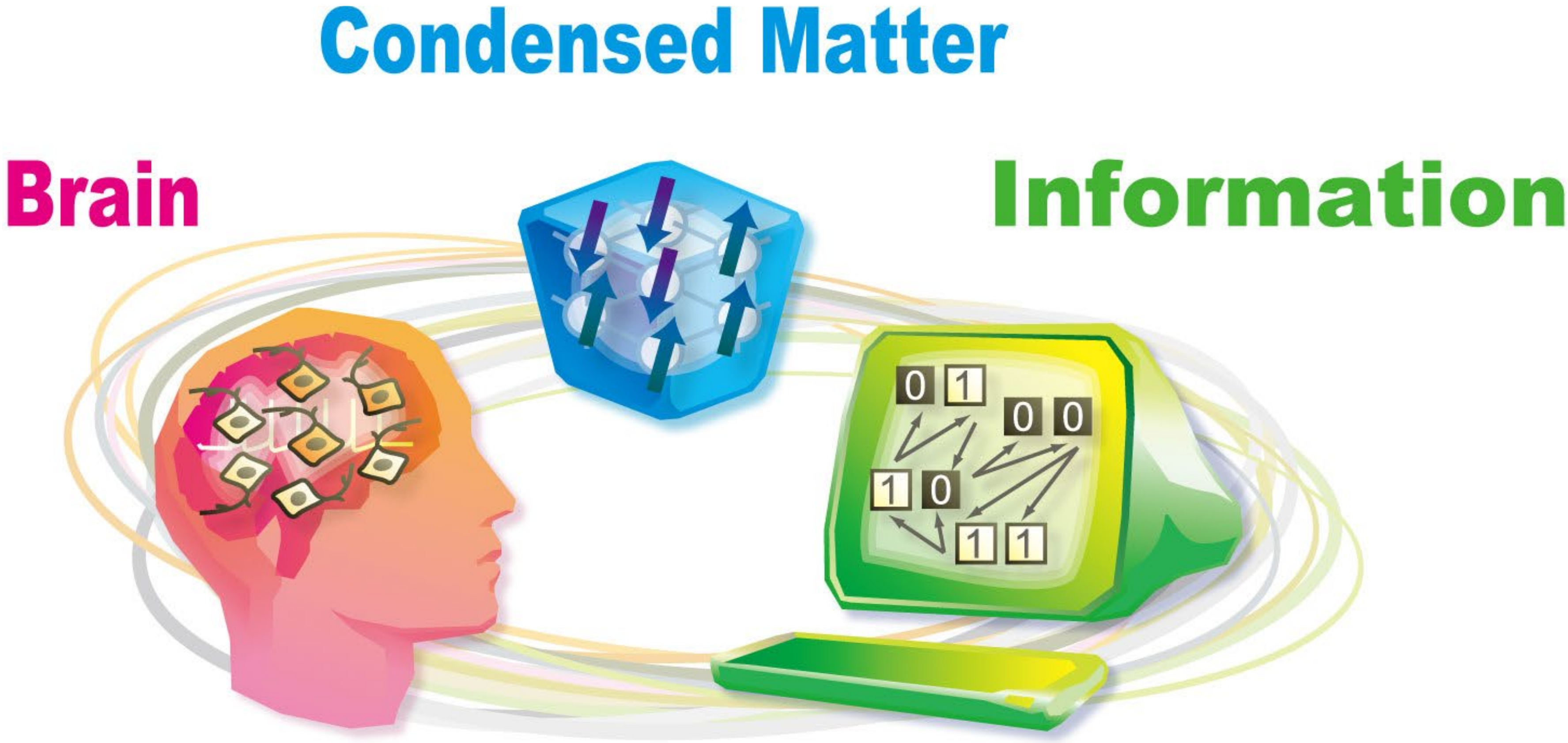
ベイズ推論と物性科学

岡田 真人
okada@k.u-tokyo.ac.jp
東京大学　大学院新領域創成科学研究科
複雑理工学専攻



Illustrated by Satohiro Tajima

自己紹介

- 大阪市立大学理学部物理学科 (1981 - 1985)
- 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 (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)
- JST さきがけ「協調と制御」 (2002 – 2006)
- 東京大学・大学院新領域創成科学研究科 複雑理工学専攻 複雑行動知能学分野 (2004/07 – 2008/8)
- 東京大学・大学院新領域創成科学研究科 複雑理工学専攻 情報認知機構分野 (2008/09 -)

修士課程の指導教官小谷章雄先生
(阪大金森研　城先生)

Journal of the Physical Society of Japan
Vol. 56, No. 2, February, 1987, pp. 798–809

Many Body Effect in Inner Shell Photoemission and Photoabsorption Spectra of La Compounds

Akio KOTANI, Masato OKADA, Takeo JO,
A. BIANCONI,[†] A. MARCELLI[†] and J. C. PARLEBAS[††]

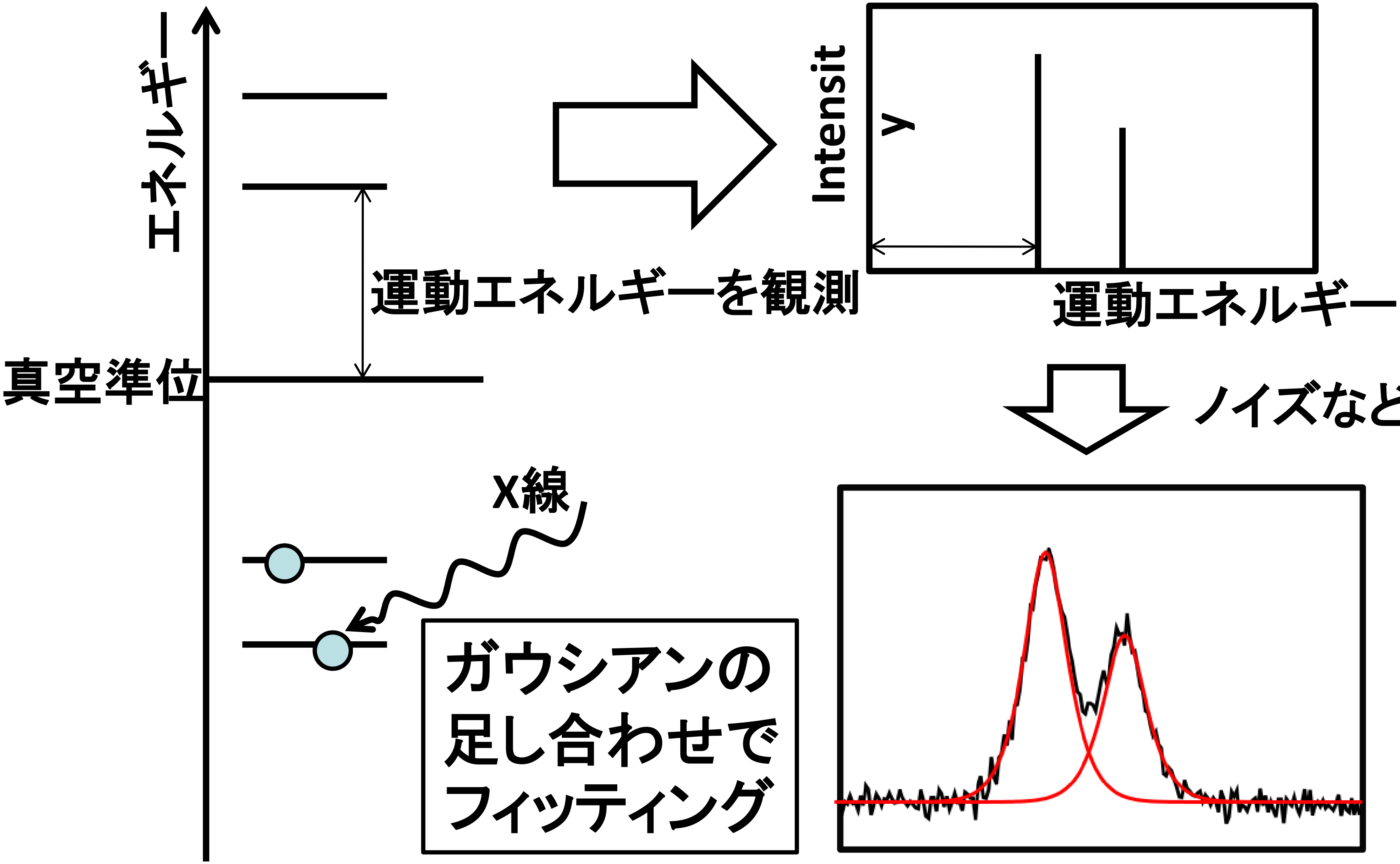
Department of Physics, Faculty of Science, Osaka University, Toyonaka 560
[†]Dipartimento di Fisica, Università di Roma "La Sapienza", 00185 Roma, Italy
[††]LMSES, Université Louis Pasteur, 67070 Strasbourg, France

(Received October 14, 1986)

REFERENCES

1. A. Kotani & Y. Toyozawa, *J. Phys. Soc. Japan* **37,** 912 (1974).
2. O. Gunnarsson & Schönhammer, *Phys. Rev.* **B27,** 4315 (1983).
3. A. Fujimori, *Phys. Rev.* **B28,** 2281 (1983).

X-ray Photoelectron Spectroscopy（XPS）とは対象となる物質の電子状態を調べる測定法



混合基底関数でスペクトルは書ける

The spectra of 3d-XPS and L_3-XAS are expressed as

$$F_{\mathrm{XPS}}(E_B) = \sum_f |\langle f|a_c|g\rangle|^2 \mathrm{L}(E_B - \underline{E_f + E_g}), \quad (2)$$

$$F_{\mathrm{XAS}}(\omega) = \frac{1}{N}\sum_f |\langle f| \sum_k a_d^{+}(k) a_c|g\rangle|^2 \times \mathrm{L}(\omega - \underline{E_f + E_g}), \quad (3)$$

where

$$\mathrm{L}(x) = \underline{\Gamma}/[\pi(x^2 + \underline{\Gamma^2})]. \quad (4)$$

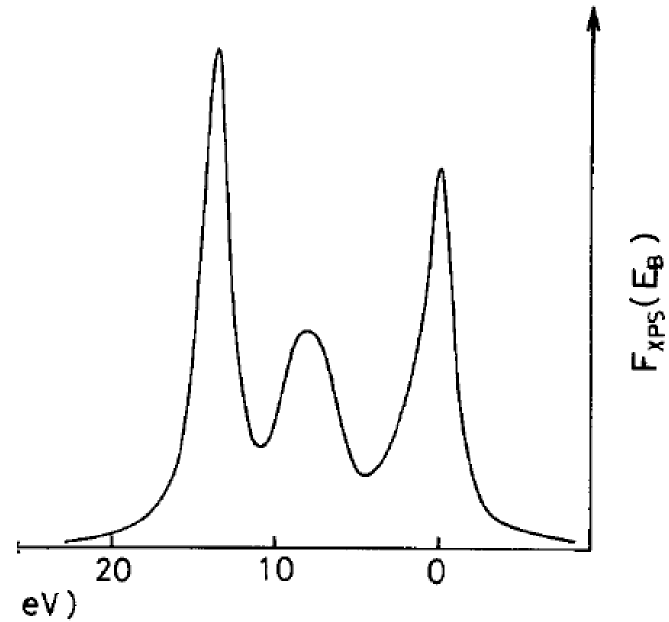


Fig. 2. Calculated result of 3d-XPS. The origin of the binding energy E_B is taken arbitrarily.

スペクトルで多体効果をみる

(A. Kotani, K. Okada and M. Okada, 1987)

ピークの位置と幅から，対象の物理的性質の考察が可能になる

順モデルからパラメータフィットしてスペクトルを再構成

混合原子価希土類化合物のL_3-XASを解析して

- 小谷先生の主張:U_{dc}が必要
- 結論
  - 絶縁体では5eV程度
  - 金属では1～2eV程度
- 当時(25年前)感じた問題点
  - 発見法的なパラメータサーチ
  - XPSを決めてからXASを解析
  - 余分なパラメータを導入?
  - オーバーフィット
  - 推定パラメータの誤差評価

→情報科学的アプローチ

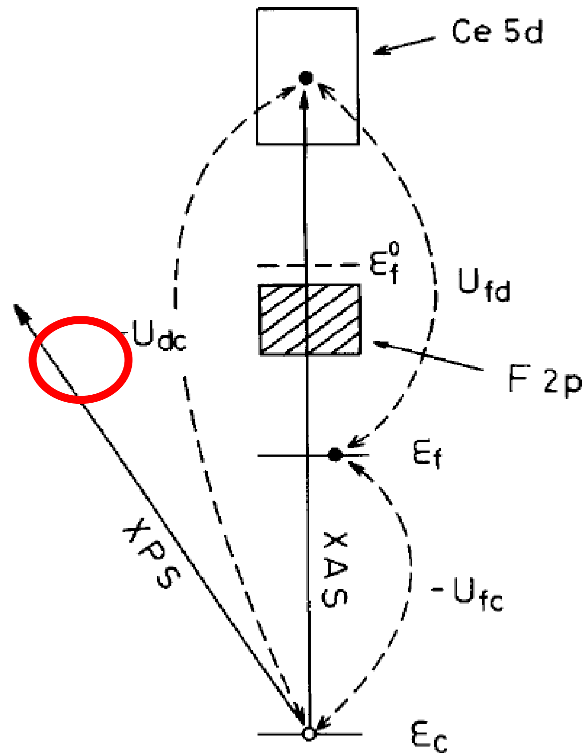


Fig. 1. Model of the present theory describing 3*d*-XPS and L_3-XAS.

KotaniとGunnarssonの主張

- Kotaniの主張：U_{dc}が必要
- 結論
  - 絶縁体には必要
  - 金属では不必要

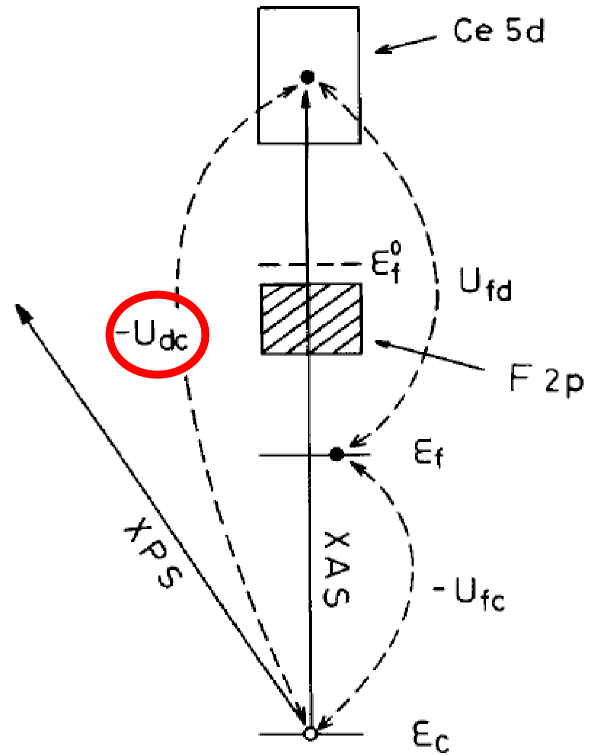
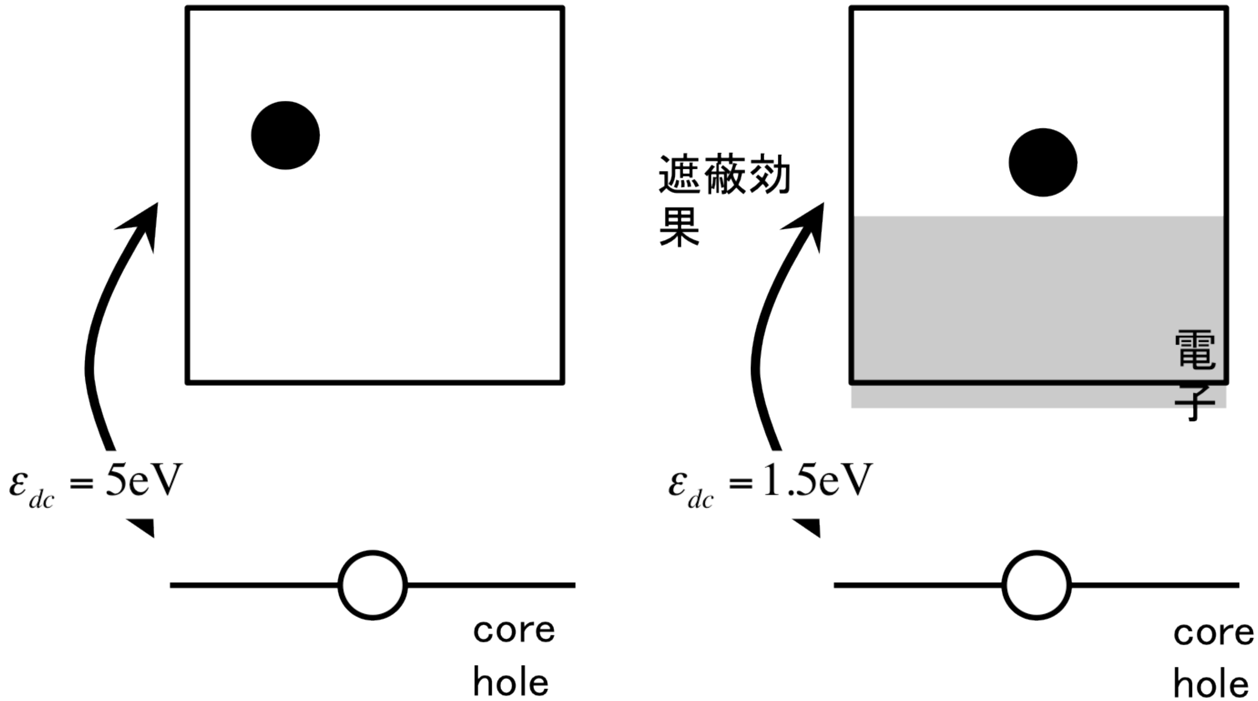


Fig. 1. Model of the present theory describing 3*d*-XPS and L_3-XAS.

