

情報科学とXAFSの融合で駆動する物質科学

JASRI/SPring-8 水牧 仁一朗

なぜ情報学との融合？

インフォマティクス

近年、めざましい発展を見た分野 ← 人工知能(AI)の導入

Alpha-Go(ゲーム) → 科学

マテリアルズインフォマティクス

他にもxxインフォマティクス

2015年 NIMSの「MI²I」の発足 **データを中心に**

キーテクノロジー

ベイズ的モデル選択、スパースモデリング、ディープラーニング



データ駆動科学

東大・岡田ら <http://sparse-modeling.jp>

Y. Igarashi, K. Nagata, M. Okadaら: J.Phys.conf 699(2016)012001

五十嵐、永田、岡田ら: 応用統計学

実験計測計算データの背後にある潜在的構造の抽出に関して、
データが対象とする学問に依存しない普遍的な学問体系

放射光とインフォマティクス

計測インフォマティクス CREST・さきがけ 2016~

計測技術と高度情報処理の融合によるインテリジェント計測・解析手法の開発と応用 放射光分野への導入がすでに始まっている

データ駆動科学G

データ駆動科学と物性理論的研究

赤井一郎(熊大教授) 専門は、データ駆動科学と光物性の実験。様々なスペクトルのベイズ分光・スパースモデリングに取り組んできました。

安仁屋勝(熊大教授) 固体中のイオン伝導、アモルファス物質の物性の物性理論研究。

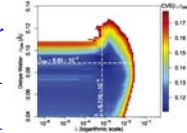
下條冬樹(熊大教授) 非晶質物質の構造、電子状態の第一原理計算による物性研究。

細川伸也(熊大教授) 3D活性サイト科学の放射光を用いた物性実験研究。

中島陽一(熊大) いた物性実験研究。

原正大(熊大准) 実験研究。

特別研究員 デ



EXAFSデータだけか

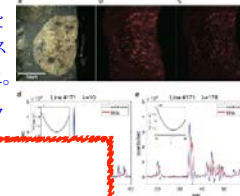
特徴量画像解析G

特徴量画像のデータ駆動科学解析

青西 亨(東工大准教授) 専門は、データ駆動科学と情報統計力学。超広視野二光子顕微鏡やLA-ICP-質量分析のイメージングデータ解析に取り組んできました。

庄野 逸(電通大教授) 専門は画像処理とデータ駆動科学。医用画像解析やテキスト画像特徴の分析に取り組んできました。

大森敏明(神戸大准教授) 専門はデータ



素イメージの解像子の検出。JAAS process(CCC)

データ駆動科学による

高次元X線吸収計測の革新

X線吸収

多角的X線吸収
妹尾与志

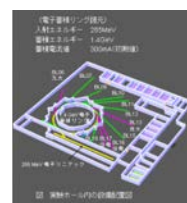
料解析に関する研究。様々な材料解析に取り組んできました。

岡島敏浩(SAGA-LS主任研究員) 放射光計測研究及びXANESスペクトル解析技術の高度化研究。

小林英一(SAGA-LS主任研究員) 表面実験のための新しい装置開発と表面物性研究。

吉村大介(SAGA-LS副主任研究員) ビームライン光学設計と軟X線領域の光電子分光研究および吸収分光研究。

瀬戸山寛之(SAGA-LS副主任研究員) 放射光を用いた機能性材料の物性研究。
特別研究員 放射光計測と解析研究。



九州シンクロトロン光研究センター(SAGA Light Source)の概略

宇留賀朋哉(JASRI副部門長) X線分光を中心とした *in situ*同時系列計測放射光マルチ分析法の開発。

関澤央輝(JASRI研究員) 専門はX線分光光学で、電池材料の化学状態イメージングの研究。

鈴木基寛(JASRI主幹研究員) ナノビーム走査型X線顕微分光法を開発し、磁性材料や触媒材料の磁性・機能イメージングの研究。

新田清文(JASRI研究員) X線顕微分光イメージング法を用いた機能性材料の特徴量イメージングを目指した装置及び解析法の開発。

特別研究員 顕微計測と解析研究。

析G

X線分光光学で

の空間構造との相関を明らかにすること。

NdFeB 焼結磁石のNd-L₂ 端 顕微 XMCDによる磁区

像。 Acta Mat. 106, 155, (2016)

5 μm

106, 155, (2016)

106, 155, (2016)

なぜ情報学との融合？

実験屋の立場から

✓ 間接測定される物理量の精度評価

物理モデル妥当性の評価 後ろめたさを一掃

✓ ノイズに埋もれたシグナルの顕在化

極限的測定の必要精度

✓ 実験条件の最適化・極限へ

実験の効率化

✓ 理論研究へのフィードバック

理論モデルの良し悪し

なぜ情報学との融合？

- ✓ 間接測定される物理量の精度評価
- ✓ ノイズに埋もれたシグナルの顕在化
- ✓ 実験条件の最適化・極限へ
- ✓ 理論研究へのフィードバック

ベイズのモデル選択

~X線光電子分光スペクトル解析~

~4f希土類元素の3d-XPS~

共同研究者

東京大・新領域 本武陽一

東京大・新領域 岡田真人

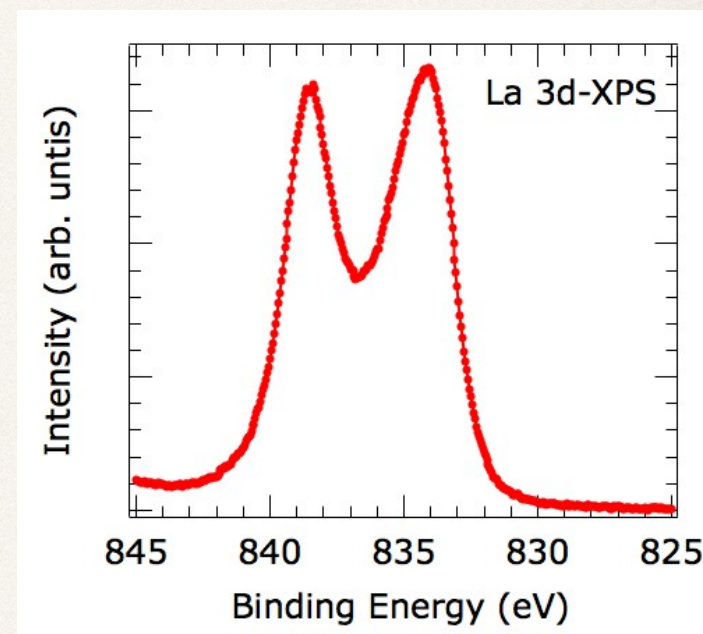
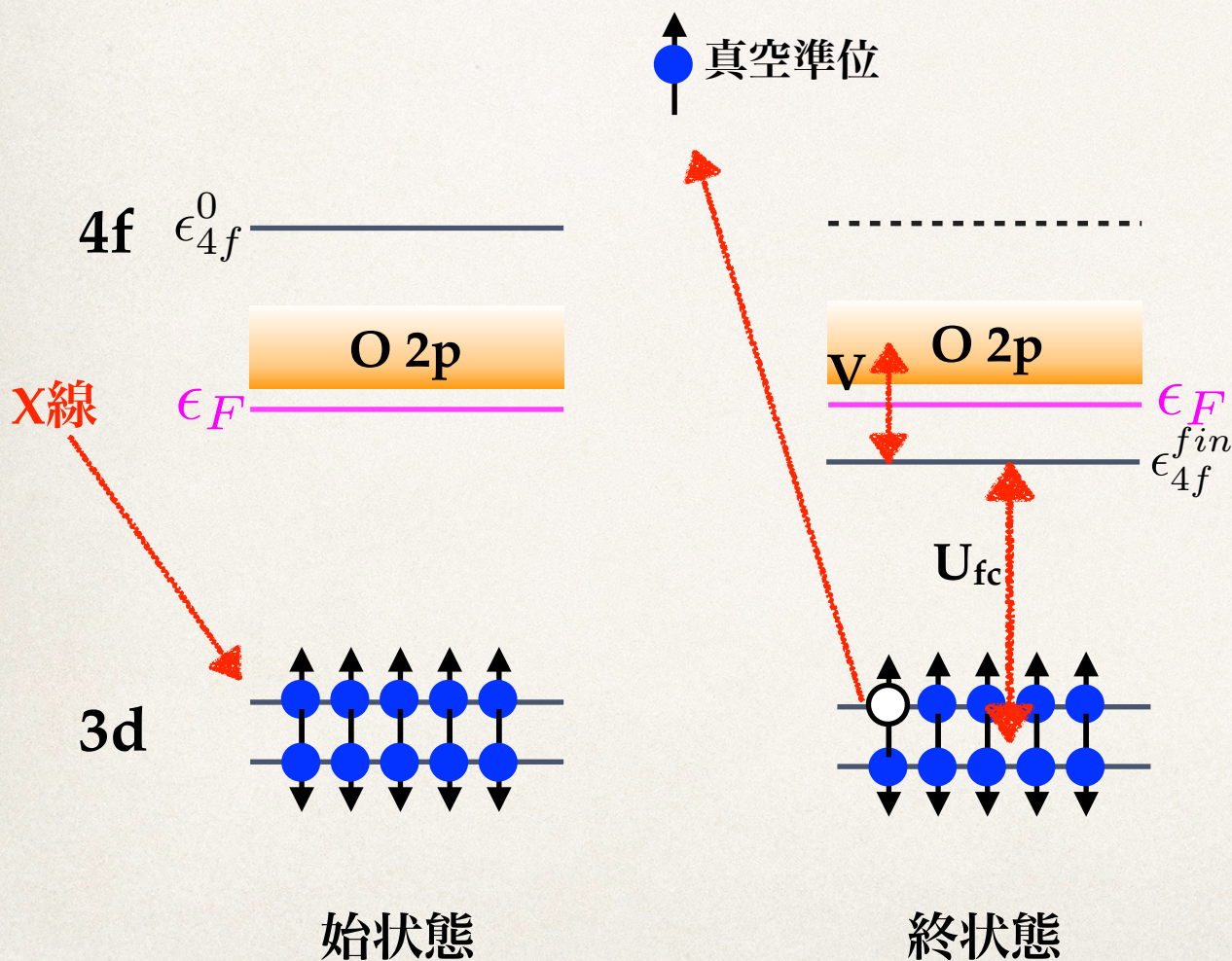
熊大・パルス研 赤井一郎

