

## EXAFS(広域 X 線吸収微細構造) のスパースモデリング

## Sparse modeling of extended X-ray absorption fine structures (EXAFS)

赤井 一郎 \*1, \*3

Ichiro Akai

岩満 一功 \*1

Kazunori Iwamitsu

五十嵐 康彦 \*2

Yasuhiko Igarashi

岡田 真人 \*2

Masato Okada

岡島 敏浩 \*3

Toshihiro Okajima

平井 康晴 \*3

Yasuharu Hirai

\*1熊本大学

Kumamoto University

\*2東京大学

The University of Tokyo

\*3佐賀県立九州シンクロトロン光研究センター

Kyushu Synchrotron Light Research Center

In x-ray absorption spectra, fine structures appear due to interference between emitted photoelectron waves just above the absorption edge of a certain atom, which are called extended x-ray absorption fine structures (EXAFS). We have applied sparse modeling (SpM) onto the EXAFS analysis to obtain local structural information nearby the atom that absorbs x-rays. Although general Fourier transformation (FT) was used for elucidating such local structure, the FT spectra are strongly deformed by extrinsic effects of thermal disturbance and of interference limitation restricted by mean-free-path of photoelectrons. In our SpM, these extrinsic effects are built into the ready-made FT formalism to eliminate these effects. After such improvement, a peak due to the nearest neighbor atom becomes sharper than that obtained by a general FT. In addition, it is possible to extend the distance range for local structure analysis beyond the nearest neighbor atoms.

## 1. はじめに

最新の情報科学を実験データの解析に適用するデータ駆動科学は、これまでのデータ解析の価値観を一変させ、既存解析法のボトルネックを解消して新しい研究展開を可能とする。ベイズ推定によるデータ駆動科学では、物性を特徴づけるパラメータの統計的精度を、1つの実験データのみから評価することが可能 [Nagata 12, Iwamitsu 16] である。

一方、スパースモデリング (SpM) [岡田 14] に基づくデータ駆動科学は、人間が物体や現象を認知・識別する際に、頭のなかで行われるデータ処理行為を定式化し、実験データの解析に適用するものである。科学で行われる実験は、実験手法を適切に選択し、明らかにしたい要素をできるだけ簡潔に抜き出せるように最適化を行う。つまり、実験データを与える物理現象は、適切に選ばれた物理モデルに基づいてより少数の主要要素で説明可能であるという前提が成り立つ。SpM では、適切な基底関数を用いて、その様な実験データに含まれるスパースな主要要素を、適切な規範に従って先入観なしに自動抽出可能である。これまで我々は、この様な SpM をコヒーレントフォノン信号 [Rosker 86] の減衰振動モード分解に適用し、既存のフーリエ・ウェーブレット変換で顕在化する諸問題を解消できる新規解析法 [赤井 16, 相原 16] を提案している。

今回はこの SpM を、物質中原子の近距離構造の解析に用いられる広域 X 線吸収微細構造 (EXAFS: Extended X-Ray Absorption Fine Structures) [Nielsen 11] 信号の解析に適用し、既存解析で用いられるフーリエ変換を凌駕する新規解析法を提案する。

## 2. EXAFS

図 1(a) に示した様にガス等の単原子の場合、注目する原子の X 線吸収端で、吸収強度  $\mu d$  が疑似ステップ状に増加し、

連絡先: 赤井 一郎, 熊本大学パルスパワー科学研究所,  
〒 860-8555 熊本市中央区黒髪 2-39-1,  
TEL/FAX: 096-342-3296, iakai@kumamoto-u.ac.jp

