TOKYO J. MATH.
VOL. 12, No. 1, 1989

The Rate of Convergence for Approximate Solutions of Stochastic Differential Equations

Shûya KANAGAWA

Yamanashi University

(Communicated by S. Koizumi)

§1. Introduction and results.

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

$$dX(t)=\sigma(t, X(t))dB(t)+b(t, X(t))dt , \tag{1.1}$$

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

$$\|\sigma(t, x)-\sigma(s, y)\|^2+|b(t, x)-b(s, y)|^2\leqq L_1^2(|x-y|^2+|t-s|^2) , \tag{1.2}$$

where

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

and $|\cdot|$ denotes the Euclidean norm. Then there exists a unique solution of the SDE (1.1) (see, for example, Ikeda-Watanabe [8]). Approximate solutions for (1.1) were constructed by Maruyama [9], and its rate of convergence was studied by Gihman-Skorokhod [2] and Shimizu [17] (see also Greenside-Helfand [4], Janković [5], Janssen [6], Milshtein [10], Platen [11], [12], Rao-Borwanker-Ramkrishna [14], Rümelin [15], Wright [18]). In [2] and [17] on the rate of convergence, approximate solutions are

Received October 14, 1987
Revised November 25, 1988

constructed from normally distributed random variables (namely, increments of a Brownian motion). In this paper we shall use approximate solutions of (1.1) defined by i.i.d. random variables with a general distribution and investigate the rate of convergence in terms of two probability metrics $l_p(\cdot, \cdot)$ and $\pi(\cdot, \cdot)$ defined below. Our result (Theorem 2) is in a sense a generalization of Borovkov [1] and Gorodetskii [3] (see Remark to Theorem 2). The advantage of defining approximate solutions by i.i.d. random variables, not by normally distributed random variables, is known for instance, in simulation on a digital computer (cf. [6]).

Let $W^d := C([0, 1] \to \boldsymbol{R}^d)$ be the space of continuous functions with the uniform norm $\|\cdot\|$, $\mathscr{B}(W^d)$ the topological σ-field of W^d and $\mathscr{P}(W^d)$ the space of all probability measures on $(W^d, \mathscr{B}(W^d))$. For any $0<p<\infty$ define a metric $l_p(\cdot, \cdot)$ on $\mathscr{P}(W^d)$ by

$$l_p(P, Q) := \left[\inf_{\mu \in \mathscr{P}_{PQ}} \int_{W^d \times W^d} \|v-w\|^p \mu(dvdw)\right]^{1/\tilde{p}}$$
$$= \inf_{\mathscr{L}(Y)=P, \mathscr{L}(Z)=Q} E[\|Y-Z\|^p]^{1/\tilde{p}}, \qquad P, Q \in \mathscr{P}(W^d),$$

where

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

Y and Z are W^d-valued random variables, $\mathscr{L}(\cdot)$ denotes the law of $\cdot$ and $\tilde{p} := \max(1, p)$. It follows from Theorem 1 of [13] that, for $P_n, P \in \mathscr{P}(W^d)$ satisfying

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

the convergence $l_p(P_n, P) \to 0$ as $n \to \infty$ is equivalent to that

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

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

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

where $G^\varepsilon(A) := \{w \in W^d;\ \|v-w\| < \varepsilon,\ v \in A\}$. A relationship between $l_p(\cdot, \cdot)$ and $\pi(\cdot, \cdot)$ is that for all $Q, R \in \mathscr{P}(W^d)$,

$$\pi(Q, R) \leqq l_p(Q, R)^{\tilde{p}/(1+p)} \tag{1.3}$$

for any posititive $p < \infty$, (Rachev [13]).

We next define approximate solutions 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 $\delta>0$. We suppose that the covariance matrix has non-zero determinant and then it is the identity without loss of generality. Define random variables $\hat{Y}_0, \hat{Y}_1, \cdots, \hat{Y}_n$ by

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

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 $k/n\leqq t\leqq (k+1)/n$, $k=0, 1, \cdots, n-1$. Maruyama (Theorem 2 in [9]) showed that

$$P^{Y_n} \Longrightarrow P^X , \qquad (n\to\infty) , \tag{1.4}$$

where P^{Y_n}, $P^X \in \mathscr{P}(W^d)$ 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 $b(\cdot, \cdot)$ is the zero vector. Our main result on the rate of convergence in (1.4) is as follows:

THEOREM 1. *Let $\{\xi_k, k\geqq 1\}$ be i.i.d. r-dimensional random variables with zero mean, regular covariance matrix and with $E[|\xi_1|^{2+\delta}]<\infty$ for some $0<\delta\leqq 1$. Assume that $\sigma(\cdot, \cdot)$ and $b(\cdot, \cdot)$ satisfy* (1.2) *and are bounded, namely,*

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

where L_2 is a positive constant independent of x, y, s and t. Under these assumptions we have for any $2\leqq p\leqq 2+\delta$,

(i) *if $d=r=1$, then for any $\varepsilon>p/2$,*

$$l_p(P^{Y_n}, P^X) = o(n^{-\delta/2(2+\delta)}(\log n)^{\varepsilon}) , \qquad (n\to\infty) , \tag{1.6}$$

(ii) *if $r>1$ and ξ_1 has a bounded or square integrable density, then* (1.6) *also holds.*

From the statement of Theorem 1 for $p=2+\delta$ and (1.3) we have the following result.

