Combinatorics for calculating expectations of stochastic differential equations

Jun Ohkubo\textsuperscript{1,2}
\textsuperscript{1} Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama, 338-8570, Japan
\textsuperscript{2} JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan
E-mail: johkubo@mail.saitama-u.ac.jp

Abstract. Combinatorial discussion is proposed and applied for calculating expectations of stochastic differential equations. Starting from the duality theory of stochastic processes, some modifications of interpretation and usages of time-ordering operators naturally lead to combinatorial discussions. As a demonstration, the first and second moments for the Ornstein-Uhlenbeck process are re-derived from the combinatorial discussion. Furthermore, two numerical methods for practical applications are proposed. One is based on a conventional exponential expansion and the Padé approximation. Another uses a resolvent of a time-evolution operator, and the Aitken series acceleration method is also employed. These two proposals recover the correct results approximately.

1. Introduction

Stochastic differential equations, or Langevin equations, are widely used in various research fields \cite{1}. Paths of stochastic differential equations are discussed based on the Ito calculus, and the corresponding Fokker-Planck equations directly deal with probability density functions. However, since analytical solutions are difficult to obtain in general, Monte Carlo simulations are employed practically. Using numerical methods such as the Euler-Maruyama approximations, probability density functions are estimated adequately \cite{2}.

In practical cases, it is common that only limited statistics, such as averages and variances, are needed. Recent studies of the duality relations in stochastic processes have been revealing that such statistical quantities in stochastic differential equations, especially moments, are also evaluated by the corresponding dual birth-death processes \cite{3,4}. In some cases, analytical solutions for the corresponding dual birth-death processes are available, and the usefulness of the duality relations was shown for population models \cite{3,7} and Brownian moment processes \cite{8,9}. Although there are many recent studies for the duality relations from mathematical viewpoints \cite{3,10,13}, our focus here
Combinatorics for calculating expectations of stochastic differential equations

is on numerical applications of the duality relations to obtain statistical quantities of analytically intractable stochastic differential equations. There are some applications with the aid of numerical calculations [14], but such numerical studies are still limited. One of the problems in the practical usage of the duality relation is the stochasticity in the dual process. For example, in the application of a filtering problem, large sample sizes were needed to construct practical filtering procedures [14].

In the present paper, a natural connection between the conventional duality relations in stochastic processes and combinatorial frameworks is given. The usage of combinatorics avoids the stochasticity in the dual processes, and it is useful for low-dimensional cases. Such combinatorial discussions are recently applied to compute the Mori-Zwanzig memory integral in generalized Langevin equations [15,16]. It was revealed that combinatorial algorithms are useful to evaluate operator exponentials; the recursive algorithm efficiently computes expansion coefficients. Some parts of these recent developments are available to construct our algorithm for evaluating statistical quantities in stochastic differential equations. In the present work, not only discussions on connections with the combinatorics but also two candidates of practical numerical techniques to evaluate statistical quantities are proposed. The first proposal is based on a conventional exponential expansion and the Padé approximation, and the second one uses a resolvent of a time-evolution operator and the Aitken series-acceleration method. Numerical comparisons with Monte Carlo simulations are also given.

This paper is organized as follows. In section 2, the brief review of previous works is given, and the first and second moments of the Ornstein-Uhlenbeck process are re-derived from the duality relations in stochastic processes. Section 3 focuses on the connection between the duality relations and combinatorics; the interpretation for the dual birth-death process is modified, and the time-ordering operator is used to make the connection. The combinatorial discussions are demonstrated using the Ornstein-Uhlenbeck process. In section 4, an analytically intractable example is discussed, and two numerical methods are proposed. The concluding remarks are given in section 5.

2. Brief review of previous works

2.1. Duality relation between stochastic differential equations and birth-death processes

Although the combinatorial discussions can be derived without using the duality relations, it is beneficial to clarify their connections. Hence, the duality in stochastic processes is briefly reviewed here.

As denoted in the Introduction, the duality relation in stochastic processes is widely used in various areas, including interacting particle systems such as simple exclusion processes [17,19]. In the present paper, we focus on the duality relation between stochastic differential equations and the birth-death process. Here, stochastic differential equations with only one variable are discussed; it is straightforward to apply the following discussions to multivariate cases.
Let $x_t \in \mathbb{R}$ be a state of a stochastic differential equation at time $t$. The corresponding dual birth-death process is a stochastic process with discrete-state and continuous-time, and its state at time $t$ is written as a state vector $n_t \in \mathbb{N}^{D_{dual}}$, where $D_{dual}$ is the number of variables in the dual process. Note that these two processes do not need to have the same dimensions, as we will see later. The process $(x_t)$ is said to be dual to $(n_t)$ with respect to a duality function $D : \mathbb{R} \times \mathbb{N}^{D_{dual}} \to \mathbb{R}$ if for all $(x_t)$, $(n_t)$ and $t \geq 0$ we have

$$
E_{n_t}[D(x_0, n_t)] = E_{x_t}[D(x_t, n_0)],
$$

where $E_{x_t}$ and $E_{n_t}$ are the expectations in the processes $(x_t)$ starting from $x_0$ and $(n_t)$ starting from $n_0$, respectively. This duality relation means that the solution of the dual birth-death process gives the expectations in the original stochastic differential equation.