絶縁体と金属



- 金属では励起電子の緩和プロセスが見えていた
- 励起の初期にはε_{dc}が絶縁体程度(Kotaniの主張)
- 緩和することで遮蔽効果がきくε_{dc} =0(Gunnarsonの主張)
- このようなことを実証するのは現在は可能でしょうか？

混合原子価希土類化合物のL_3-XASを解析して

- 小谷先生の主張：U_{dc}が必要
- 結論
  - 絶縁体では5eV程度
  - 金属では1～2eV程度
- 当時(25年前)感じた問題点
  - 発見法的なパラメータサーチ
  - XPSを決めてからXASを解析
  - 余分なパラメータを導入?
  - オーバーフィット
  - 推定パラメータの誤差評価

→情報科学的アプローチ

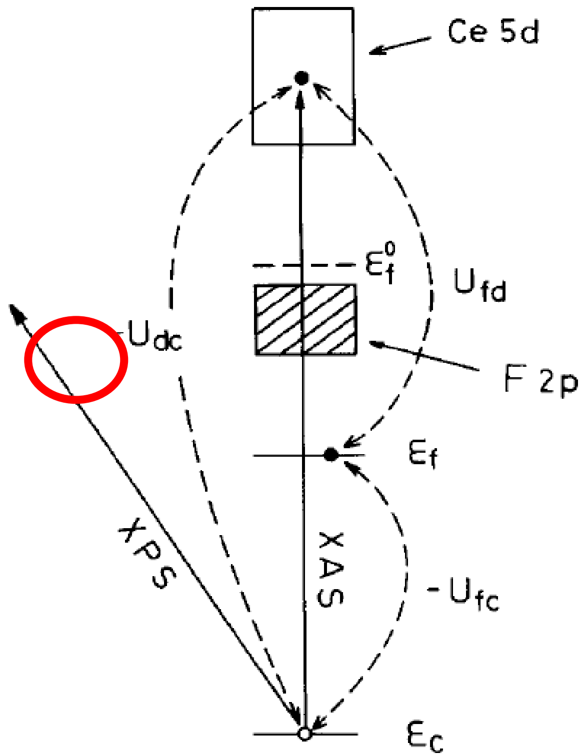


Fig. 1. Model of the present theory describing 3*d*-XPS and L_3-XAS.

ベイズ推論と物性科学

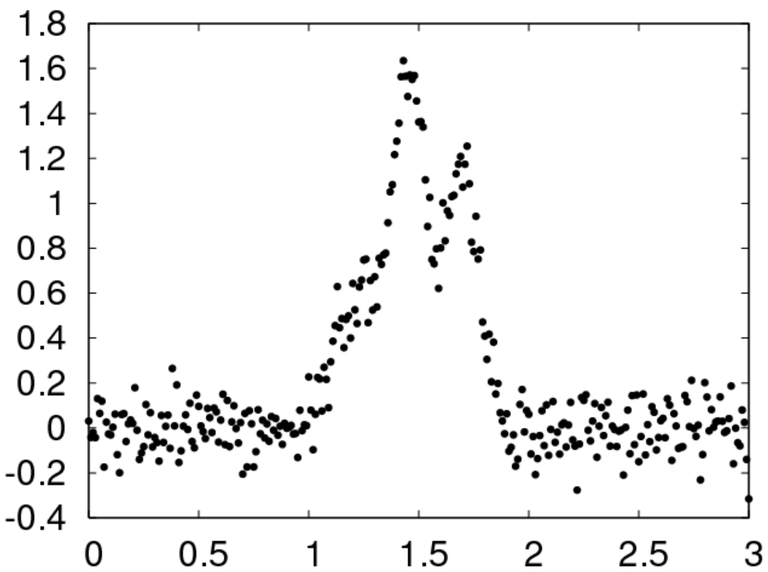
岡田 真人

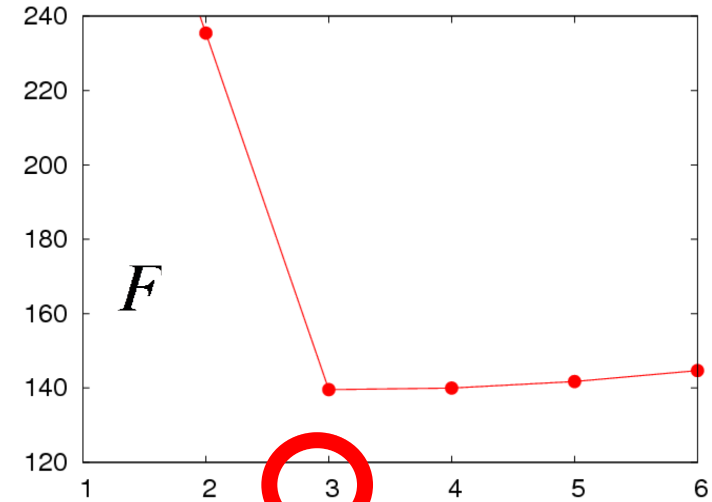
okada@k.u-tokyo.ac.jp

東京大学　大学院新領域創成科学研究科

複雑理工学専攻

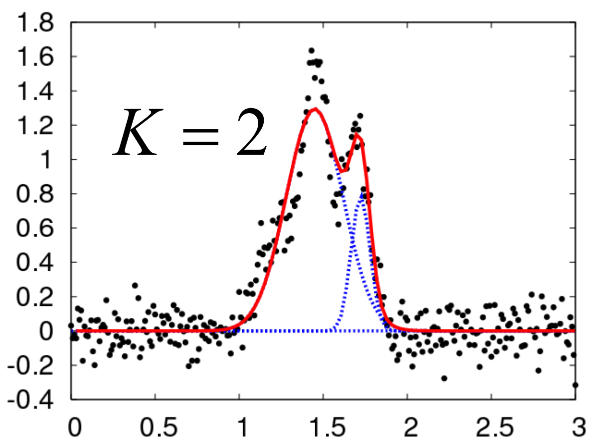
スペクトル分解: 力強い良い例題1

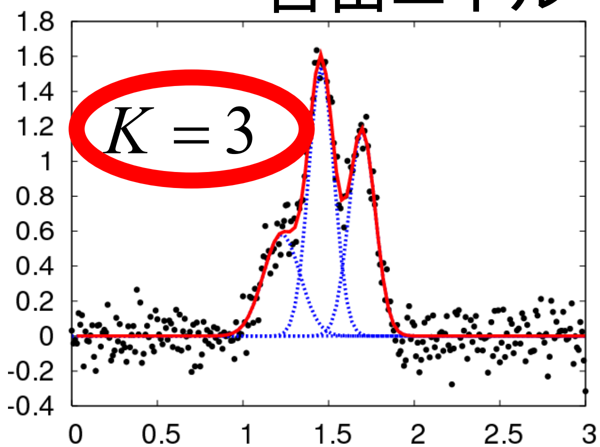


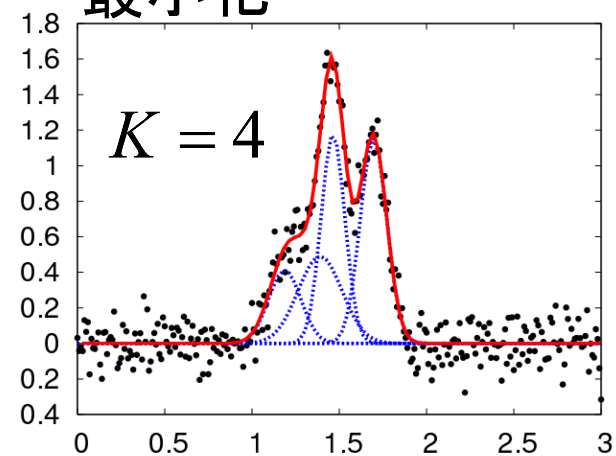


最適なKをデータだけから決める

自由エネルギー最小化

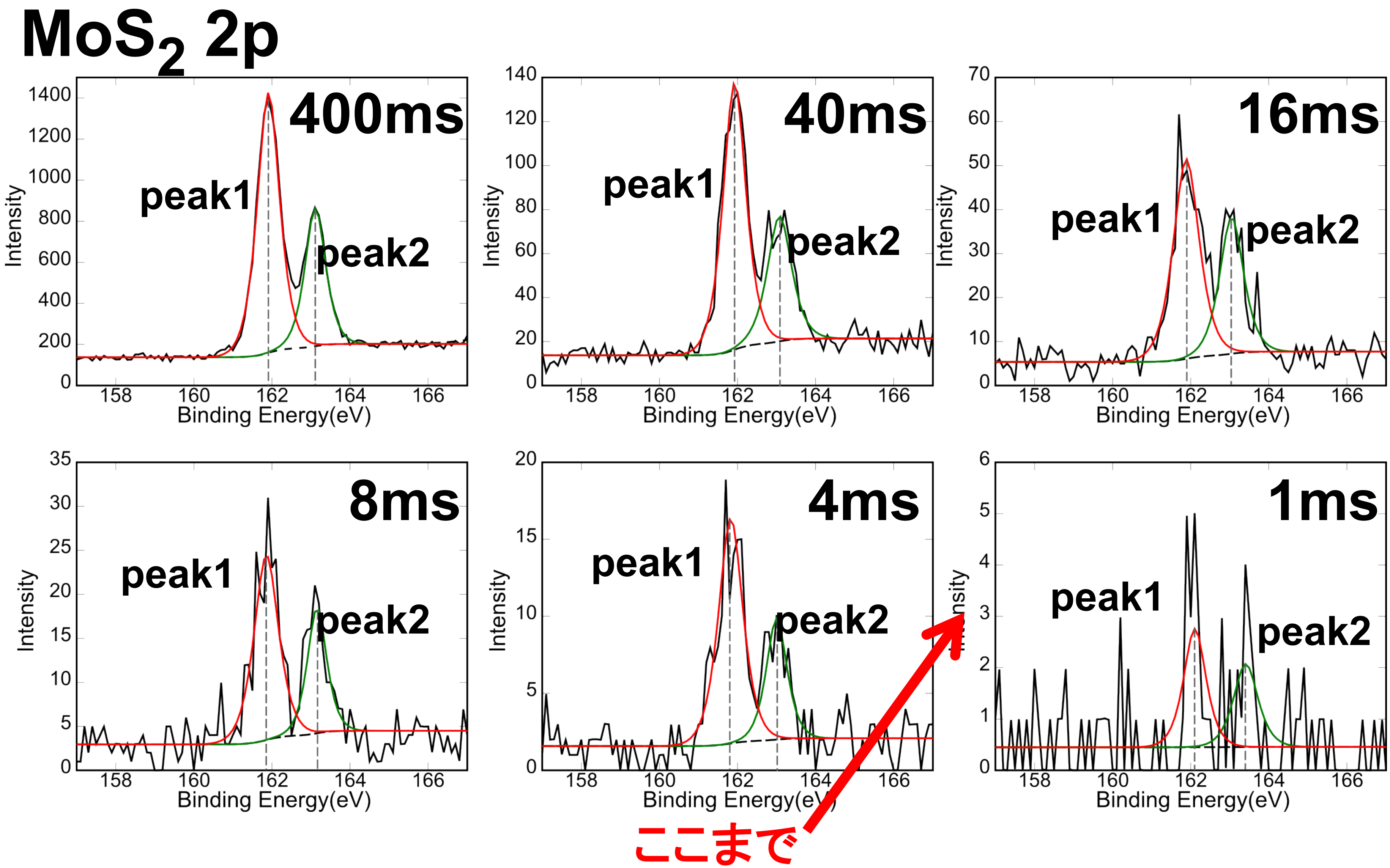






Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

例題2: 時間分解XPS(使えるかも)
どこまで時間窓を小さくできるか



共同研究者

永田賢二
岡田研
助教

村岡怜
岡田研
H23年度
修士修了

杉田精司
惑星科学

佐々木岳彦
化学

新領域創成科学研究科　複雑理工学専攻

内容

- 自己紹介:修士課程での研究
  - 希土類化合物の光電子放出・光吸収スペクトル
- 機械学習(ベイズ法)を用いたスペクトル分解
  - 確率的定式化
  - 基底関数の個数の決定
- X線光電子放出スペクトル(XPS)への適用
  - 事後分布推定→推定誤差(エラーバーが引ける!!)
  - 臨界時間窓を導出→時間分割XPSの実験計画
- 議論と今後の展開
  - 順モデルアプローチと逆モデルアプローチ
  - 超高速分光や時間分解光電子分光
  - 電子・格子ダイナミクスや固体中の非平衡電子状態

スペクトル分解の定式化

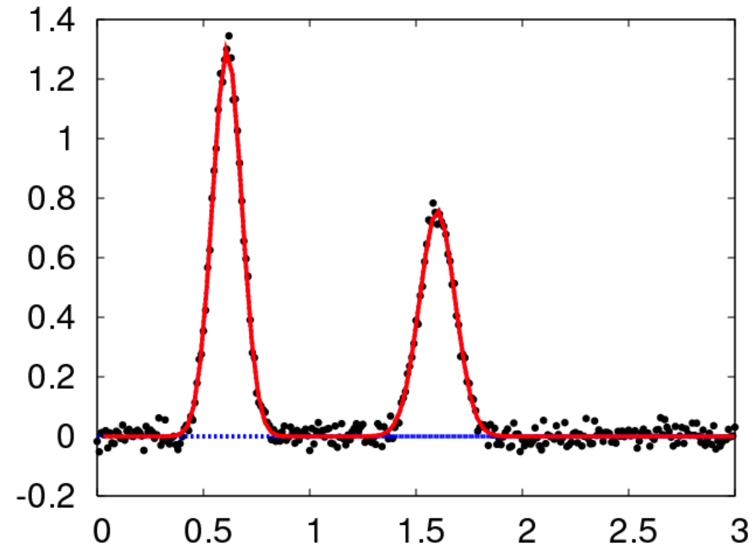
ガウス関数(基底関数)の足し合わせにより，スペクトルデータを近似

観測データ: $D = \{x_i, y_i\}_{i=1}^{n}$

x_i: **入力**　y_i: **出力**

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

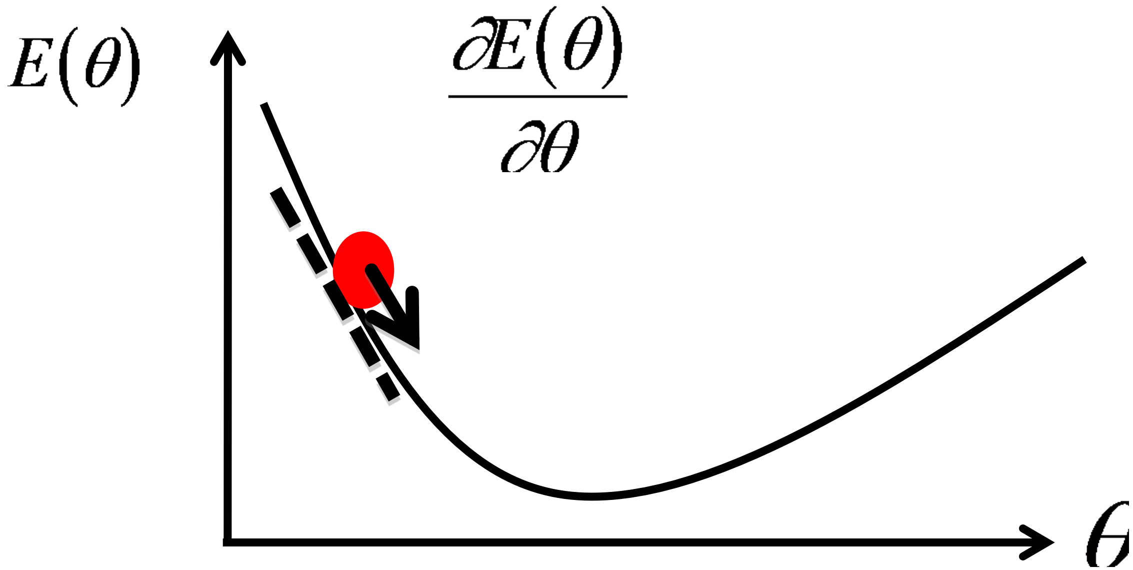
$$\theta = \{a_k, b_k, \mu_k\} \quad k = 1, \cdots, K$$



