Tohoku Math. J.
70 (2018), 65–95

THE RATES OF THE L^p-CONVERGENCE OF THE EULER-MARUYAMA AND WONG-ZAKAI APPROXIMATIONS OF PATH-DEPENDENT STOCHASTIC DIFFERENTIAL EQUATIONS UNDER THE LIPSCHITZ CONDITION

SHIGEKI AIDA, TAKANORI KIKUCHI AND SEIICHIRO KUSUOKA

(Received January 5, 2015, revised December 24, 2015)

Abstract. We consider the rates of the L^p-convergence of the Euler-Maruyama and Wong-Zakai approximations of path-dependent stochastic differential equations under the Lipschitz condition on the coefficients. By a transformation, the stochastic differential equations of Markovian type with reflecting boundary condition on sufficiently good domains are to be associated with the equations concerned in the present paper. The obtained rates of the L^p-convergence are the same as those in the case of the stochastic differential equations of Markovian type without boundaries.

1. Introduction. Solutions to path-dependent stochastic differential equations are well-defined (see e.g. Chapter IV of [9]), and the existence and uniqueness of solutions hold under the Lipschitz condition on the coefficients (see e.g. Theorem 7 of Chapter V in [12]). In the present paper, we consider the rates of the L^p-convergence of the Euler-Maruyama and Wong-Zakai approximations of such an equation, and will obtain error estimates of the approximations.

When we consider stochastic differential equations of Markovian type with reflecting boundary condition, path-dependent stochastic differential equations appear. Generally, reflected processes are constructed by Skorohod equations. The mapping from the original process to the reflected process is a mapping on the path spaces, depends only on the shape of the boundary, and is called the Skorohod map. There is an equivalence between the stochastic differential equations of Markovian type with reflecting boundary condition and the path-dependent stochastic differential equations generated by the Skorohod map. Hence, the equations considered in the present paper are the generalized version of the stochastic differential equations of Markovian type with reflecting boundary condition whose Skorohod map is Lipschitz continuous. We remark that the Skorohod map is not always Lipschitz continuous. The detail of stochastic differential equations with reflecting boundary condition is discussed in Section 4.

In the present paper, we consider the Euler-Maruyama and Wong-Zakai approximations of path-dependent stochastic differential equations. The Euler-Maruyama approximation is

2010 *Mathematics Subject Classification.* Primary 60H10; Secondary 65C30.

Key words and phrases. Stochastic differential equation, reflecting boundary condition, path-dependent coefficient, Euler-Maruyama approximation, Wong-Zakai approximation, rate of convergence.

This work was supported by JSPS KAKENHI Grant number 24340023 and 25800054.

the approximation of the solutions by the processes generated by freezing the coefficients at given times, and is one of the most standard approximations of stochastic differential equations. The almost sure convergence of the Euler-Maruyama approximation of stochastic differential equations with the reflecting boundary condition is obtained by Pettersson [11]. The Wong-Zakai approximation is the approximation to the stochastic differential equations by the ordinarily differential equations obtained by piecewise linear approximation of the driving Brownian motion, and is originally introduced by Wong and Zakai [17]. It is known that the limit equation of the Wong-Zakai approximation is the stochastic differential equation of Stratonovich type. We remark that the limit equation of the Wong-Zakai approximation is different from that of the Euler-Maruyama approximation. The almost sure convergence of the Wong-Zakai approximation of stochastic differential equations with reflecting boundary condition is obtained by Doss and Priouret [6] and Zhang [18]. The approach for the convergence in distributions of the Wong-Zakai approximation of such an equation is studied in [8] and [15]. The L^p-convergence of the Euler-Maruyama and Wong-Zakai approximations is obtained in [3]. We remark that more general approximations of path-dependent stochastic differential equations are studied in [5]. We also remark that recently the equations with reflecting boundary condition are also studied by rough-path theory (see [1] and [2]). The argument in rough paths is closely related with the Wong-Zakai approximation.

In the present paper, we obtain the rates of the L^p-convergence of the Euler-Maruyama and the Wong-Zakai approximation of path-dependent stochastic differential equations under the Lipschitz continuity and some conditions on the coefficients. The obtained rates of the convergences are as follows.

$$E\left[\|X - X^{\mathrm{EM}}\|^p_{C([0,T];\mathbb{R}^d)}\right]^{1/p} \le C|\triangle|^{1/2},$$
$$E\left[\|X - X^{\mathrm{WZ}}\|^p_{C([0,T];\mathbb{R}^d)}\right]^{1/p} \le C|\triangle|^{1/2}(1+\log N)^{1/2},$$

where $\triangle = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ is a partition of the interval $[0, T]$, $|\triangle| := \max_{k=0,1,\dots,N-1}(t_{k+1} - t_k)$, X^{EM} is the Euler-Maruyama approximation associated with $\triangle$, and X^{WZ} is the solution to the Wong-Zakai approximation equation associated with $\triangle$. Note that the limit process X is the solution to the stochastic differential equation of Itô type in the case of the Euler-Maruyama approximation and the solution to the stochastic differential equation of Stratonovich type in the case of the Wong-Zakai approximation, and is different by each approximation. The rate of the L^p-convergence of the Euler-Maruyama approximation of stochastic differential equations with reflecting boundary condition on general domains is studied by Słomiński [14]. The obtained rate of the Euler-Maruyama approximation obtained in the present paper is the same as his one for convex polyhedral domains (see Remark 4.3).

The organization of the present paper is as follows. In Section 2 we consider the rate of the convergence of the Euler-Maruyama approximation. The argument in the section will be done by the standard techniques of stochastic differential equations. In Section 3 we consider the rate of the convergence of the Wong-Zakai approximation. The section is the main part of the present paper. We will prepare some lemmas about the estimates of the oscillation of the solutions in each interval of the partition, and will obtain the L^p-norm of the difference between the solution to the original equation and the solution to the approximating equations. In Section 4 we will see the relation between the obtained results in Sections 2 and 3, and stochastic differential equations of Markovian type with reflecting boundary condition. Section 5 is an appendix, in which we prepare an upper estimate for the p-th moment of the maximum of random variables.

We prepare some notations. For $T > 0$, $w \in C([0,T];\mathbb{R}^d)$ and $t \in [0,T]$, we define $\|w\|_{C([0,t];\mathbb{R}^d)} := \sup_{s\in[0,t]} |w(s)|_{\mathbb{R}^d}$. Denote the total set of the $d \times r$-matrices by $\mathbb{R}^d \otimes \mathbb{R}^r$, and for $A = (a_{ij}) \in \mathbb{R}^d \otimes \mathbb{R}^r$ define $|A|_{\mathbb{R}^d\otimes\mathbb{R}^r} := \sqrt{\sum_{i=1}^d \sum_{j=1}^r a_{ij}^2}$. For $x \in \mathbb{R}^d$ denote the ith component of x by x^i. Let δ_{ij} be Kronecker's delta, i.e. $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$.

2. Euler-Maruyama approximation. Let $T > 0$ and let ξ be an $\mathbb{R}^d$-valued random variable. Consider the following stochastic differential equation

$$\left\{ \begin{aligned} dX_t &= \sigma(t,X)dB_t + b(t,X)dt \\ X_0 &= \xi \end{aligned} \right. \tag{1}$$

where σ is an $\mathbb{R}^d \otimes \mathbb{R}^r$-valued function on $[0,T] \times C_b([0,T];\mathbb{R}^d)$, b is an $\mathbb{R}^d$-valued function on $[0,T] \times C_b([0,T];\mathbb{R}^d)$ and B is the r-dimensional Brownian motion. We assume the Lipschitz continuity of the coefficients in the following sense.

$$\begin{aligned} |\sigma(t,w) - \sigma(t,w')|_{\mathbb{R}^d\otimes\mathbb{R}^r} + |b(t,w) - b(t,w')|_{\mathbb{R}^d} \le K_T \|w - w'\|_{C([0,t];\mathbb{R}^d)}\,, \\ t \in [0,T],\ w, w' \in C([0,T];\mathbb{R}^d) \end{aligned} \tag{2}$$

where K_T is a constant depending on T. Then, the solution X to (1) exists, and has the pathwise uniqueness (see e.g. Theorem 7 of Chapter V in [12]).

We consider the Euler-Maruyama approximation to (1). Let $\triangle := \{0 = t_0 < t_1 < \cdots < t_N = T\}$ be a partition of the interval $[0,T]$. Define the approximations $\sigma_\triangle$, $b_\triangle$ of σ, b by

$$\sigma_\triangle(t,w) := \sigma(t_k,w),\ b_\triangle(t,w) := b(t_k,w), \quad t \in [t_k, t_{k+1})$$

for $k = 0, 1, \ldots, N-1$, and $w \in C([0,T];\mathbb{R}^d)$. We consider the following stochastic differential equation.

$$\left\{ \begin{aligned} dX_t^{\mathrm{EM}} &= \sigma_\triangle(t, X^{\mathrm{EM}})dB_t + b_\triangle(t, X^{\mathrm{EM}})dt \\ X_0^{\mathrm{EM}} &= \xi\,. \end{aligned} \right. \tag{3}$$

When $t \in [t_k, t_{k+1})$, it holds that

$$X_t^{\mathrm{EM}} = \xi + \sum_{l=0}^{k} \sigma(t_l, X^{\mathrm{EM}})(B_{t\wedge t_{l+1}} - B_{t_l}) + \sum_{l=0}^{k} b(t_l, X^{\mathrm{EM}})(t \wedge t_{l+1} - t_l)\,.$$

Hence, (3) is the equation of the Euler-Maruyama approximation to (1). Our purpose of this section is to estimate the L^p-norm of $\|X^{\mathrm{EM}} - X\|_{C([0,T];\mathbb{R}^d)}$ with respect to the probability measure. To give a condition on the coefficients we introduce a class of the functions on $[0,T] \times C([0,T];\mathbb{R}^d)$, which is an analogue to the class introduced in [5]. For a Hilbert space H and a positive number K, we define a class of H-valued functions $F_K(H)$ by the total set of $h : [0, T] \times C_b([0, T]; \mathbb{R}^d) \to H$ such that

(F1) $|h(t, w)|_H \leq K$ for $t \in [0, T],\ w \in C([0, T]; \mathbb{R}^d)$,

(F2) $|h(t, w) - h(s, w)|_H \leq K(\sqrt{t-s} + \|w(\cdot + s) - w(s)\|_{C([0,t-s];\mathbb{R}^d)})$ for $s, t \in [0, T]$ such that $s < t$, and $w \in C([0, T]; \mathbb{R}^d)$,

(F3) $|h(t, w) - h(t, w')|_H \leq K\|w - w'\|_{C([0,t];\mathbb{R}^d)}$ for $t \in [0, T],\ w, w' \in C([0, T]; \mathbb{R}^d)$.

REMARK 2.1. The assumptions (F1) and (F3) are given for the boundedness and the Lipschitz continuity, respectively. The assumption (F2) is for the continuity of the functions with respect to the time. We need (F2) for the Euler-Maruyama approximation.

The result of this section is the following theorem.

THEOREM 2.2. *Let $\sigma \in F_K(\mathbb{R}^d \otimes \mathbb{R}^r)$ and $b \in F_K(\mathbb{R}^d)$. Let X and X^{EM} be the solutions to* (1) *and to the equation of the Euler-Maruyama approximation* (3)*, respectively. Then, for $p \in [1, \infty)$ there exists a constant C depending on p, K, and T such that*

$$E\left[\|X - X^{\mathrm{EM}}\|_{C([0,T];\mathbb{R}^d)}^p\right]^{1/p} \leq C|\triangle|^{1/2}.$$

PROOF. In view of the magnitude relation between L^p-norms, it is sufficient to prove the case that $p \geq 2$. By the Burkholder-Davis-Gundy inequality, there exists a constant C_p depending on p and we have for $t \in [0, T]$

$$\begin{aligned}
&E\left[\|X - X^{\mathrm{EM}}\|_{C([0,t];\mathbb{R}^d)}^p\right] \\
&\leq 2^{p-1} E\left[\sup_{s\in[0,t]}\left|\int_0^s \left(\sigma(u, X) - \sigma_\triangle(u, X^{\mathrm{EM}})\right) dB_u\right|_{\mathbb{R}^d}^p\right] \\
&\quad + 2^{p-1} E\left[\sup_{s\in[0,t]}\left|\int_0^s \left(b(u, X) - b_\triangle(u, X^{\mathrm{EM}})\right) du\right|_{\mathbb{R}^d}^p\right] \\
&\leq 2^{p-1} C_p E\left[\left(\int_0^t \left|\sigma(u, X) - \sigma_\triangle(u, X^{\mathrm{EM}})\right|_{\mathbb{R}^d\otimes\mathbb{R}^r}^2 du\right)^{p/2}\right] \\
&\quad + 2^{p-1} E\left[\left(\int_0^t \left|b(u, X) - b_\triangle(u, X^{\mathrm{EM}})\right|_{\mathbb{R}^d} du\right)^p\right] \\
&\leq 2^{p-1} T^{p/2-1} C_p \int_0^t E\left[\left|\sigma(u, X) - \sigma_\triangle(u, X^{\mathrm{EM}})\right|_{\mathbb{R}^d\otimes\mathbb{R}^r}^p\right] du \\
&\quad + 2^{p-1} T^{p-1} \int_0^t E\left[\left|b(u, X) - b_\triangle(u, X^{\mathrm{EM}})\right|_{\mathbb{R}^d}^p\right] du.
\end{aligned}$$

