A NEW SIMULATION METHOD OF DIFFUSION PROCESSES
APPLIED TO FINANCE

SHIGEO KUSUOKA AND SYOITI NINOMIYA

ABSTRACT. The authors apply a new simulation method of diffusion processes
to finance problems and show this new method realizes extremely fast calcula-
tion in some cases.

0. INTRODUCTION

Numerical calculation of $E[f(X(T,x))]$, where $X(t,x)$ denotes a diffusion process
that is a solution of some SDE and $f$ is a function with some regularity, is very
important in many fields. Pricing of derivative securities is a typical example. Many
studies have been done in this field (cf. [1] [3]). In [4] and [5], a new simulation
scheme is proposed. The method is based on Malliavin calculus and higher order
stochastic expansion and we can consider that it belongs to the higher order scheme.
[6] and [7] show that this method achieves very fast calculation when it is applied
to some finance problems. In [11], Talay and Tubaro reported that the higher
order scheme does not have practical advantages over Euler-Maruyama scheme
with acceleration methods. In this paper, we give some numerical examples and
show that this new method with the proposed sampling methods and acceleration
methods works in practice and outperforms existing methods in some cases.

0.1. Notation. $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$ are sets of all natural numbers, all integers, and all
real numbers, respectively. We also set

\[
\mathbb{R}_{>a} = \{r \in \mathbb{R} \mid r > a\}
\]

\[
\mathbb{Z}_{\geq n} = \{i \in \mathbb{Z} \mid i \geq n\}
\]

and so on. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the integer part of $x$, and $\text{fp}(x) = x - \lfloor x \rfloor$.

1. A NEW APPROXIMATION SCHEME

We introduce the new approximation scheme following [4] and [5].
1.1. The Problem. Let \((\Omega, \mathcal{F}, P)\) be a probability space. In this paper, we consider the numerical calculation of \(E[f(X(T, x))]\) where \(T \in \mathbb{R}_{>0}\) and \(X\) is a diffusion process determined by the following stochastic differential equation:

\[
X(t, x) = x + \sum_{j=0}^{d} \int_{0}^{t} V_j(X(s, x)) \circ dB^j(s)
\]

where \(\circ dB^j(s)\) denotes the Stratonovich integral, \(B^0(t) = t, (B^1(t), \ldots, B^d(t))\) is the \(d\)-dimensional standard Brownian Motion, and \(f\) is an \(\mathbb{R}\)-valued function defined on \(\mathbb{R}^n\). We often identify \(V_j \in C_\infty^0(\mathbb{R}^N; \mathbb{R}^N)\) with a section of the tangent bundle \(T\mathbb{R}^N\) in the following manner: \(V_j(x) = (V_j^1(x), \ldots, V_j^N(x)) \approx \sum_{i=1}^{N} V_j^i(x)(\partial/\partial x_i)\). Here \(C_\infty^0(\mathbb{R}^N; \mathbb{R}^N)\) denotes the set of \(\mathbb{R}^N\) valued smooth functions defined over \(\mathbb{R}^N\) whose derivatives of any order are bounded. The following equation (2) is the Ito stochastic differential equation corresponding to (1).

\[
X(t, x) = x + \sum_{j=0}^{d} \int_{0}^{t} V_j(X(s, x)) \, dB^j(s).
\]

Let \(\{P_t\}_{t \geq 0}\) be the semigroup of linear operators defined by

\[
(P_tf)(x) = E[f(X(t, x))]
\]

for \(t \geq 0\) and \(f \in C_\infty^0(\mathbb{R}^N)\).

1.2. The Euler-Maruyama approximation. The Euler-Maruyama approximation \(X^{(n)}(t, x)\) of the process \(X(t, x)\) is constructed by a difference equation \(X^{(n)}(t_{i+1}, x) = X^{(n)}(t_i, x) + \sqrt{T/n} Z, X^{(n)}(0, x) = x\), where \(t_i = iT/n\) \((i = 0, \ldots, n - 1)\), \(Z = (Z^0, Z^1, \ldots, Z^d)\), \(Z^i = \sqrt{T/n}\) \((i = 1, \ldots, d)\) are independent random variables that follow normal distribution \(N(0, 1)\). It is well known that when \(f\) is of class \(C_4^0\),

\[
|(P_T f)(x) - E[f(X^{(n)}(T, x))]| = O(n^{-1}).
\]

1.3. The new approximation scheme. We describe the new approximation scheme by using the notions of free Lie algebra [9].

