Bayesian inversion analysis of nonlinear dynamics in surface heterogeneous reactions

Toshiaki Omori,[1,][*] Tatsu Kuwatani,[2] Atsushi Okamoto,[3] and Koji Hukushima[4,5]

[1]*Department of Electrical and Electronic Engineering, Graduate School of Engineering, Kobe University,*
1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

[2]*Department of Solid Earth Geochemistry, Japan Agency for Marine-Earth Science and Technology,*
2-15 Natsushima-cho, Yokosuka 237-0061, Japan

[3]*Department of Environmental Studies for Advanced Society, Graduate School of Environmental Studies,*
Tohoku University, 6-6-20 Aramaki, Aoba-ku, Sendai 980-8579, Japan

[4]*Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo,*
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

[5]*Center for Materials Research by Information Integration, National Institute for Materials Science, 1-2-1 Sengen, Tsukuba 305-0047, Japan*

It is essential to extract nonlinear dynamics from time-series data as an inverse problem in natural sciences. We propose a Bayesian statistical framework for extracting nonlinear dynamics of surface heterogeneous reactions from sparse and noisy observable data. Surface heterogeneous reactions are chemical reactions with conjugation of multiple phases, and they have the intrinsic nonlinearity of their dynamics caused by the effect of surface-area between different phases. We adapt a belief propagation method and an expectation-maximization (EM) algorithm to partial observation problem, in order to simultaneously estimate the time course of hidden variables and the kinetic parameters underlying dynamics. The proposed belief propagation method is performed by using sequential Monte Carlo algorithm in order to estimate nonlinear dynamical system. Using our proposed method, we show that the rate constants of dissolution and precipitation reactions, which are typical examples of surface heterogeneous reactions, as well as the temporal changes of solid reactants and products, were successfully estimated only from the observable temporal changes in the concentration of the dissolved intermediate product.

I. INTRODUCTION

It is essential to extract dynamics from the observational time-series data in natural sciences. In general, the dynamics is nonlinear, which makes the inversion analysis difficult. Furthermore, observable data are restricted to only a part of changing variables, and the observed time-series data are generally sparse and noisy. It is important to develop a versatile method that can extract nonlinear dynamics directly from such incomplete and nonlinear data sets.

Kinetic parameters of chemical reactions, such as reaction-rate coefficients and diffusion coefficients, govern the nonlinear dynamical behavior of chemical systems. They are usually determined from observed time evolution of the amount of chemical species by laboratory experiments in various natural sciences [1,2]. In laboratory experiments, it is often the case that only a part of concentrations of chemical species among many species are quantitatively observable, and the observed time-series data are generally sparse and noisy [3–5].

We develop novel methodologies for estimating kinetic parameters from nonlinear dynamics of chemical reactions by using a Bayesian probabilistic approach. With an increasing demand for data-driven analysis, the Bayesian estimation has been widely applied to various fields in the natural sciences such as physics [6], brain science [7–11], and earth sciences

[12–14]. We formulate probabilistic models of the nonlinear dynamics of heterogeneous systems based on a nonlinear state space model. The sequential Monte Carlo method [15–19] and expectation-maximization (EM) algorithms [20–23] are employed to simultaneously estimate the time course of chemical systems and the kinetic parameters underlying their nonlinear dynamics. In particular, we develop a belief propagation method [24,25] for partially observable data in order to obtain the marginal posterior distribution of hidden variables from sparse and noisy data.

In this study, we focus the nonlinear dynamics of a surface heterogeneous reaction. In general, chemical reactions can be divided into two types: homogeneous reactions, which occur in a single phase, and heterogeneous reactions, which occur at the interface among two or more phases. Many of the chemical reactions important in natural sciences are heterogeneous reactions [2]. Heterogeneous reactions have an intrinsic nonlinearity of their dynamics caused by the effect of surface area; reaction rates are proportional to the reactive surface areas of phases, which change temporally according to the increase or decrease of the corresponding phases. This causes the reaction rates to change nonlinearly according to surface geometry. Moreover, by laboratory experiments, it is very difficult to measure even static surface areas, not to mention the temporal change of surface areas [1,2]. Therefore, it is difficult to determine the kinetic parameters in surface heterogeneous reactions due to their complexity, compared to those in homogeneous reactions.

Using our proposed method, the reaction rate constants of dissolution and precipitation as well as the temporal changes of solid reactants and products were successfully estimated from only the observable time-series concentration data of the dissolved intermediate product. The proposed method is

---

*[*]omori@eedept.kobe-u.ac.jp

also applicable in actual laboratory experiments investigating heterogeneous kinetics in various fields of natural sciences, including geosciences.

This paper is organized as follows. In Sec. II, we formulate a generative model of surface heterogeneous reactions by using a nonlinear state-space model and derive an estimation method by adapting belief propagation and EM algorithm to inverse problem for the surface heterogeneous reactions using sparse and noisy time-series data. In Sec. III, the validity of proposed method is evaluated by using simulated data. We show that hidden variables of the surface heterogeneous reaction are successfully estimated from observable data. Furthermore, kinetic rate constants underlying the nonlinear dynamics of the heterogeneous reaction are shown to be estimated accurately using our proposed method with an adapted EM algorithm. Moreover, the robustness and effectiveness of the proposed method for sparse and noisy data are also shown. Concluding remarks are given in Sec. IV.

II. ESTIMATION ALGORITHM

We focus on a simple surface heterogeneous reaction, called the dissolution-precipitation process. The dissolution and precipitation process can be regarded as one of the most fundamental and substantial surface heterogeneous reactions for solid-liquid interactions. For example, dissolution-precipitation processes of minerals play important roles in geoscientific and environmental phenomena in the earth's crust [2–5,26–28]. In laboratory experiments investigating solid-liquid interactions, it is often the case that only concentrations of intermediate products dissolved in liquid phase are quantitatively observable, and the observed time-series data are generally sparse and noisy [3–5].

Figure 1 shows the target dissolution-precipitation process governed by the surface heterogeneous reaction, in which the solid reactant r changes to a solid product p via an intermediate product i dissolved in a liquid phase:

$$\text{r} \xrightarrow{k_\text{r}} \text{i} \xrightarrow{k_\text{p}} \text{p},$$
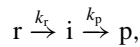
where k_r and k_p are rate constants. In this section, we propose a statistical method for inverse estimation of the dynamics underlying the surface heterogeneous reactions from partially observable noisy data: the restricted number of observations of the concentration of the intermediate product.

We first describe the kinetic equations for the dissolution-precipitation process, and then formulate a state-space model of the process. To reflect realistic experimental environments, we assume in the state-space model that only the concentration of intermediate product dissolved in liquid phase, $C^{(\text{i})}$, can be observed in noisy measurements, and that neither the solid reactant $n^{(\text{r})}$ nor the solid product $n^{(\text{p})}$ can be observed directly. Note that these notations are used to show the difference between variables in the solid phase ($n^{(\text{r})}$ and $n^{(\text{p})}$: molar amount) and the variable in the liquid phase ($C^{(\text{i})}$, concentration).