THEOREM 2. *Under the same assumptions as in Theorem 1, we have*

(i) *if $d=r=1$, then for any $\varepsilon>(2+\delta)^2/2(3+\delta)$,* (1.7) $$\pi(P^{Y_n}, P^X)=o(n^{-\delta/2(3+\delta)}(\log n)^{\varepsilon}) , \qquad (n\to\infty) ,$$

(ii) *if* $r>1$ *and* ξ_1 *has a bounded or square integrable density, then* (1.7) *also holds.*

REMARK TO THEOREM 2. The weak convergence of the sequence of approximate solutions Y_n's constructed by i.i.d. random variables with a general distribution generalized the well known Donsker's invariance principle, where $\sigma(\cdot,\cdot)\equiv 1$ and $b(\cdot,\cdot)\equiv 0$. In case that $\sigma(\cdot,\cdot)\equiv 1$ and $b(\cdot,\cdot)\equiv 0$, Borovkov [1] and Gorodetskii [3] obtained

(1.8) $$\pi(P^{Y_n}, P^X)=O(n^{-\delta/2(3+\delta)}) , \qquad (n\to\infty) .$$

In our case, the power of n is the same as their result (1.8), but the convergence is decelerated by the rate $(\log n)^{\varepsilon}$. However, Theorem 2 is the best possible in the sense that the power of n cannot be improved by a better one, since the estimate (1.8) is known in [16] to be the 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;

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

where $q=[n^{2/(2+\delta)}]$, $M=[n/q]+1\sim n^{\delta/(2+\delta)}$, $[a]$ 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 $0\leqq k\leqq M-1$ and $t_M=1$, where $\Delta:=q/n\sim n^{-\delta/(2+\delta)}$. Moreover define increments of the Brownian motion by $\eta_k:=(\eta_k^1, \eta_k^2, \cdots, \eta_k^r)=B(t_k)-B(t_{k-1})$, $1\leqq k\leqq M$. We approximate X and Y_n by the following processes $\bar{X}_n$ and $\bar{Y}_n$: 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 $1\leqq k\leqq M$ and $\tilde{X}_0=\tilde{Y}_0=X_0$. Define $D([0, 1]\to\boldsymbol{R}^d)$-valued processes $\bar{X}_n:=\{\bar{X}_n(t), 0\leqq t\leqq 1\}$ and $\bar{Y}_n:=\{\bar{Y}_n(t), 0\leqq t\leqq 1\}$ by $\bar{X}_n(t):=\tilde{X}_{k-1}$ and $\bar{Y}_n(t):=\tilde{Y}_{k-1}$ for $t_{k-1}\leqq t<t_k$, $1\leqq k\leqq M$, and $\bar{X}_n(1):=\tilde{X}_M$ and $\bar{Y}_n(1):=\tilde{Y}_M$ for $t=1$, respectively.

One of the main techniques of the proof of Theorem 1 is the following reconstruction of all random variables on a common probability space.

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

(i) *If* $d=r=1$, *then for any* $0<\varepsilon<2+\delta$ *and for each* $1\leqq k\leqq M$,

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

(ii) *If* $r>1$ *and* ξ_1 *has a bounded or square integrable density, then for any* $0<\varepsilon<2+\delta$ *and each* $1\leqq k\leqq M$,

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

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

$$\{\zeta_1, \zeta_2, \cdots, \zeta_k, \eta_1, \eta_2, \cdots, \eta_k\} \text{ is independent of } \{\zeta_{k+1}, \zeta_{k+2}, \cdots, \zeta_M, \eta_{k+1}, \eta_{k+2}, \cdots, \eta_M\}. \tag{2.3}$$

PROOF. Before proving the lemma we give several notation according to [3]. Let $x=(x^1, x^2, \cdots, x^r)\in \boldsymbol{R}^r$. For each $1\leqq k\leqq M$ and $1\leqq i\leqq r$, let $\mu_k^i(\cdot)$ be the propability measure of $((t_k-t_{k-1})^{-1/2}\zeta_k^1, (t_k-t_{k-1})^{-1/2}\zeta_k^2, \cdots, (t_k-t_{k-1})^{-1/2}\zeta_k^i)$ and $F_k^i(\cdot\,|\,x^1, \cdots, x^i)$ be the right continuous conditional distribution function defined by, for any bounded Borel function ψ,

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

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: $\boldsymbol{R}^r\to\boldsymbol{R}^1$ 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

$$\mathscr{L}\{h_1(t_1^{-1/2}\eta_1), h_2((t_2-t_1)^{-1/2}\eta_2), \cdots, h_M((t_M-t_{M-1})^{-1/2}\eta_M)\}$$
$$=\mathscr{L}\{t_1^{-1/2}\zeta_1, (t_2-t_1)^{-1/2}\zeta_2, \cdots, (t_M-t_{M-1})^{-1/2}\zeta_M\}. \tag{2.4}$$

Now, applying (2.4), we redefine processes $\{\xi_k\}$, $\{\zeta_k\}$ and $\{B(\cdot)\}$ such that (2.1)-(2.3) are satisfied as follows: Suppose that there is a Brownian motion $\hat{B}:=\{\hat{B}(t), t\geqq 0\}$ on another probability space $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$. Define

$$\text{(2.5)}\qquad \{\hat{\zeta}_1, \hat{\zeta}_2, \cdots, \hat{\zeta}_M\}:=\{t_1^{1/2}h_1(t_1^{-1/2}\hat{\eta}_1), (t_2-t_1)^{1/2}h_2((t_2-t_1)^{-1/2}\hat{\eta}_2), \cdots, (t_M-t_{M-1})^{1/2}h_M((t_M-t_{M-1})^{-1/2}\hat{\eta}_M)\},$$

where $\hat{\eta}_1, \cdots, \hat{\eta}_M$ are increments of $\hat{B}$ and $\mathscr{L}\{\hat{\eta}_k, 1\leqq k\leqq M\}=\mathscr{L}\{\eta_k, 1\leqq k\leqq M\}$. Then, from (2.4), $\mathscr{L}\{\hat{\zeta}_k, 1\leqq k\leqq M\}=\mathscr{L}\{\zeta_k, 1\leqq k\leqq M\}$. We next construct $\{\xi_k\}$ by the following method. Define probability measures U and V by, for any $A_1\in\mathscr{B}(\boldsymbol{R}^r\otimes\boldsymbol{R}^n)$, $A_2\in\mathscr{B}(\boldsymbol{R}^r\otimes\boldsymbol{R}^M)$ and $A_3\in\mathscr{B}(W^r)$,

$$U(A_1\times A_2):=P\{(\xi_1, \cdots, \xi_n)\in A_1, (\zeta_1, \cdots, \zeta_M)\in A_2\},$$
$$V(A_2\times A_3):=P\{(\hat{\zeta}_1, \cdots, \hat{\zeta}_M)\in A_2, \hat{B}\in A_3\}.$$

