Nonlinear Kalman filter based on duality relations between continuous and discrete-state stochastic processes

Jun Ohkubo

Abstract A nonlinear Kalman filter for systems described by stochastic differential equations is proposed. In the conventional ensemble Kalman filter, direct simulations for the stochastic differential equations are needed at each measurement time step. On the other hand, in the new Kalman filter, we perform Monte Carlo simulations for the dual birth-death processes ‘in advance’, and the numerical results are used to construct the new filter. The characteristics of the new Kalman filter based on the duality relations are explained, and a demonstration of the new Kalman filter is given.

Keywords Stochastic differential equation · Birth-death process · Algebraic probability · Partial measurement

1 Introduction

Data assimilation is one of the important topics in various research fields, in which both the theoretical model and experimental measurements collaborate with each other in order to estimate states from noisy data. For linear systems under Gaussian noise, the classical Kalman filter is available and gives a reasonable solution for the estimation problem. On the other hand, many phenomena in the real world are suitably modeled as nonlinear systems. As for nonlinear filtering problems, no complete solution has been known, and various methods have been proposed.

Up to now, some filtering methods based on Monte Carlo simulations have been proposed. One of the powerful methods is the particle filter, in which nonlinear systems are simulated directly by the Monte Carlo simulations (Gordon et al. 1993; Kitagawa 1996; Candy 2009). The particle filter is based on the Bayesian statistics, and it has been shown that the particle filter works well in various examples. The ensemble Kalman filter (EnKF) is also famous and popular (Evensen 1994; Gillijns et al. 2006; Evensen 2009). The EnKF is also based on the direct simulation of the nonlinear systems; using the ensemble of the simulations, the means and covariance matrix of the nonlinear systems are estimated in order to obtain the Kalman gain, which are used in the similar way with the conventional Kalman filter.

In both the particle filter and EnKF, the Monte Carlo simulations are used to predict the next states, starting from the current states. Since the measurements are performed at discrete instances in time, the Monte Carlo simulations are needed for each measurement time step. That is, we must perform a new Monte Carlo simulation with a specific initial condition at each measurement time step. In addition, if we want to obtain precise estimations, a large size of the ensemble in the Monte Carlo simulations is demanded, and the computational time is increasing with the size of the ensemble.

In the present paper, a novel filter based on the Monte Carlo simulations is proposed. The new filter can be used for filtering problems of nonlinear systems described by stochastic differential equations; the systems have continuous time and continuous states. If we apply the EnKF to the filtering problem, we need the Monte Carlo simulations at each measurement time step, as described above. On the other hand, in the new filter, Monte Carlo simulations are performed before solving the estimation problems in advance, and we reuse the
numerical results at each measurement time step in the filtering problem. In addition, the Monte Carlo simulation in advance is not for the original stochastic differential equations, but for the dual birth-death processes. The duality relations between stochastic differential equations and birth-death processes have been studied (Liggett 2005; Giardinà et al. 2009; Ohkubo 2010), and recently, a widely applicable scheme to derive a dual birth-death process from stochastic differential equations has been proposed (Ohkubo 2013b). Using the duality relation between the stochastic differential equations and the dual birth-death processes, we can obtain the information about the stochastic differential equations without solving the stochastic differential equation directly. Since there is no need to perform the Monte Carlo simulations at each measurement time step, in general, it is expected that our new filter works very rapidly compared with the EnKF.

The present paper is constructed as follows. In Sec. 2, the model used in the present paper is explained. Section 3 is a brief review of the EnKF. The main proposal in the present paper is given in Sec. 4; the derivation of the dual birth-death process and the usage of the duality relation are explained. In Sec. 5, results of an demonstration of the new filter and comparisons with the EnKF are given. Section 6 is for concluding remarks.

2 Model

2.1 Time-evolution of the state variables

In the present paper, the following Van der Pol-type model is used (Lakshmivarahan and Stensrud 2009):

\[
\begin{align*}
\frac{dx_1(t)}{dt} &= x_2(t) + w_1(t), \\
\frac{dx_2(t)}{dt} &= \epsilon(1 - x_1(t)^2)x_2(t) - x_1(t) + w_2(t),
\end{align*}
\]

where \(w_i(t) \in \mathbb{R}\) is zero-mean white Gaussian noise with a covariance matrix \(Q \in \mathbb{R}^{2 \times 2}\). Different from the original Van der Pol model, the model (1) contains the noise terms. Here, we assume that the noises in (1) are not correlated with each other, and then the covariance matrix is a diagonal matrix; \(Q = \text{diag}[Q_{11}, Q_{22}]\). In the following, the vector \(\mathbf{x}(t) = [x_1(t) \ x_2(t)]^T\) is sometimes used for notational brevity.

2.2 Measurements

The time-evolution of the state variable \(\mathbf{x}\) obeys (1). Here, suppose that only one of the state variable \(\mathbf{x}\) can be observed, and that the measurement is performed with certain time intervals. That is, although the time-evolution of the model (1) is continuous, the measurement results at each measurement time step in the Monte Carlo simulations at each measurement time step in the EnKF are given. Section 6 is for concluding remarks.

2.3 Data used in the present paper

The discrete version of the model (1) has been used in the work by Lakshmivarahan and Stensrud (2009), and hence we here employ the following parameters, which are similar to the previous work: \(\epsilon = 1.0, Q_{11} = 0.0262, Q_{22} = 0.008, R = 0.04\). (Compared with the work by Lakshmivarahan and Stensrud (2009), we use a little larger measurement noise.) Using these parameters, the data for the filtering problem is created as follows.

