Data-Driven Analysis of Nonlinear Heterogeneous Reactions through Sparse Modeling and Bayesian Statistical Approaches

Masaki Ito [1], Tatsu Kuwatani [2] , Ryosuke Oyanagi [2,3] and Toshiaki Omori [1,4,]

[1] Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan; en818.businhete2021@gmail.com

[2] Research Institute for Marine Geodynamics, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Kanagawa, Yokosuka 237-0061, Japan; kuwatani@jamstec.go.jp (T.K.); oyanagir@kokushikan.ac.jp (R.O.)

[3] School of Science and Engineering, Kokushikan University, 4-28-1 Setagaya, Setagaya-ku, Tokyo 154-8515, Japan

[4] Center for Mathematical and Data Sciences, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

[] Correspondence: omori@eedept.kobe-u.ac.jp

Abstract: Heterogeneous reactions are chemical reactions that occur at the interfaces of multiple phases, and often show a nonlinear dynamical behavior due to the effect of the time-variant surface area with complex reaction mechanisms. It is important to specify the kinetics of heterogeneous reactions in order to elucidate the microscopic elementary processes and predict the macroscopic future evolution of the system. In this study, we propose a data-driven method based on a sparse modeling algorithm and sequential Monte Carlo algorithm for simultaneously extracting substantial reaction terms and surface models from a number of candidates by using partial observation data. We introduce a sparse modeling approach with non-uniform sparsity levels in order to accurately estimate rate constants, and the sequential Monte Carlo algorithm is employed to estimate time courses of multi-dimensional hidden variables. The results estimated using the proposed method show that the rate constants of dissolution and precipitation reactions that are typical examples of surface heterogeneous reactions, necessary surface models, and reaction terms underlying observable data were successfully estimated from only observable temporal changes in the concentration of the dissolved intermediate products.

Keywords: sparse modeling; sequential Monte Carlo method; time series data analysis; heterogeneous reactions

1. Introduction

Specifying the underlying dynamical mechanism from observed time-series data sets is ubiquitously important in natural sciences [1]. Generally, the governing equations include many candidate nonlinear terms for describing complex dynamical behaviors, and the observations of the actual problems are always limited [1,2]. Therefore, in order to understand the microscopic elementary processes and predict the future macroscopic behaviors of a complex system, it is necessary to develop a data-driven method that can extract substantial elements from a number of candidates by using partially observable time-series data.

Heterogeneous reactions are chemical reactions that occur at the interface of multiple phases, and they are essential to understand the complex behavior of substances in natural systems [3]. Their dynamics have intrinsic nonlinearity caused by the effect of inevitable changes in the surface area among different phases [2]. In earth science, it is important to understand heterogeneous reactions in order to clarify the dynamics of rock formations near the surface of the earth. However, in heterogeneous reactions, there are many kinds of reaction terms with different surface models and orders of reactions [2]. Therefore, it

is difficult to extract important reactions from many candidates in order to estimate the underlying dynamics of heterogeneous reactions from the observable data.

Bayesian statistical approaches were proposed to estimate heterogeneous reactions from observable data. A Bayesian inversion analysis method with sequential Monte Carlo and expectation–maximization (EM) algorithms was proposed to simultaneously estimate the kinetic rate constants and time series of multi-dimensional hidden variables for heterogeneous reactions [4]. Moreover, a Bayesian approach, using Markov chain Monte Carlo and widely applicable information criteria, was proposed to estimate heterogeneous reaction pathways using spatial data [5]. In these previous Bayesian methods [4,5], the types of reaction terms are assumed to be known. However, it is important to establish a method for estimating reaction constants while assuming that the types of reaction terms are unknown since there exist many kinds of candidates in reaction terms in heterogeneous reactions.

Recently, sparse modeling approaches have attracted much attention owing to their ability to extract important factors from observable data [6,7]. The sparse modeling approach can extract essential elements from candidate factors that reconstruct observed data by assuming that important elements within the entire candidate elements are sparse. Sparse modeling is widely employed in various fields of natural sciences, such as physics [8,9], brain science [10–12], astronomy [13] and earth sciences [14]. In particular, applying the sparse modeling approach to heterogeneous reactions is an important subject since sparse modeling may realize the extraction of only necessary terms from many candidates of reaction terms.

This paper proposes a data-driven method to simultaneously estimate rate constants and hidden variables from nonlinear dynamics of heterogeneous reactions of fluid–rock interaction by adapting sparse modeling and sequential Monte Carlo methods. First, we establish a state-space model for nonlinear dynamics for heterogeneous reactions, including multiple candidates of the reaction terms. Next, we derive the sequential Monte Carlo method for partial observation problems in heterogeneous reactions with many candidate terms. Furthermore, we adapt the sparse modeling approach to estimation problem of dynamical system in order to estimate necessary rate constants from many candidates. We employ a sparse modeling approach with non-uniform sparsity levels and make the sparse modeling algorithm more robust in order to improve the estimation accuracy for rate constants. The estimated results show that the proposed method can extract necessary reaction terms from many candidates more accurately, compared with naïve estimation methods.

This paper is organized as follows. In Section 2, we propose a sparse modeling method for extracting heterogeneous reactions. First, we describe a general form of differential equations of heterogeneous reactions. Next, we derive a nonlinear state-space model for the heterogeneous reactions, and then we adapt a sparse modeling approach to probabilistic frameworks described by the state-space model of heterogeneous reactions. In Section 3, we verify the effectiveness of the proposed method, using simulated data obtained from a dynamical model of heterogeneous reaction while assuming many candidates of nonlinear terms. Concluding remarks are given in Section 4.

2. Method

In this section, we propose a sparse modeling method for extracting heterogeneous reactions. First, we introduce heterogeneous reactions, which depend on surface-area reactions between the liquid phase and the solid phase, and formulate a nonlinear state-space model of heterogeneous reactions with many candidate terms. Next, we employ the sequential Monte Carlo method in order to estimate the hidden variables of heterogeneous reactions from observable data. Finally, we adopt a sparse modeling approach to the framework of the sequential Monte Carlo method in order to extract important heterogeneous reactions from many candidates.

2.1. Nonlinear Dynamics of Heterogeneous Reactions of Fluid–Rock Interaction

In this study, we consider a general setting of heterogeneous reactions in fluid–rock interactions in which minerals (solid phase) dissolve and precipitate within a liquid phase. The fluid–rock interaction is one of the most important geoscientific processes in the earth. For example, serpentinization, a collection of heterogeneous reactions to generate serpentinite from peridotite and water, in the oceanic crust is known to be a key storage process of water brought into the earth's deep region. In addition, it changes the physical properties of the earth's interior significantly, and affects the occurrence of earthquakes in the subduction zone [15]. It is important to specify the kinetics of heterogeneous reactions in order to understand the detailed processes of rock formation and predict the evolution of the earth system. However, estimating the heterogeneous reactions is difficult since observation data are limited, and there are many candidates for reaction types [16–19]. In addition, dominant reaction terms generally vary according to external conditions, such as temperature and pressure, hence it is very difficult to presumably predict which terms are active or inactive. Therefore, it is imperative to establish a method for estimating the reactions based on data-driven analysis using the data sets that are obtained from laboratory experiments and natural samples [20].

Here, we consider surface heterogeneous reactions in a hydrothermal laboratory experiment as an analogue of fluid–rock interactions in natural systems, as shown in Figure 1. In the reactions, a solid reactant $N^{(\mathrm{r})}$ dissolves into a liquid to form an intermediate product C, and the intermediate product C precipitates to form a solid product $N^{(\mathrm{p})}$.
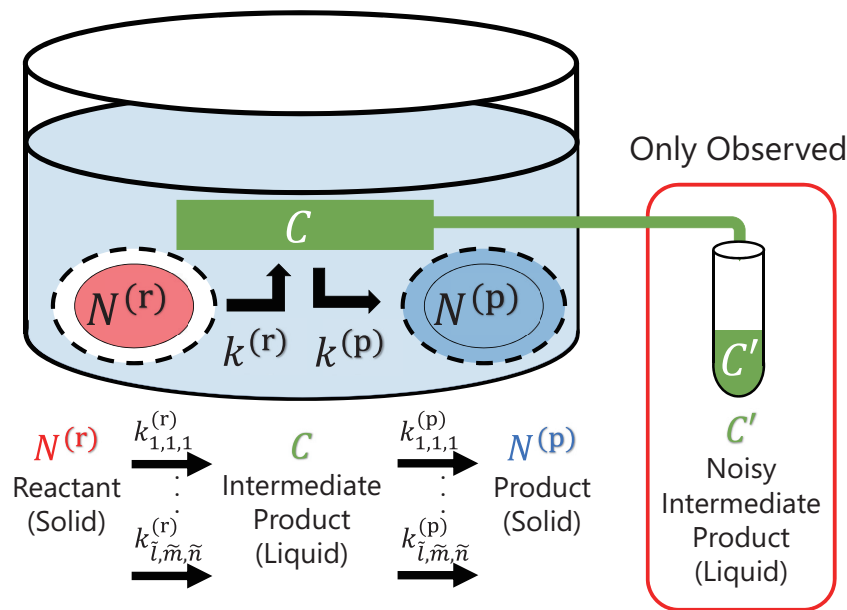
Figure 1. Schematic diagram of surface heterogeneous reactions consisting of a reactant $N^{(\mathrm{r})}$, intermediate product C and product $N^{(\mathrm{p})}$. The solid reactant dissolves into a liquid intermediate product and it precipitates into a solid product, and there are many candidate reaction terms for the surface heterogeneous reactions. These reactions depend on rate constants $k_{l,m,n}^{(\mathrm{r})}$ and $k_{l,m,n}^{(\mathrm{p})}$ for different surface-area reactions $S_l^{(\mathrm{r})}(N^{(\mathrm{r})})$ and $S_l^{(\mathrm{p})}(N^{(\mathrm{p})})$, and different orders of reaction terms m and n.

Such dissolution and precipitation via an intermediate product within fluid is considered to be the most fundamental process, which describes the essence of heterogeneous reactions in fluid–rock interactions. Since there are many kinds of heterogeneous reactions that are linear or nonlinear related to surface-area reactions and reaction terms, and we cannot presumably assume effective specific reaction terms, we consider the amounts of

reactant $N^{(\mathrm{r})}$, and those of product $N^{(\mathrm{p})}$ obey the following general differential equations with a number of candidate reaction terms [2]:

$$\frac{dN^{(\mathrm{r})}}{dt'} = \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(\mathrm{r})} S_l^{(\mathrm{r})}\left(N^{(\mathrm{r})}\right)\left\{C^m - \left(C_{eq}^{(\mathrm{r})}\right)^m\right\}^n \tag{1}$$

$$\frac{dN^{(\mathrm{p})}}{dt'} = \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(\mathrm{p})} S_l^{(\mathrm{p})}\left(N^{(\mathrm{p})}\right)\left\{C^m - \left(C_{eq}^{(\mathrm{p})}\right)^m\right\}^n \tag{2}$$

$$\frac{dN^{(\mathrm{r})}}{dt'} + \frac{dC}{dt'} + \frac{dN^{(\mathrm{p})}}{dt'} = 0 \tag{3}$$

where t' denotes time. Here, $k_{l,m,n}^{(\mathrm{r})}$ and $k_{l,m,n}^{(\mathrm{p})}$ are the rate constants governing the dynamics of heterogeneous reactions. $S_l^{(\mathrm{r})}(N)$ and $S_l^{(\mathrm{p})}(N)$ indicate the surface areas of the reactant and product, respectively. l $(l = 1, 2, \cdots, \tilde{l})$ is a type of surface area model. m $(m = 1, 2, \cdots, \tilde{m})$ and n $(n = 1, 2, \cdots, \tilde{n})$ are indices for the orders of an intermediate product C and factor $\left\{C^m - \left(C_{eq}\right)^m\right\}$ of the reaction terms, respectively. Here, $\tilde{l}$, $\tilde{m}$ and $\tilde{n}$ are the total numbers of candidate surface models and candidate orders of the two factors. Note that Equations (1) and (2) contain $\tilde{l} \times \tilde{m} \times \tilde{n}$ kinds of candidate reaction terms for each equation, and essential reaction terms are extracted from the candidates, using the framework of sparse modeling as described below.