A. Formulation of kinetics for surface heterogeneous reaction

We consider a dissolution-precipitation process from a solid reactant, r, to a solid product, p, via intermediate product, i,
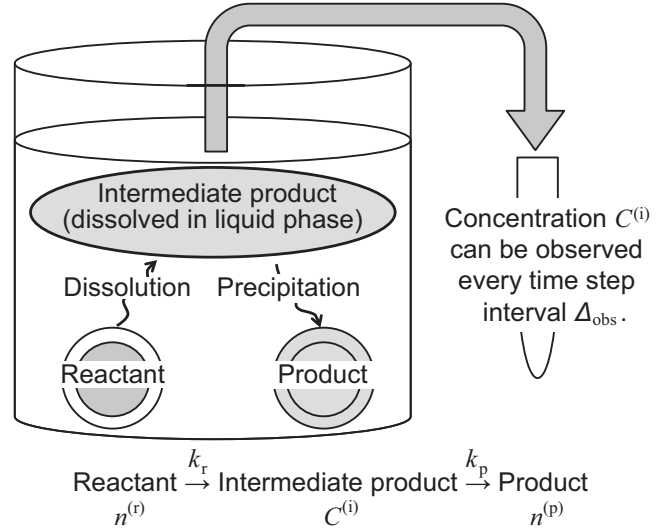

FIG. 1. Target experimental system for surface heterogeneous reactions. A solid product $n^{(\text{p})}$ is generated from a solid reactant $n^{(\text{r})}$ via an intermediate product $C^{(\text{i})}$. The reactions $n^{(\text{r})} \rightarrow C^{(\text{i})}$ and $C^{(\text{i})} \rightarrow n^{(\text{p})}$ depend on the surface areas of the solid reactant $S_\text{r}(n^{(\text{r})})$ and the solid product $S_\text{p}(n^{(\text{p})})$, respectively. In the proposed method, the rate constants k_r and k_p underlying surface heterogeneous reactions between reactant $n^{(\text{r})}$ and product $n^{(\text{p})}$ via intermediate product $C^{(\text{i})}$ are estimated by using only the intermediate product concentration $C^{(\text{i})}$ observed every time step interval Δ_obs.

dissolved in liquid phase (Fig. 1). For mathematical simplicity and general applicability, we consider simple first-order reactions with single component and multiple phases, in which reaction rates of solid phase per unit surface area are proportional to the difference between the present concentration of intermediate product ($C^{(\text{i})}$) and the equilibrium concentration with the solid reactant or the solid product ($C_{\text{r,eq}}$ or $C_{\text{p,eq}}$).

By reflecting that the reaction rates of surface heterogeneous reactions depend on the surface areas of the solids, we assume that the kinetics of the solid reactant $n^{(\text{r})}$ and product $n^{(\text{p})}$ obeys the following differential equations for first-order reactions:

$$\frac{dn^{(\text{r})}}{dt'} = k_\text{r} S_\text{r}(n^{(\text{r})})(C^{(\text{i})} - C_{\text{r,eq}}) + \xi_\text{r}(t'), \quad (1)$$

$$\frac{dn^{(\text{p})}}{dt'} = k_\text{p} S_\text{p}(n^{(\text{p})})(C^{(\text{i})} - C_{\text{p,eq}}) + \xi_\text{p}(t'), \quad (2)$$

where t' denotes time. $S_\text{r}(n^{(\text{r})})$ and $S_\text{p}(n^{(\text{p})})$ are the surface areas of the solid reactant and the solid product, respectively. k_r and k_p are reaction coefficients. $\xi_\text{r}(t')$ and $\xi_\text{p}(t')$ are fluctuation terms which obey, for any times t' and s', $\langle \xi_\text{r}(t') \rangle = \langle \xi_\text{p}(t') \rangle = 0$, $\langle \xi_\text{r}(t')\xi_\text{r}(s') \rangle = \tilde{\sigma}_\text{r}^2 \delta(t' - s')$, $\langle \xi_\text{p}(t')\xi_\text{p}(s') \rangle = \tilde{\sigma}_\text{p}^2 \delta(t' - s')$, and $\langle \xi_\text{r}(t')\xi_\text{p}(s') \rangle = 0$, where standard deviations of the fluctuation terms are denoted by $\tilde{\sigma}_\text{r}$ and $\tilde{\sigma}_\text{p}$. The parameters $\tilde{\sigma}_\text{r}$ and $\tilde{\sigma}_\text{p}$ are noise intensities that come from the randomness of the heterogeneous system itself, and the intensities are typically small. In general, the factors $S_\text{r}(n^{(\text{r})})$ and $S_\text{p}(n^{(\text{p})})$ should be zero for $n^{(\text{r})} = 0$ and $n^{(\text{p})} = 0$, respectively. Note that in the case of no fluctuating noise this system has a steady state at $(n^{(\text{r})}, C^{(\text{i})}) = (0, C_{\text{p,eq}})$ that is accessible from the initial condition and has an unphysical steady state $(n^{(\text{p})}, C^{(\text{i})}) = (0, C_{\text{r,eq}})$.
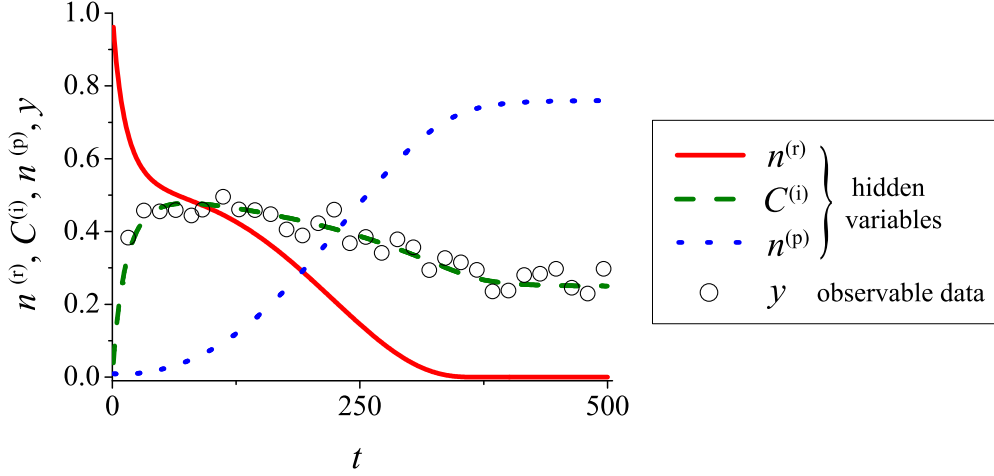
FIG. 2. Typical time course of surface heterogeneous reaction and noisy observable data. The solid product $n^{(p)}$ (blue dotted line) is generated from the solid reactant $n^{(r)}$ (red solid line) through the intermediate product dissolved in the liquid phase, $C^{(i)}$ (green dashed line). Sparse and noisy data of the intermediate product y (circles) are observed with low time-resolution.

According to conservation of mass, the intermediate product dissolved in liquid phase, $C^{(i)}$, obeys the following equation:

$$n^{(r)} + v C^{(i)} + n^{(p)} = \text{const.}, \qquad (3)$$

where v denotes the solution volume. Hereafter, we set $v = 1$ without loss of generality.

Here we assume that the rate-determining process is a surface reaction. This assumption is reasonable in laboratory experiments, which we consider here, and holds when diffusion and advection of elements in liquid phase are so fast so that the effect of transport can be negligible. As shown in Fig. 1, the rate of dissolution and precipitation reactions generally depend on surface area of solid phase, since the dissolution and precipitation reactions occur on the surface between solid and liquid phases [2]. Namely, $k_r S_r(n^{(r)})$ and $k_p S_r(n^{(p)})$ are net reaction rates depending on solid reactant and solid product, respectively. Nonlinearity emerges in Eqs. (1) and (2) since the surface area depends on hidden variable $n^{(r)}$ or $n^{(p)}$.

The reactive surface areas for the solid reactant and product in Eqs. (1) and (2) are assumed to be functions of the amounts of the corresponding substances:

$$S_r(n^{(r)}) = A_r (n^{(r)})^{\alpha_r}, \qquad (4)$$

$$S_p(n^{(p)}) = A_p (n^{(p)})^{\alpha_p}, \qquad (5)$$

where A_r and A_p are geometrical constants, and α_r and α_p are the orders of the surface model that depend on the geometrical configuration of solid substances [2,26]. For example, the order of the surface model $\alpha \in \{\alpha_r, \alpha_p\}$ is 0 if total surface area is constant regardless of the amount of the solid substance. The order α is 1 if the total surface area is linearly proportional to the amount of the solid, which is the case if the number of particles of corresponding solid increases or decreases without changing the surface area of each particle. In the case that there is each particle grows or dissolves while keeping a similar shape without changing the number of total particles, the order α becomes 2/3.

Figure 2 shows a typical behavior of the kinetics of a surface heterogeneous reaction, for which the order of the surface model is 2/3. The stochastic differential equations [Eqs. (1) and (2)] are numerically simulated with time step t where time t' is discretized with time width $\Delta_t = t'/t = 8$ throughout this study. When in the initial condition there exists only the solid reactant ($n^{(r)} = 1$ and $C^{(i)} = n^{(p)} = 0$), the solid reactant $n^{(r)}$ rapidly decreases while the intermediate product $C^{(i)}$ rapidly increases. Following these changes, the solid product $n^{(p)}$ slowly increases and converges to a positive value while the solid reactant $n^{(r)}$ converges to zero. This transient response substantially depends on nonlinearity due to the temporal change of the surface area.