J. Phys. Soc. Jpn. 88(2019)034004

内殻X選光電子分光

X線光電子分光の遷移過程

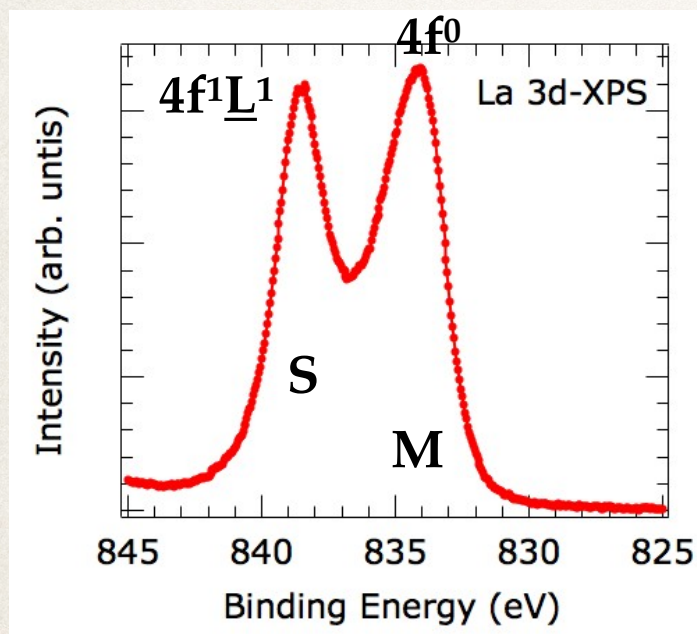
4f希土類元素の3d内殻の場合



スペクトル解析

実験屋がすると

X線光電子分光の遷移過程を表すHamiltonianを考慮



ピークがいくつあるか

ピーク位置がどこか

ピーク形状は？ローレンツ・ガウス？

ピーク幅がどれくらいか



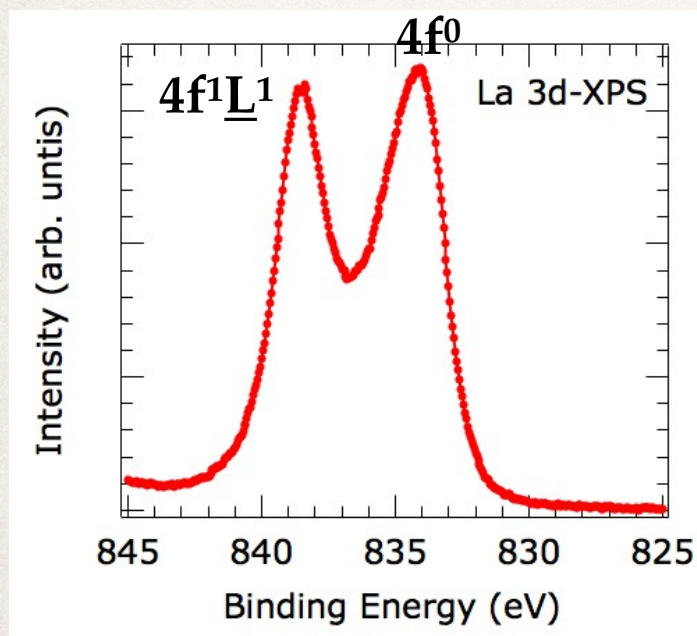
$4f^0$ poorly-screened

$4f^1 \underline{L}^1$ well-screened



電子状態を理解した

X線光電子分光の遷移過程を表すHamiltonianを考慮



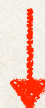
しっかりと

$$H = \sum_k \epsilon_k a_k^+ a_k + \sum_\nu \epsilon_f a_{f\nu}^+ a_{f\nu} + \epsilon_c a_c^+ a_c$$
$$+ \frac{V}{\sqrt{N_f}} \sum_{\nu,k} (a_{f\nu}^+ a_k + a_k^+ a_{f\nu}) - U_{fc} \sum_\nu a_{f\nu}^+ a_{f\nu} (1 - a_c^+ a_c)$$
$$+ U_{ff} \sum_{\nu > \nu'} a_{f\nu}^+ a_{f\nu} a_{f\nu'}^+ a_{f\nu'}$$

$$I(\omega; \theta) = \sum_f |\langle f | a_c | g \rangle|^2 \frac{\Gamma/\pi}{(\omega - E_f(\theta) + E_g(\theta))^2 + \Gamma^2}$$

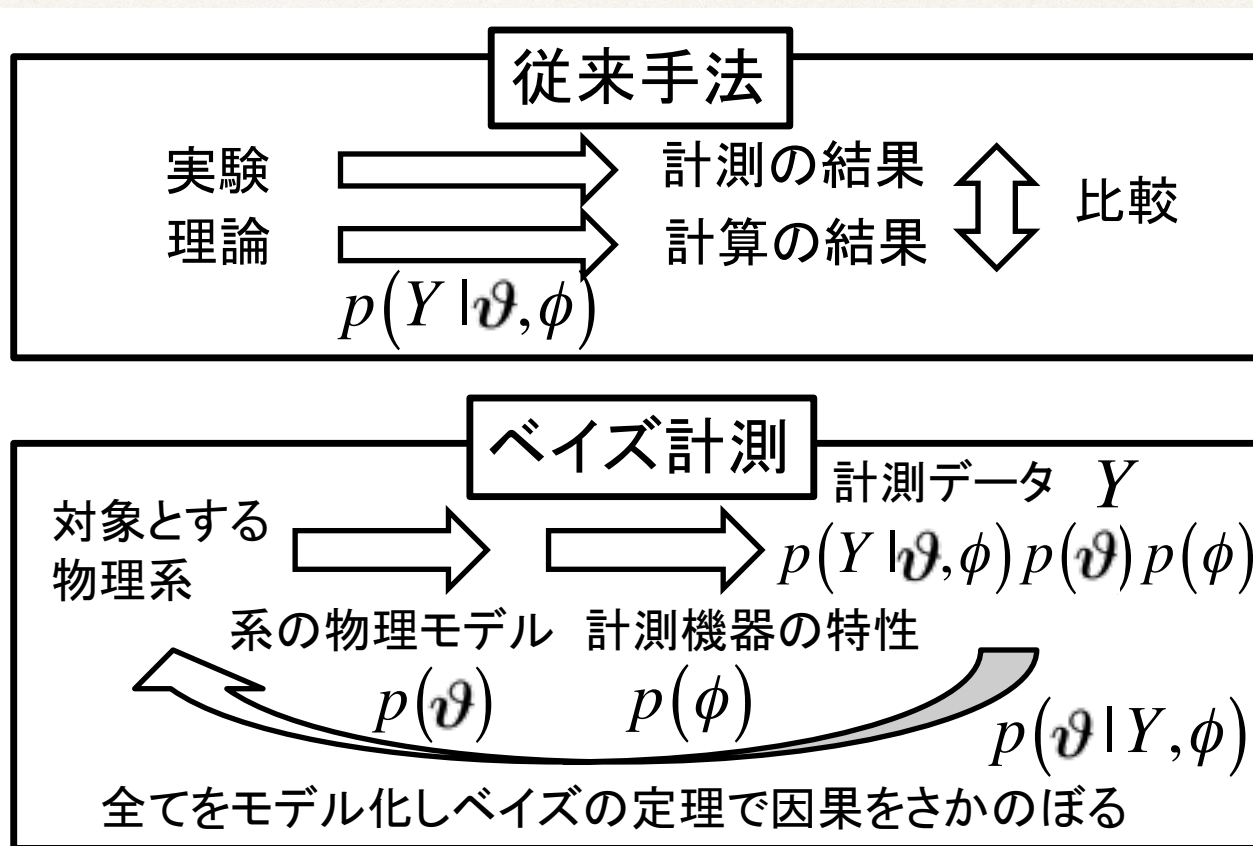
$$\theta = \{\epsilon_k, \epsilon_f, V, U_{ff}, U_{fc}, \Gamma\}$$

