MATHEMATICS OF COMPUTATION
Volume 67, Number 221, January 1998, Pages 287–298
S 0025-5718(98)00888-6

APPROXIMATION OF CONTINUOUS TIME STOCHASTIC PROCESSES BY A LOCAL LINEARIZATION METHOD

ISAO SHOJI

ABSTRACT. This paper investigates the rate of convergence of an alternative approximation method for stochastic differential equations. The rates of convergence of the one-step and multi-step approximation errors are proved to be $O((\Delta t)^2)$ and $O(\Delta t)$ in the L_p sense respectively, where Δt is discrete time interval. The rate of convergence of the one-step approximation error is improved as compared with methods assuming the value of Brownian motion to be known only at discrete time. Through numerical experiments, the rate of convergence of the multi-step approximation error is seen to be much faster than in the conventional method.

1. INTRODUCTION

It is often convenient to model the time evolution of dynamic systems by using continuous time stochastic processes whose dynamics are characterized by stochastic differential equations. However, except for a simple process, it is difficult not only to estimate parameters of continuous time stochastic processes from real data, but also to obtain a good discrete approximation of the processes. To overcome the intractability of continuous time stochastic processes, many methods have been proposed. For example, Kutoyants [4] and Yoshida [11] have proposed estimation methods for stochastic differential equations by using the maximum likelihood technique. On the other hand, Milstein [5, 6], Rümelin [9], Chang [1], and Newton [7] have proposed approximation methods by using Taylor's expansion or Runge-Kutta methods. Although these estimation and approximation methods are useful for each purpose, it seems more convenient to use a unified method for both purposes.

Ozaki [8] and Shoji and Ozaki [10] propose alternative methods, called a local linearization method, which are used for both estimation and approximation purposes. In particular, the method proposed by Shoji and Ozaki [10] shows better performance in estimation through numerical experiments than the method proposed by Ozaki [8] and the Euler method. From a theoretical point of view, it seems interesting how close the process approximated by the local linearization method is to the true process. This, however, has not been fully investigated.

The aim of this paper is to evaluate the goodness of approximation for the local linearization method proposed by Shoji and Ozaki [10] in terms of the rate of convergence. When evaluating the rate of convergence, the previous studies focus on the mean squared approximation error defined as the difference between the

Received by the editor May 19, 1996 and, in revised form, September 4, 1996.
1991 *Mathematics Subject Classification.* Primary 65D30, 65B33; Secondary 60H10.
Key words and phrases. Stochastic differential equations, discretization, rate of convergence, numerical approximation.

true process and the approximate process. Considering the approximation error, we should note that two kinds of approximation errors may be defined; one is an error on one-step approximation, and another on multi-step approximation.

The one-step approximation error has important implication in estimating parameters of a process. As is described above, estimation is usually carried out by the maximum likelihood method, and its likelihood is generally constructed from the distribution of a one-step ahead state. Since the exact distribution cannot be obtained except for a simple process, some approximation methods must be used. Thus, the better the rate of convergence of the one-step approximation error, the more efficient likelihood is expected to be obtained. However, as long as assuming that the value of Brownian motion is known only at discrete time, except for the special case, the rate of convergence of the mean squared errors of one-step approximation is shown to be no faster than $O((\Delta t)^3)$ by Rümelin [9], where Δt is discrete time interval. Instead of the above assumption, we may assume that the squared value of Brownian motion is used as well or the value of Brownian motion is continuously known. In this case, Milstein [5, 6] and Chang [1] show that the rate of convergence of the one-step approximation error can be improved. Indeed, this approach is effective in improving the rate of convergence, but is not appropriate in estimation because it is difficult to obtain the likelihood.

On the other hand, the rate of convergence of the multi-step approximation error must be considered when constructing a sample path of a stochastic process. Specifically, to obtain a sample path of a stochastic process, we first discretize the original stochastic process by an approximation method and generate step-by-step sample points of the discretized process. The step-by-step approximation of a sample path can be improved by shortening the discrete time interval and so this improvement can be evaluated from the multi-step approximation error. However, as reported by Newton [7], Clark and Cameron [2] have shown that the multi-step approximate process can converge to the true process in the root-mean-square sense, that is the L_2 sense, no faster than linearly in the discrete time interval.

In this paper, we study the rate of convergence of the one-step and multi-step approximation errors induced by the local linearization method. Unlike the previous studies, the rate of convergence is evaluated in the L_p sense instead of the L_2 sense. In addition, numerical experiments are carried out to evaluate numerically the rates of convergence of the one-step and multi-step approximation errors. In the experiments, we compare the local linearization method with the conventional discretization method.