B. Nonlinear state-space model for surface heterogeneous reactions

In order to establish a method for estimating the nonlinear dynamics of surface heterogeneous reactions from partially observable noisy data, we formulate a probabilistic model of the nonlinear dynamics of a surface heterogeneous reaction based on a nonlinear state-space model.

Figure 3 shows a graphical model of the proposed probabilistic model. The true states of the three factors in surface heterogeneous reactions, $n^{(r)}(t')$, $C^{(i)}(t')$, and $n^{(p)}(t')$, are assumed to be hidden variables. We consider the time series of states at discretized time $t = t'/\Delta_t$ as $\mathbf{x}_t = \{n_t^{(r)}, C_t^{(i)}, n_t^{(p)}\}$ ($t = 1, \cdots, T$), where Δ_t denotes the time interval for discretization.

Within these three factors, only the concentration of intermediate substance $C_t^{(i)}$ is assumed to be observed as follows:

$$y_t = C_t^{(i)} + \xi_t^{(y)}, \qquad (6)$$

where $\xi_t^{(y)}$ denotes an observation noise obeying a white Gaussian noise with $\langle \xi_t^{(y)} \rangle = 0$ and $\langle \xi_t^{(y)} \xi_s^{(y)} \rangle = \sigma_y^2 \delta(t - s)$. Here σ_y denotes the standard deviation of the observation noise. Furthermore, we assume that the observed intermediate substance y_t is available at restricted times.

Here we formulate a nonlinear state-space model of surface heterogeneous reactions. The nonlinear state-space model

consists of two models: a system model and an observation model.

The system model describes the relation between hidden variables at two sequential time steps. Due to the mass conservation relation among the three factors $n_t^{(r)}$, $C_t^{(i)}$, and $n_t^{(p)}$ [Eq. (3)], the kinetics governed by Eqs. (1) and (2) has two

$$
\begin{aligned}
p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) = {} & \frac{1}{\sqrt{2\pi\sigma_r^2}} \exp\left[-\frac{\left(n_{t+1}^{(r)} - n_t^{(r)} - \Delta_t k_r S_r\!\left(n_t^{(r)}\right)\!\left(C_t^{(i)} - C_{r,\mathrm{eq}}\right)\right)^2}{2\sigma_r^2} \right] \\
& \times \frac{1}{\sqrt{2\pi\sigma_p^2}} \exp\left[-\frac{\left(n_{t+1}^{(p)} - n_t^{(p)} - \Delta_t k_p S_p\!\left(n_t^{(p)}\right)\!\left(C_t^{(i)} - C_{p,\mathrm{eq}}\right)\right)^2}{2\sigma_p^2} \right],
\end{aligned}
\tag{7}
$$

where Δ_t denotes the time interval used for discretization of the kinetics. Note that we put $\sigma_r = \sqrt{\Delta_t}\,\tilde{\sigma}_r$ and $\sigma_p = \sqrt{\Delta_t}\,\tilde{\sigma}_p$. The derivation of the probabilistic model for the system model [Eq. (7)] is given in Appendix A.

The observation model describes the relation between hidden variables and observable data. Based on Eq. (6), the observation process is described by the following probabilistic density function:

$$
p(y_t|\boldsymbol{x}_t) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left[-\frac{\left(y_t - C_t^{(i)}\right)^2}{2\sigma_y^2} \right] \quad \text{for } t \in \mathcal{T}_{\mathrm{obs}},
\tag{8}
$$

where $\mathcal{T}_{\mathrm{obs}}$ expresses the set of all observation times.

C. Belief propagation method

According to the Bayes' theorem, a posterior distribution $p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T|Y_T)$ is proportional to a likelihood function $p(Y_T|\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)$ and a prior distribution $p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)$
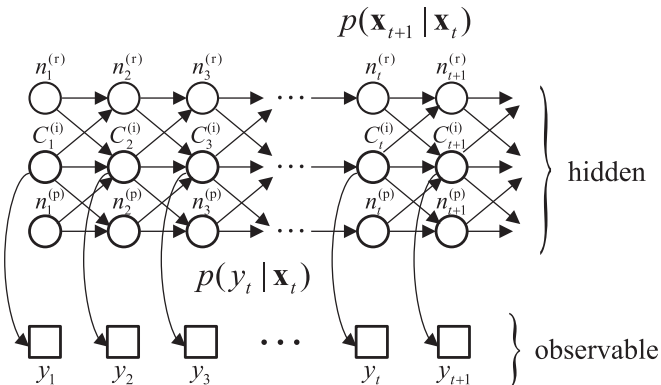

FIG. 3. Graphical model of the proposed nonlinear state-space model for surface heterogeneous reactions. Time evolution of three hidden variables $\boldsymbol{x}_t = \{n_t^{(r)}, C_t^{(i)}, n_t^{(p)}\}$ at time step t is described by the system model $p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$, whereas the observation process from hidden variables $\boldsymbol{x}_t$ to observed data y_t can be described by the observation model $p(y_t|\boldsymbol{x}_t)$. Within three hidden variables $\boldsymbol{x}_t = \{n_t^{(r)}, C_t^{(i)}, n_t^{(p)}\}$, only noisy data of intermediate product, y_t, is assumed to be observed.

independent variables. Hereafter, we express hidden variables $\boldsymbol{x}_t$ by a two-dimensional vector $\boldsymbol{x}_t = \{n_t^{(r)}, n_t^{(p)}\}$. By discretizing the differential equations [Eqs. (1) and (2)] with respect to time, the state transition probability $p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$ $(t = 1,\ldots,T-1)$ for the kinetics in surface heterogeneous reactions is expressed as follows:

as follows:

$$
\begin{aligned}
& p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T|Y_T) \\
& \quad \propto p(Y_T|\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)\,p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T).
\end{aligned}
\tag{9}
$$

Here $Y_t = \{y_1,\ldots,y_t\}$ is a set of observed data up to time t. The posterior probability density function $p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T|Y_T)$ is shown to be expressed by using the system model $p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$ and observation model $p(y_t|\boldsymbol{x}_t)$ as follows:

$$
\begin{aligned}
& p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T|Y_T) \\
& = \frac{1}{Z_T}\left(\prod_{t \in \mathcal{T}_{\mathrm{obs}}} p(y_t|\boldsymbol{x}_t)\right)\left(\prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)\right)p(\boldsymbol{x}_1),
\end{aligned}
\tag{10}
$$

where we used the independent property of observation in the likelihood function $p(Y_T|\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)$ and the Markov property of prior distribution $p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T)$ in the nonlinear state-space model. A detailed derivation of Eq. (10) is given in Appendix B. In Eq. (10), the prior distribution is expressed by the factor $p(\boldsymbol{x}_1)\prod_{t=2}^{T} p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$, whereas the likelihood is expressed by the factor $\prod_{t \in \mathcal{T}_{\mathrm{obs}}} p(y_t|\boldsymbol{x}_t)$. Z_T is a normalization factor called evidence in Bayesian statistical framework and is equivalent to the partition function in statistical physics:

$$
Z_T = \int \prod_{t=1}^{T} d\boldsymbol{x}_t \left(\prod_{t \in \mathcal{T}_{\mathrm{obs}}} p(y_t|\boldsymbol{x}_t)\right)\left(\prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)\right)p(\boldsymbol{x}_1).
\tag{11}
$$

A detailed derivation of the posterior probability density function is given in Appendix B.

In order to estimate hidden variables $\boldsymbol{x}_t$ $(t = 1,\ldots,T)$ from partially observable data Y_T, we derive marginalized posterior distributions $p(\boldsymbol{x}_t|Y_t)$ and $p(\boldsymbol{x}_{t+1}|Y_t)$. In general, marginalization corresponds to multidimensional ($2T$ dimensional) integration of $p(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_T|Y_T)$. Due to the Markov properties of the nonlinear state space model, the belief propagation method can be employed to perform this multidimensional integration, iteratively [24,25,29–31].