The remaining problem is how to derive the dual birth-death process. In the following sections, the derivation is briefly reviewed and demonstrated by using the Ornstein-Uhlenbeck process.

### 2.2. Kolmogorov backward equation and dual process

The following stochastic differential equation is a starting point:

$$
\mathrm{d}x = \mu(x)\mathrm{d}t + \sigma(x)\mathrm{d}W(t),
$$

where $\mu(x)$ is the drift coefficient, $\sigma(x)$ is the diffusion coefficient, and $W(t)$ is the Wiener process. Although the coefficients can be time-dependent, only time-independent cases are considered in the present paper. The stochastic differential equation has the following corresponding partial differential equation, i.e., the Fokker-Planck equation [1]:

$$
\frac{\partial}{\partial t} p(x, t) = \mathcal{L} p(x, t),
$$

where $p(x, t)$ is the probability density function for $x$ at time $t$, and

$$
\mathcal{L} \equiv -\frac{\partial}{\partial x} \mu(x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \sigma(x)^2
$$

is the time-evolution operator.

Different from the Fokker-Planck equation, it is known that the Kolmogorov backward equation (or the backward Fokker-Planck equation) starts from a final condition, and it is integrated backward in time [1]. The famous Feynman-Kac formula extends the discussion of the Kolmogorov backward equation, and these are starting points of the derivation of the duality relations in stochastic processes. As discussed in [20], the derivation of the corresponding dual birth-death process is also understandable only by the use of the integration-by-parts and function expansions. When we consider $m$-th moment of $x$ in the stochastic differential equation, the following rewriting is possible formally:

$$
\mathbb{E}[x^m] = \int_{-\infty}^{\infty} x^m p(x, T)\mathrm{d}x
$$
Combinatorics for calculating expectations of stochastic differential equations

\[ = \int_{-\infty}^{\infty} x^m \left( e^{LT} \delta(x - x_0) \right) dx \]

\[ = \int_{-\infty}^{\infty} \left( e^{L^\dagger T} x^m \right) \delta(x - x_0) dx \]

\[ = \int_{-\infty}^{\infty} \tilde{\varphi}(x, T) \delta(x - x_0) dx \]

\[ = \tilde{\varphi}(x_0, T), \quad (5) \]

where \( \delta(x) \) is the Dirac’s delta function, \( x_0 \) is the initial position, and \( T \) is the final time at which we evaluate the expectation. Note that the function \( \tilde{\varphi} \) is not a probability function; the time-evolution operator for \( \tilde{\varphi} \) is

\[ L^\dagger \equiv \mu(x) \frac{\partial}{\partial x} + \frac{1}{2} \sigma(x)^2 \frac{\partial^2}{\partial x^2}, \quad (6) \]

which is the adjoint operator of \( L \), and it does not satisfy the probability conservation law in general.

The time-evolution equation for \( \tilde{\varphi} \),

\[ \frac{\partial}{\partial t} \tilde{\varphi}(x, t) = L^\dagger \tilde{\varphi}(x, t), \quad (7) \]

should be performed backwardly from \( t = T \) to \( t = 0 \). This backward evolution is sometimes confusing when \( L \) depends on time \( t \). To avoid the confusion, we here rewrite \( t \) as \( T - t \) and define \( \varphi(x, t) \equiv \tilde{\varphi}(x, T - t) \); the function \( \varphi(x, t) \) is integrated forward in time from \( t = 0 \) to \( t = T \). Note that the time-evolution has the following initial condition:

\[ \varphi(x, 0) = x^m. \quad (8) \]

The function \( \varphi(x, t) \) has still a continuous variable \( x \), and the dual birth-death process is derived by the use of function expansions. Various expansions are available; Hermite polynomials \[20\] and Legendre polynomials \[21\] are used in previous works. As for the details, see \[20\]. A demonstration is helpful for understanding the discussion, and hence the derivation of the dual birth-death process is given by using the Ornstein-Uhlenbeck process in the next section.

2.3. Example: Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck process is famous and solved analytically \[1\]:

\[ dx = -\gamma x dt + \sigma dW(t), \quad (9) \]

where \( \gamma > 0 \) and \( \sigma > 0 \). The adjoint time-evolution operator \( L^\dagger \) is given as

\[ L^\dagger = -\gamma x \frac{\partial}{\partial x} + \sigma^2 \frac{\partial^2}{\partial x^2}, \quad (10) \]

Here, for later use, we introduce the coordinate transformation \( x' = x - x_c \); rewriting \( x' \) as \( x \) again, the adjoint operator is rewritten as

\[ L^\dagger = -\gamma (x + x_c) \frac{\partial}{\partial x} + \sigma^2 \frac{\partial^2}{\partial x^2} \]

\[ = -\gamma x \frac{\partial}{\partial x} - \gamma x_c \frac{\partial}{\partial x} + \sigma^2 \frac{\partial^2}{\partial x^2}. \quad (11) \]
This coordinate transformation is introduced to calculate the Taylor expansion around \( x_c \). When \( x_c \) is the initial condition of the original stochastic differential equation, it is enough to consider only \( \varphi(0, T) \) for the evaluation of expectation values.

