

ベイズ計測と SPring-8全ビームラインベイズ化計画

岡田真人

東京大学 大学院新領域創成科学研究科

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 –)

共同研究者

永田 賢二 国立研究開発法人物質・材料研究機構

村岡 恵 東京大学 大学院新領域創成科学研究所

本武 陽一 一橋大学大学院
ソーシャル・データサイエンス研究科

杉田 精司 東京大学 大学院理学系研究科

佐々木 岳彦 東京大学 大学院新領域創成科学研究所

アンケート

- ・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。

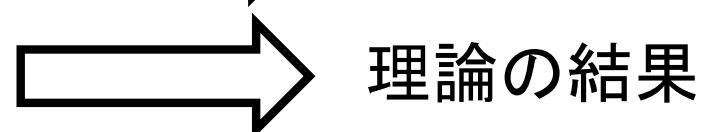
ベイズ計測

順アプローチ

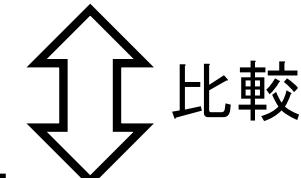
計測データ



モデル



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



逆アプローチ

対象とする
物理系

系の物理

モデル

$$p(\theta, K)$$

観測過程

計測機器の特性

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

計測データ

全てをモデル化し
ベイズの定理で因果をさかのぼる

[情報計測] 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

[◀ トップに戻る](#)

■ 領域活動についてはこちらをご覧ください。

| 戰略目標

「材料研究をはじめとする最先端研究における計測技術と高度情報処理の融合」

| 研究総括



雨宮 慶幸(（公財）高輝度光科学研究センター 理事長)

| 副研究総括



北川 源四郎(東京大学 数理・情報教育研究センター 特任教授)

[情報計測] 平成29年度採択課題

[← トップに戻る](#)

| 岡田 真人

ベイズ推論とスパースモデリングによる計測と情報の融合

| 研究代表者

岡田 真人



東京大学
大学院新領域創成科学研究科
教授

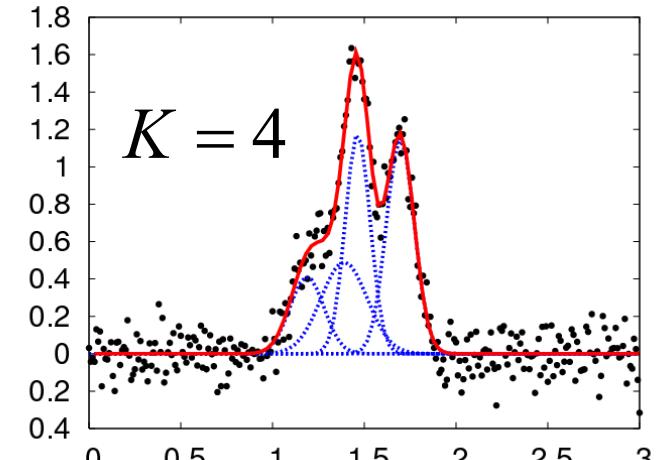
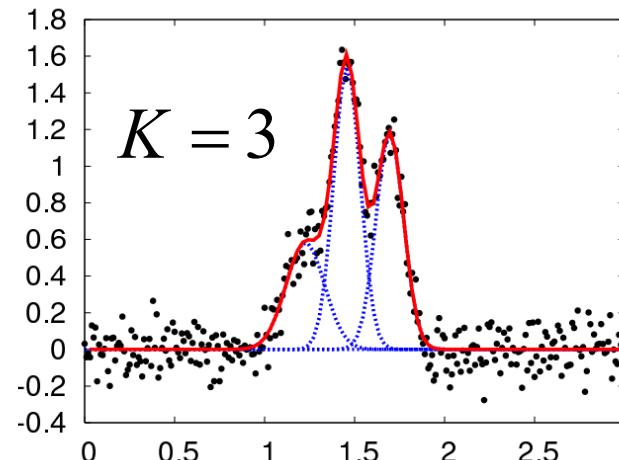
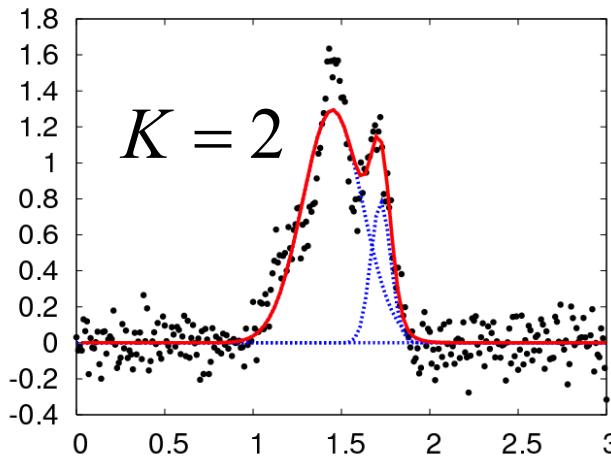
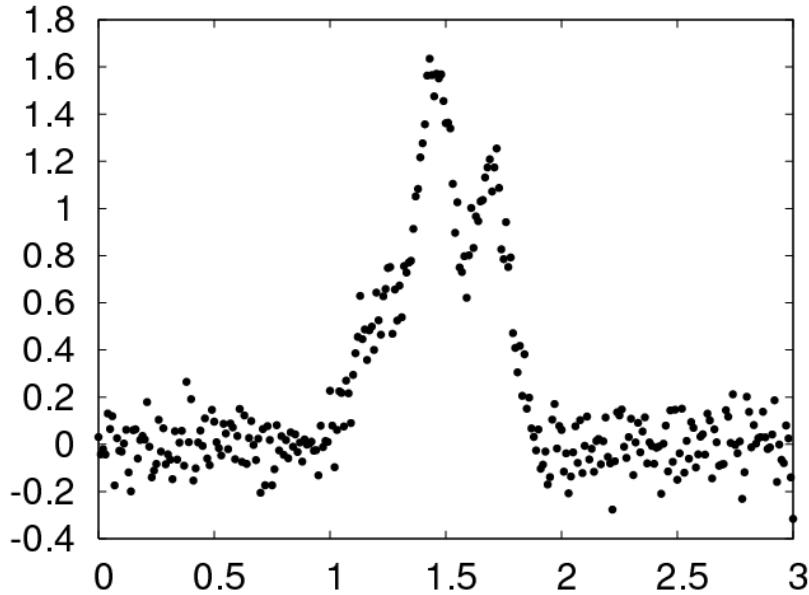
| 主たる共同研究者

桑谷 立	海洋研究開発機構 海域地震火山部門 グループリーダー代理
長尾 大道	東京大学 地震研究所 准教授
成田 憲保	東京大学 大学院総合文化研究科 教授
日野 英逸	情報・システム研究機構 統計数理研究所 教授

内容

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

例題1: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

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

スペクトル分解の定式化

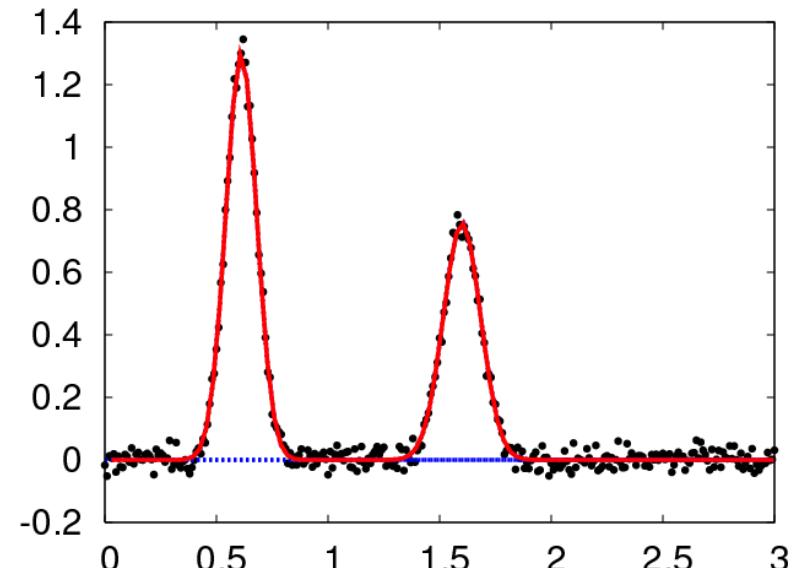
ガウス関数(基底関数)の足し合わせにより、スペクトルデータを近似

観測データ: $D = \{x_i, y_i\}_{i=1}^n$

x_i : 入力 y_i : 出力

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

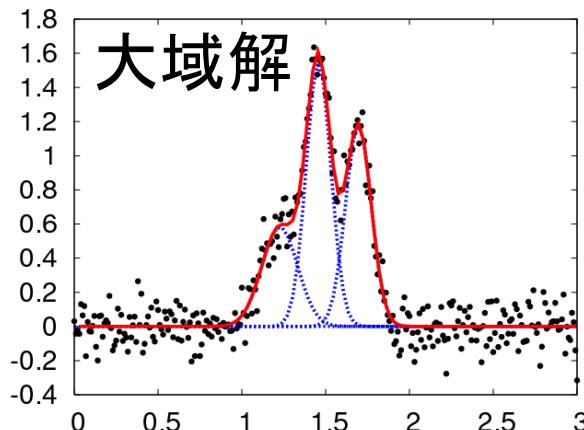
$$\theta = \{a_k, b_k, \mu_k\} \quad k = 1, \dots, K$$