Hence, it holds that

$$\begin{aligned}
&E\left[\|X - X^{\mathrm{EM}}\|_{C([0,t];\mathbb{R}^d)}^p\right] \\
&\leq C_{p,T}\left(\int_0^t E\left[\left|\sigma(u, X) - \sigma_\triangle(u, X^{\mathrm{EM}})\right|_{\mathbb{R}^d\otimes\mathbb{R}^r}^p\right] du\right.
\end{aligned} \tag{4}$$

$$+\int_0^t E\left[\left|b(u,X)-b_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d}\right]du\Big)$$

for $t \in [0,T]$, where $C_{p,T}$ is a constant depending on p and T. When $u \in [t_k, t_{k+1})$, (F2) and (F3) imply

$$\begin{aligned}
&E\left[\left|\sigma(u,X)-\sigma_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]\\
&= E\left[\left|\sigma(u,X)-\sigma(t_k,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]\\
&\le 2^{p-1}E\left[\left|\sigma(u,X)-\sigma(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]+2^{p-1}E\left[\left|\sigma(u,X^{\mathrm{EM}})-\sigma(t_k,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]\\
&\le 2^{p-1}K^pE\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]\\
&\quad +2^{p-1}K^pE\left[(\sqrt{u-t_k}+\|X^{\mathrm{EM}}(\cdot+t_k)-X^{\mathrm{EM}}(t_k)\|_{C([0,u-t_k];\mathbb{R}^d)})^p\right]\\
&\le 2^{p-1}K^pE\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]\\
&\quad +2^{2p-2}K^p\left((u-t_k)^{p/2}+E\left[\|X^{\mathrm{EM}}(\cdot+t_k)-X^{\mathrm{EM}}(t_k)\|^p_{C([0,u-t_k];\mathbb{R}^d)}\right]\right).
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&E\left[\left|\sigma(u,X)-\sigma_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]\\
&\le C_{p,K,T}\Big(E\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]\\
&\qquad +(u-t_k)^{p/2}+E\left[\|X^{\mathrm{EM}}(\cdot+t_k)-X^{\mathrm{EM}}(t_k)\|^p_{C([0,u-t_k];\mathbb{R}^d)}\right]\Big)
\end{aligned}\tag{5}$$

for $u \in [t_k, t_{k+1})$, where $C_{p,K,T}$ is a constant depending on p, K and T. On the other hand, for $u \in [t_k, t_{k+1})$

$$\begin{aligned}
&E\left[\|X^{\mathrm{EM}}(\cdot+t_k)-X^{\mathrm{EM}}(t_k)\|^p_{C([0,u-t_k];\mathbb{R}^d)}\right]\\
&= E\left[\sup_{s\in[t_k,u]}\left|\sigma(t_k,X^{\mathrm{EM}})(B_s-B_{t_k})+b(t_k,X^{\mathrm{EM}})(s-t_k)\right|^p_{\mathbb{R}^d}\right]\\
&\le 2^{p-1}K^p\left(E\left[\sup_{s\in[t_k,u]}|B_s-B_{t_k}|^p_{\mathbb{R}^d}\right]+(u-t_k)^p\right)\\
&= 2^{p-1}K^p\left((u-t_k)^{p/2}E\left[\sup_{s\in[0,1]}|B_s|^p_{\mathbb{R}^d}\right]+(u-t_k)^p\right).
\end{aligned}$$

Hence, by combining this inequality with (5) we have for $u \in [t_k, t_{k+1})$

$$\begin{aligned}
&E\left[\left|\sigma(u,X)-\sigma_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]\\
&\le 2^{p-1}K^pE\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]+C_{p,K,T}(u-t_k)^{p/2},
\end{aligned}$$

where $C_{p,K,T}$ is a constant depending on p, K and T. From this inequality we have

$$\begin{aligned}
&\int_0^t E\left[\left|\sigma(u,X)-\sigma_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]du \\
&=\sum_{k=1}^{N-1}\int_{t_k}^{t_{k+1}} E\left[\left|\sigma(u,X)-\sigma_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]du \\
&\le 2^{p-1}K^p\sum_{k=1}^{N-1}\int_{t_k}^{t_{k+1}} E\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]du + C_{p,K,T}\sum_{k=1}^{N-1}\int_{t_k}^{t_{k+1}}(u-t_k)^{p/2}du
\end{aligned}$$

for $t\in[0,T]$, where $C_{p,K,T}$ is a constant depending on p, K and T. Thus, we have

$$\begin{aligned}
&\int_0^t E\left[\left|\sigma(u,X)-\sigma_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d\otimes\mathbb{R}^r}\right]du \\
&\le 2^{p-1}K^p\int_0^t E\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]^{1/p}du + C_{p,K,T}|\triangle|^{p/2}
\end{aligned}\tag{6}$$

for $t\in[0,T]$, where $C_{p,K,T}$ is a constant depending on p, K and T. Similarly we have

$$\begin{aligned}
&\int_0^t E\left[\left|b(u,X)-b_\triangle(u,X^{\mathrm{EM}})\right|^p_{\mathbb{R}^d}\right]du \\
&\le 2^{p-1}K^p\int_0^t E\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]^{1/p}du + C_{p,K,T}|\triangle|^{p/2}
\end{aligned}\tag{7}$$

for $t\in[0,T]$, where $C_{p,K,T}$ is a constant depending on p, K and T. From (4), (6) and (7) we obtain

$$\begin{aligned}
&E\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,t];\mathbb{R}^d)}\right] \\
&\le C_{p,K,T}\int_0^t E\left[\left\|X-X^{\mathrm{EM}}\right\|^p_{C([0,u];\mathbb{R}^d)}\right]du + C_{p,K,T}|\triangle|^{p/2}
\end{aligned}$$

for $t\in[0,T]$, where $C_{p,K,T}$ is a constant depending on p, K and T. By applying Gronwall's inequality, we obtain the assertion. □

3. Wong-Zakai approximation. Let $T>0$. Let A be a mapping from $C([0,T];\mathbb{R}^d)$ to $C([0,T];\mathbb{R}^d)$ such that

(A1) $\|A(w)-A(w')\|_{C([0,t];\mathbb{R}^d)}\le K_A\|w-w'\|_{C([0,t];\mathbb{R}^d)}$ for $t\in[0,T]$, $w,w'\in C([0,T];\mathbb{R}^d)$,

(A2) $|A(w)_t-A(w)_s|_{\mathbb{R}^d}\le K_A\left(\sqrt{t-s}+\|w(\cdot+s)-w(s)\|_{C([0,t-s];\mathbb{R}^d)}\right)$ for $s,t\in[0,T]$ such that $s<t$, and $w\in C([0,T];\mathbb{R}^d)$,

(A3) $\mathrm{Var}_{[0,t]}(A(w))\le K_A(1+\|w-w(0)\|_{C([0,t];\mathbb{R}^d)})$ for $t\in[0,T]$, $w\in C([0,T];\mathbb{R}^d)$,

where $\mathrm{Var}_{[0,t]}(w)$ is the total variation of w on $[0,t]$, and let $f\in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ which has the bounded derivatives. Define the mapping $\Gamma: C([0,T];\mathbb{R}^d)\to C([0,T];\mathbb{R}^d)$ by

$$(\Gamma w)_t := f(t,w_t)+A(w)_t,\quad t\in[0,T],\ w\in C([0,T];\mathbb{R}^d)\,.\tag{8}$$

We denote the derivative of f in the time parameter by $\frac{\partial f}{\partial t}$ and the derivative of f in the l-th component of the spatial parameter by $\frac{\partial f}{\partial x^l}$. Let

$$K_f := \sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left|\frac{\partial f}{\partial t}(t,x)\right|_{\mathbb{R}^d} + \sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\sum_{l=1}^{d}\left|\frac{\partial f}{\partial x^l}(t,x)\right|_{\mathbb{R}^d}$$
$$+ \sup_{t\in[0,T]}\sup_{x,y\in\mathbb{R}^d;x\neq y}\sum_{l=1}^{d}\frac{1}{|x-y|_{\mathbb{R}^d}}\left|\frac{\partial f}{\partial x^l}(t,x)-\frac{\partial f}{\partial x^l}(t,y)\right|_{\mathbb{R}^d}.$$

From (A1) and (8), we have

$$\|\Gamma w - \Gamma w'\|_{C([0,t];\mathbb{R}^d)} \le \left(K_f + K_A\right)\|w-w'\|_{C([0,t];\mathbb{R}^d)} \tag{9}$$

for $t \in [0,T]$ and $w, w' \in C([0,T];\mathbb{R}^d)$. This implies that Γ is Lipschitz continuous in the sense of (2). From (A2) and (8), we have

$$|(\Gamma w)_t - (\Gamma w)_s|_{\mathbb{R}^d} \le \left(K_f + K_A\right)\left(\sqrt{t-s} + \|w(\cdot+s)-w(s)\|_{C([0,t-s];\mathbb{R}^d)}\right) \tag{10}$$

for $s, t \in [0,T]$ such that $s < t$, and $w \in C([0,T];\mathbb{R}^d)$.

Let $\sigma \in C_b([0,T]\times\mathbb{R}^d\times\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^r)$ such that $\sigma(t,x,y)$ is differentiable with respect to x and y and that there exists a positive constant K_σ satisfying

$$|\sigma(t,x_1,y_1)-\sigma(s,x_2,y_2)|_{\mathbb{R}^d\otimes\mathbb{R}^r} + \sum_{l=1}^{d}\left|\frac{\partial\sigma}{\partial x^l}(t,x_1,y_1)-\frac{\partial\sigma}{\partial x^l}(s,x_2,y_2)\right|_{\mathbb{R}^d\otimes\mathbb{R}^r}$$
$$+\sum_{l=1}^{d}\left|\frac{\partial\sigma}{\partial y^l}(t,x_1,y_1)-\frac{\partial\sigma}{\partial y^l}(s,x_2,y_2)\right|_{\mathbb{R}^d\otimes\mathbb{R}^r} \le K_\sigma\left(|t-s| + |x_1-x_2|_{\mathbb{R}^d} + |y_1-y_2|_{\mathbb{R}^d}\right)$$

for $s, t \in [0,T]$, and $x_1, x_2, y_1, y_2 \in \mathbb{R}^d$, where $\frac{\partial\sigma}{\partial x^l}$ is the derivative of σ in the l-th component of the first spatial parameter and $\frac{\partial\sigma}{\partial y^l}$ is the derivative of σ in the l-th component of the second spatial parameter. We denote the derivative of σ in the time parameter by $\frac{\partial\sigma}{\partial t}$. Let $b \in C_b([0,T]\times C([0,T];\mathbb{R}^d);\mathbb{R}^d)$ such that there exists a positive constant K_b satisfying

$$|b(t,w)-b(t,w')|_{\mathbb{R}^d} \le K_b\|w-w'\|_{C([0,t];\mathbb{R}^d)},$$

for $t \in [0,T]$, and $w, w' \in C([0,T];\mathbb{R}^d)$. Denote

$$M := \sup_{t\in[0,T]}\sup_{x,y\in\mathbb{R}^d}|\sigma(t,x,y)|_{\mathbb{R}^d\otimes\mathbb{R}^r} + \sup_{t\in[0,T]}\sup_{w\in C([0,T];\mathbb{R}^d)}|b(t,w)|_{\mathbb{R}^d}.$$

Let ξ be an $\mathbb{R}^d$-valued random variable. Consider the following stochastic differential equation of the Stratonovich type

$$\begin{cases} dX_t &= \sigma(t,X_t,(\Gamma X)_t)\circ dB_t + b(t,X)dt \\ X_0 &= \xi. \end{cases} \tag{11}$$

Let

$$U_t^i := \frac{1}{2}\sum_{j=1}^{r}\sum_{l=1}^{d}\int_0^t \frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s, (\Gamma X)_s)\sigma_{lj}(s, X_s, (\Gamma X)_s)ds$$

$$V_t^i := \frac{1}{2}\sum_{j=1}^{r}\sum_{l=1}^{d}\sum_{m=1}^{d}\int_0^t \frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s, (\Gamma X)_s)\frac{\partial f^l}{\partial x^m}(s, X_s)\sigma_{mj}(s, X_t, (\Gamma X)_t)ds$$

for $i = 1, 2, \ldots, d$. Then, we have

$$\sigma(t, X_t, (\Gamma X)_t) \circ dB_t = \sigma(t, X_t, (\Gamma X)_t)dB_t + dU_t + dV_t\,. \tag{12}$$

In view of this expression and the assumption on σ and b, we have the existence of the solution and the pathwise uniqueness of (11) as we have seen in Section 2.

