

第16回 日本放射光学会 若手研究会
放射光サマースクール2024
—放射光科学×キャリアパス—
S1-2「ベイズ計測と放射光科学」

東京大学・大学院新領域創成科学研究科
複雑理工学専攻

岡田真人

本講演のスライドは岡田研HPにて公開

[https://mns.k.u-
tokyo.ac.jp/lab.html#overview](https://mns.k.u-tokyo.ac.jp/lab.html#overview)

自己紹介

- 大阪市立大学理学部物理学科 (1981 - 1985)
 - アモルファスシリコンの成長と構造解析
- 大阪大学大学院理学研究科(金森研) (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(福島研) (1989 - 1996)
 - 畳み込み深層ニューラルネット
 - 情報統計力学(ベイズ推論と統計力学の数理的等価性)
- JST ERATO 川人学習動態脳プロジェクト (1996 - 2001)
 - 計算論的神経科学
- 理化学研究所 脳科学総合研究センター 甘利T(2001 - 04/06)
 - ベイズ推論, 機械学習, データ駆動型科学
- 東京大学・大学院新領域創成科学研究科 複雑理工学専攻
 - 情報統計力学、データ駆動科学 (2004/07 -)

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

SPring-8全ビームラインベイズ化計画

敬称略

情報と放射光研究者のマッチング

メスbauer

BL35XU 岡田研学生+筒井

小角散乱

BL08B2 岡田研学生+桑本

BL19B2

XAS測定

BL37XU 岡田研学生+水牧

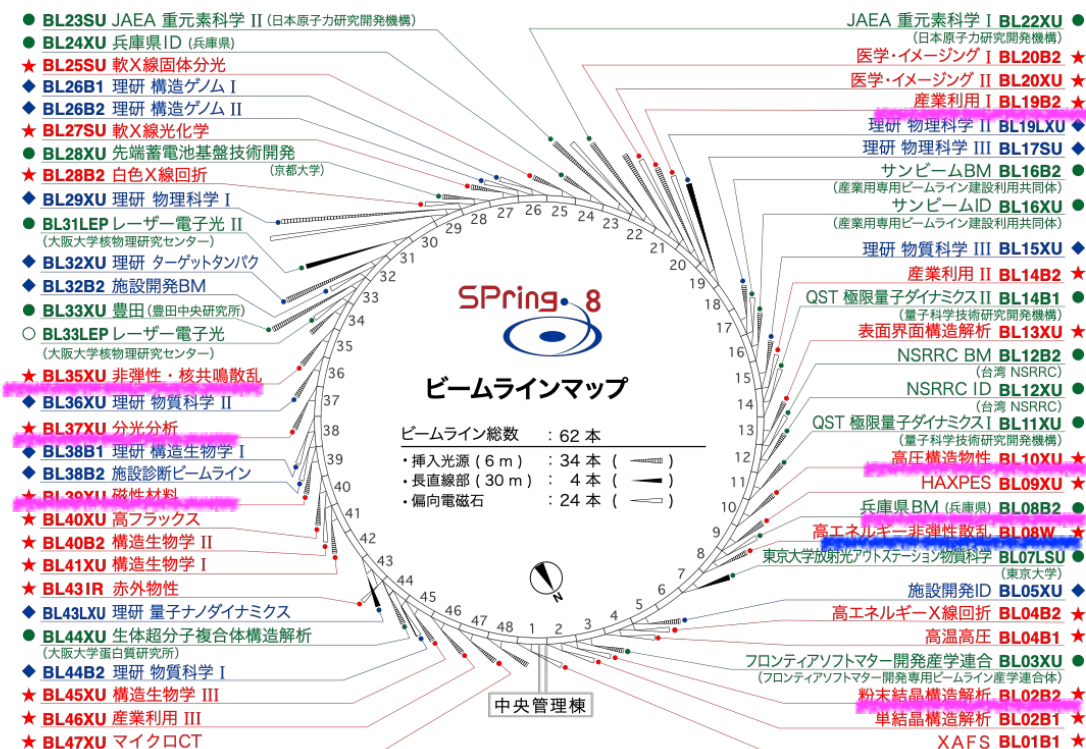
BL39XU

放射光ユーザーへの展開

時分割XRD

BL02B2 横山優一+河口彰吾、沙織

BL10XU ユーザー: 公立大、東工大



赤色BLが共用BL(JASRI担当): 計26本

全BL本数: 62本

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

SPring-8全ビームライン ベイズ化計画

- ベイズ計測を、SPring-8に導入するメリットはあるのだろうか？
- ベイズ計測は、日々行われているデータ解析のうち、ヒトにより恣意的に行われてきた、モデル選択とデータ統合を、パラメーターフィットと同じ枠組みで取り扱える
- ベイズ系をSPring-8に導入することで、これまでのデータ解析とは質的に異なるメリットを得られる

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

データ駆動科学の三つのレベル (2016)

計算理論(対象の科学, 計測科学)

データ解析の目的とその適切性を議論し, 実行可能な方法の論理(方略)を構築

モデリング(統計学, 理論物理学, 数理科学)

計算理論のレベルの目的, 適切さ, 方略を元に, 系をモデル化し, 計算理論を数学的に表現する

表現・アルゴリズム(統計学, 機械学習, 計算科学)

モデリングの結果得られた計算問題を, 実行するためのアルゴリズムを議論する.

Igarashi, Nagata, Kuwatani, Omori, Nakanishi-Ohno and M. Okada, “Three Levels of Data-Driven Science”, *Journal of Physics: Conference Series*, 699, 012001, 2016.

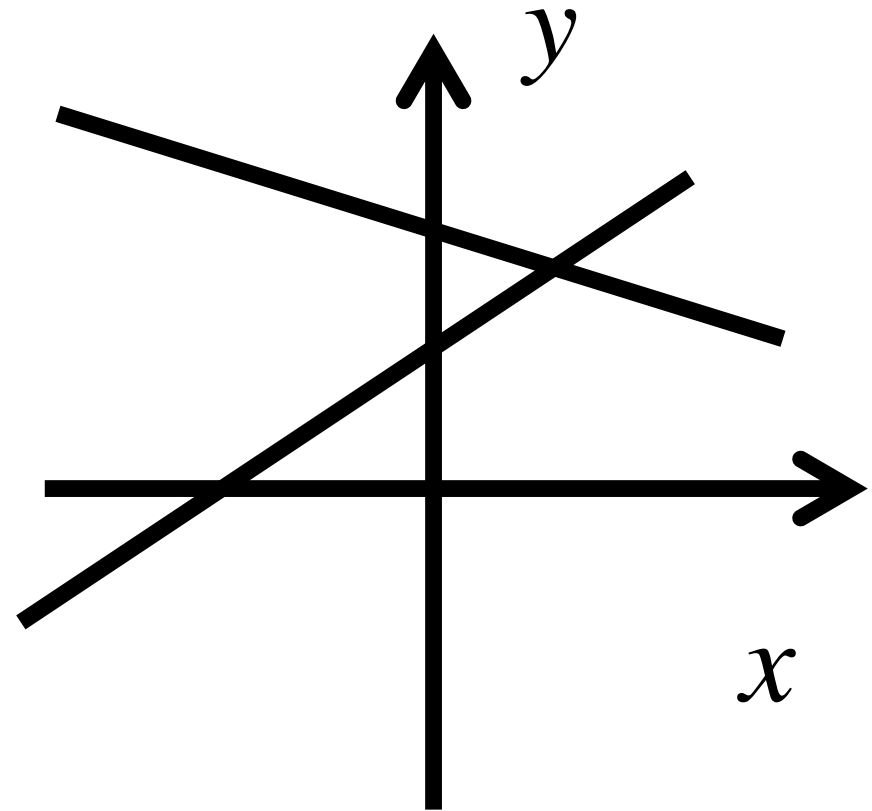
連立方程式とデータ駆動科学

連立方程式とその応用

鶴亀算 食塩水 寝坊
して追いかける問題

連立方程式への変換

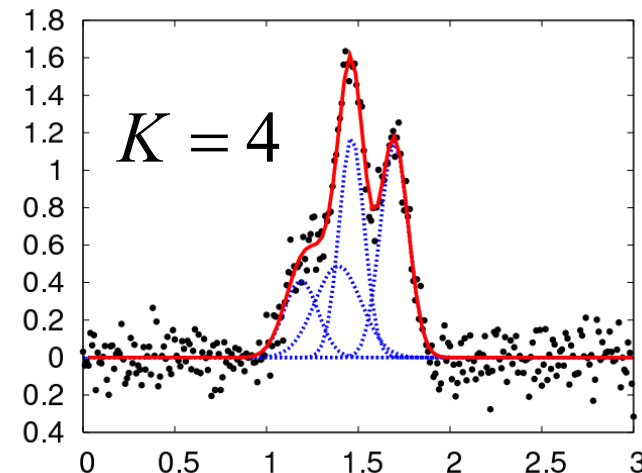
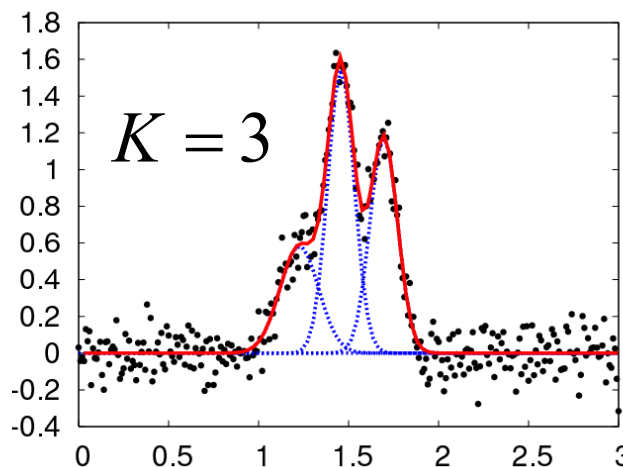
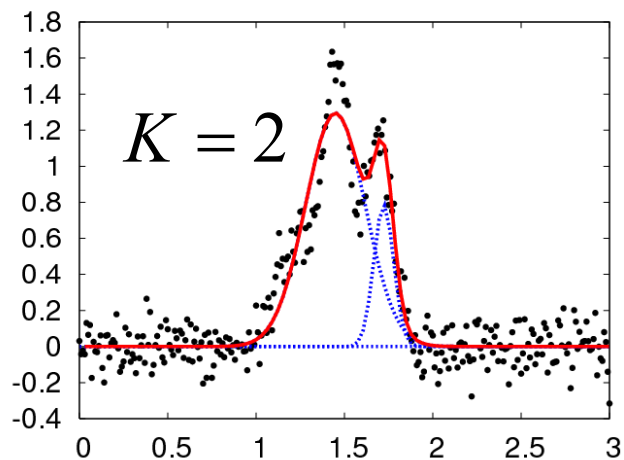
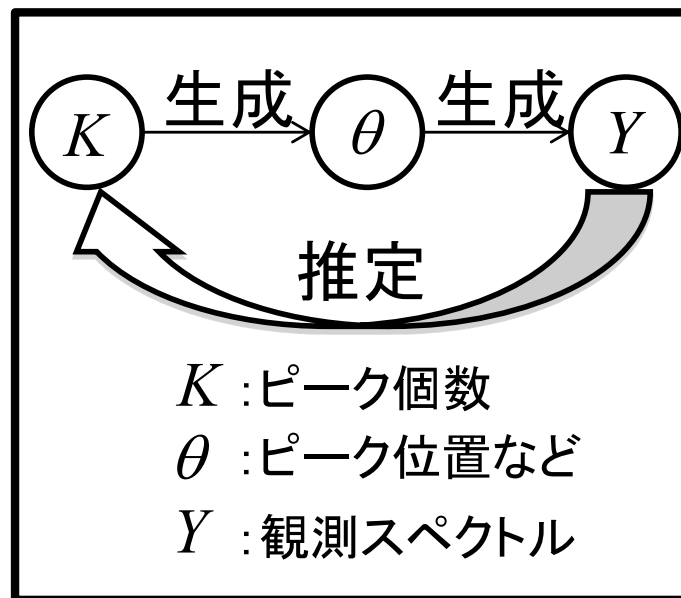
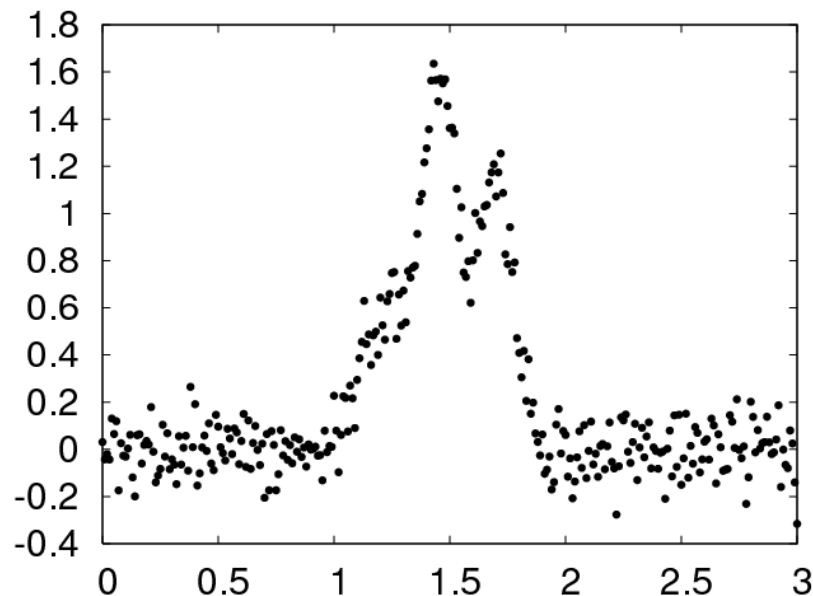
加減法, 代入法



一つの方程式が一本の線
二本の線の交点が解になる

(五十嵐, 竹中, 永田, 岡田, *応用統計学*, 2016)

ベイズ的スペクトル分解: K をどう選ぶか



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

スペクトル分解の三つのレベル (1/2)

スペクトル分解の計算理論

データ解析の目的: 多峰性スペクトルから背後にある離散電子のエネルギー準位を推定する

データ解析の適切さ: 多峰性スペクトルを単峰性関数の線形和で表し、その単峰性関数の個数を推定する。

誤差関数の最小化では、単峰性関数が多い方が誤差関数は小さい。そこで統計学の交差検証誤差やベイズ的モデル選択で単峰性関数の数 K を決める。

スペクトル分解のモデリング

多峰性スペクトルを単峰性関数の線形和に観測ノイズが付加されて生成するとモデリングする

スペクトル分解の三つのレベル (2/2)

スペクトル分解の表現・アルゴリズム

多峰性スペクトルを単峰性関数の線形和に観測ノイズが付加されて生成するとモデリングし、ベイズ推論を適用することで、 K 個の単峰性関数の大きさ、位置、幅の事後確率を求める。各 K に対して、ベイズ的自由エネルギーを求め、ベイズ的自由エネルギーを最小にする K を求める。その K 個の単峰性関数の位置を、電子のエネルギー準位とする。

Igarashi, Nagata, Kuwatani, Omori, Nakanishi-Ohno and M. Okada, “Three Levels of Data-Driven Science”, *Journal of Physics: Conference Series*, 699, 012001, 2016.

Nagata, Sugita and M. Okada, “Bayesian spectral deconvolution with the exchange Monte Carlo method”, *Neural Networks*, 28, 82-89 2012.

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- **ベイズ計測**
 - **計測科学の必須条件**
 - **ベイズ計測三種の神器**
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

計測科学の必須条件

- 数理モデルのフリーパラメータを決める
系統的枠組み
- 複数モデルをデータだけから選択する
系統的枠組み
- 同一物質に対する複数の計測データを
統合する系統的枠組み

ベイズ計測

- ベイズ推論のうち計測科学に重要な三つの要素からなる情報数理科学的体系で、その三要素は**ベイズ計測三種**の神器と呼ばれる
 1. 物理パラメータの確率分布推定
 2. 同一データを説明する複数モデルをデータのみから選べるベイズ的モデル選択
 3. 同一物質に対する複数の実験データを系統的に統合するベイズ統合
- 従来の最小二乗法によるパラメータフィットでは、1.の物理パラメータの点推定しか行えない
- **パラメータフィットを超えて**: ベイズ計測では、取り扱えることが質的に異なる

ベイズ計測の習得法

1. $y=ax+b$ への解析的計算の適用
2. $y=ax+b$ への数値計算の適用
3. 1.と2.の結果が一致するのを確認
4. ベイズ的スペクトル分解
5. 各課題に取り組む

内容

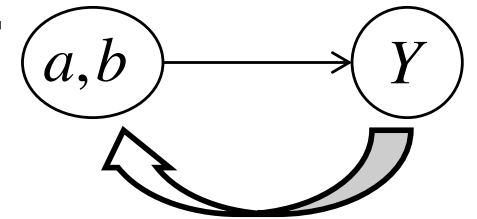
- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- **ベイズ計測**
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

ベイズ計測とは?