Firstly, the time-evolution in (1) is simulated using the first-order Euler-Maruyama scheme (Kloeden and Platen 1992; Gardina 2009); the time interval for the simulation is \(10^{-4}\), and the initial conditions are
procedure is performed; only the state variable \( x_2(t) \) is extracted and the measurement noise is added. The time interval for the measurements is \( \Delta t_{\text{obs}} = 0.2 \), and finally, we obtain the measurement data depicted in Fig. 1.

The aim of the filtering problem is to estimate the state variable \( x(t) \) (i.e., not only \( x_2(t) \), but also \( x_1(t) \)) from the partially measured data in Fig. 1. Although it is necessary to estimate the covariance matrix \( Q \) for the noise in the model and the variance \( R \) of the measurement noise in practice, in the present paper, we assume that these parameters are previously known for simplicity.

### 3 Brief review of EnKFs

Here, a brief summary of the EnKF is given (Evensen 1994; Gillijns et al. 2006; Evensen 2009). As explained in Sec. 1, the EnKF uses the Monte Carlo simulations; using an ensemble of many particles, statistical quantities such as means and covariances are evaluated, and these quantities are employed to calculate the Kalman gain. Note that we here use the problem settings in Sec. 2, and hence the measurements are performed only for the discrete times \( \{\tau_1, \tau_2, \cdots\} \).

**Algorithm for EnKF**

1. **Initialization:**
   Make the initial ensemble. Here, we choose \( n \) samples from a Gaussian distribution with mean \( \bar{x}(0) \) and covariance matrix \( P(0) \), where \( \bar{x}(0) \) and \( P(0) \) are chosen arbitrarily. We denote each sample at time \( t_0 = 0 \) as \( x_i(\tau_0) \) for \( i = 1, 2, \cdots, n \).

2. **Forecast step** (\( x_i(\tau_{k-1}) \rightarrow x_i^f(\tau_k) \)):
   Using the time-evolution of (1), simulate the path of the state variable for each sample starting from \( x_i(\tau_{k-1}) \). For the simulation, for example, the first-order Euler-Maruyama scheme is available. The simulated path for sample \( i \) is denoted as \( x_i^f(t) \) (\( \tau_{k-1} < t \leq \tau_k \)).

3. **Assimilation step** (\( x_i^f(\tau_k) \rightarrow x_i(\tau_k) \)):
   (a) Make realizations of random variables \( \{v_i(\tau_k)\}_{i=1}^{n} \) as the measurement noises. Each realization is obtained from the zero-mean white Gaussian noise with the variance \( R \).
   (b) Calculate following quantities:
   \[
   \bar{x}(\tau_k) = \frac{1}{n} \sum_{i=1}^{n} x_i(\tau_k), \tag{3}
   \]

   (error matrix \( E^f(\tau_k) \in \mathbb{R}^{2 \times n} \))
   \[
   E^f(\tau_k) = [x_1^f(\tau_k) - \bar{x}(\tau_k) \cdots x_n^f(\tau_k) - \bar{x}(\tau_k)], \tag{4}
   \]
   (unbiased covariance evaluated from the ensemble)
   \[
   \hat{P}^f(\tau_k) = \frac{1}{n-1} E^f(\tau_k)(E^f(\tau_k))^T, \tag{5}
   \]
   (mean of the measurement noises)
   \[
   \bar{y}(\tau_k) = \frac{1}{n} \sum_{i=1}^{n} y_i(\tau_k), \tag{6}
   \]
   (unbiased variance of the measurement noise)
   \[
   \hat{R}(\tau_k) = \frac{1}{n-1} \sum_{i=1}^{n} (y_i(\tau_k) - \bar{y}(\tau_k))^2, \tag{7}
   \]
   (Kalman gain)
   \[
   \hat{K}(\tau_k) = \hat{P}^f(\tau_k)H^T \left( H \hat{P}^f(\tau_k)H^T + \hat{R}(\tau_k) \right)^{-1}. \tag{8}
   \]
   (c) Modify the forecasted state variables \( \{x_i^f(\tau_k)\}_{i=1}^{n} \) using the measurement at time \( \tau_k \), i.e., \( y(\tau_k) \), as follows:
   \[
   x_i(\tau_k) = x_i^f(\tau_k) + \hat{K}(\tau_k) \left( y(\tau_k) + v_i(\tau_k) - H x_i^f(\tau_k) \right). \tag{9}
   \]

Steps 2 and 3 in the above algorithm are performed for each measurement time step.

In the EnKF, the time evolution in the forecast step is performed as the nonlinear systems, which gives the non-Gaussian distribution for \( \{x_i(t)\}_{i=1}^{n} \) even if we start from a Gaussian distribution. After the time evolution, at each assimilation step, the conventional Kalman filter is employed, which means a filter at least up to the second moment. The final filtered value of the state variable \( x(t) \) is obtained as the mean of the ensemble \( \{x_i(t)\}_{i=1}^{n} \). In addition, the estimated error could be obtained from the (co)variance of the ensemble \( \{x_i(t)\}_{i=1}^{n} \).

One of the problem in the EnKF is as follows: in order to obtain the more accurate filtering results, we need large ensemble size \( n \). That is, a small ensemble size gives inaccurate Kalman gain, and hence the filtering results would not be precise. In general, it is easy to imagine that the large ensemble size needs high computational costs. Although it has been clarified that not so large ensemble size is necessary for the EnKF in practical cases (Lakshmivarahan and Stensrud 2009), it would be preferable if we could avoid the numerical simulations of the time evolution of the ensemble at each measurement time step.
4 Filter based on duality relations

4.1 Basic concept

Here, we simply explain the basic concept of the duality relation between a stochastic differential equation and a birth-death process. For simplicity, stochastic processes with only one random variable are considered in this subsection.

