Research Report

KSTS/RR-85/017
28 Oct., 1985

The rate of convergence for approximate solutions Stochastic differential equations

by

Shuya Kanagawa

Shuya kanagawa
Department of Mathematics, Keio University

Department of Mathematics
Faculty of Science and Technology
Keio University

Hiyoshi 3-14-1, Kohoku-ku, Yokohama, 223 Japan

KSTS/RR-85/017
November 28, 1985

The rate of convergence for approximate solutions of stochastic differential equations

Shûya KANAGAWA

Department of Mathematics
Faculty of Science and Technology
Keio University, Yokohama

§1. Introduction and results

Let $(\Omega,\mathfrak{F},P)$ be a probability space and $B := \{B(t),\ t \geq 0\}$ $:= \{(B^1(t),B^2(t),\cdots,B^r(t)),\ t \geq 0\}$ an r-dimensional standard Brownian motion on it ($r \geq 1$). We consider a stochastic differential equation (abbreviated: SDE) for a d-dimensional continuous process $X := \{X(t),\ 0 \leq t \leq 1\}$ ($d \geq 1$):

$$(1.1) \qquad dX(t) = \sigma(t,X(t))dB(t) + b(t,X(t))dt,$$

with $X(0) \equiv X_0$, where $\sigma(t,x) := \{\sigma_i^j(t,x),\ 1 \leq i \leq r,\ 1 \leq j \leq d\}$ is a Borel measurable function $(t,x) \in [0,1] \times R^d \to R^d \otimes R^r$ and $b(t,x) := \{b^j(t,x),\ 1 \leq j \leq d\}$ is a Borel measurable function $(t,x) \in [0,1] \times R^d \to R^d$. Suppose that $\sigma(\cdot,\cdot)$ and $b(\cdot,\cdot)$ satisfy the following Lipschitz condition: For any $x,y \in R^r$ and $t,s \in [0,1]$ there exists a positive constant L_1 independent of x,y,t and s such that

KSTS/RR-85/017
November 28, 1985

$$(1.2)\qquad \| \sigma(t,x) - \sigma(s,y) \|^2 + |b(t,x) - b(s,y)|^2 \le L_1^2(|x - y|^2 + |t - s|^2),$$

where

$$\|a\|^2 := \sum_{i=1}^{r} \sum_{j=1}^{d} |a_i^j|^2, \quad \text{for} \quad a \in R^d \otimes R^r$$

and $|\cdot|$ denotes the Euclidian norm. Then there exists a unique solution of (1.1) (see for example Ikeda and Watanabe [4]). It is well known that the solution can be approximated by Maruyama's method (see Maruyama [5]). There are some results as to the rate of convergence in such approximation theorems. (See for example Gihman-Skorokhod [2], Platen [6],[7] and Shimizu [10].) However, approximate solutions, which are treated in these papers, are constructed from normally distributed random variables. In this paper we shall define an approximate solution of (1.1) from i.i.d. random variables with a general distribution and investigate the rate of convergence using the following two metrics $\ell_p(\cdot,\cdot)$ and $\pi(\cdot,\cdot)$.

Let $W^d := C([0,1] \to R^d)$ be the space of continuous functions with the uniform norm $\|\cdot\|$, $\mathcal{B}(W^d)$ the topological σ-field of W^d and $\mathcal{P}(W^d)$ the space of probability measures on $(W^d, \mathcal{B}(W^d))$. Define a metric $\ell_p(\cdot,\cdot)$ on $\mathcal{P}(W^d)$ by for some $0 < p < \infty$ and $P,Q \in \mathcal{P}(W^d)$,

$$\ell_p(P,Q) := [\inf_{\mu \in \mathcal{P}_{PQ}} \int_{W^d \times W^d} \|v - w\|^p \mu(dvdw)]^{1/\tilde{p}}$$

$$= \inf_{\mathcal{L}(Y)=P,\ \mathcal{L}(Z)=Q} E[\|Y - Z\|^p]^{1/\tilde{p}}$$

KSTS/RR-85/017
November 28, 1985

where

$$\mathcal{P}_{PQ} := \{\mu \in \mathcal{P}(W^d \times W^d);\ \mu(A \times W^d) = P(A),\ \mu(W^d \times A) = Q(A) \text{ for all } A \in \mathcal{B}(W^d)\},$$

Y and Z are W^d-valued random variables, $\mathcal{L}(\cdot)$ denotes the law of $\cdot$ and $\tilde{p} := \max(1,p)$. From Theorem 1 in Rachev [8], for $P_n, P \in \mathcal{P}(W^d)$ such that

$$\int_{W^d} ||w||^p\, P_n(dw) < \infty \quad \text{and} \quad \int_{W^d} ||w||^p\, P(dw) < \infty,$$

the convergence $\ell_p(P_n,P) \to 0$ as $n \to \infty$ is equivalent to the following relations: As $n \to \infty$,

$$P_n \Rightarrow P \quad \text{and} \quad \int_{W^d} ||w||^p (P_n - P)(dw) \to 0,$$

where "$\Rightarrow$" means the weak convergence in $(W^d, \mathcal{B}(W^d))$. Another metric on $\mathcal{P}(W^d)$ is the Lévy-Prokhorov metric $\pi(\cdot,\cdot)$ defined by

$$\pi(P,Q) := \inf\{\varepsilon > 0;\ P(A) \leq \varepsilon + Q(G_\varepsilon(A)) \text{ for all } A \in \mathcal{B}(W^d)\},$$

where $G_\varepsilon(A) := \{w \in W^d;\ ||v - w|| < \varepsilon,\ v \in A\}$. There is a relation between $\ell_p(\cdot,\cdot)$ and $\pi(\cdot,\cdot)$:

