YAO-ZHONG HU

Strong and weak order of time discretization schemes of stochastic differential equations

Séminaire de probabilités (Strasbourg), tome 30 (1996), p. 218-227
<http://www.numdam.org/item?id=SPS_1996__30__218_0>

STRONG AND WEAK ORDER OF TIME DISCRETIZATION SCHEMES OF STOCHASTIC DIFFERENTIAL EQUATIONS

YAOZHONG HU*

This note is taken from lectures at Oslo University based on the book [KP], *Numerical Solutions of Stochastic Differential Equations* by Kloeden and Platen. We will give a condensed presentation of *time discretization schemes, strong order estimation* and *weak order estimation.* Besides the interest of the subject itself, we would like to give a much simpler proof of the weak estimation scheme.

Thanks are due to Prof. B. Øksendal for bringing him to this subject, Profs. M. Emery, P.A. Meyer, P.E. Platen and Marta Sanz for comments and Profs. M. Emery, P.A. Meyer and Marta Sanz for editorial help.

1. General Ideas. Let $B^1,\ldots,B^m$ be m standard (real) independent Brownian motions on some time interval $[0,T]$ (bounded, and kept fixed below). Let $(\Omega,\mathcal{F},P)$ be the canonical Wiener space with the natural filtration $(\mathcal{F}_t)_{0\le t\le T}$. On $\mathbb{R}^d$ consider the following stochastic differential equation in Ito's sense

$$X_t = x + \sum_{j=0}^{m}\int_0^t b_j(s,X_s)\,dB_s^j\ ,\ t\in[0,T]\ ,\ x\in\mathbb{R}^d\ , \tag{1.1}$$

where b_j are some given regular functions from $[0,T]\times\mathbb{R}^d$ to $\mathbb{R}^d$, and we use the convention $dB_s^0 = ds$ to simplify notation.

A time discretization method consists in dividing the interval $[0,T]$ into smaller subintervals, applying the Itô–Taylor formula (to be described later) on each subinterval, keeping a given number of terms, and piecing out these approximations to get an approximate solution. We then expect these approximations will converge to the true solution when the subintervals become finer and finer.

To describe the Itô-Taylor formula, we introduce the following operators on functions $h:[0,T]\times\mathbb{R}^d\to\mathbb{R}$

$$L^j h(s,x) = \sum_{k=1}^{d} b_j^k(s,x)\,\frac{\partial h}{\partial x^k}(s,x)\ ,\quad j=1,\cdots,m\ , \tag{1.2}$$

$$L^0 h(s,x) = \frac{\partial h}{\partial s}(s,x) + \sum_{k=1}^{d} b_0^k(s,x)\,\frac{\partial h}{\partial x^k}(s,x)$$

---

* Supported by an NAVF postdoctorship, Department of Mathematics, University of Oslo, POB 1053, Blindern, N-0316 Oslo. Institute of Mathematical Sciences, Academia Sinica, WuHan, China.

$$(1.3) \qquad + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b_j^k(s,x)\, b_j^l(s,x)\, \frac{\partial^2 h}{\partial x^k \partial x^l}(s,x),$$

where b_j^k is the k-th component of the vector b_j $(k = 1, \ldots, d)$.

Consider a partition of the interval $[0, T]$, $0 = t_0 < t_1 < \cdots < t_N = T$ and put $\delta = \sup_i (t_{i+1} - t_i)$, the *step* of the partition. On each subinterval $[t_n, t_{n+1}]$ we may write (1.1) as

$$(1.4) \qquad X_t = X_{t_n} + \sum_{j=0}^{m} \int_{t_n}^{t} b_j(s, X_s)\, dB_s^j \, .$$

For a sufficiently differentiable function $h : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, an application of the Itô formula to $h(t, X_t)$ gives

$$(1.5) \qquad h(t, X_t) = h(t_n, X_{t_n}) + \sum_{j=0}^{m} \int_{t_n}^{t} L^j h(s, X_s)\, dB_s^j \, .$$

