Spectrum adapted expectation-conditional maximization algorithm for extending high-throughput peak separation method in XPS analysis

Tarojiro Matsumura, Naoka Nagamura, Shotaro Akaho, Kenji Nagata, and Yasunobu Ando

ABSTRACT
We introduced the spectrum-adapted expectation-conditional maximization (ECM) algorithm to improve the efficiency of the peak fitting of spectral data by various fitting models. The spectrum-adapted ECM algorithm can perform the peak fitting by using the Pseudo–Voigt mixture model and Doniach–Sunjić–Gauss mixture model which are generally used for the peak fitting in the X-ray photoelectron spectroscopy. Analyses of the synthetic and experimental spectral data showed that the proposed method quickly completed the calculation and estimated well-fitted curves to spectral data. This result suggests that the spectrum adapted ECM algorithm efficiently perform the peak fitting for large number of spectral data sets.

KEYWORDS
Peak fitting; X-ray photoelectron spectroscopy; expectation-conditional maximization algorithm; high-throughput analysis

1. Introduction
Analysis of spectral data is an essential part in the spectroscopy measurement for investigating electronic properties of materials and devices [1–4]. As the number of obtained spectrum data has been increasing by the improvement of the spectroscopy device, it is a serious problem in the efficiency of the spectroscopy measurement to analyze a large number of spectral data set. In the case of X-ray photoelectron spectroscopy (XPS), for example, peak assignment of core-level spectra in compounds strongly resorts to the manual trial and error, because suitable parameters of the fitting curve are explored according to previous reports and researchers’ experiences. As such procedure surely affects the efficiency of the analysis, the conventional analysis procedure is a bottleneck in the cycle of new material and device development. Therefore, the development of the analysis method is indispensable for improving the time-consuming procedure of peak fitting and for accelerating the cycle.

Recent works have proposed an analysis procedure that employs techniques developed in the field of informatics. For example, Nagata et al. [5] introduced a spectral deconvolution method based on Bayesian estimation with the exchange Monte Carlo method and demonstrated the estimation of an appropriate...
number of peaks while avoiding parameter solutions trapped into local minima. Murata et al. [6] also applied the Nagata’s framework to the time-series spectral data. Shiga et al. [7] proposed a new non-negative matrix factorization (NMF) technique to analyze spectral imaging data. This technique has support to resolve problems associated with previous NMF techniques, such as the calculation not always converging, and the number of separated peaks being specified by the manual trial and error. Very recently, Akai et al. [8] have proposed Bayesian spectroscopy for spectral analyses of optical spectra and for normal mode analyses of coherent phonon signals in solid–state materials. This Bayesian spectroscopy demonstrated high accuracy estimation of material parameters than the least-square and Fourier transform methods. Shinotsuka et al. [9] also have developed a peak fitting procedure based on the Bayesian information criterion to estimate the model parameter and appropriate number of peaks in the spectral decomposition. The methods for preprocessing process to the background subtraction have been developed by introducing some hyperparameters into the model [10–12], and these methods were demonstrated to some spectral data, such as infrared, Raman spectra, chromatography and nuclear magnetic resonance. Very recently, Ozaki et al. [13] successfully demonstrated the blackbox optimization approach [14] to the Rietveld method in order to automate the trial-and-error process for searching the hyperparameters. Moreover, a machine learning technique for material structural data have been actively studied in the classification of crystal structure by using the deep neural network e.g. [15,16].

For the high-throughput analysis of spectral data, Matsumura et al. [17] proposed the spectrum-adapted EM algorithm that focuses on performing rapid and auto peak fitting. In the spectrum adaptive EM algorithm, the peak fitting is conducted by using the intensity of spectrum as a weight for each measurement energy step. This method enables the reduction of the computational cost for optimizing the parameters of the fitting model. However, the spectrum adapted EM algorithm can only be applied to Gaussian mixture model (GMM), it is thus difficult to be applied in the analysis of various spectral data. In order to perform the high-throughput analysis method for various spectral data, it is necessary to develop a method that can extend the available fitting models such as, Pseudo–Voigt distribution and Doniach–Šunjić distribution.

In this study, we extended the spectrum-adapted EM algorithm so that it can use the Pseudo–Voigt mixture model (PVMM) and Doniach–Šunjić–Gauss mixture model (DSGMM) as the fitting model in the peak separation. The extended method is derived by applying the expectation–conditional maximization (ECM) algorithm [18]. The ECM algorithm is one of the generalized EM algorithms with the iterative calculation between the expectation (E) step and conditional maximization (CM) step. We demonstrated the application of the proposed method to the synthetic and experimental spectral data from XPS analysis.

2. The fitting models

We introduced the PVMM and DSGMM as the fitting models for the spectrum data. Each fitting model for measurement energy steps \(x_n\) is constructed of the Pseudo–Voigt distribution (Pseudo Voigt \((x_n|\mu, \sigma, \eta)\)) [19] and the Doniach–Šunjić distribution (Doniach Šunjić \((x_n|\mu, \gamma, \alpha)\)) [20] as follows:

\[
\text{PseudoVoigt}(x_n|\mu, \sigma, \eta) = \eta \text{Lorentz}(x_n|\mu, \gamma) + (1 - \eta)\text{Normal}(x_n|\mu, \sigma)
\]

(1)

and

\[
\text{Doniach Šunjić}(x_n|\mu, \gamma, \alpha) = \frac{\Gamma(1 - \alpha)}{\left\{ (x_n - \mu)^2 + \gamma^2 \right\}^{(1-\alpha)/2}} \cos \left( \frac{\pi \alpha}{2} + (1 - \alpha) \arctan \left( \frac{x_n - \mu}{\gamma} \right) \right),
\]

(2)

where

\[
\text{Lorentz}(x_n|\mu, \gamma) = \frac{\gamma}{\pi (x_n - \mu)^2 + \gamma^2}
\]

(3)

and

\[
\text{Normal}(x_n|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\frac{1}{2} \left( \frac{x_n - \mu}{\sigma} \right)^2 \right).
\]

(4)

Here, \(\mu\) is the location parameter specifying the peak of the spectrum, \(\gamma (\gamma > 0)\) is the scale parameter specifying the half-width at half-maximum (HWHM), \(\sigma\) is standard deviation \((\sigma > 0)\), \(\eta\) is mixing parameter \((0 \leq \eta \leq 1)\) between Lorentz \((x_n|\mu, \gamma)\) and Normal \((x_n|\mu, \sigma)\), and \(\alpha\) is the degree of the asymmetry of peak shape \((0 \leq \alpha \leq 1)\). The Pseudo–Voigt distribution is generally used as a fitting model to spectral data because it is empirically considered to fit well e.g. [21]. Also, some fitting models based on the Pseudo–Voigt distribution have been proposed e.g. [22,23]. The Doniach–Šunjić distribution is a peak fitting model used for the spectral data with asymmetric peak shapes. Additionally, spectral data is generally considered to have a gaussian broadening, namely the energy resolution depends on the temperature of the sample, the performance of photoelectron analyzers, spectrometers, light sources, and so on. Thus, we define the Doniach–Šunjić–Gauss distribution by using mixing parameter \((\eta, 0 \leq \eta \leq 1)\) to approximate such line shape by combining a Doniach–Šunjić distribution with a Gaussian distribution as follows:
Doniach–Sunjić_Gauss\( (x_n|\mu, \sigma, \eta, \alpha, \eta) \)
\[ = \eta \text{Doniach–Sunjić}_{\mu, \sigma, \alpha, \eta} + (1 - \eta) \text{Normal}_{\mu, \sigma} \]

