Normal mode analysis of a relaxation process with Bayesian inference

Itsumi Sakata\textsuperscript{a}, Yoshihiro Nagano\textsuperscript{b}, Yasuhiro Igarashi\textsuperscript{a,b,c}, Shin Murata\textsuperscript{b}, Kohji Mizoguchi\textsuperscript{e}, Ichiro Akai\textsuperscript{f} and Masato Okada\textsuperscript{a,b,c}

\textsuperscript{a}Graduate School of Science, The University of Tokyo, Tokyo, Japan;\textsuperscript{b}Graduate School of Frontier Science, The University of Tokyo, Chiba, Japan;\textsuperscript{c}Research and Services Division of Materials Data and Integrated System, National Institute for Material Science (NIMS), Tsukuba, Japan;\textsuperscript{d}Japan Science and Technology Agency, PRESTO, Saitama, Japan;\textsuperscript{e}Graduate School of Science, Osaka Prefecture University, Osaka, Japan;\textsuperscript{f}Institute of Pulsed Power Science, Kumamoto University, Kumamoto, Japan

ABSTRACT

Measurements of relaxation processes are essential in many fields, including nonlinear optics. Relaxation processes provide many insights into atomic/molecular structures and the kinetics and mechanisms of chemical reactions. For the analysis of these processes, the extraction of modes that are specific to the phenomenon of interest (normal modes) is unavoidable. In this study we propose a framework to systematically extract normal modes from the viewpoint of model selection with Bayesian inference. Our approach consists of a well-known method called sparsity-promoting dynamic mode decomposition, which decomposes a mixture of damped oscillations, and the Bayesian model selection framework. We numerically verify the performance of our proposed method by using coherent phonon signals of a bismuth polycrystal and virtual data as typical examples of relaxation processes. Our method succeeds in extracting the normal modes even from experimental data with strong backgrounds. Moreover, the selected set of modes is robust to observation noise, and our method can estimate the level of observation noise. From these observations, our method is applicable to normal mode analysis, especially for data with strong backgrounds.

1. Introduction

Relaxation processes are processes by which a perturbed system returns to equilibrium. Measurements of relaxation processes are essential in many fields, such as condensed matter physics \cite{1,2}, nonlinear optics \cite{3}, structural/reaction chemistry \cite{4}, astronomy \cite{5} and atmospheric science \cite{6}. Such processes provide many insights into atomic/molecular structures and the kinetics and mechanisms of chemical reactions. For example, Kim et al. \cite{3} reported evidence of a Mott phase transition by analyzing the structure of a VO\textsubscript{2} metal crystal from a coherent phonon (CP) signal, which is the relaxation of phonon vibrations. Blagg et al. \cite{7} recently reported the relaxation path and stable structure of the thermal magnetization of a lanthanide compound by analyzing X-ray...
diffraction patterns. Common to all these areas, the extraction of modes that are specific to the phenomenon of interest (normal modes) from the measured data is unavoidable.

In this study, we focus on the CP signal as a typical case in which the measurement data of a relaxation process include normal modes. Coherent phonons are waves of in-phase atomic vibrations over a macroscopic spatial range, and their signals can be observed by the pump/probe method [8,9], which uses two ultrashort pulses of light. After the excitation of a substance by the pump pulse, lattice vibrations can be observed as oscillatory changes in optical constants with a delayed probe pulse; this approach has been used to reveal the dynamics of a photoinduced structural phase transition [10–12].

Figure 1(a) shows an example of a (111)-Bi polycrystal as a reference sample [13]. The (111)-Bi polycrystal has the $A_{1g}$ and $E_c$ coherent phonon modes. Figure 1(b,c) show typical observations of the CP signal in a (111)-Bi thin film with the same experimental conditions. The difficulty in analyzing the relaxation process represented by the CP signal is that the experimental artifacts and observation noise vary greatly from measurement to measurement. Some signals are relatively easy to analyze owing to the flat background, as shown in Figure 1(b), while others have strong backgrounds and are difficult to analyze, as shown in Figure 1(c). The purpose of CP signal analysis is to extract normal modes from such signals, despite the difficulty. Conventionally, Fourier analysis and wavelet analysis have been employed for CP signal analysis [14–18]. Despite their widespread use, these methods cannot uniquely separate the normal modes since they use trigonometric and wavelet bases, which cannot express damped oscillations.

To address the aforementioned issue, Murata et al. [19] analyzed a CP signal using sparsity-promoting dynamic mode decomposition (SpDMD [20]) based on the fact that CP signals consist of a few damped oscillation modes. SpDMD is a method that decomposes time series data into a sum of damped oscillations and observation noise. The authors isolated candidates of the normal modes from experimental artifacts and observation noise by assuming that the noise was stationary and did not drift over time. Specifically, the authors estimated the amplitude of the noise from the negative time domain ($t < 0$ (ps)) in Figure 1(b) since the signal was excited at $t = 0$ (ps). However, the negative time domain of CP signals is not necessarily stationary, as shown in Figure 1(c) due to the instability of measurement instruments and lasers. The drift in such backgrounds has a significant influence on the extraction of normal modes.