The following expectation values are obtained by the use of the well-known analytical solution for the Ornstein-Uhlenbeck process \([1]\):

\[
\mathbb{E}[x_T - x_c] = x_c e^{-\gamma T} - x_c, \tag{12}
\]

\[
\mathbb{E}[(x_T - x_c)^2] = \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma T}) + x_c^2 e^{-2\gamma T} - 2x_c^2 e^{-\gamma T} + x_c^2. \tag{13}
\]

Next, we will see that the dual birth-death process recovers these two expressions. In order to recover stochasticity to the dual process, the adjoint operator \( \mathcal{L}^\dagger \) for the time-evolution is split into two parts:

\[
\mathcal{L}^\dagger = \tilde{\mathcal{L}}^\dagger + V, \tag{14}
\]

where

\[
\tilde{\mathcal{L}}^\dagger = \gamma x_c x(0) \frac{\partial}{\partial x} - \gamma x_c x \frac{\partial}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} - \frac{\sigma^2}{2} x^2 \frac{\partial^2}{\partial x^2}, \tag{15}
\]

\[
V = -\gamma x \frac{\partial}{\partial x} + \gamma x_c x \frac{\partial}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2}{\partial x^2}, \tag{16}
\]

and \( x(0) \equiv -1 \) is introduced to avoid negative transition problem \([22]\). As discussed in \([22]\), the operator \( \tilde{\mathcal{L}}^\dagger \) gives the time-evolution for dual birth-death process, and \( V \) corresponds to the Feynman-Kac term. Using the following function expansion,

\[
\varphi(x, t) = \sum_{n=0}^{\infty} P(n, n_0, t)(x(0))^{n_0} x^n, \tag{17}
\]

the time-evolution equation

\[
\frac{\partial}{\partial t} \varphi(x, t) = \tilde{\mathcal{L}}^\dagger \varphi(x, t) \tag{18}
\]

gives the following equation for the coefficients \( P(n, n_0, t) \):

\[
\frac{d}{dt} P(n, n_0, t) = \gamma x_c (n + 1) P(n + 1, n_0 - 1, t) - \gamma x_c n P(n, n_0, t) \\
+ \frac{\sigma^2}{2} (n + 2)(n + 1) P(n + 2, n_0, t) - \frac{\sigma^2}{2} n(n - 1) P(n, n_0, t). \tag{19}
\]

This master equation for \( P(n, n_0, t) \) is interpreted as the following chemical reaction system:

\[
\text{Event 1: } \ X \rightarrow X(0) \quad \text{at rate } \gamma x_c n, \\
\text{Event 2: } \ X + X \rightarrow \emptyset \quad \text{at rate } \sigma^2 n(n - 1)/2. \tag{20}
\]

Note that the initial condition of \( P(n, n_0, t) \) is \( P(n = m, n_0 = 0, t = 0) = 1 \), and otherwise zero in order to evaluate \( \mathbb{E}[x^m] \). The Feynman-Kac term in \([16]\) should be also considered; it can be interpreted in terms of the variables of the dual birth-death process as

\[
V(n) = -\gamma n + \gamma x_c n + \frac{\sigma^2}{2} n(n - 1). \tag{21}
\]
Combinatorics for calculating expectations of stochastic differential equations

Note that this term does not depend on \( n_0 \).

Using the above dual birth-death process, it is straightforward to evaluate statistics in the original stochastic differential equation. When the initial condition is set as \( n = 1 \), only Event 1 in \([20]\) is allowed, and hence we have

\[
\int_0^T e^{V(n=1)t_1} \left[ e^{-\gamma \Delta t_1} e^{x_c \Delta t_1} \right] dt_1 = - \left( x_c e^{-\gamma T} - x_c \right),
\]

where \([ \cdot ]\) corresponds to the probability density for the path with only one Event 1 \([23]\). Note that Event 1 also gives the change \( n_0 = 0 \rightarrow 1 \). Using the fact of \( n_0 = 1 \) at the final time \( T \), the negative sign in \((22)\) is cancelled, and the expectation value in \((12)\) is adequately recovered. As for \( n = 2 \) at \( t = 0 \), we have two possible paths that

(i) Event 2 occurs at once,
(ii) Event 1 occurs twice.

Path (i) gives

\[
\int_0^T e^{V(n=2)t_1} \left[ e^{-2\gamma \Delta t_1} \gamma x_c \right] dt_1 = \frac{\sigma^2}{2\gamma} \left( 1 - e^{-2\gamma T} \right),
\]

and path (ii) provides the following contribution

\[
\int_0^T dt_2 \int_0^{t_2} dt_1 e^{V(n=2)t_1} \left[ e^{-2\gamma \Delta t_1} \gamma x_c \right] e^{V(n=1)(t_2-t_1)} \left[ e^{x_c \Delta t_2} \right] = x_c e^{-2\gamma T} - 2x_c^2 e^{-\gamma T} + x_c^2.
\]

Noting \( n_0 = 0 \) for path (i) and \( n_0 = 2 \) for path (ii), the 2nd moment in \((13)\) is recovered.