(5)

Such line shape function was introduced by Joyce et al. [24], and it seems to be suitable for practical uses in the spectral data such as C 1s spectra e.g [25]. Here, the scale parameter, \( \eta \), specifies the half-width at half-maximum (HWHM) of the distribution. Also, the HWHM can be represented by using a Gaussian distribution to be \( \sqrt{2\ln 2}\sigma \). Hence, as the relationship between \( \sigma \) and \( y \) is represented as follows \( y = \sqrt{2\ln 2}\sigma \), in a formalism of Pseudo-Voigt distribution, \( y_k \) can be represented by using \( \sigma_k \) as \( y_k = \sqrt{2\ln 2}\sigma_k \). Using this relation, Equations (1) and (2) are rewritten as follows:

Pseudo Voigt\( (x_n|\mu, \sigma, \eta) = \eta \text{ Lorentz}_{\mu, \sqrt{2\ln 2}\sigma} + (1 - \eta) \text{Normal}_{\mu, \sigma} \)

(6)

and

Doniach–Sunjić_Gauss\( (x_n|\mu, \sigma, \alpha, \eta) \)
\[ = \eta \text{ Doniach–Sunjić}_{\mu, \sqrt{2\ln 2}\sigma, \alpha, \eta} + (1 - \eta) \text{Normal}_{\mu, \sigma} \]

(7)

respectively. As the number of parameters is reduced, we use Equations (6) and (7) to be the component of PVMM and DSGMM, respectively.

In the peak fitting, as these distributions have a broad tail and the observed measurement energy steps \( (x_n) \) are finite, each distribution (Equations 6 and 7) is expressed in a truncated distribution in the range of \( x_n \) as follows:

\[ \text{P}_{\text{PV}}(x_n|\mu, \sigma, \eta) = \frac{\text{Pseudo Voigt}(x_n|\mu, \sigma, \eta)}{\sum_{n=1}^{N} \text{Pseudo Voigt}(x_n|\mu, \sigma, \eta)} \]

(8)

and

\[ \text{P}_{\text{DSG}}(x_n|\mu, \sigma, \alpha, \eta) = \frac{\text{Doniach–Sunjić}_{\mu, \sigma, \alpha, \eta}}{\sum_{n=1}^{N} \text{Doniach–Sunjić}_{\mu, \sigma, \alpha, \eta}} \]

(9)

respectively. The PVMM and DSGMM are, respectively, defined as a linear superposition of \( k \)-th normalized Pseudo-Voigt distribution \( \left( \text{P}_{\text{PV}}(x_n|\mu_k, \sigma_k, \eta_k) \right) \) and normalized Doniach–Sunjić–Gauss distribution \( \left( \text{P}_{\text{DSG}}(x_n|\mu_k, \sigma_k, \alpha_k, \eta_k) \right) \) as follows:

\[ \text{PVMM}(x_n|\lambda_k, \mu_k, \sigma_k, \eta_k) = \sum_{k=1}^{K} \lambda_k \text{P}_{\text{PV}}(x_n|\mu_k, \sigma_k, \eta_k) \]

(10)

and

\[ \text{DSGMM}(x_n|\lambda_k, \mu_k, \sigma_k, \alpha_k, \eta_k) = \sum_{k=1}^{K} \lambda_k \text{P}_{\text{DSG}}(x_n|\mu_k, \sigma_k, \alpha_k, \eta_k), \]

(11)

where, \( K \) is the number of components \( (K > 0) \), and \( \lambda_k \) is the mixing coefficient of \( k \)-th component of the peaks \( (0 \leq k \leq 1) \). PVMM becomes GMM and Lorentz mixture model (LMM) in \( \eta_k = 0 \) and 1, respectively. Moreover, DSGMM becomes GMM and Doniach–Sunjić mixture model (DSMM) in \( \eta_k = 0 \) and 1, respectively. In the peak fitting based on the spectrum adapted EM algorithm, it is difficult to analytically derive the updating rule of the parameters in the M-step for PVMM and DSGMM. This difficulty also occurs in the case of GMM at \( \eta_k = 0 \) in PVMM and DSGMM because we use a truncated distribution to be the component of mixture model. In this study, we introduce the optimization procedure based on the spectrum adapted ECM algorithm to optimize the parameters of the fitting models (Equations 10 and 11)

3. Spectrum-adapted ECM algorithms

The ECM algorithm is one of the generalized EM algorithms based on maximum likelihood estimation with the iterative calculation between the expectation (E)-step, and conditional maximization (CM)-step, which conducts several conditional optimizations within each CM-step [18]. When the parameters are separated into some subsets and the maximization (M)-step is divided into multiple CM-steps, each CM-step conducts the optimization to a subset of the parameters with the other subsets are fixed [26]. This procedure addresses the problem of an intractable calculation in the M-step. It is difficult to conduct simultaneous optimization of parameters at M-step in the spectrum-adapted EM algorithm for general fitting models such PVMM and DSGMM. This difficulty is improved by individually optimizing the parameters by introducing CM-steps instead of M-step.

We apply the ECM algorithm to the peak fitting by estimating the parameters of the mixture model. In order to apply the ECM algorithm, we consider the intensity of the spectral data as a weight for each measurement step of energy [17]. When the spectral data consist of \( N \) measurement steps of energy \( (x = \{ x_1, \cdots, x_m, \cdots, x_N \}) \) corresponding to the intensity \( (w = \{ w_1, \cdots, w_m, \cdots, w_N \}) \), the spectral data \( \text{Dat} \) are represented in two dimensions:

\[ \text{Dat} = \left\{ x_1 \cdots x_N \right\} \]

(12)

As conventional EM algorithm requires to convert \( \text{Dat} \) to be one-dimensional to apply the peaks fitting,
the size of data becomes the sum of the intensity. This size of the converted data is significantly larger than the total number of measurement steps (N), and the calculation cost greatly increases. Spectrum adapted EM algorithm solved this disadvantage by using the intensity (w) as a weight for each measurement step (x) [17]. We apply this framework to the ECM algorithm for the peak fitting by PVMM and DSGMM. The flowcharts of spectrum-adapted ECM algorithm for PVMM and DSGMM were represented in Figure 1.