This is the first order Itô-Taylor formula. To define higher order Itô-Taylor formulas we introduce the following notation

$$\alpha = (\alpha_1, \cdots, \alpha_l) \quad (0 \le \alpha_i \le m) \quad \text{(a multi-index)} \quad ,$$

$$l(\alpha) = l, \quad n(\alpha) = \text{the number of zeroes among } \alpha_1, \ldots \alpha_l \ ,$$

$$(1.6) \qquad f_\alpha^k = L^{\alpha_1} \cdots L^{\alpha_{l-1}} b_{\alpha_l}^k \quad (k = 1, \cdots, d), \quad f_\alpha = (f_\alpha^1, \cdots, f_\alpha^k),$$

$$(1.7) \qquad I_\alpha[g(\cdot)]_{s,t} = \int_{s<s_1<\cdots<s_l<t} g(s_1)\, dB_{s_1}^{\alpha_1} \cdots dB_{s_l}^{\alpha_l} \ ,$$

where $g(\cdot)$ is an adapted process. We put simply $I_{\alpha,s,t} = I_\alpha[1]_{s,t}$ in the case $g(\cdot) = 1$. These are the standard multiple integrals (including dt). They replace the monomials in the classical Taylor expansion.

Now the general scheme for Itô-Taylor formulas is the following : In formula (1.4), we apply the Itô formula (1.5) to *some* processes $b_j(t, X_t)$ — usually to all, but the coefficient $b_0(t, X_t)$ may play a special role. Then in the new formula we apply again (1.5) to *some* coefficients of the stochastic integrals, etc. Then we get a general formula with the following structure of a main term plus a remainder

(1.8)

$$X_t = X_{t_n} + \sum_{\alpha \in \Gamma} f_\alpha(t_n, X_{t_n}) I_{\alpha, t_n, t} + \sum_{\alpha \in \Gamma'} I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n, t} \ , \quad (t \in [t_n, t_{n+1}]) \ .$$

The "main term" is a sum over a finite set Γ of multi-indices, which has the following property : if $\alpha = (\alpha_1, \cdots, \alpha_l) \in \Gamma$, then $-\alpha := (\alpha_2, \cdots, \alpha_l) \in \Gamma$. On the other hand, Γ' is the set $\{\alpha : \alpha \notin \Gamma, -\alpha \in \Gamma\}$. This structure comes from the fact that each term is obtained by applying (1.5) to a preceding term.

Now a so called discretization scheme is obtained by discarding the remainder in (1.8). Since X_{t_n} is not known in the recursive computation, we replace it by its approximation Y_{t_n} (Throughout this paper we will omit its explicit dependence on partition π to simplify notation), and then we get the following approximation scheme, starting at $Y_0 = x$

$$Y_t = Y_{t_n} + \sum_{\alpha\in\Gamma} f_\alpha(t_n, Y_{t_n})\, I_{\alpha,t_n,t}\ , \quad (t \in [t_n, t_{n+1}])\ . \tag{1.9}$$

This recursive formula lends itself to explicit computations (the multiple integrals can be even handled by a computer). In practice we have to choose which terms are included in Γ. The concrete choices for strong and weak convergence scheme are different. See the details below.

2. Strong approximation scheme. We give ourselves a parameter (denoted γ in [KP]) which is called the *strong order* of approximation, and which is an integer or half-integer. Our purpose is to have a norm estimate like (2.1) below. Then the number of terms to take in the Itô-Taylor formula, *i.e.* the choice of Γ, is

$$\Gamma = \mathcal{A}_\gamma = \left\{\alpha : l(\alpha) + n(\alpha) \le 2\gamma \quad \text{or} \quad l(\alpha) = n(\alpha) = \gamma + \tfrac{1}{2}\right\}\ .$$

Then we have the following theorem — denoting by C as usual some constant whose precise value doesn't interest us, and may change from line to line. It may depend on several parameters (T, γ, ...) but never on the partition (t_i).

Theorem 1. *Let γ be defined as above and let Y_t be defined by (1.9). Assume that for $\alpha \in \mathcal{A}_\gamma$ the coefficients f_α defined by (1.6) satisfy Lipschitz conditions*

$$|\, f_\alpha(t,x) - f_\alpha(t,y)\,| \le C\,|x-y|$$

and

$$|\, f_\alpha(t,x)\,| \le C(1+|x|)$$

for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$. Then we have

$$\mathbb{E}\Big(\sup_{0\le t\le T} |X_t - Y_t|^2\Big) \le C\delta^{2\gamma}\ . \tag{2.1}$$

To prove this theorem we need two lemmas.

