



第3回 SPring-8 材料構造の解析に役立つ計算科学研究会  
@ニチイ学館・神戸ポートアイランドセンター

# データ駆動科学による 放射光計測データの 新規解析法の提案

赤井一郎

熊本大学・パルスパワー科学研究所、  
九州シンクロトロン光研究センター



## 自己紹介

赤井一郎、博士(理学)

- 所属: 熊本大学 パルスパワー科学研究所  
極限物性科学部門 極限物性物理分野 教授
- これまでの専門: 光物性物理学(実験)
  - 光機能性材料のコヒーレントフォノン
  - サブKにおける酸化物半導体の励起子BEC
  - ns~fs スケール光励起状態ダイナミクス
  - 半導体ナノ構造の励起子物性
- 略歴
  - 大阪市立大学、理学部~D2中退、博士(理学) 大阪市立大学 @ 1992
  - 1989年4月 大阪市立大学 理学部助手、1996年4月 同 講師、1999年4月 同助教授
  - 2007年4月 熊本大学 衝撃・極限環境研究センター 教授
  - 2013年4月 熊本大学 パルスパワー科学研究所 教授
  - 2016年10月 九州シンクロトロン光研究センター 客員主任研究員(併任)



# 謝辞

- 共同研究者

- ✓ 岡田真人先生とそのグループ,
- ✓ 平井康晴センター長, 岡島敏浩先生,
- ✓ 細川伸也先生
- ✓ 溝口幸司先生,
- ✓ 高橋和敏先生,
- ✓ 中山正昭先生,
- 赤井研のみんな。

東京大学  
**SAGA-LS**  
熊本大学  
大阪府立大学  
佐賀大学  
大阪市立大学



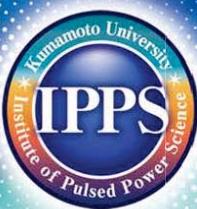
- 科研費

- ✓ 基盤研究(B)
- ✓ 新学術領域「疎性モデリング」公募研究
- ✓ 挑戦的萌芽研究

JP16H04002  
JP16H01552  
JP16K13824  
**科研費**  
KAKENHI

SPring-8 2017/02/27

3



# 目次

- ・自己紹介
- ・謝辞
- ・本日のつかみ (身近な例)
- ・データ駆動科学 / 効能
  - 間接測定される物理量の統計的精度を評価
  - 物理モデルの妥当性の評価
  - 既存の解析法を「鵜呑み」にしていいのか、それが **Best solution** か?
- ・データ駆動科学 / 方法論
  - ベイズ推定
  - Sparse Modeling (SpM)
- ・まとめ

SPring-8 2017/02/27

4



# データ駆動科学

理論家に、「あなたの実験結果がおかしい」と言わされたことありませんか？

最新の情報科学の手法をデータ解析に導入して、既存のボトルネックを解消し、データを起点として新しい研究展開を開く。

## ・ベイズ推定

ベイズの定理を用いて、因果律をさかのぼる

➤ 物理パラメータの事後確率分布を評価できる

➤ ベイズ自由エネルギーを元に、モデル選択が可能

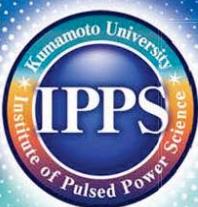
## ・スペースモデリング

仮定：“データを説明する主要因は、本来、数が少ないはず。”

➤ 実験データに含まれる主要因を「自動的」に分解可能

SPring-8 2017/02/27

5



## ベイズ推定

ある人が癌診断を受けた

- 信頼性が**95%**の癌診断を受けて「**陽性**」ですと診断された。
- 但し全人口に対する統計から、その癌の罹患率は

$$P(A) = 0.5\%.$$

- さて、その人が本当に  
罹患者である確率は?

癌診断の信頼性

	陽性 ( $C$ )	陰性 ( $\bar{C}$ )
癌である ( $A$ )	95%	5%
癌でない ( $\bar{A}$ )	5%	95%

- 本当に癌である確率  $P(A|C) = \text{約}8.7\%$ （「ホッ」！と出来ますか？）
- ベイズ推定：罹患率の推定精度が向上した  
事前確率  $P(A)=0.5\%$   
→計測(診断)→ 事後確率  $P(A|C) = \text{約}8.7\%$

SPring-8 2017/02/27

6

# データ駆動科学の効能

最新の情報科学の手法をデータ解析に導入して、既存のボトルネックを解消し、データを起点として新しい研究展開を開く。

## 実験屋こそ取り組むべき（データ解析の価値観が変わる）

- ・データを骨までしゃぶる
  - 間接測定される物理量の統計的精度を評価
  - 物理モデルの妥当性の評価
  - 既存の解析法を「鵜呑み」にしていないか、それが Best solution か？
- ・実験の最適化
  - 極限測定で、必要な精度を得るために実験条件は？
  - 実験を行う前の事前シミュレーションによる実験の効率化
- ・理論研究へのフィードバック
  - 解析に用いる物理モデルの良し悪しを判定

SPring-8 2017/02/27

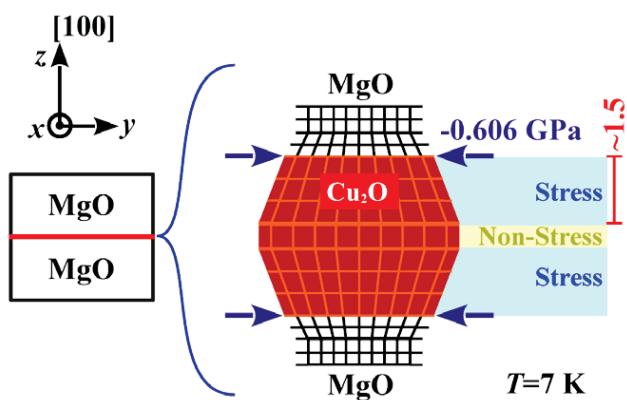
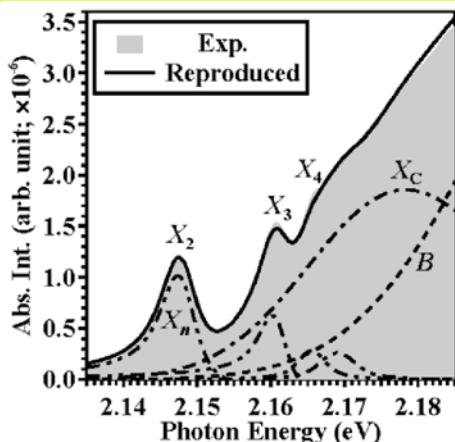
9