The organization of this paper is as follows. In section 2, the alternative approximation method is introduced, and in section 3, the rates of convergence of the one-step and multi-step approximation errors are presented. Numerical experiments are carried out in section 4 and some concluding remarks are given in section 5.

2. Discretization by the local linearization method

In this section, we describe the local linearization method proposed by Shoji and Ozaki [10] and derive a discretized process. First, we are interested in a one-dimensional stochastic process x_t which satisfies the following stochastic differential equation:

$$dx_t = f(x_t, t)dt + g(x_t)dB_t, \tag{1}$$

where $f(x_t, t)$ is twice continuously differentiable with respect to x_t, and continuously differentiable with respect to t, $g(x_t)$ is a continuously differentiable function of x_t, and B_t is a standard Brownian motion. The above stochastic differential equation can be transformed into a differential equation with a constant coefficient of the diffusion term. Let $\varphi(x)$ be a function such that $g\frac{d\varphi}{dx} = \sigma$, where σ is constant. By Ito's formula the stochastic differential equation of $y_t = \varphi(x_t)$ is

$$dy_t = \left(f\frac{d\varphi}{dx} + \frac{g^2}{2}\frac{d^2\varphi}{dx^2} \right) dt + \sigma dB_t. \tag{2}$$

Thus, we have only to consider the following stochastic differential equation:

$$dx_t = f(x_t, t)dt + \sigma dB_t, \tag{3}$$

where $f(x_t, t)$ is defined on $W \times [0, T_b)$. We assume that W is a compact set in the one-dimensional Euclidean space, T_b is fixed, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial t}$ and $\frac{\partial^2 f}{\partial x^2}$ exist and are continuous on $W \times [0, T_b)$, and that σ is constant.

The local linearization method is a method of approximation by which the drift function $f(x_t, t)$ in (3) is locally approximated by a linear function of x_t. Thus, we need to focus on the local behavior of $f(x_t, t)$. This behavior can be formulated by the differential of $f(x_t, t)$ and so this differential can be characterized by Ito's formula, which gives

$$df = \left(\frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} + \frac{\partial f}{\partial t} \right) dt + \frac{\partial f}{\partial x} dx. \tag{4}$$

Now, in order to linearize f with respect to x_t and t, we assume that $\frac{\partial^2 f}{\partial x^2}(x_t, t)$, $\frac{\partial f}{\partial x}(x_t, t)$, $\frac{\partial f}{\partial t}(x_t, t)$ are constant on the small time interval $[s, s + \Delta t)$. With this assumption, (4) can be solved as follows:

$$f(x_t, t) - f(x_s, s) = \left(\frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} + \frac{\partial f}{\partial t} \right)(t - s) + \frac{\partial f}{\partial x}(x_t - x_s). \tag{5}$$

Thus,

$$f(x_t, t) = L_s x_t + M_s t + N_s, \tag{6}$$

where

$$\begin{aligned} L_s &= \frac{\partial f}{\partial x}(x_s, s), \\ M_s &= \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}(x_s, s) + \frac{\partial f}{\partial t}(x_s, s), \\ N_s &= f(x, s) - \frac{\partial f}{\partial x}(x_s, s)x_s - \left(\frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}(x_s, s) + \frac{\partial f}{\partial t}(x_s, s) \right) s. \end{aligned}$$

Therefore, instead of (3) we have only to focus on the following linear stochastic differential equation as long as t belongs to $[s, s + \Delta t)$:

$$dx_t = (L_s x_t + M_s t + N_s)dt + \sigma dB_t. \tag{7}$$

Transforming x_t into $y_t = \mathrm{e}^{-L_s t}x_t$, the stochastic differential equation with respect to y_t can be solved as follows:

$$y_{s+\Delta t} = y_s + \int_s^{s+\Delta t} (M_s u + N_s)\mathrm{e}^{-L_s u}du + \sigma \int_s^{s+\Delta t} \mathrm{e}^{-L_s u} dB_u. \tag{8}$$

As a result, we obtain a discretized process of x_t:

$$x_{s+\Delta t} = x_s + \frac{f(x_s,s)}{L_s}(\mathrm{e}^{L_s\Delta t}-1) + \frac{M_s}{{L_s}^2}\left\{(\mathrm{e}^{L_s\Delta t}-1) - L_s\Delta t\right\} + \sigma\int_s^{s+\Delta t}\mathrm{e}^{L_s(s+\Delta t-u)}dB_u, \tag{9}$$

where

$$\begin{aligned} L_s &= \frac{\partial f}{\partial x}(x_s,s), \\ M_s &= \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}(x_s,s) + \frac{\partial f}{\partial t}(x_s,s). \end{aligned}$$

The above discretized process is easy to manipulate since the fourth term of the right-hand side of (9) follows the Normal distribution with mean 0 and variance,

$$\sigma^2 \cdot \frac{\mathrm{e}^{2L_s\Delta t}-1}{2L_s\Delta t}. \tag{10}$$

3. Rate of convergence of the local linearization method

In this section, we give the rate of convergence of the one-step and multi-step approximation errors. Here, we define the one-step and multi-step approximation errors as the difference between the state of the true process and that of the approximate process on the assumption that the current state of the true stochastic process, x_s, is known and bounded. The rate of convergence is evaluated in the L_p sense.