実験スペクトルを再現するようにパラメータセット θ をハンドチューン



電子に関する物理量抽出→電子状態を理解した

ベイズ推定による自動化



[五十嵐, 竹中, 永田, 岡田 応用統計学, 2016]

電子に関するハミルトニアンを自動選択

電子に関する物理量パラメータを自動チューニング

ベイズの定理

因果律に基づいて

$$P(\theta, D) = P(D|\theta)P(\theta)$$

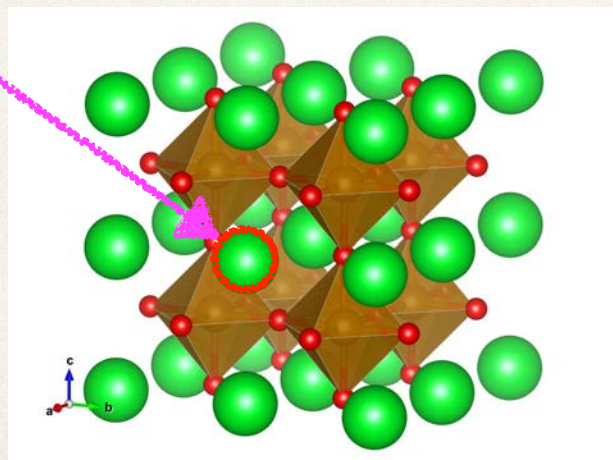
$$= P(\theta|D)P(D)$$

誤差論に基づけば、
正規分布

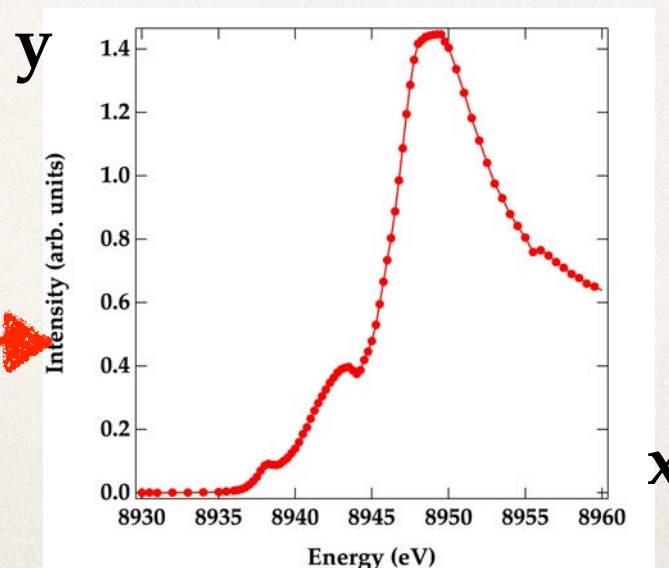
欲しいもの!

確定値(測定がすめば)

| $f(x_i; \theta)$ | θ | y_i |
|------------------|----------|-------|
| 誤差論 | 確定的 | 確率的 |
| ベイズ | 確率的 | 確定的 |



因果律



パラメータセット θ r , Partial DOS, ...

実験データセット $D = \{x, y\}$

モデル：データをどう考えるか

実験結果(データ) y_i : 出力

入力の応答とノイズの足し合わせ

$$y_i = f(x_i; \theta) + n_i$$

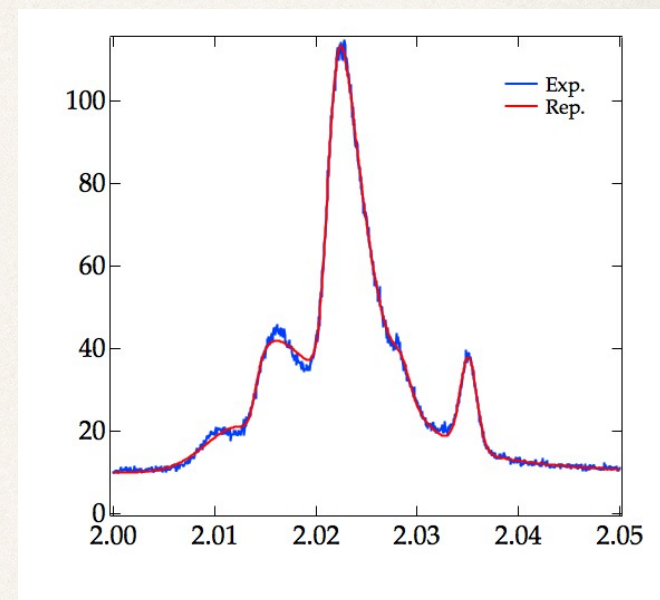
ノイズが正規分布すると仮定

$$P(y_i | \theta) \propto \exp\left(-\frac{(y_i - f(x_i; \theta))^2}{2\sigma_{data}^2}\right)$$

各点の出力 y_i が独立であるとする

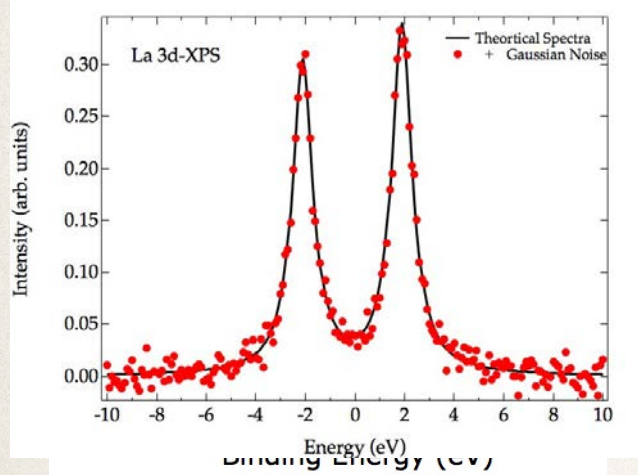
$$Y = \{y_1, \dots, y_N\}$$

$$P(Y | \theta) = \prod_{i=1}^N P(y_i | \theta) \propto \exp\left(-\frac{NE(\theta)}{\sigma_{data}^2}\right) \quad E(\theta) = \frac{1}{2N} \sum_{i=1}^N (y_i - f(x_i; \theta))^2$$



目標

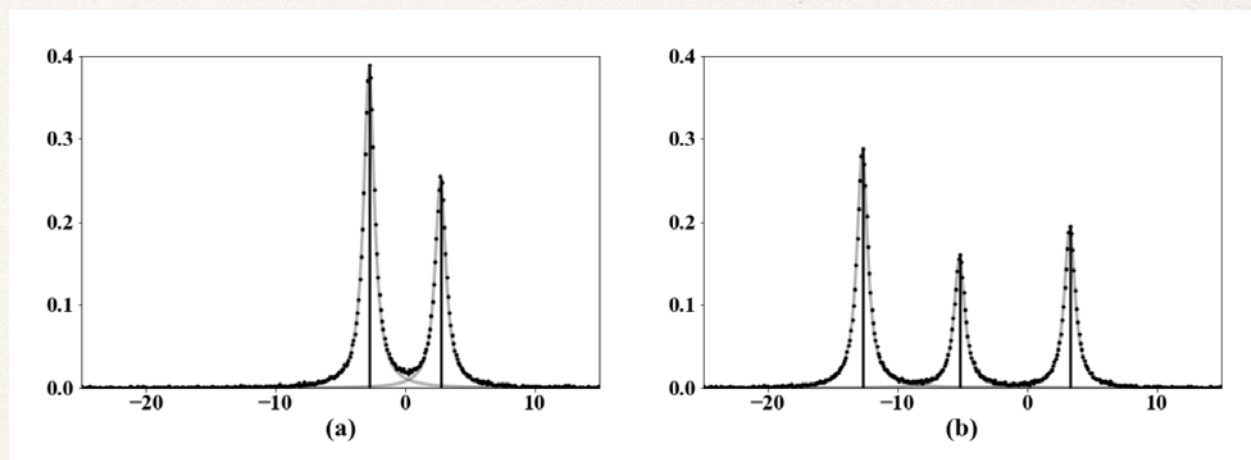
4f希土類元素の3d内殻の場合



実験データ



人工実験データ



2準位クラスター模型

3準位クラスター模型

| Parameter | Δ | V | U_{ff} | U_{fc} | Γ |
|--------------------------------|-------------|------|----------|----------|----------|
| La ₂ O ₃ | 12.5 | 0.57 | 10.5 | 12.7 | 0.5 |
| CeO ₂ | 1.6 | 0.76 | 10.5 | 12.5 | 0.7 |