Put

$$U_{A_1}(A_2):=U(A_1\times A_2), \quad V_{A_3}(A_2):=V(A_2\times A_3),$$
$$H(A_2):=U(\boldsymbol{R}^r\otimes\boldsymbol{R}^M\times A_2)=V(A_2\times W^r).$$

Since $U_{A_1}(\cdot)$ is absolute continuous with respect to $H(\cdot)$, there exists a $\mathscr{B}(\boldsymbol{R}^r\otimes\boldsymbol{R}^M)$-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 $\mathscr{B}(\boldsymbol{R}^r\otimes\boldsymbol{R}^M)$-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 $(\boldsymbol{R}^r\otimes\boldsymbol{R}^n)\times(\boldsymbol{R}^r\otimes\boldsymbol{R}^M)\times W^r$ by

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

Finally define a new probability space $(\Omega', \mathscr{F}', P')$ by $\Omega':=(\boldsymbol{R}^r\otimes\boldsymbol{R}^n)\times(\boldsymbol{R}^r\otimes\boldsymbol{R}^M)\times W^r$, where $\mathscr{F}'$ is the completion of the topological σ-field $\mathscr{B}(\Omega')$ by Q and $P':=Q$. Keeping the relation (2.5) in mind, we can redefine $\{\xi_k\}$, $\{\zeta_k\}$ and $\{B(\cdot)\}$ without changing their distributions on the common probability space $(\Omega', \mathscr{F}', P')$ by putting for each $\omega=(\omega_1, \omega_2, \omega_3)\in\Omega'$,

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

Now, from (2.5), the relation (2.3) 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. □

Finally we need the following three lemmas. In what follows, as a positive constant independent of n, we use a K which may be different in the different equations and $L:=\max(L_1, L_2)$.

LEMMA 2 (Fuk [7]). *Let $\{\nu_k, 1\leqq k\leqq n\}$ be a square integrable martingale difference sequence with respect to a reference family of σ-fields $\{\mathscr{A}_k, 0\leqq k\leqq n\}$ such that $E(\nu_k\mid\mathscr{A}_{k-1})=0$ a.s. for each k. Suppose there exist sequences of positive numbers $\{g_k, 1\leqq k\leqq n\}$ and $\{h_k, 1\leqq k\leqq n\}$ such that*

$$E(\nu_k^2\mid\mathscr{A}_{k-1})\leqq g_k \quad \text{and} \quad E(|\nu_k|^{2+\delta}\mid\mathscr{A}_{k-1})\leqq h_k \qquad \text{a.s.}$$

for some $\delta>0$ and each k. Then for any positive v

$$P\Big\{\max_{1\leqq k\leqq n}\Big|\sum_{i=1}^{k}\nu_i\Big|\geqq u\Big\}\leqq\sum_{k=1}^{n}P\{|\nu_k|\geqq v\}$$
$$+2\exp\Big\{-\beta uv^{-1}\log\Big(\frac{\beta uv^{1+\delta}}{H}\Big)+1\Big\}+2\exp\Big(\frac{-\alpha^2u^2}{2e^{\alpha}G}\Big),$$

where $\alpha=2/(4+\delta)$, $\beta=1-\alpha$, $G=g_1+\cdots+g_n$ and $H=h_1+\cdots+h_n$.

LEMMA 3. *For any $2\leqq p\leqq 2+\delta$,*

$$E\Big[\max_{1\leqq k\leqq M}\max_{1\leqq i\leqq q}\Big|\sum_{j=(k-1)q+1}^{\{(k-1)q+i\}\wedge n}\frac{\sigma((j-1)/n,\hat{Y}_{j-1})\xi_j}{n^{1/2}}\Big|^p\Big]=O(\Delta^{p/2}(\log n)^{p/2}), \qquad (n\to\infty).$$

PROOF. Put

$$A:=\max_{1\leqq k\leqq M}\max_{1\leqq i\leqq q}\Big|\sum_{j=(k-1)q+1}^{\{(k-1)q+i\}\wedge n}\frac{\sigma((j-1)/n,\hat{Y}_{j-1})\xi_j}{n^{1/2}}\Big|^p.$$

By integration by parts, we have, for any $\lambda>0$,

(2.6)
$$E[A]=\int_{\{A\leqq\lambda\Delta^{p/2}(\log n)^{p/2}\}}AdP+\int_{\{A>\lambda\Delta^{p/2}(\log n)^{p/2}\}}AdP$$
$$\leqq\lambda\Delta^{p/2}(\log n)^{p/2}+\int_{\{A>\lambda\Delta^{p/2}(\log n)^{p/2}\}}AdP$$
$$\leqq 2\lambda\Delta^{p/2}(\log n)^{p/2}+\int_{\lambda\Delta^{p/2}(\log n)^{p/2}}^{\infty}P\{A>x\}dx.$$

Putting $\Delta^{p/2}y=x$, we have

$$(2.7)\quad \int_{\lambda\Delta^{p/2}(\log n)^{p/2}}^{\infty} P\{A>x\}dx=\Delta^{p/2}\int_{\lambda(\log n)^{p/2}}^{\infty} P\{A^{1/p}>\Delta^{1/2}y^{1/p}\}dy$$
$$\leqq \Delta^{p/2}\int_{\lambda(\log n)^{p/2}}^{\infty}\sum_{k=1}^{M} P\left\{\max_{1\leq i\leq q}\left[\sum_{m=1}^{d}\left(\sum_{l=1}^{r}\sum_{j=(k-1)q+1}^{\{(k-1)q+i\}\wedge n}\frac{\sigma_m^l((j-1)/n,\hat{Y}_{j-1})\xi_j^l}{n^{1/(2+\delta)}}\right)^2\right]^{1/2}>y^{1/p}\right\}dy$$
$$\leqq \sum_{m=1}^{d}\sum_{l=1}^{r}\Delta^{p/2}\int_{\lambda(\log n)^{p/2}}^{\infty}\sum_{k=1}^{M} P\left\{\max_{1\leq i\leq q}\left|\sum_{j=(k-1)q+1}^{\{(k-1)q+i\}\wedge n}\frac{\sigma_m^l((j-1)/n,\hat{Y}_{j-1})\xi_j^l}{n^{1/(2+\delta)}}\right|>\frac{y^{1/p}}{rd}\right\}dy\ .$$

