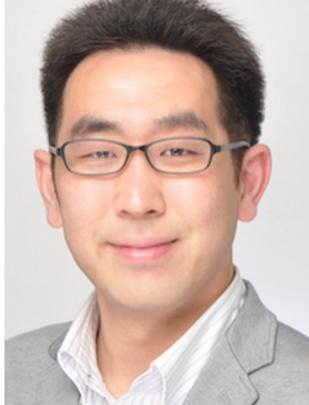


ベイズ推論とSpMに基づく 計測の情報数理基盤の構築

岡田真人

東京大学 大学院新領域創成科学研究科

共同研究者



永田 賢二
NIMS



杉田 精司
東京大学



村岡 恼
東京大学



本武陽一
東京大学



佐々木岳彦
東京大学



徳田悟
九州大学



水牧仁一朗
JASRI



赤井一郎
熊本大学



横山優一
JASRI



魚住孝幸
大阪公立大

内容

- 自己紹介
- 課題1(ベイズ計測)
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC(岡田研 片上助教)
- 本プロジェクトの到達点

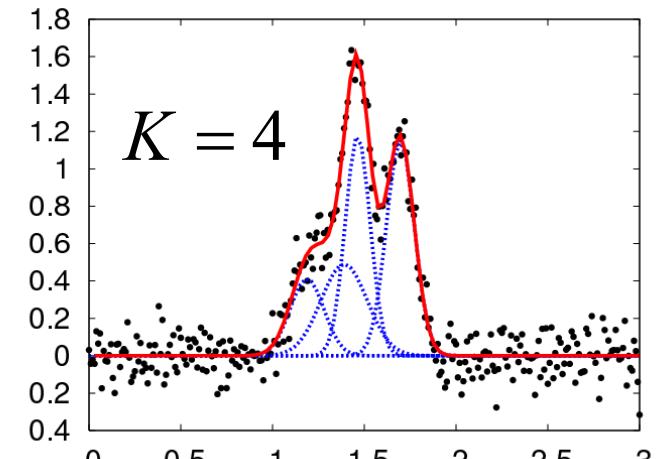
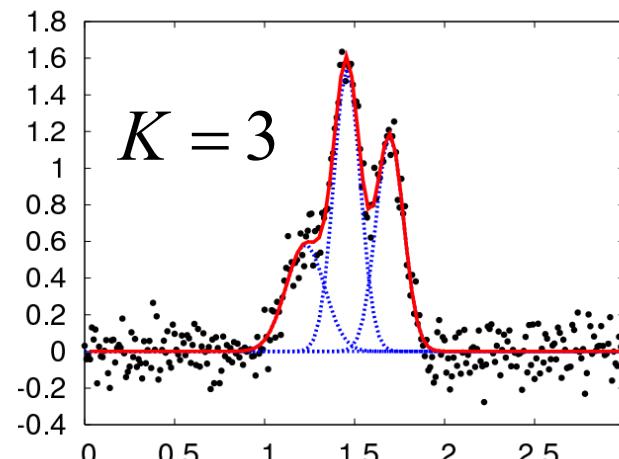
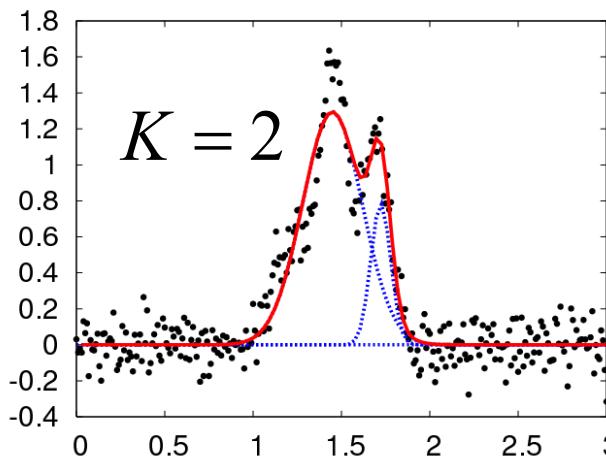
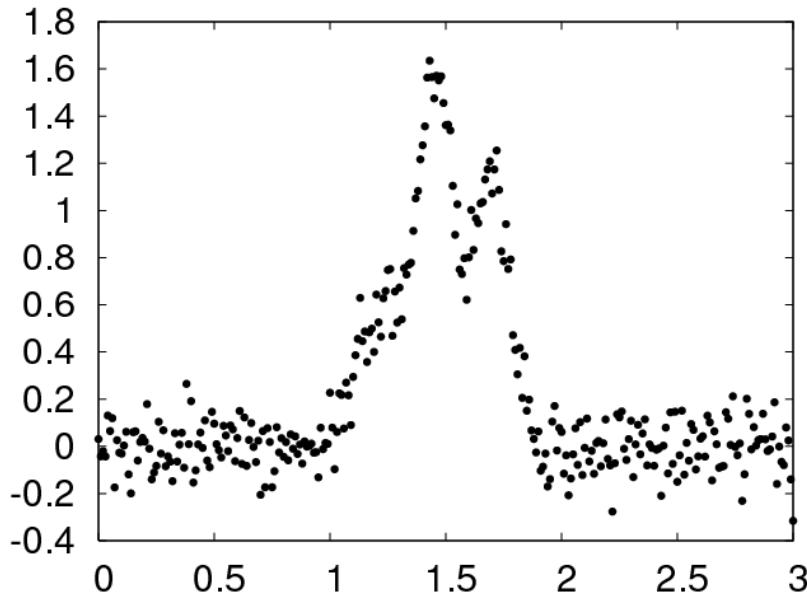
自己紹介(理論物理)

- 大阪市立大学理学部物理学学科 (1981 - 1985)
- 大阪大学大学院理学研究科物理専攻 (1985 – 1987)
 - 希土類元素の光励起スペクトルの理論
- 三菱電機 (1987 - 1989)
 - 化合物半導体(半導体レーザー)の結晶成長
- 大阪大学大学院基礎工学研究科生物工学 (1989 - 1996)
- JST ERATO 川人学習動態脳プロジェクト (1996 - 2001)
- 理化学研究所 脳科学総合研究センター (2001 - 04/06)
- 東京大学・大学院新領域創成科学研究所 複雑理工学専攻 (2004/07 –)

内容

- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

ベイズ的スペクトル分解



Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

分光学における諸問題

- ・スペクトルや画像データからフィッティングを行なっている
- ・そのフィッティングの際に、パラメータを手打ちで決めている。最急降下法などを使っているが、うまくいかない。
- ・フィッティング用のモデルが複数あって、事前にどれを使うかを決めておかないといけない。
- ・S/Nが悪いデータや欠損データをなんとかした。
- ・複数計測の統合を行いたい。

- ・そのような方は、一度ベイズ計測をお試しください。

スペクトル分解の定式化

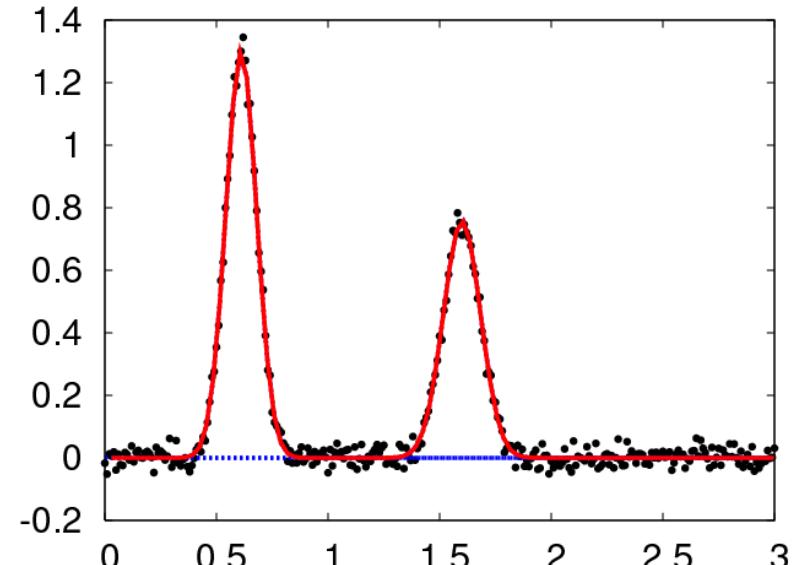
ガウス関数(基底関数)の足し合わせにより、スペクトルデータを近似

観測データ: $D = \{x_i, y_i\}_{i=1}^n$

x_i : 入力 y_i : 出力

$$f(x; \theta) = \sum_{k=1}^K a_k \exp\left(-\frac{b_k(x - \mu_k)^2}{2}\right)$$

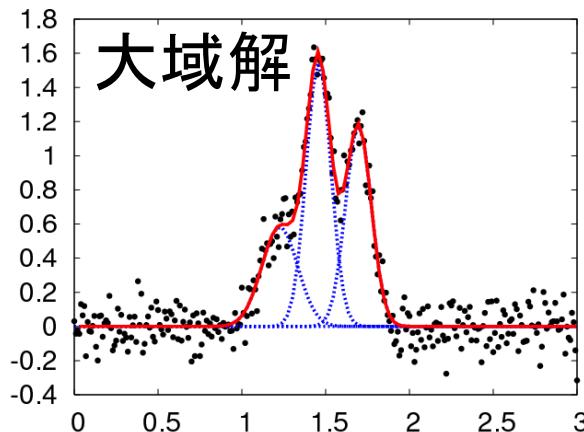
$$\theta = \{a_k, b_k, \mu_k\} \quad k = 1, \dots, K$$



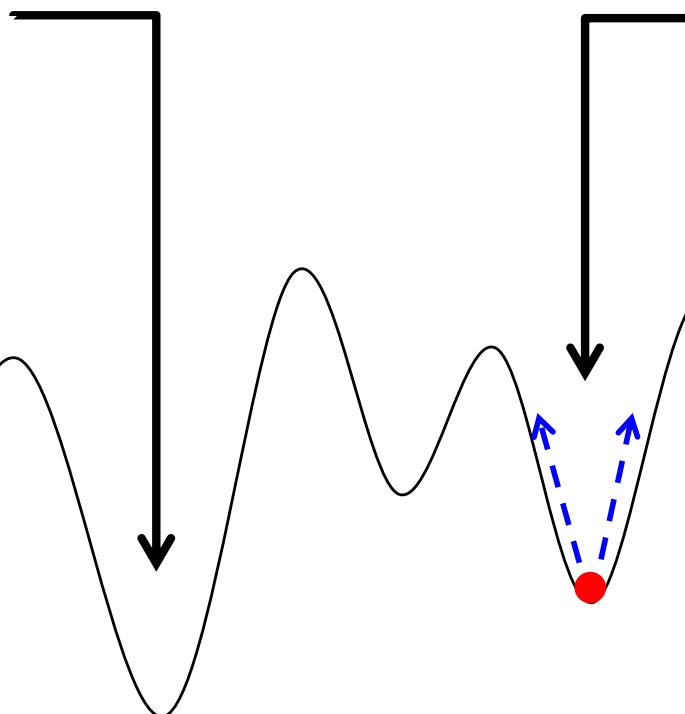
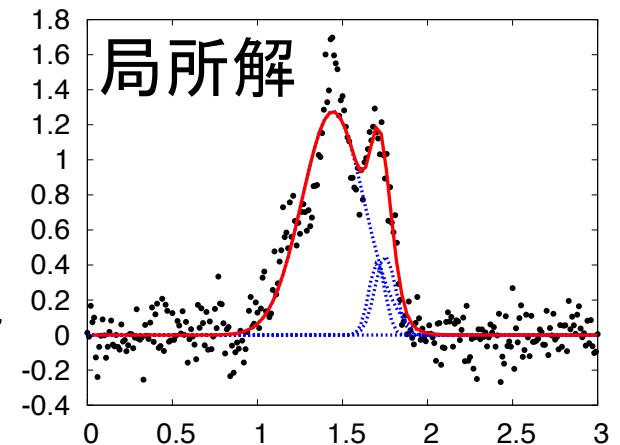
二乗誤差を最小にするようにパラメータをフィット(最小二乗法)

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

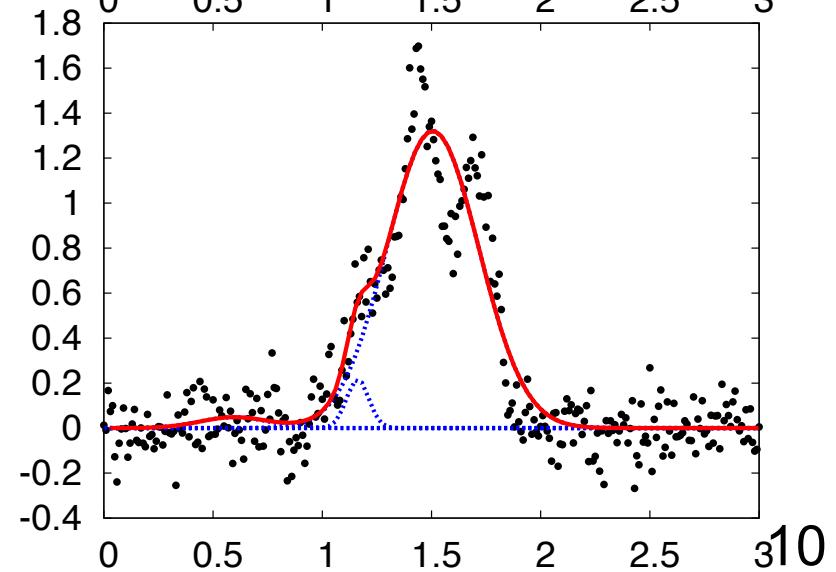
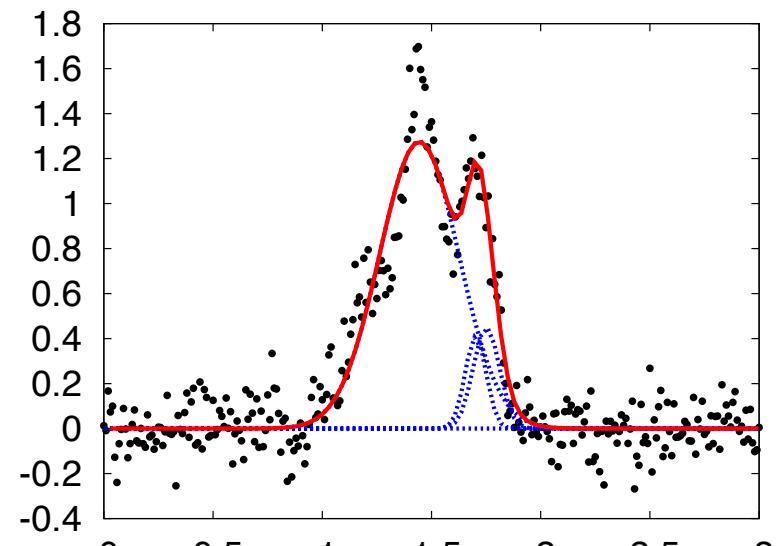
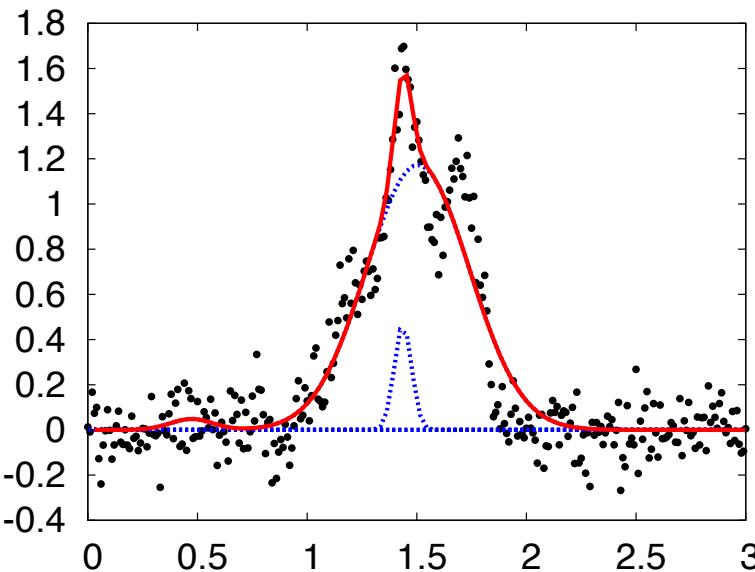
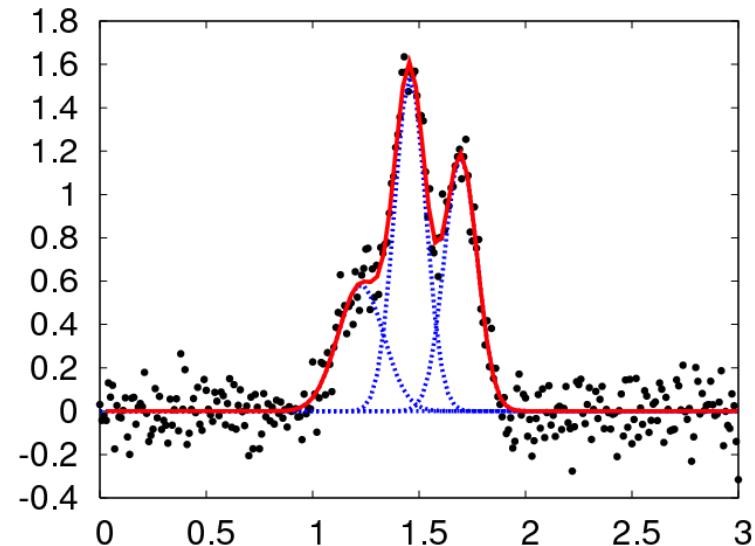
誤差関数は局所解を持つ



