J. Math. Kyoto Univ. (JMKYAZ)
22-2 (1982) 243-254

On the mollifier approximation for solutions of stochastic differential equations

By

Shu Jia GUO

(Received Jan. 24, 1981)

§ 0. Introduction

Consider the following stochastic differential equation (SDE) on $\boldsymbol{R}^d$

$$(0.1)\quad \begin{cases} dx_t^i = \sum_{\beta=1}^{r} \sigma_\beta^i(X(t))\circ dW^\beta(t) + b^i(X(t))dt \\ \quad = \sum_{\beta=1}^{r} \sigma_\beta^i(X(t))\cdot dW^\beta(t) + \left[\frac{1}{2}\sum_{j=1}^{d}\sum_{\beta=1}^{r}\left(\frac{\partial\sigma_\beta^i}{\partial x^j}\sigma_\beta^j\right)(X(t)) + b^i(X(t))\right]dt, \\ X(0) = x\in\boldsymbol{R}^d \qquad i=1,2,\dots,d \end{cases}$$

with sufficiently smooth functions $\sigma_\beta^i(x)$ and $b^i(x)$ on $\boldsymbol{R}^d$. Here, $\circ dW^\beta(t)$ and $\cdot dW^\beta(t)$ denote the stochastic differentials of the *Stratonovich type* and of the *Itô type* respectively, and $W(t)=W(t, w)=(W^\beta(t))$, where $W(t, w)=w(t)$, $w\in W_0^r$, is the canonical realization of the r-dimensional Wiener process on the r-dimensional Wiener space $(W_0^r, \boldsymbol{P}^W)$: W_0^r is the space of all continuous functions w: $[0, \infty)\rightarrow\boldsymbol{R}^d$ such that $w(0)=0$ and $\boldsymbol{P}^W$ is the r-dimensional Wiener measure on W_0^r. Introducing vector fields $A_0, A_1,\dots, A_r$ on $\boldsymbol{R}^d$ by

$$A_\beta(x) = \sum_{i=1}^{d}\sigma_\beta^i(x)\frac{\partial}{\partial x^i}, \qquad \beta=1,2,\dots,r$$

$$A_0(x) = \sum_{i=1}^{d} b^i(x)\frac{\partial}{\partial x^i},$$

the equation (0.1) is also denoted by

$$(0.1)'\quad \begin{cases} dX(t) = \sum_{\beta=1}^{r} A_\beta(X(t))\circ dW^\beta(t) + A_0(X(t))dt \\ X(0) = x. \end{cases}$$

If $\sigma_\beta^i(x)$ and $b^i(x)$ are $\boldsymbol{C}^\infty$ with bounded derivatives of all orders, the solution $X(t, x, w)$ exists globally and for $a.a.w(\boldsymbol{P}^W)$, $x\rightarrow X(t, x, w)$ is a diffeomorphism of $\boldsymbol{R}^d$ for each $t\geqq 0$ (cf. [1], [3]).

Let $W_\delta(t)=(W_\delta^\beta(t))_{\beta=1}^r(\delta>0)$ be an approximation of the Wiener process $W(t)$, i.e. the process defined on (W_0^r, P^W) which consists of smooth paths and which approximates $W(t)$ as $\delta\downarrow 0$. Then we can consider a dynamical system, i.e., an ordinary differential equation (ODE)

$$(0.2)\qquad \begin{cases} \dot{X}_\delta(t)=\sum_{\beta=1}^r A_\beta(X_\delta(t))\dot{W}_\delta^\beta(t)+A_0(X_\delta(t)) \\ X_\delta(0)=x, \qquad \left(\cdot=\frac{d}{dt}\right) \end{cases}$$

and we obtain a family $(X_\delta(t, x, w))$ of diffeomorphisms over $\boldsymbol{R}^d$ defined by the solution of (0.2). It is reasonable to expect for a class of nice approximations that $X_\delta(t, x, w)$ actually approximates $X(t, x, w)$. In fact, for the piecewise linear approximation, this approximation of diffeomorphisms was obtained by Elworthy [2], Ikeda-Watanabe [3] and Bismut [1], and for the mollifier approximation (a regularization by convolutions) it was discussed by Malliavin [4]. In particular, Malliavin called this approximation the *transfer principle* and regarded it a fundamental principle in studying the flow of diffeomorphisms $X(t, x, w)$. It seems difficult, however, to follow his proof in several points. Main objective of the present paper is to give a rigorous proof of the mollifier approximation by modifying the method of [3] in the case of piecewise linear approximation.

The author would like to express his hearty graditude to Professor S. Watanabe who kindly guided his research while he was staying at Kyoto University and gave him many advices in writing this article.