Suppose that \( (x_t)_{t \geq 0} \in \mathbb{R} \) is a sample trajectory of the stochastic differential equation, and \( p(x, t) \) is the probability density at time \( t \). For some specific stochastic differential equations, it has been known that there is the corresponding birth-death process \( (n_t)_{t \geq 0} \in \mathbb{N} \), whose probability distribution at time \( t \) is \( P(n, t) \), and the following equality is satisfied (Liggett 2005; Gardiná et al. 2009; Ohkubo 2010):

\[
E_x [x_t^n] = E_n [x_0^n],
\]

where \( E_x \) and \( E_n \) are the expectations in the stochastic differential equation \( (x_t)_{t \geq 0} \) starting from \( x_0 \) and in the birth-death process \( (n_t)_{t \geq 0} \) starting from \( n_0 \), respectively. More explicitly, we can rewrite (10) as

\[
\int_{-\infty}^{\infty} p(x, t)x^n dx = \sum_{n=0}^{\infty} P(n, t)x^n_0, \quad \text{(11)}
\]

where \( p(x, 0) = \delta(x - x_0) \) and \( P(n, t) = \delta_{n_0,n_0}. \)

Equation (10) shows that the information about the stochastic differential equation can be obtained from the solution of the birth-death process. That is, when we obtain the probability distribution of the birth-death process \( P(n, t) \), with the initial condition \( n_0 = 1 \), it is possible to evaluate the first order moment, i.e., the mean value of \( x_t \), of the stochastic differential equation, without solving the stochastic differential equation.

There are several advantage of the usage of the duality relations. It is sometimes easier to treat the birth-death process, compared with the stochastic processes. For some specific cases, the analytical solution of the birth-death process has been obtained. In addition, there are numerical algorithms to simulate the birth-death process exactly. As for the stochastic differential equations, we need some approximation; for example, the time-discretization is needed in the Euler-Maruyama scheme. On the other hand, for example, the Gillespie algorithm for the birth-death process does not need the time-discretization (Gillespie 1977). In this sense, the birth-death process would be more tractable than the stochastic differential equations.

Furthermore, only ‘a’ solution of the birth-death process can be used to obtain the information about the stochastic differential equations with ‘arbitrary’ initial conditions. That is, \( E_x [x_1] = E_n [x_0^n] = \sum_{n=0}^{\infty} P(n, t)x^n_0 \) if \( n_0 = 1 \), and \( P(n, t) \) is independent of the value of \( x_0 \). Hence, using only ‘a’ solution of the birth-death process, \( P(n, t) \), it is possible to estimate the average of \( x_t \), \( E_x [x_t] \), for the stochastic differential equations for ‘arbitrary’ initial conditions \( x_0 \). Of course, if we want to know the first and second moments of \( x_t \) in (10), two solutions of the birth-death process \( n_t \) with the different initial conditions, \( n_0 = 1 \) and \( n_0 = 2 \), are needed. However, once we have these two solutions of the birth-death process, the first and second moments, \( E_x [x_t] \) and \( E_x [x^2_t] \), with arbitrary initial conditions can be evaluated by using the duality relation (10). In contrast, when these moments are evaluated from the direct simulation of the stochastic differential equation, we need many sample trajectories with an initial condition \( x_0 \); if the initial condition \( x_0 \) is changed, we must perform many other numerical simulations.

The basic idea here is as follows; the Monte Carlo simulation in the forecast step in the EnKF is replaced with the simple numerical evaluation based on the duality relation.

The remaining problem is as follows: How should we derive the dual birth-death process from a given stochastic differential equation? In the successive subsections, we will show the method to obtain the dual birth-death process, employing a mathematical formalism called the Doi-Peliti formalism (Doi 1976a,b; Peliti 1985; Tauber et al. 2005).

4.2 Doi-Peliti formalism

We here briefly review the Doi-Peliti formalism, which is useful to obtain the duality relations. The Doi-Peliti formalism is the method similar to the second quantization method in quantum mechanics. Up to now, the Doi-Peliti formalism has been used in various contexts, mainly in order to investigate discrete systems such as chemical reactions, and it has been shown that the algebraic probability theory (Hora and Obata 2007) gives the mathematical basis of the Doi-Peliti formalism (Ohkubo 2013a).

In the Doi-Peliti formalism, creation operator \( a^\dagger \) and annihilation operator \( a \) are introduced, which satisfy the following commutation relation:

\[
[a, a^\dagger] \equiv aa^\dagger - a^\dagger a = 1, \quad [a, a] = [a^\dagger, a^\dagger] = 0. \quad \text{(12)}
\]

These operators act on a vector in the Fock space, \( |n\rangle \), as follows:

\[
a^\dagger |n\rangle = |n + 1\rangle, \quad a|n\rangle = n|n - 1\rangle, \quad \text{(13)}
\]
and the vacuum state $|0\rangle$ is characterized by $a|0\rangle$. Additionally, vectors \( \{m| \}_{n=0}^{\infty} \) satisfy the following orthogonal relation to the vectors \( \{|n\rangle \}_{n=0}^{\infty} \):

\[
\langle m|n \rangle = \delta_{m,n}!.
\]

Note that \( a^\dagger a \) corresponds to the number operator and \( a^\dagger a|n\rangle = n|n\rangle \).

It has been shown that the Doi-Peliti formalism is deeply related to the conventional generating function method, and the following correspondences can be useful to understand the usage of the Doi-Peliti formalism in the duality problem:

\[
|n\rangle \leftrightarrow x^n, \quad a^\dagger \leftrightarrow x, \quad a \leftrightarrow \partial \frac{\partial}{\partial x}.
\]

That is, the differential operator is connected to the annihilation operator, which is used in partial differential equations derived from stochastic differential equations. In addition, the annihilation operator acts on the discrete states \( |n\rangle \), which is available to construct the birth-death process, as shown later. The important point here is that the Doi-Peliti formalism can bridge continuous states with discrete states, which corresponds to the connection between a stochastic differential equation and a birth-death process.

