Bayesian sparse modeling of extended X-ray absorption fine structure to determine interstitial oxygen positions in yttrium oxyhydride epitaxial thin film

Hiroyuki Kumazoe,[1, a)] Yasuhiko Igarashi,[2, 3] Fabio Iesari,[4, 5] Ryota Shimizu,[6, 3] Yuya Komatsu,[6] Taro Hitosugi,[6] Daiju Matsumura,[7] Hiroyuki Saitoh,[8] Kazunori Iwamitsu,[9] Toshihiro Okajima,[4] Yoshiki Seno,[5] Masato Okada,[10, 11] and Ichiro Akai[1, b)]

[1)] *Institute of Industrial Nanomaterials, Kumamoto University, Kumamoto 860-8555, Japan*

[2)] *Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba 305-8573, Japan*

[3)] *Japan Science and Technology Agency, PRESTO, Saitama 332-0012, Japan*

[4)] *Aichi Synchrotron Radiation Center, Aichi 489-0965, Japan*

[5)] *Kyushu Synchrotron Light Research Center, Saga 841-0005, Japan*

[6)] *School of Materials and Chemical Technology, Tokyo Institute of Technology, Tokyo 152-8552, Japan*

[7)] *Materials Sciences Research Center, Japan Atomic Energy Agency, Hyogo 679-5148, Japan*

[8)] *Quantum Beam Science Research Directorate, National Institute for Quantum and Radiological Science and Technology, Hyogo 679-5148, Japan*

[9)] *Technical Division, Kumamoto University, Kumamoto 860-8555, Japan*

[10)] *Department of Complexity Science and Engineering, The University of Tokyo, Chiba 277-8561, Japan*

[11)] *Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science, Ibaraki, 305-0047, Japan*

(Dated: 4 November 2021)

This letter presents a Bayesian sparse modeling method to analyze extended X-ray absorption fine structure (EXAFS) data with basis functions built on two-body signals. This method does not require any structural model and allows us to evaluate regression coefficients proportional to the radial distribution functions of the respective elements and their errors and is very effective for analysis of EXAFS with weak absorption intensity and severe signal-to-noise ratio. As an application example, we used it to analyze EXAFS of an yttrium oxyhydride (YO_xH_y) epitaxial thin film. This EXAFS data shows weak absorption intensity and its signal-to-noise ratio becomes severe due to the small amount of X-ray absorption in the thin film sample. However, this approach revealed that the radial distance ratio of the second neighbor yttrium to the first neighbor oxygen coincides with that of a tetrahedral configuration. This result demonstrates that the interstitial oxygen position is tetrahedral in the YO_xH_y thin film.

The extended X-ray absorption fine structure (EXAFS)[1,2] is a primary method to study atomic-scale geometry around an element specified by the energy of the X-ray absorption edge. While the peak-fitting approach and the regularization method[3,4] have often been used to analyze EXAFS, deep-learning applications[5,6] to EXAFS would afford the key to improve the study of atomic-scale geometry. However, there is another alternative machine learning approach for measured EXAFS data analyses to avoid overfitting that attempts to reproduce even the noise. Although, of course, efforts for reducing measurement noise are important, the overfitting becomes a serious problem especially when the absorption intensity is weak, such as in thin films.

In this letter, a unified method of sparse modeling (SpM)[7–9] and Bayesian inference (BI)[10–12] is proposed to analyze EXAFS data.

Our method has the following advantages: (1) only the elemental species information is required for the analysis, but no structural information is needed; (2) the phase shift of the electron wave due to back-scattering is incorporated, so that the radial distance can be accurately estimated; (3) a model for k-dependent noise (due to k^w weight) is introduced, so that the noise intensity, which varies with k, can be properly evaluated; and (4) as a result, even EXAFS signals with a severe signal-to-noise ratio can be analyzed with high noise immunity.

We apply this method onto the EXAFS data measured on an yttrium oxyhydride YO_xH_y epitaxial thin film in order to determine the atomic-scale structure of the oxygen (O) atom around yttrium (Y) atom. This material is promising for switchable optical properties such as smart windows and optical sensors[13,14] because it exhibits a reversible light-induced insulator-to-metal transition. In addition, by fabricating thin films, the photoresponsive character is expected to be enhanced than that of polycrystalline thin films. Recent studies have shown that Y atom in YO_xH_y thin films belongs to a face-centered cu-

[a)] kumazoe@kumamoto-u.ac.jp
[b)] iakai@kumamoto-u.ac.jp

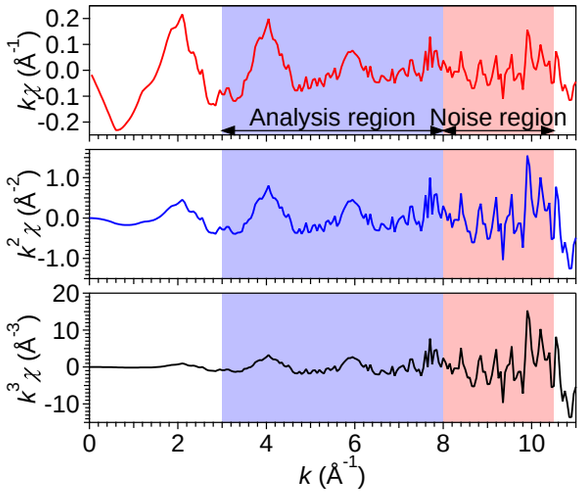


FIG. 1. Measured EXAFS data of YO_xH_y. Blue and red areas indicate the regions for analysis and noise estimation, respectively.

bic (fcc) lattice[15] and its lattice constant takes 5.34 Å[16] or 5.35 Å[15]. It is also reported that the oxygen coordinates around Gd atom for gadolinium oxyhydride thin films[17]. However, how the oxygen coordinates around Y atom for YO_xH_y thin films remain elusive.