Lemma 1. *Let $g(s)$ be an adapted process. When $l(\alpha) = n(\alpha), t \in [t_n, t_{n+1}]$,*

$$\sup_{t_n\le s\le t} |I_\alpha[g(\cdot)]_{t_n,s}| \le C(t-t_n)^{l(\alpha)-1/2}\Big(\int_{t_n}^t |g(u)\,|^2 du\Big)^{1/2} \qquad \text{a.e.} \tag{2.2}$$

and when $l(\alpha) \ne n(\alpha)$,

$$\mathbb{E} \sup_{t_n\le s\le t} (I_\alpha[g(\cdot)]_{t_n,s})^2 \le C(t-t_n)^{l(\alpha)+n(\alpha)-1} \int_{t_n}^t \mathbb{E}\,|g(u)\,|^2 du\ . \tag{2.3}$$

Here C may depend on α, but since α ranges over a finite set $\mathcal{A}_\gamma$ this dependence is not important.

Proof. 1) By the definition (1.7), $l(\alpha) = n(\alpha)$ means that there is no stochastic integral in $I_\alpha[g(\cdot)]_{t_n,s}$. Formula (2.2) is obvious.

2) When $l(\alpha) \neq n(\alpha)$, we prove (2.3) by induction on the length $l(\alpha)$ of α. We define $\alpha- := (\alpha_1, \cdots, \alpha_l)$ if $\alpha = (\alpha_1, \cdots, \alpha_l, \alpha_{l+1})$. The case $l(\alpha) = 1$ is easy by discussing $n(\alpha) = 0$ and $n(\alpha) \neq 0$ separately. For the passage from $\alpha- = (\alpha_1, \cdots, \alpha_l)$ to $\alpha = (\alpha_1, \cdots, \alpha_l, \alpha_{l+1})$ we also handle $\alpha_{l+1} = 0$ and $\alpha_{l+1} \neq 0$ differently. In the first case $I_\alpha[g(\cdot)]_{t_n,s} = \int_{t_n}^s I_{\alpha-}[g(\cdot)]_{t_n,u}du$. Thus by Hölder's inequality

$$
\begin{aligned}
\mathbb{E} \sup_{t_n \le s \le t} (I_\alpha[g(\cdot)]_{t_n,s})^2 &\le C(t-t_n) \int_{t_n}^t \mathbb{E}|I_{\alpha-}[g(\cdot)]_{t_n,u}|^2 du \\
&\overset{\text{by induction}}{\le} C(t-t_n)^{l(\alpha-)+n(\alpha-)} \int_{t_n}^t \int_{t_n}^s \mathbb{E}|g(u)|^2 duds \\
&\le C(t-t_n)^{l(\alpha-)+n(\alpha-)+1} \int_{t_n}^t \mathbb{E}|g(u)|^2 du . \qquad (2.4)
\end{aligned}
$$

But in the case $\alpha_{l+1} = 0$, $l(\alpha) = l(\alpha-) + 1$, $n(\alpha) = n(\alpha-) + 1$. So $l(\alpha-) + n(\alpha-) + 1 = l(\alpha) + n(\alpha) - 1$. This shows (2.3) in this case.

When $\alpha_{l+1} \neq 0$, $I_\alpha[g(\cdot)]_{t_n,s} = \int_{t_n}^s I_{\alpha-}[g(\cdot)]_{t_n,u}dB_s^{\alpha_{l+1}}$. By Doob's inequality, we have

$$
\begin{aligned}
\mathbb{E} \sup_{t_n \le s \le t} (I_\alpha[g(\cdot)]_{t_n,s})^2 &\le C\mathbb{E}|I_\alpha[g(\cdot)]_{t_n,t}|^2 \le C \int_{t_n}^t \mathbb{E}|I_{\alpha-}[g(\cdot)]_{t_n,u}|^2 du \\
&\le C \int_{t_n}^t (u-t_n)^{l(\alpha-)+n(\alpha-)-1} \int_{t_n}^u \mathbb{E}|g(v)|^2 dvdu \\
&\le C(t-t_n)^{l(\alpha-)+n(\alpha-)} \int_{t_n}^t \mathbb{E}|g(u)|^2 du . \qquad (2.5)
\end{aligned}
$$

