

ℓ_1 正則化と交差検定による EXAFS のスパースモデリング

宮田祐暉^A, 岩満一功^B, 瀬戸山寛之^C, 岡島敏弘^C,
五十嵐康彦^{D,E,F}, 岡田真人^{E,F}, 赤井一郎^{G,C}

^A 熊大院自然, ^B 熊大理, ^C 九州シンクロトロン研究セ, ^D JST さきがけ,
^E 東大院複雑理工, ^F NIMS MaDIS, ^G 熊大パルス研,

Sparse Modeling of Extended X-Ray Absorption Fine-Structure Spectrum by ℓ_1 -Regularization and Cross Validation

Y. Miyata^A, K. Iwamitsu^B, H. Setoyama^C, T. Okajima^C,
Y. Igarashi^{D,E,F}, M. Okada^{E,F}, I. Akai^{G,C}

^AGSST, Kumamoto Univ., ^BFac. Sci., Kumamoto Univ., ^CSAGA-LS, ^DJST, PRESTO,
^EDept. Complexity Sci. Eng. Univ. Tokyo, ^FMaDIS, NIMS, ^GIPPS, Kumamoto Univ.

Abstract

Measurements of extended X-ray absorption fine structures (EXAFSs) are an important method to research atomic-scale micro structure and structural fluctuation nearby an atom selected by X-ray absorption edge. We proposed a sparse modeling (SpM) method to analyze the EXAFS spectra using simplified basis function based on single scattering approximation. By this SpM method, we realized to estimate a Debye-Waller factor by cross validation method, and also succeeded in obtaining a radial structure function showing sparsity on atomic distance, where the sparsity is based on coordinations of neighbor atoms reflecting on chemical and crystal structures.

1 はじめに

広域 X 線吸収微細構造 (EXAFS: Extended X-ray Absorption Fine Structure)^{1, 2, 3, 4)} 計測は、原子種によって X 線吸収端エネルギーが異なることを利用して原子を選択した上で、その原子近傍の近距離構造を解析する重要な方法である。従来 EXAFS スペクトルの解析では、EXAFS 振動をフーリエ変換 (FT) して、動径構造関数 (RSF: Radial Structure Function) を得ていた (FT-RSF)。しかし非晶質物質や化学反応過程では、原子の近距離構造に関する事前情報がなく、近距離構造の構造ゆらぎを表すデバイ・ワラー因子 (σ_{DW}) の解析が困難であった。

本論文では、注目原子に対し近接原子は、本来、物質の化学構造や結晶構造によって決まる特定の原子間距離で配位、つまり原子間距離に対してスパースに配位することに注目し、EXAFS スペクトルのスパースモデリング (SpM: Sparse Modeling)⁵⁾ 解析を行なった。その結果、マイクロ構造の事前情報を必要とせず、データだけからデバイ・ワラー因子とスパースな動径構造関数を得ることに成功した。

2 EXAFS

2.1 解析対象データ

EXAFS スペクトル解析において、我々の開発した SpM 解析の妥当性を実証するために、SAGA-LS の BL11 を使用して標準試料の銅の EXAFS スペクトルを

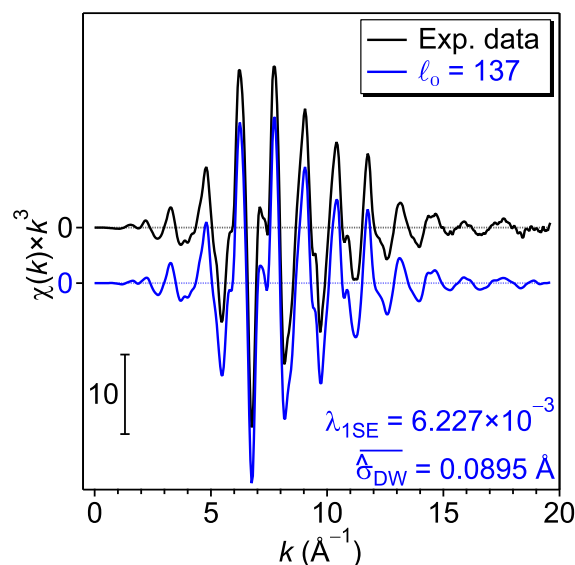


図 1: 解析対象データと再現データ

計測した。図 1 の黒線が室温で計測された EXAFS 振動 $y(k)$ で、横軸は光電子波の波数 k である。この $y(k)$ は、Athena⁶⁾ を使用し測定された X 線吸収スペクトルから EXAFS 振動を抽出し、大きな k での振動構造を強調するため k^3 をかけた^{1, 3)} ものである。

