

2019.09.30
2019年度第1回シンクロトロン光産業利用セミナー
ウインクあいち

機械学習を用いたシンクロトロン光計測・解析 の最適化

Optimized x-ray measurement and analysis
with machine learning

高エネルギー加速器研究機構
物質構造科学研究所

物質・材料研究機構
情報統合型物質・材料研究イニシアティブ

小野寛太

共同研究者

Paul Scherrer Institut

齊藤 耕太郎

量子科学技術研究開発機構 (QST)

上野 哲朗

統計数理研究所

日野 英逸

KEK

鈴木 雄太 (学生 D1)

トヨタ自動車

矢野 正雄

庄司 哲也

日立製作所

浅原 彰規

森田 秀和

Zhendong Fu, Vitaliy Pipich

(Jülich Centre for Neutron Science)

Joachim Kohlbrecher

(Paul Scherrer Institut)

武市 泰男

(KEK)



TOYOTA

HITACHI
Inspire the Next

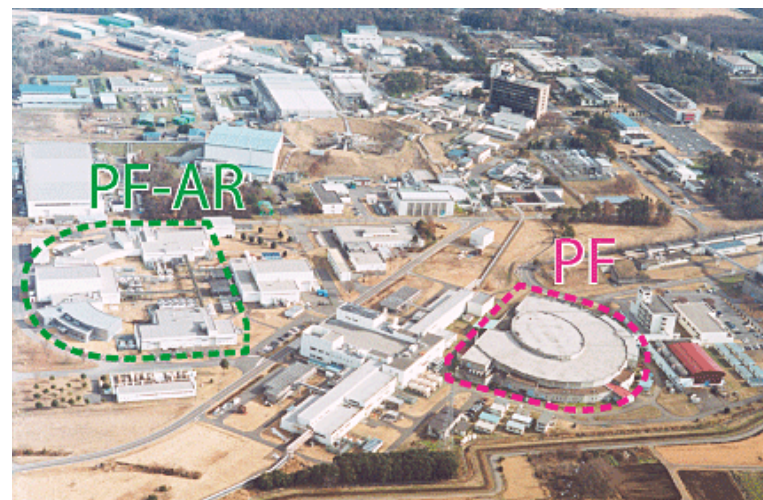


概要

物質・材料の計測をハイスループット化するためには下記の要素が重要だと考えている

- 統計的手法を用いた計測システムの効率化
- 最適な実験計画を自動で策定する方法論
 - 最適な計測時間、計測ポイント
 - 次に計測すべき探索点をデータから決める
- 計測データ解析の自動化
 - 熟練者の経験と勘に頼っていた解析の自動化
 - 次元削減：計測データ (多次元) → 物理量 (低次元)
 - パラメータ最適化
 - 知識発見