But in this case $l(\alpha) = l(\alpha-) + 1$ and $n(\alpha) = n(\alpha-)$. Thus $l(\alpha-) + n(\alpha-) = l(\alpha) + n(\alpha) - 1$. This proves (2.3) in the case $\alpha_{l+1} \neq 0$. ∎

Lemma 2. *Let $g(s)$ be an adapted process. Put $n(s) = n$ if $t_n \le s < t_{n+1}$ and*

$$
F_t^\alpha := \mathbb{E}\Big(\sup_{0 \le s \le t} \Big| \sum_{m=0}^{n(s)-1} I_\alpha[g(\cdot)]_{t_m,t_{m+1}} + I_\alpha[g(\cdot)]_{t_{n(s)},s} \Big|^2 \Big) . \qquad (2.6)
$$

Then

$$
F_t^\alpha \le \begin{cases} C\delta^{2(l(\alpha)-1)} \int_0^t R_{0,u}du & l(\alpha) = n(\alpha) \\ C\delta^{(l(\alpha)+n(\alpha)-1)} \int_0^t R_{0,u}du & l(\alpha) \neq n(\alpha), \end{cases} \qquad (2.7)
$$

where

$$R_{0,u} := \mathbb{E}\Big(\sup_{0\le s\le u} |g(s)|^2\Big) \le \infty\,. \tag{2.8}$$

Again C is independent of the subdivision but may depend on α.

Proof. The case $l(\alpha) = n(\alpha)$ is easy. We only need to discuss the case $l(\alpha) \neq n(\alpha)$. When the last index of α isn't equal to 0, in the sum of (2.6) the multiple integral is a martingale. By Doob's inequality, we have that

$$\begin{aligned} F_t^\alpha &\le C\Big(\mathbb{E}\sum_{m=0}^{m(t)-1} I_\alpha[g(\cdot)]_{t_m,t_{m+1}} + I_\alpha[g(\cdot)]_{t_{n(t)},t}\Big)^2 \\ &= C\Big(\sum_{m=0}^{n(t)-1}\Big[\mathbb{E}\,|I_\alpha[g(\cdot)]_{t_m,t_{m+1}}|^2\Big] + \mathbb{E}|I_\alpha[g(\cdot)]_{t_{n(t)},t}|^2\Big)\,. \end{aligned}$$

Estimating the second moment of each of the above multiple integrals on each interval by Lemma 1 (2.3) and then taking the sum we will get the desired inequality (2.7).

If the last index of α is 0 we have

$$|F_t^\alpha|^2 \le 2C\mathbb{E}\Big(\sup_{0\le s\le t}\sum_{n=0}^{n(s)-1} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\Big)^2 + 2C\mathbb{E}\sup_{0\le s\le t}|I_\alpha[g(\cdot)]_{t_{n_s},s}|^2.$$

We estimate separately these two terms. The first one is

$$2C\mathbb{E}\Big(\sup_{0\le r\le n(t)-1}\sum_{n=0}^{r} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}\Big)^2.$$

Now $\sum_{n=0}^{r} I_\alpha[g(\cdot)]_{t_n,t_{n+1}}$, $r = 0,\cdots,n(t)-1$ can be considered as a discrete martingale and we can then use Doob's inequality to complete the proof as in the case the last index isn't equal to 0. The second term can be handled as follows (see Kloeden and Platen's book, p.370) :

$$\begin{aligned} \mathbb{E}\sup_{0\le s\le t}|I_\alpha[g(\cdot)]_{t_s,s}|^2 &= \mathbb{E}\sup_{0\le s\le t}\int_{t_s}^{s}|I_{\alpha-}[g(\cdot)]_{t_s,u}|^2du \\ &= E\sup_{0\le s\le t}(s-n_s)\int_{t_s}^{s}|I_{\alpha-}[g(\cdot)]_{t_s,u}|^2du \\ &\le \delta\int_0^t \mathbb{E}|I_{\alpha-}[g(\cdot)]_{t_s,s}|^2ds. \end{aligned}$$

Applying Lemma 1 to estimate $\mathbb{E}|\,I_{\alpha-}[g(\cdot)]_{t_s,s}\,|^2$, we get the result. ∎

Remark. A L_p version of lemma 2 is proved in [HW], Lemma 4.1.