3. Combinatorics

This section gives the first main contribution of the present paper: Starting from the dual birth-death process, combinatorial discussions are derived. The discussion is applied to the Ornstein-Uhlenbeck process, and we will confirm that the analytical solutions are recovered adequately.

3.1. From dual birth-death process to simple combinatorics

As discussed in \([22]\), the time-evolution with \( \tilde{\mathcal{L}}^t \) acts on the state vector \( |n\rangle \) as follows:

\[
e^{\tilde{\mathcal{L}}^t|n\rangle} \simeq \left( 1 + \tilde{\mathcal{L}}^t \right) |n\rangle
\]

\[
= 1 - \sum_{r=1}^{R} a_r(n) \Delta t |n\rangle \langle n| + \sum_{r=1}^{R} a_r(n) \Delta t |n + v_r\rangle \langle n|
\]

\[
\simeq e^{-a_0(n)\Delta t} |n\rangle \langle n| + \left( \sum_{r=1}^{R} a_r(n) \Delta t |n + v_r\rangle \langle n| \right),
\]

where \( R \) is the number of events, \( a_r(n) \) is the propensity function for \( r \)-th event \([24]\), and \( a_0(n) = \sum_{r=1}^{R} a_r(n) \); the vector \( v_r \) represents stoichiometric coefficients of event
does not commute each other; for example,

Combinatorics for calculating expectations of stochastic differential equations

Although the state vector $|n\rangle$ could be an abstract one satisfying $\langle m|n\rangle = \delta_{m,n}$, it is possible to define the vector as an infinite-dimensional one or the explicit one in terms of $x$ [26, 27].

The Taylor expansion in [25] gives a natural probabilistic interpretation for the birth-death process. The first and second terms in the 3rd line in (25) is viewed as the Bernoulli trial: With probability $e^{-a_0(n)\Delta t}$ no event occurs; with probability $a_0(n) = \sum_{r=1}^{R} a_r(n)$ some event occurs. For the latter case, only an event is chosen with probability $a_r(n)/a_0(n)$; the factor $1/a_0(n)$ is compensated by considering $a_0(n) \exp(-a_0(n)\Delta t)$, which gives the exponential distribution for the event-interval time.

To make the combinatorial discussion, the interpretation of the time-evolution operator should be modified a little. That is, instead of the time-evolution operator $\exp(L_t)$, the adjoint operator $L^\dagger$ in (6) is directly considered here. Hence, we have

$$e^{L^\dagger \Delta t}|n\rangle \simeq 1 + \sum_{r=1}^{R} a_j(n)\Delta t|n + v_r\rangle \langle n|.$$ (26)

Repeated actions of $\exp(L^\dagger \Delta t)$ mean simple products of $a_j$’s as follows:

$$e^{L^\dagger T}|n_{ini}\rangle \simeq a_j^{(M)}(n_{M-1})a_j^{(M-1)}(n_{M-2}) \cdots a_j^{(1)}(n_{ini})(\Delta t)^M |n_{ini} + \sum_{m=1}^{M} v_j^{(m)}\rangle,$$ (27)

where $M$ is the number of times of selecting events in the 2nd term in (26), and $a_j^{(m)}(n)$ is the propensity function for event $j^{(m)}$. Here, the initial state is written as $n_{ini} \equiv n_1$.

Note that (27) gives a contribution of a path; from the viewpoint of the path integral, we need time-integration for all possible paths. Note that the factors $\{a_j^{(m)}\}$ does not commute each other; for example,

$$a_j^{(M)}(n_{M-1})a_j^{(M-1)}(n_{M-2}) \neq a_j^{(M)}(n_{M-2})a_j^{(M-1)}(n_{M-1})$$ (28)

because the factors depend on the state. Hence, we introduce the following time-ordering operator [28]

$$\mathcal{T} \{a_j^{(2)}(n_2)a_j^{(1)}(n_1)\} \equiv a_j^{(2)}(n_2)a_j^{(1)}(n_1)\theta(t_2 - t_1) + a_j^{(2)}(n_1)a_j^{(1)}(n_2)\theta(t_1 - t_2),$$ (29)

where $\theta(t)$ is the Heaviside function. Using the abbreviation $a^{(m)} \equiv a_j^{(m)}(n_{m-1})$, the probability, with which events occur $M$ times, is written as follows:

$$\int_0^T dt_M \int_0^{t_M} dt_{M-1} \cdots \int_0^{t_2} dt_1 a^{(M)}a^{(M-1)} \cdots a^{(1)}$$

$$= \frac{1}{M!} \int_0^T dt_M \int_0^{T} dt_{M-1} \cdots \int_0^{T} dt_1 \mathcal{T} \{a^{(M)}a^{(M-1)} \cdots a^{(1)}\}$$

$$= \frac{T^M}{M!} a^{(M)}a^{(M-1)} \cdots a^{(1)}.$$ (30)
Combinatorics for calculating expectations of stochastic differential equations

8

Figure 1. Possible combinations of paths starting from \( n = 1 \). Once we reach the state with \( n = 0 \), occurrences of events finish.