＜通常の最適化法＞
e.g., 最急降下法



ローカルミニマム



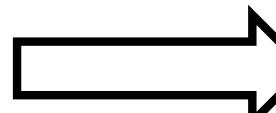
内容

- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

ベイズ計測

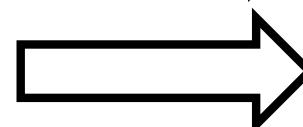
順アプローチ

計測データ

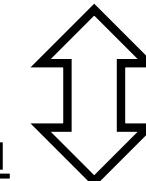


実験の結果

モデル



理論の結果

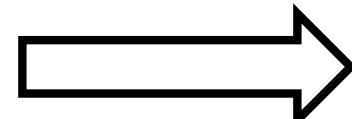


比較

$$p(Y | \theta, K) \text{ 解析計算, 数値計算}$$

逆アプローチ

対象とする
物理系



系の物理
モデル



観測過程
計測機器の特性

計測データ

$$p(\theta, K)$$

$$p(\theta, K | Y)$$

全てをモデル化し
ベイズの定理で因果をさかのぼる

確率的定式化

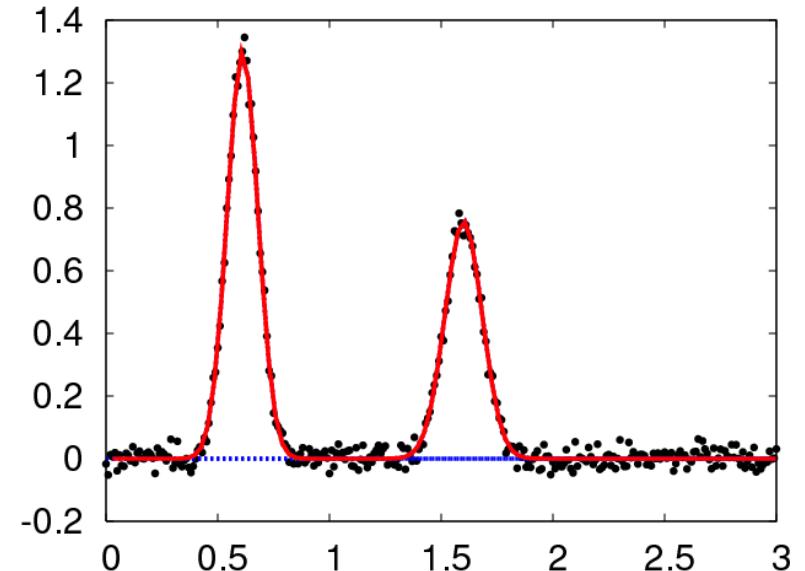
出力は、入力からの応答とノイズの足し合わせにより生成

⇒出力は、確率変数である。

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

ノイズが正規分布であるとすると、

$$p(y_i | \theta) \propto \exp\left(-\frac{1}{2}\left(y_i - f(x_i; \theta)\right)^2\right)$$



それぞれの出力 y_i が、独立であるとすると、

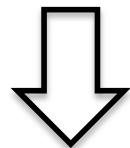
$$p(Y | \theta) = \prod_{i=1}^n p(y_i | \theta) \propto \exp(-nE(\theta)) \quad Y = \{y_1, \dots, y_n\}$$

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

ボルツマン分布

ベイズ推論：因果律を組み込んでデータ解析

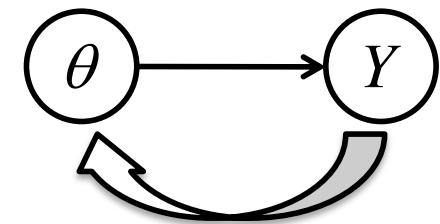
$$p(Y, \theta) = \frac{p(Y | \theta)p(\theta)}{p(Y)}$$



<ベイズの定理>

$$p(\theta | Y) = \frac{p(Y | \theta)p(\theta)}{p(Y)} \propto \exp(-nE(\theta))p(\theta)$$

生成(因果律)



$p(\theta | Y)$: 事後確率。データが与えられたもとでの、パラメータの確率。

$p(\theta)$: 事前確率。あらかじめ設定しておく必要がある。
これまで蓄積してきた科学的知見

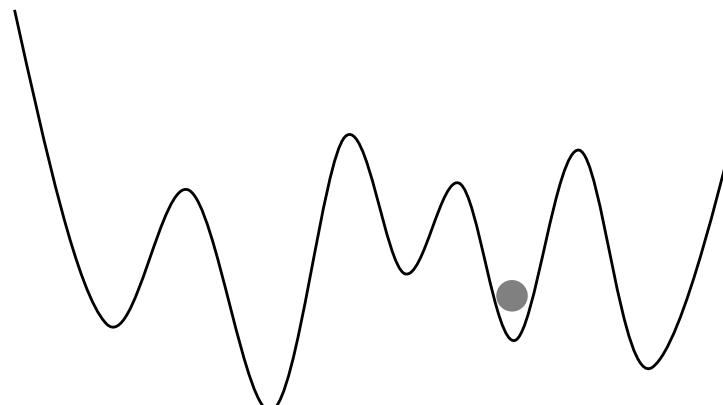
内容

- 自己紹介
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

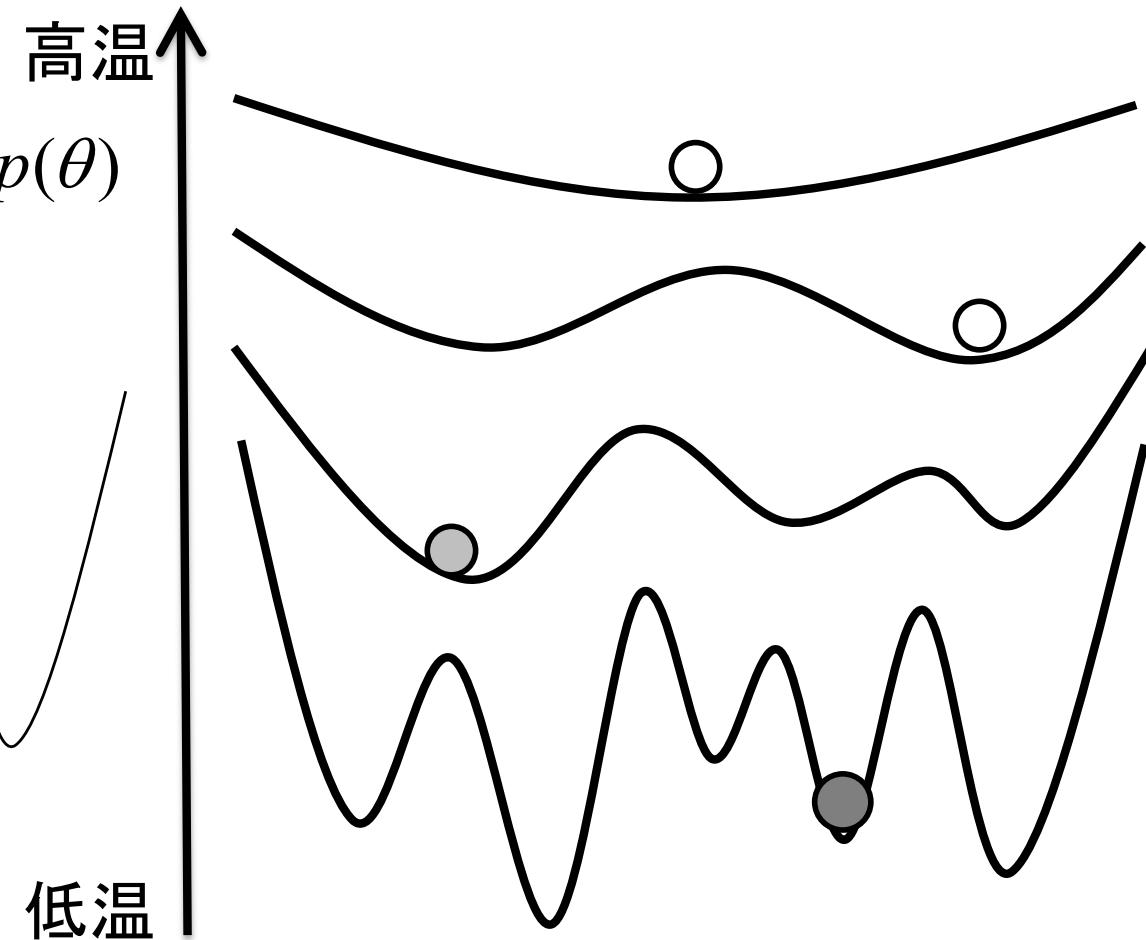
レプリカ交換モンテカルロ法 ランダムスピニ系の知見から

メトロポリス法

$$p_\beta(\theta) \propto \exp\left(-\frac{n}{\sigma^2} \beta E(\theta)\right) p(\theta)$$

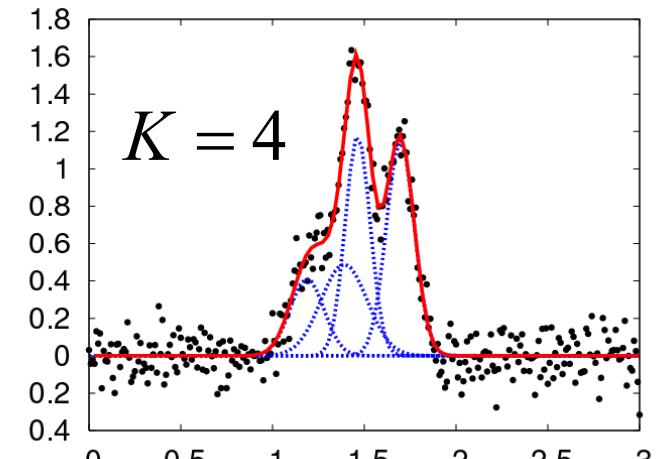
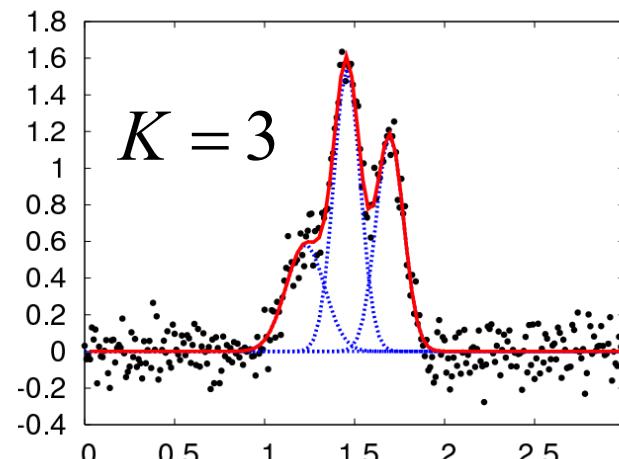
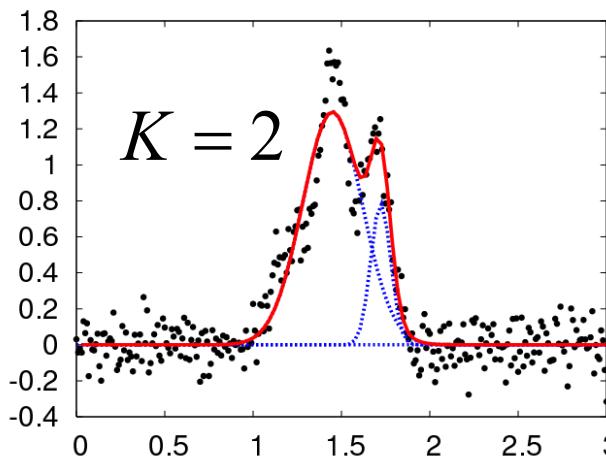
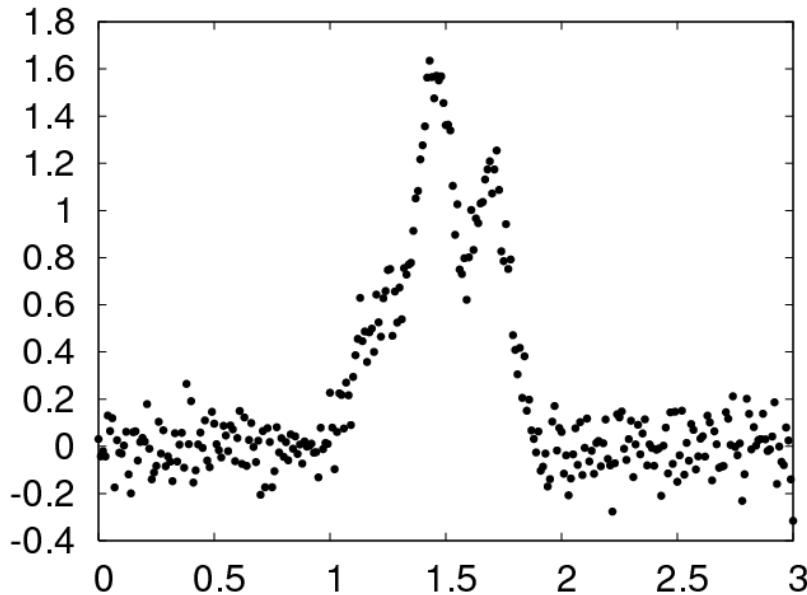