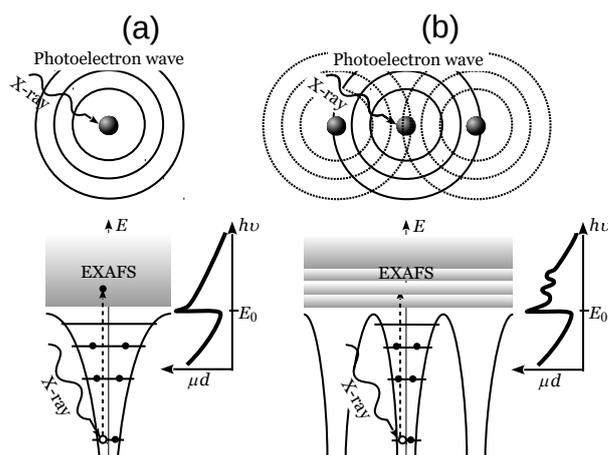


図 1: EXAFS の原理の概要

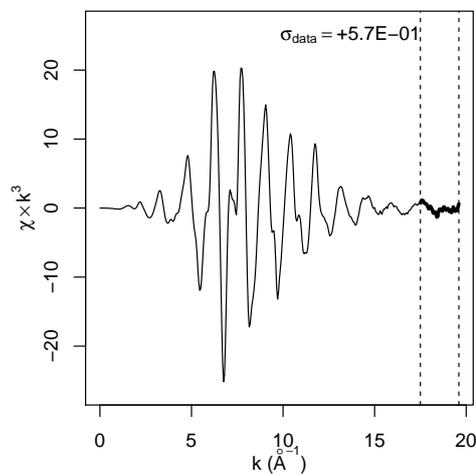


図 2: EXAFS 振動構造

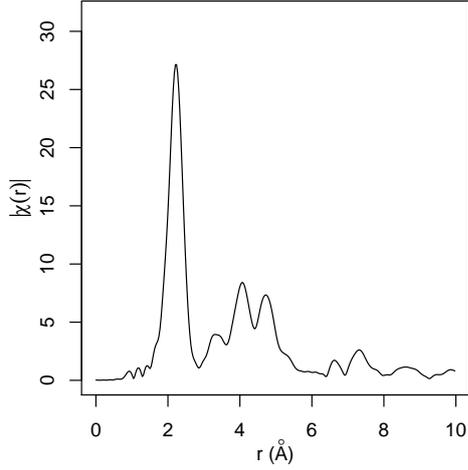


図 3: 近接原子間距離の動径分布

それより高エネルギーで一様に減少する X 線吸収スペクトルが観測される。しかし、その原子の近傍に他の原子があると、図 1(b) に示したように、吸収端の高エネルギー側の EXAFS 領域に振動波形状の微細構造が観測される。これは、吸収端  $E_0$  より高エネルギー側 ( $h\nu > E_0$ ) の X 線吸収で、原子は波数  $k = \sqrt{2m(h\nu - E_0)/\hbar^2}$  を持つ光電子を放出するが、その光電子波自身と、それが近接原子で散乱されて戻ってきた光電子波との干渉によって電子状態に変調を与えた結果である。EXAFS は、光電子波の干渉が光電子の平均自由行程  $\Lambda$  の距離程度で制限させることから、注目原子の近距離構造を解明する有効な計測法である。透過法に加えて、X 線吸収による蛍光 X 線を捉える蛍光法等 [SAGA-LS 16] を用いることにより、様々な試料の形態や状態でも測定可能で、産業利用も極めて多く、原子サイズレベルの近距離構造を計測する汎用的な方法である。

本論文では図 2 に示した様に、EXAFS 解析データの前処理は、従来法通り、吸収スペクトルからなだらかなバックグラウンドを差し引き、波数  $k$  が大きな領域の構造を強調するため  $k^3$  をかける処理を行った。このデータは、九州シンクロトロン光研究センターの BL11 [SAGA-LS 16] で測定した銅箔の結果である。これまで EXAFS 信号は、そのフーリエ変換を行い、図 3 に示した近接原子間距離の動径分布を得て、近距離構造の知見を得ていた。図 2 で最大強度を与えるピークの横軸は最近接原子間距離を表し、遠距離になる従ってピーク強度が減少する傾向があることが分かる。

### 3. EXAFS のスパースモデリング

#### 3.1 EXAFS 信号の物理モデル

この EXAFS 信号は、X 線を吸収している注目原子に対し、 $j$  番目の近接原子への距離を  $R_j$  とすると、(1) 式でモデル化 [Nielsen 11] される。