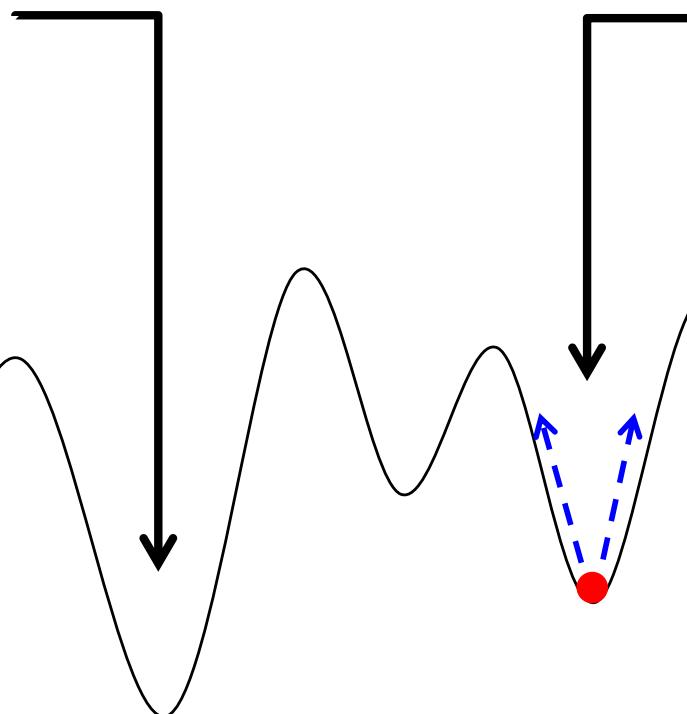
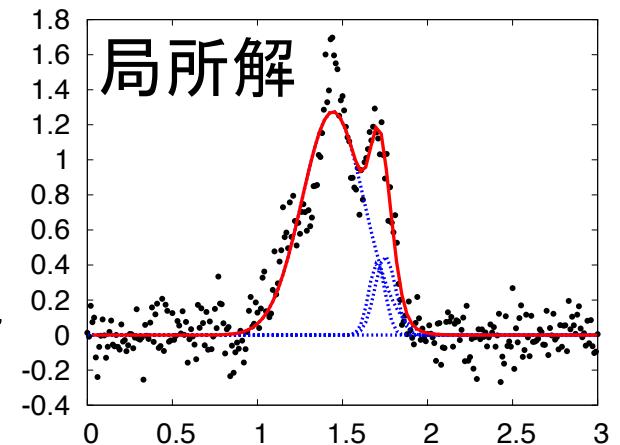
二乗誤差を最小にするようにパラメータをフィット(最小二乗法)

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

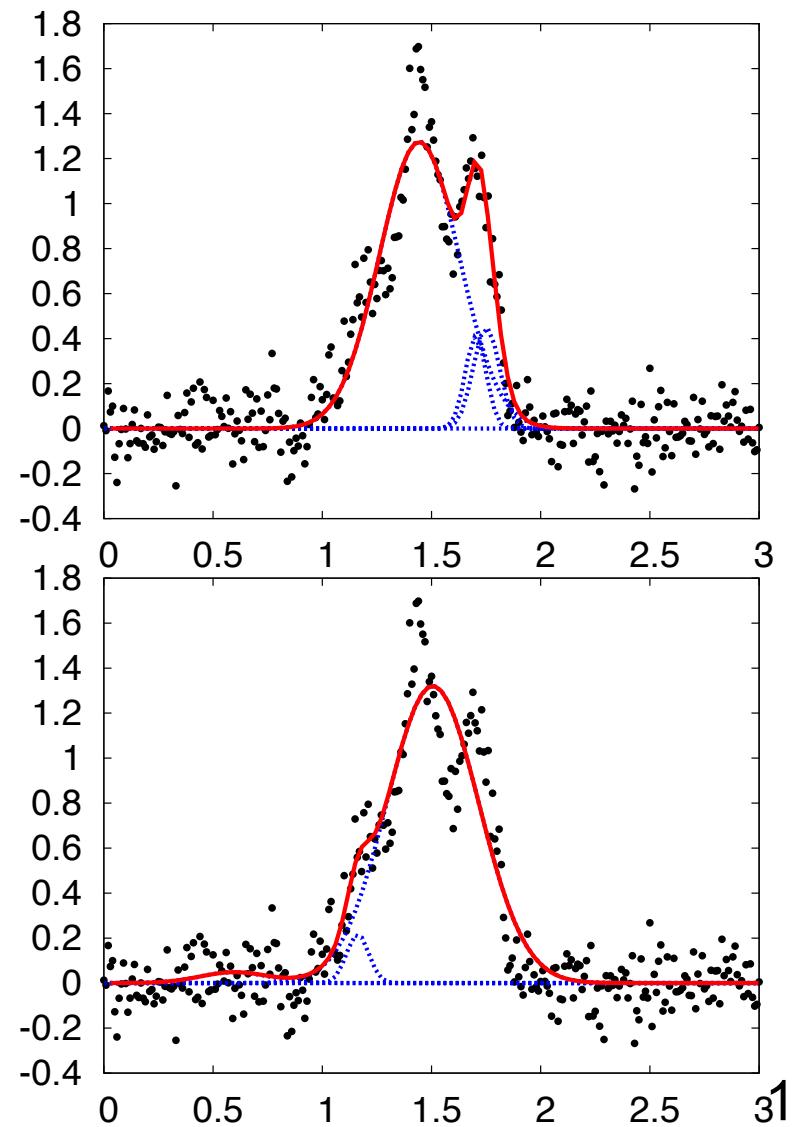
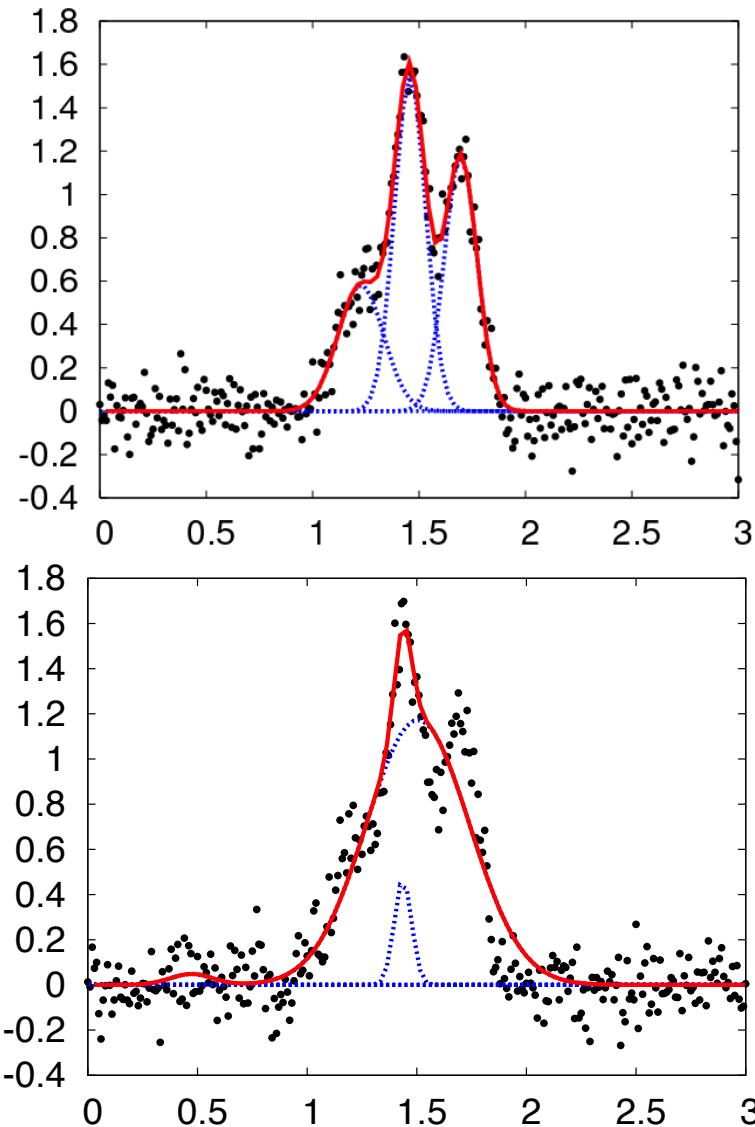
誤差関数は局所解を持つ