1.3.1. Free Lie algebra. Let \(A = \{v_0, v_1, \ldots, v_d\}, (d \geq 1)\) be an alphabet, that is to say, a set of letters. \(A^*\) denotes the free monoid on \(A\), that is to say, the set of all words on \(A\) with concatenation product. The empty word is the identity of \(A^*\) and denoted by 1. For an arbitrary multi-index \(\alpha = (\alpha_1, \ldots, \alpha_k), (\alpha_i \in \{0, 1, \ldots, d\}, i = 1, \ldots, k), v_\alpha\) denotes the word \(v_{\alpha_1} \cdots v_{\alpha_k}\) and \(|v_\alpha|\) is defined as \(k\). We also define \(|v_\alpha|\) to be \(|v_\alpha| + \text{card}(\{i|\alpha_i = 0\})\). \(A_\ast\) denotes \(\{w \in A^*| |w| = m\}\) and \(A_\ast\leq m\) denotes \(\{w \in A^*| |w| \leq m\}\). Let \(K\) be a field, \(K\langle A\rangle\) be the \(K\)-coefficients free algebra with basis \(A^\ast\), and \(K\langle\langle A\rangle\rangle\) be the set of all \(K\)-coefficients formal series with basis \(A^\ast\). \(K\langle A\rangle\) is a sub \(K\)-algebra of \(K\langle\langle A\rangle\rangle\). We call the element of \(K\langle A\rangle\), non-commutative polynomial. For all \(P \in K\langle\langle A\rangle\rangle\), we write them as

\[
P = \sum_{w \in A^\ast} (P, w) w \quad \text{or} \quad \sum_{w \in A^\ast} a_w w,
\]
where \((P, w) = a_{w} \in K\) denotes the coefficient of \(w\). The algebra structure is defined as usual, that is to say,

\[
\left( \sum_{w \in A^{*}} a_{w} w \right) \left( \sum_{w \in A^{*}} b_{w} w \right) = \sum_{w = uv \in A^{*}} a_{u} b_{v} w.
\]

The Lie bracket is defined as \([x, y] = xy - yx\) for \(x, y \in K\). We define \(L_{K}(A)\) to be the set of Lie polynomials in \(K\langle A \rangle\) and \(L_{K}(\langle A \rangle)\) to be the set of Lie series. That is to say, \(L_{K}(A)\) is the smallest \(K\)-module of \(K\langle A \rangle\) that includes \(A\) and is closed under the Lie bracket, and \(L_{K}(\langle A \rangle)\) is the set of elements of \(K\langle A \rangle\) whose every homogeneous components belong to \(L_{K}(A)\). For \(m \in \mathbb{Z}_{\geq 0}\), we define \(K\langle A \rangle_{m}, K\langle A \rangle_{\leq m}\), and \(j_{m} : K\langle \langle A \rangle \rangle \longrightarrow K\langle A \rangle_{\leq m}\) as follows:

\[
K\langle A \rangle_{m} = \{P \in K\langle A \rangle \mid (P, w) = 0, \text{ if } \|w\| \neq m\}
\]

\[
K\langle A \rangle_{\leq m} = \{P \in K\langle A \rangle \mid (P, w) = 0, \text{ if } \|w\| > m\}
\]

\[
j_{m} \left( \sum_{w \in A^{*}} a_{w} w \right) = \sum_{\|w\| \leq m} a_{w} w.
\]

For arbitrary \(P, Q \in K\langle A \rangle\), the inner product \(\langle P, Q \rangle\) is defined as follows:

\[
\langle P, Q \rangle = \sum_{w \in A^{*}} (P, w)(Q, w).
\]

For an arbitrary \(P \in K\langle \langle A \rangle \rangle\) whose constant term equals to 0, that is to say, \((P, 1) = 0\), we can define its exponential \(\exp(P)\) to be \(1 + \sum_{k=1}^{\infty} P^{k}/k!\). We can also define \(\log(Q)\) to be \(\sum_{k=1}^{\infty} (-1)^{k-1}(Q - 1)^{k}/k\) for \(Q \in K\langle \langle A \rangle \rangle\) if its constant term equals to 1, that is to say, \((Q, 1) = 1\). The following relations hold:

\[
\log(\exp(P)) = P, \quad \text{and} \quad \exp(\log(Q)) = Q.
\]

In the following part of this paper, we consider only the case in which \(K = \mathbb{R}\). We define \(\|P\|_{2} = (\langle P, P \rangle)^{1/2}\), for \(P \in K\langle A \rangle\).