シンクロトロン光を用いたハイスループット計測

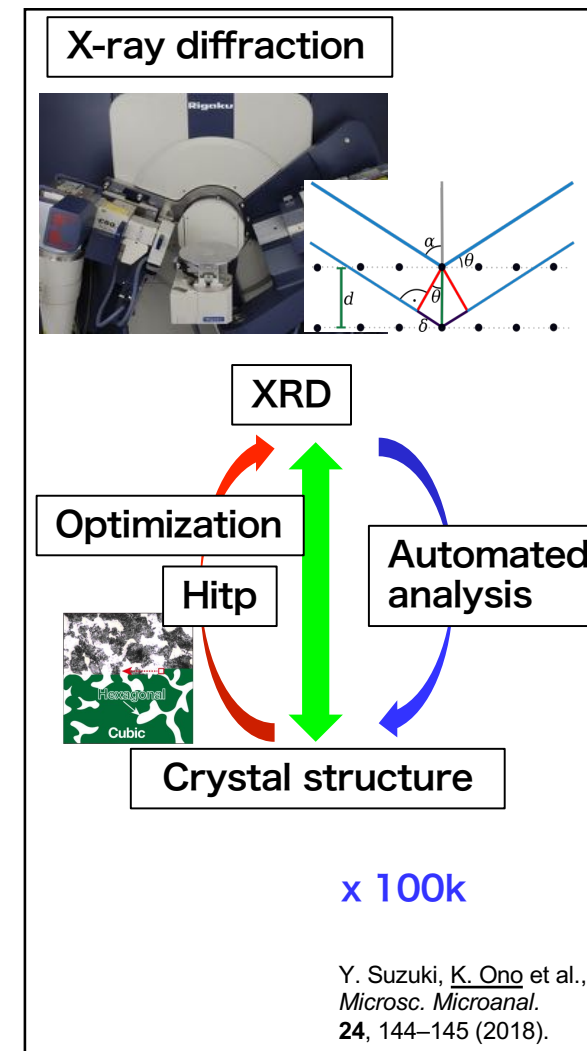
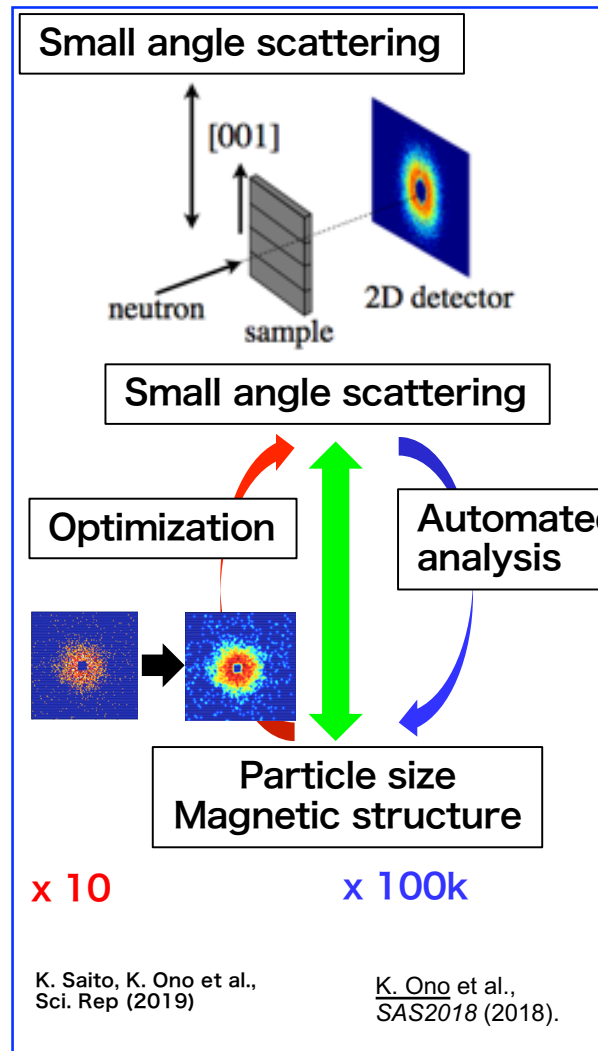
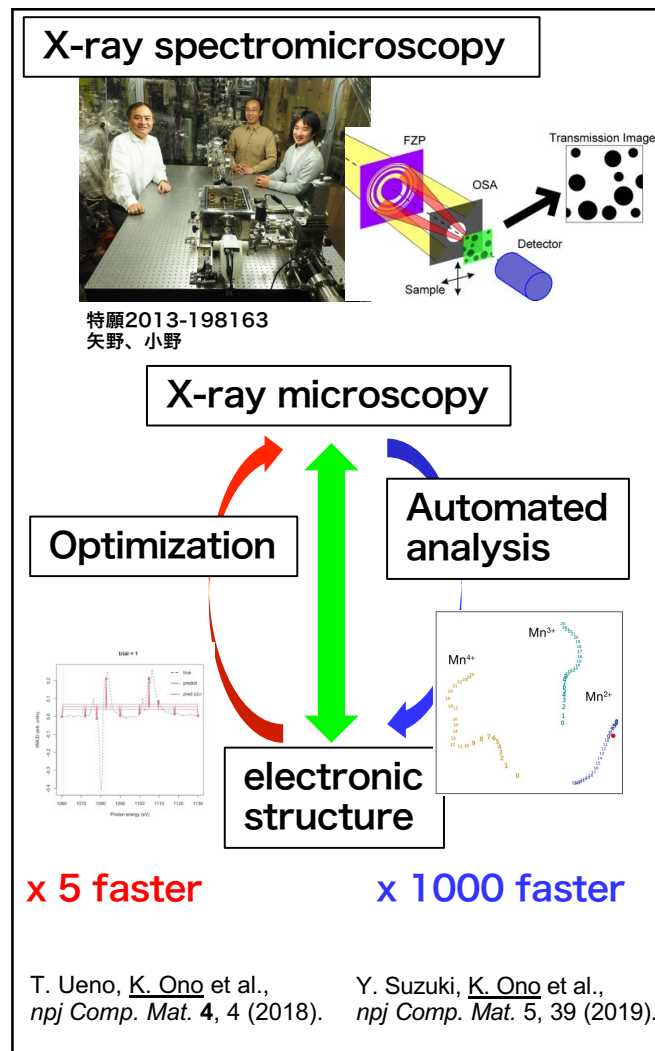