<通常の最適化法>
e.g., 最急降下法



ローカルミニマム



アンケート

- ・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。

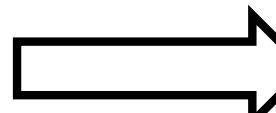
内容

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

ベイズ計測

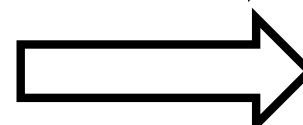
順アプローチ

計測データ

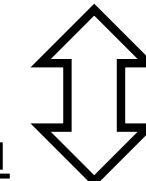


実験の結果

モデル



理論の結果

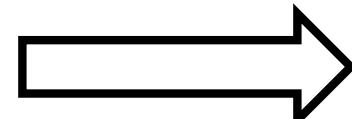


比較

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

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

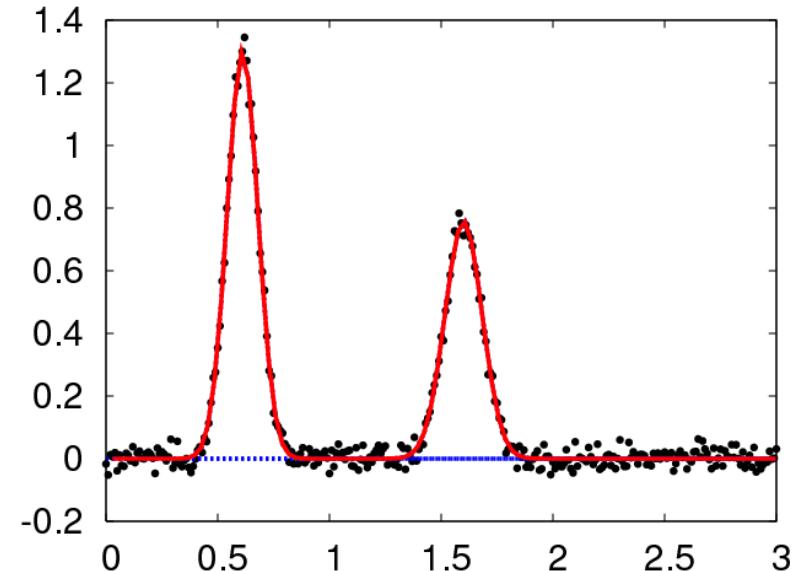
出力は、入力からの応答とノイズの足し合わせにより生成

⇒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、

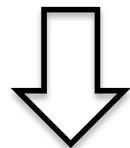
$$p(Y | \theta) = \prod_{i=1}^n p(y_i | \theta) \propto \exp(-nE(\theta)) \quad Y = \{y_1, \dots, y_n\}$$

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

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

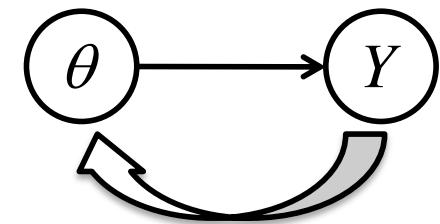
$$p(Y, \theta) = \frac{p(Y | \theta)p(\theta)}{p(Y)}$$



<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

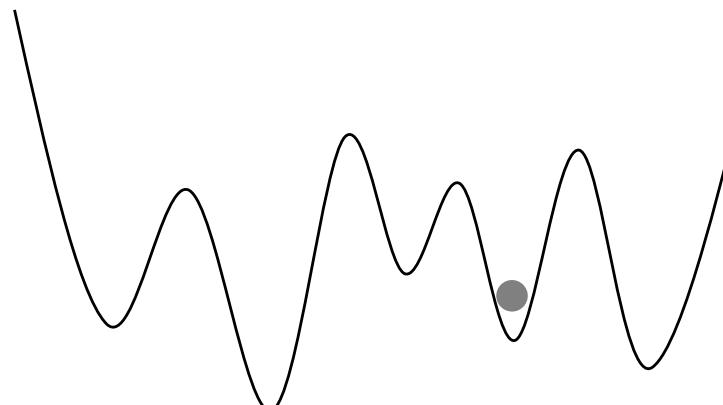
内容

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

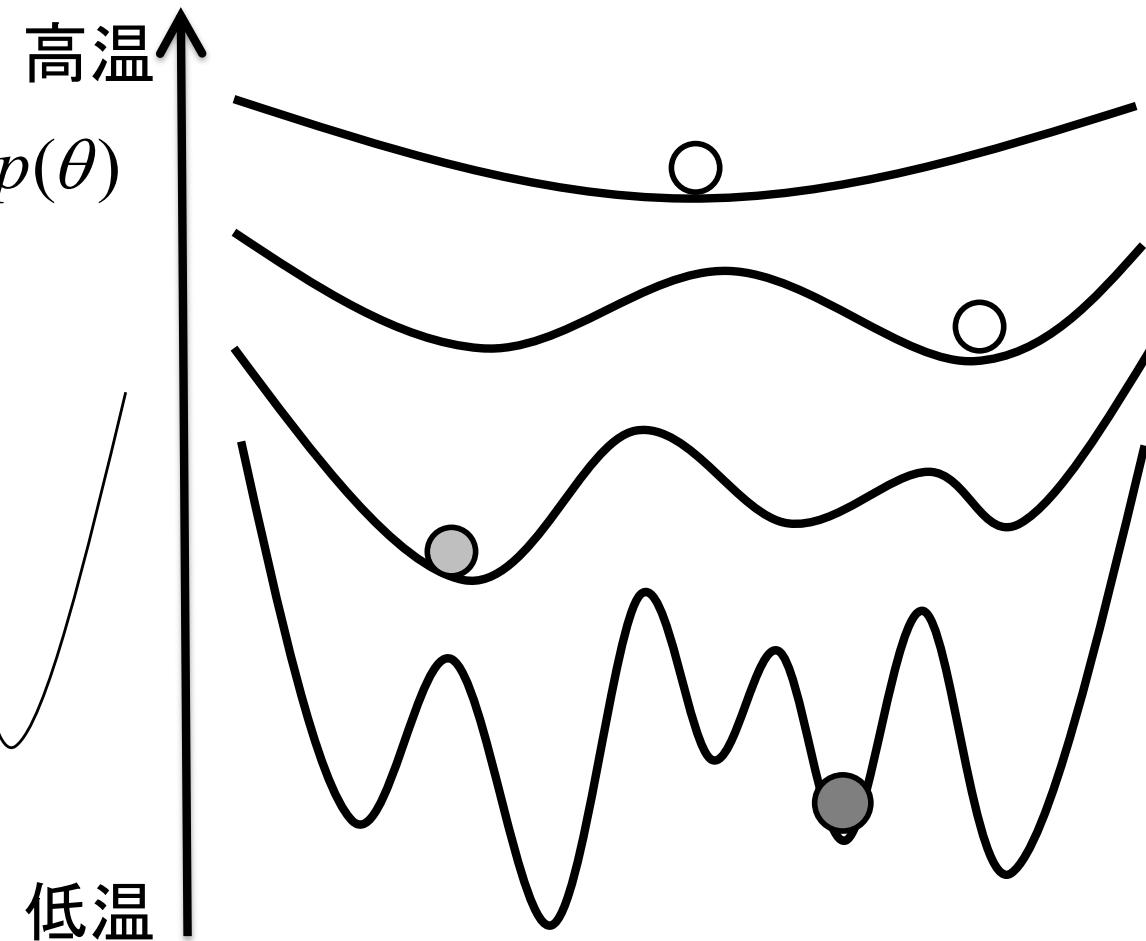
レプリカ交換モンテカルロ法 ランダムスピニ系の知見から

メトロポリス法

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

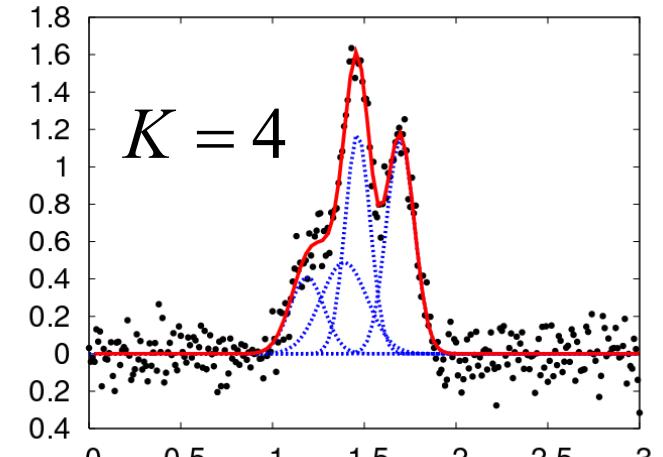
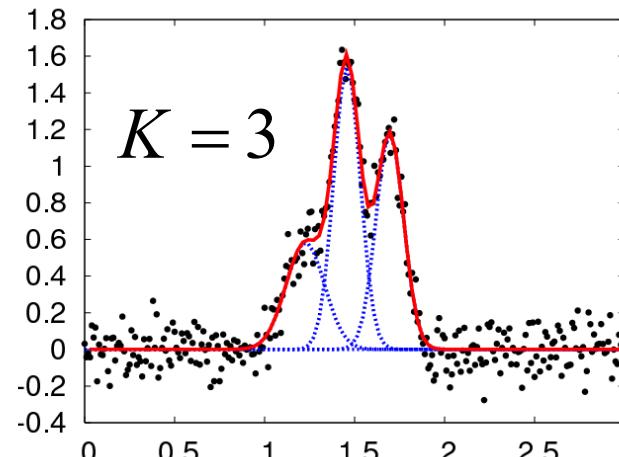
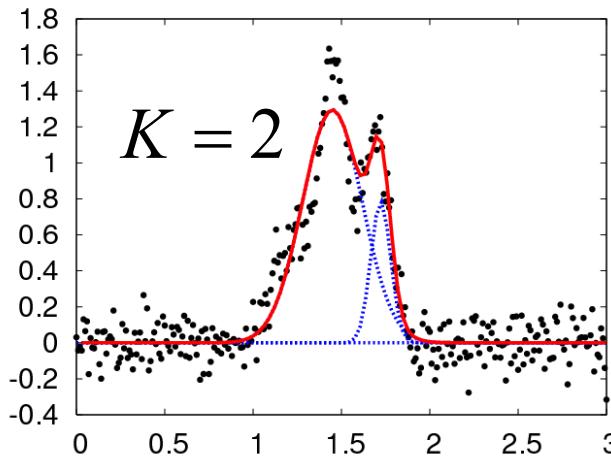
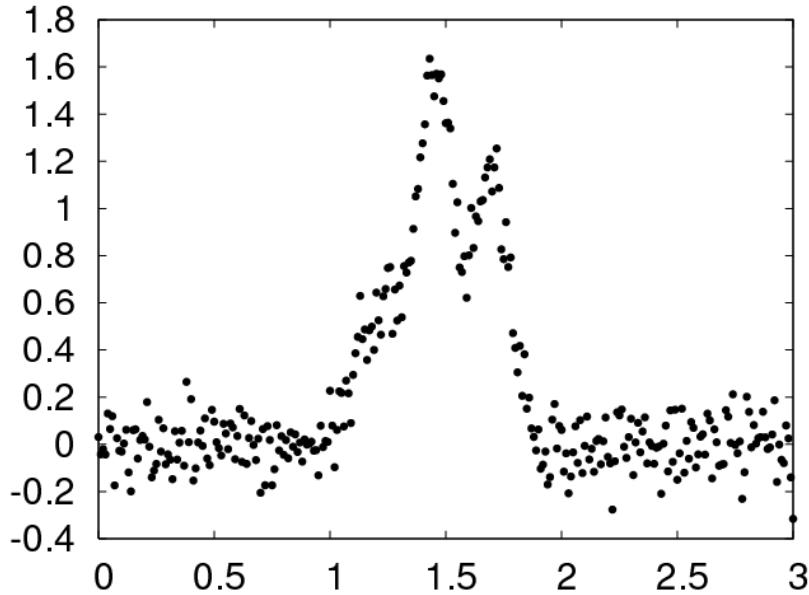