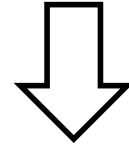
ベイズ推論

$$p(Y, a, b) = p(Y | a, b) p(a, b) = p(a, b | Y) p(Y)$$

生成(因果律)



<ベイズの定理>



$$p(a, b | Y) = \frac{p(Y | a, b) p(a, b)}{p(Y)} \propto \exp(-nE(a, b)) p(a, b)$$

$p(a, b | Y)$: 事後確率。データが与えられたもとでの
物理パラメータの確率。

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

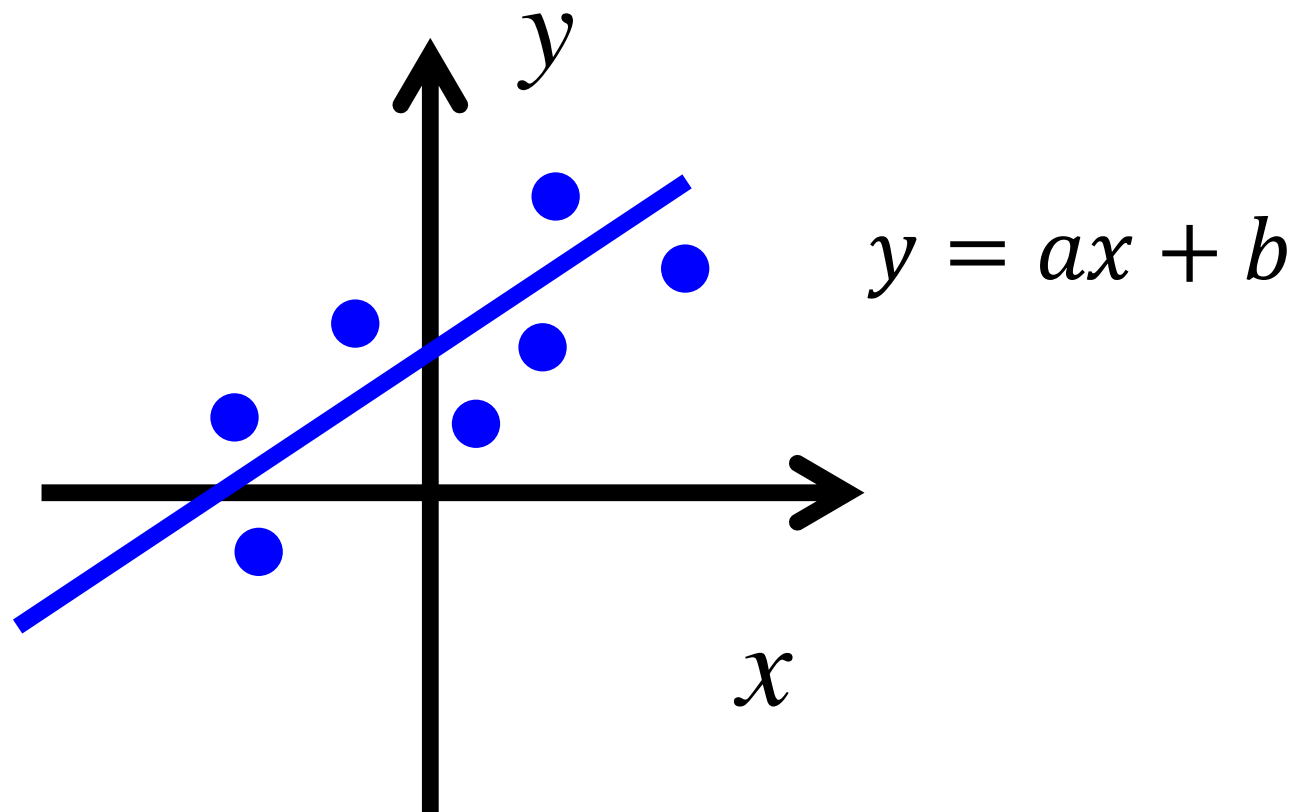
ベイズ計測三種の神器

1. 物理パラメータの事後確率分布定
2. モデル選択
3. データ統合

ベイズ計測の利点

$y=ax+b$ の取り扱いを通じて

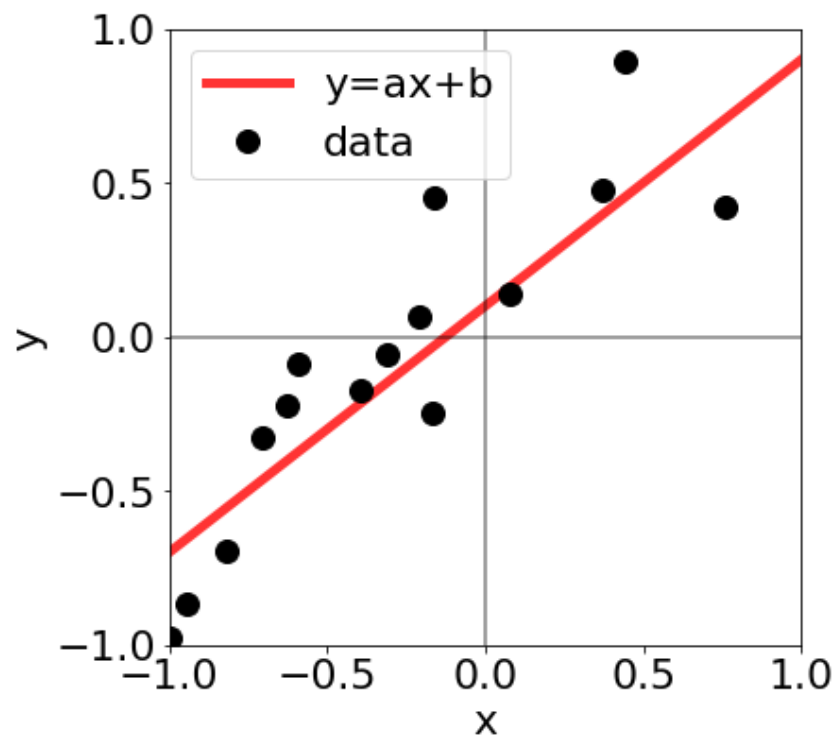
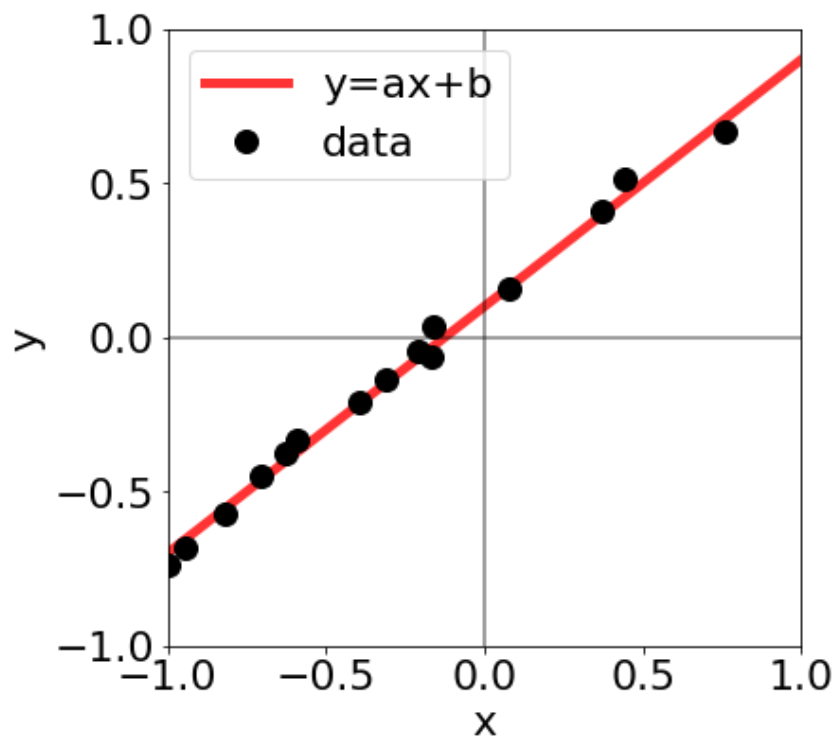
現状でも用いられている最も簡単な例



傾き a : 系の線形応答、バネ定数、電気伝導度、誘電率

ベイズ計測の利点

$y=ax+b$ の取り扱いを通じて



この二つの推定精度の違いを数学的に表現したい
準備として従来手法の最小二乗法

$y=ax+b$ の最小二乗法

$$E(a, b) = \frac{1}{N} \sum_{i=1}^N (y_i - (ax_i + b))^2$$

二乗誤差 $E(a, b)$ を最小にするようにパラメータをフィット(最小二乗法)

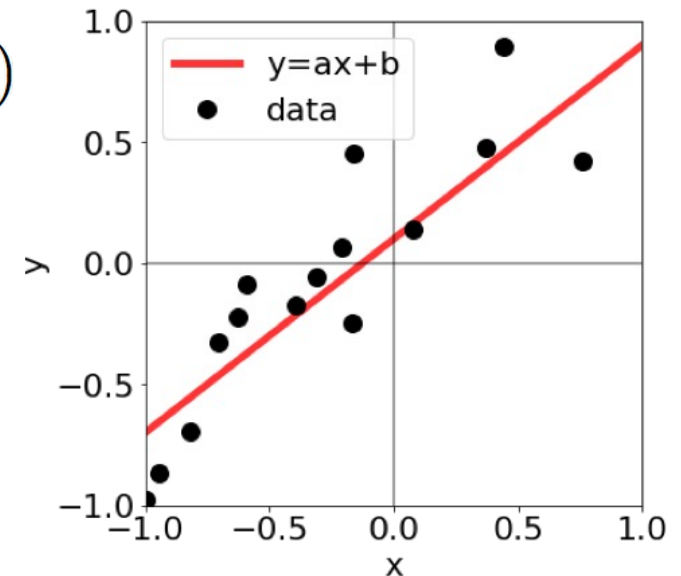
$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = 0 \text{ とする場合}$$

$$E(a, b) = \frac{1}{2} \left(\overline{x^2} \left(a - \frac{\overline{xy}}{\overline{x^2}} \right)^2 + (b - \bar{y})^2 - \frac{\overline{xy^2}}{\overline{x^2}} - \bar{y}^2 + \bar{y}^2 \right) \quad a_0 = \frac{\overline{xy}}{\overline{x^2}} \quad b_0 = \bar{y}$$

$$= \mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0, b_0) \geq E(a_0, b_0)$$

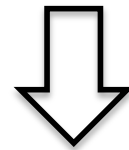
$$\text{平均: } \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i, \text{ 分散: } \overline{x^2} = \frac{1}{N} \sum_{i=1}^N x_i^2$$

$$\overline{xy} = \frac{1}{N} \sum_{i=1}^N x_i y_i$$



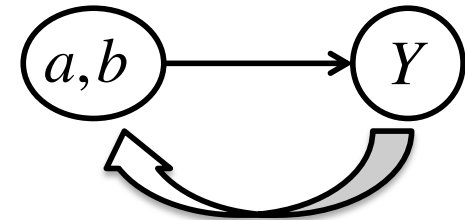
ベイズの定理による 神器1: パラメータの事後確率推定 (1/4)

$$p(Y, a, b) = p(Y | a, b) p(a, b) = p(a, b | Y) p(Y)$$



<ベイズの定理>

生成(因果律)



$$p(a, b | Y) = \frac{p(Y | a, b) p(a, b)}{p(Y)} \propto \exp(-nE(a, b)) p(a, b)$$

$p(a, b | Y)$: 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.

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

神器1: パラメータの事後確率推定 (2/4)

$$y_i = ax_i + b + n_i$$

$$p(n_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n_i^2}{2\sigma^2}\right)$$

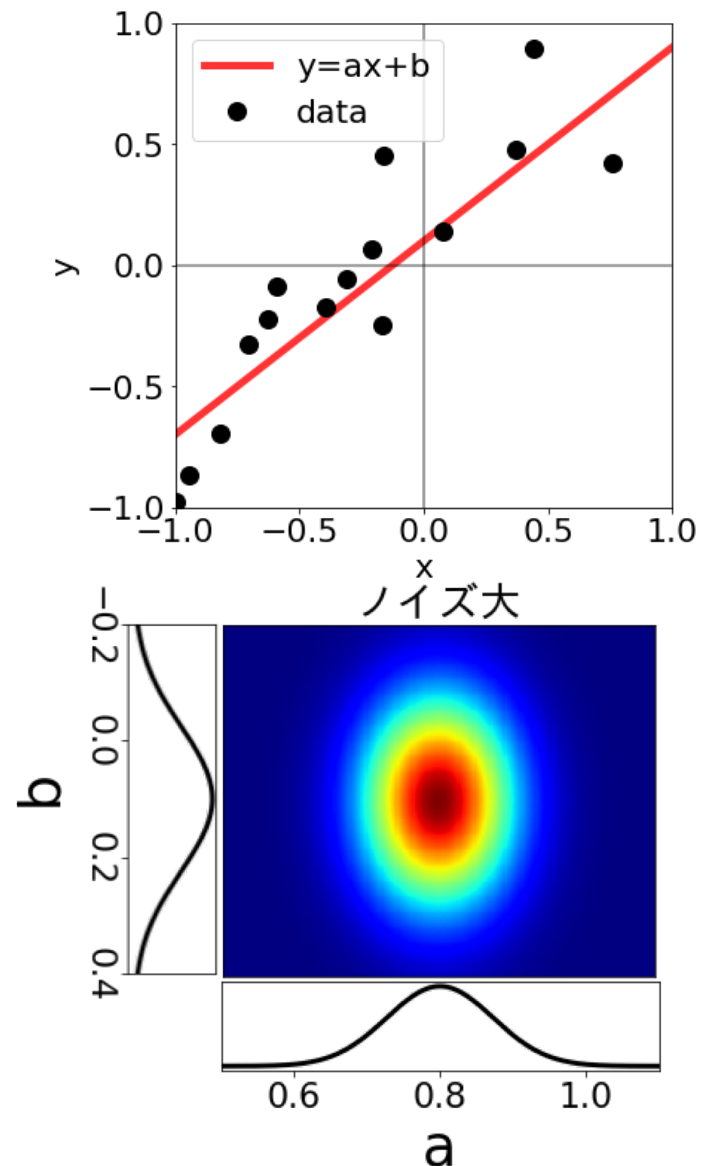
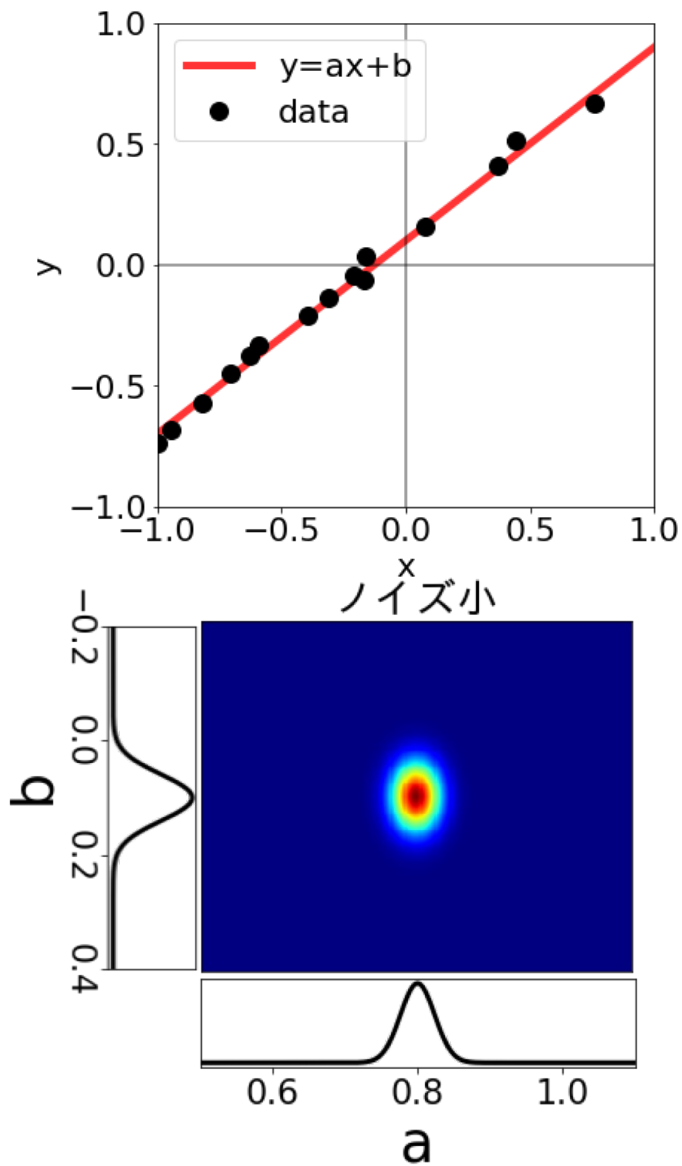
$$p(n_i) = p(y_i|a, b) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - (ax_i + b))^2}{2\sigma^2}\right)$$

$$\begin{aligned} p(Y|a, b) &= \prod_{i=1}^N p(y_i|a, b) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left(-\frac{\sum_{i=1}^N (y_i - (ax_i + b))^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left(-\frac{N}{\sigma^2} E(a, b)\right) \end{aligned}$$