The final expression in (30) gives the basis of the combinatorial discussion for calculating expectations of the original stochastic differential equation. Recall that each event changes the state \( n \) in the dual (non-stochastic) process. Additionally, different from the dual birth-death process in section 2, the term giving the Feynman-Kac contribution is also considered as an event in the present discussion. Note that the usage of the event with no state change is different from that of the Extrande algorithm \[29\] or the uniformization techniques \[30\], which are proposed for time-inhomogeneous birth-death processes. Here, the propensity functions do not depend on time.

The general discussion would be understood clearly if we see the exactly solvable cases of the Ornstein-Uhlenbeck process. The time-evolution operator \( \mathcal{L}^\dagger \) for the dual process is given in (11), and it acts on \( |n\rangle \equiv x^n \) as follows:

\[
\mathcal{L}^\dagger |n\rangle = -\gamma n|n\rangle - \gamma x_c n|n-1\rangle + \frac{\sigma^2}{2} n(n-1)|n-2\rangle.
\] (31)

That is, there are three events as follows:

(I) (from \(-\gamma x \partial_x\)) \( n \rightarrow n \) (no state change), and the factor is \(-\gamma n\).

(II) (from \(-\gamma x_c \partial_x\)) \( n \rightarrow n-1 \), and the factor is \(-\gamma x_c n\).

(III) (from \((\sigma^2/2) \partial_x^2\)) \( n \rightarrow n-2 \), and the factor is \(\sigma^2 n(n-1)/2\).

3.2. First order moment in Ornstein-Uhlenbeck process

When the first order moment \( E[x_T - x_c] \) is calculated, the initial state is set to \( n = 1 \). In this case, event III never occurs. Additionally, if event II occurs once, all three events are no longer permitted. Hence, as shown in figure 1, possible combinations of paths gives

\[
\frac{T}{1!}(-\gamma x_c) + \frac{T^2}{2!}(-\gamma)(-\gamma x_c) + \frac{T^3}{3!}(-\gamma)^2(-\gamma x_c) \cdots
\]

\[
= x_c \left[ 1 + \frac{T}{1!}(-\gamma) + \frac{T^2}{2!}(-\gamma)^2 + \frac{T^3}{3!}(-\gamma)^3 + \cdots \right] - x_c
\]

\[
= x_c e^{-\gamma T} - x_c,
\] (32)

which adequately give the same analytical result in (12).

3.3. Second order moment in Ornstein-Uhlenbeck process

When the second moment \( E[(x_T - x_c)^2] \) is evaluated, the dual process starts from \( n = 2 \). In contrast to the previous case starting from \( n = 1 \), complicated discussions are
Combinatorics for calculating expectations of stochastic differential equations

Figure 2. Possible combinations of paths, starting from \( n = 2 \) and using only events I and II in the main text. A, B, and C correspond to cases with \( M = 1 \), \( M = 2 \), and \( M = 3 \), respectively.

necessary.

Firstly, event III gives a sudden change from \( n = 2 \) to \( n = 0 \); in this case, similar to figure 1, we have

\[
\sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sigma^2 \left( 1 - e^{-2\gamma T} \right).
\]

(33)

Secondly, consider the cases with event II occurs twice. Possible cases is depicted in figure 2. Case A in figure 2 is easy to evaluate as follows:

\[
\sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sigma^2 \left( 1 - e^{-2\gamma T} \right).
\]

(33)

As seen in case B and C in figure 2, event I, which gives no state change, should be adequately inserted. Using the following three identities

\[
\sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sigma^2 \left( 1 - e^{-2\gamma T} \right).
\]

(33)

and

\[
\sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_1 m} \right] \cdot \sum_{k=0}^{\infty} \frac{T^k}{k!} \left[ \left( -2\gamma \right)^{\lambda_2 m} \right] = \sigma^2 \left( 1 - e^{-2\gamma T} \right).
\]

(33)

we obtain the same equation with (24) by setting \( \lambda_1 = -2\gamma \) and \( \lambda_2 = -\gamma \). Of course, the contribution from event II, \( (-2\gamma x_c)(-\gamma x_c) \), should be multiplied finally.
4. Numerical applications

As seen in the previous section, the expectations of the stochastic differential equation can be evaluated from the combinatorial discussion for the dual (non-stochastic) process. Although the Ornstein-Uhlenbeck process is analytically tractable, it is difficult to perform such combinatorial discussions to general cases. Hence, it is important to check the numerical applicability of the combinatorial discussions. Although it would be difficult to generate all possible paths for high-dimensional cases, some combinatorial algorithms are proposed to calculate expansion coefficients for the exponential of operators. In [15, 16], coefficients needed to calculate the Mori-Zwanzig memory kernel are efficiently evaluated by recursive algorithms. Basically, our situation in the present paper is similar to [15, 16]; the following repeated action of $L^\dagger$ should be evaluated:

$$\langle n \rangle \rightarrow L^\dagger \langle n \rangle \rightarrow (L^\dagger)^2 \langle n \rangle \rightarrow \cdots \rightarrow (L^\dagger)^M \langle n \rangle.$$  \hfill (37)

In the present paper, we only focus on the coefficient for $n = 0$ at the final time because the coefficient is enough to evaluate the expectations $E[x - x_c]$ and $E[(x - x_c)^2]$ with the initial condition $x = x_c$ at time $t = 0$. Additionally, here we only focus on a simple system with only one variable, and hence it is easy to count all possible paths. For multivariate cases, the discussions given by Zhu et al. [16] will be helpful to construct practical algorithms.