Epitaxial thin films of YO_xH_y (111)/CaF_2 (111) substrates were grown using reactive magnetron sputtering. Details for sample preparation are provided in Ref. 18. Yttrium K-edge EXAFS measurements were performed at room temperature in air at BL14B1 in SPring-8. A Si (311) double-crystal monochromator and a 36-element solid-state detector were used for EXAFS observation with fluorescence mode. Rh-coated double mirrors were applied for cutting higher-order reflection in incident X-ray.

Figure 1 shows EXAFS data $k^w\chi(k)$ ($w = 1, 2, 3$) of an YO_xH_y epitaxial thin film. Abscissa is the wavenumber k of photoelectrons. Although the EXAFS data is analyzed with the absorption intensity $\chi(k)$ multiplied by k^w to emphasize the oscillation at larger k, the weight w has been empirically selected. As seen in Fig. 1, the signal fluctuations are large in the noise region at larger k because of weak absorption intensity due to a thin film sample and measurement in fluorescence mode: in such cases, the EXAFS analysis is difficult with conventional Fourier transform methods. As discussed later, the proposed method makes possible to determine which order w should be used to interpret the data as well as to study atomic-scale geometry even in such a tough analysis with high noise intensity.

Although the EXAFS oscillation of $\chi(k)$ comes from interferences of photoelectron waves, we consider two-body terms[19] which provide dominant EXAFS oscillation:

$$\chi_\alpha(k) = \sum_\beta \int_0^\infty \gamma^{(2)}_{\alpha\beta}(r,k) n_{\alpha\beta}(r)\, dr, \tag{1}$$

where α and β mean elements of photo-absorber and scattering atoms, respectively. $n_{\alpha\beta}(r)$ is the coordination number distribution of β as a function of the distance r from α, and $\gamma^{(2)}_{\alpha\beta}(r,k)$ contains the two-body multiple scattering term corresponding to the same geometrical configuration. Here, the wavenumber k is given by $k = \sqrt{2m(E - E_0 + \Delta E)}/\hbar$ as a function of photon energy E, where E_0, ΔE and m are the absorption edge energy, the energy offset of the theoretical zero for the k-space and electron mass, respectively. In our method, this ΔE can be optimized as described below. Each signal $\gamma^{(2)}_{\alpha\beta}(r,k)$ is written as an oscillating function, $A_{\alpha\beta}(r,k)\sin\Phi_{\alpha\beta}(r,k)$[19]. The amplitude $A_{\alpha\beta}(r,k)$ and the phase shift $\Phi_{\alpha\beta}(r,k)$ are calculated by GNXAS[20–22], in which a complex Hedin-Lundqvist potential[23] is used for inelastic loss effects in the framework of the multiple scattering theory.

We convert the integral of r in Eq. (1) into a Riemann sum with N subintervals whose each length is Δr:

$$\chi_\alpha(k) = \sum_\beta \sum_{j=1}^{N} A_{\alpha\beta}(r_j,k)\sin\Phi_{\alpha\beta}(r_j,k)\Delta r\, n_{\alpha\beta}(r_j). \tag{2}$$

Let $\mathbf{y}$ be $\{y_i = k_i^w \chi_\alpha(k_i)\,|\,i = 1,\ldots,p\}$, the EXAFS analysis becomes a linear regression problem with a given design matrix $M_{ij} = k_i^w A_{\alpha\beta}(r_j,k_i)\sin\Phi_{\alpha\beta}(r_j,k_i)\Delta r$. We introduce regression coefficients $\mathbf{x}_\beta = \{x_\beta(r_j)\}$ for an element β and $\mathbf{x} = \{\mathbf{x}_\beta\}$ for all elements in proportional to the coordination numbers.

This linear regression is an ill-posed problem. However, the L_1 regularization is particularly effective in EXAFS analysis of crystalline solid materials[7] because neighbor atoms will coordinate at discrete (sparse) distances from the photo-absorbers on the basis of symmetry and chemical structure. Therefore, we employ the least absolute shrinkage and selection operator (LASSO)[24] method, and a sparse solution $\mathbf{x}_{\mathrm{SpM}}$ is obtained by the formula at some regularization parameter λ (> 0):

$$\mathbf{x}_{\mathrm{SpM}} = \arg\min_{\mathbf{x}} \left[\frac{1}{2p}\|\mathbf{y} - M\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1\right]. \tag{3}$$

Fabio et al. used this LASSO method for EXAFS signals of Cu foil, Fe foil and cuprous oxide and determined λ by cross-validation[9]. As a result, they succeeded in obtaining peak positions of the radial distribution for their materials and structural parameters, like average distances and Debye–Waller factors without using any prior information about the structure.