§ 1. Mollifier approximation

Let (W_0^r, P^W) be the r-dimensional Wiener space and $\mathscr{B}_t=\mathscr{B}_t(W_0^r)$ be the usual σ-field generated by the paths up to time t. Let ρ be a C^∞-function with support in [0, 1] such that $\rho\geqq 0$ and $\int_0^1\rho(t)dt=1$. Upon choosing such a function, we set for each $\delta>0$

$$(1.1)\qquad W_\delta^i(t)=\int_0^\delta W^i(t+s, w)\rho\left(\frac{s}{\delta}\right)\frac{ds}{\delta}, \qquad i=1, 2,\dots, r$$

and call $W_\delta(t)=(W_\delta^i(t))$ a mollifier approximation of $W(t, w)$. In order to emphasize the dependence of W_δ on w, we often denote $W_\delta(t)=W_\delta(t, w)$. It is easy to verify the following properties of the mollifier approximation:

(i) $t\to W_\delta(t)$ is C^∞ as a map: $(0, \infty)\to\boldsymbol{R}^d$ and

$\sup_{t\in[0,T]}|W_\delta(t, w)-W(t, w)|\to 0$ as $\delta\downarrow 0$ for every $T>0$ and $w\in W_0^r$,

(ii) for any $t\geqq 0$, $W_\delta(t)$ is $\mathscr{B}_{t+\delta}$-measurable,

(iii) if θ_t: $W_0^r\to W_0^r$ is defined by $(\theta_t w)(s)=w(t+s)-w(t)$, then for all t, $s\geqq 0$,

$$W_\delta(t+s, w) = W_\delta(t, \theta_s w) + W(s, w),$$

(iv) $E[W_\delta^i(t)] = 0,\ t \geqq 0,\ i = 1, 2, \dots, r$
(E denotes the expectation with respect to P^W.)

(v) $E[|W_\delta(0)|^{2m}] = e_{2m}\delta^m,\ m = 1, 2, \dots$
where e_{2m} is a positive constant depending only on $2m$,

(vi) $E\left[\left(\int_0^\delta |\dot{W}_\delta(s)|ds\right)^{2m}\right] = e'_{2m}\delta^m, \qquad m = 1, 2, \dots$
where e'_{2m} is a positive constant depending only on $2m$,

We consider SDE (0.1) and ODE (0.2) where $\sigma_\beta^i(x)$ and $b^i(x) \in C_b^\infty(R^d)$ i.e., σ_β^i and b^i together with their derivatives of all orders are continuous and bounded. Now we can state the main result of this paper as follows:

Theorem. *For all* $p \geqq 1$, $T > 0$, $N > 0$ *and multi-index* α, *we have*

$$\lim_{\delta \downarrow 0} E[\sup_{0 \leqq t \leqq T} \sup_{|x| \leqq N} |D_x^\alpha X_\delta(t, x, w) - D_x^\alpha X(t, x, w)|^p] = 0. \tag{1.2}$$

Here $D_x^\alpha = \dfrac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_d^{\alpha_d}}$, $\alpha = (\alpha_1, \dots, \alpha_d)$ and $|\alpha| = \alpha_1 + \cdots + \alpha_d$.

To prove this theorem we need the following result which has been obtained in [3] (Chapter VI, **Theorem** 7.2): *if, in the equations* (0.1) *and* (0.2), $\sigma_\beta^i \in C_b^2(R^d)$ *and* $b^i \in C_b^1(R^d)$ (*in general,* $f \in C_b^m(R^d)$ *means that* f *together with its derivatives up to the m-th order are continuous and bounded*), *then for every* $T > 0$ *and* $N > 0$,

$$\lim_{\delta \downarrow 0} \sup_{|x| \leqq N} E[\sup_{0 \leqq t \leqq T} |X_\delta(t, x, w) - X(t, x, w)|^2] = 0. \tag{1.3}$$

The theorem can be obtained by the following reasoning. First we remark that (1.2) is deduced from the following weaker estimate: for every $p \geqq 1$, $T > 0$, $N > 0$ and multi-index α,

$$\lim_{\delta \downarrow 0} \sup_{|x| \leqq N} E[\sup_{0 \leqq t \leqq T} |D_x^\alpha X_\delta(t, x, w) - D_x^\alpha X(t, x, w)|^p] = 0. \tag{1.4}$$

This can be seen by the same arguments as in Chap. V, Section 2 of [3]. So we need only to prove (1.4). For this we remark the following: if $X^{(\alpha)} = D_x^\alpha X$, then $Z = (X^{(\alpha)})_{|\alpha| \leqq m}$ is the solution of the following SDE in the matrix notation,

$$\left\{\begin{array}{l} dX(t) = \sigma_\beta(X(t)) \circ dW^\beta(t) + b(X(t))dt \\ dY(t) = (D\sigma_\beta)(X(t))Y(t) \circ dW^\beta(t) + (Db)(X(t))Y(t)dt \\ \vdots \\ X(0) = x \\ Y(0) = I \\ \vdots \end{array}\right. \tag{1.5}$$

where $X^0(t) = X(t)$, $Y(t) = (D_x^\alpha X;\ |\alpha| = 1) = \left(\dfrac{\partial X^i}{\partial x^j}\right), \dots, (D\sigma_\beta) = \left(\dfrac{\partial \sigma_\beta^i}{\partial x^j}\right), (Db) =$

$\left(\frac{\partial b^i}{\partial x^j}\right), \ldots$. If we can apply the above proposition (1.3) to this SDE and the corresponding ODE (so, of course, d should be changed to bigger one), we can obtain (1.4). Unfortunately, the coefficients of the equation (1.5) are not bounded even if we assume σ^i_β, $b^i \in C^\infty_b(R^d)$ and hence (1.3) can not be applied directly. However, if we can verify the condition: *for every* $p>0$, $T>0$ *and* $N>0$,

$$\sup_{|x|\leqq N} \sup_{\delta>0} E\,[\sup_{0\leqq t\leqq T} |D^\alpha_x X_\delta(t, x, w)|^p] < \infty, \tag{1.6}$$

then we can apply the same truncation argument as in the proof of Lemma 2.1 of Chapter V in [3] to obtain (1.4). *In conclusion, all we need for the proof of the theorem is the estimate* (1.6).