For a given partition $\triangle := \{0 = t_0 < t_1 < \cdots < t_N = T\}$ of the interval $[0, T]$, we define the piecewise linear approximation $B^\triangle$ of B by

$$B_t^\triangle := B_{t_k} + \frac{t - t_k}{t_{k+1} - t_k}(B_{t_k+1} - B_{t_k}), \quad t \in [t_k, t_{k+1})\,.$$

We define the equation of the Wong-Zakai approximation to (11) by

$$\begin{cases} dX_t^{\mathrm{WZ}} &= \sigma(t, X_t^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_t)dB_t^\triangle + b(t, X^{\mathrm{WZ}})dt \\ X_0 &= \xi\,. \end{cases} \tag{13}$$

REMARK 3.1. Since $B^\triangle$ is the function of the bounded variation, the solution X^{WZ} to (13) is uniquely-determined almost surely.

The result of this section is the following theorem.

THEOREM 3.2. *Let σ and b as above. Let X and X^{WZ} be the solutions to* (11) *and to the equation of the Wong-Zakai approximation* (13)*, respectively. Then, for $p \in [1, \infty)$ there exists a constant C depending on p, T, d, r, K_A, K_σ, K_b, K_f and M, but independent of $\triangle$ and N, such that*

$$E\left[\left\|X - X^{\mathrm{WZ}}\right\|_{C([0,T];\mathbb{R}^d)}^p\right]^{1/p} \leq C|\triangle|^{1/2}(1 + \log N)^{1/2}\,.$$

It is sufficient to prove the case that $p \geq 2$. From now on, we use C's as constants depending on p, T, d, r, K_A, K_σ, K_b, K_f and M, but independent of $\triangle$ and N, and we remark that C's can be different from line to line.

Before starting the proof of Theorem 3.2 we prepare some lemmas.

LEMMA 3.3. *We have a constant*

$$E\left[\sup_{s\in[t_k,t_{k+1}]}\left|X_s - X_{t_k}\right|_{\mathbb{R}^d}^q\right]^{1/q} \leq C_q|t_{k+1} - t_k|^{1/2}$$

$$E\left[\sup_{s\in[t_k,t_{k+1}]}\left|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^q\right]^{1/q} \leq C_q|t_{k+1} - t_k|^{1/2}$$

for $q \in [1, \infty)$ *and* $k = 0, 1, \ldots, N-1$*, where* C_q *is a constant depending on* q*,* M*,* K_σ *and* K_f*.*

PROOF. For $s \in [t_k, t_{k+1}]$ it holds that

$$X_s - X_{t_k} = \int_{t_k}^{s} \sigma(u, X_u, (\Gamma X)_u) dB_u + \int_{t_k}^{s} b(u, X) du + U_s - U_{t_k} + V_s - V_{t_k} .$$

Hence, by the standard calculation with the Burkholder-Davis-Gundy inequality, we have the first inequality.

For $s \in [t_k, t_{k+1}]$ it holds that

$$\begin{aligned}
&\sup_{s \in [t_k, t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d} \\
&\leq \sup_{s \in [t_k, t_{k+1}]} \left| \int_{t_k}^{s} \sigma(u, X_u^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_u) \frac{B_{t_{k+1}} - B_{t_k}}{t_{k+1} - t_k} du \right|_{\mathbb{R}^d} + \sup_{s \in [t_k, t_{k+1}]} \left| \int_{t_k}^{s} b(u, X^{\mathrm{WZ}}) du \right|_{\mathbb{R}^d} \\
&\leq M |B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d} + M(t_{k+1} - t_k) .
\end{aligned}$$

Hence, we have the second inequality. $\square$

LEMMA 3.4. *There exists a positive number* $\varepsilon > 0$ *depending on* $\|\sigma\|_\infty$ *such that*

$$E\left[\exp\left(\varepsilon \sup_{s \in [t_k, t_{k+1}]} \frac{|X_s - X_{t_k}|_{\mathbb{R}^d}^2}{t_{k+1} - t_k}\right)\right] \leq C$$

$$E\left[\exp\left(\varepsilon \sup_{s \in [t_k, t_{k+1}]} \frac{|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}^2}{t_{k+1} - t_k}\right)\right] \leq C .$$

PROOF. In the proof of Lemma 3.3, we have obtained

$$\sup_{s \in [t_k, t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d} \leq \|\sigma\|_\infty |B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d} + C(t_{k+1} - t_k)$$

for $k = 0, 1, \ldots, N-1$. Hence, applying the Fernique theorem, we have the second inequality.

For each $k = 0, 1, \ldots, N-1$, $i = 1, 2, \ldots, d$ and $j = 1, 2, \ldots, r$, in view of Theorem 7.2′ of Chapter II in [9], if necessary by extending the probability space, there exists a one-dimensional Brownian motion $\tilde{B}_{ijk}(t)$

$$\begin{aligned}
\left| \int_{t_k}^{s} \sigma_{ij}(u, X_u, (\Gamma X)_u) dB_u^j \right| &= \left| \tilde{B}_{ijk} \left(\int_{t_k}^{s} \sigma_{ij}(u, X_u, (\Gamma X)_u)^2 du \right) \right| \\
&\leq \sqrt{t_{k+1} - t_k} \sup_{\tilde{u} \in [0, \|\sigma\|_\infty^2]} \frac{|\tilde{B}_{ijk}((t_{k+1} - t_k)\tilde{u})|}{\sqrt{t_{k+1} - t_k}}
\end{aligned}$$

for $s \in [t_k, t_{k+1}]$ almost surely. By this inequality we have

$$\begin{aligned}
&\sup_{s\in[t_k,t_{k+1}]} |X_s - X_{t_k}|_{\mathbb{R}^d} \\
&\le \sup_{s\in[t_k,t_{k+1}]} \left(\left| \int_{t_k}^{s} \sigma(u, X_u, (\Gamma X)_u) dB_u \right|_{\mathbb{R}^d} + |U_s - U_{t_k}|_{\mathbb{R}^d} \right. \\
&\qquad\qquad \left. + |V_s - V_{t_k}|_{\mathbb{R}^d} + \int_{t_k}^{s} |b(u, X)|_{\mathbb{R}^d} du \right) \\
&\le \sqrt{t_{k+1} - t_k} \sum_{i=1}^{d} \sum_{j=1}^{r} \sup_{\tilde{u}\in[0,\|\sigma\|_\infty^2]} \frac{|\tilde{B}_{ijk}((t_{k+1} - t_k)\tilde{u})|}{\sqrt{t_{k+1} - t_k}} + C(t_{k+1} - t_k)
\end{aligned}$$

almost surely for $k = 0, 1, \dots, N-1$. Therefore, by applying the Fernique theorem, we have the first inequality. □

LEMMA 3.5. *We have*

$$E\left[\max_{k=0,1,\dots,N-1} \sup_{s\in[t_k,t_{k+1}]} |X_s - X_{t_k}|_{\mathbb{R}^d}^p \right] \le C|\triangle|^{p/2} (1 + \log N)^{p/2} ,$$

$$E\left[\max_{k=0,1,\dots,N-1} \sup_{s\in[t_k,t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}^p \right] \le C|\triangle|^{p/2} (1 + \log N)^{p/2} .$$

PROOF. In view of Lemma 3.4, we have the assertion by applying Proposition 5.2 to the sequences of the random variables

$$\left\{ \sup_{s\in[t_k,t_{k+1}]} \frac{|X_s - X_{t_k}|_{\mathbb{R}^d}^2}{t_{k+1} - t_k} ; k = 0, 1, \dots, N-1 \right\} ,$$

$$\left\{ \sup_{s\in[t_k,t_{k+1}]} \frac{|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}^2}{t_{k+1} - t_k} ; k = 0, 1, \dots, N-1 \right\} .$$

□

Now we start to prove Theorem 3.2. From (12) we have

$$\begin{aligned}
d(X_t - X_t^{\mathrm{WZ}}) &= \sigma(t, X_t, (\Gamma X)_t) dB_t + \big(b(t, X) - b(t, X^{\mathrm{WZ}})\big)\, dt \\
&\qquad + dU_t + dV_t - \sigma(t, X_t^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_t) dB_t^{\triangle} .
\end{aligned} \tag{14}$$

By the integration by parts, we have for $k = 0, 1, \dots, N-1$, $i = 1, 2, \dots, d$ and $j = 1, 2, \dots, r$

$$\begin{aligned}
&\int_{t_k}^{t_{k+1}} \sigma_{ij}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) dB_s^{\triangle, j} \\
&= \int_{t_k}^{t_{k+1}} \sigma_{ij}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) \frac{B_{t_{k+1}}^j - B_{t_k}^j}{t_{k+1} - t_k} ds
\end{aligned}$$

$$
\begin{aligned}
&= -\left[(t_{k+1}-s)\sigma_{ij}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k}\right]_{t_k}^{t_{k+1}} \\
&\quad + \int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial t}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} ds \\
&\quad + \sum_{l=1}^{d}\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} dX_s^{\mathrm{WZ},l} \\
&\quad + \sum_{l=1}^{d}\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} d(\Gamma X^{\mathrm{WZ}})_s^l \\
&= \sigma_{ij}(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B^j_{t_{k+1}} - B^j_{t_k}) \\
&\quad + \int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{t_{k+1}-t_k}\frac{\partial \sigma_{ij}}{\partial t}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)(B^j_{t_{k+1}} - B^j_{t_k}) ds \\
&\quad + \sum_{l=1}^{d}\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} dX_s^{\mathrm{WZ},l} \\
&\quad + \sum_{l=1}^{d}\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} d(\Gamma X^{\mathrm{WZ}})_s^l .
\end{aligned}
$$

Hence, for $n = 0, 1, 2, \dots, N$ it holds that

$$
X_{t_n} - X_{t_n}^{\mathrm{WZ}} = I_1(t_n) + I_2(t_n) + I_3(t_n) + I_4(t_n) + I_5(t_n) \tag{15}
$$

where

$$
\begin{aligned}
I_1^i(t_n) &:= \sum_{k=0}^{n-1}\sum_{j=1}^{r}\int_{t_k}^{t_{k+1}} \big(\sigma_{ij}(s, X_s, (\Gamma X)_s) - \sigma_{ij}(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})\big)\, dB_s^j , \\
I_2(t_n) &:= \int_0^{t_n} \big(b(s, X) - b(s, X^{\mathrm{WZ}})\big)\, ds , \\
I_3^i(t_n) &:= -\sum_{k=0}^{n-1}\sum_{j=1}^{r}\int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{t_{k+1}-t_k}\frac{\partial \sigma_{ij}}{\partial t}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)(B^j_{t_{k+1}} - B^j_{t_k}) ds , \\
I_4^i(t_n) &:= U_{t_n}^i - \sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{l=1}^{d}\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} dX_s^{\mathrm{WZ},l} , \\
I_5^i(t_n) &:= V_{t_n}^i - \sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{l=1}^{d}\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}} - B^j_{t_k}}{t_{k+1}-t_k} d(\Gamma X^{\mathrm{WZ}})_s^l
\end{aligned}
$$

for $n = 1, 2, \dots, N$ and $i = 1, 2, \dots, d$.

We consider the estimates of $I_1, I_2, \dots, I_5$.