However, the LASSO would select jointly correlated $x_\beta(r_j)$ when their features in M are highly correlated[24,25] and λ is often chosen heuristically. To overcome these

problems by BI[10], we introduced a binary vector $\mathbf{c}$ indicating nonzero components of $\mathbf{x}$ and applied linear regression $\mathbf{y} = M(\mathbf{c} \circ \mathbf{x}) + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is a noise superimposed on $\mathbf{y}$ and the operator $\circ$ means the Hadamard product between vectors. In BI, we maximize the posterior probability $P(\mathbf{c}\,|\,\mathbf{y})$, where it can be expanded as $P(\mathbf{c}\,|\,\mathbf{y}) = P(\mathbf{y}\,|\,\mathbf{c})P(\mathbf{c})/P(\mathbf{y})$ on the basis of Bayes' theorem[26]. To perform unbiased analysis, the prior probability $P(\mathbf{c})$ should be uniform. Consequently, $P(\mathbf{c}\,|\,\mathbf{y})$ is proportional to the likelihood $P(\mathbf{y}\,|\,\mathbf{c})$, and it is given by Eq. (4).

$$P(\mathbf{y}\,|\,\mathbf{c}) = \int P(\mathbf{y}\,|\,\mathbf{x},\mathbf{c})P(\mathbf{x}\,|\,\mathbf{c})\,d\mathbf{x}. \tag{4}$$

We use the Bayes free energy (BFE)[10,27,28], $F(\mathbf{c}) = -\ln P(\mathbf{y}\,|\,\mathbf{c})$ for model selection of $\mathbf{c}$.

To calculate Eq. (4), we introduce appropriate models for the prior probability $P(\mathbf{x}\,|\,\mathbf{c})$ and the likelihood $P(\mathbf{y}\,|\,\mathbf{x},\mathbf{c})$[10]. For $P(\mathbf{x}\,|\,\mathbf{c})$, we use a normal distribution $\mathcal{N}(0,\sigma_{\mathbf{x}_\beta})$ for the nonzero components in $\mathbf{x}_\beta$, where we incorporate the sparsity of $\mathbf{x}_\beta$ by being the mean to zero. On the other hand, the coordination numbers of nonzero components are expected to increase with r_j^2 because the volume of a spherical shell increases with r_j^2. So we set to $\sigma_{\mathbf{x}_\beta} = r_j^2(z_{\mathbf{x}_\beta 0})^{-1/2}$, where $z_{\mathbf{x}_\beta 0}$ of each element β will be estimated by BFE minimization.

For $P(\mathbf{y}\,|\,\mathbf{x},\mathbf{c})$, we assume a multivariate normal distribution $\mathcal{N}(M(\mathbf{c}\circ\mathbf{x}),\Sigma)$ for the superimposed noise in $\mathbf{y}$ and also introduce a model for its variance-covariance matrix Σ. Since we analyze $\{y_i = k_i^w\,\chi_\alpha(k_i)\}$, y_i is considered to include a uniform k-independent noise and a k-dependent noise in proportion to k^w, where the variance of the uniform noise and the standard deviation of the k-dependent noise are $1/z$ and $\sigma(k)$, respectively. Diagonal elements of Σ are the sum of $1/z$ and $\sigma(k)^2$, and off-diagonal elements are zero. The $1/z$ and $\sigma(k)^2$ are formulated as follows. First, the fluctuation in the noise region ($k \geq k_{\rm noise}$) of Fig. 1 is considered to come from noises because the EXAFS oscillation almost decays, and the root mean square deviation of the fluctuation is obtained as $\sigma_{\rm data}$. Subsequently, $\sigma(k)$ is given by $\sigma(k) = (k/k_{\rm noise})^w(\sigma_{\rm data}^2 - 1/z)^{1/2}$ for $k < k_{\rm noise}$. We also estimate $1/z$ by minimizing BFE.

By these modeling of $P(\mathbf{x}\,|\,\mathbf{c})$ and $P(\mathbf{y}\,|\,\mathbf{x},\mathbf{c})$, we can obtain the posterior probability and the BFE analytically. The regression coefficients $\mathbf{x}$ are optimized by the maximum a posteriori (MAP) estimation[10] under non-negative constraint. As the posterior probability $P(\mathbf{x}\,|\,\mathbf{y},\mathbf{c})$ is the Gaussian-type function, the reliability of $\mathbf{x}$ can be also estimated from the diagonal elements of the variance-covariance matrix of $P(\mathbf{x}\,|\,\mathbf{y},\mathbf{c})$.

The EXAFS oscillation shown in the analysis region of Fig. 1 is chosen for analysis because some X-ray absorption near edge structures would remain in the small k region. The analysis region is $k = 3.0$–8.0 Å^{-1}, and the region to obtain $\sigma_{\rm data}$ is $k = 8.0$–10.5 Å^{-1}. For radial distances r_j, we use a linear gird of $r_j = 2.000$–9.975 Å with the step, $\Delta r = 0.025$ Å.