1.3.2. Stochastic Taylor expansion. For the process \(X\) defined in (1) and an arbitrary \(f \in C_{b}^{\infty}(\mathbb{R}^{N})\), we obtain the following expansion:

\[
f(X(t, x)) = f(x) + \sum_{i=0}^{d} \int_{0}^{t} (V_{i}f)(X(s, x)) \circ dB^{i}(s)
\]

\[
\vdots
\]

\[
f(x) + \sum_{j=1}^{k} \sum_{i_{1}, \ldots, i_{j}=0}^{d} (V_{i_{1}} \ldots V_{i_{j}}f)(x)B(t; i_{1} \ldots i_{j}) + R_{k}^{f}(t, x)
\]

where

\[
B(t; u_{i}) = B^{i}(t), \quad \text{for} \ i \in \{0, 1, \ldots, d\},
\]

\[
B(t; uv_{i}) = \int_{0}^{t} B(s; u) \circ dB^{i}(s), \quad \text{for} \ u \in A^{*}, \ i \in \{0, 1, \ldots, d\}.
\]
We extend the definition of $B(t;P)$ to the case in which $P \in \mathbb{R}(A)$ by

$$B \left( t; \sum_{w \in A^*} a_w w \right) = \sum_{w \in A^*} a_w B(t;w).$$

1.3.3. $\mathbb{R}(\langle A \rangle)$-valued processes. By the natural identification $\mathbb{R}(\langle A \rangle) \approx \mathbb{R}^\infty$, we can induce the direct product topology into $\mathbb{R}(\langle A \rangle)$. $\mathbb{R}(\langle A \rangle)$ becomes a Polish space by the topology and we can consider its Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R}(\langle A \rangle))$, $\mathbb{R}(\langle A \rangle)$-valued random variables, their expectations, and other notions as usual. $\mathbb{R}(\langle A \rangle)$-valued continuous semi-martingale $\mathcal{X} : \mathbb{R}_{\geq 0} \times \Omega \rightarrow \mathbb{R}(\langle A \rangle)$ which satisfies the following condition that there exist continuous semi-martingale $\mathcal{X}_w : \mathbb{R}_{\geq 0} \times \Omega \rightarrow \mathbb{R}$ and satisfy $\mathcal{X}(t) = \sum_{w \in A^*} \mathcal{X}_w(t)w$. For $\mathbb{R}(\langle A \rangle)$-valued continuous semi-martingales $\mathcal{X}$ and $\mathcal{Y}$, we define $\mathbb{R}(\langle A \rangle)$-valued continuous semi-martingale $\int_0^t \mathcal{X}(s) \circ d\mathcal{Y}(s)$ and $\int_0^t \circ d\mathcal{X}(s)\mathcal{Y}(s)$ as follows:

$$\int_0^t \mathcal{X}(s) \circ d\mathcal{Y}(s) = \sum_{u,w \in A^*} \left( \int_0^t \mathcal{X}_u(s) \circ d\mathcal{Y}_w(s) \right) uw,$$

$$\int_0^t \circ d\mathcal{X}(s)\mathcal{Y}(s) = \sum_{u,w \in A^*} \left( \int_0^t \mathcal{Y}_w(s) \circ d\mathcal{X}_u(s) \right) uw,$$

where

$$\mathcal{X}(t) = \sum_{u \in A^*} \mathcal{X}_u(t)u, \text{ and } \mathcal{Y}(t) = \sum_{w \in A^*} \mathcal{Y}_w(t)w.$$

The Ito formula

$$(7) \quad \mathcal{X}(t)\mathcal{Y}(t) - \mathcal{X}(0)\mathcal{Y}(0) = \int_0^t \mathcal{X}(s) \circ d\mathcal{Y}(s) + \int_0^t \circ d\mathcal{X}(s)\mathcal{Y}(s)$$

holds.

We consider the stochastic differential equation

$$(8) \quad \mathcal{X}(t) = 1 + \sum_{i=0}^d \int_0^t \mathcal{X}(s) v_i \circ dB^i(s),$$

in $\mathbb{R}(\langle A \rangle)$, which corresponds to the stochastic differential equation (1). From (7), it is easy to see that (8) has a unique solution

$$(9) \quad \mathcal{X}(t) = \sum_{w \in A^*} B(t;w)w.$$

We have the following theorem.

**Theorem 1.1** ([5][10]). $\log(\mathcal{X}(t)) \in \mathcal{L}_{\mathbb{R}(\langle A \rangle)}$.

This theorem is proved in [5] and [10] by using Friedrich's theorem.

