Thorough Search Analysis of Extended X-ray Absorption Fine Structure Data for Complex Molecules and Nanomaterials Applications

Daiki Kido, a  Yohei Uemura, b, †  Yuki Wakisaka, a  Hiroko Ariga-Miwa, a  Satoru Takakuasgi, a  Kiyotaka Asakura a, ‡

a Institute for Catalysis and Division of Quantum Science and Engineering, Graduate School of Engineering, Hokkaido University, Kita 21-10, Sapporo, Hokkaido 001-0021, Japan 
b Institute for Molecular Science, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan 
† Present address: Laboratory for Environmental Chemistry, Paul Scherrer Institute, FLB 106 (PSI Ost), PSI, Forschungsstrasse 111, 5232 Villigen, Switzerland 
‡ Corresponding author: askr@cat.hokudai.ac.jp

Received: 9 June, 2020; Accepted: 21 July, 2020; Published: 1 August, 2020

Extended X-ray absorption fine structure (EXAFS) plays an important role in the surface science and nanotechnology to characterize the non-crystalline materials using the curve fitting (CF) analysis. However, the CF has problems such as correlation between the structural parameters, the dependence on initial parameters, and the limitation of degree of freedom when EXAFS is applied to the complex system. In this paper, we propose a thorough search (TS) method to solve these problems. We analyzed EXAFS data for molybdenum oxide ($\alpha$-MoO$_3$) using the TS method. $\alpha$-MoO$_3$ possesses a well-defined but complex local structure in which a central molybdenum (Mo) atom is surrounded by six oxygen (O) atoms. In CF analysis, the correlations of these six Mo–O bonds make it very difficult to derive reliable structural parameters from EXAFS data. In the TS analysis, the structural parameters regarded as a point ($P$) were surveyed thoroughly over a certain range. The goodness of fit was evaluated by $R$-factor. All $P$ with $R$-factors less than a certain value were accepted. The accepted points $P$ made a domain in which it was assumed that all points $P$ in the domain should occur with equal probability and consequently their averages were used as representative structural parameters. If multiple independent domains were obtained, they were all regarded as possible candidates and the TS analysis provided possible structural parameters. The feasibility and advantages of the TS method were compared with the CF analysis and the micro reverse Monte Carlo method.

Keywords: Extended X-ray absorption fine structure (EXAFS); Platinum; Molybdenum oxide; Data analysis

I. INTRODUCTION

Extended X-ray absorption fine structure (EXAFS) spectroscopy refers to the oscillation appearing from 50 to 1000 eV above an X-ray absorption edge. EXAFS is an element-specific spectroscopy and provides information about the local structure around the X-ray-absorbing atoms of materials without long-range order such as surfaces, nanomaterials and catalysts [1]. EXAFS spectra are usually analyzed with background removal and Fourier transformation (FT) into $r$-space, followed by non-linear least-square curve fitting (CF) analysis to search for the lowest $R$-factor [defined by Eq. (3) shown below] in $k$- or $r$-space [2]. The CF analysis provides structural parameters such as the coordination number ($N$), correction of the zero kinetic energy ($\Delta E_0$), the bond length ($r$), and the Debye-Waller factor ($\sigma^2$) for each coordination shell.

When an EXAFS spectrum of a complex system where several shells are present in an inseparably small range is
analyzed by the CF method, the following three problems arise.

1. The strong correlation between structural parameters makes the CF analysis unstable and prevents it from converging to give physically meaningful results [3, 4]. Even if the CF results give a reasonably low R-factor, it would be hard to exclude the other structural parameters.
2. The fitting results were dependent on the choice of the initial structural parameters.
3. The number of total parameters for CF analysis is limited by the finite degree of freedom determined by the Nyquist theory [5].

\[
\frac{2\Delta k \cdot \Delta r}{\pi} + \alpha, \quad \alpha = 0, 1, 2
\]  

(1)

Here, \(\Delta k\) is the fitting \(k\)-range and \(\Delta r\) is the FT \(r\)-range. Because typical \(\Delta k\) and \(\Delta r\) are 12 Å\(^{-1}\) and 1 Å, respectively, the degrees of freedom are eight to ten. Thus, the maximum number of coordination shells is two or three for every 1-Å interval.

The \(\alpha\)-molybdenum oxide (\(\alpha\)-MoO\(_3\)) has a distorted octahedral structure with five different bond lengths, as shown in Figure 1 [6]. Chun et al. [7] reported the CF analysis of EXAFS data for \(\alpha\)-MoO\(_3\) powder. It is still quite difficult to obtain the bond lengths of MoO\(_3\) correctly by the conventional CF method. In their paper, they measured polarization dependent and temperature dependent XAFS of \(\alpha\)-MoO\(_3\) single crystal and carefully chose the initial parameters. They fixed several parameters using the relationships between the bond length and the Debye-Waller factor to decrease the total number of fitting parameters. Finally, they obtained good results. However, the above-mentioned careful and particular experiments and analyses were necessary.

To overcome the problems associated with the CF analysis, Fujikawa et al. [8] developed the micro reverse Monte Carlo (m-RMC) method and successfully determined the structure of \(\alpha\)-MoO\(_3\) without any assumptions of initial structural parameters. The m-RMC is a kind of reverse Monte Carlo approach [9–14]. The reverse Monte Carlo approach tries to find the structure with the random walk of atoms in a large ensemble to seek the low R-factor using the Metropolis method. The m-RMC method uses a large number of ensembles of clusters, for which the EXAFS oscillations are calculated and determines the local structure [8].

