}}{2}} \right]. \end{aligned}$$

Then (4.10) follows immediately from the Burkholder–Davis–Gundy inequality. ■

Proof of Proposition 4.2. First, by (4.5) we have, for $\delta > 0$,

$$R_m(u, t, \delta) = \sum_{|\theta|=m+1} \mathcal{I}_{t, t_\delta}^\theta (\mathcal{D}_\omega^\theta u) + \sum_{|\theta|=m} \left[\mathcal{I}_{t, t_\delta}^\theta \left(\int_{t_0}^{\cdot} \partial_t \mathcal{D}_\omega^\theta u_s ds \right) - \mathcal{I}_{t, t_\delta}^\theta (1) \int_{t_0}^t \partial_t \mathcal{D}_\omega^\theta u_s ds \right].$$

Denote $D_+^\varepsilon := \{(t, \delta) : 0 < \delta \leq \varepsilon, t_0 \leq t \leq t_0 + \varepsilon\}$ and note that $\partial_\omega \int_{t_0}^t \partial_t \mathcal{D}_\omega^\theta u_s ds = 0$. Then, combining Lemma 4.5 and (4.8) and recalling $T \geq t_0 + 2\varepsilon$, we have

$$\begin{aligned} &\mathbb{E} \left[\sup_{(t, \delta) \in D_+^\varepsilon} |R_m(u, t, \delta)|^p \right] \\ &\leq C \sum_{|\theta|=m+1} \mathbb{E} \left[\sup_{(t, \delta) \in D_+^\varepsilon} |\mathcal{I}_{t, t_\delta}^\theta (\mathcal{D}_\omega^\theta u)|^p \right] \\ &\quad + C \sum_{|\theta|=m} \mathbb{E} \left[\sup_{(t, \delta) \in D_+^\varepsilon} \left| \mathcal{I}_{t, t_\delta}^\theta \left(\int_{t_0}^{\cdot} \partial_t \mathcal{D}_\omega^\theta u_s ds \right) \right|^p \right] \\ &\quad + C \sum_{|\theta|=m} \left(\mathbb{E} \left[\left(\int_{t_0}^{t_0+2\varepsilon} |\partial_t \mathcal{D}_\omega^\theta u_s| ds \right)^{p_0} \right] \right)^{\frac{p}{p_0}} \left(\mathbb{E} \left[\left(\sup_{(t, \delta) \in D_+^\varepsilon} |\mathcal{I}_{t, t_\delta}^\theta (1)| \right)^{\frac{pp_0}{p_0-p}} \right] \right)^{\frac{p_0-p}{p_0}} \\ &\leq C \sum_{|\theta|=m+1} \left[\|\mathcal{D}_\omega^\theta u\|_{p, T}^p \varepsilon^{\frac{p(m+1)}{2}} + \|\partial_\omega \mathcal{D}_\omega^\theta u\|_{p, T}^p \varepsilon^{\frac{p(m+2)}{2}} \right] \end{aligned}$$

$$\begin{aligned}
 & + C \varepsilon^{\frac{pm}{2}} \sum_{|\theta|=m} \sup_{t_0 \leq t \leq t_0 + 2\varepsilon} \mathbb{E} \left[\left| \int_{t_0}^t \partial_t \mathcal{D}_\omega^\theta u_s ds \right|^p \right] + C \sum_{|\theta|=m} \left(\|\partial_t \mathcal{D}_\omega^\theta u\|_{p_0, T}^p \varepsilon^p \right) \varepsilon^{\frac{pm}{2}} \\
 \leq & C \sum_{|\theta|=m+1} \left[\|\mathcal{D}_\omega^\theta u\|_{p_0, T}^p \varepsilon^{\frac{p(m+1)}{2}} + \|\partial_t \mathcal{D}_\omega^\theta u\|_{p_0, T}^p \varepsilon^{\frac{p(m+2)}{2}} \right] \\
 & + C \sum_{|\theta|=m} \left[\|\partial_t \mathcal{D}_\omega^\theta u\|_{p_0, T}^p \varepsilon^{\frac{p(m+2)}{2}} + \|\partial_t \mathcal{D}_\omega^\theta u\|_{p_0, T}^p \varepsilon^{\frac{p(m+1)}{2}} \right] \\
 \leq & C \|u\|_{m+2, p_0, T} \varepsilon^{\frac{p(m+1)}{2}}. \tag{4.11}
 \end{aligned}$$

Next, for $\theta = (\theta_1, \tilde{\theta}) = (\theta_1, \theta_2, \dots, \theta_n)$, applying the integration by parts formula and recalling (2.24) and (2.25), we have, for $\varphi := \mathcal{D}_\omega^\theta u$,

$$\begin{aligned}
 \int_{t_\delta^-}^t (\varphi_s \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}}) d\theta_1 s & = \left(\int_{t_\delta^-}^t \varphi_s d\theta_1 s \right) \mathcal{I}_{t_\delta^-, t}^{-\tilde{\theta}} - \int_{t_\delta^-}^t \left(\int_{t_\delta^-}^s \varphi_r d\theta_1 r \right) \circ d\mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}} \\
 & = \mathcal{I}_{t_\delta^-, t}^{(\theta_1)}(\varphi) \mathcal{I}_{t_\delta^-, t}^{(\theta_2, \dots, \theta_n)}(1) - \int_{t_\delta^-}^t \left(\mathcal{I}_{t_\delta^-, s}^{(\theta_1)}(\varphi) \mathcal{I}_{t_\delta^-, s}^{(\theta_2, \dots, \theta_n)} \right) d\theta_2 s.
 \end{aligned}$$

Repeating the above arguments we obtain

$$\int_{t_\delta^-}^t (\varphi_s \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}}) d\theta_1 s = \sum_{i=1}^n (-1)^{i-1} \mathcal{I}_{t_\delta^-, t}^{(\theta_1, \dots, \theta_i)}(\varphi) \mathcal{I}_{t_\delta^-, t}^{(\theta_{i+1}, \dots, \theta_n)}(1). \tag{4.12}$$

Then, by changing variable $t - \delta$ to t and denoting $q := \frac{p+p_0}{2}$, we have

$$\begin{aligned}
 & \mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \vee \delta \leq t \leq t_0 + \varepsilon} \left| \int_{t_\delta^-}^t (\mathcal{D}_\omega^\theta u_s \mathcal{I}_{t_\delta^-, s}^{-\tilde{\theta}}) d\theta_1 s \right|^p \right] \\
 \leq & C \sum_{i=1}^n \mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{(t_0 - \varepsilon)^+ \leq t \leq t_0 + \varepsilon} |\mathcal{I}_{t, t_\delta}^{(\theta_1, \dots, \theta_i)}(\mathcal{D}_\omega^\theta u) \mathcal{I}_{t, t_\delta}^{(\theta_{i+1}, \dots, \theta_n)}(1)|^p \right] \\
 \leq & C \sum_{i=1}^n \left(\mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{(t_0 - \varepsilon)^+ \leq t \leq t_0 + \varepsilon} |\mathcal{I}_{t, t_\delta}^{(\theta_1, \dots, \theta_i)}(\mathcal{D}_\omega^\theta u)|^q \right] \right)^{\frac{p}{q}} \\
 & \times \left(\mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{(t_0 - \varepsilon)^+ \leq t \leq t_0 + \varepsilon} |\mathcal{I}_{t, t_\delta}^{(\theta_{i+1}, \dots, \theta_n)}(1)|^{\frac{2pq}{p_0 - p}} \right] \right)^{\frac{p_0 - p}{2q}} \\
 \leq & C \sum_{i=1}^n \varepsilon^{\frac{p(|(\theta_1, \dots, \theta_i)| - 1)}{2}} I^p(\mathcal{D}_\omega^\theta u, p_0, (t_0 - \varepsilon)^+, t_0 + 2\varepsilon) \varepsilon^{\frac{p(|(\theta_{i+1}, \dots, \theta_n)|)}{2}} \\
 = & C \varepsilon^{\frac{p(|\theta| - 1)}{2}} I^p(\mathcal{D}_\omega^\theta u, p_0, (t_0 - \varepsilon)^+, t_0 + 2\varepsilon), \tag{4.13}
 \end{aligned}$$

thanks to Lemma 4.5. Now following similar arguments as in (4.11), one may easily derive from (4.6) that

$$\mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \vee \delta \leq t \leq t_0 + \varepsilon} |R_m(u, t, -\delta)|^p \right] \leq C \|u\|_{m+2, p_0, T} \varepsilon^{\frac{p(m+1)}{2}}. \tag{4.14}$$

Finally, note that

$$\begin{aligned} & \sup_{(t, \delta) \in D_{[t_0, t_0 + \varepsilon]}^\varepsilon} |R_m(u, t; \delta)| \\ & \leq \sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \leq t \leq t_0 + \varepsilon} |R_m(u, t; \delta)| + \sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \vee \delta \leq t \leq t_0 + \varepsilon} |R_m(u, t; -\delta)|. \end{aligned}$$

Combining (4.11) and (4.14) we obtain (4.4). ■

5. Pathwise Taylor expansion for random fields

In this section we extend our results to Itô random fields. Again, denote $t_\delta := t + \delta$ and we emphasize that δ could be negative. Let us denote

$$\begin{aligned} \hat{D} & := \left\{ (t, x, \delta, h) \in [0, \infty) \times \mathcal{O} \times \mathbb{R} \times \mathbb{R}^{d'} : (t_\delta, x + h) \in Q \right\}; \\ \hat{D}_N & := \left\{ (t, x, \delta, h) : (t, x) \in Q_N, |\delta| + |h|^2 \leq \frac{1}{N(N+1)}, t_\delta \geq 0 \right\}. \end{aligned} \tag{5.1}$$

We remark that $(t_\delta, x + h) \in Q_{N+1}$ for any $(t, x, \delta, h) \in \hat{D}_N$. Furthermore, for any $m \geq 0$ and $u \in \mathcal{H}_2^{[m]}(\hat{\Lambda})$, in light of (4.2), and noting that the spatial derivative ∂_x commutes with all the path derivatives, we shall define the m -th order Taylor expansion by:

$$u(t + \delta, x + h) = \sum_{|(\theta, \ell)| \leq m} \frac{1}{\ell!} (\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u)(t, x) h^\ell \mathcal{I}_{t, t_\delta}^\theta + R_m(u, t, x, \delta, h), \tag{5.2}$$

for any $(t, x, \delta, h) \in D$ and $\omega \in \Omega$. Clearly, if $\delta = 0$, then we recover the standard Taylor expansion in x ; and if $h = \mathbf{0}$, then we recover the Taylor expansion (4.2) for Itô processes. Moreover, if $m = 2$, then we have

$$\begin{aligned} u(t + \delta, x + h) & = \left[u + \partial_t u \delta + \partial_x u \cdot h + \partial_\omega u \cdot B_{t, t_\delta} + \frac{1}{2} \partial_{xx}^2 u : hh^T \right. \\ & \quad \left. + \partial_{x\omega} u : h[B_{t, t_\delta}]^T + \partial_{\omega\omega}^2 u : \underline{B}_{t, t_\delta} \right] (t, x) + R_2(u, t, x, \delta, h). \end{aligned} \tag{5.3}$$

Again, we begin with the following simple recursive relations for the remainders.