### 4.3 Derivation of the dual birth-death process

The derivation of the simple duality relation, such as (10), has been discussed by Giardiná et al. (2009), and the derivation based on the Doi-Peliti formalism has also been proposed (Ohkubo 2010). However, in order to treat the filtering problem, only the simple duality relation (10) is not enough; the simple duality relation (10) is not enough to construct the birth-death process, as shown later. The important point here is that the Doi-Peliti formalism can bridge continuous states with discrete states, which corresponds to the connection between a stochastic differential equation and a birth-death process.

The adjoint operator of (18), \( L \), is as follows:

\[
L^* \left( x_1, \frac{\partial}{\partial x_1}, x_2, \frac{\partial}{\partial x_2} \right) = x_2 \frac{\partial}{\partial x_1} - \epsilon(1 - x_1^2)x_2 \frac{\partial}{\partial x_2} - x_1 \frac{\partial}{\partial x_2} + \frac{1}{2} \left[ Q_{11} \frac{\partial^2}{\partial x_1^2} + Q_{22} \frac{\partial^2}{\partial x_2^2} \right],
\]

and using the correspondence between operators in the Doi-Peliti formalism and the differential operators (16), we have

\[
L \left( a_1^\dagger, a_1, a_2^\dagger, a_2 \right) = a_2^\dagger a_1 + \epsilon(1 - a_1^\dagger a_1) a_2^\dagger a_2 - a_1^\dagger a_2 + \frac{1}{2} \left[ Q_{11} a_1^\dagger a_1 + Q_{22} a_2^\dagger a_2 \right],
\]

where the following correspondences are used:

\[
a_1^\dagger \leftrightarrow x_1, \quad a_1 \leftrightarrow \frac{\partial}{\partial x_1}, \quad a_2^\dagger \leftrightarrow x_2, \quad a_2 \leftrightarrow \frac{\partial}{\partial x_2}.
\]

In addition, as discussed later, it is convenient to introduce the time scaling \( t = r_{ts} \tilde{t} \); due to the time scaling, the original Fokker-Planck equation (19) is rewritten as

\[
\frac{\partial}{\partial \tilde{t}} p(x_1, x_2, r_{ts} \tilde{t}) = r_{ts} L^* \left( x_1, \frac{\partial}{\partial x_1}, x_2, \frac{\partial}{\partial x_2} \right) p(x_1, x_2, r_{ts} \tilde{t}).
\]
We here focus on the fact that continuous variables, \(x_1\) and \(x_2\), are replaced with creation operators in (22). Hence, if we reinterpret a constant as a creation operator, the following replacement is available:

\[
r_{ts} \leftrightarrow a_0^\dagger,
\]

and therefore the following linear operator is obtained:

\[
L \left( a_0^\dagger, a_1^\dagger, a_1, a_2^\dagger, a_2 \right)
\equiv r_{ts} L \left( a_1^\dagger, a_1, a_2^\dagger, a_2 \right)
= a_0^\dagger L \left( a_1^\dagger, a_1, a_2^\dagger, a_2 \right)
= a_0^\dagger a_0 a_1^\dagger + e a_0^\dagger a_1^\dagger a_2^\dagger a_2 - e a_0^\dagger a_1^\dagger a_1^\dagger a_2 - 2 a_0^\dagger a_1^\dagger a_2^\dagger a_2
+ \frac{1}{2} \left[ Q_{11} a_0^\dagger a_1 + Q_{22} a_2 a_2^\dagger \right].
\]

Note that all creation operators in each term in (25) are placed on the left side of the annihilation operators; if not, we must replace the term using the commutation relation (12).

Since the linear operator (25) acts on the Fock space, i.e., the discrete state space \(\{|n\}_n \in \mathbb{N}_0\), one may expect that the linear operator simply gives the time-evolution for a birth-death process. That is, defining the state vector |\(\psi(t)\rangle\) as

\[
|\psi(t)\rangle \equiv \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} P(n_0, n_1, n_2, t) |n_0, n_1, n_2\rangle,
\]

where \(P(n_0, n_1, n_2, t)\) is a probability distribution of a birth-death process, and considering the time-evolution equation

\[
\frac{\partial}{\partial t} |\psi(t)\rangle = L(a_0^\dagger, a_1^\dagger, a_1, a_2^\dagger, a_2) |\psi(t)\rangle,
\]

we may have a time-evolution equation for \(P(n_0, n_1, n_2, t)\) by comparing the coefficient of a state vector |\(n_0, n_1, n_2\rangle\) on the right and left hand sides. However, as discussed in the previous work (Ohkubo 2013b), it is impossible to simply interpret the linear operator (25) as a time-evolution operator for a birth-death process; the linear operator does not satisfy the probability conservation law, and, in addition, some terms seem to correspond to ‘negative’ transition rates.

In order to construct an adequate birth-death process, we need additional operator \(b\), which satisfies the following relations:

\[
b |+\rangle = \langle -|, \quad b |-\rangle = |+\rangle,
\]

where \(+\) and \(-\) are orthonormal state vectors satisfying \(\langle +|+\rangle = \langle -|-\rangle = 1\) and \(\langle +|-\rangle = \langle -|+\rangle = 0\). The following procedure is needed to construct an adequate time-evolution operator for a birth-death process (although the following procedure might seem complicated, a concrete example will be given soon):

For each term in the linear operator \(L\), apply the following procedures.