Here, we propose a new analysis procedure for EXAFS data, which is called thorough search (TS) analysis. In TS analysis, EXAFS spectra and their R-factors are calculated thoroughly over a certain range of structural parameters. Although TS analysis is a time-consuming and inefficient method, the appropriate choice of constraints for fitting parameters makes it feasible to determine the local structures of materials, even those with complex structures like \(\alpha\)-MoO\(_3\). In this paper, the TS analysis is applied to two systems. The first is platinum (Pt) foil as a simple example to illustrate details of TS analysis. The second is \(\alpha\)-MoO\(_3\) as a more complex case. The advantages and disadvantages of the TS analysis are discussed through comparison with the CF analysis [7] and the m-RMC method [8].

II. EXPERIMENTAL

A. EXAFS measurements

A Pt L\(_3\)-edge EXAFS spectrum of the Pt foil was measured at the BL-12C beam line of the Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK-IMSS-PF) using a Si(111) double-crystal monochromator. The higher harmonics were rejected by a bent cylindrical focusing mirror. The Mo K-edge EXAFS spectrum of the \(\alpha\)-MoO\(_3\) powder was accumulated in a transmission mode at the BL-10B beam line of the Photon Factory using a channel cut Si(311) monochromator [7]. Both data can be downloaded from the XAFS database of the Institute for Catalysis, Hokkaido University (https://www.cat.hokudai.ac.jp/catdb/index.php?action=xafs_login_form&opnid=2), where the data for the Pt foil and \(\alpha\)-MoO\(_3\) have been uploaded as JXAFS-1203210001 and JXAFS-1206270007, respectively.

Figure 1: Structure of \(\alpha\)-MoO\(_3\) [6].
B. Data processing

Background removals were carried out using REX2000 [2, 15]. EXAFS oscillations $\chi(k)$ of the Pt foil and the α-MoO$_3$ powder were Fourier transformed in the range of 3.0–16.0 and 3.0–13.0 Å$^{-1}$, respectively, with a Hanning window function. The inverse FT data were analyzed by the CF and TS methods, independently. The inverse FT ranges were 2.0–3.0 and 1.0–2.0 Å for the Pt foil and the α-MoO$_3$ powder, respectively. Only the first neighboring atoms were considered for both cases. The CF analysis was carried out in the REX2000 and Larch programs using the following equation [1, 2, 15, 16],

$$\chi(k) = \sum_{i}^{M} \frac{S_0^2 N_i F_i(k)}{kr_i^2} \exp\left(-\frac{2r_i}{\lambda}\right) \exp(-2\sigma_i^2 k^2) \times \sin(2kr_i + \phi_i(k)), \quad (2)$$

where $S_0^2$, $F_i(k)$, $\phi_i(k)$, $\lambda$, $m$, $h$, and $E_0$ are the inelastic energy loss, backscattering amplitude and phase shift functions, inelastic mean free path, mass of an electron, X-ray energy, and absorption edge energy, respectively. The subscript $i$ indicates the shell number and $M$ is the total number of shells to be analyzed. $F_i(k)$ and $\phi_i(k)$ for the $i$th shell in Eq. (2) were calculated using the FEFF8.2 program (see SM 1 in Supplementary Material) [17].

Goodness of fit was evaluated using the R-factor for all results,

$$\text{R-factor} = \frac{\sum[k^n \cdot \chi_{\text{data}}(k) - k^n \cdot \chi_{\text{cal}}(k)]^2}{\sum[k^n \cdot \chi_{\text{data}}(k)]^2} \quad (3)$$

Here, $\chi_{\text{data}}(k)$ and $\chi_{\text{cal}}(k)$ are the observed EXAFS data and calculated values, respectively. The error for CF analysis was estimated by the Hamilton method [18]. The degree of freedom was calculated using Eq. (1) with $\alpha = 2$.

III. OVERVIEW OF TS ANALYSIS

We considered the structural parameter space in the TS and CF analyses. This parameter space was a multidimensional metric space consisting of structural parameters such as $N$, $\Delta E_0$, $r$, and $\sigma^2$, each of which corresponded to one axis in the parameter space. In other words, the dimensionality of the structural parameter space was equal to the number of structural parameters. A set of the structural parameters was represented by a point $P$ with components of $(N, \Delta E_0, r, \sigma^2)$. The point $P$ provided the EXAFS oscillation $\chi(k, P)$ and its R-factor. This parameter space had a function mapping from $P$ onto the R-factor axis expressed as $R = f(P)$, as depicted in Figure 2 in two-dimensional (2D) parameter space as an example. In the CF method, the initial point $P_0$ was provided first and then $P$ moved in the parameter space to follow the steepest (highest gradient) direction to search for the point $P'$ with the lowest $R$-factor as shown in Figure 2(a). The curve fitting routine in the Larch program and the REX2000 were equipped with the Levenberg-Marquardt algorithm to find $P'$ effectively and efficiently [19]. The CF method might reach local minima such as $P''$ and $P'''$ shown in Figure 2(a) instead of the global minimum $P'$. As the number of structural parameters increased, the shape of function $f$ became broader with several dips that might trap point $P$ at a local minimum. In this case, even though the CF analysis gave the converged value, it did not necessarily guarantee the obtained result was the unique and accurate solution.