レプリカ交換モンテカルロ法



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

例題1:スペクトル分解



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

アンケート

- ・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。

アンケート

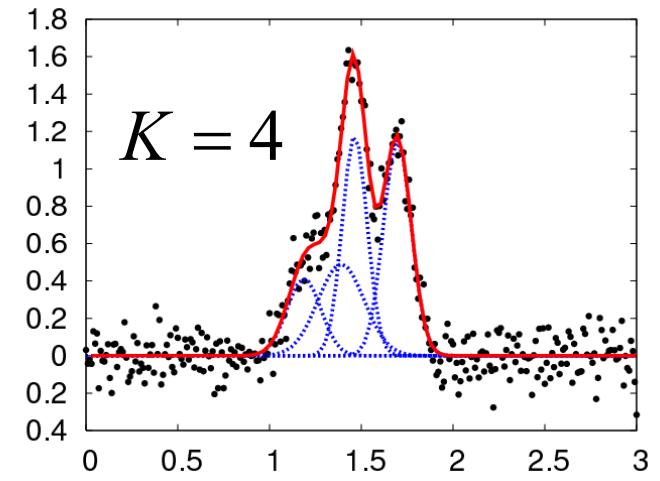
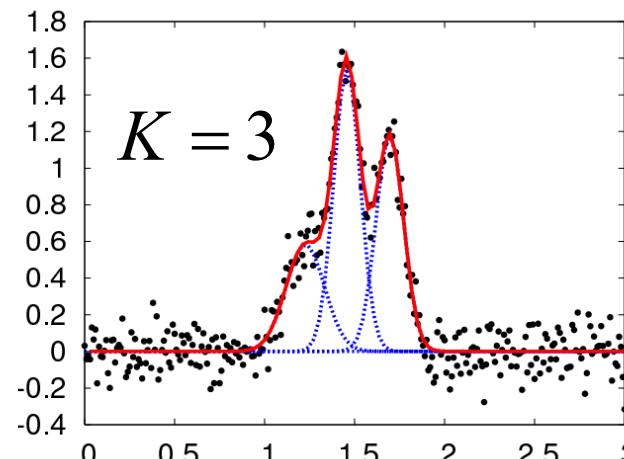
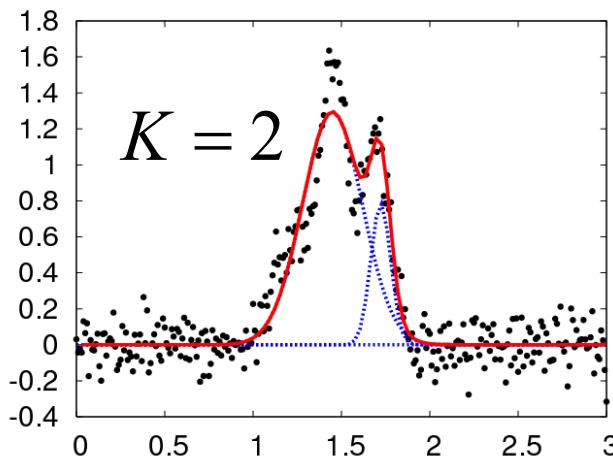
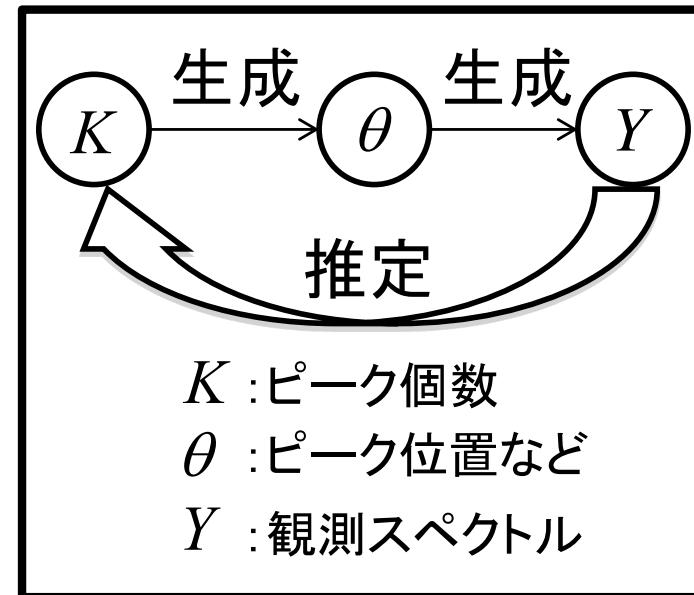
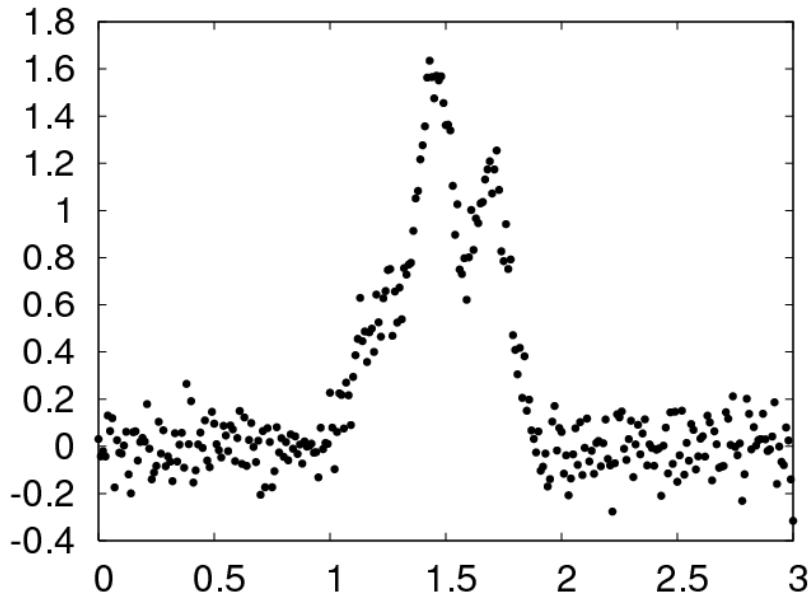
- ・スペクトルや画像データからフィッティングを行なっている
- ・そのフィッティングの際に、パラメータを手打ちで決めている。最急降下法などを使っているが、うまくいかない。
- ・フィッティング用のモデルが複数あって、事前にどれを使うかを決めておかないといけない。
- ・S/Nが悪いデータや欠損データをなんとかした。
- ・複数計測の統合を行いたい。