3.1. Spectrum-adapted ECM algorithm in PVMM

At first of the calculation, the parameters of the PVMM (λ_k, σ_k and η_k) are initialized, and the initial value of weighted log-likelihood is calculated as follows.

\[ L_{PVMM} = \sum_{n=1}^{N} w_n \ln \left\{ \sum_{k=1}^{K} \lambda_k P_{PV}(x_n|\mu_k, \sigma_k, \eta_k) \right\} \]  (13)

In the E-step, the responsibilities (y_{zk}) are calculated by using current parameter values (λ_k^{old}, μ_k^{old}, σ_k^{old}, and η_k^{old}) as follows.

\[ y_{zk} = \frac{\lambda_k^{old} P_{PV}(x_n|\mu_k^{old}, \sigma_k^{old}, \eta_k^{old})}{\sum_{k=1}^{K} \lambda_k^{old} P_{PV}(x_n|\mu_k^{old}, \sigma_k^{old}, \eta_k^{old})} \]  (14)

where *y_{zk} is the posterior probability when the measurement steps (x) are observed (Bishop 2006). Moreover, z_{nk} is a latent variable associated with x_n. In PVMM, x_n is assumed to be generated from one of the Pseudo-Voigt components. z_{nk} represents the component that generated x_n; i.e. z_{nk} is equal to 1 when x_n is generated from k-th component otherwise z_{nk} is equal to 0. Theoretical details are described by Bishop [26] and McLachlan and Krishnan [27].

Then, in the CM-step for λ, the parameters of λ_k are updated from λ_k^{old} to λ_k^{new} by using current responsibilities (y_{zk}) and intensities (w) that correspond to the measurement steps (x) as follows:

\[ \lambda_k^{new} = \frac{N_{wy_{zk}}}{\sum_{n=1}^{N} w_n}, \]  (15)

where

\[ N_{wy_{zk}} = \sum_{n=1}^{N} w_n y_{zk} \]  (16)

In CM-step for μ, the parameters of μ_k are updated from μ_k^{old} to μ_k^{new}, μ_k^{new} are obtained by the maximization of the expectation of the complete-data log-likelihood with the weight of the intensity as follows

\[ Q_{wPVMM} = \sum_{n=1}^{N} \sum_{k=1}^{K} \left\{ w_n y_{zk} \ln \lambda_k + \ln P_{PV}(x_n|\mu_k, \sigma_k, \eta_k) \right\}. \]  (17)

Here, λ_k, σ_k and η_k are fixed to be λ_k^{new}, σ_k^{old} and η_k^{old}, respectively. This maximization is conducted by using the Brent’s method [28] in this study. In CM-step for σ, the parameters of σ_k are updated from σ_k^{old} to σ_k^{new} by maximizing the Equation (17) in which λ_k, μ_k and η_k are fixed to be λ_k^{new}, μ_k^{new} and η_k^{old}, respectively. In CM-step for η, the parameter of η_k are updated from η_k^{old} to η_k^{new} by maximizing the Equation (17) in which λ_k, μ_k and σ_k are fixed to be λ_k^{new}, μ_k^{new} and σ_k^{old}, respectively.

After the CM-steps, when the distance of the weighted log-likelihood values (Equation 13) after

![Flowcharts](attachment:flowchart.png)

Figure 1. Flowcharts of the spectrum adapted ECM algorithm for (a) PVMM and (b) DSGMM. Details are shown in the manuscript.
the update and that immediately before the update is more than $1 \times 10^{-8}$, the calculation is returned to the E-step. Otherwise, the calculation is determined to be converged and the parameters at that time are adopted to be the solution.

### 3.2. Spectrum-adapted ECM algorithm in DSGMM

The parameters of DSGMM ($\lambda_k, \mu_k, \sigma_k, \alpha_k$ and $\eta_k$) are estimated by the same procedure as that for PVMM. At first, the initial value of weighted log-likelihood is calculated by using initialized parameters as follows.

$$L_{\text{DSGMM}} = \sum_{n=1}^{N} \sum_{k=1}^{K} w_n \ln \left( \sum_{k=1}^{K} \lambda_k \mathcal{P}_{\text{DSG}}(x_n; \mu_k, \sigma_k, \alpha_k, \eta_k) \right)$$  \hspace{0.5cm} \text{(18)}

In the E-step, $y_{n,k}$ are calculated by using current parameter values ($\lambda_k^{old}, \mu_k^{old}, \sigma_k^{old}, \alpha_k^{old}$ and $\eta_k^{old}$) as follows.

$$y_{n,k} = \frac{\lambda_k^{old} \mathcal{P}_{\text{DSG}}(x_n; \mu_k^{old}, \sigma_k^{old}, \alpha_k^{old}, \eta_k^{old})}{\sum_{k=1}^{K} \lambda_k^{old} \mathcal{P}_{\text{DSG}}(x_n; \mu_k^{old}, \sigma_k^{old}, \alpha_k^{old}, \eta_k^{old})}$$  \hspace{0.5cm} \text{(19)}

In DSGMM, $x_n$ is assumed to be generated from one of the Doniach–Šunjić-Gauss components. $z_{nk}$ represents the component that generated $x_n$.

Then, in CM-step for $\lambda$, the parameters of $\lambda_k$ are updated from $\lambda_k^{old}$ to $\lambda_k^{new}$ by using current responsibilities and intensities ($w$) that correspond to the measurement steps of energy ($x$) as follows:

$$\lambda_k^{new} = \frac{\sum_{n=1}^{N} w_n y_{n,k}}{\sum_{n=1}^{N} w_n},$$  \hspace{0.5cm} \text{(20)}

where

$$N w_{y,n,k} = \sum_{n=1}^{N} w_n y_{n,k}.\hspace{0.5cm} \text{(21)}$$

In CM-step for $\mu$, the parameter of $\mu_k$ are updated from $\mu_k^{old}$ to $\mu_k^{new}$, $\mu_k^{new}$ are obtained by maximizing the following expectation of the complete-data log-likelihood with the weight of the intensity:

$$Q_{\mu \text{DSGMM}} = \sum_{n=1}^{N} \sum_{k=1}^{K} \left( w_n y_{n,k} \right) \ln \lambda_k + \ln \mathcal{P}_{\text{DSG}}(x_n; \mu_k, \sigma_k, \alpha_k, \eta_k) \right\}.$$  \hspace{0.5cm} \text{(22)}

Here, $\lambda_k, \mu_k, \sigma_k$ and $\eta_k$ are fixed to be $\lambda_k^{new}, \mu_k^{new}, \sigma_k^{old}$ and $\eta_k^{old}$, respectively. In CM-step for $\sigma$, the parameter of $\sigma_k$ are updated from $\sigma_k^{old}$ to $\sigma_k^{new}$ by maximizing the Equation (22) in which $\lambda_k, \mu_k, \alpha_k$ and $\eta_k$ are fixed to be $\lambda_k^{new}, \mu_k^{new}, \sigma_k^{old}$ and $\eta_k^{old}$, respectively. In CM-step for $\alpha$, the parameter of $\alpha_k$ are updated from $\alpha_k^{old}$ to $\alpha_k^{new}$ by maximizing the Equation (22) in which $\lambda_k, \mu_k, \sigma_k$ and $\eta_k$ are fixed to be $\lambda_k^{new}, \mu_k^{new}, \sigma_k^{new}$ and $\eta_k^{old}$, respectively. In CM-step for $\eta$, the parameter of $\eta_k$ are updated from $\eta_k^{old}$ to $\eta_k^{new}$ by maximizing the Equation (22) in which $\lambda_k, \mu_k, \sigma_k$ and $\alpha_k$ are fixed to be $\lambda_k^{new}, \mu_k^{new}, \sigma_k^{new}$ and $\alpha_k^{new}$, respectively.