Rachev's result ([8]). For any $Q, R \in \mathcal{P}(W^d)$,

$$\pi(Q,R) \leq (\ell_p(Q,R))^{\tilde{p}/(1+p)}. \tag{1.3}$$

We next define an approximate solution of the SDE (1.1). Let $\{\xi_k, k \geqq 1\} := \{(\xi_k^1, \xi_k^2, \cdots, \xi_k^r), k \geqq 1\}$ be i.i.d. r-dimensional random variables with zero mean and finite $2+\delta$-th absolute moment for some $\delta > 0$. Without loss of generality we suppose that the covariance matrix is the identity. Define random variables $\hat{Y}_0, \hat{Y}_1, \cdots, \hat{Y}_n$ by

$$\hat{Y}_k := X_0 + \sum_{j=1}^{k} \sigma(\tfrac{j-1}{n}, \hat{Y}_{j-1})\xi_j / n^{\frac{1}{2}} + \sum_{j=1}^{k} b(\tfrac{j-1}{n}, \hat{Y}_{j-1})/n,$$

and $\hat{Y}_0 := X_0$. Let $Y_n := \{Y_n(t);\ 0 \leqq t \leqq 1\}$ be a continuous polygonal line defined by

$$Y_n(t) := \hat{Y}_k + (nt - k)(\hat{Y}_{k+1} - \hat{Y}_k),$$

for $k/n \leq t \leq (k+1)/n$, $k = 0,1,\cdots,n-1$. In 1955 Maruyama (Theorem 2 in [5]) showed an invariance principle: As $n \to \infty$

$$P^{Y_n} \Rightarrow P^X, \tag{1.4}$$

where $P^{Y_n}, P^X \in \mathcal{P}(W^d)$ are the probability measures of Y_n and X, respectively. (1.4) includes classical Donsker's invariance principle as a trivial case where $\sigma(\cdot,\cdot)$ is the identical matrix and $b(\cdot,\cdot)$ the zero vector. In the special case where ξ_1 has the standard normal distribution, we have the following result on the rate of convergence (1.4) (see Gihman-skorokhod [2]):

Let $\{\xi_k, k \geq 1\}$ be i.i.d. r-dimensional random variables with the standard normal distribution. Assume that $\sigma(\cdot,\cdot)$ and $b(\cdot,\cdot)$ satisfy the Lipschitz condition (1.2) and they are bounded, namely, there exists a positive constant L_2 uniformly in $0 \leq t \leq 1$ and $x \in R^d$ such that

$$\|\sigma(t,x)\|^2 + |b(s,x)|^2 \leq L_2^2. \tag{1.5}$$

Under these assumptions we have as $n \to \infty$,

$$\ell_2(P^{Y_n}, P^X) = O(n^{-\frac{1}{2}}).$$

We shall extend the above result to the case where ξ_1 has a general distribution with the $2+\delta$-th absolute moment for some $\delta > 0$.

Theorem 1. Let $\{\xi_k, k \geqq 1\}$ be i.i.d. r-dimensional random variables with zero mean, regular covariance matrix and $E[|\xi_1|^{2+\delta}] < \infty$ for some $0 < \delta \leqq 1$. Assume that $\sigma(\cdot,\cdot)$ and $b(\cdot,\cdot)$ satisfy conditions (1.2) and (1.5). Under these assumptions we have for any $2 \leqq p \leqq 2+\delta$,

(i) if $d = r = 1$, then as $n \to \infty$

$$\ell_p(P^{Y_n}, P^X) = O(n^{-\frac{\delta}{2(2+\delta)}}), \tag{1.6}$$

(ii) if $r > 1$ and ξ_1 has a bounded or square integrable density, then as $n \to \infty$

$$\ell_p(P^{Y_n}, P^X) = O(n^{-\frac{\delta}{2(2+\delta)}}(\log n)^{\frac{1}{2}}). \tag{1.7}$$

From Theorems 1 and (1.3) we have

Theorem 2. We suppose all assumptions of Theorem 1. Then we have

(i) if $d = r = 1$, then as $n \to \infty$

$$\pi(P^{Y_n}, P^X) = O(n^{-\frac{\delta}{2(3+\delta)}}), \tag{1.8}$$

(ii) if $r > 1$ and ξ_1 has a bounded or square integrable density, then as $n \to \infty$

$$(1.9) \qquad \pi(P^{Y_n}, P^X) = O\left(n^{-\frac{\delta}{2(3+\delta)}} (\log n)^{\frac{2+\delta}{2(3+\delta)}}\right).$$

Remark. If $\sigma(\cdot,\cdot) \equiv 1$ and $b(\cdot,\cdot) \equiv 0$, then Theorem 2 yields the results of Borovkov [1] and Gorodetskii [3]. Even in this special case it is known that (1.8) is the best possible result (see Sahanenko [9]). Thus (1.6), from which (1.8) was derived immediately, may also be regarded as best possible.