レプリカ交換モンテカルロ法



K. Hukushima, K. Nemoto, *J. Phys. Soc. Jpn.* **65** (1996).

ベイズ的スペクトル分解

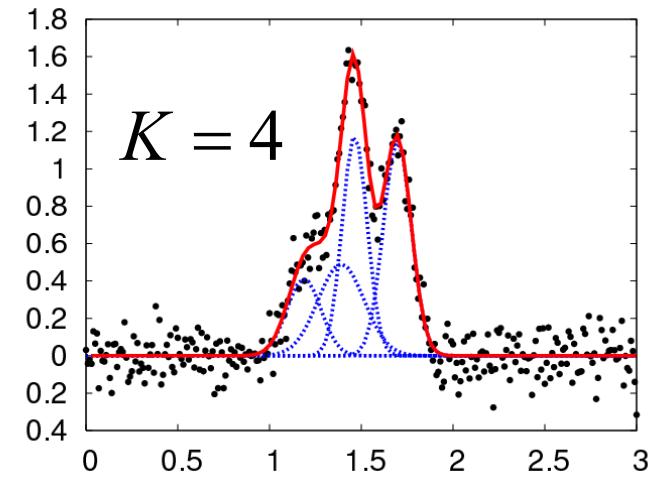
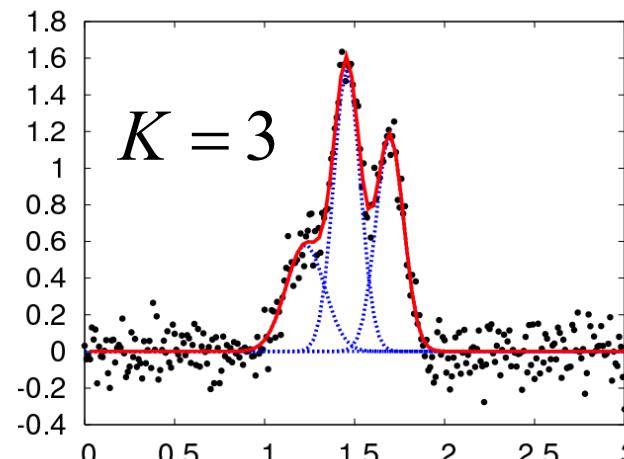
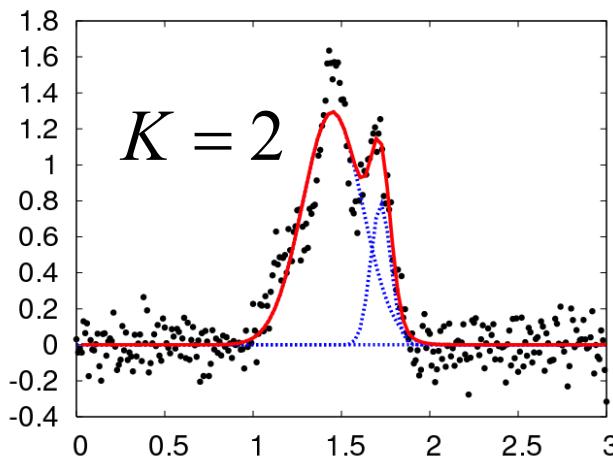
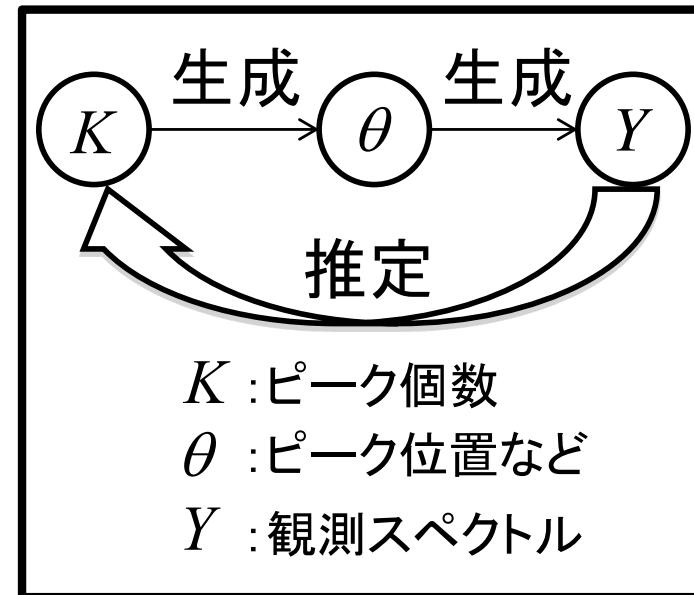
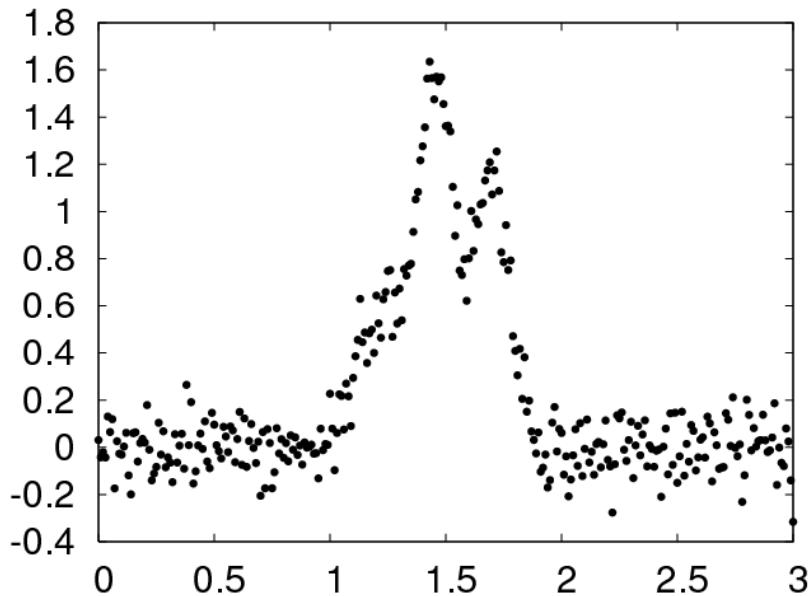


Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

内容

- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

より深い構造をさぐる: モデル選択



Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

モデル選択

- 欲しいのは $p(K|Y)$
- θ がないぞ
- $p(K, \theta, Y)$ の存在を仮定

$$p(K, \theta, Y) = p(Y|\theta, K)p(K)$$

$$p(Y|\theta, K) = \prod_{i=1}^n p(y_i|\theta) \propto \exp(-nE(\theta))$$

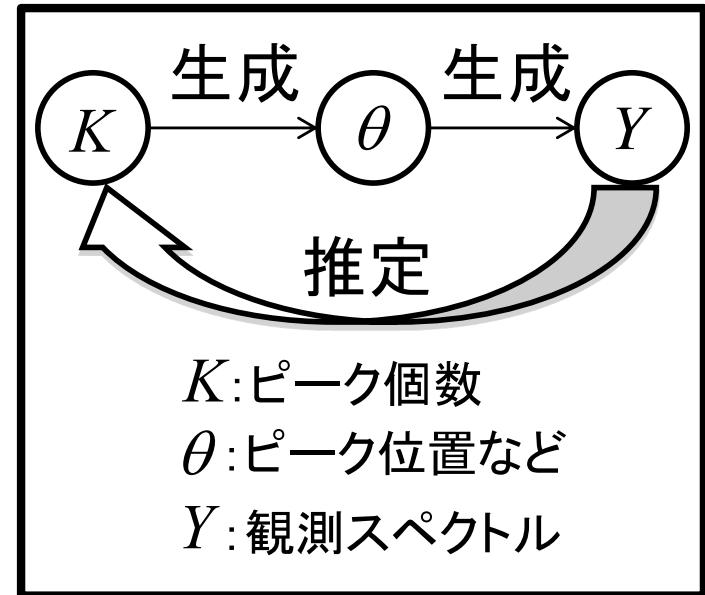
- 無駄な自由度の系統的消去: 周辺化, 分配関数

$$p(K, Y) = \int p(K, \theta, Y) d\theta$$

$$p(K|Y) = \frac{p(Y|K)p(K)}{p(Y)} \propto p(K) \int \exp(-nE(\theta)) p(\theta) d\theta$$

$$F(K) = -\log \int \exp(-nE(\theta)) p(\theta) d\theta$$

自由エネルギーを最小にする個数 K を求める。



自由エネルギーの数値的計算法 レプリカ交換法の性質を巧妙に使う

$$F = -\log \int \exp\left(-\frac{n}{\sigma^2} E(\theta)\right) p(\theta) d\theta$$

自由エネルギー:

以下のように、補助変数 β を導入する。 β : 逆温度

$$F_\beta = -\log \int \exp\left(-\frac{n}{\sigma^2} \beta E(\theta)\right) p(\theta) d\theta \quad (F_{\beta=0} = 0)$$

$$F = F_{\beta=1} = \int_0^1 d\beta \frac{\partial F_\beta}{\partial \beta}$$

たくさんの温度でのシミュレーションが必要

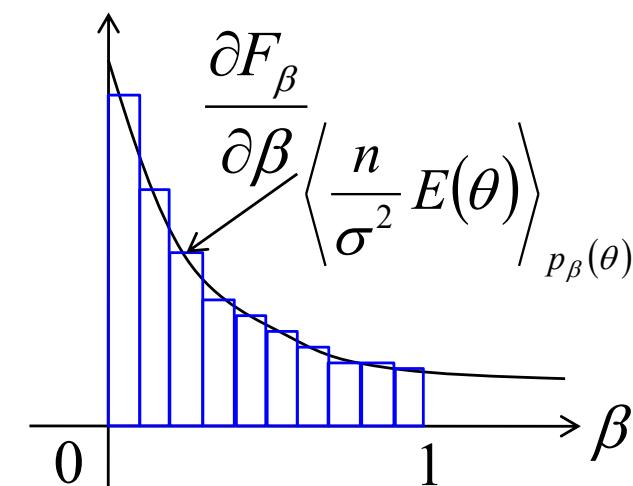
\rightarrow 各温度でのエネルギー平均(すでにやってる)

$$\frac{\partial F_\beta}{\partial \beta} \dots$$

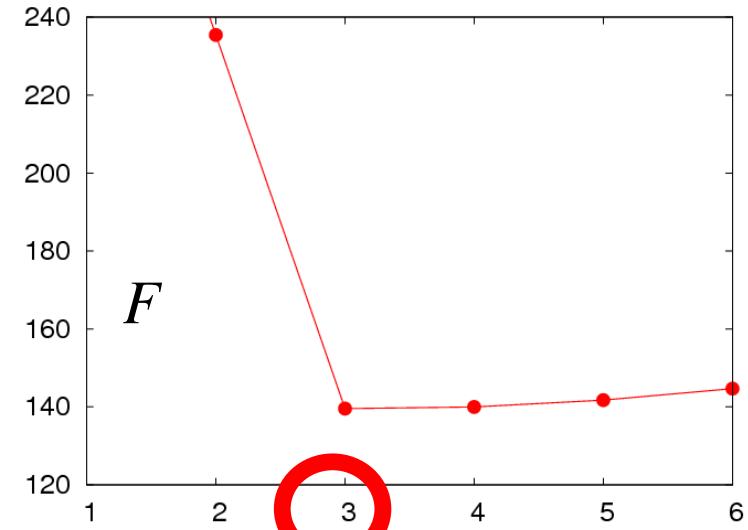
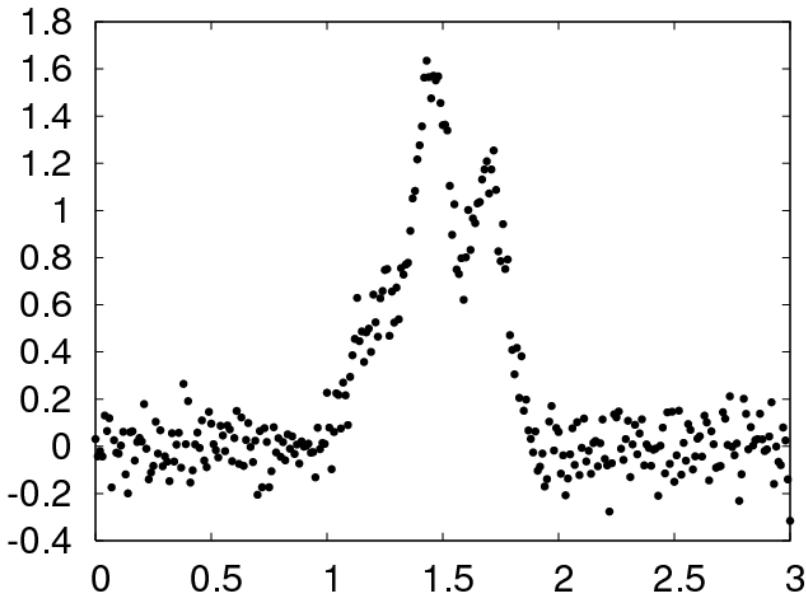
確率分布 $p(\theta; \beta)$ に従う

二乗誤差 $\frac{n}{\sigma^2} E(\theta)$ の期待値

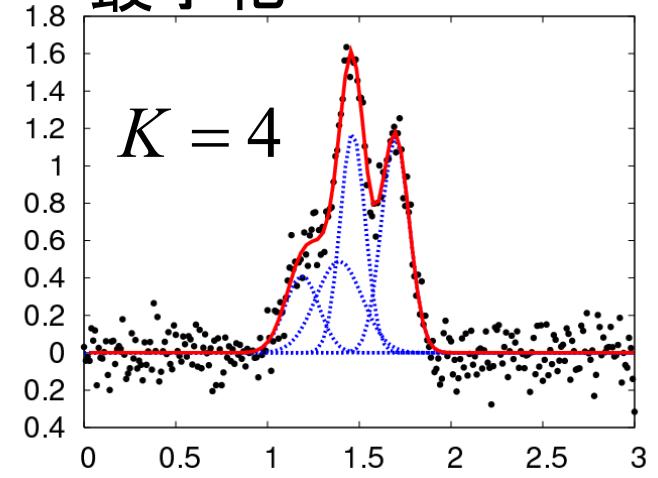
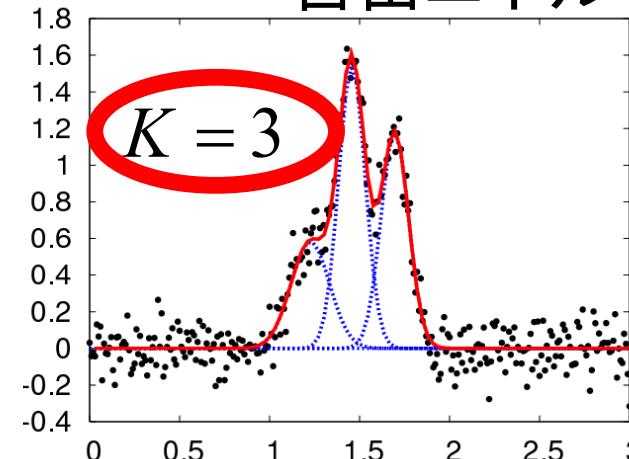
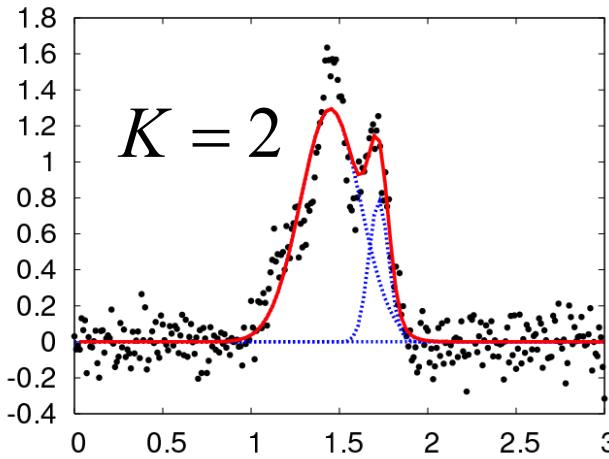
$$p_\beta(\theta) \propto \exp\left(-\frac{n}{\sigma^2} \beta E(\theta)\right) p(\theta)$$



スペクトル分解



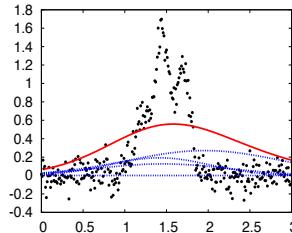
最適な K をデータだけから決める
自由エネルギー最小化



Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

ベイズ的スペクトル分解(Nagata *et al.* 2012)