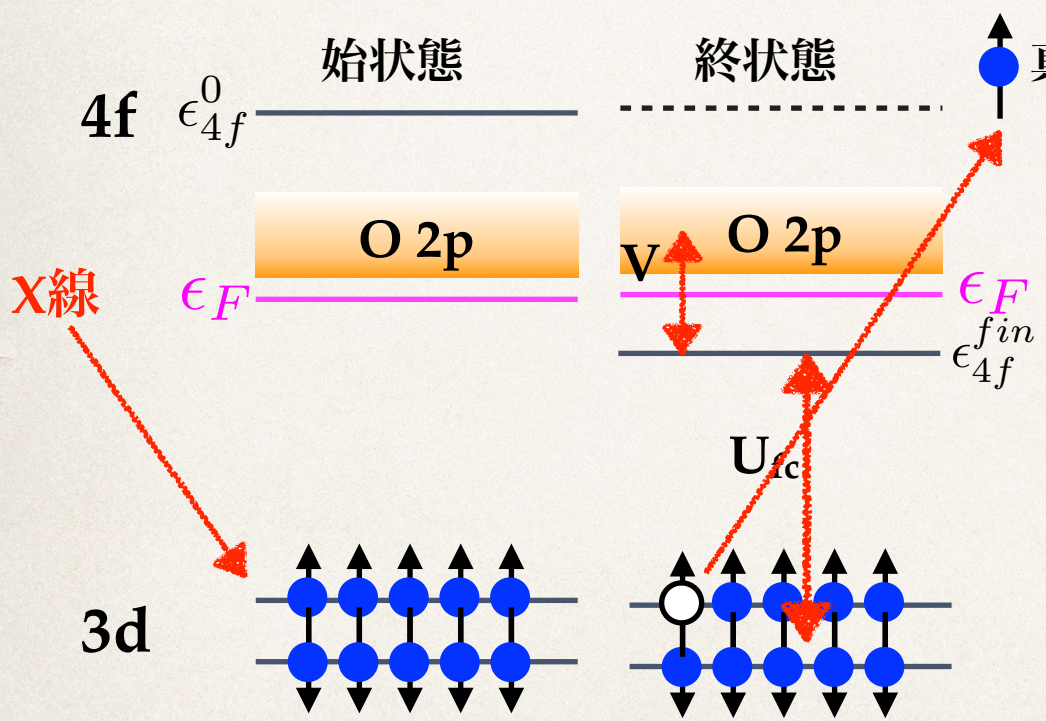
ハミルトニアンを自動選択

パラメータを自動チューニング

2準位クラスターモデル

La₂O₃を模擬
$$H_2^I = \sum_k \epsilon_k a_k^+ a_k + \sum_\nu \epsilon_f a_{f\nu}^+ a_{f\nu} + \epsilon_c a_c^+ a_c$$

$$+ \frac{V}{\sqrt{N_f}} \sum_{\nu,k} (a_{f\nu}^+ a_k + a_k^+ a_{f\nu}) - U_{fc} \sum_\nu a_{f\nu}^+ a_{f\nu} (1 - a_c^+ a_c)$$



$$I(\omega; \theta_2) = \sum_{j=0}^1 |\langle f_j | a_c | g \rangle|^2 \frac{\Gamma_j / \pi}{(\omega - E_f(\theta_2) + E_g(\theta_2))^2 + \Gamma_j^2}$$

$$|g\rangle = |f^0\rangle \quad |f\rangle = c_0 |d^9 f^0\rangle + c_1 |d^9 f^1 \underline{L}^1\rangle$$

2準位クラスターモデル

La₂O₃を模擬

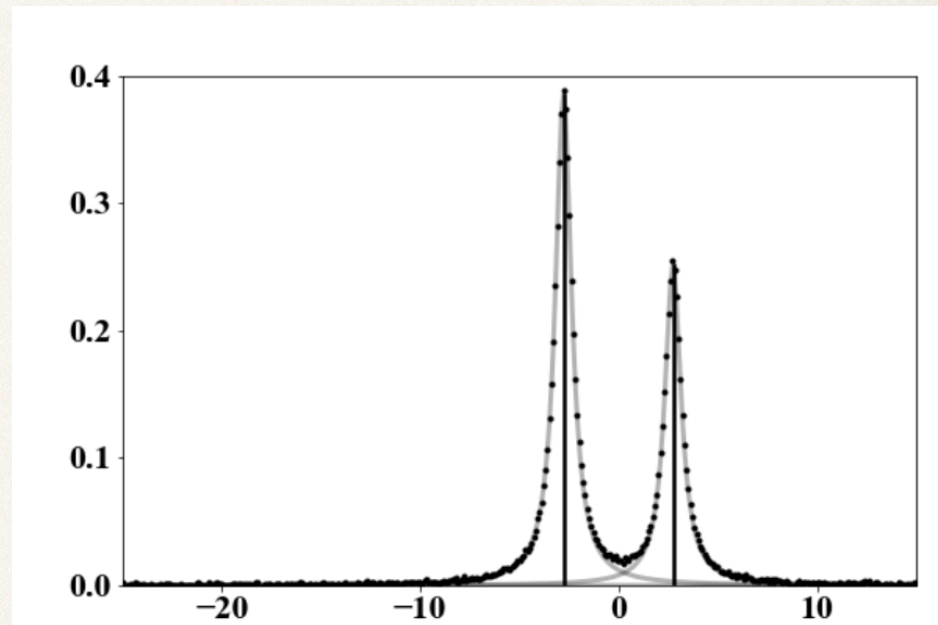
始状態 $|g\rangle = |f^0\rangle$

始状態エネルギー: $E_g=0$

終状態 $|f\rangle = c_0|d^9 f^0\rangle + c_1|d^9 f^1 \underline{L}^1\rangle$

$$H_2^F = \begin{pmatrix} 0 & \sqrt{N_f}V \\ \sqrt{N_f}V & \epsilon_f^0 - \epsilon_L - U_{fc} \end{pmatrix}$$

終状態エネルギー: E_{f1} E_{f2} の2個



Fermiの黄金律とローレンツ関数で畳み込み

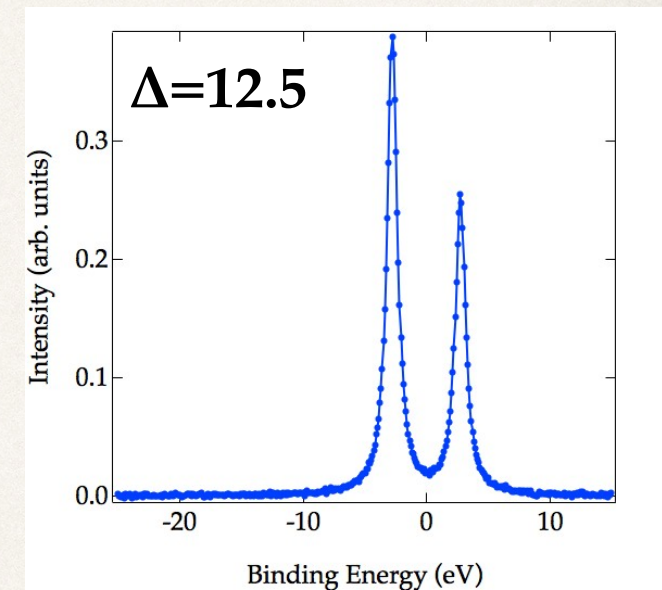
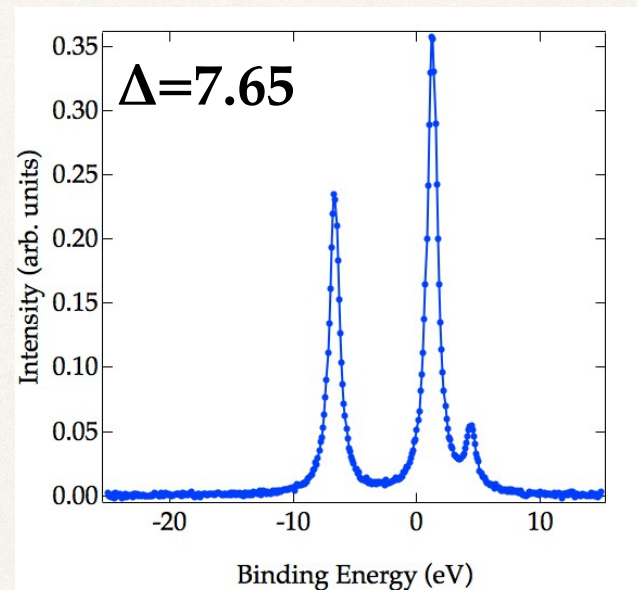
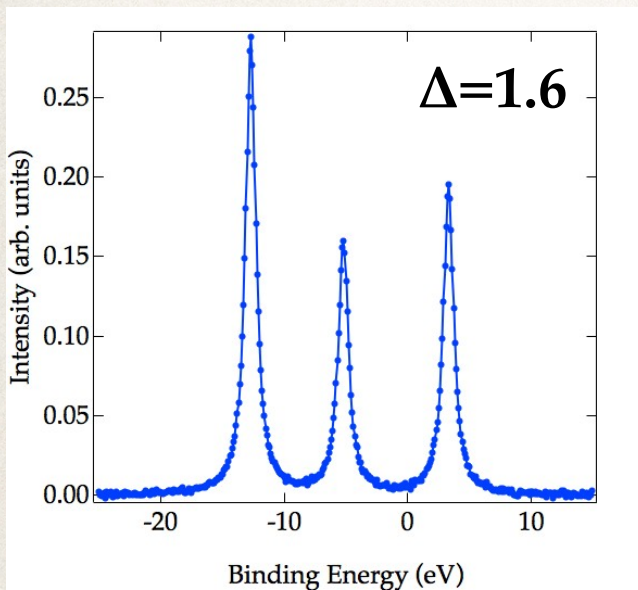
$$I(\omega; \theta_2) = \sum_{j=0}^1 |\langle f_j | a_c | g \rangle|^2 \frac{\Gamma_j / \pi}{(\omega - E_f(\theta_2) + E_g(\theta_2) - b)^2 + \Gamma_j^2}$$