LEMMA 3.6. *For $n = 0, 1, \ldots, N$ we have*

$$E\left[\max_{k=0,1,\ldots,n} |I_1(t_k)|^p_{\mathbb{R}^d}\right] \le C\left(|\triangle|^{p/2} + \int_0^{t_n} E\left[\sup_{u\in[0,s]} \left|X_u - X_u^{\mathrm{WZ}}\right|^p_{\mathbb{R}^d}\right] ds\right),$$

$$E\left[\max_{k=0,1,\ldots,n} |I_2(t_k)|^p_{\mathbb{R}^d}\right] \le C\int_0^{t_n} E\left[\sup_{u\in[0,s]} \left|X_u - X_u^{\mathrm{WZ}}\right|^p_{\mathbb{R}^d}\right] ds,$$

$$E\left[\max_{k=0,1,\ldots,n} \left|I_3^i(t_k)\right|^p\right] \le C|\triangle|^{p/2}.$$

PROOF. Let

$$\Phi_{ij}(s) := \sigma_{ij}(s, X_s, (\Gamma X)_s) - \sigma_{ij}(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}), \quad s \in [t_k, t_{k+1})$$

for $i = 1, 2, \ldots, d$, $j = 1, 2, \ldots, r$ and $k = 0, 1, \ldots, N-1$. Then, Φ_{ij} is adapted process and we have

$$I_1^i(t_n) = \sum_{j=1}^r \int_0^{t_n} \Phi_{ij}(s) dB_s^j$$

for $i = 1, 2, \ldots, d$ and $n = 0, 1, 2, \ldots, N$. When $s \in [t_k, t_{k+1})$, in view of the Lipschitz continuity of σ, (9) and (10) it holds that

$$\begin{aligned}
|\Phi_{ij}(s)| &\le C\left(s - t_k + \left|X_s - X_{t_k}^{\mathrm{WZ}}\right|_{\mathbb{R}^d} + \left|(\Gamma X)_s - (\Gamma X^{\mathrm{WZ}})_{t_k}\right|_{\mathbb{R}^d}\right)\\
&\le C\Big(|\triangle|^{1/2} + \left|X_s - X_s^{\mathrm{WZ}}\right|_{\mathbb{R}^d} + \left|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}\right|_{\mathbb{R}^d}\\
&\qquad + \left|(\Gamma X)_s - (\Gamma X^{\mathrm{WZ}})_s\right|_{\mathbb{R}^d} + \left|(\Gamma X^{\mathrm{WZ}})_s - (\Gamma X^{\mathrm{WZ}})_{t_k}\right|_{\mathbb{R}^d}\Big)\\
&\le C\left(|\triangle|^{1/2} + \sup_{u\in[0,s]}\left|X_u - X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d} + \sup_{s\in[t_k,t_{k+1}]}\left|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}\right|_{\mathbb{R}^d}\right)
\end{aligned}$$

for $i = 1, 2, \ldots, d$ and $j = 1, 2, \ldots, r$. From this inequality and Lemma 3.3, we have

$$E\left[|\Phi_{ij}(s)|^p\right] \le C|\triangle|^{p/2} + CE\left[\sup_{u\in[0,s]}\left|X_u - X_u^{\mathrm{WZ}}\right|^p_{\mathbb{R}^d}\right]$$

for $s \in [0, T]$, $i = 1, 2, \ldots, d$ and $j = 1, 2, \ldots, r$. Hence, by the Burkholder-Davis-Gundy inequality

$$\begin{aligned}
E\left[\max_{k=0,1,\ldots,n}\left|I_1^i(t_k)\right|^p\right] &\le E\left[\sup_{t\in[0,t_n]}\left|\sum_{j=1}^r \int_0^t \Phi_{ij}(s) dB_s^j\right|^p\right]\\
&\le CE\left[\left(\sum_{j=1}^r \int_0^{t_n} |\Phi_{ij}(s)|^2 ds\right)^{p/2}\right]
\end{aligned}$$

$$\leq CE\left[\sum_{j=1}^{r}\int_0^{t_n}|\Phi_{ij}(s)|^p ds\right]$$
$$\leq C\left(|\triangle|^{p/2}+\int_0^{t_n}E\left[\sup_{u\in[0,s]}\left|X_u-X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right]ds\right)$$

for $i=1,2,\ldots,d$ and $n=0,1,2,\ldots,N$. Thus, we have the first inequality.

An explicit calculation implies

$$\begin{aligned}E\left[\max_{k=0,1,\ldots,n}|I_2(t_k)|_{\mathbb{R}^d}^p\right]&\leq E\left[\left(\int_0^{t_n}\left|b(s,X)-b(s,X^{\mathrm{WZ}})\right|_{\mathbb{R}^d}ds\right)^p\right]\\&\leq CE\left[\left(\int_0^{t_n}\sup_{u\in[0,s]}\left|X_u-X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}ds\right)^p\right]\\&\leq C\int_0^{t_n}E\left[\sup_{u\in[0,s]}\left|X_u-X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right]ds\end{aligned}$$

for $n=0,1,\ldots,N$. Hence, we have the second inequality.

Since Jensen's inequality implies

$$\begin{aligned}\left(\sum_{k=0}^{n-1}|B_{t_{k+1}}^j-B_{t_k}^j|(t_{k+1}-t_k)\right)^p&=T^p\left(\sum_{k=0}^{n-1}|B_{t_{k+1}}^j-B_{t_k}^j|\frac{t_{k+1}-t_k}{T}\right)^p\\&\leq T^p\sum_{k=0}^{n-1}|B_{t_{k+1}}^j-B_{t_k}^j|^p\frac{t_{k+1}-t_k}{T}\\&=T^{p-1}\sum_{k=0}^{n-1}|B_{t_{k+1}}^j-B_{t_k}^j|^p(t_{k+1}-t_k)\end{aligned}$$

for $j=1,2,\ldots,r$, we have

$$\begin{aligned}&E\left[\max_{k=0,1,\ldots,n}\left|I_3^i(t_k)\right|^p\right]\\&\leq E\left[\left(\sum_{k=0}^{n-1}\sum_{j=1}^{r}\left|\int_{t_k}^{t_{k+1}}\frac{t_{k+1}-s}{t_{k+1}-t_k}\frac{\partial\sigma_{ij}}{\partial s}(s,X_s^{\mathrm{WZ}},(\Gamma X^{\mathrm{WZ}})_s)(B_{t_{k+1}}^j-B_{t_k}^j)ds\right|\right)^p\right]\\&\leq C\sum_{j=1}^{r}E\left[\left(\sum_{k=0}^{n-1}|B_{t_{k+1}}^j-B_{t_k}^j|(t_{k+1}-t_k)\right)^p\right]\\&\leq C\sum_{j=1}^{r}\sum_{k=0}^{n-1}(t_{k+1}-t_k)E\left[|B_{t_{k+1}}^j-B_{t_k}^j|^p\right]\\&\leq C|\triangle|^{p/2}\end{aligned}$$

for $n=0,1,\ldots,N$. Hence, the last inequality holds. □

Before estimating I_4 and I_5, we prepare the following. For $i = 1, 2, \dots, d$ and $j, m = 1, 2, \dots, r$, let

$$\mu_t^{ijm} := \sum_{k=0}^{N-1} \sum_{l=1}^{d} \left(\frac{\partial \sigma_{ij}}{\partial x^l} \sigma_{lm} \right) (t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})$$
$$\times \left(\delta_{jm}(t \wedge t_{k+1} - t \wedge t_k) - (B_{t\wedge t_{k+1}}^j - B_{t \wedge t_k}^j)(B_{t\wedge t_{k+1}}^m - B_{t\wedge t_k}^m) \right)$$
$$\nu_t^{ijm} := \sum_{k=0}^{n-1} \sum_{l=1}^{d} \sum_{q=1}^{d} \left(\frac{\partial \sigma_{ij}}{\partial y^l} \sigma_{qm} \right) (t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \frac{\partial f^l}{\partial x^q}(t_k, X_{t_k}^{\mathrm{WZ}})$$
$$\times \left(\delta_{jm}(t \wedge t_{k+1} - t \wedge t_k) - (B_{t\wedge t_{k+1}}^j - B_{t \wedge t_k}^j)(B_{t\wedge t_{k+1}}^m - B_{t\wedge t_k}^m) \right) .$$

Then, we have the following lemma.

LEMMA 3.7. *For $n = 0, 1, 2, \dots, N$ and $i = 1, 2, \dots, d$, we have*

$$E\left[\max_{k=0,1,\dots,n} \left| \sum_{j=1}^{r} \sum_{m=1}^{r} \mu_{t_k}^{ijm} \right|^p \right] \le C|\triangle|^{p/2},$$
$$E\left[\max_{k=0,1,\dots,n} \left| \sum_{j=1}^{r} \sum_{m=1}^{r} \nu_{t_k}^{ijm} \right|^p \right] \le C|\triangle|^{p/2}.$$

PROOF. For $i = 1, 2, \dots, d$ and $j, m = 1, 2, \dots, r$, let

$$\widetilde{\mu}_t^{jm} := \sum_{k=0}^{N-1} \left(\delta_{jm}(t \wedge t_{k+1} - t \wedge t_k) - (B_{t\wedge t_{k+1}}^j - B_{t \wedge t_k}^j)(B_{t\wedge t_{k+1}}^m - B_{t\wedge t_k}^m) \right) ,$$
$$\Phi_{ijm}(t) := \sum_{l=1}^{d} \left(\frac{\partial \sigma_{ij}}{\partial x_l} \sigma_{lm} \right) (t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \quad t \in [t_k, t_{k+1}) .$$

Then, $\widetilde{\mu}_t^{jm}$ is a martingale and it holds that

$$\mu_t^{ijm} = \int_0^t \Phi_{ijm}(s) d\widetilde{\mu}_s^{jm}, \quad t \in [0, T]$$

for $i = 1, 2, \dots, d$ and $j, m = 1, 2, \dots, r$. Let $\langle \widetilde{\mu}^{jm} \rangle$ be the quadratic variation of $\widetilde{\mu}^{jm}$ for $j, m = 1, 2, \dots, r$. Since Itô's formula implies

$$\widetilde{\mu}_t^{jm} = -\sum_{k=0}^{N-1} \left(\int_{t\wedge t_k}^{t \wedge t_{k+1}} (B_s^j - B_{t_k}^j) dB_s^m + \int_{t\wedge t_k}^{t\wedge t_{k+1}} (B_s^m - B_{t_k}^m) dB_s^j \right), \quad t \in [0, T]$$

for $j, m = 1, 2, \dots, r$, we have

$$
\begin{aligned}
E\left[\left\langle\widetilde{\mu}^{jm}\right\rangle_{t_n}^{p/2}\right]^{2/p} &\le E\left[\left\{\sum_{k=0}^{n-1}\left(2\int_{t_k}^{t_{k+1}}(B_s^j - B_{t_k}^j)^2 ds + 2\int_{t_k}^{t_{k+1}}(B_s^m - B_{t_k}^m)^2 ds\right)\right\}^{p/2}\right]^{2/p} \\
&\le \sum_{k=0}^{n-1} E\left[\left(2\int_{t_k}^{t_{k+1}}(B_s^j - B_{t_k}^j)^2 ds + 2\int_{t_k}^{t_{k+1}}(B_s^m - B_{t_k}^m)^2 ds\right)^{p/2}\right]^{2/p} \\
&\le C \max_{j=1,2,\dots,r} \sum_{k=0}^{n-1}(t_{k+1} - t_k) E\left[\sup_{s\in[t_k,t_{k+1}]}|B_s^j - B_{t_k}^j|^p\right]^{2/p} \\
&\le C|\triangle|
\end{aligned}
$$

for $n = 0, 1, 2, \dots, N$ and $j, m = 1, 2, \dots, r$. Hence, by the Burkholder-Davis-Gundy inequality we have

$$
\begin{aligned}
E\left[\max_{k=0,1,\dots,n}\left|\sum_{j=1}^{r}\sum_{m=1}^{r}\mu_{t_k}^{ijm}\right|^p\right] &\le C\sum_{j=1}^{r}\sum_{m=1}^{r} E\left[\max_{t\in[0,t_n]}\left|\int_0^t \Phi_{ijm}(s) d\widetilde{\mu}_s^{jm}\right|^p\right] \\
&\le C\sum_{j=1}^{r}\sum_{m=1}^{r} E\left[\left(\int_0^{t_n}|\Phi_{ijm}(s)|^2 d\left\langle\widetilde{\mu}^{jm}\right\rangle_s\right)^{p/2}\right] \\
&\le C\sum_{j=1}^{r}\sum_{m=1}^{r} E\left[\left\langle\widetilde{\mu}^{jm}\right\rangle_{t_n}^{p/2}\right] \\
&\le C|\triangle|^{p/2}
\end{aligned}
$$

for $n = 0, 1, 2, \dots, N$ and $i = 1, 2, \dots, d$. Similarly we obtain the proof of the estimate for ν^{jm}. □

Now we have estimates of I_4 and I_5, as follows.

LEMMA 3.8. *It holds that*

$$
\begin{aligned}
E\left[\max_{k=0,1,\dots,n}\left|I_4^i(t_k)\right|^p\right] &\le C\left(|\triangle|^{p/2} + \int_0^{t_n} E\left[\sup_{u\in[0,s]}\left|X_u - X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right] ds\right), \\
E\left[\max_{k=0,1,\dots,n}\left|I_5^i(t_k)\right|^p\right] &\le C\left(|\triangle|^{p/2}(1+\log N)^{p/2} + \int_0^{t_n} E\left[\sup_{u\in[0,s]}\left|X_u - X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right] ds\right)
\end{aligned}
$$

for $n = 0, 1, \dots, N$ *and* $i = 1, 2, \dots, d$.