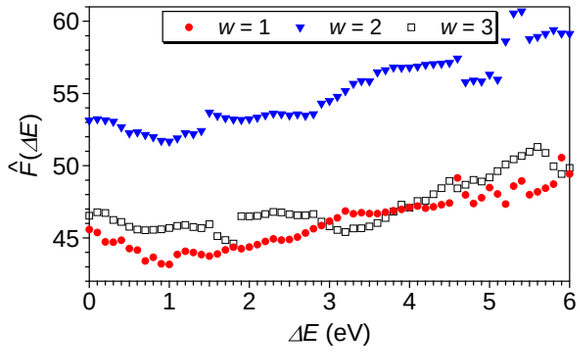


FIG. 2. BFE $\hat{F}(\Delta E)$ for $w = 1, 2, 3$.

In BI, the number of combinations for the binary vector $\mathbf{c}$ that represents the nonzero components of $\mathbf{x}$ is enormous, and their exhaustive search is impractical. So we first use Eq. (3) to get a sparse solution $\mathbf{x}_{\rm SpM}$, i.e. $\mathbf{c}_{\rm SpM}$. The sparsity control parameter λ is set to be in the range of $\lambda_{\max}$ to $\lambda_{\max}/100$ in a geometric sequence, where all $\mathbf{x}_{\rm SpM}$ are zero for $\lambda \geq \lambda_{\max}$[29].

Since the design matrix M contains the parameter ΔE to be optimized, we obtain the BFE as a function of ΔE for each w as follows: (i) Obtain $\mathbf{x}_{\rm SpM}$ ($\mathbf{c}_{\rm SpM}$) by solving Eq. (3) at given ΔE and λ and calculate the MAP solution $\mathbf{x}_{\rm BI}$ for the obtained $\mathbf{c}_{\rm SpM}$; (ii) Exhaustive search in the power set, $\mathcal{C} := {\rm Power}(\mathbf{c}_{\rm SpM})$ of $\mathbf{c}_{\rm SpM}$ through the minimization of the BFE: $\hat{\mathbf{c}}_{\rm BI}(\Delta E,\lambda) = \arg\min_{\mathbf{c}\subseteq\mathcal{C}} F(\mathbf{c};\Delta E,\lambda)$. The exhaustive search is performed under $\|\mathbf{c}\|_0 < 20$ for $\mathbf{c} \subseteq \mathcal{C}$ because of the computational cost; (iii) Finally, we obtain the minimized BFE $[\hat{F}(\Delta E) = \min_\lambda F(\hat{\mathbf{c}}_{\rm BI}(\Delta E,\lambda))]$ for each ΔE. In these procedures, the exhaustive search is necessary because the basis functions are not strictly orthogonal each other and there is concern about excess extraction of correlated components.

Figure 2 shows $\hat{F}(\Delta E)$ for $w = 1, 2, 3$ in the range of $\Delta E = 0.0 - 10.0$ eV with 0.1 eV step. The information criterion minimizing the BFE selects $\Delta E = 1.0$ eV and $w = 1$. Although the same ΔE is selected for $w = 2$, ΔE for $w = 3$ differs from others. This is because the noise intensity is overestimated for large exponent for high k-weighting. In Fig. 1, the amplitude of $k^3\chi(k)$ in the analysis region is smaller than them in the noise region.

The red curve in Fig. 3(a) is the regression curve for the selected model $\hat{\mathbf{x}}_{\rm BI}$ with $\Delta E = 1.0$ eV and $w = 1$, and it is found that the red curve explains appropriately the measured data indicated by open circles without overfitting. In EXAFS analysis, we have to pay attention to the contributions from three- or more-many body multiple scattering effects[30]. However, as seen from the scattered open circles in Fig. 3(a), it is considered that such contributions are not important in the case of severe signal-to-noise ratio because of a thin film sample. On the other