## ベイズ推定： $\text{Cu}_2\text{O}$ 薄膜結晶における圧力効果

### $\text{Cu}_2\text{O}$ の励起子状態：

励起子ボーズ・AINシュタイン凝縮が期待される

- ・高密度励起子を空間的に閉じ込めるトラップポテンシャルの形成
- ・ $\text{Cu}_2\text{O}/\text{MgO}$ :  $E_g < E_g^{\text{bulk}}$  or  $E_g > E_g^{\text{bulk}}$





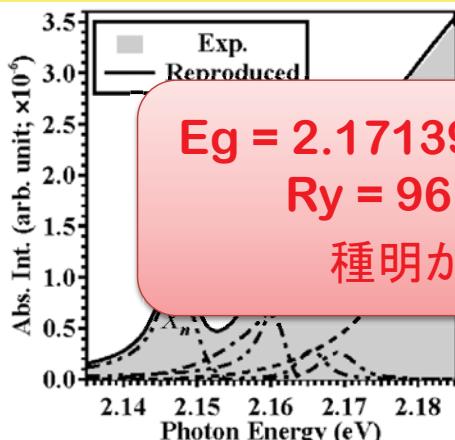
# ベイズ推定: $\text{Cu}_2\text{O}$ 薄膜結晶における圧力効果

$\text{Cu}_2\text{O}$ の励起子状態:

励起子ボーズ・AINシュタイン凝縮が期待される

- 高密度励起子を空間的に閉じ込めるトラップポテンシャルの形成

- $\text{Cu}_2\text{O}/\text{MgO}: E_g < E_g^{\text{bulk}}$  or  $E_g > E_g^{\text{bulk}}$



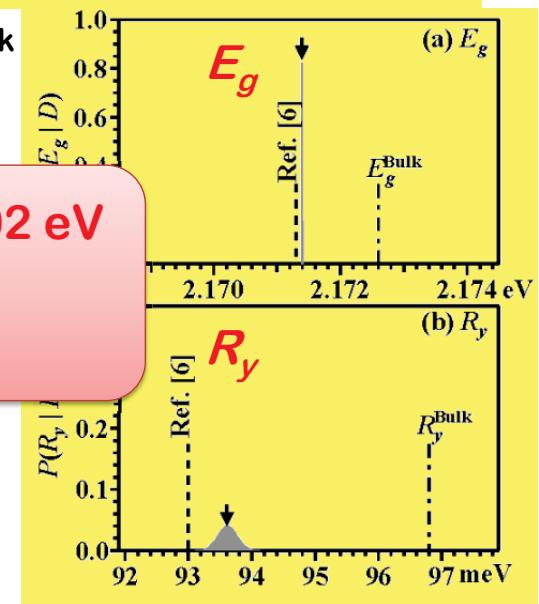
$$E_g = 2.171392 \pm 0.000002 \text{ eV}$$

$$R_y = 96.6 \pm 0.1 \text{ meV}$$

種明かしは後ほど!

K. Iwamitsu et.al, JPSJ 85 094716 (2016)

SPring-8 2017/02/27

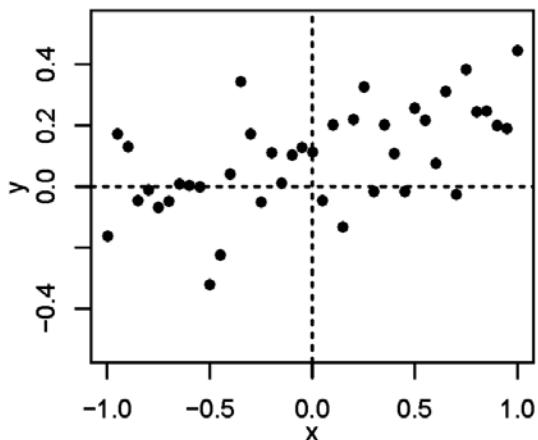


## Bayesian 自由エネルギーによる モデル選択

どのモデルで解析すべきか?