PROOF. For $k = 0, 1, \dots, N$, $i, l = 1, 2, \dots, d$ and $j = 1, 2, \dots, r$, it holds that

$$
\begin{aligned}
&\int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}}-B^j_{t_k}}{t_{k+1}-t_k}dX_s^{\mathrm{WZ},l} \\
&= \sum_{m=1}^r \int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}}-B^j_{t_k}}{t_{k+1}-t_k} \\
&\qquad\qquad\qquad \times \sigma_{lm}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^m_{t_{k+1}}-B^m_{t_k}}{t_{k+1}-t_k}ds \\
&\quad + \int_{t_k}^{t_{k+1}} (t_{k+1}-s)\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B^j_{t_{k+1}}-B^j_{t_k}}{t_{k+1}-t_k}b^l(s, X^{\mathrm{WZ}})ds \\
&= \sum_{m=1}^r (B^j_{t_{k+1}}-B^j_{t_k})(B^m_{t_{k+1}}-B^m_{t_k})\int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{(t_{k+1}-t_k)^2} \\
&\qquad \times \left[\left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) - \left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})\right]ds \\
&\quad + \sum_{m=1}^r \left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \\
&\qquad\qquad\qquad \times (B^j_{t_{k+1}}-B^j_{t_k})(B^m_{t_{k+1}}-B^m_{t_k})\int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{(t_{k+1}-t_k)^2}ds \\
&\quad + (B^j_{t_{k+1}}-B^j_{t_k})\int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{t_{k+1}-t_k}\frac{\partial \sigma_{ij}}{\partial x^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)b^l(s, X^{\mathrm{WZ}})ds\,.
\end{aligned}
$$

Hence, by (10), we have for $n = 0, 1, \dots, N$ and $i = 1, 2, \dots, d$

$$
\begin{aligned}
&\left|I_4^i(t_n)\right| \\
&\le \frac{1}{2}\left|\sum_{j=1}^r\sum_{l=1}^d \int_0^t \left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lj}\right)(s, X_s, (\Gamma X)_s)ds\right. \\
&\qquad \left. - \sum_{k=0}^{n-1}\sum_{j=1}^r\sum_{l=1}^d\sum_{m=1}^r \left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B^j_{t_{k+1}}-B^j_{t_k})(B^m_{t_{k+1}}-B^m_{t_k})\right| \\
&\quad + \sum_{k=0}^{n-1}\sum_{j=1}^r\sum_{l=1}^d\sum_{m=1}^r |B^j_{t_{k+1}}-B^j_{t_k}||B^m_{t_{k+1}}-B^m_{t_k}|\int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{(t_{k+1}-t_k)^2} \\
&\qquad \times \left|\left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) - \left(\frac{\partial \sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})\right|ds \\
&\quad + C\sum_{k=0}^{n-1} |B^j_{t_{k+1}}-B^j_{t_k}|\int_{t_k}^{t_{k+1}} \frac{t_{k+1}-s}{t_{k+1}-t_k}ds
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{1}{2}\left|\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{l=1}^{d}\sum_{m=1}^{r}\left(\delta_{jm}(t_{k+1}-t_k)\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X_{t_k}, (\Gamma X)_{t_k})\right.\right.\\
&\qquad\left.\left.-\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B^j_{t_{k+1}}-B^j_{t_k})(B^m_{t_{k+1}}-B^m_{t_k})\right)\right|\\
&\quad+\frac{1}{2}\left|\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{l=1}^{d}\int_{t_k}^{t_{k+1}}\left[\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lj}\right)(s, X_s, (\Gamma X)_s)-\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lj}\right)(t_k, X_{t_k}, (\Gamma X)_{t_k})\right]ds\right|\\
&\quad+C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{l=1}^{d}\sum_{m=1}^{r}|B^j_{t_{k+1}}-B^j_{t_k}||B^m_{t_{k+1}}-B^m_{t_k}|\\
&\qquad\times\int_{t_k}^{t_{k+1}}\frac{t_{k+1}-s}{(t_{k+1}-t_k)^2}\left(s-t_k+|X^{\mathrm{WZ}}_s-X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d}+|(\Gamma X^{\mathrm{WZ}})_s-(\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right)ds\\
&\quad+C\sum_{j=1}^{r}\sum_{k=0}^{n-1}|B^j_{t_{k+1}}-B^j_{t_k}|(t_{k+1}-t_k)\\
&\le \frac{1}{2}\left|\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{l=1}^{d}(t_{k+1}-t_k)\right.\\
&\qquad\qquad\left.\times\left[\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lj}\right)(t_k, X_{t_k}, (\Gamma X)_{t_k})-\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lj}\right)(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k})\right]\right|\\
&\quad+\frac{1}{2}\left|\sum_{j=1}^{r}\sum_{m=1}^{r}\sum_{k=0}^{n-1}\sum_{l=1}^{d}\left(\frac{\partial\sigma_{ij}}{\partial x^l}\sigma_{lm}\right)(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k})\right.\\
&\qquad\qquad\qquad\qquad\left.\times\left(\delta_{jm}(t_{k+1}-t_k)-(B^j_{t_{k+1}}-B^j_{t_k})(B^m_{t_{k+1}}-B^m_{t_k})\right)\right|\\
&\quad+C\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}\left(s-t_k+|X_s-X_{t_k}|_{\mathbb{R}^d}+|(\Gamma X)_s-(\Gamma X)_{t_k}|_{\mathbb{R}^d}\right)ds\\
&\quad+C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r}|B^j_{t_{k+1}}-B^j_{t_k}||B^m_{t_{k+1}}-B^m_{t_k}|\\
&\quad\times\int_{t_k}^{t_{k+1}}\frac{t_{k+1}-s}{(t_{k+1}-t_k)^2}\left(s-t_k+|X^{\mathrm{WZ}}_s-X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d}+|(\Gamma X^{\mathrm{WZ}})_s-(\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right)ds\\
&\quad+C\sum_{j=1}^{r}\sum_{k=0}^{n-1}|B^j_{t_{k+1}}-B^j_{t_k}|(t_{k+1}-t_k)\\
&\le C\sum_{k=0}^{n-1}(t_{k+1}-t_k)\left(|X_{t_k}-X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d}+|(\Gamma X)_{t_k}-(\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}\left|\sum_{j=1}^{r}\sum_{m=1}^{r}\mu_{t_n}^{ijm}\right| + C\sum_{k=0}^{n-1}(t_{k+1}-t_k)\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_k+1]}|X_s - X_{t_k}|_{\mathbb{R}^d}\right)\\
&+C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r}|B_{t_{k+1}}^j - B_{t_k}^j||B_{t_{k+1}}^m - B_{t_k}^m|\\
&\qquad\times\sup_{s\in[t_k,t_{k+1}]}\left(s-t_k + |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d} + |(\Gamma X^{\mathrm{WZ}})_s - (\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right)\\
&+C\sum_{j=1}^{r}\sum_{k=0}^{n-1}|B_{t_{k+1}}^j - B_{t_k}^j|(t_{k+1}-t_k)\\
\le{}& C\int_0^{t_n}\sup_{u\in[0,s]}|X_u - X_u^{\mathrm{WZ}}|_{\mathbb{R}^d}ds + \frac{1}{2}\max_{k=0,1,\dots,n}\left|\sum_{j=1}^{r}\sum_{m=1}^{r}\mu_{t_k}^{ijm}\right|\\
&+C\sum_{k=0}^{n-1}(t_{k+1}-t_k)\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]}|X_s - X_{t_k}|_{\mathbb{R}^d}\right)\\
&+C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r}|B_{t_{k+1}}^j - B_{t_k}^j||B_{t_{k+1}}^m - B_{t_k}^m|\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]}|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}\right)\\
&+C\sum_{j=1}^{r}\sum_{k=0}^{n-1}|B_{t_{k+1}}^j - B_{t_k}^j|(t_{k+1}-t_k)\,.
\end{aligned}
$$

By taking the L^p-norm of the both sides and applying Lemmas 3.3 and 3.7, Example 5.3 and Hölder's inequality, we obtain

$$
\begin{aligned}
&E\left[\max_{k=0,1,\dots,n}\left|I_4^i(t_k)\right|^p\right]^{1/p}\\
&\le C|\triangle|^{1/2} + C\int_0^{t_n}E\left[\sup_{u\in[0,s]}|X_u - X_u^{\mathrm{WZ}}|_{\mathbb{R}^d}^p\right]^{1/p}ds\\
&\quad+C\sum_{k=0}^{n-1}(t_{k+1}-t_k)\left(|\triangle|^{1/2} + E\left[\sup_{s\in[t_k,t_{k+1}]}|X_s - X_{t_k}|_{\mathbb{R}^d}^p\right]^{1/p}\right)\\
&\quad+C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r}E\left[|B_{t_{k+1}}^j - B_{t_k}^j|^p|B_{t_{k+1}}^m - B_{t_k}^m|^p\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]}|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}\right)^p\right]^{1/p}\\
&\quad+C\sum_{j=1}^{r}\sum_{k=0}^{n-1}E\left[|B_{t_{k+1}}^j - B_{t_k}^j|^p\right]^{1/p}(t_{k+1}-t_k)
\end{aligned}
$$

$$
\begin{aligned}
&\le C|\triangle|^{1/2} + C\int_0^{t_n} E\left[\sup_{u\in[0,s]} |X_u - X_u^{\mathrm{WZ}}|_{\mathbb{R}^d}^p\right]^{1/p} ds \\
&\quad + C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r} E\left[|B_{t_{k+1}}^j - B_{t_k}^j|^{3p}\right]^{1/(3p)} E\left[|B_{t_{k+1}}^m - B_{t_k}^m|^{3p}\right]^{1/(3p)} \\
&\qquad\qquad \times E\left[\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}\right)^{3p}\right]^{1/(3p)} \\
&\le C|\triangle|^{1/2} + C\int_0^{t_n} E\left[\sup_{u\in[0,s]} |X_u - X_u^{\mathrm{WZ}}|_{\mathbb{R}^d}^p\right]^{1/p} ds\,.
\end{aligned}
$$

The desired estimate for I_4 follows from this inequality.

Next we consider the estimate for I_5. Note that for $t \in [0, T]$

$$
\begin{aligned}
(\Gamma X^{\mathrm{WZ}})_t - (\Gamma X^{\mathrm{WZ}})_0 &= \int_0^t \frac{\partial f}{\partial t}(s, X_s^{\mathrm{WZ}})ds \\
&\quad + \sum_{j=1}^{d}\sum_{l=1}^{r}\int_0^t \frac{\partial f}{\partial x^j}(s, X_s^{\mathrm{WZ}})\sigma_{jl}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)dB_s^{\triangle,l} \\
&\quad + \sum_{j=1}^{d}\int_0^t \frac{\partial f}{\partial x^j}(s, X_s^{\mathrm{WZ}})b^j(s, X^{\mathrm{WZ}})ds + A(X^{\mathrm{WZ}})_t - A(X^{\mathrm{WZ}})_0\,.
\end{aligned}
$$

To simplify the notation, let

$$
g_{ijm}(t, x, y) := \sum_{l=1}^{d}\sum_{q=1}^{d}\frac{\partial \sigma_{ij}}{\partial y^l}(t, x, y)\sigma_{qm}(t, x, y)\frac{\partial f^l}{\partial x^q}(t, x), \quad t \in [0, T], x, y \in \mathbb{R}^d
$$

for $i = 1, 2, \ldots, d$ and $j, m = 1, 2, \ldots, r$. For $k = 0, 1, \ldots, N-1$, $i, l = 1, 2, \ldots, d$ and $j = 1, 2, \ldots, r$, it holds that

$$
\begin{aligned}
&\int_{t_k}^{t_{k+1}} (t_{k+1} - s)\frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B_{t_{k+1}}^j - B_{t_k}^j}{t_{k+1} - t_k} d(\Gamma X^{\mathrm{WZ}})_s^l \\
&= \sum_{q=1}^{d}\sum_{m=1}^{r}\int_{t_k}^{t_{k+1}} (t_{k+1} - s)\frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B_{t_{k+1}}^j - B_{t_k}^j}{t_{k+1} - t_k} \\
&\qquad\qquad \times \frac{\partial f^l}{\partial x^q}(s, X_s^{\mathrm{WZ}})\sigma_{qm}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B_{t_{k+1}}^m - B_{t_k}^m}{t_{k+1} - t_k} ds \\
&\quad + \int_{t_k}^{t_{k+1}} (t_{k+1} - s)\frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)\frac{B_{t_{k+1}}^j - B_{t_k}^j}{t_{k+1} - t_k}
\end{aligned}
$$

$$
\times \left(\frac{\partial f^l}{\partial t}(s, X_s^{\mathrm{WZ}}) + \sum_{q=1}^{d} \frac{\partial f^l}{\partial x^q}(s, X_s^{\mathrm{WZ}}) b^q(s, X^{\mathrm{WZ}}) \right) ds
$$

$$
+ \int_{t_k}^{t_{k+1}} (t_{k+1} - s) \frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) \frac{B_{t_{k+1}}^j - B_{t_k}^j}{t_{k+1} - t_k} dA^l(X^{\mathrm{WZ}})_s
$$

$$
= \sum_{q=1}^{d} \sum_{m=1}^{r} (B_{t_{k+1}}^j - B_{t_k}^j)(B_{t_{k+1}}^m - B_{t_k}^m) \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{(t_{k+1} - t_k)^2}
$$

$$
\times \left[\left(\frac{\partial \sigma_{ij}}{\partial y^l} \sigma_{qm} \right)(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) \frac{\partial f^l}{\partial x^q}(s, X_s^{\mathrm{WZ}}) \right.
$$

$$
\left. - \left(\frac{\partial \sigma_{ij}}{\partial y^l} \sigma_{qm} \right)(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \frac{\partial f^l}{\partial x^q}(t_k, X_{t_k}^{\mathrm{WZ}}) \right] ds
$$

$$
+ \sum_{q=1}^{d} \sum_{m=1}^{r} \left(\frac{\partial \sigma_{ij}}{\partial y^l} \sigma_{qm} \right)(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \frac{\partial f^l}{\partial x^q}(t_k, X_{t_k}^{\mathrm{WZ}})
$$

$$
\times (B_{t_{k+1}}^j - B_{t_k}^j)(B_{t_{k+1}}^m - B_{t_k}^m) \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{(t_{k+1} - t_k)^2} ds
$$

$$
+ (B_{t_{k+1}}^j - B_{t_k}^j) \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{t_{k+1} - t_k} \frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s)
$$

$$
\times \left(\frac{\partial f^l}{\partial t}(s, X_s^{\mathrm{WZ}}) + \sum_{q=1}^{d} \frac{\partial f^l}{\partial x^q}(s, X_s^{\mathrm{WZ}}) b^q(s, X^{\mathrm{WZ}}) \right) ds
$$

$$
+ (B_{t_{k+1}}^j - B_{t_k}^j) \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{t_{k+1} - t_k} \frac{\partial \sigma_{ij}}{\partial y^l}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) dA^l(X^{\mathrm{WZ}})_s .
$$