Let $\mathscr{A}_k:=\sigma\{\xi_1,\cdots,\xi_k\}$ and $\nu_k^{lm}:=\sigma_m^l((j-1)/n,\hat{Y}_{j-1})\xi_j^l/n^{1/(2+\delta)}$ for each k, l and m. Let us agree to write ν_k, $\sigma(\cdot,\cdot)$ and ξ_k for ν_k^{lm}, $\sigma_m^l(\cdot,\cdot)$ and ξ_k^l. Since, from (1.5),

$$E(\nu_j^2\mid\mathscr{A}_{j-1})=\frac{\sigma((j-1)/n,\hat{Y}_{j-1})^2E[\xi_j^2]}{n^{2/(2+\delta)}}\leqq Kn^{-2/(2+\delta)}\quad \text{a.s.}\ ,$$

$$E(|\nu_j|^{2+\delta}\mid\mathscr{A}_{j-1})=\frac{|\sigma((j-1)/n,\hat{Y}_{j-1})|^{2+\delta}E[|\xi_j|^{2+\delta}]}{n}\leqq Kn^{-1}\quad \text{a.s.}\ ,$$

we can take $G=Kqn^{-2/(2+\delta)}=K$ and $H=qKn^{-1}=Kn^{-\delta/(2+\delta)}$. Furthermore we put $u=y^{1/p}/rd$ and $v=cy^{1/p}$ for some $c>0$ and apply Lemma 2 as follows;

$$(2.8)\quad P\left\{\max_{1\leq i\leq q}\left|\sum_{j=(k-1)q+1}^{\{(k-1)q+i\}\wedge n}\frac{\sigma((j-1)/n,\hat{Y}_{j-1})\xi_j}{n^{1/(2+\delta)}}\right|>\frac{y^{1/p}}{rd}\right\}$$
$$\leqq\sum_{j=(k-1)q+1}^{\{kq\}\wedge n}P\left\{\left|\frac{\sigma((j-1)/n,\hat{Y}_{j-1})\xi_j}{n^{1/(2+\delta)}}\right|>cy^{1/p}\right\}$$
$$+\exp\left[-\frac{\beta}{crd}\log\left(\frac{\beta y^{(2+\delta)/p}c^{1+\delta}n^{\delta/(2+\delta)}}{Krd}\right)+1\right]$$
$$+2\exp\left(\frac{-\alpha^2y^{2/p}}{2e^{\alpha}Kr^2d^2}\right)$$
$$\leqq\sum_{j=(k-1)q+1}^{\{kq\}\wedge n}P\left\{|\xi_j|>\frac{n^{1/(2+\delta)}y^{1/p}}{L}\right\}$$
$$+Ky^{-\beta(2+\delta)/crpd}n^{-\beta\delta/crd(2+\delta)}+2\exp\left(\frac{-\alpha^2y^{2/p}}{2e^{\alpha}Lr^2d^2}\right)\ .$$

We can easily see that

$$(2.9)\quad M\Delta^{p/2}\int_{\lambda(\log n)^{p/2}}^{\infty}\exp\left(\frac{-\alpha^2y^{2/p}}{2e^{\alpha}Lr^2d^2}\right)dy\leqq K\Delta^{p/2}(\log n)^{p/2}$$

for sufficiently large $\lambda>0$, and

$$(2.10)\quad M\Delta^{p/2}\int_{\lambda(\log n)^{p/2}}^{\infty}y^{-\beta(2+\delta)/crpd}n^{-\beta\delta/crd(2+\delta)}dy$$
$$\leqq K\Delta^{p/2}(\log n)^{p/2}$$

for sufficiently small $c>0$. Furthermore,

$$\begin{aligned}(2.11)\qquad &\Delta^{p/2}\int_{\lambda(\log n)^{p/2}}^{\infty}\sum_{k=1}^{M}\sum_{j=(k-1)q+1}^{\{kq\}\wedge n}P\{|\xi_j|>n^{1/(2+\delta)}y^{1/p}/L\}dy\\ &\leqq n\Delta^{p/2}\int_1^{\infty}P\{|\xi_1|^{2+\delta}>Kny\}dy\\ &\leqq(\Delta^{p/2}/K)\int_{Kn}^{\infty}P\{|\xi_1|^{2+\delta}>x\}dx\\ &\leqq(\Delta^{p/2}/K)E[|\xi_1|^{2+\delta}]\,.\end{aligned}$$

Hence we finish the proof from (2.6)-(2.11). □

LEMMA 4. *For any positive T,*

$$(2.12)\qquad E[\max_{0\leqq s\leqq T}\max_{s\leqq t\leqq s+T\Delta}|B(t)-B(s)|^p]=O(\Delta^{p/2}(\log n)^{p/2})\,,\qquad (n\to\infty)\,.$$

PROOF. Let $u_k:=Tt_k$ for each k. For any s with $u_{l-1}\leqq s\leqq u_l$,

$$\begin{aligned}(2.13)\qquad &\max_{s\leqq t\leqq s+T\Delta}|B(t)-B(s)|=\max\{\max_{s\leqq t\leqq u_l}|(B(u_{l-1})-B(t))-(B(u_{l-1})-B(s))|\,,\\ &\qquad\max_{u_l\leqq t\leqq s+T\Delta}|(B(u_{l-1})-B(u_l))+(B(u_l)-B(t))-(B(u_{l-1})-B(s))|\}\\ &\leqq\max\{2\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|\,,\\ &\qquad 2\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|+\max_{u_l\leqq t\leqq u_{l+1}}|B(t)-B(u_l)|\}\\ &=2\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|+\max_{u_l\leqq t\leqq u_{l+1}}|B(t)-B(u_l)|\,.\end{aligned}$$