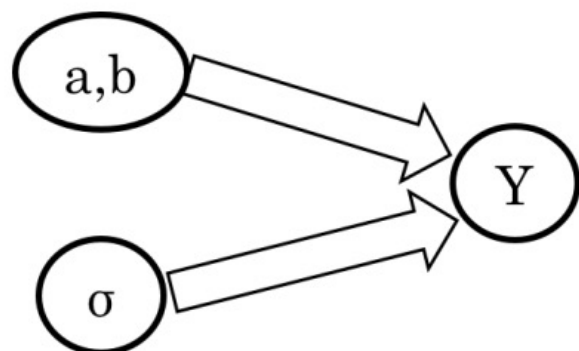
神器1: パラメータの事後確率推定 (3/4)

$$\begin{aligned} p(a, b|Y) &= \frac{p(Y|a, b)p(a, b)}{p(Y)} \propto p(Y|a, b) \\ &= \exp \left\{ -\frac{N}{\sigma^2} \left(\mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0, b_0) \right) \right\} \\ &\propto \exp \left\{ -\frac{N}{\sigma^2} \left(\mathcal{E}_a(a) + \mathcal{E}_b(b) \right) \right\} \\ &= \exp \left\{ -\frac{N\bar{x}^2}{2\sigma^2} (a - a_0)^2 + \frac{N}{2\sigma^2} (b - b_0)^2 \right\} \end{aligned}$$

神器1: パラメータの事後確率推定 (4/4)



神器1: パラメータの事後確率推定 ノイズ分散推定

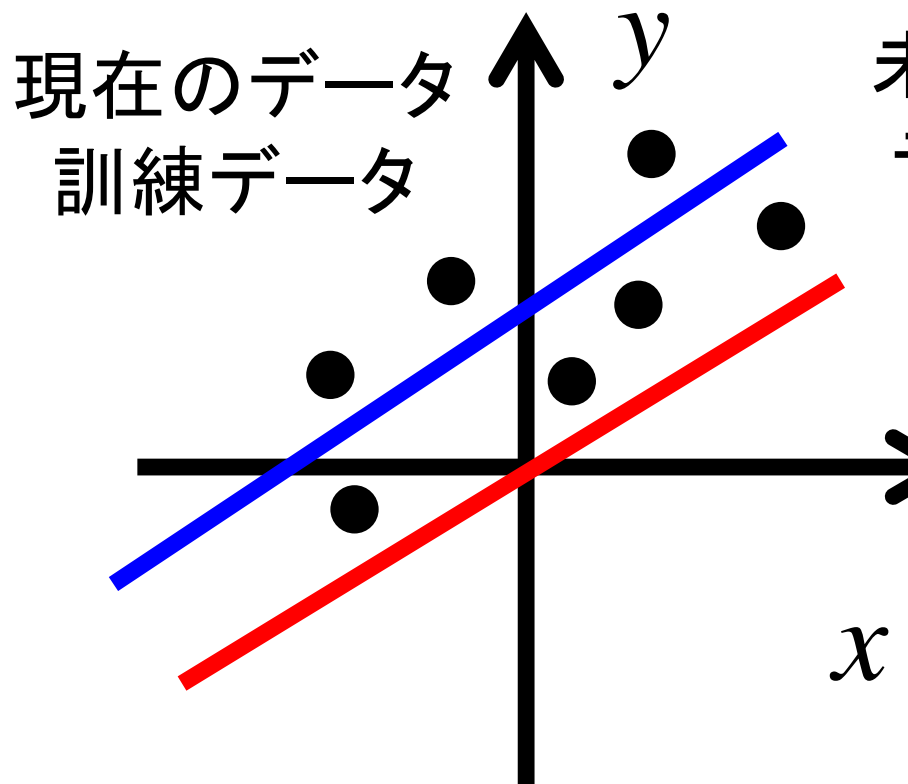


$$\begin{aligned} p(\sigma^2|Y) &\propto \int dadb p(Y|a, b, \sigma^2) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \int dadb \exp\left\{-\frac{N}{\sigma^2}E(a, b)\right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \left\{ \exp\left(-\frac{N}{\sigma^2}E(a_0, b_0)\right) + \int da \exp\left(-\frac{N\bar{x}^2}{2\sigma^2}(a - a_0)^2\right) + \int db \exp\left(-\frac{N}{2\sigma^2}(b - b_0)^2\right) \right\} \\ &= (2\pi\sigma^2)^{-\frac{N-2}{2}} (N^2\bar{x}^2)^{\frac{1}{2}} \exp\left(-\frac{N}{\sigma^2}E(a_0, b_0)\right) \end{aligned} \quad (25)$$

$$\sigma^2 = \frac{NE(a_0, b_0)}{N-2} = \frac{1}{N-2} \sum_{i=1}^N \{y_i - (a_0x_i + b_0)\}^2$$

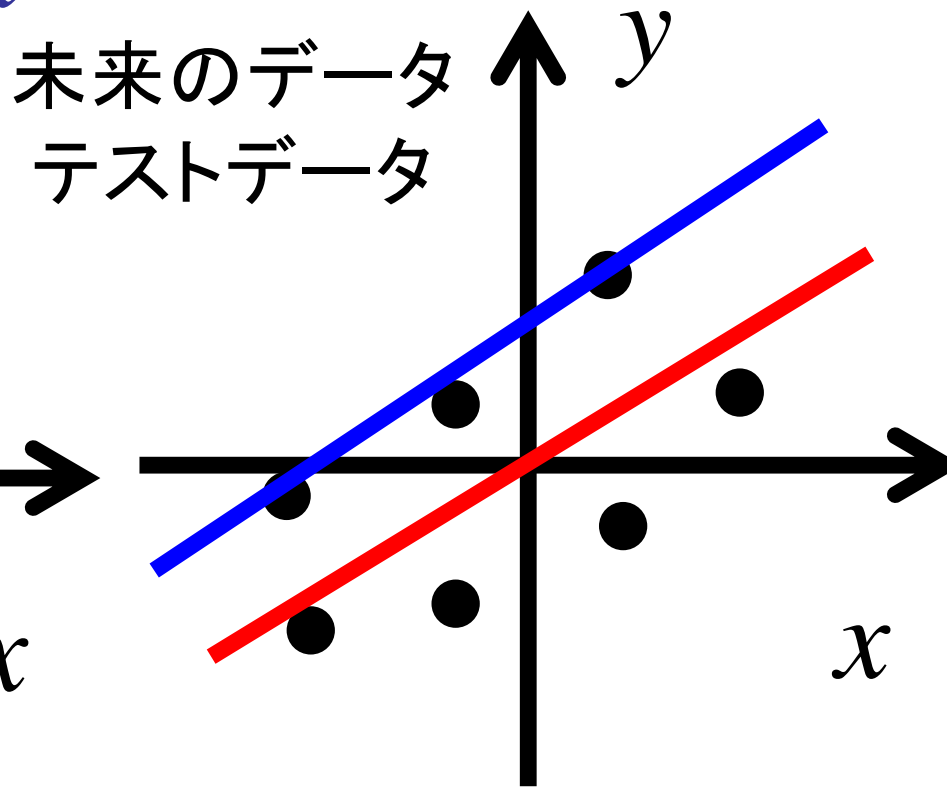
問題意識 神器2: ベイズ的モデル選択

$y=ax$ か $y=ax+b$ か?



$y=ax+b$: 訓練誤差小
訓練誤差

ノイズに過学習



$y=ax$: 訓練誤差小
汎化誤差

モデル選択できる理由: 汎化誤差は観測ノイズに依存する

神器2: ベイズ的モデル選択

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

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

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

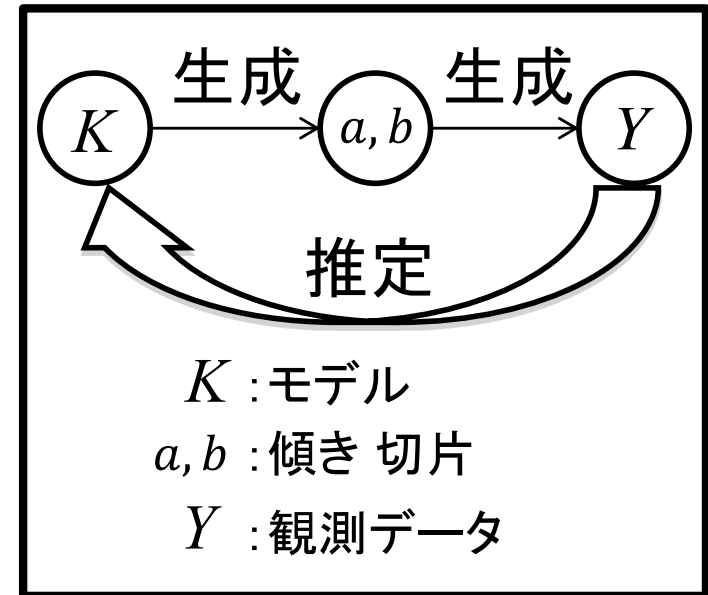
4. **無駄な自由度の系統的消去**: 周辺化, 分配関数

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

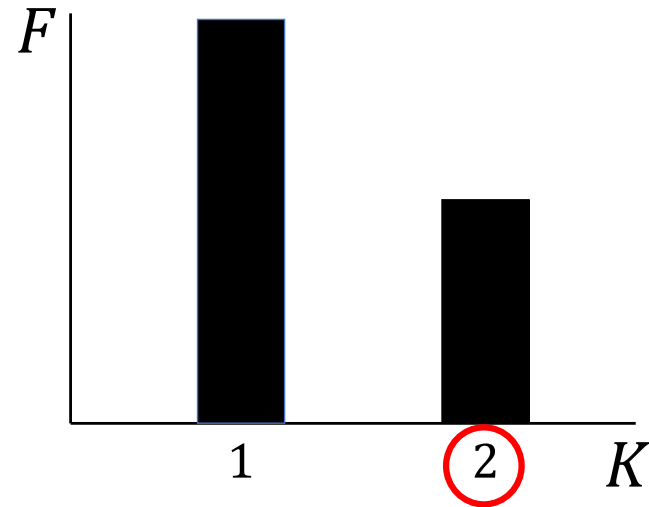
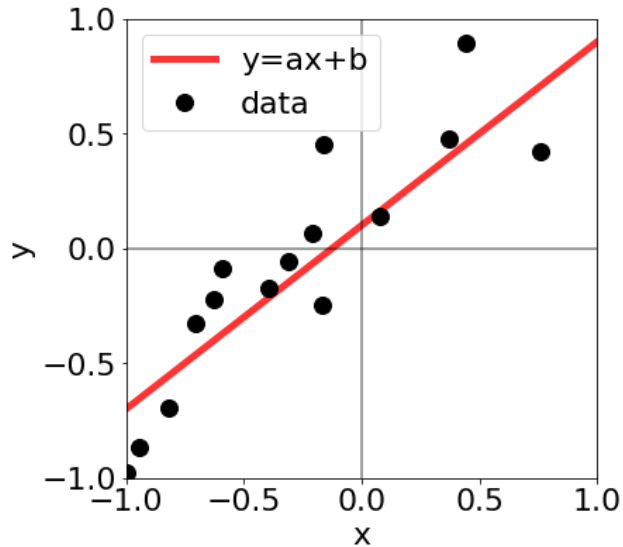
$$p(K | Y) = \frac{p(Y | K) p(K)}{p(Y)} \propto p(K) \int \exp(-nE(\theta)) p(\theta) d\theta$$

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

自由エネルギーを最小にするモデル K を求める.



モデル選択: 自由エネルギー差



- $K = 1 : y = ax$
- $K = 2 : y = ax + b$

$$F(K=1) = N \left\{ \frac{1}{\sigma^2} E(a_0) + \frac{\log N}{2N} \right\}$$

$$F(K=2) = N \left\{ \frac{1}{\sigma^2} E(a_0, b_0) + \frac{\log N}{N} \right\}$$

データのみからモデルを選択できる

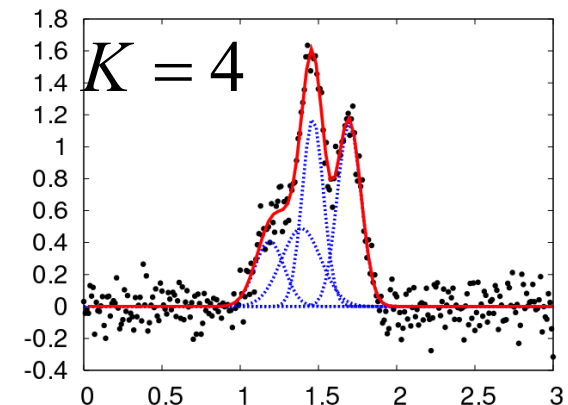
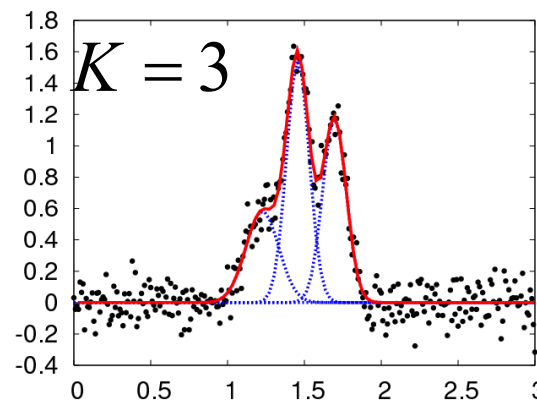
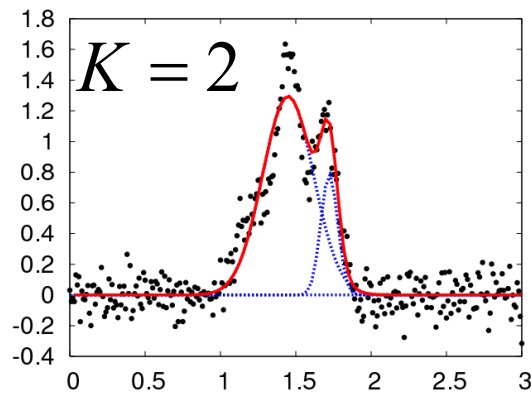
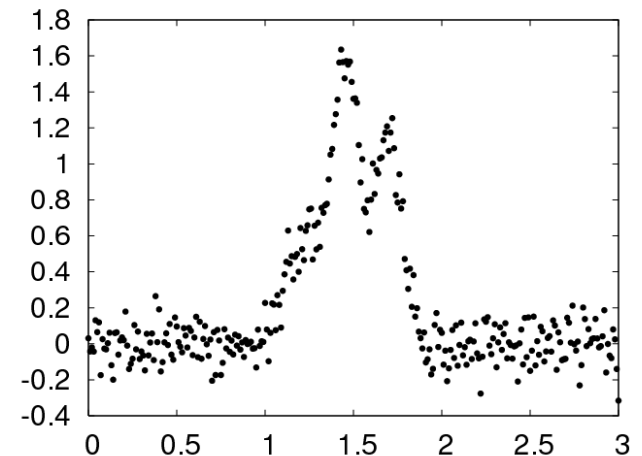
まとめ: ベイズ計測三種の神器 $y=ax+b$ の解析取り扱いを通じて

- 従来の最小二乗法
 - 1. 物理パラメータの点推定
- ベイズ計測
 1. 物理パラメータの確率分布推定
 2. データからのベイズ的モデル選択
 3. ベイズ統合: 今回は説明を省略
 - 水牧先生の基調講演

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

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

スペクトル分解の定式化

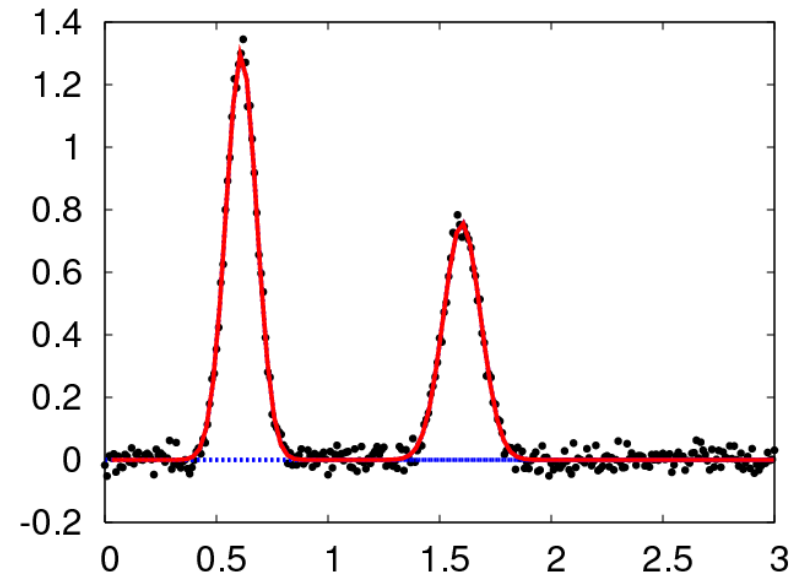
ガウス関数(基底関数)の足し合わせにより, スペクトルデータを近似

観測データ: $D = \{x_i, y_i\}_{i=1}^n$

x_i : 入力 y_i : 出力

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

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

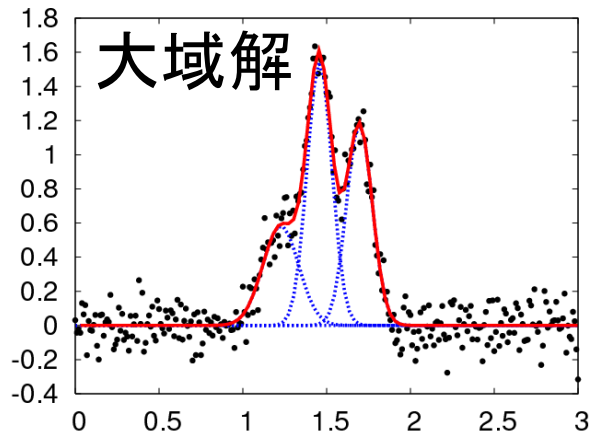


二乗誤差を最小にするようにパラメータをフィット(最小二乗法)