- ・そのような方は、一度ベイズ計測をお試しください。

内容

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

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

アンケート

- ・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。

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

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

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

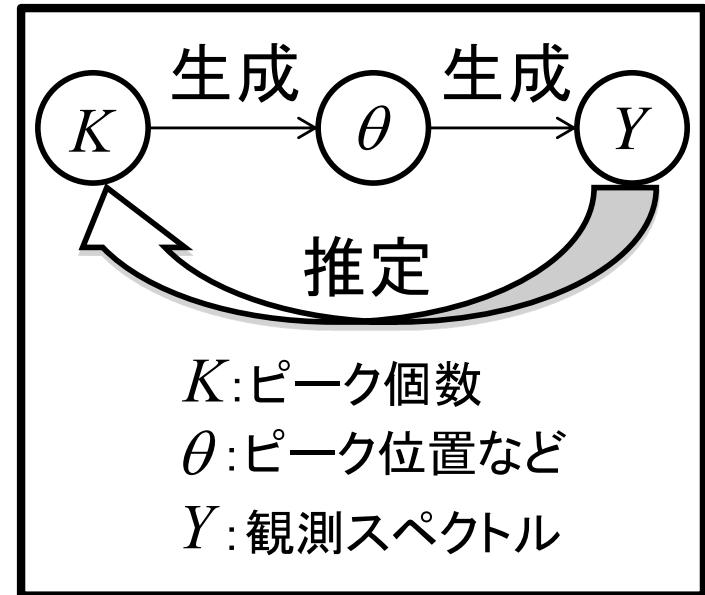
- 無駄な自由度の系統的消去: 周辺化, 分配関数

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

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

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

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

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

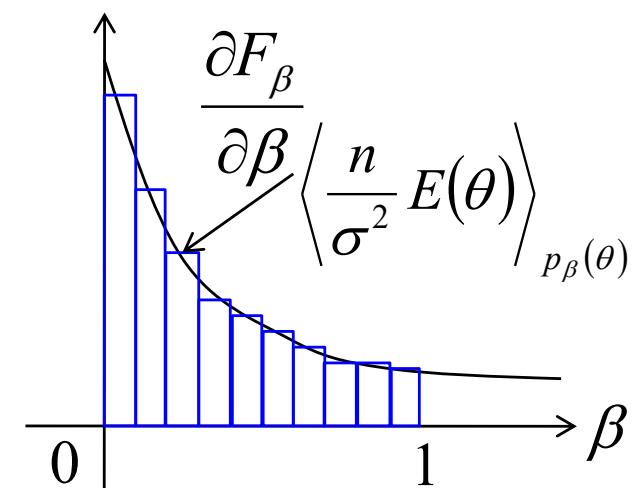
$$F = F_{\beta=1} = \int_0^1 d\beta \frac{\partial F_\beta}{\partial \beta}$$

たくさんの温度でのシミュレーションが必要

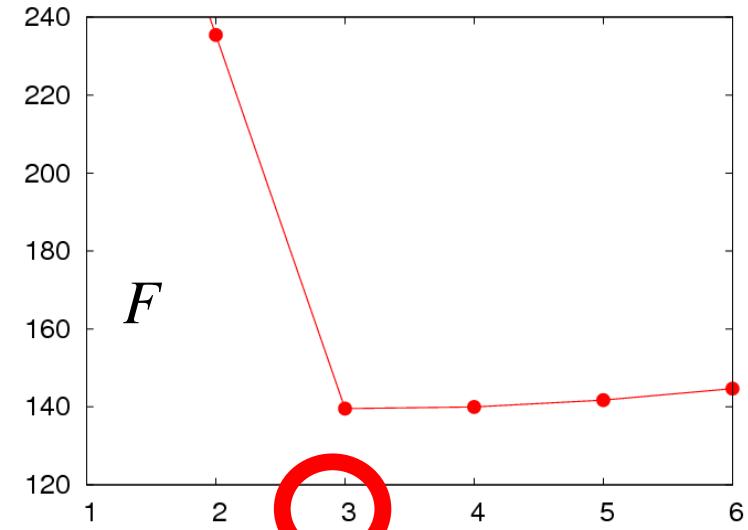
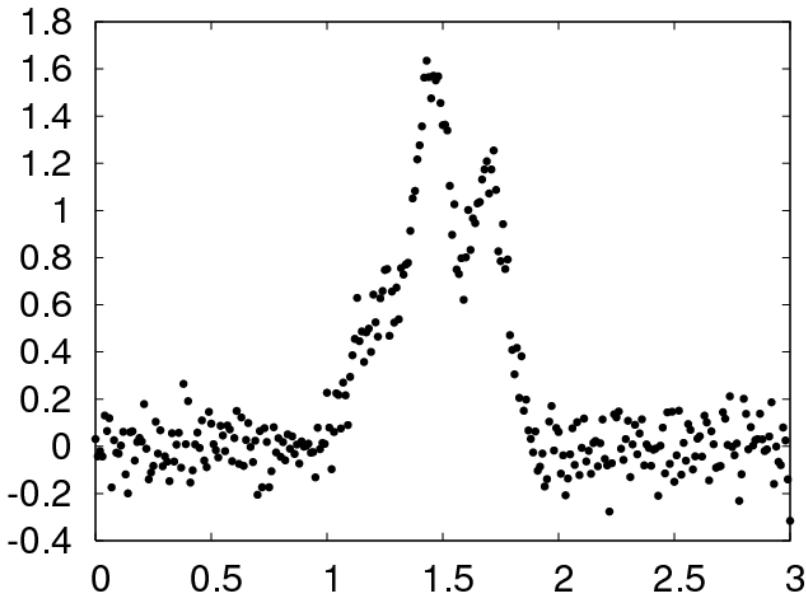
\rightarrow 各温度でのエネルギー平均(すでにやってる)

$\frac{\partial F_\beta}{\partial \beta}$... 確率分布 $p(\theta; \beta)$ に従う
 $\frac{\partial F_\beta}{\partial \beta}$... 二乗誤差 $\frac{n}{\sigma^2} E(\theta)$ の期待値

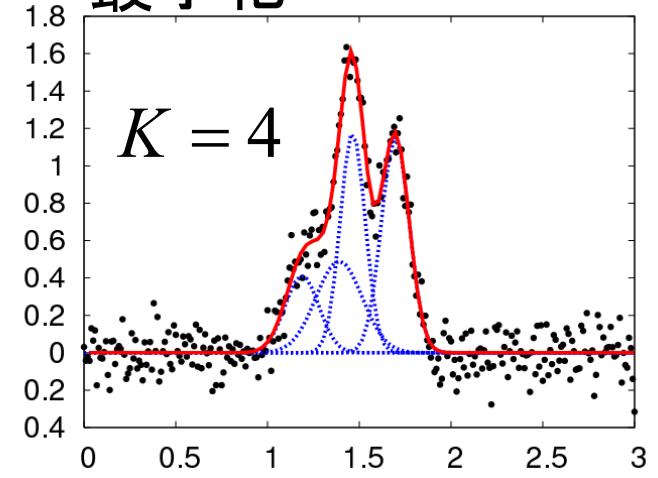
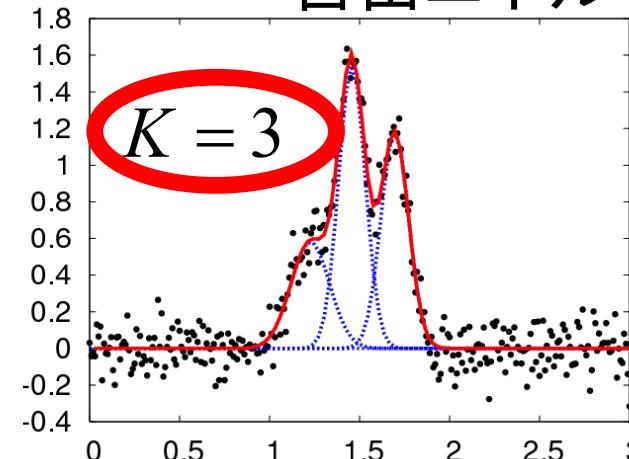
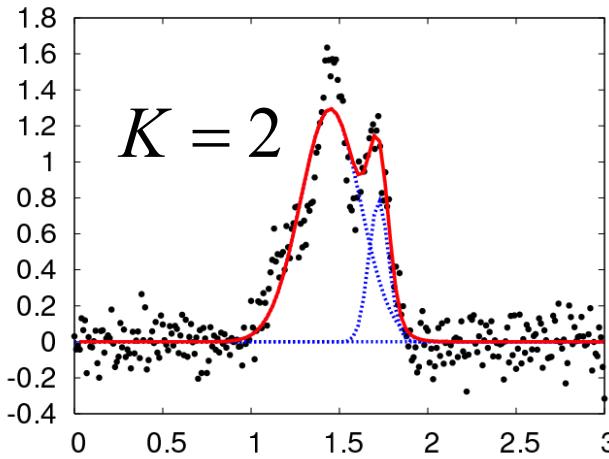
$$p_\beta(\theta) \propto \exp\left(-\frac{n}{\sigma^2} \beta E(\theta)\right) p(\theta)$$