1. If the coefficient of the term has a negative sign, replace the negative sign ‘\(-\)’ with ‘\(+b\)’ using the operator \(b\).
2. Act the operators on \(|n\rangle\), and evaluate the coefficient including the effects of the number operators \(a_0^\dagger a_1\).
3. Subtract a term in order to guarantee the probability conservation law; the term gives the same coefficient with Step 2, and additionally, the term consists of the same number of creation and annihilation operators for the same sub-index.
4. In order to compensate the subtracted term, add the same term with Step 3.

For example, the first term in (25) is \(a_0^\dagger a_1^\dagger a_1, a_0\), and then we do not need Step 1. Since \(a_0^\dagger a_1^\dagger a_1|n_0, n_1, n_2\rangle = n_1|n_0 + 1, n_1 − 1, n_2 + 1\rangle\),

\[
\text{the coefficient is } n_1, \text{ and therefore the term, } a_0^\dagger a_1, \text{ must be subtracted in Step 2. On the other hand, the third term in (25), } −e a_0^\dagger a_1^\dagger a_2^\dagger a_2, \text{ has the negative sign, and hence Step 1 gives } +b a_0^\dagger a_1^\dagger a_2^\dagger a_2.
\]

The subtracted term in step 3 is \(ba_2^\dagger a_2\) because

\[
a_0^\dagger a_1^\dagger a_2^\dagger a_2|n_0, n_1, n_2\rangle = n_2|n_0 + 1, n_1 + 2, n_2\rangle,
\]

which gives the coefficient \(n_2\).

Using the above procedure, the linear operator \(L\) in (25) is rewritten as follows:

\[
L \left( a_0^\dagger, a_0, a_1^\dagger, a_1, a_2^\dagger, a_2, b \right)
= L' \left( a_0^\dagger, a_0, a_1^\dagger, a_1, a_2^\dagger, a_2, b \right) + V \left( a_1^\dagger a_1, a_2^\dagger a_2 \right),
\]

where

\[
L' \left( a_0^\dagger, a_0, a_1^\dagger, a_1, a_2^\dagger, a_2, b \right)
= (a_0^\dagger a_2^\dagger a_1 - a_1^\dagger a_1) + (ea_0^\dagger a_2^\dagger a_2 - ea_1^\dagger a_2)
+ (eba_0^\dagger a_1^\dagger a_2 a_2 - ea_2^\dagger a_2) + (ba_2^\dagger a_2 - a_3^\dagger a_2)
+ \left( \frac{1}{2} Q_{11} a_1 a_1 - \frac{1}{2} Q_{11} a_1^\dagger a_1^\dagger a_1 \right)
+ \left( \frac{1}{2} Q_{22} a_2^\dagger a_2^\dagger a_2 - \frac{1}{2} Q_{22} a_2^\dagger a_2^\dagger a_2^\dagger a_2 \right)
\]

and

\[
V \left( a_1^\dagger a_1, a_2^\dagger a_2 \right)
= a_1^\dagger a_1 + 2ca_2 a_2 + a_2^\dagger a_2
+ \frac{1}{2} Q_{11} \left( a_1^\dagger a_1 (a_1^\dagger a_1) - a_1^\dagger a_1 \right)
+ \frac{1}{2} Q_{22} \left( a_2^\dagger a_2 (a_2^\dagger a_2) - a_2^\dagger a_2 \right).
\]
Note that we used \( a^\dagger a^\dagger a a = (a^\dagger a)(a^\dagger a) - a a \), which stems from \([a, a^\dagger] = aa^\dagger - a^\dagger a = 1\).

As shown in the previous work (Ohkubo 2013b), it is possible to interpret the linear operator \( L' \) in (30) as a time-evolution operator for a birth-death process. Because of the operator \( b \), we must consider an additional state variable, which takes only two states \((+ or -)\), in addition to the state variables \( n_0, n_1 \) and \( n_2 \).

See the first term in (30); the action of the first term on \([n]\) (27), and hence we can interpret this term as an elementary birth-death process \( X_1 \rightarrow X_2 \rightarrow X_0 \) with rate \( n_1 \), where \( n_0, n_1 \) and \( n_2 \) is the number of particles \( X_0, X_1 \), and \( X_2 \), respectively. On the other hand, the third term in (30) gives \( X_2 \rightarrow 2X_1 + X_2 + X_0 \) and the state change with \(+ \rightarrow - \) or \(- \rightarrow + \) with rate \( cn_2 \); the state \((+ or -)\) is changed due to an event corresponding to the third term. Repeating the similar discussions, finally the following birth-death process is obtained:

\[
\begin{align*}
(i) & \quad X_1 \xrightarrow{n_1} X_2 + X_0, \\
(ii) & \quad X_2 \xrightarrow{cn_2} X_2 + X_0, \\
(iii) & \quad X_2 \xrightarrow{cn_2} 2X_1 + X_2 + X_0 \text{ with S.C.,} \\
(iv) & \quad X_2 \xrightarrow{n_2} X_1 + X_0 \text{ with S.C.,} \\
(v) & \quad 2X_1 \xrightarrow{n_1(n_1-1)} \emptyset, \\
(vi) & \quad 2X_2 \xrightarrow{n_2(n_2-1)} \emptyset, 
\end{align*}
\]

where ‘S.C.’ means the state change \(+ \rightarrow - \) or \(- \rightarrow + \).

The term \( V(a_1^\dagger a_1, a_2^\dagger a_2) \) is called a Feynman-Kac term, and the important point is that this term is written only in terms of the number operators \( a_1^\dagger a_1 \). Since the number operators \( a_1^\dagger a_1 \) does not affect the state vectors, we can simply replace the number operators with the random variables in the birth-death process; i.e., \( V(a_1^\dagger a_1, a_2^\dagger a_2) = V(n_1, n_2) \).

