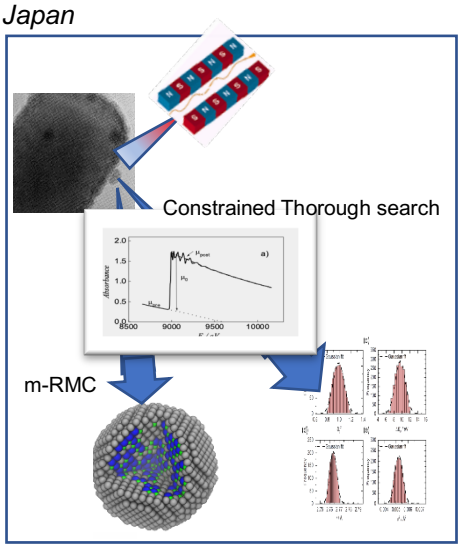
Accounts of
Materials & Surface
Research

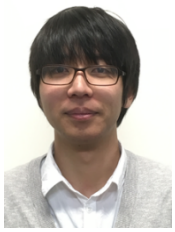
Problems in EXAFS analysis and its future prospects

Daiki Kido and Kiyotaka Asakura *
Institute for Catalysis, Hokkaido University
Kita 21-10 Sapporo, Hokkaido Japan
askr@cat.hokudai.ac.jp



Extended X-ray absorption fine structure (EXAFS) spectroscopy is a strong tool for investigating the local structure of materials, even without crystal-like long range order. X-rays have a large penetration power so that catalytic reactions or electrochemical processes can be followed in the presence of reactants under reaction conditions. However, this provides only one-dimensional information and a limited number of data points. This can make the EXAFS approach to the analysis of complex systems difficult. In this short review we describe the present status of EXAFS analysis and its problems. We also discuss future directions and new possibilities.

Keyword: EXAFS, Curve fitting, Reverse Monte Carlo, Constrained Thorough Search, Hamilton Method.



Daiki Kido received BS and MS degrees from Hokkaido University in 2015 and 2017, respectively. He is now in the process of submitting a doctoral thesis (2020) in the Department of Quantum Science and Engineering, Hokkaido University, under the supervision of Professor Kiyotaka Asakura. He received a best poster award at the Korea-Japan Symposium on Catalysis in 2017.

Dr. Kiyotaka Asakura FRSC, Professor at Hokkaido University, graduated from the Department of Chemistry, University of Tokyo, in 1981. He received a PhD in 1987, and was a Research Associate in the Department of Chemistry, University of Tokyo from 1984, and was promoted to Lecturer in 1992. He joined Prof. Gerhard Ertl's group at the Fritz-Haber Institute in 1993. In 1994, he became an associate professor at the Research Center for Spectrochemistry, University of Tokyo, and a full Professor at the Institute for Catalysis at Hokkaido University in 1999. He was a director of the Institute for Catalysis in Hokkaido University from 2014 to 2018. He was a chairman of the XAFS Society of Japan from 2009 to 2015. He is now an associate editor of the journal *Physical Chemistry Chemical Physics*, vice president of the International XAFS Society, and president of the Japanese Society of Synchrotron Radiation.

Problems in EXAFS analysis and its future prospects

Daiki Kido and Kiyotaka Asakura *
Institute for Catalysis, Hokkaido University
Kita 21-10 Sapporo, Hokkaido Japan

1. Introduction

Extended X-ray absorption fine structure (EXAFS) spectroscopy is a powerful method for analyzing the structure of materials with no long range order. [1,2] Figure 1 shows an X-ray spectrum near the Pt L_3 edge. The absorption coefficient suddenly increases at the binding energy of core electrons (in this case L_3 or $2p_{3/2}$ electrons), which is termed the absorption edge. At this energy, core electrons are excited to photoelectrons and leave the atom in a spherical wave. The modulations of the absorption coefficient near and in the higher energy region of the absorption edge are called X-ray absorption fine structure (XAFS). The XAFS is divided into two regions. One is the X-ray absorption near-edge structure (XANES) that arises from transitions to unoccupied bound states or from multiple scattering of photoelectrons by the surrounding atoms. EXAFS appears at an energy 50 eV higher energy than the absorption edge. The EXAFS is explained by single scattering of photoelectrons by surrounding atoms due to interference between the outgoing photoelectrons and the electrons scattered by the surrounding atoms as shown in inset of Fig.1 where the outgoing photoelectron and scattered electrons interfere at the x-ray absorbing atom.[3] The EXAFS oscillations are expressed by the following equation.[4]

$$\chi(k) = \sum \mathrm{C}N_j F_j(k)\exp\left(-2r_j/\lambda_j\right) \exp\left(-2{\sigma_j}^2k^2\right)\sin\left(2kr_j + \phi_j(k)\right)/kr_j{}^2 \quad (1)$$

where $CN_j, r_j,\ \lambda_j\ and\ {\sigma_j}^2$ are the coordination number, the bond distance, the inelastic mean free path and the Debye-Waller factor for the j-th surrounding atom, respectively. $F_j(k)$ and $\phi_j(k)$ are the backscattering amplitude and the phase shift functions for the j-th atom, respectively. k is the wavenumber of the photoelectron given by

$$k = \sqrt{2m/\hbar^2\ (h\upsilon - E_0)}, \quad (2)$$

where m is the mass of an electron and $\hbar = h/2\pi$. $h\upsilon$ and E_0 are the X-ray photon energy and the binding energy of core electron or photoelectron kinetic energy zero point, respectively. E_0 can be selected to be equal to the absorption edge. However, it should be optimized because the wavenumber in the theoretically calculated result is different from the experimental wavenumber calculated from Eq. (2). After $F_j(k)$ and $\phi_j(k)$ are determined either theoretically or empirically, $CN_j, r_j, {\sigma_j}^2$ and E_0 are then optimized by curve fitting to give the structure.

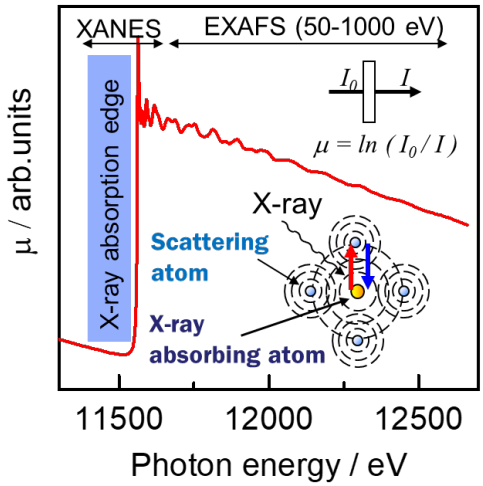


Figure 1. Pt L_3-edge EXAFS. The abrupt increase at 11560 eV is referred to as the absorption edge. The edge peak is called a white line. The modulation of the absorption coefficient is referred to as the XAFS. The higher energy region corresponds to the EXAFS. Inset is the shematic explanation for the mechanism of EXAFS .

Since EXAFS is a type of X-ray absorption spectroscopy, it can be applied to non-crystalline systems.[1] The high penetrating power of X-rays enables *in situ* experiments to be conducted. These features have made it possible to perform structural determination of catalysts under working conditions.[2,5-10]

Some drawbacks are that EXAFS only provides one-dimensional information and that the number of data points obtained is limited by Nyquist theory, as discussed later. It can be difficult to determine the structure of complex molecules or materials with different bond lengths by curve fitting.

In this short review, we present the current status of EXAFS analysis and the remaining problems, especially with respect to curve fitting. We propose several possible solutions to overcome these problems.

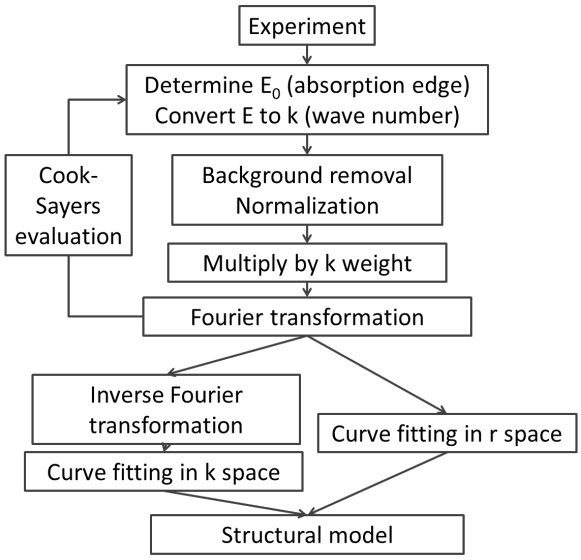


Figure 2. Flow chart of EXAFS analysis

2. EXAFS data treatment and analysis

Figure 2 shows a flowchart of the steps involved in EXAFS analysis.[11,12] The photon energy is converted to k (wavenumber) using Eq. (**2**). The background is removed from the raw data; the background is calculated using spline smoothing using Cook/Sayers criteria.[13] The oscillations are then normalized with respect to the edge height. The edge height is estimated by extrapolation of the pre-edge region to higher energy and calculation of the difference between the extrapolated pre-edge curve and the post-edge background determined by spline smoothing.[12] The edge height has an energy dependence expressed by the Victoreen or McMaster equations.[14,15] The EXAFS oscillations are multiplied by a weighting factor or k^n (n is typically 3) to compensate for damping of the EXAFS oscillations in high-k regions. EXAFS oscillations are expressed as a sine function; therefore, a Fourier transform gives peaks that correspond to the bond

lengths. The term coordination shell is often used. Atoms of the same type of element coordinated to the central atom with the same bond length are classified as one coordination shell. The coordination shells can be separated using a Fourier transform if the bond lengths are sufficiently different from each other. Each peak is inversely Fourier transformed to k space again and curve fitting is performed using Eq.(**1**). $F_j(k)$ and $\phi_j(k)$ can be derived theoretically using FEFF, a theoretical EXAFS calculation code.[16] Non-linear least squares curve fitting is performed, in which the initial parameters are based on the scattering atoms. To separate several coordination shells, the Fourier transform peaks are selected and the data are inversely Fourier transformed to k-space again, and curve fitting is conducted. The residual between the observed and calculated spectra is obtained and the parameters are then changed to reduce the residual. This residual is discussed later as R^2 (the R-factor), which is given by Eq. (4). Several routines can be used to determine the minimum, such as the Newton-Raphson, Powell, and Levenberg–Marquardt methods.[17]

When the difference between two coordination shells is small, the shells must be analyzed at the same time. This multi-shell analysis causes many serious problems and is the main topic of this paper. If the different shells consist of the same element but have slightly different bond lengths, they can be analyzed as a single shell with static disordering. If the bond length distribution is symmetric or Gaussian, the static disordering can be expressed using the Debye-Waller term. If the distribution is asymmetric, then it cannot be exactly expressed using the Debye-Waller term, and this is a problem for non-crystalline systems for which EXAFS is the only method for structural analysis. Such multi-shell and asymmetric scenarios are the two main problems in EXAFS analysis.