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

アンケート

- ・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。

アンケート

- ・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。

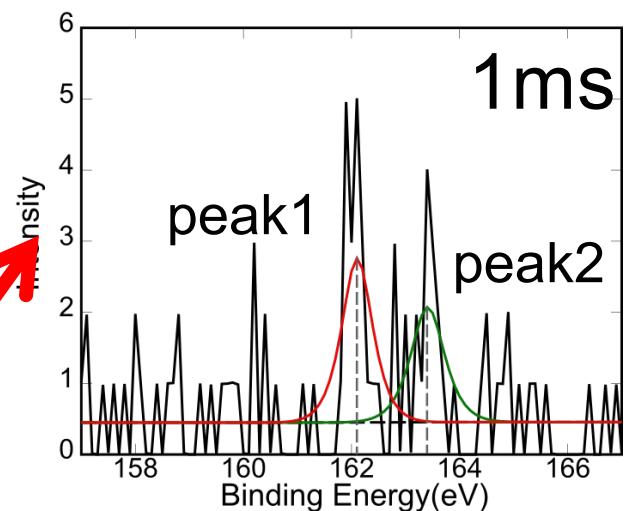
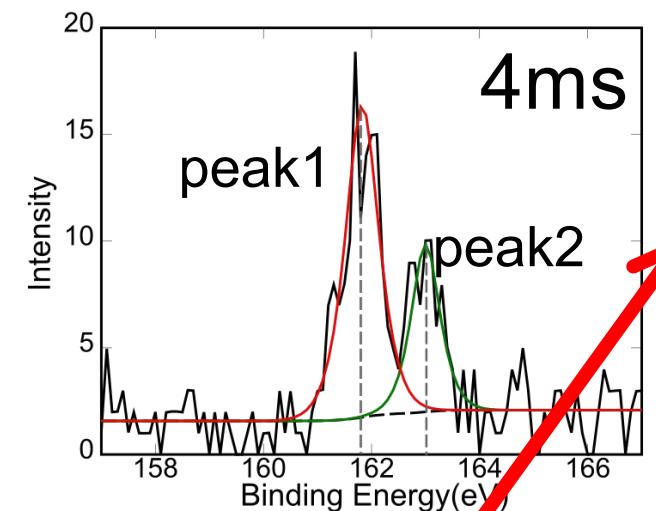
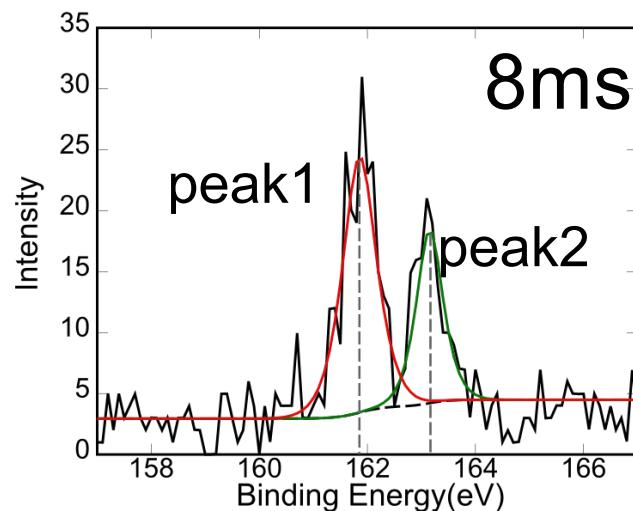
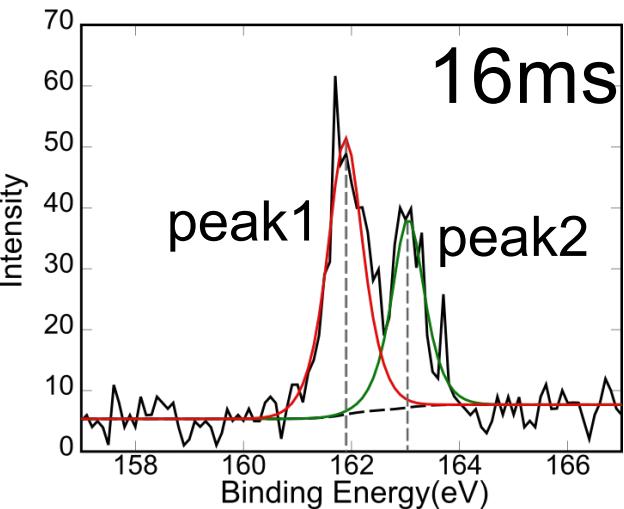
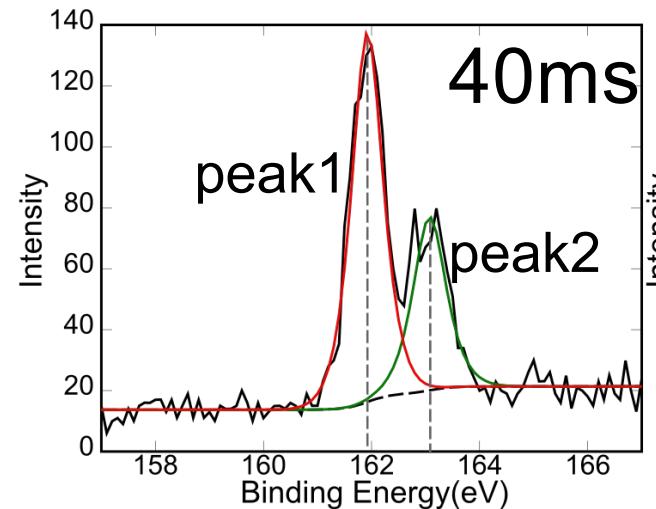
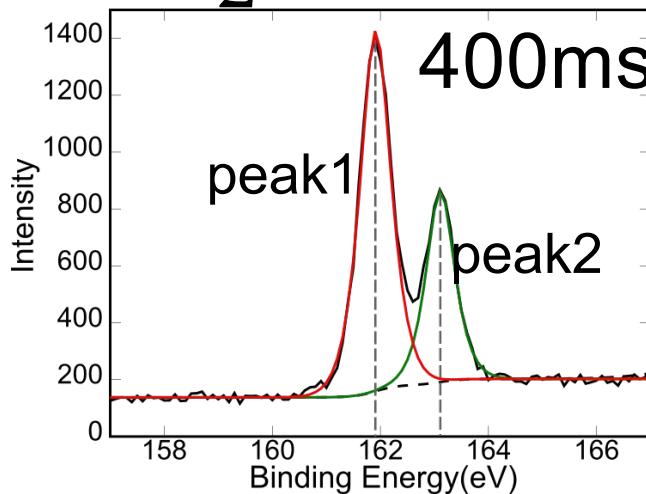
内容

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

例題2: XPS

どこまで時間窓を小さくできるか

MoS₂ 2p



ここまで

アンケート

- ・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。

ベイズ推論の拡張性

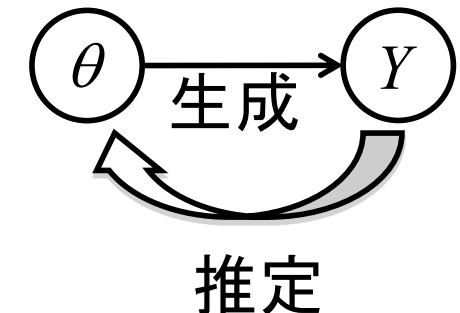
光電子の量子性を考慮する(ポアソン分布)

■ 事後確率: $p(\theta | Y) = \frac{p(Y | \theta)p(\theta)}{p(Y)}$ $Y = \{y_i\}_{i=1}^n$

$$p(\theta | Y) = \frac{1}{p(Y)} \exp\left(-\frac{n}{\sigma^2} E(\theta)\right) p(\theta)$$

これまでの

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



θ : ピーク位置など

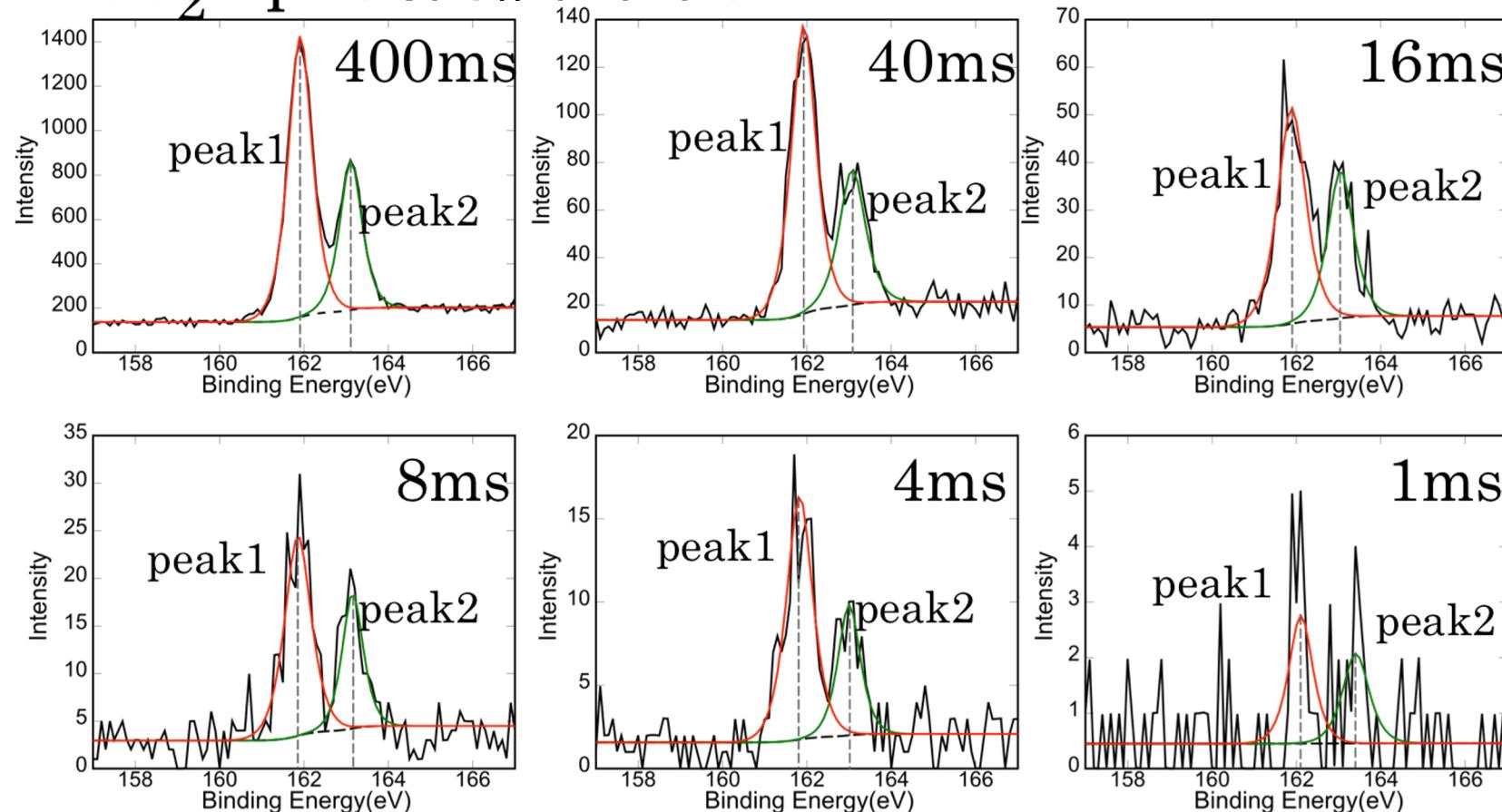
Y : 観測スペクトル

$$E(\theta) = \sum_{i=1}^n \left(y_i \log f(x_i; \theta) + f(x_i; \theta) + \sum_{j=1}^{y_i} \log(j) \right)$$