Hence, by (10), we have for $n = 0, 1, \dots, N$ and $i = 1, 2, \dots, d$

$$
\left| I_5^i(t_n) \right|
$$

$$
\leq \frac{1}{2} \left| \sum_{j=1}^{r} \int_0^t g_{ijj}(s, X_s, (\Gamma X)_s) ds \right.
$$

$$
\left. - \sum_{k=0}^{n-1} \sum_{j=1}^{r} \sum_{m=1}^{r} g_{ijm}(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B_{t_{k+1}}^j - B_{t_k}^j)(B_{t_{k+1}}^m - B_{t_k}^m) \right|
$$

$$
+ \sum_{k=0}^{n-1} \sum_{j=1}^{r} \sum_{m=1}^{r} |B_{t_{k+1}}^j - B_{t_k}^j||B_{t_{k+1}}^m - B_{t_k}^m| \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{(t_{k+1} - t_k)^2}
$$

$$
\times \left| g_{ijm}(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) - g_{ijm}(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \right| ds
$$

$$
\begin{aligned}
&+ C \sum_{j=1}^{r} \sum_{k=0}^{n-1} |B^j_{t_{k+1}} - B^j_{t_k}| \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{t_{k+1} - t_k} ds \\
&+ C \sum_{j=1}^{r} \sum_{k=0}^{n-1} |B^j_{t_{k+1}} - B^j_{t_k}| \mathrm{Var}_{[t_k, t_{k+1}]}(A(X^{\mathrm{WZ}})) \\
\le{}& \frac{1}{2} \Bigg| \sum_{k=0}^{n-1} \sum_{j=1}^{r} \sum_{m=1}^{r} \big[\delta_{jm}(t_{k+1} - t_k) g_{ijm}(t_k, X_{t_k}, (\Gamma X)_{t_k}) \\
&\qquad - g_{ijm}(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B^j_{t_{k+1}} - B^j_{t_k})(B^m_{t_{k+1}} - B^m_{t_k}) \big] \Bigg| \\
&+ \frac{1}{2} \Bigg| \sum_{k=0}^{n-1} \sum_{j=1}^{r} \int_{t_k}^{t_{k+1}} \big[g_{ijj}(s, X_s, (\Gamma X)_s) - g_{ijj}(t_k, X_{t_k}, (\Gamma X)_{t_k}) \big] ds \Bigg| \\
&+ C \sum_{k=0}^{n-1} \sum_{j=1}^{r} \sum_{m=1}^{r} |B^j_{t_{k+1}} - B^j_{t_k}| |B^m_{t_{k+1}} - B^m_{t_k}| \\
&\quad \times \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{(t_{k+1} - t_k)^2} \big(s - t_k + |X^{\mathrm{WZ}}_s - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d} + |(\Gamma X^{\mathrm{WZ}})_s - (\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d} \big) ds \\
&+ C \sum_{j=1}^{r} \sum_{k=0}^{n-1} |B^j_{t_{k+1}} - B^j_{t_k}| (t_{k+1} - t_k) + C \sum_{j=1}^{r} \sum_{k=0}^{n-1} |B^j_{t_{k+1}} - B^j_{t_k}| \mathrm{Var}_{[t_k, t_{k+1}]}(A(X^{\mathrm{WZ}})) \\
\le{}& \frac{1}{2} \sum_{k=0}^{n-1} \sum_{j=1}^{r} (t_{k+1} - t_k) \big| g_{ijj}(t_k, X_{t_k}, (\Gamma X)_{t_k}) - g_{ijj}(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \big| \\
&+ \frac{1}{2} \Bigg| \sum_{j=1}^{r} \sum_{m=1}^{r} \sum_{k=0}^{n-1} g_{ijm}(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k}) \\
&\qquad \times \Big(\delta_{jm}(t_{k+1} - t_k) - (B^j_{t_{k+1}} - B^j_{t_k})(B^m_{t_{k+1}} - B^m_{t_k}) \Big) \Bigg| \\
&+ C \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(s - t_k + |X_s - X_{t_k}|_{\mathbb{R}^d} + |(\Gamma X)_s - (\Gamma X)_{t_k}|_{\mathbb{R}^d} \big) ds \\
&+ C \sum_{k=0}^{n-1} \sum_{j=1}^{r} \sum_{m=1}^{r} |B^j_{t_{k+1}} - B^j_{t_k}| |B^m_{t_{k+1}} - B^m_{t_k}| \\
&\quad \times \int_{t_k}^{t_{k+1}} \frac{t_{k+1} - s}{(t_{k+1} - t_k)^2} \big(s - t_k + |X^{\mathrm{WZ}}_s - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d} + |(\Gamma X^{\mathrm{WZ}})_s - (\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d} \big) ds
\end{aligned}
$$

$$
\begin{aligned}
&\quad + C\sum_{j=1}^{r}\sum_{k=0}^{n-1}|B^j_{t_{k+1}} - B^j_{t_k}|\left(t_{k+1} - t_k + \mathrm{Var}_{[t_k,t_{k+1}]}(A(X^{\mathrm{WZ}}))\right) \\
&\le C\sum_{k=0}^{n-1}(t_{k+1} - t_k)\left(|X_{t_k} - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d} + |(\Gamma X)_{t_k} - (\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right) + \frac{1}{2}\left|\sum_{j=1}^{r}\sum_{m=1}^{r}\nu^{ijm}_{t_n}\right| \\
&\quad + C\sum_{k=0}^{n-1}(t_{k+1} - t_k)\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]}|X_s - X_{t_k}|_{\mathbb{R}^d}\right) \\
&\quad + C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r}|B^j_{t_{k+1}} - B^j_{t_k}||B^m_{t_{k+1}} - B^m_{t_k}| \\
&\qquad \times \sup_{s\in[t_k,t_{k+1}]}\left(s - t_k + |X^{\mathrm{WZ}}_s - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d} + |(\Gamma X^{\mathrm{WZ}})_s - (\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right) \\
&\quad + C\sum_{j=1}^{r}\sum_{k=0}^{n-1}|B^j_{t_{k+1}} - B^j_{t_k}|\left(t_{k+1} - t_k + \mathrm{Var}_{[t_k,t_{k+1}]}(A(X^{\mathrm{WZ}}))\right) \\
&\le C\int_0^{t_n}\sup_{u\in[0,s]}|X_u - X^{\mathrm{WZ}}_u|_{\mathbb{R}^d}ds + \frac{1}{2}\max_{k=0,1,\dots,n}\left|\sum_{j=1}^{r}\sum_{m=1}^{r}\nu^{ijm}_{t_k}\right| \\
&\quad + C\sum_{k=0}^{n-1}(t_{k+1} - t_k)\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]}|X_s - X_{t_k}|_{\mathbb{R}^d}\right) \\
&\quad + C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r}|B^j_{t_{k+1}} - B^j_{t_k}||B^m_{t_{k+1}} - B^m_{t_k}|\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]}|X^{\mathrm{WZ}}_s - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d}\right) \\
&\quad + C\sum_{j=1}^{r}\left(\max_{k=0,1,\dots,n-1}|B^j_{t_{k+1}} - B^j_{t_k}|\right)\left(T + \mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right).
\end{aligned}
$$

By taking the L^p-norm of the both sides and applying Lemmas 3.3 and 3.7, Example 5.3 and Hölder's inequality, we obtain

$$
\begin{aligned}
&E\left[\max_{k=0,1,\dots,n}\left|I^i_5(t_k)\right|^p\right]^{1/p} \\
&\le C|\triangle|^{1/2} + C\int_0^{t_n}E\left[\sup_{u\in[0,s]}|X_u - X^{\mathrm{WZ}}_u|^p_{\mathbb{R}^d}\right]^{1/p}ds \\
&\quad + C\sum_{k=0}^{n-1}(t_{k+1} - t_k)\left(|\triangle|^{1/2} + E\left[\sup_{s\in[t_k,t_{k+1}]}|X_s - X_{t_k}|^p_{\mathbb{R}^d}\right]^{1/p}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r} E\Bigg[|B^j_{t_{k+1}} - B^j_{t_k}|^p |B^m_{t_{k+1}} - B^m_{t_k}|^p \\
&\qquad\qquad \times \left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]} |X^{\mathrm{WZ}}_s - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d}\right)^p\Bigg]^{1/p} \\
&+ C\sum_{j=1}^{r}\sum_{k=0}^{n-1} E\left[\left(\max_{k=0,1,\dots,n-1} |B^j_{t_{k+1}} - B^j_{t_k}|^p\right)\left(T + \mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right)^p\right]^{1/p} \\
\le{}& C|\triangle|^{1/2} + C\int_0^{t_n} E\left[\sup_{u\in[0,s]} |X_u - X^{\mathrm{WZ}}_u|^p_{\mathbb{R}^d}\right]^{1/p} ds \\
&+ C\sum_{k=0}^{n-1}\sum_{j=1}^{r}\sum_{m=1}^{r} E\left[|B^j_{t_{k+1}} - B^j_{t_k}|^{3p}\right]^{1/(3p)} E\left[|B^m_{t_{k+1}} - B^m_{t_k}|^{3p}\right]^{1/(3p)} \\
&\qquad\qquad \times E\left[\left(|\triangle|^{1/2} + \sup_{s\in[t_k,t_{k+1}]} |X^{\mathrm{WZ}}_s - X^{\mathrm{WZ}}_{t_k}|_{\mathbb{R}^d}\right)^{3p}\right]^{1/(3p)} \\
&+ C\sum_{j=1}^{r} E\left[\max_{k=0,1,\dots,n-1} |B^j_{t_{k+1}} - B^j_{t_k}|^{2p}\right]^{1/(2p)} E\left[\left(T + \mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right)^{2p}\right]^{1/(2p)} \\
\le{}& C|\triangle|^{1/2} + C\int_0^{t_n} E\left[\sup_{u\in[0,s]} |X_u - X^{\mathrm{WZ}}_u|^p_{\mathbb{R}^d}\right]^{1/p} ds \\
&+ C|\triangle|^{1/2}(1+\log N)^{1/2}\left(T + E\left[\left(\mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right)^{2p}\right]^{1/(2p)}\right).
\end{aligned}
$$

Therefore, once it is shown that

$$
E\left[\left(\mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right)^{2p}\right]^{1/(2p)} \le C, \tag{16}
$$

the desired estimate for I_5 is obtained. Let

$$
\Psi(t) := \sigma(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k}), \quad t \in [t_k, t_{k+1})
$$

for $k = 0, 1, \dots, N-1$. The Burkholder-Davis-Gundy inequality implies

$$
\begin{aligned}
&E\left[\sup_{t\in[0,T]} \left|\sum_{k=0}^{N-1} \frac{t\wedge t_{k+1} - t\wedge t_k}{t_{k+1} - t_k}\sigma(t_k, X^{\mathrm{WZ}}_{t_k}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B_{t_{k+1}} - B_{t_k})\right|^{2p}_{\mathbb{R}^d}\right]^{1/(2p)} \\
&= E\Bigg[\max_{n=0,1,\dots,N-1} \sup_{t\in[t_n,t_{n+1}]} \left|\frac{t - t_n}{t_{n+1} - t_n}\sigma(t_n, X^{\mathrm{WZ}}_{t_n}, (\Gamma X^{\mathrm{WZ}})_{t_n})(B_{t_{n+1}} - B_{t_n})\right.
\end{aligned}
$$

$$
\begin{aligned}
&\qquad\qquad + \sum_{k=0}^{n-1} \sigma(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B_{t_{k+1}} - B_{t_k}) \Bigg|_{\mathbb{R}^d}^{2p} \Bigg]^{1/(2p)} \\
&\le CE\left[\max_{n=0,1,\dots,N-1} |B_{t_{n+1}} - B_{t_n}|_{\mathbb{R}^d}^{2p}\right]^{1/(2p)} + CE\left[\max_{n=0,1,\dots,N-1} \left|\int_0^{t_n} \Psi(s)\,dB_s\right|_{\mathbb{R}^d}^{2p}\right]^{1/(2p)} \\
&\le C\left(\sum_{n=0}^{N-1} E\left[|B_{t_{n+1}} - B_{t_n}|_{\mathbb{R}^d}^{2p}\right]\right)^{1/(2p)} + CE\left[\left(\int_0^T |\Psi(s)|_{\mathbb{R}^d \otimes \mathbb{R}^r}^2\,ds\right)^p\right]^{1/(2p)} \\
&\le C\left(\sum_{n=0}^{N-1} (t_{n+1} - t_n)^p\right)^{1/(2p)} + C \\
&\le C\,.
\end{aligned}
$$