§2. Preliminaries

Define new random variables $\zeta_1, \zeta_2, \cdots, \zeta_M$ which are sums of blocks of ξ_k's as follows:

$$\zeta_k := (\zeta_k^1, \zeta_k^2, \cdots, \zeta_k^r) := \sum_{i=(k-1)q+1}^{kq \wedge n} \xi_i / n^{\frac{1}{2}}, \quad 1 \leq k \leq M,$$

where $q = [n^{\frac{2}{2+\delta}}]$, $M := [n/q] + 1 \sim n^{\frac{\delta}{2+\delta}}$, $[a]$ being the integral part of a, and $a \wedge b$ means $\min(a,b)$. Let $\{t_k, k = 0,1,\cdots,M\}$ be a partition of the interval $[0,1]$ which is defined by $t_k = k\Delta$, for $0 \leq k \leq M-1$ and $t_M = 1$, where $\Delta := q/n \sim n^{-\delta/(2+\delta)}$. Moreover define increments of the Brownian motion by $\eta_k := (\eta_k^1, \eta_k^2, \cdots, \eta_k^r) := B(t_k) - B(t_{k-1})$, $1 \leq k \leq M$. We approximate X and Y_n by the following processes $\overline{X}_n$ and $\overline{Y}_n$: Let $\{\tilde{X}_k, k = 0,1, \cdots,M\}$ and $\{\tilde{Y}_k, k = 0,1,\cdots,M\}$ be random variables defined by

$$\tilde{X}_k := X_0 + \sum_{j=1}^{k} \sigma(t_{j-1}, \tilde{X}_{j-1})\eta_j + \sum_{j=1}^{k} b(t_{j-1}, \tilde{X}_{j-1})(t_j - t_{j-1})$$

$$\tilde{Y}_k := X_0 + \sum_{j=1}^{k} \sigma(t_{j-1}, \tilde{Y}_{j-1})\zeta_j + \sum_{j=1}^{k} b(t_{j-1}, \tilde{Y}_{j-1})(t_j - t_{j-1}),$$

for each $1 \leq k \leq M$ and $\tilde{X}_0 = \tilde{Y}_0 := X_0$. Denote $\overline{X}_n := \{\overline{X}_n(t), 0 \leq t \leq 1\}$ and $\overline{Y}_n := \{\overline{Y}_n(t), 0 \leq t \leq 1\}$ $D([0,1] \to R^d)$-valued processes defined by $\overline{X}_n(t) := \tilde{X}_{k-1}$ and $\overline{Y}_n(t) := \tilde{Y}_{k-1}$ for $t_{k-1} \leq t < t_k$, $1 \leq k \leq M$, and $\overline{X}_n(1) := \tilde{X}_M$ and $\overline{Y}_n(1) := \tilde{Y}_M$ for $t = 1$, respectively.

One of the main techniques of the proof of Theorem 2 is the following reconstruction of all random variables on a common probability space.

Lemma 1. Without changing distributions of $\{\xi_k, 1 \leq k \leq n\}$ and $\{\zeta_k, 1 \leq k \leq M\}$, we can redefine them on a richer probability space with a Brownian motion $\{B(t), t \geqq 0\}$ and its increments $\{\eta_k, 1 \leq k \leq M\}$ such that the following properties holds.

(i) If $d = r = 1$, then for any $0 < p < 2+\delta$ and for each $1 \leq k \leq M$, as $n \to \infty$

$$(2.1) \qquad E[|\zeta_k - \eta_k|^p] = O(\Delta^{(p-\delta)/2} n^{-\delta/2}).$$

(ii) If $r > 1$ and ξ_1 has a bounded or square integrable density, then for any $0 < p < 2+\delta$ and each $1 \leq k \leq M$, as $n \to \infty$

$$(2.2) \qquad E[|\zeta_k - \eta_k|^p] = O(\Delta^{(p-\delta)/2} n^{-\delta/2}(\log n)^{p/2}).$$

(iii) For each $1 \leq k \leq M-1$,

$$(2.3) \qquad \{\eta_1, \eta_2, \cdots, \eta_k\} \text{ is independent of } \{\zeta_{k+1}, \zeta_{k+2}, \cdots, \zeta_M\},$$

$$(2.4) \qquad \{\eta_{k+1}, \eta_{k+2}, \cdots, \eta_M\} \text{ is independent of } \{\zeta_1, \zeta_2, \cdots, \zeta_k\}.$$

For the proof of this lemma we give several notations (according to Gorodetskii [3]). Let $x := (x^1, x^2, \cdots, x^r) \in R^r$. For each $1 \leq k \leq M$ and $1 \leq i \leq r$, let $\mu_k^i(\cdot)$ be the probability measure of $((t_k - t_{k-1})^{-\frac{1}{2}}\zeta_k^1, (t_k - t_{k-1})^{-\frac{1}{2}}\zeta_k^2, \cdots, (t_k - t_{k-1})^{-\frac{1}{2}}\zeta_k^i)$ and $F_k^i(\cdot|x^1, \cdots, x^{k-1})$ the right continuous conditional distribution function defined by for any bounded Borel function ψ,