The intuitive understanding of the duality relation is written as follows: we here abbreviate a state related to the Fokker-Planck equation at time \( t \) as \( FP(t) \) and to that of the birth-death process as \( BD(t) \). In addition, for simplicity, set \( r_{ts} = 1 \) and hence \( t = \tau \); furthermore, we neglect the Feynman-Kac term here. Then, formally, time-evolution of \( BD(t) \) is given by \( BD(t) = e^{L \tau} BD(0) \). In contrast, that of \( FP(t) \) is written as \( FP(t) = FP(0)e^{L \tau} \), when we consider the left-action of \( L \). Hence, the adjoint operator \( L^* \) gives the actual time-evolution of the Fokker-Planck equation. In addition, we have formally \( FP(t) BD(0) = FP(0)e^{L \tau} BD(0) = FP(0)BD(t) \); this corresponds to the duality relation. The linear operator \( L \) can be written in terms of both the differential operators and the creation-annihilation operators, and hence the stochastic differential equation (1) and the birth-death process (32) are connected naturally. Of course, the above discussion is just an intuitive one, and for the mathematical explanation of the duality relations, see the previous work (Ohkubo 2013b).

Here we consider a time-evolution from time \( 0 \) to \( \tau \). Because we consider the time-scaling factor \( r_{ts} \), the time interval, \( \tau \), in the stochastic differential equations correspond to the time interval \( \tilde{\tau} = \tau/r_{ts} \) in the dual birth-death process. Using the duality relation, we can obtain the following identity:

\[
\mathbb{E}_x \left[ x_1(\tau)^{n_1(0)} x_2(\tau)^{n_2(0)} \right] \\
= \mathbb{E}_{n,+} \left[ \exp \left\{ \int_0^\tau V(n_1(s), n_2(s)) ds \right\} \right. \\
\times r_{ts}^{n_2(\tilde{\tau})} x_1(0)^{n_1(\tilde{\tau})} x_2(0)^{n_2(\tilde{\tau})} \right] \\
- \mathbb{E}_{n,-} \left[ \exp \left\{ \int_0^\tau V(n_1(s), n_2(s)) ds \right\} \right. \\
\times r_{ts}^{n_1(\tilde{\tau})} x_1(0)^{n_1(\tilde{\tau})} x_2(0)^{n_2(\tilde{\tau})} \right], 
\]

where the abbreviations \( x = (x_1, x_2) \) and \( n = (n_0, n_1, n_2) \) are used; we set \( n_0(0) = 0 \); \( \mathbb{E}_{n,+} \) and \( \mathbb{E}_{n,-} \) are the conditional expectations for the states \(+ \) and \(- \) in the birth-death process (32), respectively. Using the duality relation (33), we can evaluate various information for the stochastic differential equation (1) from the solution of the birth-death process (32). In order to evaluate the probability distribution and the corresponding Feynman-Kac terms for the birth-death process (32), Monte Carlo simulations with the Gillespie algorithm are available.

As noted in Sec. 4.1, the initial conditions for the birth-death process correspond to the order of moments in the stochastic differential equation. When we use \( n_1(0) = 2 \) and \( n_2(0) = 0 \), the time-evolution of the second moment for \( x_1 \), i.e., \( \mathbb{E}_x [x_1(\tau)^2] \), is evaluated.

Using the above procedures, the time evolution of the stochastic differential equation for arbitrary initial conditions can be replaced with that of the dual birth-death process for specific initial conditions. In addition, if changing the time-scaling variable \( r_{ts} \), we can evaluate information for various time interval in the stochastic differential equations from only a result for a single fixed time interval, \( \tilde{\tau} \), in the birth-death process. For example, assume that \( \tilde{\tau} = 1 \); if we want to know the information of the stochastic differential equation at time \( t = 1 \), we should set \( r_{ts} = 1 \). On the other hand, the information at time \( t = 0.9 \) can be evaluated from the same results of the birth-death process with \( \tilde{\tau} = 1.0 \) and setting \( r_{ts} = 0.9 \). Hence, there is no need to perform the Monte Carlo simulations for various different
time intervals; this is also one of the advantages of the new method.

4.4 Moment evaluation

The duality relation in (33) can be used for evaluating the moments in stochastic differential equation (1) starting from the Dirac-delta-type initial conditions, as discussed in Sec. 4.1. However, in the EnKF, we use an ensemble of samples. In the EnKF, only the mean and covariance matrix of the ensemble are needed, so we can consider that the ensemble is essentially characterized by a Gaussian distribution. Hence, the Dirac-delta-type initial conditions are not enough.

We here write the expectation value, which is taken by using a Gaussian distribution with mean \(\langle \cdots \rangle\) and covariance \(V\), as \(\langle \cdots \rangle\). Hence, the first term in the r.h.s. in (33) should be replaced with

\[
\langle E_{n,+} \left[ e^{\int_{0}^{\tau} V ds} r_{ts}^{n_0(\tau)} x_1(0)^{n_1(\tau)} x_2(0)^{n_2(\tau)} \right] \rangle = \langle E_{n,+} \left[ e^{\int_{0}^{\tau} V ds} r_{ts}^{n_0(\tau)} x_1(0)^{n_1(\tau)} x_2(0)^{n_2(\tau)} \right] \rangle.
\]

We need a little additional calculations in order to evaluate the expectation with the Gaussian distribution; introducing the following notations,

\[
\mu_1 = \langle x_1 \rangle, \mu_2 = \langle x_2 \rangle,
\]

it is possible to obtain the following recursion formula (Willink 2005):

\[
\langle x_1^{n_1} x_2^{n_2} \rangle = \mu_1 \langle x_1^{n_1-1} x_2^{n_2} \rangle + V_{11}(n_1-1) \langle x_1^{n_1-2} x_2^{n_2} \rangle + V_{12}(n_2) \langle x_1^{n_1-1} x_2^{n_2-1} \rangle
\]

or

\[
\langle x_1^{n_1} x_2^{n_2} \rangle = \mu_2 \langle x_1^{n_1} x_2^{n_2-1} \rangle + V_{21}(n_1) \langle x_1^{n_1-1} x_2^{n_2-1} \rangle + V_{22}(n_2-1) \langle x_1^{n_1} x_2^{n_2-2} \rangle.
\]