2.2 一体散乱近似

EXAFS 振動には、近接する複数の原子による散乱を経た多重散乱の信号も重畳するが、多重散乱過程は電子波の平均自由行程と比較して経路長が長くなるた

め、一般的に EXAFS 振動への寄与は顕著ではない¹⁾。そこで我々は、EXAFS スペクトルの主要な振動構造を支配し、(1) 式で定式化される一体散乱過程^{1,3)}をベースに SpM 法を適用する。

$$y(k) = \chi(k)k^3 \propto \sum_j N_j t_j(k) \frac{k^2}{R_j^2} \exp \left\{ -2 \left[k^2 \sigma_j^2 + \frac{R_j}{\Lambda(k)} \right] \right\} \times \sin[2kR_j + \delta_j(k)] \quad (1)$$

ここで N_j は原子間距離 R_j で配位する近接原子数、 σ_j は R_j の位置にある原子のデバイ・ワラー因子^{1,2)}である。 $t_j(k)$, $\Lambda(k)$, $\delta_j(k)$ は、後方散乱振幅、電子波の平均自由行程、位相シフトである。

(1) 式の $\exp[-2R_j/\Lambda(k)]$ の項は、平均自由行程以遠の R_j で干渉現象が抑制されることを表すもので、この項のため、FT-RSF では、 R_j の増加に伴い構造ピーク強度が減少してしまう。よって正しい RSF を得るためには、平均自由行程 $\Lambda(k)$ が推定できるのであれば、補償すべきである。

近接原子は、化学構造や結晶構造を反映して、原子間距離 R_j においてスパースな位置で配位する。しかしフーリエ変換で得られる FT-RSF は、 R_j の連続関数として動径構造関数を評価し、(1) 式中の k^2 項と $\exp(-2k^2\sigma_j^2)$ 項によって、ピークにスペクトル幅を与えてしまう。 k^2 項は大きな k の振動構造を強調するための項であり、図 1 の $k = 0 \sim 7 \text{ \AA}^{-1}$ の領域でみられるように EXAFS 振動振幅が強くなる特徴をもたらす。一方、デバイ・ワラー因子 σ_j を含む $\exp(-2k^2\sigma_j^2)$ 項は、原子間距離の熱的・構造的揺らぎや原子の可動性によって、電子波の干渉が抑制されるデコヒーレンス効果を示し、図 1 の $k > 8 \text{ \AA}^{-1}$ の領域で EXAFS 振動振幅が減衰する特徴をもたらす。

この様に EXAFS 振動の包絡関数に重要な物性情報を含んでいるにもかかわらず、FT-RSF では、振幅が変化しない正弦・余弦関数でモード展開を行っている。動径構造関数におけるスパース解を得るためには、 k^2 項や $\exp(-2k^2\sigma_j^2)$ 項を振動関数に組み込んだ基底関数を用いなければならない。

2.3 一体散乱近似の単純化基底関数

そこで我々は、(2) 式を用いて EXAFS 振動のスパースモデリングを行った。

$$y(k) = \sum_j \frac{k^2}{R_j^2} \exp \left[-2 \left(k^2 \sigma_{\text{DW}}^2 + \frac{R_j}{\Lambda} \right) \right] \times (a_j \sin 2kR_j + b_j \cos 2kR_j) \quad (2)$$

(2) 式で用いる基底関数は、平均自由行程 Λ を補償する項、 k^2 項、デバイ・ワラー因子項を含んでいるが、 $\Lambda(k)$ の k 依存性とデバイ・ワラー因子 σ_j の R_j 依存性を無

