Bayesian spectroscopy of synthesized soft X-ray absorption spectra showing magnetic circular dichroism at the Ni-L\textsubscript{3}, -L\textsubscript{2} edges

Taiga Yamasaki\textsuperscript{a}, Kazunori Iwamitsu\textsuperscript{b}, Hiroyuki Kumazoe\textsuperscript{c}, Masato Okada\textsuperscript{d,e}, Masaichiro Mizumaki\textsuperscript{f} and Ichiro Akai\textsuperscript{c}

\textsuperscript{a}Graduate School of Science and Technology, Kumamoto University, Kumamoto, Japan; \textsuperscript{b}Technical Division, Kumamoto University, Kumamoto, Japan; \textsuperscript{c}Institute of Industrial Nanomaterials, Kumamoto University, Kumamoto, Japan; \textsuperscript{d}Department of Complexity Science and Engineering, The University of Tokyo, Chiba, Japan; \textsuperscript{e}Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Ibaraki, Japan; \textsuperscript{f}Spectroscopy Division, Japan Synchrotron Radiation Research Institute, Hyogo, Japan

\section{Introduction}

Bayesian spectroscopy \cite{ refs1, refs2} is a data-driven science approach of that we apply Bayesian inference \cite{ refs3} onto spectral analysis to achieve highly challenging spectral decomposition. In the Bayesian inference of the data-driven science \cite{ refs4}, Bayes’ theorem \cite{ refs3} is applied to the joint probability of causes and results in the causality, and evaluation of posterior probability distributions for causes becomes available based on the resultant data. Such evaluation is one of major advantages in Bayesian spectroscopy and it makes possible to provide a statistical guarantee for a particularly-difficult spectral decomposition. Such advantage has been illustrated by the Bayesian spectroscopy of admixed photoluminescence spectra with exciton, biexciton and electron-hole droplet (EHD) states in a highly-excited GaAs/AlAs type-II superlattice \cite{ refs5}, where we have demonstrated that the EHD state becomes stable from the evaluation of the posterior probability distributions for a chemical potential of the EHD, an energy of excitonic system and their effective temperature \cite{ refs6}.

Another important advantage of Bayesian spectroscopy is the ability to select an optimal model for explaining the data without preconceptions by using Bayes free energy \cite{ refs1} (BFE) as an information criterion. This model selection is a key data-driven methodology in analyses for various measured data, and Bayesian Hamiltonian selection is also available to explain physical phenomena \cite{ refs7}. This model selection also works properly on decomposing complicated spectra \cite{ refs2, refs8, refs9, refs10, refs11, refs12}. Applying Bayesian spectroscopy, we have succeeded in decomposing weak pre-edge structures from an X-ray absorption near edge structure (XANES) spectrum of α-Fe\textsubscript{2}O\textsubscript{3} \cite{ refs13}.

In this study, we apply Bayesian spectroscopy to decompose X-ray absorption (XA) spectra and an X-ray magnetic circular dichroism \cite{ refs14, refs15, refs16} (XMCD) spectrum, which were synthesized to imitate the L\textsubscript{3,2}-absorption edges of Ni ions in NiFe\textsubscript{2}O\textsubscript{4} \cite{ refs17} under magnetic fields with $\pm$ helicities. This material belongs to the class of strongly correlated 3d system under an octahedral crystal field, and so that, versatile split X-ray transitions are expected in the respective $\pm$ helicity polarized XA spectra. In addition, at the L\textsubscript{3,2} edges of such magnetic materials having d-electrons, it is possible to obtain the spin, orbital and total...
magnetic moments from the spectral integration of the XMCD spectrum based on the sum rule [18].

To sensitively detect changes on opposite helicities, we have always obtained the XMCD spectrum as a difference spectrum of $\mp$ helicity $\text{XA}$ spectra by eliminating non-magnetic spectral components. From differential spectrum by modulation technique, one can detect spectral changes with high sensitivity [19]. However, in the case of the spectrum obtained by subtraction (difference spectrum), the data point number becomes half of that in the original spectra ($\mp$ helicity $\text{XA}$ spectra), and the noise variance included in the difference spectrum becomes the sum of the noise variances superimposed on these original spectra. Therefore, it is concerned that information missing might occur on extracting physical properties for the target material.

In order to overcome such difficulties in difference spectral analyses, we apply the Bayesian spectroscopy with the BFE-based model selection [1,20] to decompose the original $\mp$ helicity spectral components from an XMCD spectrum, and we demonstrate the advantage of the Bayesian spectroscopy comparing the decomposition results between the XMCD and individual $\mp$ helicity $\text{XA}$ spectra. In addition, we evaluate the estimation accuracies of the spin, orbital and total magnetic moments obtained by the decomposed spectral components based on the posterior probability distributions of the decomposed spectral components.

2. Synthesized spectra

In this section, we detail the synthesized spectra to analyze by Bayesian spectroscopy. The target spectra was prepaired to imitate an XMCD spectrum of the $L_{2,3}$-absorption edges of Ni$^{2+}$ ions in NiFe$_2$O$_4$ [17,21].

2.1. Effective Hamiltonian

The XMCD spectrum was calculated as a difference spectra of $\mp$ helicity $\text{XA}$ spectra, while the $\text{XA}$ spectra were synthesized on the basis of the X-ray transitions at the $L_{2,3}$ edges among electronic states derived from an effective Hamiltonian in Equation (1) [22].

$$
\begin{align}
\mathcal{H} &= \sum_{\Gamma, \sigma} N_{\Gamma, \sigma} \frac{1}{2} \mathcal{D}_{\Gamma, \sigma} \mathcal{D}_{\Gamma, \sigma}^{\dagger} + \sum_{m, \sigma} \sum_{m', \sigma'} c_{m, \sigma}^{\dagger} c_{m', \sigma'} + \sum_{\Gamma, \sigma} \frac{1}{2} \sum_{m, \sigma} \frac{1}{\mathcal{N}} \left( \mathcal{D}_{\Gamma, \sigma} \mathcal{D}_{\Gamma, \sigma}^{\dagger} + \mathcal{D}_{\Gamma, \sigma}^{\dagger} \mathcal{D}_{\Gamma, \sigma} \right) + U_{dd} \sum_{(\Gamma, \sigma) \neq (\Gamma', \sigma')} \left| \mathcal{D}_{\Gamma, \sigma} \right|^2 \left| \mathcal{D}_{\Gamma', \sigma'} \right|^2 \\
&+ \sum_{\Gamma, \sigma} \frac{1}{\mathcal{N}} \left( \mathcal{D}_{\Gamma, \sigma} \mathcal{D}_{\Gamma, \sigma}^{\dagger} + \mathcal{D}_{\Gamma, \sigma}^{\dagger} \mathcal{D}_{\Gamma, \sigma} \right) - U_{dd} (2p) \sum_{(\Gamma, m, \sigma, \sigma')} \mathcal{D}_{\Gamma, m, \sigma}^{\dagger} \mathcal{D}_{\Gamma, m, \sigma} \left( 1 - \mathcal{P}_{m, \sigma, \sigma'} \right) + \mathcal{H}_{\text{multiplet}}.
\end{align}
$$