By this inequality, (A3) and (10), we have

$$
\begin{aligned}
&E\left[\left(\mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right)^{2p}\right]^{1/(2p)} \\
&\le CE\left[\left(1 + \|X^{\mathrm{WZ}} - \xi\|_{C([0,T];\mathbb{R}^d)}\right)^{2p}\right]^{1/(2p)} \\
&\le C + CE\left[\sup_{t\in[0,T]} \left|\sum_{k=0}^{N-1} \int_{t\wedge t_k}^{t\wedge t_{k+1}} \sigma(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) \frac{B_{t_{k+1}} - B_{t_k}}{t_{k+1} - t_k}\,ds\right|_{\mathbb{R}^d}^{2p}\right]^{1/(2p)} \\
&\qquad\qquad + CE\left[\sup_{t\in[0,T]} \left|\int_0^t b(s, X^{\mathrm{WZ}})\,ds\right|_{\mathbb{R}^d}^{2p}\right]^{1/(2p)} \\
&\le C + CE\left[\sup_{t\in[0,T]} \left|\sum_{k=0}^{N-1} \frac{t\wedge t_{k+1} - t\wedge t_k}{t_{k+1} - t_k} \sigma(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})(B_{t_{k+1}} - B_{t_k})\right|_{\mathbb{R}^d}^{2p}\right]^{1/(2p)} \\
&\quad + CE\Bigg[\sup_{t\in[0,T]} \Bigg|\sum_{k=0}^{N-1} \int_{t\wedge t_k}^{t\wedge t_{k+1}} \left(\sigma(s, X_s^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_s) - \sigma(t_k, X_{t_k}^{\mathrm{WZ}}, (\Gamma X^{\mathrm{WZ}})_{t_k})\right) \\
&\qquad\qquad\qquad \times \frac{B_{t_{k+1}} - B_{t_k}}{t_{k+1} - t_k}\,ds\Bigg|_{\mathbb{R}^d}^{2p}\Bigg]^{1/(2p)} \\
&\le C + CE\Bigg[\sup_{t\in[0,T]} \Bigg(\sum_{k=0}^{N-1} \frac{|B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d}}{t_{k+1} - t_k} \\
&\qquad \times \int_{t\wedge t_k}^{t\wedge t_{k+1}} \left(s - t_k + |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d} + |(\Gamma X^{\mathrm{WZ}})_s - (\Gamma X^{\mathrm{WZ}})_{t_k}|_{\mathbb{R}^d}\right) ds\Bigg)^{2p}\Bigg]^{1/(2p)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq C + CE\left[\left(\sum_{k=0}^{N-1} |B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d}\left(\sqrt{t_{k+1}-t_k} + \sup_{s\in[t_k,t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}\right)\right)^{2p}\right]^{1/(2p)} \\
&\leq C + C\sum_{k=0}^{N-1} E\left[|B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d}^{2p}\left(\sqrt{t_{k+1}-t_k} + \sup_{s\in[t_k,t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}\right)^{2p}\right]^{1/(2p)} \\
&\leq C + C\sum_{k=0}^{N-1} E\left[|B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d}^{4p}\right]^{1/(4p)} \\
&\qquad\qquad \times E\left[\left(\sqrt{t_{k+1}-t_k} + \sup_{s\in[t_k,t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}\right)^{4p}\right]^{1/(4p)} \\
&\leq C + C\sum_{k=0}^{N-1} E\left[|B_{t_{k+1}} - B_{t_k}|_{\mathbb{R}^d}^{4p}\right]^{1/(4p)} \\
&\qquad\qquad \times \left(\sqrt{t_{k+1}-t_k} + E\left[\sup_{s\in[t_k,t_{k+1}]} |X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}^{4p}\right]^{1/(4p)}\right).
\end{aligned}
$$

Hence, by Lemma 3.3 we obtain

$$
\begin{aligned}
&E\left[\left(\mathrm{Var}_{[0,T]}(A(X^{\mathrm{WZ}}))\right)^{2p}\right]^{1/(2p)} \\
&\leq C + C\sum_{k=0}^{N-1} \sqrt{t_{k+1}-t_k}\left(\sqrt{t_{k+1}-t_k} + C\sqrt{t_{k+1}-t_k}\right) \\
&\leq C.
\end{aligned}
$$

Thus, (16) is proved. □

Now we are going to finish the proof of Theorem 3.2 by using the estimates above. From (15), Lemmas 3.6 and 3.8 we have

$$
\begin{aligned}
&E\left[\max_{k=0,1,\ldots,n} |X_{t_k} - X_{t_k}^{\mathrm{WZ}}|_{\mathbb{R}^d}^p\right] \\
&\leq C\left(|\triangle|^{p/2}(1+\log N)^{p/2} + \int_0^{t_n} E\left[\sup_{u\in[0,s]} \left|X_u - X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right] ds\right)
\end{aligned}
\tag{17}
$$

for $n = 0, 1, \ldots, N$.

For $t \in [t_n, t_{n+1}]$, (17) and Lemma 3.5 imply

$$
E\left[\sup_{s\in[0,t]} |X_s - X_s^{\mathrm{WZ}}|_{\mathbb{R}^d}^p\right]
$$

$$
\begin{aligned}
&\le E\left[\max_{k=0,1,\dots,n}\sup_{s\in[t_k,t_{k+1}]}\left|X_s - X_s^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right] \\
&\le CE\left[\max_{k=0,1,\dots,n}\sup_{s\in[t_k,t_{k+1}]}\left|X_s - X_{t_k}\right|_{\mathbb{R}^d}^p\right] + CE\left[\max_{k=0,1,\dots,n}\sup_{s\in[t_k,t_{k+1}]}\left|X_s^{\mathrm{WZ}} - X_{t_k}^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right] \\
&\qquad + CE\left[\max_{k=0,1,\dots,n}\left|X_{t_k} - X_{t_k}^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right] \\
&\le C\left(|\triangle|^{p/2}(1+\log N)^{p/2} + \int_0^{t_n} E\left[\sup_{u\in[0,s]}\left|X_u - X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right]ds\right).
\end{aligned}
$$

Hence, for $t \in [0, T]$ it holds that

$$
E\left[\sup_{s\in[0,t]}|X_s - X_s^{\mathrm{WZ}}|_{\mathbb{R}^d}^p\right] \le C\left(|\triangle|^{p/2}(1+\log N)^{p/2} + \int_0^t E\left[\sup_{u\in[0,s]}\left|X_u - X_u^{\mathrm{WZ}}\right|_{\mathbb{R}^d}^p\right]ds\right).
$$

Applying Gronwall's inequality, we obtain the estimate of Theorem 3.2.

4. Remark on the stochastic differential equations with the reflecting boundary condition. In this section, we discuss the relation between stochastic differential equations of the Markovian type with the reflecting boundary condition and path-dependent stochastic differential equations, and see that the results in the present paper include some stochastic differential equations of Markovian type with the reflecting boundary condition.

To study the reflection of stochastic processes, we usually consider the Skorohod equation. Let $T > 0$ and let D be a connected domain in $\mathbb{R}^d$. Denote the closure of D by $\overline{D}$ and the boundary of D by ∂D. The solution to the Skorohod equation on D is defined as follows.

DEFINITION 4.1. For given $w \in C([0, T]; \mathbb{R}^d)$ with $w_0 \in \overline{D}$, a pair $(\xi, \phi) \in C([0, T]; \overline{D}) \times C([0, T]; \mathbb{R}^d)$ is called a solution of the Skorohod equation on D, if $\phi_0 = 0$, ϕ has the bounded variation on $[0, T]$,

$$
\begin{aligned}
\xi_t &= w_t + \phi_t, && t \in [0, T], \\
\mathrm{Var}_{[0,t]}(\phi) &= \int_0^t \mathbb{I}_{\partial D}(\xi_s)d\mathrm{Var}_{[0,s]}(\phi), && t \in [0, T],
\end{aligned}
$$

and there exists $\mathbf{n} \in C([0, T]; \mathbb{R}^d)$ such that

$$
\begin{aligned}
&\mathbf{n}_t \in \bigcup_{r>0}\{\tilde{\mathbf{n}} \in \mathbb{R}^d; |\tilde{\mathbf{n}}| = 1, B(\xi_t - r\tilde{\mathbf{n}}, r) \cap D = \emptyset\}, && t \in \{s \in [0, T]; \xi_s \in \partial D\} \\
&\phi_t = \int_0^t \mathbf{n}_s d\mathrm{Var}_{[0,s]}(\phi), && t \in [0, T]
\end{aligned}
$$

where $B(x, r) := \{y \in \mathbb{R}^d; |x - y| < r\}$ for $x \in \mathbb{R}^d$ and $r > 0$.

We remark that for $t \in [0, T]$ $\mathbf{n}_t$ is an inward unit normal vector of ∂D at ξ_t. It is known that; if the Skorohod equation has the existence and the uniqueness of the solution (ξ, ϕ) with respect to w, ξ is a sufficiently nice process to be regarded as "the reflected process of w", and actually the reflected process of w is defined by ξ. When the Skorohod equation has the existence and the uniqueness of the solution, we call the mapping $\Gamma : C([0,T];\mathbb{R}^d) \to C([0,T];\overline{D})$ given by $\Gamma : w \mapsto \xi$ the Skorohod map.

Originally the problem of stochastic differential equations with the reflecting boundary condition was considered on half spaces. In the case that $\overline{D}$ is a half space $\mathbb{R}^d_+ := \{x = (x^1, x^2, \dots, x^d) \in \mathbb{R}^d; x^1 \geq 0\}$, the Skorohod equation has the existence and the uniqueness of the solution and the Skorohod map is given by

$$(\Gamma w)_t = \left(w_t^1 - \inf_{s\in[0,t]} (w_s^1 \wedge 0), w_t^2, \dots, w_t^d \right)$$

(see e.g. [4], [6] or [7].) In particular, the Skorohod map is Lipschitz continuous. The case of general domains was studied by Tanaka [16]. He showed the existence and the uniqueness of the solution in the case that the domain is in a certain class, which includes the convex domains. Later, Tanaka's result was extended to the cases of more general domains by Lions and Sznitman [10] and Saisho [13]. It is known that even in the cases of such a general domain the Skorohod map is $(1/2)$-Hölder continuous (see Theorem 1.1 in [10]). On the other hand, it is also known that, even if the Skorohod equation has the existence and the uniqueness of the solution, there exist some domains such that the Skorohod map is not Lipschitz continuous (see Proposition 4.1 in [7]). Note that a necessary and sufficient condition for the Skorohod map to be Lipschitz continuous is also obtained in [7]. The boundary condition mentioned above is of a simple type of the reflection. We remark that the cases of the more general boundary conditions, for example the oblique reflecting boundary condition and the boundary condition of the Wentzell type, are also studied (see [10] and Section 7 of Chapter IV in [9]).

Now we see the relation between stochastic differential equations with the reflecting boundary condition and path-dependent stochastic differential equations. Let D be a connected domain in $\mathbb{R}^d$ such that the Skorohod equation on D has the existence and the uniqueness of the solution, and denote the Skorohod map by Γ. Consider a stochastic differential equation with the reflecting boundary condition,

$$\begin{cases} dX_t &= \sigma(t, X_t)dB_t + b(t, X_t)dt + d\Phi_t \\ X_0 &= x \in \overline{D} \end{cases} \tag{18}$$

where Φ plays the role of the reflection of X on ∂D, i.e. $\Gamma(X - \Phi) = X$. On the other hand, consider a path-dependent stochastic differential equation

$$\begin{cases} dY_t &= \sigma(t, (\Gamma Y)_t)dB_t + b(t, (\Gamma Y)_t)dt \\ Y_0 &= x \in \overline{D}. \end{cases} \tag{19}$$

Then, there is a one-to-one correspondence between solutions to (18) and (19). Indeed, if Y is a solution to (19), X defined by $X = \Gamma Y$ satisfies (18). While, if X is a solution to (18), then Y defined by

$$Y_t = x + \int_0^t \sigma(s, X_s)dB_s + \int_0^t b(s, X_s)ds, \quad t \in [0,T]$$

satisfies (19). Hence, there is the equivalence between solving (18) and (19). The equivalence is originally introduced in the case of half spaces by Anderson and Orey (see Proposition 1 of [4]).