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）

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

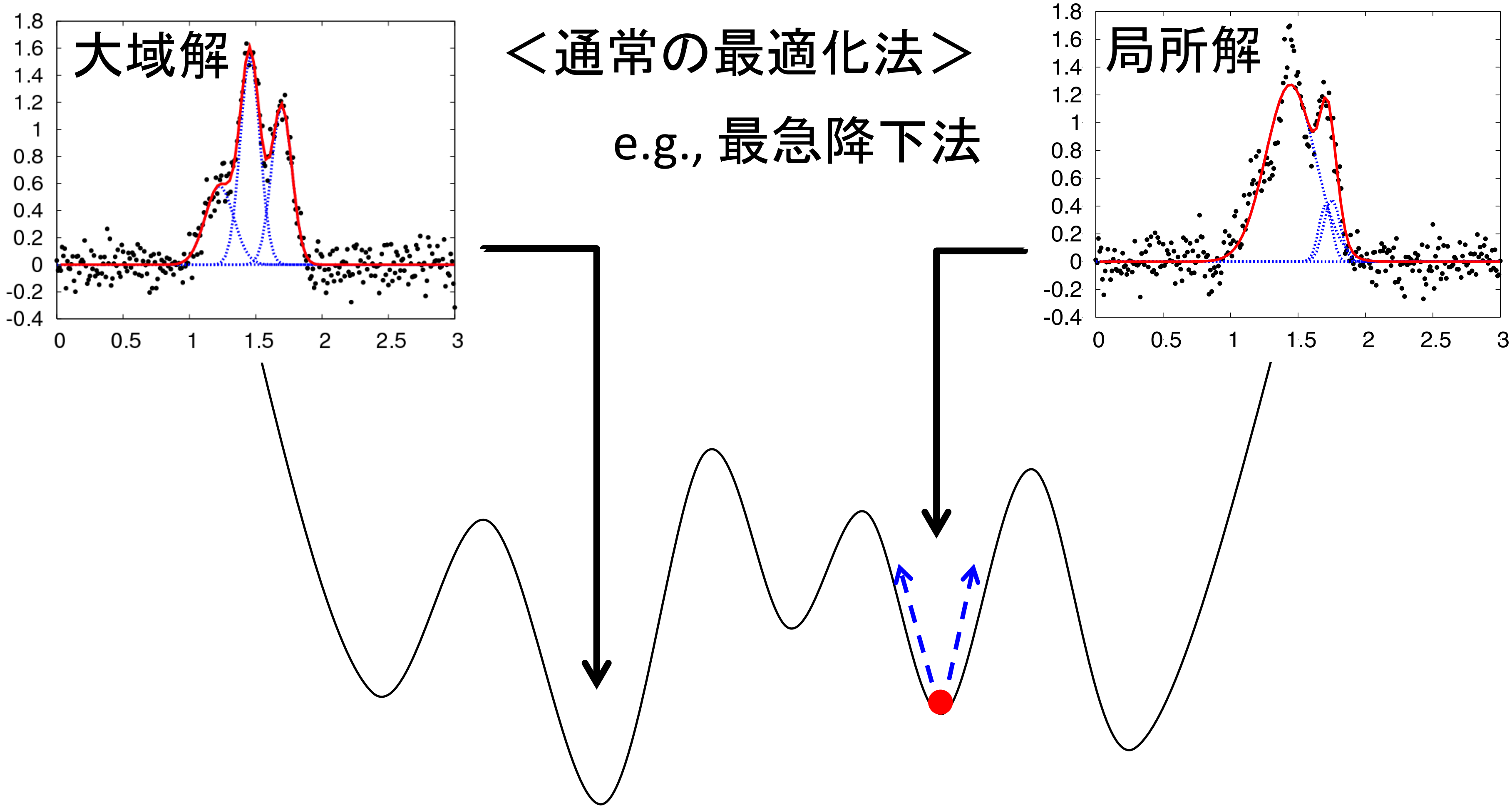
学習法:最急降下法



$$\frac{d\theta}{dt} = -\frac{\partial E(\theta)}{\partial \theta}$$

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

誤差関数は局所解を持つ

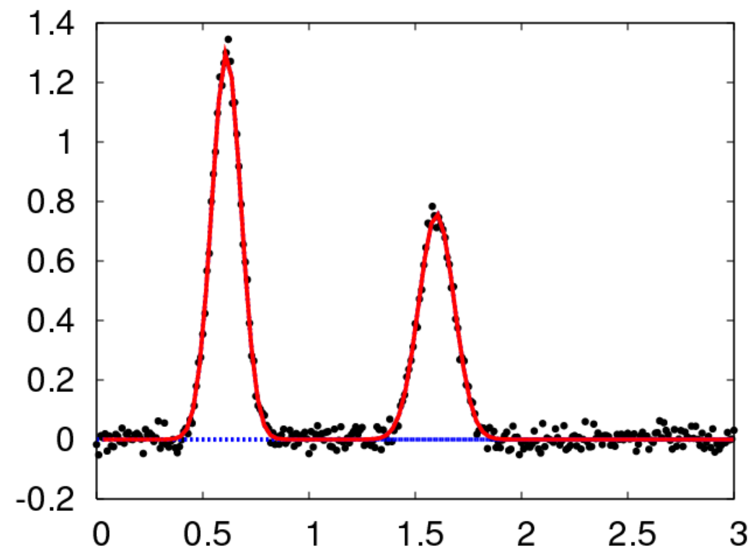


確率的定式化

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，

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

それぞれの出力y_iが，独立であるとすると，

$$p(Y \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta) \propto \exp(-nE(\theta)) \qquad Y = \{y_1, \cdots, y_n\}$$

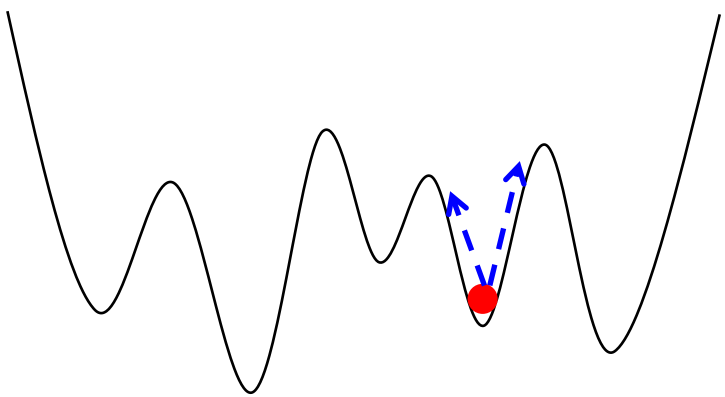
$$E(\theta) = \frac{1}{2n}\sum_{i=1}^{n}\left(y_i - f(x_i;\theta)\right)^2$$

ボルツマン分布

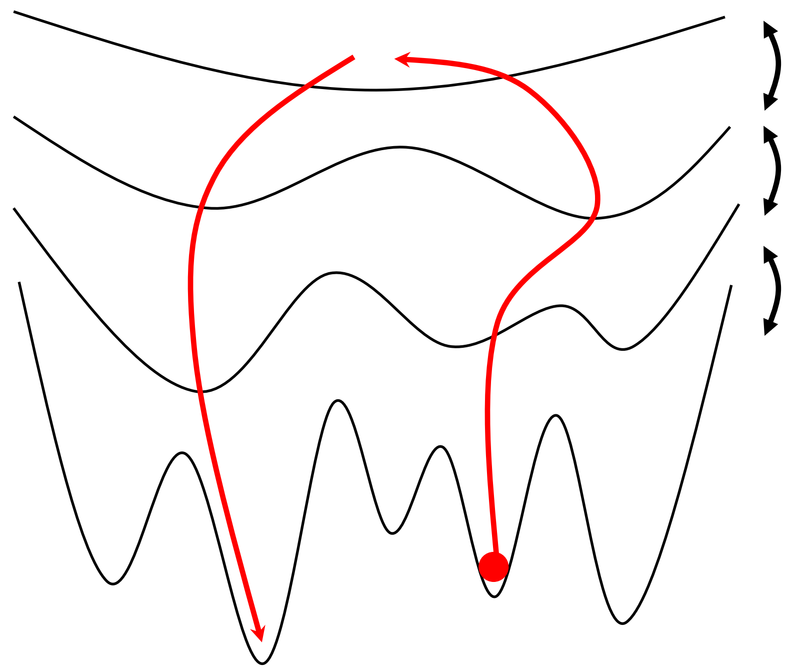
レプリカ交換モンテカルロ法
ランダムスピン系から

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

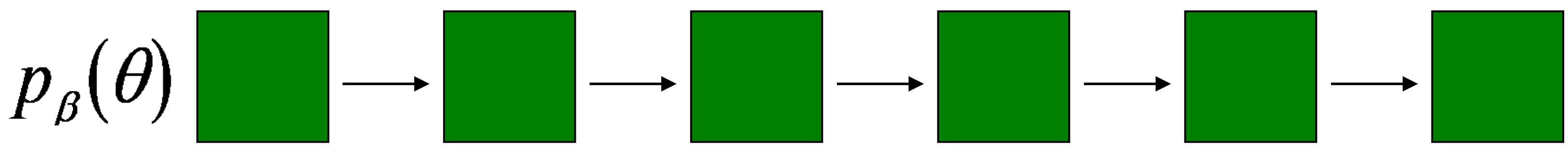


K. Hukushima, K. Nemoto, *J. Phys. Soc. Jpn.* **65** (1996).

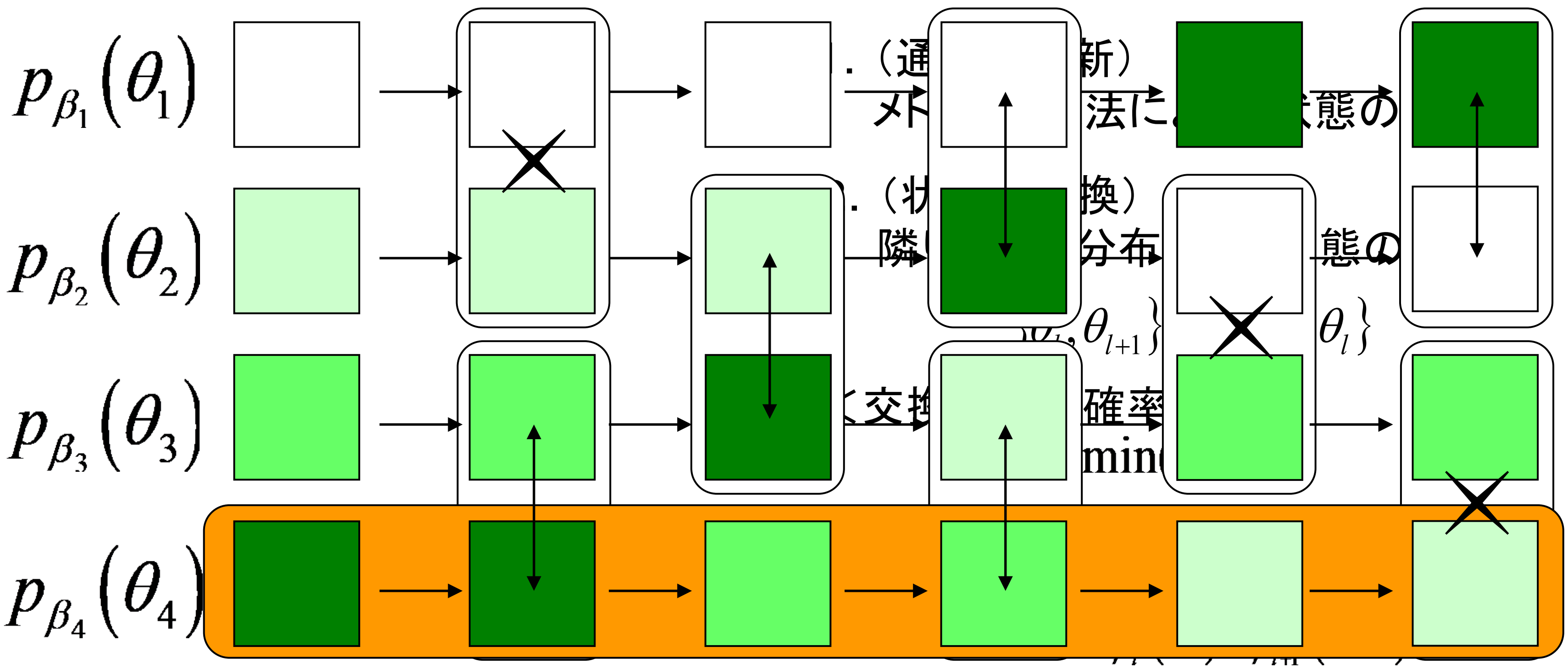
情報統計力学(Statistical Mechanical Informatics)へ