パラ相

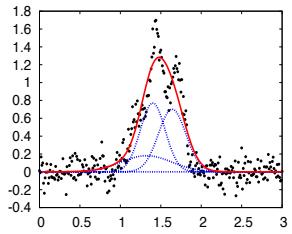


$$f(x; \theta) = \sum_{k=1}^K a_k \exp\left(-\frac{b_k(x - \mu_k)^2}{2}\right)$$

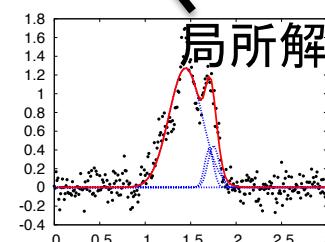
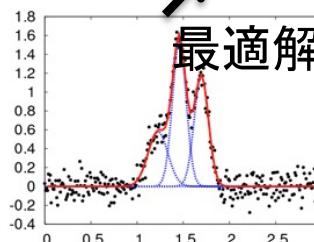
$$E(\theta) = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i; \theta))^2$$

交換モンテカルロ法

フェロ相



ノイズレベル β



因果律

$p(Y | \theta)$

θ : ピーク位置など

Y : 観測スペクトル

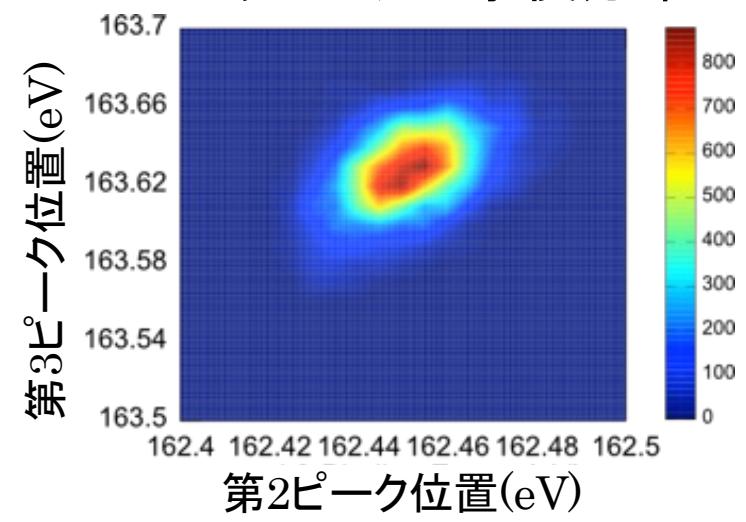
推論

$$p(\theta | Y) = \frac{p(Y | \theta)p(\theta)}{p(Y)}$$

ボルツマン分布

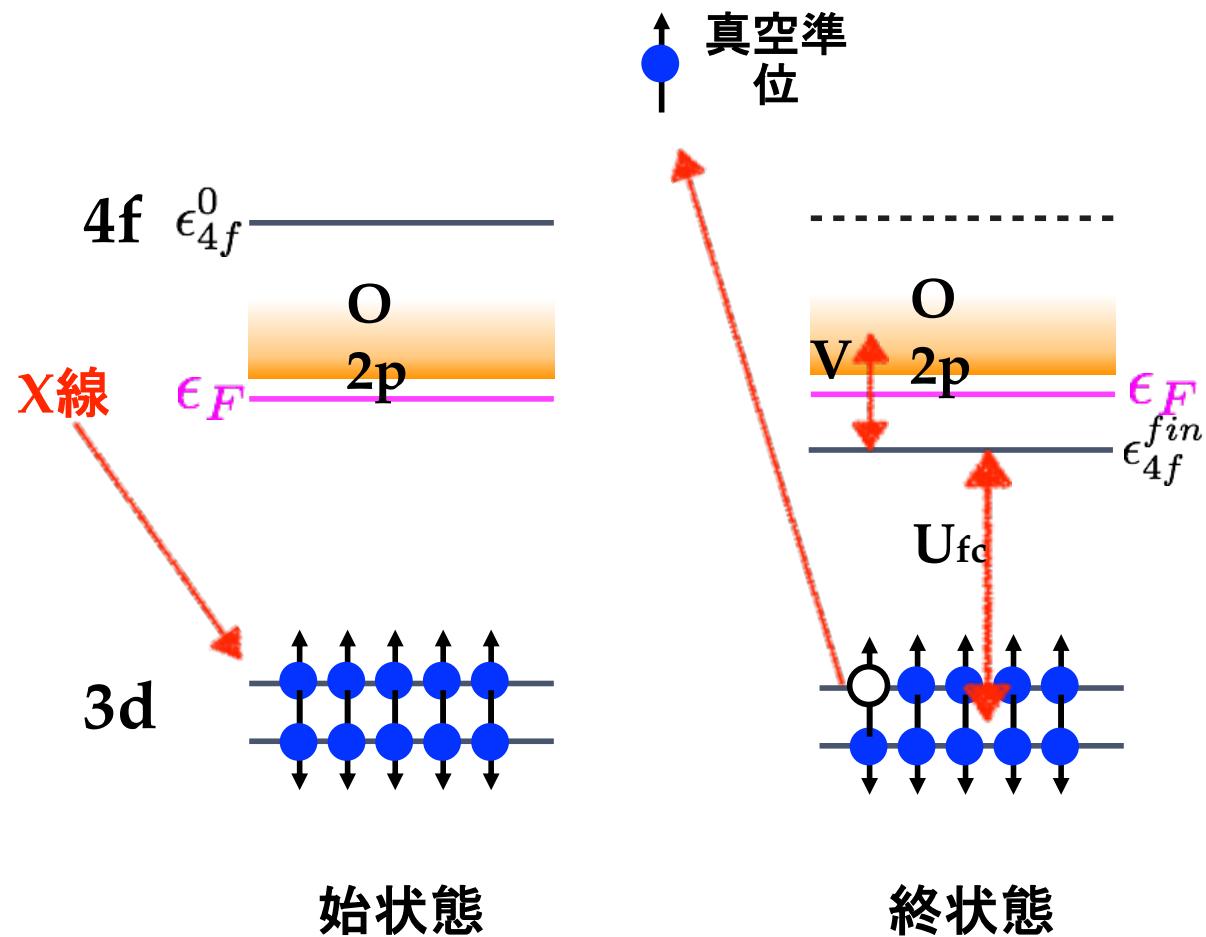
$$\propto \exp(-nE(\theta))p(\theta)$$

パラメータの事後分布

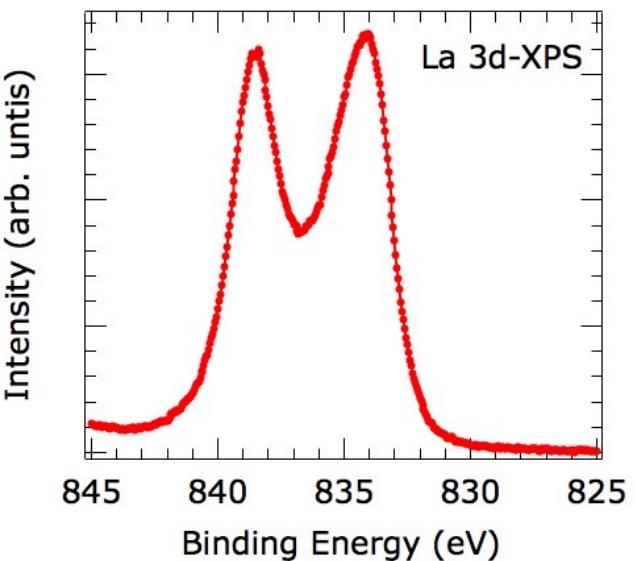


内殻X線光電子分光 (Mototake *et al.*, 2019)

X線光電子分光の遷移過程

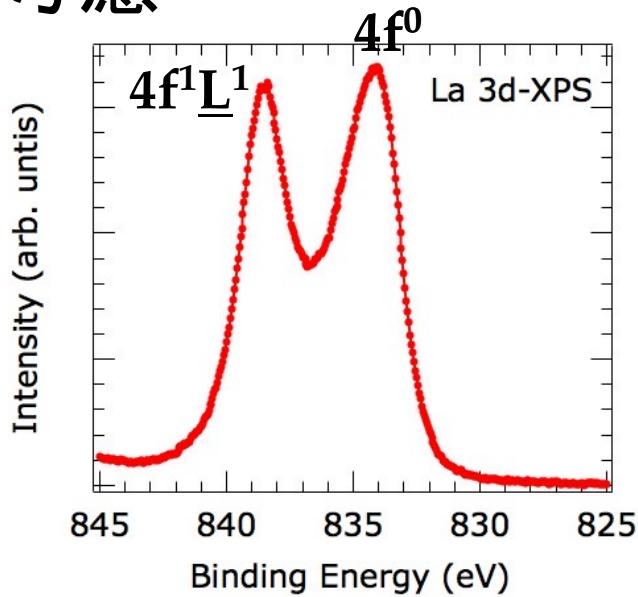


4f希土類元素の3d内殻の場合



スペクトル解析

X線光電子分光の遷移過程を表すHamiltonianを考慮



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

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

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

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

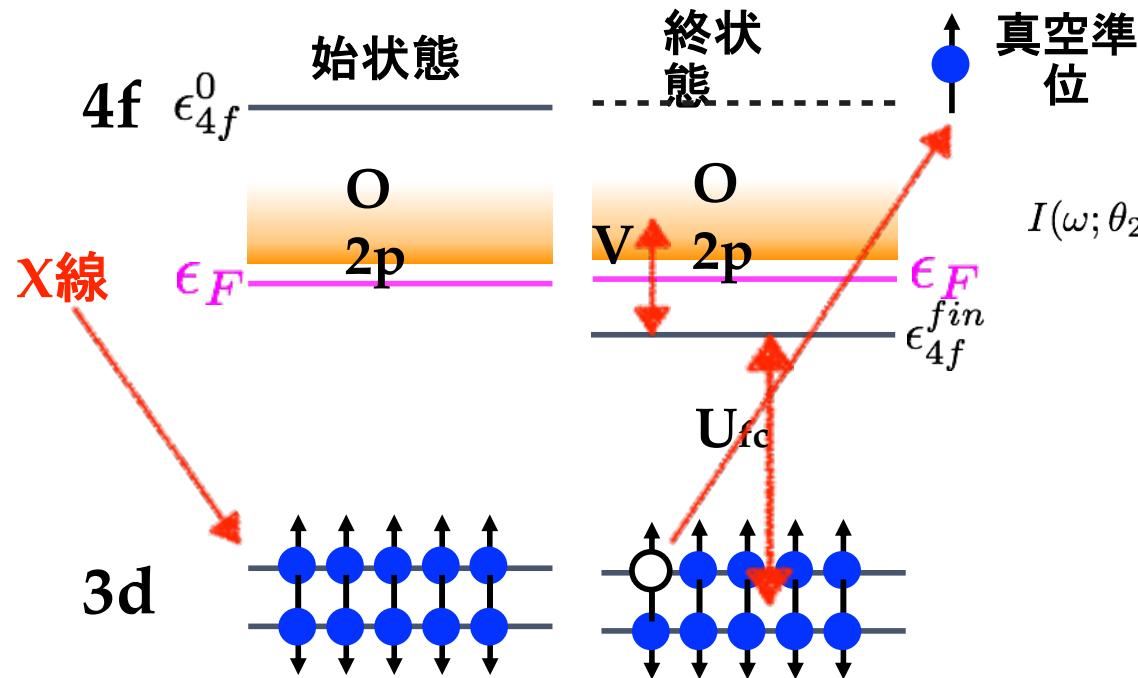


電子に関する物理量抽出→電子状態を理解した

2準位クラスター モデル

La₂O₃を模擬

$$H_2^I = \sum_k \epsilon_k a_k^+ a_k + \sum_\nu \epsilon_f a_{f\nu}^+ a_{f\nu} + \epsilon_c a_c^+ a_c \\ + \frac{V}{\sqrt{N_f}} \sum_{\nu, k} (a_{f\nu}^+ a_k + a_k^+ a_{f\nu}) - U_{fc} \sum_\nu a_{f\nu}^+ a_{f\nu} (1 - a_c^+ a_c)$$



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

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

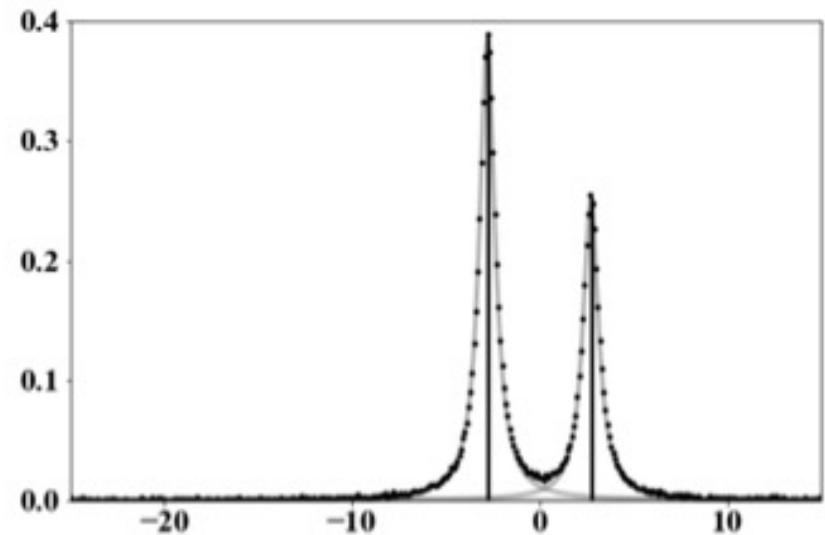
2準位クラスター モデル

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



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

△ 終状態エネルギー: Ef1 Ef2 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準位クラスター模型

CeO₂を模擬

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

$$H_3^I = \begin{pmatrix} 0 & 0 \\ \sqrt{N_f} & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{array}{c} \text{図} \\ \text{V} \\ T_{ff} \end{array}$$

終状態 $|f\rangle = c_f$

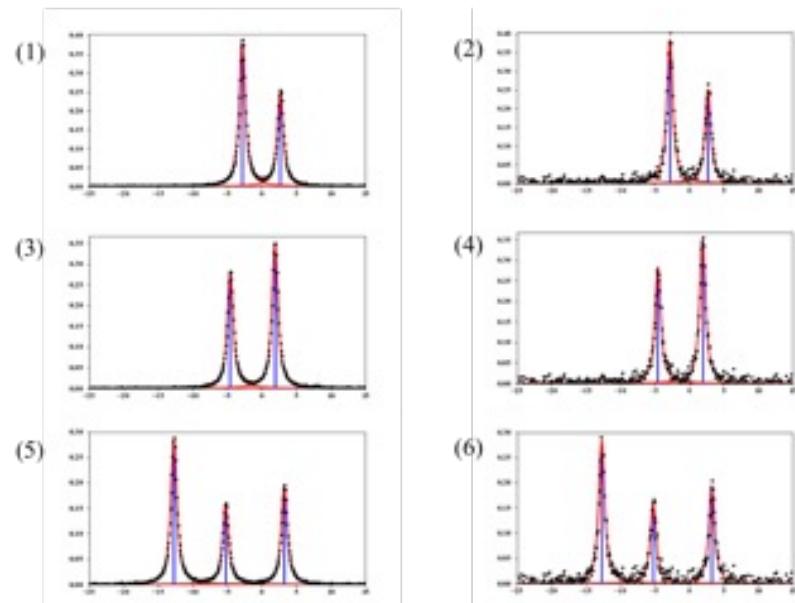
$$H_3^F = \begin{pmatrix} 0 & 0 \\ \sqrt{N_f} & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{array}{c} \text{図} \\ \overline{V} \\ \overline{T} \end{array}$$