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\psi(x^1,\cdots,x^i)F_i(dx^i|x^1,\cdots,x^{i-1})\mu_k^{i-1}(dx^1,\cdots,dx^{i-1})$$

$$= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\psi(x^1,\cdots,x^i)\mu_k^i(dx^1,\cdots,dx^i).$$

Define the inverse function of $F_k^i(\cdot|x^1,\cdots,x^{i-1})$ by

$$(F_k^i)^{-1}(u|x^1,\cdots,x^{i-1}) := \sup_{F_k^i(v|x^1,\cdots,x^{i-1}) \leqq u} v.$$

Let $\Phi(\cdot)$ be the one dimensional standard normal distribution function. Furthermore define transformations $h_k^i, h_k : R^r \to R^1$ by

$$h_k^i(x) := (x^1,\cdots,x^{i-1},(F_k^i)^{-1}(\Phi(x^i)|x^1,\cdots,x^{i-1}),x^{i+1},\cdots,x^r),$$

and $h_k := h_k^r \circ h_k^{r-1} \circ \cdots \circ h_k^1$. Then we have

$$(2.5)\quad \mathcal{L}\{h_1(t_1^{-\frac{1}{2}}\eta_1),h_2((t_2 - t_1)^{-\frac{1}{2}}\eta_2),\cdots,h_M((t_M - t_{M-1})^{-\frac{1}{2}}\eta_M)\}$$

$$= \mathcal{L}\{t_1^{-\frac{1}{2}}\zeta_1,(t_2 - t_1)^{-\frac{1}{2}}\zeta_2,\cdots,(t_M - t_{M-1})^{-\frac{1}{2}}\zeta_M\}.$$

Now, applying (2.5), we redefine processes $\{\xi_k\}$, $\{\zeta_k\}$ and $\{B(\cdot)\}$ such that (2.1)-(2.4) are satisfied as follows: Suppose that there is a Brownian motion $B^* := \{B^*(t),\ t \geqq 0\}$ on another probability space $(\Omega^*,\mathcal{F}^*,P^*)$. Define

(2.6) $\{\zeta_1^*, \zeta_2^*, \cdots, \zeta_M^*\} := \{t_1^{\frac{1}{2}} h_1(t_1^{-\frac{1}{2}}(t_1^{-\frac{1}{2}}\eta_1^*)),$

$(t_2 - t_1)^{-\frac{1}{2}} h_2((t_2 - t_1)^{-\frac{1}{2}}\eta_2^*), \cdots, (t_M - t_{M-1})^{\frac{1}{2}} h_M((t_M - t_{M-1})^{-\frac{1}{2}}\eta_M^*)\}$

where $\eta_1^*, \cdots, \eta_M^*$ are increments of B^* and $\mathcal{L}\{\eta_k^*, 1 \leq k \leq M\} = \mathcal{L}\{\eta_k, 1 \leq k \leq M\}$. Then, from (2.5), $\mathcal{L}\{\zeta_k^*, 1 \leq k \leq M\} = \mathcal{L}\{\zeta_k, 1 \leq k \leq M\}$. We next reconstruct $\{\xi_k\}$ by the following method. Define probability measures U and V by for any $A_1 \in \mathcal{B}(R^r \otimes R^n)$, $A_2 \in \mathcal{B}(R^r \otimes R^M)$ and $A_3 \in \mathcal{B}(W^r)$,

$$U(A_1 \times A_2) := P\{(\xi_1, \cdots, \xi_n) \in A_1, (\zeta_1, \cdots, \zeta_M) \in A_2\}$$

and

$$V(A_2 \times A_3) := P^*\{(\zeta_1^*, \cdots, \zeta_M^*) \in A_2, B^* \in A_3\}.$$

Put

$$U_{A_1}(A_2) := U(A_1 \times A_2), \quad V_{A_3}(A_2) := V(A_2 \times A_3)$$

and

$$H(A_2) := U(R^r \otimes R^M \times A_2) = V(A_2 \times W^r).$$

Since $U_{A_1}(\cdot)$ is absolute continuous with respect to $H(\cdot)$, there exists a $\mathcal{B}(R^r \otimes R^M)$-measurable function $p_{A_1}(\cdot)$ such that

$$U_{A_1}(A_2) = \int_{A_2} p_{A_1}(y) H(dy).$$

Furthermore there also exists a $\mathcal{B}(R^r \otimes R^M)$-measurable function $q_{A_3}(\cdot)$ such that

$$V_{A_3}(A_2) = \int_{A_2} q_{A_3}(y) H(dy).$$

Define a probability measure Q on $(R^r \otimes R^n) \times (R^r \otimes R^M) \times W^r$ by

$$(2.7) \qquad Q(A_1 \times A_2 \times A_3) := \int_{A_3} p_{A_1}(y) q_{A_3}(y) H(dy).$$