4.1. Problem settings

Here, the following one variable system is considered:

$$dx = -\gamma x^3 dt + \sigma dW(t),$$  \hfill (38)

which is similar to the Ornstein-Uhlenbeck process, but the dependency on $x$ in the drift term is different. As in section 2.3, the two expectations, $E[x - x_c]$ and $E[(x - x_c)^2]$, are evaluated in this section. Hence, the adjoint operator of the dual process is given as

$$L^\dagger = -\gamma x^3 \frac{\partial}{\partial x} - 3\gamma x_c x^2 \frac{\partial}{\partial x} - 3\gamma x_c^2 x \frac{\partial}{\partial x} - \gamma x_c^3 \frac{\partial}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2}.$$  \hfill (39)

There are five events:

(I) (from $-\gamma x^3 \partial_x$) $n \rightarrow n + 2$, and the factor is $-\gamma n$.

(II) (from $-3\gamma x_c x^2 \partial_x$) $n \rightarrow n + 1$, and the factor is $-3\gamma x_c n$.

(III) (from $-3\gamma x_c^2 x \partial_x$) $n \rightarrow n$ (no state change), and the factor is $-3\gamma x_c^2 n$.

(IV) (from $-\gamma x_c^3 \partial_x$) $n \rightarrow n - 1$, and the factor is $-\gamma x_c^3 n$.

(V) (from $(\sigma^2/2) \partial_x^2$) $n \rightarrow n - 2$, and the factor is $\sigma^2 n(n - 1)/2$.

4.2. Taylor series and Padé approximation

As discussed in section 2.2, the time-evolution operator is formally expressed by the exponential, and then the following Taylor series is a first candidate to construct
Figure 3. Numerical results for (38). (a) and (b) correspond to expectation values of $E[x - x_c]$ and $E[(x - x_c)^2]$, respectively. Points with error bars are given by the Monte Carlo simulations. Solid and dotted lines correspond to $[7/7]$ and $[8/8]$ Padé approximations, respectively, and the dashed line is obtained by the simple Taylor series up to 16th terms.

Numerical evaluation:

$$e^{L^\dagger t} \simeq \sum_{m=0}^{M} \frac{t^m}{m!} (L^\dagger)^m.$$  \hfill (40)

As an example, we here employ the following parameters: $\gamma = 2.0$, $\sigma = 2.0$, and $x_c = 1.0$. Let $f_{1\to0}(t)$ be the contribution of the state change from $n = 1$ to $n = 0$. Then, by evaluating the combinatorics of the possible paths numerically, we have

$$f_{1\to0}(t) \simeq -2t - 6t^2 + 68t^3 - 218t^4 - 1653t^5 + 23562.4t^6 + \cdots.$$ \hfill (41)

The behavior of the coefficients has the increasing oscillations with sign changes. Although we expect that the coefficients are finally decreasing because of the existence of $m!$ in the denominator in (40), such $m$ become very large in general, and it is not practical to calculate the combinatorics up to such $m$.

It is easy to see that the simple summation of the Taylor series does not work, and hence here the Padé approximation is used; the order $[m/n]$ Padé approximation is defined as [31]

$$f(t) \simeq \frac{P_m(t)}{Q_n(t)} = \frac{p_0 + p_1 t + p_2 t^2 + \cdots + p_m t^m}{1 + q_1 t + q_2 + \cdots + q_n t^n}. \hfill (42)$$

That is, the function is approximately expressed as a rational function in the Padé approximation.

Figure 3 shows the numerical results. For comparison, Monte Carlo simulations used the Euler-Maruyama approximation are performed; the time-discretization with $10^{-4}$ is employed, and averages for 1000 samples are taken. To depict the error bars, the same simulations with different random-number seeds are performed 10 times. [7/7]
and [8/8] Padé approximation results, and that of the simple Taylor summation up to 16th terms are depicted for $\mathbb{E}[x - x_c]$ and $\mathbb{E}[(x - x_c)^2]$. It is easy to see that the simple Taylor summation gives the sudden diverging behavior when $t$ increases. On the other hand, the [8/8] Padé approximation provides reasonable results even for the large $t$ cases.

Note that it is not guaranteed that higher-order approximations give always better results; the denominator $Q_n(t)$ sometimes takes a small value, which causes unstable results. However, the following facts are clarified:

- The simple Taylor summation is not applicable in practical cases.
- Additional techniques, such as Padé approximations, are necessary.
- For small time-interval cases, the algorithm gives enough accurate estimations. Although the accuracy of the approximations would not be enough for large time-interval cases, rough estimations are possible.