Fermiの黄金律とローレンツ関数で置み込み

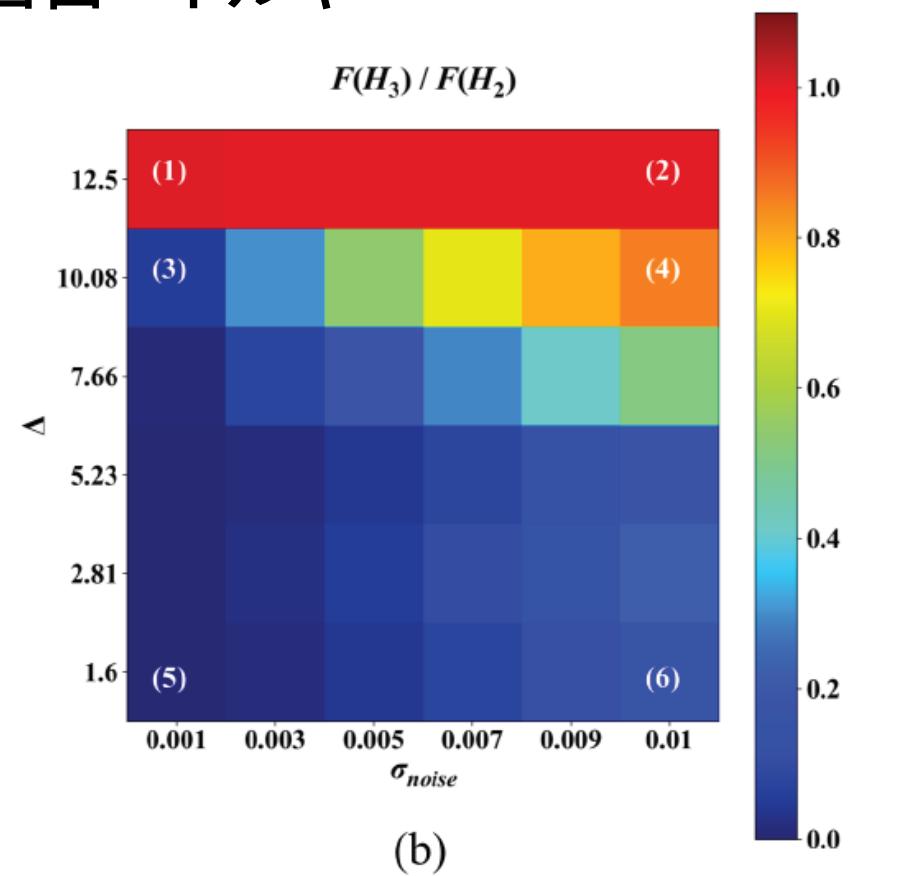
$$I(\omega; \theta_3) = \sum_{j=0}^2 |\langle f_j | a_c | g \rangle|^2 \frac{\Gamma_j / \pi}{(\omega - E_f(\theta_3) + E_g(\theta_3))^2 + \Gamma_j^2}$$

ハミルトニアン選択

スペクトル形状



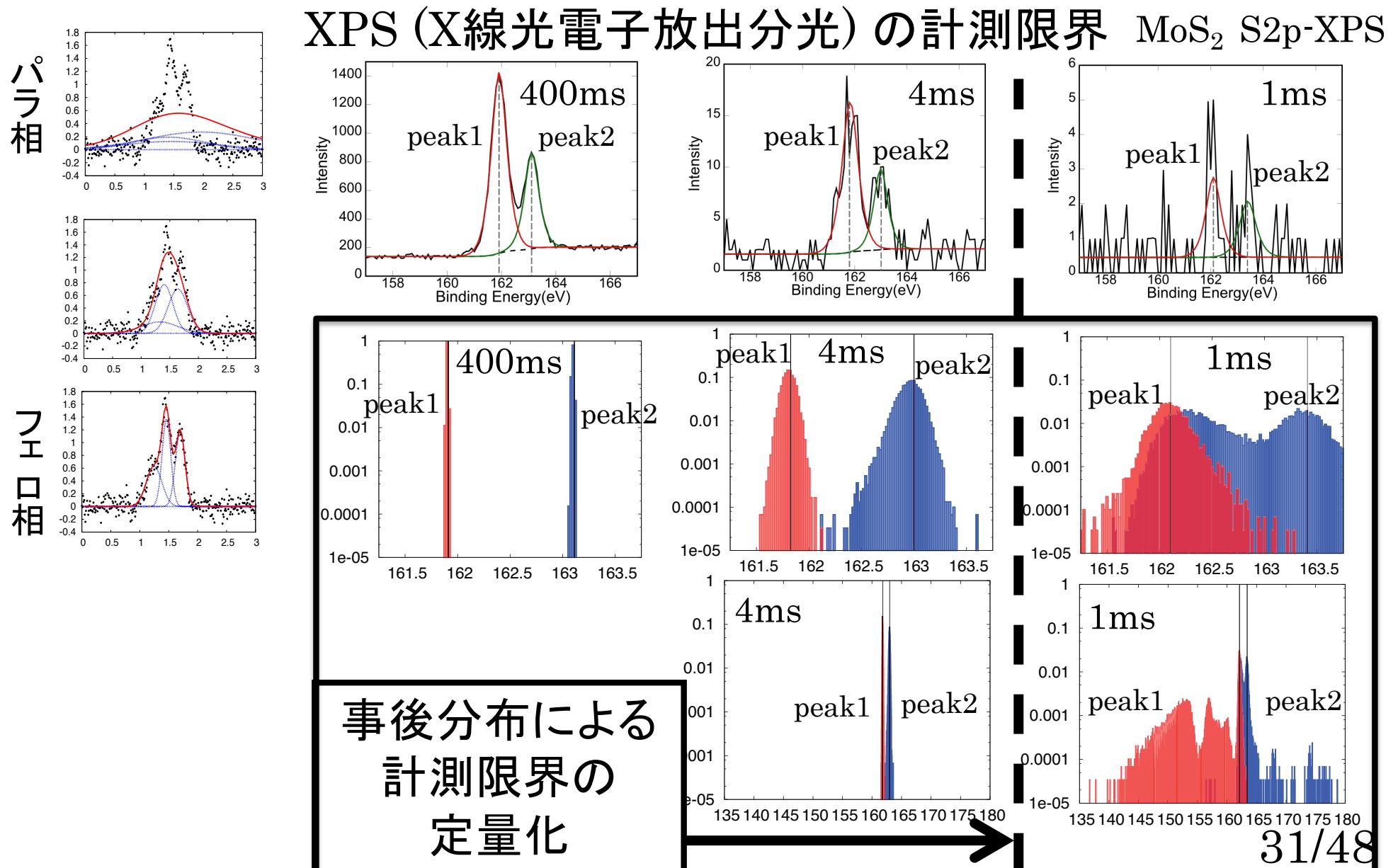
自由エネルギー



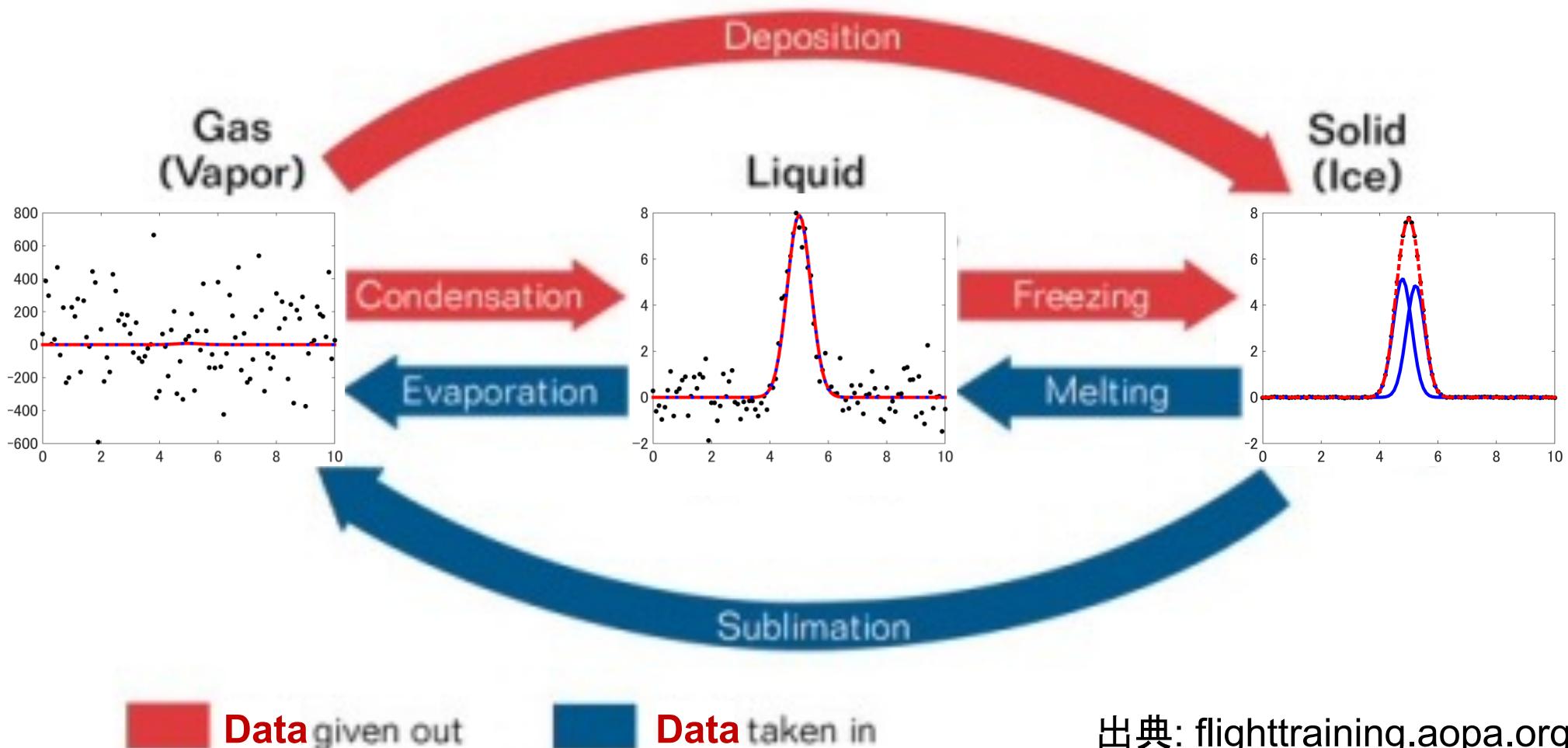
内容

- 自己紹介
- 課題1(ベイズ計測)
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

ベイズ的スペクトル分解(Nagata *et al.* 2019)



計測限界を相転移として捉える



出典: flighttraining.aopa.org

ベイズ比熱

定義

$$C_n(\beta) = \frac{\partial \langle nL_n(w) \rangle_\beta}{\partial \beta^{-1}}$$

$$C_n(\beta) = \lim_{\beta_1 \rightarrow \beta_2} \frac{\langle nL_n(w) \rangle_{\beta_1} - \langle nL_n(w) \rangle_{\beta_2}}{\beta_1^{-1} - \beta_2^{-1}} = \lambda + O_p\left(\frac{1}{\sqrt{\log n}}\right)$$

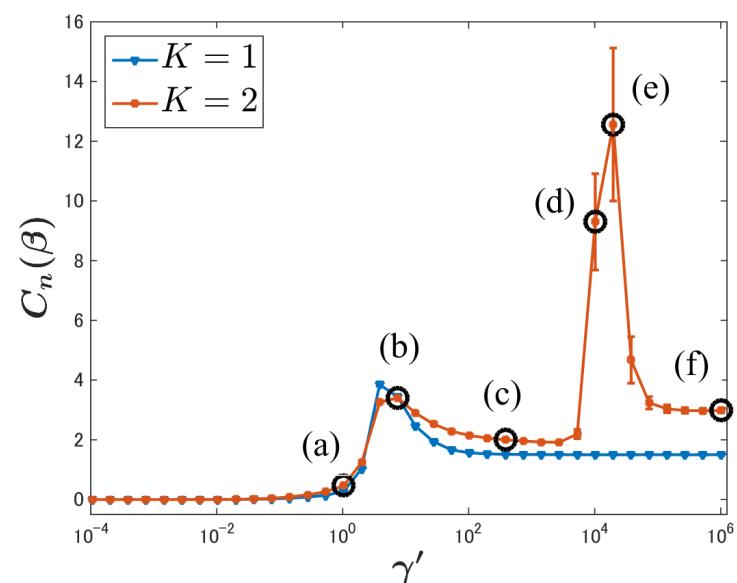
特に、尤度が正規分布のとき

$$C_n(\beta) = \gamma'^2 \left(\left\langle E_n(w)^2 \right\rangle_\beta - \left\langle E_n(w) \right\rangle_\beta^2 \right),$$

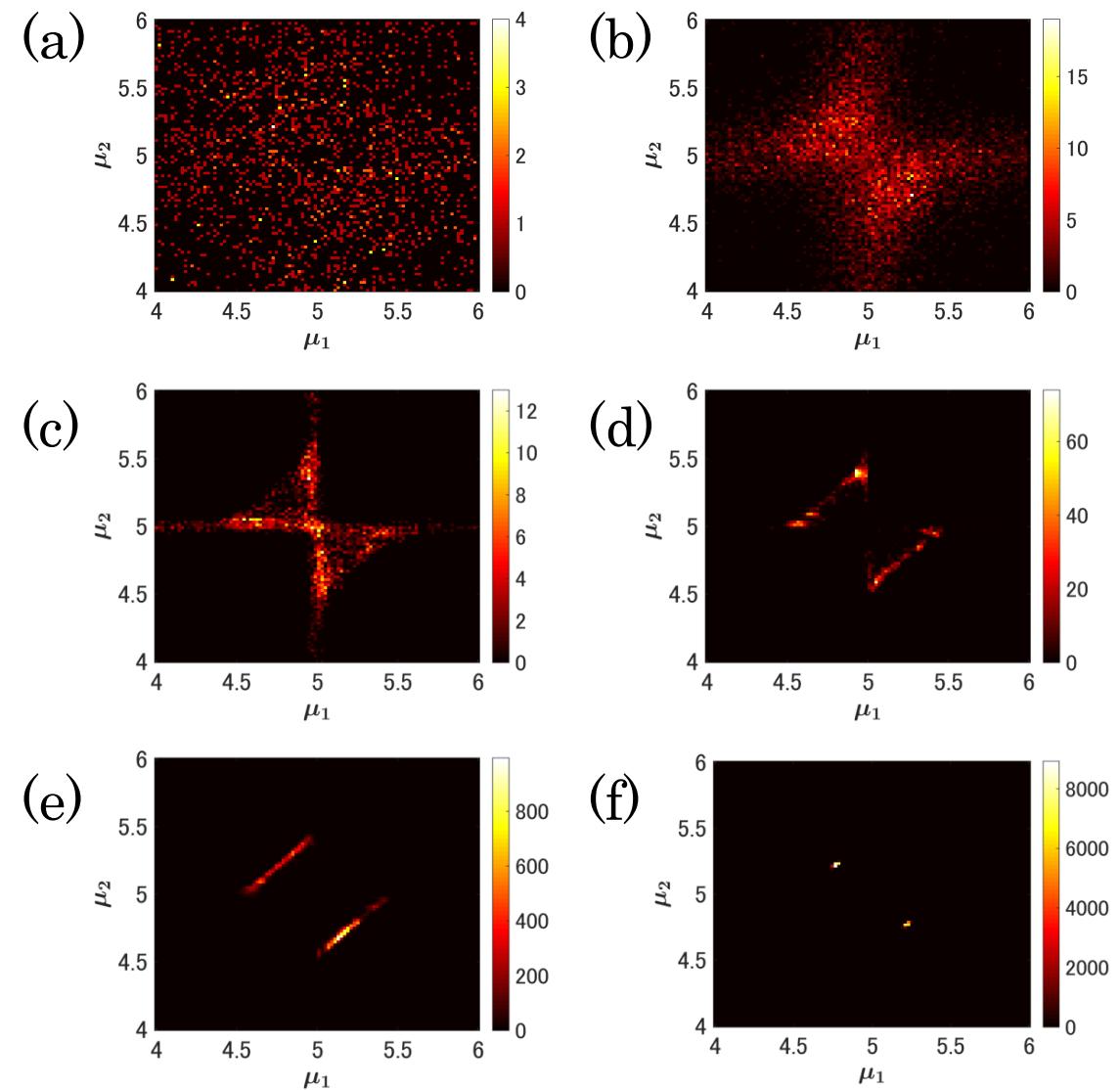
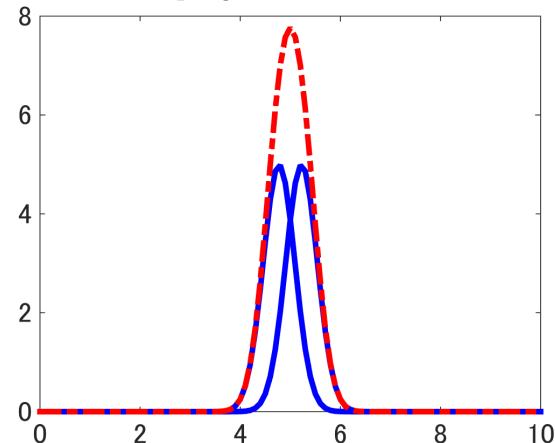
$$\gamma' = n\beta\gamma, \quad E_n(w) = \frac{1}{2n} \sum_{i=1}^n (y_i - g(x_i; w))^2$$

n : サンプル数
 γ : サンプル精度

ベイズ比熱と事後確率密度 (Tokuda, Nagata and Okada, 2022)