Ni$^{2+}$ in NiFe$_2$O$_4$ is located in the O$_8$ crystal field where oxygen atoms are octahedrally coordinated, and the number of $d$-electrons is $d^8$. In this effective Hamiltonian, we considered X-ray transitions from $|d^8\rangle + |d^9 L\rangle + |d^{10} L\rangle$ states to $|p^5 d^9 \rangle + |p^5 d^{10} \rangle$ states, where $d^{n+1} \frac{1}{2} \mathcal{D}$ means charge transfer of $\lambda$ electrons from the $p$-orbitals in the ligand (L) oxygen to $d$-orbitals in Ni [22]. In Equation (1), the first, second, and third terms are electron energies of the Ni-3d, Ni-2p and the oxygen 2p orbitals, respectively; the fourth term is a hybridization energy between the Ni-3d and oxygen 2p-orbitals; the fifth term is a Coulomb repulsion energy among the Ni-3d orbital electrons; and the sixth term is a Coulomb binding energy between holes in Ni-2p orbital and electrons in Ni-3d orbital; and the last term $\mathcal{H}_{\text{multiplet}}$ means effects of crystal field and spin-orbit interaction.

By using the physical quantities in a previous study for NiFe$_2$O$_4$ [17], 23 and 31 transition components are obtained for $\mp$ helicities, respectively. Vertical lines in Figure 1(a,b) denote these transition components for the respective helicities, where ordinates are in logarithmic scale and the ordinate for $+$ helicity is inverted upside down.

2.2. Synthesized spectra

Synthesized spectra were prepared on the basis of the transition components in Figure 1(a,b) with spectral convolution of Lorentz shapes $\mathcal{L}(x; \Theta)$ to mimic measured XA spectra. The $\mathcal{L}(x; \Theta)$ is expressed in Equation (2) as a function of energy $x$ and is characterized by a parameter set $\Theta (\Rightarrow (\mathcal{I}_x, \mathcal{E}_j, \mathcal{G}_0))$ with integral intensity $\mathcal{I}_x$, transition energy $\mathcal{E}_j$ and full-width at half maximum $\mathcal{G}_0$.

$$
\mathcal{L}(x; \Theta) := \frac{\mathcal{I}_x}{2\pi} \frac{\mathcal{G}_0}{(x - \mathcal{E}_j)^2 + (\mathcal{G}_0/2)^2}.
$$

XA spectra $\mathbf{D}^\mp \Rightarrow \{ (\cdots, (x_i, y_i^\mp), \cdots) \}$ for $\mp$ helicities were prepared by Equation (3), and an XMCD spectrum $\mathbf{D}^\text{XA} \Rightarrow \{ (\cdots, (x_i, y_i^\text{XMCD}), \cdots) \}$ was obtained by Equation (4).

$$
y_i^\mp := \mathcal{F}^\mp (x_i; \{ \Theta_0^\mp \}) + \mathcal{N}_i(0, \sigma_\text{noise}^\text{XA}),
$$

$$
y_i^\text{XMCD} := y_i^\mp - y_i^+, \quad y_i^+ \Rightarrow \mathcal{F}^+ (x_i; \{ \Theta_0^+ \}) := \sum_{j=1}^{f^+} \mathcal{L}(x_i; \Theta_j^+),
$$

where $\mathcal{F}^\mp (x; \{ \Theta_0^\mp \})$ are the true XA spectra for $\mp$ helicities and they consist of $f^- = 23$ and $f^+ = 31$ transition components, respectively, as described in Section 2.1. Energies $x_i$ were prepared from $-30$ to $+30 \text{ eV}$ with $N = 601 (i = 1, \cdots, N)$ data points in an arithmetic series, and all Lorentz shapes were given a common 1.00 eV spectral width ($\mathcal{G}_0$). The respective blue and red lines in Figure 1(a,b) are the true spectra $\mathcal{F}^\mp (x; \{ \Theta_0^\mp \})$ for $\mp$ helicity XA spectra. $\mathcal{N}_i(0, \sigma_\text{noise}^\text{XA})$ in Equation (3) means random noises in a normal
Figure 1. (a) and (b) Vertical lines denote transition components derived from the effective Hamiltonian in Equation (1). Their intensity is displayed in logarithmic scales and +/− helicity components are shown upside down. Blue and red curves indicate the true XA spectra for +/− helicity without normal noises. Gray spectra are +/− helicity XA spectra (D+) to analyze, and dashed lines mean the standard deviation of superimposed normal noises in the synthesized XA spectra. Gray spectra in (c), (d) and (e) are +/− helicity XA spectra (D+) and an XMCD spectrum (DXMCD) to analyze. The solid curves indicate reproduced spectra obtained by Bayesian spectroscopy.

distribution with mean and standard deviation being 0 and σ_noise², respectively. The gray areas in Figure 1(a,b) denote the +/− helicity XA spectra D+ to analyze, and in D−, normal noises with σ_noise² = 5.00 × 10⁻⁴ in absorption intensity scale were superimposed independently. As a result, the standard deviation of noises superimposed on the XMCD spectrum becomes √2σ_noise² and is larger than that of the +/− helicity XA spectra, since the XMCD spectrum was obtained by Equation (4). The synthesized spectra D−, D+ and DXMCD are shown by gray spectra in Figure 1(c-e), respectively, where the ordinate in Figure 1(d) is inverted for the + helicity XA spectrum.

3. Bayesian spectroscopy

In this section, we outline the formulation of Bayesian spectroscopy and a replica exchange Monte Carlo method [23] used for sampling.

3.1. Bayesian spectroscopy

Let D := {\cdots, (x_i, y_i), \cdots} (i = 1, \cdots, N) be a spectral dataset to analyze and f_k(x_i; \Theta) be a phenomenological model for describing D, then a posterior probability distribution P(\Theta|D, K, b) for the parameter set \Theta in the model f_k(x_i; \Theta) can be expand as Equation (6) by applying the Bayes’ theorem [3] between D and \Theta [1,2].

\[ P(\Theta|D, K, b) = \frac{P(D|\Theta, K, b)P(\Theta|K, b)}{P(D|K, b)}, \]

where K is a model identifier and b is quasi-inverse temperature [20] defined as b := σ_noise⁻² with the standard deviation σ_noise of the superimposed noises in y_i. P(\Theta|K, b) is a prior probability for \Theta, and P(D|\Theta, K, b) is a conditional probability of D under the condition \Theta given, and the denominator P(D|K, b) is a normalization factor.