Using the above recursion formula, once we have the probability distribution \(P(n_0, n_1, n_2, t)\) for the birth-death process (32) with adequate initial values, it is possible to evaluate various moments for the stochastic differential equation (1) with a Gaussian initial distribution, using the following duality relation:

\[
\begin{align*}
\mathbb{E}_{x, \text{Gaussian initial}} \left[ x_1(\tau)^{n_1(\tau)} x_2(\tau)^{n_2(\tau)} \right] &= \mathbb{E}_{n,+} \left[ \exp \left\{ \int_{0}^{\tau} V(n_1(s), n_2(s)) ds \right\} \right. \\
&\times r_{ts}^{n_0(\tau)} \langle x_1(0)^{n_1(\tau)} x_2(0)^{n_2(\tau)} \rangle \\
&\left. - \mathbb{E}_{n,-} \left[ \exp \left\{ \int_{0}^{\tau} V(n_1(s), n_2(s)) ds \right\} \right. \right. \\
&\times r_{ts}^{n_0(\tau)} \langle x_1(0)^{n_1(\tau)} x_2(0)^{n_2(\tau)} \rangle \right].
\end{align*}
\]

4.5 Algorithm

We here consider general cases with variable time intervals, and denote the maximum of the time intervals, \(\tau_{k-1} - \tau_k\), as \(\Delta \tau\). The new Kalman filter based on the duality relation, called the DuKF here, is as follows:

**Algorithm for DuKF**

1. Preparation for the duality relations:
   - Simulate the dual birth-death process (32). We need simulations with five different initial conditions;
     - (c1) \(n_0 = 0, n_1 = 1, n_2 = 0\),
     - (c2) \(n_0 = 0, n_1 = 0, n_2 = 1\),
     - (c3) \(n_0 = 0, n_1 = 2, n_2 = 0\),
     - (c4) \(n_0 = 0, n_1 = 0, n_2 = 2\),
     - (c5) \(n_0 = 0, n_1 = 1, n_2 = 1\),
   - which are necessary to evaluate \(\mathbb{E}_x[x_1(t)], \mathbb{E}_x[x_2(t)], \mathbb{E}_x[x_1(t)^2], \mathbb{E}_x[x_2(t)^2], \mathbb{E}_x[x_1(t)x_2(t)]\), respectively. For all cases, the additional state variable is set to ‘+’ initially. The Monte Carlo simulations are performed from \(t = 0\) to \(t = \Delta \tau\), and the integral of the Feynman-Kac term \(V(n_1, n_2)\) and the final probability distribution \(P(n_0, n_1, n_2, \Delta \tau)\) are evaluated numerically.

2. Initialization:
   - Set an initial Gaussian distribution with mean \(\mathbf{x}(0)\) and covariance matrix \(P(0)\).

3. Forecast step at \(\tau_{k-1}\):
   - Using the duality relation (37), evaluate various moments, which are necessary to characterize a Gaussian distribution. The ensemble average in (37), \(\langle \cdots \rangle\), is taken for the Gaussian distribution with mean \(\mathbf{x}(\tau_{k-1})\) and covariance matrix \(P(\tau_{k-1})\). The above procedure gives the mean \(\mathbf{x}(\tau_k)\) and the covariance \(P(\tau_k)\) of the nonlinear systems at time \(\tau_k\). Note that the time-scaling variable \(r_{ts}\) must be selected adequately according to the ratio between \(\tau_k - \tau_{k-1}\) and \(\Delta \tau\).
4. Assimilation step at $\tau_k$:
Calculate the Kalman gain
\[
\hat{K}(\tau_k) = \tilde{P}^f(\tau_k)H^T \left( H\tilde{P}^f(\tau_k)H^T + R \right)^{-1},
\]
and update the following quantities:
\[
\tilde{P}^a(\tau_k) = \tilde{P}^f(\tau_k) + \hat{K}(\tau_k)(g(\tau_k) - H\tilde{P}^f(\tau_k))
\]
and
\[
P^a(\tau_k) = (1 - \hat{K}(\tau_k)H)\tilde{P}^f(\tau_k).
\]
Steps 3 and 4 are performed for each measurement time step.

5 Results

In order to demonstrate the DuF, we employ the algorithm in Sec. 4.5 to the problem in Sec. 2. For simplicity, we assume here that the parameters in the stochastic differential equations, $Q$ and $R$, are previously known, as explained in Sec. 2. We use the following parameters for the initial Gaussian distribution; $\mathbf{x}(0) = (0.1, 0.1)$, $[P(0)]_{11} = 0.1$, $[P(0)]_{12} = [P(0)]_{21} = 0$, and $[P(0)]_{22} = 0.1$. Firstly, for the DuKF, we performed the Monte Carlo simulations for the dual birth-death process using the Gillespie algorithm (Gillespie 1977). For each initial condition, $10^{12}$ sample paths were generated, and the Feynman-Kac term and the probability distribution were evaluated.

Figure 2 shows the results of the state estimation. Note that only the state variable $x_2$ is observed at discrete times, as depicted in Fig. 1. Although the observation contains the measurement noises, as seen in Fig. 1, both the EnKF and DuKF give adequate estimations. Especially, the non-observed state, $x_1$, can be estimated reasonably, as shown in Fig. 2. From Fig. 2, it may be difficult to judge whether the DuKF gives better results than the EnKF or not, but we confirmed that the DuKF gives a slightly better mean squared error. If we want to obtain the similar mean squared error, we need larger ensemble for the EnKF, and the larger ensemble needs more computational time. On the other hands, the forecast step in the DuKF does not need any Monte Carlo simulation, and hence the DuKF works rapidly.