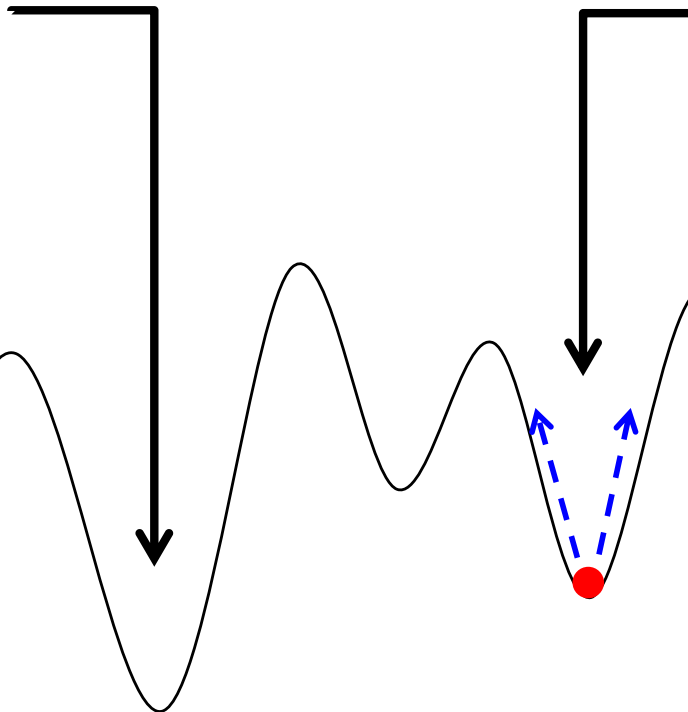
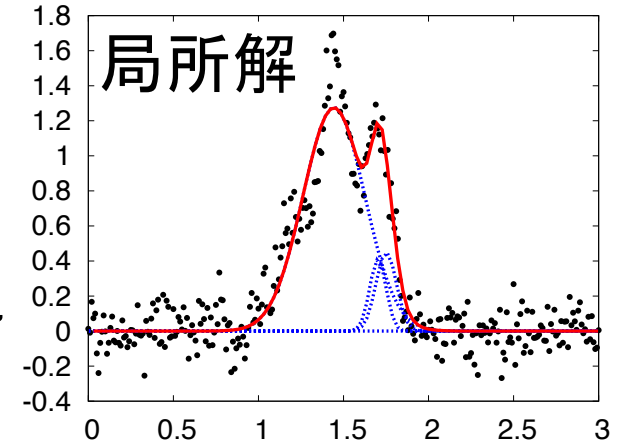
$$E(\theta) = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i; \theta))^2$$

スペクトル分解従来法: 最急降下法

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

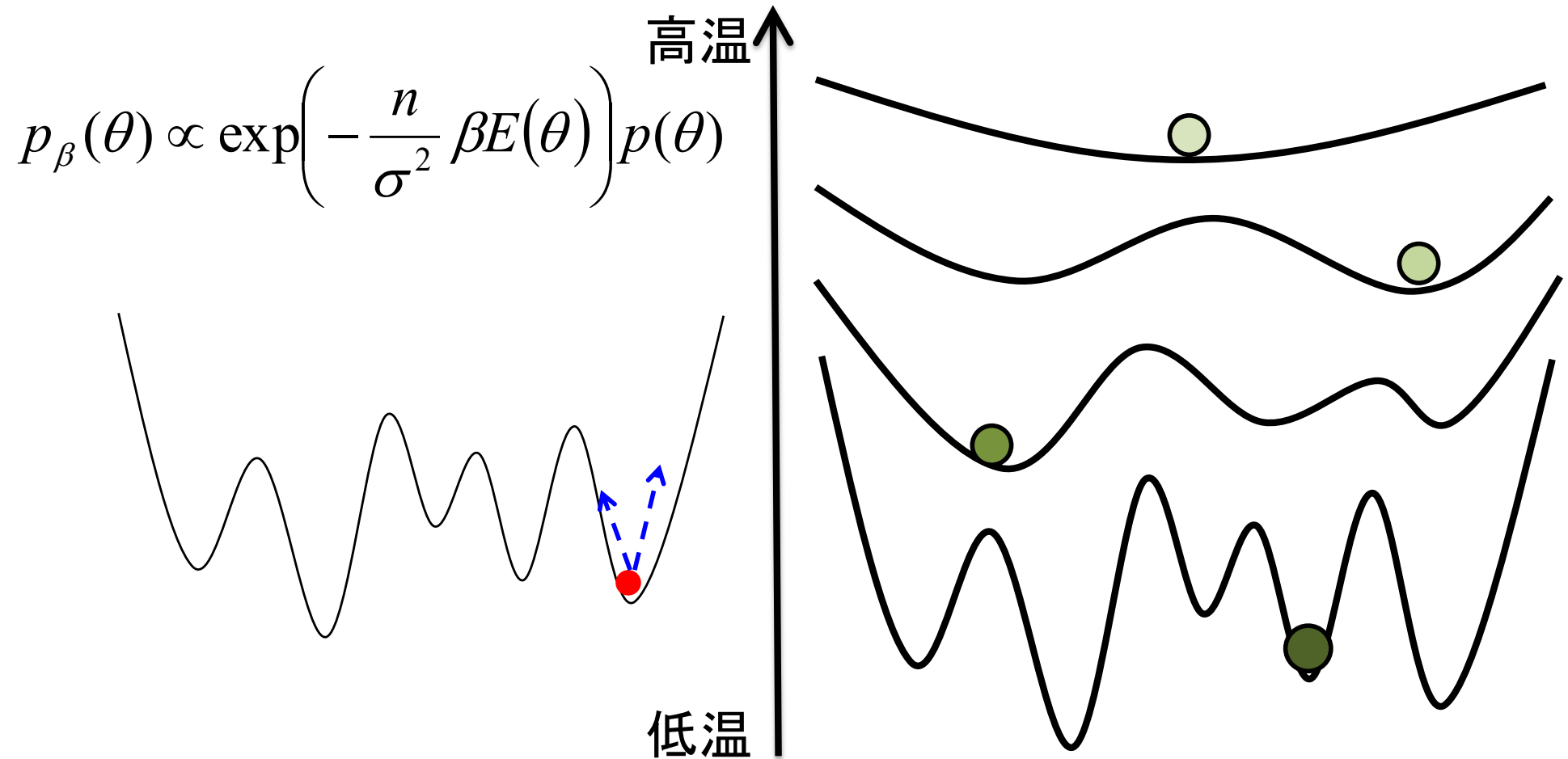


モンテカルロ法の適用

レプリカ交換モンテカルロ法

メトロポリス法

レプリカ交換モンテカルロ法



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

モデル選択: 自由エネルギーの導入

1. 欲しいのは $p(K|Y)$

2. θ がないぞ

3. $p(K, \theta, Y)$ の存在を仮定

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

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

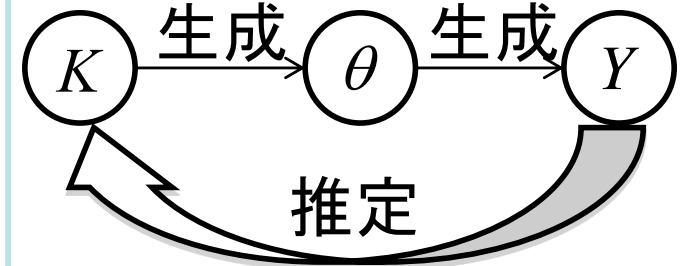
4. **無駄な自由度の系統的消去**: 周辺化, 分配関数

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

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

$$F(K) = -\log \int \exp(-nE(\theta)) p(\theta) d\theta = \boxed{E - TS}$$

自由エネルギーを最小にする個数 K を求める.

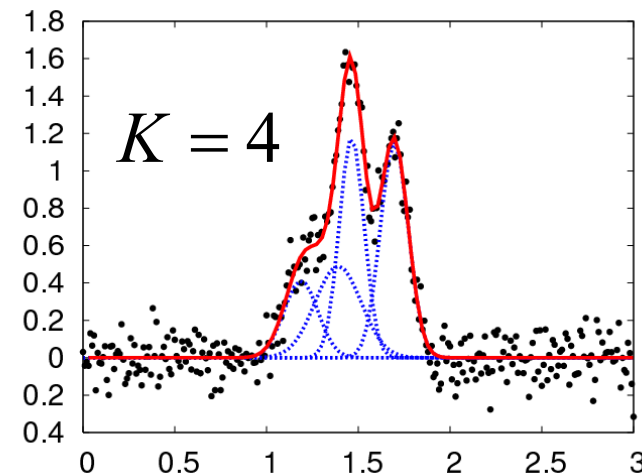
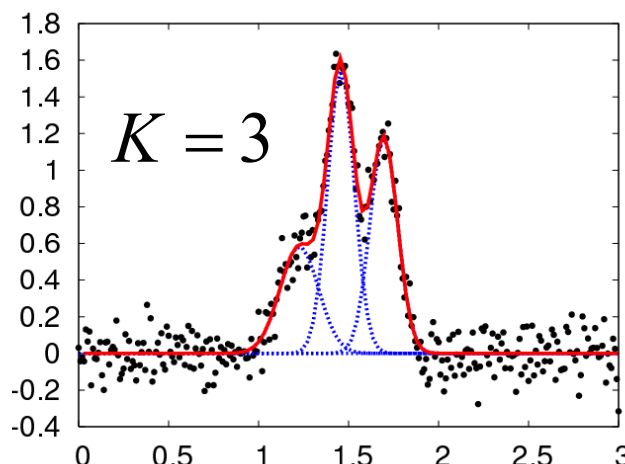
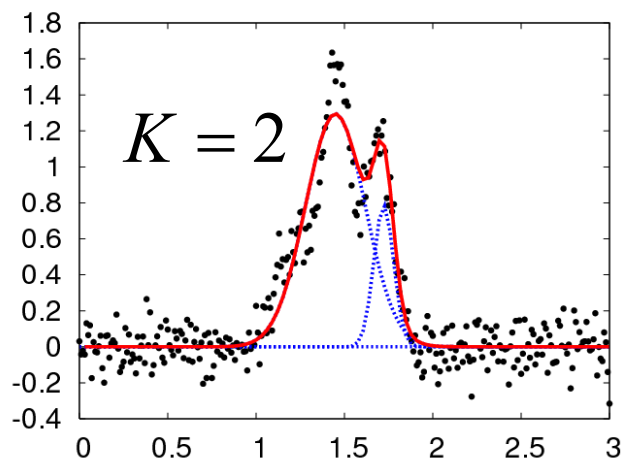
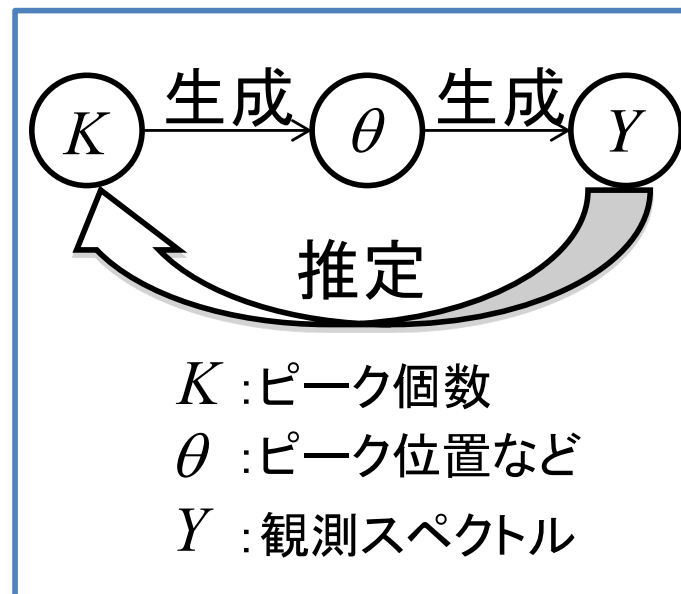
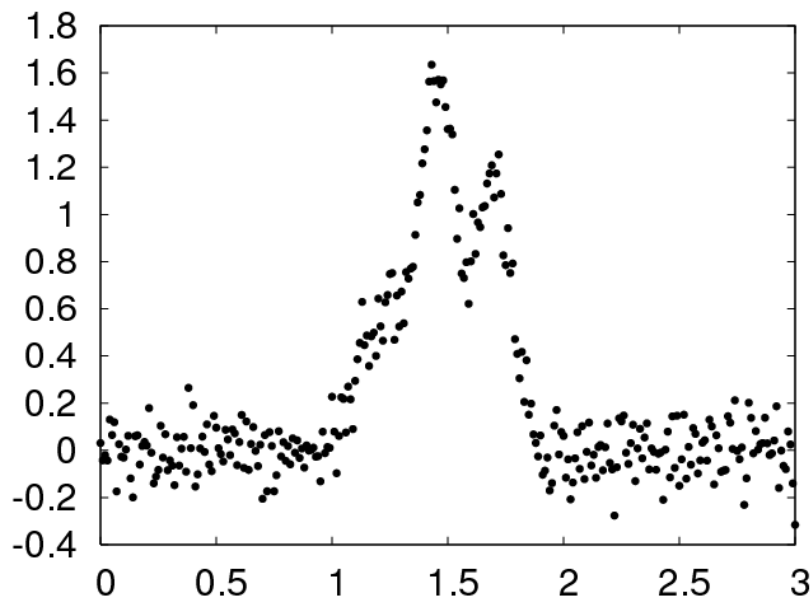


K : ピーク個数

θ : ピーク位置など

Y : 観測スペクトル

モデル選択: K をどう選ぶか

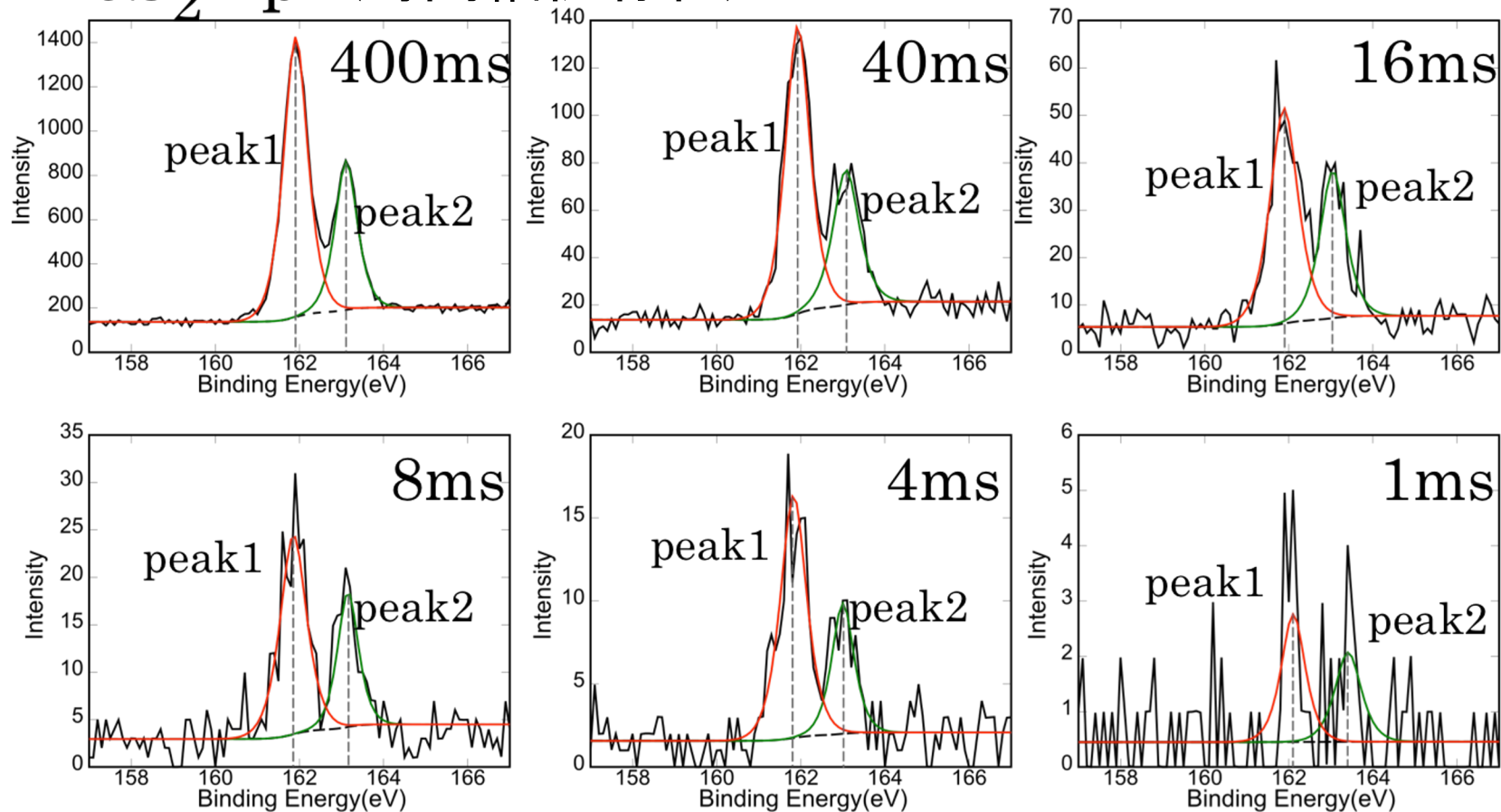


Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

計測限界の理論的取り扱い (4/9)