§2. The proof of the estimate (1.6)

The proof of (1.6) can be carried over in a similar way as in the proof of Lemma 7.2, Chapter VI of [3]. Since $W_\delta(t)$ is $\mathscr{B}_{t+\delta}$-measurable for every t, however, it needs to be modified in several points. The term involving the drift coefficients b^i do not cause any difficulty and so, just by the reason of simplicity, we assume $b^i=0$ in the following discussions. Thus, instead of (0.1) and (0.2) we consider the equations

$$\begin{cases} dX(t) = \sum_{\beta=1}^{r} \sigma_\beta(X(t)) \circ dW^\beta(t) \\ X(0) = x \end{cases} \tag{2.1}$$

and

$$\begin{cases} \dot{X}_\delta(t) = \sum_{\beta=1}^{r} \sigma_\beta(X_\delta(t))\, \dot{W}^\beta_\delta(t). \\ X_\delta(0) = x. \end{cases} \tag{2.2}$$

First we consider the case $\alpha=(0,\ldots,0)$, i.e., $D^\alpha X_\delta = X_\delta$, and assume that $T>0$, $N>0$ and $p\geqq 2$ are given arbitrarily. Denoting by $[x]$ the largest integer not exceeding x as usual, we have for any $t\geqq 0$

$$\begin{aligned} &X_\delta(t) - x \\ &= \sum_{k=0}^{[t/\delta]-1} \sigma(X_\delta((k-1)\delta))[W_\delta((k+1)\delta) - W_\delta(k\delta)] \\ &\quad + \sum_{k=0}^{[t/\delta]-1} \int_{k\delta}^{(K+1)\delta} [\sigma(X_\delta(s)) - \sigma(X_\delta((k-1)\delta))]\, \dot{W}_\delta(s)\, ds \\ &\quad + \sigma(X_\delta(([t/\delta]-1)\delta))[W_\delta(t) - W_\delta([t/\delta]\delta)] \\ &\quad + \int_{[t/\delta]\delta}^{t} [\sigma(X_\delta(s)) - \sigma(X_\delta(([t/\delta]-1)\delta))]\, \dot{W}_\delta(s)\, ds \\ &\triangleq I_1(t) + I_2(t) + I_3(t) + I_4(t). \end{aligned} \tag{2.3}$$

Defining the function $\phi_\delta(s)$ by

$$\phi_\delta(s)=\begin{cases}(k-1)\delta & \text{if } k\delta<s\leqq(k+1)\delta\\ 0 & \text{if } s=0\end{cases}$$

and setting $X_\delta(-\delta)=0$, we have

$$\begin{aligned}
&\boldsymbol{E}[\sup_{0\leqq t\leqq T}|I_1(t)|^p]\\
&=\boldsymbol{E}[\sup_{0\leqq t\leqq T}|\sum_{k=0}^{[t/\delta]-1}\int_0^\delta \rho_\delta(t')\{\sigma(X_\delta((k-1)\delta))[W((k+1)\delta+t')\\
&\quad -W(k\delta+t')]\}dt'|^p]\\
&=\boldsymbol{E}\left[\sup_{0\leqq t\leqq T}\left|\left\{\int_0^\delta \rho_\delta(t')\left\{\int_0^{[t/\delta]\delta}\sigma(X_\delta(\phi_\delta(s)))dW(s+t')\right\}dt'\right|^p\right]\\
&\leqq\boldsymbol{E}\left[\sup_{0\leqq t\leqq T}\left\{\int_0^\delta \rho_\delta(t')\left|\int_0^{[t/\delta]\delta}\sigma(X_\delta(\phi_\delta(s)))dW(s+t')\right|dt'\right\}^p\right]\\
&\leqq\boldsymbol{E}\left[\sup_{0\leqq t\leqq T}\left\{\int_0^\delta \rho_\delta(t')\left|\int_0^{[t/\delta]\delta}\sigma(X_\delta(\phi_\delta(s)))\,dW(s+t')\right|^p dt'\right\}\right]\\
&\leqq\int_0^\delta \rho_\delta(t')\boldsymbol{E}\left[\sup_{0\leqq t\leqq T}\left|\int_0^{[t/\delta]\delta}\sigma(X_\delta(\phi_\delta(s)))dW(s+t')\right|^p\right]dt'.
\end{aligned}$$