In contrast, in TS analysis, all points $P$ were thoroughly surveyed to produce the whole R-factor map, as shown in Figure 2.
Figure 2(b). In other words, the overall picture of function f was obtained first. When the R-factor was less than a certain value [as shown by the black line in Figure 2(b)], all \( \mathcal{P} \) are accepted as candidate parameter points.

Each candidate parameter point \( \mathcal{P} \) was not an isolated point but instead formed an \( n \)-dimensional domain with other candidate points \( \mathcal{P} \). Histograms of each parameter were obtained, as illustrated in Figure 2(b). The average of each parameter in the domain was the representative structural parameter of the domain of interest, as indicated by arrows in Figure 2(b). The width of the histogram corresponded to the error or precision of the corresponding structural parameter. In more complex systems, several domains were obtained that include points \( \mathcal{P} \) that were below the R-factor threshold. The TS analysis gave all possible structures. The domains with the R-factor less than certain value were then explored more precisely by other methods. The TS analysis was supposed to involve higher cost (time and machine power) than the CF analysis. However, the real costs were not so high as we anticipated and the TS analysis provided more information than the CF analysis.

IV. RESULTS

A. TS analysis of EXAFS data for Pt foil

The strategy of the TS analysis is to survey a selected area in the structural parameter space thoroughly and then calculate the EXAFS oscillation and the R-factor of each point \( \mathcal{P} \). The TS analysis was carried out using a customized Python program borrowing many functions from the Larch library \([16]\). Figure 3 shows a flow chart of the TS analysis procedure including data processing. The survey range and step for EXAFS analysis of the Pt foil are shown in Table 1. A total of 33201 points \( \mathcal{P} \) were calculated. A domain was created that consisted of \( \mathcal{P} \) satisfying R-factor < 0.0052. The maximum R-factor (0.0052) was determined by the criterion that the null hypothesis “calculated spectrum is equivalent to that of the best fit one” could not be rejected at a significance level of 0.05. The domain was composed of eight points \( \mathcal{P} \) of the structural parameters. Next, the fine TS analysis was carried out for the confined survey area, as shown in Table 1. In this case, 1814 points \( \mathcal{P} \) satisfied R-factor < 0.0052. The average values of the structural parameters in the fine TS analysis were \( S_0^2 = 0.99 \pm 0.09 \), \( \Delta E_0 = 9.6 \pm 1.3 \) eV, \( r = 2.76 \pm 0.004 \AA \), and \( \sigma^2 = 0.0052 \pm 0.0003 \AA^2 \), respectively. Figure 4(a) shows the EXAFS oscillation provided by the average structural parameters together with the observed one. Both oscillations agreed well with an R-factor of 0.0018. These parameters determined by the TS analysis corresponded well with those of the CF results within errors and those obtained from m-RMC results, as shown in Table 2 \([8]\). Figure 4(b–e) presents histograms of the structural parameters \( S_0^2 \), \( \Delta E_0 \), \( r \), and \( \sigma \). In Figure 4(b–e), the width of each histograms corresponded to the error assessed from the related Gaussian distribution fit (shown as dashed curves).

B. TS analysis of EXAFS data for \( \alpha \)-MoO\(_3\)

\( \alpha \)-MoO\(_3\) has an orthorhombic-type crystal structure with
lattice constants $a = 3.96$ Å, $b = 13.86$ Å, and $c = 3.70$ Å [6].

$\alpha$-MoO$_3$ has six Mo–O bonds. Crystallographic data revealed the presence of five different bond lengths in $\alpha$-MoO$_3$ (1.67, 1.73, 2.195, 2.25, and 2.33 Å), as shown in Figure 1. The short bonds of 1.67 and 1.73 Å correspond to the M=O double bonds, while the bonds with a length of 1.95 Å are the M−O single bonds. The question is “can we determine six bond lengths by EXAFS?” Figure 5 shows the FT of the observed $\alpha$-MoO$_3$ EXAFS oscillation. The peaks appeared around 1.5 and 3.2 Å. The strong peak at 3.2 Å was assigned to the Mo–Mo bonds and not considered in this paper. The peak around 1.5 Å was assigned to the Mo–O bonds. This peak was split into two because of the interference of the EXAFS oscillations of several Mo–O bonds with different lengths.

The TS analysis of the $\alpha$-MoO$_3$ powder was carried out as follows. It was assumed that $\alpha$-MoO$_3$ possessed an octahedral structure so that the surveyed parameter space had dimensionality of 24 ($= 6 \times 4$). The number of parameters was greater than the degrees of freedom. Therefore, it was necessary to decrease the number of survey parameters. $S_0^2$ and $\Delta E_0$ were fixed to those of Na$_2$MoO$_4$ because these two parameters mainly depended on the bonding pair and valence state. Their transferability from the reference compound was often assumed and confirmed in the literature [7]. When $S_0^2$ and $\Delta E_0$ were roughly adjusted, the those obtained from the reference compound provide the least $R$-factor as shown in SM 2 in Supplementary Material.