Recalling that a Brownian motion does not have finite variation for any time interval, the process x_t satisfying (3) or (7) cannot be expected to be bounded. In a practical situation, however, no infinite state of a process can be treated. Therefore, by introducing a stopping time, we establish a bounded process induced by x_t. First, for each $M \geq 1$, a stopping time T_M is defined as follows:

$$T_M = \begin{cases} s; & |x_s| \geq M, \\ \inf\{t | t \geq s, |B_t - B_s| > M\}; & |x_s| < M. \end{cases} \tag{11}$$

Clearly, T_M is non-decreasing and $\lim_{M\to\infty} T_M = \infty$, almost surely. By taking x_t as $x_{T_M \wedge t}$, we may assume x_t to be a bounded process. In addition, making M sufficiently large, we may assume $s < T_M$. Now, T is fixed such that $s < T < \infty$ and $T < T_b$. In this setting, we can easily get the following known results.

Lemma 1. *If $\phi(u) = O(u^q)$, then $\int_0^u \phi(w)dw = O(u^{q+1})$.*

Lemma 2 (Gronwall's inequality, see Karatzas and Shreve [3])**.** *Suppose*

$$0 \leq g(t) \leq \alpha(t) + \beta\int_s^t g(u)du,$$

then,

$$g(t) \leq \alpha(t) + \beta\mathrm{e}^{\beta t}\int_s^t \alpha(u)\mathrm{e}^{-\beta u}du.$$

Lemma 3. *Suppose $s \leq t \leq T$, then there exists a function ϕ such that,*

$$E_s\,|x_t - x_s|^q \leq \phi(t-s),$$

where E_s is designated as conditional expectation on time s and $\phi(u) = O(u^{\frac{q}{2}})$ or $O(u^{\frac{q+1}{2}})$ depending upon whether q is even or odd.

Lemma 4. *Let x_t and $\tilde{x}_t$ be the true stochastic process and the approximate stochastic process derived from the proposed local linearization method. Suppose that $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial t}$, $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial x \partial t}$ exist and are continuous on $W \times [0,T]$. Then, there exists a constant C such that*

$$\begin{aligned}
&\int_s^t |f(x_u,u) - (L_s\tilde{x}_u + M_s u + N_s)|^p \, du \\
&\leq C\Bigg\{\left((t-s) + \int_s^t |x_u - x_s|^p du\right)|x_s - \tilde{x}_s|^p + \int_s^t |x_u - x_s|^{2p} du \\
&\quad + \left(2 + \frac{\sigma^2}{2}\right)^p \frac{(t-s)^{p+1}}{p+1} + \frac{1}{2^p}\frac{(t-s)^{2p+1}}{2p+1} + \int_s^t |x_u - \tilde{x}_u|^p du\Bigg\}
\end{aligned}$$

where $s \leq t \leq T$.

Using the above lemmas, we prove the following theorem which presents the rate of convergence of the one-step approximation error in the L_p sense.

Theorem 1. *Suppose that x_t and $\tilde{x}_t$ are the true stochastic process and the approximate process derived from the proposed local linearization method, respectively. Let a $p-th$ order error of one-step ahead prediction be defined as $E_s|x_t - \tilde{x}_t|^p$ with $s \leq t \leq T$. Then, the rate of convergence of $(E_s|x_t - \tilde{x}_t|^p)^{1/p}$ is 2, in other words,*

$$E_s|x_t - \tilde{x}_t|^p = O((t-s)^{2p}),$$

where E_s is designated as conditional expectation on time s.