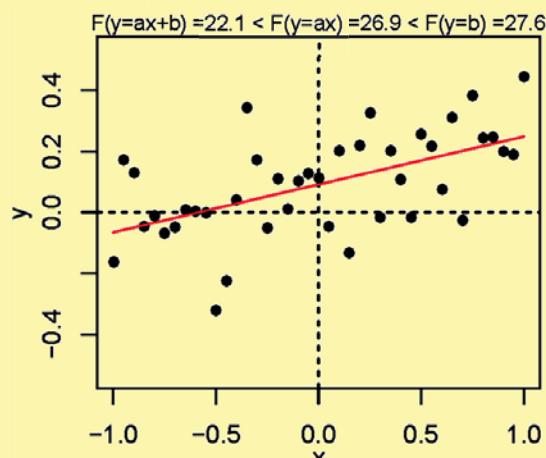
- 実験データだけから判定する

$$y = b, y = ax, y = ax + b$$



Bayesian 自由エネルギー

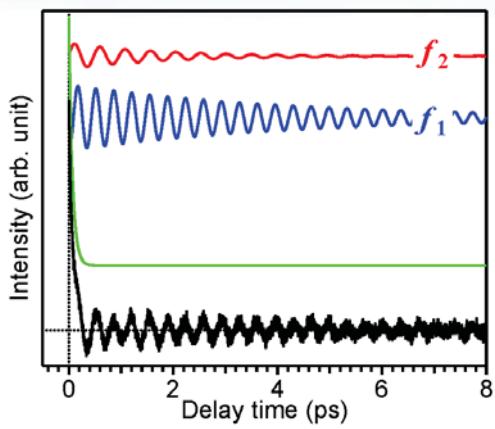
$$F(y=b) > F(y=ax) > F(y=ax+b)$$



SPring-8 2017/02/27

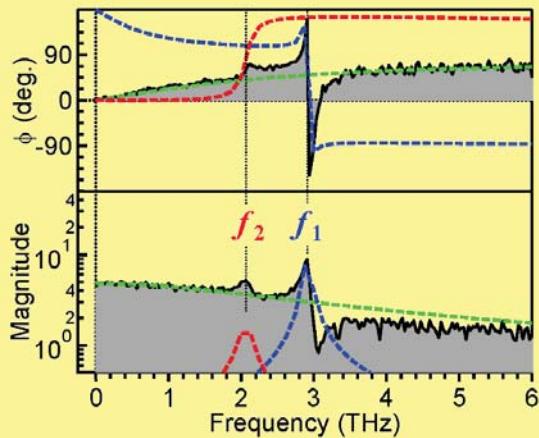
# 既存解析法の諸問題

多モード減衰振動波形



既存解析法

➤ フーリエ変換、Wavelet変換



基底関数 = 平面波

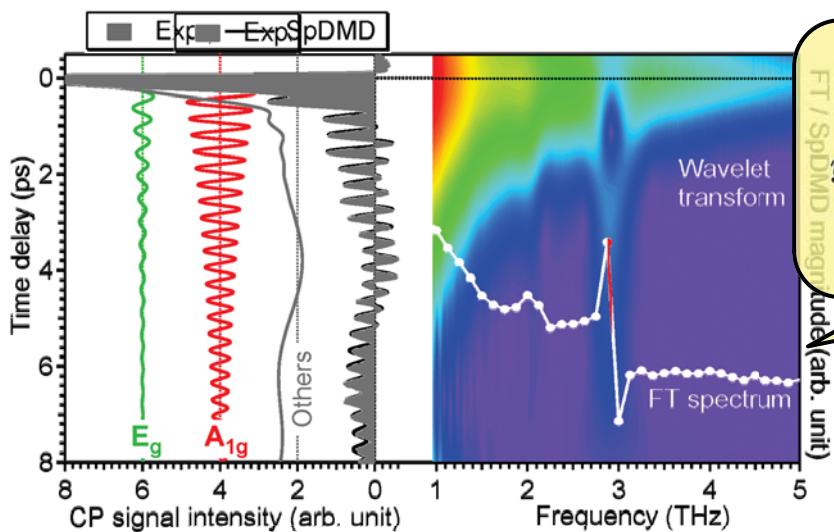
$$y(t) = \sum_i c(f_i) \cos(2\pi f_i t + \phi_i)$$

過剰な成分によるモード展開

- 不確定性幅
- 他成分が含まれると大きく変形
- 位相情報は得られない

# SpDMDによる少数モードによる成分分解

• FT / Wavelet → 離散的な振動モードのスペクトルを大きく変形



基底関数 = 減衰振動

$$y(t) = \sum_{i=1}^6 c_i \exp\left(-\frac{t}{\tau_i}\right) \cos(2\pi f_i t + \phi_i)$$

スパースなモード分解

**SpDMD は 2つのモードを正しく抽出できる**



## 目次

- 自己紹介
- 謝辞
- 本日のつかみ (身近な例)
- データ駆動科学 / 効能
  - 間接測定される物理量の統計的精度を評価
  - 物理モデルの妥当性の評価
  - 既存の解析法を「鵜呑み」にしていいか、それが Best solution か?
- データ駆動科学 / 方法論
  - ベイズ推定
  - Sparse Modeling (SpM)
- まとめ

SPring-8 2017/02/27

20



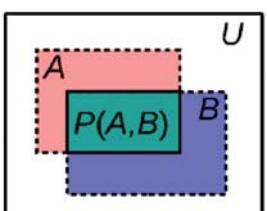
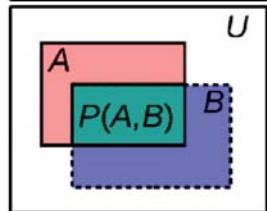
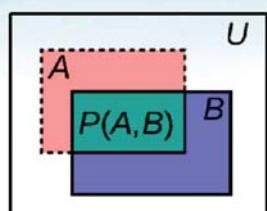
## ベイズの定理

- 同時確率( $A$ と $B$ が同時に起こる確率)  $P(A, B)$
- $B$ 条件下の  $A$ の確率  $P(A|B) = \frac{P(A, B)}{P(B)}$
- $A$ 条件下の  $B$ の確率  $P(B|A) = \frac{P(A, B)}{P(A)}$
- **ベイズの定理** Phil. Trans. Roy. Soc. **53**, 370 (1763).

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A),$$

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)},$$

$$P(B|A) = P(A|B) \frac{P(B)}{P(A)}$$



SPring-8 2017/02/27

21

# ベイズ推定

ベイズの定理を用いて、因果律をさかのぼる

- ベイズの定理

$$P(\theta, D) = \underbrace{P(D|\theta)}_{\text{Measure}} P(\theta) = \underbrace{P(\theta|D)}_{\text{What we want}} P(D)$$

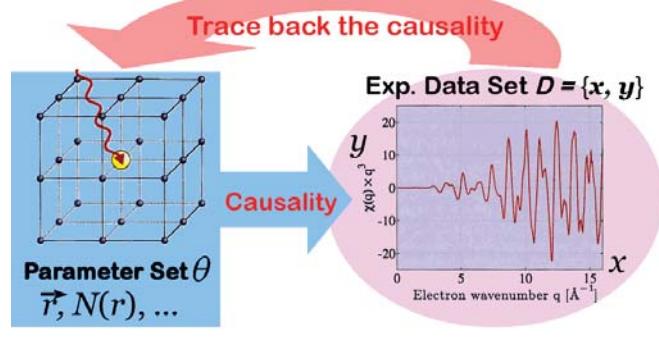
- 物理モデル + ランダムノイズ

$$y_i = \underbrace{f(x_i; \theta)}_{\text{Physical model}} + \underbrace{n_i}_{\text{Gaussian noise}}$$

- 事後確率

$$E(\theta) = \frac{1}{2N} \sum_{i=1}^N [y_i - f(x_i; \theta)]^2$$

$$P(\theta|D) \propto \underbrace{\exp\left(-\frac{N}{\sigma_{\text{data}}^2} E(\theta)\right)}_{\text{Posterior likelihood}} \underbrace{P(\theta)}_{\text{Prior}}$$



- MCMC法で、 $P(\theta|D)$  分布を得る←ノイズの大きさ  $\sigma_{\text{data}}$  で決まる

SPring-8 2017/02/27

22

## ベイズ推定を実装するアルゴリズム

- グリッドサーチ (全パラメータの総スキャン)

➤ 例えば5つのパラメータで有効桁2~3桁、 $P(\theta|D)$ の計算が1 msかかるとすると、2,500年 by (i5-CPU)

- マルコフ連鎖モンテカルロ法

(MCMC: Markov chain Monte Carlo methods)

P. J. Green,

Biometrika **82**, 711 (1995).