に変更するだけ

ベイズ計測 スペクトル分解のベイズ理論(Nagata *et al.* 2019)

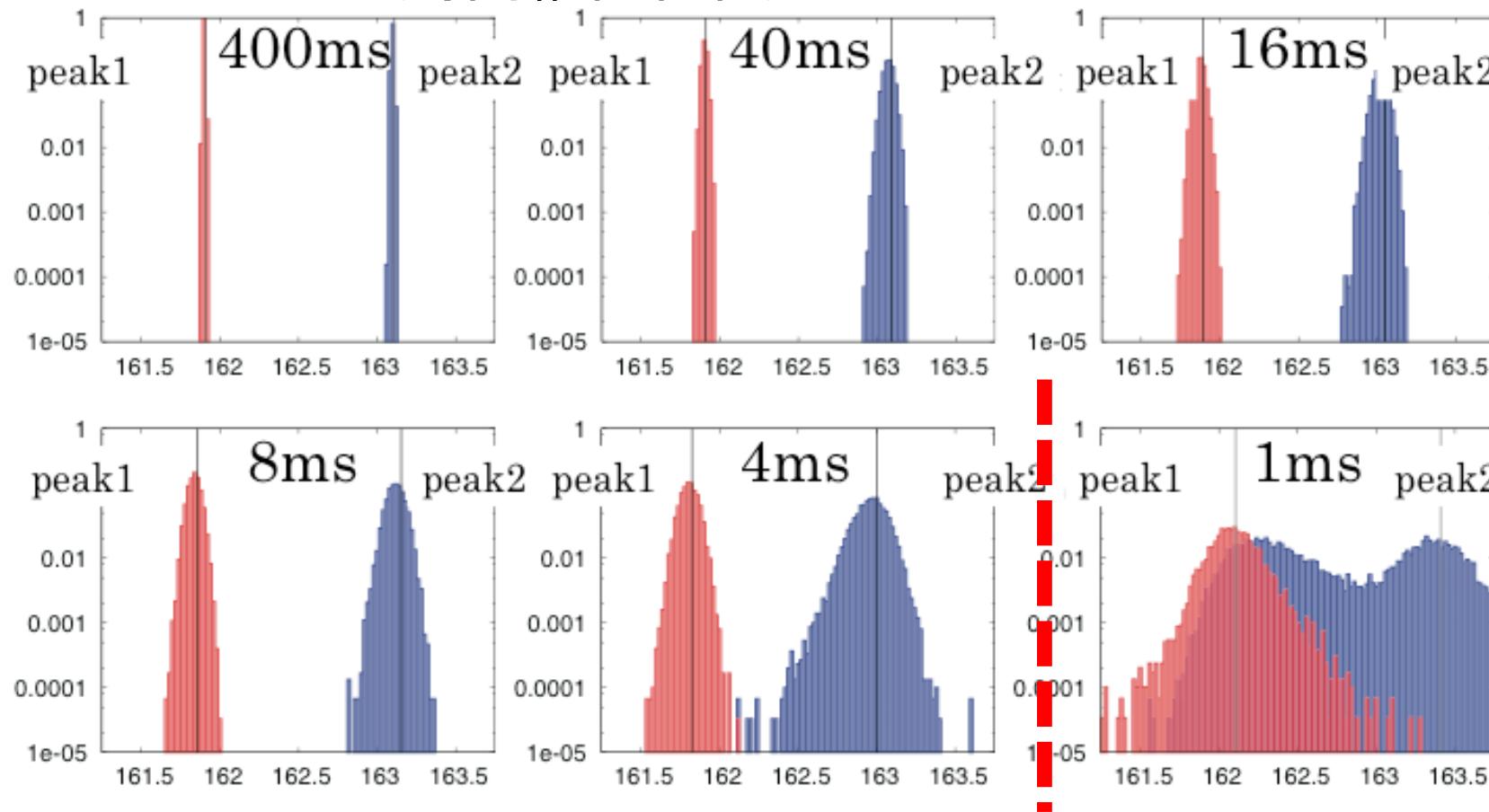
MoS₂ 2p (時間幅依存性)



ベイズ計測 スペクトル分解のベイズ理論(Nagata *et al.* 2019)

ベイズ計測: ベイズ推論によって、ピーク位置のベイズ事後確率を計算

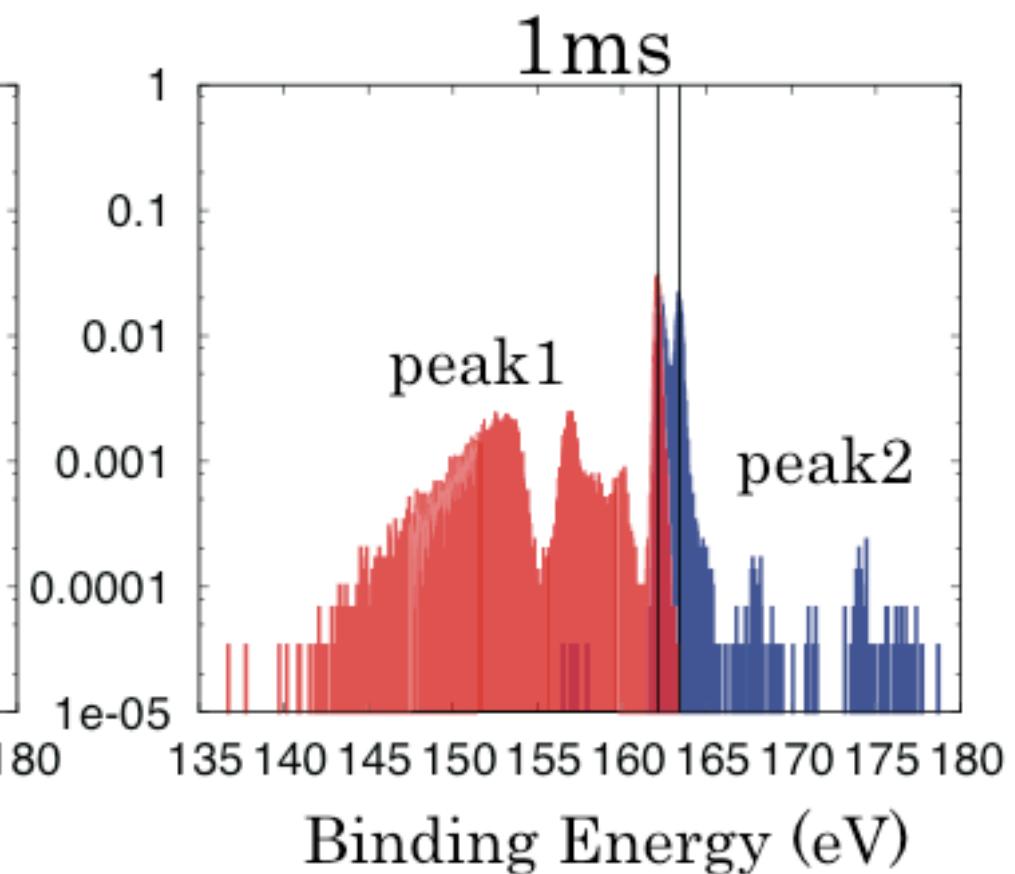
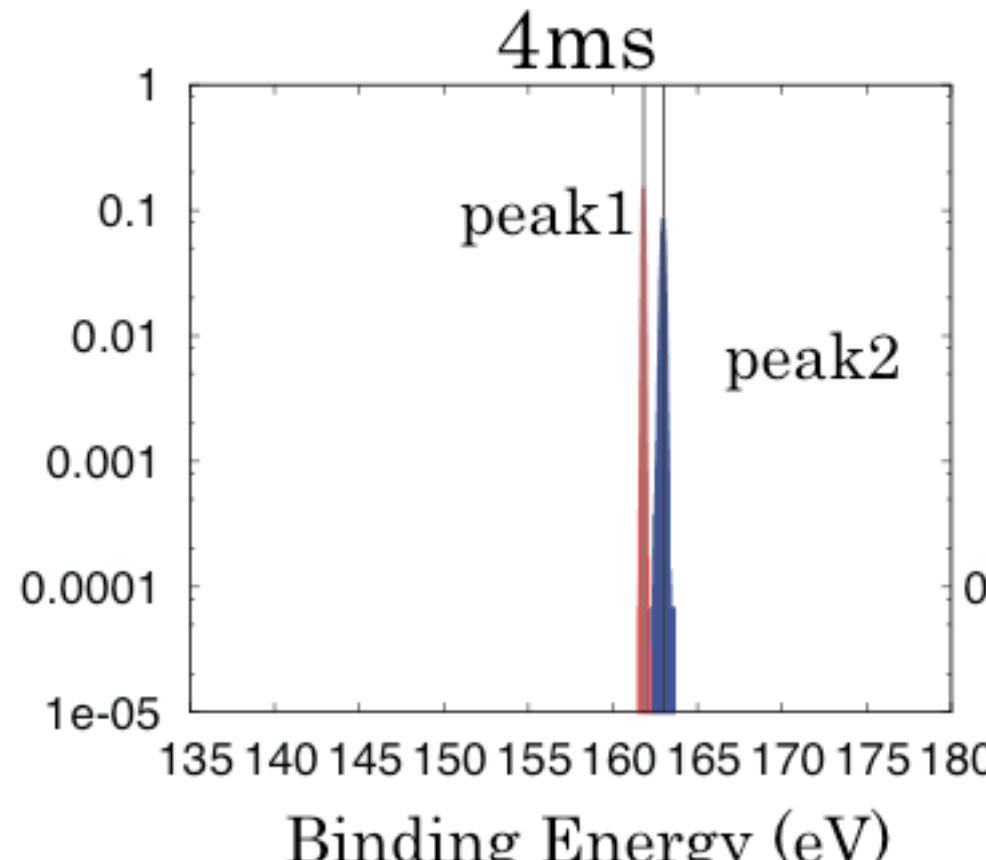
(時間幅依存性)