The dynamics of heterogeneous reactions shown in Equations (1)–(3) depend on the rate constants $k_{l,m,n}^{(\mathrm{r})}$ and $k_{l,m,n}^{(\mathrm{p})}$. The constants indicate how fast the reactants $N^{(\mathrm{r})}$ dissolve and the products $N^{(\mathrm{p})}$ precipitate. Therefore, rate constants $k_{l,m,n}^{(\mathrm{r})}$ and $k_{l,m,n}^{(\mathrm{p})}$ are important factors to determine what kind of mineral dissolves into a liquid intermediate product and precipitates into a solid product. Figure 2 shows our problem setting for extracting important reactions from many reaction candidates. In Figure 2, surface area is simply expressed as $S_l(N)$ rather than $S_l^{(\mathrm{r})}(N)$ and $S_l^{(\mathrm{p})}(N)$. $S_l(N)$ indicates the surface area between two solid and liquid phases, and there are many types of surface-area reactions that highly influence heterogeneous surface reactions. Generally, a surface area $S_l(N)$ is expressed as follows:

$$S_l(N) \propto N^\alpha \tag{4}$$

where α is the order of the surface-area reaction. Note that the surface area $S_l(N)$ depends on the number N of moles of the solid phase. As shown in Equation (4), $S_l(N)$ is proportional to the α-th power of N. It is necessary to consider the kind of surface-area reaction that occurs since heterogeneous reactions change at the interface between the solid and liquid phases.

In this study, we consider three typical surface models (Figure 2, horizontal line) as follows:

$$S_1(N) = \mathrm{const.} \tag{5}$$

$$S_2(N) \propto N^{\frac{2}{3}} \tag{6}$$

$$S_3(N) \propto N \tag{7}$$

In Figure 2, the blue area indicates the solid phase, whereas the red lines represent the surface between the solid phase (blue area) and liquid phase (white area). For $\alpha = 0$, the surface area does not change with an increase or decrease in the amount of the solid phase. For $\alpha = 2/3$, the geometrical shape of the solid phase does not change with an increase or decrease in the amount of the solid phase. For $\alpha = 1$, the bulk reaction occurs, and the surface area changes in proportion to the amount of solid phase.
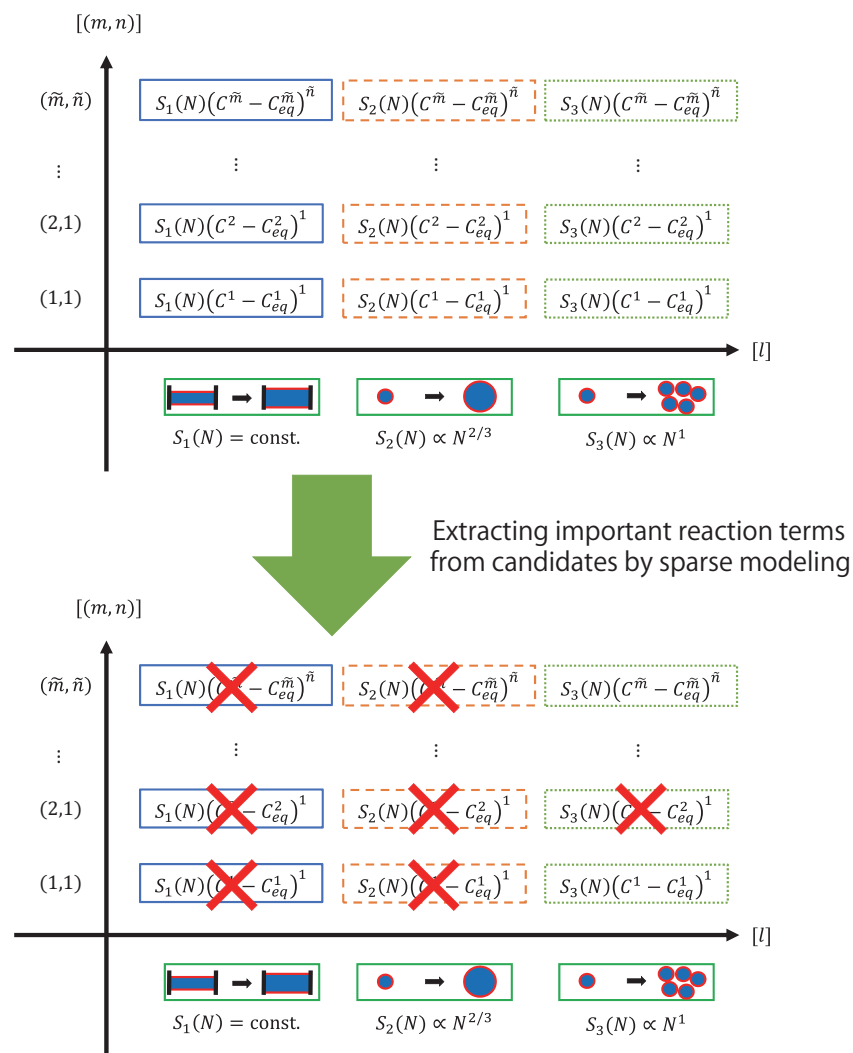
Figure 2. Extraction of important reaction terms from many reaction candidates. (Top) reaction candidates in heterogeneous reactions. The horizontal line shows candidates for surface models, whereas the vertical line shows candidates for the orders of reaction terms. We consider $M \times N \times L$ candidate terms with different surface models ($S_1(N) = \text{const.}$, $S_2(N) \propto N^{2/3}$, and $S_3(N) \propto N$) and different orders m and n in factors of reaction terms $\left\{C^m - \left(C_{eq}\right)^m\right\}^n$. (Bottom) example of extraction of important reaction terms.

As shown in the vertical line of Figure 2, there are candidates for indices of nonlinearity m and n in the reaction terms, where m ($m = 1, 2, \cdots, \tilde{m}$) is a multiplier of the intermediate product C. n ($n = 1, 2, \cdots, \tilde{n}$) is a multiplier of the reaction terms $\left\{C^m - \left(C_{eq}\right)^m\right\}$. Both C^m and $\left\{C^m - \left(C_{eq}\right)^m\right\}^n$ are important for the dynamics of heterogeneous reactions.

Figure 3 shows a typical time course of the heterogeneous reaction when the type of surface-area model is $l = 3$ and the multipliers $m = 1$, $n = 1$ in Equations (1) and (2). We simulate differential equations, Equations (1)–(3), with discretized time step t and set the initial values $N^{(\mathrm{r})}(0) = 0.99$, $C(0) = 0.005$ and $N^{(\mathrm{p})}(0) = 0.005$. As shown in Figure 3, when the time is zero, the solid reactant $N^{(\mathrm{r})}$ exists and dissolves into a liquid intermediate product C with time. When the intermediate product C dissolves sufficiently, it precipitates into a solid product $N^{(\mathrm{p})}$. After that, the solid reactant $N^{(\mathrm{r})}$ almost disappears, and the intermediate product C and solid product $N^{(\mathrm{p})}$ become constant.
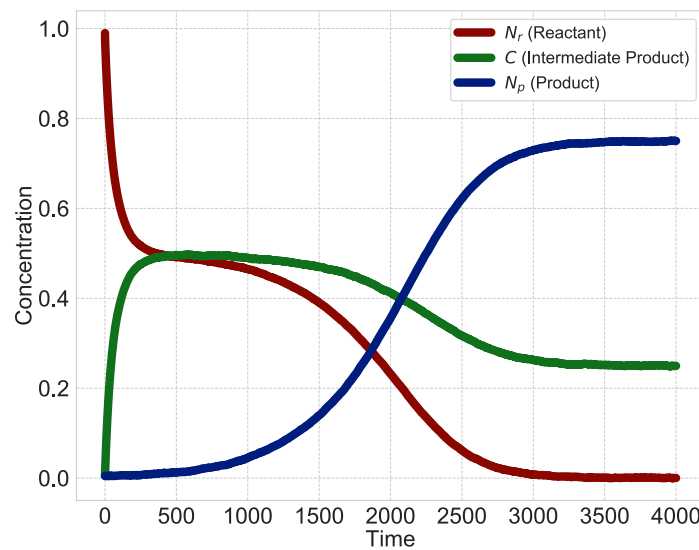
Figure 3. Typical behavior of heterogeneous reactions. At $t = 0$, a solid reactant $N^{(\mathrm{r})}$ only exists and dissolves into a liquid intermediate product C as time passes. When the intermediate product C dissolves sufficiently, it precipitates into a solid product $N^{(\mathrm{p})}$. With more time, the solid reactant $N^{(\mathrm{r})}$ almost disappears, and the intermediate product C and the solid product $N^{(\mathrm{p})}$ become constant.

2.2. Estimating Heterogeneous Reaction Dynamics Based on Sequential Monte Carlo Method

In order to estimate the rate constants $\boldsymbol{k} = \left\{ k^{(\mathrm{r})}_{l,m,n}, k^{(\mathrm{p})}_{l,m,n} \right\}$, it is necessary to estimate the solid reactant $N^{(\mathrm{r})}$, liquid intermediate product C and solid product $N^{(\mathrm{p})}$. We consider how the reactants $N^{(\mathrm{r})}$, intermediate products C and products $N^{(\mathrm{p})}$ change over time.

Figure 4 shows the graphical model of surface heterogeneous reactions governed by Equations (1)–(3). Here, we consider discretized time steps for the states of surface heterogeneous reactions: the reactant $N^{(\mathrm{r})}(t+1)$ at time step $t+1$ depends on the reactant $N^{(\mathrm{r})}(t)$ and the intermediate product $C(t)$ at the preceding time step t. Additionally, the reactant $N^{(\mathrm{r})}(t)$, the intermediate product $C(t)$ and the product $N^{(\mathrm{p})}(t)$ affect the intermediate product $C(t+1)$ and the product $N^{(\mathrm{p})}(t+1)$. Moreover, since the intermediate product is assumed to be observable, the observed intermediate product $C'(t)$ has a relationship with the intermediate product $C(t)$ at the respective time.
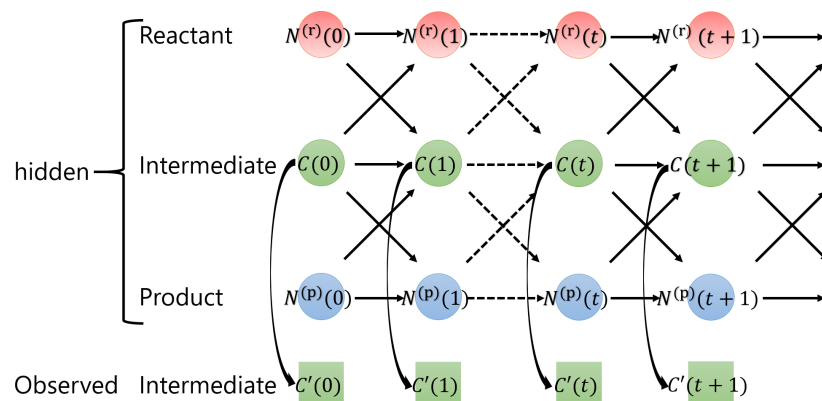
Figure 4. Graphical model of a nonlinear state-space model for heterogeneous reactions. The reactant $N^{(\mathrm{r})}(t+1)$ at time $t+1$ depends on the reactant $N^{(\mathrm{r})}(t)$ and the intermediate product $C(t)$ at time t. In addition, the reactant $N^{(\mathrm{r})}(t)$, intermediate product $C(t)$ and product $N^{(\mathrm{p})}(t)$ affect the intermediate product $C(t+1)$ and product $N^{(\mathrm{p})}(t+1)$. Moreover, since only the intermediate product $C(t)$ is observable, an observable variable $C'(t)$ is assumed to be generated from intermediate product $C(t)$ at each time.