3準位クラスターモデル

3準位クラスターモデルの特徴

| Parameter | Δ | V | U_{ff} | U_{fc} | Γ |
|--------------------------------|----------|------|----------|----------|----------|
| La ₂ O ₃ | 12.5 | 0.57 | 10.5 | 12.7 | 0.5 |
| CeO ₂ | 1.6 | 0.76 | 10.5 | 12.5 | 0.7 |

Δ (=4fレベルとリガンドのエネルギー差)を変化させると



スペクトル本数が増える 3本から2本へ

Ce化合物だけど
La的な電子構造

2準位模型と同様になってくる

ベイズ推論

我々が知っていること $P(I|\theta_k, H_k)$

有効模型(Hamiltonian)とその回帰式

$$I(\omega; \theta_3) = \sum_{j=0}^2 |\langle f_j | a_c | g \rangle|^2 \frac{\Gamma_j / \pi}{(\omega - E_f(\theta_3) + E_g(\theta_3))^2 + \Gamma_j^2}$$

$H_k \rightarrow$ パラメータ $\rightarrow I$

\downarrow ベイズの定理

我々が知りたいこと $P(\theta_k|I, H_k)$

スペクトルデータIに対する H_k の尤もらしさ

$$P(H_k|I)$$

スペクトルデータIに対する H_k モデルの推定値

$$P(\theta_k|I, H_k)$$

$H_k \leftarrow$ パラメータ $\leftarrow I$

ベイズの定理

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

知りたいこと

$\theta \rightarrow D$ 因果律

\downarrow 逆転

$D \rightarrow \theta$

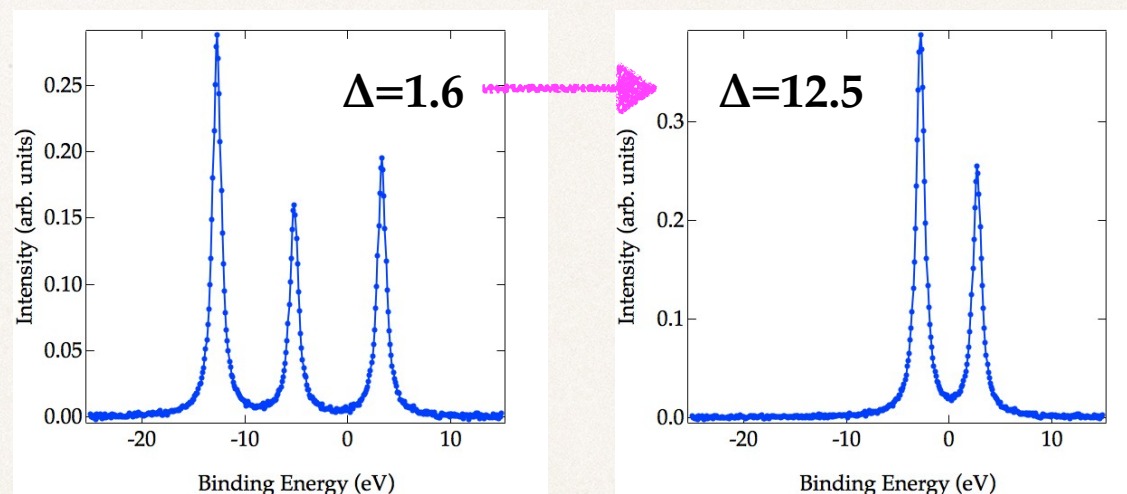
我々が行いたいプロセス

データ駆動科学

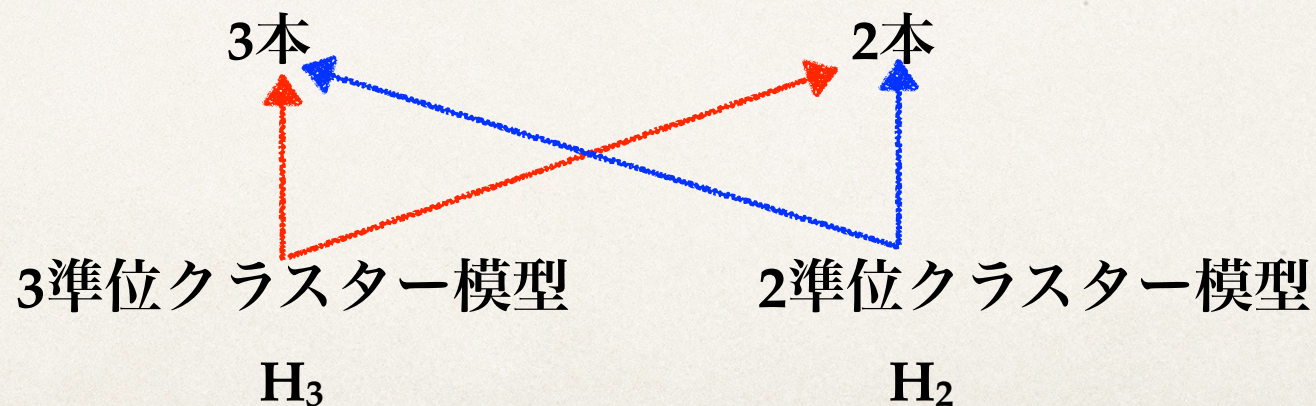
問題設定

3準位クラスター模型にて人工データを生成

生成モデル



認識モデル



提案手法

我々が知っていること $P(I|\theta_k, H_k)$

$$P(I|\theta_k, H_k) = \prod_{i=1}^N P(I(\omega_i)|\theta_k, H_k)$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left\{-\sum_{i=1}^N \left[\frac{1}{2\sigma_{noise}^2} (I(\omega_i; \theta) - I(\omega_i; \theta_k))^2\right]\right\}$$

我々が知りたいこと $P(\theta_k|I, H_k)$

$$\propto P(\theta_k|I, H_k)$$

$$P(H_k|I) \propto \int_{-\infty}^{\infty} P(I|H_k, \theta_k) P(\theta_k|H_k) d\theta_k$$

メトロポリス法により

$$= \left(\frac{1}{2\pi\sigma_{noise}^2}\right)^{N/2} \int_{-\infty}^{\infty} \exp[-NE(\theta_k)] P(\theta_k|H_k) d\theta_k$$

分配関数

最大事後確率推定

ここで $E(\theta_k) = \frac{1}{2N\sigma_{noise}^2} \sum_{i=1}^N [I(\omega_i; \theta) - I_k(\omega_i; \theta_k)]^2$

交換レプリカモンテカルロ法

$$P(H_k|I) \propto \left(\frac{1}{2\pi\sigma_{noise}^2} \right)^{N/2} \int_{-\infty}^{\infty} \exp[-NE(\theta_k)] P(\theta_k|H_k) d\theta_k$$