In view of this fact, the results obtained in the present paper are applicable to the stochastic differential equations of Markovian type with the reflecting boundary condition whose Skorohod map is Lipschitz continuous. Indeed, the assumptions on the coefficients in Section 2 are checked as follows.

PROPOSITION 4.2. *Let* Γ *be the Skorohod map and assume that* Γ *is Lipschitz continuous. If* σ *is an* $\mathbb{R}^d \otimes \mathbb{R}^r$*-valued bounded Lipschitz continuous function on* $[0, T] \times \mathbb{R}^d$*, then* $h(t, w) := \sigma(t, (\Gamma w)_t)$ *satisfies* (F1)*,* (F2) *and* (F3) *in Section* 2 *with a certain constant* K.

PROOF. From the boundedness of σ and the Lipschitz continuity of σ and Γ, (F1) and (F3) are immediately obtained with a certain constant K. To see (F2), let $s \in [0, T]$ and fix s. Let (ξ, ϕ) be the solution to the Skorohod equation with respect to w. Then, it is easy to see that $(\xi(\cdot \wedge s), \phi(\cdot \wedge s))$ is the solution to the Skorohod equation with respect to $w(\cdot \wedge s)$. This fact implies that for $t \in [s, T]$

$$\begin{aligned} |\xi_t - \xi_s|_{\mathbb{R}^d} &= |(\Gamma w)_t - (\Gamma w(\cdot \wedge s))_t|_{\mathbb{R}^d} \\ &\leq \|\Gamma\| \|w - w(\cdot \wedge s)\|_{C([0,t];\mathbb{R}^d)} \end{aligned}$$

where $\|\Gamma\|$ is the Lipschitz constant of Γ. Hence we have

$$|(\Gamma w)_t - (\Gamma w)_s|_{\mathbb{R}^d} \leq \|\Gamma\| \|w(\cdot + s) - w(s)\|_{C([0,t-s];\mathbb{R}^d)} \tag{20}$$

for $s, t \in [0, T]$ such that $s < t$, and $w \in C([0, T]; \mathbb{R}^d)$. From this inequality and the Lipschitz continuity of σ we obtain (F2). □

REMARK 4.3. Theorem 2.2 in [7] implies that the Skorohod map is Lipschitz continuous if the domain is a convex polyhedron. Hence, this fact, Theorem 2.2 and Proposition 4.2 yields the same rate of the convergence as Theorem 3 in [14].

Now we think of the assumptions on the coefficients in Section 3.

PROPOSITION 4.4. *Let* Γ *be the Skorohod map and assume that* Γ *is Lipschitz continuous. Then,* $A(w)_t := (\Gamma w)_t - w_t$ *satisfies* (A1) *and* (A2) *in Section* 3 *with a certain constant* K_A*, and satisfies* (8) *with* $f(t, x) := x$ *for* $t \in [0, T]$ *and* $x \in \mathbb{R}^d$.

PROOF. It is easy to see that A satisfies (A1). From (20) we have for $s < t$, and $w \in C([0, T]; \mathbb{R}^d)$

$$|A(w)_t - A(w)_s|_{\mathbb{R}^d} \leq (1 + \|\Gamma\|) \|w(\cdot + s) - w(s)\|_{C([0,t-s];\mathbb{R}^d)} .$$

Hence, (A2) also holds. □

Even if the Skorohod map is Lipschitz continuous, it is difficult to see that the map A defined in Proposition 4.4 satisfies (A3). However, some sufficient conditions on the domains for (A3) have been concerned (see Lemma 2.6 in [16]). In particular, (A3) is satisfied when

$\overline{D}$ is a half space. This fact immediately follows from the explicit form of the Skorohod map. These cases are to be examples of the result in Section 3.

5. Appendix. In this section, we consider an upper estimate for the p-th moment of the maximum of random variables.

PROPOSITION 5.1. *Let ϕ be a non-negative, strictly increasing and convex function on $[0, \infty)$, and $\{X_k\}$ be a sequence of random variables such that $E[\phi(|X_k|)] < \infty$ for $k \in \mathbb{N}$. Then,*

$$E\left[\max_{k=1,2,\dots,n} |X_k|\right] \leq \phi^{-1}\left(E\left[\max_{k=1,2,\dots,n} \phi\left(|X_k|\right)\right]\right) \leq \phi^{-1}\left(\sum_{k=1}^{n} E\left[\phi\left(|X_k|\right)\right]\right)$$

for $n \in \mathbb{N}$ where $\phi^{-1}(t) := \inf\{s \in [0, \infty); \phi(s) > t\}$.

PROOF. Since

$$\phi\left(\max_{k=1,2,\dots,n} |X_k|\right) = \max_{k=1,2,\dots,n} \phi(|X_k|),$$

by Jensen's inequality we have

$$\begin{aligned}\phi\left(E\left[\max_{k=1,2,\dots,n} |X_k|\right]\right) &\leq E\left[\phi\left(\max_{k=1,2,\dots,n} |X_k|\right)\right] \\ &= E\left[\max_{k=1,2,\dots,n} \phi\left(|X_k|\right)\right].\end{aligned}$$

Therefore, we obtain the first inequality. The second inequality is obvious. □

PROPOSITION 5.2. *Let $\{X_k\}$ be a sequence of random variables. Then,*

$$\begin{aligned}E\left[\max_{k=1,2,\dots,n} |X_k|^p\right] &\leq \max\{0, p-1\}^p + \left[\log\left(E\left[\max_{k=1,2,\dots,n} e^{|X_k|}\right]\right)\right]^p \\ &\leq \max\{0, p-1\}^p + \left[\log\left(\sum_{k=1}^{n} E\left[e^{|X_k|}\right]\right)\right]^p\end{aligned}$$

for $n \in \mathbb{N}$ and $p \in (0, \infty)$.

PROOF. Define a function f on $[0, \infty)$ by

$$f(x) := \exp\left(x^{1/p}\right),$$

and let $x_0 := \max\{0, p-1\}^p$. Then, it is easy to see that f is strictly increasing on $[0, \infty)$ and convex on $[x_0, \infty)$. Hence, by Proposition 5.1 we have for $n \in \mathbb{N}$

$$
\begin{aligned}
&E\left[\left(\max_{k=1,2,\dots,n}|X_k|\right)^p ; \max_{k=1,2,\dots,n}|X_k| \geq x_0^{1/p}\right] \\
&= E\left[\max_{k=1,2,\dots,n}|X_k|^p \, \mathbb{I}_{\left[x_0^{1/p},\infty\right)}\left(\max_{k=1,2,\dots,n}|X_k|\right)\right] \\
&\leq f^{-1}\left(E\left[f\left(\max_{k=1,2,\dots,n}|X_k|^p \, \mathbb{I}_{\left[x_0^{1/p},\infty\right)}\left(\max_{k=1,2,\dots,n}|X_k|\right)\right)\right]\right) \\
&\leq f^{-1}\left(E\left[f\left(\max_{k=1,2,\dots,n}|X_k|^p\right)\right]\right) \\
&\leq f^{-1}\left(E\left[\max_{k=1,2,\dots,n} f\left(|X_k|^p\right)\right]\right) \\
&= \left[\log\left(E\left[\max_{k=1,2,\dots,n} e^{|X_k|}\right]\right)\right]^p .
\end{aligned}
$$

On the other hand, for $n \in \mathbb{N}$

$$
E\left[\left(\max_{k=1,2,\dots,n}|X_k|\right)^p ; \max_{k=1,2,\dots,n}|X_k| < x_0^{1/p}\right] \leq \max\{0, p-1\}^p .
$$

Therefore, we obtain the first inequality. The second inequality is obvious. □

EXAMPLE 5.3. Let B be the one-dimensional Brownian motion, $\triangle = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ be a partition of the interval $[0, T]$, and $p \in [2, \infty)$ Then, it holds that

$$
E\left[\max_{k=0,1,\dots,N-1}|B_{t_{k+1}} - B_{t_k}|^p\right] \leq 2^p |\triangle|^{p/2}\left[\left(\frac{p}{2}-1\right)^{p/2} + \left(\frac{1}{2}\log 2 + \log N\right)^{p/2}\right]
$$

for $n \in \mathbb{N}$ and $p \in (0, \infty)$. Indeed, applying Proposition 5.2 to the sequence of random variables $\left\{\frac{|B_{t_{k+1}} - B_{t_k}|^2}{4(t_{k+1}-t_k)}\right\}$, we have

$$
\begin{aligned}
&E\left[\max_{k=0,1,\dots,N-1}\left(\frac{|B_{t_{k+1}} - B_{t_k}|^2}{4(t_{k+1}-t_k)}\right)^{p/2}\right] \\
&\leq \left(\frac{p}{2}-1\right)^{p/2} + \left[\log\left(\sum_{k=0}^{N-1} E\left[\exp\left(\frac{|B_{t_{k+1}} - B_{t_k}|^2}{4(t_{k+1}-t_k)}\right)\right]\right)\right]^{p/2} \\
&= \left(\frac{p}{2}-1\right)^{p/2} + \left(\log(\sqrt{2}N)\right)^{p/2} .
\end{aligned}
$$

REFERENCES

[1] S. AIDA, Reflected rough differential equations, Stochastic Process. Appl. 125 (2015), no. 9, 3570–3595.

[2] S. AIDA, Wong-Zakai approximation of solutions to reflecting stochastic differential equations on domains in Euclidean spaces II, Stochastic Analysis and Applications 2014, 1–23, Springer Proc. Math. Stat. 100, Springer, Cham, 2014.

[3] S. AIDA AND K. SASAKI, Wong-Zakai approximation of solutions to reflecting stochastic differential equations on domains in Euclidean spaces, Stochastic Process. Appl. 123 (2013), no. 10, 3800–3827.

[4] R. F. ANDERSON AND S. OREY, Small random perturbations of dynamical systems with reflecting boundary, Nagoya Math. J. 60 (1976), 189–216.

[5] V. BALLY, Approximation for the solutions of stochastic differential equations. I: L^p-convergence, Stochastics Stochastics Rep. 28 (1989), no. 3, 209–246.

[6] H. DOSS AND P. PRIOURET, Support d'un processus de réflexion, Z. Wahrsch. Verw. Gebiete 61 (1982), no. 3, 327–345.

[7] P. DUPUIS AND H. ISHII, On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications, Stochastics Stochastics Rep. 35 (1991), no. 1, 31–62.

[8] L. C. EVANS AND D. W. STROOCK, An approximation scheme for reflected stochastic differential equations, Stochastic Process. Appl. 121 (2011), no. 7, 1464–1491.

[9] N. IKEDA AND S. WATANABE, Stochastic differential equations and diffusion processes, volume 24 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam, second edition, 1989.

[10] P.-L. LIONS AND A.-S. SZNITMAN, Stochastic differential equations with reflecting boundary conditions, Comm. Pure Appl. Math. 37 (1984), no. 4, 511–537.

[11] R. PETTERSSON, Approximations for stochastic differential equations with reflecting convex boundaries, Stochastic Process. Appl. 59 (1995), no. 2, 295–308.

[12] P. E. PROTTER, Stochastic integration and differential equations, volume 21 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, Second edition, Version 2.1, Corrected third printing, 2005.

[13] Y. SAISHO, Stochastic differential equations for multidimensional domain with reflecting boundary, Probab. Theory Related Fields 74 (1987), no. 3, 455–477.

[14] L. SŁOMIŃSKI, On approximation of solutions of multidimensional SDEs with reflecting boundary conditions, Stochastic Process. Appl. 50 (1994), no. 2, 197–219.

[15] L. SŁOMIŃSKI, On Wong-Zakai type approximations of reflected diffusions, Electron. J. Probab. 19 (2014), no. 118, 15 pp.

[16] H. TANAKA, Stochastic differential equations with reflecting boundary condition in convex regions, Hiroshima Math. J. 9 (1979), no. 1, 163–177.

[17] E. WONG AND M. ZAKAI, On the relation between ordinary and stochastic differential equations, Internat. J. Engrg. Sci. 3 (1965), 213–229.

[18] T. ZHANG, Strong convergence of Wong-Zakai approximations of reflected SDEs in a multidimensional general domain, Potential Anal. 41 (2014), no. 3, 783–815.

GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
THE UNIVERSITY OF TOKYO
3–8–1 KOMABA, MEGURO-KU
TOKYO 153–8914
JAPAN

E-mail address: aida@ms.u-tokyo.ac.jp

RESEARCH INSTITUTE FOR INTERDISCIPLINARY SCIENCE
OKAYAMA UNIVERSITY
3–1–1 TSUSHIMANAKA, KITA-KU
OKAYAMA 700–8530
JAPAN

E-mail address: kusuoka@okayama-u.ac.jp

HITACHI POWER SOLUTIONS CO., LTD
3–2–2 SAIWAI-CHO
HITACHI 317–0073
JAPAN

E-mail address: takanori.kikuchi@hotmail.com