Bayesian Spectroscopy with a Replica Exchange Monte Carlo Method on an Excitonic Absorption Spectrum of a Cu$_2$O Thin Crystal

Kazunori Iwamitsu$^1$, Masato Okada$^{2,3}$, Ichiro Akai$^4$

$^1$Faculty of Science, Kumamoto University, 2-39-1 Kurokami Chuo-ku, Kumamoto 860-8555, Japan.
$^2$Department of Complexity Science and Engineering, The University of Tokyo, 5-1-5 Kashiwanoha Kashiwa, Chiba 277-8561, Japan.
$^3$Research and Service Division of Materials Data and Integrated System, National Institute for Material Science, 1-2-1 Sengen, Tsukuba, Ibaraki, 305-0047, Japan.
$^4$Institute of Pulsed Power Science, Kumamoto University, 2-39-1 Kurokami Chuo-ku, Kumamoto 860-8555, Japan.

E-mail: iwamitsu@kumamoto-u.ac.jp

Abstract. To study a bi-axial stress effect in a Cu$_2$O thin crystal sandwiched by paired MgO substrates, we have applied Bayesian spectroscopy to decompose an absorption spectrum with employing a replica exchange Monte Carlo (RXMC) method. The absorption spectrum includes broad bands of an inter-band transition and an exciton continuous state simultaneously with excitonic resonance transitions. However, it had been difficult to decompose correctly because of a bottleneck in optimal solution search efficiency. To relieve the bottleneck, we introduced the RXMC method and succeeded in decomposing all spectral components. Posterior probabilities of the band-gap energy and the exciton binding energy distribute at lower energy side than those in stress-free bulk crystals, which provides statistical evidence to show that the bi-axial stress remains in the Cu$_2$O thin crystal sandwiched by MgO substrates.

1. Introduction

The exciton in Cu$_2$O is an important system for studying excitonic Bose-Einstein condensation. In an experiment of Cu$_2$O bulk crystals performed by Yoshioka et al. [1], a uniaxial stress was imposed to form a trapping potential for saving cold excitons spatially. On the other hand, we prepared Cu$_2$O thin crystals, which were recrystallized epitaxially in the small gap between paired MgO substrates [2]. In such thin crystals, although a bi-axial stress will be imposed at their interfaces due to a lattice mismatch between Cu$_2$O and MgO, the energy shift of the yellow exciton state would be small because the lattice mismatch is sufficiently small. So that, we applied Bayesian inference [3] to evaluate the small energy shift of the yellow exciton state [4]. This absorption spectrum includes not only the absorption lines $X_n$ of the yellow-exciton series but also the absorption bands of an inter-band transition $B$ and an exciton continuous state $X_C$. As a result, search efficiency of an optimal solution through the Markov chain Monte Carlo (MCMC) samplings deteriorates, so we had been forced to fix the spectral parameters for $B$ and $X_C$ [4].
In this paper, in order to improve on the analysis performed in Ref. [4], here we employ a replica exchange Monte Carlo (RXMC) method [3, 5] to improve the search efficiency of the optimal solution and to obtain posterior probability distributions of all spectral parameters also including B and X_C bands.

2. Bayesian Spectroscopy and Spectral Functions

To study the bi-axial stress effect in Cu_2O thin crystals sandwiched by paired MgO substrates, the small changes of the band-gap energy \( E_g \) and the exciton binding energy \( R_y \) will be discussed, so that we have to decompose a complicated absorption spectrum. Bayesian spectroscopy [6], in which Bayesian inference is utilized to perform spectral decomposition, can be used to evaluate probability distributions of all spectral parameters and provide us a statistical evidence of their small changes \( \Delta E_g \) and \( \Delta R_y \). Let’s consider an analysis of a measured spectrum \( D := \{(x_1, y_1), \cdots, (x_N, y_N)\} \) explained by a physical model \( \{G(x_i; \theta)\} \), where \( \theta \) is a parameter set for all spectral components. Since random noises are superposed in the measured spectrum \( y_i \), an error function is defined as \( E(\theta) := \sum_{i=1}^N [y_i - G(x_i; \theta)]^2 / (2N) \). When the superposed noises are Gaussian random noises distributed with a standard deviation \( \sigma_{\text{noise}} := b^{-0.5} \), the conditional probability of \( D \) under \( \theta \) determined can be written as \( P(D|\theta, b) \propto \exp \left[ -b \frac{N}{2} E(\theta) \right] \) by using \( b \) and \( E(\theta) \) [4]. The Bayesian spectroscopy enables us to obtain posterior probability distributions \( P(\theta|D) \) of the spectral parameters \( \theta \). However, when we use a Metropolis method [8, 4] for MCMC sampling, it is necessary to know the noise intensity beforehand because \( \sigma_{\text{noise}} \) determines the distribution width of \( P(\theta|D) \). On the other hand, the RXMC method makes it possible to evaluate \( P(\theta|D) \) as well as the noise intensity only from the measured data \( D \) [7].