Applying a standard moment inequality for martingales ([3], Chapter III. Section 3), we obtain

$$\begin{aligned}
&\boldsymbol{E}[\sup_{0\leqq t\leqq T}|I_1(t)|^p]\\
&\leqq K_1\int_0^\delta \rho_\delta(t')\boldsymbol{E}\left[\left(\int_0^{[t/\delta]\delta}\|\sigma(X_\delta(\phi_\delta(s))\|^2ds\right)^{\frac{p}{2}}\right]dt' \qquad (2.4)\\
&\leq K_2.
\end{aligned}$$

(Here and in the following, $K_1, K_2,\ldots,$ are constants independent of $\delta>0$). $I_2(t)$ can be estimated as follows:

$$\begin{aligned}
&\boldsymbol{E}[\sup_{0\leqq t\leqq T}|I_2(t)|^p]\\
&\leqq\boldsymbol{E}\left[\sup_{0\leqq t\leqq T}\left(\sum_{k=0}^{[t/\delta]-1}\int_{k\delta}^{(k+1)\delta}\|\sigma(X_\delta(s))-\sigma(X_\delta((k-1)\delta))\|\cdot\|\dot W_\delta(s)\|ds\right)^p\right]\\
(2.5)\quad&\leqq K_3\delta^{-(p-1)}\sum_{K=0}^{[t/\delta]-1}\boldsymbol{E}\left[\left(\int_{k\delta}^{(k+1)\delta}\left\|\int_{(k-1)\delta}^s\sigma(X_\delta(u))\,\dot W_\delta(u)du\right\|\,\|\dot W_\delta(s)\|ds\right)^p\right]\\
&\leqq K_4\delta^{-(p-1)}\sum_{k=0}^{[T/\delta]-1}\boldsymbol{E}\left[\left(\int_{(k-1)\delta}^{(k+1)\delta}\|\dot W_\delta(u)\|ds\right)^{2p}\right]\\
&\leqq K_5
\end{aligned}$$

by the property (vi) in the section 1. The proof of the estimates

$$E[\sup_{0\leq t\leq T}|I_3(t)|^p]\leqq K_6$$

and

$$E[\sup_{0\leq t\leq T}|I_4(t)|^p]\leqq K_7$$

can be given similarly as (actually even more easily than) (2.4) and (2.5) and hence

$$E[\sup_{0\leq t\leq T}\|X(t, x, w)\|^p]\leqq K_8(1+|x|^p)$$

completing the proof of (1.6) in the case of $\alpha=(0, 0,\dots, 0)$.

Next, we consider the case of $|\alpha|=1$, i.e., the case of the first order derivatives. Setting $Y_\delta(t, x, w)=\left(\frac{\partial}{\partial x^j}X_\delta^i(t, x, w)\right)$ and $D\sigma=\left(\frac{\partial}{\partial x^j}\sigma_\beta^i\right)$, we have

$$Y_\delta(t, x, w)=I+\int_0^t D\sigma(X_\delta(s))\,Y_\delta(s, x, w)\,\dot{W}_\delta(s)ds \tag{2.6}$$

(in the matrix notation: to be precise, $(D\sigma(X_\delta(s))Y_\delta(s, x, w)\dot{W}_\delta(s))^i_j=\sum_{k=1}^d\sum_{\beta=1}^r\frac{\partial}{\partial x^k}\sigma_\beta^i(X_\delta(s))\frac{\partial}{\partial x^j}X_\delta^k(s, x, w)\dot{W}_\delta^\beta(s)$). Consequently we have for any $t\geqq 0$,

$$\begin{aligned}
&Y_\delta(t)-I\\
&=\sum_{k=0}^{[t/\delta]-1}D\sigma(X_\delta((k-1)\delta))\,Y_\delta((k-1)\delta)[W_\delta((k+1)\delta)-W_\delta(k\delta)]\\
&\quad+\sum_{k=0}^{[t/\delta]-1}\int_{k\delta}^{(k+1)\delta}[D\sigma(X_\delta(s))\,Y_\delta(s)-D\sigma(X_\delta((k-1)\delta))\,Y_\delta((k-1)\delta)]\,\dot{W}_\delta(s)ds\\
&\quad+D\sigma(X_\delta(([t/\delta]-1)\delta))\,Y_\delta(([t/\delta]-1)\delta)[W_\delta(t)-W_\delta([t/\delta]\delta)]\\
&\quad+\int_{[t/\delta]\delta}^t[D\sigma(X_\delta(s))\,Y_\delta(s)-D\sigma(X_\delta(([t/\delta]-1)\delta)\,Y_\delta(([t/\delta]-1)\delta)]\,\dot{W}_\delta(s)ds\\
&\triangleq J_1(t)+J_2(t)+J_3(t)+J_4(t)
\end{aligned} \tag{2.7}$$

where we set $Y_\delta(-\delta)=I$. By the same estimate as for $I_1(t)$, we obtain for any $t_1\in[0, T]$

$$E[\sup_{0\leq t\leq t_1}|J_1(t)|^p]\leqq K_9\int_0^{t_1}E[\sup_{0\leq s\leq t}\|Y(s)\|^p]dt. \tag{2.8}$$