Proof.

$$\begin{aligned}
|x_t - \tilde{x}_t|^p &= \left|x_s - \tilde{x}_s + \int_s^t (f(x_u,u) - (L_s\tilde{x}_u + M_s u + N_s))du\right|^p \\
&\leq 2^{p-1}\left(|x_s - \tilde{x}_s|^p + \left|\int_s^t (f(x_u,u) - (L_s\tilde{x}_u + M_s u + N_s))du\right|^p\right) \\
&\leq 2^{p-1}\left(|x_s - \tilde{x}_s|^p + (t-s)^{p-1}\int_s^t |f(x_u,u) - (L_s\tilde{x}_u + M_s u + N_s)|^p \, du\right).
\end{aligned}$$

The second line and the third line are derived from Jensen's inequality and Hölder's inequality, respectively. From Lemma 4,

$$\begin{aligned}
&\int_s^t |f(x_u,u) - (L_s\tilde{x}_u + M_s u + N_s)|^p \, du \\
&\leq C\Bigg\{\left((t-s) + \int_s^t |x_u - x_s|^p du\right)|x_s - \tilde{x}_s|^p \\
&\quad + \int_s^t |x_u - \tilde{x}_u|^p du + \int_s^t |x_u - x_s|^{2p} du \\
&\quad + \left(2 + \frac{\sigma^2}{2}\right)^p \frac{(t-s)^{p+1}}{p+1} + \frac{1}{2^p}\frac{(t-s)^{2p+1}}{2p+1}\Bigg\}.
\end{aligned}$$

Since the one-step ahead prediction error is considered, we may assume $x_s = \tilde{x}_s$. Therefore,

$$E_s|x_t - \tilde{x}_t|^p \leq 2^{p-1}(t-s)^{p-1}C\left\{E_s\left[\int_s^t |x_u - \tilde{x}_u|^p du\right] + E_s\left[\int_s^t |x_u - x_s|^{2p} du\right] + \left(2 + \frac{\sigma^2}{2}\right)^p \frac{(t-s)^{p+1}}{p+1} + \frac{1}{2^p}\frac{(t-s)^{2p+1}}{2p+1}\right\}.$$

From Lemma 1 and Lemma 3,

$$E_s\left[\int_s^t |x_u - x_s|^{2p} du\right] \leq \phi(t-s), \tag{12}$$

where $\phi(u) = O(u^{p+1})$. Thus,

$$E_s|x_t - \tilde{x}_t|^p \leq \alpha(t-s) + \beta E_s\left[\int_s^t |x_u - \tilde{x}_u|^p du\right], \tag{13}$$

where,

$$\alpha(u) = 2^{p-1}C\left(u^{p-1}\phi(u) + \left(2 + \frac{\sigma^2}{2}\right)^p \frac{u^{2p}}{p+1}\frac{1}{2^p}\frac{u^{3p}}{3p+1}\right),$$
$$\beta = 2^{p-1}C(T-s)^{p-1}.$$

From Gronwall's inequality,

$$E_s|x_t - \tilde{x}_t|^p \leq \alpha(t-s) + \beta e^{\beta(t-s)} \int_s^t \alpha(u-s)e^{-\beta(u-s)} du. \tag{14}$$

Clearly, the right-hand side of the above inequality has order $2p$. □

Next, we prove a theorem which presents the rate of convergence of the multi-step approximation error. The multi-step approximation error exhibits the cumulative error of the local linearization.

Theorem 2. *Let t be fixed at $s \leq t \leq T$ and $\{t_k\}_{0 \leq k \leq n}$ be an $n-$partition of the time interval $[s,t]$ with $t_k - t_{k-1} = \Delta t$. Consider a step-by-step approximation of integration*

$$x_t - x_s = \int_s^t f(x_u, u)du + \sigma \int_s^t dB_u,$$

by using the proposed local linearization method. Then the convergence of the step-by-step approximation is $O(\Delta t)$ in the L_p sense.

Proof. For $s = t_0 < t_1 < \cdots < t_n = t$ with $t_k - t_{k-1} = \Delta t$, the step-by-step approximation of the local linearization is given by

$$\sum_{k=1}^n \int_{t_{k-1}}^{t_k} d\tilde{x}_u = \sum_{k=1}^n \int_{t_{k-1}}^{t_k} (L_{k-1}\tilde{x}_u + M_{k-1}u + N_{k-1})du + \sum_{k=1}^n \sigma \int_{t_{k-1}}^{t_k} dB_u,$$

where L_{k-1}, M_{k-1} and N_{k-1} are equivalent to $L_{t_{k-1}}$, $M_{t_{k-1}}$ and $N_{t_{k-1}}$ that are defined in (6). Since the left-hand side of the above equation is reduced to $\tilde{x}_t - \tilde{x}_s$,

$$\begin{aligned}
|x_t - \tilde{x}_t|^p &= \left| \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} (f(x_u, u) - (L_{k-1}\tilde{x}_u + M_{k-1}u + N_{k-1}))du \right|^p \\
&\le n^{p-1} \sum_{k=1}^{n} \left| \int_{t_{k-1}}^{t_k} (f(x_u, u) - (L_{k-1}\tilde{x}_u + M_{k-1}u + N_{k-1}))du \right|^p \\
&\le n^{p-1} \sum_{k=1}^{n} (t_k - t_{k-1})^{p-1} \int_{t_{k-1}}^{t_k} |f(x_u, u) - (L_{k-1}\tilde{x}_u + M_{k-1}u + N_{k-1})|^p \, du \\
&= (t-s)^{p-1} \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} |f(x_u, u) - (L_{k-1}\tilde{x}_u + M_{k-1}u + N_{k-1})|^p \, du.
\end{aligned}$$