When the noises in D are distributed in a normal distribution with the inverse variance b, P(D|\Theta, K, b) in Equation (6) can be written as Equation (7).

\[ P(D|\Theta, K, b) = \left(\frac{b}{2\pi}\right)^{\frac{N}{2}} \exp\left[-bN\mathcal{E}_K(\Theta)\right], \]

where \mathcal{E}_K(\Theta) is an error function of f_k(x_i; \Theta) for D as follows:

\[ \mathcal{E}_K(\Theta) = \frac{1}{2N} \sum_{i=1}^{N} [y_i - f_k(x_i; \Theta)]^2. \]

The denominator P(D|K, b) in Equation (6) is named a Bayesian partition function Z(K, b) [1] and is obtained by marginalization of numerator terms with \Theta as follows:

\[ Z(K, b) := P(D|K, b) = \left(\frac{b}{2\pi}\right)^{\frac{N}{2}} \int \exp\left[-bN\mathcal{E}_K(\Theta)\right] P(\Theta|K, b) d\Theta. \]

With this Z(K, b), BFE F(K, b) is defined as F(K, b) := −ln Z(K, b) [1]. By minimizing F(K, b) as Equation (8), we can simultaneously perform model selection and estimate the noise intensity superimposed in D [20].
\[ \hat{K}, \hat{b} = \text{argmin}_{K,b} F(K,b), \quad (8) \]

where \( \hat{K} \) is the model identifier for selected one, and a standard deviation \( \hat{\sigma}_{\text{noise}} \) of the normal noises in \( D \) is estimated by \( \hat{\sigma}_{\text{noise}} := \hat{b}^{-1/2} \).

Furthermore, the posterior probability of each candidate model can be evaluated. Although \( P(K|D) \) is expanded as \( P(K|D) = P(D|K)P(K)/P(D) \) on the basis of the Bayes' theorem [3], in unbiased model selection, \( P(K|D) \) is simply proportional to \( P(D|K) \) because the prior probabilities \( P(K) \) of all candidate models should be the same. Since the conditional probability \( P(D|K) \) can be evaluated by marginalization of \( P(D|K,b) \) with \( b \), the \( P(K|D) \) is obtained as follows:

\[ P(K|D) \propto \int P(D|K,b) db = \int Z(K,b) db \quad (9) \]

Finally, the posterior probability distribution of the parameter set \( \Theta \) in the selected model \( \hat{K} \) can be sampled with the estimated quasi-inverse temperature \( \hat{b} \) by the following equation:

\[ P(\Theta|D, \hat{K}, \hat{b}) \propto \exp\left[ -bN_{\mathcal{E}_K(\Theta)} \right] P(\Theta|\hat{K}, \hat{b}). \quad (10) \]

### 3.2. Replica exchange Monte Carlo method

To perform model selection [1,20], a replica exchange Monte Carlo (RXMC) method is necessary to sample the parameter space \( \Theta \) on multiple replicas with different quasi-inverse temperatures \( b \). In the RXMC method, a state of \( \Theta \) is updated by a Metropolis method [24–26] in each replica, and simultaneously the state \( \Theta \) is exchanged stochastically between neighbor replicas. Although many replicas are prepared for a wide range of \( b \), convexo-concave structures of the error function \( \mathcal{E}_K(\Theta) \) are suppressed on the replicas with small \( b \) since \( b \) is multiplied on \( \mathcal{E}_K(\Theta) \) in the exponential function of Equation (7). As a result, rapid search in the wide parameter space is realized on the replica having small \( b \), and through state exchanges, release of the state \( \Theta \) from local minima of \( \mathcal{E}_K(\Theta) \) is also realized. Consequently, rapid and effective samplings become possible nearby the global optimal solution by the RXMC.

We prepared commonly eighty replicas \( (L := 80) \) for both analyses of the \( \mp \) helicity XA and XMCD spectra. Although the smallest \( b \) being \( b_1 = 0 \), other replicas were prepared with geometric sequences on \( b_k \) \( (k = 2, \cdots, L) \). Although this range of \( b_k \) should be adjusted to include the inverse variance of the normal noises in \( D \), we set \( b_1 = 7.2 \times 10^3 \) and \( b_2 = 7.2 \times 10^6 \) for the XA spectra and \( b_1 = 2.0 \times 10^4 \) and \( b_2 = 2.0 \times 10^7 \) for the XMCD spectrum, respectively. The corresponding inverse-variances of the superimposed noises are \( 4.0 \times 10^6 \) \( = (\sigma_{\text{noise}}^{XA})^{-2} \) for the \( \mp \) helicity XA spectra and \( 2.0 \times 10^6 \) \( = (\sqrt{2} \sigma_{\text{noise}}^{XMCD})^{-2} \) for the XMCD spectrum, and they are included in the respective ranges of \( b_k \) \( ( = 2, \cdots, L) \). In both cases, the ratio \( b_{k+1}/b_k \) between neighbor replicas is 1.0926, and since it is close to unity, rapid and wide-range searching on the replicas having small \( b \) will be promoted through highly efficient state exchange between neighbor replicas. The RXMC samplings were performed at 200,000 times after sufficient burn-in phase of 200,000 times.

### 4. Results

Our purpose is to demonstrate the effectiveness of Bayesian spectroscopy in difference spectral analysis such as the XMCD spectra. While the spectral structure of XMCD becomes more complex, there is a concern that some components may be missing when one try to extract the original \( \mp \) helicity components from the XMCD spectrum. Therefore, in this section, we first describe the model selection on the individual \( \mp \) helicity XA spectra. After that, the result of model selection for the XMCD spectrum is presented. In the last part, we show posterior probability distributions of spin, orbital, and total magnetic moments to evaluate their estimation accuracies.

#### 4.1. Model selections for \( \mp \) helicity XA spectra

The gray spectra in Figure 1(a–d) are \( \mp \) helicity XA spectra, respectively. To perform the spectral decomposition of these spectra, we employ phenomenological models \( f^\mp(x; \Theta^\mp) \) with the Lorentz shapes \( \mathcal{L}(x; \Theta) \) as follows:

\[ f^\mp(x; \Theta^\mp) := \sum_{k=1}^{K^\mp} \mathcal{L}(x; \Theta_k^\mp), \]

where \( K^\mp \) are the numbers of spectral components in the \( \mp \) helicity XA spectra. In this case, the model selection based on Equation (8) is equivalent to estimating the number of spectral components \( K^\mp \) in the model \( f^\mp(x; \Theta^\mp) \). The parameter sets for the respective Lorentz shapes are \( \Theta_k^\mp := \{ I_k^\mp, E_k^\mp, \Gamma_k^\mp \} \), where \( I_k^\mp, E_k^\mp \) and \( \Gamma_k^\mp \) mean integral intensity, transition energy and spectral width, respectively, and for \( I_k^\mp \) and \( \Gamma_k^\mp \), we introduce non-negative constraints. The spectral widths \( \Gamma_k^\mp \) were free parameters for each spectral component to estimate independently although the spectral widths of all Lorentz shapes were set to a common width \( (= 1.0\text{eV}) \) in the synthesized spectra.