As for $J_2(t)$,

$$\begin{aligned}
&E[\sup_{0\leq t\leq t_1}|J_2(t)|^p]\\
&\leqq E\Big[\sup_{0\leq t\leq t_1}\Big(\sum_{k=0}^{[t/\delta]-1}\int_{k\delta}^{(k+1)\delta}\|D\sigma(X_\delta(s))\,Y_\delta(s)\\
&\quad-D\sigma(X_\delta((k-1)\delta))\,Y_\delta((k-1)\delta)\|\cdot\|\dot{W}_\delta(s)\|ds)^p\Big]
\end{aligned}$$

$$\leqq K_{10}\Big\{\boldsymbol{E}\Big[\Big(\sum_{k=0}^{[t_1/\delta]-1}\int_{k\delta}^{(k+1)\delta}\|D\sigma(X_\delta(s))Y_\delta(s)-D\sigma(X_\delta(k\delta))Y_\delta(k\delta)\|\cdot \|\dot{W}_\delta(s)\|ds\Big)^p\Big]+\boldsymbol{E}\Big[\Big(\sum_{k=0}^{[t_1/\delta]-1}\int_{k\delta}^{(k+1)\delta}\|D\sigma(X_\delta(k\delta))Y_\delta(k\delta) - D\sigma(X_\delta((k-1)\delta))Y_\delta((k-1)\delta)\|\|\dot{W}_\delta(s)\|ds\Big)^p\Big]\Big\}. \tag{2.9}$$

If $k\delta\leqq s\leqq(k+1)\delta$,

$$\begin{aligned}&\|D\sigma(X_\delta(s))Y_\delta(s)-D\sigma(X_\delta(k\delta))Y_\delta(k\delta)\|\\ &\leqq\int_{k\delta}^{s}\|Y_\delta(u)\|\|D^2\sigma(X_\delta(u))\|\|\sigma(X_\delta(u))\|\|\dot{W}_\delta(u)\|du\\ &\quad+\int_{k\delta}^{s}\|D\sigma(X_\delta(u))\|^2\|Y_\delta(u)\|\|\dot{W}_\delta(u)\|du \qquad (2.10)\\ &\leqq K_{11}\int_{k\delta}^{s}\|Y_\delta(u)\|\|\dot{W}_\delta(u)\|du\\ &\leqq K_{11}\Big[\int_{k\delta}^{s}\|Y_\delta(k\delta)\|\|\dot{W}_\delta(u)\|du+\int_{k\delta}^{s}\|Y_\delta(u)-Y_\delta(k\delta)\|\|\dot{W}_\delta(u)\|du\end{aligned}$$

and if $k\delta\leqq u\leqq(k+1)\delta$

$$\begin{aligned}&\|Y_\delta(u)-Y_\delta(k\delta)\|\\ &\leqq\int_{k\delta}^{u}\|D\sigma(X_\delta(s'))\|\|Y_\delta(s')\|\|\dot{W}_\delta(s')\|ds' \qquad (2.11)\\ &\leqq K_{12}\Big\{\|Y_\delta(k\delta)\|\int_{k\delta}^{(k+1)\delta}\|\dot{W}_\delta(s)\|ds+\int_{k\delta}^{u}\|Y_\delta(s)-Y_\delta(k\delta)\|\|\dot{W}_\delta(s)\|ds\Big\}.\end{aligned}$$

Set $b_k=K_{12}\int_{k\delta}^{(k+1)\delta}\|\dot{W}_\delta(s)\|ds$, From the integral inequality (2.11), we can conclude as usual the following:

$$\|Y_\delta(u)-Y_\delta(k\delta)\|\leqq\|Y_\delta(k\delta)\|C_k \tag{2.12}$$

where $C_k=b_ke^{b_k}$. We may assume that $K_{11}\leqq K_{12}$, then substituting (2.12) into (2.11) yields the following estimate:

$$\|D\sigma(X_\delta(s))Y_\delta(s)-D\sigma(X_\delta(k\delta))Y_\delta(k\delta)\|\leqq\|Y_\delta(k\delta)\|(C_k+C_k^2). \tag{2.13}$$

By substituting (2.13) into (2.9) (we also assume as we may that $K_{10}\leqq K_{12}$), we obtain

$$\begin{aligned}&\boldsymbol{E}[\sup_{0\leqq t\leqq t_1}|J_2(t)|^p]\\ &\leqq K_{13}\{\boldsymbol{E}[(\sum_{k=0}^{[t_1/\delta]-1}\|Y_\delta(k\delta)\|\cdot C_k^2)^p]\\ &\quad+\boldsymbol{E}[(\sum_{k=0}^{[t_1/\delta]-1}\|Y_\delta(k\delta)\|\cdot C_k^3)^p] \qquad (2.14)\end{aligned}$$

$$+\boldsymbol{E}[(\sum_{k=0}^{[t_1/\delta]-1}\|Y_\delta((k-1)\delta)\|C_{k-1}C_k)^p]$$

$$+\boldsymbol{E}[(\sum_{k=0}^{[t_1/\delta]-1}\|Y_\delta((k-1)\delta)\|C_{k-1}^2C_k)^p]\}$$