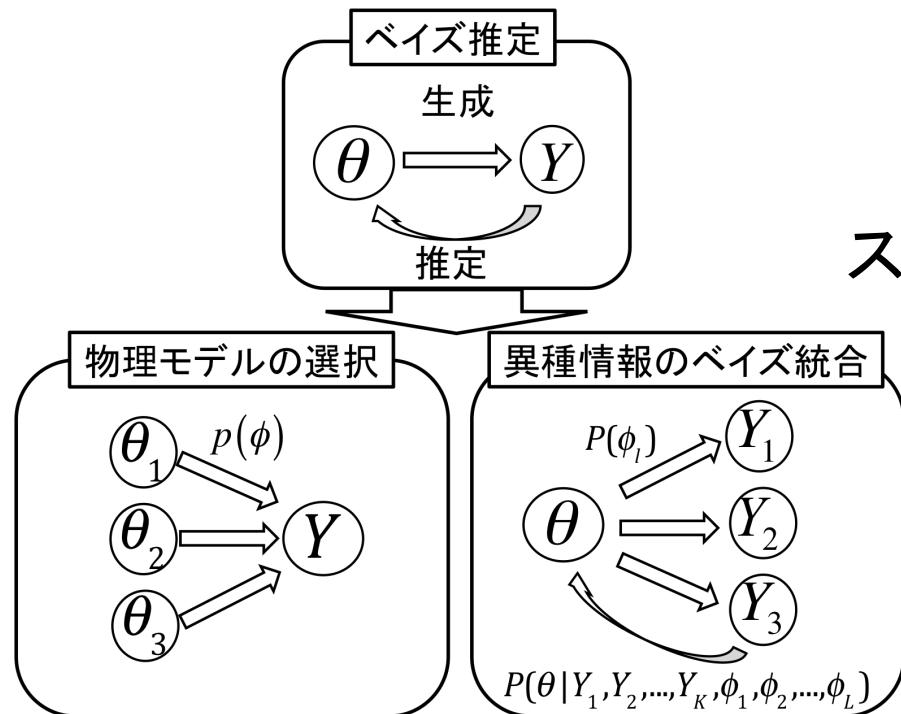
学習データ



内容

- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

【課題1】ベイズ計測



θ : 物理モデル ϕ : 計測手法のモデル Y : 観測データ

スペクトル

フィッティング関数

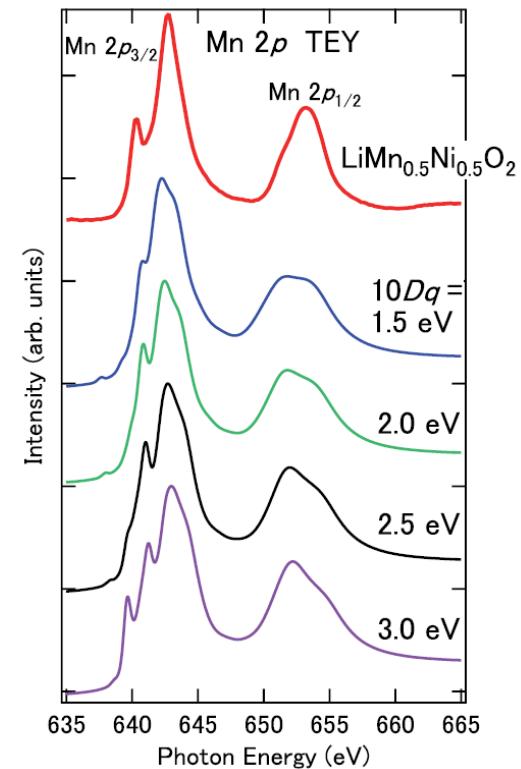
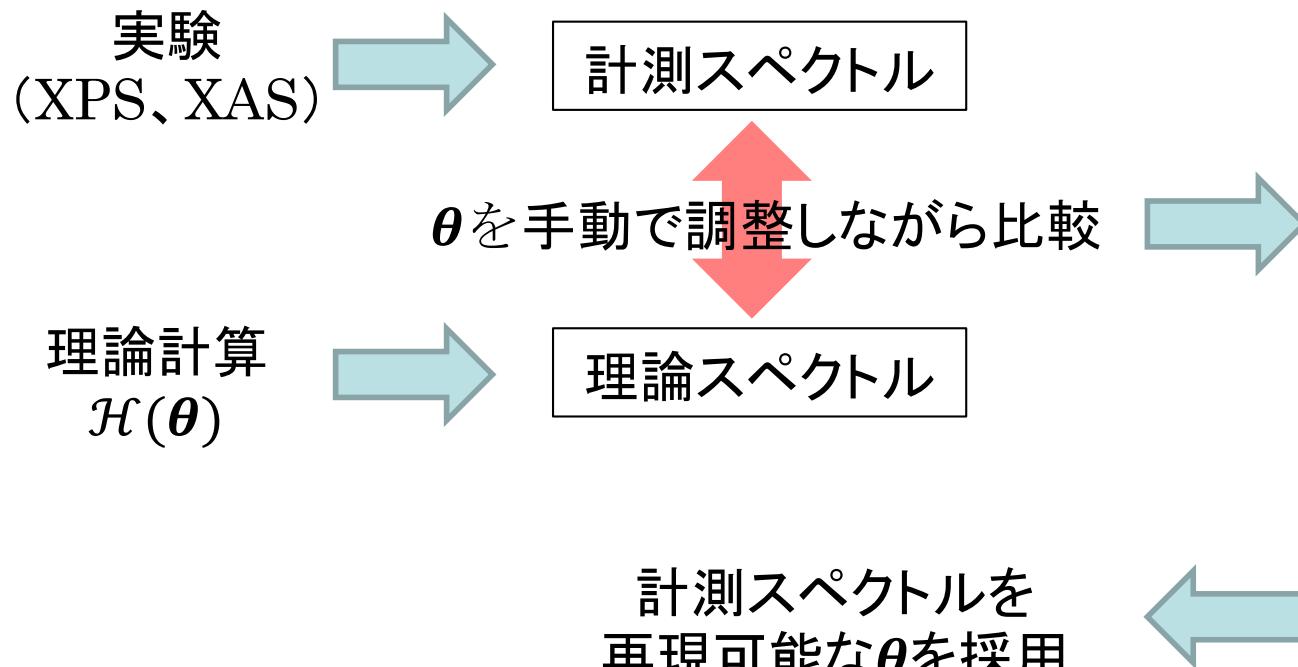
一つの物質に複数の仮説(モデル)

一つの物質に複数のスペクトル

有効ハミルトニアンのモデル選択

- 計測限界を定量的に評価できる枠組みの提案
 - S/N比の低いスペクトル, より少ないデータ
 - マシンタイムの効率化, 大型施設とインラボ装置の役割分担
- 異種情報統合
- 直接観測が困難な情報の取得

XPSとXASの従来型の解析法

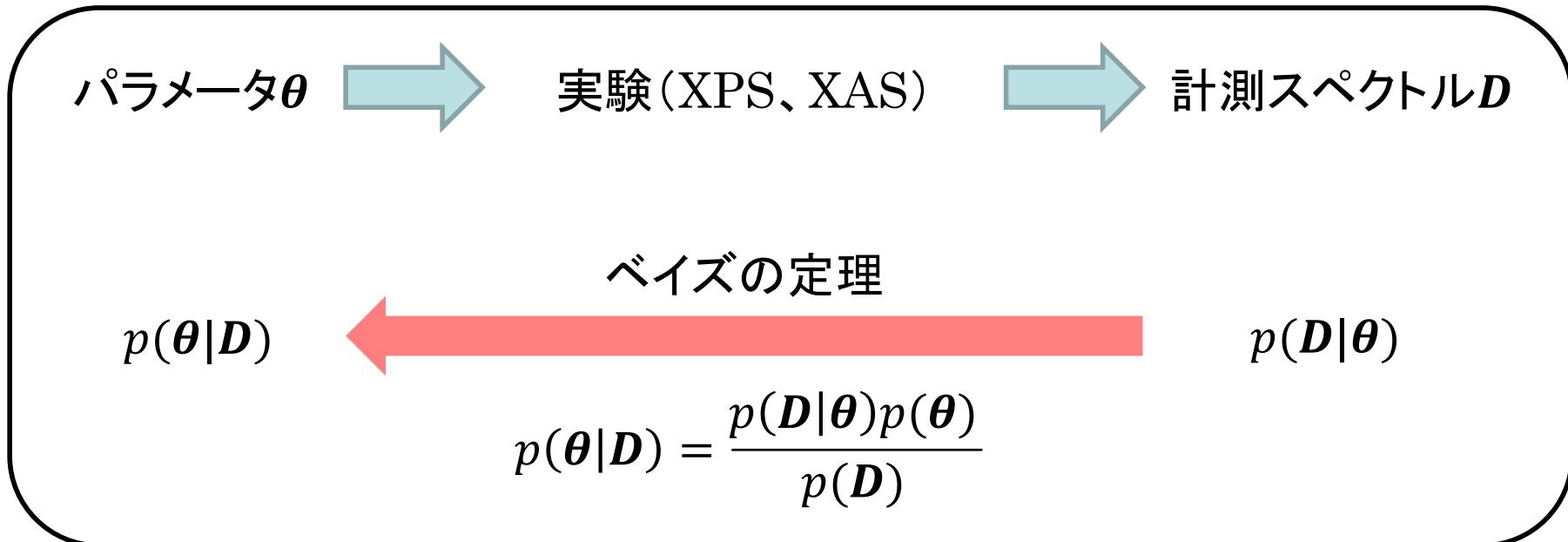


Y. Y et al., APL 107, 033903 (2015).

従来法の課題点

- ・点推定であるため、推定精度を評価することが困難である。
- ・パラメータは手動で調整されるため、解析者の主観が介在する余地がある。
- ・XPSとXASを同時解析する場合、客観的な統合手法が存在しない。

XPSとXASのベイズ推定による解析法



ベイズ推定に基づく解析では、

- ・パラメータの値と精度の両方を推定することが可能になる。
- ・全パラメータ空間を自動で探索するため、客観的な解析が可能になる。
- ・XPSとXASを統合的にベイズ推定することで客観的な同時解析が可能になる。

⇒本研究では、ハミルトニアンのパラメータをベイズ推定の枠組みで統合させ
XPSとXASの統合的解析法(ベイズ統合)を実現させた。

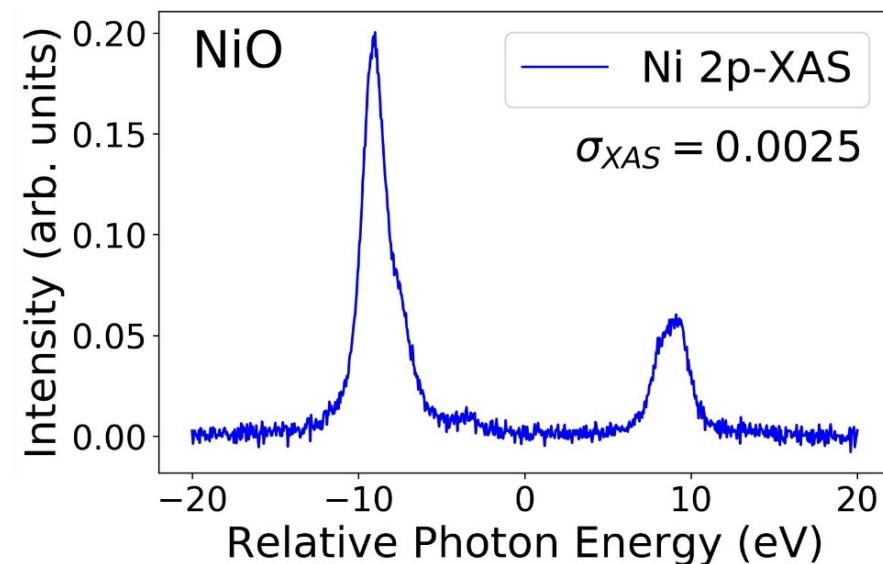
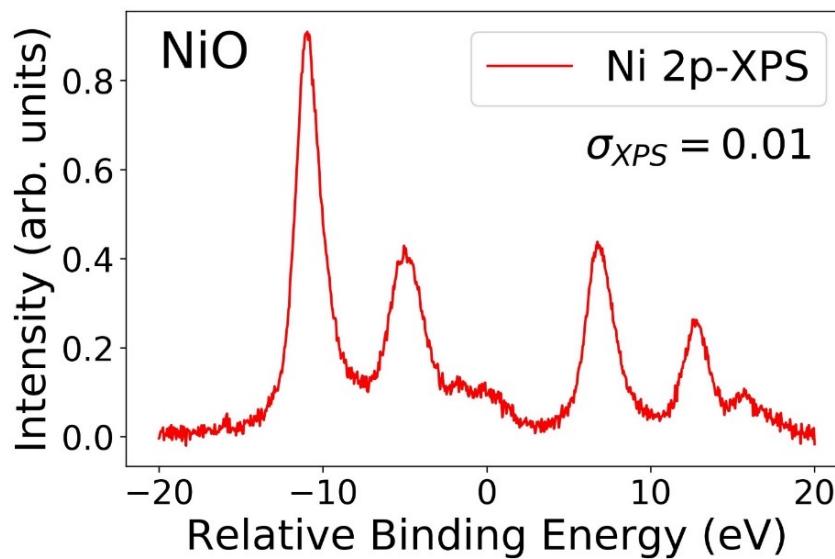
NiOを想定したXPSとXASの理論スペクトル (Yokoyama *et al.*, 2021)

以下のハミルトニアンに基づくクラスターモデル計算

$$H = \sum_{\gamma,\sigma} \varepsilon_{3d}(\gamma) d_{\gamma\sigma}^\dagger d_{\gamma\sigma} + \sum_{m,\sigma} \varepsilon_P p_{m\sigma}^\dagger p_{m\sigma} + \sum_{\gamma,\sigma} \varepsilon_{p_L} a_{\gamma\sigma}^\dagger a_{\gamma\sigma} + \sum_{\gamma,\sigma} V(\gamma) (d_{\gamma\sigma}^\dagger a_{\gamma\sigma} + a_{\gamma\sigma}^\dagger d_{\gamma\sigma})$$