In this study, we propose a data-driven framework to systematically extract normal modes, even when the background noise is unstable. The isolation of background noise and the extraction of normal modes from measured signals can be recast as a mode selection problem for CP signal estimation. We extend the previously proposed Bayesian LARS-OLS framework [21,22] to be used with SpDMD. Bayesian LARS-OLS is the mode selection framework from the viewpoint of data-driven approach in Bayesian inference. Bayesian LARS-OLS can analytically calculate the performance of a set of SpDMD modes in terms of the Bayesian free energy (FE) [23], which assumes that the measured data can be decomposed into a sum of amplified oscillations and observation noise. We numerically compare the performance of our proposed method with that of the previous method proposed by Murata et al. [19]. Our results consist of the following three points. First, our method succeeds in extracting normal modes from experimental data with strong backgrounds. We then use virtual data to examine the effect of noise. The set of selected modes is robust to artificially induced

**Figure 1.** (a) Schematic diagram of coherent lattice vibrations in a (111)-Bi thin film generated by ultrashort optical pulses. In the (111)-Bi thin film, the coherent lattice vibrations corresponding to two different symmetric point groups of $A_{1g}$ and $E_c$ are observed by the pump-probe method. (b), (c) Signal data from lattice vibrations in the (111)-Bi thin film. The lattice vibrations start from the initial time 0. The signal contains two damped oscillations, $A_{1g}$ and $E_c$. We measured two experimental datasets, Exp. 1 in (b) and Exp. 2 in (c), at different times, so their backgrounds are different, but the signal contains the same damped oscillations.
observation noise. Moreover, our proposed method can estimate not only the normal modes but also the level of observation noise. From these observations, our method applies to both mode selection and noise estimation, even for a strong background. These results suggest that our proposed data-driven method is applicable to the relaxation process analysis of material science.

Conventionally, a lines of studies conducted Fourier analysis on the relaxation process in material science [1–4]. Our method provides a new way to decompose normal modes by incorporating the physical properties of the relaxation process with Bayesian inference.

2. Method

In this section, we introduce a method to extract modes from observation signals of one-dimensional time series, as shown in Figure 1. First, we introduce DMD [24] in Section 2.1. Then, SpDMD [20] is introduced in Section 2.2 as a method that can extract a normal mode expressing the lattice vibrations of a small (sparse) number of modes. Finally, we introduce a framework that extracts the normal mode from the mode extracted by Bayesian LARS-OLS [22] by optimizing the sparse parameter in SpDMD even if the value of experimental noise is unknown in Section 2.3.

2.1. Dynamic mode decomposition

DMD is a method of extracting modes with damped or amplified vibrations from data [24–26]. DMD has been utilized in various fields of science and engineering, including fluid mechanics [27], analyses of power systems [28], epidemiology [29], robotic control [30], neuroscience [31], image processing [32], and non-linear system identification [33].

CP signals are measured as one-dimensional time series \((y_0, ..., y_N)\) with a uniform time interval \(\delta t\), where \(N\) is the number of data points. To apply DMD and extract modes from observation signals of one-dimensional time series \((y_0, ..., y_N)\), we map the signals \(y_t\) to an \(M\)-dimensional vector as

\[
\psi_t = (y_t; y_{t+1}; ..., y_{t+M-1})^T,
\]

where the \(M\)-dimensional vector is called a state vector. We then obtain \(m + 1\) snapshots with a fixed time interval \(\delta t\), as shown in Figure 2, and define the \(t\)-th state vector \(\psi_t\) as the value at \(t \times \delta t\). Assuming discrete linear dynamics between time \(t \times \Delta t\) and \((t-1) \times \Delta t\), we obtain

\[
\psi_t = A\psi_{t-1}, \quad t = 1, ..., m,
\]

where \(A \in \mathbb{C}^{M \times M}\). The objective is to estimate the matrix \(A\) from measured data that describe the discrete linear dynamical system. For this objective, we briefly explain the DMD algorithm according to Jovanović et al. [20]. As shown in Figure 2, we form two data matrices from the snapshot sequence

\[
\Psi_0 = (\psi_0; ..., \psi_{m-1}) \in \mathbb{R}^{M \times m},
\]

\[
\Psi_1 = (\psi_1; ..., \psi_m) \in \mathbb{R}^{M \times m},
\]

where \(m = N - M + 1\). Using Equation (2), we obtain the following relationship between \(\Psi_0\) and \(\Psi_1\):

\[
\Psi_1 = A\Psi_0.
\]

The data matrix \(\Psi_0\) is a nonsquare matrix; there is no guarantee of regularity even in the case of a square matrix, so the inverse matrix \(\Psi_0^{-1}\) does not exist in general. We represent the matrix \(A\) that satisfies the above equation by using the Moore-Penrose pseudoinverse matrix \(\Psi^+\). The best-fit linear operator \(A\) that
satisfies Equation 5 is the solution to the following least-squares optimization:

\[
A = \arg\min_A \| \Psi_1 - A\Psi_0 \|_F = \Psi_1\Psi_0^\dagger \approx \Psi_1V\Sigma^{-1}U^*.
\]

(6)

Here, \( U \in \mathbb{R}^{M \times r}, \Sigma \in \mathbb{R}^{r \times r}, V \in \mathbb{R}^{m \times r} \) are matrices of a singular value decomposition (SVD): \( \Psi_0 = U\Sigma V^* \). The eigenvalues and eigenvectors of the matrix \( A \) correspond to eigenvalues and eigenmodes of time evolution. The size of the matrix \( A \in \mathbb{R}^{M \times M} \) is very large. To solve the eigenvalue problem efficiently, instead of the eigenvalue problem of the matrix \( A \), we consider the low rank matrix \( \tilde{A} = U^*AU = U^*\Psi_1V\Sigma^{-1} \),