Let $J:=\max_{1\leqq l\leqq M}\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|$ and $I\{\cdot\}$ be the indicator function of $\cdot$. Then the right hand side of (2.13) is bounded by

$$\begin{aligned}(2.14)\quad 3E[J^p]&=3\int_{\{J>2\Delta^{1/2}(\log n)^{1/2}\}}J^p dP+3\int_{\{J\leqq 2\Delta^{1/2}(\log n)^{1/2}\}}J^p dP\\ &\leqq 3\int J^p\cdot I\Big[\bigcup_{l=1}^{M}\{\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|>2\Delta^{1/2}(\log n)^{1/2}\,,\\ &\qquad\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|\geqq\max_{k\neq l}\max_{u_{k-1}\leqq t\leqq u_k}|B(t)-B(u_{k-1})|\}\Big]dP\\ &\qquad+6\Delta^{p/2}(\log n)^{p/2}\\ &\leqq 3\sum_{l=1}^{M}\int I\{\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|>2\Delta^{1/2}(\log n)^{1/2}\}\\ &\qquad\times\max_{u_{l-1}\leqq t\leqq u_l}|B(t)-B(u_{l-1})|^p dP+6\Delta^{p/2}(\log n)^{p/2}\,.\end{aligned}$$

Now we can easily see that

$$(2.15)\quad \int I\{\max_{u_{l-1}\leq t\leq u_l}|B(t)-B(u_{l-1})|>2\Delta^{1/2}(\log n)^{1/2}\}\times\max_{u_{l-1}\leq t\leq u_l}|B(t)-B(u_{l-1})|^p dP$$

$$\leq 2\int I\{\max_{u_{l-1}\leq t\leq u_l}(B(t)-B(u_{l-1}))>2\Delta^{1/2}(\log n)^{1/2}\}$$

$$\times\max_{u_{l-1}\leq t\leq u_l}\{(B(t)-B(u_{l-1}))\vee 0\}^p dP$$

$$\leq K\Delta^{p/2}\int_{2(\log n)^{1/2}}^{\infty} x^p\Phi(dx)\leq K\Delta^{p/2}n^{-1}\ .$$

Hence (2.12) follows from (2.13)-(2.15). □

§ 3. Proof of Theorem 1.

Note $\|X-Y_n\|\leq\|X-\bar{X}_n\|+\|\bar{X}_n-\bar{Y}_n\|+\|\bar{Y}_n-Y_n\|$. To estimate the left hand side, we consider three terms on the right hand side, each of which will be estimated in the following three lemmas.

LEMMA 5. *For any* $2\leq p\leq 2+\delta$,

$$E[\|X-\bar{X}_n\|^p]=O(\Delta^{p/2}(\log n)^{p/2})\ ,\qquad (n\to\infty)\ .$$

PROOF. For simplicity, we prove only the case $r=1$. The case $r>1$ can be treated similarly. 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, for $t_k\leq t<t_{k+1}$, we have

$$(3.1)\quad X(t)-\bar{X}_n(t)=\int_0^t(\sigma(s, X(s))-\sigma_n(s))dB(s)+\int_{t_k}^t\sigma_n(s)dB(s)$$

$$+\int_0^t(b(s, X(s))-b_n(s))ds+\int_{t_k}^t b_n(s)ds$$

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

$$+\int_0^t(\sigma(s, \bar{X}_n(s))-\sigma_n(s))dB(s)+\int_{t_k}^t\sigma_n(s)dB(s)$$

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

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

We first estimate the third term of the right hand side of (3.1).

$$(3.2)\quad E\left[\max_{1\leq l\leq k+1}\max_{t_{l-1}\leq s<t_l}\left|\int_{t_{l-1}}^s\sigma_n(u)dB(u)\right|^p\right]$$

$$=E[\max_{1\leq l\leq k+1}\max_{t_{l-1}\leq s<t_l}|B(A(s))-B(A(t_{l-1}))|^p]$$

where $A(t):=\int_0^t \sigma_n(s)^2 ds$ is the quadratic variation process of the martingale $N(t):=\int_0^t \sigma_n(s)dB(s)$. By assumption (1.5) we have

$$A(t)\leqq L^2 t \quad \text{and} \quad A(t+\Delta)-A(t)\leqq L^2\Delta \qquad \text{a.s.}$$

Thus the right hand side of (3.2) is bounded by

$$\begin{aligned}(3.3)\qquad & E[\max_{0\leqq s\leqq t_k}\max_{s\leqq t\leqq s+\Delta}|B(A(t))-B(A(s))|^p] \\ &= E[\max_{0\leqq s\leqq t_k}\max_{A(s)\leqq t\leqq A(s+\Delta)}|B(t)-B(A(s))|^p] \\ &\leqq E[\max_{0\leqq s\leqq L^2 t_k}\max_{s\leqq t\leqq s+L^2\Delta}|B(t)-B(s)|^p] .\end{aligned}$$

Hence, combining (3.2), (3.3) with (2.12) of Lemma 4, we have

$$(3.4)\qquad E\left[\max_{1\leqq l\leqq k+1}\max_{t_{l-1}\leqq s<t_l}\left|\int_{t_{l-1}}^s \sigma_n(u)dB(u)\right|^p\right]\leqq K\Delta^{p/2}(\log n)^{p/2} .$$