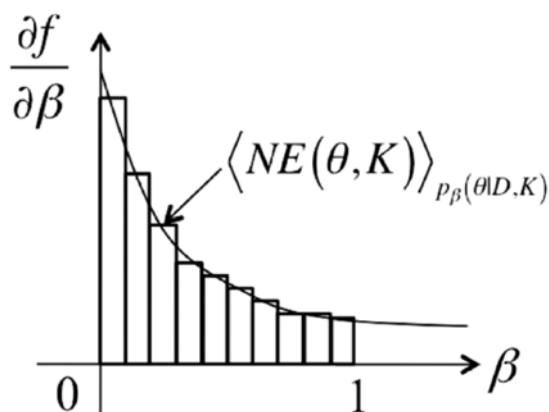
分配関数

この多重積分は計算が困難 メトロポリス法→局所解に落ち込む

(ベイズ)自由エネルギー

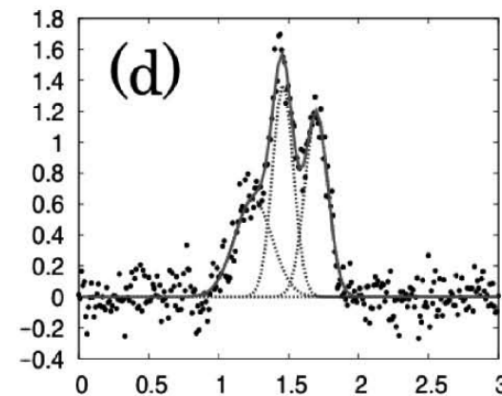
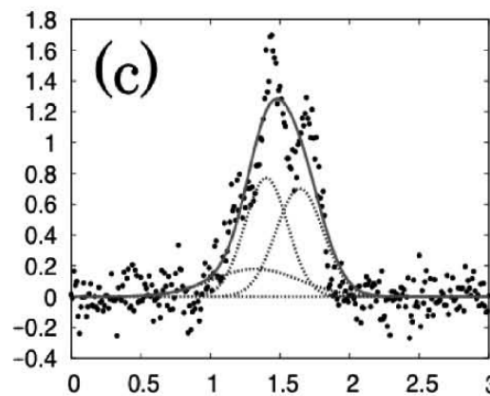
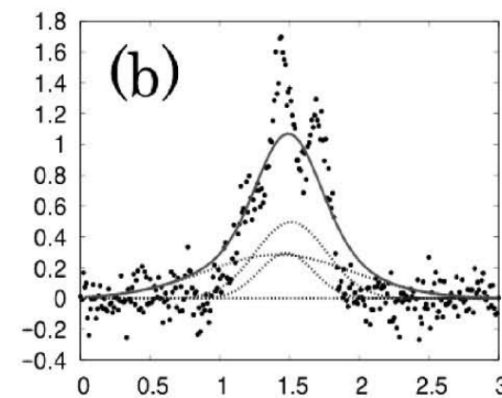
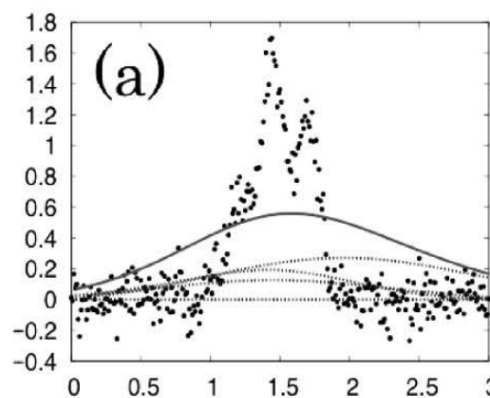
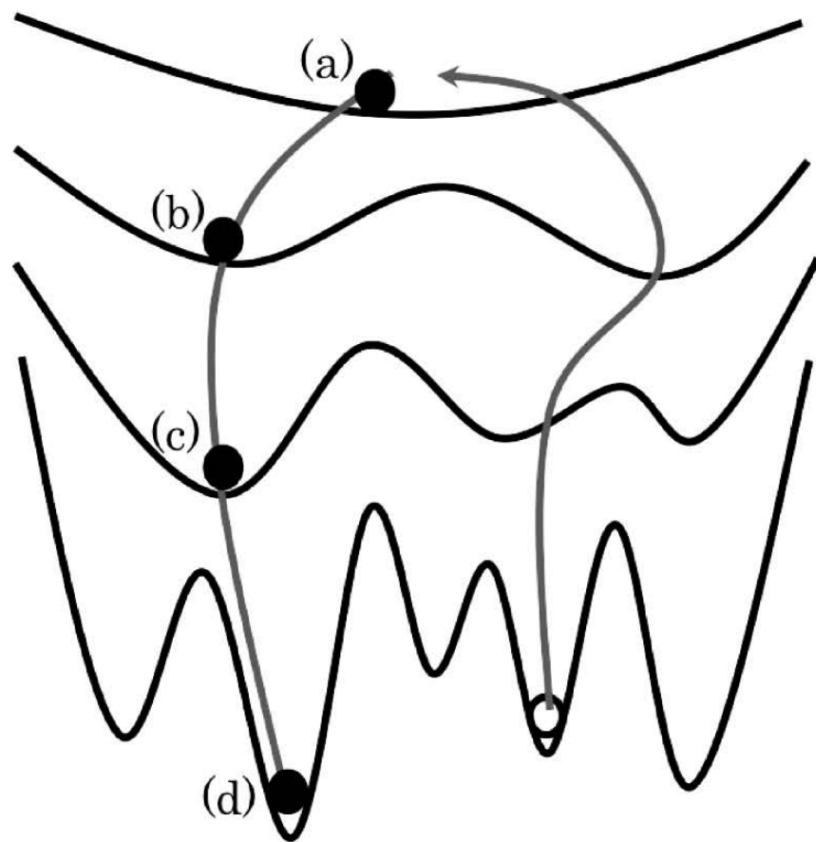
$$\begin{aligned} F(H_k) &= \log \int_{-\infty}^{\infty} \exp[-NE(\theta_k)] P(\theta_k|H_k) d\theta_k \\ &= \int_0^1 \frac{\partial}{\partial \beta} \left\{ \int_{-\infty}^{\infty} \exp[-N\beta E(\theta_k)] P(\theta_k|H_k) d\theta_k \right\} d\beta \\ &= \int_0^1 \int_{-\infty}^{\infty} \underline{NE(\theta_k) P(\theta_k|I, \beta)} d\theta_k d\beta \\ &= \int_0^1 \int_{-\infty}^{\infty} \langle NE(\theta_k) \rangle_{P(\theta_k|I, \beta)} d\beta \end{aligned}$$

平均二乗誤差
の期待値



$$P(\theta_k|I, \beta) = \frac{\exp[-N\beta E(\theta_k)] P(\theta_k|H_k)}{\int_{-\infty}^{\infty} \exp[-N\beta E(\theta_k)] P(\theta_k|H_k) d\theta_k}$$

交換レプリカモンテカルロ法



heating & annealing を繰り返す

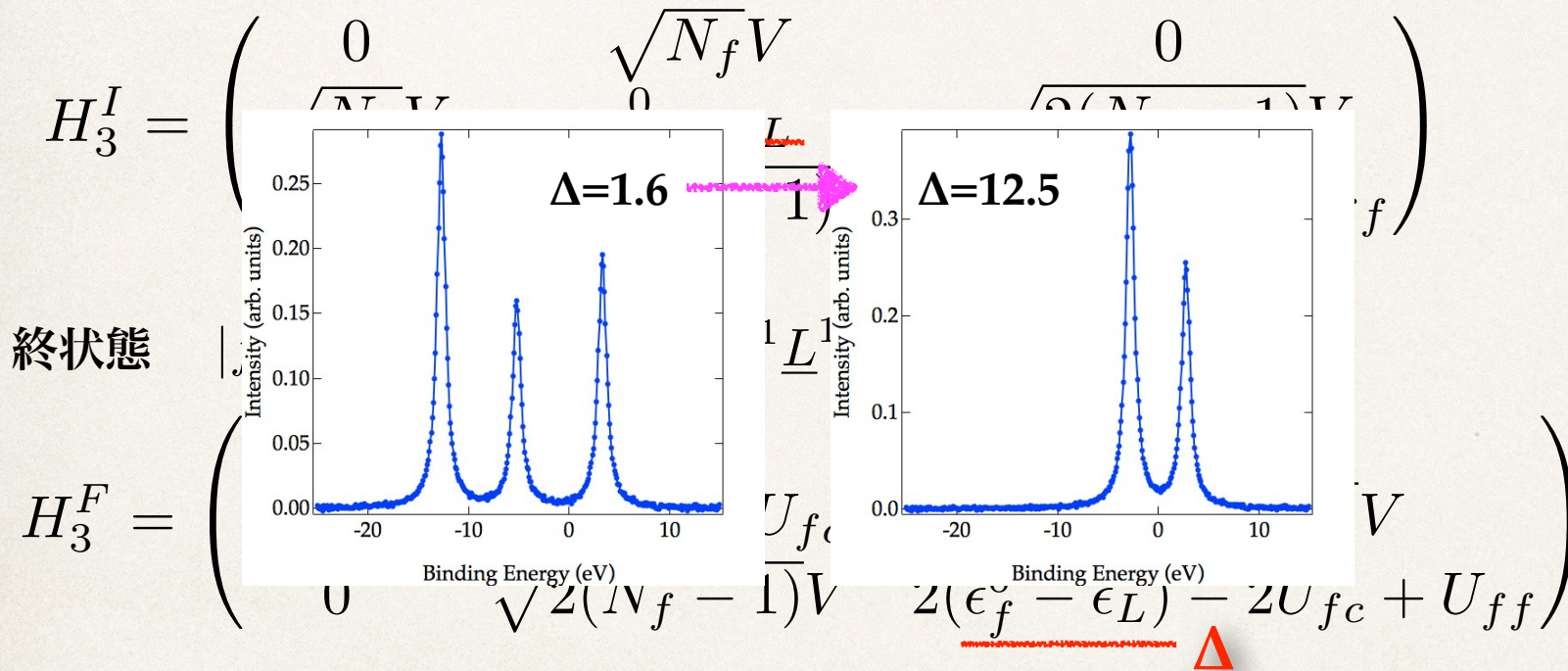
メトロポリス法→局所解に落ち込むことを防ぐ

3準位クラスターモデル

CeO₂を模擬

| Parameter | Δ | V | U_{ff} | U_{fc} | Γ |
|--------------------------------|----------|------|----------|----------|----------|
| La ₂ O ₃ | 12.5 | 0.57 | 10.5 | 12.7 | 0.5 |
| CeO ₂ | 1.6 | 0.76 | 10.5 | 12.5 | 0.7 |

始状態 $|g\rangle = c_0|f^0\rangle + c_1|f^1 \underline{L}^1\rangle + c_2|f^2 \underline{L}^2\rangle$