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

ここで、 $t_j(k)$ ,  $\sigma_j$ ,  $\delta_j(k)$  は、 $R_j$  に位置する原子の後方散乱振幅、デバイ・ワーラー因子、位相シフトである。

(1) 式において、 $N_j$  は注目原子から距離  $R_j$  にいる他原子の占有確率で、 $N_j(R_j)$  が得たい情報である。しかし従来用いられてきたフーリエ変換では、 $\sin(2kR_j + \delta_j(R_j))$  を基底関数として展開するため、得られる動径分布関数 (図 3) には、 $N_j(R_j)$  以外に、後方散乱因子、温度揺らぎを反映するデバイ・ワーラー因子、光電子の平均自由行程  $\Lambda$  の情報が残存し、 $N_j(R_j)$  の理解を困難にさせてきた。

#### 3.2 LASSO 法

スパースモデリング (SpM) には、Least Absolute Shrinkage and Selection Operator (LASSO) 法 [Tibshirani 96] を用いた。LASSO 法は、推定したいパラメータを  $\vec{x}$ 、解析対象データを  $\vec{y}$  とした時、(2) 式のように、 $\vec{y} = A\vec{x}$  の線形回帰問題において、残差二乗平均  $\|\vec{y} - A\vec{x}\|_2^2$  に、 $\vec{x}$  の 1 次ノルム  $\|\vec{x}\|_1$  を、非スパース性のペナルティ項として付加して最小化 (L1 正則化) させ、 $\vec{x}$  のスパース解  $\hat{\vec{x}}$  を得る。

$$\hat{\vec{x}} = \arg \min_{\vec{x}} (\|\vec{y} - A\vec{x}\|_2^2 + \lambda \|\vec{x}\|_1) \quad (2)$$

ここで  $\lambda$  はスパース性の制御パラメータで、 $\lambda = 0$  で (2) 式は最小二乗法の解と一致し、データに重畳するノイズまでも合わせにいく過剰フィッティング結果を与える。一方  $\lambda \rightarrow \infty$  の極限では、 $\vec{x}$  の全ての要素がゼロとなり、データ全てがノイズである解析結果を与える。LASSO 法では、適切な規範で  $\lambda$  を選択し、ノイズを刈り込んで、データの特徴構造を適切に抽出したスパース解を得る [Donoho 06, Candes 06] ことが可能である。

#### 3.3 EXAFS のスパースモデリング

今回は、線形回帰の写像行列  $A = \{A_{k,R_j}\}$  に、従来のフーリエ変換の基底である平面波項  $\sin(2kR_j + \delta_j(k))$  に、(1) 式に含まれるデバイ・ワーラー因子  $\sigma_j$  と光電子の平均自由行程  $\Lambda$  が関与する項を組み込んだ

$$A_{k,R_j} = \frac{k^2}{R_j^2} \exp \left[ -2 \left( k^2 \sigma_j^2 + \frac{R_j}{\Lambda} \right) \right] \times [a(R_j) \sin 2kR_j + b(R_j) \cos 2kR_j]$$

を用い、スパース解を得る  $\vec{x}$  を、

$$\vec{x} = \{a(R_1), b(R_1), \dots, a(R_j), b(R_j), \dots\}$$

とした。これで X 線吸収端で特定される注目原子から距離  $R_j$  にいる原子の占有確率  $N(R_j)$  に比例する物理量は、(3) 式で得られ、(1) 式において  $\sum N_j t_j(k)$  による動径分布関数に対応する。

$$\hat{N}(R_j) \propto \sqrt{a^2(R_j) + b^2(R_j)} \quad (3)$$

SpM 法の実装は、R の glmnet ライブラリ [Friedman 10, Simon 11] を使い、原子の熱揺動を表すデバイ・ワーラー因子は  $\sigma_j = 0.113 \text{ \AA}$ 、光電子の平均自由行程は  $\Lambda = 10 \text{ \AA}$  とした。スパース化させる  $\vec{x}$  の  $R_j$  の刻み間隔は、図 3 の動径分布関数のプロット間隔と同じにとり、 $0 \sim 9.97 \text{ \AA}$  を 326 分割した。LASSO 法による解析結果を図 4, 5 に示した。