Lemma 5.1. *Let $u \in \mathcal{H}^{[m+1]}(\hat{\Lambda})$ for some $m \geq 1$. Then,*

$$R_m(u, t, x, \delta, h) = R_m(u, t, x, \delta, \mathbf{0}) + \sum_{i=1}^{d'} h_i \int_0^1 R_{m-1}(\partial_{x_i} u, t, x, \delta, h^i(\kappa)) d\kappa, \tag{5.4}$$

where $h^i(\kappa) := (h_1, \dots, h_{i-1}, \kappa h_i, 0, \dots, 0)$.

Proof. Given $(t, x, \delta, h) \in D$, we write

$$u(t_\delta, x + h) - u(t, x) = E_0 + \sum_{i=1}^{d'} E_i, \tag{5.5}$$

where $E_0 := u(t_\delta, x) - u(t, x)$ and $E_i := u(t_\delta, x + h^i(1)) - u(t_\delta, x + h^i(0))$. Note that by applying the temporal expansion we have

$$\begin{aligned}
 E_0 &= \sum_{1 \leq |\theta| \leq m} \mathcal{D}_\omega^\theta u(t, x) \mathcal{I}_{t, t_\delta}^\theta + R_m(u, t, x, \delta, \mathbf{0}). \\
 &= \sum_{1 \leq |(\theta, \ell)| \leq m, |\ell| = 0} \frac{1}{\ell!} \mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u(t, x) h^\ell \mathcal{I}_{t, t_\delta}^\theta + R_m(u, t, x, \delta, \mathbf{0}).
 \end{aligned} \tag{5.6}$$

On the other hand, for $i = 1, \dots, d'$, using the Taylor expansion for $\partial_x u$ we can write

$$\begin{aligned}
 E_i &= h_i \int_0^1 \partial_{x_i} u(t_\delta, x + h^i(\kappa)) d\kappa \\
 &= h_i \int_0^1 \left[\sum_{|(\theta, \ell)| \leq m-1} \frac{1}{\ell!} \mathcal{D}_x^\ell \mathcal{D}_\omega^\theta (\partial_{x_i} u)(t, x) (h^i(\kappa))^\ell \mathcal{I}_{t, t_\delta}^\theta + R_{m-1}(\partial_{x_i} u, t, x, \delta, h^i(\kappa)) \right] d\kappa \\
 &= \sum_{|(\theta, \ell)| \leq m-1, \ell_{i+1} = \dots = \ell_{d'} = 0} \frac{h_1^{\ell_1} \dots h_{i-1}^{\ell_{i-1}} h_i^{\ell_i+1}}{\ell_1 \dots \ell_{i-1} (\ell_i + 1)} \partial_{x_i} \mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u(t, x) \mathcal{I}_{t, t_\delta}^\theta \\
 &\quad + h_i \int_0^1 R_{m-1}(\partial_{x_i} u, t, x, \delta, h^i(\kappa)) d\kappa \\
 &= \sum_{|(\theta, \ell)| \leq m, \ell_i \geq 1, \ell_{i+1} = \dots = \ell_{d'} = 0} \frac{h^\ell}{\ell!} \mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u(t, x) \mathcal{I}_{t, t_\delta}^\theta \\
 &\quad + h_i \int_0^1 R_{m-1}(\partial_{x_i} u, t, x, \delta, h^i(\kappa)) d\kappa,
 \end{aligned}$$

where we replaced $\ell = (\ell_1, \dots, \ell_i, 0, \dots, 0)$ with $(\ell_1, \dots, \ell_{i-1}, \ell_i + 1, 0, \dots, 0)$. Then

$$\sum_{i=1}^{d'} E_i = \sum_{|(\theta, \ell)| \leq m, |\ell| \geq 1} \frac{h^\ell}{\ell!} \mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u(t, x) \mathcal{I}_{t, t_\delta}^\theta + \sum_{i=1}^{d'} h_i \int_0^1 R_{m-1}(\partial_{x_i} u, t, x, \delta, h^i(\kappa)) d\kappa.$$

This, together with (5.5) and (5.6), implies that

$$\begin{aligned}
 u(t_\delta, x + h) - u(t, x) &= \sum_{1 \leq |(\theta, \ell)| \leq m} \frac{h^\ell}{\ell!} \mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u(t, x) \mathcal{I}_{t, t_\delta}^\theta + R_m(u, t, x, \delta, \mathbf{0}) \\
 &\quad + \sum_{i=1}^{d'} h_i \int_0^1 R_{m-1}(\partial_{x_i} u, t, x, \delta, h^i(\kappa)) d\kappa.
 \end{aligned}$$

Now (5.4) follows immediately from (5.2). ■

Our main result of this section is the following pathwise estimate for the remainders, extending Theorem 4.1.

Theorem 5.2. Assume $u \in \mathcal{H}_{p_0}^{[m+2]}$ for some $m \geq 0$ and $p_0 > p_* := (m + 1)d' + 2$. Then for any $p < p_0$ and $0 < \alpha < 1 - \frac{p_*}{p_0}$, it holds that, for any $N > 0$,

$$\mathbb{E} \left\{ \sup_{(t, x, \delta, h) \in \hat{D}_N} \left| \frac{R_m(u, t, x; \delta, h)}{(|\delta| + |h|^2)^{\frac{m+\alpha}{2}}} \right|^p \right\} < \infty. \tag{5.7}$$

Proof. We fix N and let $0 < \varepsilon < \frac{1}{4N^2(N+1)^2}$. Then for any $(t, x) \in Q_N$ (recall (2.11)), and $|\delta| \leq \varepsilon, |h| \leq \sqrt{\varepsilon}$, we have $((t_\delta)^+, x+h) \in Q_{N+1}$ and $(t+2\varepsilon, x+h) \in Q_{N+1}$. In what follows our generic constant C will depend on $\|u\|_{m+2, p_0, N+1}$. Denote, for $(t_0, x_0) \in Q_N$ and $n \geq 0$,

$$\hat{D}_n^\varepsilon(t_0, x_0) := \{(t, x, \delta, h) : |\delta| + |h|^2 \leq \varepsilon, t_0 \leq t \leq t_0 + \varepsilon, t_\delta \geq 0, |x - x_0| \leq \varepsilon^{\frac{n+1}{2}}\}. \tag{5.8}$$

We split the proof into the following steps.

(i) We first show that, for any $(t_0, x_0) \in Q_N$ and $p < p_0$,

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_0^\varepsilon(t_0,x_0)} |R_0(u, t, x, \delta, h)|^p \right] \leq C\varepsilon^{\frac{p}{2}}. \tag{5.9}$$

Indeed, note that

$$|R_0(u, t, x, \delta, h)| = |u(t_\delta, x+h) - u(t, x)| \leq R_{0,1} + R_{0,2},$$

where $R_{0,1} := |u(t_\delta, x+h) - u(t_\delta, x_0)| + |u(t, x_0) - u(t, x)|$, and $R_{0,2} := |u(t_\delta, x_0) - u(t, x_0)|$.

Note that $R_{0,2}$ is for fixed x_0 . Applying Proposition 4.2 we get

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_0^\varepsilon(t_0,x_0)} |R_{0,2}|^p \right] \leq C\varepsilon^{\frac{p}{2}}. \tag{5.10}$$

Moreover, note that

$$R_{0,1} = \left| (x - x_0 + h) \cdot \int_0^1 \partial_x u(t_\delta, x_0 + \kappa(x - x_0 + h)) d\kappa \right| + \left| (x - x_0) \cdot \int_0^1 \partial_x u(t, x_0 + \kappa(x - x_0)) d\kappa \right|,$$

and $x_0 + \kappa(x - x_0), x_0 + \kappa(x - x_0 + h) \in K_{N+1}$, thanks to (2.12). Then, since $\partial_x u \in \mathcal{H}_{p_0}^{[m+1]} \subset \mathcal{H}_{p_0}^{[0]}$ and $p < p_0$, we have

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_0^\varepsilon(t_0,x_0)} |R_{0,1}|^p \right] \leq C\varepsilon^{\frac{p}{2}}.$$

This, together with (5.10), proves (5.9).

(ii) We next show by induction on m that, for any $m \geq 0$, any $(t_0, x_0) \in D_N$ and $p < p_0$,

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_m(u, t, x, \delta, h)|^p \right] \leq C\varepsilon^{\frac{p(m+1)}{2}}. \tag{5.11}$$

Indeed, by (5.9) we have (5.11) for $m = 0$. Assume (5.11) holds true for $m - 1$. Applying Lemma 5.1, we have

$$\begin{aligned} \mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_m(u, t, x, \delta, h)|^p \right] &\leq C \mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_m(u, t, x, \delta, \mathbf{0})|^p \right] \\ &+ C\varepsilon^{\frac{p}{2}} \sum_{i=1}^{d'} \int_0^1 \mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_{m-1}(\partial_{x_i} u, t, x, \delta, \omega, h^i(\kappa))|^p \right] d\kappa. \end{aligned}$$

Note that $\partial_{x_i} u \in \mathcal{H}_{p_0}^{[m+1]}$ and $(t, x, \delta, h^i(\kappa)) \in \hat{D}_m^\varepsilon(t_0, x_0) \subset \hat{D}_{m-1}^\varepsilon(t_0, x_0)$. Then by induction assumption we have

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_{m-1}(\partial_{x_i} u, t, x, \delta, \omega, h^i(\kappa))|^p \right] \leq C\varepsilon^{\frac{pm}{2}}.$$

So it suffices to prove

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_m(u, t, x, \delta, \mathbf{0})|^p \right] \leq C\varepsilon^{\frac{p(m+1)}{2}}. \tag{5.12}$$

To this end, we note that

$$R_m(u, t, x, \delta, \mathbf{0}) = R_m(u, t, x, \delta, \mathbf{0}) - R_m(u, t, x_0, \delta, \mathbf{0}) + R_m(u, t, x_0, \delta, \mathbf{0}). \tag{5.13}$$