1.3.4. The new approximation scheme. We construct a family of linear operators in $C_b(\mathbb{R}^N)$ that approximate $\{P_t\}_{t \in \mathbb{R}_{\geq 0}}$ defined in (3). We define $\Phi$ to be a homomorphism between $\mathbb{R}(A)$ and the $\mathbb{R}$-algebra which consists of smooth differential operators over $\mathbb{R}^N$ defined by:

$$\Phi(1) = \mathrm{Id},$$

$$(10) \quad \Phi(v_{i_1} \cdots v_{i_n}) = V_{i_1} \cdots V_{i_n}, \; i_1, \ldots, i_n \in \{0, 1, \ldots, d\}.$$
A NEW SIMULATION METHOD OF DIFFUSION PROCESSES APPLIED TO FINANCE

In this paper, we assume that the stochastic differential equation (1) satisfies the following condition (UFG):

(UFG): There exist an integer $l$, and $\varphi_{u,u'} \in C_0^\infty(\mathbb{R}^N)$ which satisfy

$$\Phi(l(u)) = \sum_{u' \in A_{\leq l}^s \setminus \{1,v_0\}} \varphi_{u,u'} \Phi(l(u')),$$

for any $u \in A^* \setminus \{1,v_0\}$.

**Definition 1.1.** A family $\{Q(s)\}_{0 \leq s < 1}$ of linear operators in $C_b(\mathbb{R}^N)$ is called $m$-similar ($m \in \mathbb{Z}_{\geq 1}$), if there are constants $C \in \mathbb{R}_{>0}$ and $M \in \mathbb{R}_{\geq m+1}$ which satisfy

$$||P_s f - Q(s) f||_\infty \leq C \left( \sum_{k=m+1}^M s^{k/2} ||f||_{V,k} + s^{(m+1)/2} \|\text{grad}(f)\|_\infty \right),$$

$$||Q(s)f - P_s f||_\infty \leq C s^{1/2} \|\text{grad}(f)\|_\infty,$$

and

$$||Q(s)f||_\infty \leq \exp(Cs)||f||_\infty$$

for any $s \in (0,1]$ and $f \in C_b^\infty(\mathbb{R}^N;\mathbb{R})$, where

$$||f||_{V,n} = \sum_{k=1}^n \sum_{u_1,..,u_k \in A^* \setminus \{1,v_0\}} ||\Phi(l(u_1)\cdots(l(u_k)))f||_\infty.$$ 

**Definition 1.2.** An $L_\mathbb{R}(\mathbb{A})$-valued random variable $\xi$ is called $m$-$\mathbb{L}$-moment similar ($m \in \mathbb{Z}_{\geq 2}$), if it satisfies the following conditions:

(i) $E[||j_n \xi||_2^p] < +\infty$ for any $n \in \mathbb{Z}_{\geq 1}$ and $p \in \mathbb{R}_{>1}$,

(ii) $(\xi,v_0) = 1$ with probability 1,

(iii) $E[j_m(\exp(\xi))] = E[j_m(\mathcal{X}(1))]$,

where $\mathcal{X}$ is the $\mathbb{R}(\mathbb{A})$-valued random variable defined by (9).

**Example 1.1.** $\xi = \log(\mathcal{X}(1))$ is a trivial example of $m$-$\mathbb{L}$-similar random variable.

**Example 1.2.** In the case of $d = 1$, an $\mathbb{R}(\mathbb{A})$-valued random variable $\xi$ defined as:

$$(\xi,v_0) = 1, \quad (\xi,v_1) = \eta, \quad (\xi,v_1v_1) = \frac{\eta^2}{2},$$

$$(\xi,v_1v_0) = (\xi,v_0v_1) = \frac{\eta}{2}, \quad (\xi,v_1v_1v_1) = \frac{\eta^3}{6},$$

$$(\xi,v_1v_1v_0) = (\xi,v_0v_1v_1) = \frac{1}{4}, \quad (\xi,v_0v_0) = \frac{1}{2}, \quad (\xi,v_1v_1v_1v_1) = \frac{1}{8},$$

$$(\xi,v) = 0 \quad \text{otherwise},$$

where $\eta$ is a random variable which satisfies $P(\eta = 0) = 2/3$, $P(\eta = \pm\sqrt{3}) = 1/6$ becomes a 5-$\mathbb{L}$-similar random variable.
Example 1.3. In the case of $d = 2$, an $\mathbb{R} \langle \langle A \rangle \rangle$-valued random variable $\xi$ defined as:

$$
(\xi, v_0) = 1, \quad (\xi, v_i) = \eta_i,
$$

$$
(\xi, v_1 v_2) = \frac{1}{2} (\eta_1 \eta_2 + \sigma), \quad (\xi, v_2 v_1) = \frac{1}{2} (\eta_1 \eta_2 - \sigma), \quad (\xi, v_i v_i) = \frac{\eta_i^2}{2},
$$

$$
(\xi, v_1 v_0) = (\xi, v_0 v_0) = \frac{\eta_1}{2}, \quad (\xi, v_i v_i v_i) = \frac{\eta_i^3}{6},
$$