Proof of Theorem 1. From (1.8) and (1.9) we have

$$Z(t) = \mathbb{E}\Big(\sup_{0\le s\le t} |X_s - Y_s|^2 \Big) \le C\Big(\sum_{\alpha\in\mathcal{A}_\gamma} R_t^\alpha + \sum_{\alpha\in\mathcal{A}'_\gamma} U_t^\alpha \Big) , \tag{2.9}$$

where R_t^α and U_t^α are defined and estimated as follows :

$$\begin{aligned} R_t^\alpha := \mathbb{E}\Big(\sup_{0\le s\le t} \Big| & \sum_{m=0}^{n(s)-1} I_\alpha [f_\alpha(t_m, X_{t_m}) - f_\alpha(t_m, Y_{t_m})]_{t_m,t_{m+1}} \\ & + I_\alpha [f_\alpha(t_{n(s)}, X_{n(s)}) - f_\alpha(t_{n(s)}, Y_{t_{n(s)}})]_{t_{n(s)},s} \Big|^2 \Big) \\ \overset{\text{Lemma 2}}{\le} C \int_0^t \mathbb{E} \sup_{0\le s\le u} & | f_\alpha(t_{n(s)}, X_{n(s)}) - f_\alpha(t_{n(s)}, Y_{t_{n(s)}}) |^2 du \le C \int_0^t Z(u) du . \end{aligned} \tag{2.10}$$

$$\begin{aligned} U_t^\alpha &:= \mathbb{E}\Big(\sup_{0\le s\le t} \Big| \sum_{m=0}^{n(s)-1} I_\alpha [f_\alpha(\cdot, X_\cdot)]_{t_m,t_{m+1}} + I_\alpha [f_\alpha(\cdot, X_\cdot)]_{t_{n(s)},s} \Big|^2 \Big) \\ &\overset{\text{Lemma 2}}{\le} C\delta^{\varphi(\alpha)}, \end{aligned} \tag{2.11}$$

where $\alpha \in \mathcal{A}'_\gamma$ and

$$\varphi(\alpha) = \begin{cases} 2(l(\alpha)-1) & : \quad l(\alpha) = n(\alpha) \\ l(\alpha) + n(\alpha) - 1 : & \quad l(\alpha) \ne n(\alpha) \end{cases} . \tag{2.12}$$

Since $\alpha \in \mathcal{A}'_\gamma$ which implies $2(l(\alpha)-1) \ge 2\gamma$ if $l(\alpha) = n(\alpha)$ and $l(\alpha)+n(\alpha)-1 \ge 2\gamma$ if $l(\alpha) \ne n(\alpha)$, we have

$$U_t^\alpha \le C\delta^{2\gamma}. \tag{2.13}$$

From (2.9), (2.10) and (2.13) we have

$$Z(t) \le C \int_0^t Z(u)\, du + C\delta^{2\gamma} .$$

Then we deduce the theorem from Gronwall's inequality. ∎

Remark. Let us return to Lemma 1, Formula (2.3). In next section we will need the following extension to higher moments : If $\mathbb{E} \sup_{0\le t\le T} |g(t)|^p < \infty$, then

$$\mathbb{E} \sup_{t_n\le s\le t} (I_\alpha [g(\cdot)]_{t_n,s})^p \le C(t-t_n)^{p[l(\alpha)+n(\alpha)]/2} . \tag{2.14}$$

The easy proof is left to the reader.