At the end of this section, where we present an overview of EXAFS analysis, the errors involved in EXAFS will be discussed, specifically those related to the experimental procedure and data processing.

Concerning errors related to the experimental procedure, synchrotron radiation is a typical light source and provides a high X-ray flux and good statistics. The EXAFS instruments installed in synchrotron facilities are well tuned and maintained by the synchrotron facility staff members. Errors in experiments may arise from sample preparation, which is carried out by the users themselves. The important point in sample preparation is how to obtain a homogenous sample with an appropriate thickness. A powder sample must be well ground,[18] and then mixed with boron nitride (BN) or other diluents and made into a disk with uniform thickness. The appropriate thickness is that which provides a total absorption coefficient of less than 4, an edge height of more than 0.2 and ideally close to 1.0. If it is difficult to obtain such values, then the sample can be measured using other methods, such as fluorescence yield and electric current yield modes. When a sample is carefully prepared, the experimental error is less than 1% for the coordination number, which is much smaller than the errors created in data processing.

Concerning the data processing, its most important point is how to correctly estimate the normalization factor. Its incorrect and careless evaluation can lead to an error of up to 10%. Bian et al. discussed the possible errors in EXAFS data processing for a PdCu bimetallic cluster and found that the estimation of the normalization factor can produce a 3% error in the coordination number, even after its careful selection.[19] More attention should therefore be paid to the normalization process.

Larger errors may occur in the curve fitting process for a complex system. If the sample is a simple system with one type of element in well-

separated coordination shells, then EXAFS analysis can determine the bond lengths with a precision of 0.001 nm, and the coordination number with an accuracy of ±0.3 for the first shell. However, for multi-shell analysis of a complex structure, large errors or no reliable fitting results will be obtained. In what follows, we discuss difficulties and problems with curve fitting.

3. Problems with curve fitting

The problems and difficulties with curve fitting to determine the structural parameters in a complex system can be summarized as follows:

1. Number of data points
2. Goodness of fit
3. Parameter correlations
4. Initial parameter dependence
5. Error estimation
6. Asymmetric distribution

We discuss each of these separately, although the six items are correlated with each other. The key words are "multi-shell fitting" and "asymmetric distribution". The sources of the difficulties are the limited number of data points and parameter correlation.

3-1. Number of data points

The number of data points, ν, which can be extracted from the EXAFS is limited by the Nyquist theory:[20]

$$\nu = 2(k_{max} - k_{min})(r_{max} - r_{min})/\pi + 2,$$

or

$$\nu = (2\,\Delta k\ \Delta r\ /\pi) + 2, \tag{3}$$

where

$$\Delta k = (k_{max} - k_{min}),$$
$$\Delta r = (r_{max} - r_{min}).$$

If $\Delta k = 120(= 150 - 30)nm^{-1}$ and $\Delta r = 0.15\,nm$, then $\nu \approx 14$. The number of fitting parameters, p, must be smaller than ν. If four parameters are necessary for each shell, then only three shells at most can be determined. Therefore, it is quite difficult to investigate a complex system with different bond lengths and types of coordinating atoms.

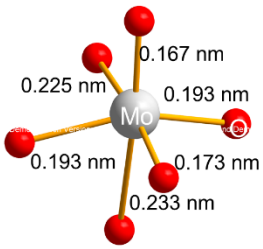


Figure 3. Local structure of MoO_3

For example, consider the structure of MoO_3. Figure 3 shows that MoO_3 has 6 different bonds with lengths in the range of 0.160-0.240 nm. The bonds cannot typically be distinguished by the Fourier transform approach and the data is thus analyzed using multi-shell curve fitting. If each bond requires four parameters (*CN*, *r*, σ, and E_0), then 24 fitting parameters (4×6) are needed. This is larger than the number of data points (ν=14) obtained by Eq. (**3**), so it is difficult to determine the exact structure. The number of fitting parameters must therefore be reduced. One way to achieve this is to classify the bonds into three categories. One is short bonds in the range of 0.16-0.18 nm, which corresponds to Mo=O double bonds. The second is in the range of 0.19-0.21 nm, which corresponds to Mo-O single bonds, and the third is in the range of 0.22-0.24 nm. E_0 can then be set to the same value, which is equal to that for the reference compounds. Consequently, the number of fitting parameters can be reduced to six and reasonable fitting results can be obtained, as shown in Table 1, if the initial structure determined by crystallography was used.[21] The coordination number can be well reproduced within the error margin. However, a priori information regarding the MoO_3 structure is required,

without which, the coordination number cannot be obtained easily.

Table 1. Three-shell curve fitting results for MoO_3.[21]

	CN	*r* / nm	σ^2 / 0.01 nm^2	R^2
r_1	1.6	0.170	0.0024	
r_2	2.2	0.196	0.0037	0.0004
r_3	2.4	0.228	0.0079	

The number of data points and the number of parameters for the EXAFS spectrum must thus be given significant attention. The way to increase the number of data points is to measure as large a *k*-range as possible. EXAFS oscillations are damped in the high *k*-region because the scattering ability of surrounding atoms for electrons with high wavenumber, *k*, (or large momentum *p*) is small, and because of the reduction factor 1/*k* in Eq.(**1**), and the Debye-Waller factor due to the thermal and static disorder. Thermal disorder can be reduced by low-temperature measurement. XAFS experiments at low temperature are thus recommended in order to measure the high *k*-region signal with good quality.

Figure 4 shows EXAFS oscillations for Fe foil measured at room temperature (RT) and at 70 K. The EXAFS at 70 K is approximately 1.2 times stronger than that at room temperature. The *k*-range for the Fourier transform can be extended to 200 nm^{-1}. The first peak around 0.217 nm is split in the Fourier transformed data for the 70 K measurement. Fe has a body-centered cubic (BCC) crystal structure, so that the first peak contains two shells of eight (r=0.246 nm) Fe-Fe + 6 (r=0.284 nm) Fe-Fe bonds. The peak at RT is not split clearly where data up to 170 nm^{-1} is available, but it begins to split at 70 K due to the increase of the k-range (Δk=30–200 nm^{-1} at 70 K) and the increase in the oscillation intensity in the high *k*-region.

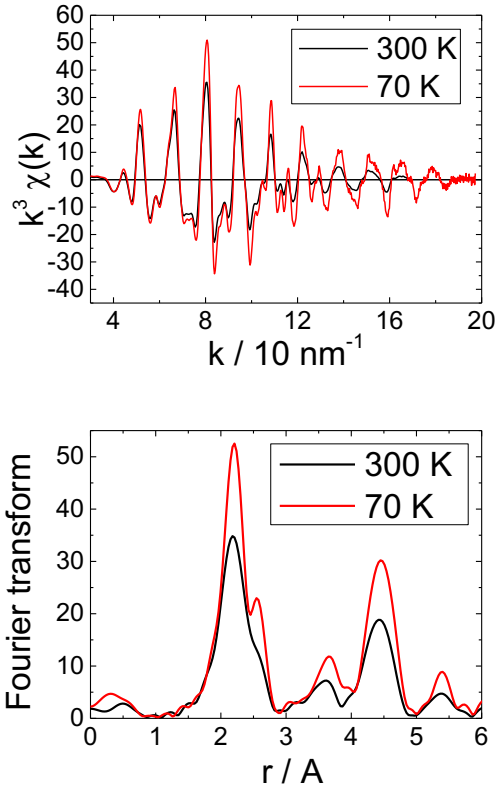


Figure 4. EXAFS oscillations and their Fourier transforms for Fe foil measured at different temperatures.

The fitting results are shown in Table 2. When the Fe K-edge EXAFS measured at RT was analyzed by two-shell fitting, a good fitting result was obtained. In this analysis, E_0 for the two shells were common. The coordination number for the second shell (at 0.285 nm) was larger than expected at RT, while that for the first shell was well reproduced. On the other hand, the EXAFS results measured at 70 K gave coordination numbers of near 8 and 6, which are in good agreement with the crystal data. The numbers of data points were $\nu = 13.4$ for RT and $\nu = 16.2$ for 70 K. Therefore, the fit is more reliable for the 70 K data and the best fit value corresponds well with the actual value. On the other hand, the RT data gave the best fit values with large errors. The second shell Fe-Fe coordination number (8.6) is larger than 6. When the parameters were fixed to values determined at 70 K, except σ^2,

which expresses the thermal disorder, a structure with a reasonable degree of fitting (R^2= 8×10^{-3}) could be obtained. R^2 is defined by Eq. (4). R^2 is insensitive to the fitting parameters when a large number of fitting parameters are used with a small number of data points, or when curve fitting is conducted for a smaller value of ν-*p*.

Table 2. Curve fitting results for Fe foil measured at RT and 70 K.