ベイズ計測 スペクトル分解のベイズ理論(Nagata *et al.* 2019)

ベイズ計測: ベイズ推論によって、ピーク位置のベイズ事後確率を計算

(時間幅依存性)

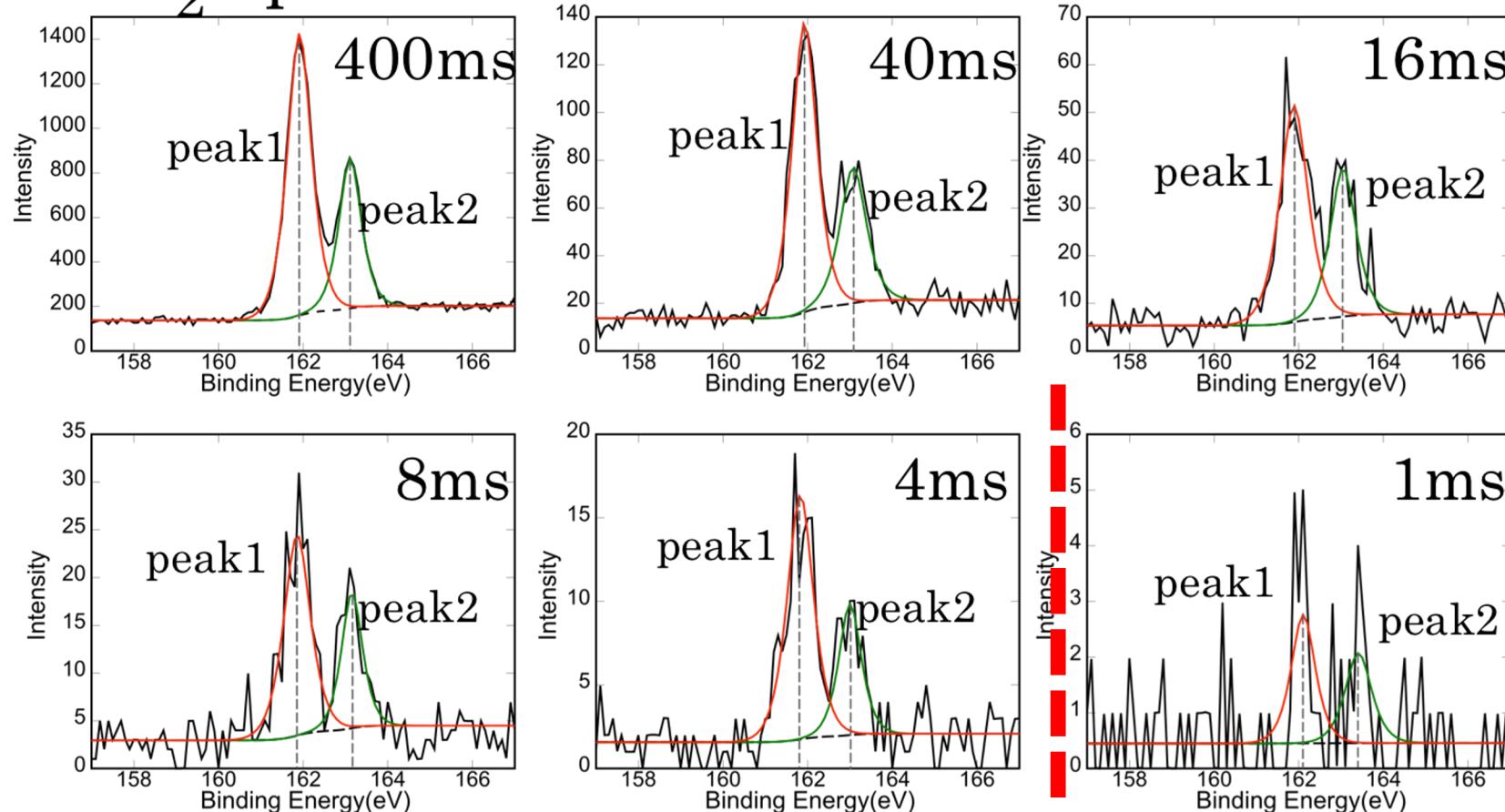


ベイズ計測 スペクトル分解のベイズ理論(Nagata *et al.* 2019)

ベイズ計測: ベイズ推論によって、ピーク位置のベイズ事後確率を計算

戦略目標: 計測限界を定量的に評価できる枠組みの提案

MoS₂ 2p (時間幅依存性)



アンケート

- ・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。

アンケート

- ・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。

内容

- アンケート
- 自己紹介
- スペクトル分解
 - 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
- SPring-8全ビームラインベイズ化計画

SPring-8

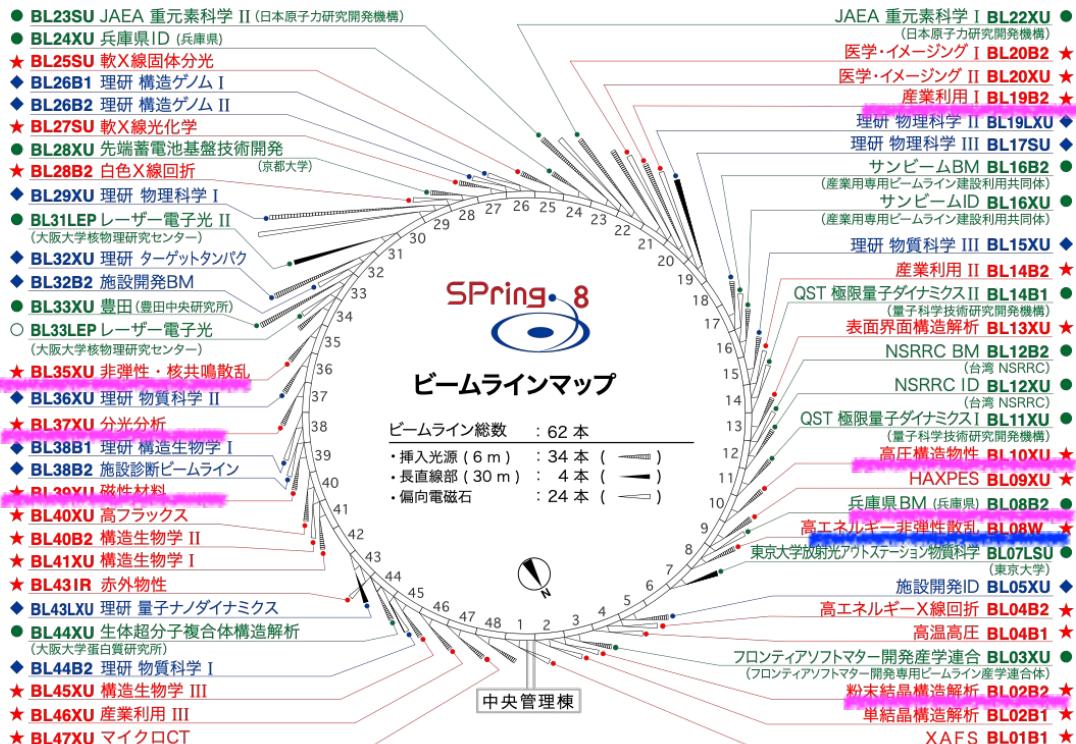
- ・ アメリカのAdvanced Photon Source (APS),ヨーロッパのEuropean Synchrotron Radiation Facility(ESRF) と合わせて, **世界3大放射光施設**.
- ・ 理研はSPring-8を「データ創出基盤」であると言っている. **年間延べ1万人**が利用.
- ・ APSやESRFにおいてベイズ計測は導入されていない.
- ・ 放射光におけるベイズ計測に関しては**日本が最先端**である.



水牧 仁一朗
熊本大学大学院先端科学研究所
JASRI客員主席研究員

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

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

15