$$
(\xi, v_1 v_1 v_2) = (\xi, v_2 v_1 v_1) = \frac{\eta_2}{4}, \quad (\xi, v_2 v_2 v_1) = (\xi, v_1 v_2 v_2) = \frac{\eta_1}{4}, \quad (\xi, v_0 v_0) = \frac{1}{2},
$$

$$
(\xi, v_1 v_1 v_2 v_2) = (\xi, v_2 v_2 v_1 v_1) = \frac{1}{8}, \quad (\xi, v_0 v_0 v_i) = (\xi, v_i v_0 v_i) = \frac{1}{4}, \quad i \in \{1, 2\},
$$

$$(\xi, v, 0) = 0 \quad \text{otherwise},$$

where $\sigma, \eta_1$, and $\eta_2$ are independent random variables which satisfy $P(\sigma = \pm 1) = 1/2, P(\eta_i = 0) = 2/3, \text{and } P(\eta_i = \pm \sqrt{3}) = 1/6 \text{ (} i = 1, 2\text{)}$ becomes a 5-$\mathcal{L}$-similar random variable.

For $s \in \mathbb{R}_{>0}$, we define $\Psi_s : \mathbb{R} \langle \langle A \rangle \rangle \rightarrow \mathbb{R} \langle \langle A \rangle \rangle$ as follows:

$$
\Psi_s \left( \sum_{m=0}^{\infty} P_m \right) = \sum_{m=0}^{\infty} s^{m/2} P_m, \text{ where } P_m \in \mathbb{R} \langle A \rangle_m.
$$

Following two theorems give the new simulation scheme.

Theorem 1.2. Let $m \in \mathbb{Z}_{\geq 1}$ and $\xi$ be an $m$-$\mathcal{L}$-similar random variable. We also let $Y : (0, 1] \times \mathbb{R}^N \times \Omega \rightarrow \mathbb{R}^N$ be a measurable map that satisfies the following conditions:

(i) $Y(s, \cdot, \omega) : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is continuous for any $s \in (0, 1]$ and $\omega \in \Omega$,

(ii) $\sup_{0<s\leq 1, x \in \mathbb{R}^N} s^{(m+1)/2} E[|Y(s, x)|] < \infty$.

Then, the family of linear operators $\{Q(s)\}_{0 \leq s < 1}$ defined by

$$
(Q(s)f)(x) = E[f(\exp(\Phi(j_m(\Psi_s(\xi))))(x) + Y(s, x))],
$$

where $f \in C_b(\mathbb{R}^N)$, becomes $m$-similar.

This theorem shows how to construct $m$-similar operators and the following theorem shows the new approximation scheme.

Theorem 1.3. We assume (1) satisfies the condition (UFG). Let $m \geq 1$ and $\{Q(s)\}_{0 \leq s < 1}$ be an $m$-similar family of linear operators as defined in Definition 1.1.

Let $t_k^{(n)} = k^n n^{-\gamma} T$ where $n \in \mathbb{N}, \gamma \in \mathbb{R}_{>0}$ and $k = 0, 1, \ldots, n$, and let $s_k = t_k^{(n)} - t_{k-1}^{(n)}$. For $f \in C_b^\infty(\mathbb{R}^N)$, we define

$$
epsilon(f) = \|P_{r}f - Q(s_n)Q(s_{n-1}) \cdots Q(s_1)f\|_{\infty}.
$$

Then we have the following statements. If $0 < \gamma < m - 1$, there exists a constant $C > 0$ which satisfies

$$
epsilon(f) \leq C n^{-\gamma/2} \|\operatorname{grad}(f)\|_{\infty} \quad \text{for all } f \in C_b^\infty(\mathbb{R}^N) \text{ and } n \in \mathbb{N}.
$$

If $\gamma = m - 1$, there exists a constant $C > 0$ which satisfies

$$
epsilon(f) \leq C n^{-(m-1)/2} \log(n + 1) \|\operatorname{grad}(f)\|_{\infty} \quad \text{for all } f \in C_b^\infty(\mathbb{R}^N) \text{ and } n \in \mathbb{N}.
223

A NEW SIMULATION METHOD OF DIFFUSION PROCESSES APPLIED TO FINANCE

If \( m - 1 < \gamma \), there exists a constant \( C > 0 \) which satisfies
\[
\epsilon(f) \leq Cn^{-(m-1)/2} \| \text{grad}(f) \|_\infty \quad \text{for all } f \in C_0^\infty(\mathbb{R}^N) \text{ and } n \in \mathbb{N}.
\]

2. Numerical examples: Application to Finance

Taking Theorem 1.3 and (4) into account, we can reasonably expect that the operator \( Q(s_n)Q(s_{n-1}) \cdots Q(s_1) \) realizes more accurate approximation than the Euler-Maruyama approximation does. We confirm this expectation in the following. We give some numerical examples and compare the new approximation scheme with Euler-Maruyama scheme. In the following, we call the new approximation scheme based on the results in the previous section K-method.