T	Bond	*CN*	*r* / nm	E_0/ eV	σ/ 0.0 1 nm^2	R^2/ 10^{-3}
70K	Fe-Fe	7.5	0.248	1.1	0.045	10
	Fe-Fe	6.2	0.285	1.1	0.059	
RT	Fe-Fe	7.2	0.247	-3.6	0.063	2
	Fe-Fe	8.6	0.283	-5.0	0.085	
RT*	Fe-Fe	(7.5)	(0.248)	(1.1)	0.062	8
	Fe-Fe	(6.2)	(0.285)	(1.1)	0.075	
70K	Fe-Fe	10.5	0.250	7.3	0.058	120

*1 structure parameter, except σ, is fixed at values for 70 K. Numbers in parentheses show the fixed parameters.

In summary, if curve fitting is performed for a complex system, then care must be taken with the number of data points; p must be less than ν. Moreover, a large value of $\nu - p$ is preferable because curve fitting is more reliable and sharp around the best fitted values.

3-2. Goodness of fit

The goodness of fit is evaluated by R^2, which is defined as

$$R^2 = \sum(\chi_{\mathrm{cal}}(k) - \chi_{\mathrm{obs}}(k))^2 / \sum(\chi_{\mathrm{obs}}(k))^2 . \quad (4)$$

A smaller R^2 indicates better fitting results. However, there are questions with respect to the meaning of R^2 and the criterion for good fitting results.

If $\chi_{\mathrm{cal}}(k) \approx \chi_{\mathrm{obs}}(k)$, then R^2 would be 0. If $\chi_{\mathrm{cal}}(k) = 0$ for all k, $R^2 = 1$. If $\chi_{\mathrm{cal}}(k)$ is $\frac{\pi}{2}$ rad different from $\chi_{\mathrm{obs}}(k)$, then $R^2 \approx 2$. Although we have no definite criterion to show if a fit is sufficient, it is often said that when R^2 is less than 0.1, the model is acceptable. If R^2 for the reference compounds is always more than 0.1, then the validity of the phase shift and amplitude functions derived from theoretical calculation such as the FEFF program could be accepted. Care must be taken to increase the number of fitting parameters to obtain a good fitting result. An increase in the number of fitting parameters often gives a smaller R^2, and R^2<0.1 can easily be obtained. However, a smaller R^2 is not always better, even if the number of parameters is less than the number of data points. In addition, we cannot say that the model structure that gives the smallest R^2 is superior to the others, even if the number of fitting parameters is the same.

In summary, R^2 is a simple indicator of the goodness of fit; however, there are no definite criteria regarding sufficient fitting and caution is required with respect to R^2 if too many parameters are needed to obtain better fitting.

The committee of XAFS Standards and Criteria have suggested the use of χ^2 instead of R^2:[22,23]

$$\chi^2 = \frac{\nu}{\nu - p}\frac{1}{\mathrm{N_{data}}}\sum\left(\chi_{\mathrm{cal}}(k_i) - \chi_{\mathrm{obs}}(k_i)\right)^2 / {\varepsilon_i}^2, \quad (5)$$

where $\mathrm{N_{data}}$ is the number of data points and ε_i is the error for each data point.

χ^2 follows a χ^2 −distribution, which should take a maximum when $\chi^2 = 1$. If $\chi^2 \gg 1$ or $\chi^2 \ll 1$, then the fitting results are less reliable. Note that $\chi^2 = 0$ is much less probable and not acceptable from statistical reasons alone.

If many parameters are used, then the χ^2 value will increase for a certain number of fitting parameters, even though R^2 is much lower. The

effect of p can be taken into account by including the $\nu - p$ term in the denominator. The estimation of parameter errors is easier. The XAFS Standards and Criteria suggested in 1990 that the error bar for a parameter should be estimated by an increase of $(\nu - p)\chi^2$ by 1 when the parameter was fixed at a value that deviated from the best fit with other parameters optimized.[22] This value corresponds to the significance level of one σ in the Gaussian distribution (or 32%).[24]

Despite these advantages, χ^2 has a disadvantage in the estimation of ε_i, the error at each data point. ε_i not only includes the probability (statistical) error, which can be estimated from the statistical treatment of the data, but it also contains systematic (non-statistical) errors.[22] Systematic errors arise from many uncontrollable sources, so that correct estimations are difficult. If the probability errors are larger than the systematic errors, then the error ε_i^2 could be represented by the probability error. However, the standard data quality obtained in synchrotron facilities in transmission mode is very good and the probability errors can be neglected. If only the probability error is used, then χ^2 will be much larger than unity. The analysis of systematic errors is too difficult to estimate ε_i^2 correctly; therefore, we cannot use χ^2 directly, which is why many people still use R^2 instead of χ^2, although the meaning of R^2 is not so clear. This disadvantage of R^2 will be discussed later.

3-3. Correlation problem

R^2 is adopted for estimation of the degree of fit. Non-linear curve fitting is conducted to determine the minimum of R^2. Here, the p dimensional parameter vector $\boldsymbol{\mathcal{P}}$ is defined, of which the components are fitting parameters $(p_1, p_2, p_3, \cdots, p_p)$, where the number of fitting parameters is p. Now we define the residual $\delta\varepsilon_r(k_i)$:

$$\chi_{\mathrm{cal}}(k_i, \boldsymbol{\mathcal{P}}) - \chi_{\mathrm{obs}}(k_i) = \delta\varepsilon_r(k_i)$$

$$\sum \delta\varepsilon_r(k_i)^2 = \sum |\chi_{\mathrm{cal}}(k_i, \boldsymbol{\mathcal{P}}) - \chi_{\mathrm{obs}}(k_i)|^2$$

$$= R^2(\boldsymbol{\mathcal{P}}) \sum |\chi_{\mathrm{obs}}(k_i)|^2$$

$$= R^2(\boldsymbol{\mathcal{P}}) * Const = R'^2(\boldsymbol{\mathcal{P}}) \quad (6)$$

After curve fitting, the optimal parameters are determined as $\boldsymbol{\mathcal{P}}_f$. Around $\boldsymbol{\mathcal{P}}_f$, $R'^2(\boldsymbol{\mathcal{P}})$ can be expanded in a Taylor series. $\frac{\partial R'^2(\boldsymbol{\mathcal{P}}_f)}{\partial p_i} = 0 (i = 1,2,3,\cdots,p)$; therefore,

$$R'^2(\boldsymbol{\mathcal{P}}) = R'^2(\boldsymbol{\mathcal{P}}_f) + 1/2 \sum_{i,j}^{\mathrm{p}} \frac{\partial^2 R'^2(\boldsymbol{\mathcal{P}}_f)}{\partial p_i \, \partial p_j} \Delta p_i \Delta p_j, \quad (7)$$

$$R'^2(\boldsymbol{\mathcal{P}}) = R'^2(\boldsymbol{\mathcal{P}}_f) + \sum_{i,j}^{p} B_{ij} \Delta p_i \Delta p_j, \quad (8)$$

If the non-diagonal term $B_{ij} = 0, (i \neq j)$, then there is no correlation between the fitting parameters. In contrast, when $B_{ij} \neq 0, (i \neq j)$, there is a correlation. In a single-shell analysis, amplitude and phase terms can be separated by a ratio method, as discussed later.[25,26] The two parameters, coordination number (*CN*) and Debye-Waller factor

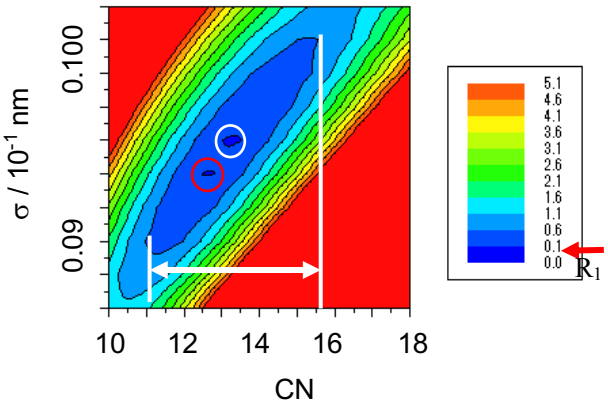


Figure 5. 2-dimensional mapping, where white and red circles indicate the locations of the best and next best R^2 points. The white double arrowed line indicates the error range.

(σ), in the amplitude term are correlated strongly, while the bond length (r) and binding energy or photoelectron kinetic energy zero (E_0) are correlated in the phase term. 2-dimensional (2D) mapping is performed to study the correlation problem, as shown in Figure 5.

In the 2D mapping, two parameters (CN and σ) are fixed and the others are optimized. The two fixed parameters are varied to cover the 2D range for the R^2 mapping, or R^2 is plotted against the two fixed parameters. The minimum corresponds to the best fit value. Each R^2 contour curve may be in an ellipse. If the major and minor axes of the ellipse are parallel to the x- and y-axes, then the two parameters are independent. The axes are typically tilted and there is a correlation between the two parameters. The error may be large if a large correlation is present. The error can be evaluated using the contour curve. The error bar in one parameter is express by the R^2 less than a certain value, R^2_1, as shown by the arrow at the color map in Figure 5. If the 2D mapping is considered and a correlation is present, then the error is as indicated by the white line. In addition, 2D mapping gives another local minimum near the solution, as shown in the circles in Figure 5. The second best R^2 value (red circle) is slightly larger than the best R^2 value. The second best R^2 could be a candidate for the true structure parameters. In this case, the second best R^2 is located near the best R^2 within the error bar region, so that the second best is included in the error bar. However, if it is located far away from the solution, the next local minimum may be missed, or even if it is determined, we do not know how to deal with that minimum, which has some possibility to correspond to the real structure. In multi-shell fitting, the correlation becomes more complex and gives more local minima in the mapping. The coordination number and bond length are correlated through interference of the two shells. The complex interdependence between parameters reduces the stability and reliability of the curve fitting. The minimum R^2 does not give the correct answer as in the case of the Fe foil measured at RT, where the coordination number in the second shell was too large. The correlation between parameters is a serious problem and makes curve fitting useless, especially for a complex multi-shell system because the number of fitting parameters, p, becomes large and ν-p becomes small. This point will be discussed in a thorough search analysis in Section 4-1. The correlation problem would be less serious if EXAFS spectrum could be measured up to a large k-range, as shown in Figure 4.