The $\sigma^2$ values were calculated from the spring constants derived from Raman spectroscopy [7, 20] using an Equation of Motion (EM) method [21]. Because spring constants are related to bond lengths, the Debye-Waller factor can be estimated as a function of bond length. We assumed the linear relation described by Eq. (4), as shown in Figure S2 in Supplementary Material.

$$\sigma^2 = 0.0082r - 0.0124 \quad (4)$$

This process decreased the dimensionality from 24 to six. Next, the TS analysis was carried out in this subspace. The analysis was performed under the condition of $r_1 \leq r_2 \leq$.
The shortest bond was a leading factor in determining the EXAFS amplitude. Decreases with oscil-
lation. When the CF analysis was carried out with one influential structural parameter that determined the EXAFS
obtained, as shown in Table S2 (Supplementary Material).

\( r_3 \leq r_4 \leq r_5 \leq r_6 \) to avoid the calculation of duplicated structures. We carried out the TS analysis in two steps. The first step was a rough TS analysis with low resolution, which was followed by a fine TS analysis limited to the candidate parameter domains derived from the rough analysis. The rough TS analysis was performed with a survey step of 0.05 Å for each parameter in the range of 1.60–2.40 Å. Table 1 shows detail range of each parameter. The number of surveyed points \( \mathcal{P} \) was \( 7 \times 10^4 \) (74614). From these, 40 \( \mathcal{P} \) satisfied \( R\text{-factor} < 0.10 \). Figure 6 shows the distributions for the Mo–O bond lengths \( r_i \) (\( i = 1–6 \)), \( r_1, r_3, \) and \( r_6 \) were determined independently of the other bond lengths. \( r_1 \) appeared around 1.70 Å as a single sharp peak. \( r_3 \) and \( r_6 \) also consisted of single peaks around 1.95 and 2.35 Å, respectively. The peak widths of \( r_3 \) and \( r_6 \) were larger than that of \( r_1 \). Thus, \( r_1 \) could be determined with less error than \( r_3 \) and \( r_6 \). This meant that the shortest bond (\( r_1 \)) was the most influential structural parameter that determined the EXAFS oscillation. When the CF analysis was carried out with one shell, \( N, \Delta E_0, r, \) and \( \sigma^2 \) as fitting parameters, \( r = 1.68 \) Å was obtained, as shown in Table S2 (Supplementary Material).

The shortest bond was a leading factor in determining the EXAFS oscillation for following three reasons: 1) The EXAFS amplitude decreases with \( r \) according to \( 1/r^2 \) and the inelastic mean free path term described as \( \exp(-2r/\lambda) \) in

\[ \mathcal{P} = \{ P \mid r_1 \leq r_2 \leq r_3 \leq r_4 \leq r_5 \leq r_6 \} \]

\[ r_3 \leq r_4 \leq r_5 \leq r_6 \]

Table 3: Results of the TS analysis for the Mo K-edge EXAFS data for \( \alpha \)-MoO\(_3\). The average of each bond from the models which satisfied \( R\text{-factor} < 0.10 \) under the condition of \( r_1 \leq r_2 \leq r_3 \leq r_4 \leq r_5 \leq r_6 \).

| Parameter | Structure 1 | Structure 2 | Structure 3 |
|-----------|-------------|-------------|-------------|
| \( r_1 / \text{Å} \) | 1.65        | 1.70        | 1.70        |
| \( r_2 / \text{Å} \) | 1.75        | 1.96        | 1.70        |
| \( r_3 / \text{Å} \) | 1.94        | 2.00        | 1.91        |
| \( r_4 / \text{Å} \) | 1.97        | 2.25        | 1.95        |
| \( r_5 / \text{Å} \) | 2.30        | 2.30        | 1.98        |
| \( r_6 / \text{Å} \) | 2.34        | 2.34        | 2.23        |

Figure 7: (a) 3D mapping of \( R\text{-factor} \) distribution. Size and color indicate the frequency of the points \( (r_2, r_4, \) and \( r_5) \) including different parameters \( (r_1, r_3, \) and \( r_6) \). Three domains projected to three planes were found (shown in gray). Projections of the three planes for (b) \( r_2-r_4, \) (c) \( r_2-r_5, \) and (d) \( r_4-r_5). Color indicates frequency.

\[ \lambda \approx 2.31 \]
\[ \text{Table 4: Parameters of the fine TS analysis for } \alpha\text{-MoO}_3. \text{ Each domain is shown in Table 3. All bond lengths were searched in the searched range and steps.} \]

| Parameter | Structure 1 | Structure 2 | Structure 3 |
|-----------|-------------|-------------|-------------|
| \( S_0^2 \) | 1.00        | 1.00        | 1.00        |
| \( \Delta E_0 / \text{eV} \) | 0           | 0           | 0           |
| \( r_1 / \text{\AA} \) | 1.60–1.70 / 0.01 | 1.65–1.75 / 0.01 | 1.65–1.75 / 0.01 |
| \( r_2 / \text{\AA} \) | 1.70–1.80 / 0.01 | 1.90–2.04 / 0.02 | 1.65–1.75 / 0.01 |
| \( r_3 / \text{\AA} \) | 1.90–2.06 / 0.02 | 1.90–2.04 / 0.02 | 1.85–2.05 / 0.02 |
| \( r_4 / \text{\AA} \) | 1.90–2.06 / 0.02 | 2.15–2.45 / 0.03 | 1.85–2.05 / 0.02 |
| \( r_5 / \text{\AA} \) | 2.15–2.45 / 0.03 | 2.15–2.45 / 0.03 | 1.85–2.05 / 0.02 |
| \( r_6 / \text{\AA} \) | 2.15–2.45 / 0.03 | 2.15–2.45 / 0.03 | 1.85–2.05 / 0.02 |
| \( \sigma^2 / \text{\AA}^2 \) | 0.0082 r–0.0124 | 1.60–1.75 / 0.01 | 1.65–1.75 / 0.01 |

\( \Delta r \approx 2.30 \text{ \AA} \) and the third around \( r_3 \approx 1.70 \text{ \AA} \), \( r_4 \approx 1.95 \text{ \AA} \), and \( r_5 \approx 1.98 \text{ \AA} \). In other words, two or three bonds always appeared around 2.00 Å.