We perform model selection to estimate \( K^\mp \) in the respective XA spectra for \( \mp \) helicities using BFE as an information criterion, and the results of the \( \mp \) helicity XA spectra are shown in Figure 2(a,b), respectively. Line graphs connecting open squares show the variations of the minimized BFE \( = F(K^\mp, \hat{b}_{K^\mp}) \) with the numbers \( K^\mp \) of spectral components, in which the minimized
values are determined from the variation of BFE with respect to \(b_{K^\pm}\). The values of \(b_{K^\pm}\) that minimize BFE are \(b_{K^-=8} = 4.23 \times 10^6\), \(b_{K^+=8} = 3.55 \times 10^6\) for the \(\mp\) helicity XA spectra, respectively, and they hardly change in the whole ranges of \(K^- = 7\sim9\) and \(K^+ = 6\sim8\). From \(b_{K^\pm}\), it is possible to estimate the standard deviations of the noises superimposed on the target spectra, and these standard deviations are \(4.86 \times 10^{-4}\) and \(5.31 \times 10^{-4}\) in signal intensity scale for the \(\mp\) helicity XA spectra, respectively, which are in close agreement with the standard deviation \(\sigma_{\text{noise}}\) \((= 5.00 \times 10^{-4})\) of the noises in the synthesized XA spectra.

On the other hand, bar graphs in Figure 2(a,b) show the posterior probabilities \(P(K^\pm|D^\mp)\) for the model selection of the respective models \((K^- = 7\sim9, K^+ = 6\sim8)\), where \(P(K^\pm|D^\mp)\) are calculated by Equation (9) and displayed in logarithmic scales. As seen in these figures, models containing eight spectral components are coincidentally selected in the both spectral decompositions of \(\mp\) helicity XA spectra, and it is obvious that the selected models \((K^- = 8, K^+ = 8)\) have high posterior probabilities \(P(K^\pm|D^\mp)\) for model selection.

Although 23 and 31 transition components were derived for the \(\mp\) helicity components from the effective Hamiltonian, respectively, 8 spectral components \((K^\mp = 8)\) were extracted for both helicities in the spectral decomposition of the respective \(\mp\) helicity XA spectra. However, these decreases in the number of decomposed components are reasonable under consideration of that the merging of the near transition components within the spectral width and the competition between the peak intensity of each spectral component and the superimposed noise intensity. A detailed discussion will be given in Section 5.1.

In the selected models \((K^- = 8, K^+ = 8)\), we obtained the posterior probability distributions \(P(\Theta_{k^\pm}|D^\mp, K^\pm, b_{K^\pm})\) according to Equation (10) through the RXMC samplings, and evaluated the mean values \(\bar{\Theta}_{k^\pm}\) of the respective parameters and the standard deviations \(\sigma_{\Theta_{k^\pm}}\) of their posterior probability distributions. These results are summarized in the center-column group of Table 1 in the form of \(\bar{\Theta}_{k^\mp} \pm \sigma_{\Theta_{k^\pm}}\), in which (a) and (b) show the results for the \(\mp\) helicity XA spectra, respectively. As categorized in the first column of Table 1, the spectral components having negative transition energies \((E_{k^\mp})\) are the \(L_3\)-absorption edge, and the subsequent components are the \(L_2\)-absorption edge.

Solid curves in Figure 1(c,d) show reproduced \(\mp\) helicity XA spectra with the mean values \(\bar{\Theta}_{k^\mp}\) in Table 1, respectively, and it is found that they explain well the respective gray spectra \(D^\mp\). Root-mean-square
deviations (RMSDs) of these reproduced spectra to the synthesized spectra \( D^\mp \) are shown in the legends of Figure 1(c,d). Although the RMSDs are sufficiently small, they are about 11−38% larger than the standard deviation \( \sigma_{\text{noise}} = 5.00 \times 10^{-4} \) of the normal noises superimposed on \( D^\mp \). This increase in RMSDs is considered to be due to the fact that the synthesized spectra contain weaker transition components than the magnitude of noises, while Bayesian spectroscopy may lose such weak components.

![Figure 3](image-url)

**Figure 3.** (a) and (b) Blue and red curves denote the regressive XA spectra for \( \mp \) helicities, respectively. Dashed curves are the spectral components \( \mathcal{L}(x; \theta^\mp_\kappa) \) included in these regressive spectra. (c) and (d) Blue and red curves denote the \( \mp \) helicity XA spectra synthesized with the respective parameter sets \( \{ \theta^\mp_\kappa \} \) and \( \{ \theta^\mp_\kappa \} \), which are obtained by the spectral decomposition of the XMCD spectrum. Dashed curves are the spectral components \( \mathcal{L}(x; \theta^\mp_\kappa) \). The ordinate is displayed in a logarithmic scale, and horizontal dashed lines mean the standard deviation of the superimposed normal noises in the respective spectra. Transition components and \( \mp \) helicity spectra (\( \theta^\mp \)) are also indicated in the same manner with Figure 1(a) and (b).
To show the respective spectral components decomposed by Bayesian spectroscopy, Figure 3 was prepared. In this figure, the spectral intensities are shown on the logarithmic scale, and the superimposed noise intensities $\hat{\sigma}_{\text{noise}}^{XA}$ are depicted by horizontal dashed lines. The spectra depicted by dashed curves are the respective spectral components $L(x_i; \hat{\Theta}_k^{\pm})$, and Figure 3(a,b) are for the XA spectra of $\mp$ helicities, respectively. The spectra shown in gray area are the $\mp$ helicity XA spectra with noise superimposed, and are the same as those shown in Figure 1(a,b). Considering the noise intensity $\hat{\sigma}_{\text{noise}}^{XA}$, the reproduced spectra indicated in blue and red curves are good reproductions of the $\mp$ helicity XA spectra, respectively.

In Section 5, we will discuss the mapping of these decomposed spectral components by Bayesian spectroscopy to the transition components derived from the effective Hamiltonian in Equation (1). Here, we would like to focus on the values of the estimated spectral widths. In Table 1, it is found that the spectral widths $\tilde{\Gamma}_k^\pm$ for the spectral components having intense transition intensities ($I_k^\pm > 10 \times 10^{-3}$) are around 1.0 eV, and the ground truth of the broadening factor $\Gamma_0$ ($= 1.00$ eV) in the XA spectra is estimated correctly.

### 4.2. Model selections for XMCD

In this study, we attempt to extract the original $\mp$ helicity XA spectral components from the XMCD spectrum. To realize this, we employ a phenomenological model in Equation (11) for the XMCD spectrum $D^{\text{XMCD}}$. 

$$f^{\text{XMCD}}(x_i; \Theta^{\text{XMCD}}) : = \sum_{k=-1}^{K^-} L(x_i; \Theta_k^-) + \sum_{k=1}^{K^+} L(x_i; \Theta_k^+),$$ 

(11)

where $K^-$ and $K^+$ are the numbers of $-$ helicity and $+$ helicity spectral components needed to describe the XMCD spectrum, and in this case, the model selection means the simultaneous estimation both of $K^-$ and $K^+$, and where we used the notations of $K^\mp$ and $K^\pm$ for the numbers of spectral components and the indexes for the respective components, respectively, to avoid confusion with the case of individual XA spectra. In the model of Equation (11), we introduce non-negative ($I_k^+ > 0$) and negative ($I_k^- < 0$) constrains on the integral intensities to distinguish $\mp$ helicity spectral components. The other settings of the prior probabilities are the same as in Section 4.1.