(7)

Next, we computed the eigendecomposition of \( \tilde{A} \):

\[
\tilde{A} = \Phi D \Phi^T = \Phi D_\mu.
\]

(8)

The diagonal matrix \( D_\mu = \text{diag}(\mu_1, \mu_2, \ldots, \mu_r) \) means the eigenvalues of \( A \), which are same in \( A \). The corresponding eigenvectors of \( A \) is represented by:

\[
\Phi = UW,
\]

(9)

where \( W \in \mathbb{C}^{r \times r} \) is a matrix of \( \tilde{A} \)'s eigenvectors. Corresponding to DMD eigenvalues \( \mu \), we called the columns \( \{ \phi_1, \phi_2, \ldots, \phi_r \} \in \mathbb{C}^{M} \) as DMD eigenmodes. Applying Equation (2) repeatedly, the state vector \( \psi_i \) at step \( t \) from the initial vector \( \psi_0 \) can be represented as

\[
\psi_t = A^t\psi_0 \approx (\Phi D_\mu \Phi^T)^t \psi_0 = \sum_{i=1}^{r} \alpha_i \phi_i \mu_i^t.
\]

(10)

Here, we define \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r)^T = \Phi^T \psi_0 \in \mathbb{C}^r \) as the mode amplitude, which corresponds to the coefficients of each mode. The vector corresponding to the eigenvalue is \( \mu = (\mu_1, \mu_2, \ldots, \mu_r)^T \in \mathbb{C}^r \).

We need to estimate \( \alpha \) from the data, which is the solution to an optimization problem. From the above expressions, the data matrix \( \Psi_0 \) can be represented as

\[
\Psi_0 \approx \Phi D_\mu V_{\text{and}}.
\]

(11)

Here, \( D_\mu = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_r) \) and a \( r \times m \) Vandermonde matrix representing the latent structure in the temporal direction

\[
V_{\text{and}} = \begin{bmatrix}
1 & \mu_1 & \ldots & \mu_1^{m-1} \\
1 & \mu_2 & \ldots & \mu_2^{m-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \mu_r & \ldots & \mu_r^{m-1}
\end{bmatrix} \in \mathbb{C}^{r \times m}.
\]

(12)

The coefficient \( \alpha \) of each mode is expressed using the pseudo-inverse matrix \( \Phi^+ \) of the dynamic mode \( \Phi \). Here, we obtain \( \alpha \) by solving the least squares problem.

\[
\alpha_{LS} = \arg\min_{\alpha} J(\alpha),
\]

(13)

\[
J(\alpha) = \| \Psi_0 - \Phi D_\mu V_{\text{and}} \|_{Fro}^2.
\]

(14)

\( \|X\|_{Fro} \) represents the Frobenius norm of the matrix \( X \). \( J(\alpha) \) is the objective function that needs to be minimized to obtain the mode amplitude \( \alpha \). In Appendix A, we present the correspondence between the DMD modes and the relaxation process.

### 2.2. Sparse estimation of DMD

We assume prior knowledge that the normal modes representing the lattice oscillations of a substance among modes extracted by DMD are few (sparse). We introduce SpDMD [20] to obtain the sparse solution of the mode amplitude. SpDMD is a combination of DMD and the least absolute shrinkage and selection operator (Lasso) [36], which automatically extracts a small number of components for the purpose of explaining data. We solve the following \( \ell_1 \)-regularized optimization problem.

\[
\alpha_{Sp}(y) = \arg\min_{\alpha} J(\alpha) + y \sum_{i=1}^{r} |\alpha_i|,
\]

(15)

Here, \( y \) represents the weight of the \( \ell_1 \) regularization term. If \( y \) takes a significant value, \( \alpha \) becomes a more sparse solution. If SpDMD gives the regularization parameter \( y \), we obtain a unique mode set.

#### 2.3. Mode selection by Bayesian LARS-OLS

We deal with mode selection in Bayesian inference by considering the prior of \( \alpha \). Although we optimized \( \alpha \) with \( \ell_1 \) regularization as shown in above, the \( \ell_1 \) regularization develops a bias of the basis coefficient. In this section, we present a well-known technique for preventing the bias, called ‘polishing’ [37], within the framework of Bayesian inference. Such framework is known as Bayesian LARS-OLS [22] and we extend this framework to the mode selection problem in SpDMD. Bayesian LARS-OLS framework uses Bayesian free energy [23] to evaluate the goodness of model fit. We can determine the optimal sparsity \( y \) and select the set of DMD modes based on the criteria. In Section 2.3.1, we first introduce the ‘polishing’ method. Next, in Section 2.3.2, we introduce the Bayesian LARS-OLS and Bayesian free energy (FE).