レプリカ交換モンテカルロ法のイメージ

<メトロポリス法>

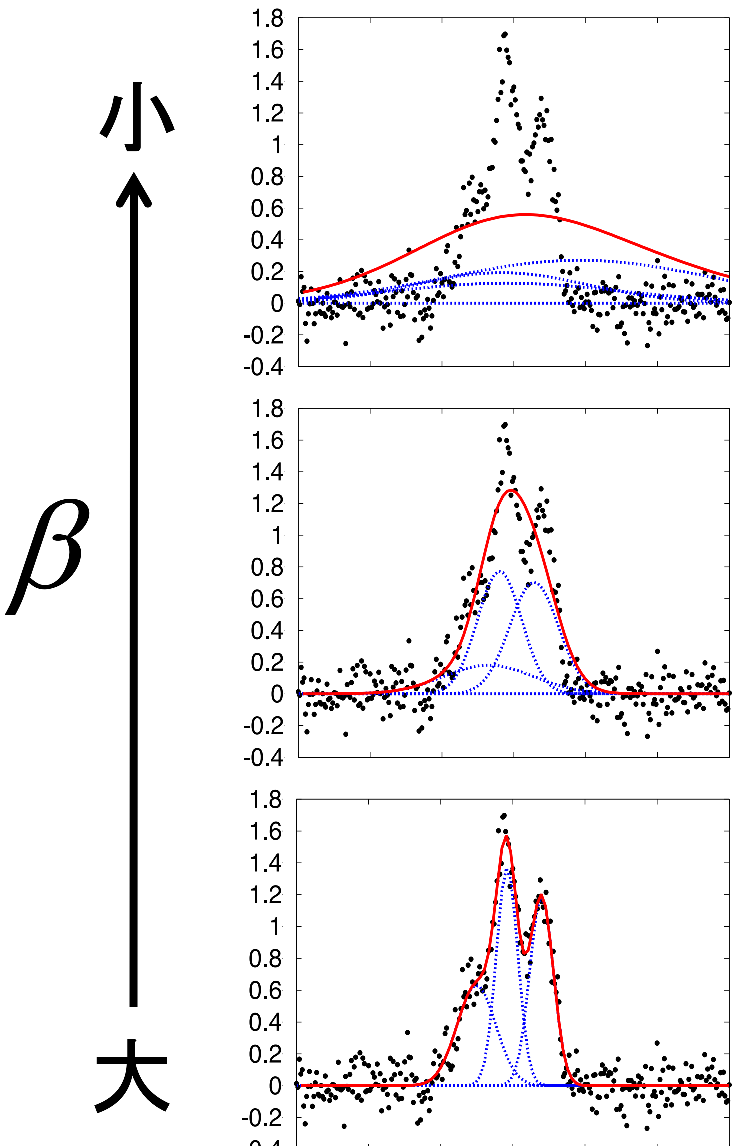


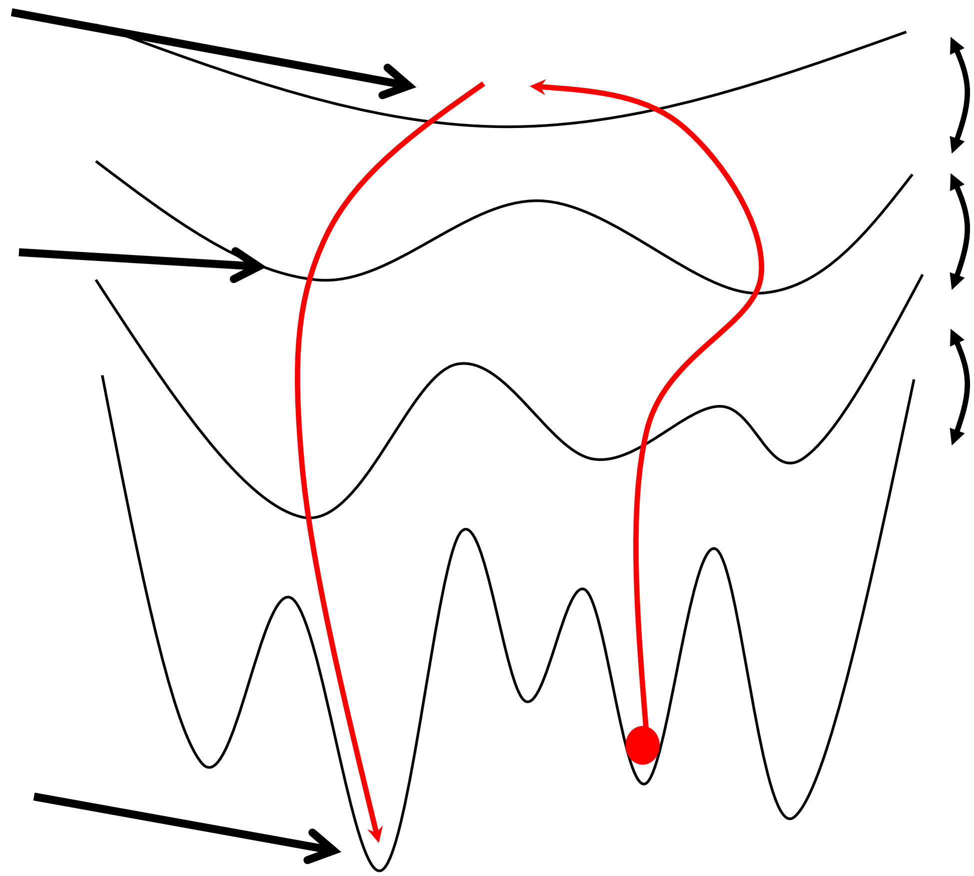
<レプリカ交換モンテカルロ法>



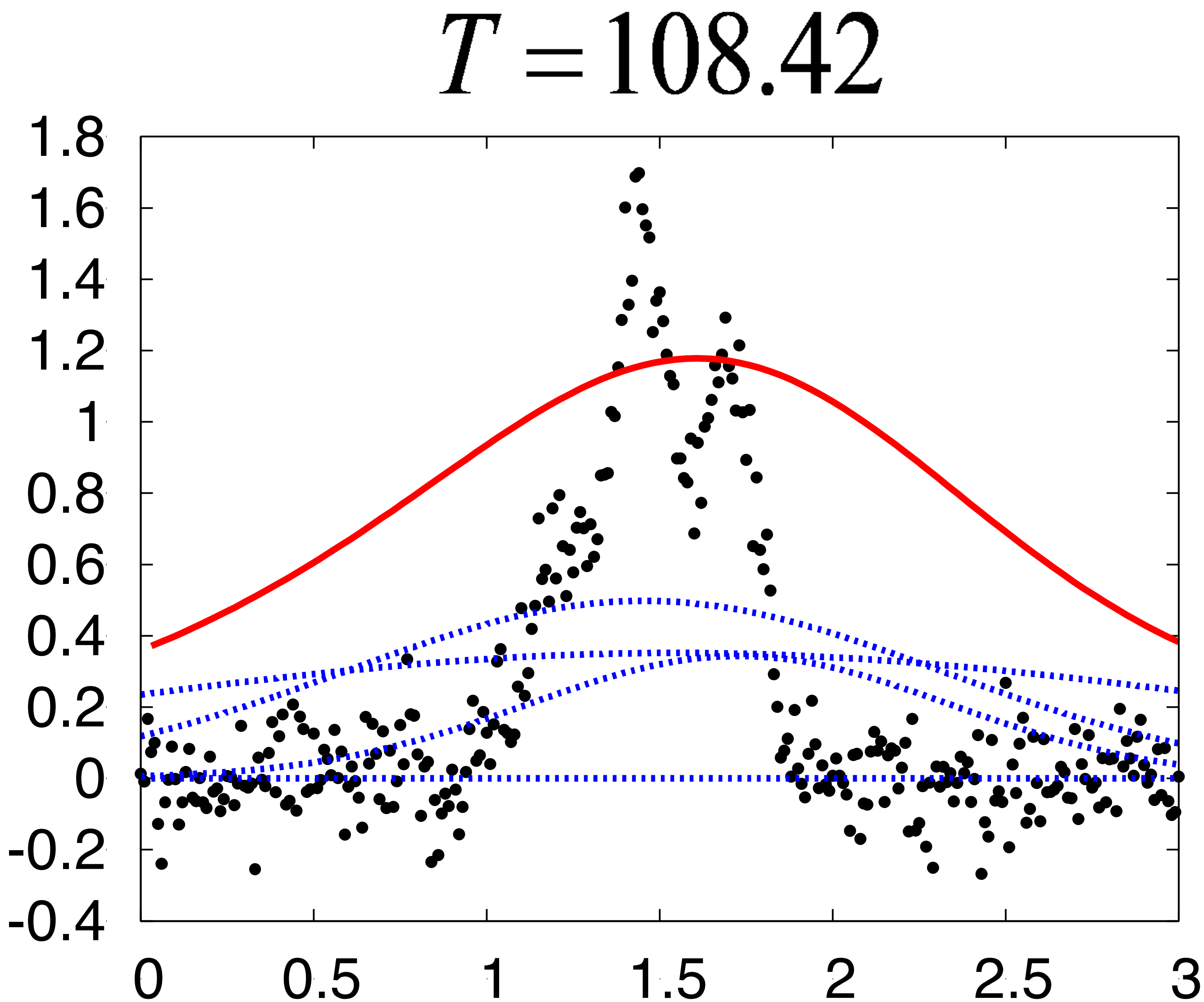
レプリカ交換モンテカルロ法

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

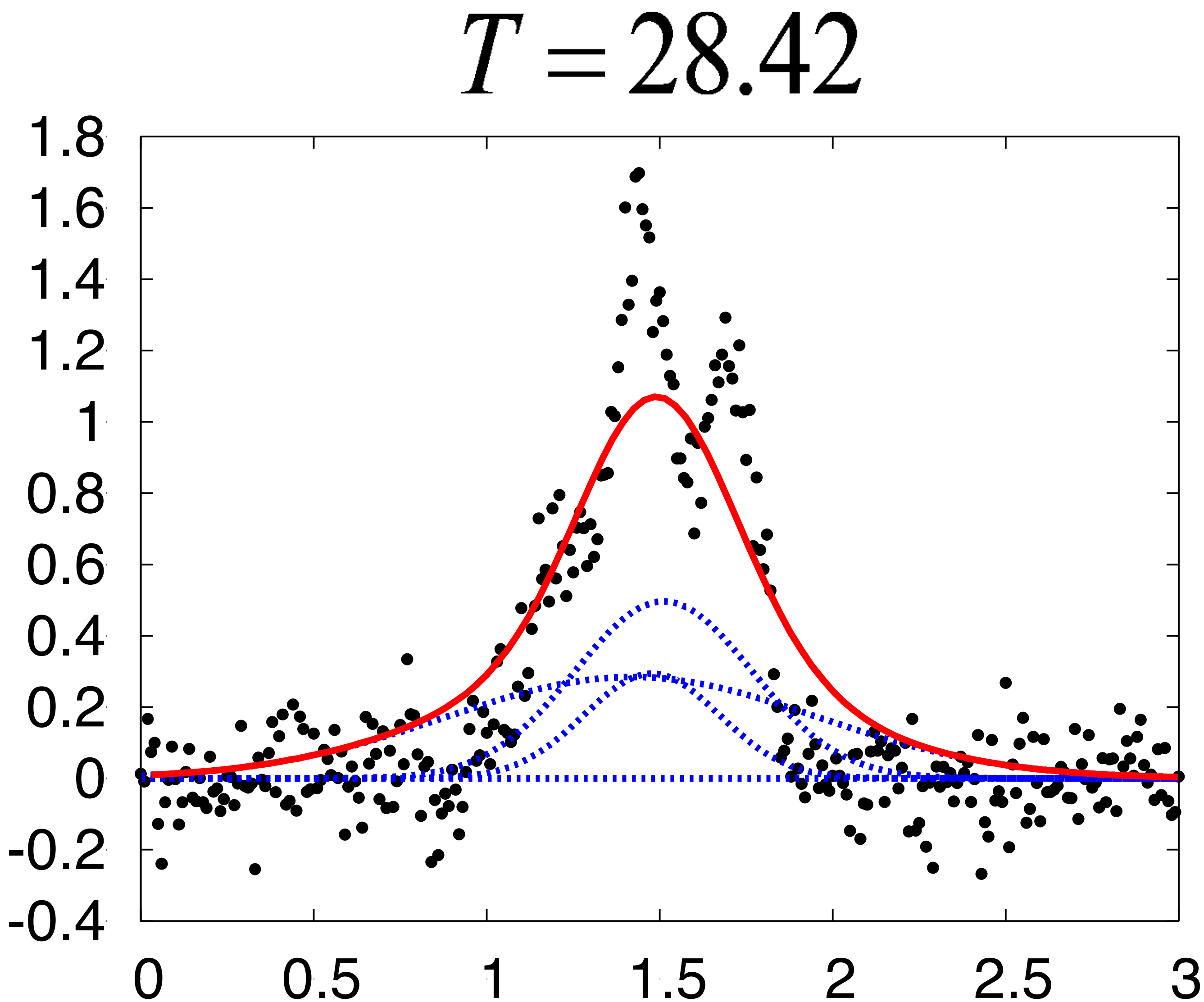




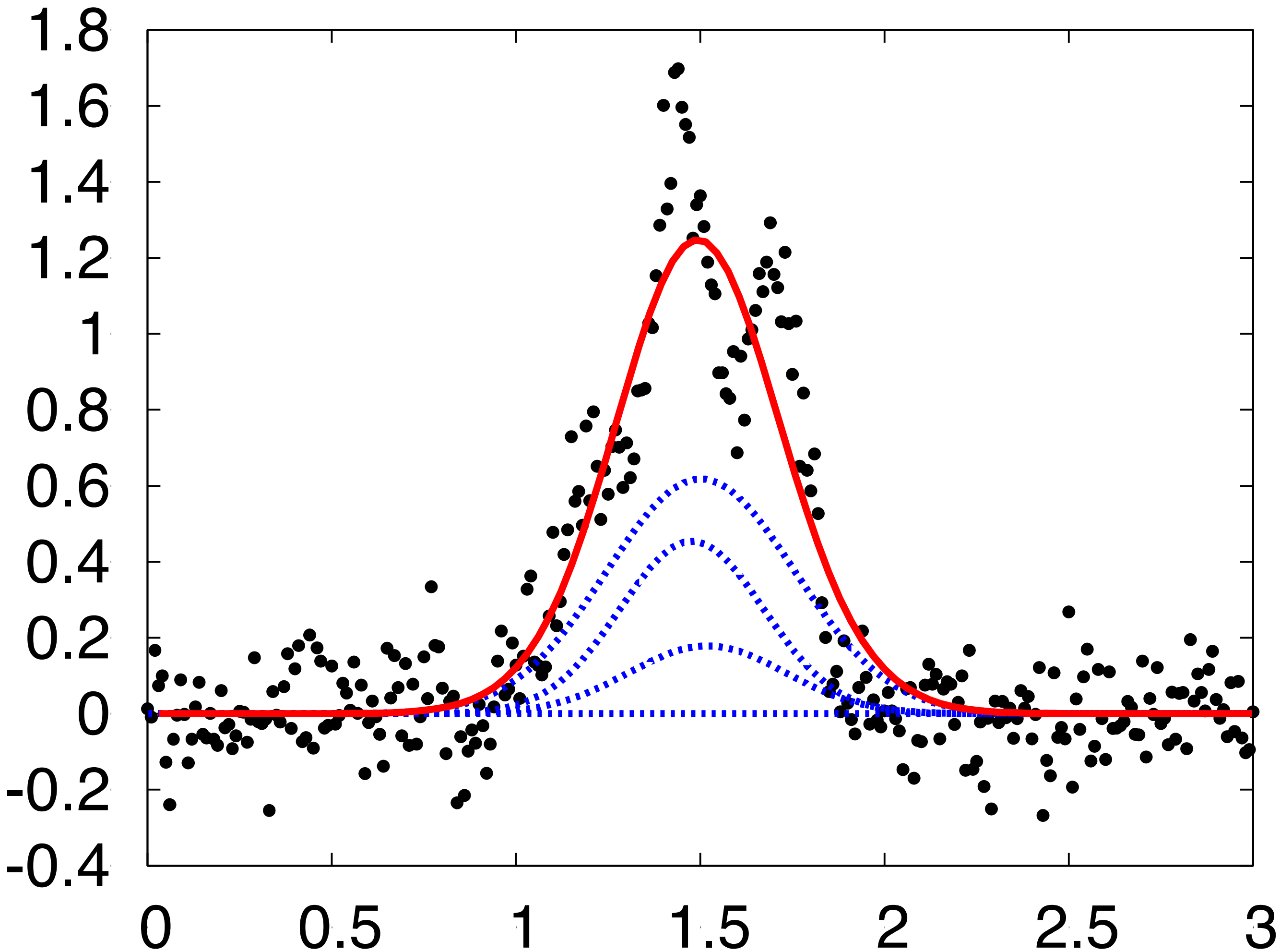
[Hukushima and Nemoto, 1996]

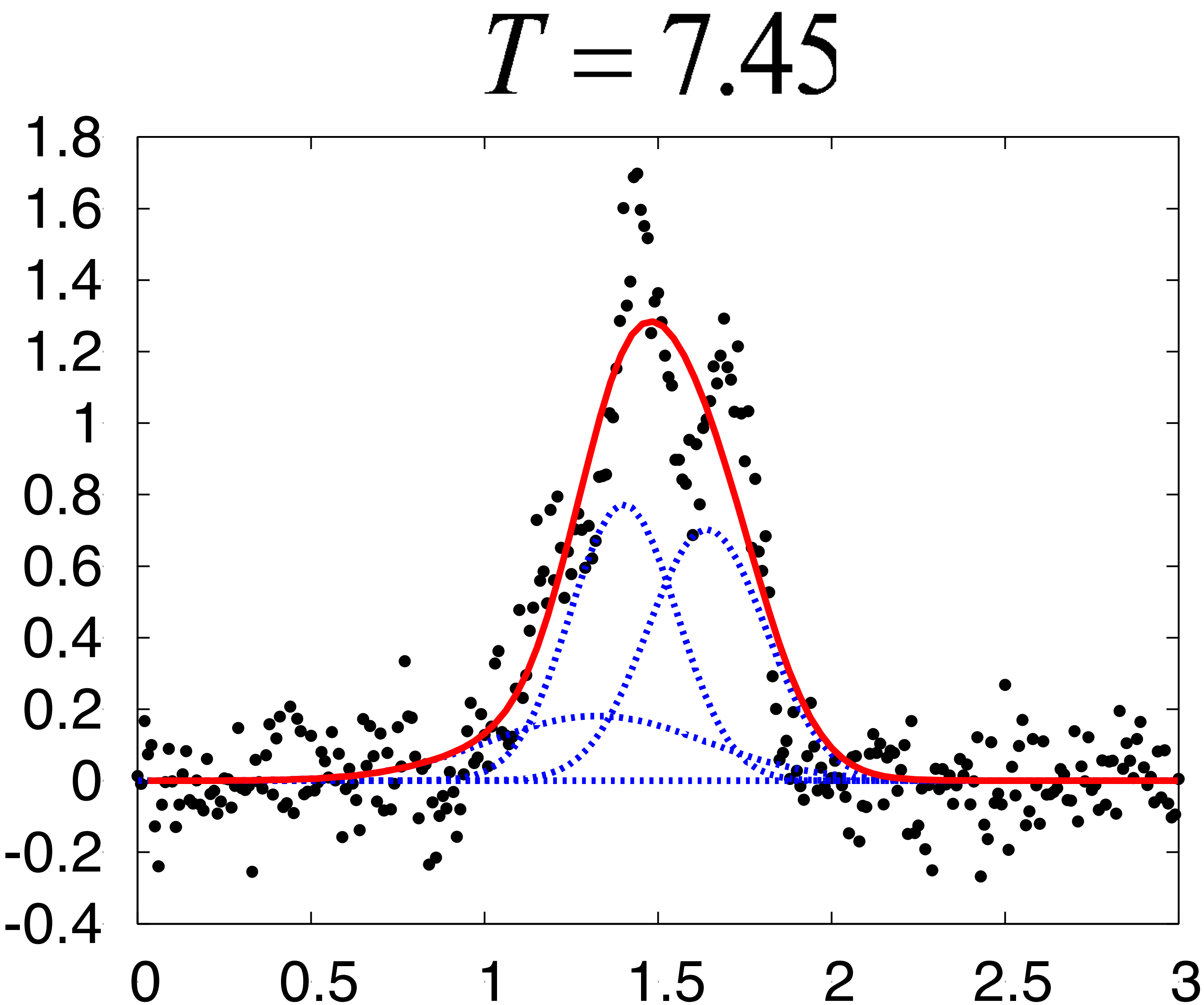


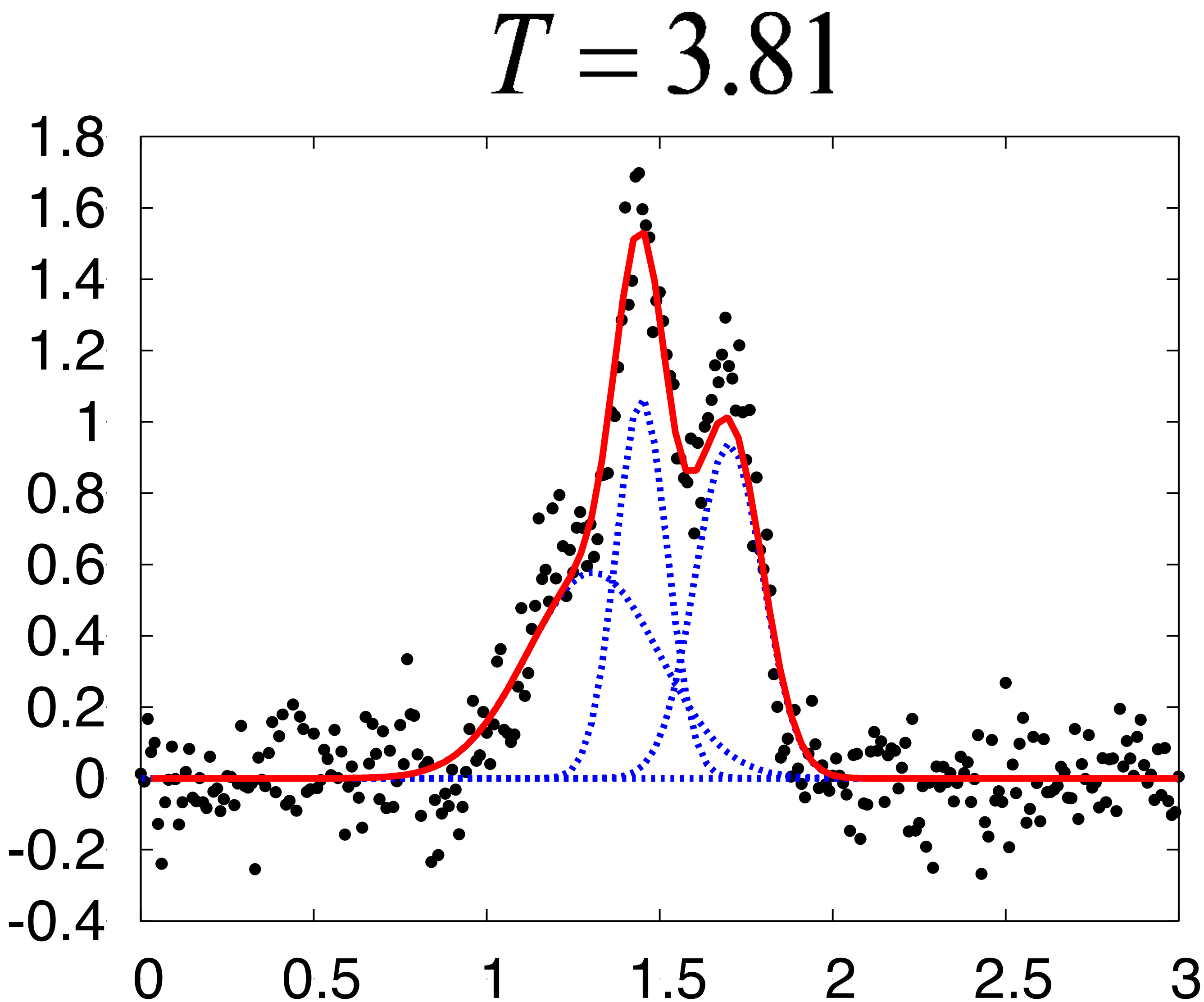
$T = 108.42$
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.5
1
1.5
2
2.5
3