視し、それぞれ Λ と σ_{DW} の一定値と簡略化⁷⁾した。この簡略化は、 k によって変化する $\Lambda(k)$ の内、EXAFS 振動の振動振幅が一番顕著な波数 k を持つ電子の平均自由行程の値を選択したことになる。一方デバイ・ワラー因子では、 R_j に依存する σ_j の内、EXAFS 振動の主成分となっている原子間距離 R_j (一般的に最近接) の原子のデバイ・ワラー因子を推定することに対応する。また (2) 式の $(a_j \sin 2kR_j + b_j \cos 2kR_j)$ は振動波形を表し、擬動径構造関数 $\tilde{N}(R_j)$ は $\tilde{N}(R_j) = \sqrt{a_j^2 + b_j^2}$ で評価される。

(2) 式で表した基底関数を用いたスパースモデリングで、原子配位のスパース性を表現する $\tilde{N}(R_j)$ と、一切の事前情報なしにデータだけから、デバイ・ワラー因子 σ_{DW} を推定することが出来る。

3 スパースモデリング

3.1 ℓ_1 正則化

ℓ_1 正則化法である LASSO 法⁸⁾を用いてスパースモデリングを行った。我々は (2) 式の a_j, b_j を要素を持つ $\vec{x} = (a_1, b_1, \dots, a_j, b_j, \dots)$ を用いて、EXAFS 振動データ \vec{y} が $\vec{y} = A(\sigma_{\text{DW}})\vec{x}$ の様に線形写像で表すことが出来るとして、(3) 式より \vec{x} のスパース解 $\hat{\vec{x}}$ を得た。

$$\hat{\vec{x}}(\lambda; \sigma_{\text{DW}}) = \arg \min_{\vec{x}} \left(\left\| \vec{y} - \overrightarrow{A(\sigma_{\text{DW}})} \vec{x} \right\|_2^2 + \lambda \|\vec{x}\|_1 \right) \quad (3)$$

ここで線形写像行列 $\overrightarrow{A(\sigma_{\text{DW}})}$ は、デバイ・ワラー因子の推定を実現するため σ_{DW} を変数として含む。 λ は、解のスパース性を制御するハイパーパラメータ⁸⁾である。

3.2 交差検定

線形写像行列にデバイ・ワラー因子 σ_{DW} が含まれているため、 σ_{DW} を適切な範囲でスキャンしてスパースモデリングと交差検定 (CV: Cross Validation)⁹⁾を行った。交差検定は、データを訓練データとテストデータに分割し、まず訓練データを対象としてスパースモデリングを用いて機械学習を行い、得られたスパース解を用いてテストデータの予測が出来るかを検証する方法である。この交差検定では、予測誤差 (CVE: Cross Validation Error) を最小化する条件で、スパース性を制御する λ と、物理モデル $\overrightarrow{A(\sigma_{\text{DW}})}$ に含まれる物理量の推定を実現できる。

(x_i, y_i) のデータ点で構成されるデータセット $D = \{\dots, (x_i, y_i), \dots\}$ を、訓練データセット D_f^{tr} とテストデータセット D_f^{te} に分割 ($D = \{D_f^{\text{tr}}, D_f^{\text{te}}\}$) する。

$$D = \{D_f^{\text{tr}}, D_f^{\text{te}}\} \begin{cases} D_f^{\text{tr}} \equiv \{\vec{x}_f^{\text{tr}}, \vec{y}_f^{\text{tr}}\}, \vec{y}_f^{\text{tr}} \equiv \{y_i\}, & i \notin D_f^{\text{te}} \\ D_f^{\text{te}} \equiv \{\vec{x}_f^{\text{te}}, \vec{y}_f^{\text{te}}\}, \vec{y}_f^{\text{te}} \equiv \{y_i\}, & i \in D_f^{\text{te}} \end{cases}$$

この際、全データ点をランダムな位置で分割し、分割数を F とするなら、 F ($f = 1, \dots, F$) 組の $D_f^{\text{tr}}, D_f^{\text{te}}$ を用