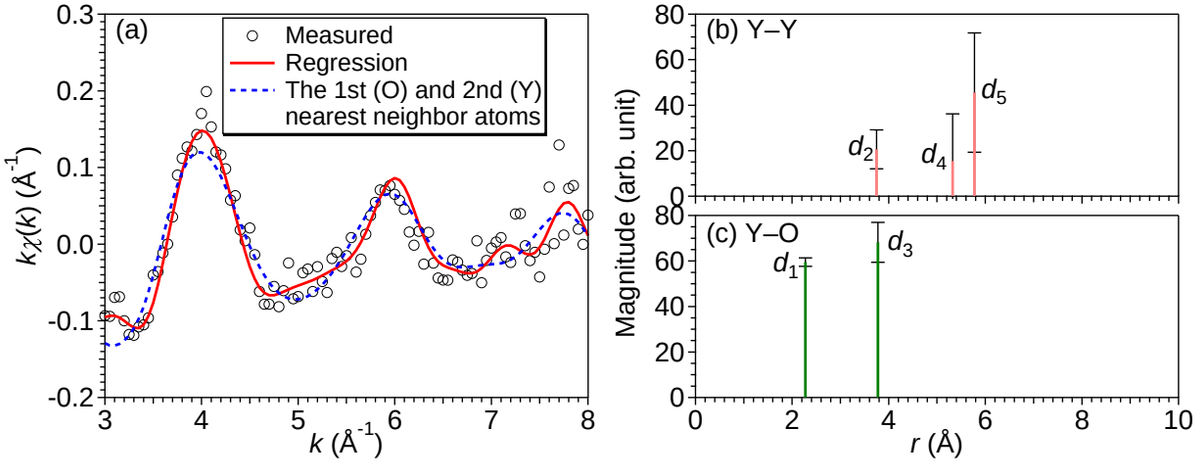


FIG. 3. (a) Open circles indicate the measured EXAFS data $k\chi(k)$. Red and blue curves are the regression results with $\hat{\mathbf{x}}_{\mathrm{BI}}$ at $\Delta E = 1.0$ eV and only with the first and second nearest neighbor atoms in the $\hat{\mathbf{x}}_{\mathrm{BI}}$, respectively. (b) and (c) Magnitude of the regression coefficients for Y and O atoms, respectively, as functions of r, where the coefficients are labeled by d_i in order of distance of the photo-absorber Y atom.

hand, the blue-dashed curve is drawn only by components of the two nearest neighbor atoms (O: 2.275 Å, Y: 3.750 Å) and presents a good overview of the EXAFS oscillation. In our method, the elements can be treated separately as β in Eq. (1), and the radial distribution of each element can be obtained separately as shown in Figs. 3(b) and (c), which are for Y and O atoms, respectively. Error bars show the reliability estimated from the variance-covariance matrix of $P(\mathbf{x}\,|\,\mathbf{y}, \mathbf{c})$. Since we do not consider the effect of the reduction factor S_0, the regression coefficients will be underestimated. Nevertheless, the coefficients take large values. This is because that thermal effect is excluded by the exhaustive search. As the EXAFS measurements were performed at room temperature, each peak in Figs. 3(b) and (c) should be distributed with a finite width near the equilibrium atomic position. However, it is very difficult to estimate the width of the radial distribution function from the data with a severe signal-to-noise ratio. Because the exhaustive search excludes the correlated coefficients and the contributions of the thermal effect due to a tradeoff between excluding correlated coefficients and considering the thermal effect. As a result, the coefficients stand only at the equilibrium position despite the finite temperature, and become large.

In the used basis functions, the phase shift due to photoelectron wave scatterings are incorporated. Therefore, the distance d_i for the ith neighbor atom can be evaluated correctly, and we can estimate the atomic-scale three dimensional structure based on the distance ratio between the neighbor atoms. When we focus on Y atoms in Fig. 3(b), $d_4/d_2 = 5.325/3.750 \approx 1.420$ is very close to $\sqrt{2} \approx 1.414$ of that characterizes the fcc structure. This result confirms a previous work[15] that Y atom takes the fcc structure. Additionally, a lattice constant of fcc is also obtained from $\sqrt{2}d_2$ and d_4, and their weighted-mean value with the weights of the inverse squares of the error bars in Fig. 3(b) is 5.306 ± 0.025 Å. It is slightly smaller than the previous studies[15,16] (5.34−5.35 Å), and this difference might come from the epitaxial growth.

The coordination structure of O atoms around Y atom can be determined from the ratio d_2/d_1, where there are two candidates for the interstitial O-sites: tetrahedral and octahedral sites. If the interstitial sites form a regular tetrahedron, the ratio d_2/d_1 is $\sqrt{8/3} \approx 1.633$. On the other hand, d_2/d_1 for octahedral becomes $\sqrt{2} \approx 1.414$. In our analysis, $d_2/d_1 \approx 1.648$ $(= 3.750/2.275)$ is quite close to the tetrahedral one, and we can conclude that the interstitial O sites in $\mathrm{YO}_x\mathrm{H}_y$ are tetrahedral. A similar result has been obtained in a previous study[17] for EXAFS of Gadolinium oxyhydride thin films, in which the conventional Fourier transform was used with a structure model predicted by the X-ray diffraction. In contrast, the advantage of our method is obvious, since it is based on the accurate atomic distances compensated with the phase shift and any structure models are not required.

In summary, we proposed a Bayesian SpM for analysis of EXAFS data with GNXAS basis functions incorporating two-body multiple scatterings. Model selection for the sparse solution of the regression coefficients proportional to the radial distribution function was performed by BFE minimization through the exhaustive search within nonzero components obtained by L_1 regularization. Using this method, we succeeded in determining that the sublattice of Y atom is fcc and the interstitial O position is tetrahedral in the $\mathrm{YO}_x\mathrm{H}_y$ epitaxial thin film.