Finally define new probability space $(\Omega', \mathcal{F}', P')$ by $\Omega' := (R^r \otimes R^n) \times (R^r \otimes R^M) \times W^r$, $\mathcal{F}'$ being the completion of the topological σ-field $\mathcal{B}(\Omega')$ by Q and $P' := Q$. Then without changing distributions of $\{\xi_k\}$, $\{\zeta_k\}$ and $\{B(\cdot)\}$ and keeping the relation (2.6), we can redefine them on the common probability space $(\Omega', \mathcal{F}', P')$ by putting for each $\omega := (\omega_1, \omega_2, \omega_3) \in \Omega'$,

$$(\xi_1, \xi_2, \cdots, \xi_n)(\omega) := \omega_1, \quad (\zeta_1, \zeta_2, \cdots, \zeta_M)(\omega) := \omega_2 \text{ and } B(\cdot, \omega) := \omega_3.$$

Now, from (2.6), the relations (2.3) and (2.4) in Lemma 1 can be easily shown. Moreover (2.1) and (2.2) are proved by Borovkov (Lemma 1 in [1]) and Gorodetskii (Lemma 2 in [3]), respectively. Thus we can conclude the proof of Lemma 1.

In what follows, as an absolute positive constant, we use a K which may be different in the different equations.

§3. Proof of Theorem 2

Since

$$d(X,Y_n) \leq d(X,\overline{X}_n) + d(\overline{X}_n,\overline{Y}_n) + d(\overline{Y}_n,Y_n),$$

we shall give the following three lemmas. The first lemma is due to Shimizu[10].

Lemma 2. As $n \to \infty$, for any $2 \leq p \leq 2+\delta$,

$$E[d(X,\overline{X}_n)^p] = O(\Delta^{p/2}).$$

Proof. For $t_k \leq t < t_{k+1}$, $k = 0,1,\cdots,M-1$, let $\sigma_n(t) := \sigma(t_{k-1},\tilde{X}_{k-1})$ and $b_n(t) := b(t_{k-1},\tilde{X}_{k-1})$. Obviously

$$(3.1) \qquad X(t) - \overline{X}_n(t) = \int_0^t (\sigma(s,X(s)) - \sigma_n(s))dB(s) + \int_0^t (b(s,X(s)) - b_n(s))ds$$

$$= \int_0^t (\sigma(s,X(s)) - \sigma(s,\overline{X}_n(s))dB(s) + \int_0^t (\sigma(s,\overline{X}_n(s)) - \sigma_n(s))dB(s)$$

$$+ \int_0^t (b(s,X(s)) - b(s,\overline{X}_n(s))ds + \int_0^t (b(s,\overline{X}_n(s)) - b_n(s))ds.$$

From a moment inequality for martingales (see for example Theorem 3.1 in Chapter III of Ikeda-Watanabe [4]), Jensen's inequality and condition (1.2),

$$(3.2) \qquad E[\max_{0 \leqq s \leqq t} |\int_0^s (\sigma(u,X(u)) - \sigma(u,\overline{X}_n(u))dB(u)|^p]$$

$$\leqq K E[|\int_0^t \|\sigma(u,X(u)) - \sigma(u,\overline{X}_n(u)\|^2 du]^{p/2}$$

$$\leqq K \int_0^t E[\|\sigma(u,X(u)) - \sigma(u,\overline{X}_n(u))\|^p]du$$

$$\leqq KL_1^p \int_0^t E[|X(u) - \overline{X}_n(u)|^p]du$$

$$\leqq KL_1^p \int_0^t E[\max_{0 \leqq u \leqq s} |X(u) - \overline{X}_n(u)|^p]ds.$$

For $t_k \leqq t < t_{k+1}$ we have from the conditions (1.2) and (1.5),

$$E[\|\sigma(t,\overline{X}_n(t)) - \sigma_n(t)\|^p] \leqq L_1^p E[(|t - t_{k-1}|^p + |\tilde{X}_k - \tilde{X}_{k-1}|^p)]$$

$$\leqq K\Delta^p + K E[|\sigma(t_{k-1},\overline{X}_{k-1})\eta_k|^p] + K E[b(t_{k-1},\overline{X}_{k-1})(t_k - t_{k-1})|^p]$$

$$\leqq K\Delta^p + K\Delta^{p/2}$$

Hence, from Theorem 3.1 in [4] and (1.2),

$$(3.3) \qquad E[\max_{0 \leqq s \leqq t} \int_0^s (\sigma(u,\overline{X}_n(u)) - \sigma_n(u))dB(u)|^p]$$

$$\leqq K E[|\int_0^t (\sigma(u,\overline{X}_n(u)) - \sigma_n(u))dB(u)|^p]$$

$$\leqq K\int_0^t E[\|\sigma(u,\overline{X}_n(u)) - \sigma_n(u)\|^p]du \leqq K\Delta^{p/2}.$$

Furthermore it follows from Jensen's inequality that

$$(3.4)\qquad E[\max_{0\leq s\leq t}|\int_0^s(b(u,X(u)) - b(u,\overline{X}_n(u)))du|^p]$$

$$\leq K\int_0^t E[\max_{0\leq u\leq s}|X(u) - \overline{X}_n(u)|^p]ds,$$

and

$$(3.5)\qquad E[\max_{0\leq s\leq t}|\int_0^s(b(u,\overline{X}_n(u)) - b_n(u))du|^p]$$

$$\leq K\int_0^t E[|b(u,\overline{X}_n(u)) - b_n(u)|^p]du \leq K\Delta^{p/2}.$$

Combining (3.1) - (3.5) we conclude the proof of the lemma from Gronwall's inequality. Q.E.D.