$T = 28.42$
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.5
1
1.5
2
2.5
3

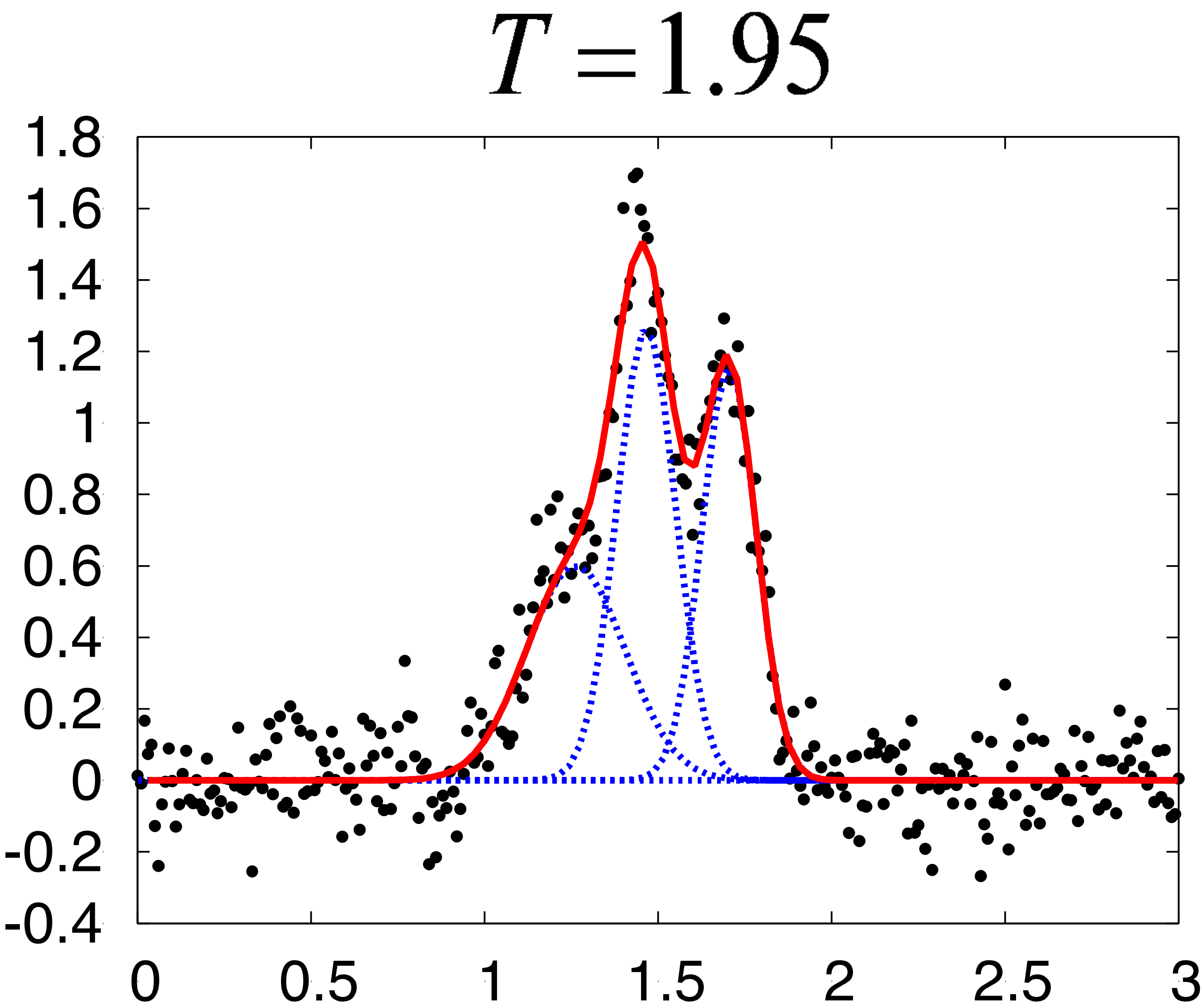
$T = 14.55$



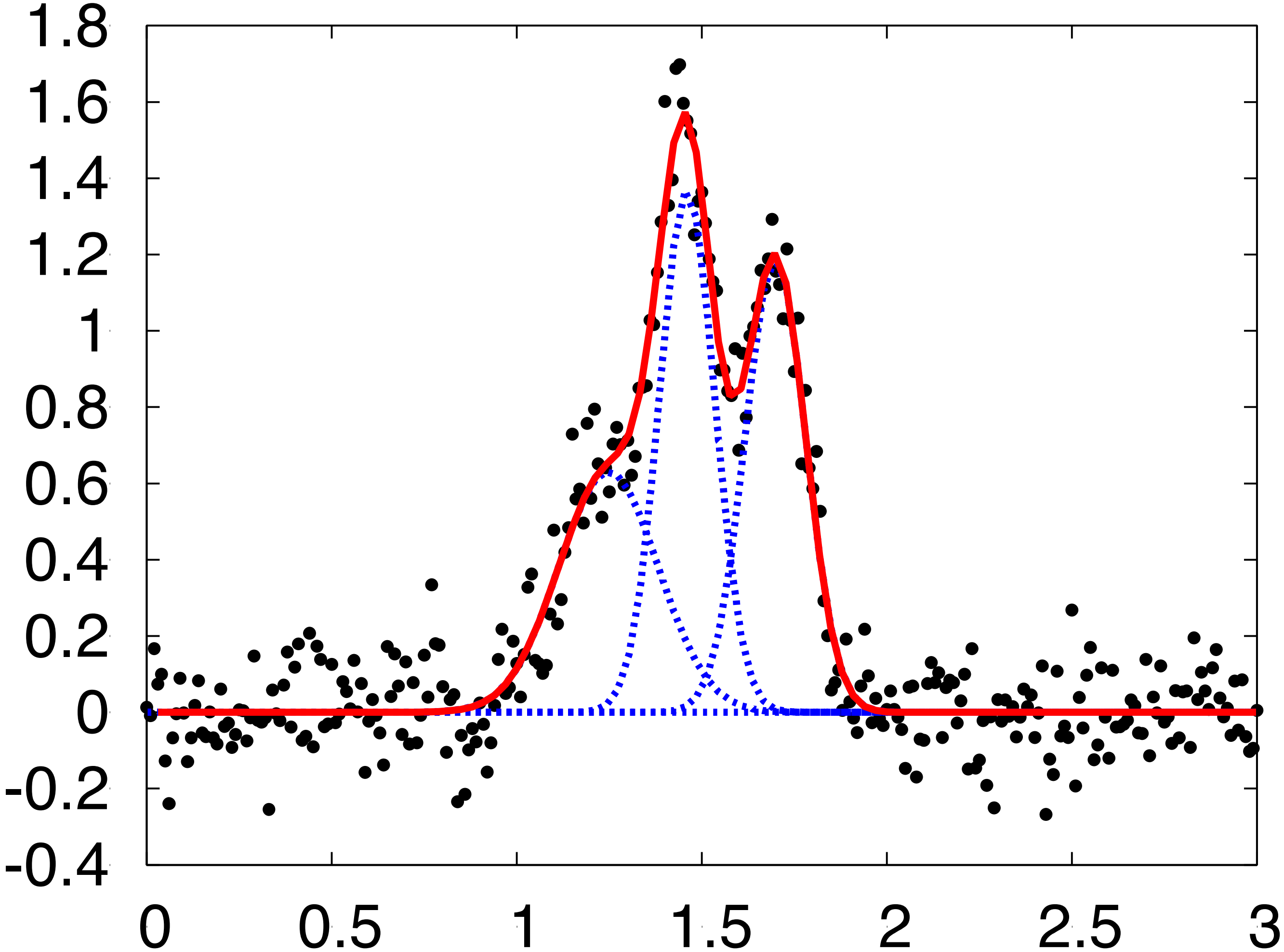

$T = 7.45$
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.5
1
1.5
2
2.5
3



$T = 3.81$
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.5
1
1.5
2
2.5
3


$T = 1.95$
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
0
0.5
1
1.5
2
2.5
3

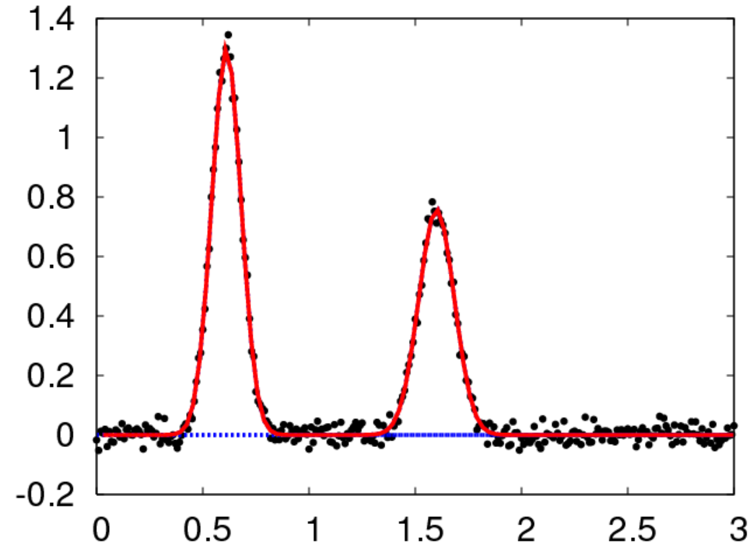
$$T = 1.00$$



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，

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

それぞれの出力 y_i が，独立であるとすると，$p(Y \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta)$

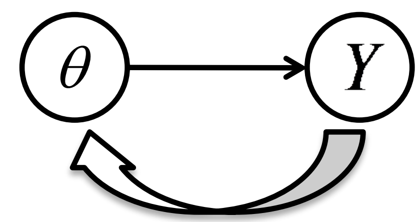
$$Y = \{y_1, \cdots, y_n\}$$

ベイズ推論:因果律を組み込んでデータ解析

観測データ Y だけでなく，パラメータ θ も確率変数とみなす．

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

生成（因果律）



$$\Downarrow$$

＜ベイズの定理＞

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

$p(\theta\,|\,Y)$: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。
これまで蓄積されてきた科学的知見

対数損失関数とエネルギー
ベイズ推論と統計物理は数理的に等価

対数損失関数：パラメータ θ に対する評価関数

$$
\begin{aligned}
L(\theta) &= -\log p(Y \mid \theta) \\
&= \frac{1}{2}\sum_{i=1}^{n}\left(y_i - f(x_i;\theta)\right)^2 - \frac{n}{2}\log(2\pi) \\
&= nE(\theta) + \text{Const.}
\end{aligned}
$$

$$
\Rightarrow \quad p(Y \mid \theta) \propto \exp(-nE(\theta))
$$

確率的情報処理への統計力学的アプローチ(2002-2005)
東北大　田中准教授(教授)

Statistical-Mechanical Approach to Probabilistic Information Processing

共通の数理

$$m_i = \tanh(h_i + \sum_j J_{ij} m_j) \qquad \beta = \frac{1}{2}\ln\left(\frac{1-p}{p}\right)$$

ボルツマン分布

統計力学

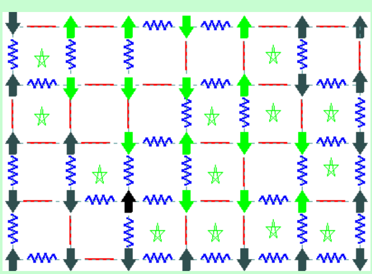
確率分布

情報科学

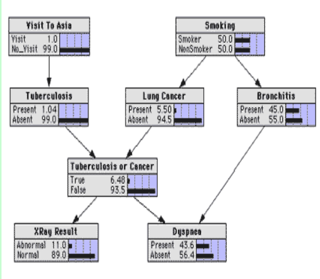
$$P(\boldsymbol{s}|\boldsymbol{h}) = \frac{P(\boldsymbol{h}|\boldsymbol{s})P(\boldsymbol{s})}{P(\boldsymbol{h})}$$

ベイズ統計

物性の理解・予言



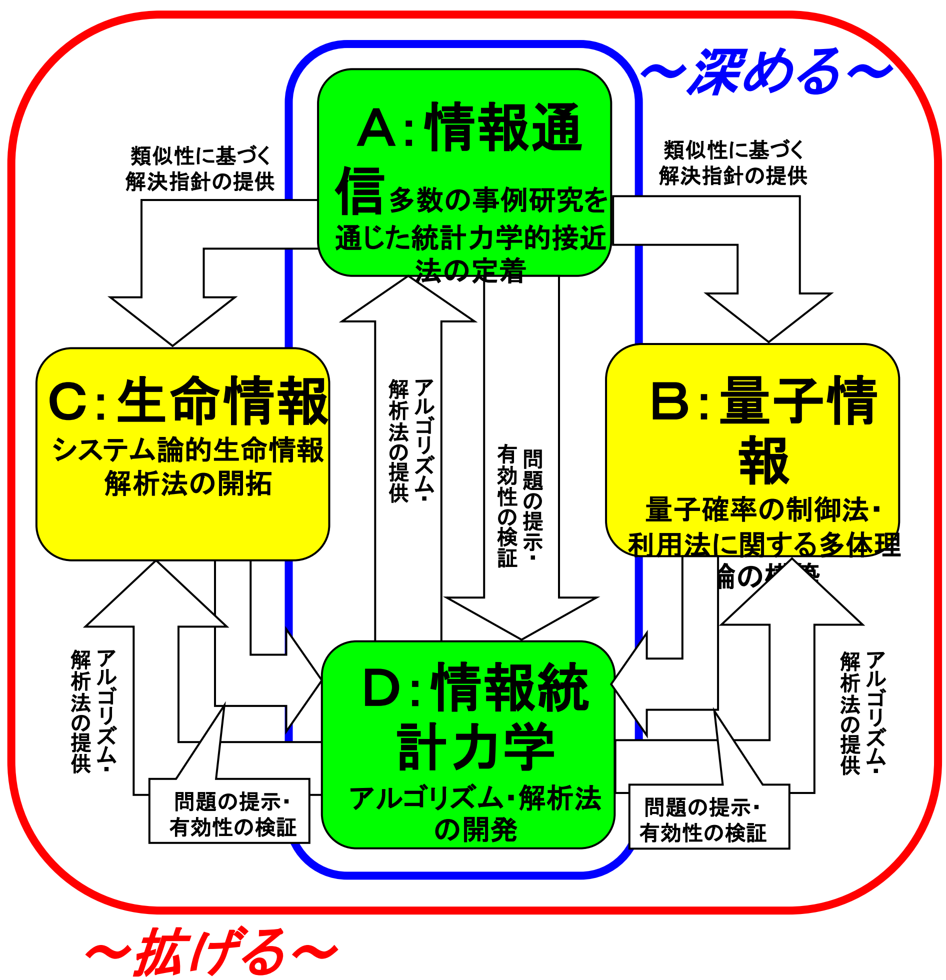
情報の抽出・加工



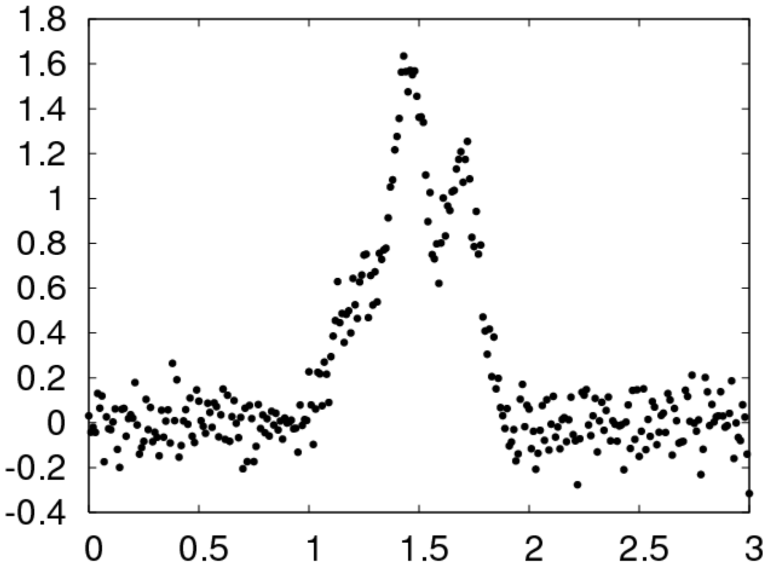
情報統計力学の深化と展開
東工大　樺島教授　(2006-2009)