The second line and the third line are derived from Jensen's inequality and Hölder's inequality, respectively. From Lemma 4,

$$\begin{aligned}
&\int_{t_{k-1}}^{t_k} |f(x_u, u) - (L_{k-1}\tilde{x}_u + M_{k-1}u + N_{k-1})|^p \, du \\
&\quad \le C \left(\Delta t + \int_{t_{k-1}}^{t_k} |x_u - x_{t_{k-1}}|^p du \right) |x_{t_{k-1}} - \tilde{x}_{t_{k-1}}|^p \\
&\quad + C \int_{t_{k-1}}^{t_k} |x_u - x_{t_{k-1}}|^{2p} du \\
&\quad + C \left(2 + \frac{\sigma^2}{2} \right)^p \frac{(\Delta t)^{p+1}}{p+1} + \frac{C}{2^p} \frac{(\Delta t)^{2p+1}}{2p+1} \\
&\quad + C \int_{t_{k-1}}^{t_k} |x_u - \tilde{x}_u|^p du.
\end{aligned} \tag{15}$$

From Lemma 3,

$$\begin{aligned}
\int_{t_{k-1}}^{t_k} |x_u - x_{t_{k-1}}|^p du &\le \varphi_1(\Delta t), \\
\int_{t_{k-1}}^{t_k} |x_u - x_{t_{k-1}}|^{2p} du &\le \varphi_2(\Delta t),
\end{aligned}$$

where $\varphi_1(\Delta t) = O((\Delta t)^{\frac{p}{2}+1})$ or $O((\Delta t)^{\frac{p+1}{2}+1})$ depending upon whether q is even or odd and $\varphi_2(\Delta t) = O((\Delta t)^{p+1})$. For convenience of the proof, $\varphi_1(\Delta t)$ and $\varphi_2(\Delta t)$ are replaced by $\Delta t \phi_1(\Delta t)$ and $\Delta t \phi_2(\Delta t)$, respectively, where $\phi_1(\Delta t) = O((\Delta t)^{\frac{p}{2}})$ or $O((\Delta t)^{p+\frac{1}{2}})$ and $\phi_2(\Delta t) = O((\Delta t)^p)$. Summing up the right-hand side of (15)

from t_0 to t_n, we get

$$\begin{aligned}|x_t - \tilde{x}_t|^p \leq\ & \tilde{C}(t-s)\Delta t(1+\phi_1(\Delta t))\sum_{k=1}^{n}|x_{t_{k-1}} - \tilde{x}_{t_{k-1}}|^p + \tilde{C}(t-s)\phi_2(\Delta t) \\ & + \tilde{C}(t-s)\left(2+\frac{\sigma^2}{2}\right)^p \frac{(\Delta t)^p}{p+1} + \frac{\tilde{C}(t-s)}{2^p}\frac{(\Delta t)^{2p}}{2p+1} \\ & + \tilde{C}\int_s^t |x_u - \tilde{x}_u|^p du,\end{aligned}$$

where $\tilde{C} = C(t-s)^{p-1}$. From Gronwall's inequality,

$$E_s|x_t - \tilde{x}_t|^p \leq \alpha(t-s) + \beta e^{\beta(t-s)}\int_s^t \alpha(u-s)e^{-\beta(u-s)}du, \tag{16}$$

where

$$\begin{aligned}\alpha(u) &= u\psi_1(\Delta t)\rho_n + u\psi_2(\Delta t), \\ \beta &= \tilde{C}, \\ \rho_n &= E_s\left[\sum_{k=1}^{n}|x_{t_{k-1}} - \tilde{x}_{t_{k-1}}|^p\right], \\ \psi_1(\Delta t) &= \tilde{C}\Delta t(1+\phi_1(\Delta t)), \\ \psi_2(\Delta t) &= \tilde{C}\phi_2(\Delta t) + \tilde{C}\left(2+\frac{\sigma^2}{2}\right)^p \frac{(\Delta t)^p}{p+1} + \frac{\tilde{C}}{2^p}\frac{(\Delta t)^{2p}}{2p+1}.\end{aligned}$$