Here, the belief propagation is extended in the proposed method in order to consider partial observation process. We consider two cases of observations: the one for an unobservable data point and the other for an observable data point. Detailed derivations of belief propagation algorithm

for the partial observation processes in these two cases are given in Appendix C. In the first case where we do not have any observation data point at time t, a recursive equation for predictive distribution $p(\boldsymbol{x}_{t+1}|Y_t)$ is derived as follows:

$$
\begin{aligned}
p(\boldsymbol{x}_{t+1}|Y_t) &= \int d\boldsymbol{x}_t \, p(\boldsymbol{x}_{t+1}, \boldsymbol{x}_t|Y_t) \\
&= \int d\boldsymbol{x}_t \, p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) p(\boldsymbol{x}_t|Y_{t-1}).
\end{aligned} \tag{12}
$$

Note that the predictive distribution $p(\boldsymbol{x}_{t+1}|Y_t)$ at time $t+1$ on the left-hand side of Eq. (12) is recursively expressed by the one $p(\boldsymbol{x}_t|Y_{t-1})$ at the preceding time t on the right-hand side. In the second case, where we have an observable data point at time t, we consider two probability distributions: predictive distribution and filtering distribution [15–17]. The predictive distribution $p(\boldsymbol{x}_{t+1}|Y_t)$ of hidden variables $\boldsymbol{x}_{t+1}$ at time $t+1$, given the data Y_t up to preceding time t, is derived as follows:

$$
p(\boldsymbol{x}_{t+1}|Y_t) = \int d\boldsymbol{x}_t \, p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) p(\boldsymbol{x}_t|Y_t), \tag{13}
$$

whereas the filtering distribution $p(\boldsymbol{x}_t|Y_t)$ of hidden variables $\boldsymbol{x}_t$ at time t given the data Y_t up to the same time t is derived as the following expression:

$$
p(\boldsymbol{x}_t|Y_t) = \frac{p(y_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|Y_{t-1})}{p(y_t|Y_{t-1})}. \tag{14}
$$

Here, $p(y_t|Y_{t-1})$ denotes the distribution for observation data expressed as follows:

$$
p(y_t|Y_{t-1}) = \int d\boldsymbol{x}_t \, p(y_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|Y_{t-1}). \tag{15}
$$

Note that the predictive distribution $p(\boldsymbol{x}_t|Y_{t-1})$ and the filtering distribution $p(\boldsymbol{x}_t|Y_t)$ are obtained iteratively from $t=2$ to $t=T$, while substituting the system model $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ and observation model $p(y_t|\boldsymbol{x}_t)$ into these recursive update equations.

As shown in this subsection, the belief propagation algorithm derives the recursive update equations for probabilistic models, corresponds to the transfer matrix method in statistical physics, and realizes the integration for high-dimensional random variables in probabilistic models without loops [24,25,29–31]. That is, this multidimensional integration is

reduced to a multiple number of two-dimensional integrations by means of belief propagation.

D. Estimation of hyperparameters

We employ an EM algorithm [21–23] to estimate the parameters of the nonlinear state-space model. The EM algorithm is a statistical method used to estimate parameters in probabilistic model based on maximum likelihood estimation. In the present study, we adapt the EM algorithm to inverse problem of heterogeneous reactions obeying the nonlinear state-space model from sparse time-series data, and we focus on the estimation of the rate constant $\{k_{\mathrm{r}}, k_{\mathrm{p}}\}$.

The nonlinear state-space model for the surface heterogeneous reaction has the following several parameters:

$$
\{k_{\mathrm{r}}, k_{\mathrm{p}}, A_{\mathrm{r}}, A_{\mathrm{p}}, C_{\mathrm{r,eq}}, C_{\mathrm{p,eq}}\}. \tag{16}
$$

The equilibrium constants $C_{\mathrm{r,eq}}$ and $C_{\mathrm{p,eq}}$ are assumed to be known, since these constants are observable directly in the usual experimental setting. Furthermore, we can set $A_{\mathrm{r}} = A_{\mathrm{p}} = 1$ without loss of generality.

We propose a method to extract a set of the rate constants $\Theta = \{k_{\mathrm{r}}, k_{\mathrm{p}}\}$ based on the EM algorithm. The EM algorithm is an iterative method consisting of an expectation step (E-step) and a maximization step (M-step). In the EM algorithm, we first prepare an initial set of parameters Θ. In the E-step, we consider the expectation of log-likelihood $\log p(X_T, Y_T|\Theta)$ with respect to posterior distribution:

$$
\mathcal{Q}(\Theta|\Theta') = \int dX_T \, p(X_T|Y_T, \Theta') \ln p(X_T, Y_T|\Theta), \tag{17}
$$

where X_t denotes a time series of hidden variables up to time t ($X_t = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_t\}$). In the M-step, we obtain a set of parameters Θ' as the one that maximizes the $\mathcal{Q}$ function as follows:

$$
\Theta' = \arg\max_{\Theta} \mathcal{Q}(\Theta|\Theta'), \tag{18}
$$

where Θ' is obtained by setting the gradients of $\mathcal{Q}$ function with respect to Θ to zero. We iterate Eqs. (17) and (18) until the condition $|\Theta - \Theta'| < \epsilon$ is satisfied. Here ϵ is set to be a sufficiently small positive constant.

Here we derive the EM algorithm for the nonlinear state space model for heterogeneous reactions using sparse observation data. The log-likelihood function in the $\mathcal{Q}$ function can be obtained from the nonlinear state-space model as follows:

$$
\ln p(X_T, Y_T|\Theta) = \ln p(Y_T|X_T) p(X_T) \tag{19}
$$

$$
= -\frac{1}{2\sigma_{\mathrm{r}}^2} \sum_{t=1}^{T-1} \left(n_{t+1}^{(\mathrm{r})} - n_t^{(\mathrm{r})} - \Delta_t k_{\mathrm{r}} S_{\mathrm{r}}\left(n_t^{(\mathrm{r})}\right)\left(C_t^{(\mathrm{i})} - C_{\mathrm{r,eq}}\right) \right)^2
$$

$$
- \frac{1}{2\sigma_{\mathrm{p}}^2} \sum_{t=1}^{T-1} \left(n_{t+1}^{(\mathrm{p})} - n_t^{(\mathrm{p})} - \Delta_t k_{\mathrm{p}} S_{\mathrm{p}}\left(n_t^{(\mathrm{p})}\right)\left(C_t^{(\mathrm{i})} - C_{\mathrm{p,eq}}\right) \right)^2 - \frac{1}{2\sigma_y^2} \sum_{t \in \mathcal{T}_{\mathrm{obs}}} \left(y_t - C_t^{(\mathrm{i})} \right)^2
$$