- 目覚しい発展を遂げている情報科学への統計力学的接近法（以下，情報統計力学）を深化させ，有望な新領域に展開する．
  - 深化軸：情報通信，情報統計力学
  - 展開軸：量子情報，生命情報
- 以上を通じて，当該分野における世界的拠点としての揺るぎない地位を我が国に確立する．

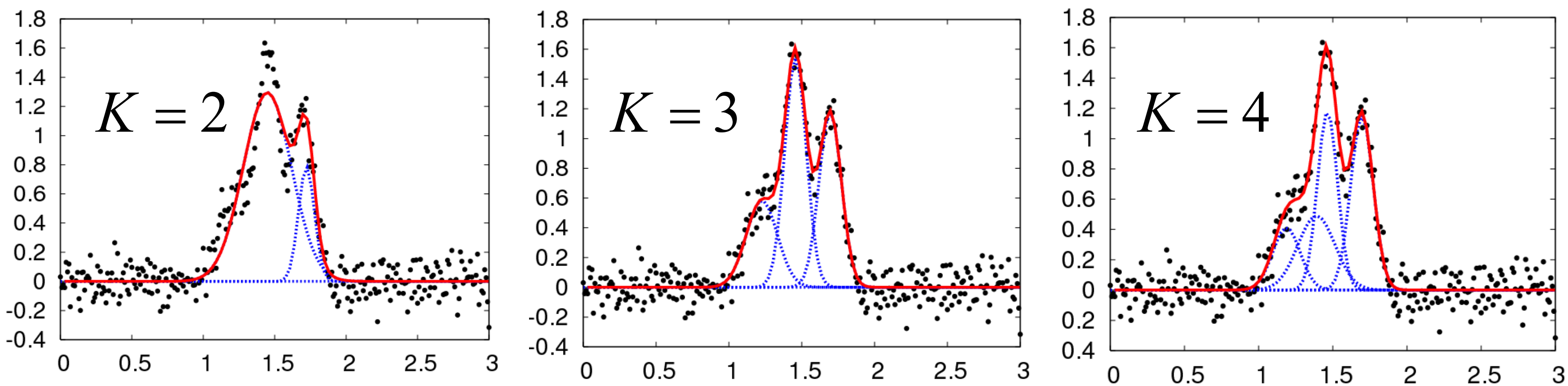
本領域のねらい



それぞれのKについて最適化完了



Kをどう選ぶか



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

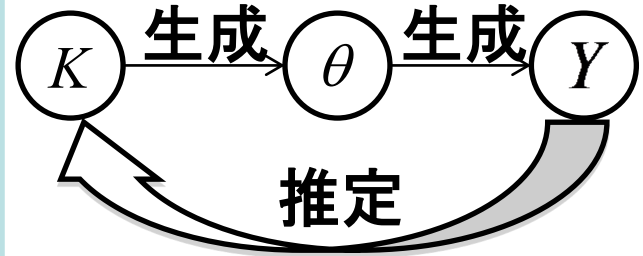
ベイズ推論による逆アプローチ

目的: $p(K|Y)$ を最大にする K の選択

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

$$\propto p(K)\int p(Y|K,\theta)p(\theta)d\theta$$

$$\propto p(K)\int \exp(-nE(\theta))p(\theta)d\theta$$



K:ピーク個数
θ:ピーク位置など
Y:観測スペクトル

分配関数と等価(情報統計力学) $E(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - f(x_i;\theta)\right)^2$

「自由エネルギー，確率的複雑さ」

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

自由エネルギーを最小にする個数 K を求める.

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

自由エネルギー: $F=-\log\int\exp\left(-\frac{n}{\sigma^2}E(\theta)\right)p(\theta)d\theta$

以下のように、補助変数 β を導入する。　　　β: 逆温度

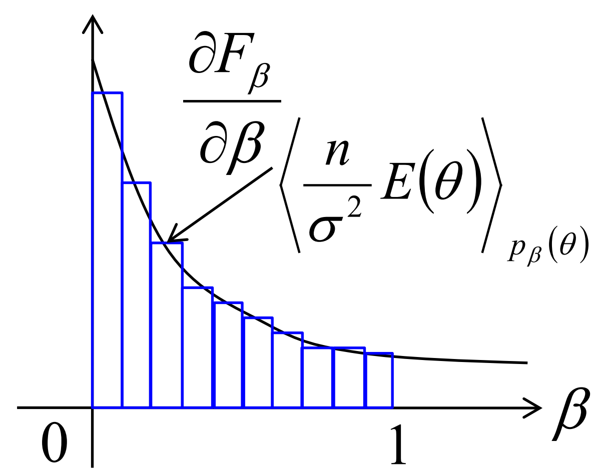
$$F_\beta=-\log\int\exp\left(-\frac{n}{\sigma^2}\beta E(\theta)\right)p(\theta)d\theta\ \left(F_{\beta=0}=0\right)$$

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

→ **たくさんの温度でのシミュレーションが必要**
各温度でのエネルギー平均(すでにやってる)

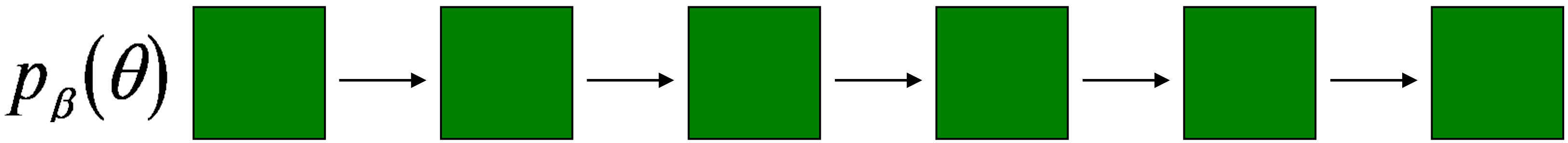
$\frac{\partial F_\beta}{\partial\beta}\cdots$ 確率分布 $p(\theta;\beta)$ に従う二乗誤差 $\frac{n}{\sigma^2}E(\theta)$ の期待値

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

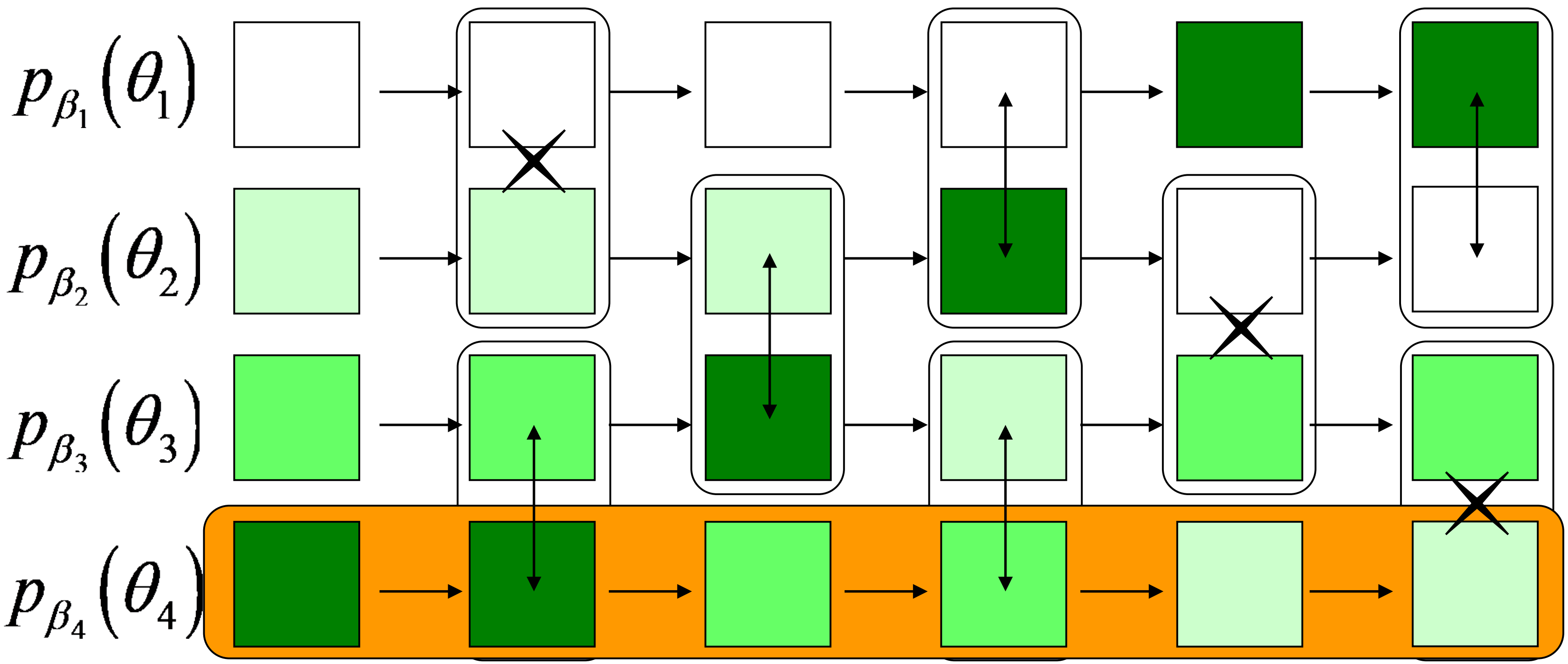


レプリカ交換モンテカルロ法のイメージ

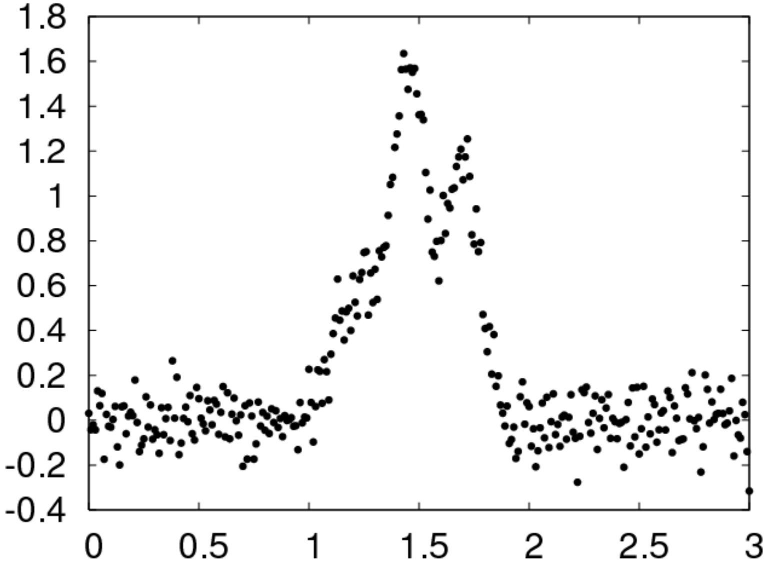
<メトロポリス法>

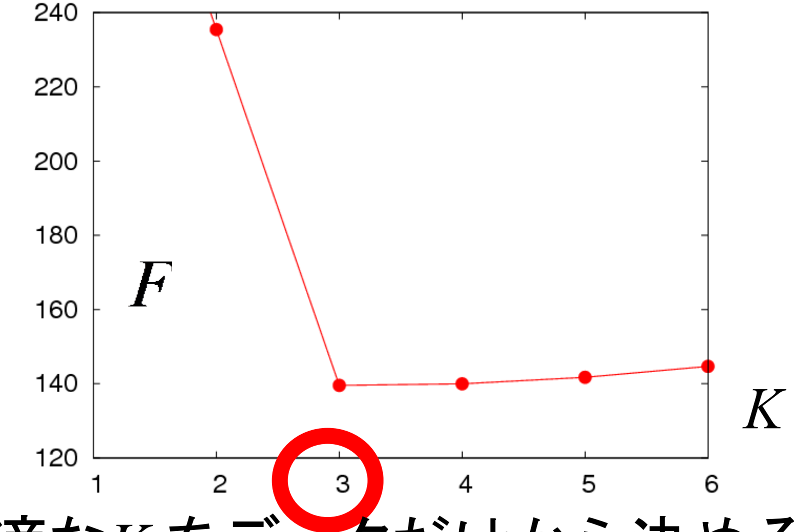


<レプリカ交換モンテカルロ法>

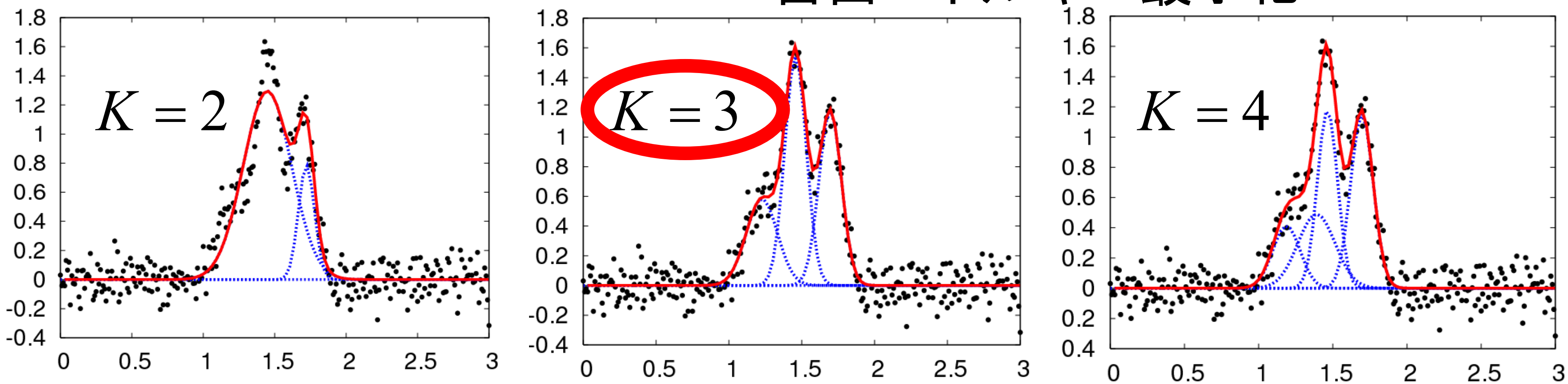


問題1 マルチピークスペクトルの最適分解





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



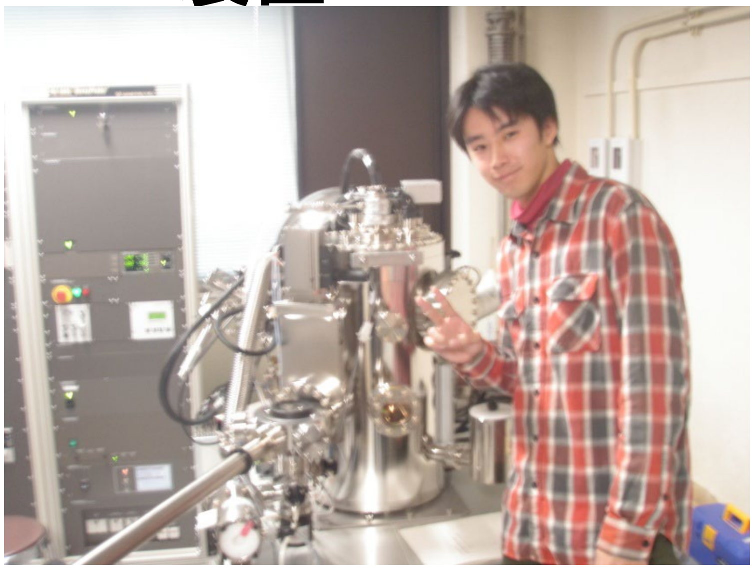
Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

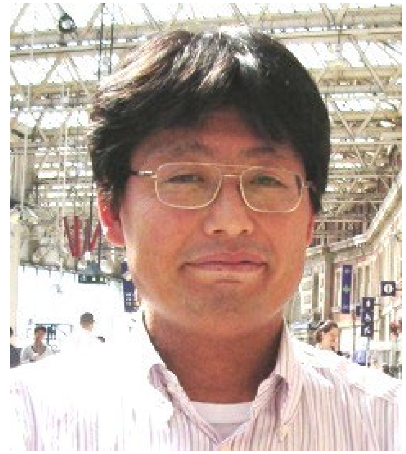
内容

- 自己紹介：修士課程での研究
  - 希土類化合物の光電子放出・光吸収スペクトル
- 機械学習（ベイズ法）を用いたスペクトル分解
  - 確率的定式化
  - 基底関数の個数の決定
- X線光電子放出スペクトル(XPS)への適用
  - 事後分布推定→推定誤差(エラーバーが引ける!!)
  - 臨界時間窓を導出→時間分割XPSの実験計画
- 議論と今後の展開
  - 順モデルアプローチと逆モデルアプローチ
  - 超高速分光や時間分解光電子分光
  - 電子・格子ダイナミクスや固体中の非平衡電子状態