Fermiの黄金律とローレンツ関数で畳み込み

$$I(\omega; \theta_3) = \sum_{j=0}^2 |\langle f_j | a_c | g \rangle|^2 \frac{\Gamma_j / \pi}{(\omega - E_f(\theta_3) + E_g(\theta_3))^2 + \Gamma_j^2}$$

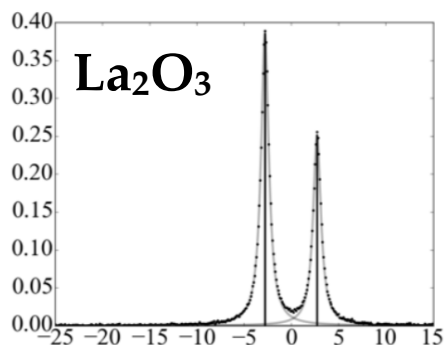
人工光電子スペクトル

ノイズ小

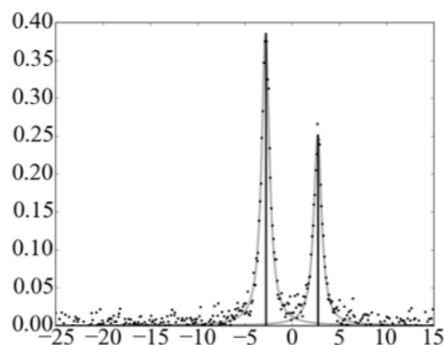
ノイズ大

ノイズ小

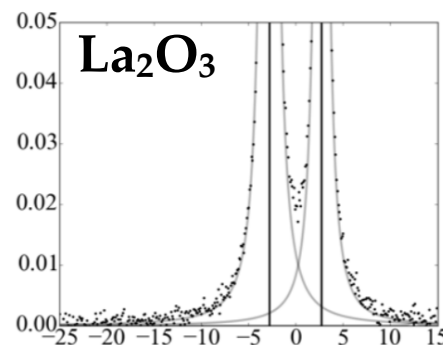
ノイズ大



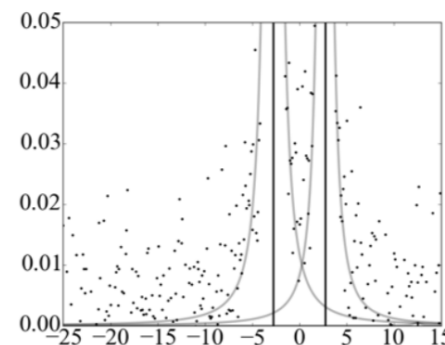
(1)



(2)

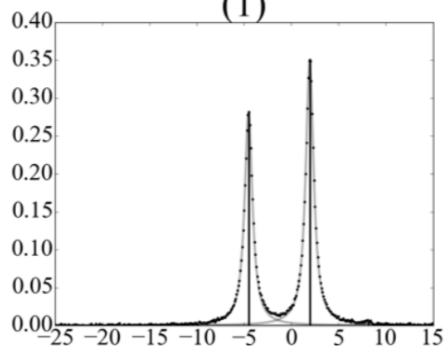


(3)

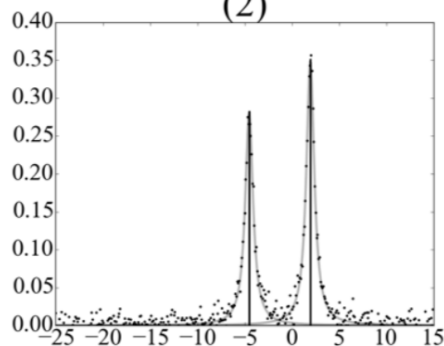


(4)

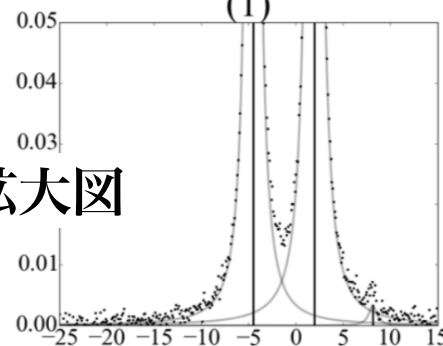
拡大図



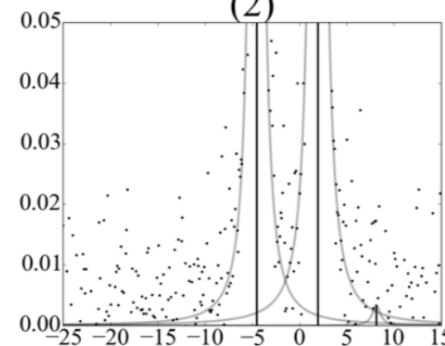
(5)



(6)



(7)



(8)

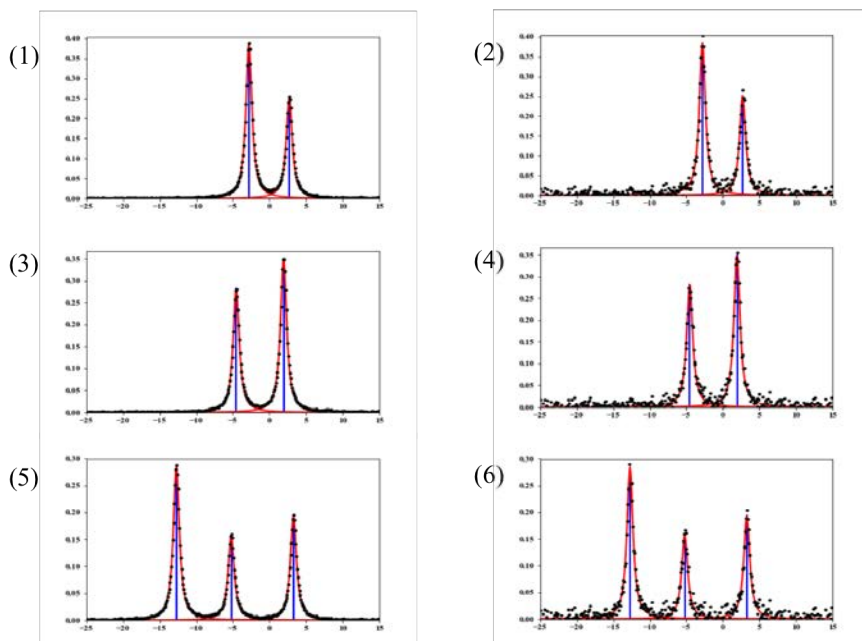
CeO₂

(5)

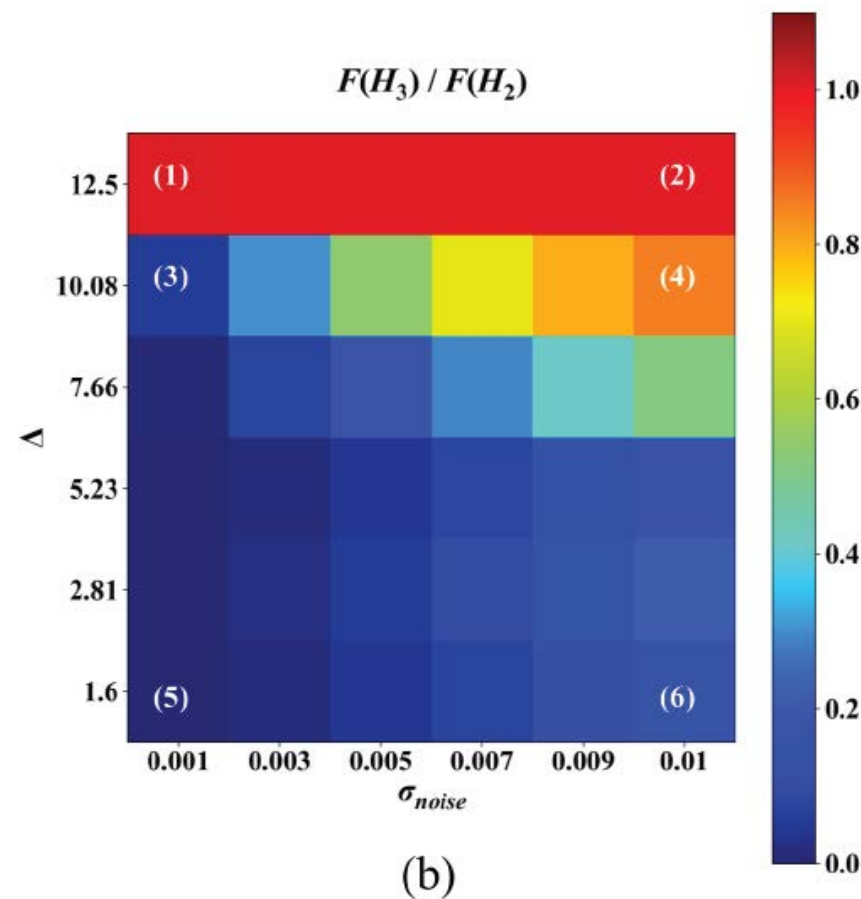
(6)