$$
- \frac{T-1}{2} \ln\left(2\pi\sigma_{\mathrm{r}}^2\right) - \frac{T-1}{2} \ln\left(2\pi\sigma_{\mathrm{p}}^2\right) - \frac{|\mathcal{T}_{\mathrm{obs}}|}{2} \ln\left(2\pi\sigma_y^2\right). \tag{20}
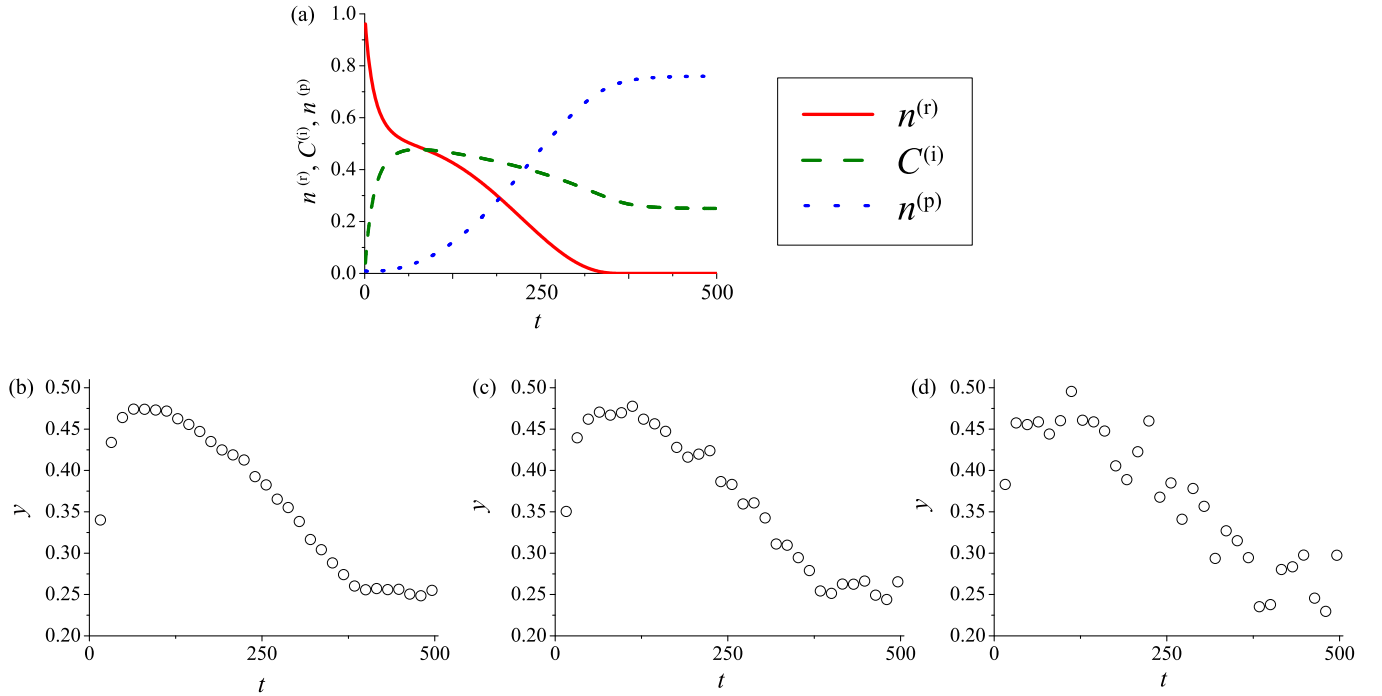$$

FIG. 4. (a) Time courses of true hidden variables: solid reactant $n_t^{(r)}$ (solid line), intermediate product $C_t^{(i)}$ (dashed line), and solid product $n_t^{(p)}$ (dotted line). (b)–(d) Observed data y_t generated by adding Gaussian noise to intermediate product $C_t^{(i)}$. The standard deviation of the observation noise is set to be $\sigma = 10^{-2.5}$ for (b), $\sigma = 10^{-2}$ for (c), and $\sigma = 10^{-1.5}$ for (d), whereas the number of observation points is set to be $N_{obs} = 31$ for (b)–(d).

Note that we assumed that observation times are restricted to the set, $\mathcal{T}_{obs}$, in order to reflect sparse observation. By differentiating $\mathcal{Q}$ function with respect to parameters $\Theta = \{k_r, k_p\}$, we obtain the following expressions:

$$k_r = \frac{\sum_{t=1}^{T-1} \langle (n_{t+1}^{(r)} - n_t^{(r)}) S_r(n_t^{(r)})(C_t^{(i)} - C_{r,eq}) \rangle}{\Delta_t \sum_{t=1}^{T-1} \langle (S_r(n_t^{(r)}))^2 (C_t^{(i)} - C_{r,eq})^2 \rangle}, \quad (21)$$

$$k_p = \frac{\sum_{t=1}^{T-1} \langle (n_{t+1}^{(p)} - n_t^{(p)}) S_p(n_t^{(p)})(C_t^{(i)} - C_{p,eq}) \rangle}{\Delta_t \sum_{t=1}^{T-1} \langle (S_p(n_t^{(p)}))^2 (C_t^{(i)} - C_{p,eq})^2 \rangle}, \quad (22)$$

where $\langle \cdot \rangle$ is an expectation with respect to posterior distribution $p(X_t | Y_t, \Theta')$:

$$\langle \mathcal{X} \rangle = \int dX_t \, p(X_t | Y_t, \Theta') \mathcal{X}. \quad (23)$$

Similarly, parameters for noise intensities in the nonlinear state-space model are estimated by maximizing the $\mathcal{Q}$ function as follows:

$$\sigma_r^2 = \frac{1}{T-1} \sum_{t=1}^{T-1} \langle (n_{t+1}^{(r)} - n_t^{(r)} - \Delta_t k_r S_r(n_t^{(r)})(C_t^{(i)} - C_{r,eq}))^2 \rangle, \quad (24)$$

$$\sigma_p^2 = \frac{1}{T-1} \sum_{t=1}^{T-1} \langle (n_{t+1}^{(p)} - n_t^{(p)} - \Delta_t k_r S_p(n_t^{(p)})(C_t^{(i)} - C_{p,eq}))^2 \rangle, \quad (25)$$

$$\sigma_y^2 = \frac{1}{|\mathcal{T}_{obs}|} \sum_{t \in \mathcal{T}_{obs}} \langle (y_t - C_t^{(i)})^2 \rangle. \quad (26)$$

III. RESULTS

In this section, we validate the proposed method by using simulated data. As described in the previous section, we assume that only the noisy intermediate product y_t is observable at every time step interval. We estimate the time course of hidden variables and the kinetic parameters underlying nonlinear dynamics in the dissolution-precipitation process. We also clarify how the estimation accuracy depends on the number of observations N_{obs} and the level of observation noise σ_y.

The observable data with N_{obs} data points are fed into the proposed method. We assume that data y_t is available at constant time-step intervals within total time steps T. The true kinetic parameters are set to be $k_r = 9.66 \times 10^{-3}$ and $k_p = 4.83 \times 10^{-3}$, and the equilibrium concentrations are set as $C_{r,eq} = 0.5$, $C_{p,eq} = 0.25$. The orders of the surface model are set to be $\alpha_r = \alpha_p = 2/3$.

A. Estimation of hidden state variables from sparse and noisy data of intermediate product

First, we employ the proposed method to estimate the hidden state variables from sparse and noisy data of intermediate product. As shown in Fig. 4, we consider a situation in which only the noisy intermediate product y_t can be observed at every time interval. We assume that we have a time series of y_t with only 31 data points ($N_{obs} = 31$) through this subsection.

Figure 5 shows the hidden variables estimated from the observation data with noise level $\sigma_y = 10^{-2}$ shown in Fig. 4(c). We find that the time trace of intermediate product $C^{(i)}$ can be estimated with high time resolution from noisy low resolution
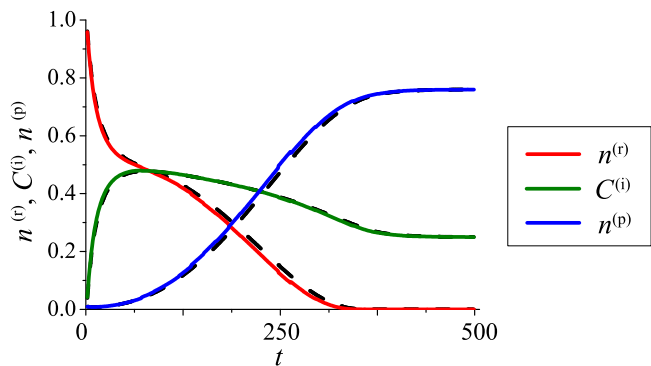
FIG. 5. Estimated hidden variables (solid lines): solid reactant $n_t^{(r)}$ (red line), intermediate product $C_t^{(i)}$ (green line), and solid product $n_t^{(p)}$ (blue line). Dashed lines show the true time courses of the hidden variables. Parameters are set to be $N_{obs} = 31$ and $\sigma_y = 1.0 \times 10^{-2}$.

data. Furthermore, the solid reactant and product are found to be estimated successfully, even though the solid reactant and product are assumed not to be observable directly. We see in Fig. 5 that the proposed method can simultaneously extract a variable observed in a noisy process and other hidden variables in the surface heterogeneous reactions from partially observable noisy data.

We also estimate the reaction rate constants k_r and k_p underlying the nonlinear dynamics of the hidden variables from partially observable data. The red solid line in Fig. 6

shows the reaction rate constants k_r and k_p estimated from the noisy intermediate product shown in Figs. 4(b)–4(d). The initial values of parameters k_r and k_p are set to be away from the true values (initial values $k_r = 0.05$ and $k_p = 0.001$, whereas true values $k_r^{true} \sim 0.01$ and $k_p^{true} \sim 0.005$). Red lines in Fig. 6 show results of parameter estimation from observation data for $\sigma_y = 10^{-2}$. We see in Fig. 6 that estimated kinetic constants converge to true values as the number of iterations in the proposed algorithm increases, showing that the proposed method estimates the true values of kinetic constants accurately.