3. Weak Itô-Taylor scheme. Now we want to treat the *weak convergence rate problem*, that is to say, to estimate $|\mathbb{E}[h(X_T) - h(Y_T)]|$ for a continuous function h of polynomial growth. Note that the absolute value sign is outside the expectation. We could also estimate $\sup_{0\le t\le T}|\mathbb{E}[h(X_t) - h(Y_t)]|$ without essential modifications. To get a weak convergence rate of order γ (here, an integer), we take from now on

$$\Gamma = \mathcal{B}_\gamma = \{\alpha, l(\alpha) \le \gamma\} .$$

Then we put $\mathcal{B}'_\gamma = \{\alpha : \alpha \notin \mathcal{B}_\gamma, -\alpha \in \mathcal{B}_\gamma\} = \{\alpha : l(\alpha) = \gamma + 1\}$.

We have the following theorem.

THEOREM 2. *Let γ be an positive integer. Assume that all coefficients b_j ($j = 0, 1, \dots, m$) are Lipschitz continuous and their components belong to $C_p^{2(\gamma+1)}$ (the space of functions from $\mathbb{R}^d$ to $\mathbb{R}$ whose derivatives of order $\le 2(\gamma + 1)$ are continuous and of polynomial growth). Assume that for all $\alpha \in \Gamma$, $f_\alpha = L^{\alpha_1} \cdots L^{\alpha_{l-1}} b_{\alpha_l}$, define by (1.6), is of linear growth :*

$$|f_\alpha(x)| \le C(1 + |x|) .$$

Then for each $h \in C_p^{2(\gamma+1)}$ there is a constant C_h independent of δ such that

$$|\mathbb{E}[h(X_T) - h(Y_T)]| \le C_h\, \delta^\gamma . \tag{3.1}$$

We need two lemmas to prove this theorem. We introduce the following notation : $X_\cdot^{s,x}$ is the solution of the s.d.e.

$$X_t^{s,x} = x + \sum_{j=0}^{m} \int_s^t b_j(X_r^{s,x})\, dB_r^j , \quad s \le t \le T \tag{3.2}$$

and put $\widetilde{X}_s = X_s^{t_n, Y_{t_n}}$ for $t_n \le s \le t_{n+1}$.

LEMMA 3. Let $g(\cdot)$ be adapted and $\sup_{0\le s\le T} \mathbb{E}|g(s)|^p < \infty$ for any $1 \le p < \infty$. Then for $t_n \le t \le t_{n+1}$,

$$|\mathbb{E}\{I_\alpha[g(\cdot)]_{t_n,t}|\mathcal{F}_{t_n}\}| \le C(\omega)(t - t_n)^{l(\alpha)}, \tag{3.3}$$

where and in what follows $C(\omega)$ denotes a positive generic random constant independent of partition (which may vary from line to line) such that $\mathbb{E}C(\omega)^p < \infty$ for any $1 \le p < \infty$.

PROOF. Easy. ∎

LEMMA 4. Under the assumption of Lemma 3, we have for $t_n \le t \le t_{n+1}$,

$$|M_\alpha| := |\mathbb{E}\{[h(\tilde{X}_t) - h(Y_{t_n})] I_\alpha[g(\cdot)]_{t_n,t}|\mathcal{F}_{t_n}\}| \le C(\omega)(t - t_n)^{l(\alpha)}. \tag{3.4}$$

Proof. We shall prove this lemma by induction on $l(\alpha)$. It is easy to see that (3.4) is true when $l(\alpha) = 1$. Let (3.4) be true for $l(\alpha) \le k$. We are going to prove that it is true for $l(\alpha) = k+1$.

Let $l(\alpha) = k+1$. Applying the Itô formula (1.5), we have

$$h(\tilde{X}_t) - h(Y_{t_n}) = \sum_{j=0}^{m} \int_{t_n}^{t} L^j h(\tilde{X}_s) dB_s^j, \quad t_n \le t \le t_{n+1}.$$