ハミルトニアン選択

スペクトル形状



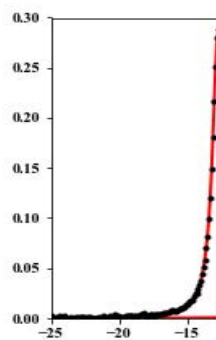
自由エネルギー



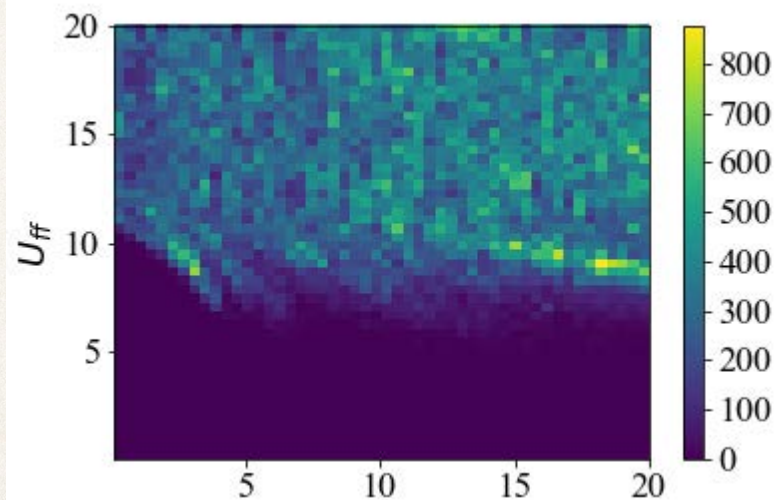
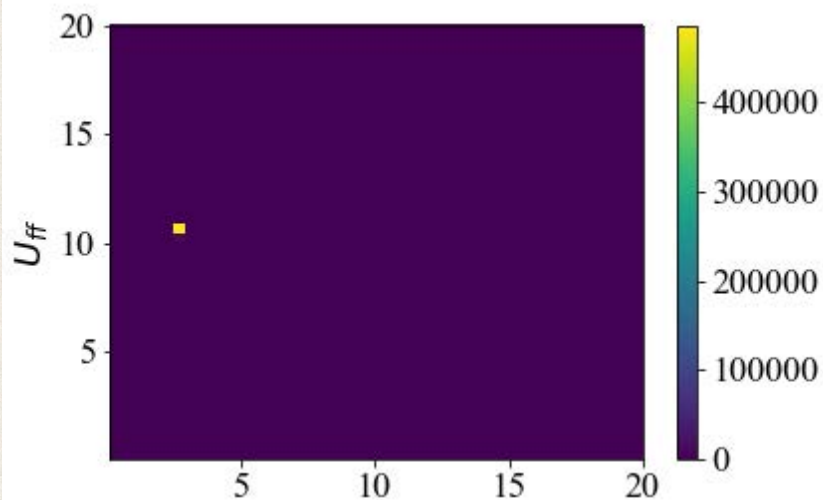
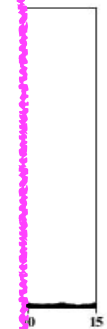
パラメータ推定

CeO₂

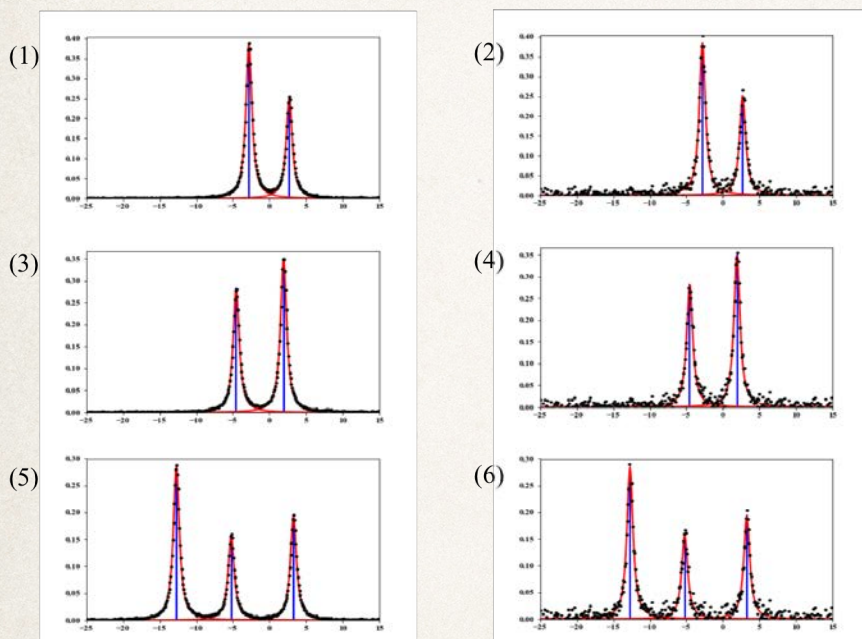
La₂O₃



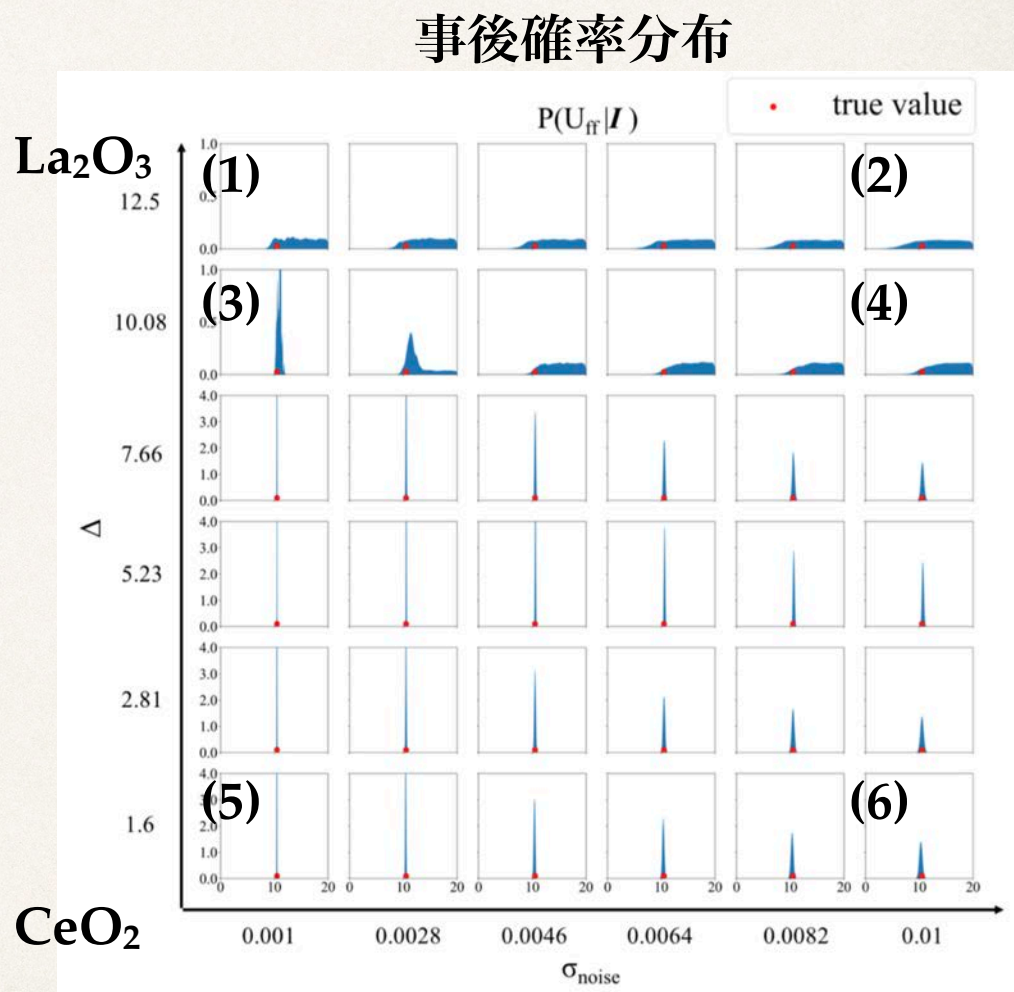
$$H = \sum_k \epsilon_k a_k^\dagger a_k + \sum_\nu \epsilon_f a_{f\nu}^\dagger a_{f\nu} + \epsilon_c a_c^\dagger a_c$$
$$+ \frac{V}{\sqrt{N_f}} \sum_{\nu,k} (a_{f\nu}^\dagger a_k + a_k^\dagger a_{f\nu}) - U_{fc} \sum_\nu a_{f\nu}^\dagger a_{f\nu} (1 - a_c^\dagger a_c)$$
$$+ U_{ff} \sum_{\nu > \nu'} a_{f\nu}^\dagger a_{f\nu} a_{f\nu'}^\dagger a_{f\nu'}$$



パラメータ推定



推定パラメータを用いると
生成モデルスペクトルを再現



電子物理量が誤差付きで評価できた

✓ ベイズ推定の分光学への応用を紹介

ベイズ推定により精度が評価された
物理パラメータが得られた



実測データの
適用を検討

しかも自動チューニング

ベイズ的モデル選択によりハミルトニアン
を選択した

J. Phys. Soc. Jpn. 88(2019)034004

CREST

計測技術と高度情報処理の融合によるインテリジェント計測・
解析手法の開発と応用