The data observed in actual laboratory experiments may have noise elements due to the observation process. In order to find out how the observation noise affects the estimation performance in the proposed method, we investigate the estimation accuracy for the observation data with different observation noise levels $\sigma_y = 10^{-2.5}$, 10^{-2}, and $10^{-1.5}$ (corresponding to 0.63%, 2%, and 6.3% deviations relative to the maximal concentration of intermediate product $C^{(i)}$). We see in Figs. 6(a) and 6(b) that the estimated rate constants k_r and k_p converge to the true ones even for different levels of observation noise, while the convergence becomes slow with increasing observation noise. We also see in Fig. 6(c) that trajectory of estimated parameters tend to take a circuitous route in the case of strong observation noise $\sigma_y = 10^{-1.5}$. These results show that our proposed method can estimate from partially observable noisy data not only the hidden variables but also the rate constants underlying the behavior of the hidden variables.
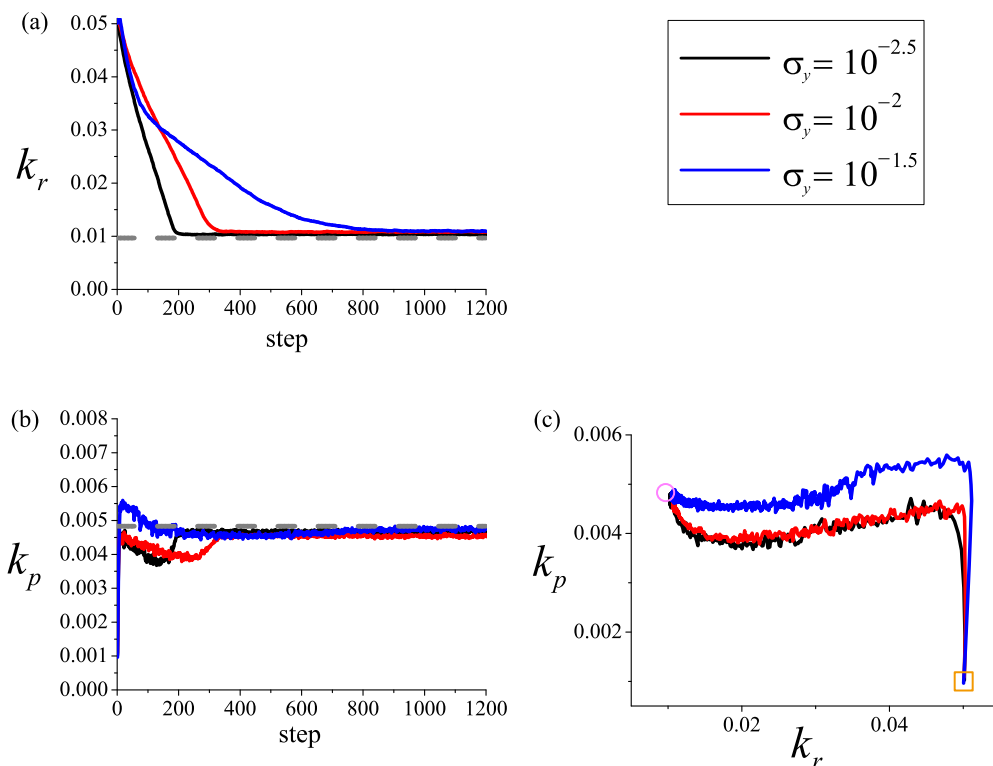

FIG. 6. Estimation of rate constants k_r and k_p from partially observable data. Estimated values for different levels of noise [$\sigma_y = 10^{-2.5}$ (black), 10^{-2} (red), and $10^{-1.5}$ (blue)] are shown. (a), (b) The estimated rate constants (solid lines) converge to true rate constants as the iteration steps proceed, even when we set the initial values are far from the true values (dashed lines). (c) Trajectory of estimated rate constants in two-dimensional space (k_r, k_p). The square shows initial values and the circle shows true values.
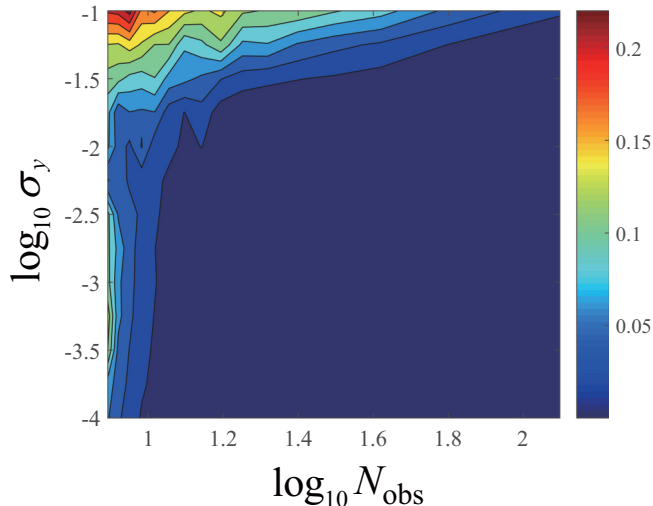
FIG. 7. Dependence of discrepancy between true and estimated kinetic parameters k_r and k_p on the number of observation N_{obs} and the observation noise σ_y. The noisy data of intermediate products are set to be observable at every time step interval $\Delta_{obs} \sim T/N_{obs}$ within total time steps $T = 500$.

B. Dependence of estimation accuracy on time resolution and observation noise

Here we investigate how robustly the rate constants for the surface heterogeneous reaction k_r and k_p can be estimated for different numbers of observations with different levels of observation noise. Figure 7 shows how the discrepancy between true and estimated kinetic parameters k_r and k_p depends on the number of observations N_{obs} and the observation noise σ_y. The discrepancy is calculated as a squared error $\varepsilon = (k_r^{(true)} - k_r^{(est)})^2/(k_r^{(true)})^2 + (k_p^{(true)} - k_p^{(est)})^2/(k_p^{(true)})^2$. For each setting of N_{obs} and σ_y, estimated results in time series data for twenty trials are evaluated.

We see in Fig. 7 that the accuracy levels of the estimation of the rate constants ($\varepsilon \leqslant 0.05$) shown in Fig. 6 [for $N_{obs} = 31$ ($\log_{10} N_{obs} \sim 1.5$); $\sigma_y = 10^{-2.5}$, 10^{-2}, and $10^{-1.5}$] can be held for wide ranges of the number of observations and the observation noise intensity (blue area of the Fig. 7). Even when the number of observations is very limited ($N_{obs} = 10$), the rate constants are estimated with $\varepsilon \sim 0.05$ for moderate noise levels, as we see such trajectory of rate constants in the blue lines of Fig. 6 (for $\log_{10} \sigma_y = -1.5$ and $\log_{10} N_{obs} \sim 1.5$). From these results, we see that our proposed method has robustness to sparseness and noise in the observation; kinetic parameters are found to be estimated accurately for wide ranges of the number of observations N_{obs} and noise intensity σ_y. Moreover, in the estimation accuracy, there is a tradeoff relation between the observation noise and the sparseness of the observation.

IV. CONCLUDING REMARKS

We have proposed a Bayesian estimation algorithm for inverse problem of nonlinear dynamics in surface heterogeneous reactions. We first formulated a nonlinear state-space model of the surface heterogeneous reaction and then derived the marginalized posterior distribution of the heterogeneous

reactions in order to estimate hidden variables. In particular, we developed a belief propagation algorithm for partially observable data to realize estimation of nonlinear dynamics from sparse data. We also adapted the EM algorithm to the partial observation problem. Using the proposed method, we have shown that hidden variables (true solid reactant and product as well as intermediate product dissolved in the liquid phase) are simultaneously estimated from partially observable noisy data. Furthermore, kinetic parameters were found to be successfully extracted as well as the hidden variables. Our verification using simulated data showed that the proposed method provides robust estimation for wide ranges of observation noise levels and for data with low temporal resolution.