Clearly, $\psi_1(\Delta t) = O(\Delta t)$ and $\psi_2(\Delta t) = O((\Delta t)^p)$. Rewriting the above inequality by using $\psi_1(\cdot)$, $\psi_2(\cdot)$ and $G(\cdot)$,

$$E_s|x_t - \tilde{x}_t|^p \leq \psi_1(\Delta t)G(t-s)\rho_n + \psi_2(\Delta t)G(t-s), \tag{17}$$

where

$$G(t-s) = (t-s) + \beta e^{\beta(t-s)}\int_s^t (u-s)e^{-\beta(u-s)}du.$$

Then,

$$\rho_{n+1} - \rho_n \leq \psi_1 G\rho_n + \psi_2 G. \tag{18}$$

Noticing that $\psi_1 G$ is non-negative and that $\rho_1 = |x_s - \tilde{x}_s|^p = 0$,

$$\rho_n \leq (1+\psi_1 G)^{n-1}\frac{\psi_2}{\psi_1} - \frac{\psi_2}{\psi_1}. \tag{19}$$

Using (17) and (19), we get,

$$E_s|x_t - \tilde{x}_t|^p \leq (1+\psi_1 G)^{n-1}\psi_2 G. \tag{20}$$

Since $n = (t-s)/\Delta t$,

$$\begin{aligned}\log\left((1+\psi_1 G)^{n-1}\right) &= \frac{n-1}{n}n\psi_1 G\frac{\log(1+\psi_1 G)}{\psi_1 G} \\ &= \left(1-\frac{\Delta t}{t-s}\right)\frac{(t-s)\psi_1 G}{\Delta t}\frac{\log(1+\psi_1 G)}{\psi_1 G}.\end{aligned}$$

Since $\psi_1(\Delta t) \to 0$ $(\Delta t \to 0)$ and $\psi_1(\Delta t)/\Delta t \to \tilde{C}$, the right-hand side of the above equality converges to a constant (depending on $t-s$). Noticing that $\psi_2(\Delta t) =$

$O((\Delta t)^p)$ and G is independent of Δt, $(1+\psi_1 G)^{n-1}\psi_2 G/(\Delta t)^p$ also converges to a constant. □

4. Numerical experiments

In this section, numerical experiments are carried out to evaluate the rate of convergence of the one-step and multi-step approximation errors induced by the proposed local linearization method. We compare the local linearization method (LL) with the Euler method (Euler) in the L_1 and L_2 sense.

4.1. **Method of experiments.** For numerical experiments, the following nonlinear stochastic differential equation is used:

$$dx_t = -x_t(x_t^2 - 1)dt + dB_u. \tag{21}$$

To evaluate numerically the rate of convergence of the one-step ahead approximation error, we first calculate a Δt ahead state of the true process, $x_{t+\Delta t}$. Since it is difficult to obtain the exact discretized process of (21), we use an approximation method to get a realization of $x_{t+\Delta t}$. Considering that almost all numerical experiments are carried out by the Euler method, we also use this method in our experiments. This approximation is reasonable if the discrete time interval is sufficiently small. Hence, applying repeatedly the Euler method with discrete time interval Δt_g and setting $\Delta t = m\Delta t_g$, we calculate a Δt ahead state and assume it to be $x_{t+\Delta t}$. This calculation is carried out as follows. Let x_t be an initial state and $\{s_i\}_{0\le i\le m}$ be an $m-$partition of time interval $[t, t+\Delta t]$, where $s_i - s_{i-1} = \Delta t_g$. Applying the Euler method to (21), we get,

$$x_{s_i} = x_{s_{i-1}} + f(x_{s_{i-1}})(s_i - s_{i-1}) + (B_{s_i} - B_{s_{i-1}}), \tag{22}$$

where $f(x) = -x(x^2-1)$ and $B_{s_i} - B_{s_{i-1}}$ is a realization of a Normal random variable with mean 0 and variance Δt_g. Then we assume x_{s_m} to be a realization of $x_{t+\Delta t}$. In our experiments, we set $\Delta t_g = 1/4000$ and use the four alternative x_t and Δt; $x_t = 0.0$, 1.0, 2.0, 3.0 and $\Delta t = 0.005$, 0.01, 0.02, 0.04.