From Figure 4, we propose a nonlinear state-space model for surface heterogeneous reactions. The state space model is a probabilistic model that describes the time evolution and observation process of the dynamical system, and was used in the time series analysis for estimating and predicting hidden variables in dynamical system from observable time-series data [21,22]. The state-space model consists of observation and system models [23]. The observation model represents the relationship between hidden and observation variables, whereas the system model represents the relationship between hidden variables at two adjacent times. In this study, the solid reactant $N^{(r)}$, liquid intermediate product C and solid product $N^{(p)}$ are hidden variables, and the observed liquid intermediate product C' is an observation variable.

First, we formulate the observation model by assuming that the observed liquid intermediate product $C'(t)$ is obtained as a sum of the true liquid intermediate product $C(t)$ and an additive noise as follows:

$$C'(t) = C(t) + \xi^C(t) \tag{8}$$

where $\xi^C(t)$ denotes an observation noise. Assuming that the observation noise obeys the Gaussian distribution, the observation model is expressed by a probabilistic model as follows:

$$p\big(C'(t)|C(t)\big) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(C'(t) - C(t))^2}{2\sigma_y^2}\right) \tag{9}$$

where σ_y shows a standard deviation of the observation noise.

Next, we derive the system model that expresses the relationship among hidden variables. By discretizing Equations (1)–(3) at the time interval Δ_t, we obtain the following difference equation for each time step t:

$$\frac{N^{(r)}(t+1) - N^{(r)}(t)}{\Delta_t} = \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(r)} S_l^{(r)}\big(N^{(r)}(t)\big)\left\{C^m(t) - \left(C_{eq}^{(r)}\right)^m\right\}^n \tag{10}$$

$$\frac{N^{(p)}(t+1) - N^{(p)}(t)}{\Delta_t} = \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(p)} S_l^{(p)}\big(N^{(p)}(t)\big)\left\{C^m(t) - \left(C_{eq}^{(p)}\right)^m\right\}^n \tag{11}$$

$$\frac{N^{(r)}(t+1) - N^{(r)}(t)}{\Delta_t} + \frac{C(t+1) - C(t)}{\Delta_t} + \frac{N^{(p)}(t+1) - N^{(p)}(t)}{\Delta_t} = 0 \tag{12}$$

Furthermore, by assuming a system noise, we obtain the following equations:

$$N^{(r)}(t+1) = N^{(r)}(t) + \Delta_t \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(r)} S_l^{(r)}\big(N^{(r)}(t)\big)\left\{C^m(t) - \left(C_{eq}^{(r)}\right)^m\right\}^n + \sqrt{\Delta_t}\,\xi^{(r)}(t) \tag{13}$$

$$N^{(p)}(t+1) = N^{(p)}(t) + \Delta_t \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(p)} S_l^{(p)}\big(N^{(p)}(t)\big)\left\{C^m(t) - \left(C_{eq}^{(p)}\right)^m\right\}^n + \sqrt{\Delta_t}\,\xi^{(p)}(t) \tag{14}$$

$$N^{(r)}(t+1) - N^{(r)}(t) + C(t+1) - C(t) + N^{(p)}(t+1) - N^{(p)}(t) = 0 \tag{15}$$

where $\xi^{(r)}(t)$ and $\xi^{(p)}(t)$ denote the system noise obeying the white Gaussian noise: for time steps t and s, $\langle \xi^{(r)}(t) \rangle = \langle \xi^{(p)}(t) \rangle = \langle \xi^{(r)}(t)\xi^{(p)}(s) \rangle = 0$, $\langle \xi^{(r)}(t)\xi^{(r)}(s) \rangle = \sigma_r^2 \delta_{t,s}$, and $\langle \xi^{(p)}(t)\xi^{(p)}(s) \rangle = \sigma_p^2 \delta_{t,s}$. Here, $\delta_{t,s}$ is a Kronecker delta; $\delta_{t,s} = 1$ for $t = s$ and $\delta_{t,s} = 0$ for $t \neq s$. σ_r and σ_p express the standard deviations of the system noise. Since Equations (13)–(15) have additive Gaussian noise, a probabilistic model of system model for heterogeneous reactions is derived as follows:

$$p\left(N^{(\mathrm{r})}(t+1)|N^{(\mathrm{r})}(t), C(t), N^{(\mathrm{p})}(t)\right) =$$

$$\frac{1}{\sqrt{2\pi\tilde{\sigma}_{\mathrm{r}}^2}} \exp\left(-\frac{\left(N^{(\mathrm{r})}(t+1) - N^{(\mathrm{r})}(t) - \Delta_t \sum_{l,m,n} k_{l,m,n}^{(\mathrm{r})} S_l^{(\mathrm{r})}\left(N^{(\mathrm{r})}(t)\right)\left\{C^m(t) - \left(C_{eq}^{(\mathrm{r})}\right)^m\right\}^n\right)^2}{2\tilde{\sigma}_{\mathrm{r}}^2}\right) \tag{16}$$

$$p\left(N^{(\mathrm{p})}(t+1)|N^{(\mathrm{r})}(t), C(t), N^{(\mathrm{p})}(t)\right) =$$

$$\frac{1}{\sqrt{2\pi\tilde{\sigma}_{\mathrm{p}}^2}} \exp\left(-\frac{\left(N^{(\mathrm{p})}(t+1) - N^{(\mathrm{p})}(t) - \Delta_t \sum_{l,m,n} k_{l,m,n}^{(\mathrm{p})} S_l^{(\mathrm{p})}\left(N^{(\mathrm{p})}(t)\right)\left\{C^m(t) - \left(C_{eq}^{(\mathrm{p})}\right)^m\right\}^n\right)^2}{2\tilde{\sigma}_{\mathrm{p}}^2}\right) \tag{17}$$

where we put $\tilde{\sigma}_{\mathrm{r}} = \sqrt{\Delta_t}\sigma_{\mathrm{r}}$ and $\tilde{\sigma}_{\mathrm{p}} = \sqrt{\Delta_t}\sigma_{\mathrm{p}}$.

Here, we employ a sequential Monte Carlo method based on the state-space model of heterogeneous reactions. The sequential Monte Carlo method is a statistical method for estimating hidden variables of nonlinear dynamical systems by using a set of particles for posterior distributions of hidden variables [22]. We derive a method for estimating time series of hidden variables $x_t = [N^{(\mathrm{r})}(t), C(t), N^{(\mathrm{p})}(t)]$ $(t = 0, \cdots, T)$ from the observed data $y_{0:T} = [y_0, \cdots, y_T]$ where $y_t = [C'(t)]$.

The predictive distribution of a hidden variable at time step t, given time-series of observed data up to preceding time step $t-1$, $p(x_t|y_{0:t-1})$, is expressed as follows [21–24]:

$$p(x_t|y_{0:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{0:t-1})dx_{t-1} \tag{18}$$

where $p(x_t|x_{t-1})$ is the system model, and $p(x_{t-1}|y_{0:t-1})$ is the filtering distribution at time step $t-1$. Therefore, if we obtain the filtering distribution $p(x_{t-1}|y_{0:t-1})$ at the time step $t-1$, we can calculate the prediction distribution $p(x_t|y_{0:t-1})$ by integrating with respect to x_{t-1}. Moreover, the filtering distribution at time step t, $p(x_t|y_{0:t})$, is expressed using the predictive distribution at time step $t-1$ as follows:

$$p(x_t|y_{0:t}) = \frac{p(y_t|x_t)p(x_t|y_{0:t-1})}{p(y_t|y_{0:t-1})} \tag{19}$$

where $p(y_t|y_{0:t-1})$ is expressed as follows:

$$p(y_t|y_{0:t-1}) = \int p(y_t|x_t)p(x_t|y_{0:t-1})dx_t \tag{20}$$

Note that $p(y_t|x_t)$ is the observation model. Thus, hidden variables can be estimated with forward recursion by alternatively calculating the predictive distribution and filtering distribution. Furthermore, the smoothing distribution of hidden variables at time $t(<T)$ given the time-series of observation data up to time T is expressed as follows:

$$p(x_t|y_{0:T}) = p(x_t|y_{0:t}) \int \frac{p(x_{t+1}|x_t)}{p(x_{t+1}|y_{0:t})} p(x_{t+1}|y_{0:T})dx_{t+1} \tag{21}$$

where $p(x_t|y_{0:t})$ indicates the filtering distribution at time step t. Note that the smoothing distribution at a time step t is obtained by that at time step $t+1$. Therefore, the smoothing distribution can be calculated using backward recursion.

To obtain the above conditional distributions, we express the distributions by using particle approximation. The particle approximation is employed to estimate predictive, filtering, and smoothing distributions numerically since it is difficult to conduct their updates shown in Equations (19)–(21) analytically when nonlinear dynamical systems are assumed [22]. Each conditional distribution for the hidden variables

$x_t = [N^{(\mathrm{r})}(t), C(t), N^{(\mathrm{p})}(t)]$ $(t = 0, \cdots, T)$ is assumed to be expressed by many particles. In the particle approximation, the predictive distribution $p(x_t|y_{1:t-1})$ and the filtering distribution $p(x_t|y_{1:t})$ are expressed as follows:

$$p(x_t|y_{1:t-1}) \simeq \frac{1}{I} \sum_{i=1}^{I} \delta\left(x_t - x_{t|t-1}^{(i)}\right) \tag{22}$$

$$p(x_t|y_{1:t}) \simeq \frac{1}{I} \sum_{i=1}^{I} \delta\left(x_t - x_{t|t}^{(i)}\right) \tag{23}$$

where $\delta(t)$ is the Dirac's delta function and I is the total number of particles. We substitute the approximated predictive and filtering distributions into Equations (18) and (19). By using the recursive relationship between the predictive and filtering distributions, we obtain a set of particles for the filtering distribution $X_{t|t} = [x_{t|t}^{(i)}]_{i=1}^{I}$ from a set of particles for the predictive distribution $X_{t|t-1} = [x_{t|t-1}^{(i)}]_{i=1}^{I}$, and also obtain a set of particles for the predictive distribution $X_{t+1|t}$ from that for the filtering distribution $X_{t|t}$. Based on the sequential Monte Carlo method, we estimate hidden variables by using particles $[x_{t|t}^{(i)}]_{i=1}^{I} = [N_i^{(\mathrm{r})}(t), C_i(t), N_i^{(\mathrm{p})}(t)]_{i=1}^{I}$ from the partially observed data $y_t = [C'(t)]$.