##### 2.3.1. Polishing

We regress the mode amplitude \( \alpha \) again in terms of the candidate modes selected by SpDMD to prevent the bias of the coefficient. The technique is called ‘polishing’ [36]. We deal with the ‘polishing’ in mathematics and introduce the prior of \( \alpha \) condition in Bayesian inference by considering parameter \( c(y) \). We define
an indicator \( c(y) \) representing whether the modes is selected or not by SpDMD:

\[
  c(y) = \{ i | a_i y \neq 0 \}, \tag{16}
\]

where \( i \) is the index of the modes. In other words, the indicator \( c(y) \) is the set of indices of which DMD mode was estimated as nonzero. We solve the constrained optimization problem of Equation (17) using the set \( c(y) \) with suffix \( i \) depending on the \( \ell_1 \) regularization term \( y \) in SpDMD.

\[
  \text{minimize } J(\alpha) \\
  \text{subject to } a_i = 0, \forall i \notin c(y) \tag{17}
\]

This equation means that we minimize \( J(\alpha) \) under the constraint \( a_i = 0 \), and \( \alpha \) depends on \( y \) only through \( c \).

### 2.3.2. Evaluation of decomposed modes

In this section, we deal with the aforementioned ‘polishing’ method in a Bayesian inference. The ‘polishing’ has two different purposes: mode search by \( \ell_1 \) regularization and regression. Least angle regression and ordinary least squares (LARS-OLS) [37] is a framework of regression analysis that performs basis search and regression in a stepwise manner. Igarashi et al. considered the LARS-OLS as Bayesian inference (called Bayesian LARS-OLS) and proposed a method that achieves the basis search in the Bayesian free energy criterion [21,22]. We extend their method to be able to apply the material’s dynamics such as a relaxation process. In Bayesian LARS-OLS framework, we consider DMD as a problem of linear regression and express it by a probability model.

The model of DMD as described above can be written as \( \hat{y} = X^* \alpha \) where \( X^* \) is a matrix determined from the eigenvalues \( \mu \) and eigenvectors \( \psi \) of DMD. Please see Appendix B for the detail derivation. The optimization of \( \ell_2 \) loss (Equation 14) for above linear model corresponds to the maximum likelihood estimation under the assumption that the observation noise is the independent and identically distributed (i.i.d.) Gaussian. Overall, we can write the DMD estimator as follows:

\[
  y = X^* \alpha + \epsilon. \tag{18}
\]

When we estimate the modes \( X \) and sparsity \( y \) as \( \hat{X} \) and \( \hat{y} \), we define a combination of modes as \( \hat{\epsilon} = c(\hat{y}) \). We note that \( c(\cdot) \) computes \( \hat{\epsilon} \) through SpDMD by using \( \hat{X} \) and \( \hat{y} \). To evaluate \( \hat{\epsilon} \), we consider a posterior distribution \( P(\hat{\epsilon}|y) \). From the Bayesian theorem, the posterior is

\[
P(\hat{\epsilon}|y) = \frac{P(y|\hat{\epsilon})P(\hat{\epsilon})}{P(y)}. \tag{19}
\]

Here, \( P(y) \) is constant. We define the prior distribution \( P(\hat{\epsilon}) \) to be uniform and use the following relationship:

\[
P(\hat{\epsilon}|y) \propto P(y|\hat{\epsilon})
\]

\[
= \int P(y|\alpha, X, \hat{\epsilon}) P(\alpha|X, \hat{\epsilon}) P(X|\hat{\epsilon}) d\alpha dX. \tag{20}
\]

Based on the above assumption (Equation 18), the following holds: \( P(y|\alpha, X, \hat{\epsilon}) \sim N(X^* \alpha, \sigma^2) \). We assume that the matrix \( X \), which characterizes modes, does not depend on the combination \( \hat{\epsilon} \), so we define

\[
P(X|\hat{\epsilon}) = \delta(X). \tag{21}
\]

This equation means that the combination of modes does not change either the eigenvalue of the time evolution \( \mu \) or the dynamic mode \( \phi \), where \( \delta(\cdot) \) is the Dirac delta function. Substituting Equation (21) into Equation (20), we obtain

\[
P(y|\hat{\epsilon}) = \int P(y|\alpha, X, \hat{\epsilon}) P(\alpha|X, \hat{\epsilon}) d\alpha. \tag{22}
\]

Here, the \( P(\alpha|X, \hat{\epsilon}) \) is a prior probability and the \( P(y|\alpha, X) \) is a distribution of data generation. The \( P(y|\hat{\epsilon}) \), called the marginal likelihood, represents the likelihood of the indicator \( \epsilon \) estimated from the data. The negative marginal log-likelihood,

\[
\text{FE}(\hat{\epsilon}) = -\log P(y|\hat{\epsilon}), \tag{23}
\]

is often referred to as the Bayesian free energy (FE) [23]. The FE criterion evaluates the goodness of model fit. In our criteria, the indicator \( \epsilon \), which means whether the DMD mode was selected or not, determines the models and FE. The indicator \( \epsilon \) with the smallest FE represents the best basis set. To calculate FE analytically, we define the prior distribution \( P(\alpha|X, \hat{\epsilon}) \) as follows in the cases of \( a_i = 0 \) and \( a_i \neq 0 \).

\[
P(\alpha_i|X, i \in \hat{\epsilon}) = \mathcal{N}(0, \sigma_{\text{pri}}^2)
\]

\[
= \frac{1}{\pi \sigma_{\text{pri}}^2} \exp \left( -\frac{1}{\sigma_{\text{pri}}^2} \alpha_i^2 \right),
\]

\[
P(\alpha_i|X, i \notin \hat{\epsilon}) = \delta(\alpha_i) = \delta(0). \tag{25}
\]