Applying Proposition 4.2 again we have

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_m(u, t, x_0, \delta, \mathbf{0})|^p \right] \leq C\varepsilon^{\frac{p(m+1)}{2}}. \tag{5.14}$$

On the other hand, recall that $R_m(u, t, x, \delta, \mathbf{0}) = u(t_\delta, x) - \sum_{|\theta| \leq m} \mathcal{D}_\omega^\theta u(t, x) \mathcal{I}_{t,t_\delta}^\theta$. Then

$$\begin{aligned} &R_m(u, t, x, \delta, \mathbf{0}) - R_m(u, t, x_0, \delta, \mathbf{0}) \\ &= (x - x_0) \cdot \int_0^1 \left[\partial_x u(t_\delta, x_0 + \kappa(x - x_0)) - \sum_{|\theta| \leq m} \partial_x \mathcal{D}_\omega^\theta u(t, x_0 + \kappa(x - x_0)) \mathcal{I}_{t,t_\delta}^\theta \right] d\kappa. \end{aligned}$$

Notice that $\partial_x u, \partial_x \mathcal{D}_\omega^\theta \in \mathcal{H}_{p_0}^{[1]}$ for $|\theta| \leq m$, and $|x - x_0| \leq \varepsilon^{\frac{m+1}{2}}$ for $(t, x, \delta, h) \in \hat{D}_m^\varepsilon(t_0, x_0)$. Then, by Lemma 4.5,

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_0,x_0)} |R_m(u, t, x, \delta, \mathbf{0}) - R_m(u, t, x_0, \delta, \mathbf{0})|^p \right] \leq C\varepsilon^{\frac{p(m+1)}{2}}.$$

Plugging this and (5.11) into (5.13) we obtain (5.12), which in turn implies (5.11).

(iii) We now claim that

$$\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_N, |\delta| + |h|^2 \leq \varepsilon} |R_m(u, t, x, \delta, h)|^p \right] \leq C\varepsilon^{\frac{(p-d')(m+1)}{2} - 1}. \tag{5.15}$$

Indeed, set $t_i := i\varepsilon, i = 0, \dots, [\frac{N}{\varepsilon}] + 1$, and let $x_j \in K_N, j = 1, \dots, [(2N\varepsilon^{-\frac{m+1}{2}})^{d'}] + 1$ be discrete grids such that the union of their $\varepsilon^{\frac{m+1}{2}}$ -neighborhood covers K_N . Then we have

$$\left\{ (t, x, \delta, h) \in \hat{D}_N : |\delta| + |h|^2 \leq \varepsilon \right\} \subset \bigcup_{i,j} \hat{D}_m^\varepsilon(t_i, x_j).$$

Thus, by (5.11),

$$\begin{aligned} &\mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_N, |\delta| + |h|^2 \leq \varepsilon} |R_m(u, t, x, \delta, h)|^p \right] \\ &\leq \sum_{i,j} \mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_m^\varepsilon(t_i,x_j)} |R_m(u, t, x, \delta, h)|^p \right] \\ &\leq C \sum_{i,j} \varepsilon^{\frac{p(m+1)}{2}} \leq C\varepsilon^{-1 - \frac{d'(m+1)}{2}} \varepsilon^{\frac{p(m+1)}{2}} = C\varepsilon^{\frac{(p-d')(m+1)}{2} - 1}. \end{aligned}$$

(iv) Finally, without loss of generality, we may assume $\frac{p_*}{1-\alpha} < p < p_0$. By (5.15) we have

$$\begin{aligned} & \mathbb{E} \left[\sup_{(t,x,\delta,h) \in \hat{D}_N} \left| \frac{R_m(u, t, x, \delta, h)}{(\delta + |h|^2)^{\frac{m+\alpha}{2}}} \right|^p \right] \\ &= \mathbb{E} \left[\sup_{n \geq 0} \sup_{2^{-(n+1)} \leq |\delta| + |h|^2 \leq 2^{-n}} \sup_{(t,x,\delta,h) \in \hat{D}_N} \left| \frac{R_m(u, t, x, \delta, h)}{(\delta + |h|^2)^{\frac{m+\alpha}{2}}} \right|^p \right] \\ &\leq \sum_{n=0}^{\infty} \mathbb{E} \left[\sup_{2^{-(n+1)} \leq |\delta| + |h|^2 \leq 2^{-n}} \sup_{(t,x,\delta,h) \in \hat{D}_N} \left| \frac{R_m(u, t, x, \delta, h)}{(\delta + |h|^2)^{\frac{m+\alpha}{2}}} \right|^p \right] \\ &\leq \sum_{n=0}^{\infty} 2^{\frac{p(m+\alpha)(n+1)}{2}} \mathbb{E} \left[\sup_{|\delta| + |h|^2 \leq 2^{-n}} \sup_{(t,x,\delta,h) \in \hat{D}_N} |R_m(u, t, x, \delta, h)|^p \right] \\ &\leq C \sum_{n=0}^{\infty} 2^{\frac{p(m+\alpha)(n+1)}{2}} 2^{-n \left(\frac{(p-d')(m+1)}{2} - 1 \right)} \\ &= C 2^{\frac{p(m+\alpha)}{2}} \sum_{n=0}^{\infty} 2^{-\frac{n}{2} [p(1-\alpha) - p_*]} < \infty, \end{aligned}$$

completing the proof. ■

6. Extension to Hölder continuous case

In this section we weaken the requirement $u \in \mathcal{H}_{p_0}^{[m+2]}(\Lambda)$ in Theorem 4.1 slightly, by replacing the highest order differentiability with a certain Hölder continuity. First recall the space $\mathcal{H}_p^{[n]}(\Lambda)$ defined in (2.14). We shall now prove (2.15).

Lemma 6.1. *Let $\alpha \in (0, 1)$ and $p > \frac{2}{1-\alpha}$. Then $\mathcal{H}_p^{[n+2]}(\Lambda) \subset \mathcal{H}_p^{[n]+\alpha}(\Lambda)$ for any $n \geq 0$.*

Proof. Without loss of generality, we shall only prove the case that $n = 0$. Let $u \in \mathcal{H}_p^{[2]}(\Lambda)$. First, for any $0 \leq t_0 \leq T$ and $0 < \varepsilon < 1$, applying functional Itô formula (2.4) and then Lemma 4.5, we have

$$\mathbb{E} \left[\sup_{0 \leq \delta \leq \varepsilon} \sup_{t_0 \leq t \leq t_0 + \varepsilon} |u_{t+\delta} - u_t|^p \right] \leq C \varepsilon^{\frac{p}{2}}.$$

Then the lemma follows exactly the same arguments as in the proof of Theorem 4.1. ■

In what follows we shall assume $u \in \mathcal{H}_{p_0}^{[m+1]}(\Lambda) \cap \mathcal{H}_{p_0}^{[m]+\alpha_0}(\Lambda)$ for some appropriate m, p_0 and α_0 . Note that in this case R_m is still well-defined by (4.2), however, we cannot use the representations (4.5) and (4.6) anymore because of their involvement of the $(m + 2)$ -th order derivatives.

In order to estimate R_m in this case, we first establish the following recursive representation. Recall that $\varphi_{s,t} := \varphi_t - \varphi_s$.

Lemma 6.2. *Let $u \in \mathcal{H}_2^{[m+1]}(\Lambda)$ for some $m \geq 1$. Then for any $\delta > 0$, it holds that:*

$$\begin{aligned} R_m(u, t, \delta) &= \sum_{|\theta| \leq m-2, \theta_i \neq 0} \int_t^{t+\delta} R_{m-2-|\theta|}(\partial_t \mathcal{D}_\omega^\theta u, t, s-t) \mathcal{I}_{s,t+\delta}^\theta ds \\ &+ \sum_{|\theta|=m-1, \theta_i \neq 0} \int_t^{t+\delta} \partial_t \mathcal{D}_\omega^\theta u_s \mathcal{I}_{s,t+\delta}^\theta ds + \sum_{|\theta|=m, \theta_i \neq 0} \mathcal{I}_{t,t+\delta}^\theta ([\mathcal{D}_\omega^\theta u]_{t,\cdot}); \end{aligned} \tag{6.1}$$

$$R_m(u, t, -\delta) = - \sum_{|\theta| \leq m} (-1)^{|\theta|_0} R_{m-|\theta|}(\mathcal{D}_\omega^\theta u, t_\delta^-, \delta) \mathcal{I}_{t_\delta^-, t}^{-\theta}, \quad t \geq \delta. \tag{6.2}$$

Proof. (i) We first prove (6.1). We claim that, for $m \geq 2$,

$$R_m(u, t, \delta) = \int_t^{t_\delta} R_{m-2}(\partial_t u, t, s-t) ds + \sum_{i=1}^d \int_t^{t_\delta} R_{m-1}(\partial_{\omega^i} u, t, s-t) \circ dB_s^i. \tag{6.3}$$

Indeed, denote by $\tilde{R}_m(u, t, \delta)$ the right side above, and notice the simple fact:

$$\begin{aligned} & \sum_{|\theta| \leq m-2} \mathcal{D}_\omega^\theta \partial_t u_t \int_t^{t_\delta} \mathcal{I}_{t,s}^\theta ds + \sum_{|\theta| \leq m-1} \mathcal{D}_\omega^\theta \partial_{\omega^i} u_t \sum_{i=1}^d \int_t^{t_\delta} \mathcal{I}_{t,s}^\theta \circ dB_s^i \\ &= \sum_{1 \leq |\theta| \leq m} \mathcal{D}_\omega^\theta u_t \mathcal{I}_{t,t_\delta}^\theta. \end{aligned} \tag{6.4}$$

Applying the functional Itô formula (2.5) and by (4.2), we have

$$\begin{aligned} u_{t_\delta} &= u_t + \int_t^{t_\delta} \partial_t u_s ds + \sum_{i=1}^d \int_t^{t_\delta} \partial_{\omega^i} u_s \circ dB_s^i \\ &= u_t + \int_t^{t_\delta} \left[\sum_{|\theta| \leq m-2} \mathcal{D}_\omega^\theta \partial_t u_t \mathcal{I}_{t,s}^\theta + R_{m-2}(\partial_t u, t, s-t) \right] ds \\ &\quad + \sum_{i=1}^d \int_t^{t_\delta} \left[\sum_{|\theta| \leq m-1} \mathcal{D}_\omega^\theta \partial_{\omega^i} u_t \mathcal{I}_{t,s}^\theta + R_{m-1}(\partial_{\omega^i} u, t, s-t) \right] \circ dB_s^i \\ &= u_t + \sum_{|\theta| \leq m-2} \mathcal{D}_\omega^\theta \partial_t u_t \int_t^{t_\delta} \mathcal{I}_{t,s}^\theta ds \\ &\quad + \sum_{|\theta| \leq m-1} \mathcal{D}_\omega^\theta \partial_{\omega^i} u_t \sum_{i=1}^d \int_t^{t_\delta} \mathcal{I}_{t,s}^\theta \circ dB_s^i + \tilde{R}_m(u, t, \delta). \end{aligned}$$

Then (6.3) follows immediately from (6.4) and (4.2).