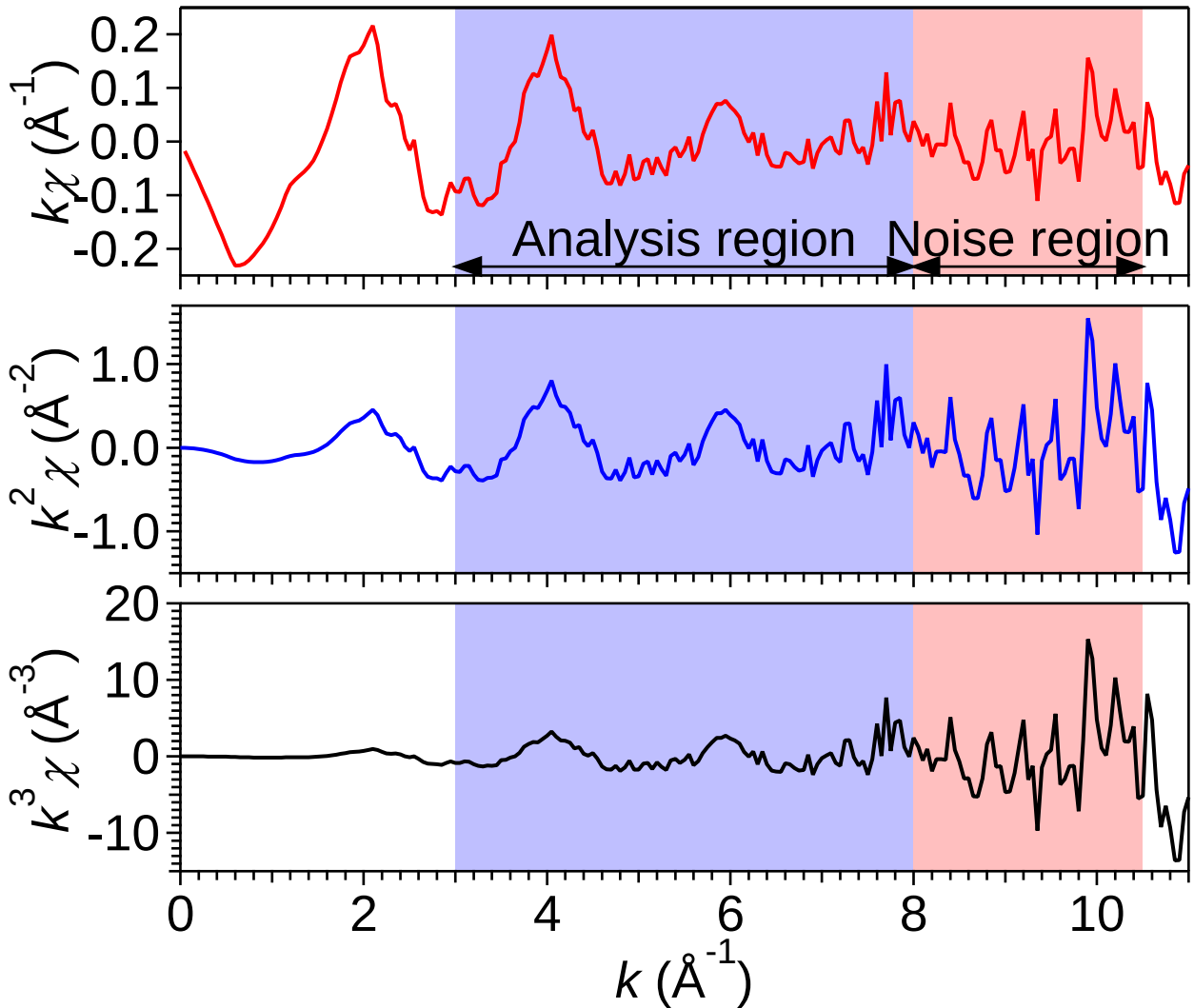
ACKNOWLEDGMENTS

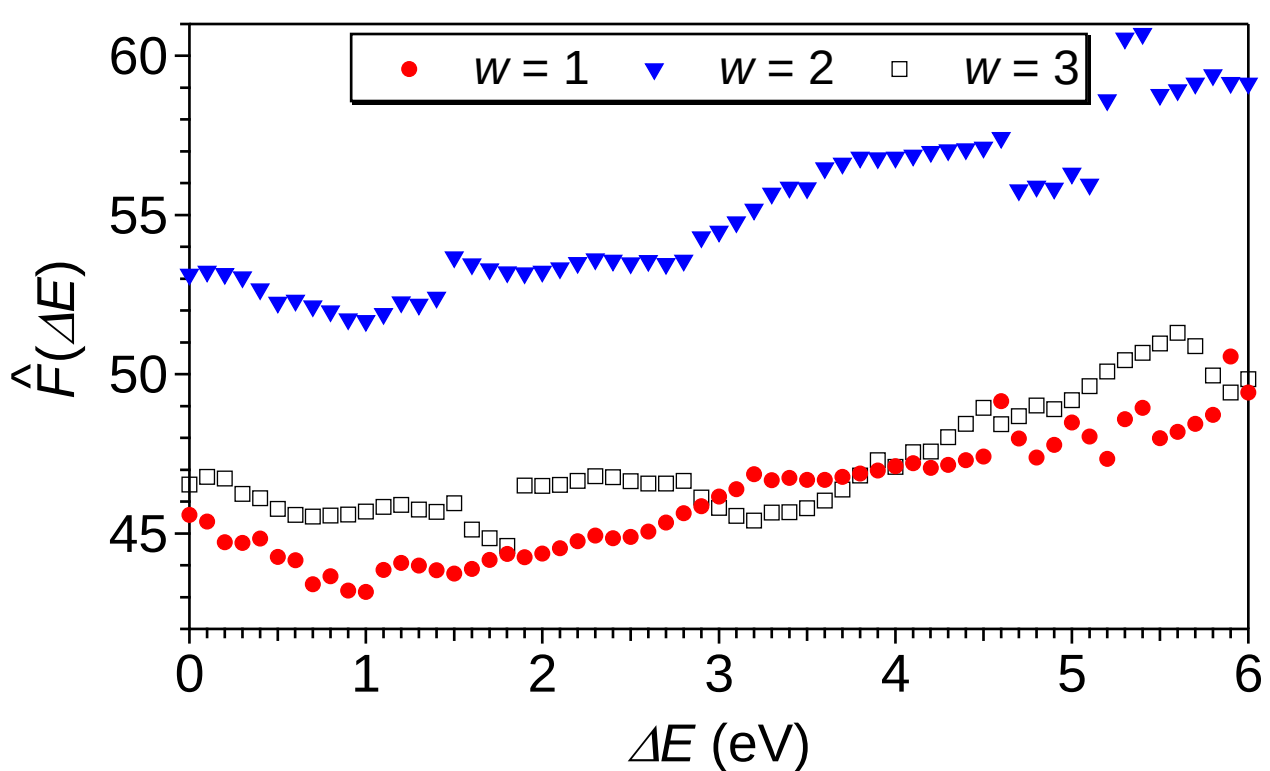
The synchrotron radiation experiments were performed using a JAEA experimental station at JAEA (QST) beamline BL14B1, SPring-8, with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2020A3648); The JAEA Advanced Characterization Nanotechnology Platform under the remit of "Nanotechnology Platform" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan (Grant No. JPMXP09A20AE0006). This work was supported by JST CREST: JPMJCR1861 and JPMJCR1761, Japan, JST PRESTO: JPMJPR17N2 and JPMJPR17N6, Japan, JSPS KAKENHI: JP19H02596, JP19H04689, JP18H05513, and JP18H05514, Japan, and AGC Inc.: KC31AGC08p.