When the distance of the weighted log-likelihood values (Equation 18) after the update and that immediately before the CM-steps is less than $1 \times 10^{-8}$, the calculation is determined to be converged and the parameters at that time are adopted to be the solution. Otherwise, the calculation is returned to the E-step.

### 4. Synthetic data analysis

#### 4.1. Generation of the synthetic spectral data

We demonstrate the peak fitting by spectrum-adapted ECM algorithm to the synthetic spectral data set. The synthetic spectral data are generated from three-component PVMM and DSGMM, respectively. The data is represented as follows

$$\text{Dat}(\text{spv}) = \left( \begin{array}{c} x_1^{(\text{spv})} \\ \vdots \\ x_N^{(\text{spv})} \end{array} \right) \hspace{0.5cm} \text{(23)}$$

and

$$\text{Dat}(\text{sdg}) = \left( \begin{array}{c} x_1^{(\text{sdg})} \\ \vdots \\ x_N^{(\text{sdg})} \end{array} \right), \hspace{0.5cm} \text{(24)}$$

where $w_n^{(\text{spv})}$ and $w_n^{(\text{sdg})}$ are the intensity of the synthetic spectral data corresponding to $x_n$ and spv and sdg mean the synthetic Pseudo–Voigt distribution and synthetic Doniach–Šunjić–Gauss distribution, respectively. In the XPS, as the intensity is obtained from the count data of the number of photoelectrons at each measurement energy step, we assume that $w_n^{(\text{spv})}$ and $w_n^{(\text{sdg})}$ are randomly generated from the Poisson distribution (Poisson$(X|\xi)$) as follows

$$\text{Poisson}(X|\xi) = \frac{\xi^X \exp(-\xi)}{X!}, \hspace{0.5cm} \text{(25)}$$

where $X$ is a discrete random variable that is natural numbers starting from 0 (0, 1, 2, \ldots). $\xi$ is the parameter representing the mean of the distribution. We defined $\xi^{(\text{spv})}$ and $\xi^{(\text{sdg})}$ for $w_n^{(\text{spv})}$ and $w_n^{(\text{sdg})}$, respectively. $\xi^{(\text{spv})}$ and $\xi^{(\text{sdg})}$ were calculated from PVMM and DSGMM as follows

$$\xi^{(\text{spv})} = \sum_{k=1}^{3} \lambda_k^{(\text{spv})} \mathcal{P}_{\text{PV}}(x_n; \mu_k^{(\text{spv}), \sigma_k^{(\text{spv})}, \eta_k^{(\text{spv})}) \times 10^d, \hspace{0.5cm} \text{(26)}$$

and

$$\xi^{(\text{sdg})} = \sum_{k=1}^{3} \lambda_k^{(\text{sdg})} \mathcal{P}_{\text{DSG}}(x_n; \mu_k^{(\text{sdg})}, \sigma_k^{(\text{sdg})}, \eta_k^{(\text{sdg})}) \times 10^d. \hspace{0.5cm} \text{(27)}$$

respectively. $d$ is a constant ($d \geq 0$) for the number of intensity related to a measurement time. In $\text{Dat}(\text{spv})$ and $\text{Dat}(\text{sdg})$, the steps ($x_n$) were collected from the range [0:100] in steps of 0.5, so that the total number of steps
was 201. The parameters for Dat\((spv)\) were \(\mu_{k}^{(spv)} =\{35, 50, 65\}\), \(\sigma_{k}^{(spv)} =\{3, 3, 3\}\), \(\eta_{k}^{(spv)} =\{0.4, 0.6, 0.1\}\), \(\lambda_{k}^{(spv)} =\{1/3, 1/3, 1/3\}\), respectively. The parameters for Dat\((sdp)\) were \(\mu_{k}^{(sdp)} =\{20, 50, 80\}\), \(\sigma_{k}^{(sdp)} =\{2, 1, 3\}\), \(\eta_{k}^{(sdp)} =\{0.1, 0.3, 0.1\}\), \(\lambda_{k}^{(sdp)} =\{0.4, 0.6, 0.1\}\). The synthetic spectrum data were generated in the case of \(d =\{3, 4, 5\}\) from Equations (26) and (27). Moreover, the noise-free synthetic data \((w_{n}^{(spv)}\) and \(w_{n}^{(sdp)}\)) are generated from PVMM and DSGMM as follows:

\[
\begin{align*}
    w_{n}^{(spv)} = \sum_{k=1}^{3} \lambda_{k}^{(spv)} \mathbf{P}_{PV}(x_{n} | \mu_{k}^{(spv)}, \sigma_{k}^{(spv)}, \eta_{k}^{(spv)}) \times 10^{7}
\end{align*}
\]

and

\[
\begin{align*}
    w_{n}^{(sdp)} = \sum_{k=1}^{3} \lambda_{k}^{(sdp)} \mathbf{P}_{DSG}(x_{n} | \mu_{k}^{(sdp)}, \sigma_{k}^{(sdp)}, \eta_{k}^{(sdp)}) \times 10^{7},
\end{align*}
\]
respectively.

At the calculation, the number of peaks \((K)\) of PVMM and DSGMM were \(K = 3\), and the initial values of the parameters \((\mu, \sigma, \eta, \alpha, \lambda)\) were randomly collected from the range \([20, 80], [1, 10], [0, 0.5], [0, 0.5], [0, 1]\) satisfying \(\sum_{k=1}^{K} \lambda_{k} = 1\), respectively (Tables 1 and 2). Calculation was conducted by using our own source code developed in R (http://cran.r-project.org/) [29]. R is an open-source programing language and software environment for the statistical analysis. Also the computer carrying out the calculations had an Intel(R) Core(TM) i7 CPU with four cores at 2.9 GHz with 16 GB memory.