In summary, the correlation problem is the most serious for multi-shell analysis, and thus requires many parameters and the value of ν-p becomes small. As a result, fitting becomes unreliable. This is the fundamental problem in curve fitting analysis.

3-4. Initial parameter dependence

In relation to multi-shell fitting, we should consider the initial parameter effect on the curve fitting results. Non-linear least squares curve fitting analysis requires initial parameters, from which the algorithm begins to reduce R^2 efficiently. Once the minimum R^2 is determined in the fitting process, the parameters are trapped at that minimum and it is difficult to escape that minimum to search for another possible minimum. The fitting result could be trapped at a wrong or meaningless local minimum, and thus it is necessary to change the initial parameters to confirm that the obtained minimum is not due to a local minimum but the global minimum, or to search for other possible structures. It is not a serious problem for a simple system with one coordination shell present and a symmetric distribution. The fitting results often give the global minimum because the function $R^2(\mathcal{P})$ has a single minimum. However, a serious problem arises for a complicated system where

multi-shell fitting is necessary. In this case, the parameter dependence of R^2 becomes small or B_{ij} in Eq. (**8**) is small, so that the shape of R^2 becomes broad and a tiny perturbation can easily change the R^2 minimum position. Moreover, the correlation between the parameters may give a physically unreasonable minimum. Kido et al. demonstrated that a curve fitting analysis of MoO_3 gave three possible structures with good R^2 values by changing the initial curve fitting parameters, as shown in Table 3.[21] We cannot say which is better from only the R^2 comparison, because the difference in R^2 is not large. Other characterization methods must thus be used to determine which is the most probable model. In addition, we cannot find other possible structure candidates by only a random selection of initial parameters.

Table 3. Curve fitting results for Mo K-edge EXAFS of α-MoO_3 with different initial parameters.

	Initial	Best	Initial	Best	Initial	Best
r_1	0.170	0.169	0.170	0.167	0.170	0.165
r_2	0.196	0.198	0.170	0.172	0.170	0.175
r_3	0.200	0.198	0.191	0.188	0.194	0.195
r_4	0.225	0.227	0.196	0.198	0.194	0.195
r_5	0.230	0.227	0.198	0.198	0.230	0.226
r_6	0.234	0.232	0.223	0.227	0.230	0.235
R^2		0.028		0.033		0.025

In this context, it should be noted that curve fitting does not always distinguish the coordinating element correctly based on R^2. We always assume the coordinating element first and calculate the backscattering amplitude and phase shift for that element. Even if we assume a different scattering atom, rather good fitting results can be obtained. Figure 6 shows the curve fitting results for K_2PtCl_4. Good fitting results can be obtained if we assume Cl as the first nearest neighbor. The Pt-Cl distance is 0.232 nm, which corresponds well to the reported value of 0.231 nm.[27] However, if O or Fe are assumed as the scattering atoms, then good fitting results with bond lengths of 0.242 nm and 0.244 nm, respectively, can be obtained. The fitting results for Pt-O and Pt-Fe are unusual due to the abnormality in fitting parameters such as the bond length, Debye-Waller factor and E_0. The Pt-O bond length is typically found to be less than 0.220 nm in the reference compounds and the Debye-Waller factor for Pt-O is too small. In the same way, the Pt-Fe bond length should be 0.250-0.260 nm. This example indicates that care should be taken with the coordinating element, even if R^2 gives a better fitting result.

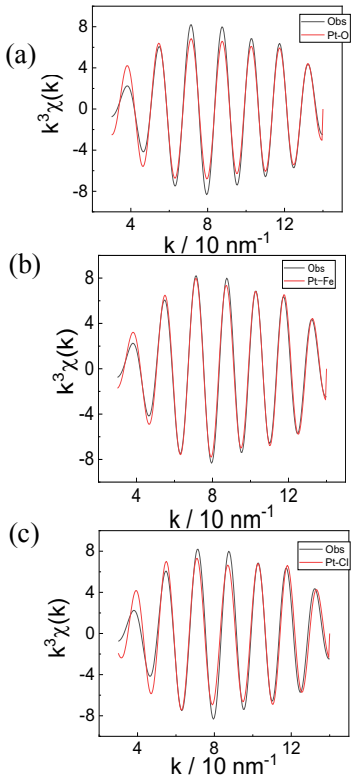


Figure 6. Curve fitting results for Pt L_3-edge EXAFS spectra for K_2PtCl_4 assuming different coordinating atoms; a) Pt-O, b) Pt-Fe, and c) Pt-Cl.

In summary, care must be taken to select the initial parameters before the curve fitting analysis is started.

Moreover, attention should be paid to the fitting results with a comparison to the known data for the material to confirm that the obtained results are reasonable. Sometimes the validity of the results must also be checked using other characterization methods.

Table 4. Curve fitting results for K_2PtCl_4 with different coordination atoms

	CN	*r* / nm	σ / 0.1 nm	E_0 / eV	R^2
O	8.2	0.243	0.000	-1.5	0.06
Fe	2.4	0.244	0.030	-26	0.02
Cl	3.1	0.232	0.018	9.8	0.03

3-5. Error estimation

As discussed earlier, error estimation would be possible if χ^2 could be determined more exactly. The problem was how to estimate the error for each data point, ε_i^2. R^2 does not have an appropriate value to indicate that it is a sufficiently good fit or not.

Here, we propose the Hamilton test for error estimation using R^2. The Hamilton test may be an approach to use R^2 efficiently.

3-6. Hamilton test

We have stated that χ^2 is better than R^2; however, the difficulty in error estimation prevents the use of χ^2. The solution to this dilemma is to use the Hamilton test.[28] The Hamilton test uses an *F*-test. Hamilton proposed the method in 1965 for the analysis of diffraction data, where R^2 is used to estimate the goodness of fit. If R^2 is approximately distributed in a χ^2-like way, then the ratio of R^2's of two models exhibits an *F*-distribution. This assumption can be rationalized if the error for each data point ε_i^2, is constant. The superiority of model A over model B can be determined from the R^2 ratio of the two models. The error in the fitting parameters can be estimated in the same way. The shape of an *F*-distribution is dependent on ν-*p*, difference between the number of data points and the number of parameters; therefore, the Hamilton test can judge the appropriate number of fitting parameters as described in section 3-6-3.

First we show the procedure used in the Hamilton test. The Hamilton test requires R^2 for both model A and model B, which are given as R_A^2 and R_B^2, respectively. ν is the number of data, and p_A and p_B are the numbers of fitting parameters to obtain models A and B, respectively. Also, b is the increase in the number of parameters in model A from model B or

$$b = p_A - p_B. \tag{9}$$

The R^2 ratio is defined as

$$\mathcal{R}^e = R_B^2/R_A^2. \tag{10}$$

$F = (\nu - p_A)(\mathcal{R}^e - 1)/b$ follows an *F*-distribution with two degrees of freedoms, b and $\nu - p_A$. The important concept is the significance level, which is the probability of whether the null hypothesis can be rejected or not. For example, when model A has better fitting results than model B, the point we want to claim is that model A is better than model B. The null hypothesis is defined as "Model B can reproduce the EXAFS as well as model A". The significance level is set at 5%. $\mathcal{R}^e$ and *F* are calculated first using R_A^2, R_B^2, b, and $\nu - p_A$. The occurrence probability for the null hypothesis is calculated from the *F*-distribution. If the occurrence probability is more than the 5% significance level, we cannot reject the null hypothesis. But if it is less than 5%, we can reject the null hypothesis and can safely say that model A is superior to model B. If the error is estimated by one σ of a Gaussian distribution, then the

significance level is 32%. The Hamilton test can then be used to conduct an error estimation and compare two models using R^2 as shown next.

3-6-1. Error estimation

Here we present examples of error estimation using the Hamilton method. The value of R_A^2 for the best fit is 0.05 with four fitting parameters. The fitting range and Fourier transform range are 30-160 nm^{-1} and the inverse Fourier transform range in r-space is 0.1-0.2 nm. The best fit bond length is 0.200 nm. When the bond length is fixed at 0.201 nm and the best fit is obtained with the other three parameters optimized, this corresponds to model B where one parameter is fixed. If R_B^2 = 0.055, then, χ^2 is 1.1, $\nu = 2\Delta k\Delta r/\pi + 2 = 2 \times (160 - 30) \times (0.200 - 0.100)/\pi + 2 \approx 10.3 \approx 10$, p_A=4, and b=p_A-p_B=1. The null hypothesis is that R_B^2 is the same as R_A^2. The null hypothesis cannot be rejected at the probability of 75%, which is more than 32%. When the bond length is fixed at 0.203 nm and we obtain R_B^2 = 0.09, the same hypothesis is rejected at the 32% significance level, since the occurrence probability of the null hypothesis is 31.8 % less than 32 %. We can say the error should be 0.003 nm.

3-6-2. Determining the number of fitting parameters

We cannot obtain a meaningful result when p_A is more than ν. If $\nu - p_A > 0$ is satisfied, then it is allowable to fit the multi-shell data with model A using p_A parameters. We have a one-shell fitting result using model B. We want to demonstrate that model A is superior to model B. The goodness of fit for model B is 0.05 (R_B^2) with the four (p_B) fitting parameters, and then two-shell fitting is performed based on model A, where a total of p_A =8 parameters are used. The best fitting for eight parameters (model A) is 0.01 (R^2_A). The fitting range and inverse Fourier transform range are the same and $\nu = 10.3$, which gives χ^2 =5.0, b=4 and $\nu - p_A =$ 2. The null hypothesis is that the two fitting models are identical. In the Hamilton test, the occurrence probability is calculated to be 36%, and model B cannot be rejected at the one σ level. If the Fourier transform range is increased to Δk =190 nm^{-1}, and $\nu - p_A \approx 4$, then the occurrence probability is 10.4% less than the 32% significance level. Consequently, fitting using two shells can be improved if we use the wide k-range of 30-190 nm^{-1} while 30-150 nm^{-1} is still small, and we cannot say that two-shell fitting is better than one-shell fitting.

3-6-3. Comparison of model structure.