- ✓ メトロポリス法 (Metropolis-Hastings method)

N. Metropolis, *et.al.*,

J. Chem. Phys. **21**, 1087 (1953)

C. Andrieu, *et.al.*,

Machine Learning **50**, 5 (2003).

- レプリカ交換モンテカルロ法

R. H. Swendsen, *et.al.*,

Phys. Rev. Lett. **57**, 2607 (1986).

K. Hukushima, *et.al.*,

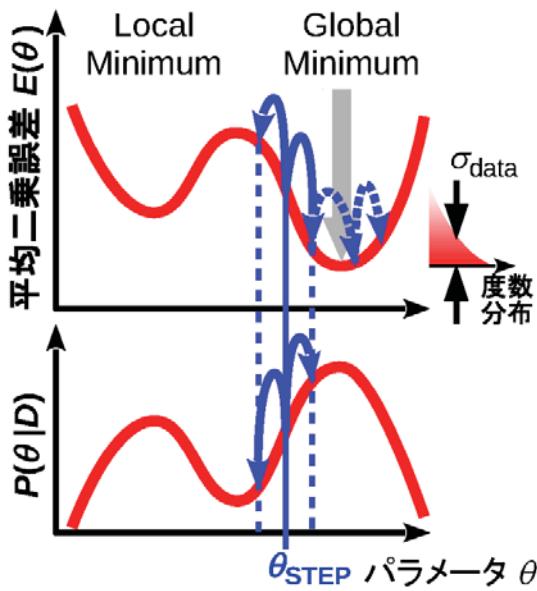
J. Phys. Soc. Japan **65**, 1604 (1996).

K. Nagata, *et.al.*,

Neural Networks **28**, 82 (2012).

# メトロポリス法

適当な初期値を与え、直前の値から乱数で生成した候補値の確率を評価して、採択・非採択を決定し、その履歴から値の確率分布得る。



## メリット

- 履歴を取るため確率分布を効率よく得ることが出来る。
- 採択・非採択での詳細つりあいを保証するため、平衡確率分布が得られる。
- 十分な履歴を取れば、パラメータの全空間を探索することが可能で、エルゴード性を保証する。

## デメリット

- 十分な履歴 = 天文学的時間

SPRING-8 2017/02/27

24

# 物理モデルを導入する

## スペクトル要素

•  $X_n$ : 励起子共鳴 
$$X_n = \left( \frac{f_n}{\pi} \right) \frac{\Gamma_n + 2A_n(E - E_n)}{(E - E_n)^2 + \Gamma_n^2}$$

•  $X_C$ : 励起子連続帯 
$$X_C = f_{X_C} \pi \beta (1 + \beta^2) \frac{\exp(\pi\beta)}{\sinh(\pi\beta)}, \beta \equiv \sqrt{\frac{R_y^*}{E - E_g}}$$

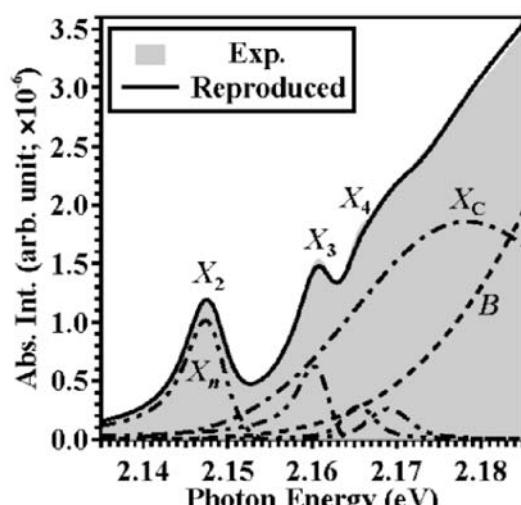
•  $B$ : バンド間吸収 
$$B = f_B (E - E_g)^{3/2}$$

それぞれで不均一幅を考慮

$X_n$ 遷移の共鳴エネルギー  
 $E_n$ の従う物理法則

$$E_n(T) = E_g(T) - \frac{R_y^*}{n^2}$$

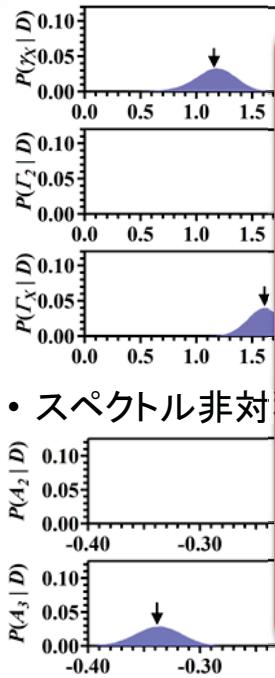
複数のピークから  $E_g, R_y$  を推定  
→ 推定精度向上





# 全てのパラメータ値の確率分布を得られる

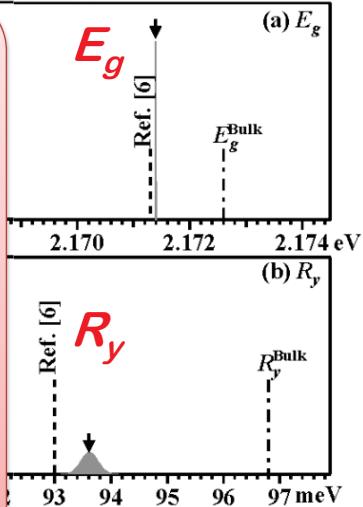
## ・スペクトル幅



## ・スペクトル強度

Parameters	Mean	$\epsilon_\theta$
$E_g$ (eV)	2.171392	0.000002
$R_y$ (meV)	93.6	0.1
$A_2$	-0.201	0.006
$A_3$	-0.34	0.01
$f_0$ ( $\times 10^{-8}$ )	8.8	0.1
$f_3$ ( $\times 10^{-8}$ )	0.355	0.008
$f_4$ ( $\times 10^{-8}$ )	0.179	0.008
$\gamma_X$ (meV)	1.2	0.1
$\Gamma_2$ (meV)	2.37	0.09
$\Gamma_X$ (meV)	1.6	0.1