### 4.2. Result of the synthetic data analysis

Figure 2 shows the fitting curves of PVMM to the synthetic data. The calculation time required about 4 seconds to satisfy the convergence criterion in each synthetic data (Figure 2). In this analysis, over 500 iterations were required to reach the convergence criterion (Figure S1). The parameters of fitting curve at \(d =\{4, 5\}\) and noise-free were estimated similar values to the true parameters (Table 3). In contrast, the synthetic data at \(d = 3\), values of estimated \(\eta\) were deviated from the true value, although \(\mu\) and \(\lambda\) were close to the true values. The parameters significantly fluctuated at the early stage during iterative calculation, and then the values of \(\mu\), \(\sigma\) and \(\lambda\) hardly changed when iterative calculation was reached about 100 iterations, whereas values of \(\eta\) were still unstable at those iterations (Figures S2–S5). Thus, values of \(\eta\) required more iterative calculations than other parameters to be stable.

Figure 3 shows the fitting curves of DSGMM to the synthetic data. The fitting curve of DSGMM successfully estimated the asymmetric shape of the synthetic data. In this analysis, the calculation time required less than 9 seconds to satisfy the convergence criterion, and over 400 iterations were required to reach the convergence criterion (Figure S6). Each parameter of the DSGMM fluctuated at the beginning of the iterative calculation, and then these values of \(\mu\), \(\sigma\) and \(\lambda\) hardly changed when iterative calculation is reached about 100–200 iterations, whereas values of \(\alpha\) and \(\eta\) were unstable over 200 iterations (Figures S7–S10). Thus, the values of \(\eta\) and \(\alpha\) required more iterative calculations than other parameters to be stable. Table 4 shows that the values of \(\mu\), \(\sigma\) and \(\lambda\) were estimated to be close values to the true values in each synthetic data. In contrast, the values of \(\alpha\) and \(\eta\) were deviated from the true values at \(d =\{3, 4\}\). This result suggests that the accuracy in the estimation of \(\alpha\) and \(\eta\) were affected by noise level.

In order to compare the proposed method with the conventional peak fitting procedure, we performed synthetic data analysis by an adaptive non-linear least-square algorithm based on the least-squares method [30,31]. In the analysis, synthetic data \((d = 4)\) and initial values were generated by the same procedure in the analysis by the proposed method. The peak fitting by the conventional method was terminated without estimating an appropriate parameter solution at various initial values because the loss function diverges to infinity. Thus, as the conventional methods became unstable depending on the initial value, it is necessary to search for the appropriate initial values by

---

**Table 1.** The initial values of the parameters by using PVMM via the spectrum adapted ECM algorithm.

| Parameter | PVMM noise free | d = 3 | d = 4 | d = 5 |
|-----------|----------------|------|------|------|
| \(\mu_1\) | 31.59 | 32.02 | 22.19 | 31.40 |
| \(\mu_2\) | 49.55 | 57.10 | 39.24 | 32.70 |
| \(\mu_3\) | 58.26 | 79.11 | 70.11 | 21.78 |
| \(\sigma_1\) | 1.01 | 8.73 | 1.98 | 4.12 |
| \(\sigma_2\) | 2.28 | 3.33 | 7.34 | 6.78 |
| \(\sigma_3\) | 6.44 | 1.65 | 2.05 | 7.40 |
| \(\eta_1\) | 0.48 | 0.28 | 0.09 | 0.15 |
| \(\eta_2\) | 0.04 | 0.27 | 0.39 | 0.36 |
| \(\eta_3\) | 0.02 | 0.44 | 0.18 | 0.32 |
| \(\lambda_1\) | 0.37 | 0.39 | 0.62 | 0.15 |
| \(\lambda_2\) | 0.57 | 0.47 | 0.16 | 0.82 |
| \(\lambda_3\) | 0.05 | 0.15 | 0.22 | 0.02 |

**Table 2.** The initial values of the parameters by using DSGMM via the spectrum adapted ECM algorithm.

| Parameter | DSGMM noise free | d = 3 | d = 4 | d = 5 |
|-----------|----------------|------|------|------|
| \(\mu_1\) | 32.30 | 43.12 | 20.78 | 31.49 |
| \(\mu_2\) | 50.05 | 48.36 | 50.00 | 50.79 |
| \(\mu_3\) | 74.96 | 54.34 | 76.13 | 76.04 |
| \(\sigma_1\) | 6.76 | 5.91 | 1.31 | 5.97 |
| \(\sigma_2\) | 9.48 | 7.10 | 1.98 | 2.52 |
| \(\sigma_3\) | 7.74 | 8.51 | 8.93 | 3.12 |
| \(\eta_1\) | 0.46 | 0.38 | 0.34 | 0.29 |
| \(\eta_2\) | 0.26 | 0.40 | 0.36 | 0.45 |
| \(\eta_3\) | 0.25 | 0.08 | 0.41 | 0.13 |
| \(\eta_4\) | 0.46 | 0.34 | 0.45 | 0.17 |
| \(\eta_5\) | 0.15 | 0.31 | 0.28 | 0.48 |
| \(\lambda_1\) | 0.33 | 0.16 | 0.35 | 0.34 |
| \(\lambda_2\) | 0.30 | 0.02 | 0.16 | 0.22 |
| \(\lambda_3\) | 0.49 | 0.64 | 0.43 | 0.37 |
| \(\lambda_4\) | 0.28 | 0.34 | 0.41 | 0.41 |
trial-and-error in the analysis. Such trial-and-error is unsuitable in the high-throughput analysis of the large amount of spectral data. When the calculations were converged in conventional method, the parameter solution and calculation time were summarized in Table S1–S4 for proposed method and conventional method. The parameter values estimated by both methods did not show a significant difference. In contrast, the calculations by the conventional method were completed 5–18 and 11–24 times faster than the proposed method in PVMM and DSGMM, respectively. The differences in the calculation time would be because the proposed method requires more calculation steps than conventional method to conduct the conditional optimization of each parameter in CM-step.

5. Experimental data analysis

5.1. Collection of the experimental data

Proposed method was applied to the W 4f core-level spectrum and the C 1s core-level photoelectron spectrum. The W 4f core-level spectrum was taken

---

Table 3. The estimated parameters of PVMM via the spectrum adapted ECM algorithm.

| PVMM  | True value | noise free | d = 3 | d = 4 | d = 5 |
|-------|------------|------------|-------|-------|-------|
| μ₁    | 35         | 35.00      | 35.16 | 35.05 | 34.95 |
| μ₂    | 50         | 50.00      | 50.05 | 49.93 | 50.00 |
| μ₃    | 65         | 65.00      | 65.12 | 65.00 | 65.02 |
| σ₁    | 3          | 3.00       | 3.36  | 3.11  | 3.01  |
| σ₂    | 3          | 3.00       | 3.23  | 3.14  | 2.96  |
| σ₃    | 3          | 3.00       | 2.87  | 3.02  | 2.98  |
| η₁    | 0.4        | 0.40       | 0.25  | 0.37  | 0.38  |
| η₂    | 0.6        | 0.60       | 0.43  | 0.57  | 0.68  |
| η₃    | 0.1        | 0.10       | 0.27  | 0.11  | 0.06  |
| λ₁    | 1/3        | 0.33       | 0.30  | 0.33  | 0.33  |
| λ₂    | 1/3        | 0.33       | 0.35  | 0.34  | 0.34  |
| λ₃    | 1/3        | 0.33       | 0.35  | 0.33  | 0.33  |

---

Figure 2. Fitting curves of PVMM to the synthetic spectral data at $d = \{3, 4, 5\}$ and noise free. The open black circles are the spectral data. Red solid curve is the fitting curve, and black dotted curve is each Pseudo-Voigt distribution.