$$=K_{13}\{J_{21}+J_{22}+J_{23}+J_{24}\}.$$

Now we need the following estimates for the moments of random variables C_k. First the constants e_q' in (vi) of the section 1 have the following estimates:

$$e_q'=\boldsymbol{E}\Big[\Big(\int_0^\delta\|\dot W_\delta(s)\|ds\Big)^q\Big]/\delta^{\frac{q}{2}}$$

$$\leqq r^q\boldsymbol{E}\Big[\Big(\int_0^\delta|\dot{\tilde W}_\delta(s)|ds\Big)^q\Big]/\delta^{\frac{q}{2}}$$

$$=r^q\boldsymbol{E}\Big[\Big(\int_0^1\Big|\int_0^1\tilde W(s+\xi)\rho'(\xi)\Big|d\xi\,ds\Big)^q\Big]$$

$$\leqq(2rM)^q d_q$$

where $\tilde W$ is a 1-dimensional Brownian motion, $M=\|\rho'\|$ and

$$d_q=\begin{cases}\frac{1}{\sqrt{2\pi}}2^{m'}m'! & \text{if } q=2m'-1\\ (2m'-1)!! & \text{if } q=2m'\end{cases}$$

$$=\boldsymbol{E}(|\tilde W(1)|^q).$$

Then for every $p'\geqq 2$ and $k=0,1,2,\ldots$

$$\boldsymbol{E}[C_k^{p'}]\leqq\sum_{l=0}^\infty\frac{(p'K_{12})^l}{l!}(2rM)^{p'+l}d_{p'+l}\delta^{\frac{1}{2}(l+p')}$$

$$\leqq(2rM)^{p'}\sum_{l=0}^\infty(\sqrt{\delta}\,2rMp'K_{12})^l d_{2p'}^{\frac{1}{2}}d_{2l}^{\frac{1}{2}}\delta^{\frac{p'}{2}}/l!$$

(2.15)
$$\leqq(2rM)^{p'}d_{2p'}^{\frac{1}{2}}[\sum_{l=0}^{\infty}(4rMp'K_{12}\sqrt{\delta}))^l]\delta^{\frac{p'}{2}}$$

$$\leqq K_{14}\delta^{\frac{p'}{2}}.$$

By (2.12),

$$\|Y_\delta(k\delta)\|C_k^2$$

$$\leqq\|Y_\delta((k-1)\delta)\|C_k^2+\|Y_\delta((k-1)\delta)\|C_{k-1}C_k^2$$

$$\leqq\|Y_\delta((k-1)\delta)\|C_k^2+\frac{1}{2}\|Y_\delta((k-1)\delta)\|C_k^4+\frac{1}{2}\|Y_\delta((k-1)\delta)\|C_{k-1}^2$$

and continuing this, we obtain

$$\|Y_\delta(k\delta)\|C_k^2$$

(2.16)
$$\leqq \sum_{m=0}^{k} \frac{1}{2^{k-m}} \|Y_\delta((m-1)\delta)\|C_m^2 + \sum_{m=0}^{k} \frac{1}{2^{k+1-m}} \|Y_\delta((m-1)\delta)\|C_m^4.$$

From (2.15) and (2.16) we see that

$$\begin{aligned} J_{21} &= \boldsymbol{E}\Big[\Big(\sum_{k=0}^{[t_1/\delta]-1} \|Y_\delta(k\delta)\|C_k^2\Big)^p\Big] \\ &\leqq K_{15}\left\{\boldsymbol{E}\left[\left(\sum_{k=0}^{[t_1/\delta]-1} \sum_{m=0}^{k} \frac{1}{2^{k-m}} \|Y_\delta((m-1)\delta)\|C_m^2\right)^p\right]\right. \\ &\quad \left. + \boldsymbol{E}\left[\left(\sum_{k=0}^{[t_1/\delta]-1} \sum_{m=0}^{k} \frac{1}{2^{k+1-m}} \|Y_\delta((m-1)\delta)\|C_m^4\right)^p\right]\right\} \\ &\leqq K_{16}\delta^{-(p-1)} \sum_{m=0}^{[t_1/\delta]-1} \boldsymbol{E}(\|Y_\delta((m-1)\delta)\|^p)\boldsymbol{E}(C_m^{2p}+C_m^{4p}) \\ &\leqq K_{17}\delta \sum_{m=0}^{[t_1/\delta]-1} \boldsymbol{E}(\|Y_\delta((m-1)\delta)\|^p) \\ &\leqq K_{18}\int_0^{t_1} \boldsymbol{E}[\sup_{0\leqq s\leqq t} \|Y_\delta(s)\|^p]\,dt. \end{aligned}$$

In a similar way, we have

$$J_{22}+J_{23}+J_{24} \leqq K_{19}\int_0^{t_1} \boldsymbol{E}[\sup_{0\leqq s\leqq t} \|Y_\delta(s)\|^p]\,dt$$

and hence we obtain

(2.17)
$$\boldsymbol{E}[\sup_{0\leqq t\leqq t_1} |J_2(t)|^p] \leqq K_{20}\int_0^{t_1} \boldsymbol{E}[\sup_{0\leqq s\leqq t} \|Y_\delta(s)\|^p]\,dt.$$