2.1. Comparison to Euler-Maruyama approximation (1): 1 dimensional case. We calculate the price of an Asian call option with maturity \( T \) and strike price \( K \) written on a stock. The payoff at \( T \) of this option is
\[
\max \left( 0, \frac{1}{T} \int_0^T X_1(t, x) dt - K \right)
\]
where \( X_1(t, x) \) denotes the price process of the underlying stock. We assume that \( X_1(t, x) \) follows the Black-Scholes equation \( dX_1(t, x) = X_1(t, x)(r' dt + \sigma dB(t)) \)
where \( r' \) and \( \sigma \) are constants and \( B(t) \) is the 1-dimensional standard Brownian motion with respect to the equivalent martingale measure. In this example, \( r' = 0.05, \sigma = 0.3, T = 0.25 \) and \( K = 1.05 \). Let \( X_2(t, x) = \int_0^t X_1(s, x) ds \) and \( X(t, x) = (X_1(t, x), X_2(t, x)) \). The price of this Asian option is \( e^{-r'T} E[f(X(T, x))] \). We use the 5-L-similar random variable defined in Example 1.2. Figure 1 shows the relation between the number of partition and the approximation error that comes from the discretization. We compare following four methods:

(i.) Euler-Maruyama approximation.
(ii.) Euler-Maruyama approximation with Romberg extrapolation [11][3].
(iii.) K-method.
(iv.) K-method with Romberg extrapolation.

Here we set \( \gamma = 2 \). Theorem 1.3 suggests that greater \( \gamma \) is better as far as convergence order is concerned. In reality, the constant term is very important and in this case, \( \gamma = 2 \) realizes the best results. In the case of Euler-Maruyama approximation, we calculate the expectation by using Monte Carlo method. In the case of K-method, we proceed exhaustive calculation. Overall results from the point of CPU time required for 4-digits accuracy are shown in Figure 2.1 and Figure 2.1.

2.2. Partial sampling problem. As we see in Examples 1.2 and 1.3, we construct \( \xi \), the \( m \)-L-similar random variable included in the new approximation scheme to be discrete. In the cases of Example 1.2 and Example 1.3, the numbers of states of \( \xi \) are 3 and \( 2 \times 3 \times 3 = 18 \), respectively. Then, if we discretize the time by partitioning the simulation period \([0, T]\) into \( n \) parts, the number of points to be calculated becomes \( 3^n \) in the case of Example 1.2 and \( 18^n \) in the case of Example 1.3. In the latter case, \( n = 7 \) is enough for us to giving up the exhaustive calculation. Even in the former case, it is not possible to calculate all points in practice when \( n \) is 15 or 16. As a result, there are many cases in which we cannot do without partial sampling from finite set of points.
Discretization Error: Asian option $(S=1.0, T=1.0, K=1.05, \Delta t=0.05, v=0.3)$

\[ \text{Euler-Maruyama Approx.} \quad \text{Euler-Maruyama + Extrapolation} \quad \text{K-method} \quad \text{K-method + Extrapolation} \]

\[ r^{-0.05}, v=0.3 \]

Figure 1. Approximation Error and number of partition. Asian option, BS-model

<table>
<thead>
<tr>
<th>Method</th>
<th>#Partition</th>
<th>#Sample</th>
<th>CPU time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-M</td>
<td>2000</td>
<td>$3.6 \times 10^7$</td>
<td>$20563.2$</td>
</tr>
<tr>
<td>E-M + Extrptln</td>
<td>32 + 64</td>
<td>$3.24 \times 10^8$</td>
<td>$856.6$</td>
</tr>
<tr>
<td>K-method</td>
<td>20</td>
<td>$5.76 \times 10^6$</td>
<td>$573.69$</td>
</tr>
<tr>
<td>K-method + Extrptln</td>
<td>4 + 8</td>
<td>6642</td>
<td>$0.0170$</td>
</tr>
</tbody>
</table>

Figure 2. #Partition, #Sample, and CPU time required for 4-digits accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>#Partition</th>
<th>#Sample</th>
<th>CPU time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-M</td>
<td></td>
<td>0.0279</td>
<td></td>
</tr>
<tr>
<td>E-M + Extrptln</td>
<td>0.67</td>
<td>1.0</td>
<td>$3.37 \times 10^4$</td>
</tr>
</tbody>
</table>