(Nagata *et al.* 2019)

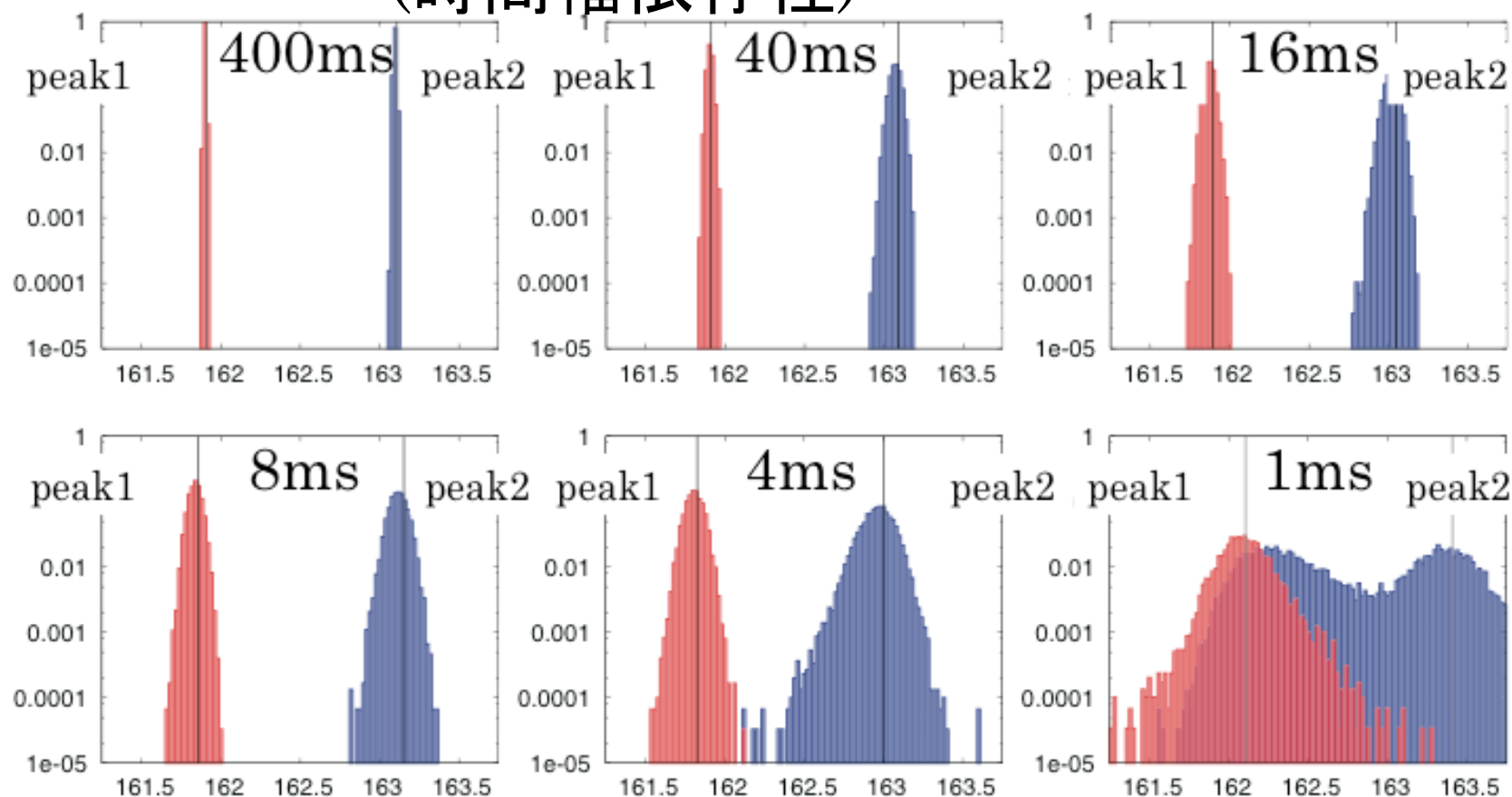
MoS₂ 2p (時間幅依存性)



計測限界の理論的取り扱い (6/9)

(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)

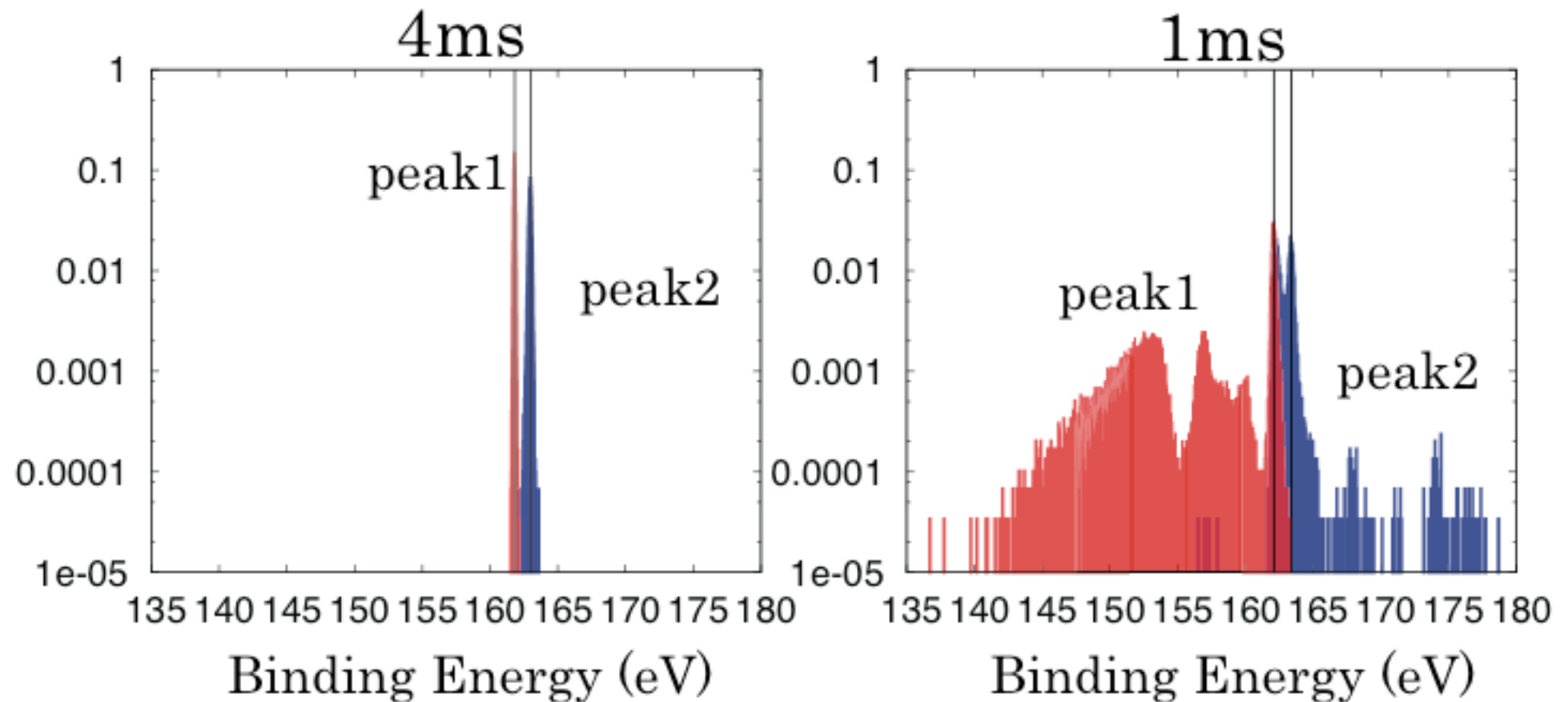


計測限界の理論的取り扱い (7/9)

(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

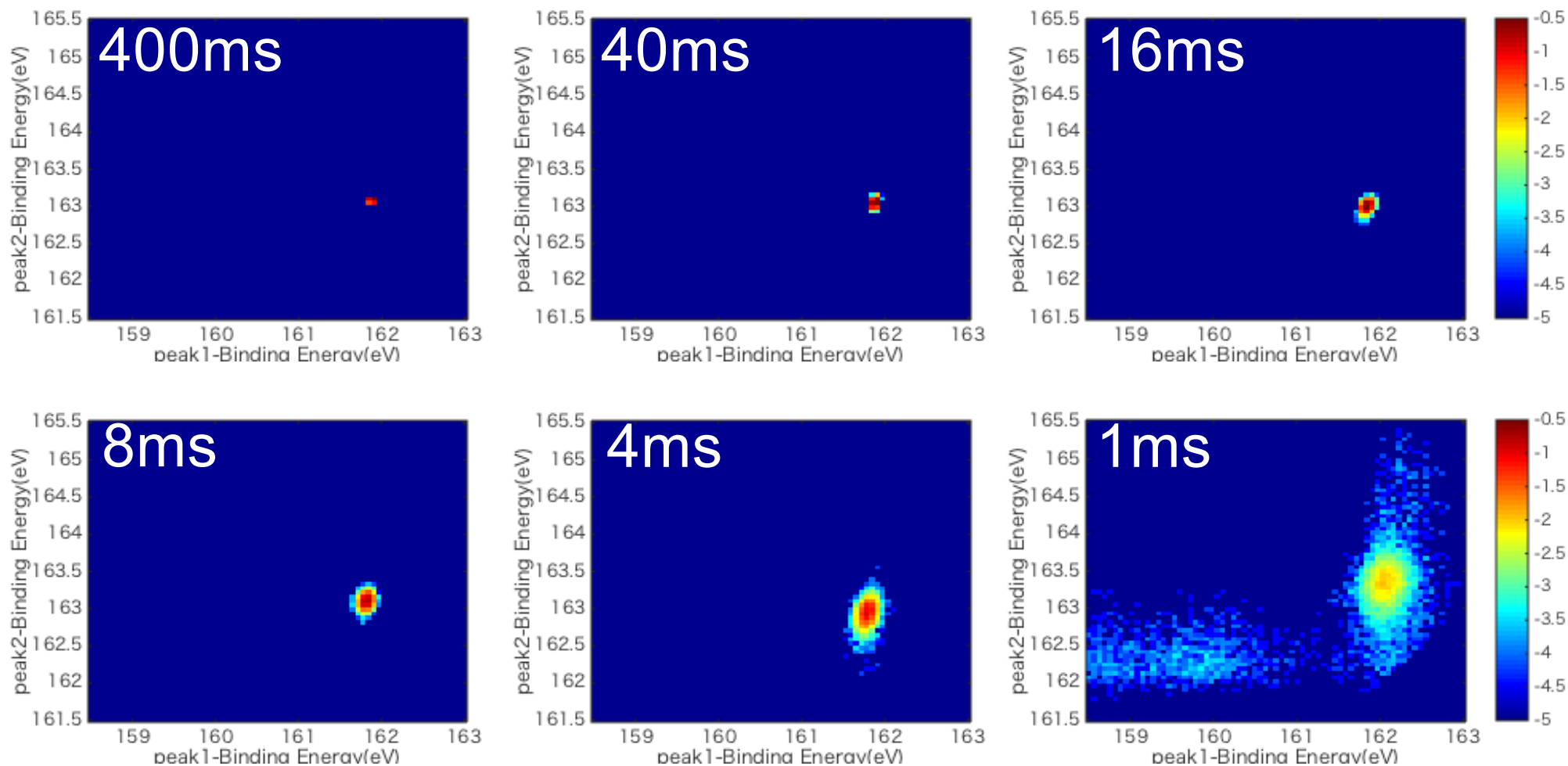
(時間幅依存性)



計測限界の理論的取り扱い (8/9)

(Nagata *et al.* 2019)

MoS₂ 2p



ここまで

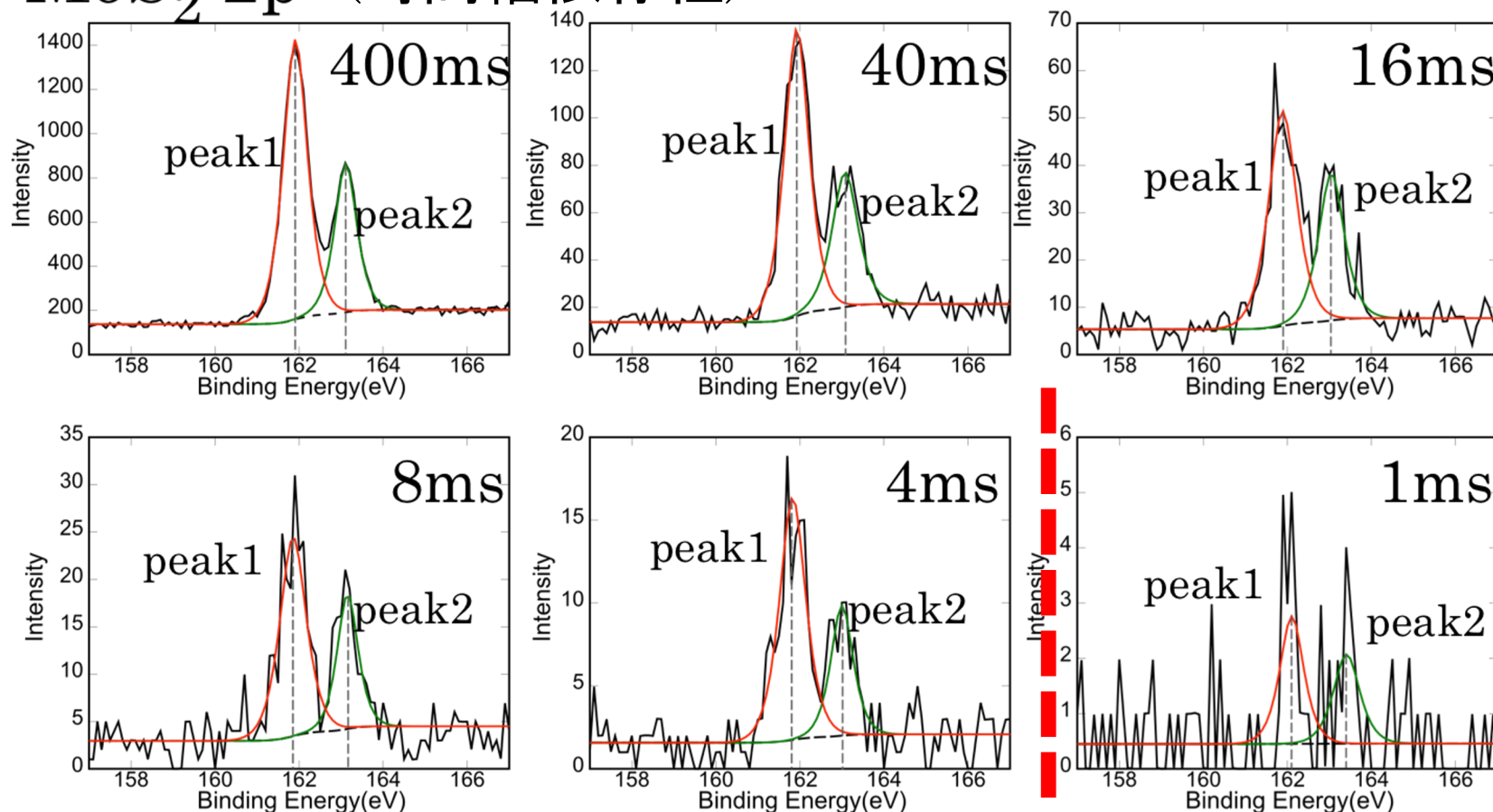
計測限界の理論的取り扱い (9/9)

(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)