Here, the coefficient \( \alpha \) is complex number, we define \( P(\alpha|X, \hat{\epsilon}) \) as a complex Gaussian distribution [38]. From the obtained FE (\( \hat{\epsilon} \)), we estimate the variance \( \nu = \sigma^2 \) of the noise and the variance \( \nu_{\text{pri}} = \sigma_{\text{pri}}^2 \) of the prior distribution as values that minimize the FE (\( \hat{\epsilon} \)). By repeatedly solving the self-consistent equations until convergence, the values of \( \nu \) and \( \nu_{\text{pri}} \) that minimize the FE (\( \hat{\epsilon} \)) are obtained. In this way, we can analytically calculate the FE. The reason why we can analytically calculate the FE (\( \hat{\epsilon} \)) is that the parameter \( X \) of the DMD’s eigenvalues and eigenvectors has already been determined.
3. Results and discussion

We focus on experimental signals of a bismuth (Bi) thin film as typical examples of relaxation processes. We deposited a Bi thin film on a sapphire substrate with a thickness of 150 nm and used a reflective pump-probe spectroscopy method [16] to obtain the CP signals. The signal background varied from measurement to measurement due to experimental instabilities. We analyzed two typical signals. One signal has a background close to steady-state, as shown in Figure 1(b), and the other signal has a strong background, as shown in Figure 1(c). In Sections 3.1 and 3.2, we explain the advantage of our proposed method. Finally, in Section 3.3, by using virtual data, we describe the robustness to noise and the noise estimation performance.

3.1. Background close to steady-state

We analyzed the weak-background data in Figure 1(b). The experimental data had N = 2433 data points from t = 0.00 to 8.00 ps at an interval of Δt = 8/2433 ps. The data matrices W₀ and W₁ were generated according to Figure 2 with M = 1000. We prepared the weight of the sparsity term γ in Equation (15) as 600 units with equal intervals on a log scale between γ = 10⁻⁹ and 10⁻².

Figure 3(a) shows the reconstruction of the original signal with the optimal modes selected by the previous method. As shown by the red and green lines in Figure 3(a), the previously proposed method successfully extracted the normal vibration modes A₁g and E₁g. Moreover, the other selected modes are denoted by the black dotted lines and correspond to a weak background in the experimental data [19]. By using the selected modes, the recovered signals coincide with the experimental data, as shown in the pink filled area and the blue line of Figure 3(a). The residual signal between the original signal and the reconstructed signal is close to zero, as shown by the solid bottom lines of Figure 3(a). In Murata et al. [19], assuming prior knowledge that the normal modes representing lattice oscillations of the substance of interest were sparse, the authors selected the modes extracted by DMD as explained in Section 2.1. The previous method determined the optimal number of modes to be nine, where the curve of the MSE curve crosses the estimated noise variance σ²_est = 4.00 x 10⁻¹⁴, as shown in Figure 3(c). Note that the authors estimated the noise variance from the variance of the negative time signal [19]. On the other hand, the proposed method also succeeded in extracting the normal modes from the experimental data without using speculative noise information, as shown in Figure 3(b). In the proposed method, we estimated the noise variance in objective data using Bayesian inference and optimized the number of normal modes as 19 by minimizing the Bayesian FE, as shown in Figure 3(d). Our proposed method can obtain an estimate of the observed noise level from the Bayesian FE. We estimated that the standard deviation of the noise was σ_FE = 3.56 x 10⁻⁸, while the value of the noise estimated in the previous study was σ_est = 2.00 x 10⁻⁷.

To qualitatively investigate the selected normal modes, we compare the results of a frequency spectrum analysis of the selected normal modes with the results of a previous study [16]. Hase et al. [16] reported that the two normal modes of (111)-Bi, A₁g and E₁g, have frequencies of 2.91 and 2.07 THz, respectively. Figure 4 quantitatively shows the frequency of each mode, where the horizontal axis represents the frequency calculated from the eigenvalue of the extraction mode, and the vertical axis represents the absolute value of the mode amplitude. Figure 4(b,c) show the modes selected by the previous method [19] and the proposed method, respectively, where both methods assumed that the original signal consisted of a few modes. The figures show that both methods successfully extracted the normal modes at 2.90 and 2.09 THz. These modes correspond well to A₁g (2.91 THz) and E₁g (2.07 THz), respectively.

3.2. Strong background

In contrast to the previous section, the assumption regarding the noise estimation by Murata et al. [19] does not hold when the data have a strong background, which can cause the normal mode selection to fail. Let us consider the strong-background case. The experimental data had N = 9740 data points from t = 0.00 to 7.50 ps at an interval of Δt = 7.5/9740 ps. The data matrices W₀ and W₁ were generated according to Figure 2 with M = 4970. We prepared the weight of the sparsity term γ in Equation (15) as 600 units with equal intervals on a log scale between γ = 10⁻⁷ and 10⁻².

Figures 5 and 6 show the results for a strong background. The previous method determined seven modes as shown in Figure 5(c). Although the selected modes included the A₁g mode, they did not include the E₁g mode. On the other hand, our method selected 85 modes, as shown in Figure 5(d), which successfully included both the A₁g and E₁g modes. Our method also has an advantage in terms of the residuals, as shown by the solid black line in Figure 5(a,b). The residual signal extracted by our method is steady, whereas the previous method includes a vibration component in the residual signal.