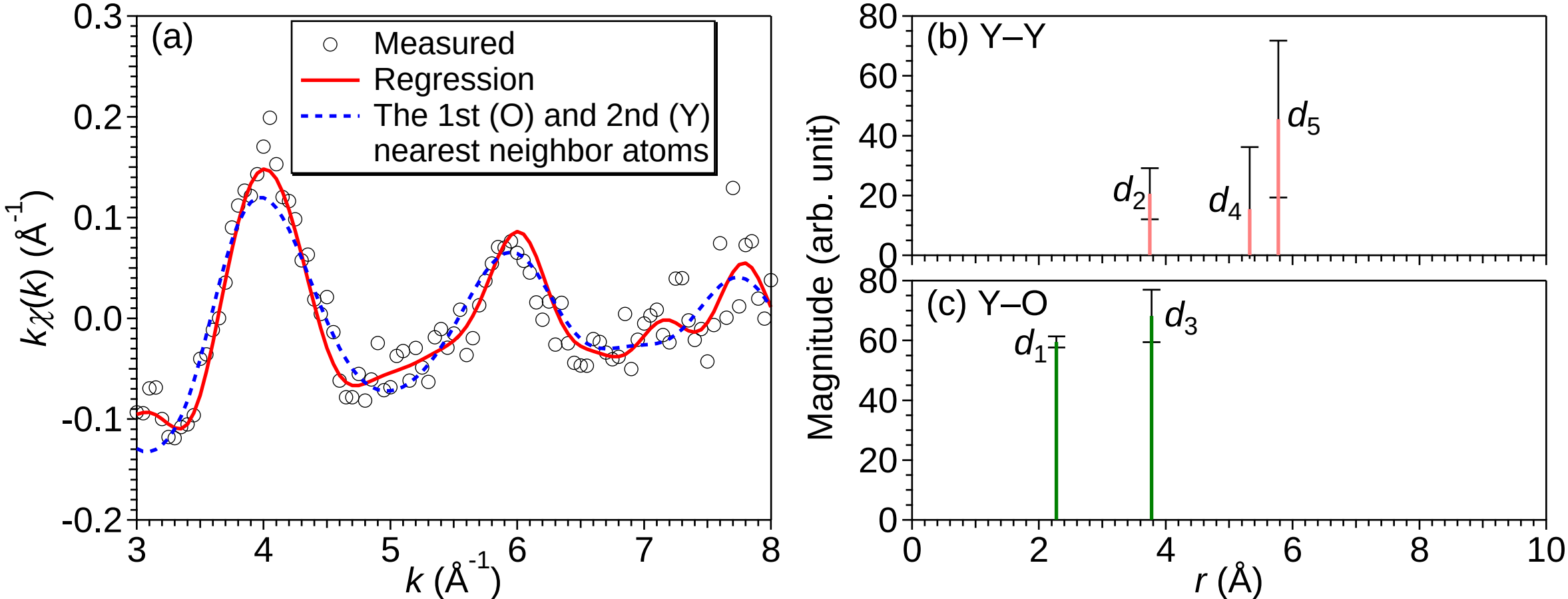
DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