Similarly as for $J_1(t)$ and $J_2(t)$ (actually even more easily) we can obtain

(2.18)
$$\boldsymbol{E}[\sup_{0\leqq t\leqq t_1} |J_3(t)|^p] \leqq K_{21}\int_0^{t_1} \boldsymbol{E}[\sup_{0\leqq s\leqq t} \|Y_\delta(s)\|^p]\,dt$$

and

(2.19)
$$\boldsymbol{E}[\sup_{0\leqq t\leqq t_1} |J_4(t)|^p] \leqq K_{22}\int_0^{t_1} \boldsymbol{E}[\sup_{0\leqq s\leqq t} \|Y_\delta(s)\|^p]\,dt.$$

By (2.7), (2.8), (2.17), (2.18) and (2.19) we have

(2.20)
$$\boldsymbol{E}[\sup_{0\leqq t\leqq t_1} \|Y_\delta(t)\|^p] \leqq K_{23}\Big(1+\int_0^{t_1} \boldsymbol{E}[\sup_{0\leqq s\leqq t} \|Y_\delta(s)\|^p]\,dt\Big)$$

and we can conclude from this inequality

(2.21)
$$\sup_{x,\delta>0} \boldsymbol{E}(\sup_{0\leqq t\leqq T} \|Y_\delta(t)\|^p) \leqq K_{24}.$$

Next we proceed to the case of $|\alpha|=2$.

Set

$$Y^{i,\delta}_{j_1,j_2}(t,x,w)=\frac{\partial^2}{\partial x_{j_1}\partial x_{j_2}}X^i_\delta(t,x,w), \qquad 1\leqq i,j_1,j_2\leqq d.$$

Then

$$Y^{i,\delta}_{j_1,j_2}(t)=\sum_{k=1}^{d}\sum_{\beta=1}^{r}\int_0^t \sigma'_\beta(X_\delta(s))^i_k Y^{k,\delta}_{j_1,j_2}(s)\dot W^\beta_\delta(s)ds+\alpha^{i,j_1,j_2}_\delta(t) \tag{2.22}$$

where

$$\alpha^{i,j_1,j_2}_\delta(t)=\int_0^t\sum_{k,l=1}^{d}\sum_{\beta=1}^{r}\sigma''_\beta(X_\delta(s))^i_{k,l}Y_\delta(s)^k_{j_1}Y_\delta(s)^l_{j_2}\dot W^\beta_\delta(s)ds \tag{2.23}$$

with

$$\sigma'_\beta(x)^i_j=\frac{\partial}{\partial x_j}\sigma^i_\beta(x) \quad\text{and}\quad \sigma''_\beta(x)^i_{k,l}=\frac{\partial^2}{\partial x_k\partial x_l}\sigma^i_\beta(x).$$

If we denote $\alpha^{i,j_1,j_2}_\delta(t)$ as

$$\begin{aligned}\alpha^{i,j_1,j_2}_\delta(t)=&\sum_{k,l=1}^{d}\sum_{\beta=1}^{r}\Big(\sum_{m=0}^{[t/\delta]-1}\sigma''_\beta(X_\delta((m-1)\delta))^i_{k,l}Y_\delta((m-1)\delta)^k_{j_1}\\&\times Y_\delta((m-1)\delta)^l_{j_2}[W^\beta_\delta((m+1)\delta)-W^\beta_\delta(m\delta)]\Big)\\&+\sum_{k,l=1}^{d}\sum_{\beta=1}^{r}\Big(\sum_{m=0}^{[t/\delta]-1}\int_{m\delta}^{(m+1)\delta}[\sigma''_\beta(X_\delta(s))^i_{k,l}Y_\delta(s)^k_{j_1}Y_\delta(s)^l_{j_2}\\&-\sigma''_\beta(X_\delta((m-1)\delta))^i_{k,l}Y_\delta((m-1)\delta)^k_{j_1}Y_\delta((m-1)\delta)^l_{j_2}]\dot W^\beta_\delta(s)ds\Big)\\&+\sum_{k,l=1}^{d}\sum_{\beta=1}^{r}\int_{[t/\delta]}^{t}\sigma''_\beta(X_\delta(s))^i_{k,l}Y_\delta(s)^k_{j_1}Y_\delta(s)^l_{j_2}\dot W^\beta_\delta(s)ds\\&\triangleq H_1(t)+H_2(t)+H_3(t),\end{aligned} \tag{2.24}$$

$H_1(t)$ can be estimated by the method used in the estimate of $I_1(t)$ as follows:

$$\begin{aligned}E[\sup_{0\leqq t\leqq T}|H_1(t)|^p]&\leqq K_{25}E\Big[\Big(\int_0^T\|Y_\delta(\phi_\delta(s))\|^4ds\Big)^{\frac{p}{2}}\Big]\\&\leqq K_{26}\int_0^T E[\sup_{0\leqq t\leqq T}\|Y_\delta(t)\|^{2p}]ds\leq K_{27}<\infty.\end{aligned} \tag{2.25}$$