両図の (a) は、スパース性制御パラメータ  $\lambda$  を自然対数で横軸にとり、○と●は、(3) 式の子非ゼロ要素の数と、得られたスパース解による再現データと実験データの平均二乗誤差根 (RMSD: Root-Mean-Square Deviation)  $\tau$  の変化を示したものである。 $\lambda$  の減少とともに、非ゼロ要素の数が増加しつつ、RMSD  $\tau$  が減少することが分かる。

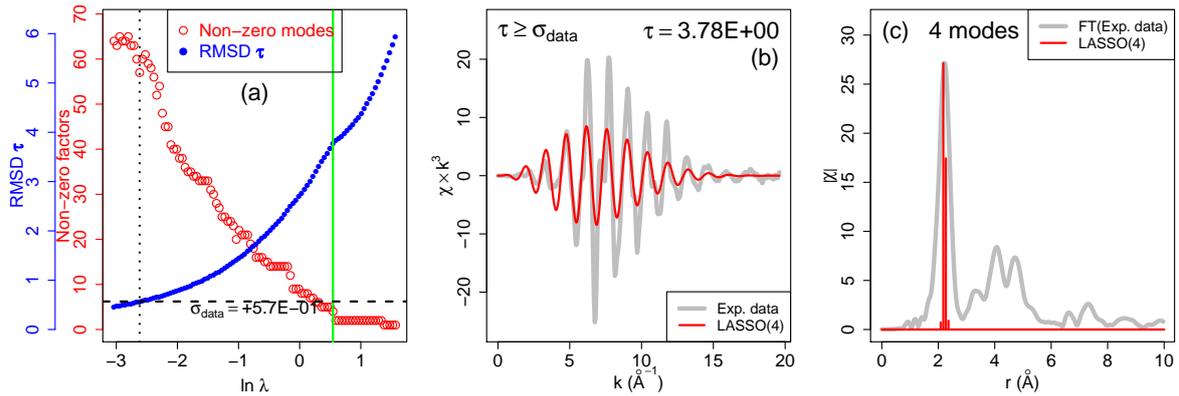


図 4: EXAFS 振動構造の SpM 解析結果:  $\lambda$  が大きい場合。

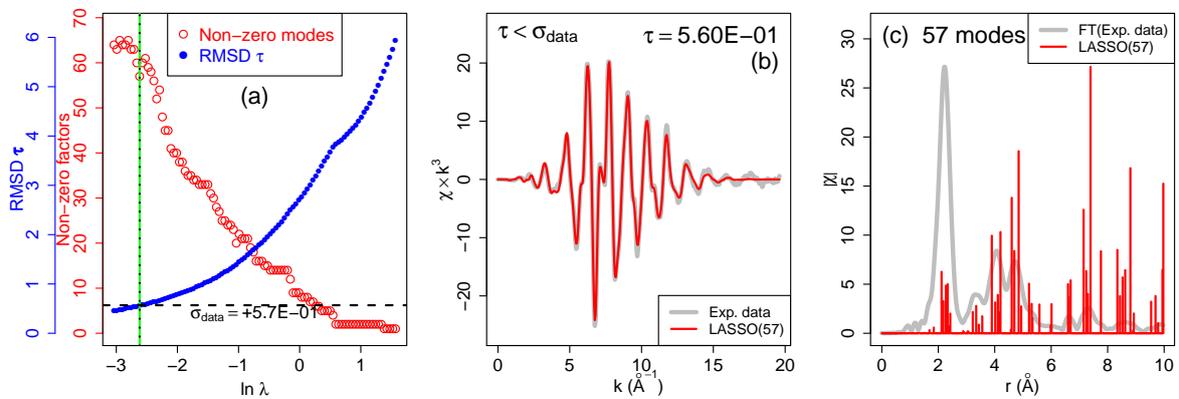


図 5: EXAFS 振動構造の SpM 解析結果: 平均二乗誤差根  $\tau$  がデータのノイズの大きさ  $\sigma_{\text{data}}$  と均衡する場合。