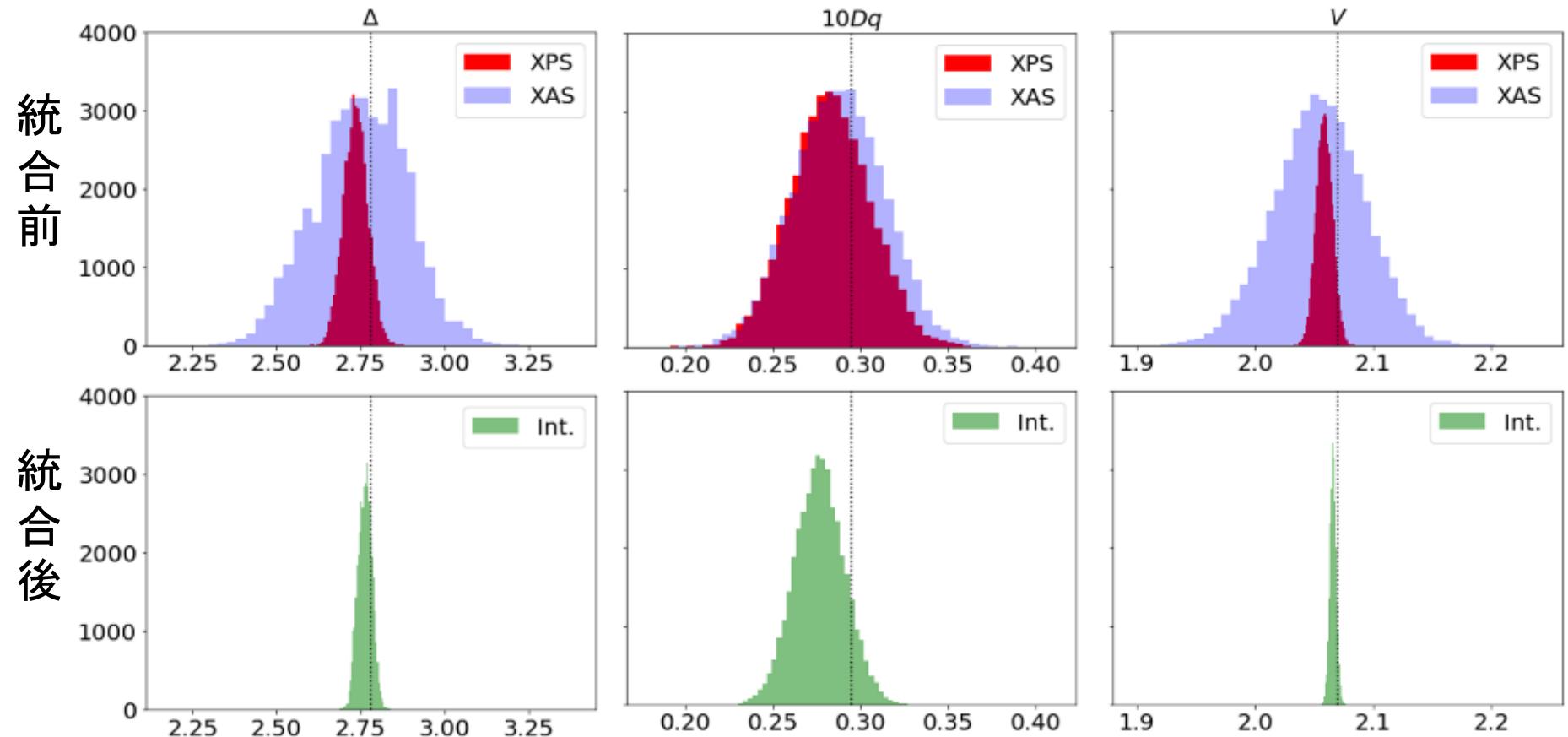
$$+ U_{dd} \sum_{(\gamma,\sigma)>(\gamma',\sigma')} \left(d_{\gamma\sigma}^\dagger d_{\gamma\sigma} d_{\gamma'\sigma'}^\dagger d_{\gamma'\sigma'} \right) - U_{dc} \sum_{\gamma,m,\sigma,\sigma'} d_{\gamma\sigma}^\dagger d_{\gamma\sigma} (1 - p_{m\sigma'}^\dagger p_{m\sigma'}) + H_{multiplet}$$

A 10Dq



※ S/Nが同程度(約1%)になるようにガウシアンノイズの大きさを設定した。

ベイズ推定とベイズ統合の比較



ベイズ統合前後を比較すると、

- ・全てのパラメータで統合後の推定精度が向上した。 \Rightarrow 情報の統合に成功。

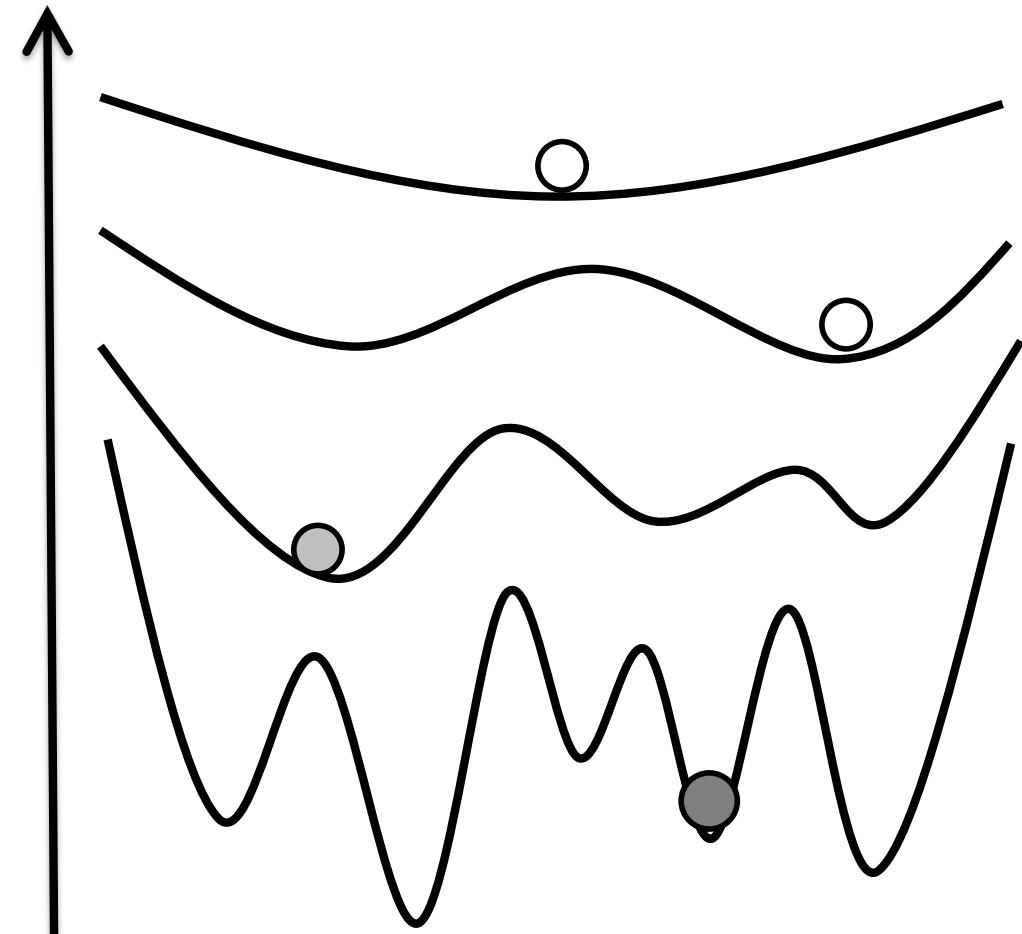
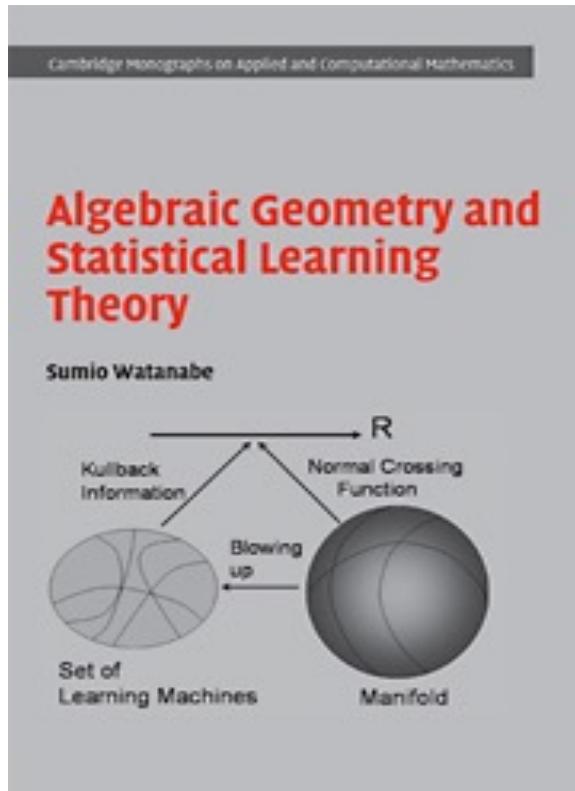
内容

- 自己紹介
- 課題1(ベイズ計測)
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

SPring-8

- ・ アメリカのAdvanced Photon Source (APS),ヨーロッパのEuropean Synchrotron Radiation Facility(ESRF) と合わせて, **世界3大放射光施設**.
- ・ 理研はSPring-8を「データ創出基盤」であると言っている. **年間延べ1万人**が利用.
- ・ APSやESRFにおいてベイズ計測は導入されていない.
- ・ 放射光におけるベイズ計測に関しては**日本が最先端**である.

APSやESRFにおいてベイズ計測が導入されていない理由



多くのスペクトル解析が統計的特異モデルであることを知らない

レプリカ交換モンテカルロ法の知見の欠如

内容

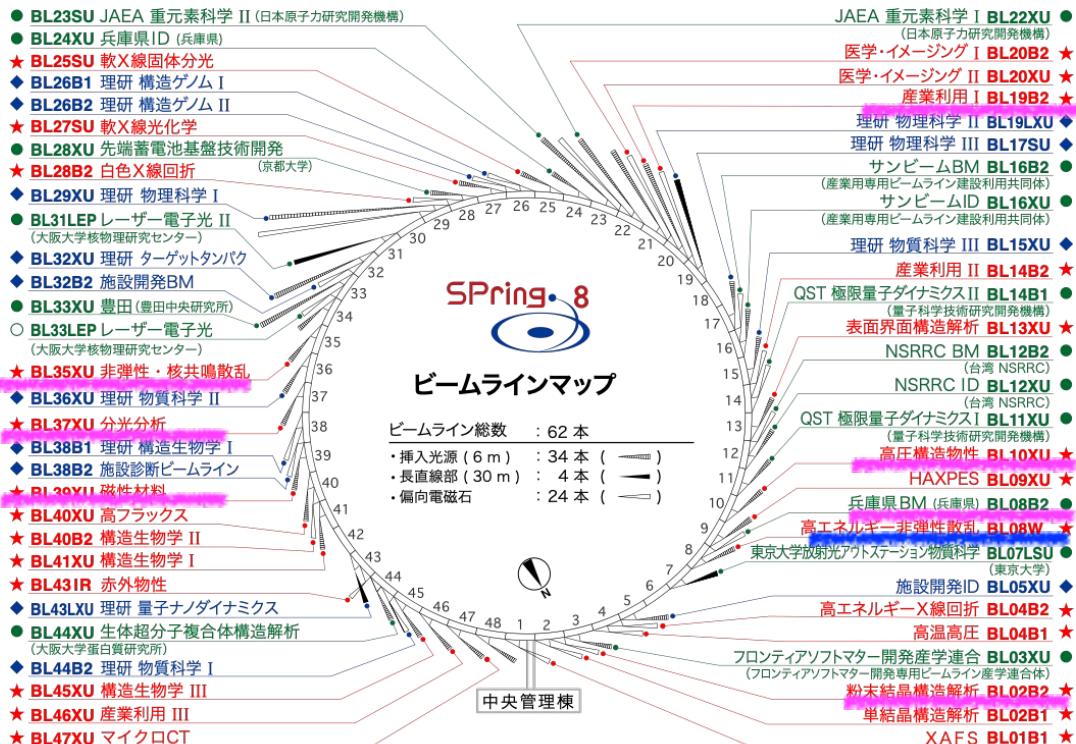
- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - **ベイズ化計画**
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点



水牧 仁一朗
公益財団法人高輝度光科学研究中心
放射光利用研究基盤センター
コーディネータ
4月1日より熊本大学大学院先端科学研究院

SPring-8全ビームラインベイズ化計画

敬称略



赤色BLが共用BL(JASRI担当): 計26本
全BL本数: 62本

来年度には過半数をこえる予定

情報と放射光研究者のマッチング

メスバウアー
BL35XU

岡田研学生+筒井

小角散乱
BL08B2
BL19B2

岡田研学生+桑本

XAS測定
BL37XU
BL39XU

岡田研学生+水牧

放射光ユーザーへの展開

時分割XRD
BL02B2
BL10XU

横山優一+河口彰吾、沙織
ユーザー: 公立大、東工大

年度	2021	2022	2023
導入	2	8	14
全BL	26	26	26

内容

- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

第4回データ駆動科学によるデータ解析 高度化～ベイズ計測～

1. 日程: 2023年03月07日 13時00分～15時40分
2. 開催形式:ハイブリッド開催 (使用ツール:Zoom):
場所: SPring-8上坪講堂
3. 主催: 公益財団法人高輝度光科学研究中心
(JASRI)
4. 後援(予定): SPring-8ユーザー協同体(SPRUC)
データ駆動科学研究会
SPring-8利用推進協議会 データ科学研究会

第4回データ駆動科学によるデータ解析 高度化～ベイズ計測～

5. 趣旨

データ駆動科学によるデータ解析高度化における活動として、「第3回データ駆動科学によるデータ解析高度化(2021年10月28日JASRIにて開催済)」では、データ駆動科学技術であるスペースモデリングについて放射光測定・解析への適用に関する最先端研究例を紹介した。その際にデータ駆動科学技術が様々な放射光測定で利用可能であるとの報告を行った。

これまで単発的な導入であったが、達組織データ駆動科学グループを立ち上げ、データ駆動科学の導入を組織的に行うことができるようになった。東京大学・岡田真人教授を本グループ客員研究員として迎え、ベイズ計測をSPring-8供用ビームラインの解析に導入を試みている。本ワークショップにおいては岡田教授に基調講演としてベイズ計測についてご紹介いただき、各ビームラインで行われている測定・解析への導入例について、各講演者に紹介していただき、データ駆動科学のさらなる多くのビームラインへの導入の契機としたい。

6. 放射光測定・計測データ解析にデータ駆動科学の技術の導入を検討しているユーザー。

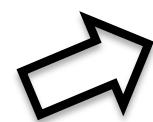
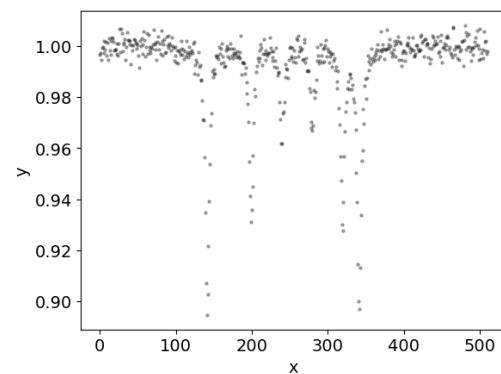
7. 参加者 206名(参加登録 246名)

内容

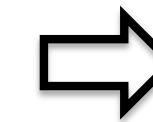
- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - **プロジェクト紹介**
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

メスバウアーハミルトニアン推定

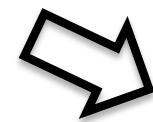
- ・ベイズ計測により、データ駆動でハミルトニアン推定やモデル推定が可能
- ・物理量の事後分布が得られることにより、解析結果の議論の幅が広がる