When $\alpha_l = 0$ ($l(\alpha) = k+1$), by the Itô formula (1.5), Lemma 3 and the induction assumption we have

$$\begin{aligned}
|M_\alpha| &\le \int_{t_n}^{t} |\mathbb{E}\{[h(\tilde{X}_s) - h(Y_{t_n})] I_{\alpha-}[g(\cdot)]_{t_n,s} | \mathcal{F}_{t_n}\}| ds \\
&\quad + \int_{t_n}^{t} |\mathbb{E}\{[L^0 h(\tilde{X}_s) - L^0 h(Y_{t_n})] I_\alpha[g(\cdot)]_{t_n,s} | \mathcal{F}_{t_n}\}| ds \\
&\quad + \int_{t_n}^{t} \mathbb{E}|L^0 h(Y_{t_n})| |\mathbb{E}\{I_\alpha[g(\cdot)]_{t_n,s} | \mathcal{F}_{t_n}\}| ds \\
&\le C(\omega)(t-t_n)^{k+1} + \int_{t_n}^{t} |\mathbb{E}\{[L^0 h(\tilde{X}_s) - L^0 h(Y_{t_n}] I_\alpha[g(\cdot)]_{t_n,s} | \mathcal{F}_{t_n}\}| ds. \qquad (3.5)
\end{aligned}$$

Applying (3.5) repeatedly, we obtain

$$\begin{aligned}
|M_\alpha| \le C(\omega)(t-t_n)^{k+1} + \int_{t_n \le s_1 < \cdots < s_{k+1} < t} &|\mathbb{E}\{[(L^0)^{k+1} h(\tilde{X}_{s_1}) \\
&- (L^0)^{k+1} h(Y_{t_n})] I_\alpha[g(\cdot)]_{t_n,s_1} | \mathcal{F}_{t_n}\}| ds_1 \cdots ds_{k+1}.
\end{aligned}$$

Now it is easy to see that the conditional expectation inside the above multiple integral is in L^p for any $1 \le p < \infty$. This proves (3.4) for $\alpha_l = 0$. In the same way we can prove (3.4) for $\alpha_l \ne 0$ ∎

Proof of Theorem 2. Set $u(s,x) = \mathbb{E}[h(X_T^{s,x})]$ (see (3.2)). Then for $h \in C_p^{(2\gamma+1)}$ we have $u(s,\cdot) \in C_p^{(2\gamma+1)}$ (this can be shown easily by Malliavin calculus for example). We have $\mathbb{E}[h(X_T)] = u(0, X_0)$ and

$$\mathbb{E}u(t_n, X_{t_n}^{t_{n-1}, Y_{t_{n-1}}}) = \mathbb{E}u(t_{n-1}, Y_{t_{n-1}}), \quad n \ge 1 .$$

We compute the following expectation

$$\begin{aligned}
&|\,\mathbb{E}[\,h(Y_T) - h(X_T)\,]\,| = |\,\mathbb{E}[\,u(T, Y_T) - u(0, X_0)\,]\,| = |\,\mathbb{E}[\,u(T, Y_T) - u(0, Y_0)\,]\,| \\
&\le |\,\mathbb{E}\sum_{n=1}^{N} [u(t_n, Y_{t_n}) - u(t_{n-1}, Y_{t_{n-1}})\,]| = |\,\mathbb{E}\sum_{n=1}^{N} [u(t_n, Y_{t_n}) - u(t_n, X_{t_n}^{t_{n-1} Y_{n-1}})\,]\,|
\end{aligned}$$

$$(3.6)\qquad = \Big| \mathbb{E}\sum_{n=1}^{N} u'(t_n, X_{t_n}^{t_{n-1}Y_{n-1}})\big(Y_{t_n} - X_{t_n}^{t_{n-1}Y_{n-1}}\big) + \frac{1}{2}\sum_{n=1}^{N} \mathbb{E}u''(t_n, Z_{\theta,n})\big(X_{t_n}^{t_{n-1}Y_{n-1}} - Y_{t_n}\big)^2 \Big|,$$

where u' and u'' are derivatives of $u(s,x)$ w.r.t. x, $Z_{\theta,n} := X_{t_n}^{t_{n-1}Y_{n-1}} + \theta(Y_{t_n} - X_{t_n}^{t_{n-1}Y_{n-1}})$ and $0 \le \theta \le 1$. By (2.14) we know that the last term is dominated by the sum of