## ・ $E_g$ , $R_y$



事後確率分布の平均値を用いて  
実験データを再現

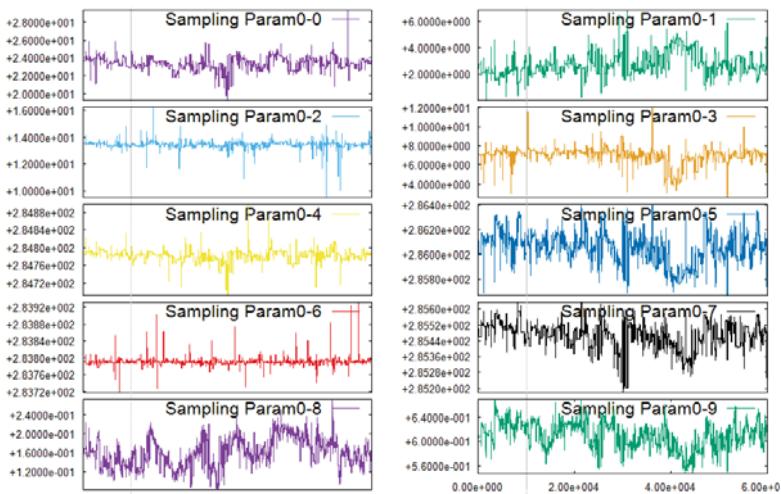
SPring-8 2017/02/27  
K. Iwamitsu et.al, JPSJ 85 094716 (2016)

26

## ノイズの大きさによって推定精度は決まる

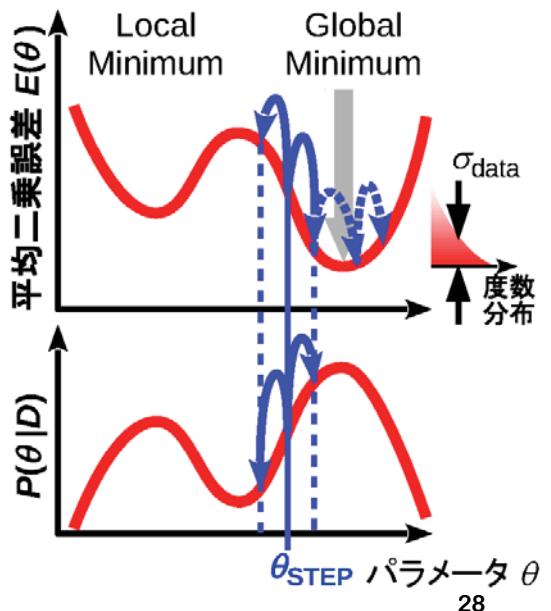


## ・MCMC法によるBayesianサンプリング



➤ データに重畳するノイズの大きさ  $\sigma_{\text{data}}$  が事後確率分布幅を決定する。

$$\underbrace{P(\theta|D)}_{\text{Posterior}} \propto \underbrace{\exp\left(-\frac{N}{\sigma_{\text{data}}^2} E(\theta)\right)}_{\text{likelihood}} \underbrace{P(\theta)}_{\text{Prior}}$$



SPring-8 2017/02/27

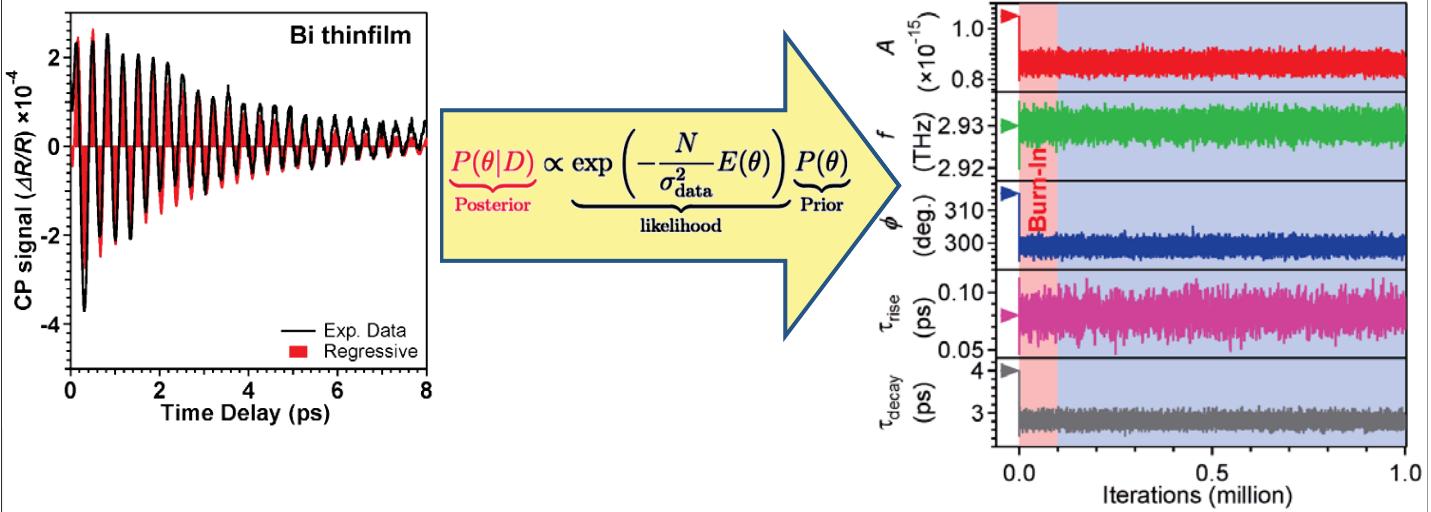
28

# 1モード減衰振動のベイズ推定

## 物理モデル

$$Y(t) = \frac{a}{\tau^d - \tau^r} \left[ \exp\left(-\frac{t}{\tau^d}\right) - \exp\left(-\frac{t}{\tau^r}\right) \right] \sin(2\pi ft + \phi)$$