意し、全データ点を $f = 1, \dots, F$ で、一度だけ、かつ、必ずどれかの D_f^e として選択する。

まず、 f 組目の訓練データ D_f^t を対象として、ある λ におけるスパース解 $\vec{x}_f^t(\lambda; \sigma_{\text{DW}})$ を、

$$\vec{x}_f^t(\lambda; \sigma_{\text{DW}}) = \arg \min_{\vec{x}_f^t} \left(\left\| \vec{y}_f^t - A_f^t(\sigma_{\text{DW}}) \vec{x}_f^t \right\|_2^2 + \lambda \left\| \vec{x}_f^t \right\|_1 \right)$$

で得る。次にこのスパース解が、訓練データに含まれなかった残りのテストデータ D_f^e を予測できるかを評価するため、学習結果 \vec{x}_f^t を用いて、予測誤差を意味する次式の検定誤差 $\text{VE}_f(\lambda)$ を評価する。

$$\text{VE}_f(\lambda) \equiv \left\| \vec{y}_f^e - A_f^e(\sigma_{\text{DW}}) \vec{x}_f^t \right\|_2^2$$

交差検定誤差 $\text{CVE}(\lambda)$ は、 $\text{VE}_f(\lambda)$ の全てのデータ点(点数 n) の平均値として、(4) 式で定義される。

$$\text{CVE}(\lambda) \equiv \frac{1}{n} \sum_{f=1}^F \text{VE}_f(\lambda) \quad (4)$$

また、各 $f (= 1, \dots, F)$ における $\text{VE}_f(\lambda)$ の揺らぎより、 $\text{CVE}(\lambda)$ の誤差⁹⁾ に対応する標準偏差 $\text{SE}(\lambda)$ も評価することが出来る。

4 結果と考察

4.1 デバイ・ワラー因子の推定

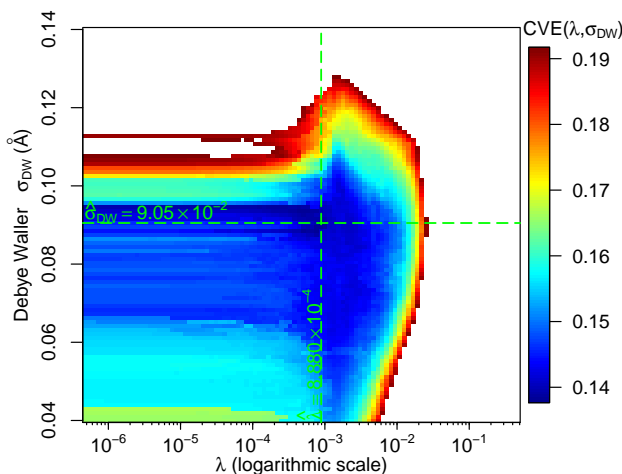


図2: CVE 最小化によるデバイ-ワラー因子の自動推定

スパース性を制御する λ とデバイ・ワラー因子 σ_{DW} を変えながら評価した交差検定誤差 $\text{CVE}(\lambda, \sigma_{\text{DW}})$ の例をヒートマップとして図2に示した。横軸は λ を対数で示し、縦軸は σ_{DW} である。 $\text{CVE}(\lambda, \sigma_{\text{DW}})$ は、 λ と σ_{DW} に対し顕著に変化し、図2の例では $\hat{\sigma}_{\text{DW}} = 9.05 \times 10^{-2} \text{ \AA}$ で交差検定誤差は最小化する。交差検定では、ランダムな位置でデータ D からテストデータセット D_f^e を抽出する

ため、定まった値として $\hat{\sigma}_{\text{DW}}$ の値は決定できない。そのため我々は、交差検定を4回行って、それらの交差検定で得られた $\hat{\sigma}_{\text{DW}}$ の平均値を、 $\overline{\hat{\sigma}_{\text{DW}}} = 0.0895 \pm 0.0028 \text{ \AA}$ と求めた。この誤差は平均値の誤差である。この得られた結果は、銅の面心立方構造を事前情報として用いてFEFFソフトウェア¹⁰⁾によって求めた最近接原子のデバイ・ワラー因子の値 $0.092 \pm 0.011 \text{ \AA}$ と誤差の範囲内でよく一致する。

この結果は、FEFFソフトウェアを用いた従来法では、デバイ・ワラー因子の決定には構造の事前情報が、本末転倒的に必要としたのに対し、我々のスパースモデリングでは、一切の事前情報を必要とせず、データだけから正しい値を推定できたことを意味する。