### 4.3. Usage of resolvent and Aitken acceleration

In section 4.2, we employed the following simple Taylor series to interpret the time-evolution operator:

$$e^{Lt} = \sum_{m=0}^{\infty} \frac{t^m}{m!} (L^\dagger)^m.$$  \hspace{1cm} (43)

However, it is known that the definition based on the Taylor series is difficult and numerically not useful if $L^\dagger$ is unbounded operator in a Banach space [32]. Instead, the following formula is available:

$$e^{Lt} = \lim_{M \to \infty} \left[ \left( 1 - \frac{t}{M} L^\dagger \right)^{-1} \right]^M,$$  \hspace{1cm} (44)

where $(1 - \frac{t}{M} L^\dagger)^{-1}$ is a resolvent of $L^\dagger$, apart from a constant factor [32]. In this section, this definition for the expansion of the time-evolution operator is employed and investigated.

Since $L^\dagger$ is expressed as an infinite-dimensional matrix, it is difficult to obtain the inverse matrix of $1 - \frac{t}{M} L^\dagger$. However, neglecting mathematical rigorous discussions, the conventional Gauss elimination method is employed here as follows: Consider the infinite dimensional vector

$$|n\rangle = (0 \cdots 0 1 0 \cdots)^T,$$  \hspace{1cm} (45)

where only $n$-th element has 1. Then, the infinite-dimensional matrix corresponding to
Combinatorics for calculating expectations of stochastic differential equations

Figure 4. The dependence of the factor at \( n = 0 \) with \( M \). The parameters \( \gamma = 2.0, \sigma = 2.0, x_c = 1.0, \) and \( t = 1.0 \) are used. As \( M \) increases, the factor shows convergence behavior; however, larger \( M \) makes the value unstable because of the approximation of the resolvent. The dotted horizontal line corresponds to the mean value of the corresponding Monte Carlo simulations.

\[ 1 - \frac{t}{M^L} \] is written in the following form:

\[
\begin{pmatrix}
  \cdots & \cdots & \cdots & \cdots & \cdots \\
  \cdots & e_{n-1} & 0 & 0 & \cdots \\
  \cdots & d_{n-1} & e_n & 0 & \cdots \\
  \cdots & 1 - c_{n-1} & d_n & e_{n+1} & \cdots \\
  \cdots & b_{n-1} & 1 - c_n & d_{n+1} & \cdots \\
  \cdots & a_{n-1} & b_n & 1 - c_{n+1} & \cdots \\
  \cdots & 0 & a_n & b_{n+1} & \cdots \\
  \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]

where \( a_n, \ldots, e_n \) correspond to factors of the above five events, respectively. Note that each factor is multiplied by \( t/M \). Hence, if \( M \) is large enough, the \( n \)-th column in the inverse matrix, \( (1 - \frac{t}{M^L})^{-1} \), would be approximately expressed as follows:

\[
\begin{pmatrix}
  \vdots \\
  0 \\
  (-e_n)(1 - c_{n-2})^{-1}(1 - c_n)^{-1} \\
  (-d_n)(1 - c_{n-1})^{-1}(1 - c_n)^{-1} \\
  (1 - c_n)^{-1} \\
  (-b_n)(1 - c_{n+1})^{-1}(1 - c_n)^{-1} \\
  (-a_n)(1 - c_{n+2})^{-1}(1 - c_n)^{-1} \\
  0 \\
  \vdots
\end{pmatrix}
\]
Combinatorics for calculating expectations of stochastic differential equations

Figure 5. Numerical results obtained by the use of resolvent in (46). (a) and (b) correspond to expectation values of $E[x - x_c]$ and $E[(x - x_c)^2]$, respectively. Points with error bars are given by the Monte Carlo simulations. Dotted line corresponds to the raw value of the factor at $M = 100$. Dashed line is obtained by the Aitken acceleration method using the sequence from $M = 0$ to $M = 100$.

That is, if $M$ is large, $a_n, \ldots, e_n$ are small enough, and then, for example, $b_n e_{n+1}$ could be small enough to be neglected. Hence, using the combinatorial algorithm, it is possible to evaluate

$$e^{L^\dagger t}|n\rangle \simeq \left[ \left( 1 - \frac{t}{M L^\dagger} \right)^{-1} \right]^M |n\rangle$$

approximately. At the final step, the factor at the state $n = 0$ corresponds to the target statistics.

To evaluate $E[x - x_c]$, the initial state is set to $n = 1$. The dependence of the factor at $n = 0$ with $M$ is depicted in figure 4 in which the time is $t = 1.0$, and the same parameters in section 4.2, i.e., $\gamma = 2.0$, $\sigma = 2.0$, and $x_c = 1.0$, are used. When $M$ increases, the factor seems to converge to a certain value. However, larger $M$ gives the unstable and diverging behavior as shown in the region $M \geq 150$ in figure 4. The resolvent used here is the approximate one, and hence the diverging behavior could be caused.

To avoid this unstable numerical behavior, let us consider the use of only the sequences in the stable regions. A simple choice could be to find the convergent point from figure 4. However, if we change the parameters, for example, the time $t$, the diverging behavior is changed from that of figure 4. This detailed analysis is cumbersome, and therefore the following method is proposed here: Only using the series in the stable region, the Aitken acceleration method is applied [31]. The Aitken acceleration is performed as

$$z^{(k+1)}_m = z^{(k)}_m - \frac{\left( z^{(k+1)}_m - z^{(k)}_m \right)^2}{z^{(k)}_m - 2z^{(k+1)}_m + z^{(k)}_m},$$

(47)
and here the series from $M = 0$ to $M = 100$ was used. The Aitken acceleration is repeatedly applied 10 times, i.e., $k = 10$ cases are evaluated. The evaluated results for $\mathbb{E}[x - x_c]$ and $\mathbb{E}[(x - x_c)^2]$ are shown in figure 5. Even for large $t$ regions, the Aitken acceleration method works well.