We now prove (6.1). When $m = 1$, the right side of (6.1) becomes

$$\begin{aligned} & \int_t^{t_\delta} \partial_t u_s ds + \sum_{i=1}^d \int_t^{t_\delta} [\partial_{\omega^i} u]_{t,s} \circ dB_s^i \\ &= \int_t^{t_\delta} \partial_t u_s ds + \sum_{i=1}^d \int_t^{t_\delta} \partial_{\omega^i} u_s \circ dB_s^i - \sum_{i=1}^d \partial_{\omega^i} u_t B_{t,t_\delta}^i \\ &= u_{t_\delta} - u_t - \sum_{i=1}^d \partial_{\omega^i} u_t B_{t,t_\delta}^i = R_1(u, t, \delta). \end{aligned}$$

For $m \geq 2$, applying (6.3) repeatedly on the stochastic integral terms in (6.3) we obtain

$$\begin{aligned} R_m(u, t, \delta) &= \sum_{|\theta| \leq m-2, \theta_i \neq 0} \mathcal{I}_{t,t_\delta}^\theta \left(\int_t^{t_\delta} R_{m-2-\theta}(\partial_t \mathcal{D}_\omega^\theta u, t, s-t) ds \right) \\ &\quad + \sum_{|\theta| = m-1, \theta_i \neq 0} \mathcal{I}_{t,t_\delta}^\theta \left(R_1(\mathcal{D}_\omega^\theta u, t, \cdot - t) \right). \end{aligned} \tag{6.5}$$

Applying stochastic Fubini theorem repeatedly, we have

$$\mathcal{I}_{t,t_\delta}^\theta \left(\int_t^\cdot R_{m-2-\theta}(\partial_t \mathcal{D}_\omega^\theta u, t, s-t) ds \right) = \int_t^{t_\delta} R_{m-2-\theta}(\partial_t \mathcal{D}_\omega^\theta u, t, s-t) \mathcal{I}_{s,t_\delta}^\theta ds.$$

Moreover, note that $\mathcal{D}_\omega^\theta u \in \mathcal{H}_2^{[2]}(\Lambda)$ for $|\theta| = m - 1$. Then

$$\begin{aligned} R_1(\mathcal{D}_\omega^\theta u, t, s-t) &= [\mathcal{D}_\omega^\theta u]_{t,s} - \sum_{j=1}^d \partial_{\omega_j} \mathcal{D}_\omega^\theta u_t B_{t,s}^j \\ &= \int_t^s \partial_t \mathcal{D}_\omega^\theta u_r dr + \sum_{j=1}^d \int_t^s \partial_{\omega_j} \mathcal{D}_\omega^\theta u_r \circ dB_r^j - \sum_{j=1}^d \partial_{\omega_j} \mathcal{D}_\omega^\theta u_t B_{t,s}^j \\ &= \int_t^s \partial_t \mathcal{D}_\omega^\theta u_r dr + \sum_{j=1}^d \int_t^s [\partial_{\omega_j} \mathcal{D}_\omega^\theta u]_{t,r} \circ dB_r^j. \end{aligned}$$

Thus, (6.5) leads to

$$\begin{aligned} R_m(u, t, \delta) &= \sum_{|\theta| \leq m-2, \theta_i \neq 0} \int_t^{t_\delta} R_{m-2-\theta}(\partial_t \mathcal{D}_\omega^\theta u, t, s-t) \mathcal{I}_{s,t_\delta}^\theta ds \\ &\quad + \sum_{|\theta| = m-1, \theta_i \neq 0} \mathcal{I}_{t,t_\delta}^\theta \left(\int_t^\cdot \partial_t \mathcal{D}_\omega^\theta u_r dr + \sum_{j=1}^d \int_t^\cdot [\partial_{\omega_j} \mathcal{D}_\omega^\theta u]_{t,r} \circ dB_r^j \right), \end{aligned}$$

which, together with stochastic Fubini theorem again, implies (6.1) immediately.

(ii) We next prove (6.2). By applying (4.2) twice we have

$$\begin{aligned} R_m(u, t, -\delta) &= u_{t_\delta^-} - \sum_{|\theta| \leq m} (-1)^{|\theta|_0} \mathcal{D}_\omega^\theta u_t \mathcal{I}_{t_\delta^-,t}^{-\theta} \\ &= u_{t_\delta^-} - \sum_{|\theta| \leq m} (-1)^{|\theta|_0} \left[\sum_{|\tilde{\theta}| \leq m-|\theta|} \mathcal{D}_\omega^{\tilde{\theta}} \mathcal{D}_\omega^\theta u_{t_\delta^-} \mathcal{I}_{t_\delta^-,t}^{\tilde{\theta}} + R_{m-|\theta|}(\mathcal{D}_\omega^\theta u, t_\delta^-, \delta) \right] \mathcal{I}_{t_\delta^-,t}^{-\theta}. \end{aligned}$$

We now define

$$\begin{aligned} \Delta_m &:= R_m(u, t, -\delta) + \sum_{|\theta| \leq m} (-1)^{|\theta|_0} R_{m-|\theta|}(\mathcal{D}_\omega^\theta u, t_\delta^-, \delta) \mathcal{I}_{t_\delta^-,t}^{-\theta} \\ &= u_{t_\delta^-} - \sum_{|\theta| \leq m} (-1)^{|\theta|_0} \sum_{|\tilde{\theta}| \leq m-|\theta|} \mathcal{D}_\omega^{\tilde{\theta}} \mathcal{D}_\omega^\theta u_{t_\delta^-} \mathcal{I}_{t_\delta^-,t}^{\tilde{\theta}} \mathcal{I}_{t_\delta^-,t}^{-\theta}. \end{aligned} \tag{6.6}$$

Denote $\hat{\theta} := (\tilde{\theta}, \theta) = (\hat{\theta}_1, \dots, \hat{\theta}_n)$, then one can check that

$$\Delta_m = \sum_{1 \leq |\hat{\theta}| \leq m} (-1)^{|\hat{\theta}|_0+1} \mathcal{D}_\omega^{\hat{\theta}} u_{t_\delta^-} a(\hat{\theta}), \quad \text{where } a(\hat{\theta}) := \sum_{i=0}^n (-1)^i \mathcal{I}_{t_\delta^-,t}^{(\hat{\theta}_1, \dots, \hat{\theta}_i)} \mathcal{I}_{t_\delta^-,t}^{(\hat{\theta}_{i+1}, \dots, \hat{\theta}_n)}.$$

By setting $\varphi := 1$ in (4.12) we see that $a(\hat{\theta}) = 0$ for all $1 \leq |\hat{\theta}| \leq m$. Then $\Delta_m = 0$ and we complete the proof. ■

With the above representations, we can now extend Theorem 4.1.

Theorem 6.3. Assume that $u \in \mathcal{H}_{p_0}^{[m+1]}(\Lambda) \cap \mathcal{H}_{p_0}^{[m]+\alpha_0}(\Lambda)$ for some $m \geq 0, \alpha_0 \in (0, 1)$, and $p_0 > \frac{2}{\alpha_0}$. The for any $p < p_0$ and $0 < \alpha < \alpha_0 - \frac{2}{p_0}$, it holds that, for any $T > 0$,

$$\mathbb{E} \left\{ \sup_{(t, \delta) \in D_{[0, T]}^1} \left| \frac{R_m(u, t; \delta)}{|\delta|^{\frac{m+\alpha}{2}}} \right|^p \right\} < \infty. \tag{6.7}$$

Proof. Applying the representations (6.1) and (6.2), one can easily prove by induction on m that

$$\mathbb{E} \left[\sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \leq t \leq t_0 + \varepsilon} |R_m(u, t; \delta)|^p + \sup_{0 < \delta \leq \varepsilon} \sup_{t_0 \vee \delta \leq t \leq t_0 + \varepsilon} |R_m(u, t; -\delta)|^p \right] \leq C\varepsilon^{\frac{p(m+\alpha_0)}{2}}, \tag{6.8}$$

extending (4.4). The result then follows from very similar arguments as those in Section 4. We leave the details to interested reader. ■

Along the similar lines of arguments we can also weaken the assumptions of Theorem 5.2, the case for random field, to the Hölder conditions. Since the proof is a routine combination of the previous results, we omit it.

Theorem 6.4. Assume that $u \in \mathcal{H}_{p_0}^{[m+1]}(\Lambda) \cap \mathcal{H}_{p_0}^{[m]+\alpha_0}(\Lambda)$ for some $m \geq 0, \alpha_0 \in (0, 1)$, and $p_0 > p_* := \frac{1}{\alpha_0}[(m + \alpha_0)d' + 2]$. Then for any $p < p_0$ and $0 < \alpha < \alpha_0[1 - \frac{p_*}{p_0}]$, (5.7) holds for any $N > 0$.