In order to see the difference between the EnKF and DuKF more explicitly, we show the covariance $[P(t)]_{12}$ calculated in the EnKF and DuKF in Fig. 3. As shown in Fig. 3, the larger ensemble size in the EnKF gives the similar covariance with that of the DuKF. The covariance is used to evaluate the Kalman gain, and hence more precise covariance is necessary to have better es-
6 Concluding remarks

The duality relation between stochastic processes has been still developing, and hence it is expected that new applications of the duality relations give completely novel algorithms for various research fields. We demonstrate one of the applications to the filtering problem. As a result, preliminary Monte Carlo simulations for the dual birth-death processes enable us to avoid the time-consuming Monte Carlo simulations for each forecast step, as in the EnKF.

Of course, we do not claim that the proposed DuKF is always superior to EnKF for any filtering problem. Actually, the implementation of the EnKF is simpler than that of the DuKF. We imagine that the DuKF would be suitable for a kind of embedded systems or chips, in which only simple numerical calculations are performed. As demonstrated in the present paper, the DuKF does not require any Monte Carlo simulation in the filtering steps; it is enough to perform the Monte Carlo simulations in high-performance computers in advance, and to preserve only the probability distribution and the Feynman-Kac terms on the memory in the embedded systems.

In order to use the DuKF for practical purposes, more studies are necessary, but we expect that the present paper opens a new way in the filtering problems. Finally, we give some comments for the practical aspects. In order to construct the DuKF, very precise probability distributions for the dual birth-death processes are needed. In our demonstrations, the probability distributions and the Feynman-Kac terms were evaluated from the Monte Carlo simulations with \(10^{12}\) sample paths. Actually, simulations with \(10^{10}\) sample paths were not enough; imprecise probability distributions give terrible estimations for the moments, and the filtering results become disastrous. This is the reason why we set the time interval for the measurement \(\Delta t_{\text{obs}} = 0.2\) in our demonstrations; \(10^{12}\) sample paths are not enough for larger time intervals. In order to avoid this problem, we will employ simulations based on the important sampling methods (Asmussen and Glynn 2007; Landau and Binder 2009). It is beyond the scope of the present paper to develop efficient important sampling methods for the birth-death processes. In addition, as in the same reason above, the time scaling variable \(r_{ts}\) should be \(0 \leq r_{ts} \leq 1\) in practice.

Acknowledgements This work was supported in part by grant-in-aid for scientific research (No. 25870339) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

References

Asmussen, S., Glynn, P.W.: Stochastic simulation. Springer, New York (2007)
Candy, J.V.: Bayesian signal processing: classical, modern, and particle filtering methods. Wiley, Hoboken, N.J. (2009)
Doi, M.: Second quantization representation for classical many-particle system. J. Phys. A: Math. Gen. 9, 1465–1477 (1976a)
Doi, M.: Stochastic theory of diffusion-controlled reaction. J. Phys. A: Math. Gen. 9, 1479–1495 (1976b)
Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99, 10143–10162 (1994)
Evensen, G.: The ensemble Kalman filter for combined state and parameter estimation. IEEE Control Syst. 29, 83–104 (2009)
Gardiner, C.: Stochastic methods, 4th ed. Springer, Heidelberg (2009)
Giardina, C., Kurchan, J., Redig, F., Vafayi, K.: Duality and hidden symmetries in interacting particle systems. J. Stat. Phys. 135, 25–55 (2009)
Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81, 2340–2361 (1977)
Gillijns, S., Mendoza, O.B., Chandrasekar, J., De Moor, B.L.R., Bernstein, D.S., Ridley, A.: What is the ensemble Kalman filter and how well does it work? Proc. 2006 American Control Conference (2006). doi:10.1109/ACC.2006.1657419
Gordon, N.J., Salmond D.J., Smith, A.F.M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. Radar and Signal Process. 140, 107–113 (1993)
Hora, A., Obata, N.: Quantum probability and spectral analysis of graphs. Springer, Berlin (2007)
Kitagawa, G.: Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. J. Compt. Graph. Stat. 5, 1–25 (1996)
Kloeden, P.E., Platen, E.: Numerical solution of stochastic differential equations. Springer, Berlin (1992)
Lakshmivarahan, S., Stensrud, D.J.: Ensemble Kalman filter. IEEE control syst. 29, 34–46 (2009)
Landau, D.P., Binder, K.: A guide to Monte Carlo simulations in statistical physics, 3rd ed. Cambridge University Press, Cambridge (2009)
Liggett, T.M.: Interacting particle systems (Classics in Mathematics), Springer, Berlin (2005)
Ohkubo, J.: Duality in interacting particle systems and boson representation. J. Stat. Phys. 139, 454–465 (2010)
Ohkubo, J.: Algebraic probability, classical stochastic processes, and counting Statistics. J. Phys. Soc. Jpn. 82, 084001 (2013a)

Ohkubo, J.: Extended duality relations between birth-death processes and partial differential equations. J. Phys. A: Math. Theor. 46, 375004 (2013b)

Peliti, L.: Path integral approach to birth-death processes on a lattice. J. Physique 46, 1469–1483 (1985)

Täuber, U.C., Howard, M., Vollmayr-Lee, B.P.: Applications of field-theoretic renormalization group methods to reaction-diffusion problems. J. Phys. A: Math. Gen. 38, R79-R131 (2005)

Willink, R.: Normal moments and Hermite polynomials. Stat. Prob. Lett. 73, 271–275 (2005)