On the basis of Equation (11), we perform the model selection to estimate $K^-$ and $K^+$ using the BFE $F(K^-, K^+, \hat{b}_{K^-, K^+})$ as the information criterion, and then, the model of $\hat{K}^+ = 8$ and $\hat{K}^- = 9$ is selected, in which the quasi-inverse temperature minimizing the BFE is $\hat{b}_{K^+, K^-} = 2.00 \times 10^6$. Figure 2(c) shows a heat map of the selection posterior probabilities $P(K^+, K^- | D^{\text{XMCD}})$ in the range of $K^+ = 6$–8 and $K^- = 7$–9, and it is confirmed that the selected model ($\hat{K}^+ = 8$, $\hat{K}^- = 9$) has the highest posterior probability. From the value of $\hat{b}_{K^+, K^-}$, the standard deviation of the superimposed noises in the XMCD spectrum is estimated to be $7.07 \times 10^{-1}$ in the XMCD signal intensity scale, and it is completely coincide with $\sqrt{2} \hat{\sigma}_{\text{noise}}^{XA}$ (see Section 2.2).

Compared to the model selection results ($\hat{K}^+ = 8$, $\hat{K}^- = 8$) for individual XA spectra in Section 4.1, the selected model in the XMCD spectrum has one more $-$ helicity component. Although this result seems strange at first glance, we will discuss the details in Section 5.2.

As similar with Section 4.1, we sampled the posterior probability distributions of $\Theta^{\text{XMCD}}$, and in the center-column group of Table 2, their mean values $\bar{\Theta}_k^\mp$ and the standard deviations $\sigma_{\Theta_k^\mp}$ of the posterior probability distributions are summarized in the form of $\bar{\Theta}_k^\mp \pm \sigma_{\Theta_k^\mp}$. Table 2(a,b) are for the $\mp$ helicity spectral components, respectively, and the $L_3$- and $L_2$-absorption edges are also categorized in the first column of Table 2.

A reproduced spectrum using $\bar{\Theta}_k^\mp$ is depicted by a solid curve in Figure 1(e), and it explains well the gray-colored XMCD spectrum. The RMSD of the reproduced spectrum to $D^{\text{XMCD}}$ is indicated in the legend of Figure 1(e). Although the RMSD is larger than the RMSDs for $\mp$ helicity XA spectra [see Figure 1(c,d)], in the case of the XMCD spectrum, it is reasonable because the standard deviation of the superimposed normal noises is $\sqrt{2}$ times larger ($= 7.07 \times 10^{-1}$) than the respective XA spectra, as described in Section 2.2.

### 4.3. Posterior probability distributions of magnetic moments

Another advantage of Bayesian spectroscopy is that, only from one dataset, we can evaluate posterior probability distributions of physical quantities. According to the sum rule [18], measurement of XMCD spectra covering the $L_{3,2}$-absorption edge allow us to evaluate orbital $m_{\text{orb}}$, spin $m_{\text{spin}}$ and total $m_{\text{tot}}$ magnetic moments as well as their ratio $m_{\text{ratio}}$ ($:= m_{\text{orb}} / m_{\text{spin}}$) on the basis of spectral integral of the $L_{3,2}$ edge. However, the conventional method by spectral integration is point estimation, and the evaluation of estimation accuracy is quite difficult.

On the other hand, as demonstrated in Section 4.2, Bayesian spectroscopy enables us to decompose...
Table 2. Decomposed spectral components from the XMCD spectrum are summarized in the center-column group. The right-column group means the corresponding transition components (discussed in Section 5.1). In the left column, $L_2$ and $L_3$ of absorption edges are categorized, and (a) and (b) are the $\uparrow$ helicity spectral components in the XMCD spectrum, respectively.

| $L_{1,2}$ | $\kappa^\pm$ | $\epsilon_\pm \pm \sigma_\pm$ | $\epsilon_\mp \pm \sigma_\mp$ | $\sum \epsilon_\pm^\pm$ | $f_\pm$ | $f_\mp$ |
|-----------|---------------|--------------------------------|--------------------------------|----------------|-------|-------|
| $L_3$     |               |                                |                                | $\times 10^{-3}$ | $\times 10^{-3}$ | $\times 10^{-3}$ |
| i         | +28 ± 10      | $-8.92 \pm 0.15$               | $1.10 \pm 0.41$               | +27.21          | -9.063 | -9.063 |
| ii        | +56 ± 12      | $-8.965 \pm 0.087$             | $1.10 \pm 0.28$               | +54.41          | -9.020 | -9.020 |
| iii       | +344 ± 3.2    | $-7.302 \pm 0.017$             | $1.109 \pm 0.065$             | +36.63          | -7.297 | -7.297 |
| iv        | +7.3 ± 7.2    | $-4.7 \pm 5.6$                 | $8.5 \pm 8.0$                 | +1.466          | -3.343 | -3.343 |
| v         | +7.5 ± 4.2    | $-3.00 \pm 0.94$               | $2.4 \pm 2.3$                 | +4.610          | -2.816 | -2.816 |
| $L_2$     |               |                                |                                | $\times 10^{-3}$ | $\times 10^{-3}$ | $\times 10^{-3}$ |
| i         | -114 ± 0.14   | $-8.962 \pm 0.059$             | $1.00 \pm 0.11$               | -118.27         | -8.976 | -8.976 |
| ii        | -92 ± 13      | $-8.819 \pm 0.042$             | $0.99 \pm 0.11$               | -88.87          | -8.846 | -8.846 |
| iii       | -3.8 ± 4.8    | $-7.8 \pm 3.3$                 | $9.1 \pm 6.4$                 | -2.923          | -7.463 | -7.463 |
| iv        | -5.8 ± 4.5    | $-6.7 \pm 6.2$                 | $9.7 \pm 7.8$                 | -1.361          | -6.972 | -6.972 |
| v         | -4.7 ± 3.9    | $-6.02 \pm 0.41$               | $1.3 \pm 1.0$                 | -3.127          | -5.955 | -5.955 |
| vi        | -6.7 ± 4.5    | $-3.7 \pm 1.2$                 | $2.4 \pm 2.5$                 | -5.080          | -3.278 | -3.278 |
| $L_3$     |               |                                |                                | $\times 10^{-3}$ | $\times 10^{-3}$ | $\times 10^{-3}$ |
| i         | -28 ± 11      | $+8.48 \pm 0.17$               | $1.30 \pm 0.33$               | -31.35          | +8.405 | +8.405 |
| ii        | -6.4 ± 7.3    | $10.4 \pm 6.5$                 | $11 \pm 11$                   | -1.571          | +10.298 | +10.298 |

separately the $\uparrow$ helicity spectral components from the XMCD spectrum, and samplings of the posterior probability distributions $P(m_i|x)$ (x = orb., spin, tot., and ratio) become available from the integral intensities ($I_k$ and $I_{k'}$) of the decomposed $\uparrow$ helicity spectral components since the magnetic moments $m_{\text{orb}}$ and $m_{\text{spin}}$ can be evaluated by Equations (12) and (13) [18] in a unit of $\mu_B$/atom, respectively.