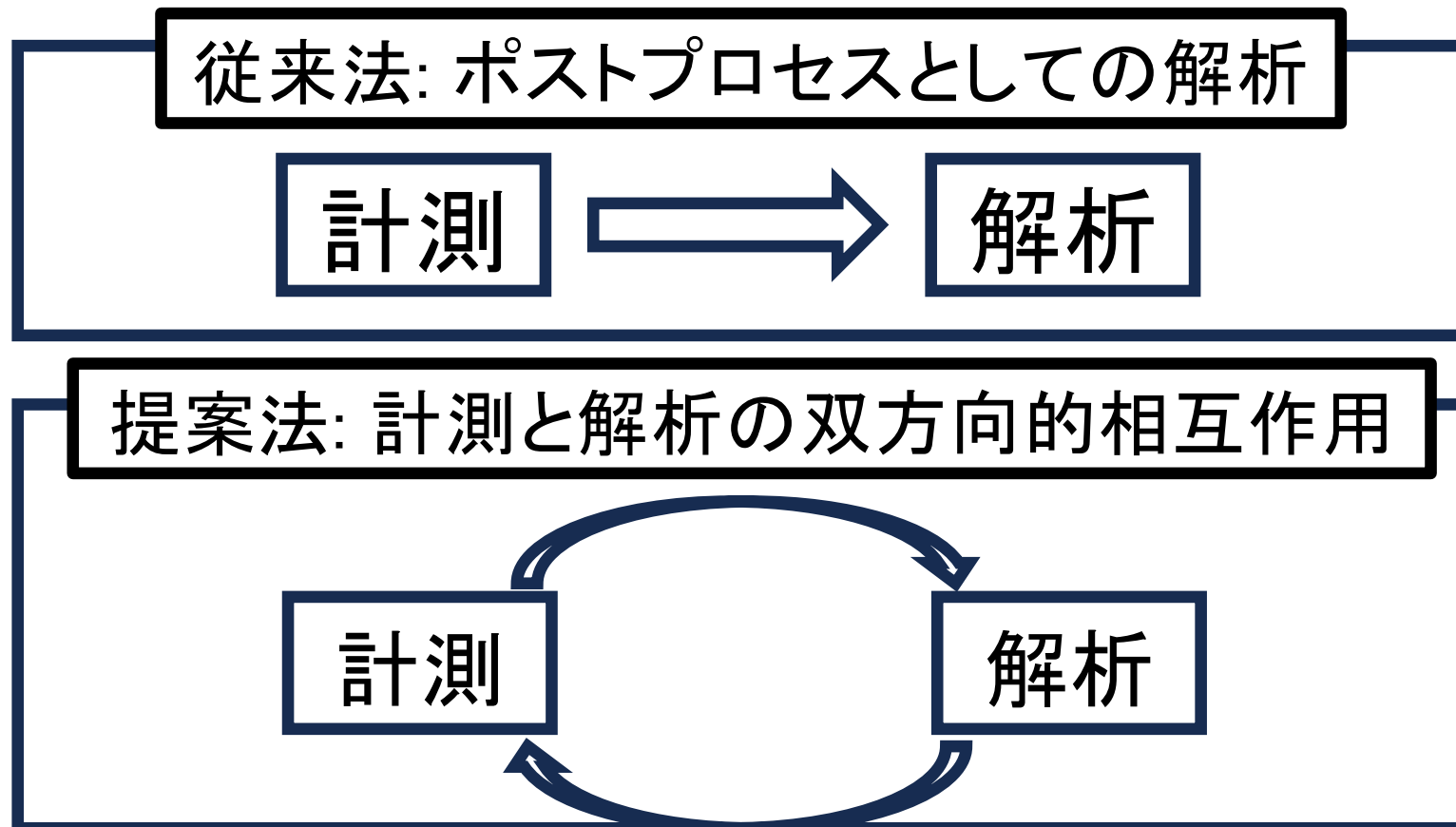
JASRIデータ駆動科学グループの役割

計測と解析の双方向相互作用

計測限界の向上と計測時間の短縮 ⇨

実験計画へのフィードバック

Y. Yokoyama et al., *Sci. Rep.* 13, 14349 (2023).



内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

マグネタイトにおける メスバウアースペクトルのベ イズ的解析

東京大学大学院 理学系研究科 森口 椋太

公益財団法人高輝度光科学研究センター 筒井 智嗣

東京大学大学院 新領域創成科学研究科 片上 舜

国立研究開発法人物質材料研究機構 永田 賢二

熊本大学大学院 先端科学研究部 水牧 仁一朗

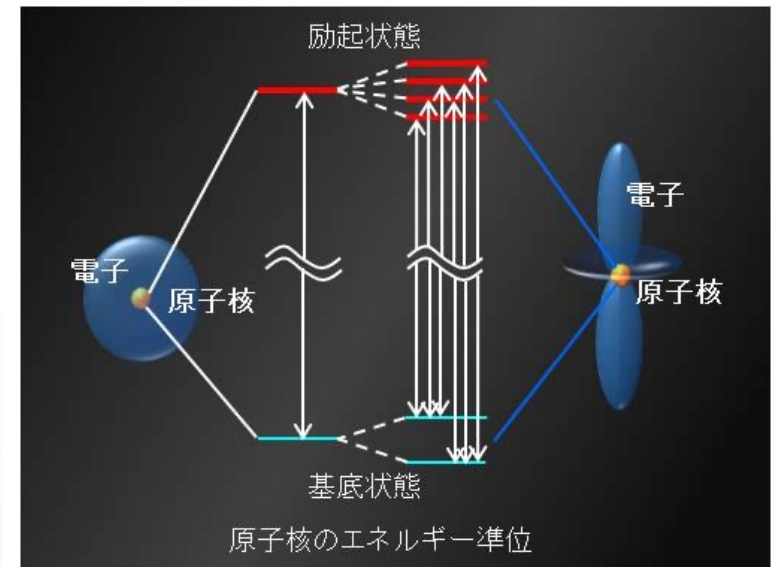
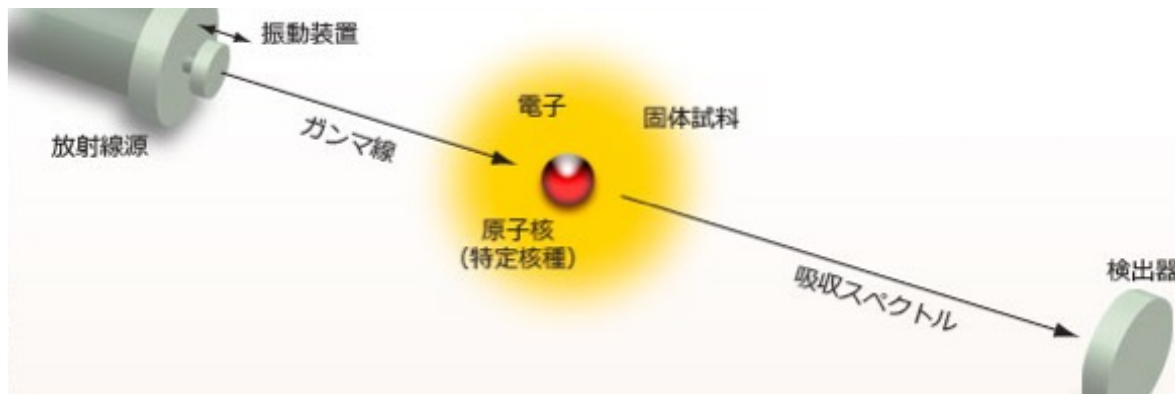
東京大学大学院 新領域創成科学研究科 岡田 真人

Moriguchi, Tsutsui, Katakami, Nagata, Mizumaki and Okada,
Journal of the Physical Society of Japan, 91, W104002 (2022)

メスバウアー分光 (1/7)

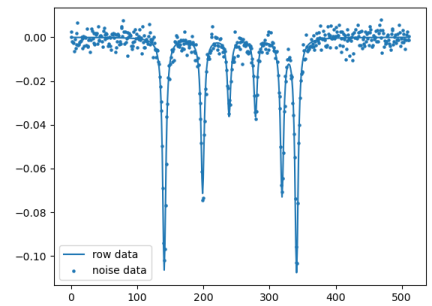
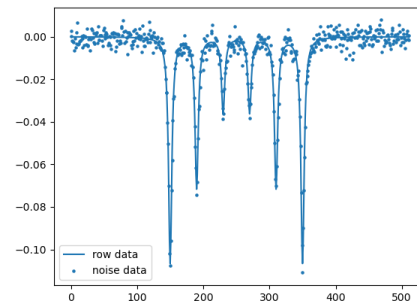
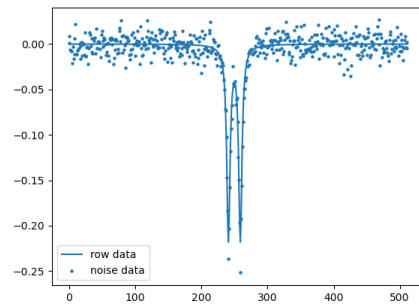
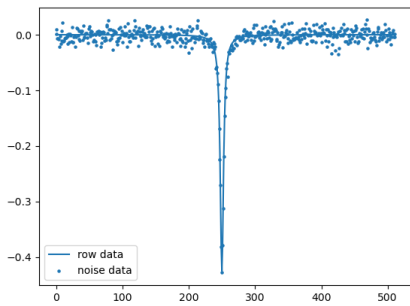
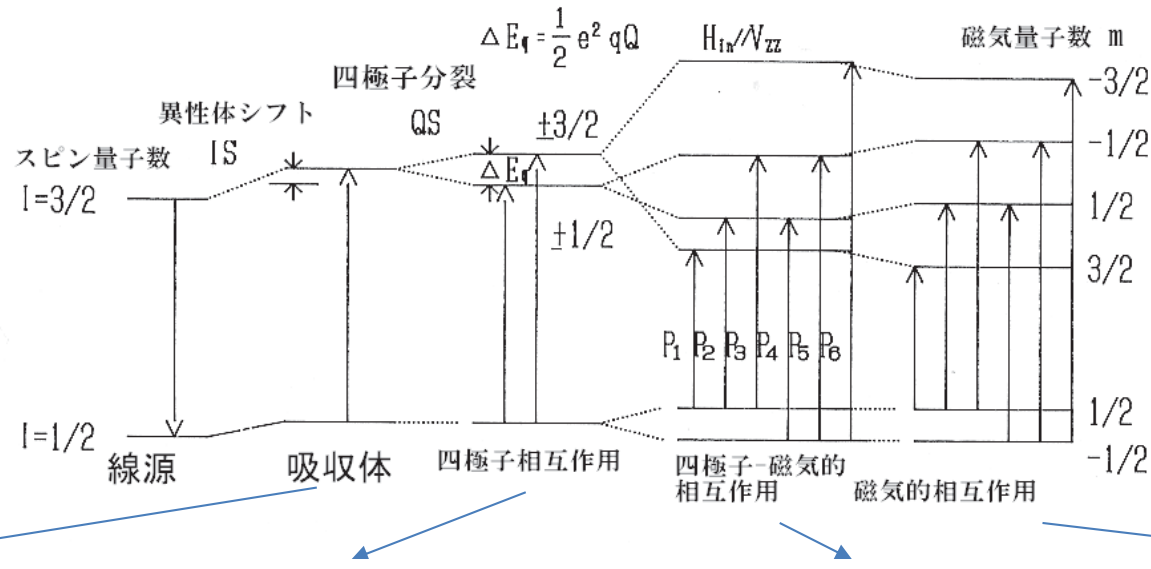
メスバウアー分光: 物質中の原子核の吸収スペクトルを測定

→物質中の原子核周りの内部磁場や電子状態を測定



メスバウアー分光 (2/7)

相互作用による吸収スペクトルの変化

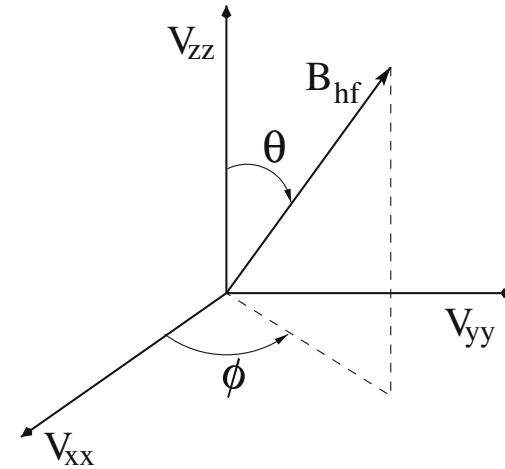


従来はスペクトルの形から相互作用を考えていたため専門家でないと解析が困難

メスバウアー分光 (3/7)

スペクトルに関する3つのハミルトニアン

- 核ゼーマン相互作用: $H_{M3/2}, H_{M1/2}$
- 四極子相互作用: $H_{Q3/2}, H_{Q1/2}$
- 異性体シフト: H_c



各ハミルトニアン

$$H_{M3/2} = \begin{pmatrix} \frac{3}{2}\alpha \cos \theta & -\frac{\sqrt{3}}{2}\alpha \sin \theta e^{i\phi} & 0 & 0 \\ -\frac{\sqrt{3}}{2}\alpha \sin \theta e^{-i\phi} & \frac{\alpha}{2} \cos \theta & -\alpha \sin \theta e^{i\phi} & 0 \\ 0 & -\alpha \sin \theta e^{-i\phi} & -\frac{\alpha}{2} \cos \theta & -\frac{\sqrt{3}}{2}\alpha \sin \theta e^{i\phi} \\ 0 & 0 & -\frac{\sqrt{3}}{2}\alpha \sin \theta e^{-i\phi} & -\frac{3}{2}\alpha \cos \theta \end{pmatrix}$$

$$H_{Q3/2} = \begin{pmatrix} 3A & 0 & \sqrt{3}A\eta & 0 \\ 0 & -3A & 0 & \sqrt{3}A\eta \\ \sqrt{3}A\eta & 0 & -3A & 0 \\ 0 & \sqrt{3}A\eta & 0 & 3A \end{pmatrix}$$

$$H_{M1/2} = \begin{pmatrix} -\frac{\beta}{2} \cos \theta & -\frac{\beta}{2} \sin \theta e^{-i\phi} \\ -\frac{\beta}{2} \sin \theta e^{i\phi} & +\frac{\beta}{2} \cos \theta \end{pmatrix}$$

$$H_{Q1/2} = 0$$

$$H_c = E_{center}$$

$$\begin{aligned} \alpha &= g_{\frac{3}{2}} \mu_N B_{hf} \\ \beta &= g_{1/2} \mu_N B_{hf} \end{aligned}$$

メスバウアー分光 (4/7)

尤度と事後分布の設定

物理モデル

$$F(x; \Theta) = \sum_k^K f(x; \Theta_k)$$

事後分布

$$p(\Theta|D) \propto \exp\left(-\frac{n}{\sigma^2}\beta E(\Theta)\right) \varphi(\Theta)$$

事前分布

$$\varphi(\Theta) \quad E(\Theta) = \sum_{i=1}^n \frac{1}{2n} (y_i - F(x_i; \Theta))^2$$

$$f(x; \Theta_k) := \sum_{i,j} r_k \times \frac{1}{\pi} \frac{INT_{i,j,k} \times \gamma_k}{(x - E_{i,j,k} - E_{shift,k})^2 + \gamma_k}$$

事前分布詳細

Γ の事前分布はガンマ分布, その他は一様分布

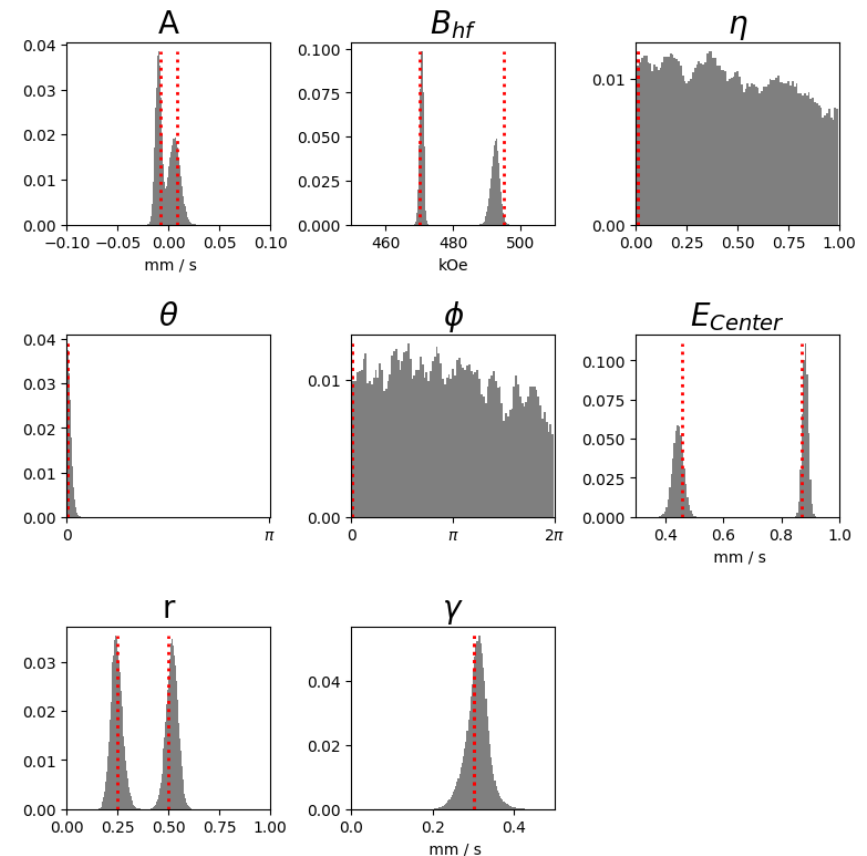
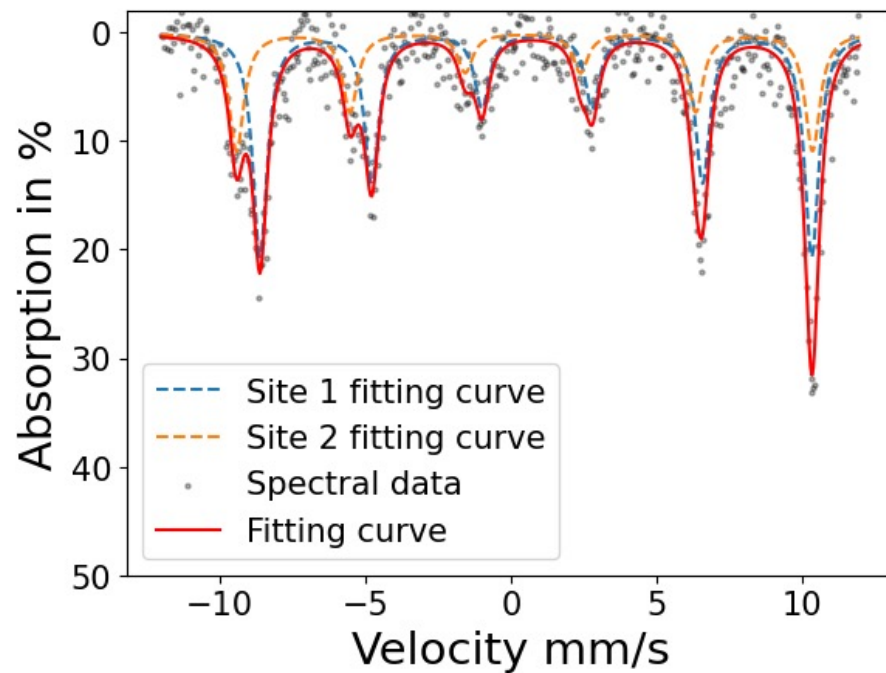
$A : \text{Uniform}(-1, 1), B_{hf} : \text{Uniform}(0, 10), \eta : \text{Uniform}(0, 1), \theta : \text{Uniform}(0, \pi),$
 $\phi : \text{Uniform}(0, 2\pi), E_{\text{center}} : \text{Uniform}(-1.0, 2.5), r : \text{Uniform}(0.0, 1.0), \Gamma : \text{Gamma}(1.5, 1.5)$