Secondly, we construct a Δt ahead state, $\tilde{x}_{t+\Delta t}$, induced by the LL method. This is obtained from (9). While the deterministic part of (9) is easy to calculate, we use the following approximation to calculate the stochastic part:

$$\begin{aligned}\int_t^{t+\Delta t} \mathrm{e}^{L_t(t+\Delta t-u)}dB_u &= \sum_{i=1}^{m}\int_{s_{i-1}}^{s_i} \mathrm{e}^{L_t(t+\Delta t-u)}dB_u \\ &\approx \sum_{i=1}^{m} \mathrm{e}^{L_t(t+\Delta t-s_{i-1})}(B_{s_i} - B_{s_{i-1}}).\end{aligned} \tag{23}$$

Lastly, averaging N $(= 10000)$ realizations of $|x_{t+\Delta t} - \tilde{x}_{t+\Delta t}|^p$ derived from the above procedures, we get $E_t|x_{t+\Delta t} - \tilde{x}_{t+\Delta t}|^p$ $(p = 1 \text{ or } 2)$.

Next, to evaluate numerically the rate of convergence of the multi-step approximation error, we first set up an $n-$partition of time interval $[t, t+1]$; that is, $t = t_0 < t_1 < \cdots < t_n = t+1$ and $t_k - t_{k-1} = \Delta t$. For each $k-$step, the above procedure is used to get a realization of the Δt ahead approximation error, where $x_{t_{k-1}}$ and $\tilde{x}_{t_{k-1}}$ are used as the initial states of the true process x_t and the approximate process $\tilde{x}_t$, respectively. Like the evaluation of the one-step approximation

TABLE 1. One-step approximation error

		$E_t\lvert x_{t+\Delta t} - \tilde{x}_{t+\Delta t}\rvert$		$E_t\lvert x_{t+\Delta t} - \tilde{x}_{t+\Delta t}\rvert^2$	
		LL	Euler	LL	Euler
$x_t = 0.0$	Δt				
	0.005	1.50938e-05	0.000153496	3.72941e-10	3.71496e-08
	0.01	2.54511e-05	0.000440247	1.24931e-09	3.02164e-07
	0.02	6.18635e-05	0.00124171	1.08285e-08	2.38862e-06
	0.04	0.000235754	0.00346198	2.09957e-07	1.79897e-05
$x_t = 1.0$	Δt				
	0.005	3.79232e-05	0.000312567	2.7568e-09	1.56451e-07
	0.01	0.000117147	0.000904327	2.8217e-08	1.31725e-06
	0.02	0.000446495	0.00255695	4.05045e-07	1.09742e-05
	0.04	0.00176649	0.00717693	6.22168e-06	9.34308e-05
$x_t = 2.0$	Δt				
	0.005	0.000178436	0.0278133	5.65982e-08	0.000777987
	0.01	0.000377579	0.0557419	2.94886e-07	0.00314094
	0.02	0.00119972	0.107707	3.20086e-06	0.011835
	0.04	0.00443204	0.198674	4.4913e-05	0.0408887
$x_t = 3.0$	Δt				
	0.005	0.000628459	0.10717	6.66681e-07	0.0115074
	0.01	0.00192129	0.207632	6.17473e-06	0.0432641
	0.02	0.00874093	0.378569	0.000109119	0.144213
	0.04	0.0413803	0.637447	0.00208985	0.410556

error, averaging N $(= 10000)$ realizations of $|x_{t+1} - \tilde{x}_{t+1}|^p$, we get $E_t|x_{t+1} - \tilde{x}_{t+1}|^p$ $(p = 1$ or $2)$. We use the same values of x_t and Δt as in evaluation of the one-step approximation error.

Similar to the LL case, the rate of convergence of the one-step and multi-step approximation errors induced by the Euler method is evaluated.

4.2. **Results of experiments.** Results of experiments for the one-step and multi-step approximation errors are shown in Table 1 and Table 2 respectively. For all cases, the errors induced by the LL method are much smaller than the errors induced by the Euler method. In particular, the differences between the two methods for multi-step approximation are very pronounced. The efficiency of the LL method in multi-step approximation mainly follows from the construction of a stochastic integration. In the LL method, the stochastic integration (23) uses the information of increments of a Brownian motion with time interval Δt_g. In the Euler method, however, the sum of increments of a Brownian motion is reduced to $B_{t+1} - B_t$, and so the intermediate information is lost as a result. Thus, the LL method shows much better performance than the Euler method. Of course, in practice, the information of increments of a Brownian motion cannot be expected to be used in constructing such a stochastic integration as (23). However, since we know that the stochastic integration of (23) follows the Normal distribution, a sample path of the approximate process can be generated without numerical error caused by constructing the stochastic integration.