$$m_{\text{orb}} = \frac{4 \sum_{l_1, l_2} (I_k - I_{k'}) \times n_h}{3 \sum_{l_1, l_2} (I_k + I_{k'})} \times n_h,$$

Equation (12)

$$m_{\text{spin}} = \frac{6 \sum_{l_1} (I_k - I_{k'}) - 4 \sum_{l_1, l_2} (I_k - I_{k'})}{\sum_{l_1, l_2} (I_k + I_{k'})} \times n_h,$$

Equation (13)

where $n_h$ is the occupancy number of 3d-holes of Ni ions and $n_h = 1.832$ [27], since the effective occupation number of 3d-electrons is evaluated to be 8.168 from the effective Hamiltonian in Equation (1).

Posterior probability distributions of $m_{\text{orb}}$, $m_{\text{spin}}$, $m_{\text{tot}}$, and $m_{\text{ratio}}$ are shown by color-filled graphs in Figure 4(a-d), respectively. Black lines in Figure 4 indicate the true values of these quantities, which are obtained from the effective Hamiltonian. On the other hand, red lines and error bars in Figure 4 are the mean values of these quantities and the standard deviations of their posterior probability distributions, and it is found that the true values of all these quantities are contained within the respective posterior probability distributions. This result demonstrates that the

Figure 4. Posterior probability distributions of (a) orbital, (b) spin and (c) total magnetic moments. Black lines denote the true values obtained by the effective Hamiltonian. Red lines and error bars means the mean values of the respective magnetic moments and their standard deviation of the posterior probability distributions.
magnetic moments are correctly estimated, and the evaluation of estimation accuracy can be realized with the distribution widths of their posterior probabilities.

Such evaluation of the posterior probability distributions is meaningful when measurements with a high signal-to-noise (S/N) ratio are difficult. Although XMCD microscopic measurements [28], which capture magnetic domain in magnetic materials, become available with recent developments [29] in synchrotron radiation measurement techniques, it is still difficult to obtain a high S/N ratio for in situ measurements that capture the dynamics of magnetic domain formation. Of course, although the estimation accuracy will degrade with bad S/N ratios, the evaluation of posterior probability distributions by Bayesian spectroscopy becomes more important in such situations.

5. Discussion

In this section, we describe the attribution of the decomposed spectral components paying attention both to the noise intensity and the spectral widths in the target spectra. Subsequently, we will discuss the difference in the spectral decomposition results between the individual XA and XMCD spectra.

5.1. Attribution of decomposed spectral components

When performing spectral decomposition to extract weak transition components and to separate components with close transition energies, the magnitude of noises superimposed on the target spectrum and the spectral width of each transition component become obstacles. In fact, as described in Section 2.2, normally distributed noises with a standard deviation of \( \sigma_{\text{noise}}^{\text{XA}} = 5.00 \times 10^{-4} \) were superimposed on the synthesized \( \mp \) helicity XA spectra, and a spectral width of \( \Gamma_0 = 1.00 \text{eV} \) was convolved with each transition component derived from the effective Hamiltonian.

The first issue is the noise intensity, and competition with the noise intensity is considered to be characterized by the peak intensity of each spectral component. In the case of the Lorentzian shape [see Equation (2)] with a width of \( \Gamma_0 \), the spectral component with an integral intensity of \( I_{k^\mp} \) has a peak intensity of \( 2I_{k^\mp} / (\pi \Gamma_0) \). From all transition components derived from the effective Hamiltonian, the components having peak intensities beyond the noise intensity were summarized selectively in the right-column group of Table 1, where the transition intensity \( I_{j^\mp} \) being \( I_{j^\mp} > (\pi \Gamma_0 / 2) \times \sigma_{\text{noise}}^{\text{XA}} \) (\( \approx 7.85 \times 10^{-4} \)). The horizontal dashed lines in Figure 1(a,b) are the levels of transition intensities corresponding to the noise intensity. Although the index \( j \) in Equation (5) is scanned on all transition components derived from the effective Hamiltonian, and hereafter, the index \( j \) is used for the subset that satisfies this condition.

The second issue is that each transition component has a finite spectral width. Although, as seen in the right-column group of Table 1, there are thirteen \( (j^- = 1 \sim 13) \) and fourteen \( (j^+ = 1 \sim 14) \) transition components in \( \mp \) helicities, respectively, it is found that there are components whose transition energies \( E_{j^\mp} \) are in close proximity within the spectral width \( \Gamma_0 \). Such adjacent transition components are considered to be merged to one spectral component in the spectral decomposition by Bayesian spectroscopy, such merging is also shown in Table 1. For example, in Table 1(a), the transition components of \( j^-= 1, 2 \) are merged to the spectral component of \( k^- = i \) at the \( L_3 \)-absorption edge. In order to distinguish between the transition components specified by index \( j \), the spectral components decomposed by Bayesian spectroscopy are numbered using Roman numerals in Table 1 and in Table 2 (discussed later). Columns of \( \sum I_{j^\mp} \) and \( \sum E_{j^\mp} \) in Table 1 indicate the merged transition intensities and the weighted–mean values of the transition energies based on the transition intensity.

In Table 1, comparing the columns of \( I_{k^\mp} \pm \sigma_{k^\mp} \) and \( \sum I_{j^\mp} \), and of \( E_{k^\mp} \pm \sigma_{k^\mp} \) and \( \sum E_{j^\mp} \), it is clear that \( \sum I_{j^\mp} \) and \( \sum E_{j^\mp} \) are included within the respective posterior probability distributions of \( I_{k^\mp} \) and \( E_{k^\mp} \) of each spectral component decomposed by Bayesian spectroscopy. This result demonstrates that Bayesian spectroscopy is able to extract the transition components without excess or deficiency under consideration of the noise intensity and finite spectral width.

Such precise spectral decomposition by Bayesian spectroscopy is also realized by the spectral decomposition of XMCD although the noise intensity in the XMCD spectrum is \( \sqrt{2} \) times larger than that of the XA spectra. In Table 2, taking the noise intensity into account, we summarized the attribution of the decomposed spectral components to the transition components obtained from the effective Hamiltonian. This result also demonstrates that, from the XMCD spectrum, original spectral components of \( \mp \) helicities are extracted separately without excess or deficiency by Bayesian spectroscopy.