As for $H_2(t)$, we estimate it by the method used in the estimate of $J_2(t)$ and obtain

$$\begin{aligned}&E[\sup_{0\leqq t\leqq T}|H_2(t)|^p]\\&\leqq K_{28}E\Big[\Big(\sum_{m=0}^{[T/\delta]-1}\big(\sup_{0\leqq t\leqq T}\|Y_\delta(t)\|^2\big)(C_m^2+C_m^3+C_{m-1}C_m+C_{m-1}^2C_m)\Big)^p\Big]\\&\leqq K_{29}<\infty.\end{aligned} \tag{2.26}$$

In a similar way, we can obtain

$$(2.27)\qquad E[\sup_{0\leq t\leq T}|H_3(t)|^p]\leq K_{33}<\infty.$$

Thus we have proved the estimate for $\alpha_\delta(t, x, w)=(\alpha_\delta^{i,j_1,j_2}(t, x, w))$:

$$(2.28)\qquad \sup_{\delta,x} E[\sup_{0\leq t\leq T}\|\alpha_\delta(t, x, w)\|^p]<\infty.$$

Also we remark that for any $0\leq m\leq[t/\delta]$

$$(2.29)\qquad E[\sup_{(m-1)\delta\leq s\leq m\delta}\|\alpha_\delta(s)-\alpha_\delta((m-1)\delta)\|^p]\leq K_{34}\delta^{\frac{p}{2}}$$

as is easily seen from (2.23) and (2.21). If we set

$$\tilde{Y}_{j_1,j_2}^{i,\delta}(t)=Y_{j_1,j_2}^{i,\delta}(t)-\alpha_\delta^{i,j_1,j_2}(t),$$

then

$$(2.30)\qquad \tilde{Y}_{j_1,j_2}^{i,\delta}(t)=\sum_{k=1}^{d}\sum_{\beta=1}^{r}\int_0^t \sigma'_\beta(X_\delta(s))_k^i Y_{j_1,j_2}^{k,\delta}(s)\dot{W}_\delta(s)ds$$

and from this we can deduce that

$$(2.31)\qquad \|\tilde{Y}_{j_1,j_2}^{(\delta)}(u)-\tilde{Y}_{j_1,j_2}^{(\delta)}(m\delta)\|\leq\|\tilde{Y}_{j_1,j_2}^{(\delta)}(m\delta)\|\tilde{C}_m+\tilde{d}_m\tilde{C}_m$$

$$\text{for}\quad m\delta\leq u\leq(m+1)\delta$$

where $\tilde{Y}_{j_1,j_2}^{(\delta)}=(\tilde{Y}_{j_1,j_2}^{i,\delta})_{i=1}^d$, $\tilde{b}_m=K_{35}\int_{m\delta}^{(m+1)\delta}\|\dot{W}_\delta(s)\|ds$, $\tilde{C}_m=\tilde{b}_m e^{\tilde{b}_m}$ and $\tilde{d}_m=\sup_{m\delta\leq t\leq(m+1)\delta}\|\alpha_\delta(t)-\alpha_\delta(m\delta)\|$. Using this, the estimate

$$(2.32)\qquad \sup_{\delta,x}E[\sup_{0\leq t\leq T}\|Y_{j_1,j_2}^{(\delta)}(t)\|^p]<\infty$$

can be proved in the similar way as for $Y_\delta(t)$. Since the proof is almost a repetition of that for (2.21), we omit the details.

The proof of (1.6) for higher derivatives can be given in a similar way. This completes the proof of (1.6) and hence that of the theorem in section 1.

As a corollary, we can obtain the following from the theorem by a usual trucation argument.

Corollary. *Suppose only that $\sigma_\beta^i(x)$ and $b^i(x)\in C^\infty(R^d)$ but also that the global solutions $X(t, x, w)$ and $X_\delta(t, x, w)$ of* (0.1) *and* (0.2) *exist. Then for any $\varepsilon>0$, we have*

$$(2.33)\qquad \lim_{\delta\downarrow 0}P^W(\sup_{0\leq t\leq T}\sup_{|x|\leq N}|D_x^\alpha X_\delta(t, x, w)-D_x^\alpha X(t, x, w)|>\varepsilon)=0$$

for all $T>0$, $N>0$ and multi-index α.

DEPARTMENT OF MATHEMATICS, KYOTO UNIVERSITY.
PRESENT ADDRESS: DEPARTMENT OF MATHEMATICS,
FUDAN UNIVERSITY,
SHANGHAI, CHINA.

References

[1] J-M. Bismut, Principes de mécanique aléatoire, to appear.
[2] K. D. Elworthy, Stochastic dynamical systems and their flows, Stochastic Analysis (ed. by A. Friedman and M. Pinsky), 79–95, Academic Press, New York, 1978.
[3] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, Kodansha, 1981.
[4] P. Malliavin, Stochastic calculus of variation and hypoelliptic operators, Proc. Int. Symp. SDE. Kyoto 1976 (ed. by K. Itô), 195–263, Kinokuniya, 1978.