As in Figure 4, we performed a frequency spectrum analysis of each mode to quantitatively evaluate the difference from the normal modes obtained in a previous study [16] and to ensure that normal
modes can be extracted quantitatively (Figure 6). As in Figure 4(a–c), Figure 6(a) shows the modes extracted from DMD, and Figure 6(b,c) show the modes selected by the previous method [19] and the proposed method, respectively. Figure 6(b) shows that the previous method successfully extracted the mode at 2.96 THz, which corresponds to the $A_{1g}$ (2.91 THz) mode band; however, it failed to extract the $E_{g}$ (2.07 THz) mode band. On the other hand, Figure 6(c) shows that the proposed method successfully extracted the normal modes at 2.96, and 2.34 THz. These modes correspond to the $A_{1g}$ (2.91 THz) and the $E_{g}$ (2.07 THz) modes bands, respectively. The estimated errors were 0.05 and 0.27 THz, respectively. This latter error seemed to be influenced by the intense fluctuating background components. To decompose such phonon modes, which have a weak amplitude and short decay time, we should improve the experimental conditions to reduce such experimental artifacts.

Let us consider the noise estimation results. We estimated that the standard deviation of the noise was $\sigma_{FE} = 8.08 \times 10^{-8}$, while the value of the noise estimated in the previous research was $\sigma_{est} = 5.33 \times 10^{-7}$. In both the strong- and weak-background data, $\sigma_{est}$ is approximately six times larger than $\sigma_{FE}$. The error of the noise estimation has a critical effect on the normal mode
Figure 4. Frequencies of each decomposed mode for data with a weak background. The vertical axes represent the absolute values of the mode amplitudes, and the horizontal axes represent the mode frequencies. The red lines correspond to the frequency of $A_{1g}$, and the green lines correspond to the frequency of $E_g$. (a): All decomposed modes. (b): The modes selected by the previous method for the estimated noise. (c): The modes selected by our method for the Bayesian FE.

Figure 5. Comparison of our method and the previous method for data with a strong background. The figures on the left side show the result of the previous method, and the figures on the right side show the result of our method. (a-b): The intensity of the signals according to the delay time. (c-d): The value of the metric used for the mode selection according to the number of modes. The meaning of each color and line style is the same as in Fig. 3.
selection in the data with a strong background. The previous method failed to select the $E_g$ mode, because the noise level estimated by the previous method was larger than the noise level estimated by the proposed method. Because the amplitude of the $E_g$ mode was smaller than the $A_{1g}$ mode, it was considered that the $E_g$ mode deviated from the selected mode. However, since our method estimates noise without a stationary assumption for an experimental artifact, we succeeded in selecting normal modes with small amplitudes. Our method is effective, especially for data with a strong background. By applying our method to CP analysis, we can select normal modes from measurement signals without considering the experimental situation. Since the normal modes were known for (111)-Bi, it was possible to find normal modes from candidates of selected modes obtained by mode selection.

3.3. Virtual CP data

Finally, we investigate the performance of the previous method and the proposed method against virtual CP data to evaluate the effect of noise. Let us consider the results of the virtual CP data. We created virtual data by injecting Gaussian noise ($\sigma \sim 8.0 \times 10^{-8}$) into the reconstructed signal from nine modes chosen by the previous method in the weak-background (Figure 3(a)). We prepared 100 virtual datasets with different samples of noise.

Figure 7 shows the number of selected modes. We plotted the distributions of counts as a box plot. The previous method selected $191 \pm 114$ modes, while our method selected $14 \pm 11$ modes. We note that the number of modes selected by proposed method was always greater than the correct number of modes of nine. Since there is no negative time domain in the generated virtual data, the noise level used to determine the mode in the previous method was the generated noise level ($\sigma \sim 8.0 \times 10^{-8}$). In the result of the previous method, the selected numbers varied considerably according to the noise, and the mode number tended to be much larger than the correct number of modes. On the other hand, our method has minimal dispersion and can select a number that is closer to the correct mode number from noisy data. These results show that our method is much more robust to noise than the previous method.

Let us consider the noise estimation results. The estimated noise level by the FE was $\sigma \sim 9.4 \pm 1.5 \times 10^{-8}$. Since the true noise level is $\sigma \sim 8.0 \times 10^{-8}$, our method can estimate the amplitude of the original noise from an observed signal whose noise value is unknown. This result indicates that it is possible to quantitatively evaluate the amplitude of the noise without using a heuristic technique.
4. Conclusions

In this study, we extended Bayesian LARS-OLS to be used in SpDMD for the selection of normal modes in a relaxation process. We numerically compared the performance of our proposed method to the previous method proposed by Murata et al. [19]. The following three results showed the effectiveness of the proposed method based on data of the CP signal of a (111)-Bi thin film. First, our method succeeded in extracting a small number of modes, including the normal modes, from experimental data even with significant background trends. Then, we used virtual data to examine the effect of noise. The selected set of modes was more robust to observation noise than the conventional method. Finally, our proposed method allowed to estimate not only normal modes but also the level of observation noise quantitatively. From these observations, our proposed method is applicable to normal mode analysis of relaxation processes, especially for data with strong backgrounds, which broadens the applicability of data-driven approach in analyzing of relaxation phenomena in material science.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Japan Science and Technology Agency [CREST(JP)MCR1761), CREST (JP)MCR1861), PREST(JP)MCR17N2)].