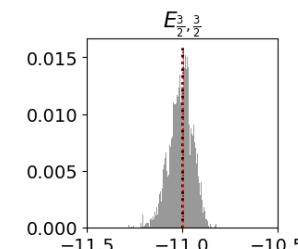
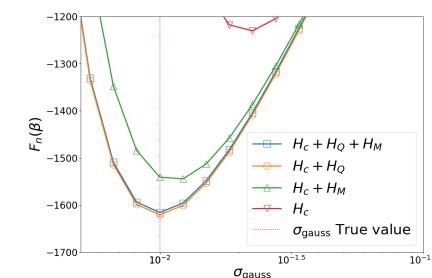
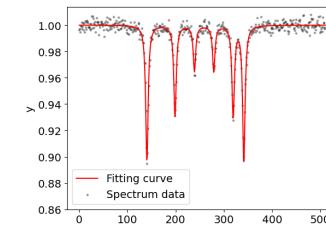
フィッティング



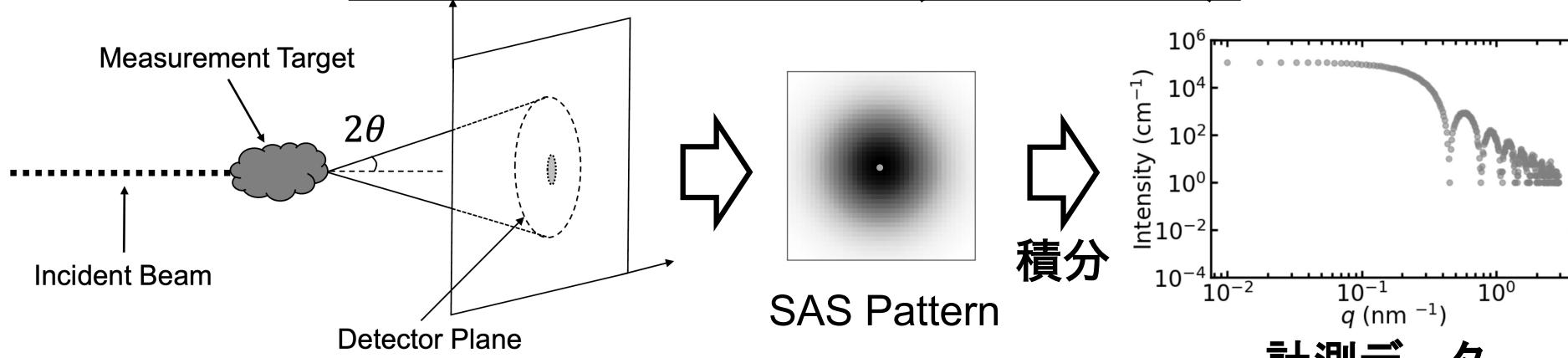
ハミルトニアン推定



パラメータ分布

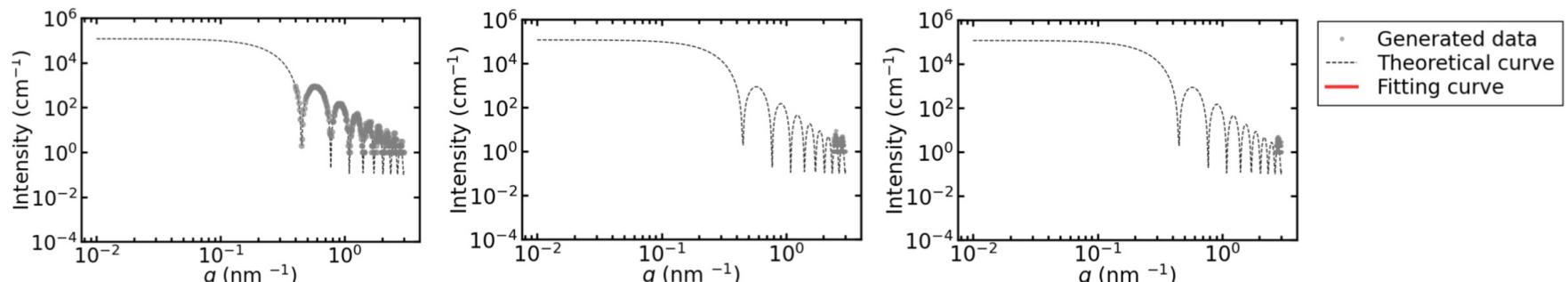


X線小角散乱(SAXS)

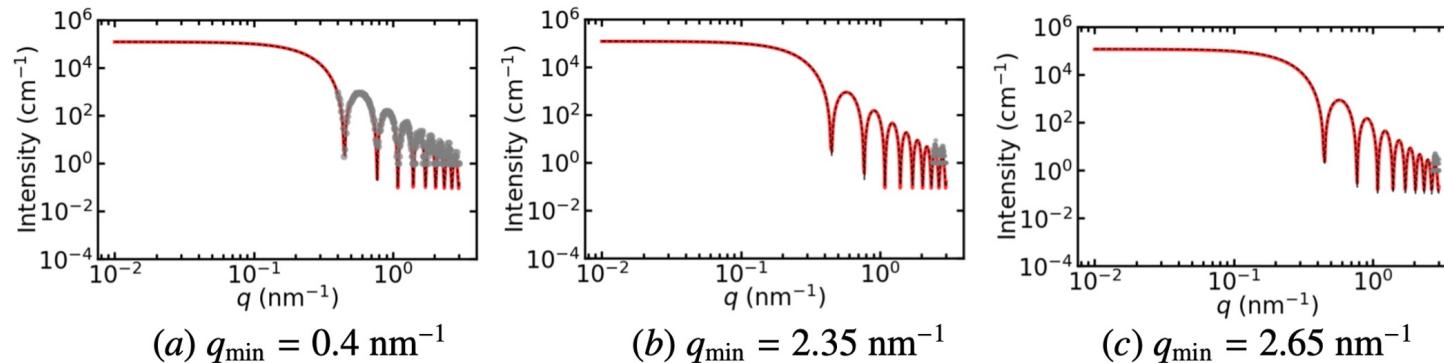


ベイズ計測

人工データ



フィッティング

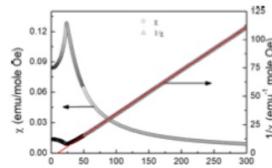


将来展望 真のマルチモーダル測定へ

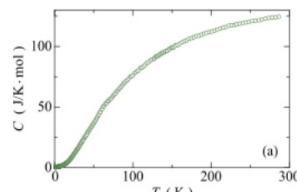


研究室

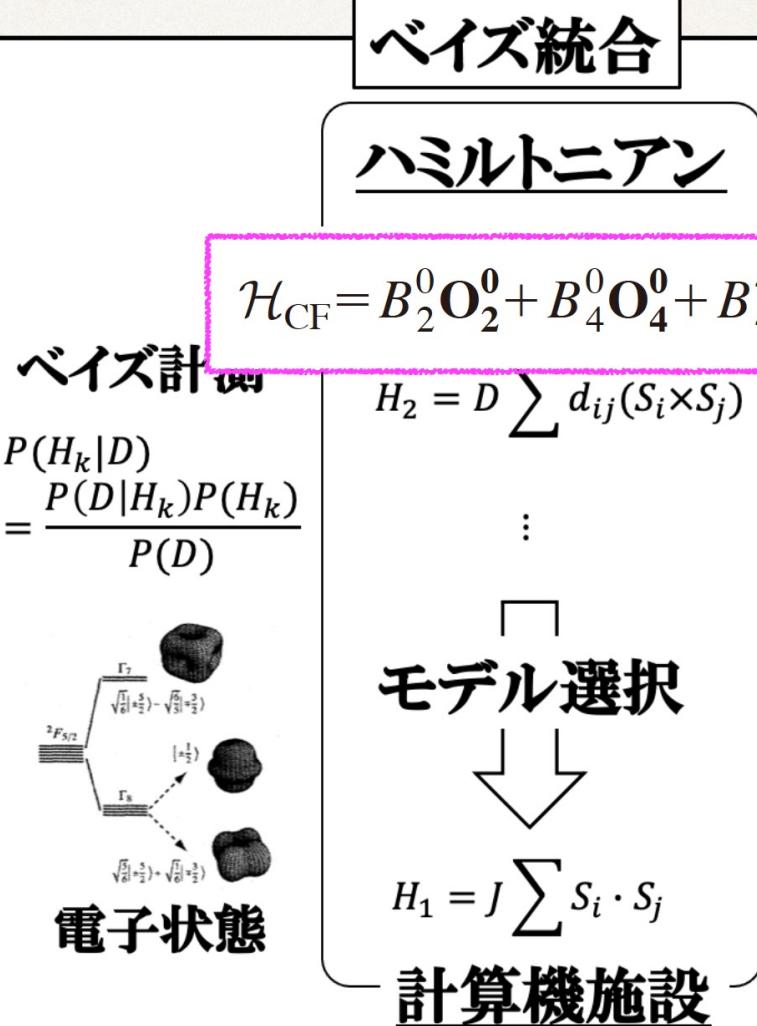
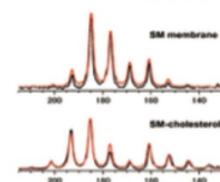
磁化測定



比熱測定

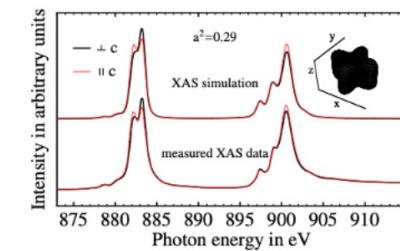


NMR測定

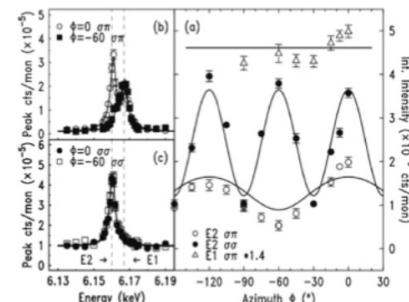
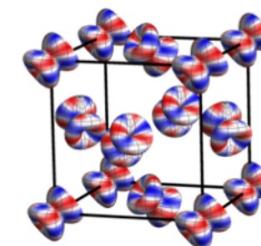


大型測定施設

分光測定



回折測定



例えば、物性物理で考えると、結晶場や磁気相互作用パラメータなど

内容

- 自己紹介
- 課題1（ベイズ計測）
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- 本プロジェクトの到達点

第88回SPring-8先端利用技術ワークショップ
第4回データ駆動科学によるデータ解析高度化～ベイズ計測～

2023.3.7



The University of Tokyo

ベイズ推定のスキーム

～全てのユーザーのために～

東京大学
岡田研究室

片上 舜





ベイズ計測オープンソースソフトウェアの構築





01 既存のベイズ推定ライブラリ

マルコフ連鎖モンテカルロ法 (MCMC) のライブラリ

- Stan
- PyMC3
- JAGS
- emcee
- BUGS

交換モンテカルロ法のライブラリ

- Tensorflow
- emcee
- ptmcmc sampler

01

計測科学における理想的なベイズ推論ツール



求められる機能性

- ・ 基本的なベイズ推論ワークフローを完備
- ・ 迅速に実装可能かつ柔軟なモデル構築が可能なUI
- ・ 実行解析結果の可視化
- ・ ベイズ推論の高速な実行

ベイズ推論

$$P(\theta|D) \propto P(D|\theta) \times P(\theta)$$

事後分布 尤度(モデル) 事前分布

θ : 物理量 (モデルパラメータ)

D : 計測データ

モデルの実装



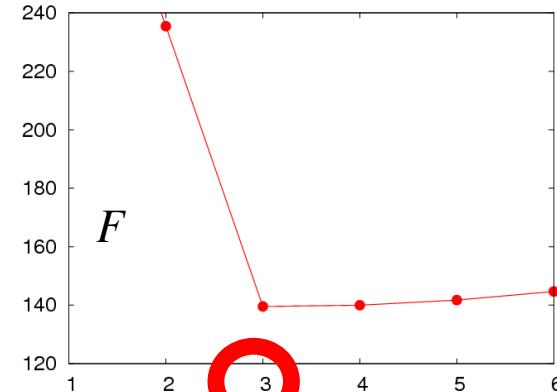
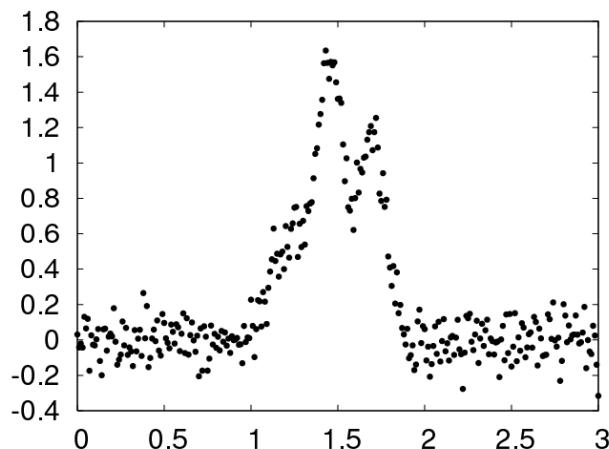
データの取り込み



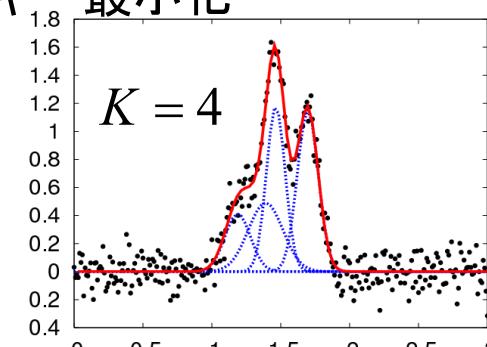
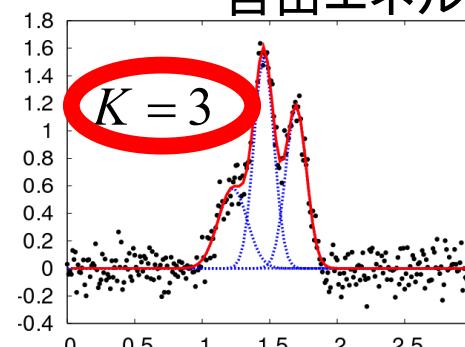
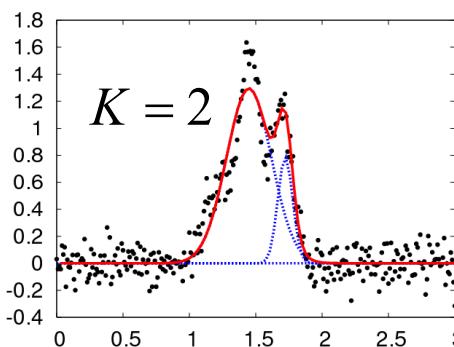
解析結果の確認

ベイズ推論ワークフロー

まずはスペクトル分解



最適な K をデータだけがら決める
自由エネルギー最小化



Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

内容

- 自己紹介
- 課題1(ベイズ計測)
 - スペクトル分解の従来法
 - ベイズ計測
 - レプリカ交換モンテカルロ法
 - モデル選択
 - 計測限界の定量的評価
 - 異種計測の情報統合
- SPring-8全ビームラインベイズ化計画
 - SPring-8とは
 - ベイズ化計画
 - 「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」報告
 - プロジェクト紹介
 - ベイズ計測用ライブラリ EMC (岡田研 片上助教)
- **本プロジェクトの到達点**

本プロジェクトの到達点

- 目的
 - ベイズ推論を計測科学に導入した**ベイズ計測**により、計測科学がどう変わるかの具体例を提示し、その具体例の中から、ベイズ計測の情報数理科学的な学理の構築を目指す。
- 成果
 - ベイズ推論とレプリカ交換モンテカルロ法で計測科学は変わることを示した
 - CREST領域の目的を遂行
 - ベイズ計測による情報計測領域への横断的展開の一つとして、SPring-8全ビームラインベイズ化計画を推進
 - あいちSRでも同様のシンポジウムを行う予定
 - 東大新領域複雑理工学専攻内でのベイズ計測の水平展開