Figure 7(b) shows the positive correlation between \( r_2 \) and \( r_4 \). When \( r_2 \) was shorter than 1.80 Å, \( r_4 \) appeared near 2.00 Å, while when \( r_2 \) appeared near 1.95 Å, \( r_4 \) became longer than 2.25 Å. Weak positive correlations were also found between \( r_2 \) and \( r_5 \) or between \( r_4 \) and \( r_5 \). These domains corresponded to the three candidate structures named Structure 1, Structure 2, and Structure 3. Structure 1 was composed of \( r_1 \) and \( r_2 \) with a range of 1.60 to 1.75 Å, which corresponded to the Mo=O double bonds, while \( r_3 \) and \( r_4 \) appeared around 1.95 Å, corresponding to the Mo−O single bonds. Meanwhile, \( r_5 \) and \( r_6 \) were located around 2.20–2.40 Å. Similarly, Structure 2 had one Mo=O double bond and two Mo−O single bonds. Structure 3 had two Mo=O double bonds and three Mo−O single bonds.

Next, the fine TS analysis was carried out in each domain using smaller survey ranges and finer steps, as shown in Table 4. The numbers of accepted \( \mathcal{P} \) that satisfied R-factor < 0.10 were 5089, 2013, and 4114 out of about 3.5 × 10^5, 1.4 × 10^4, and 2.0 × 10^4 searching processes, respectively, in domains 1, 2, and 3, respectively. Figure 8 shows the histograms of bond lengths obtained from the fine TS analysis of the three domains. Figure S3 (Supplementary Material) displays the 3D frequency plots for \( r_2 \), \( r_4 \), and \( r_5 \). The colors of the dots corresponded to the frequencies. Table 5 lists the averaged bond lengths over each domain. The error of each parameter was estimated from the width of each histogram. Figure 9 compares the EXAFS oscillations obtained experimentally with the results of the TS analysis shown in Table 5. The bond lengths of Structure 1 had the structure corresponding to crystallographic data within the error bars, although the differences in the R-factors among the three structures were not large enough to claim the superiority of Structure 1 over the other structures based on the R-factor.

We then performed the CF analysis with six Mo−O bonds as fitting parameters. As expected, the results strongly depended on the initial structural parameters. When the CF analysis started with all bond lengths at 1.95 Å, it converged to 1.69, 1.98, 1.98, 2.27, 2.27, and 2.32 Å, corresponding to Structure 2. When the CF analysis started with six bond lengths of 1.69 Å, it converged to 1.16, 1.23, 1.34, 1.64, 1.99, and 2.27 Å with an R-factor of 0.13. Only when the CF analysis was started with bond lengths of 2 × 1.70 Å, 2 × 2.00 Å, and 2 × 2.20 Å did it converge as reported in the literature [6]. Therefore, the appropriate choice of the initial value was crucial to obtain accurate parameters using the CF analysis. For reference, two-shell fitting was carried out and it provided two different bond lengths of 1.71 and 1.98 Å, as shown in Table S2 (Supplementary Material).

V. DISCUSSION

A. Features of TS analysis

We could find the methods similar to the TS analysis in the paper, which were called as brute force approach [24–28] or grid search method [29]. These methods were adjuvantive estimation or the refinement of fitted parameters after curve fitting analyses. We used the TS analysis from the first without the CF analysis. Namely, we carried out the rough survey in the wide range and the fine survey in the limited range. In our TS analysis the structural parameters were surveyed thoroughly throughout a certain region in the \( n \)-dimensional space, as shown in Figure 2(b), where the sets of structural parameters were denoted as points \( \mathcal{P} \). The dimensions of parameters in the TS analysis are limited by the degrees of freedom. In the analysis of the Pt foil, each \( \mathcal{P} \) consisted of four parameters (\( S_0^2 \), \( \Delta E_0 \), \( r \), and \( \sigma^2 \)). The average values of the structural parameters were reported. These parameters obtained from the TS analysis corresponded well with those determined from the CF analysis. The number of structural parameters in the CF analysis of the Pt foil was much lower than the degrees of freedom. Consequently, the result of the CF analysis was reliable.