Lemma 3. As $n \to \infty$, for any $2 \leq p \leq 2+\delta$,

$$E[d(Y_n,\overline{Y}_n)^p] = O(\Delta^{p/2}).$$

Proof. For $t_k \leq t < t_{k+1}$,

$$\max_{t_k\leq s\leq t}|Y_n(s) - \overline{Y}_n(s)| = \max_{kq<i\leq[nt]}|\hat{Y}_i - \tilde{Y}_k|$$

$$\leq |\sum_{j=1}^{kq}\sigma(\tfrac{j-1}{n},\hat{Y}_{j-1})\xi_j/n^{\frac{1}{2}} - \sum_{\ell=1}^{k}\sigma(t_{\ell-1},\tilde{Y}_{\ell-1})\zeta_\ell|$$

$$+ \max_{kq<i\leq[nt]}|\sum_{j=kq+1}^{i}\sigma(\tfrac{j-1}{n},\hat{Y}_{j-1})\xi_j/n^{\frac{1}{2}}|$$

$$+ |\sum_{j=1}^{kq}b(\tfrac{j-1}{n},\hat{Y}_{j-1})/n - \sum_{\ell=1}^{k}b(t_{\ell-1},\tilde{Y}_{\ell-1})q/n|$$

$$+ \max_{kq<i\leq[nt]}|\sum_{j=kq+1}^{i}b(\tfrac{j-1}{n},\hat{Y}_{j-1})/n|.$$

$$\leq |\sum_{\ell=1}^{k}\sum_{j=(\ell-1)q+1}^{\ell q}(\sigma(\tfrac{j-1}{n},\hat{Y}_{j-1}) - \sigma(t_{\ell-1},\tilde{Y}_{\ell-1}))\xi_j/n^{\frac{1}{2}}|$$

$$+ \max_{kq \le i \le [nt]} \Big| \sum_{j=kq+1}^{i} \sigma(\tfrac{j-1}{n}, \hat{Y}_{j-1})\xi_j / n^{\frac{1}{2}} \Big|$$

$$+ \Big| \sum_{\ell=1}^{k} \sum_{j=(\ell-1)q+1}^{\ell q} (b(\tfrac{j-1}{n}, \hat{Y}_{j-1}) - b(t_{\ell-1}, \tilde{Y}_{\ell-1}))/n \Big|$$

$$+ \sum_{j=kq+1}^{[nt]} |b(\tfrac{j-1}{n}, \hat{Y}_{j-1})|/n.$$

Thus, by Doob's inequality,

$$(3.6) \qquad E[\max_{0 \le s \le t} |Y_n(s) - \overline{Y}_n(s)|^p]$$

$$\le K\,E[\Big| \sum_{\ell=1}^{k} \sum_{j=(\ell-1)q+1}^{\ell q} (\sigma(\tfrac{j-1}{n}, \hat{Y}_{j-1}) - \sigma(t_{\ell-1}, \tilde{Y}_{\ell-1}))\xi_j / n^{\frac{1}{2}} \Big|^p]$$

$$+ K\,E[\Big| \sum_{j=kq+1}^{[nt]} \sigma(\tfrac{j-1}{n}, \hat{Y}_{j-1})\xi_j / n^{\frac{1}{2}} \Big|^p]$$

$$+ K\,E[(\sum_{\ell=1}^{k} \sum_{j=(\ell-1)q+1}^{\ell q} |b(\tfrac{j-1}{n}, \hat{Y}_{j-1}) - b(t_{\ell-1}, \tilde{Y}_{\ell-1})|/n)^p]$$

$$+ K\,E[(\sum_{j=kq+1}^{[nt]} |b(\tfrac{j-1}{n}, \hat{Y}_{j-1})|/n)^p]$$

$$=: H_1 + H_2 + H_3 + H_4,$$

say. Since $\{(\sigma(\frac{j-1}{n}, \hat{Y}_{j-1}) - \sigma(t_{\ell-1}, \tilde{Y}_{\ell-1}))\xi_j\}_j$ and $\{\sigma(\frac{j-1}{n}, \hat{Y}_{j-1})\xi_j\}_j$ are martingale difference sequences, we have from Theorem 3.1 in [4], (1.2) and (1.5) that

$$(3.7)\ H_1 \le Kn^{-\frac{1}{2}p} E[\Big\{ \sum_{\ell=1}^{k} \sum_{j=(\ell-1)q+1}^{\ell q} E(|(\sigma(\tfrac{j-1}{n}, \hat{Y}_{j-1}) - \sigma(t_{\ell-1}, \tilde{Y}_{\ell-1}))\xi_j|^2 | \mathcal{G}_{j-1}) \Big\}^{p/2}]$$

$$\le K\,E[\Big\{ \sum_{\ell=1}^{k} \sum_{j=(\ell-1)q+1}^{\ell q} (|\hat{Y}_{j-1} - \tilde{Y}_{\ell-1}|^2 + \Delta^2)/n \Big\}^{p/2}]$$

$$\le K \int_0^t E[\max_{0 \le u \le s} |Y_n(u) - \overline{Y}_n(u)|^p] ds + K\Delta^p,$$

where $\mathcal{G}_j$ is the σ-field generated by $\xi_1, \cdots, \xi_j$ for each j, and

$$(3.8)\qquad H_2 \leqq K\Big(\sum_{j=kq+1}^{[nt]} L_1^2 E[|\xi_j|^2]/n\Big)^{p/2} \leqq K\Delta^{p/2}.$$