5. Concluding remarks

In the present paper, the combinatorial discussion for calculating expectations of stochastic differential equations was given. The discussion naturally connects the duality relation in stochastic processes and the combinatorics. The combinatorics was useful to recover the well-known results of the Ornstein-Uhlenbeck process. Furthermore, to seek numerical methods is important in the practical use of combinatorial discussions, and two candidates were proposed. One is the use of Padé approximation in the conventional Taylor type expansion of the time-evolution operator. Another is based on the resolvent of the time-evolution operator, and the Aitken acceleration is also employed. These two methods give reasonable approximations, and especially, the usage of the resolvent and the Aitken acceleration seems to work well. Of course, numerical checks have been performed only for limited cases; further studies, including higher-dimensional cases, should be performed in future.

Finally, there are some comments. Mathematically rigorous discussions should be performed in the future works, especially for the usage of the approximated resolvent and its complicated divergent behavior. The repeated applications of the resolvent in (46) seem to be a simple Bernoulli trial; its limit with $M \to \infty$ might give the conventional Poisson process. Such consideration might be important to construct more stable numerical algorithms.

Acknowledgments

This work was supported by JST, PRESTO Grant Number JPMJPR18M4, Japan.

[1] Gardiner C 2009 *Stochastic methods: A handbook for the natural and social sciences, 4th edition.* (Berlin Heidelberg: Springer)
[2] Kloeden P E and Platen E 1992 *Numerical Solution of Stochastic Differential Equations* (Berlin: Springer)
[3] Liggett T M 2005 *Interacting Particle Systems (Classics in Mathematics)* (Berlin: Springer) Reprint of the 1985 edition
[4] Jansen S and Kurt N 2014 *Probab. Surveys* **11** 59
[5] Shiga T and Uchiyama K 1986 *Probab. Th. Rel. Fields* **73** 87
[6] Möhle M 1999 *Bernoulli* **5** 761
[7] Carinci G, Giardinà C, Giberti C and Redig F 2015 *Stochastic Processes and their Applications* **125** 941
[8] Giardinà C, Kurchan J and Redig F 2007 *J. Math. Phys.* **48** 033301
[9] Carinci G, Giardinà C, Giberti C and Redig F 2013 *J. Stat. Phys.* **152** 657
[10] Giardinà C, Kurchan J, Redig F and Vafayi K 2009 *J. Stat. Phys.* **135** 25
Combinatorics for calculating expectations of stochastic differential equations

[11] Franceschini C and Giardinà C 2019 *Sojourns in Probability Theory and Statistical Physics - III* ed V Sidoravicius (Singapore: Springer) pp 187-214
[12] Redig F and Sau F 2018 *J. Stat. Phys.* **172** 980
[13] Groenevelt W 2019 *J. Stat. Phys.* **174** 97
[14] Ohkubo J 2015 *Phys. Rev. E* **92** 043302
[15] Amati G, Meyer H and Schilling T 2019 *J. Stat. Phys.* **174** 219
[16] Zhu Y and Venturi D 2020 *J. Stat. Phys.* **178** 1217
[17] Schütz G M 1997 *J. Stat. Phys.* **86** 1265
[18] Imamura T and Sasamoto T 2011 *J. Stat. Phys.* **142** 919
[19] Ohkubo J 2017 *J. Phys. A: Math. Gen.* **50** 095004
[20] Ohkubo J and Arai Y 2019 *J. Stat. Mech.* 063202
[21] Ohkubo J 2020 *J. Phys. Soc. Jpn.* **89** 044004
[22] Ohkubo J 2013 *J. Stat. Phys.* **174** 219
[23] Gillespie D T 1977 *J. Phys. Chem.* **81** 2340
[24] Wilkinson D J 2006 *Stochastic Modelling for Systems Biology* (Boca Raton: CRC Press)
[25] Warne D J, Baker R E and Simpson M J 2019 *J. R. Soc. Interface* **16** 20180943
[26] Ohkubo J 2010 *J. Stat. Phys.* **139** 454
[27] Weber M F and Frey E 2017 *Rep. Prog. Phys.* **80** 046601
[28] Kleinert H 2006 *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets, 4th Edition* (Danvers: World Scientific Publishing)
[29] Voliotis M, Thomas P, Grima R and Bowsher C G 2016 *PLoS Comput. Biol.* **12** e1004923
[30] Beentjes C H L and Baker R E 2019 *J. Chem. Phys.* **150** 154107
[31] Press W H, Teukolsky S A, Vetterling W T and Flannery B P *Numerical Recipes in C, 2nd edition* (Cambridge: Cambridge University Press)
[32] Kato T 1966 *Perturbation Theory for Linear Operators* (Berlin: Springer)