To evaluate the correlation of the degrees of freedom and reliability of the TS analysis, the TS analysis of the Pt foil...
was carried out using the EXAFS data in the limited and shorter $k$-range of $3.0−6.0$ Å$^{-1}$ with the four structural parameters, as shown in Table S4 (Supplementary Material). There were three degrees of freedom, which was less than the number of fitting parameters (4). The TS analysis determined the three parameters $S_0^2$, $\Delta E_0$, and $r$ with large

![Figure 8](image-url) Distribution of bond lengths from significance models that satisfied $R$-factor $< 0.10$ for around Structures 1, 2, and 3. Numbers indicate structures and letters indicate $r_1$ to $r_6$. Dashed lines show the fitting results using a Gaussian function.
errors, as illustrated in Figure S4 and Table S5 (Supplementary Material), and we found that the frequency of $\sigma^2$ increased with $\sigma^2$, as shown in Figure S4(d). The function $R = f(\mathcal{P})$ appeared to be a flat function around the minimum for these parameters. When the survey region was increased to that summarized in Table S6 (Supplementary Material), the two parameters of $N$ and $\sigma^2$ could not be determined by the TS method, as shown in Figure S5 (Supplementary Material). Consequently, we concluded that the number of degree of freedom was also important in the TS analysis. In other words, we can not obtain the candidates parameters domain when the number of survey parameters are more than degrees of freedom. The TS automatically prevents the excess of survey parameters over the degrees of freedom.

In the analysis of $\alpha$-MoO$_3$, the values of $S_0$ and $\Delta E_0$ were estimated using those of the reference compound Na$_2$MoO$_4$. In addition, the relationship between the Debye-Waller factor and the bond length was used to decrease the number of structural parameters to six. We now consider the results if the Debye-Waller factor are varied as free parameters. It cannot be evaluated directly based on the six independent bond length model because of the limitation of degree of freedom. However, the three-shell TS analysis with coordination numbers $2+ 2+ 2$ for each shell was carried out with three bond lengths and three Debye-Waller factors as survey parameters. As the results, the structure corresponding to the Structure 1 was obtained as shown in Table S7 (Supplementary Material). When other coordination number combination of three-shell analysis provided the Structures 2 ($1+2+3$) and 3 ($2+3+1$). The survey range was limited to 1.60−2.40 Å, which corresponded to the Mo−O bond lengths usually found in Mo oxide compounds [30]. As a result of parameter estimation, the TS analysis revealed three possible domains. It is an important feature of the TS that the domains were visualized in the 3D graph such as Figure 7. The structural parameters were surveyed more precisely by limiting the survey ranges for the candidate domains (the fine survey process), as shown in Figure 8. In $\alpha$-MoO$_3$, it was difficult to conclude which structure was superior to the others only from the EXAFS results or $R$-factor because these structures could not be distinguished only by the $R$-factor. Therefore, the combination of the TS analysis results with other experimental evidence or theoretical calculations is necessary. In the theoretical approach, molecular dynamics (MD) and density functional theory (DFT) calculations would be good choices to help select the appropriate candidate [31–42]. But, here, we applied the bond valence theory [43] to the three structure candidates identified by the TS analysis. The bond valence sums obtained for Structures 1, 2, and 3 were 6.00, 4.48, and 6.62, respectively. Because the Mo valence in $\alpha$-MoO$_3$ is 6+, Structure 1 was the most relevant structure. Structure 1 was the same structure as that of the crystallographic data for $\alpha$-MoO$_3$. Thus, the TS analysis can provide the several possible structure candidates with the determined six different bonds.

Figure 9: Comparison of experimental data with the precise TS analysis results for the Mo K-edge EXAFS data shown in Table 5. Results for Structures 1, 2, and 3 are shown in (a), (b) and (c), respectively.

Table 5: Results of the TS analysis of $\alpha$-MoO$_3$. The average of each bond from the models that satisfied $R$-factor < 0.10 under the condition $r_1 \leq r_2 \leq r_3 \leq r_4 \leq r_5 \leq r_6$, are shown. The results of curve fitting analysis [7] and m-RMC analysis [8] are also shown for comparison.

| Parameter | Structure 1 | Structure 2 | Structure 3 | CF analysis [7] | m-RMC [8] |
|-----------|-------------|-------------|-------------|----------------|-----------|
| $r_1$ / Å | 1.66 ± 0.01 | 1.69 ± 0.01 | 1.67 ± 0.01 | 1.65 ± 0.02 | 1.70 |
| $r_2$ / Å | 1.73 ± 0.02 | 1.96 ± 0.02 | 1.72 ± 0.02 | 1.75 ± 0.02 | |
| $r_3$ / Å | 1.94 ± 0.02 | 2.00 ± 0.02 | 1.90 ± 0.03 | 1.95 ± 0.02 | 1.97 |
| $r_4$ / Å | 1.97 ± 0.02 | 2.25 ± 0.03 | 1.95 ± 0.02 | |
| $r_5$ / Å | 2.28 ± 0.04 | 2.30 ± 0.03 | 2.00 ± 0.03 | 2.26 ± 0.02 | 2.37 |
| $r_6$ / Å | 2.36 ± 0.06 | 2.36 ± 0.04 | 2.27 ± 0.04 | 2.35 ± 0.02 | |
| $R$-factor | 0.0292 | 0.0369 | 0.0401 | - | - |

e-J. Surf. Sci. Nanotechnol. 18, 249–261 (2020) | DOI: 10.1380/ejssnt.2020.249
Next, we discuss the maximum value of $R$-factor for acceptance/rejection conditions. The maximum value of the $R$-factor was 0.10 in the case of $\alpha$-MoO$_3$. This value corresponded to the Hamilton test criterion with the null hypothesis that “the candidate calculated spectrum is equivalent to that of the best fit one” could not be rejected at a significance level of 0.05 [18]. In this paper, all the parameter points $\mathcal{P}$ that satisfied the condition $R$-factor < 0.10 were treated equally. If $\chi^2$-test could be used, where $\chi^2$ is defined as