4.2 擬動径構造関数

我々は、平均自由行程以遠の原子では光電子波干渉が抑制されることを表す項 $\exp(-2R_j/\Lambda)$ を補償するために、 Λ として 10 \AA を用いた。平均自由行程の k 依存性 $\Lambda(k)$ は、構造の事前情報があれば、FEFFソフトウェア¹⁰⁾ を用いて評価することが可能で、その解析によると、EXAFS 振動振幅が最大になる $k = 6.75 \text{ \AA}^{-1}$ (図1を参照) で、平均自由行程は 10 \AA である。

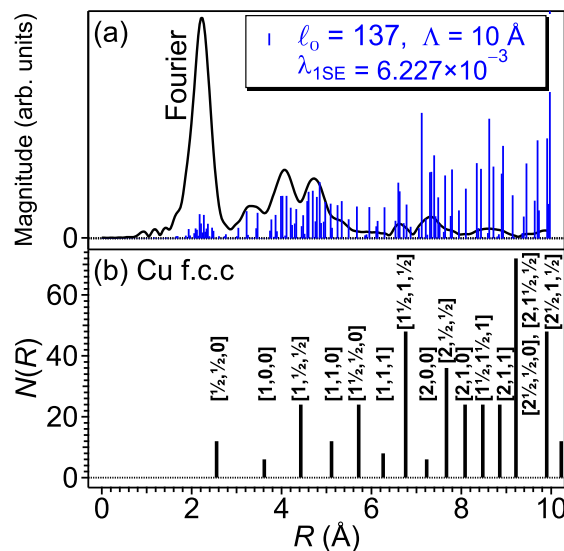


図3: スパース解と動径構造関数

得られたデバイ・ワラー因子 $\overline{\hat{\sigma}_{\text{DW}}}$ とこの平均自由行程 Λ を用いて、スパースな擬動径構造関数 $\tilde{N}(R_j)$ を求める。

$\overline{\hat{\sigma}_{\text{DW}}}$ の推定に用いた $\text{CVE}(\lambda, \sigma_{\text{DW}})$ 最小化は、予測誤差の最小化¹¹⁾ である。一方、本当の物理モデルを復元する目的、つまり実験データを解釈する目的で λ を選択する行為は違う¹¹⁾ ものであり、より正則(シンプル)な物理モデルを選択する方法として、交差検定誤差 $\text{CVE}(\lambda, \overline{\hat{\sigma}_{\text{DW}}})$ とその標準偏差 $\text{SE}(\lambda, \overline{\hat{\sigma}_{\text{DW}}})$ を用いた、one standard error rule^{11,9)} (1SE 則) が、経験的によく用

いられる。

ISE 則によって得られた擬動径構造関数 $\tilde{N}(R_j)$ を図 3(a) の縦棒で示した。この $\tilde{N}(R_j)$ のスパース解に非ゼロ要素が 137 個含まれるが、その個数は基底として用いた原子間距離 R_j の点数の約 1/3 である。これは、残りの約 2/3 の R_j の異なる基底は、計測された EXAFS 振動の再現には不必要であることを意味し、データに重畳するノイズを刈り込んだスパース解を得たことを意味する。この $\tilde{N}(R_j)$ の R_j スパース性は、図 3(a) の実線で示した EXAFS 振動データのフーリエ変換スペクトルと対照的で、フーリエ変換スペクトルは R_j に関して連続関数として得られる。この $\tilde{N}(R_j)$ に見られるスパース性は、(2) 式に示した基底関数に、フーリエ変換では基底関数に含まれていなかった、 $\chi(k)k^3$ 由来の k^2 項とデバイ・ワラー因子項 $\exp(-2k^2\sigma_{DW}^2)$ を基底関数に組み込んだ結果である。