7. Application to (forward) stochastic PDEs

One of the main purposes of our study on the pathwise Taylor expansion is to lay the foundation for the notion of (pathwise) viscosity solution for stochastic PDEs and the associated forward path-dependent PDEs, which will be articulated in our accompanying paper [4]. More precisely, we are particularly interested in the case when the random field u is a (classical) solution (in standard sense) of the following SPDE:

$$du(t, x) = f(t, x, u, \partial_x u, \partial_x^2 u)dt + g(t, x, u, \partial_x u) \circ dB_t, \quad t \geq 0, \mathbb{P}_0\text{-a.s.} \tag{7.1}$$

where $f(t, x, y, z, \gamma)$ and $g(t, x, y, z)$ are \mathbb{F} -progressively measurable and taking values in \mathbb{R} and \mathbb{R}^d , respectively, with the variable ω omitted as usual. Clearly, the SPDE (7.1) can be rewritten as the following system of (forward) path dependent PDE (PPDE):

$$\partial_t u - f(t, x, u, \partial_x u, \partial_x^2 u) = 0; \quad \partial_\omega u(t, x) = g(t, x, u, \partial_x u). \tag{7.2}$$

We will be particularly interested in the version of Theorem 5.2, applied to the solutions of (7.1) (or equivalently (7.2)) in the case $m = 2$. To this end, we first assume that u is a solution of SPDE (7.1) that is smooth enough in our sense. We shall also assume that g is sufficiently smooth. It is then clear that

$$\partial_\omega u(t, x) = g(t, x, u(t, x), \partial_x u(t, x)). \tag{7.3}$$

Differentiating both sides above in x we get (suppressing variables and noting that $g = (g_1, \dots, g_d)^T$): for $i = 1, \dots, d'$, and $j = 1, \dots, d$,

$$\partial_{x_i \omega^j}^2 u = \partial_{x_i} (g_j(t, x, u, \partial_x u)) = \partial_{x_i} g_j + \partial_y g_j \partial_{x_i} u + \sum_{k=1}^{d'} [\partial_{z_k} g_j] \partial_{x_i x_k}^2 u, \tag{7.4}$$

or in matrix form:

$$\partial_{x\omega}^2 u = [\partial_{x_i \omega_j} u]_{1 \leq i \leq d', 1 \leq j \leq d} = \partial_x g + \partial_x u [\partial_y g]^T + \partial_{xx}^2 u \partial_z g \in \mathbb{R}^{d' \times d}. \tag{7.5}$$

Similarly, we can easily derive

$$\begin{aligned} \partial_{\omega\omega}^2 u &= \partial_{\omega} g + \partial_{\omega} u [\partial_y g]^T + [\partial_{x\omega}^2 u]^T [\partial_z g] \\ &= \partial_{\omega} g + g [\partial_y g]^T + [\partial_x g + \partial_x u [\partial_y g]^T + \partial_{xx}^2 u \partial_z g]^T [\partial_z g]. \end{aligned} \tag{7.6}$$

In light of (5.2) with $m = 2$, we can now formally write down the pathwise Taylor expansion:

$$\begin{aligned} &u(t + \delta, x + h, \omega) - u(t, x, \omega) \\ &= \sum_{1 \leq |(\theta, \ell)| \leq 2} \frac{1}{\ell!} (D_x^\ell D_\omega^\theta u)(t, x) h^\ell I_{t, t+\delta}^\theta + R(u, t, x, \delta, h) \\ &= f(t, x, u, \partial_x u, \partial_{xx}^2 u) \delta + \partial_x u \cdot h + g \cdot \omega_{t+\delta}^t + \frac{1}{2} \partial_{xx}^2 u : hh^T \\ &\quad + \left[\partial_x g + \partial_x u [\partial_y g]^T + \partial_{xx}^2 u \partial_z g \right] : h [B_{t, t+\delta}]^T \\ &\quad + \left[\partial_{\omega} g + g [\partial_y g]^T + [\partial_x g + \partial_x u [\partial_y g]^T + \partial_{xx}^2 u \partial_z g]^T [\partial_z g] \right] : \underline{B}_{t, t+\delta} \\ &\quad + R(u, t, x, \delta, h), \end{aligned} \tag{7.7}$$

for any $(t, x, \delta, h) \in \hat{D}$, where the right hand side above is evaluated at (t, x, ω) , and $\underline{B}_{t, t+\delta}$ is defined by (3.15). Applying Theorem 5.2 we then obtain the following result.

Theorem 7.1. Assume that SPDE (7.1) has a solution $u \in \mathcal{H}_{p_0}^{[4]}(\hat{\Lambda})$ for some $p_0 > p_* := 3d' + 2$, and that the coefficient g is regular enough so that all the derivatives in (7.6) are well-defined. Let $R(u, t, x, \delta, h)$ be determined (7.7). Then for any $p < p_0$ and $0 < \alpha < 1 - \frac{p_*}{p_0}$, the remainder R satisfies (5.7) with $m = 2$.

Remark 7.2. The SPDE (7.1) can be written as the following Itô form:

$$du(t, x) = F(t, x, u, \partial_x u, \partial_{xx}^2 u) dt + g(t, x, u, \partial_x u) \cdot dB_t, \quad t \geq 0, \mathbb{P}_0\text{-a.s.} \tag{7.8}$$

where

$$F(t, x, y, z, \gamma) := f + \frac{1}{2} \text{tr}(\partial_{\omega\omega} g + g [\partial_y g]^T + [\partial_x g + z [\partial_y g]^T + \gamma \partial_z g]^T [\partial_z g]).$$

It is thus natural to define the *parabolicity* of the PPDE (7.2) as

$$\partial_\gamma f = \partial_\gamma F - \frac{1}{2} [\partial_z g] [\partial_z g]^T \geq \mathbf{0}. \tag{7.9}$$

Clearly, this is exactly the coercivity condition in the SPDE literature (see for example Rozovskii [17] and Ma–Yong [16] in linear cases). ■

We should note that the requirement $u \in \mathcal{H}_{p_0}^{[4]}(\hat{\Lambda})$ in Theorem 7.1 is much stronger than u being a classical solution, due to the involvement of path derivatives. In the rest of this section we shall establish the connection between the two concepts. To begin with let us recall some Sobolev spaces: for any $m \geq 0$ and $p \geq 1$,

$$W_p^m(\hat{\Lambda}) := \{u \in \mathbb{L}^0(\hat{\Lambda}) : \|u\|_{W_p^m, N} < \infty, \forall N > 0\}, \quad W_*^m(\hat{\Lambda}) := \bigcap_{p \geq 1} W_p^m(\hat{\Lambda}), \tag{7.10}$$

where $\|u\|_{W_p^m, N}^p := \sup_{0 \leq t \leq T} \sum_{|\ell| \leq m} \mathbb{E}[\int_{K_N} |\mathcal{D}_x^\ell u(t, x)|^p dx]$.

Next, we extend the spatial derivatives slightly to those involving (x, y, z, γ) . To this end, recall the multi-index set Θ and the norm $|(\theta, \ell)|$ on Θ defined by (2.29). We define

$$\mathcal{D}_{\omega, (x, y, z)}^{(\theta, \ell, \tilde{\ell})} \varphi := \mathcal{D}_x^\ell \mathcal{D}_{(y, z)}^{\tilde{\ell}} \mathcal{D}_\omega^\theta \varphi; \quad \text{and} \quad \mathcal{D}_{\omega, (x, y, z, \gamma)}^{(\theta, \ell, \hat{\ell})} \varphi := \mathcal{D}_x^\ell \mathcal{D}_{(y, z, \gamma)}^{\hat{\ell}} \mathcal{D}_\omega^\theta \varphi, \tag{7.11}$$

where, by a slight abuse of notation, $(\theta, \ell, \tilde{\ell}) \in \Theta \times \mathbb{N}^{1+d'}$ and $(\theta, \ell, \hat{\ell}) \in \Theta \times \mathbb{N}^{1+d'+d' \times d'}$, respectively; and $\varphi = \varphi(t, x, \omega, y, z, \gamma)$ is any random field such that these derivatives exist. Moreover, we define $|\tilde{\ell}|$ and $|\hat{\ell}|$ in an obvious way.

We now state the main result of this section.

Proposition 7.3. *Let $m \geq 1$, $p > d'$, and denote $p_m := p(1 + \frac{m(m+1)}{2})$. Assume that*

- (i) *for any $|\theta| \leq m - 1$, $\mathcal{D}_{\omega, (x, y, z)}^{(\theta, \ell, \tilde{\ell})} g$ exists for all $|\ell| + |\tilde{\ell}| \leq m - |\theta|$, and is uniformly bounded when $|\tilde{\ell}| \geq 1$. Moreover, $\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta g$ is uniformly Lipschitz continuous in (y, z) and $\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta g(\cdot, 0, \mathbf{0}) \in W_{p_m}^0(\hat{\Lambda})$ for $|\ell| \leq m - |\theta|$.*
- (ii) *for any $|\theta| \leq m - 2$, $\mathcal{D}_{\omega, (x, y, z, \gamma)}^{(\theta, \ell, \hat{\ell})} f$ exists for all $|\ell| + |\hat{\ell}| \leq m - 1 - |\theta|$, and is uniformly bounded when $|\hat{\ell}| \geq 1$. Moreover, $\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta f$ is uniformly Lipschitz continuous in (y, z, γ) and $\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta f(\cdot, 0, \mathbf{0}, \mathbf{0}) \in W_{p_m}^0(\hat{\Lambda})$ for $|\ell| \leq m - 1 - |\theta|$.*

Let $u \in W_{p_m}^{m+1}(\hat{\Lambda})$ be a classical solution (in the standard sense with differentiability in x only) to SPDE (7.1), then $u \in \mathcal{H}_p^{[m]}(\hat{\Lambda})$. ■

We remark that the requirement $u \in W_{p_m}^{m+1}(\hat{\Lambda})$ is quite strong. We refer for example to Krylov [13, Theorem 5.1] for some results along this direction. The problem is in general very challenging, especially in fully nonlinear case, and we shall leave it for future research.

To prove Proposition 7.3 we need a technical lemma that would transform all path derivatives to the x -derivatives. Let us first introduce some notations. For any random fields $\varphi = \varphi(t, x, \omega, y, z, \gamma)$ and $u = u(t, x, \omega)$, we define (suppressing variables)

$$\widehat{\varphi}(t, x, \omega) := \varphi(t, x, \omega, u, \partial_x u, \partial_{xx} u). \tag{7.12}$$

Next, for a given $(\bar{\theta}, \bar{\ell}) \in \Theta$, we define, with $(\theta, \ell, \tilde{\ell}) \in \Theta \times \mathbb{N}^{1+d'}$, $(\theta, \ell, \hat{\ell}) \in \Theta \times \mathbb{N}^{1+d'+d' \times d'}$,

$$\begin{aligned} A_1 &:= A_1(\bar{\theta}, \bar{\ell}) := \{\ell : |\ell| \leq |(\bar{\theta}, \bar{\ell})|\}; \\ A_2 &:= A_2(\bar{\theta}, \bar{\ell}) := \{(\theta, \ell, \tilde{\ell}) : |\theta| \leq |\bar{\theta}| - 1, |(\theta, \ell, \tilde{\ell})| \leq |(\bar{\theta}, \bar{\ell})| - 1\}; \\ A_3 &:= A_3(\bar{\theta}, \bar{\ell}) := \{(\theta, \ell, \hat{\ell}) : |\theta| \leq |\bar{\theta}| - 2, |(\theta, \ell, \hat{\ell})| \leq |(\bar{\theta}, \bar{\ell})| - 2\}. \end{aligned} \tag{7.13}$$

Let us now consider the random fields of the following form:

$$\psi := \prod_{\ell \in A_1} (\mathcal{D}_x^\ell u)^{a_\ell^1} \prod_{(\theta, \ell, \tilde{\ell}) \in A_2} [\widehat{\mathcal{D}}_{\omega, (x, y, z)}^{(\theta, \ell, \tilde{\ell})} g]^{a_{(\theta, \ell, \tilde{\ell})}^2} \prod_{(\theta, \ell, \hat{\ell}) \in A_3} [\widehat{\mathcal{D}}_{\omega, (x, y, z, \gamma)}^{(\theta, \ell, \hat{\ell})} f]^{a_{(\theta, \ell, \hat{\ell})}^3} \tag{7.14}$$

where, $a_\ell^1, a_{(\theta, \ell, \tilde{\ell})}^3 \in \mathbb{N}$ and $a_{(\theta, \ell, \tilde{\ell})}^2 \in \mathbb{N}^d$. We note that by definition the derivatives in (7.11) have the same dimension as the function φ . In particular, since g is \mathbb{R}^d -valued, the meaning of $[\widehat{\mathcal{D}}_{\omega, (x, y, z)}^{(\theta, \ell, \tilde{\ell})} g]^a$, $a \in \mathbb{N}^d$, in (7.14) should be understood as that of x^ℓ defined in (2.28).