$$C\big(\mathbb{E}|u''(Z_{\theta,n})|^2\big)^{1/2}\big(\mathbb{E}(X_{t_n}^{t_{n-1}Y_{n-1}} - Y_{t_n})^4\big)^{1/2} \le C\sum_{\alpha\in\mathcal{B}'_\gamma}\big(\mathbb{E}|\, I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_n,t_{n+1}}|^4$$

$$\le C\sum_{\alpha\in\mathcal{B}'_\gamma}\Big((t_{n+1}-t_n)^{2(l(\alpha)+n(\alpha))}\Big)^{1/2} \le C\sum_{\alpha\in\mathcal{B}'_\gamma}(t_{n+1}-t_n)^{l(\alpha)}.$$

But when $\alpha \in \mathcal{B}'_\gamma$ we have $l(\alpha) \ge \gamma + 1$, so the last term of (3.6) is at most $C(t_{n+1}) - t_n)^\gamma$.

As for the first term of (3.6), first we note that by the assumptions of Theorem 2, we have $\sup_{0\le t\le T}|f_\alpha(t, X_t)|^p < \infty$ for any $1 \le p < \infty$ and $\alpha \in \mathcal{B}'_\gamma$. We have

$$\Big| \mathbb{E}\sum_{n=1}^{N} u'(t_n, X_{t_n}^{t_{n-1}Y_{n-1}})\big(Y_{t_n} - X_{t_n}^{t_{n-1}Y_{n-1}}\big)\Big|$$

$$(3.7)\qquad \le \Big| \mathbb{E}\sum_{n=1}^{N} u'(t_n, Y_{t_{n-1}})\,\big(Y_{t_n} - X_{t_n}^{t_{n-1}Y_{n-1}}\big)\Big| + \Big| \mathbb{E}\sum_{n=1}^{N}\big[u'(t_n, X_{t_n}^{t_{n-1}Y_{n-1}}) - u'(t_n, Y_{t_{n-1}})\big]\,\big(Y_{t_n} - X_{t_n}^{t_{n-1}Y_{n-1}}\big)\Big|.$$

By Lemma 4, the second term of (3.7) is dominated by

$$\sum_{\alpha\in\mathcal{B}'_\gamma}\sum_{n=1}^{N}\Big|\mathbb{E}\Big\{\mathbb{E}\Big(\big[u'(t_n, X_{t_n}^{t_{n-1},Y_{n-1}}) - u'(t_n, Y_{n-1})\big]I_\alpha[f_\alpha(\cdot, X_\cdot)]_{t_{n-1},t_n}|\mathcal{F}_{t_{n-1}}\Big)\Big\}\Big|$$

$$\le \sum_{\alpha\in\mathcal{B}'_\gamma}\sum_{n=1}^{N}\mathbb{E}[C(\omega)](t_n - t_{n-1})^{l(\alpha)} \le C\sum_{n=1}^{N}(t_n - t_{n-1})^{\gamma+1} \le C\delta^\gamma.$$

This gives the necessary estimate for the second term of (3.7). It is easy to see that the first term of (3.7) is also dominated by $C\delta^\gamma$. This proves the thoerem. ∎

REFERENCES

[Be] G. BEN AROUS, Flots et séries de Taylor stochastiques, *Prob. Th. Rel. Fields,* 81 (1989), 29-77.

[Hu] Y.Z. HU, Séries de Taylor stochastiques et formule de Campbell-Hausdorff, d'après Ben Arous, *Sem. Prob. XXVI,* Lect. notes in Math. 1526, Springer, 1992, 587-594.

[HW] Y.Z. HU and S. WATANABE, Donsker's delta functions and approximation of heat kernels by time discretization method, preprint, 1995.

[KP] P. E. KLOEDEN and E. PLATEN, *Numerical Solutions of Stochastic Differential Equations,* Springer-Verlag, 1992.

[Me] P.A. MEYER, Sur deux estimations d'intégrales multiples, *Sem. Prob. XXV,* Lecture Notes in Mathematics 1458, Springer 1991, 425-426.

[Øk] B. ØKSENDAL, *Stochastic Differential Equations,* Springer, 1985.