Belief propagation [24,25] and EM algorithm [21–23] have been used for estimating hidden variables and parameters in probabilistic model, respectively. Although these existing methods deal with estimation from observations of a part of the variables, they do not consider switches between unobservable and observable cases of the same variables. In order to realize inverse estimation from temporally sparse data, we derived posterior distributions used to estimate hyperparameters for both unobservable and observable cases by adapting belief propagations and an EM algorithm to the inverse problem of surface heterogeneous reactions. To conduct the derived belief propagation method in nonlinear dynamical systems, we employed sequential Monte Carlo algorithm. Evaluation using simulated data showed that our algorithm can estimate hyperparameters including rate constants accurately even from sparsely observable data.

In the present paper, we employed the sequential Monte Carlo method to extract hidden variables since it can deal with arbitrary distributions. There exist other methods (Kalman-filter-based algorithm) for extracting nonlinear dynamics such as unscented Kalman filters, and extended Kalman filters [20]. Nonlinearity and non-Gaussian distribution are reflected precisely for sufficient sampling points in the sequential Monte Carlo method, whereas linear and Gaussian approximation is employed in the Kalman-filter-based algorithms [20]. We leave for future work comparison of the estimation performance of different Kalman-filter-based algorithms with that of the proposed method.

The present study has proposed a Bayesian framework for estimating complex dynamical systems such as heterogeneous reactions from noisy and sparse measurements. Previous studies using compressive sensing (CS) proposed reconstruction methods for complex dynamical systems from sparse measurements using noiseless CS framework [32–35] and a noisy CS framework [36]. In the framework based on the Bayesian inference, the parameters including noise intensities can be automatically estimated as shown in the present study, whereas the noisy CS framework may require some tuning of constants related to regularization for sparsity while these studies using the CS framework have shown some robustness for weak noise.

In this study, we considered a simple surface heterogeneous reaction with three variables, two of which are independent variables. In general, many variables and reaction terms would affect the nonlinear dynamics of surface heterogeneous reactions [5,26,37]. Since our proposed method based on

Bayesian statistics can estimate nonlinear dynamics in surface heterogeneous reaction from restricted kinds of chemical species, our framework may be applicable to complex chemical reactions with multiple components and multiple phases if we can observe sufficient kinds of chemical species in solution. Evaluation of the applicability of the proposed method to more complex chemical reactions and extension of the proposed method are also important. We also leave them for a future work.

Our proposed method is applicable to a wide range of estimation problems in dynamical systems governed by Markov processes. Estimation for a non-Markov process is also an important problem. Although in this study we assumed white Gaussian noise for simplicity, our proposed method is applicable to wider ranges of noise types, such as colored noise. We leave these problems as a future work.

APPENDIX A: DERIVATION OF PROBABILISTIC MODEL FOR SYSTEM MODEL

Here we derive probabilistic models of the system model from kinetics for the surface heterogeneous reaction.

By discretizing a stochastic differential equation [Eq. (1)] with respect to time [38], we obtain the following difference equation:

$$\frac{n_{t+1}^{(\mathrm{r})} - n_t^{(\mathrm{r})}}{\Delta_t} = k_\mathrm{r} S_\mathrm{r}\big(n_t^{(\mathrm{r})}\big)\big(C_t^{(\mathrm{i})} - C_\mathrm{r,eq}\big) + \frac{1}{\sqrt{\Delta_t}}\xi_t^{(\mathrm{r})}, \quad (\mathrm{A1})$$

where t denotes a time step and Δ_t is a time interval used for time discretization. Here $\xi_t^{(\mathrm{r})}$ shows a Gaussian noise with $\langle \xi_t^{(\mathrm{r})} \rangle = 0$ and $\langle \xi_t^{(\mathrm{r})} \xi_s^{(\mathrm{r})} \rangle = \tilde{\sigma}_\mathrm{r}^2 \delta(t,s)$. From this expression, we obtain the following recursion equation for $n_t^{(\mathrm{r})}$:

$$n_{t+1}^{(\mathrm{r})} = n_t^{(\mathrm{r})} + \Delta_t k_\mathrm{r} S_\mathrm{r}\big(n_t^{(\mathrm{r})}\big)\big(C_t^{(\mathrm{i})} - C_\mathrm{r,eq}\big) + \sqrt{\Delta_t}\xi_t^{(\mathrm{r})}. \quad (\mathrm{A2})$$

Since $\xi_t^{(\mathrm{r})}$ is assumed to be a Gaussian noise, probabilistic model for the reactant $p(n_{t+1}^{(\mathrm{r})}|n_t^{(\mathrm{r})})$ is expressed as follows:

$$p\big(n_{t+1}^{(\mathrm{r})}\big|n_t^{(\mathrm{r})}\big) = \frac{1}{\sqrt{2\pi\sigma_\mathrm{r}^2}} \exp\left[-\frac{\big(n_{t+1}^{(\mathrm{r})} - n_t^{(\mathrm{r})} - \Delta_t k_\mathrm{r} S_\mathrm{r}\big(n_t^{(\mathrm{r})}\big)\big(C_t^{(\mathrm{i})} - C_\mathrm{r,eq}\big)\big)^2}{2\sigma_\mathrm{r}^2}\right], \quad (\mathrm{A3})$$

where we put $\sigma_\mathrm{r} = \sqrt{\Delta_t}\tilde{\sigma}_\mathrm{r}$. Similary, a probabilistic model for the product is derived as follows:

$$p\big(n_{t+1}^{(\mathrm{p})}\big|n_t^{(\mathrm{p})}\big) = \frac{1}{\sqrt{2\pi\sigma_\mathrm{p}^2}} \exp\left[-\frac{\big(n_{t+1}^{(\mathrm{p})} - n_t^{(\mathrm{p})} - \Delta_t k_\mathrm{p} S_\mathrm{p}\big(n_t^{(\mathrm{p})}\big)\big(C_t^{(\mathrm{i})} - C_\mathrm{p,eq}\big)\big)^2}{2\sigma_\mathrm{p}^2}\right]. \quad (\mathrm{A4})$$

By using these probabilistic models for each hidden variable, we obtain the following probabilistic model $p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$:

$$\begin{aligned} p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) &= p\big(n_{t+1}^{(\mathrm{r})}, n_{t+1}^{(\mathrm{p})}\big|n_t^{(\mathrm{r})}, n_t^{(\mathrm{p})}\big) \\ &= p\big(n_{t+1}^{(\mathrm{r})}\big|n_t^{(\mathrm{r})}\big) p\big(n_{t+1}^{(\mathrm{p})}\big|n_t^{(\mathrm{p})}\big), \quad (\mathrm{A5}) \end{aligned}$$

where fluctuating noises $\xi_t^{(\mathrm{r})}$ and $\xi_t^{(\mathrm{p})}$ for reactant and product, respectively, are assumed to be independent as $\langle \xi_t^{(\mathrm{r})} \xi_t^{(\mathrm{p})} \rangle = 0$.

APPENDIX B: POSTERIOR PROBABILITY DENSITY FUNCTION FOR NONLINEAR STATE SPACE MODEL

From Bayes's theorem, a posterior density function $p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T | Y_T)$ is proportional to a product of likelihood function $p(Y_T | \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$ and prior probability $p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$ as follows:

$$\begin{aligned} &p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T | Y_T) \\ &\propto p(Y_T | \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T) p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T). \quad (\mathrm{B1}) \end{aligned}$$

By assuming that each observation of y_t is independent of other observations $y_{s \neq t}$, the likelihood function is expressed as a product of observation models $p(\boldsymbol{y}_t | \boldsymbol{x}_t)$ at observable time steps as follows:

$$p(Y_T | \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T) = \prod_{t \in \mathcal{T}_\mathrm{obs}} p(\boldsymbol{y}_t | \boldsymbol{x}_t), \quad (\mathrm{B2})$$

where $\mathcal{T}_\mathrm{obs}$ denotes a set of sparse observable time within total time step T ($|\mathcal{T}_\mathrm{obs}| < T$).