Furthermore, from Jensen's inequality, we have

$$(3.9)\qquad H_3 \leqq K\,E\Big[\Big(\sum_{\ell=1}^{k}\sum_{j=(\ell-1)q+1}^{\ell q}(|\hat{Y}_{j-1} - \tilde{Y}_{\ell-1}| + \Delta)/n\Big)^p\Big]$$

$$\leqq K\sum_{\ell=1}^{k}\sum_{j=(\ell-1)q+1}^{\ell q} E[|\hat{Y}_{j-1} - \tilde{Y}_{\ell-1}|^p]/n$$

$$+ K\sum_{\ell=1}^{k}\sum_{j=(\ell-1)q+1}^{\ell q} \Delta^p/n$$

$$\leqq K\int_0^t E[\max_{0\leqq u\leqq s}|Y_n(u) - \overline{Y}_n(u)|^p]ds + K\Delta^p,$$

and

$$(3.10)\qquad H_4 \leqq K(qL_2/n)^p \leqq K\Delta^p.$$

Combining (3.6) - (3.10) and Gronwall's inequality the proof of this lemma is finished. Q.E.D.

Lemma 4. We can redefine the processes $\overline{X}_n$ and $\overline{Y}_n$ on a richer probability space such that the following relation holds:

(i) If $d = r = 1$, then as $n \to \infty$, for any $2 \leqq p \leqq 2+\delta$,

$$E[d(\overline{X}_n,\overline{Y}_n)^p] = O(\Delta^{p/2}),$$

(ii) if $r > 1$ and ξ_1 has a bounded or square integrable density, then as $n \to \infty$, for any $2 \leqq p \leqq 2+\delta$,

$$E[d(\overline{X}_n,\overline{Y}_n)^p] = O(\Delta^{p/2}(\log n)^{p/2}).$$

Proof. For $t_k \leq t < t_{k+1}$,

$$(3.11)\qquad E[\max_{0\leq s\leq t}|\overline{X}_n(s) - \overline{Y}_n(s)|^p] = E[\max_{1\leq i\leq k}|\tilde{X}_j - \tilde{Y}_j|^p]$$

$$\leq K E[\max_{1\leq i\leq k}|\sum_{j=1}^{i}\sigma(t_{j-1},\tilde{X}_{j-1})\eta_j - \sum_{j=1}^{i}\sigma(t_{j-1},\tilde{Y}_{j-1})\zeta_j|^p]$$

$$+ K E[\max_{1\leq i\leq k}|\sum_{j=1}^{i}b(t_{j-1},\tilde{X}_{j-1})(t_j - t_{j-1}) - \sum_{j=1}^{i}b(t_{j-1},\tilde{Y}_{j-1})(t_j - t_{j-1})|^p]$$

$$\leq K E[\max_{1\leq i\leq k}|\sum_{j=1}^{i}(\sigma(t_{j-1},\tilde{X}_{j-1}) - \sigma(t_{j-1},\tilde{Y}_{j-1}))\eta_j|^p]$$

$$+ K E[\max_{1\leq i\leq k}|\sum_{j=1}^{i}\sigma(t_{j-1},\tilde{Y}_{j-1})(\eta_j - \zeta_j)|^p]$$

$$+ K E[\max_{1\leq i\leq k}|\sum_{j=1}^{i}(b(t_{j-1},\tilde{X}_{j-1}) - b(t_{j-1},\tilde{Y}_{j-1}))(t_j - t_{j-1})|^p]$$

$$=: I_1 + I_2 + I_3,$$

say. We first deal with I_1. Let $\sigma_n'(s) := \sigma(t_{j-1},\tilde{X}_{j-1}) - \sigma(t_{j-1},\tilde{Y}_{j-1})$ for $t_j \leq s < t_{j+1}$, $j = 0,1,\cdots,M-1$. Since $\{\sigma_n'(s),\ s < t_k\}$ is independent of $\{\eta_k, \eta_{k+1},\cdots,\eta_M\}$ by the relation (2.4) of Lemma 1, I_1 is represented by

$$I_1 = K E[\max_{1\leq i\leq k}|\int_0^{t_i}\sigma_n'(s)dB(s)|^p].$$

Thus from Theorem 3.1 in [4] and the condition (1.2),

$$(3.12)\qquad I_1 \leq K E[|\int_0^{t_k}\sigma_n'(s)dB(s)|^p] \leq K\int_0^{t_k}E[\|\sigma_n'(s)\|^p]ds$$

$$\leq K\int_0^{t}E[\max_{0\leq u\leq s}|\overline{X}_n(u) - \overline{Y}_n(u)|^p]ds.$$