Moreover, for each such ψ , we define its “index”, $\lambda(\psi)$, by:

$$\lambda(\psi) := \sum_{\ell \in A_1} a_\ell^1 + \sum_{\substack{(\theta, \ell, \tilde{\ell}) \in A_2, \\ |\tilde{\ell}|=0}} |a_{(\theta, \ell, \tilde{\ell})}^2| + \sum_{\substack{(\theta, \ell, \hat{\ell}) \in A_3, \\ |\hat{\ell}|=0}} a_{(\theta, \ell, \hat{\ell})}^3. \tag{7.15}$$

We remark that in the above we do not include the exponents when $|\tilde{\ell}| > 0$ or $|\hat{\ell}| > 0$, since the (y, z, γ) -derivatives are assumed to be bounded in Proposition 7.3.

The following lemma will be crucial to the proof of Proposition 7.3. Since its proof is rather lengthy, we defer it to the end of the section in order not to disturb our discussion.

Lemma 7.4. *Assume f and g are smooth enough with respect to all variables $(t, \omega, x, y, z, \gamma)$, and u is a classical solution (in standard sense) to SPDE (7.1) with sufficient regularity in x . Then, for any $(\bar{\theta}, \bar{\ell}) \in \Theta$, $\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u$ is a linear combination of the terms in the form (7.14).*

Moreover, for each term ψ , the following estimate holds for its index:

$$\lambda(\psi) \leq 1 + |\bar{\theta}|_0 |\bar{\ell}| + \frac{|\bar{\theta}|(|\bar{\theta}| - 1)}{2}. \tag{7.16}$$

Assuming this lemma we now prove Proposition 7.3.

Proof of Proposition 7.3. First, we recall so-called Morrey’s inequality (see, e.g., [10]) which states: for any $\varphi : O \rightarrow \mathbb{R}$ that is in $W_p^1(O)$ (namely the generalized derivative $\partial_x \varphi$ is in $\mathbb{L}^p(O)$), where $O \subset \mathbb{R}^{d'}$ is a bounded domain with C^1 boundary, and any $p > d'$ and $0 < \gamma < 1 - \frac{d'}{p}$, it holds that

$$\sup_{x \in O} |\varphi(x)|^p + \sup_{x, x' \in O} \left(\frac{|\varphi(x) - \varphi(x')|}{|x - x'|^\gamma} \right)^p \leq C \int_O [|\varphi|^p + |\partial_x \varphi|^p] dx. \tag{7.17}$$

Now for $N > 0$, recall the set K_N defined by (2.11). Let O be a domain with C^1 boundary such that $K_N \subset O \subset K_{N+1}$. From Morrey’s inequality (7.17) we deduce that

$$\sup_{x \in K_N} |u(t, x)|^p \leq C_N \int_{K_{N+1}} [|u(t, x)|^p + |\partial_x u(t, x)|^p] dx.$$

Thus, to prove the proposition it suffices to show that $\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta u \in W_p^1(\hat{\Lambda})$ for all $|(\theta, \ell)| \leq m$, that is,

$$\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u \in W_p^0(\hat{\Lambda}) \quad \text{for all } |\bar{\theta}| \leq m, \quad |(\bar{\theta}, \bar{\ell})| \leq m + 1. \tag{7.18}$$

To this end we fix $(\bar{\theta}, \bar{\ell})$ as in (7.18). If $m = 1$, then $\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u$ is either $\mathcal{D}_x^{\bar{\ell}} u$ for $|\bar{\ell}| \leq 2$, or $\mathcal{D}_x^{\bar{\ell}} \partial_i u = \widehat{\mathcal{D}_x^{\bar{\ell}} g^i}$ for some $i = 1, \dots, d$ and $|\bar{\ell}| \leq 1$, and one can check (7.18) immediately.

We thus assume $m \geq 2$. Denote, for $A_i := A_i(\bar{\theta}, \bar{\ell})$, $i = 1, 2, 3$, we see that

$$\xi := \sum_{\ell \in A_1} |\mathcal{D}_x^\ell u| + \sum_{(\theta, \ell, \tilde{\ell}) \in A_2, |\tilde{\ell}|=0} |\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta g(\cdot, 0, \mathbf{0})| + \sum_{(\theta, \ell, \hat{\ell}) \in A_3, |\hat{\ell}|=0} |\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta f(\cdot, 0, \mathbf{0}, \mathbf{0})|.$$

Note that,

$$\begin{aligned} |\widehat{\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta g}| &\leq |\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta g(\cdot, 0, 0)| + C[|u| + |\partial_x u|] \leq C\xi, \quad (\theta, \ell, \tilde{\ell}) \in A_2; \\ |\widehat{\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta f}| &\leq |\mathcal{D}_x^\ell \mathcal{D}_\omega^\theta f(\cdot, 0, 0, 0)| + C[|u| + |\partial_x u| + |\partial_{xx} u|] \leq C\xi, \quad (\theta, \ell, \hat{\ell}) \in A_3. \end{aligned}$$

Note that $1 + |\bar{\theta}|_0|\bar{\ell}| + \frac{|\bar{\theta}|(|\bar{\theta}|-1)}{2} \leq 1 + \frac{m(m+1)}{2}$ for any $(\bar{\theta}, \bar{\ell})$ in (7.18). Applying Lemma 7.4, one can then check that $|\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u| \leq C\xi^{1+\frac{m(m+1)}{2}}$, which leads to (7.18) immediately. ■

We now complete this section by proving Lemma 7.4.

Proof of Lemma 7.4. For simplicity, in this proof we assume $d = d' = 1$. In particular, in this case $\bar{\ell} \in \mathbb{N}$ and thus $|\bar{\ell}| = \bar{\ell}$. We first remark that if ψ_1, \dots, ψ_n are the terms taking form of (7.14), then so is $\prod_{i=1}^n \psi_i$. Furthermore, it holds that

$$\lambda\left(\prod_{i=1}^n \psi_i\right) = \sum_{i=1}^n \lambda(\psi_i). \tag{7.19}$$

We shall proceed in two steps.

Step 1. We first prove by induction on $|\bar{\ell}|$ that $\mathcal{D}_x^{\bar{\ell}}(\widehat{f})$ is a linear combination of terms:

$$\psi := \prod_{|\ell| \leq |\bar{\ell}|+2} (\mathcal{D}_x^\ell u)^{a_\ell^1} \prod_{|(\ell, \hat{\ell})| \leq |\bar{\ell}|} (\widehat{\mathcal{D}_{(x,y,z,\gamma)}^{\ell, \hat{\ell}} f})^{a_{(\ell, \hat{\ell})}^3}; \tag{7.20}$$

and the index $\lambda(\psi)$ satisfies the estimate

$$\lambda(\psi) := \sum_{|\ell| \leq |\bar{\ell}|+2} a_\ell^1 + \sum_{|(\ell, \hat{\ell})| \leq |\bar{\ell}|} a_{(\ell, \hat{\ell})}^3 \leq 1 + |\bar{\ell}|. \tag{7.21}$$

Indeed, when $|\bar{\ell}| = 0$, we have $\psi := \widehat{f}$. Then all a_ℓ^1 's and $a_{(\ell, \hat{\ell})}^3$'s are equal to 0 except $a_{(\ell, \hat{\ell})}^3 = 1$ for $|\ell, \hat{\ell}| = 0$, and thus $\lambda(\psi) = 1$.

Assume the results hold true for m and $\bar{\ell} := \bar{\ell}' + 1$ with $|\bar{\ell}'| = m$. Let ψ' be a term in (7.20) corresponding to $\bar{\ell}'$. Then a typical term ψ of $\mathcal{D}^{\bar{\ell}}(\widehat{f}) = \partial_x[\mathcal{D}^{\bar{\ell}'}(\widehat{f})]$ should come from $\partial_x \psi'$. By (7.19), we now check the x -derivative of each factor of ψ' and see its impact on λ .

First, for $|\ell| \leq |\bar{\ell}'|+2$, we have $\partial_x(\mathcal{D}_x^\ell u) = \mathcal{D}_x^{\ell+1} u$, we see that $|\ell+1| \leq |\bar{\ell}'|+2+2 = |\bar{\ell}|+2$ and the corresponding $\lambda(\psi) = \lambda(\psi')$. Next, for $|(\ell', \hat{\ell}')| \leq |\bar{\ell}'|$,

$$\begin{aligned} \partial_x \left[\widehat{\mathcal{D}_{(x,y,z,\gamma)}^{\ell', \hat{\ell}'}} f \right] &= \left[\partial_x \widehat{\mathcal{D}_{(x,y,z,\gamma)}^{\ell', \hat{\ell}'}} f \right] + \left[\partial_y \widehat{\mathcal{D}_{(x,y,z,\gamma)}^{\ell', \hat{\ell}'}} f \right] \partial_x u \\ &\quad + \left[\partial_z \widehat{\mathcal{D}_{(x,y,z,\gamma)}^{\ell', \hat{\ell}'}} f \right] \partial_{xx} u + \left[\partial_\gamma \widehat{\mathcal{D}_{(x,y,z,\gamma)}^{\ell', \hat{\ell}'}} f \right] \partial_{xxx} u. \end{aligned}$$

The derivatives of the f terms are up to the order $|(\ell', \hat{\ell}')| + 1 \leq |\bar{\ell}'| + 1 \leq |\bar{\ell}|$, and those of the u terms are up to the order $3 \leq |\bar{\ell}| + 2$, so each term is still in the form of (7.20). Moreover, the first three terms do not increase λ , while the last term increase λ by 1. Summarizing, we see that each term ψ of $\partial_x \psi'$ is in the form of (7.20) and $\lambda(\psi) \leq \lambda(\psi') + 1$. Then we prove (7.21) for $|\bar{\ell}|$.