Here, we consider a joint probability \( P(\theta, D, b) \) among \( \theta \), \( D \) and \( b \). According to the causality, \( P(\theta, D, b) \) is expressed as \( P(D|\theta, b)P(\theta|b)P(b) \). By using a Bayesian theorem [9], \( P(\theta, D, b) \) can be reversely factorized as \( P(\theta, D, b) = P(\theta|D, b)P(D|b)P(b) \). Consequently, the conditional probability \( P(\theta|D, b) \) of \( \theta \) under \( D \) and \( b \) given can be written as Eq. (1).

\[
P(\theta|D, b) = \frac{P(D|\theta, b)P(\theta|b)}{P(D|b)} \propto \exp \left[ -b \frac{N}{2} E(\theta) \right] P(\theta|b),
\]

where the denominator for normalization is given by marginalization of the numerator and it is defined as a partition function \( Z(b) := \int d\theta P(D|\theta, b)P(\theta|b) \) on the analogy of statistical mechanics. Subsequently, a \( b \)-dependent Bayes free energy is defined as \( F(b) := -\ln Z(b) \) [7, 10].

To determine the noise intensity \( b = \sigma_{\text{noise}} \) without any assumptions, we have to estimate a \( \hat{b} \) that maximizes \( P(b|D) \). Although the conditional probability \( P(b|D) \) of \( b \) under \( D \) is expressed as \( P(D|b)P(b)P(D) \), \( P(b|D) \) is proportional to \( P(D|b) \), where, since the measurement has been already completed and any assumptions are not introduced for \( b \), the prior probabilities \( P(D) \) and \( P(b) \) become constant. Consequently, the \( \hat{b} \) can be estimated by minimization of \( F(b) \) as \( \hat{b} := \arg \max_b P(b|D) = \arg \min_b F(b) \) [7, 10]. By using \( \hat{b} \), the conditional probability of \( \theta \) can be sampled by the RXMC method on the basis of \( P(\theta|D, \hat{b}) \propto \exp \left[ -\frac{N}{2} E(\theta) \right] P(\theta|b) \).

Open circles in Fig. 1 show the measured absorption spectrum in a Cu_2O thin crystal, and it includes spectral components of excitonic series \( X_n \), an excitonic continuous band \( X_C \) and an inter-band transition \( B \). For \( X_C \) and \( B \), we used spectral functions for direct-gap and dipole-forbidden transition described in [11, 4] with considering the respective inhomogeneous broadening factors \( \gamma_{X_C} \) and \( \gamma_B \). For \( X_n \), Voigt functions based on an asymmetric Lorentzian function [12, 4] were employed, in which the same inhomogeneous broadening factor \( \gamma_X \) was introduced. Asymmetric factors \( A_2 \) and \( A_3 \) were considered only for \( X_2 \) and \( X_3 \), respectively. Although a common homogeneous broadening factor \( \Gamma_X \) was used for \( X_n \), a different homogeneous broadening factor \( \Gamma_2 \) was required for \( X_2 \). For \( X_n \), the transition energy...
Figure 1. (a) Absorption spectra of the measured, reproduced and the respective spectral components. (b) $b$-dependent Bayes free energy. (c) Posterior probability distributions of energy shifts for $E_g$ and $R_y$.