Now from a moment inequality for martingales (see for example Theorem 3.1 in Ikeda-Watanabe [8], Chapter III), Jensen's inequality and condition (1.2), we have

$$\begin{aligned}(3.5)\qquad & E\left[\max_{0\leqq s\leqq t}\left|\int_0^s\{\sigma(u, X(u))-\sigma(u, \bar{X}_n(u))\}dB(u)\right|^p\right] \\ &\leqq KE\left[\int_0^t|\sigma(u, X(u))-\sigma(u, \bar{X}_n(u))|^2 du\right]^{p/2} \\ &\leqq K\int_0^t E[|\sigma(u, X(u))-\sigma(u, \bar{X}_n(u))|^p]du \\ &\leqq KL^p\int_0^t E[|X(u)-\bar{X}_n(u)|^p]du \\ &\leqq KL^p\int_0^t E[\max_{0\leqq u\leqq s}|X(u)-\bar{X}_n(u)|^p]ds .\end{aligned}$$

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

$$\begin{aligned} E[|\sigma(t, \bar{X}_n(t))-\sigma_n(t)|^p] &\leqq L^p E[(|t-t_{k-1}|^p+|\tilde{X}_k-\tilde{X}_{k-1}|^p)] \\ &\leqq K\Delta^p+KE[|\sigma(t_{k-1}, \tilde{X}_{k-1})\eta_k|^p]+KE[|b(t_{k-1}, \tilde{X}_{k-1})(t_k-t_{k-1})|^p] \\ &\leqq K\Delta^p+K\Delta^{p/2} .\end{aligned}$$

Hence, in the same way as in (3.5), we have

$$\begin{aligned}(3.6)\qquad & E\left[\max_{0\leqq s\leqq t}\left|\int_0^s\{\sigma(u, \bar{X}_n(u))-\sigma_n(u)\}dB(u)\right|^p\right] \\ &\leqq K\int_0^t E[|\sigma(u, \bar{X}_n(u))-\sigma_n(u)|^p]du\leqq K\Delta^{p/2} .\end{aligned}$$

Furthermore from Jensen's inequality,

$$(3.7)\qquad E\left[\max_{0\leq s\leq t}\left|\int_0^s\{b(u,\ X(u))-b(u,\ \bar{X}_n(u)\}du\right|^p\right] \leq KL^p\int_0^t E[\max_{0\leq u\leq s}|X(u)-\bar{X}(u)|^p]ds\ ,$$

$$(3.8)\qquad E\left[\max_{0\leq s\leq t}\left|\int_0^s\{b(u,\ \bar{X}_n(u))-b_n(u)\}du\right|^p\right] \leq K\int_0^s E[|b(u,\ \bar{X}_n(u))-b_n(u)|^p]du\leq K\Delta^{p/2}\ .$$

Moreover, from (1.5),

$$(3.9)\qquad E\left[\max_{1\leq l\leq k+1}\max_{t_{l-1}\leq s<t_l}\left|\int_{t_{l-1}}^s b_n(u)du\right|^p\right] \leq E\left[\max_{1\leq l\leq k+1}\max_{t_{l-1}\leq s<t_l}\left\{\int_{t_{l-1}}^s Ldu\right\}^p\right]\leq L^p\Delta^p\ .$$

Combining (3.1) and (3.4)-(3.9) and using Gronwall's inequality, we complete the proof of the lemma. □

LEMMA 6. *For any* $2\leq p\leq 2+\delta$,

$$E[\|Y_n-\bar{Y}_n\|^p]=O(\Delta^{p/2}(\log n)^{p/2})\ ,\qquad (n\to\infty)\ .$$

PROOF. We prove only the case $r=1$ again. For $t_k\leq t<t_{k+1}$,

$$\begin{aligned}
\max_{t_k\leq s\leq t}|Y_n(s)-\bar{Y}_n(s)| &= \max_{kq<i\leq[nt]+1}|\hat{Y}_i-\tilde{Y}_k| \\
&\leq \left|\sum_{j=1}^{kq}\frac{\sigma((j-1)/n,\ \hat{Y}_{j-1})\xi_j}{n^{1/2}}-\sum_{l=1}^{k}\sigma(t_{l-1},\ \tilde{Y}_{l-1})\zeta_l\right| \\
&\quad+\max_{kq<i\leq[nt]+1}\left|\sum_{j=kq+1}^{i}\frac{\sigma((j-1)/n,\ \hat{Y}_{j-1})\xi_j}{n^{1/2}}\right| \\
&\quad+\left|\sum_{j=1}^{kq}\frac{b((j-1)/n,\ \hat{Y}_{j-1})}{n}-\sum_{l=1}^{k}\frac{b(t_{l-1},\ \tilde{Y}_{l-1})q}{n}\right| \\
&\quad+\max_{kq<i\leq[nt]+1}\left|\sum_{j=kq+1}^{i}\frac{b((j-1)/n,\ \hat{Y}_{j-1})}{n}\right| \\
&\leq \left|\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{\{\sigma((j-1)/n,\ \hat{Y}_{j-1})-\sigma(t_{l-1},\ \tilde{Y}_{l-1})\}\xi_j}{n^{1/2}}\right| \\
&\quad+\max_{kq<i\leq[nt]+1}\left|\sum_{j=kq+1}^{i}\frac{\sigma((j-1)/n,\ \hat{Y}_{j-1})\xi_j}{n^{1/2}}\right| \\
&\quad+\left|\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{\{b((j-1)/n,\ \hat{Y}_{j-1})-b(t_{l-1},\ \tilde{Y}_{l-1})\}}{n}\right| \\
&\quad+\sum_{j=kq+1}^{[nt]+1}\frac{|b((j-1)/n,\ \hat{Y}_{j-1})|}{n}\ .
\end{aligned}$$