ORCID

Yasuhiro Igarashi https://orcid.org/0000-0003-1042-6657
Masato Okada https://orcid.org/0000-0002-9040-8784

References

[1] Casalini R, Roland CM. Aging of the secondary relaxation to probe structural relaxation in the glassy state. Phys Rev Lett. 2009 Jan;102(3):035701. ISSN 0031-9007.
[2] Case DA. Molecular dynamics and NMR spin relaxation in proteins. Acc Chem Res. 2002 Jun;35 (6):325–331. ISSN 0001-4842.
[3] Kim H-T, Lee YW, Kim B-J, et al. Monoclinic and correlated metal phase in VO 2 as evidence of the mott transition: coherent phonon analysis. Phys Rev Lett. 2006 Dec;97(26):266401. ISSN 0031-9007.
[4] Voth GA, Hochstrasser RM. Transition state dynamics and relaxation processes in solutions: a frontier of physical chemistry. J Phys Chem. 1996 Jan;100(31):13034–13049. ISSN 0022-3654.
[5] D’Ercole A, Vesperini E, D’Antona F, et al. Formation and dynamical evolution of multiple stellar generations in globular clusters. Mon Not R Astron Soc. 2008 Dec;391(2):825–843. ISSN 00358711.
[6] Salma I, Ockskar R, Varga I, et al. Surface tension of atmospheric humic-like substances in connection with relaxation, dilution, and solution pH. J Geophys Res Atmos. 2006 Dec;111(D23):n/a–n/a. ISSN 01480227.
[7] Blagg RJ, Ungur L, Tuna F, et al. Magnetic relaxation pathways in lanthanide single-molecule magnets. Nat Chem. 2013 Aug;5(8):673–678. ISSN 1755-4330.
[8] Rosker MJ, Wise FW, Tang CL. Femtosecond relaxation dynamics of large molecules. Phys Rev Lett. 1986 Jul;57:321–324.
[9] Cheng TK, Brorson SD, Kazeroonian AS, et al. Impulsive excitation of coherent phonons observed in reflection in bismuth and antimony. Appl Phys Lett. 1990;57(10):1004–1006.
[10] Dumitric T, Garcia ME, Jeschke HO, et al. Selective cap opening in carbon nanotubes driven by laser-induced coherent phonons. Phys Rev Lett. 2004 Mar;92(11):117401. ISSN 0031-9007.
[11] Iwai S, Okamoto H. Ultrafast phase control in one-dimensional correlated electron systems. J Phys Soc Jpn. 2006;75(1):011007.
[12] Wall S, Wegkamp D, Foglia L, et al. Ultrafast changes in lattice symmetry probed by coherent phonons. Nat Commun. 2012 Mar;3:721.
[13] Madelung O. Semiconductors: data handbook. Berlin, Heidelberg: Springer Berlin Heidelberg; 2004. ISBN 978-3-642-62332-5.
[14] Mizoguchi K, Morishita R, Oohata G. Generation of coherent phonons in a CdTe single crystal using ultrafast two-photon laser-excitation process. Phys Rev Lett. 2013 Feb;110:077402.
[15] Hase M, Mizoguchi K, Harima H, et al. Optical control of coherent optical phonons in bismuth films. Appl Phys Lett. 1996 Oct;69(17):2474–2476. ISSN 0003-6951.
[16] Hase M, Kitajima M, Nakashima S, et al. Dynamics of coherent anharmonic phonons in bismuth using high density photoexcitation. Phys Rev Lett. 2002 Jan;88(6):067401. ISSN 0031-9007.
[17] Hase M, Kitajima M, Constantinescu AM, et al. The birth of a quasiparticle in silicon observed in time-frequency space. Nature. 2003 Nov;426(6962):51–54. ISSN 0028-0836.
[18] Yoshino S, Oohata G, Mizoguchi K. Dynamical fano-like interference between rabi oscillations and coherent phonons in a semiconductor microcavity system. Phys Rev Lett. 2015 Oct;115(15):157402. ISSN 0031-9007.
[19] Murata S, Aihara S, Tokuda S, et al. Analysis of relaxation processes of optical control over coherent phonons in a semiconductor microcavity system. Opt Express. 2018 Dec;26(26):35656–35664. ISSN 1094-4087.
[20] Schmid PJ. Dynamic mode decomposition. J Phys Soc Jpn. 2013 Apr;82(4):044802. ISSN 0031-9015.
[21] Mototake Y, Igarashi Y, Takenaka H, et al. Exhaustive search for sparse variable selection in linear regression. J Phys Soc Jpn. 2018 Apr;87(4):044802. ISSN 0031-9015.
[22] Watanabe S. Mathematical theory of Bayesian statistics. CRC Press; 2018. ISBN 1315355698.
[23] Schmid PJ. Dynamic mode decomposition of numerical and experimental data. J Fluid Mech. 2010;656:528.
\[ \mu_i^n = R_i \exp \left( \frac{\ln r_i}{\Delta t} t \right) \exp \left( \frac{\theta_i}{\Delta t} t \right). \]

where \( R_i \) is \( R_i = \mu_i^{\nu_i/\Delta t} \). We rewrite this equation in the form of damped oscillation as follows

\[ \mu_i^n = R_i \exp \left( -\frac{t}{\tau_j} \right) \left( \cos 2\pi f_i t + i \sin 2\pi f_i t \right). \]

Then, \( \tau_j \) and \( f_j \) are

\[ \tau_j = -\frac{\Delta t}{\ln \tau_j}, f_j = \frac{\theta_j}{2\pi \Delta t}. \]

\( \tau_j \) is the damping time constant and \( f_j \) is the frequency. Thus, the (10) formula can be rewritten in terms of time evolution from a discrete form to a continuous form as follows

\[ \psi_n = \sum_{j=1}^{\nu} R_j \exp \left( -\frac{t}{\tau_j} \right) \left( \cos 2\pi f_j t + i \sin 2\pi f_j t \right). \]

where \( R_j \) is \( a_j R_j \phi_j \).