We next estimate I_2. By the relation (2.4),

$$E[\sigma(t_{j-1},\tilde{Y}_{j-1})(\eta_j - \zeta_j)|\mathcal{H}_{j-1}] = 0 \quad \text{a.s.},$$

for each j, where $\mathcal{H}_j$ is the σ-field generated by $\eta_1,\cdots,\eta_j,\zeta_1,\cdots,\zeta_j$ for each $1 \leq j \leq M$. Thus, from Theorem 3.1 in [4] and (1.5), (2.1), (2.2) and (2.3), we have

$$(3.13)\qquad I_2 \leq KE[\sum_{j=1}^{k} E(|\sigma(t_{j-1},\tilde{Y}_{j-1})(\eta_j - \zeta_j)|^2|\mathcal{H}_{j-1})]^{p/2}$$

$$\leq KE[\sum_{j=1}^{k} \|\sigma(t_{j-1},\tilde{Y}_{j-1})\|^2 E[|\eta_j - \zeta_j|^2]]^{p/2}$$

$$\leq KL_2^p(\sum_{j=1}^{k} E[|\eta_j - \zeta_j|^2])^{p/2}$$

$$\leq \begin{cases} K(k\Delta^{(2-\delta)/2}n^{-\delta/2})^{p/2} & \text{if } d = r = 1, \\ K(k\Delta^{(2-\delta)/2}n^{-\delta/2}\log n)^{p/2} & \text{if } r > 1, \end{cases}$$

$$\leq \begin{cases} Kt(\Delta^{-\delta/2}n^{-\delta/2})^{p/2} & \text{if } d = r = 1, \\ Kt(\Delta^{-\delta/2}n^{-\delta/2}\log n)^{p/2} & \text{if } r > 1, \end{cases}$$

$$\leq \begin{cases} Kt\Delta^{p/2} & \text{if } d = r = 1, \\ Kt\Delta^{p/2}(\log n)^{p/2} & \text{if } r > 1. \end{cases}$$

As for I_3, letting $b_n'(s) = b(t_{j-1},\tilde{X}_{j-1}) - b(t_{j-1},\tilde{Y}_{j-1})$ for $t_j \leqq s < t_{j+1}$, $j = 0,1,\cdots,M-1$, we have from (2.1) that

$$(3.14)\quad I_3 = KE[\max_{1\leqq i\leqq k}|\int_0^{t_j} b_n'(s)ds|^p] \leqq K\int_0^{t_k} E[|b_n'(s)|^p]ds$$
$$\leqq K\int_0^t E[\max_{0\leqq u\leqq s}|\overline{X}_n(u) - \overline{Y}_n(u)|^p]ds.$$

Combining (3.11)-(3.14) we have

$$E[\max_{0\leqq s\leqq t}|\overline{X}_n(s) - \overline{Y}_n(s)|^p]$$
$$\leqq K\int_0^t E[\max_{0\leqq u\leqq s}|\overline{X}_n(u) - \overline{Y}_n(u)|^p]ds$$
$$+\begin{cases} Kt\Delta^{p/2} & \text{if } d = r = 1, \\ Kt\Delta^{p/2}(\log n)^{p/2} & \text{if } r > 1, \end{cases}$$

for any $0 \leqq t \leqq 1$. Consequently the lemma is proved by Gronwall's inequality. Q.E.D.

Without changing distributions we can reconstruct W^d-valued processes X and Y_n on the common probability space $(\Omega',\mathcal{F}',P')$ by Lemmas 1-4 such that the conclusion of Theorem 1 holds:

$$\ell_p(P^X,P^{Y_n}) \leqq E[d(X,Y_n)^p]^{1/p} \leqq \begin{cases} K\Delta^{\frac{1}{2}} & \text{if } d = r = 1, \\ K\Delta^{\frac{1}{2}}(\log n)^{\frac{1}{2}} & \text{if } r > 1. \end{cases}$$

Therefore we finish the proof of Theorem 1.

Acknowledgement

The author would like to express his gratitude to Professors H. Tanaka and M. Maejima for their helpful suggestions. He also thanks Professors S. T. Rachev and S. A. Utev for their valuable comments.

References

[1] A. A. Borovkov, On the rate of convergence for the invariance principle, Theory Prob. Appl. 18 (1973) 207-225.

[2] I. I. Gihman and A. V. Skorokhod, The Theory of Stochastic Processes III, (Springer-Verlag, Berlin, 1979).

[3] V. V. Gorodetskii, On the rate of convergence for the multidimensional invariance principle, Theory Prob. Appl. 20 (1975) 631-638.

[4] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, (North-Hollanda/Kodansha, Amsterdam-Tokyo, 1981).

[5] G. Maruyama, Continuous Markov processes and stochastic equations, Rend. Circ. Mate. Palermo 4 (1955) 48-90.

[6] E. Platen, An approximation method for a class of Ito processes, Liet. Matem. Rink. 21 (1981) 121-133.

[7] E. Platen, An approximation method for a class of Ito processes with jump component, Liet. Matem. Rink. 22 (1982) 124-136.

[8] S. T. Rachev, On a class of minimal function on a space of probability measures, Theory Prob. Appl. 29 (1984) 41-49.

[9] A. I. Sahanenko, Estimates of rate of convergence in the invariance principle, Dokl. Akad. Nauk USSR 219 (1974) 1752-1755.

[10] A. Shimizu, Approximate solutions for stochastic differential equations, Proc. Symp. on Stochastic Differential Equations for Population Genetics (1984) 201-210, (in Japanese).