Thus, by Doob's inequality,

$$(3.10)\qquad E[\max_{0\leq s\leq t}|\bar{Y}_n(s)-Y_n(s)|^p]$$
$$\leq KE\Big[\Big|\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{\{\sigma((j-1)/n,\hat{Y}_{j-1})-\sigma(t_{l-1},\tilde{Y}_{l-1})\}\xi_j}{n^{1/2}}\Big|^p\Big]$$
$$+KE\Big[\max_{1\leq l\leq k}\max_{1\leq i\leq q}\Big|\sum_{j=(l-1)q+1}^{\{(l-1)q+i\}\wedge n}\frac{\sigma((j-1)/n,Y_{j-1})\hat{\xi}_j}{n^{1/2}}\Big|^p\Big]$$
$$+KE\Big[\Big|\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{b((j-1)/n,\hat{Y}_{j-1})-b(t_{l-1},\tilde{Y}_{l-1})}{n}\Big|^p\Big]$$
$$+KE\Big[\max_{1\leq l\leq k}\max_{1\leq i\leq q}\Big\{\sum_{j=(l-1)q+1}^{\{(l-1)q+i\}\wedge n}\frac{|b((j-1)/n,\hat{Y}_{j-1})|}{n}\Big\}^p\Big]$$
$$=:D_1+D_2+D_3+D_4\ ,$$

say. Since $\{(\sigma((j-1)/n,\hat{Y}_{j-1})-\sigma(t_{l-1},\tilde{Y}_{l-1}))\xi_j,\ (l-1)q+1\leq j\leq lq\}$ and $\{\sigma((j-1),\hat{Y}_{j-1})\xi_j),\ 1\leq j\leq n\}$ are martingale differences, we have from Theorem 3.1 in [5], (1.2) and (1.5) that

$$(3.11)\qquad D_1\leq Kn^{-p/2}E\Big[\Big\{\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}E[\{\sigma((j-1)/n,\hat{Y}_{j-1})-\sigma(t_{l-1},\tilde{Y}_{l-1})\}^2\xi_j^2|\mathscr{G}_{j-1}]\Big\}^{p/2}\Big]$$
$$\leq KE\Big[\Big\{\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{|\hat{Y}_{j-1}-\tilde{Y}_{l-1}|^2+\Delta^2}{n}\Big\}^{p/2}\Big]$$
$$\leq K\int_0^t E[\max_{0\leq u\leq s}|Y_n(u)-\bar{Y}_n(u)|^p]ds+K\Delta^p\ ,$$

where $\mathscr{G}_j$ is the σ-field generated by $\xi_1,\cdots,\xi_j$ for each j. Now, using Lemma 3, we have

$$(3.12)\qquad D_2\leq K\Delta^{p/2}(\log n)^{p/2}\ .$$

Furthermore, from Jensen's inequality and (1.5), we have

$$(3.13)\qquad D_3\leq KE\Big[\Big\{\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{|\hat{Y}_{j-1}-\tilde{Y}_{l-1}|+\Delta}{n}\Big\}^p\Big]$$
$$\leq K\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{E[|\hat{Y}_{j-1}-\tilde{Y}_{l-1}|^p]}{n}+K\sum_{l=1}^{k}\sum_{j=(l-1)q+1}^{lq}\frac{\Delta^p}{n}$$
$$\leq K\int_0^t E[\max_{0\leq u\leq s}|Y_n(u)-\bar{Y}_n(u)|^p]du+K\Delta^p\ ,$$

$$(3.14)\qquad D_4\leq K\Big(\frac{Lq}{n}\Big)^p\leq K\Delta^p\ .$$

Combining (3.10)-(3.14) and using Gronwall's inequality, we conclude the lemma. □

LEMMA 7. *We can redefine the processes* $\bar{X}_n$ *and* $\bar{Y}_n$ *on a richer probability space such that the following relations hold:*

(i) *If* $d=r=1$, *then for any* $2\leqq p\leqq 2+\delta$,

$$E[\|\bar{X}_n-\bar{Y}_n\|^p]=O(\Delta^{p/2}), \qquad (n\to\infty).$$

(ii) *If* $r>1$ *and* ξ_1 *has a bounded or square integrable density, then for any* $2\leqq p\leqq 2+\delta$,

$$E[\|\bar{X}_n-\bar{Y}_n\|^p]=O(\Delta^{1/2}(\log n)^{p/2}), \qquad (n\to\infty).$$

PROOF. Let $d=r=1$. For $t_k\leqq t<t_{k+1}$,

$$\begin{aligned}(3.15)\quad \max_{0\leqq s\leqq t}|\bar{X}_n(s)-\bar{Y}_n(s)| &= \max_{1\leqq i\leqq k}|\tilde{X}_i-\tilde{Y}_i| \\ &\leqq KE\Big[\max_{1\leqq i\leqq k}\Big|\sum_{j=1}^{i}\sigma(t_{j-1},\tilde{X}_{j-1})\eta_j-\sum_{j=1}^{i}\sigma(t_{j-1},\tilde{Y}_{j-1})\zeta_j\Big|^p\Big] \\ &\quad +KE\Big[\max_{1\leqq i\leqq k}\Big|\sum_{j=1}^{i}b(t_{j-1},\tilde{X}_{j-1})(t_j-t_{j-1})-\sum_{j=1}^{i}b(t_{j-1},\tilde{Y}_{j-1})(t_j-t_{j-1})\Big|^p\Big] \\ &\leqq KE\Big[\max_{1\leqq i\leqq k}\Big|\sum_{j=1}^{i}\{\sigma(t_{j-1},\tilde{X}_{j-1})-\sigma(t_{j-1},\tilde{Y}_{j-1})\}\eta_j\Big|^p\Big] \\ &\quad +KE\Big[\max_{1\leqq i\leqq k}\Big|\sum_{j=1}^{i}\sigma(t_{j-1},\hat{Y}_{j-1})(\eta_j-\zeta_j)\Big|^p\Big] \\ &\quad +KE\Big[\max_{1\leqq i\leqq k}\Big|\sum_{j=1}^{i}\{b(t_{j-1},\tilde{X}_{j-1})-b(t_{j-1},\tilde{Y}_{j-1})\}(t_j-t_{j-1})\Big|^p\Big] \\ &=: I_1+I_2+I_3,\end{aligned}$$

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 $t_j\leqq s<t_{j+1}$, $j=0,1,\cdots,M-1$. Since $\{\sigma_n'(s), s<t_k\}$ is $\sigma\{\eta_1,\eta_2,\cdots,\eta_{k-1}\}$-measurable because of relation (2.4) of Lemma 1, I_1 is represented by