2.3. Sparse Modeling Algorithm for Estimating Rate Constants

Rate constants $k = \left\{k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})}\right\}$ play a important role as determinants for reaction rates and surface-area reactions. Moreover, rate constants determine linearity or nonlinearity of reaction term $\left\{C^m - \left(C_{eq}\right)^m\right\}^n$ and intermediate product C in Equations (1) and (2). According to Equations (1) and (2), there are $\tilde{l} \times \tilde{m} \times \tilde{n}$ types of reactant and product reactions. Therefore, there are $2^{2 \times \tilde{l} \times \tilde{m} \times \tilde{n}}$ reaction candidates of heterogeneous reactions.

We employ a sparse modeling approach in order to estimate rate constants $k = \left\{k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})}\right\}$ from many candidates. The sparse modeling is a framework for extracting only essential elements from candidates by assuming that the essential elements are *sparse* compared with a number of candidates [25,26]. Based on Equations (1) and (2), there are $2^{2 \times \tilde{l} \times \tilde{m} \times \tilde{n}}$ reaction candidates. By using the sparse modeling approach, we estimate each value of rate constants $k = \left\{k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})}\right\}$ and determine as zero unnecessary terms or non-zero for necessary terms. Then, we extract only the necessary heterogeneous reactions.

Figure 5 shows the schematic diagram of the estimation method for surface-area reaction in heterogeneous reactions by sparse modeling. The difference equations for heterogeneous reactions can be regarded as the following linear equation with respect to the rate constants $k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})}$; each heterogeneous reaction is multiplied by the reaction rate coefficient $k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})}$ as follows:

$$\begin{aligned}
-(C(t+1) - C(t)) = \Delta_t \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(\mathrm{r})} S_l^{(\mathrm{r})} \left(N^{(\mathrm{r})}(t)\right) \left\{C^m(t) - \left(C_{eq}^{(\mathrm{r})}\right)^m\right\}^n \\
+ \Delta_t \sum_{l=1}^{\tilde{l}} \sum_{m=1}^{\tilde{m}} \sum_{n=1}^{\tilde{n}} k_{l,m,n}^{(\mathrm{p})} S_l^{(\mathrm{p})} \left(N^{(\mathrm{p})}(t)\right) \left\{C^m(t) - \left(C_{eq}^{(\mathrm{p})}\right)^m\right\}^n
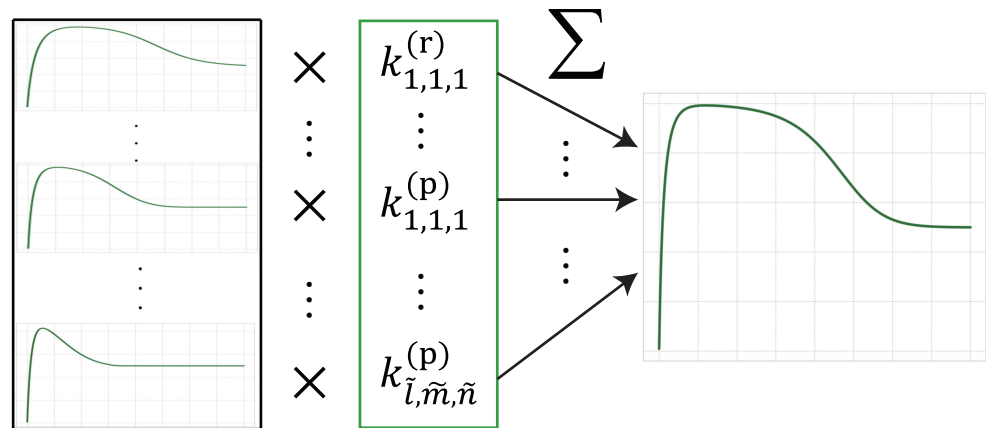\end{aligned} \tag{24}$$

Figure 5. Conceptual diagram of sparse modeling approach for nonlinear dynamics in heterogeneous reactions. The nonlinear differential equation for heterogeneous reactions (Equations (10)–(12)) can be expressed by as a linear sum of the candidate reaction terms. The proposed method, using the sparse modeling approach, extracts essential reaction terms from candidates by estimating the rate constants $k_{l,m,n}^{(\mathrm{r})}$ and $k_{l,m,n}^{(\mathrm{p})}$: zero and non-zero values for unnecessary and necessary terms, respectively.

Notably, the right-hand side of Equation (24) is expressed as a linear sum of the reaction term $\{S_l(N_i(t))\}\left\{C_i^m(t) - \left(C_{eq}\right)^m\right\}^n$ with coefficients $k_{l,m,n}$ corresponding to rate constants. To extract rate constant terms from a number of candidates, we use the Lasso framework (least absolute shrinkage and selection operator), which employs a sparse modeling approach [6,7,27,28]. By considering the L_1 regularization term for rate constants, sparse modeling of heterogeneous reactions is formulated based on the Lasso framework as follows:

$$\underset{k}{\arg\min} \, ||z_{1:I,0:T} - A_{1:I,0:T}k||_2^2 + \lambda||k||_1 \tag{25}$$

$$||k||_1 = |k_{1,1,1}^{(\mathrm{r})}| + \cdots + |k_{1,1,1}^{(\mathrm{p})}| + \cdots + |k_{\tilde{l},\tilde{m},\tilde{n}}^{(\mathrm{p})}| \tag{26}$$

where $z_{1:I,0:T}$, $A_{1:I,0:T}$ and k are represented by the following matrix and vectors:

$$z_{1:I,0:T} = \frac{1}{I}\sum_{i=1}^{I} \begin{bmatrix} -C_i(1)+C_i(0) \\ -C_i(2)+C_i(1) \\ \vdots \\ -C_i(T)+C_i(T-1) \end{bmatrix} = \begin{bmatrix} \overline{-C(1)+C(0)} \\ \overline{-C(2)+C(1)} \\ \vdots \\ \overline{-C(T)+C(T-1)} \end{bmatrix} \tag{27}$$

$$A_{1:I,0:T} = \Delta_t \begin{bmatrix} \overline{S_1^{(\mathrm{r})}(N^{(\mathrm{r})}(0))\left\{C(0)-\left(C_{eq}^{(\mathrm{r})}\right)\right\}} & \cdots & \overline{S_{\tilde{l}}^{(\mathrm{p})}\left(N^{(\mathrm{p})}(0)\right)\left\{C^{\tilde{m}}(0)-\left(C_{eq}^{(\mathrm{p})}\right)^{\tilde{m}}\right\}^{\tilde{n}}} \\ \overline{S_1^{(\mathrm{r})}(N^{(\mathrm{r})}(1))\left\{C(1)-\left(C_{eq}^{(\mathrm{r})}\right)\right\}} & \cdots & \overline{S_{\tilde{l}}^{(\mathrm{p})}\left(N^{(\mathrm{p})}(1)\right)\left\{C^{\tilde{m}}(1)-\left(C_{eq}^{(\mathrm{p})}\right)^{\tilde{m}}\right\}^{\tilde{n}}} \\ \vdots & \ddots & \vdots \\ \overline{S_1^{(\mathrm{r})}(N^{(\mathrm{r})}(T-1))\left\{C(T-1)-\left(C_{eq}^{(\mathrm{r})}\right)\right\}} & \cdots & \overline{S_{\tilde{l}}^{(\mathrm{p})}\left(N^{(\mathrm{p})}(T-1)\right)\left\{C^{\tilde{m}}(T-1)-\left(C_{eq}^{(\mathrm{p})}\right)^{\tilde{m}}\right\}^{\tilde{n}}} \end{bmatrix} \tag{28}$$

$$k = \begin{bmatrix} k_{1,1,1}^{(\mathrm{r})} \\ \vdots \\ k_{1,1,1}^{(\mathrm{p})} \\ \vdots \\ k_{\tilde{l},\tilde{m},\tilde{n}}^{(\mathrm{p})} \end{bmatrix} \tag{29}$$

In Equation (25), λ is the regularization coefficient that controls the relative importance of a data-dependent error $||z_{1:I,0:T} - A_{1:I,0:T}k||_2^2$ and the regularization term $||k||_1$. The regularization term $||k||_1$ is the sum of the absolute values of each rate constants $|k_{l,m,n}|$. When the regularization term $||k||_1$ becomes smaller, the rate constants approach zero, and we can obtain a sparse solution. Here, $z_{1:I,0:T}$ is a vector consisting of the averages of the differences $\overline{-C(t+1) + C(t)}$ between intermediate products' adjacent time steps, which are obtained by particles as the averages of the difference $-C_i(t+1) + C_i(t)$ $(i = 1, \cdots, I)$. $A_{1:I,0:T}$ denotes a matrix consisting of the average of reaction terms, which are obtained by particles $(i = 1, \cdots, I)$. k indicates a vector consisting of rate constants. We calculate the matrix $A_{1:I,0:T}$ and vector $z_{1:I,0:T}$ by using hidden variables $[N_i^{(r)}]_{i=1}^I$, $[C_i]_{i=1}^I$ and $[N_i^{(p)}]_{i=1}^I$. To obtain an appropriate value of λ, we employ nested cross-validation error since time-series data are used [29].

To compare with the Lasso framework, we consider the Ridge framework, expressed by the following equation [30]:

$$\underset{k}{\arg\min} \, ||z_{1:I,0:T} - A_{1:I,0:T}k||_2^2 + \lambda ||k||_2^2 \tag{30}$$

$$||k||_2 = \left(|k_{1,1,1}^{(r)}|^2 + \cdots + |k_{1,1,1}^{(p)}|^2 + \cdots + |k_{\tilde{l},\tilde{m},\tilde{n}}^{(p)}|^2 \right)^{1/2} \tag{31}$$

Namely, the Ridge framework introduces the L_2 regularization term. Ridge and Lasso frameworks are different in terms of regularization terms. Figure 6 shows a conceptual diagram of the Lasso and Ridge frameworks in two dimensions. Lasso yields a sparse solution because the part where the contour line of the error term and the boundary of the constraint condition meet is likely to be a corner since the constraint condition of Lasso is given by $||k||_1$. In Figure 6, k_2 is estimated to be a zero value by using Lasso. However, when Ridge is used, both k_1 and k_2 are estimated to be non-zero values. Therefore, the Lasso framework can obtain a sparse solution more effectively than the Ridge framework. Thus, we assume that the sparse modeling approach extracts only essential reactions from candidates in heterogeneous reactions by estimating each rate constant as either zero or non-zero.
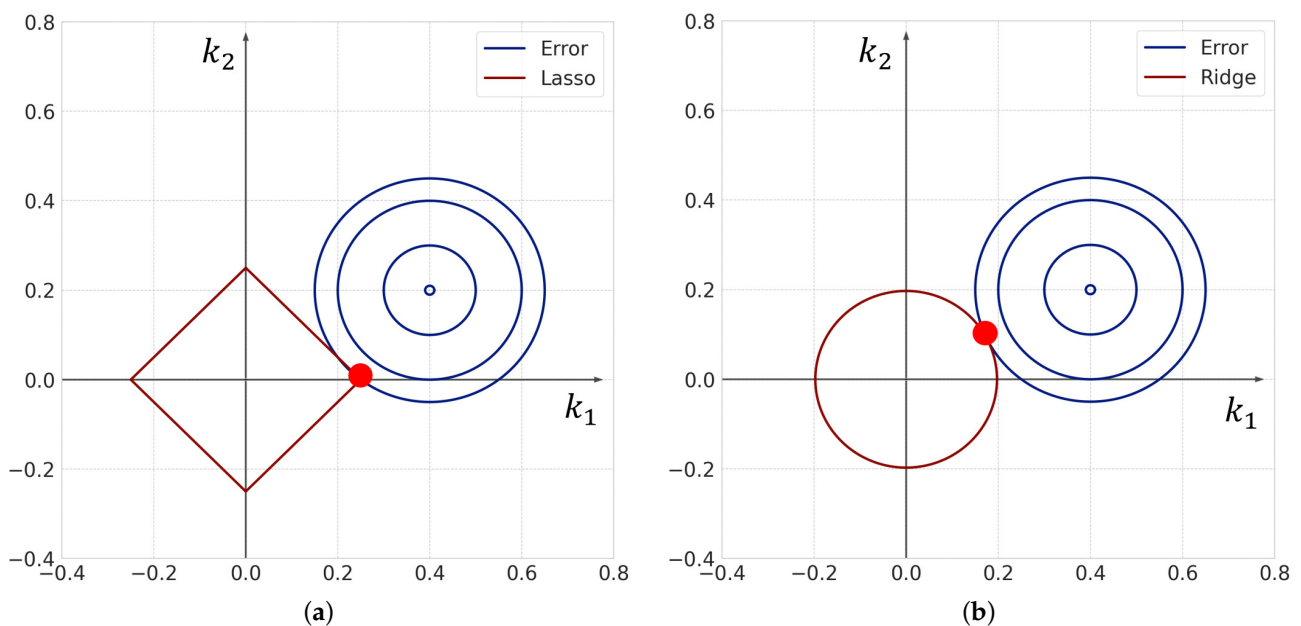

(a)   **(b)**

Figure 6. Conceptual diagram of (**a**) Lasso and (**b**) Ridge frameworks in two dimensions. The red dot in each figure indicates the estimated values in the corresponding algorithm. A sparse solution is obtained in the case of Lasso, whereas a non-sparse solution is obtained in the case of Ridge.