$$\chi^2 = \frac{N_{\text{ dof}}}{N_{\text{ dof}} - N_{\text{ tsp}}} \sum_k \frac{[\chi_{\text{ data}}(k) - \chi_{\text{ cal}}(k, \mathcal{P})]^2}{[\epsilon(k)]^2}, \quad (5)$$

the occurrence probability of each point $\mathcal{P}$ can be decided exactly by the $\chi^2$-distribution. In Eq. (5), $N_{\text{ dof}}, N_{\text{ tsp}}, \chi_{\text{ data}}(k), \chi_{\text{ cal}}(k, \mathcal{P}), \epsilon(k)$, and $\epsilon(k)$ are the degree of freedom, number of TS analysis parameters, number of data points, observed and calculated EXAFS oscillations, and error, respectively. The calculation of $\chi^2$ requires errors $\epsilon(k)$, but it was difficult to determine the systematic error exactly. Thus, in this study, the $R$-factor was used instead of $\chi^2$ value. Although this assumption decreased the precision of the determined structural parameters, it decreased the risk of reporting the wrong values.

Finally, we should discuss the meaning of the histograms. Each histogram is proportional to the occurrence frequency of each structural parameter. Consequently, the average is the most probable value of the structural parameter and the width of the histogram corresponds to the error or reliability of the structural parameter. In the present study, the structural parameters were six bond lengths. The histograms looked like the radial distribution curve, as shown in Figure 6(g). The peak areas were 1.65, 2.15, and 2.25, which corresponded well to the coordination numbers of 1.60, 2.20, and 2.24 determined from the three-shell CF analysis, as shown in Table S3 (Supplementary Material). However, the histograms of the TS analysis are not the same as the radial distribution curve. That is, the histogram of the TS analysis is just the frequency of the structural parameter and not the real radial distribution curve.

B. Comparison of TS analysis with CF and m-RMC analyses

Because $\alpha$-MoO$_3$ has already been analyzed by the CF analysis [7] and the m-RMC method [8], the three analysis methods can be compared directly in the case of MoO$_3$. It is difficult to obtain meaningful results using the CF analysis for materials with complex local structures like $\alpha$-MoO$_3$ if the initial structural parameters are not appropriate. As mentioned in Section IV.B, when the CF analysis started with all bond lengths at 1.95 Å, it converged to the Structure 2. In addition, when the CF analysis started with the results of rough TS, each initial structure converges to the structures which corresponds to the results of fine TS as shown in Table S8 (Supplementary Material). Only when the CF analysis starts with the correct structural parameters, does it give an answer that corresponds well to the crystallographic data. This means that the results derived from the CF analysis are not necessarily accurate structures even if rather good $R$-factors are obtained in the case of complex structures. If the number of the fitting shells is decreased, the most influential structural parameters are obtained in the CF analysis as shown in Tables S2 and S3 (Supplementary Material). Note that the value is not equal to the average of all Mo–O bonds but the most influential shortest one. In the two-shell fitting, the CF analysis provided values of 1.72 and 2.09 Å with coordination numbers of 1.5 and 2.1, respectively. The result corresponded to two bond lengths, which mostly explained the EXAFS oscillation. Two-shell fitting was carried out in the analysis of Mo oxide catalysts, which thus corresponded to the shells with the largest and second largest contributions [44].

Unlike CF analysis, the TS and m-RMC methods do not require the exact initial structures, although the TS analysis requires the number of survey parameters or the dimensionality of the parameter space to be less than the degrees of freedom. What we assumed are the following three items: 1) The Mo oxide in 6+ has mainly distorted octahedral and sometimes tetrahedral. 2) The Mo–O bond lengths are present in the range of 1.6–2.4 Å. 3) The Mo–O bond length and spring constants have relations. The TS analysis provides several possible candidates structural parameters that reproduce the observed spectra well. This seems to be a disadvantage of the TS analysis because the CF analysis gives a unique solution with a low $R$-factor. However, the unique solution of the CF analysis is not always reliable for complex systems. In this paper, three possible structure candidates were identified by the TS analysis. This is beneficial to those who are not so familiar with the EXAFS analysis and may consider the obtained CF result as a unique solution. In addition, these candidates are visually shown like Figure 7. Of course, such visualization is allowed only when the correlation of parameters is less than 3 because it is impossible to show the six-dimensional graph. Even though there is a limitation in visualization, we can definitively say that there is no other possibility based on the EXAFS data. The final structure can be selected from the candidates by other experimental or theoretical methods. In this study, bond valence analysis of the three candidate structures allowed us to select the correct one consistent with crystallographic data. As mentioned above the histograms of the TS analysis provides the error of each parameter though the m-RMC method gives the radial distribution of bond lengths as shown in Figure S8 (Supplementary Material).

A disadvantage of the TS analysis compared to the CF analysis is that it requires more costs. Therefore, the CF analysis should be carried out for simple systems. The increased cost of the TS analysis compared to that of the CF analysis is within a tolerable range. The m-RMC method
requires more resources than the CF and TS analyses because the longer calculation time increases the reliability of the m-RMC approach.

We compare the precision of the three types of analyses. The precision of the CF and TS analyses are at the same level if the system is simple. In a complex system, the CF analysis may converge to an inaccurate local minimum. Generally, m-RMC analysis gives a worse R-factor than those obtained from the CF and TS analyses because the EXAFS oscillation is calculated based on the ensemble of a large number of model structures, which may happen to include inappropriate model structures [8].