XPS測定

XPS装置





佐々木岳彦

試料

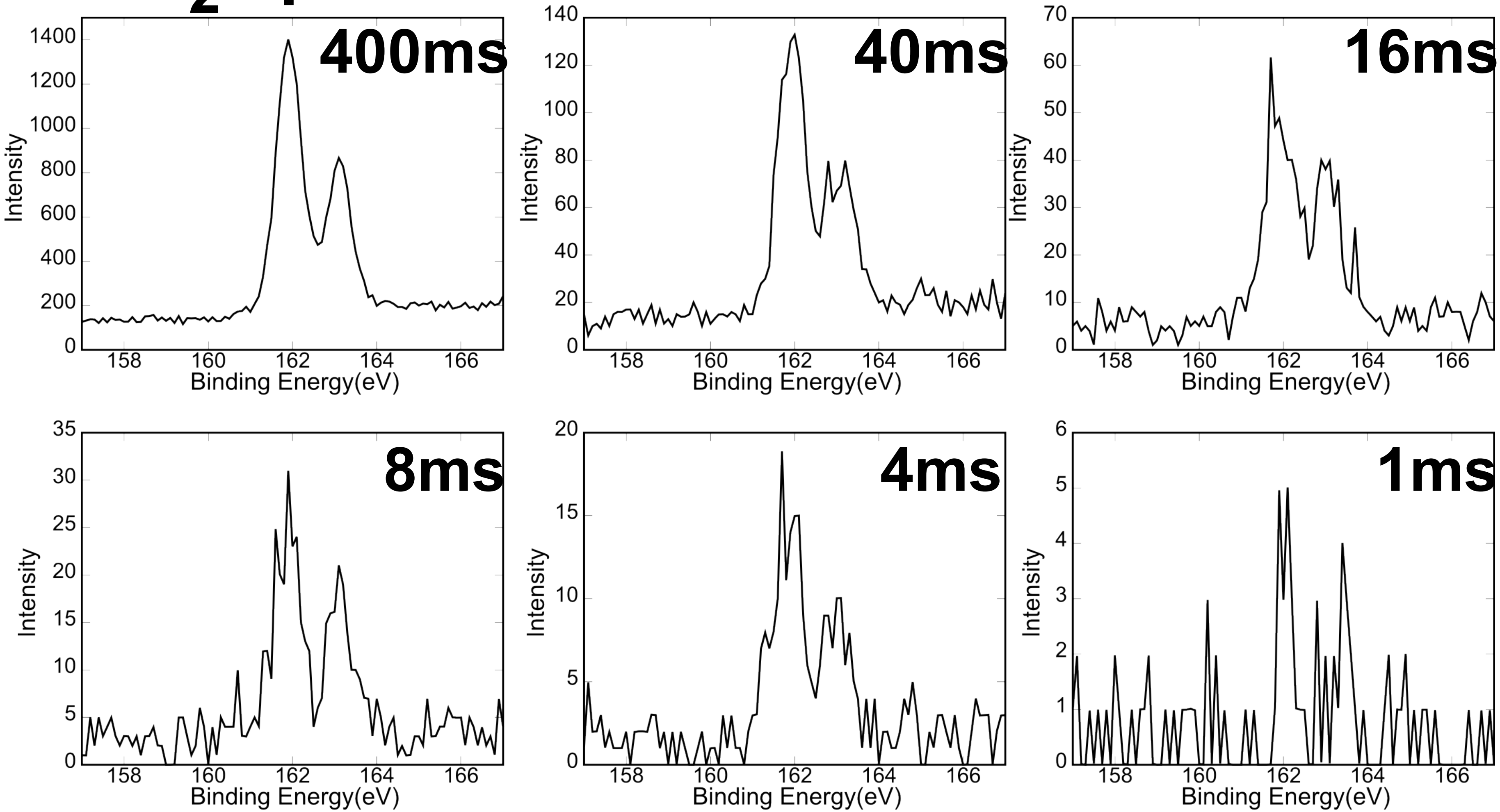
試料は粉末の二硫化モリブデン(MoS_2)
カーボンテープ上に固定
X線源はモノクロAlKα
電子銃による帯電補正

1ステップあたりの測定時間は 400ms,40ms,16ms,8ms,4ms,1ms の6条件で測定

東大工学部9号館・強力X線実験室

測定時間とスペクトル
どこまで時間窓を小さくできるか？

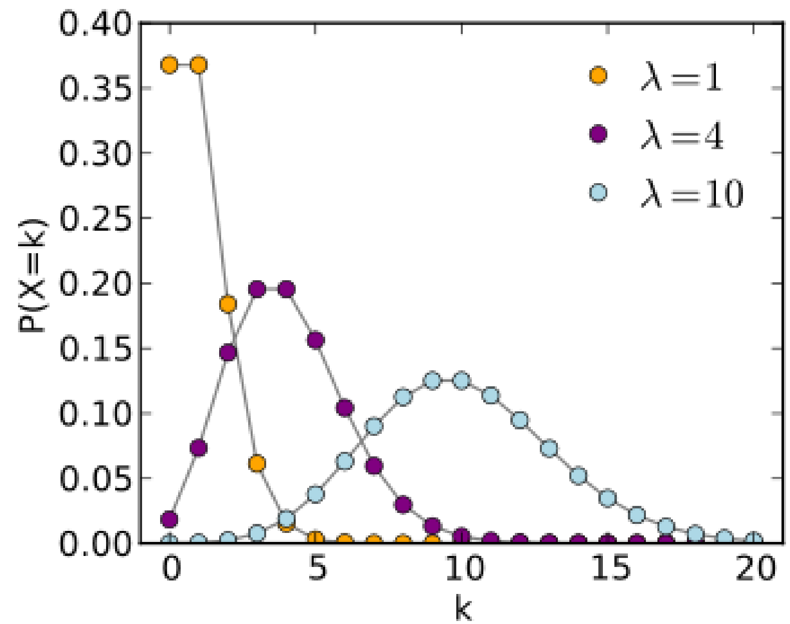
MoS_2 2p



XPS測定におけるノイズ

XPS測定で，データに乗るノイズはポアソン分布に従う

$$P(I=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$



ポアソン分布では，平均＝分散＝λ

$$\Rightarrow \frac{\text{信号強度}}{\text{ノイズ}}=\frac{\mathrm{I(E)}}{\sqrt{\mathrm{I(E)}}}$$

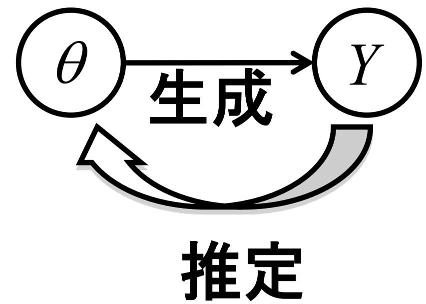
となるので，信号強度が小さくなるほどノイズの大きさが相対的に大きくなる

では，どこまで信号強度が下がっても，真のピークを推定できるのかということが，今回明らかにしたい点である

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: $p(\theta|Y)=\dfrac{p(Y|\theta)p(\theta)}{p(Y)}$ $\quad Y=\{y_i\}_{i=1}^{n}$