Figure 3. Performance ratio (K-method = 1, greater = faster)

2.2.1. Tree based branching algorithm (TBBA). The tree based branching algorithm (TBBA) is introduced by Crisan and Lyons in [2] and it is shown in the paper that the algorithm realized the best sampling in some sense. In [7], it is reported that when we apply this algorithm to the K-method, we can reduce the number of sampling points by using the algorithm called tree based branching algorithm (TBBA). Here we give a short explanation of TBBA. See [2] and [7] for details.
Let $I$ be a countable set and $\mathcal{M}^p(I)$ be the set of probability measures on $I$ with the product topology. $B(\mathcal{M}^p(I))$ is the Borel $\sigma$-algebra on $\mathcal{M}^p(I)$ generated by the cylinder sets as usual. We prepare another probability space $(\Omega, \mathcal{F}, P)$ with no atoms. Let $a \in \mathcal{M}^p(I)$ be a probability with finite entropy $H(a) = \sum_{i \in I} a(i) \log a(i) < \infty$ and $A^a_n$, $n \in \mathbb{N}$ be a set of $\mathcal{M}^p(I)$-valued random variables of $(\Omega, \mathcal{F}, P)$ defined as follows:

$$A^a_n = \left\{ \tilde{a} : \Omega \to \mathcal{M}^p(I) \mid \exists V_k : \Omega \to I, \mathcal{F} \text{-measurable s.t.} \right\},$$

where $\delta_x$ denotes the Dirac distribution concentrated at $x$. We can regard $A^a_n$ to be a set of unbiased estimators that takes $n$ samples, or a set of measures realized by $n$ samples. What we want is to find a $\tilde{a} \in A^a_n$ which has small discrepancy with respect to $a$. For measuring the discrepancy between an element $\tilde{a}$ of $A^a_n$ and $a$, we introduce $H(\tilde{a}|a)$, the mean relative entropy of $\tilde{a}$ with respect to $a$, as follows:

$$H(\tilde{a}|a) = E^P_{\Omega} \left[ \sum_{i \in I} \tilde{a}(i) \log \left( \frac{\tilde{a}(i)}{a(i)} \right) \right] = \sum_{i \in I} E^P_{\Omega} [\tilde{a}(i) \log(\tilde{a}(i))] - \sum_{i \in I} a(i) \log(a(i)).$$

If, for any $i \in I$, $\tilde{a} \in A^a_n$ satisfies the following condition:

$$\tilde{a}(i) = \begin{cases} \lfloor n a(i) \rfloor/n & \text{with probability } 1 - \text{fp}(na(i)), \\ \lfloor n a(i) \rfloor + 1)/n & \text{with probability } \text{fp}(na(i)), \end{cases}$$

then we say that $\tilde{a}$ has property $MV(n)$.

**Theorem 2.1** ([2]). Let $\tilde{a} \in A^a_n$, then $H(\tilde{a}|a)$ is minimal if and only if $\tilde{a}$ has property $MV(n)$.

According to this theorem, our object is deduced to construction of a $\tilde{a} \in A^a_n$ which has the property $MV(n)$. In [2], Crisan and Lyons give an algorithm by which we can construct a random measure in $A^a_n$ satisfying the property $MV(n)$. We sketch out this algorithm. In the first step, we construct a filtration $\{\mathcal{F}_k\}_{k \geq 0}$ of $I$ which satisfy $\{0, I\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_k \subset \mathcal{F}_{k+1} \subset \cdots \subset 2^I$ and $\sigma(\cap_k \mathcal{F}_k) = 2^I$. Because $I$ is discrete, we can consider a partitioning $\mathcal{P}_k^\mathcal{F}$ of $I$ which corresponds to $\mathcal{F}_k$. For $i \in I$, $\mathcal{P}_k^\mathcal{F}(i)$ denotes the element of $\mathcal{P}_k^\mathcal{F}$ which includes $i$. In the second step, we construct a sequence $\{\tilde{a}^k\}_{k \geq 0}$ of elements of $A^a_n$ inductively from $\tilde{a}^0$ in the way that the equality

$$\tilde{a}^k(i) = \sum_{j \in \mathcal{P}_k^\mathcal{F}(i)} \tilde{a}(j)$$

holds for all $i \in I$ and $k \geq 0$. We compare naive Monte Carlo sampling with TBBA under K-method. The problem is the same as before, that is, pricing of an Asian option. We use the natural filtration

$$\mathcal{F}_t = \sigma(Q(s) ; s \leq t)$$

as the filtration in the first step of TBBA. Figure 2.2.1 shows the square root of the variance of 10 Monte Carlo trials and 10 TBBA trials around the true value obtained by exhaustive calculation. From this figure we can see that TBBA achieves much faster calculation than the naive Monte Carlo sampling does. If we want to make the standard deviation smaller than $10^{-5}$, TBBA needs about $2 \times 10^5$ samples and the naive Monte Carlo about $10^7$ samples, therefore TBBA is about 50 times faster in this case. If we want more accuracy, this ratio becomes larger. Figure 2.2.1 shows
Figure 4. Convergence speed of MC and TBBA with K-method: n=12