図 3(b) は、面心立方構造である銅の動径構造関数 $N(R_j)$ である。この $N(R_j)$ は、面心立方構造を反映して離散化するが、 R_j の増加とともに強度が増加することが分かる。これは結晶においては、その対称性を反映して、遠距離になるに従って、同一距離で配位する原子数が立体的に増えていくためである。それに対し、図 3(a) に示したフーリエスペクトルは、遠距離になるに従ってピーク強度が減少する。これは、平均自由行程程度で電子波干渉が抑制されるためである。そのため、これまで EXAFS では、最近接や第 2、第 3 近接原子程度の近距離のマイクロ構造の解析しか出来ないと考えられていた。しかし、我々のスパースモデリングでは、適切な平均自由行程 Λ を選択することによって、その干渉抑制効果を補償することが可能である。その結果、図 3(a) のスパース解は、 R_j の増加とともに強度が増加しており、それは図 3(b) の傾向をうまく説明する。

図 3(b) に示した $N(R_j)$ と、図 3(a) のスパース解 $\tilde{N}(R_j)$ を比較すると、最近接である動径距離が 0.34 \AA ずれていることが分かる。この R_j のオフセットは、(1) 式の位相項 $\delta_j(k)$ が、 $\delta_j(k) \approx \delta_j^0 + \alpha_j k$ で表される^{1,7)} ように、電子波数 k に対して勾配を持って変化するためである。

5 まとめ

我々は、 ℓ_1 正則化と交差検定を用いて、標準試料である銅の EXAFS スペクトルのスパースモデリングを行った。その結果、計測データだけから、一切の事前情報なしで、注目原子に配位する近接原子の熱的な揺らぎや、マイクロ構造の揺らぎ、原子の可動性を表すデバイ・ワラー因子を推定することに成功した。さらに、電子波の平均自由行程以遠での干渉の抑制効果を補償した上で、近接原子の原子間距離においてスパースな動径構造関数を得ることに成功した。これらは従来用

いられていたフーリエ変換だけでは得られなかった物性情報で、我々は、スパースモデリングを用いて、データ駆動科学的に重要な物性情報を抽出することに成功したことを意味する。

今後我々の開発した方法は、マイクロ構造が未知な新規材料や、多結晶やアモルファス等の機能性材料、化学反応系等、マイクロ構造の事前情報が得難い物質・材料の EXAFS 計測で、新しい知見を得ることが期待できる。特に、デバイ・ワラー因子が重要な熱電材料¹²⁾ や超イオン伝導材料¹³⁾ での大きな発展が期待できる。

謝辞

本研究の一部は、JPSJ 科研費、JP16H04002 と、JST, CREST, JPMJCR1861 の支援を受けたものである。

参考文献

- 1) B. K. Teo (ed.), *EXAFS: Basic Principles and Data Analysis* (Springer-Verlag, Berlin Heidelberg, 1986).
- 2) P. Fornasini and R. Grisenti, *J. Synchrotron Radiat.* **22**, 1242 (2015).
- 3) S. Calvin (ed.), *XAFS for Everyone* (CRC Press, Boca Ration FL, 2013).
- 4) E. A. Stern, D. E. Sayers, and F. W. Lytle, *Phys. Rev.* **11**, 4836 (1975).
- 5) T. Kuwatani, R. Nakata, T. Hori, and M. Okada, *J. Inst. Electron. Inf. Commun. Eng.* **99**, 406 (2016).
- 6) B. Ravel and M. Newville, *J. Synchrotron Radiat.* **12**, 537 (2005).
- 7) I. Akai *et al.*, *J. Phys. Soc. Jpn.* **87**, 074003 (2018)
- 8) R. Tibshirani, *J. R. Stat. Soc. B* **58**, 267 (1996).
- 9) K. P. Murphy (ed.), *Machine Learning: A Probabilistic Perspective* (The MIT Press, Cambridge MA, 2012).
- 10) The University of Washington, The FEFF project, <http://leonardo.phys.washington.edu/index-feffproject.html> (2012).
- 11) R. Tibshirani, *Data Mining: Spring 2013*, 18. Model election and validation 1: Cross-validation, <http://www.stat.cmu.edu/~ryantibs/datamining/lectures/18-val1.pdf> (2013).
- 12) J. B. Goodenough, F. Rivadulla, E. Winkler, and J.-S. Zhou, *Europhys. Lett.* **61**, 527 (2003).
- 13) A. Yoshiasa, K. Koto, H. Maeda, and T. Ishii, *Jpn. J. Appl. Phys.* **36**, 781 (1997).