However, in the case where we can observe few data or many candidates can be assumed, it may not be possible to extract only important reactions. Therefore, instead of uniform sparsity levels, we introduce non-uniform ones in order to extract only the important reactions more accurately. Here, a sparse modeling approach based on adaptive Lasso is introduced [31]. The sparse modeling approach with non-uniform sparsity levels is expressed by the following equation:

$$\underset{k}{\operatorname{argmin}} \, ||z_{1:I,0:T} - A_{1:I,0:T}k||_2^2 + \lambda \sum_{j=1}^{2\tilde{I}\tilde{m}\tilde{n}} \omega_j |k_j| \tag{32}$$

$$\omega_j = \frac{1}{|\beta_j|^\gamma} \tag{33}$$

where $\boldsymbol{\beta} = \{\beta_j\}$ is a consistent estimator, and it is determined by the least-squares method and the Ridge framework. Here, k_j is j-th element of vector $\boldsymbol{k}$ shown in Equation (29). According to Equation (33), ω_j is the reciprocal of the absolute value of β_j to the γ ($\gamma > 0$) power. Here, we employ the Ridge framework represented by Equation (30) to obtain $\boldsymbol{\beta}$ and put $\gamma = 1$. If the value of $|\beta_j|$ is small, the values of $|\omega_j|$ become large. Therefore, the sparsity tends to increase to realize $|k_j| = 0$. In contrast, if the value of $|\beta_j|$ is large, the value of $|\omega_j|$ becomes small, and the sparsity tends to decrease to realize $|k_j| \neq 0$.

Figure 7 shows a conceptual diagram of the sparse modeling approach with uniform and non-uniform sparsity levels for a two-dimensional vector. For the sparse modeling approach with non-uniform sparsity levels, the regularization term has non-uniformity, due to the weight vector obtained by the consistent estimator. The intersection of the contour line of the error term and regularization term of non-uniform sparsity is $(k_1, k_2) = (k_1', 0)$. Therefore, non-uniform sparsity yields a sparse solution. However, for the sparse modeling approach with uniform sparsity levels, the intersection of the contour line of the error term and regularization term of uniform sparsity is $(k_1, k_2) = (k_1', k_2')$ where both elements of the weight vector become non-zero ($k_1' \neq 0$ and $k_2' \neq 0$). Therefore, sparse modeling with uniform sparsity may not yield a sparse solution. Thus, we employ the sparse modeling approach with non-uniform sparsity levels in order to extract only the important reactions more accurately.
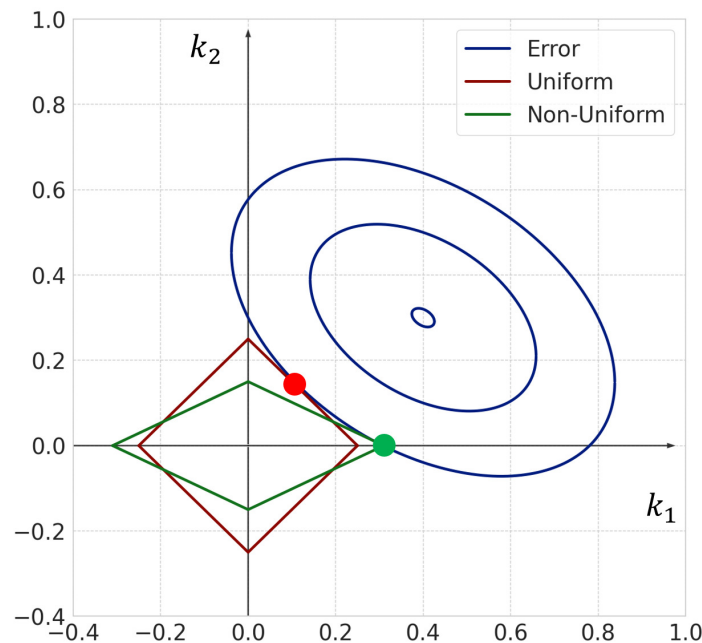

Figure 7. Conceptual diagram of sparse modeling with uniform and non-uniform sparsity levels. A sparse solution is obtained using sparse modeling with non-uniform sparsity levels.

3. Results

In this section, we evaluate the effectiveness of our proposed method using simulation data. The proposed method is a combined method based on sparse modeling and sequential Monte Carlo approaches. Multi-dimensional hidden variables consisting of solid reactant $N^{(\mathrm{r})}$, liquid intermediate product C, and solid product $N^{(\mathrm{p})}$ are estimated from a partially observed liquid intermediate product C' by using the sequential Monte Carlo method. We employ the sparse modeling approach to estimate the rate constants $k = \left\{ k^{(\mathrm{r})}_{l,m,n}, k^{(\mathrm{p})}_{l,m,n} \right\}$. By conducting these two calculations alternately, we simultaneously estimate $N^{(\mathrm{r})}, C, N^{(\mathrm{p})}$ and $k = \left\{ k^{(\mathrm{r})}_{l,m,n}, k^{(\mathrm{p})}_{l,m,n} \right\}$.

We assume that the data $C'(t)$ are available at constant time step intervals within total time steps T. In Equation (8), the observation noise $\xi^C(t)$ is assumed to be a Gaussian noise with zero means and standard deviation 1×10^{-3}. Further, in Equations (13) and (14), the system noise $\xi^{(\mathrm{r})}(t)$ and $\xi^{(\mathrm{p})}(t)$ are also assumed to be Gaussian noise with zero mean and standard deviation 1×10^{-4}. Since the reaction time is $0 \le t \le 4000$ and the time interval is $\Delta_t = 1$, the maximum number of observed data is 4001 points. We set $\tilde{l} = 3, \tilde{m} = 2, \tilde{n} = 2$ in Equations (13) and (14). Note that $\tilde{l}$ denotes the number of candidates for the type of surface-area model, $\tilde{m}$ is the number of candidates for a multiplier of the intermediate product C, and $\tilde{n}$ is the number of candidates for a multiplier of the reaction term $\left\{ C^m - \left(C_{eq} \right)^m \right\}$. Therefore, the number of rate constants k is $2 \times \tilde{l} \times \tilde{m} \times \tilde{n} = 2 \times 3 \times 2 \times 2 = 24$ and we assume 2^{24} nonlinear or linear candidates. The value of each parameter used to obtain the simulation data is shown in Table 1.

Table 1. Parameter values for simulations.

Parameter	Value	Parameter	Value
$C^{(\mathrm{r})}_{eq}$	0.5	$C^{(\mathrm{p})}_{eq}$	0.25
$k^{(\mathrm{r})}_{1,1,1}$	0	$k^{(\mathrm{p})}_{1,1,1}$	0
$k^{(\mathrm{r})}_{1,1,2}$	0	$k^{(\mathrm{p})}_{1,1,2}$	0
$k^{(\mathrm{r})}_{2,1,1}$	0	$k^{(\mathrm{p})}_{2,1,1}$	0
$k^{(\mathrm{r})}_{2,1,2}$	0	$k^{(\mathrm{p})}_{2,1,2}$	0
$k^{(\mathrm{r})}_{3,1,1}$	1.93×10^{-2}	$k^{(\mathrm{p})}_{3,1,1}$	9.66×10^{-3}
$k^{(\mathrm{r})}_{3,1,2}$	0	$k^{(\mathrm{p})}_{3,1,2}$	0
$k^{(\mathrm{r})}_{1,2,1}$	0	$k^{(\mathrm{p})}_{1,2,1}$	0
$k^{(\mathrm{r})}_{1,2,2}$	0	$k^{(\mathrm{p})}_{1,2,2}$	0
$k^{(\mathrm{r})}_{2,2,1}$	0	$k^{(\mathrm{p})}_{2,2,1}$	0
$k^{(\mathrm{r})}_{2,2,2}$	0	$k^{(\mathrm{p})}_{2,2,2}$	0
$k^{(\mathrm{r})}_{3,2,1}$	0	$k^{(\mathrm{p})}_{3,2,1}$	0
$k^{(\mathrm{r})}_{3,2,2}$	0	$k^{(\mathrm{p})}_{3,2,2}$	0

We set the initial values of the hidden variables $N^{(\mathrm{r})}(0) = 0.99$, $C(0) = 0.005$, $N^{(\mathrm{p})}(0) = 0.005$ and the initial values of rate constants k random values near zero. From Equations (13)–(15) and Table 1, we generate the simulation data, and the reaction is consistent with the typical heterogeneous reaction dynamics shown in Figure 3. We employ the sparse modeling and sequential Monte Carlo method approaches alternately for a sufficient number of times, and use values with which the generalization error, using cross-validation, is the smallest as the estimated results.

3.1. Simultaneous Estimation of Hidden Variables and Rate Constant

Here, we employ the proposed method to estimate the hidden variables and the rate constants by using 4000 observation points. The observation values under the conditions are shown in Figure 8. We employ the sparse modeling with uniform sparsity levels and that with non-uniform sparsity levels, which are our proposed methods, in order to estimate rate constants $\boldsymbol{k}$. We compare results with those of two other methods: Ridge in Equation (30) and the least-squares method.