Figure 3. Fitting curves of DSGMM to the synthetic spectral data at $d = \{3, 4, 5\}$ and noise free. The open black circles are the spectral data. Red solid curve is the fitting curve, and black dashed curve is each Doniach–Sunjić-Gauss distribution.
by a mechanically exfoliated WSe₂ sheet in a 2D heterostructure tunnel field effect transistor [32]. Also, the C 1s core-level photoelectron spectrum was taken by a graphene sheet on a 90 nm SiO₂/p⁺-Si(100) substrate [33]. Both spectral data were acquired using a VG Scienta R3000 spectrometer in the scanning photoelectron microscopy system (3D nano-ESCA) installed at the University of Tokyo outstation beamline, BL07LSU at Spring-8. The energy resolution of the spectrometer was set to 300 meV. The photon energy of the synchrotron radiation beam used for measurements was 1000 eV. For convenience, these spectral data are called TFET (tunnel field effect transistor)-data and GFET (graphene field effect transistor)-data, respectively. The background of TFET-data and GFET-data were removed as a linear and Shirley background, respectively.

In the calculation, TFET-data and GFET-data were analyzed by PVMM and DSGMM, respectively. For TFET-data, the number of peaks were \( K = 3 \), and the initial values of the parameters \( (\mu, \sigma, \eta, \alpha, \lambda) \) were randomly collected from the range \([29.35:43.35, 1:10], [0:1], [0:1] \) satisfying \( \sum_{k=1}^{K} \lambda_k = 1 \), respectively. For GFET-data, the number of peaks were \( K = 2 \), and the initial values of the parameters \( \mu, \sigma, \alpha, \lambda \) were randomly collected from the range \([279.95:290.45, 1:10], [0:0.5], [0:1] \) satisfying \( \sum_{k=1}^{K} \lambda_k = 1 \), respectively. Also, the initial values of \( \eta_1 \) and \( \eta_2 \) were set 0 and 0.5, respectively. The parameter \( \eta_1 \) was fixed to be 0.

### 5.2. Result of experimental data analysis

Figure 4 shows the result of the peak fitting by using PVMM and DSGMM, and Tables 5 and 6 summarizes the initial and estimated values of the parameters. The fitting curves of PVMM and DSGMM were well-fitted to the spectral data. Calculation time of the peak fitting were 1.47 and 11.27 sec in TFET-data and GFET-data, respectively. Also, the values of \( L_{PVMM} \) and \( L_{DSGMM} \) were monotonically increased during the iterative calculation, and about 140 and 1000 iterations were required to satisfy the convergence criterion in the analysis of TFET-data and GFET-data, respectively (Figure S11). The parameters of PVMM and DSGMM fluctuated at the beginning of the iterative calculation, and then these values were stable after 60 and 500 iterations, respectively (Figures S12 and S13). The peaks (\( \mu_1, \mu_2 \) and \( \mu_3 \)) estimated from the TFET-data were 37.82 eV, 34.89 eV and 32.74 eV, respectively (Table 5). These peaks represent the contaminant, W 4f₁/2 and W 4f₃/₂ component, respectively. As the mixture ratio (\( \lambda_1, \lambda_2 \) and \( \lambda_3 \)) were 0.16, 0.37 and 0.48, respectively, the mixture ratio between W 4f₁/2 and W 4f₃/₂ component were 1:1.3. In the GFET-data, the peaks (\( \mu_1 \) and \( \mu_2 \)) were 285.23 eV and 284.28 eV, respectively (Table 6). These peaks represent the contaminant, and graphene \( sp^2 \) component, respectively. The component of contamination would arise lithographic processing [33,34]. The peak of graphene \( sp^2 \) component (284.28 eV) was very close to the value reported in Nagamura et al. [33] (284.23 eV). As the mixture ratio (\( \lambda_1 \) and \( \lambda_2 \)) were 0.63 and 0.38, respectively, the contaminant component occupies larger part of the data than that of the graphene \( sp^2 \) component. Such relatively large contaminant component was also observed in Nagamura et al [33]. The degree of the asymmetry of the graphene \( sp^2 \) component (\( \alpha_1 \)) was 0.29, whereas the asymmetric parameter of the graphene C 1s peak is generally assumed to be about 0.1 [35–37]. Thus, the value of \( \alpha_1 \) would be overestimated due to the influence of noise and contamination components in this analysis.

### 6. Discussion and implication

The spectrum-adapted ECM algorithm was introduced to extend the spectrum-adapted EM algorithm [17] and was successfully applied to synthetic and experimental data analyses. The advantage of the peak fitting method based on the EM algorithm is its stable and fast calculation [17]. As the parameters converged to a local optimal solution by monotonically increasing log-likelihood value in the iterative calculation (Figures S1, S6 and S11) and calculations were completed less than 12 sec regardless of the spectral data (Figures 2–4), the proposed method can stably and quickly conduct the peak fitting without the manual trial and error to obtain the parameter solution. As the proposed method requires non-negative intensity values to run the calculation appropriately, it can perform peak fitting whenever the spectral data satisfies such conditions regardless of the experimental method. Therefore, the application of this proposed method is not limited to XPS.