$E_n$ and absorption intensity $f_n$ were parameterized as Eq. (2) with $E_g$, $R_y$, $f_0$ and the quantum number $n$ [13].

$$E_n = E_g - \frac{R_y}{n^2} \quad (\text{for } n = 2, \cdots, 20), \quad f_n = f_0 \left( \frac{n^2 - 1}{n^3} \right) \quad (\text{for } n = 2, 5, \cdots, 20). \quad (2)$$

As exceptions for $f_n$ in Eq. (2), independent absorption intensities $f_3$ and $f_4$ were required as described in [4]. As the result, the parameter set $\Theta$ for $G(x; \Theta)$ has fourteen spectral parameters: $\Theta = \{E_g, R_y, f_0, f_3, f_4, \Gamma_2, \Gamma_X, A_2, A_3, \gamma_X, f_{X_C}, \gamma_{X_C}, f_B, \gamma_B\}$.

3. Results and Discussion

Figure 1(b) shows the variation of Bayes free energy as a function of $b$. $F(b)$ minimizes at $b$ as marked by a downward arrow, and this $b$ value corresponds to the standard deviation of random noise distribution ($\hat{\sigma}_{\text{noise}} = b^{-0.5}$) being $\hat{\sigma}_{\text{noise}} = 2.5 \times 10^{-8}$ in optical density (O.D.).

Table 1. $\Theta$: mean values of $\Theta$ weighted by $P(\Theta|D, \hat{b})$, $\sigma_\Theta$: distribution widths of $P(\Theta|D, \hat{b})$ shown by the standard deviation. †: $f_{X_C}, \gamma_{X_C}, f_B$ and $\gamma_B$ were fixed in [4].

| $\Theta \pm \sigma_\Theta$ | $E_g \pm \sigma_{E_g}$ (meV) | $R_y \pm \sigma_{R_y}$ (meV) | $f_{0} \pm \sigma_{f_0}$ $\times 10^{-8}$ | $f_{3} \pm \sigma_{f_3}$ $\times 10^{-8}$ | $f_{4} \pm \sigma_{f_4}$ $\times 10^{-8}$ |
|-----------------------------|-------------------------------|-------------------------------|---------------------------------|---------------------------------|---------------------------------|
| RXMC                        | 2171.36 ± 0.01                | 93.3 ± 0.1                    | 8.9 ± 0.1                       | 0.385 ± 0.009                   | 0.21 ± 0.01                     |
| Ref. [4]                    | 2171.392 ± 0.003              | 93.6 ± 0.1                    | 8.8 ± 0.2                       | 0.36 ± 0.01                     | 0.18 ± 0.01                     |

| $\Theta \pm \sigma_\Theta$ | $T_2 \pm \sigma_{T_2}$ (meV) | $T_X \pm \sigma_{T_X}$ (meV) | $A_2 \pm \sigma_{A_2}$          | $A_3 \pm \sigma_{A_3}$          | $\gamma_X \pm \sigma_{\gamma_X}$ (meV) |
|-----------------------------|-------------------------------|-------------------------------|---------------------------------|---------------------------------|---------------------------------|
| RXMC                        | 2.35 ± 0.07                   | 1.69 ± 0.08                   | -0.197 ± 0.008                 | -0.34 ± 0.02                   | 1.2 ± 0.1                       |
| Ref. [4]                    | 2.4 ± 0.1                     | 1.6 ± 0.2                     | -0.201 ± 0.008                 | -0.34 ± 0.02                   | 1.2 ± 0.2                       |