Figure 5 demonstrates the attributions of the decomposed spectral components by Bayesian spectroscopy and the transition components derived from the effective Hamiltonian. Figure 5(a,b) are for the results of the individual \( \mp \) helicity XA spectra, respectively, and Figure 5(c,d) are for the \( \mp \) helicity components decomposed from the XMCD spectrum, respectively, where the ordinates are in logarithmic scales and are inverted upside down for \( \mp \) helicity ones. The horizontal dashed lines indicate the noise-intensity levels in the integral intensity scale.
Vertical lines in Figure 5 mean the merged transition components in the right-column group of Tables 1 and Tables 2, which are displayed with $E_{f}^\pm$ and $\sum I_r^\pm$. Open squares in Figure 5(a,b) are the points of $\{E^\pm, I_r^\pm\}$ for the decomposed $\mp$ helicity components from the $\mp$ helicity XA spectra, respectively. On the other hand, the $\{E^\pm, I_r^\pm\}$ are plotted by open lozenges in Figure 5(c,d) for the respective $\mp$ helicity components decomposed from the XMCD spectrum. The error bars accompanying those open marks are standard deviations of the respective posterior probability distributions and indicate the estimation accuracy of each mean value. In Figure 5, one can confirm that most of the points indicated by the vertical lines are included in the ranges of the respective posterior probability distributions. This result clearly implies that, even in the case of XMCD, the original $\mp$ helicity components can be extracted separately without excess or deficiency. Such challenging spectral decomposition can be realize with Bayesian spectroscopy based on the model selection using BFE as the information criterion.

The results of such challenging spectral decomposition can be confirmed in Figure 3(c,d). The dashed curves are the respective spectral components $\mathcal{L}(\tau, \Theta_{X})$ for $\mp$ helicities decomposed from the XMCD spectrum, and the blue and red spectra are the spectra where we attempt to reconstruct the $\mp$ helicity XA spectra as the sum of those components. Amazingly, it is found that the blue and red reconstructed spectra well explain the original $\mp$ helicity XA spectra, which are depicted by gray spectra for the guides to the eyes, under consideration of the superimposed noise intensity $\sigma_{\text{noise}}^{\text{XMCD}}$. This result implies that Bayesian spectroscopy has successfully reproduced the original $\pm$-helicity spectra from the XMCD spectrum, even though they were hidden from the human eye in the XMCD spectrum. However, comparing the upper and lower panels in Figure 3, it is found that the spectra (XMCD) in the lower panel do not completely reproduce the upper ones ($\mp$ XA). Although this difference might be due to the difference in noise intensities ($\sigma_{\text{noise}}^{\text{XMCD}}, \sigma_{\text{noise}}^{\text{XA}}$), in order to solve this problem, Bayesian-integration [30] study of $\mp$ helicity XA and XMCD spectra is expected in the future.

The decomposition of XMCD spectra and the evaluation of magnetic moments based on Bayesian spectroscopy are expected to be widely applied to various magnetic materials such as nickel ferrite [17]. As described in this section, spectral components weaker than the noise intensity are impossible to be decomposed even with Bayesian spectroscopy, and a low signal-to-noise ratio results in a broadening of the posterior probability distribution of magnetic moments and a decrease in estimation accuracy. Conversely, however, it can be claimed that Bayesian spectroscopy is the method to decompose the spectral components without deficiency when it is difficult to improve the signal-to-noise ratio such as the case of in situ and operand measurements. The same analysis was performed using synthesized data with twice the noise intensity, and the spectral components exceeding the noise intensity were appropriately decomposed, and this result confirms the advantage of Bayesian spectroscopy for XMCD spectral analyses.
5.2. Difference between the spectral decomposition results for XA and XMCD spectra

As described in Section 2.2, the standard deviation of the superimposed noises in the XMCD spectrum is $\sqrt{2}$ times larger than those in the XA spectra. Such degradation of the S/N ratio may be unfavorable for the extraction of weak spectral components. Such concern occurs at the $f^- = 7$, $f^+ = 13$ components in the right-column group of Table 2. Although these transition components have been treated as the components to be merged for the XA spectra in Table 1 because of $I_{f^+} > \sigma_{\text{noise}}^{\text{XA}}$, they are excluded from the merged transition components for the XMCD spectrum as commented in Table 2 because of $I_{f^-} < \sigma_{\text{noise}}^{\text{XMCD}}$. Fortunately, there is a strong component in the vicinity of those excluded components as seen in Table 2, so the numbers of the spectral components did not decrease in this spectral decomposition of the XMCD spectrum.

However, in the actual results (see Section 4.2), the model that includes one more $-$ helicity components ($K^- = 9$) was selected by the information criterion of BFE than the result of the spectral decomposition for the XA spectrum. This difference can be understood by comparing Figure 5(a,c), where the same spectral components are shown in the same color for the upper (XA) and lower (XMCD) panels. Let’s focus on the energy region around $-9.0$eV, we see that only a brown component appears for the $-$ helicity XA spectrum in Figure 5(a). On the other hand, for the $-$ helicity components of the XMCD spectrum in Figure 5(c), brown and red components are recognized in close proximity.

It is considered that such highly resolved spectral decomposition can be realized by Bayesian spectroscopy because the XMCD spectrum will contain material informations of such physical properties. In fact, in the energy region around $-9.0$eV, two components of $K^+ = i$ and $ii$ for $+$ helicity and two components of $K^- = i$ and $ii$ for $-$ helicity are decomposed in close proximity for the XMCD spectrum. In other words, the result of the XMCD-spectrum decomposition means that, by using the BFE as the information criterion, the necessity of these four components of opposite signs is proposed by the approach of data-driven science to explain the asymmetric spectral shape of the $L_3$ edge shown in the XMCD spectrum of Figure 1(e).

To begin with, the XMCD spectra have been measured to elucidate spin-selected electronic states under magnetic fields. In order to detect such precise splitting of the electronic states, difference spectral measurements such as XMCD spectra are advantageous. In addition, it is clear that Bayesian spectroscopy based on the model selection is important method to make the sense of such advantage in the XMCD measurements.

6. Conclusion

In order to demonstrate the advantage of Bayesian spectroscopy in difference spectrum analysis, spectral decomposition of $\mp$ helicity XA and XMCD spectra was performed based on the model selection by using BFE as an information criterion. The target spectra were prepared and synthesized by transition components derived from an effective Hamiltonian to imitate the $L_{3,2}$-absorption edge of Ni$^{2+}$ in NiFe$_2$O$_4$. In addition, to mimic measured spectra, random noises of normal distribution and a spectral width of $1.00$ eV were embedded in these synthesized spectra.

In the respective $\mp$ helicity XA spectra, 8 spectral components ($K^\mp = 8$) were decomposed in both cases. Although these numbers of spectral components are less than the numbers of transition components derived from the effective Hamiltonian, we have succeeded in decomposing significant spectral components without excess or deficiency considering the noise intensity superimposed on the data and the merging of transition components being close to each other within the spectral width.