Model A assumes an M-A bond and model B assumes an M-B bond with R_A^2= 0.01 and R_B^2=0.016, respectively. Can we say that model A is better than model B? The k-range and the inverse Fourier transform range are 100 nm^{-1} and 0.1 nm, respectively. $p_A = 4, p_B = 4$, $\nu - p_A \approx 8$ and b=p_B=4, so that χ^2 =1.6. The null hypothesis is "the goodness of fit for model B is equivalent to that for model A". The occurrence probability is 38%, so that it cannot be rejected at a significance level of 32%. On the other hand, if R^2_B=0.02, then the occurrence probability is 19%, so that model B can be rejected at a significance level of 32% (or one σ in the Gaussian distribution). In the case of $PtCl_4$ discussed previously, Pt-Cl, Pt-O and Pt-Fe gave R^2 = 0.03, 0.06 and 0.02, respectively. ν-p_A and b were 5 and 4, respectively, for the three analyses. The result for Pt-Fe is the best fit value. The hypothesis is "the result for Pt-Cl(O) has the same goodness of fi as that for Pt-Fe". The occurrence probability for Pt-Cl is 69% >32%, while the occurrence probability for null hypothesis for Pt-O is 20 % and can be rejected at the 32% significance level. The Pt-O coordination can be rejected from an EXAFS perspective. However, this does not mean that Pt-Fe and Pt-Cl are the only candidates. There are possibilities that another model gives a lower R^2 and is a valid model.

3-7. Asymmetric distributions

The curve fitting method assumes the radial distribution function to be Gaussian or symmetric. Therefore, an asymmetric distributed system with a large static disorder cannot be analyzed by curve fitting using Eq. (**1**). One approach is to use an asymmetric distribution function to include the effect of asymmetry. The coordination number multiplied by the Debye-Waller factor term $CN\mathrm{e}^{-2\sigma^2k^2}$ can be replaced with the distribution function, $g_j(r)$:[5,29,30]

$$\chi(k) = \sum \int g_j(r)F_j(k)e^{(-2r_j/\lambda_j)} \sin\left(2kr_j + \phi_j(k)\right)/kr_j^2\, dr \quad (11)$$

where j indicates the j-th element (O, C, N, Fe, etc.). If a single shell and the following asymmetric function are assumed, then:

$$g(r)\begin{cases} = A(r-r_0)^2e^{-B(r-r_0)} & r \geq r_0 \\ = 0 & r \leq r_0 \end{cases}. \quad (12)$$

$g(r)$ has a sharp rise and a gentle descent, as shown in Figure 7. $g(r)$ has a peak at $r = r_0 + \frac{B}{2}$. If the $F_j(k)e^{(-2r_j/\lambda_j)}/r_j^2$ and $\phi_j(k)$ terms can be neglected and $\chi(k)$ is simply expressed as $k\chi(k) = \int g(r)\sin 2krdr$, then

$$k\chi(k) = \frac{AB^2}{(B^2+4k^2)^3}[B(B^2-12k^2)\sin(2kr_0) + 2k(3B^2-4k^2)\cos(2kr_0)] \quad (13)$$

Figure 7 (b) shows the Fourier transform of Eq. (**13**). The black curve is obtained by the Fourier transform of Eq. (**13**) in the *k*-range of 0–150 nm^{-1}. The asymmetry of the original function is well reproduced. The red broken line in Figure 7 (b) is the Fourier transform of Eq. (**13**) over the *k*-range of 30-150 nm^{-1}. The low-*k* region is typically cut off in an EXAFS analysis because this region includes the effects of edge and multiple scattering, which make the EXAFS analysis complicated. The red broken line has a smaller peak and is shifted to the lower *r* direction because of the asymmetric distribution and low *k*-region cut off. In the asymmetric distribution, the oscillation $\chi(k)$ is composed of many components with different wavelengths that are dependent on the bond length, *r*. The steep rise in the radial distribution function gives a larger contribution than the gentle descent where oscillations with various wavelengths destructively interfere with each other, so that the EXAFS oscillations are damped quickly in the low-*k* range. The Fourier transform thus reflects the steeply rising part of the radial distribution curve more strongly than the gently descending part when the low-*k* region is cut off. The amplitude decreases and the peak position is shifted to the lower direction. Moreover, the Fourier transform becomes rather symmetric. Figure 7(c) shows the normalized Fourier transforms where the peak of the two Fourier transforms coincide with different Fourier transform ranges to show the asymmetric distribution effects more clearly. The Fourier transform (red broken line) over *k*=30-150 nm^{-1} is multiplied by 2.95 and shifted by 0.0052 nm for ease of comparison. The Fourier transform peak over 30-150 nm^{-1} has a more symmetric structure than that over the range of 0-150 nm^{-1}. The asymmetry distribution effect can be expressed by a model radial distribution function; however, the model distribution is not so useful analytically. If the asymmetric distribution is not so large, then the cumulant expansion method can be applied:

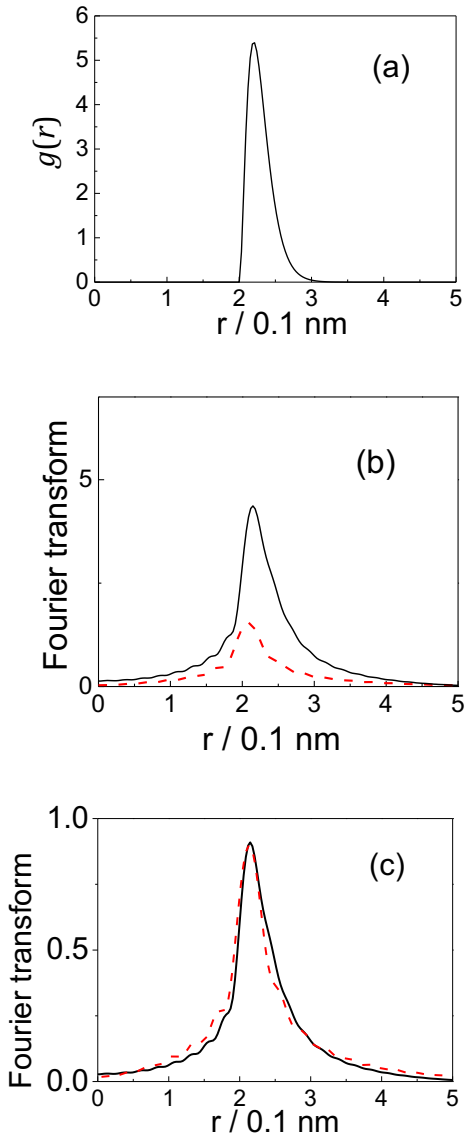


Figure 7. (a) Asymmetric function given by Eq. (9) (A=100, B=10) and (b) the Fourier transforms of Eq. (10) with different k-ranges; the black solid line is for *k*=30-150 nm^{-1}, and the red dotted line is for *k*=30-150 nm^{-1}. (c) Normalized Fourier transforms of (b). Two peaks are normalized by the peak height to be unity.

$$< e^{i2k(r-R)} > = \int P(r) e^{i2kr} dr$$

$$= e^{-W+i\phi}, \tag{14}$$

$$W + i\phi = \sum_{n=0}^{\infty} (2ik)^n / n!\ C_n, \tag{15}$$

$$W = \ln\left(\frac{kr^2\chi(k)}{CNS_0^2F(k)}\right) = C_0 - C_2(2k)^2/2! + C_4(2k)^4/4! + \cdots, \tag{16}$$

$$\phi = \phi_0 + 2kC_1 - C_3(2k)^3/3! + C_5(2k)^5/5! + \cdots. \tag{17}$$

In the cumulant expansion, if the distribution is Gaussian shaped, then $C_1 = r$ and $C_2 = \sigma^2$ and $C_n = 0;\ n \geq 3$.

If the radial distribution function is expressed as Eq. (**12**), then $C_{n(n>1)} = \frac{3(n-1)!}{B^n}$.

If the asymmetry is not large and the cumulant expansion converges rapidly, then the EXAFS oscillation can be expressed with C_n ($n \leq 4$), as

$$\chi(k) = \sum CN_jF_j(k)\exp(-2r_j/\lambda_j) \exp\left(-2\sigma_j^2k^2 - \frac{2}{3}C_4k^4\right) \sin\left(2kr_j + \phi_j(k) - \frac{4}{3}C_3k^3\right)/kr_j^2 \tag{18}$$

Although C_3 and C_4 can be determined by the curve fitting method, it is better to use the ratio method proposed by Bunker[25] and Yokoyama[26] to determine them more easily and accurately. In the ratio method, the amplitude term and the phase term can be separated using an inverse Fourier transform. Once the amplitude term and phase term are separated, the data can be fitted using the following equations derived from Eqs. (**16**) and (**17**):

$$\ln\left[\frac{A_i(k)}{A_s(k)}\right] = \ln\left(\frac{CN_i}{CN_s}\right) + \ln\left(\frac{r_s^2}{r_i^2}\right) - 2(\sigma_i^2 - \sigma_s^2)k^2 - \frac{2}{3}(C_{4.i} - C_{4,s})k^4 \quad (19)$$

$$\phi_i - \phi_s = 2k(r_i - r_s) - \frac{4}{3}(C_{3.i} - C_{3.s})k^3, \quad (20)$$

where subscripts i and s indicate the unknown compound and standard, respectively. In this ratio method, the transferability of the amplitude and phase shift functions for the same atomic pair are presumed. When $\ln\left[\frac{A_i(k)}{A_s(k)}\right]$ is plotted against k^2 as shown in Fig.8(a), the y-intercept of Eq. (**19**) corresponds to the coordination number ratio and the slope corresponds to the Debye-Waller factor difference. The second-order term corresponds to the C_4 parameters. Similarly, $(\phi_i - \phi_s)$ is plotted against k shown in Figure 8(b), where the slope corresponds to the bond length and the C_3 difference is given in the second order term. However, this ratio method cannot be applied to a system where the Fourier transform peak is composed of two or more inseparable coordination shells. In this case, a nonlinear least-squares curve fitting analysis must be applied with C_3 and C_4 parameters, so that the number of fitting parameters increases and the correlation problems become more serious. The cumulant expansion method is sometimes less effective if two or more inseparable coordination shells are present.