| Table 4. The estimated parameters of DSGMM via the spectrum adapted ECM algorithm. |
|--------------------------------------------|--------|--------|--------|--------|
| DSGMM | True value | noise free  | \( d = 3 \) | \( d = 4 \) | \( d = 5 \) |
| \( \mu_1 \) | 20 | 20.00 | 19.86 | 20.20 | 20.04 |
| \( \mu_2 \) | 50 | 50.00 | 48.82 | 49.54 | 50.01 |
| \( \mu_3 \) | 80 | 80.00 | 79.78 | 80.05 | 80.02 |
| \( \sigma_1 \) | 3 | 3.00 | 3.09 | 2.96 | 3.00 |
| \( \sigma_2 \) | 3 | 3.00 | 2.97 | 3.05 | 3.08 |
| \( \sigma_3 \) | 3 | 3.00 | 2.86 | 2.99 | 3.00 |
| \( \eta_1 \) | 0.1 | 0.10 | 0.10 | 0.15 | 0.10 |
| \( \eta_2 \) | 0.3 | 0.30 | 0.18 | 0.21 | 0.33 |
| \( \eta_3 \) | 0.1 | 0.10 | 1.54E-06 | 0.12 | 0.10 |
| \( \lambda_1 \) | 1/3 | 0.33 | 0.37 | 0.34 | 0.33 |
| \( \lambda_2 \) | 1/3 | 0.33 | 0.30 | 0.31 | 0.34 |
| \( \lambda_3 \) | 1/3 | 0.33 | 0.33 | 0.33 | 0.33 |
The spectrum-adapted ECM algorithm enables the conduction of high-throughput analysis by using various fitting models. The synthetic and experimental data analyses showed that the spectrum-adapted ECM algorithm estimated well-fitted curves using PVMM and DSGMM, both of which are general fitting models in XPS. As PVMM and DSGMM involve the component of Gauss and Lorentz distribution, and Doniach-Šunjić and Gauss distribution, respectively, the fitting curves can fit the spectral data more flexibly than the Gaussian mixture model in the spectrum adapted EM algorithm [17]. As the experimental data analysis suggests that the calculation can be completed in several to several tens of seconds for the fitting of 2–3 peaks, the calculation cost is sufficiently low in practical application. Moreover, as the method based on the EM algorithm has an initial value dependence [17], computational cost and accuracy of the analysis can be improved by appropriate initial values of the parameters instead of randomly generated initial values. An appropriate procedure of the initial value generation may improve the calculation cost and accuracy in the analysis of large number of spectral data sets. In

![Figure 4. Fitting curves for TFET-data (a) and GFET-data (b). The open black circles are the spectral data. Red solid curve is a fitting curve. Black dashed curve is each component of the fitting model and the background of the spectrum. The proposed method estimated a fitting curve that fitted the experimental data well.](image-url)

**Table 5.** Summary of the peak fitting for the spectra of TFET. Initial and estimated values of the parameters of PVMM are represented.

| Parameter | Initial | Estimated |
|-----------|---------|-----------|
| \(\mu_1\) | 32.84   | 37.82     |
| \(\mu_2\) | 41.87   | 34.89     |
| \(\mu_3\) | 39.70   | 32.74     |
| \(\sigma_1\) | 9.04    | 1.56      |
| \(\sigma_2\) | 3.61    | 0.43      |
| \(\sigma_3\) | 7.13    | 0.40      |
| \(\eta_1\) | 0.00    | 0.96      |
| \(\eta_2\) | 0.35    | 2.09E-07  |
| \(\eta_3\) | 0.62    | 3.74E-09  |
| \(\lambda_1\) | 0.62    | 0.16      |
| \(\lambda_2\) | 0.07    | 0.37      |
| \(\lambda_3\) | 0.31    | 0.48      |

**Table 6.** Summary of the peak fitting for the spectra of GFET. Initial and estimated values of the parameters of DSGMM are represented.

| Parameter | Initial | Estimated |
|-----------|---------|-----------|
| \(\mu_1\) | 290.2   | 285.23    |
| \(\mu_2\) | 284.15  | 284.28    |
| \(\sigma_1\) | 1.65    | 0.97      |
| \(\sigma_2\) | 8.67    | 0.48      |
| \(\sigma_3\) | 0.45    | 1.00      |
| \(\eta_1\) | 0.15    | 0.29      |
| \(\eta_2\) | 0.5     | 0.07      |
| \(\eta_3\) | 0.61    | 0.63      |
| \(\lambda_1\) | 0.39    | 0.38      |
| \(\lambda_2\) | 0.39    | 0.38      |
contrast, the conventional method by the adaptive non-linear least-square algorithm is unsuitable for the high-throughput analysis because the behaviour of calculation can be unstable depending on the initial values. As the ECM algorithm promises the convergence to the local optimum solution in the iterative calculations, such a behaviour does not occur. Therefore, the proposed method is robust in the peak fitting of large amounts of spectral data.

The proposed method has a limitation in the analysis of spectral data including large noise because the accuracy of estimation in some parameters, such as $\alpha$ and $\eta$, may be affected by the noise. Moreover, the number of components ($K$) and the distribution that consists the fitting model need to be selected before the calculation. In contrast, Bayesian spectrum deconvolution with the exchange Monte Carlo method performs highly accurate analysis and an appropriate fitting model is selected by calculating the marginal likelihood value [5]. However, this method requires high computational cost, and it is not easy to set a suitable prior distribution and inverse temperature by a non-expert. As such a setting is unnecessary in the proposed method, the peak fitting can be performed easily in the high-throughput analysis. Thus, these methods need to be selectively applied depending on the task.

Further development of the peak fitting method requires to develop an automated pre-processing method for the background subtraction. In the case of XPS, the peak assignment requires the proper subtraction of the background component in the pre-processing. This procedure is strongly relies on the manual trial and error and is inefficient when analyzing large amounts of spectral data sets. As the proposed method in this study extends the available fitting model, the background subtraction process will be automated in further extension of the method by defining a model that represents the background. Moreover, the blackbox optimization approach has been demonstrated in the selection of appropriate background function from some candidates during the optimization loop [13]. We consider that such optimization technique is useful for an automatic background subtraction. It will be possible to develop an advanced high-throughput peak fitting method by combing the proposed method and the BBO approach. Such advanced method will enable us to conduct the optimization of the fitting and background model automatically and quickly in the high throughput spectral data analysis.

7. Conclusions

Spectrum adapted ECM algorithm was proposed to use various fitting models, such as PVMM and DSGMM, in the high-throughput peak fitting. The synthetic and experimental data analyses suggest that the proposed method is applicable to peak fitting problems for the investigation of electronic properties of materials and devices at a sufficiently small computational cost. The advantage of the proposed method is that stable and quick peak fitting can be conducted to obtain the parameter solution without the manual trial and error. This advantage enables the high-throughput analysis of large amount of spectral data to be performed by using various fitting models.

Acknowledgements

This paper is based on results from a project (P16010) commissioned by the New Energy and Industrial Technology Development Organization (NEDO), JST CREST (JPMJCR1761), the “Materials Research by Information Integration” Initiative project and PRESTO (Grant number: JPMJPR17NB), commissioned by the Japan Science and Technology Agency (JST), and the Research Program for CORE lab (2016002) of “Dynamic Alliance for Open Innovation Bridging Human, Environment and Materials” of the “Network Joint Research Center for Materials and Devices”. We thank Prof. Kosuke Nagashio of the University of Tokyo and Prof. Hirokazu Fukidome of Tohoku University for measuring the presented data. The spectral datasets were obtained with the support of the University of Tokyo outstation beamline at SPring-8 (Proposal Numbers: 2012B7402, 2013A7443, 2013B7451, 2014B7472, and 2015A7482).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the New Energy and Industrial Technology Development Organization [JPNP16010].