メスバウアー分光 (5/7)

物理パラメータの事後確率

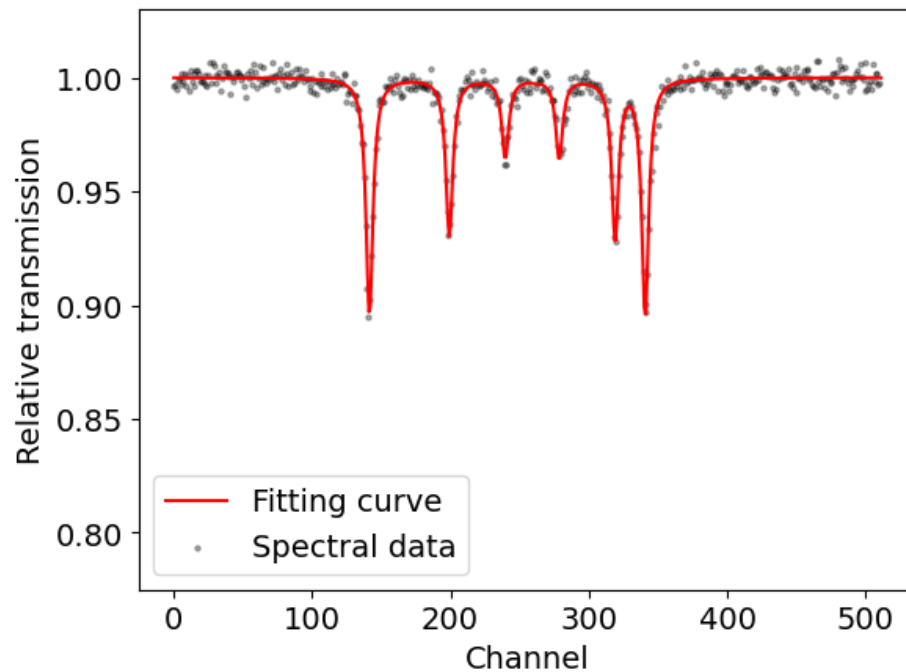
Posterior probability
(Red line: true values)

Result of fitting

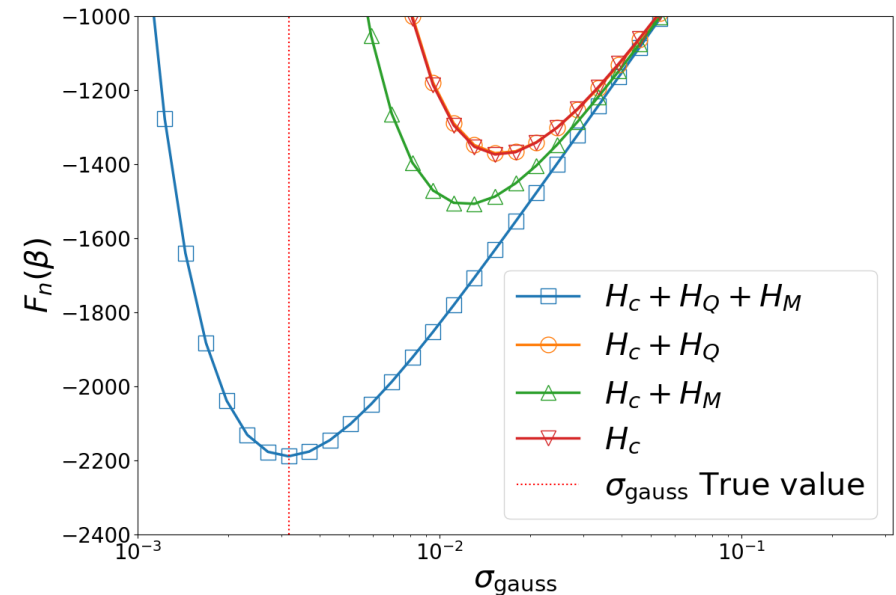


メスバウアー分光 (6/7) ハミルトニアン選択

核ゼーマン相互作用なし ($H_c+H_Q+H_M$) の数値実験

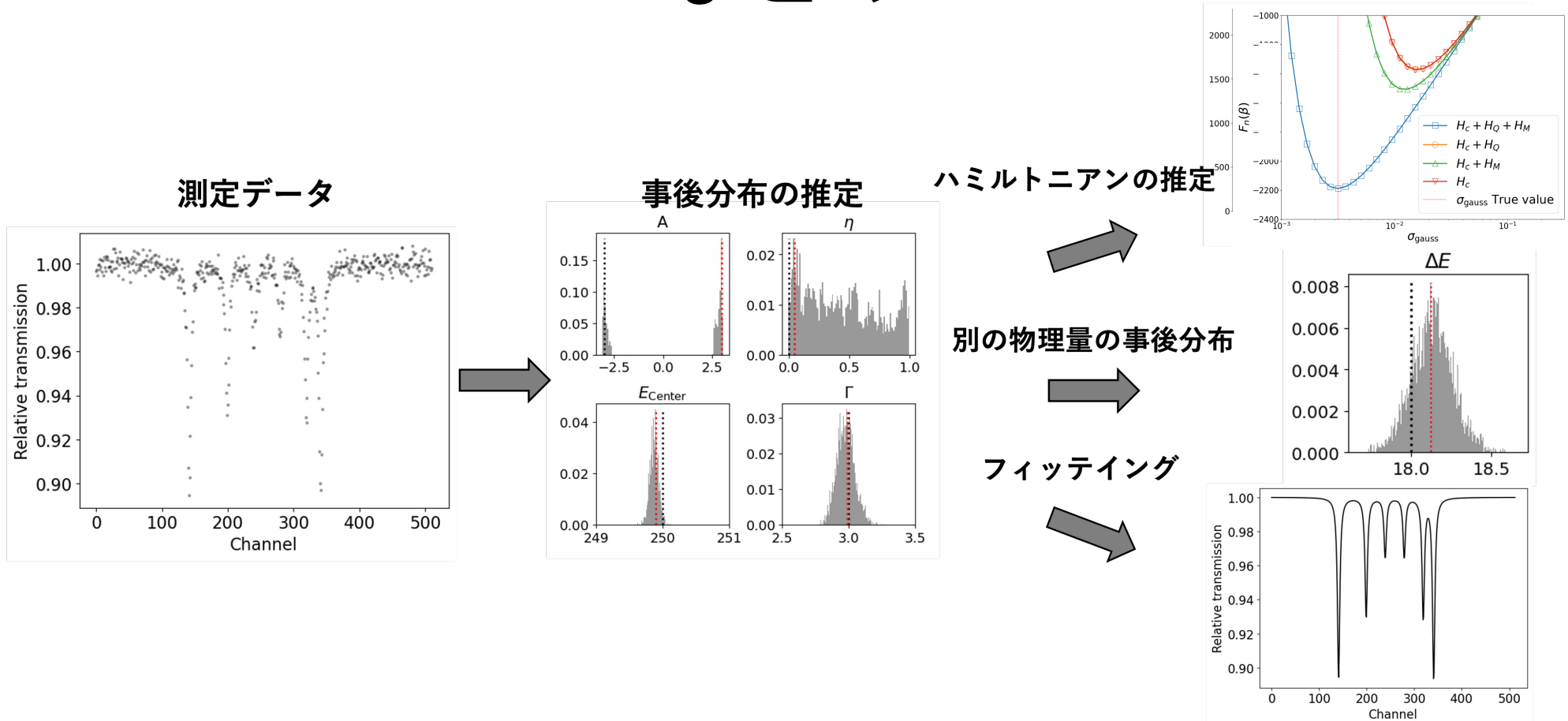


フィッティング



各ハミルトニアンでの自由エネルギー

メスバウアー分光 (7/7) まとめ



Moriguchi, Tsutsui, Katakami, Nagata, Mizumaki and Okada,
JPSJ, 91, W104002 (2022)

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- まとめと今後の展開

小角散乱法データを用いた 試料パラメータのベイズ推論

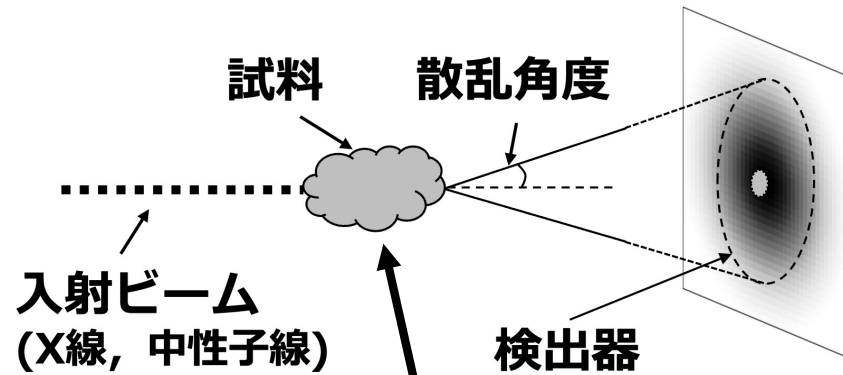
林悠偉^A, 片上舜^A, 桑本滋生^B, 永田賢二^C,
水牧仁一郎^B, 岡田真人^A

東大新領域^A 高輝度光科学研究セ^B 物材機構^C

Hayashi, Katakami, Kuwamoto, Nagata, Mizumaki, and Okada, "Bayesian Inference for Small-Angle Scattering Data", *Journal of the Physical Society of Japan* 92(9) (2023).

小角散乱 (1/6)

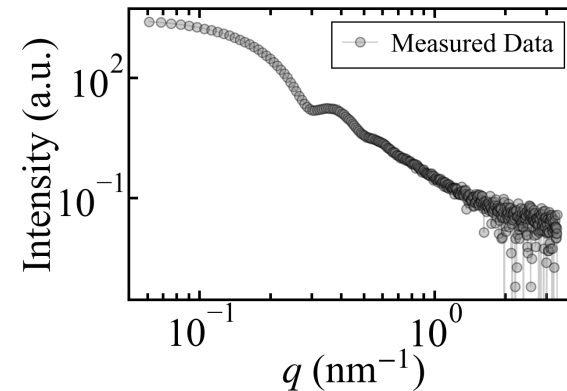
試料にビームを照射し、小さい角度の散乱強度から
ナノ構造解析を行う。



不均一系試料

- 高分子溶液
- タンパク質
- コロイド粒子
- 金属, セラミック
- 繊維 ...

散乱強度-散乱ベクトル

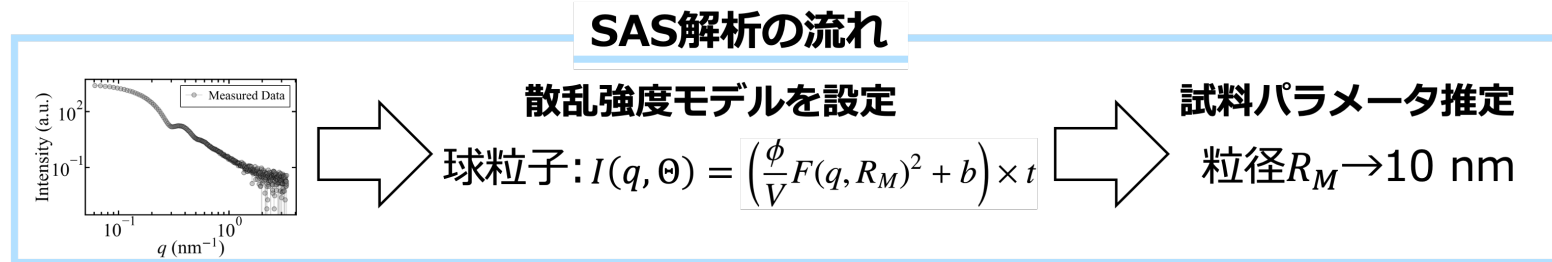


データ解析

1. 散乱強度モデルを設定
2. モデルパラメータ推定

小角散乱 (2/6)

従来法とその課題



- 試料パラメータ推定に勾配法などを用いているために、局所解が得られることが多い。
→ 何度もパラメータの初期値を変えて対処している。

- 解析に用いる散乱強度モデルが予め分からないと解析できない。
→ フィッティングの二乗誤差や経験則を基にモデル選択している。
本研究の目的

2つの従来法の課題をベイズ計測を応用し解決する。

小角散乱 (3/6)

- ベイズ計測による定式化

データ点 y_i の確率分布 : $p(y|q, \Theta, K) = \frac{I_K(q, \Theta)^y \exp(-I_K(q, \Theta))}{y!}$

(∵散乱強度は, 光子のカウントデータとして計測)

データ \mathcal{D} の生成確率 : $p(\mathcal{D}|\Theta, K) = \prod_{i=1}^N p(y_i|q_i, \Theta, K)$
 $= \exp(-NE(\Theta, K))$

→誤差関数 : $E(\Theta, K) = \frac{1}{N} \sum_{i=1}^N \left\{ I_K(q_i, \Theta) - y_i \log I_K(q_i, \Theta) + \sum_{j=1}^{y_i} \log j \right\}$

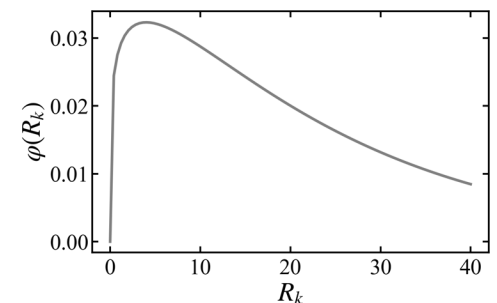
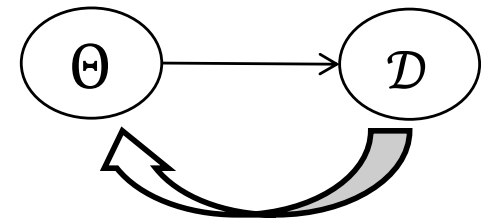
パラメータ θ の事後確率 : $p(\Theta|\mathcal{D}) = \frac{p(\mathcal{D}|\Theta)\varphi(\Theta)}{\int p(\mathcal{D}, \Theta)d\Theta}$
 $\propto \exp(-NE(\Theta))\varphi(\Theta)$

モデル K の事後確率 : $p(K|\mathcal{D}) = \frac{\int p(\mathcal{D}, \Theta, K)d\Theta}{\sum_K \int p(\mathcal{D}, \Theta, K)d\Theta}$
 $= \frac{\exp(-F(K))\varphi(K)}{\sum_K \exp(-F(K))\varphi(K)}$