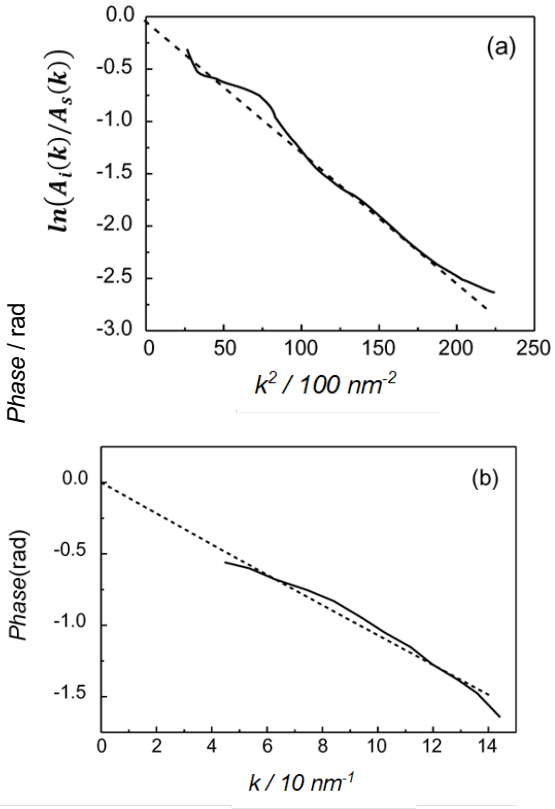


Figure 8. The ratio method; (a) amplitude, and (b) phase difference.

4. Proposal for future analysis methods to overcome curve fitting problems.

Finally, we would like to propose methods to overcome the problems with curve fitting analysis, which we again summarize:

1. Number of data points
2. Goodness of fit
3. Parameter correlations
4. Initial parameter dependence
5. Error estimation
6. Asymmetric distribution

We propose two methods to overcome these difficulties: a constrained thorough search and the reverse Monte Carlo method.

4-1. Constrained thorough search method

In the curve fitting analysis, the observed data are analyzed using the EXAFS Eq.(**1**) by optimization of the fitting parameters. Parameters are adjusted in the

parameter space to minimize R^2. This parameter space is a multidimensional metric space that consists of structural parameters such as CN, E_0, r, and σ^2, each of which correspond to one axis in the parameter space, i.e., the dimensionality of the structural parameter space is equal to the number of structural parameters. A set of structural parameters is represented by a point $\boldsymbol{\mathcal{P}}$ with components of (CN, E_0, r, and σ^2) in the parameter space. The point $\boldsymbol{\mathcal{P}}$ provides the EXAFS oscillation $\chi_{\mathrm{cal}}(k,\boldsymbol{\mathcal{P}})$ and its unique R^2. This parameter space has a function, $R^2 = f(\boldsymbol{\mathcal{P}}) = \sum(\chi_{\mathrm{cal}}(k) - \chi_{\mathrm{obs}})^2 / \sum(\chi_{\mathrm{obs}})^2$. In the curve fitting analysis, the initial point $\boldsymbol{\mathcal{P}}_0$ is given first near the solution and then R^2 is optimized, searching for the final point $\boldsymbol{\mathcal{P}}_f$ with the lowest R^2. If $\nu - p$ is small, then the gradient of $R^2 = f(\boldsymbol{\mathcal{P}})$ is not steep and the shape of the $f(\boldsymbol{\mathcal{P}})$ function around the minimum, $\boldsymbol{\mathcal{P}}_f$, becomes broad. $\boldsymbol{\mathcal{P}}_f$ may have a large error bar or the other points may be the real solutions. However, we do not know the exact shape of $f(\boldsymbol{\mathcal{P}})$, only by the curve fitting analysis.

In contrast, in the constrained thorough search analysis, all points $\boldsymbol{\mathcal{P}}$ are thoroughly surveyed to produce the entire p-dimensional R^2 map, where p is the number of parameters. We may find positions of minimum R^2 and we understand the shape of the R^2 distribution around the minimum, $\boldsymbol{\mathcal{P}}_f$. If the shape of $f(\boldsymbol{\mathcal{P}})$ is sharp, then the curve fitting analysis is reliable and gives only a small error. However, if it is not sharp, or if several local minima are determined in addition to $\boldsymbol{\mathcal{P}}_f$, then the fitting process is unreliable or the possibility of other solutions exists. Kido et al. proposed a simple estimation of the results and errors using the constrained thorough search method.[21] The occurrence probability is related to R^2 for each parameter. Exact estimation of the occurrence probability is difficult, as discussed previously, using only R^2. Another approach that can be applied is the F-test to obtain the relative occurrence probability compared to the minimum, and the occurrence probability can then be summed over the entire parameter space, which requires another large number of calculations. They suggested that the occurrence probability is unity when the result gives R^2 less than a certain threshold R^2_{c}. If R^2_{c} is sufficiently small and the assumption that all $\boldsymbol{\mathcal{P}}$'s have the same occurrence probability, the result can then be equally represented by averages of the parameters. Kido et al. drew histograms for each parameter; the histogram is a projection of the R^2 map onto the parameter axis. They approximated the histogram for each parameter, CN, r, E_0 or σ, by a Gaussian function. The Gaussian peak position was almost the same as the average value for each parameter, and the width of the Gaussian peak was treated as the error.

The confined thorough search requires a tremendous number of calculations. The number of parameters is p and one of them is p_A. If the survey range is Δp_{A} for parameter p_{A}, and its survey step is δp_{A}, then the total number of fitting parameters can be expressed as $\prod_A^p N_{P_A}$, where $N_{P_A} = \Delta p_A / \delta p_A$.

The advantage is to overcome the initial parameter dependence problem to avoid trapping at meaningless local minima. The constrained thorough search method can give several possible candidates for structures that have identically small R^2 values, which are all possible candidates. How can we treat these local minima? Kido et al. proposed the domain-averaging method.[21] They defined the domain around the possible structures and could determine possible structure candidates for different domains. The peak position and the width of the Gaussian distribution can correspond to the estimated value of the parameter, and the error or precision of the parameter in each domain. How can we determine which structure is better than the others? The domain with the smallest R^2 may not be a good criterion when the difference in R^2 for each domain is not large. We thus rely on other methods or knowledge regarding the system. Kido et al.

proposed that the use of bond valence theory[31] was effective to select one good result. Curve fitting analysis can also provide candidate structures by changing the initial parameters. However, it is unclear if these sets of candidate structures are unique and that there are no more possibilities in parameter space from the curve fitting method. The thorough search method can guarantee that the obtained sets of structures are unique and no more structures are possible, because the entire space has been surveyed.

The thorough search method in itself cannot overcome the limitation of the number of data points. It is possible to perform the thorough search method with the number of parameters p, or a dimension of parameter space that intentionally exceeds ν. Kido et al. showed that the distributions of the fitting parameters with R^2 less than R^2_c became wider and non-Gaussian when p exceeded ν. The obtained parameters were thus less meaningful. Therefore, the thorough search method cannot overcome the limitation of the number of data points, and it is important to reduce the number of dimensions or the number of fitting parameters. Kido et al. proposed a "constrained" thorough search method where the number of fitting parameters was reduced by establishing a relationship between the fitting parameters.[21] For example, the Debye-Waller factor can be expressed as a function of bond length because both parameters are related to the bond strength. A linear relation between the bond length and Debye-Waller factor was used to conduct the thorough search processes.

The number of calculations in the thorough search method is dependent on the search range Δp_A, and the search step, δp_A. A larger Δp_A and smaller δp_A are preferable to obtain precise and accurate solutions for the p_A parameter. Kido et al. proposed that a two steps calculation method, i.e., analysis first using a wide range with a rough search step followed by a limited range with a fine search. Once some candidate domains were determined, they limited the search region to each domain with a better resolution. Sets of parameters were selected that gave R^2 less than a certain value. The sets of parameters made a domain. Finally, the averages of each parameter after the fine search were obtained using a histogram fitted with a Gaussian for each domain, and this was termed the domain averaging method. Finally, the constrained thorough search analysis would identify all possible structures.

The thorough search analysis requires higher cost (time and machine power) than curve fitting analysis. However, the constrained thorough search method and the two-step analysis save much cost and the calculations can be performed within a reasonable time scale, even using a desktop PC (CPU=Intel® Core™ i7). Further development in hardware computing speeds and effective data mining software are expected to make the constrained thorough search method a standard analysis technique for EXAFS in the future.

4-2. Reverse Monte Carlo method

Monte Carlo simulation is named after the town named Monte Carlo in Monaco, famous for its casinos. Monte Carlo simulation is based on a random walk to obtain stable complex structures by minimizing the energy of the system. The energy E_0 is calculated with one structure. The atomic positions are then randomly moved and the energy of the new structure E_1, is calculated. When the energy becomes smaller than the previous value, the new structure is accepted and the structure is then moved again from the new structure. On the other hand, when the energy becomes larger than the previous value, the new structure is accepted with a probability proportional to $\exp(-(E_1 - E_0)\beta)$, which is called the Metropolis method. β is inversely proportional to the temperature: $\beta = 1/k_BT$ (where k_B is the Boltzmann constant). If

the new structure is rejected, then the atom positions are returned to the previous positions, which is the starting point for another new random walk. The system will reach a stable state or no further change will occur in the energy. This corresponds to the equilibrium at β. The structure at this equilibrium corresponds to the stable structure for the given system.