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

θ :ピーク位置など
Y :観測スペクトル

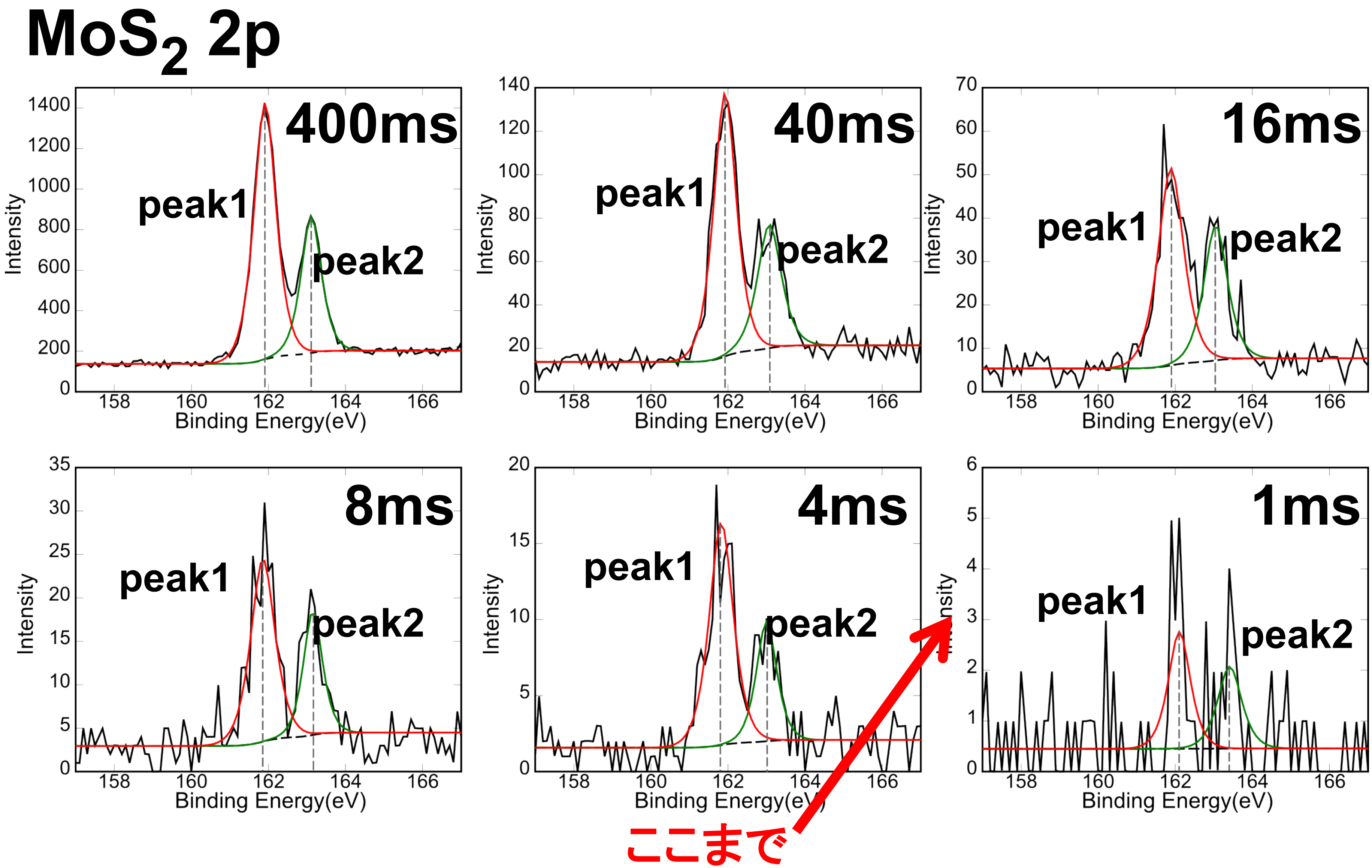
これまでの

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

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

に変更するだけ

問題2 時間分解XPS
どこまで時間窓を小さくできるか

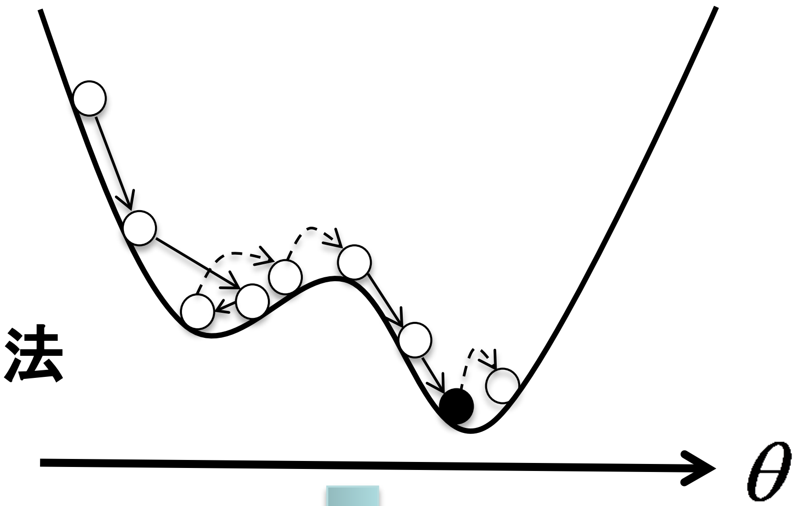


MCMC法によるヒストグラム作成

誤差関数

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

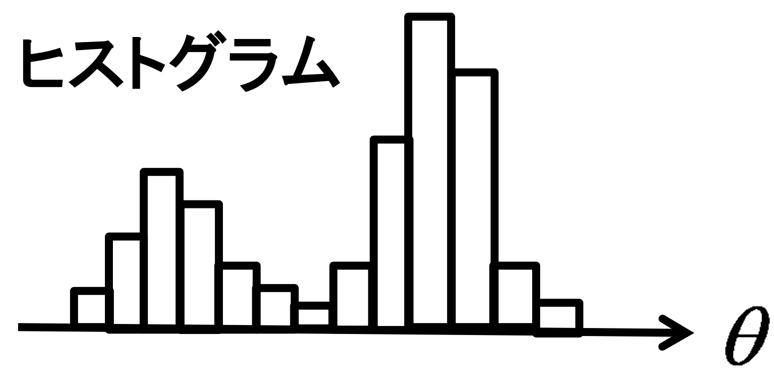
MCMC法



サンプルを保持

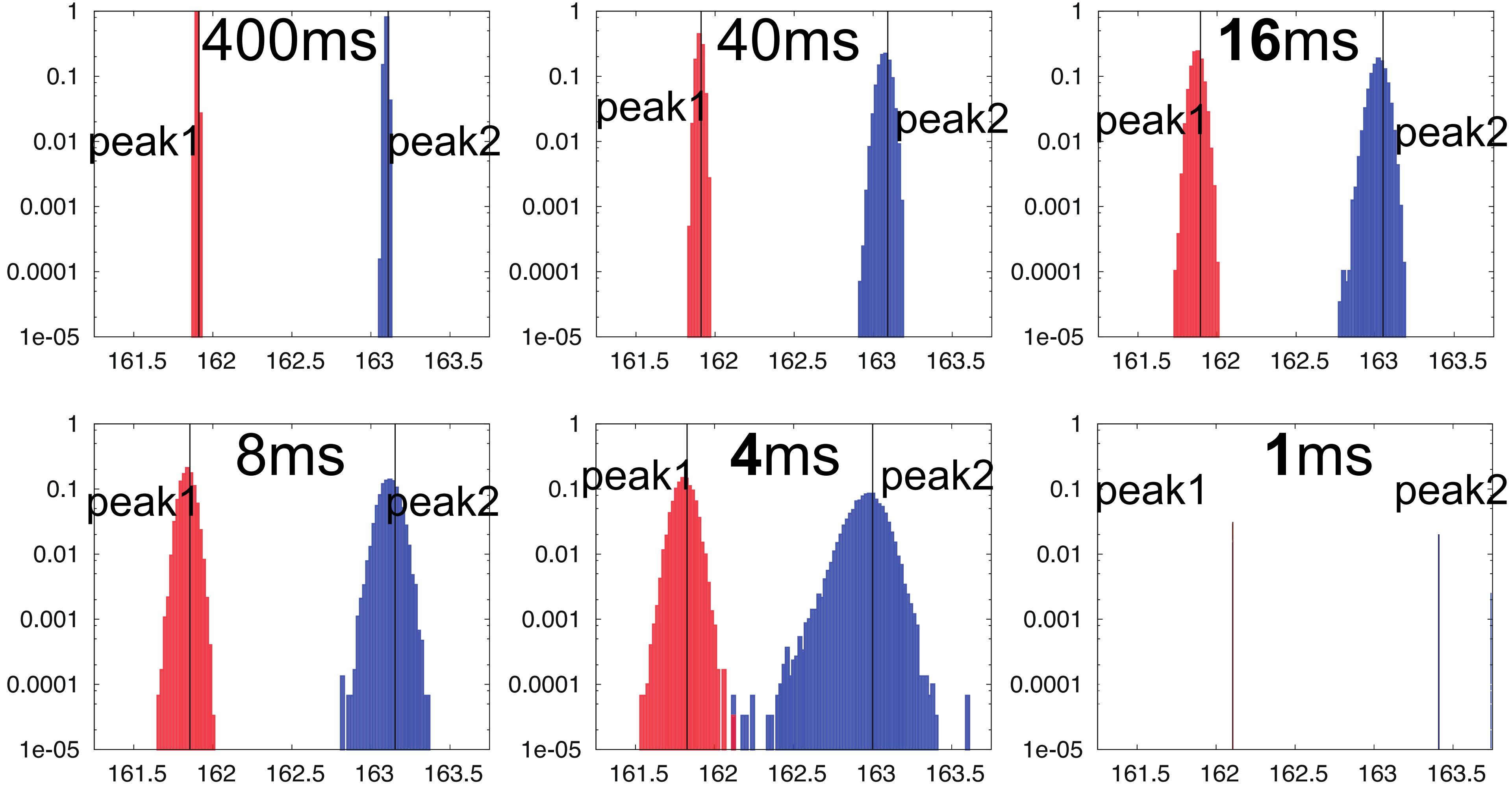
事後分布からのサンプリング

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

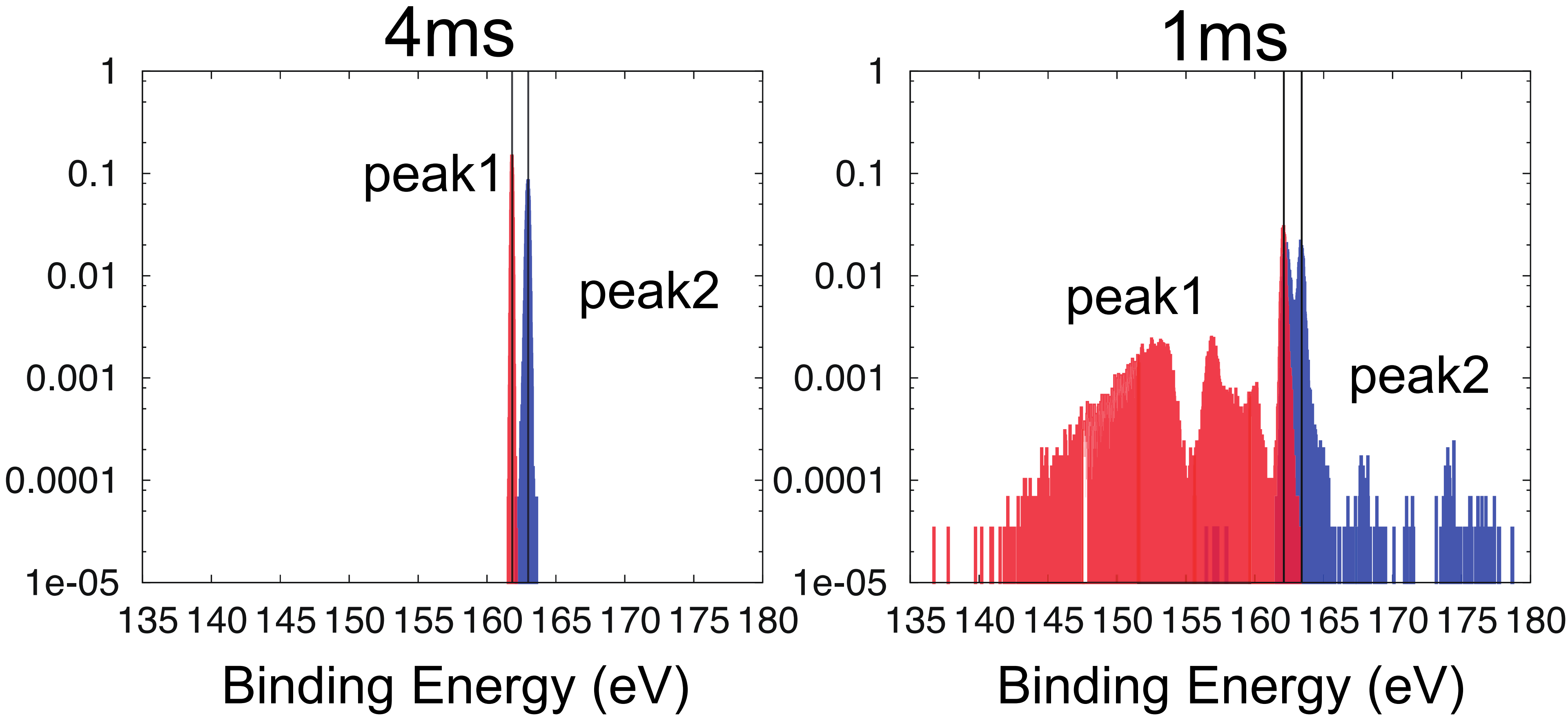


パラメータの推定精度を示すことができる

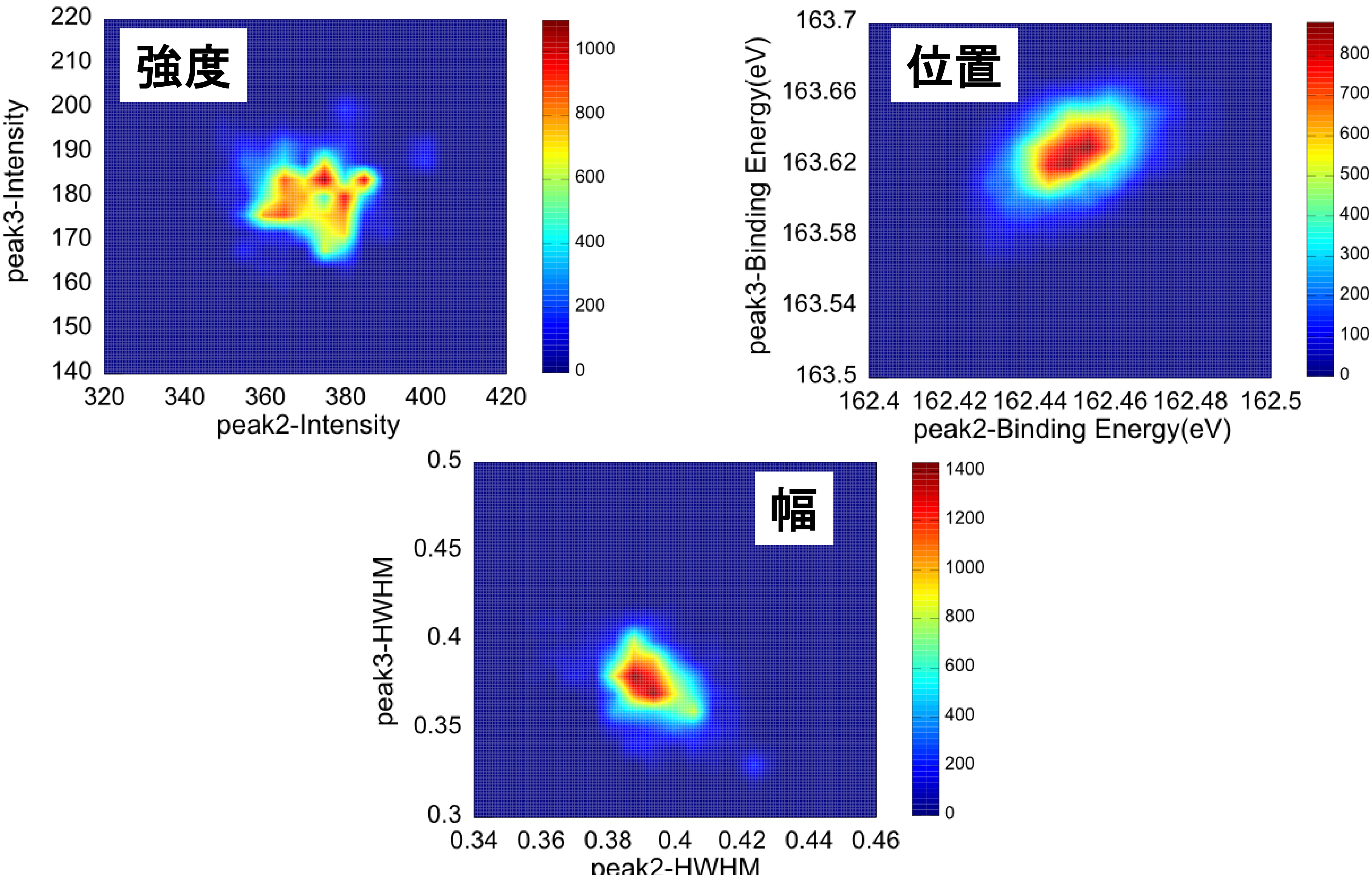
ピーク位置パラメータの推定精度



4msと1msでの相転移の様子



標準試料
二つのピーク間でのパラメータ相関



内容

- 自己紹介：修士課程での研究
  - 希土類化合物の光電子放出・光吸収スペクトル
- 機械学習（ベイズ法）を用いたスペクトル分解
  - 確率的定式化
  - 基底関数の個数の決定
- X線光電子放出スペクトル(XPS)への適用
  - 事後分布推定→推定誤差(エラーバーが引ける!!)
  - 臨界時間窓を導出→時間分割XPSの実験計画
- 議論と今後の展開
  - 順モデルアプローチと逆モデルアプローチ
  - 超高速分光や時間分解光電子分光
  - 電子・格子ダイナミクスや固体中の非平衡電子状態

順アプローチと逆アプローチ

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

$p(Y \mid \theta, K)$ 解析計算，数値計算

逆アプローチ

対象とする物理系 → 計測データ

系の物理モデル

$p(\theta, K)$

観測過程
計測機器の特性

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

全てをモデル化し
ベイズの定理で因果をさかのぼる

系の複雑性と適したアプローチ

複雑さの度合い

物性物理
化学

惑星科学　脳科学
地球科学　生命科学

単結晶　表面　多結晶
粉体

順アプローチ　逆アプローチ

混合原子価希土類化合物のL$_3$-XASを解析して

- 主張：$U_{dc},\ U_{fd}$が必要
- 結論
  - 絶縁体では5eV程度
  - 金属では1～2eV程度
- 当時(25年前)感じた動機
  - 人手によるパラメータサーチ
  - XPSを決めてからXASを解析
  - 余分なパラメータを導入?
  - オーバーフィット
  - 推定パラメータの誤差評価

→情報科学的アプローチ　機械学習

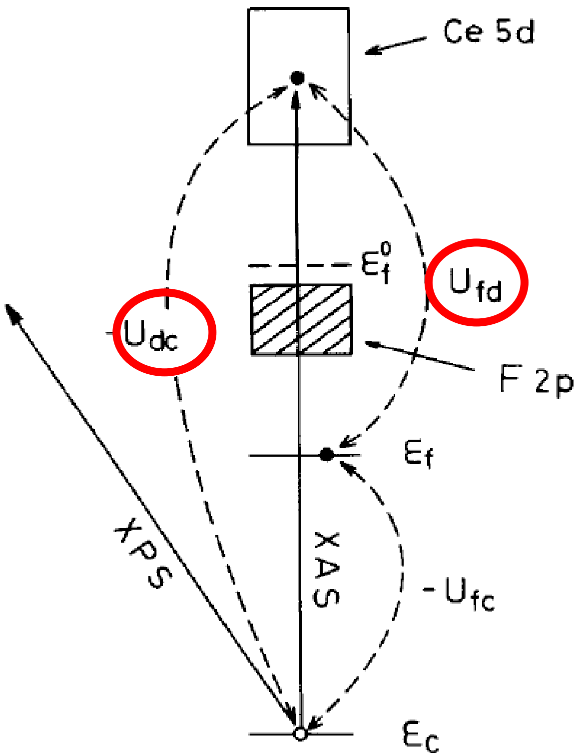
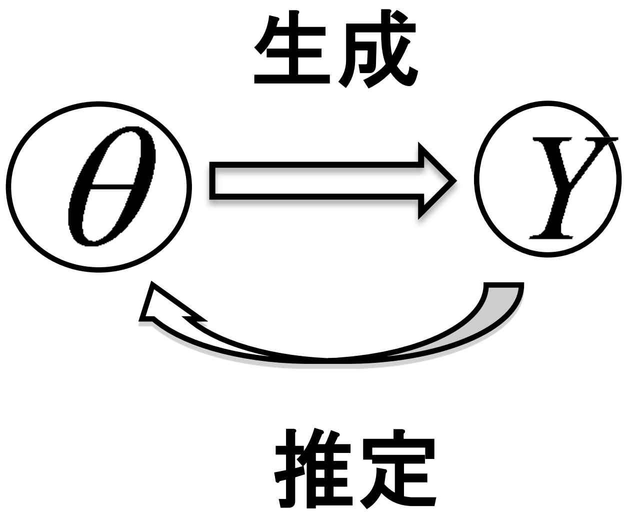
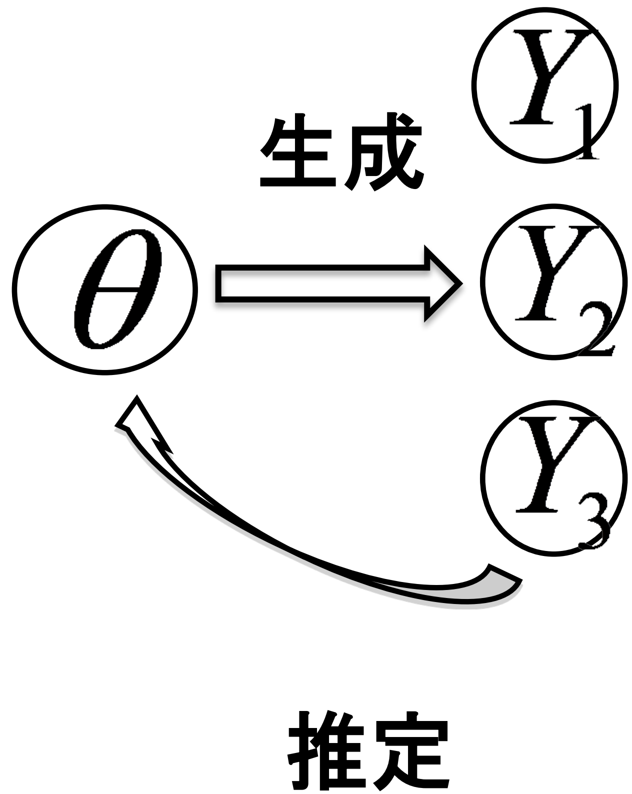


Fig. 1. Model of the present theory describing 3d-XPS and L$_3$-XAS.

スペクトル統合　LASOR





θ ：系の物理

Y ：観測スペクトル

ベイズ推論と物性科学(1/2)

これまで実績のある物理モデルを因果律として内包できる数理的に枠組み

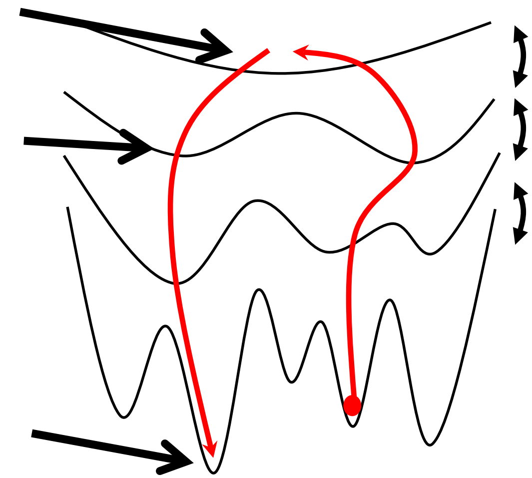
$p(\boldsymbol{x})$　$p(\boldsymbol{y}\mid\boldsymbol{x})$

ベイズの定理で因果律を遡る

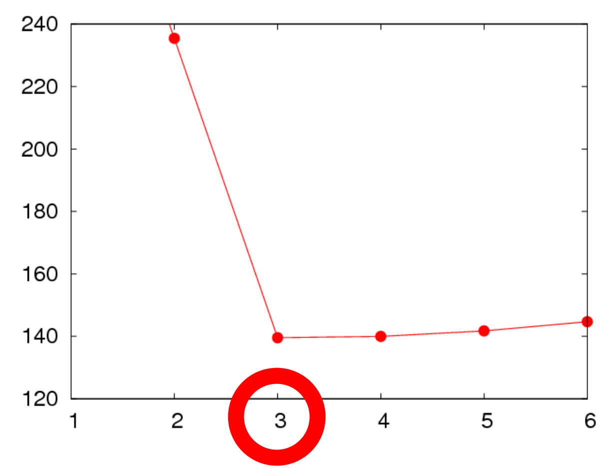
$$p(\boldsymbol{x}\mid\boldsymbol{y})=\frac{p(\boldsymbol{y}\mid\boldsymbol{x})\,p(\boldsymbol{x})}{p(\boldsymbol{y})}$$

(1)ベイズ推論がこれまで実績のある物理モデルを因果律として内包できる数理的に枠組みであることである．この特性により，これまで蓄積された順アプローチでの知見を全て，ベイズ推論の中に取り込むことができて，現状より状況が悪化することはあり得ない．

ベイズ推論と物性科学(2/2)



(2) これまで手作業や近似計算で求めていた物理パラメータを，最先端の数理的最適化手法を用いて，データのみから恣意性無く，系統的・客観的に決定できる



(3) どこまでの相互作用を考慮すべきか等，研究者が直感による思索や試行錯誤できめていたモデル設定をも自動的に行える可能性を秘めた，パラメータ数等のモデルのサイズを決めるモデル選択が行える．

混合原子価希土類化合物のL_3-XASを解析して

- 小谷先生の主張: U_{dc}, U_{fd},が必要
- 結論
  - 絶縁体では5eV程度
  - 金属では1～2eV程度
- 当時(25年前)感じた問題点
  - 発見法的なパラメータサーチ
  - XPSを決めてからXASを解析
  - 余分なパラメータを導入?
  - オーバーフィット
  - 推定パラメータの誤差評価

→情報科学的アプローチ

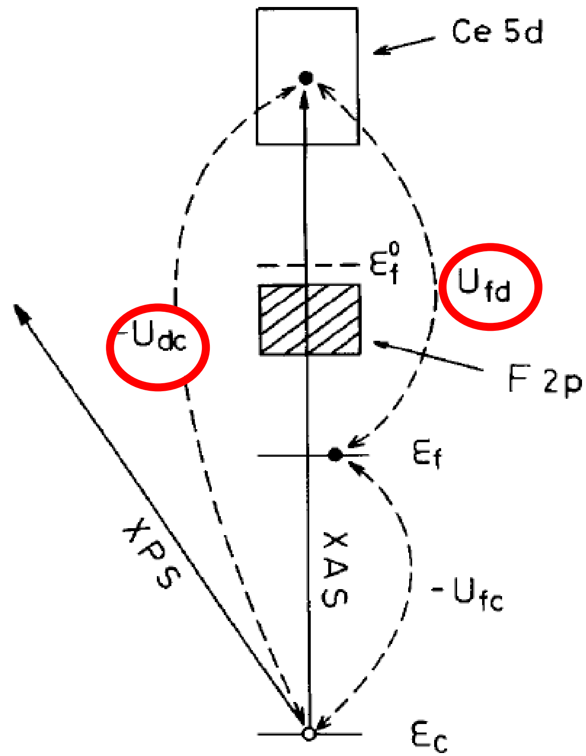


Fig. 1. Model of the present theory describing 3*d*-XPS and L_3-XAS.

まとめと今後の展開

- 情報統計力学，スピングラス
  - 知るを知り，知らざるを知る.
- ベイズ推論による逆アプローチは物性物理全般に有効
- スペクトル解析(解析へのモデルの導入，モデル選択)
  - 時間分割，角度分解スペクトル
  - 有効ハミルトニアンのパラメータの自動決定
- 第一原理計算と逆アプローチの比較融合
  - 生命科学，脳科学，地球惑星科学にない特徴
  - 第一原理計算を推論ループに組み込む