Table 2 lists the estimation results of each rate constants obtained using different methods. We find that by using the least-squares method, the estimated values based on the sequential Monte Carlo method diverge. Note that the values for the least-squares method listed in Table 2 are estimated rate constants just before the values diverge. The reaction rate constants estimated by sparse modeling with uniform sparsity levels and that with non-uniform sparsity levels are also shown in Figure 9. Figure 10 shows the estimated hidden variables based on the sequential Monte Carlo method obtained by using the estimated rate constants in Table 2. Table 2 also shows the mean squared error (MSE) between the estimated hidden variables and simulation data.
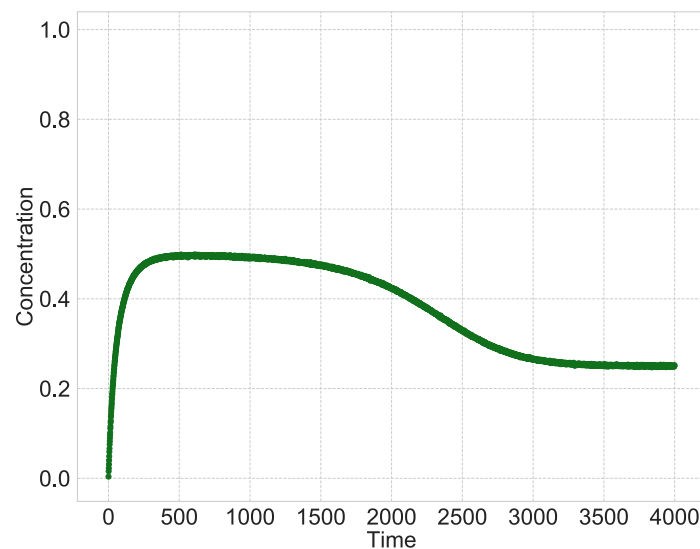

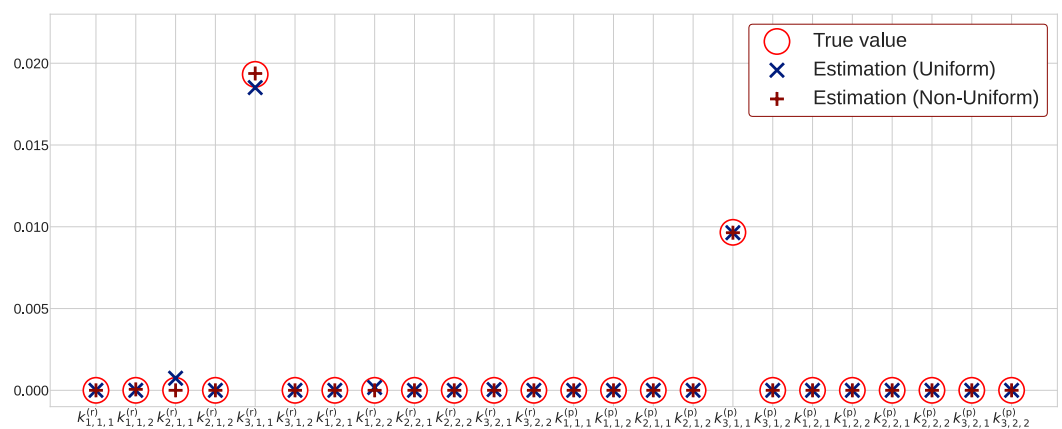**Figure 8.** Observation variable $C'(t)$ for observation points $T = 4000$.

Figure 9. Estimated rate constants for observation points $T = 4000$.

Table 2. Estimated rate constants for the number of data points $T = 4000$.

	True Value	Least Squares	Ridge	Uniform	Non-Uniform
MSE		—	5.55×10^{-3}	2.70×10^{-6}	2.42×10^{-6}
$k_{1,1,1}^{(r)}$	0	-5.39×10^{8}	-2.85×10^{-3}	0	0
$k_{1,1,2}^{(r)}$	0	1.72×10^{9}	-1.86×10^{-4}	0	5.82×10^{-5}
$k_{2,1,1}^{(r)}$	0	4.91×10^{8}	6.95×10^{-3}	7.40×10^{-4}	0
$k_{2,1,2}^{(r)}$	0	4.91×10^{8}	-2.02×10^{-3}	0	0
$k_{3,1,1}^{(r)}$	1.93×10^{-2}	-8.04×10^{8}	8.28×10^{-3}	1.85×10^{-2}	1.94×10^{-2}
$k_{3,1,2}^{(r)}$	0	-8.04×10^{8}	-2.28×10^{-3}	0	0
$k_{1,2,1}^{(r)}$	0	1.29×10^{9}	-3.03×10^{-3}	0	0
$k_{1,2,2}^{(r)}$	0	-2.53×10^{7}	4.94×10^{-4}	1.70×10^{-4}	0
$k_{2,2,1}^{(r)}$	0	-4.91×10^{8}	4.93×10^{-3}	0	0
$k_{2,2,2}^{(r)}$	0	5.74	-7.74×10^{-4}	0	0
$k_{3,2,1}^{(r)}$	0	8.04×10^{8}	6.00×10^{-3}	4.58×10^{-5}	0
$k_{3,2,2}^{(r)}$	0	-6.28	-9.66×10^{-4}	0	0
$k_{1,1,1}^{(p)}$	0	4.12×10^{9}	5.21×10^{-4}	0	0
$k_{1,1,2}^{(p)}$	0	3.73×10^{9}	-7.67×10^{-4}	0	0
$k_{2,1,1}^{(p)}$	0	3.59×10^{8}	2.63×10^{-3}	0	0
$k_{2,1,2}^{(p)}$	0	7.18×10^{8}	-1.60×10^{-4}	0	0
$k_{3,1,1}^{(p)}$	9.66×10^{-3}	3.51×10^{7}	2.84×10^{-3}	9.64×10^{-3}	9.64×10^{-3}
$k_{3,1,2}^{(p)}$	0	7.01×10^{7}	4.75×10^{-5}	0	0
$k_{1,2,1}^{(p)}$	0	-6.74×10^{9}	-5.06×10^{-4}	0	0
$k_{1,2,2}^{(p)}$	0	2.53×10^{7}	-1.69×10^{-4}	0	0
$k_{2,2,1}^{(p)}$	0	-7.18×10^{8}	1.15×10^{-3}	0	0
$k_{2,2,2}^{(p)}$	0	6.67×10^{-1}	-1.58×10^{-4}	0	0
$k_{3,2,1}^{(p)}$	0	-7.01×10^{7}	1.47×10^{-3}	0	0
$k_{3,2,2}^{(p)}$	0	-1.94	-3.80×10^{-5}	0	0

First, we evaluate the estimation result of the rate constants. Table 2 shows that the rate constants estimated by the least squares method are quite different from true values. Although the rate constants estimated with Ridge does not deviate as much as those with the least-squares method, they have many non-zero rate constants. Therefore, it is difficult to determine which reaction is the most important. On the other hand, the proposed methods based on uniform sparsity levels and non-uniform sparsity levels can obtain many zero values for rate constants, and it is possible to extract important reactions. Here, we compare the two sparse modeling methods regarding the number of non-zero rate constants. Table 2 reveals that the method with non-uniform sparsity levels provides three non-zero elements, whereas the method with uniform sparsity levels provides five non-zero elements. Therefore, we find that the proposed method with non-uniform sparsity levels can extract important reactions better than that with uniform sparsity levels. Notably, the proposed method can be applied to other cases with different true reaction terms in order to extract only essential reaction terms from the candidate reaction terms.
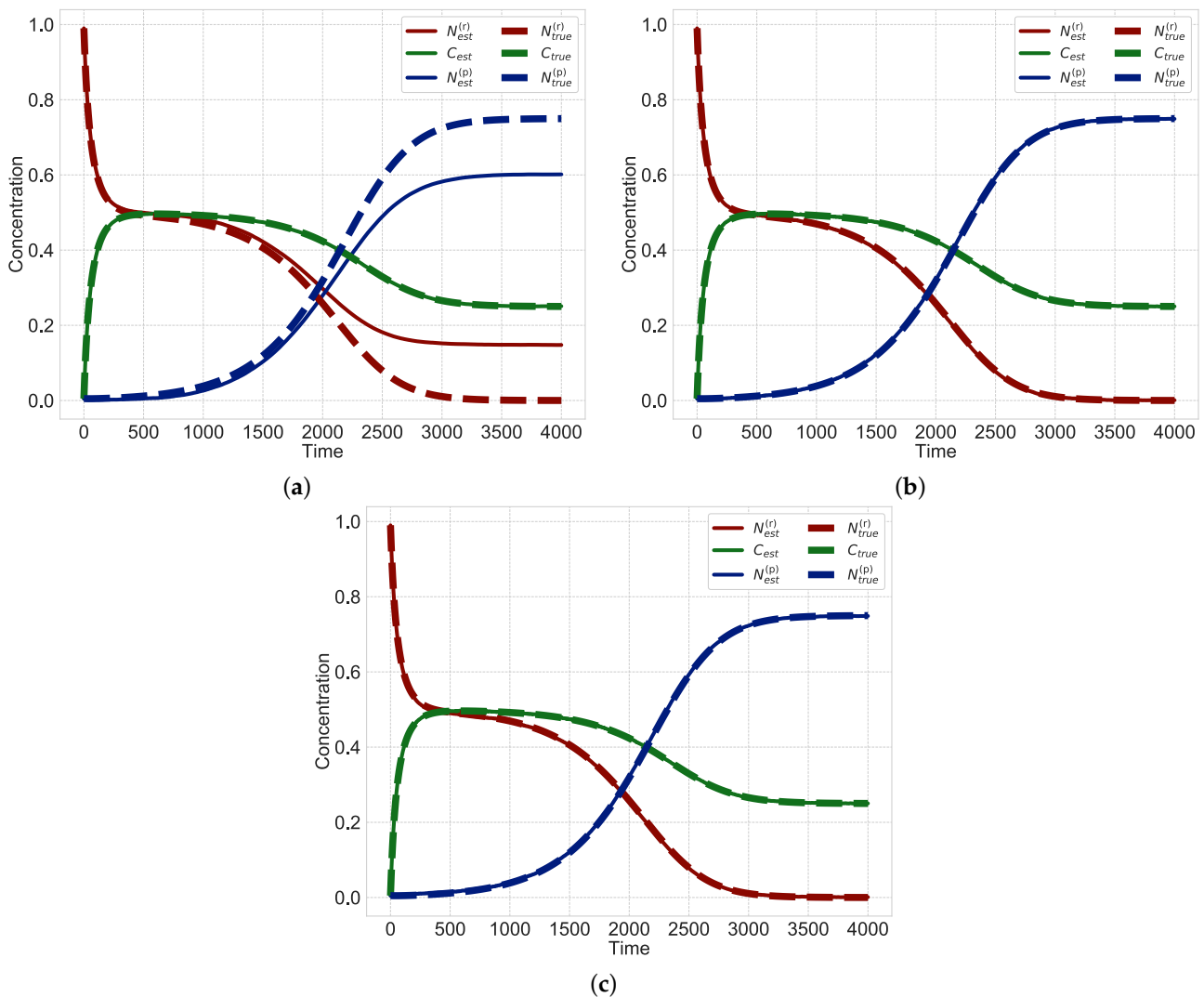
Figure 10. Estimated hidden variables $N^{(\mathrm{r})}$, C and $N^{(\mathrm{p})}$ for data points $T = 4000$. The dotted and solid lines represent the true and estimated values, respectively. (**a**) Ridge; (**b**) Uniform sparsity; (**c**) Non-uniform sparsity.

Next, we evaluate hidden variables estimated by using the sequential Monte Carlo method based on the obtained rate constants. In Figure 10, the dotted and solid lines represent the true and estimated values, respectively. Since the dotted and solid lines in Figure 10a do not align, hidden variables cannot be well-estimated using Ridge. In contrast, the dotted and solid lines in the results obtained from sparsity modeling with uniform sparsity levels and non-uniform sparsity levels almost overlap (Figure 10b,c). These results imply that the proposed method can estimate three hidden variables $N^{(\mathrm{r})}$, C and $N^{(\mathrm{p})}$ from observed data C'. We also evaluate the mean squared error (MSE) in Table 2. Since the MSE obtained from the method with uniform sparsity levels is larger than that with non-uniform sparsity levels, the estimated results obtained from the method with non-uniform sparsity levels are found to be better than those with uniform sparsity levels.