両図の (a) 中の緑色垂直実線は解析に用いられている  $\lambda$  の位置を示している。また水平破線は、図 2 の実験データに重畳するノイズの標準偏差  $\sigma_{\text{data}}$  を表し、垂直点線は、平均二乗誤差根  $\tau$  が  $\sigma_{\text{data}}$  と交わるところである。ここで  $\sigma_{\text{data}}$  は、図 2 で太線で示した、波数  $k$  の大きい端の領域の実験データより決定した。

図 4, 5 の (b) は、灰色が図 2 に示した解析対象データ、赤色が得られたスパース解による再現データである。また図中の  $\tau$  値は、再現データの平均二乗誤差根である。図 4, 5 の (c) は、灰色が図 3 と同一で、従来解析法であるフーリエ変換から得られる動径分布関数で、赤色がスパース解の (3) 式を表したもので、縦軸は最大強度がフーリエスペクトルのピーク強度に合うように規格化してある。

図 4 は、スパース性制御パラメータ  $\lambda$  が比較的大きい場合で、(b) で見られる一番顕著な振動構造が自動的に選択され、それが、(c) のフーリエ変換動径分布関数の最大強度を与える  $R_j$  に対応する 4 つスパース解として抽出されている事がわかる。

一方図 5 は、その (a) が示すように、平均二乗誤差根  $\tau$  が、重畳するノイズの大きさ  $\sigma_{\text{data}}$  と均衡した時の解析結果である。(b) より、実験データは極めてよく再現されていることが分かる。一方 (c) では、灰色で示した動径分布関数と大きく異なるスパース解が得られる。灰色のフーリエ変換動径分布関数が、 $R_j$  の全領域 ( $0 \sim 9.97 \text{ \AA}$ ) で非ゼロ (326 個) の強度をもっ

ているのに対し、スパース解では、約 1/5 の 57 個の非ゼロ要素しか持たず、それ以外は実験データの再現には寄与していないことが分かる。この結果より、この SpM 解析により、データに重畳するノイズを適切に刈り込んだスパース解が得られていることが分かる。

#### 4. 考察

図 3 で示したフーリエ変換動径分布関数は、基底関数として平面波を用いている。しかし、図 2 に示した EXAFS 振動波形は、振動構造が、小さな  $k$  領域で立上り、大きな  $k$  領域で減衰することから、 $k$  に対して平坦な平面波で展開するのは不適切であることは明白である。ここで、小さな  $k$  領域での立上りは、(1) 式の左辺で大きな  $k$  での振動構造を強調するためにかけられる  $k^3$  による。一方、大きな  $k$  領域における振動構造の減衰は、(1) 式で原子の熱揺動を表すデバイ・ワラー因子  $\sigma_j$  を含む項  $\exp(-2k^2\sigma_j^2)$  による。これらの項は、図 3 に見られるように、フーリエ変換動径分布関数のピーク構造に顕著な幅を与えてしまう。

更に、フーリエ変換動径分布関数では、図 3 でも見られるように、 $R_j$  の小さな近距離で主要ピーク構造が現れ、遠距離でピーク強度が減少するため、EXAFS 解析において、最近接より以遠の中距離構造の高精度解析の障害となっている。これは、(1) 式において、光電子の平均自由行程  $\Lambda$  程度の距離で干

渉が制限されることを表す指数関数減衰項  $\exp(-2R_j/\Lambda)$  が、フーリエ変換動径分布関数に重畳しているためである。今回対象とした結晶などの場合を考えると、本来その対称性から、距離が遠ざかるにつれ、同距離にいる原子の占有確率  $N(R_j)$  は増加するはずである。

今回の EXAFS 信号の SpM では、3.3 節に述べたとおり、これら 2 つの要因は線形写像行列  $A$  に含まれており、この 2 つの問題を除去した上で、より本質的な近距離構造の解明ができたと考えられる。具体的には、図 5(c) で  $\sim 2.2 \text{ \AA}$  に現れ、最近接原子間距離を表すピーク構造が、灰色で示されたフーリエ変換動径分布関数と比較して、赤縦棒で示したスパース解で得られるピーク構造の方がより先鋭である。これは、今回の SpM 解析で、デバイ・ワーラー因子で表される原子の熱揺動を除去した本質的な近距離構造が解明できていることを意味している。更に図 5(c) で、スパース解が、最近接より以遠で原子の占有確率が増加する傾向を如実に表していることも明らかである。