- MCMCによるサンプリング履歴 (100万回, Burn-In = 10万回)



SPring-8 2017/02/27

29

# パラメータの推定精度

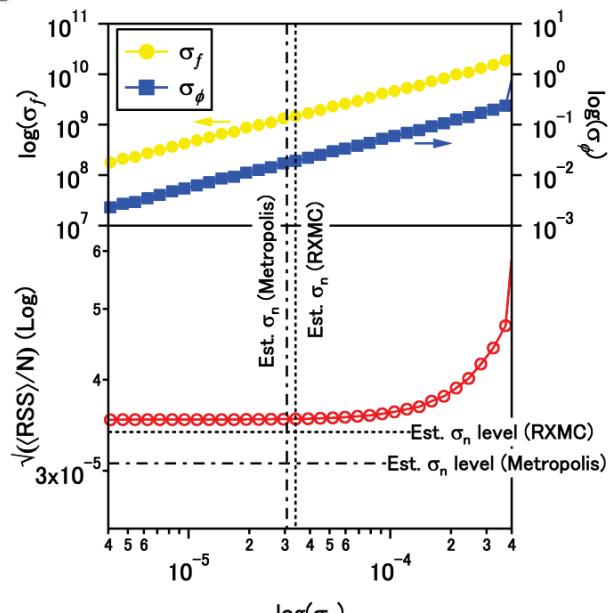
## 物理モデル

$$Y(t) = \frac{a}{\tau^d - \tau^r} \left[ \exp\left(-\frac{t}{\tau^d}\right) - \exp\left(-\frac{t}{\tau^r}\right) \right] \sin(2\pi ft + \phi)$$

- パラメターの推定精度

Bayesian inference			
$\theta$	mean value	$\pm \epsilon_\theta$	unit
$A$	8.63	$\pm 0.10$	
$f$	2930.04	$\pm 0.93$	GHz
$\phi$	5.219	$\pm 0.012$	radian
$\tau_r$	80.9	$\pm 4.8$	fs
$\tau_d$	2811	$\pm 51$	fs

S/N~16dB でも、 $f$  は不確定性幅の100倍精度で推定可能



SPring-8 2017/02/27

30

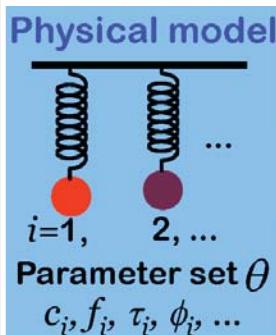
# スペースモデリング : SpM

データを数の少ない(スペースな)主要因で説明したい!

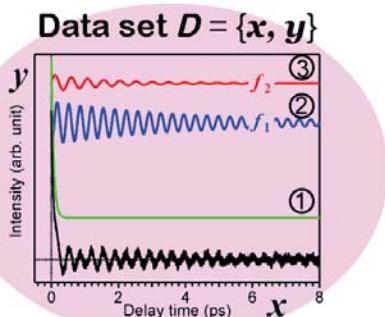
- ・どんな現象だ?  
→ 減衰振動だ!
- ・データには少ないと仮定!
- ・振幅の大きい要素から!

- ・実験データ  $\vec{y}$
- ・スペースな主要素  $\vec{x}$
- ・線形写像,  
現象の因果律に対応  $\vec{y} = \vec{A} \vec{x}$

実験結果を説明できて、なおかつ、 $\vec{x}$  の非ゼロ要素数( $\|\vec{x}\|_0$ )を最小化する  $\vec{x}$  の解を探査する。



Causality



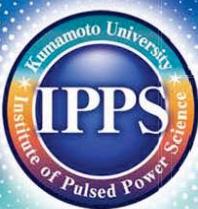
$p$ -th Norm  $\|\vec{x}\|_p \equiv \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}$

1<sup>st</sup> Norm  $\|\vec{x}\|_1 \equiv \left( \sum_{i=1}^n |x_i| \right)$

0<sup>th</sup> Norm  $\|\vec{x}\|_0 = \sum_{i=1}^n |x_i|^0$

SPring-8 2017/02/27

32



## LASSO 法 (Least Absolute Shrinkage and Selection Operator)

R. Tibshirani, *J. Royal Stat. Soc. B* **58**, 267 (1996).

平均二乗誤差  $E(\theta)$  と ペナルティー項として 1<sup>st</sup> ノルム  
(0<sup>th</sup> ノルムの代わりに)との和を最小化する

$$E(\theta) = \frac{1}{2N} \sum_{i=1}^N [y_i - f(x_i; \theta)]^2$$

ベクトルの非スペース性  
のペナルティー項

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

$\lambda$ : スペース性を制御するハイパーパラメータ

- $\lambda = 0$ : 最小二乗法と同じ解が得られる  
(ノイズまで合わせに行く過剰フィッティング結果)
- $0 < \lambda < \infty$ : ある規範に基づいたスペース解
- $\lambda = \infty$ : 主要素が全くない  
(全てのデータはノイズである)

SPring-8 2017/02/27

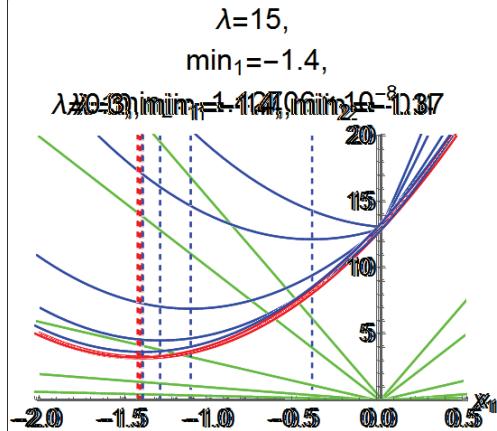
33

# 1次ノルムのペナルティー項で スパース解を得る ( $K=1$ )

- 簡単な例

$$\underbrace{\begin{pmatrix} \vec{y} \\ -2 \\ 3 \end{pmatrix}}_{\vec{A}\vec{x}} = \underbrace{\begin{pmatrix} 2 & -1 & 3 \\ -1 & 1 & 2 \end{pmatrix}}_{\vec{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\vec{x}}$$