Similarly, we can prove that $\mathcal{D}_x^{\bar{\ell}}(\widehat{g})$ is a linear combination of terms:

$$\psi := \prod_{|\ell| \leq |\bar{\ell}|+1} (\mathcal{D}_x^\ell u)^{a_\ell^1} \prod_{|(\ell, \hat{\ell})| \leq |\bar{\ell}|} (\widehat{\mathcal{D}_{(x,y,z)}^{\ell, \hat{\ell}} g})^{a_{(\ell, \hat{\ell})}^2}; \tag{7.22}$$

and the index satisfies the estimate: $\lambda(\psi) := \sum_{|\ell| \leq |\bar{\ell}|+1} a_\ell^1 + \sum_{|(\ell, \hat{\ell})| \leq |\bar{\ell}|} a_{(\ell, \hat{\ell})}^2 \leq 1 + |\bar{\ell}|$.

Step 2. We now prove the lemma by induction on $|\bar{\theta}|_0$. When $|\bar{\theta}|_0 = 0$, the results are obvious. Assume the results hold true for n , and $\bar{\theta} = (\theta_1, \bar{\theta}')$ with $|\bar{\theta}'|_0 = n$. Note that

$\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u = \partial_{\theta_1} (\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}'} u)$. Let ψ' be a term in the form of (7.14) corresponding to $(\bar{\theta}', \bar{\ell})$, then a typical term of $\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u$ should come from $\partial_{\theta_1} \psi'$. We show that

$$\psi \text{ is in the form of (7.14) corresponding to } (\bar{\theta}, \bar{\ell}), \text{ and } \lambda(\psi) \leq \lambda(\psi') + |(\bar{\theta}', \bar{\ell})|. \tag{7.23}$$

This clearly implies (7.16) for $(\bar{\theta}, \bar{\ell})$.

We prove (7.23) in two cases. Denote $m := |(\bar{\theta}, \bar{\ell})|$ and $m' := |(\bar{\theta}', \bar{\ell})|$.

Case 1. $\theta_1 = 0$. Then $|\bar{\theta}| = |\bar{\theta}'| + 2$, $m = m' + 2$ and $\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}} u = \partial_t (\mathcal{D}_x^{\bar{\ell}} \mathcal{D}_\omega^{\bar{\theta}'} u)$. By (7.19), we now check the t -derivative for each factor of ψ' and see its impact on λ .

First, for $\ell \in A_1(\bar{\theta}', \bar{\ell})$, we have $\partial_t (\mathcal{D}_x^\ell u) = \mathcal{D}_x^\ell (\widehat{f})$. Note that $|\ell| + 2 \leq m' + 2 = m$ and $|\ell| \leq m' = m - 2$, then (7.20) implies that each term of $\mathcal{D}_x^\ell (\widehat{f})$ is in the form of (7.14). Moreover, by (7.21), this differentiation increases the index λ from 1 up to $1 + |\ell| \leq 1 + m'$. Then $\lambda(\psi) \leq \lambda(\psi') + m'$.

Next, for $(\theta, \ell, \tilde{\ell}) \in A_2(\bar{\theta}', \bar{\ell})$ with $m' \geq 1$ ($A_2(\bar{\theta}', \bar{\ell})$ is empty when $m' = 0$), we have

$$\begin{aligned} \partial_t [\widehat{\mathcal{D}_{\omega, (x, y, z)}^{(\theta, \ell, \tilde{\ell})} g}] &= [\widehat{\mathcal{D}_{(x, y, z)}^{(\ell, \tilde{\ell})} \partial_t \mathcal{D}_\omega^\theta g}] + [\partial_y \widehat{\mathcal{D}_{(x, y, z)}^{(\ell, \tilde{\ell})} \mathcal{D}_\omega^\theta g}] \partial_t u + [\partial_z \widehat{\mathcal{D}_{(x, y, z)}^{(\ell, \tilde{\ell})} \mathcal{D}_\omega^\theta g}] \partial_{xt} u \\ &= [\widehat{\mathcal{D}_{(x, y, z)}^{(\ell, \tilde{\ell})} \partial_t \mathcal{D}_\omega^\theta g}] + [\partial_y \widehat{\mathcal{D}_{(x, y, z)}^{(\ell, \tilde{\ell})} \mathcal{D}_\omega^\theta g}] \widehat{f} + [\partial_z \widehat{\mathcal{D}_{(x, y, z)}^{(\ell, \tilde{\ell})} \mathcal{D}_\omega^\theta g}] \partial_x [\widehat{f}]. \end{aligned}$$

The derivatives of g are up to the order $|(\theta, \ell, \tilde{\ell})| + 2 \leq m' - 1 + 2 = m - 1$, and its path derivatives are up to the order $|\theta| + 2 \leq |\bar{\theta}'| - 1 + 2 = |\bar{\theta}| - 1$, then these terms are in the form of (7.14). Moreover, since $m \geq 3$, by (7.20) one can easily see that all the terms of \widehat{f} and $\partial_x [\widehat{f}]$ are in the form of (7.14). Furthermore, all the g -terms do not increase λ ; the term \widehat{f} increases λ up to $1 \leq m'$. When $m' \geq 2$, the term $\partial_x [\widehat{f}]$ increase λ up to $1 + 1 \leq m'$. When $m' = 1$, we must have $|(\theta, \ell, \tilde{\ell})| = 0$, then one can check straightforwardly that the λ increases from 1 to 2, namely the increase is $1 = m'$. So in all the cases we have $\lambda(\psi) \leq \lambda(\psi') + m'$.

Finally, for $(\theta, \ell, \hat{\ell}) \in A_3(\bar{\theta}', \bar{\ell})$ with $m' \geq 2$ ($A_3(\bar{\theta}', \bar{\ell})$ is empty when $m' \leq 1$), we have

$$\begin{aligned} \partial_t [\widehat{\mathcal{D}_{\omega, (x, y, z, \gamma)}^{(\theta, \ell, \hat{\ell})} f}] &= [\widehat{\mathcal{D}_{(x, y, z, \gamma)}^{(\ell, \hat{\ell})} \partial_t \mathcal{D}_\omega^\theta f}] + [\partial_y \widehat{\mathcal{D}_{(x, y, z, \gamma)}^{(\ell, \hat{\ell})} \mathcal{D}_\omega^\theta f}] \widehat{f} \\ &\quad + [\partial_z \widehat{\mathcal{D}_{(x, y, z, \gamma)}^{(\ell, \hat{\ell})} \mathcal{D}_\omega^\theta f}] \partial_x [\widehat{f}] + [\partial_\gamma \widehat{\mathcal{D}_{(x, y, z, \gamma)}^{(\ell, \hat{\ell})} \mathcal{D}_\omega^\theta f}] \partial_{xx} [\widehat{f}]. \end{aligned}$$

The derivatives of f are up to the order $|(\theta, \ell)| + 2 \leq m' - 2 + 2 = m - 2$, and its path derivatives are up to the order $|\theta| + 2 \leq |\bar{\theta}'| - 2 + 2 = |\bar{\theta}| - 2$, then these terms are in the form of (7.14). Moreover, since $m \geq 4$, by (7.20) one can easily see that all the terms of \widehat{f} , $\partial_x [\widehat{f}]$, and $\partial_{xx} [\widehat{f}]$ are in the form of (7.14). Furthermore, similarly to the g -case above, one can show that $\lambda(\psi) \leq \lambda(\psi') + m'$.

Case 2. $\theta_1 = 1$. Then $m = m' + 1$ and $\partial_x^{\bar{\ell}} \partial_\omega^{\bar{\theta}} u = \partial_\omega (\partial_x^{\bar{\ell}} \partial_\omega^{\bar{\theta}'} u)$. By using (7.22) and following similar arguments as in Case 1 we can easily prove the result. ■

8. Consistency with [2]

In this section we compare our stochastic Taylor expansions (7.7) with those in our previous works [3,2] (in particular, the one in [2]), and consequently unify them under the language of our path-derivatives. To be consistent with [3,2], we assume in what follows that $d = 1$, $\mathcal{O} = \mathbb{R}^d$, and that the coefficients f and g in (7.1) are *deterministic*. We should note that in this case we have $A_{t+\delta}^t = 0$, and $\underline{B}_{t,t+\delta} = \frac{1}{2}(\omega_{t+\delta}^t)^2$.

We begin by recalling the definition of the “ n -fold derivatives” introduced in [2].

Definition 8.1. A random field $\zeta \in C^{0,n}(\mathbb{F}^B, [0, T] \times \mathbb{R}^{d'})$ is called “ n -fold” differentiable in the spatial variable x if there exist n random fields $\zeta_i \in C^{0,n}(\mathbb{F}, [0, T] \times \mathbb{R}^{d'}; \mathbb{R}^{d_i}), 2 \leq i \leq n + 1$, with $d_1 = d'$ and $d_i \in \mathbb{N}, 2 \leq i \leq n + 1$, and functions $F_i, G_i : [0, T] \times \mathbb{R}^{d'} \times \mathbb{R}^{d_{i+1}} \rightarrow \mathbb{R}^{d'}, i = 1, \dots, n$, such that, denoting $\zeta_1 := \zeta$, the following properties are satisfied:

- (1) $F_i, G_i \in C_{\ell,p}^\infty, i = 1, \dots, n$;
- (2) For $1 \leq i \leq n$, it holds that

$$\zeta_i(t, x) = \zeta_{i,0}(x) + \int_0^t F_i(s, x, \zeta_{i+1}(s, x))ds + \int_0^t G_i(s, x, \zeta_{i+1}(s, x))dB_s, \tag{8.1}$$

for all $(t, x) \in [0, T] \times \mathbb{R}^{d'}$, with $\zeta_{i,0} \in C^2(\mathbb{R}^{d'}; \mathbb{R}^{d_i}), 1 \leq i \leq n$.

- (3) For any $1 \leq i \leq n + 1, N \geq 1$, and $|\ell| \leq n$, it holds that

$$\sup\{\mathcal{D}_x^\ell \zeta_i(t, x), t \in [0, T], |x| \leq N\} \in \cap_{p>1} L^p(\Omega, \mathcal{F}, \mathbb{P}_0).$$

We shall call $\zeta_i, i = 2, \dots, n + 1$ the “generalized derivatives” of $\zeta = \zeta_1$, with “coefficients” $(F_i, G_i), i = 1, \dots, n$.