Furthermore, we have succeeded in extracting separately the original $\mp$ helicity components without excess or deficiency from the difference spectrum, XMCD spectrum, by Bayesian spectroscopy based on appropriate model selection. The selected model contains 9 $-$ helicity and 8 $+$ helicity components, and one more $-$ helicity component was extracted than in the individual $-$ helicity XA spectrum at the $L_3$-absorption edge. Although the $L_3$-absorption edge originally consists of many transition components with opposite signs in close proximity, this is a result of the data-driven science approach using Bayesian spectroscopy, which determines that an additional $-$ helicity component is necessary to explain the asymmetric spectral profile at the $L_3$ edge in the XMCD spectrum.

In addition, we proposed to use the posterior probability distribution obtained by Bayesian spectroscopy for a method to estimate the magnetic moments. Bayesian spectroscopy allows us to decompose separately the $\mp$ helicity components over the $L_{3,2}$- absorption edges, and on the basis of the respective integral intensities, the samplings of the spin $m_{\text{spin}}$, orbital $m_{\text{orb}}$, and total $m_{\text{tot}}$ magnetic moments and the ratio $m_{\text{orb}}/m_{\text{spin}}$ become available to obtain their posterior probability distributions. Such evaluation of posterior probability distributions, i.e. estimation accuracies, is expected to become important in severe analyses with lower S/N ratios, and its application to $\mu$-XMCD and in situ measurement data is strongly desired in the future.

Disclosure statement

No potential conflict of interest was reported by the authors.
Funding
This study was supported by JST, CREST, JPMJCR 1861 and 1761.

References
[1] Nagata K, Sugita S, Okada M. Bayesian spectral deconvolution with the exchange Monte Carlo method. Neural Netw. 2012;28:82.
[2] Akai I, Iwamitsu K, Okada M. Bayesian spectroscopy in solid-state photo-physics. J Phys: Conf Series. 2018;1036:012022.
[3] Bayes T, Price R. An essay towards solving a problem in the doctrine of chances. Phil Trans Roy Soc A. 1763;53:370.
[4] Ramprasad R, Batra R, Pilania G, et al Machine learning in materials informatics: recent applications and prospects. Npj Comput Mat. 2017;3:1.
[5] Furukawa Y, Nakayama M. Dynamical formation process of electron-hole droplets in a GaAs/AlAs type-II superlattice. J Phys Soc Jpn. 2016;85:034701.
[6] Iwamitsu K, Furukawa Y, Nakayama M, et al Bayesian spectroscopy of admixed photoluminescence spectra with exciton, bixcition and electron hole droplet states in a GaAs/AlAs type-II superlattice. J Lumin. 2018;197:18.
[7] Mototake Y, Mizumaki M, Akai I, et al The Bayesian hamiltonian selection in x-ray photoelectron spectroscopy. J Phys Soc Jpn. 2019;88:034004.
[8] Aihara S, Hamamoto M, Iwamitsu K, et al High precision modeling of a damped oscillation in coherent phonon signals by bayesian inference. AIP Adv. 2017;7:045107.
[9] Kiridoshi A, Aihara S, Arishima S, et al Bayesian spectroscopy on polarization dependent photoluminescence spectra of doubly-split excitons in a Cu$_2$O thin-crystal sandwiched by MgO substrates. Phys Status Solidi (B). 2018;255:1800136.
[10] Iwamitsu K, Okada M, Akai I. Bayesian spectroscopy with a replica exchange monte carlo method for study of a biaxial stress effect on excitons in a Cu$_2$O thin crystal. J Phys: Conf Series. 2019;1220:012009.
[11] Yakura S, Iwamitsu K, Hira S, et al Strained lateral structure and its relaxation in a Cu$_2$O thin crystal epitaxially grown on MgO surface. Jpn J Appl Phys. 2020;59:025506.
[12] Iwamitsu K, Okada M, Akai I. Spectral decomposition of components weaker than noise intensity by bayesian spectroscopy. J Phys Soc Jpn. 2020;89:104004.
[13] Iwamitsu K, Yokota T, Murata K, et al Spectral analysis of X-ray absorption near edge structure in α-Fe$_2$O$_3$ based on Bayesian spectroscopy. Phys Status Solidi (B). 2020;257:2000107.
[14] Schütz G, Wagner W, Wilhelm W, et al Absorption of circularly polarized X rays in iron. Phys Rev Lett. 1987;58:737.
[15] Thole BT, Carra P, Sette F, et al X-ray circular dichroism as a probe of orbital magnetization. Phys Rev Lett. 1992;68:1943.
[16] Carra P, Thole BT, Altarelli M, et al X-ray circular dichroism and local magnetic fields. Phys Rev Lett. 1993;70:694.
[17] van der Laan G, Henderson CMB, Patrick RAD, et al Orbital polarization in NiFe$_2$O$_4$ measured by Ni-2p x-ray magnetic circular dichroism. Phys Rev B. 1999;59:4314.
[18] Chen CT, Iderza YU, Lin HJ, et al Experimental confirmation of the X-ray magnetic circular dichroism sum rules for iron and cobalt. Phys Rev Lett. 1995;75:152.
[19] Iwamitsu K, Aihara S, Shimamoto T, et al A wavelength modulation system for highly sensitive absorption spectroscopy. Rev Sci Instrum. 2012;83:073101.
[20] Tokuda S, Nagata K, Okada M. Simultaneous estimation of noise variance and number of peaks in Bayesian spectral deconvolution. J Phys Soc Jpn. 2017;86:024001.
[21] Richter MC, Mariot JM, Heckmann O, et al NiFe$_2$O$_4$ and Fe$_3$O$_4$ studied by XMCD and resonant photoemission. Eur Phys J Special Topics. 2009;169:175.
[22] de Groot F, Kotani A. Core level spectroscopy of solids. Florida (USA): CRC Press; 2008. Chapter 4.4, 7.2, Advances in Condensed Matter Science Book 6; p. 133–144, 288–299.
[23] Hukushima K, Nemoto K. Exchange Monte Carlo method and application to spin glass simulations. J Phys Soc Jpn. 1996;65:1604.
[24] Metropolis N, Rosenbluth AW, Rosenbluth MN, et al Equation of state calculations by fast computing machines. J Chem Phys. 1953;21:1087.
[25] Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 1995;82:711.
[26] Andrieu C, Freitas ND, Doucet A, et al An introduction to MCMC for machine learning. Machine Learning. 2003;50:5.
[27] Wu R, Freeman AJ. Limitation of the magnetic-circular-dichroism spin sum rule for transition metals and importance of the magnetic dipole term. Phys Rev Lett. 1994;73:1994.
[28] Nakamura T, Suzuki M. Recent progress of the x-ray magnetic circular dichroism technique for element-specific magnetic analysis. J Phys Soc Jpn. 2013;82:021006.
[29] Terada Y, Tanida H, Uruga T, et al High-resolution X-ray microprobe using a spatial filter and its application to micro-XAFS measurements. AIP Conf Proc. 2011;1365:172.
[30] Yokoyama Y, Uozumi T, Nagata K, et al Bayesian integration for hamiltonian parameters of x-ray photoemission and absorption spectroscopy. J Phys Soc Jpn. 2021;90:034703.