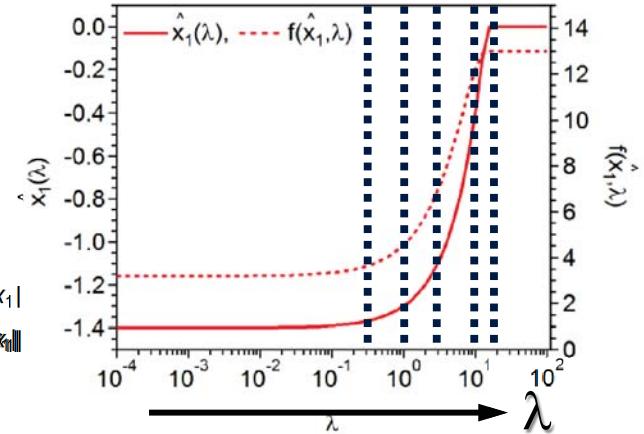
$x_1 \neq 0, x_2 = 0, x_3 = 0$  の場合



LASSO法

$$f(\vec{x}, \lambda) \text{ の極小点}$$

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



SPring-8 2017/02/27

34

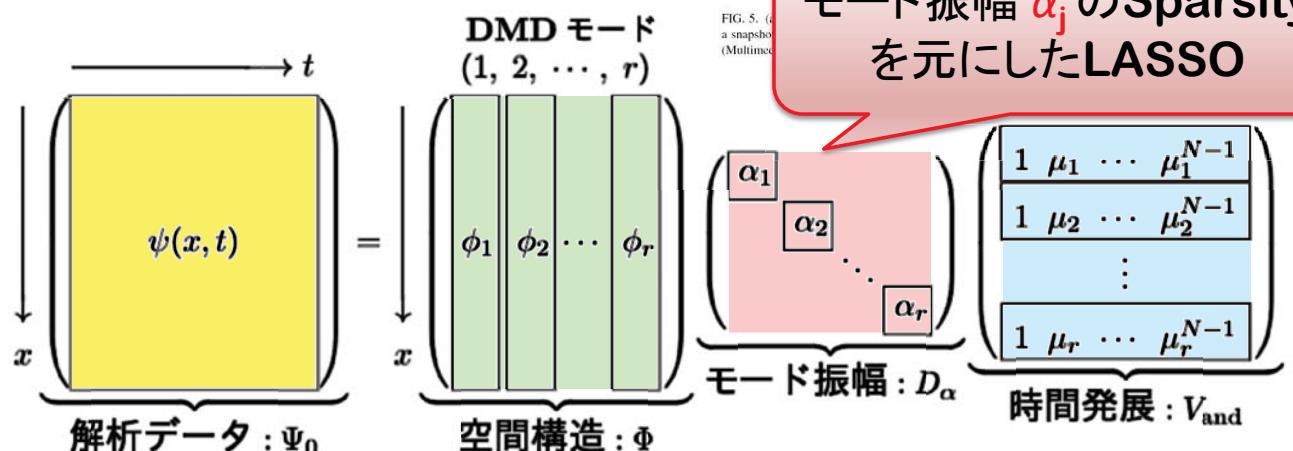
## SpDMD (Sparsity-promoting Dynamic Mode Decomposition)

DMD:流体振動モード解析に用いられる。 Phys. Fluids 26, 024103(2014)

- 空間・時間変化する信号  $\Psi_0, \Psi_1$

$$\psi(x, t) = \sum_{j=1}^r \alpha_j \phi_j(x) e^{\sigma_j t} (\cos \omega_j t + i \sin \omega_j t)$$

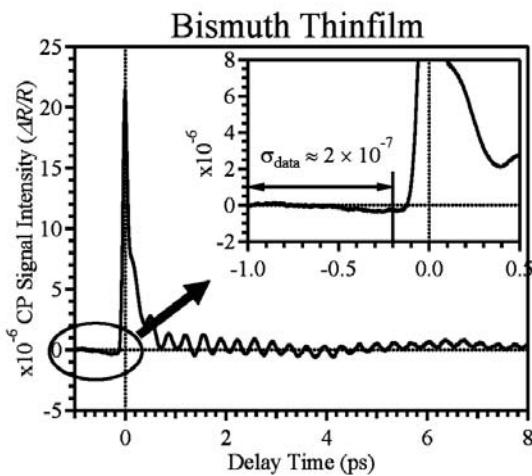
基底関数 =  $\exp(\sigma_j t + i \omega_j t)$   
指数減衰 / 減衰振動によるモード展開



SPring-8 2017/02/27

35

# SpDMD 対象とする実験データ



ビスマス薄膜 @ R.T.

- $\lambda \sim 750 \text{ nm}, 60 \text{ fs}$
- $I_{\text{pump}} = 185 \text{ mW}, I_{\text{probe}} = 3 \text{ mW}$
- $A_{1g} \sim 2.9 \text{ THz}, \phi_{A_{1g}} = ?$
- $E_g \sim 2.1 \text{ THz}, \phi_{E_g} = ?$

- 実験的アーティファクト
  - $t \sim 0 \text{ ps}$ : 指数関数減衰
  - Backgroundの長周期ゆらぎ
- データに重畳するノイズ

$$\sigma_{\text{data}} = 2 \times 10^{-7}$$

SPring-8 2017/02/27

36

# SpDMD 推定されたパラメータ

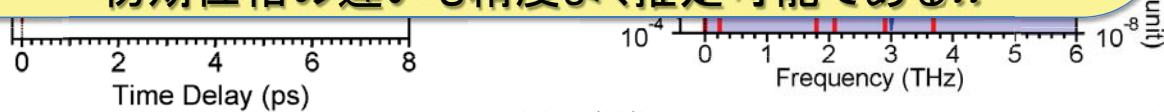
- 6つの指数減衰/減衰振動で実験データをよく再現する!!