Reverse Monte Carlo (RMC) simulation uses the degree of fit R^2, instead of energy, E.[32-34] RMC has been used for the analysis of X-ray and neutron scattering data where medium-range interactions are important. First, a large number of atoms are prepared in a large box (referred to as a cell) and R_0^2 is then calculated for the initial atom positions in the large cell. The number of atoms in the box is more than 10^2-10^3. After the random walk, the new R_1^2 is calculated for the new model. If $R_1^2 < R_0^2$, then the new model is accepted for the next trial. Even if $R_1^2 > R_0^2$, the new model is accepted with a probability of $\exp(-(R_1^2 - R_0^2)\beta)$, and the next calculation then starts. The system reaches equilibrium when there is no further improvement in R^2 after many cycles. The structure at equilibrium is the final structure for this system. Thus, a large number of atoms are contained in a large cell where the atoms move with the calculation of X-ray or neutron scattering intensity and comparison with the experimental data. The RMC method has been applied to EXAFS analysis.[35-40] In the case of EXAFS analysis, the nearest-neighbour interactions are the most important, so that it is not necessary to calculate all of the atomic pairs in a large cell as shown by the arrows in Figure 9(a). Fujikawa et al. proposed a micro RMC (m-RMC) method[41] that, in principle, follows the conventional RMC procedure. In m-RMC, a large cell is divided into independent replica files, each of which includes a small chemical species such as a molecule, a metal complex, or a metal cluster, as shown in Figure 9(b). The initial model structure in each replica file is assumed, and the m-RNC method is then performed using the set of replica files.

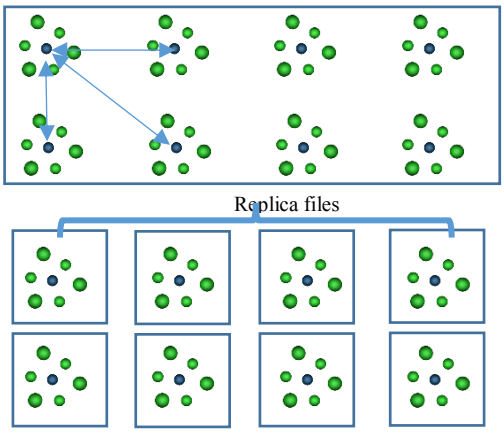


Figure 9. Calculated cells in (a) reverse Monte Carlo (RMC) and (b) micro-reverse Monte Carlo (m-RMC) methods.

Many replica files are prepared. $\chi_{\mathrm{cal}}^{i}(k)$ for the i-th file is calculated based on the structure in each replica file. For a molecule, the central atom is regarded as an X-ray absorbing atom and the others are treated as scattering atoms. In a metal cluster that has several absorbing atoms, for example in a Au cluster, all Au atoms in the cluster are simultaneously both absorbers and scatterers. $\chi_{\mathrm{cal}}^{i}(k)$ in the replica file is obtained using Eq.(**1**). The $\chi_{\mathrm{cal}}^{i}(k)$ values are averaged over all the absorbing atoms in the replica files to obtain the averaged $\chi_{\mathrm{cal}}(k)$. The averaged $\chi_{\mathrm{cal}}(k)$ is then compared with $\chi_{\mathrm{obs}}(k)$ to estimate the goodness of fit, R^2. To avoid being trapped at a local minimum and to decrease the calculation time, β is first set to a large value of approximately 10^2. β is then gradually decreased to obtain an accept/reject ratio of almost unity. After 10^4–10^5 steps of cycles, R^2 reaches equilibrium, as shown in Figure 10. The radial distribution function can be calculated using all replica files, as shown in Figure 11. The cumulant coefficients from the

final configurations are given by the following equations:

$$r = (C_1) =< r >,$$

$$C_2 =< (r - \bar{r})^2 >, \text{ and}$$

$$C_3 =< (r - \bar{r})^3 > \quad (21)$$

Figure 12 shows the results of a Au L_3-edge EXAFS analysis of Au foil by the m-RMC method. The first peak in the Fourier transform data for Au foil was analyzed after the inverse Fourier transform. The initial parameters gave ill-fitting results. R^2 rapidly decreased at the first calculation and then it gradually decreased, as shown in Figure 10. After approximately 100000 iterations, the fit became much better, as shown in Figure 12(b). β was modified with R^2; β was first selected as 1×10^4 when $R^2 > 10^{-4}$. When the R^2 became less than 10^{-4}, β was set to 10^5. In a complex case, β should sometimes be reduced to escape the minimum. Fujikawa et al. showed that 10 files were necessary to obtain good R^2 results.[41] More than 100 replica files were sufficient to give a stable structure. The results are obtained as $<r>$=0.287 nm, $\sqrt{<\sigma^2>}$=0.00914 nm, and $<C_3>= 9\times10^{-7}$ nm^3. A very small peak appeared around 0.330 nm, as shown in the inset of Figure 11; this was negligibly small but always appeared. The R^2 difference with and without the small peak was very small; therefore, it was considered to be a type of unremovable noise. However, it was possible to remove such noise peaks by setting a higher limit for the bond length.

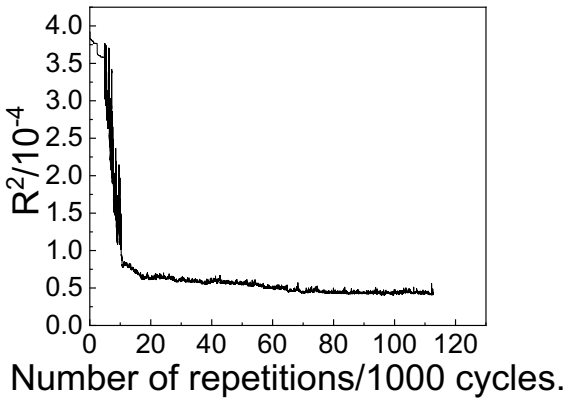


Figure 10. Change in R^2 as function of number of repetitions.

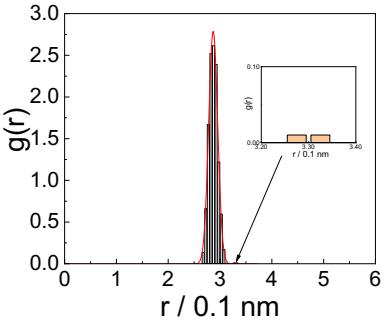
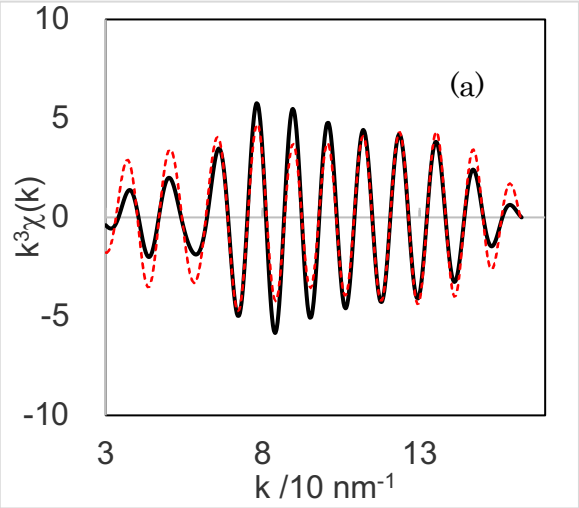


Figure 11. Radial distribution function for Au foil obtained from m-RMC analysis.



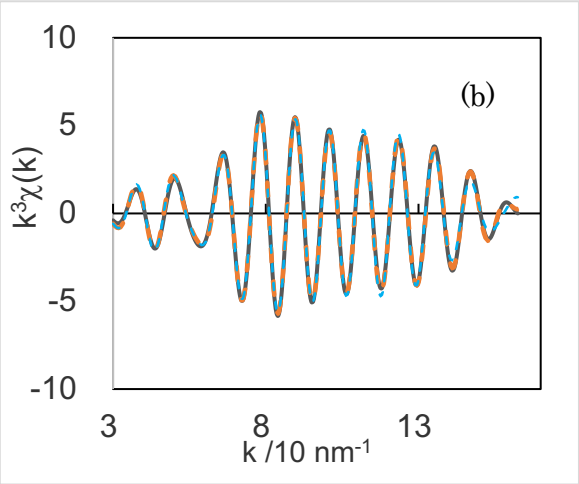


Figure 12. EXAFS analysis by m-RMC method; (a) before starting m-RMC calculation, and (b) after R^2 reached equilibrium.

In the m-RMC method, the total R^2 is reduced during the repetition. This does not mean that $\chi(k)$ for each replica file necessarily follows $\chi_{obs}(k)$, giving a small R^2, i.e., each replica file does not necessarily represent the real molecular structure. Only the final radial distribution curve can reproduce the EXAFS oscillation. For example, the real molecular structure is MO(L)2O(S)2, where two types of oxygen atoms are coordinated. One type, O(S), is the oxygen with a short bond length, while the other type of O(L) has a larger bond length. The total ratio of O(S) to O(L) is 1. The m-RMC method can give two solutions. One is MO(S)2O(L)2 or the structure with two short and long O atoms coordinated at the same time. The other is a mixture of the two types of molecular structures, each of which is composed either of the MO(S)4 or MO(L)4 structure, and each type is present with a 1:1 ratio. Both give the same averaged $\chi_{cal}(k)$. We cannot distinguish which is the real structure from only the EXAFS. Other characterization information is thus required to reach a conclusion.

In the case of MoO_3 analysis, the m-RMC method can provide the distribution function, as shown in Figure 13. Three peaks that correspond to two short Mo=O, two medium length Mo-O and two long Mo-O bonds are identified. Note that it is not necessary to select the correct initial parameters. In addition to Mo-O, the Mo-Mo bonds in the higher shells can be analyzed.[41] Compared to the constrained thorough search method, the resolution of m-RMC is not so good, because the two short bonds in m-RMC analysis cannot be distinguished, although they can with the confined thorough search.

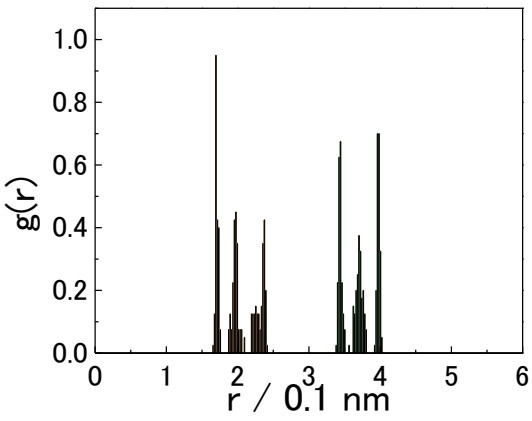


Figure 13. Radial distribution histogram for MoO_3 obtained by m-RMC method.