The notion of the n -fold derivatives is particularly motivated by the structure of SPDE (7.1), or more precisely, (7.8). In fact, if we define $\zeta_1 = u, \zeta_2 = (u, \partial_x u, \partial_{xx} u), F_1 = F$, and $G_1 = g$, then (8.1) holds for $i = 1$. Moreover, if we assume that the coefficients f, g are sufficiently smooth so that the solution u is 3-fold differentiable, then all the coefficients of the 3-fold generalized derivatives can be determined by differentiating the Eq. (7.8) (in x) repeatedly. Clearly, to compare the Taylor expansion we need only compare the derivatives in Definition 8.1 and the path-derivatives defined in Definition 2.1. For notational simplicity in what follows we shall assume $d' = 1$.

To begin with, we note that by Definition 8.1 we have $\zeta_1 = u, \zeta_2 = (u, \partial_x u, \partial_{xx} u), \zeta_3 = (u, \partial_x u, \partial_{xx} u, \partial_{xxx} u, \partial_{xxxx} u)$. Thus we have $(F_1, G_1) = (F, g)$ and the coefficients of the 2-fold derivative are $F_2 = (F_2^{(1)}, F_2^{(2)}, F_2^{(3)})$ and $G_2 = (G_2^{(1)}, G_2^{(2)}, G_2^{(3)})$, where $F_2^{(1)} = F, G_2^{(1)} = g$, and $(F_2^{(2)}, G_2^{(2)})$ can be determined by differentiating Eq. (7.8) with respect to x , namely,

$$\partial_x u(t, x) = u'_0(x) + \int_0^t [F_2^{(2)}(t, x, \zeta_3(t, x))]ds + \int_0^t G_2^{(2)}(t, x, \zeta_3(t, x))dB_s, \tag{8.2}$$

where, by direct calculation, it is readily seen that

$$G_2^{(2)}(t, x, \zeta_3) = \partial_x g + \partial_y g \partial_x u + \partial_z g \partial_{xx}^2 u. \tag{8.3}$$

Note that g is deterministic, we have $\partial_\omega g = 0$ and, by (8.2), (8.3), and the definition of path-derivative,

$$\partial_{\omega x} u = \partial_\omega(\partial_x u) = G_2^{(2)}(t, x, \zeta_3) = \partial_x g + \partial_y g \partial_x u + \partial_z g \partial_{xx}^2 u = \partial_x(\partial_\omega u). \tag{8.4}$$

Now, applying functional Itô’s formula (or “chain rule”) to $g(\cdot, u, \partial_x u)$ we have

$$dg(t, x, u, \partial_x u) = F^g(t, x, \zeta_3)dt + [\partial_y g \partial_\omega u + \partial_z g \partial_{\omega x} u]dB_s, \tag{8.5}$$

where we simply denote the drift by F^g as it is irrelevant to our argument. Note that, denoting $\mathbf{z} = (y, z, \gamma)$, we can write $G_1(t, x, \mathbf{z}) = G_2^{(1)}(t, x, \mathbf{z}) = g(t, x, \mathbf{z}) = g(t, x, y, z)$, and

$D_{\mathbf{z}}g = (g_y, g_z, 0)$. Therefore

$$\partial_y g \partial_\omega u + \partial_z g \partial_{\omega x} u = \partial_y g g + \partial_z g G_2^{(2)}(t, x, \zeta_3) = \langle D_{\mathbf{z}}g(t, x, \zeta_2), G_2(t, x, \zeta_3) \rangle. \tag{8.6}$$

This, together with the fact $\partial_\omega u = g$ and (8.5), shows that

$$\partial_{\omega\omega}^2 u = \partial_\omega g = \langle D_{\mathbf{z}}g, G_2(t, x, \zeta_3) \rangle; \tag{8.7}$$

$$\partial_t u = f(t, x, \zeta_2) - \frac{1}{2} \partial_{\omega\omega}^2 u = f(t, x, \zeta_2) - \frac{1}{2} \langle D_{\mathbf{z}}g, G_2(t, x, \zeta_3) \rangle. \tag{8.8}$$

We can now recast the pathwise Taylor expansion of Buckdahn–Bulla–Ma [2] in the new path-derivative language. Recall that $B_s^t(\omega) := \omega_s - \omega_t$, for $s \geq t$.

Theorem 8.2. *Let u be the classical solution to the SPDE (7.1) with deterministic coefficients f and g , and assume that it is 3-fold differentiable. Then, for every $\alpha \in (\frac{1}{3}, \frac{1}{2})$ and $m \in \mathbb{N}$, there exist a subset $\tilde{\Omega}_{\alpha,m} \subset \Omega$ with $P\{\tilde{\Omega}_{\alpha,m}\} = 1$, such that, for all $(t, y, \omega) \in [0, T] \times \bar{B}_m(0) \times \tilde{\Omega}_{\alpha,m}$, the following Taylor expansion holds:*

$$u(t+h, y+k) - u(t, y) = ah + bB_{t+h}^t + \frac{c}{2}(B_{t+h}^t)^2 + pk + \frac{1}{2}Xk^2 + qkB_{t+h}^t + (|h| + |k|^2)^{3\alpha} R_{\alpha,m}(t, t+h, y, y+k), \tag{8.9}$$

for all $(t+h, y+k) \in [0, T] \times \bar{B}_m(0)$. Here, the coefficients (a, b, c, p, q, X) are given by

$$\begin{cases} a = \partial_t u(t, x, \cdot), & b = \partial_\omega u(t, x, \cdot), & c = \partial_{\omega\omega} u(t, x, \cdot); \\ p = \partial_x u(t, x, \cdot), & q = \partial_{\omega x} u(t, x, \cdot), & X = \partial_{xx} u(t, x, \cdot). \end{cases} \tag{8.10}$$

Furthermore, the remainder of Taylor expansion $R_{\alpha,m} : [0, T]^2 \times (\mathbb{R}^d)^2 \times \Omega \mapsto \mathbb{R}$ is a measurable random field such that

$$\bar{R}_{\alpha,m} := \sup_{t,s \in [0,T]; y,z \in \bar{B}_m(0)} |R_{\alpha,m}(t, s, y, z)| \in \cap_{p>1} L^p(\Omega, \mathcal{F}, P). \tag{8.11}$$

Proof. Again, we assume $d' = 1$. Since u satisfies (7.1) and is 3-fold differentiable, by a direct application of Theorem 2.3 in [2], with $\zeta = \zeta_1 = u$, $\zeta_2 = (u, u_x, u_{xx})$, and $F_1 = F$, $G_1 = g$, we obtain a stochastic Taylor expansion (8.9) with the following coefficients

$$\begin{aligned} a &= f(t, x, \zeta_2) = F(t, x, \zeta_2) - \frac{1}{2} \langle D_{\mathbf{z}}g(t, x, \zeta_2), G_2(t, x, \zeta_3) \rangle, \\ b &= g(t, x, \zeta_2), & c &= \langle D_{\mathbf{z}}g(t, x, \zeta_2), G_2(t, x, \zeta_3) \rangle, \\ p &= \partial_x u(t, x), & X &= \partial_{xx} u(t, x), \\ q &= \partial_x g(t, x, \zeta_2) + \langle D_{\mathbf{z}}g(t, x, \zeta_2), D_x \zeta_2 \rangle. \end{aligned} \tag{8.12}$$

Combining (7.3), (8.4), (8.7) and (8.8) we have

$$\begin{aligned} a &= f(t, x, \zeta_2) = F(t, x, \zeta_2) - \frac{1}{2} \langle D_{\mathbf{z}}g(t, x, \zeta_2), G_2(t, x, \zeta_3) \rangle = \partial_t u; \\ b &= g(t, x, \zeta_2) = \partial_\omega u; \\ c &= \langle D_{\mathbf{z}}g(t, x, \zeta_2), G_2(t, x, \zeta_3) \rangle = \partial_{\omega\omega}^2 u; \\ q &= \partial_x g(t, x, \zeta_2) + \langle D_{\mathbf{z}}g(t, x, \zeta_2), D_x \zeta_2 \rangle = \partial_x g + \partial_y g \partial_x u + \partial_z g \partial_{xx}^2 u = \partial_{x\omega}^2 u. \end{aligned}$$

This proves (8.10), whence the theorem. ■

Acknowledgments

The second author is supported in part by US NSF grant #1106853. The third author is supported in part by NSF grant #1008873.

References

- [1] K. Bichteler, Stochastic integration and L^p -theory of semi-martingales, *Ann. Probab.* 9 (1981) 48–89.
- [2] R. Buckdahn, I. Bulla, J. Ma, Pathwise Taylor expansions for Itô random fields, *Math. Control Relat. Fields* 1 (4) (2011) 437–468.
- [3] R. Buckdahn, J. Ma, Pathwise stochastic Taylor expansions and stochastic viscosity solutions for fully nonlinear stochastic PDEs, *Ann. Probab.* 30 (3) (2002) 1131–1171.
- [4] R. Buckdahn, J. Ma, J. Zhang, Pathwise viscosity solutions for stochastic PDEs and forward path dependent PDEs, Preprint. arXiv:1501.06978.
- [5] R. Cont, D. Fournie, Functional Itô calculus and stochastic integral representation of martingales, *Ann. Probab.* 41 (1) (2013) 109–133.
- [6] B. Dupire, Functional Itô calculus, papers.ssrn.com.
- [7] I. Ekren, N. Touzi, J. Zhang, Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part I, *Ann. Probab.* (2015) in press. arXiv:1210.0006.
- [8] P. Friz, M. Hairer, *A Course on Rough Paths: with an Introduction to Regularity Structures*, Springer Universitext, 2014.
- [9] P. Friz, N. Victoir, *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, Cambridge University Press, 2010.
- [10] D. Gilbarg, N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, second ed., Springer, 1998.
- [11] R. Karandikar, On pathwise stochastic integration, *Stochastic Process. Appl.* 57 (1995) 11–18.
- [12] P.E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, Heidelberg, 1992.
- [13] N.V. Krylov, An analytic approach to SPDEs, in: *Stochastic Partial Differential Equations: Six Perspectives*, in: *Math. Surveys Monogr.*, vol. 64, Amer. Math. Soc., Providence, RI, 1999, pp. 185–242.
- [14] C. Litterer, H. Oberhauser, A Chen-Fliess approximation for diffusion functionals, Preprint. arXiv:1110.2481.
- [15] T. Lyons, M. Caruana, T. Lévy, *Differential Equations Driven by Rough Paths*. Lectures from the 34th Summer School on Probability Theory Held in Saint-Flour, July 6–24, 2004, in: *Lecture Notes in Mathematics*, vol. 1908, Springer, Berlin, 2007.
- [16] J. Ma, J. Yong, On linear backward stochastic partial differential equations, *Probab. Theory Related Fields* 113 (1999) 135–170.
- [17] B.L. Rozovskii, *Stochastic Evolution Systems: Linear Theory and Applications to Non-linear Filtering*, Kluwer Academic Publishers, Boston, 1990.