$$y(t) = \sum_{\nu=1}^6 R_{\nu} \exp\left(-\frac{t}{\tau_{\nu}}\right) \cos(2\pi f_{\nu} t + \phi_{\nu})$$

mping1  
background  
mping3  
mping2

Mode	Frequency	Decay Time	Amplitude	Phase deg cos
$A_{1g}$	2.90318	4.35434	9.49199e-007	175.829
$E_g$	2.09007	1.81861	3.75284e-007	108.183
Damping 1	0	0.267535	1.23342e-005	0
Damping 2	3.69751	0.0247972	2.34792e-005	62.6367
Damping 3	1.7961	0.502478	7.91816e-007	-97.3564
Background	0.233482	7.58546	4.34694e-007	149.961

初期位相の違いも精度よく推定可能である!!

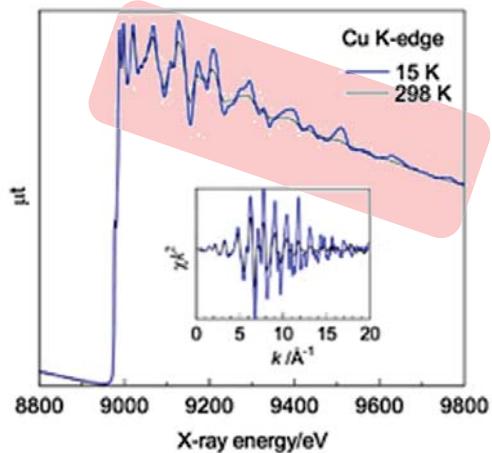
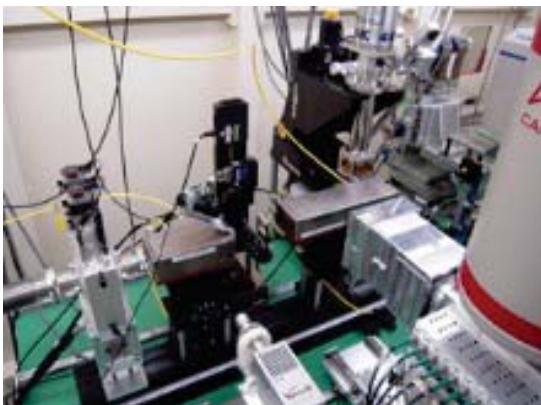


SPring-8 2017/02/27

38

# SpM on EXAFS

- EXAFS (Extended x-ray absorption fine structures: 広域X線吸収微細構造)



選択した原子種(X線吸収端)近傍の原子距離レベルのミクロ構造  
試料形態は選ばず、汎用的で産業レベルでも利用が極めて多い

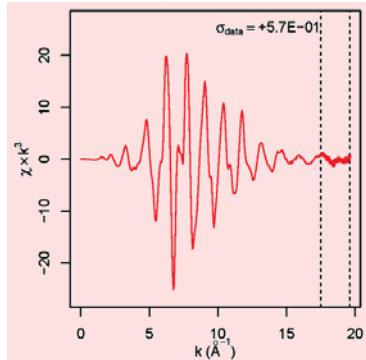
SPring-8 2017/02/27

39

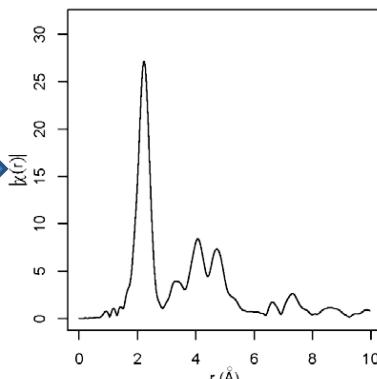
## 従来の解析法: フーリエ変換

- $\chi(k)k^3$ をフーリエ変換

$$\chi(k)k^3$$



フーリエ変換



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

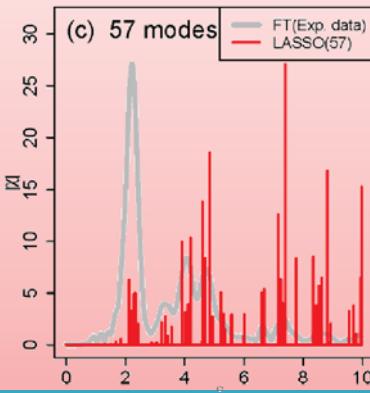
Radial Distribution

# EXAFSのSpM解析

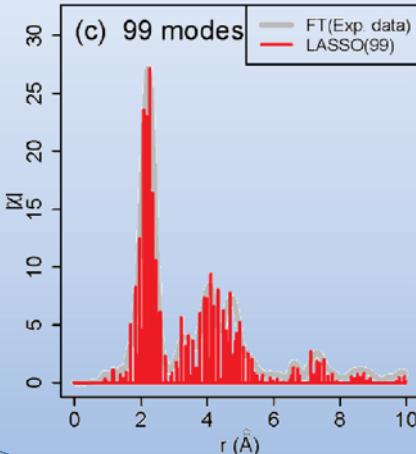
$$\hat{\vec{x}} = \arg \min_{\vec{x}} \left( \|\vec{y} - \vec{A} \vec{x}\|_2^2 + \lambda \|\vec{x}\|_1 \right)$$

## • EXAFS

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



基底関数の改善で、  
より本質が見えてきた！



SPring-8 2017/02/27

49

# まとめ

## データ駆動科学

### • ベイズ推定

ベイズの定理を用いて、因果律をさかのぼる！

パラメータの推定精度を評価できる

- 励起子スペクトル: Eg, Ry
- XPS: 結合エネルギーの化学シフト
- 過渡特性: 振動数

### • スパースモデリング

我々の頭脳が行う認識プロセスを定式化！

- 超高速ダイナミクス: SpDMD
- EXAFS: Sp-Adv. FT
- XFH: SpM on XFH

データ解析の価値観が一変する！