The m-RMC method can include an asymmetric distribution. Regulations of multi-edge analysis are automatically included. For example, the A-B bond

length must be the same when it is observed from edge A or from edge B because the same cluster structure is used for the analysis of both edges; therefore, the A-B bond length is always the same. The m-RMC method is a promising analysis technique, even if the resolution is not as good as curve fitting or the constrained thorough search method.

5. Conclusions.

We have reviewed EXAFS analysis, especially curve fitting analysis. In a simple system where one type of bond is involved in the EXAFS spectrum, the curve fitting analysis provides good results with little errors. However, for a complicated or asymmetric system, curve fitting analysis encounters several problems:

1. Number of data points
2. Goodness of fit
3. Parameter correlations
4. Initial parameter dependence
5. Error estimation
6. Asymmetric distribution

The goodness of fit and error estimation can be overcome through the use of R^2 and the Hamilton test. We have proposed solutions for the other problems. The constrained thorough search method can solve the initial parameter dependence and partly address correlation problems. The RMC or m-RMC methods can address the initial parameter, the correlation and the asymmetric distribution problems. The number of data points is limited in EXAFS analysis; therefore, EXAFS measurements at low temperature are preferable to increase the number of data points. If polarization dependence can be applied to systems that involve impurities in a crystalline material or metal species adsorbed on a single crystal surface, then the number of data points is increased.[42] Multi-edge analysis is also useful.[19] Another possibility to increase the number of data points is to combine the data obtained from other methods such as Raman spectroscopy, transmission electron microscopy (TEM), X-ray diffraction (XRD) and density functional theory (DFT). However, it may not be so easy to include such information directly in curve fitting analysis. On the other hand, the m-RMC method can use such data directly because the real structure model can be shared by the other methods, so that calculated and observed data from the same sets of replica files can be compared for many characterizations. The goodness of fit R^2 obtained from each characterization technique can be summed with appropriate weight. The summed R^2 is used to find the minimum by randomly changing the common replica files in the Metropolis method. In this sense, structure determination for the system can be increasingly improved to be more precise. Even if EXAFS is just one characterization method, it is still a powerful technique in the analysis of non-crystalline materials under working conditions. Combination with other characterization methods is thus expected to empower EXAFS spectroscopy for applications in material science.

------- Acknowledgement ----------

We would like to express our thanks to the Photon Factory staff members, especially Prof. Masaharu Nomura, and SPring-8 staff members, especially Prof Tomoya Uruga for their continuous assistance. We are also indebted to Prof. Yasuhiro Iwasawa (The University of Electro-Communications) for fruitful suggestions. We are grateful to Dr. Yohei Umemura, Mr. Keisuke Fujikawa, and Mr. Yuya Iwasaki for their technical support. The work was supported by the Core Research for Evolutional Science and Technology (CREST) program, the New Energy and Industrial Technology Development Organization (NEDO) and the Japan Society for the Promotion of Science (JSPS; Grant-in-Aid for Scientific Research A, No. 20H00367).

Reference

1 D. E. Sayers, E. A. Stern, F. W. Lytle, *Physical Review Letters* **1971**, *27*, 1204-1207.

2 Y. Iwasawa, K. Asakura, M. Tada, *XAFS Techniques for Catalysts, Nanomaterials, and Surfaces*. Editor, Springer Nature, **2016**.

3 P. Lee, J. Pendry, *Physical Review B* **1975**, *11*, 2795-2811.

4 E. A. Stern, *Physical Review B* **1974**, *10*, 3027-3037.

5 E. D. Crozier, J. J. Rehr, R. Ingalls, in *EXAFS, XANES and SEXAFS*, ed. by D. C. Koningsberger, R. Prins, **1988**.

6 Y. Iwasawa, *X-ray absorption fine structure for catalysts and surfaces*. Editor, World Scientific, **1996**, Vol. 2.

7 U. Kashaboina, Y. Nishikawa, Y. Wakisaka, N. Sirisit, S. Nagamatsu, D. Bao, H. Ariga-Miwa, S. Takakusagi, Y. Inami, F. Kuriyama, A. L. Dipu, H. Ogihara, S. Iguchi, I. Yamanaka, T. Wada, K. Asakura, *Chemistry Letters* **2019**, *48*, 1145-1147.

8 K. K. Bando, T. Wada, T. Miyamoto, K. Miyazaki, S. Takakusagi, Y. Koike, Y. Inada, M. Nomura, A. Yamaguchi, T. Gott, S. Ted Oyama, K. Asakura, *Journal of Catalysis* **2012**, *286*, 165-171.

9 T. Wada, K. K. Bando, T. Miyamoto, S. Takakusagi, S. T. Oyama, K. Asakura, *Journal of Synchrotron Radiation* **2012**, *19*, 205-209.

10 T. Wada, K. K. Bando, S. T. Oyama, T. Miyamoto, S. Takakusagi, K. Asakura, *Chemistry Letters* **2012**, *41*, 1238-1240.

11 E. A. Stern, D. E. Sayers, F. W. Lytle, *Physical Review B* **1975**, *11*, 4836-4846.

12 K. Asakura, in *X-ray Absorption Fine Structure for Catalysts and Surfaces*, 1 ed., ed. by Y. Iwasawa, World Scientific, Singapore, **1996**, Vol. 2, pp. 33-58.

13 J. J. W. Cook, D. E. Sayers, *Journal of Applied Physics* **1981**, *52*, 5024-5031.

14 J. A. Victoreen, *Journal of Applied Physics* **1949**, *20*, 1141-1147.

15 W. H. McMaster, N. K. Del Grande, J. H. Mallett, J. H. Hubbell, **1969**.

16 J. J. Rehr, R. Albers, *Reviews of Modern Physics* **2000**, *72*, 621.

17 W. H. Press, S. A. Teukolsky, B. P. Flannery, W. T. Vetterling, *Numerical recipes in Fortran 77: volume 1, volume 1 of Fortran numerical recipes: the art of scientific computing*. Editor, Cambridge university press, **1992**.

18 K.-q. Lu, E. Stern, *Nuclear Instruments and Methods in Physics Research* **1983**, *212*, 475-478.

19 C. R. Bian, S. Suzuki, K. Asakura, L. Hei, N. Toshima, *The Journal of Physical Chemistry B* **2002**, *106*, 8587-8598.

20 E. Stern, *Phys.Rev.B* **1993**, *48*, 9825.

21 D. Kido, Y. Uemura, Y. Wakisaka, H. Ariga-Miwa, S. Takakuasgi, K. Asakura, *e-Journal of Surface Science and Nanotechnology* **2020**, *18*, 249-261.

22 F. W. Lytle, D. Sayers, E. Stern, *Physica B: Condensed Matter* **1989**, *158*, 701-722.

23 D. E. Sayers, in *Book Report of the International XAFS Society, Standards and Criteria Committee*, ed., ed. by Editor, **2000**, Chap. Chapter.

24 H. J. C. Berendsen, *A Student's Guide to Data and Error Analysis by Herman J. C. Berendsen*. Editor, CAMBRIDGE UNIVERSITY PRESS, **2015**.

25 G. Bunker, *Nuclear Instruments and Methods in Physics Research* **1983**, *207*, 437-444.

26 T. Yokoyama, N. Kosugi, A. Asakura, Y. Iwasawa, H. Kuroda, *Le Journal de Physique Colloques* **1986**, *47 C8*, 273-276.

27 S. Ohba, S. Sato, Y. Saito, K.-I. Ohshima, J. Harada, *Acta Crystallographica Section B* **1983**, *39*, 49-53.

28 W. C. Hamilton, *Acta Crystallographica* **1965**, *18*, 502.

29 E. D. Crozier, A. J. Seary, *Canadian Journal*

of Physics **1980**, *58*, 1388-1399.

30 E. D. Crozier, in *EXAFS*, ed. by T. a. Joy, **1982**.

31 I. D. Brown, *Chemical Society Reviews* **1978**, *7*, 359-376.

32 R. L. McGreevy, L. Pusztai, *Molecular Simulation* **1988**, *1*, 359-367.

33 S. J. Gurman, R. L. McGreevy, *Journal of Physics: Condensed Matter* **1990**, *2*, 9463-9473.

34 D. A. Keen, R. L. McGreebvy, *Nature* **1990**, *344*, 423.

35 S. J. Gurman, *Journal of Synchrotron Radiation* **1995**, *2*, 56-63.

36 W. Bras, R. Xu, J. D. Wicks, F. van der Horst, M. Oversluizen, R. L. McGreevy, W. van der Lugt, *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment* **1994**, *346*, 394-8.

37 L. Mikael, W. Kjartan Thor, G. M. P. Lars, *Journal of Physics: Condensed Matter* **2010**, *22*, 135001.

38 J. Timoshenko, A. Kuzmin, J. Purans, *Computer Physics Communications* **2012**, *183*, 1237-1245.

39 J. Timoshenko, A. Anspoks, A. Kalinko, A. Kuzmin, *Zeitschrift für Physikalische Chemie* **2016**, *230*, 551-568.

40 M. Harada, R. Ikegami, L. Kumara, R. Singgappulige, S. Kohara, O. Sakata, *RSC Advances* **2019**, *9* 29511-29521.

41 K. Fujikawa, H. Ariga, S. Takakusagi, H. Uehara, T. Ohba, K. Asakura, *e-Journal of Surface Science and Nanotechnology* **2014**, *12*, 322-329.

42 K. Asakura, in *Book Polarization-dependent Total Reflection Fluorescence Extended X-ray Absorption Fine Structure and its Application to Supported Catalysis*, ed., ed. by Editor, RSC publishing, City, **2012**, Vol. 24, Chap. Chapter, pp. 281-322.