Figure 5. Approximation Error and number of partition. Levy area
another example. This is a example of \( d = 2 \) case. In this example, we calculate
\[
E[\max(S(1), 0)],
\]
where \( S(t) \) is the Levy area defined as follows:
\[
S(t) = \frac{1}{2} \left( \int_0^t B^1(s) \circ dB^2(s) - \int_0^t B^2(s) \circ dB^1(s) \right).
\]
We can see how \( \gamma \) effects in this case.

2.2.2. Exotic filter with TBBA. In the previous example, we see that TBBA realize
very effective partial sampling. The property \( MV(n) \) does not give enough explanation. If \( a \) is uniform over \( I \), TBBA coincides with the random sampling without
replacement, which behaves almost the same as the naive Monte Carlo provided
that the number of samples is not so large as \( \text{card}(I) \). In the previous example,
the discrete set \( I \) is constructed by connecting identical trinomial tree unit and we
can regard this set to be a set with almost uniform probability measure. Therefore,
\( MV(n) \) property of \( \tilde{a} \) dose not explain why TBBA show such superior performance
to the naive Monte Carlo. The property \( MV(n) \) relates only \( \tilde{a} \) and the filtration
\( \{\mathcal{F}_k\}_{k \geq 0} \) used in the first step of TBBA is perfectly ignored. Then, we may
consider that \( \{\mathcal{F}_k\}_{k \geq 0} \) is the essential factor. In [8], it is reported that by changing the
filtration used in the first step of TBBA, faster calculations are realized in some
cases. It seems this report supports our hypothesis. In the following, we introduce
the exotic filter method briefly and see some numerical examples. For details of the
algorithm, see [8]. We consider an exotic filter \( \{G_k\}_{k \geq 0} \) of \( I \) defined by:
\[
(15)
\begin{align*}
G_0 &= \{\emptyset, I\}, \\
G_k &= \sigma \{B_{(1-r(n))}, G_{k-1}\},
\end{align*}
\]
where \( r(n) \) is the \( n \)-th element of the van der Corput sequence and \( B_s \) is the
standard Brownian motion which drives the SDE (1). The van der Corput sequence
is a sequence of real numbers defined as:
\[
r(n) = \sum_{l=0}^{\infty} \alpha(n, l)2^{-l-1},
\]
\[
n = \sum_{l=0}^{\infty} \alpha(n, l)2^l, \quad \alpha(n, l) \in \{0, 1\} \quad \text{for all } n, l \in \mathbb{Z}_{\geq 0}.
\]

Figure 2.2.2, shows the result of calculation of the same Asian option pricing
example \( (d = 1) \) as we see before. Here we compare the following three methods:
K-method with natural filter TBBA, K-method with exotic filter TBBA, and K-
method with Monte Carlo.

REFERENCES
1. Vlad Bally and Denis Talay, The law of the Euler scheme for stochastic differential equations
 I. Convergence rate of the distribution function, Probability theory and related fields 104
 (1996), 43-60.
2. Dan Crisan and Terry Lyons, Minimal Entropy Approximations and Optimal Algorithms for
 the Filtering Problem, Monte Carlo Methods and Applications (to appear).
3. Peter E. Kloeden and Eckhard Platen, Numerical Solution of Stochastic Differential Equations,
4. Shigeo Kusuoka, Approximation of Expectation of Diffusion Process and Mathematical Fi-
Black Scholes model (Tree Depth = 16, $2^{16} = 65536$)

Figure 6. K-method with 3 sampling method: Convergence speed.

5. ———, Approximation of Expectation of Diffusion Processes based on Lie Algebra and Malliavin Calculus, University of Tokyo Graduate School of Mathematical Sciences Preprint series (2003).


7. ———, A partial sampling method applied to the Kusuoka approximation, Monte Carlo Methods and Applications 9 (2003), 27–38.


Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan
E-mail address: kusuoka@ms.u-tokyo.ac.jp

Center for Research in Advanced Financial Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552 Japan
E-mail address: ninomiya@craft.titech.ac.jp