TABLE 2. Multi-step approximation error

		$E_t\lvert x_{t+1} - \tilde{x}_{t+1}\rvert$		$E_t\lvert x_{t+1} - \tilde{x}_{t+1}\rvert^2$	
		LL	Euler	LL	Euler
$x_t = 0.0$					
Δt	n				
0.005	200	0.000283681	0.170943	1.30954e-07	0.0500009
0.01	100	0.000660692	0.178169	7.13403e-07	0.0557444
0.02	50	0.00179558	0.181939	5.25284e-06	0.0589751
0.04	25	0.00495413	0.183869	4.08175e-05	0.0606949
$x_t = 1.0$					
Δt	n				
0.005	200	0.000340106	0.342291	1.85408e-07	0.219675
0.01	100	0.000771964	0.372901	9.5541e-07	0.270491
0.02	50	0.00214957	0.390301	7.46015e-06	0.302821
0.04	25	0.00612755	0.399641	6.07739e-05	0.321292
$x_t = 2.0$					
Δt	n				
0.005	200	0.000450752	0.835303	3.46061e-07	0.8527
0.01	100	0.000817214	0.968199	1.17836e-06	1.16991
0.02	50	0.00213726	1.052	7.83151e-06	1.40262
0.04	25	0.00605888	1.10001	6.46092e-05	1.54841
$x_t = 3.0$					
Δt	n				
0.005	200	0.000535419	1.3057	5.21037e-07	1.79508
0.01	100	0.00091965	1.60859	1.61988e-06	2.75481
0.02	50	0.00234248	1.82524	1.07743e-05	3.5837
0.04	25	0.00687399	1.96071	0.000105844	4.16684

5. CONCLUDING REMARKS

In this paper we studied the rate of convergence of the one-step and multi-step approximation errors induced by the alternative approximation method called a local linearization method. In particular, unlike previous studies, the rate of convergence was evaluated in the L_p sense instead of in the L_2 sense.

For discrete time interval Δt, the rate of convergence of the one-step approximation error is $O((\Delta t)^2)$ in the L_p sense, which is faster than any other approximate processes derived from the assumption that the value of Brownian motion is known only at discrete time. On the other hand, the rate of convergence of the multi-step approximation error is $O(\Delta t)$ which is optimal as shown by Clark and Cameron [2].

Through the numerical comparison of the rate of convergence, it was shown that the approximate process derived from the local linearization method was much closer to the true process in one-step and multi-step approximation than the approximate process derived from the Euler method. In particular, the efficiency of the local linearization method in multi-step approximation has important implication for generating a discrete sample path of a diffusion process. The process discretized by the local linearization method follows the Normal distribution and we know how to simulate a random variable following the Normal distribution. Thus,

there is no approximation error resulting from the construction of a stochastic integration of the discretized process. To a large extent, samples generated by the local linearization method can be treated as a good approximation of the realization of the true process.

Acknowledgement

The author thanks Professor Ozaki and Professor Kitagawa of the Institute of Statistical Mathematics for many helpful comments and suggestions. The author also thanks the referee and the editor for their helpful comments.

References

1. Chang, C. C., *Numerical solutions of stochastic differential equations with constant diffusion coefficients*, Mathematics of Computation **49** (1987), 523–542. MR **88j:**65305
2. Clark, J. M. C. and Cameron, R. J., *The maximum rate of convergence of discrete approximations for stochastic differential equations*, Lecture Notes in Control and Information Science **25** (1980), Springer-Verlag, Berlin, New York. MR **82f:**60133
3. Karatzas, I. and Shreve, S. E., *Brownian Motion and Stochastic Calculus*: 2nd edition, (1992), (Springer-Verlag). MR **92h:**60127
4. Kutoyants, Y. A., *Parameter estimation for stochastic processes*, (1984), (Heldermann). MR **86b:**62145
5. Milstein, G. N., *Approximate integration of stochastic differential equations*, Theory of Probability and its Application **19** (1974), 557–562. MR **50:**8696
6. ———, *A method of second-order accuracy integration of stochastic differential equations*, Theory of Probability and its Application **23** (1978), 396–401. MR **58:**24549
7. Newton, N. J., *Asymptotically efficient Runge-Kutta methods for a class of Ito and Storatonovich equations*, *SIAM* Journal on Applied Mathematics **51**, (1991), 542–567. MR **92e:**60121
8. Ozaki, T., *Statistical identification of storage models with application to stochastic hydrology*, Water Resources Bulletin **21** (1985), 663–675.
9. Rümelin, W., *Numerical treatment of stochastic differential equations*, *SIAM* Journal on Numerical Analysis **19** (1982), 604–613. MR **83i:**60075
10. Shoji, I. and Ozaki, T., *Estimation for nonlinear stochastic differential equations by a local linearization method*, forthcoming in Stochastic Analysis and Applications.
11. Yoshida, N., *Estimation for diffusion processes from discrete observation*, Journal of Multivariate Analysis **41** (1992), 220–242. MR **93g:**62113

Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba Ibaraki 305, Japan

E-mail address: shoji@shako.sk.tsukuba.ac.jp