\( \psi_n \) is the sum of elements for which \( f_j = 0 \) and \( \sum_\nu \) is the sum of elements for which \( f_j \neq 0 \). Each decomposed form is as follows

\[ \psi_n = \sum_{j=1}^{\nu} \exp \left( -\frac{t}{\tau_j} \right) \left( \cos 2\pi f_j t + i \sin 2\pi f_j t \right), \]

where \( \sum_\nu \) is the sum of elements for which \( f_j = 0 \), and \( \sum_\nu \) is the sum of elements for which \( f_j \neq 0 \). \( y(t) \) is observation data and contains only real components as follows.

\[ \psi_1 = \sum_{j=1}^{\nu} C_j \exp \left( -\frac{t}{\tau_j} \right) \]

\[ \sum_\nu \exp \left( -\frac{t}{\tau_j} \right) \left( a_j \cos 2\pi f_j t + b_j \sin 2\pi f_j t \right), \]

where \( C_j = R_j, a_j = 2 \Re(R_j), b_j = -2 \Im(R_j) \). Since only the real component is obtained by discarding the pure imaginary component, the degree of freedom of the complex conjugate remains. As a result, the vibration mode of \( f_j \neq 0 \) is obtained by two modes of the same value in the space of only real numbers.

Appendix A. Correspondence between the DMD basis and the form of damped oscillations

This chapter explains how the form of the damped oscillations representing the relaxation process can be expressed using the DMD basis. The time evolution of the DMD basis is \( V_{\text{and}} \) in the (12) expression. Here, when we converted each eigenvalue \( \mu_j \) of DMD into \( \mu_j = r_j \exp(i\theta_j) \) in the complex polar coordinate system, each element \( \mu_j^n \) of \( V_{\text{and}} \) becomes \( \mu_j^n = \exp(n \ln r_j) \exp(in\theta_j) \).

Assuming that the initial time of the time series data is \( t_0 \) and that the time interval \( \Delta t \) is constant, the relationship between the time step \( n \) and the time \( t \) is \( n = t - t_0/\Delta t \). Now, \( \mu_j^n \) is expressed using \( t \) as follows

Appendix B. Transformation to original signal shape

In this chapter, we explain the operation to restore the data matrix obtained by SpDMD to the original one-dimensional time series signal. From the result of (1), the original signal data \( y = y_0, \ldots, y_N \) can be expressed as follows
\[ 2y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_M \\ y_{M+1} \\ \vdots \\ y_N \end{pmatrix} + \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-M} \\ y_{N-M+1} \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \Psi_0^T \\ \chi_0^T \\ \Psi_{N-M+1}^T \end{pmatrix}. \] (B1)

Here, \( \chi, \chi' \) is a column vector which is a pair of state vectors \( \psi \) of DMD introduced for the reconstruction of the original time series and can be expressed using

\[ V'_{\text{and}} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_r \\ \mu_1^2 \\ \mu_2^2 \\ \vdots \\ \mu_r^2 \\ \vdots \\ \mu_1^{N-M+1} \\ \mu_2^{N-M+1} \\ \vdots \\ \mu_r^{N-M+1} \end{pmatrix} \in \mathbb{C}^{r \times (N-M+1)}, \] (B2)

\[ T_i = \begin{pmatrix} \phi_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \phi_{i2} & \cdots & 0 \\ 0 & 0 & \cdots & \phi_{ir} \end{pmatrix} \in \mathbb{C}^{r \times r}, \] (B3)

and \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r)^T \) as

\[ X_i = \sum_{j=1}^{r} \alpha_j \phi_{ij} y_j = \alpha^T T_i V'_{\text{and}} \in \mathbb{C}^{1 \times N-M+1}. \] (B4)

\[ \chi' = \sum_{j=1}^{r} \alpha_j \phi_{ij}' y_j = \alpha^T T_i V'_{\text{and}} \in \mathbb{C}^{1 \times N-M+1}. \] (B5)

Therefore, it becomes \( \chi_0^T = V'_{\text{and}} T_0 \alpha; \chi_M^T = V'_{\text{and}} T_M \alpha. \)

The next thing to do is to display matrix with \( y = X' \alpha \). By using

\[ \Psi_0 \text{ corresponds to } \Phi \alpha \in \mathbb{C}^M; \Psi_N \text{ corresponds to } \Phi \alpha \in \mathbb{C}^M. \]

Eventually, it becomes

\[ y = \frac{1}{2} \left( V'_{\text{and}} T_{\text{and}} \right) \alpha + \frac{1}{2} \left( V'_{\text{and}} T_0 \right) \alpha = X' \alpha. \] (B7)

Here, we note

\[ X' = \frac{1}{2} \left( V'_{\text{and}} T_{\text{and}} \right) + \frac{1}{2} \left( V'_{\text{and}} T_0 \right). \] (B8)