ORCID

Tarojiro Matsumura http://orcid.org/0000-0001-9950-509X
Naoka Nagamura http://orcid.org/0000-0002-7697-8983
Shotaro Akaho http://orcid.org/0000-0002-4623-2718
Kenji Nagata http://orcid.org/0000-0001-9894-4461
Yasunobu Ando http://orcid.org/0000-0003-3702-034X

References

[1] Hufner S, editor. Very high resolution photoelectron spectroscopy. Berlin (GE): Springer; 2007.
[2] Hufner S. Photoelectron spectroscopy: principles and applications. 3rd ed. Berlin (GE): Springer; 2013.
[3] Suga S, Sekiyama A. Photoelectron spectroscopy. Berlin (GE): Springer; 2014.
[4] Attwood D, Sakdinawat A. X-rays and extreme ultraviolet radiation: principles and applications.
2nd ed. Cambridge (UK): Cambridge university press; 2017.

[5] Nagata K, Sugita S, Okada M. Bayesian spectral deconvolution with the exchange Monte Carlo method. Neural Netw. 2012;28:82–89.

[6] Murata S, Nagata K, Uemura M, et al Extraction of latent dynamical structure from time-series spectral data. J Phys Soc Jpn. 2016;85(10):140403.

[7] Shiga M, Tatsumi K, Muto S, et al Sparse modeling of EELS and EDX spectral imaging data by nonnegative matrix factorization. Ultramicroscopy. 2016;170:43–59.

[8] Akai I, Iwamitsu K, Okada M. Bayesian spectroscopy in solid-state photo-physics. J Phys Conf. 2018;1036(1):012022.

[9] Shinotsuka H, Nagata K, Yoshikawa H, et al Development of spectral decomposition based on Bayesian information criterion with estimation of confidence interval. Sci Technol Adv Mater. 2020;21(1):402–419.

[10] Mazet V, Carteret C, Brie D, et al Background removal from spectra by designing and minimising a non-quadratic cost function. Chemom Intell Lab Syst. 2005;76(2):121–133.

[11] Zhang ZM, Chen S, Liang YZ. Baseline correction using adaptive iteratively reweighted penalized least squares. Analyst. 2010;135(5):1138–1146.

[12] Ning X, Selesnick IW, Duval L. Chromatogram baseline estimation and denoising using sparsity (BEADS). Chemom Intell Lab Syst. 2014;139:156–167.

[13] Ozaki Y, Suzuki Y, Hawai T, et al Automated crystal structure analysis based on blackbox optimisation. NPJ Comput Mater. 2020;6(1):1–7.

[14] Ziletti A, Kumar D, Scheffler M, et al Insightful classification of crystal structures using deep learning. Nat Commun. 2018;9(1):1–10.

[15] Audet C, Hare W. Derivative-free and blackbox optimization. Berlin (GE): Springer; 2017.

[16] Park WB, Chung J, Jung J, et al Classification of crystal structure using a convolutional neural network. IUCrJ. 2017;4(4):486–494.

[17] Matsumura T, Nagamura N, Akaho S, et al Spectrum adapted expectation-maximization algorithm for high-throughput peak shift analysis. Sci Technol Adv Mater. 2019;20(1):733–745.

[18] Meng XL, Rubin DB. Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika. 1993;80(2):267–278.

[19] Wertheim GK, Butler MA, West KA, et al Determination of the Gaussian and Lorentzian content of experimental line shapes. Rev Sci Instrum. 1974;45(11):1369–1371.

[20] Doniach S, Sunjic M. Many-electron singularity in X-ray photoemission and X-ray line spectra from metals. J Phys C. 1970;3(2):285–291.

[21] Hesse R, Streubel P, Szargan R. Product or sum: comparative tests of Voigt, and product or sum of Gaussian and Lorentzian functions in the fitting of synthetic Voigt-based X-ray photoelectron spectra. Surf Interface Anal. 2007;39(5):381–391.

[22] Ida T, Ando M, Toraya H. Extended pseudo-Voigt function for approximating the Voigt profile. J Appl Crystallogr. 2000;33(6):1311–1316.

[23] Schmid M, Steinrück HP, Gottfried JM. A new asymmetric Pseudo-Voigt function for more efficient fitting of XPS lines. Surf Interface Anal. 2014;46(8):505–511.

[24] Joyce JJ, Del Giudice M, Weaver JH. Quantitative analysis of synchrotron radiation photoemission core level data. J Electron Spectrosc Relat Phenomena. 1989;49(1):31–45.

[25] Apostol NG, Lizzit D, Lungu GA, et al Resistance hysteresis correlated with synchrotron radiation surface studies in atomic sp² layers of carbon synthesized on ferroelectric (001) lead zirconate titanate in an ultrahigh vacuum. RSC Adv. 2020;10(3):1522–1534.

[26] Bishop CM. Pattern recognition and machine learning. New York (NY): Springer; 2006. Chapter 9, mixture models and EM; p. 423–455.

[27] McLachlan G, Krishnan T. Chapter 1, general introduction. In: The EM algorithm and its extensions. 2nd ed. New York (NY): Wiley; 2008. p. 1–37.

[28] Brent RP. Algorithms for minimization without derivatives. New Jersey (NJ): Prentice-Hall, Englewood Cliffs; 1973.

[29] An analysis package named “EMPeaks” is developing. This package will execute the proposed method together with the method of Matsumura et al. [17] on R and Python environment.

[30] Bates DM, Watts DG. Nonlinear regression analysis and its applications. New York (US): Wiley; 1988.

[31] Bates DM, Chambers JM. Chapter 10, nonlinear models. In: Hastie TJ, editor. Statistical models in S. Yonkers (NY): Wadsworth & Brooks/Cole; 1992. p. 421–453.

[32] Nakamura K, Nagamura N, Ueno K, et al All 2D heterostructure tunnel field-effect transistors: impact of band alignment and heterointerface quality. ACS Apple Mater Interfaces. 2020;12(42):51598–51606.

[33] Nakamura N, Fukidome H, Nagashio K, et al Influence of interface dipole layers on the performance of graphene field effect transistors. Carbon. 2019;152:680–687.

[34] Nakamura N, Horiba K, Toyota S, et al Direct observation of charge transfer region at interfaces in graphene devices. Appl Phys Lett. 2013;102(24):241604.

[35] Yang DQ, Sacher E. Carbon 1s X-ray photoemission line shape analysis of highly oriented pyrolytic graphite: the influence of structural damage on peak asymmetry. Langmuir. 2006;22(3):860–862.

[36] Smith M, Scudiero L, Espinal J, et al Improving the deconvolution and interpretation of XPS spectra from chars by ab initio calculations. Carbon. 2016;110:155–171.

[37] Peltekis N, Kumar S, McEvoy N, et al The effect of downstream plasma treatments on graphene surfaces. Carbon. 2012;50(2):395–403.