Finally, the advantages and disadvantages of the TS analysis compared with the CF analysis are summarized. The important advantages are 1) to search the whole parameter space, or independent of the selection of initial parameters, 2) to visualize the distribution of the parameter space and their correlation, 3) to pick up the local minima to confirm that the obtained structure is unique or not in the parameter space, and 4) to automatically consider the limitation of degree of freedom. These advantages overcome the problems of the CF analysis, or parameter correlation and dependence of initial parameter, as mentioned in Section I. On the other hand, disadvantages are 1) to take much time to search the whole parameter space, and 2) to treat huge data. However, a recent development of computer hardware has lessened these disadvantages and may not be the practical problems. Although the CF analysis is useful for the simple system, the TS analysis especially shows its large ability for analyzing complex systems on surface and in nanomaterials.

C. Possibility of the practical application of TS analysis

The TS analysis can be effectively applied to nanoparticle structure analysis during reaction processes because the structure change occurs continuously during the reaction so that the survey range can be effectively narrowed. We are now analyzing WO₃ nanoparticle, which has also a complex structure, during the photoexcitation process [45–47]. Some of the parameters can be estimated from the structure in the ground state. Consequently, the excited-state structure can be revealed by the TS analysis. Alloy nanoparticle system is another application direction of the TS analysis. Nanoparticles of a binary alloy (MM’) have two bonds (M–M and M–M’) at similar positions. The eight parameters for two different bonds should be optimized for each edge at the same time. The correlation between M–M and M–M’ deteriorates the accuracy of the fitting results [48–56]. The TS analysis will provide possible candidate domains for the MM’ alloys. The final application is metal species deposited on the single crystal surfaces where we can study the polarization dependent EXAFS [57–59]. The 3D surface structures of single atom site or nano particles can be exactly determined with definite error bars [60–63].

We anticipate that the TS analysis will become a standard analysis method for complex systems. The TS analysis is still under development and we hope that this paper will invoke discussion about the scope of the TS analysis and spur the EXAFS investigation of complex systems using the TS analysis.

VI. CONCLUSION

In this paper, we proposed the TS method for analysis of EXAFS spectral data. The TS analyses of the Pt L₃-edge EXAFS data for the Pt foil and the Mo K-edge EXAFS data for the α-MoO₃ powder were presented as a simple case and an example of a complex structure, respectively. The TS analysis provided possible structural candidates that reproduced the experimental data well without the selection of accurate initial parameters. Two step TS analyses save the time and provide enough precisions in the results. In the case of α-MoO₃, three candidate structures were obtained and visualized in the 3D figure. Our results reveal that the TS analysis is a powerful tool to analyze EXAFS data obtained for complex systems on surfaces and in nanomaterials.

Acknowledgments

We thank the New Energy and Industrial Technology Development Organization (NEDO) Polymer Electrolyte Fuel Cell project for their financial support and JSPS for their Grant in Aid for Scientific Research A (20H00367). The EXAFS experiments were carried out with the approval of PF-PAC (Proposal Nos. 98G303 and 2016G035). The authors also acknowledge Prof. Matthew Newville (University of Chicago) for development of the Larch package and fruitful discussion. We thank Dr. Natasha Lund from Edanz Group for her kind proofreading of this manuscript.

Appendix

The details of parameter estimation and other results of the analysis is available in Supplementary Material at https://doi.org/10.1380/ejssnt.2020.249.

References

[1] Y. Iwasawa, K. Asakura, and M. Tada (Eds.), XAFS Techniques for Catalysts, Nanomaterials, and Surfaces (Springer, Cham, 2017).
[2] K. Asakura, in: X-ray Absorption Fine Structure for Catalysts and Surfaces, edited by Y. Iwasawa (World Scientific, Singapore, 1996) p. 33.
[3] H. Maeda, J. Phys. Soc. Japan 56, 2777 (1987).
[4] S. J. Gurman, J. Synchrotron Radiat. 2, 56 (1995).
[5] E. A. Stern, Phys. Rev. B 48, 9825 (1993).
[6] L. Kihlborg, Arkiv Kemi 21, 357 (1963).
[7] W.-J. Chun, K. Iijima, Y. Ohminami, S. Suzuki, and K. Asakura, J. Synchrotron Radiat. 11, 291 (2004).
[8] K. Fujikawa, H. Ariga, S. Takakusagi, H. Uehara, T. Ohba, and K. Asakura, e-J. Surf. Sci. Nanotechnol. 12, 322 (2014).
[9] R. L. McGreavy and L. Pusztai, Mol. Simul. 1, 359 (1988).

DOI: 10.1380/ejssnt.2020.249
[63] Q. Yuan, Y. Wakisaka, Y. Uemura, T. Wada, H. Ariga-Miwa, S. Takakusagi, K. Asakura, and S. R. Brankovic, J. Phys. Chem. C \textbf{122}, 16664 (2018).

All articles published on e-J. Surf. Sci. Nanotechnol. are licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0). You are free to copy and redistribute articles in any medium or format and also free to remix, transform, and build upon articles for any purpose (including a commercial use) as long as you give appropriate credit to the original source and provide a link to the Creative Commons (CC) license. If you modify the material, you must indicate changes in a proper way.

Published by The Japan Society of Vacuum and Surface Science