事前分布 $\varphi(\Theta)$: 事前知識を基に設定

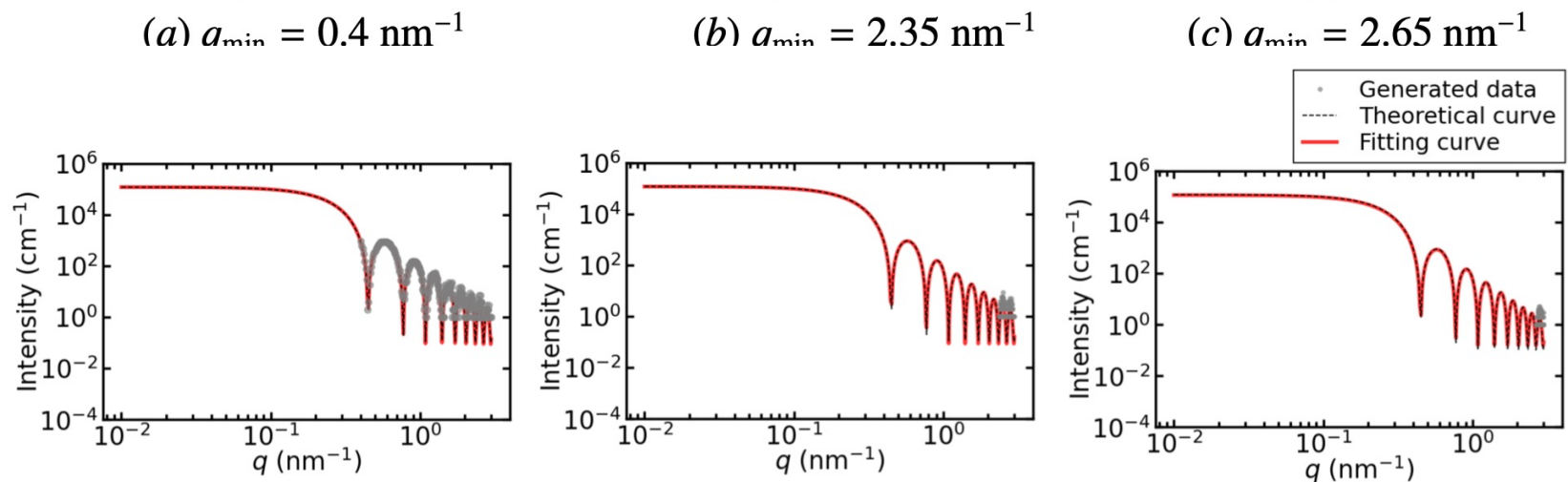
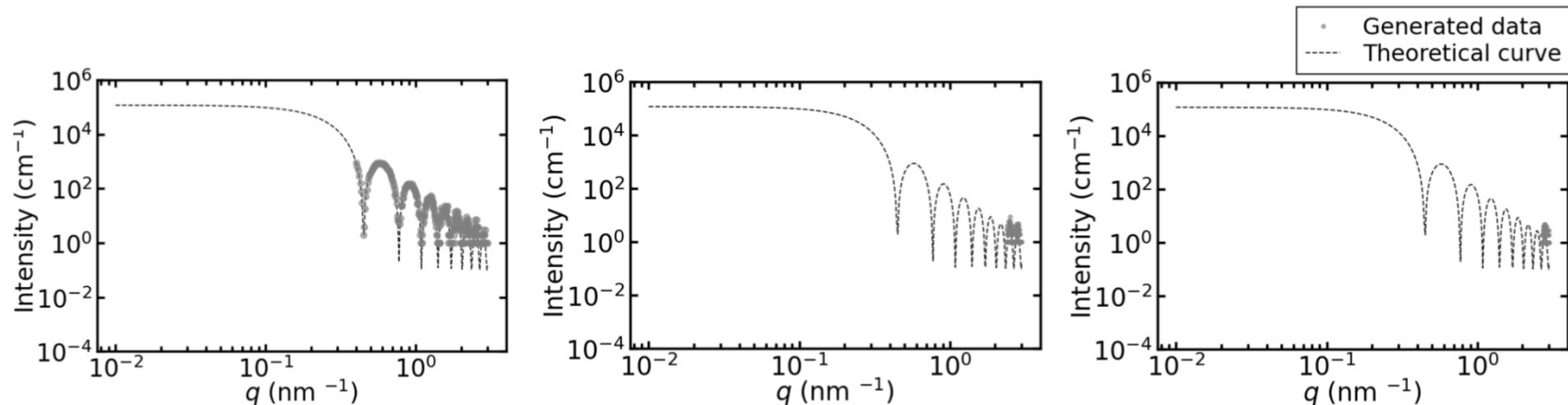
散乱強度の計測値 : y_i
 散乱ベクトル : q_i
 試料モデル : $I(q_i; \Theta)$

生成(因果律)



小角散乱 (4/6)

パラメータ推定の数値実験

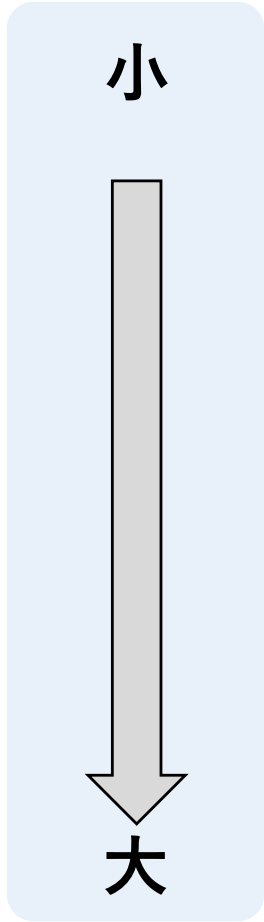


Hayashi, Katakami, Kuwamoto, Nagata, Mizumaki, and Okada, *JPSJ*. 92, 094002 (2023)

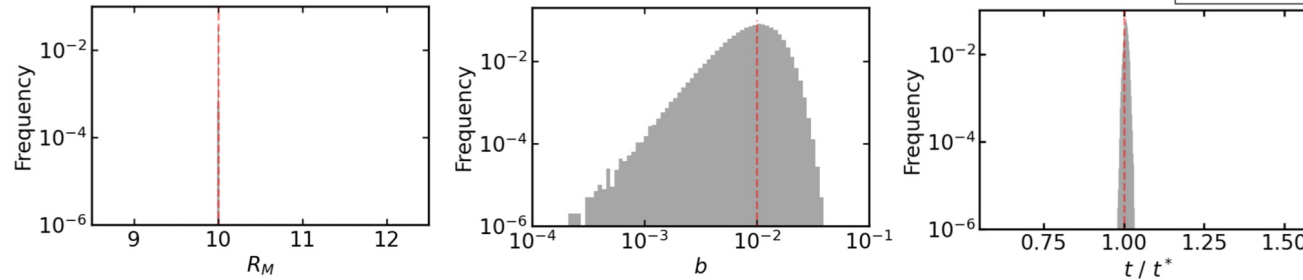
小角散乱 (5/6)

パラメータの事後分布

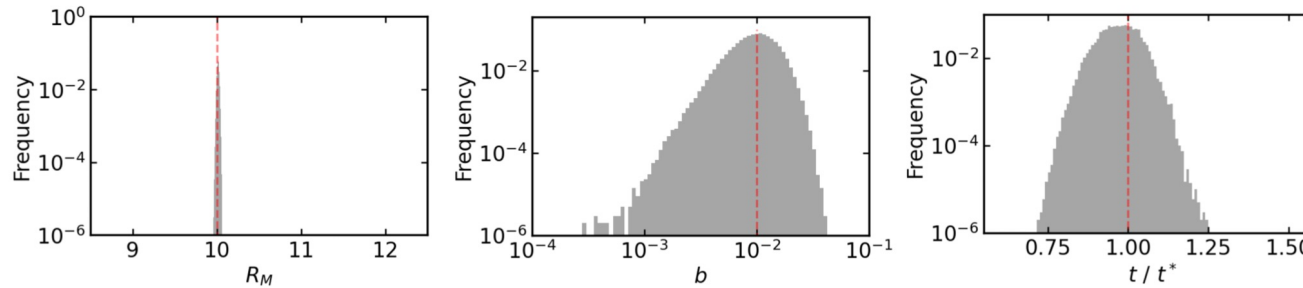
データ欠損



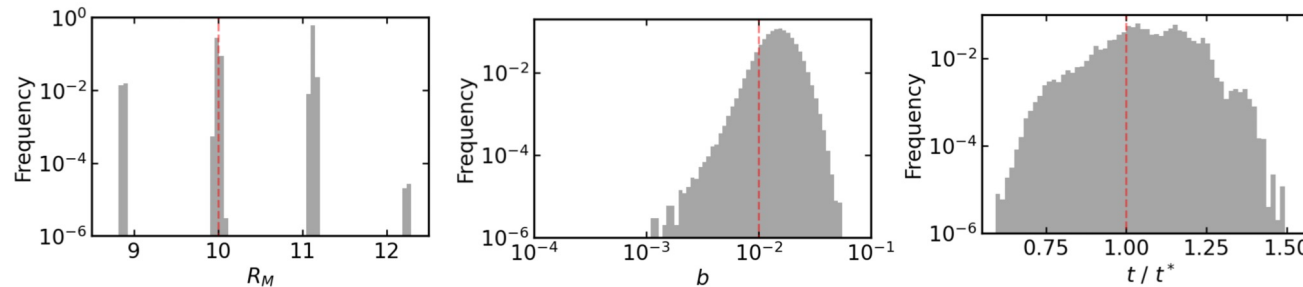
(a) $q_{\min} = 0.4$



(b) $q_{\min} = 2.35$

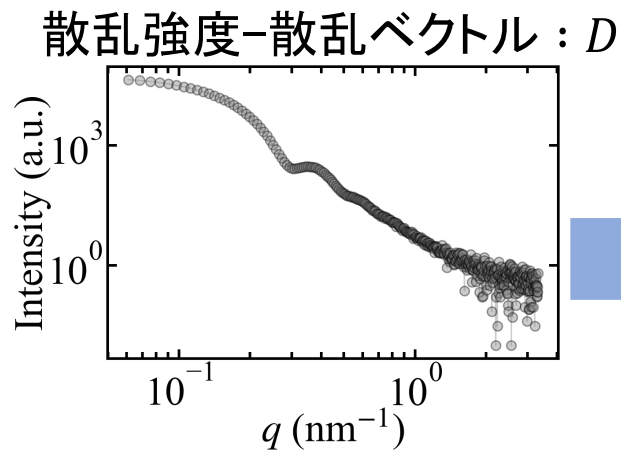


(c) $q_{\min} = 2.65$



(c) に複数のピーク → $q_{\min} = 2.35 - 2.65 \text{ nm}^{-1}$ で推定限界
Hayashi, Katakami, Kuwamoto, Nagata, Mizumaki, and
Okada, *JPSJ*. 92, 094002 (2023)

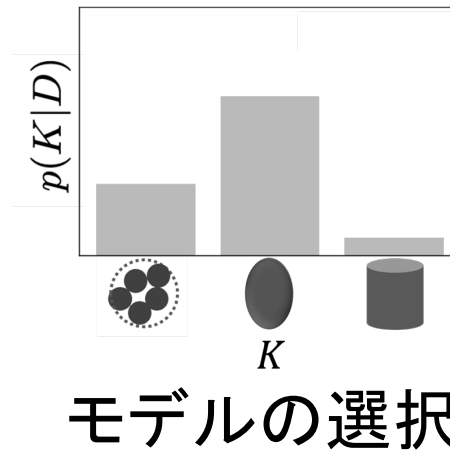
小角散乱 (6/6)



モデル選択

$$p(K|D) = \frac{\int p(D, \Theta, K) d\Theta}{\sum_K \int p(D, \Theta, K) d\Theta}$$

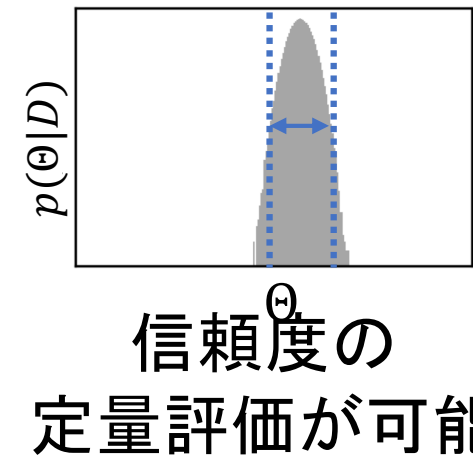
試料構造を表す散乱
強度モデルの選択



パラメータ推定

$$p(\Theta|D) = \frac{p(D|\Theta)p(\Theta)}{\int p(D, \Theta) d\Theta}$$

試料の大きさや密度
のパラメータを推定



ベイズ計測の適用例

東京大学 岡田研究室

- 事後分布推定 and/or モデル選択
 1. スペクトル分解
 2. X線光電子放出スペクトル(XPS)
 3. X線吸収スペクトル(XAS)
 4. メスバウアー分光
 5. X線小角散乱スペクトル
 6. NMR
 7. 中性子非弾性散乱スペクトル
 8. 比熱
 9. 帯磁率
- ベイズ統合
 1. XPSとXAS
 2. 比熱と帯磁率

熊本大学 赤井研究室

- 事後分布推定 and/or モデル選択
 1. フォトルミネッセンススペクトル

熊本大学 水牧研究室

- 事後分布推定 and/or モデル選択
 1. XRD

ベイズ計測の枠組みは様々な計測データに適用できる
(シミュレーションはノートPCで実行できる場合が多い)

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- **ベイズ計測とキャリアパス**
- まとめと今後の展開

ベイズ計測とキャリアパス (1/3)

アカデミア編

物理学科/各学科に

データ駆動科学—講座導入

理論物理

実験物理

素粒子論

物性理論

光物性

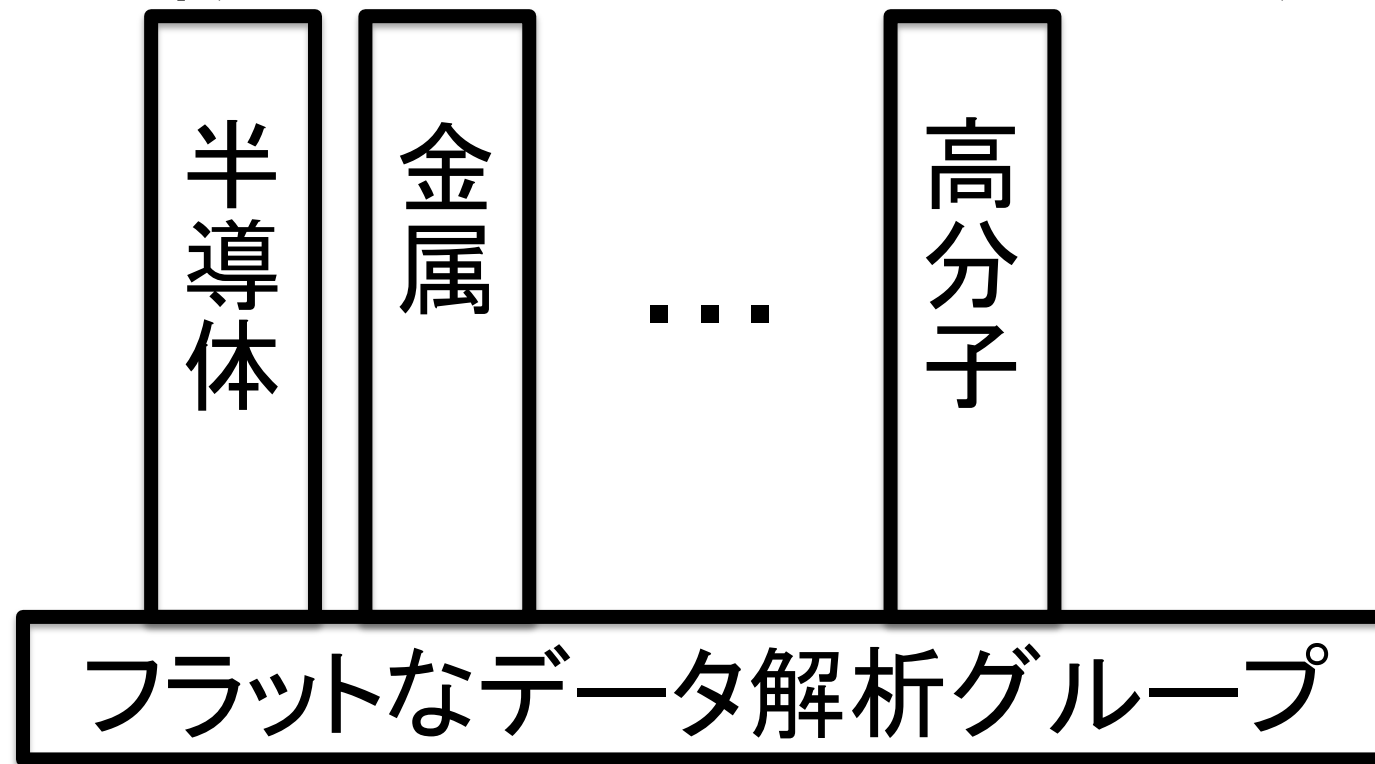
磁性物理

データ駆動科学講座

ベイズ計測とキャリアパス (2/3)

民間企業編

R & D組織のフラット化と人材の流動化



ある縦組織が**リストラ**されても、フラットなデータ解析グループに沿って、**他の縦組織にソフトランディング**でき、**人材の流動化が加速**される。

ベイズ計測とキャリアパス (3/3)

放射光科学関連

- SPring-8全ンビームラインベイズ化計画により、アカデミアだけでなく民間企業でもベイズ計測を取り入れる動きが急速に進んでいる。民間企業は動き出すとアカデミアより早い。
- SPring-8/JASRIのような従来の計測サービス部門シーズと民間企業のニーズの乖離
- そこを埋めるために、新たな市場が形成
 - a.s.ist (東大岡田研の学生が起業)
<https://www.a-s-ist.com/>
- **ベイズ計測を用いれば**、自分のポジションは、自分が創れる。

内容

- 自己紹介
- SPring-8全ビームラインベイズ化計画
- データ駆動科学の三つのレベル
- ベイズ計測
 - 計測科学の必須条件
 - ベイズ計測三種の神器
 - $y=ax+b$ のベイズ計測の解析的取り扱い
- 非線形系にするベイズ計測
 - スペクトル分解、メスバウアー分光、小角散乱
- ベイズ計測とキャリアパス
- **まとめと今後の展開**

まとめと今後の展望

- $y=ax+b$ を例にベイズ計測とベイズ計測三種の神器を解説。一般的な非線形の計測の紹介
- ベイズ計測の物性物理学への導入が大きなビジネスチャンス。
- ベイズ計測の普及戦略としてのSPring-8全ビームラインベイズ化計画
- 物理学全般を取り扱うことができるデータ駆動科学の枠組みの紹介
- 物理学科/各学科にデータ駆動科学—講座導入
- 民間企業のR&D組織の新陳代謝を加速
- 自分のポジションは、自分が創る。

JASRIデータ駆動科学グループの役割

計測と解析の双方向相互作用

計測限界の向上と計測時間の短縮 ⇨

実験計画へのフィードバック

Y. Yokoyama et al., *Sci. Rep.* 13, 14349 (2023).