## 5. まとめ

銅箔試料の EXAFS 信号の SpM 解析を行った。今回は EXAFS 信号の物理モデルに基づいた線形回帰解析において、原子の熱揺動を表すデバイ・ワーラー因子項と、光電子の平均自由行程で光電子波干渉現象が制限される要因を取り込んだ線形写像を用い、L1 正則化を実装した LASSO 法で解析を行った。

その結果、最近接原子においてはより先鋭な原子占有確率分布が得られ、最近接より以遠の原子では、結晶の対称性を反映したより正しい原子占有確率分布が得られることを示した。これらは既存のフーリエ変換における諸問題を解決したもので、EXAFS 信号の新規解析法として SpM が有効であることを明確に示すものである。

## 参考文献

- [Candes 06] Candes, E. J. and Tao, T.: Near-optimal signal recovery from random projections: Universal encoding strategies?, *IEEE Transactions, Information Theory*, Vol. 52, pp. 5406–5425 (2006)
- [Donoho 06] Donoho, D. L.: Compressed sensing, *IEEE Transactions, Information Theory*, Vol. 52, pp. 1289–1306 (2006)
- [Friedman 10] Friedman, J., Hastie, T., and Tibshirani, R.: Regularization Paths for Generalized Linear Models via Coordinate Descent, *J. Stat. Software*, Vol. 33, pp. 1–22 (2010)
- [Iwamitsu 16] Iwamitsu, K., Aihara, S., Okada, M., and Akai, I.: Bayesian Analysis of an Excitonic Absorption Spectrum in a  $\text{Cu}_2\text{O}$  Thin Film Sandwiched by Paired  $\text{MgO}$  Plates, *J. Phys. Soc. Jpn.*, Vol. **85**, p. 094716 (2016)
- [Nagata 12] Nagata, K., Sugita, S., and Okada, M.: Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks*, Vol. **28**, p. 82 (2012)
- [Nielsen 11] Nielsen, J. A. and McMorrow, D.: §7.2, *Elements of Modern X-ray Physics 2nd Ed.*, John Wiley & Sons Ltd. (2011)
- [Rosker 86] Rosker, M. J., Wise, F. W., and Tang, C. L.: Femtosecond Relaxation Dynamics of Large Molecules, *Phys. Rev. Lett.*, Vol. 57, pp. 321–324 (1986)
- [SAGA-LS 16] SAGA-LS 九州シンクロトロン光研究センター：県有ビームライン BL11 装置概要 (2016), <http://www.saga-ls.jp/?page=200>
- [Simon 11] Simon, N., Friedman, J., Hastie, T., and Tibshirani, T.: Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent, *J. Stat. Software*, Vol. 39, pp. 1–13 (2011)
- [Tibshirani 96] Tibshirani, R.: Regression shrinkage and selection via the lasso, *J. Roy. Stat. Soc. B*, Vol. 58, pp. 267–288 (1996)
- [岡田 14] 岡田 真人：新学術領域研究「スパースモデリングの深化と高次元データ駆動科学の創成」(2014), <http://sparse-modeling.jp/>
- [赤井 16] 赤井 一郎, 村田 伸, 相原 慎吾, 徳田 悟, 岩満 一功, 岡田 真人：SpDMD によるコヒーレントフォノン信号のモード分解解析 (I) (2016), 日本物理学会「2016 年秋季大会」14pAL-3, 4
- [相原 16] 相原 慎吾, 村田 伸, 徳田 悟, 岩満 一功, 溝口 幸司, 岡田 真人, 赤井 一郎：SpDMD によるコヒーレントフォノン信号のモード分解解析 (II) (2016), 日本物理学会「2016 年秋季大会」14pAL-4