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

Thus, from Theorem 3.1 in [5] and condition (1.2),

$$\begin{aligned}(3.16)\quad I_1 &\leqq KE\Big[\Big|\int_0^{t_k}\sigma_n'(s)dB(s)\Big|^p\Big]\leqq K\int_0^{t_k}E[|\sigma_n'(s)|^p]ds \\ &\leqq K\int_0^t E[\max_{0\leqq u\leqq s}|\bar{X}_n(u)-\bar{Y}_n(u)|^p]du.\end{aligned}$$

We next estimate I_2. By (2.3),

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

for each j, where $\mathscr{H}_j$ is the σ-field generated by $\eta_1,\cdots,\eta_j,\zeta_1,\cdots,\zeta_j$ for each $1\leqq j\leqq M$. Thus, from Theorem 3.1 in [8] and (1.5), (2.1), we have

$$I_2\leqq KE\Big[\sum_{j=1}^{k}E(|\sigma(t_{j-1},\tilde{Y}_{j-1})(\eta_j-\zeta_j)|^2|\mathscr{H}_{j-1})\Big]^{p/2} \tag{3.17}$$
$$\leqq KE\Big[\sum_{j=1}^{k}|\sigma(t_{j-1},\tilde{Y}_{j-1})|^2E[|\eta_j-\zeta_j|^2]\Big]^{p/2}$$
$$\leqq KL^p\Big\{\sum_{j=1}^{k}E[|\eta_j-\zeta_j|^2]\Big\}^{p/2}$$
$$\leqq K(k\Delta^{(2-\delta)/2}n^{-\delta/2})^{p/2}\leqq Kt\Delta^{p/2}\,.$$

When $r>1$, we can similarly prove that

$$I_2\leqq KL^p\Big\{\sum_{j=1}^{k}E[|\eta_j-\zeta_j|^2]\Big\}^{p/2}\,,$$

and by using (2.2) instead of (2.1), we have

$$I_2\leqq K(k\Delta^{(2-\delta)/2}n^{-\delta/2}\log n)^{p/2}\leqq Kt\Delta^{p/2}(\log n)^{p/2}\,. \tag{3.18}$$

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 (1.2) that

$$I_3=KE\Big[\max_{0\leqq j\leqq k}\Big|\int_0^{t_j}b_n'(s)ds\Big|^p\Big]\leqq K\int_0^{t_k}E[|b_n'(s)|^p]ds \tag{3.19}$$
$$\leqq K\int_0^t E[\max_{0\leqq u\leqq s}|\bar{X}_n(u)-\bar{Y}_n(u)|]^p du\,.$$

Combining (3.15)-(3.19) we have

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

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

PROOF OF THEOREM 1. Without changing distributions we can reconstruct W^d-valued processes X and Y_n on the common probability space $(\Omega',\mathscr{F}',P')$ by Lemmas 1 and 4-6 such that the conclusion of Theorem 1 holds, namely, for any $\varepsilon>p/2$,

$$l_p(P^X,P^{Y_n})\leqq E[\|X-Y_n\|^p]^{1/p}\leqq K\Delta^{p/2}(\log n)^{p/2}=o(\Delta^{p/2}(\log n)^\varepsilon)$$

as $n\to\infty$. Therefore we finish the proof of Theorem 1. □

ACKNOWLEDGEMENT. The author would like to express his gratitude to Professors H. Tanaka, M. Maejima and S. T. Rachev for their valuable suggestions and comments. He also thanks the referee for his helpful comments and remarks.

References

[1] A. A. BOROVKOV, On the rate of convergence for the invariance principle, Theory Probab. 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 Probab. Appl., **20** (1975), 631–638.

[4] H. S. GREENSIDE and E. HELFAND, Numerical integration of stochastic differential equations—II, Bell System Techn. J., **60** (1981), 1927–1940.

[5] S. JANKOVIĆ, Iterative procedure for solving stochastic differential equations, Math. Balkanica, **1** (1987), 64–71.

[6] R. JANSSEN, Discretization of the Wiener-process in difference-methods for stochastic differential equations, Stochastic Process. Appl., **18** (1984), 361–369.

[7] D. KH. FUK, Some probabilistic inequalities for martingales, Siberian Math. J., **14** (1973), 131–137.

[8] N. IKEDA and S. WATANABE, *Stochastic Differential Equations and Diffusion Processes*, North-Holland/Kodansha, Amsterdam/Tokyo, 1981.

[9] G. MARUYAMA, Continuous Markov processes and stochastic equations, Rend. Circ. Mat. Palermo, **4** (1955), 48–90.

[10] G. N. MILSHTEIN, A method of second-order accuracy integration of stochastic differential equations, Theory Probab. Appl., **23** (1978), 396–401.

[11] E. PLATEN, An approximation method for a class of Ito Processes, Litovsk. Math. Sb., **21** (1981), 121–133.

[12] E. PLATEN, An approximation method for a class of Ito Processes with jump component, Ibid., **22** (1982), 124–136.

[13] S. T. RACHEV, On a class of minimal function on a space of probability measures, Theory Probab. Appl., **29** (1984), 41–49.

[14] N. J. RAO, J. D. BORWANKER and D. RAMKRISHNA, Numerical solution of Ito integral equations, SIAM J. Control, **12** (1974), 124–139.

[15] W. RÜMELIN, Numerical treatment of stochastic differential systems, Technical Report No. 12, Forschungsschwerpunkt Dynamische Systeme der Universität Bremen, 1980.

[16] A. I. SAHANENKO, Estimates of rate of convergence in the invariance principle, Dokl. Akad. Nauk USSR, **219** (1974), 1752–1755.

[17] A. SHIMIZU, Approximate solutions for stochastic differential equations, *Proc. Symp. on Stochastic Differential Equations for Population Genetics*, (1984), 201–210, (in Japanese).

[18] D. J. WRIGHT, The digital simulation of stochastic differential equations, IEEE Trans. Automatic Control, **19** (1974), 75–76.

Present Address:
DEPARTMENT OF MATHEMATICS, YAMANASHI UNIVERSITY
KOFU 400, JAPAN