The prior probability is expressed as follows:

$$\begin{aligned} &p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T) \\ &= p(\boldsymbol{x}_T | \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{T-1}) p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{T-1}) \\ &= p(\boldsymbol{x}_T | \boldsymbol{x}_{T-1}) p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{T-1}), \quad (\mathrm{B3}) \end{aligned}$$

where we used the Markov property of the nonlinear state-space model $p(\boldsymbol{x}_T | \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{T-1}) = p(\boldsymbol{x}_T | \boldsymbol{x}_{T-1})$. Note that Eq. (B3) is a recursive equation for $p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t)$. By using this recursive equation, the prior probability is obtained as a product of system models $p(\boldsymbol{x}_{t+1} | \boldsymbol{x}_t)$ and initial probability $p(\boldsymbol{x}_1)$ as follows:

$$p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T) = \left(\prod_{t=1}^{T-1} p(\boldsymbol{x}_{t+1} | \boldsymbol{x}_t)\right) p(\boldsymbol{x}_1). \quad (\mathrm{B4})$$

Substituting Eqs. (B2) and (B4) into Eq. (B1), we obtain a posterior probability density function [Eq. (9)].

APPENDIX C: BELIEF PROPAGATION FOR PARTIALLY OBSERVABLE DATA

In this section, we derive the marginal posterior distribution $p(\boldsymbol{x}_{t+1}|Y_t)$ of the proposed method [Eqs. (11)–(14)]. To derive the marginal posterior distribution for partially observable data, we need to consider two cases; in the first case we assume that observable data is not available at time t, and in the second case we assume that observable data is available at time t.

1. Nonobservable case

First, we consider the case in which we do not have any observable point y_t at time t. Therefore, we cannot assume observation model $p(y_t|\boldsymbol{x}_t)$ in this case. A predictive distribution $p(\boldsymbol{x}_{t+1}|Y_t)$ can be derived using integral representation of the probability as follows:

$$
\begin{aligned}
p(\boldsymbol{x}_{t+1}|Y_t) &= \int p(\boldsymbol{x}_{t+1},\boldsymbol{x}_t|Y_t)d\boldsymbol{x}_t \\
&= \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)p(\boldsymbol{x}_t|Y_t)d\boldsymbol{x}_t \\
&= \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)p(\boldsymbol{x}_t|Y_{t-1})d\boldsymbol{x}_t, \quad \text{(C1)}
\end{aligned}
$$

where we used $Y_t = \{y_1,\ldots,y_t\} = Y_{t-1}$ since we assume that there exists no observation for y_t at time t. Note that Eq. (C1) is a recursive equation for the predictive distribution $p(\boldsymbol{x}_{t+1}|Y_t)$.

2. Observable case

Next, we consider the case in which we have an observable data point at time t. A predictive distribution $p(\boldsymbol{x}_{t+1}|Y_t)$ is expressed by using integration with respect to $\boldsymbol{x}_t$ as follows:

$$
\begin{aligned}
p(\boldsymbol{x}_{t+1}|Y_t) &= \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t,Y_t)p(\boldsymbol{x}_t|Y_t)d\boldsymbol{x}_t \\
&= \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)p(\boldsymbol{x}_t|Y_t)d\boldsymbol{x}_t, \quad \text{(C2)}
\end{aligned}
$$

where we used a Markov property $p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t,Y_t) = p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t)$. On the other hand, a filtering distribution $p(\boldsymbol{x}_t|Y_t)$ is derived as follows:

$$
\begin{aligned}
p(\boldsymbol{x}_t|Y_t) &= p(\boldsymbol{x}_t|\boldsymbol{y}_t,Y_{t-1}) \\
&= \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_t,Y_{t-1})p(\boldsymbol{x}_t|Y_{t-1})}{p(\boldsymbol{y}_t|Y_{t-1})} \\
&= \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_t)p(\boldsymbol{x}_t|Y_{t-1})}{p(\boldsymbol{y}_t|Y_{t-1})}, \quad \text{(C3)}
\end{aligned}
$$

where we have used a Markov property $p(\boldsymbol{y}_t|\boldsymbol{x}_t,Y_{t-1}) = p(\boldsymbol{y}_t|\boldsymbol{x}_t)$ in the final equality.

[1] P. Atkins and J. D. Paula, *Physical Chemistry*, 10th ed. (Oxford University Press, Oxford, 2014).

[2] A. C. Lasaga, *Kinetic Theory in the Earth Science* (Princeton University Press, Princeton, 1998).

[3] A. Okamoto, Y. Ogasawara, Y. Ogawa, and N. Tsuchiya, Chem. Geol. **289**, 245 (2011).

[4] V. A. Alekseyev, L. S. Medvedeva, and N. P. Starshinova, Geochem. Intl. **47**, 731 (2009).

[5] C. Zhu, P. Lu, Z. Zheng, and J. Ganor, Geochim. Cosmochim. Acta **74**, 3963 (2010).

[6] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences* (Cambridge University Press, Cambridge, 2010).

[7] Y. Naruse, K. Takiyama, M. Okada, and T. Murata, Phys. Rev. E **82**, 011912 (2010).

[8] T. Omori, T. Aonishi, and M. Okada, Adv. Cognit. Neurodyn. **3**, 649 (2013).

[9] X. Liu and Q. Gao, Phys. Rev. E **88**, 042905 (2013).

[10] T. Omori, Neural Inf. Proc. Lec. Not. Comput. Sci. **8834**, 27 (2014).

[11] T. Omori and K. Hukushima, J. Phys. Conf. Ser. **699**, 012011 (2016).

[12] T. Kuwatani, K. Nagata, M. Okada, and M. Toriumi, Contrib. Mineral. Petrol. **163**, 547 (2012).

[13] T. Kuwatani, K. Nagata, M. Okada, and M. Toriumi, Earth Planets Space **66**, 5 (2014).

[14] T. Kuwatani, K. Nagata, M. Okada, and M. Toriumi, Phys. Rev. E **90**, 042137 (2014).

[15] A. Doucet, S. Godsill, and C. Andrieu, Stat. Comput. **10**, 197 (2000).

[16] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice* (Springer-Verlag, Berlin, 2001).

[17] G. Kitagawa, J. American Stat. Assoc. **82**, 1032 (1987).

[18] R. E. Johnson and S. Ranganathan, Phys. Rev. E **75**, 056706 (2007).

[19] Z. Ghahramani and G. E. Hinton, University of Toronto Technical Report CRG-TR-96-2 (1996).

[20] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer-Verlag, Berlin, 2006).

[21] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Its Extensions* (Wiley, New York, 1997).

[22] A. P. Dempster, N. M. Laird, and D. B. Rubin, J. R. Stat. Soc. B **39**, 1 (1977).

[23] C. F. J. Wu, Annal. Stat. **11**, 95 (1983).

[24] J. Pearl, *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufman, New York, 1988).

[25] J. S. Yedidia, W. T. Freeman, and Y. Weiss, in *Exploring Artificial Intelligence in the New Millenium*, edited by G. Lakemeyer and B. Nebel (Morgan Kaufmann, New York, 2002), p. 239.

[26] S. Brantley and J. Kubicki, *Kinetics of Water-Rock Interaction* (Springer-Verlag, Berlin, 2007).

[27] A. Okamoto, H. Saishu, N. Hirano, and N. Tsuchiya, Geochim. Cosmochim. Acta **74**, 3692 (2010).

[28] A. Okamoto, T. Kuwatani, T. Omori, and K. Hukushima, Phys. Rev. E **92**, 042130 (2015).

[29] K. Tanaka, J. Phys. A: Math. Gen. **35**, R81 (2002).

[30] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic Press, San Diego, 1982).

[31] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).

[32] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, Phys. Rev. Lett. **106**, 154101 (2011).

[33] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and M. A. F. Harrison, Europhys. Lett. **94**, 48006 (2011).

[34] W.-X. Wang, Y.-C. Lai, C. Grebogi, and J. Ye, Phys. Rev. X **1**, 021021 (2011).

[35] R.-Q. Su, W.-X. Wang, and Y.-C. Lai, Phys. Rev. E **85**, 065201(R) (2012).

[36] R.-Q. Su, Y.-C. Lai, X. Wang, and Y. Do, Sci. Rep. **4**, 3944 (2014).

[37] C. M. Bethke, *Geochemical Modeling: Concepts and Applications* (Oxford University Press, Oxford, 1996).

[38] H. Risken, *Fokker-Planck Equation*, 2nd Ed. (Springer, Berlin, 1989).