Therefore, the proposed method can simultaneously estimate three hidden variables $N^{(\mathrm{r})}$, C and $N^{(\mathrm{p})}$ as well as rate constants $\boldsymbol{k} = \left\{ k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})} \right\}$. Moreover, the sparse modeling method with non-uniform sparsity levels is better than that with uniform sparsity levels in estimating hidden variables and rate constants.

3.2. Simultaneous Estimation of Hidden Variables and Rate Constant for a Small Number of Observation Points

Here, we consider a situation where the number of observation points is limited. We validate the proposed method under the condition that the number of observation points is less than that in Section 3.1. We set the number of observation points to 100. The observation values under the conditions are shown in Figure 11. The number of observation points is small and sparse compared to that in Figure 8, which has 4000 observation points. We estimate heterogeneous reactions using the proposed method with uniform and non-uniform sparsity levels in order to estimate the rate constants k. The estimated rate constants are shown in Table 3. The reaction rate constants estimated using the method with uniform sparsity levels and that with non-uniform sparsity levels are also shown in Figure 12. We further use estimated rate constants in Table 3 to calculate the hidden variables based on the sequential Monte Carlo method, and the results are shown in Figure 13. Additionally, Table 3 shows the MSE between the estimated hidden variables and simulation data.
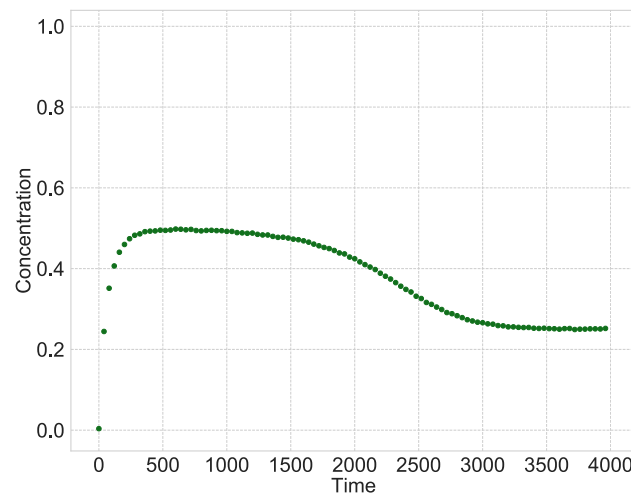

Figure 11. Observation variable $C'(t)$ for observation points $T = 100$.
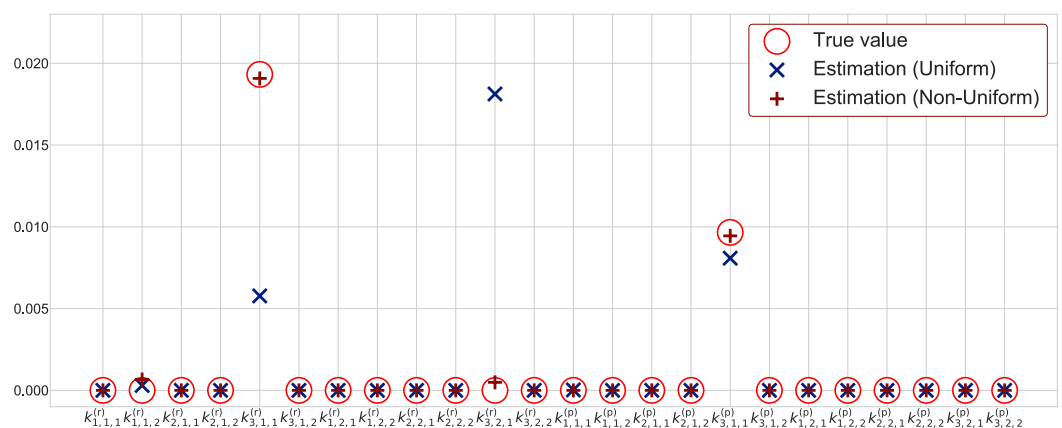

Figure 12. Estimated rate constants for observation points $T = 100$.

First, we evaluate the estimation result of the rate constants. From Figure 12, non-zero values of rate constants estimated by uniform sparsity show $k_{3,2,1}^{(\mathrm{r})}$, $k_{3,1,1}^{(\mathrm{P})}$ and $k_{3,1,1}^{(\mathrm{r})}$ in descending order. However, the rate constant $k_{3,2,1}^{(\mathrm{r})}$ must be zero. Therefore, the method with uniform sparsity levels cannot estimate rate constants accurately when the number of observation points is 100. On the other hand, the proposed methods based on non-uniform sparsity levels can well estimate the rate constants at such a condition since the true and

estimated values almost overlap, as shown in Figure 12. Additionally, we compare the proposed two sparse modeling methods in terms of the number of non-zero values of the rate constants. Table 3 shows that the method with non-uniform sparsity levels provides four non-zero elements, and that with uniform sparsity levels provides five non-zero elements. Therefore, even when the number of observation points is small, the method with non-uniform sparsity levels can extract important reactions better than that with uniform sparsity levels.

Table 3. Estimated rate constants for observation points $T = 100$.

	True Value	**Uniform**	**Non-Uniform**
MSE		3.18×10^{-4}	1.52×10^{-5}
$k_{1,1,1}^{(\mathrm{r})}$	0	0	0
$k_{1,1,2}^{(\mathrm{r})}$	0	3.02×10^{-4}	6.63×10^{-4}
$k_{2,1,1}^{(\mathrm{r})}$	0	0	0
$k_{2,1,2}^{(\mathrm{r})}$	0	0	0
$k_{3,1,1}^{(\mathrm{r})}$	1.93×10^{-2}	5.77×10^{-3}	1.91×10^{-2}
$k_{3,1,2}^{(\mathrm{r})}$	0	0	0
$k_{1,2,1}^{(\mathrm{r})}$	0	0	0
$k_{1,2,2}^{(\mathrm{r})}$	0	0	0
$k_{2,2,1}^{(\mathrm{r})}$	0	0	0
$k_{2,2,2}^{(\mathrm{r})}$	0	0	0
$k_{3,2,1}^{(\mathrm{r})}$	0	1.81×10^{-2}	4.90×10^{-4}
$k_{3,2,2}^{(\mathrm{r})}$	0	0	0
$k_{1,1,1}^{(\mathrm{p})}$	0	2.53×10^{-5}	0
$k_{1,1,2}^{(\mathrm{p})}$	0	0	0
$k_{2,1,1}^{(\mathrm{p})}$	0	0	0
$k_{2,1,2}^{(\mathrm{p})}$	0	0	0
$k_{3,1,1}^{(\mathrm{p})}$	9.66×10^{-3}	8.08×10^{-3}	9.45×10^{-3}
$k_{3,1,2}^{(\mathrm{p})}$	0	0	0
$k_{1,2,1}^{(\mathrm{p})}$	0	0	0
$k_{1,2,2}^{(\mathrm{p})}$	0	0	0
$k_{2,2,1}^{(\mathrm{p})}$	0	0	0
$k_{2,2,2}^{(\mathrm{p})}$	0	0	0
$k_{3,2,1}^{(\mathrm{p})}$	0	0	0
$k_{3,2,2}^{(\mathrm{p})}$	0	0	0

Next, we evaluate the hidden variables estimated by using the sequential Monte Carlo method, based on the estimated rate constants. Since time courses of the estimated variables (the dotted line) and that of true variables (the solid line) are slightly different in Figure 13a, we find that the proposed method with uniform sparsity levels cannot well estimate hidden variables. In contrast, as shown in Figure 13b, the estimation results of the proposed method with non-uniform sparsity levels show that the time courses of the estimated variables (dotted line) and true variables (solid line) almost overlap. These results show that the proposed method with non-uniform sparsity levels can estimate three hidden variables $N^{(\mathrm{r})}$, C and $N^{(\mathrm{p})}$ from observed data C'. We then evaluate the MSE in Table 3. Since the MSE of the method with uniform sparsity levels is higher than that with non-uniform sparsity levels, we see that the method with non-uniform sparsity levels

is better than that with uniform sparsity levels when the number of observation points is small.
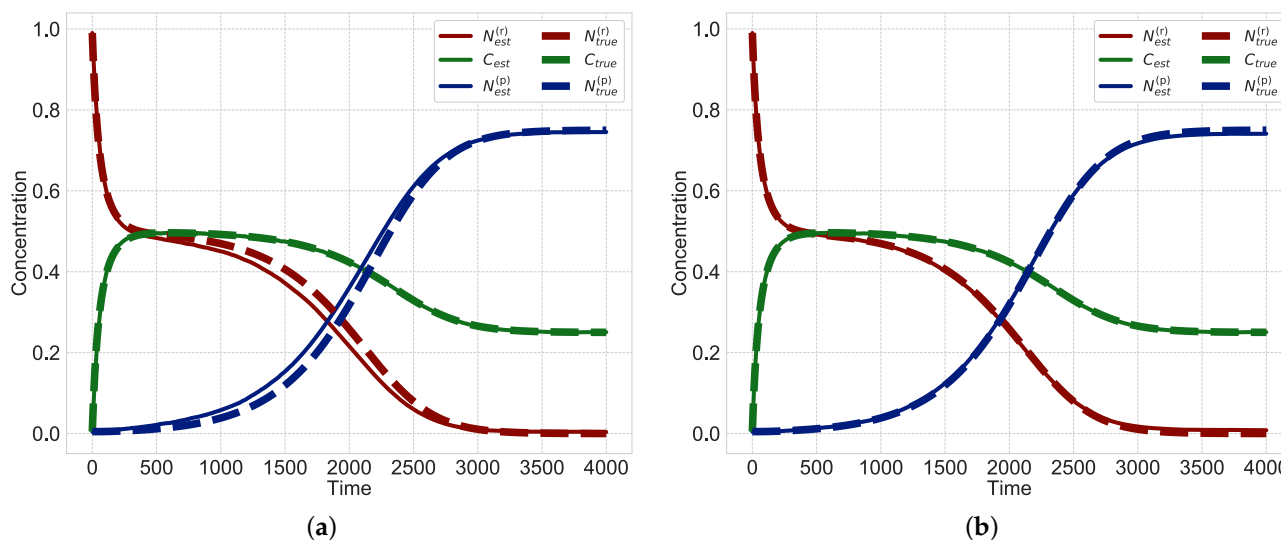

(a)

(b)

Figure 13. Estimated hidden variables for the number of observation points $T = 100$. The dotted lines represent true values and solid lines represent estimated values. (**a**) Uniform sparsity; (**b**) Non-uniform sparsity.

These results show that even when the number of observation points is small, the proposed method with non-uniform sparsity levels can estimate simultaneously three hidden variables $N^{(\mathrm{r})}$, C and $N^{(\mathrm{p})}$ and rate constants $\boldsymbol{k} = \left\{ k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})} \right\}$.

3.3. Dependence of the Estimation Accuracy on the Number of Observation Points

Here, we investigate how accurately the rate constants $\boldsymbol{k} = \left\{ k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})} \right\}$ can be estimated for different numbers of observations. We estimate the rate constants for different numbers of observation points, including $T = 4000, 1000, 400, 100$ and 40, to consider the effectiveness of the proposed method with uniform sparsity levels and that with non-uniform sparsity levels. We focus on the difference between the true and estimated rate constants. Figure 14 shows the degree of non-zero and zero agreement, and Figure 15 shows the MSE. Since we assume 24 kinds of rate constants, the maximum degree of agreement is 24. Figure 16 shows the MSE between the estimated hidden variables and simulation data for different numbers of observation points. Both MSE in Figures 15 and 16 are expressed using the logarithms to base 10.