[1]B. Teo, *EXAFS: Basic Principles and Data Analysis*, Inorganic Chemistry Concepts (Springer Berlin Heidelberg, 1986).
[2]G. Bunker, *Introduction to XAFS* (Cambridge University Press, 2009).
[3]Y. A. Babanov, V. V. Vasin, A. L. Ageev, and N. V. Ershov, physica status solidi (b) **105**, 747 (1981).
[4]N. V. Ershov, A. L. Ageev, V. V. Vasin, and Y. A. Babanov, physica status solidi (b) **108**, 103 (1981).
[5]J. Timoshenko, A. Anspoks, A. Cintins, A. Kuzmin, J. Purans, and A. I. Frenkel, Phys. Rev. Lett. **120**, 225502 (2018).
[6]F. Iesari, H. Setoyama, and T. Okajima, Symmetry **13**, 1070 (2021).
[7]I. Akai, K. Iwamitsu, Y. Igarashi, M. Okada, H. Setoyama, T. Okajima, and Y. Hirai, J. Phys. Soc. Jpn. **87**, 074003 (2018).
[8]H. Setoyama, I. Akai, K. Iwamitsu, Y. Miyata, S. Yakura, Y. Igarashi, M. Okada, and T. Okajima, J. Phys. Soc. Jpn. **89**, 074602 (2020).
[9]F. Iesari, H. Setoyama, Y. Igarashi, M. Okada, H. Kumazoe, K. Iwamitsu, I. Akai, Y. Seno, and T. Okajima, arXiv/physics.data-an , 2104.01805 (2021).
[10]Y. Igarashi, F. Iesari, H. Setoyama, T. Okajima, H. Kumazoe, I. Akai, and M. Okada, arXiv/physics.data-an , 2105.02341 (2021).
[11]K. Nagata, S. Sugita, and M. Okada, Neural Netw. **28**, 82 (2012).
[12]I. Akai, K. Iwamitsu, and M. Okada, J. Phys.: Conf. Ser. **1036**, 012022 (2018).
[13]J. N. Huiberts, R. Griessen, J. H. Rector, R. J. Wijngaarden, J. P. Dekker, D. G. de Groot, and N. J. Koeman, Nature **380**, 231 (1996).
[14]D. Moldarev, M. V. Moro, C. C. You, E. M. Baba, S. Z. Karazhanov, M. Wolff, and D. Primetzhofer, Phys. Rev. Mat. **2**, 115203 (2018).
[15]T. Mongstad, C. Platzer-Björkman, S. Karazhanov, A. Holt, J. Maehlen, and B. Hauback, J. Alloys. Compd. **509**, S812 (2011).
[16]C. C. You, T. Mongstad, E. S. Marstein, and S. Z. Karazhanov, Materialia **6**, 100307 (2019).
[17]G. Colombi, S. Cornelius, A. Longo, and B. Dam, J. Phys. Chem. C **124**, 13541 (2020).
[18]R. Shimizu, H. Oguchi, and T. Hitosugi, J. Phys. Soc. Jpn. **89**, 051012 (2020).
[19]A. Filipponi, J. Phys.: Cond. Mat. **6**, 8415 (1994).
[20]A. Di Cicco, *GNXAS Extended Suite of Programs for Advanced X-ray Absorption Data-analysis: Methodology and Practice* (TASK publishing, Gdansk, Poland, 2009).
[21]A. Filipponi, A. Di Cicco, and C. R. Natoli, Phys. Rev. B **52**, 15122 (1995).
[22]A. Filipponi and A. Di Cicco, Phys. Rev. B **52**, 15135 (1995).
[23]L. Hedin and B. I. Lundqvist, J. Phys. C: Solid State Phys. **4**, 2064 (1971).
[24]R. Tibshirani, J. Roy. Stat. Soc. B **58**, 267 (1996).
[25]H. Zou and T. Hastie, J. Roy. Stat. Soc. B **67**, 301 (2005).
[26]T. Bayes and R. Price, Phil. Trans. Roy. Soc. **53**, 370 (1763).
[27]Y. Igarashi, H. Takenaka, Y. Nakanishi-Ohno, M. Uemura, S. Ikeda, and M. Okada, J. Phys. Soc. Jpn. **87**, 044802 (2018).
[28]Y. Mototake, Y. Igarashi, H. Takenaka, K. Nagata, and M. Okada, J. Phys. Soc. Jpn. **87**, 114004 (2018).
[29]J. H. Friedman, T. Hastie, and R. Tibshirani, J. Stat. Software **33**, 1 (2010).
[30]I. Jonane, K. Lazdins, J. Timoshenko, A. Kuzmin, J. Purans, P. Vladimirov, T. Gräning, and J. Hoffmann, J. Synchrotron. Rad. **23**, 510 (2016).


0.2
0.1
0.0
-0.1
-0.2
kχ (Å-1)
Analysis region
Noise region
1.0
0.0
-1.0
k2 χ (Å-2)
20
10
0
-10
k3 χ (Å-3)
0
2
4
6
8
10
k (Å-1)


w = 1
w = 2
w = 3
60
55
50
45
0
1
2
3
4
5
6
$\hat{F}(\Delta E)$
ΔE (eV)


(a)
Measured
Regression
The 1st (O) and 2nd (Y) nearest neighbor atoms
$k\chi(k)$ (Å^{-1})
k (Å^{-1})
(b) Y–Y
d_2
d_4
d_5
(c) Y–O
d_1
d_3
Magnitude (arb. unit)
r (Å)