| $\Theta \pm \sigma_\Theta$ | $f_{X_C} \pm \sigma_{f_{X_C}}$ $\times 10^{-8}$ | $\gamma_{X_C} \pm \sigma_{\gamma_{X_C}}$ (meV) | $f_B \pm \sigma_{f_B}$ $\times 10^{-4}$ | $\gamma_B \pm \sigma_{\gamma_B}$ (meV) |
|-----------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| RXMC                        | 0.97 ± 0.01                     | 13.9 ± 0.1                     | 8.79 ± 0.09                    | 15.06 ± 0.05                   |
| Ref. [4]                    | 0.95†                          | 14†                            | 9.0†                           | 15†                            |
By using $\hat{b}$, posterior probability distributions $P(\theta|D, \hat{b})$ for all spectral parameters were obtained through RXMC samplings. In Table 1, mean values $\overline{\theta}$ and posterior probability distribution widths $\sigma_\theta$ of $\theta$ are summarized, where $\overline{\theta}$ are weighted means by $P(\theta|D, \hat{b})$ and $\sigma_\theta$ are indicated in standard deviations of $P(\theta|D, \hat{b})$ distributions. The first lines of the respective groups are the results by the RXMC method. The second lines in upper two groups are the results by the Metropolis method [4], and the second lines in the bottom group is the fixed values used in [4]. The optimal solutions $\overline{\theta}$ obtained by the RXMC method coincide approximately with the previous result in [4], in which the spectral parameters for $X_C$ and $B$ were fixed to be assumed ones. In contrast, it became possible for the first time to obtain optimal solutions $\overline{\theta}$ of fourteen spectral parameters by using RXMC method, although it was difficult with the Metropolis method [4]. By using $\overline{\theta}$, a reproduced spectrum and the respective spectral components are displayed in Fig. 1(a), and the reproduced one can explain well the measured spectrum.

In Cu$_2$O thin crystals sandwiched by paired MgO substrates, we have offered a statistical evidence of that $E_g$ and $R_y$ become smaller than those in stress-free Cu$_2$O crystals in [4]. This result is completely confirmed as seen in Fig. 1(c), which shows posterior probability distributions of their energy shifts $\Delta E_g$ and $\Delta R_y$ from bulk crystals [14]. These $P(\theta|D, \hat{b})$ have narrow distribution widths and appear in the negative shift region, and have no significant distribution at the point where the shift amount is zero.

4. Summary

By using RXMC method, Bayesian spectroscopy was applied to decompose the absorption spectrum for study of the bi-axial stress effect in Cu$_2$O thin crystals sandwiched by paired MgO substrates. Since the spectrum includes the absorption bands $B$ and $X_C$ as well as the exciton resonances $X_n$, there are fourteen spectral parameters by considering the inhomogeneous broadening factors. Although such spectral decomposition is difficult for the Metropolis method, we succeeded by the RXMC method, and obtained the posterior probability distributions of all spectral parameters. The measured spectrum was reproduced well by the optimal solution. On the basis of these probability distributions, we confirmed that $E_g$ and $R_y$ in our Cu$_2$O thin crystals decreases evidently from those in the stress-free crystals.

Acknowledgments

This work was supported by KAKENHI Grants-in-Aid (Nos. 16H01552, 16H04002, and 16K13824) and Grants-in-Aid for Scientific Research on Innovative Areas (JP25120009) from the Japan Society for the Promotion of Science (JSPS), and CREST (JPMJCR1761) from the Japan Science and Technology Agency (JST).

References

[1] Yoshioka K, Chae E and Kuwata-Gonokami M 2011 Nature Commun. 2 328
[2] Iwamitsu K, Aihara S, Ota A, Ichikawa F, Shimamoto T and Akai I 2014 J. Phys. Soc. Jpn. 83 124714
[3] Nagata K, Sugita S and Okada M 2012 Neural Networks 28 82
[4] Iwamitsu K, Aihara S, Okada M and Akai I 2016 J. Phys. Soc. Jpn. 85 094716
[5] Hukushima K and Nemoto K, 1996 J. Phys. Soc. Jpn. 65 1604
[6] Akai I, Iwamitsu K and Okada M, 2018 J. Phys.: Conf. Ser. 1036 012022
[7] Tokuwa S, Nagata K and Okada M 2017 J. Phys. Soc. Jpn. 86 024001
[8] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 J. Chem. Phys. 21 1087
[9] Bayes T and Price R 1763 Phil. Trans. Roy. Soc. 53 370
[10] Iwamitsu K, Furukawa Y, Nakayama M, Okada M and Akai I 2018 J. Lumin. 197 18
[11] Sturge M D 1987 in Excitons ed. Rashba E I and Sturge M D (Elsevier, Amsterdam)
[12] Toyozawa Y 1964 J. Phys. Chem. Solids 25 59
[13] Elliott R J 1957 Phys. Rev. 108 1384
[14] Gross E F 1956 Nuovo Cimento Suppl. 3 672