First, we evaluate the estimated rate constants. Figure 14 shows that when the number of observation points is varied from 4000 to 100, there is not much difference in the degree of agreement of rate constants between the proposed method with non-uniform sparsity levels and that with uniform sparsity levels. This result indicates that we can select non-zero and zero rate constants. Therefore, we can extract important reactions from many candidates when the number of observable points is sufficiently large by means of both the proposed methods. On the other hand, Figure 15 shows that the proposed method with non-uniform sparsity levels has a smaller MSE of rate constants than that with uniform sparsity levels for all observation points. Figures 14 and 15 reveal that the proposed method with non-uniform sparsity levels shows small discrepancy between the estimated and true rate constants $\boldsymbol{k} = \left\{ k_{l,m,n}^{(\mathrm{r})}, k_{l,m,n}^{(\mathrm{p})} \right\}$ compared with that with uniform sparsity levels. When the number of observation points is 40, both methods have low accuracy in the degree of coincidence and the MSE. Therefore, it is difficult to estimate rate constants when the number of observation points is 40, and we consider around 100 points as the limit to estimate rate constants when candidates of heterogeneous reaction terms assumed in our study are given.
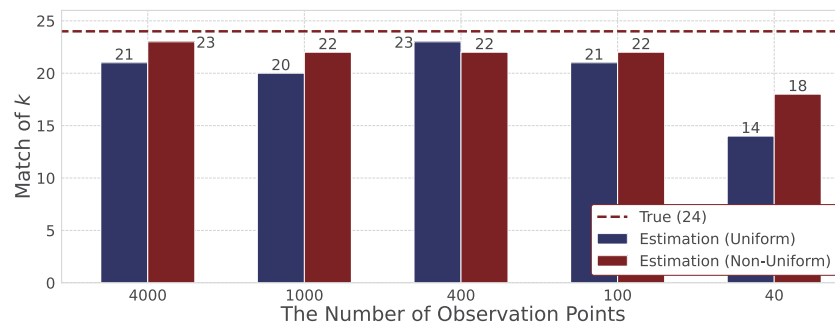
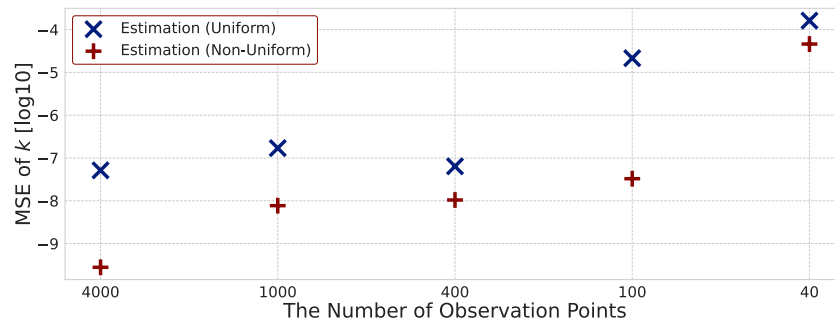**Figure 14.** The degree of non-zero and zero agreement for different numbers of observation points.

Figure 15. Mean squared error between true and estimated rate constants for different numbers of observation points.
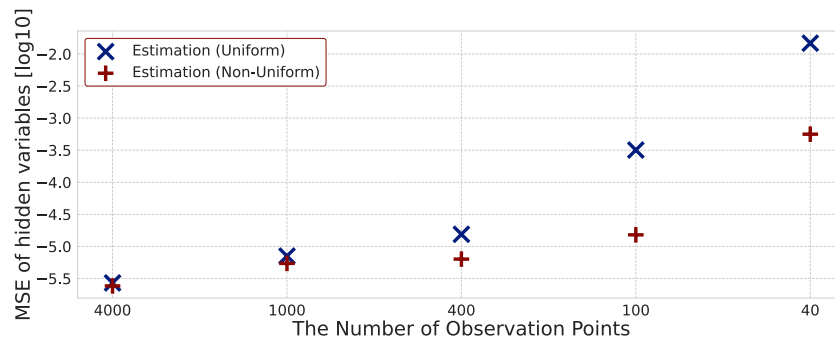

Figure 16. Mean squared error between true and estimated hidden variables for different numbers of observation points.

Next, we consider the MSE between the simulation data and estimated hidden variables. Figure 16 shows that the MSE increases as the number of observation points decreases with both non-uniform and uniform sparsity levels. This result indicates that the proposed method can estimate hidden variables more accurately as the number of observation points increases. In addition, as the number of observation points decreases, the difference in the MSE between the two methods increases. Therefore, we find that when the number of observation points decreases, the proposed method using non-uniform sparsity levels is more effective.

To summarize, the proposed method with non-uniform sparsity levels can estimate rate constants accurately from partially observable data; Using time series of observed data points of an intermediate product, we have shown that the proposed method can extract essential reaction terms from candidates; the rate constants of necessary terms are estimated to be non-zeros, whereas those of unnecessary terms are estimated to be zeros.

4. Concluding Remarks

We have proposed an algorithm based on the sparse modeling and sequential Monte Carlo approaches for estimating heterogeneous reaction dynamics. We focused on heterogeneous reactions, which depend on the surface-area reactions between solid and liquid phases, and we developed a nonlinear state-space model for the heterogeneous reactions. We considered three aspects: l $(l = 1, 2, \cdots, \tilde{l})$ types of surface-area reactions, the multiplier of the intermediate product C^m $(m = 1, 2, \cdots, \tilde{m})$ and the multiplier of the reaction term $\left\{ C^m - \left(C_{eq}\right)^m \right\}^n$ $(n = 1, 2, \cdots \tilde{n})$. The results in this study showed that the proposed sparse modeling method can extract essential reaction terms from candidates; the rate constants of necessary terms are estimated to be non-zeros and those of unnecessary terms are estimated to be zeros. By introducing the sequential Monte Carlo algorithm, we estimated the multi-dimensional hidden variables consisting of the solid reactant, liquid intermediate product, and solid product from the partially observed data of the liquid intermediate product. The results have shown that the proposed method with non-sparsity levels simultaneously estimates the time course of hidden variables and rate constants accurately by extracting sufficient reaction terms from the candidates.

Author Contributions: Conceptualization, T.O.; investigation, M.I., T.K., R.O. and T.O.; formal analysis, M.I. and T.O.; methodology, M.I. and T.O.; software, M.I. and T.O.; supervision, T.O.; writing—original draft preparation, M.I.; writing—review and editing, M.I., T.K., R.O. and T.O.; All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Strogatz, S.H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2019.
2. Lasaga, A.C. *Kinetic Theory in the Earth Sciences*; Princeton University Press: Princeton, NJ, USA, 1998.
3. Atkins, P.; Paula, J.D. *Physical Chemistry*, 10th ed.; Oxford University Press: Oxford, UK, 2014.
4. Omori, T.; Kuwatani, T.; Okamoto, A.; Hukushima, K. Bayesian inversion analysis of nonlinear dynamics in surface heterogeneous reactions. *Phys. Rev. E* **2016**, *94*, 33305. [CrossRef] [PubMed]
5. Oyanagi, R.X.; Kuwatani, T.; Omori, T. Exploration of nonlinear parallel heterogeneous reaction pathways through Bayesian variable selection. *Eur. Phys. J. B* **2021**, *94*, 42. [CrossRef]
6. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. [CrossRef]
7. Rish, I.; Grabarnik, G. *Sparse Modeling: Theory, Algorithms, and Applications*; CRC Press: Boca Raton, FL, USA, 2014.
8. Gregory, P. *Bayesian Logical Data Analysis for the Physical Sciences*; Cambridge University Press: Cambridge, UK, 2005.
9. Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **2010**, *19*, 2861–2873. [CrossRef] [PubMed]
10. Omori, T.; Hukushima, K. Extracting nonlinear spatiotemporal dynamics in active dendrites using data-driven statistical approach. *J. Phys. Conf. Ser.* **2016**, *699*, 12011. [CrossRef]
11. Otsuka, S.; Omori, T. Estimation of neuronal dynamics based on sparse modeling. *Neural Netw.* **2019**, *109*, 137–146. [CrossRef] [PubMed]
12. Yokoi, M.; Omori, T. Sparse modeling approach for estimating odor pleasantness from multi-dimensional sensor data. In Proceedings of the IEEE 2nd Global Conference on Life Sciences and Technologies, Kyoto, Japan, 10–12 March 2020; pp. 187–188.

13. Honma, M.; Akiyama, K.; Tazaki, F.; Kuramochi, K.; Ikeda, S.; Hada, K.; Uemura, M. Imaging black holes with sparse modeling. *J. Physics: Conf. Ser.* **2016**, *699*, 012006. [CrossRef]

14. Kuwatani, T.; Yoshida, K.; Ueki, K.; Oyanagi, R.; Uno, M.; Akaho, S. Sparse isocon analysis: A data-driven approach for material transfer estimation. *Chem. Geol.* **2020**, *532*, 119345. [CrossRef]

15. Rüpke, L.H.; Morgan, J.P.; Hort, M.; Connolly, J.A. Serpentine and the subduction zone water cycle. *Earth Planet. Sci. Lett.* **2004**, *223*, 17–34. [CrossRef]

16. Normand, C.; Williams-Jones, A.E.; Martin, R.F.; Vali, H. Hydrothermal alteration of olivine in a flow-through autoclave: Nucleation and growth of serpentine phases. *Am. Mineral.* **2002**, *87*, 1699–1709. [CrossRef]

17. Williams-Jones, A. Experimental water-rock interaction: Applications to ore-forming hydrothermal systems. In *Alteration and Alteration Processes Associated with Ore-Forming Systems*; Geological Assn of Canada: St. John's, NL, Canada, 1994; pp. 131–160.

18. Oyanagi, R.; Okamoto, A.; Tsuchiya, N. Silica controls on hydration kinetics during serpentinization of olivine: Insights from hydrothermal experiments and a reactive transport model. *Geochim. Cosmochim. Acta* **2020**, *270*, 21–42. [CrossRef]

19. Okamoto, A.; Ogasawara, Y.; Ogawa, Y.; Tsuchiya, N. Progress of hydration reactions in olivine–H_2O and orthopyroxenite–H_2O systems at 250 °C and vapor-saturated pressure. *Chem. Geol.* **2011**, *289*, 245–255. [CrossRef]

20. Oyanagi, R.; Okamoto, A.; Tsuchiya, N. Multiple kinetic parameterization in a reactive transport model using the exchange Monte Carlo method. *Minerals* **2018**, *8*, 579. [CrossRef]

21. Kitagawa, G. Non-gaussian state-space modeling of nonstationary time series. *J. Am. Stat. Assoc.* **1987**, *82*, 1032–1041.

22. Doucet, A.; de Freitas, N.; Gordon, N. *Sequential Monte Carlo Methods in Practice*; Springer: Berlin/Heidelberg, Germany, 2001.

23. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Applications*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2005.

24. West, M.; Harrison, J. *Bayesian Forecasting and Dynamic Models*; Springer Science + Business Media: Berlin/Heidelberg, Germany, 2006.

25. Elad, M. *Sparse and Redundant Representations*; Springer: Berlin/Heidelberg, Germany, 2010.

26. Starck, J.L.; Murtagh, F.; Fadili, J.M. *Sparse Image and Signal Processing*; Cambridge University Press: Cambridge, UK, 2010.

27. Natarajan, B.K. Sparse approximate solutions to linear systems. *SIAM J. Comput.* **1995**, *24*, 227–234. [CrossRef]

28. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.

29. Varma, S.; Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* **2006**, *7*, 91. [CrossRef] [PubMed]

30. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]

31. Zou, H. The adaptive lasso and its oracle properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429. [CrossRef]