シンクロトロン光を用いた材料研究で必要とされていること

- 低コスト・高精度で計測可能にする方法論
- 放射光・中性子施設の大量の計測データの自動解析
- 広範な自然科学の分野に適用可能な計測・解析の方法論

合理的な計測・解析システム

物質・材料計測のハイスループット化



Case 1 : 統計手法を用いた小角散乱実験の高速化

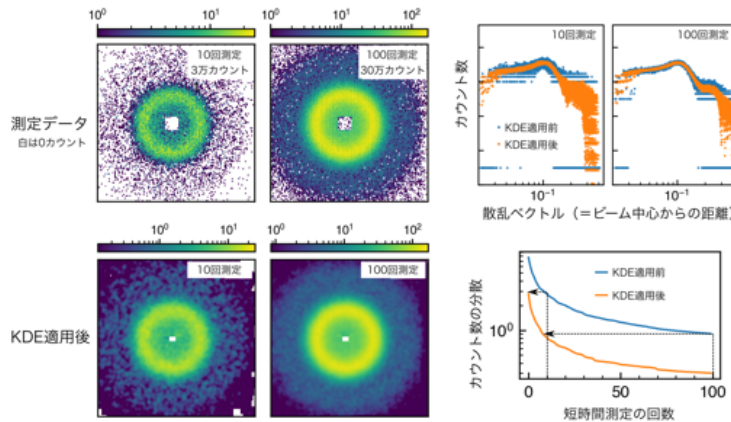
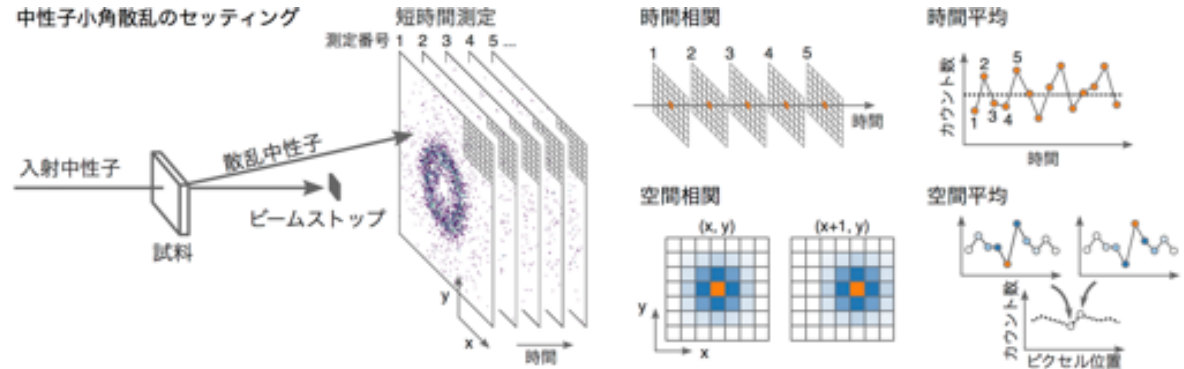
SCIENTIFIC REPORTS

OPEN Accelerating small-angle scattering experiments on anisotropic samples using kernel density estimation

Received: 16 August 2018
Accepted: 5 December 2018
Published online: 06 February 2019

Kotaro Saito¹, Masao Yano², Hideitsu Hino³, Tetsuya Shoji², Akinori Asahara⁴, Hidekazu Morita⁵, Chiharu Mitsumata⁵, Joachim Kohlbrecher⁶ & Kanta Ono⁶

K. Saito, K. Ono et al.,
Sci. Rep (2019)



解析に使う一次元データ
でのノイズ抑制の様子

測定回数 = 時間と分散の関係
KDEによって分散が約10分の1
になっている

KDEによるノイズ抑制効果。短時間測定10回分にKDEを適用すると
100回分の生データとほぼ同等の統計品質 (分散) が得られる。

長時間測定 : 時間相関を使った統計品質向上手法

(誰もが使っている)

カーネル密度推定 : 空間相関

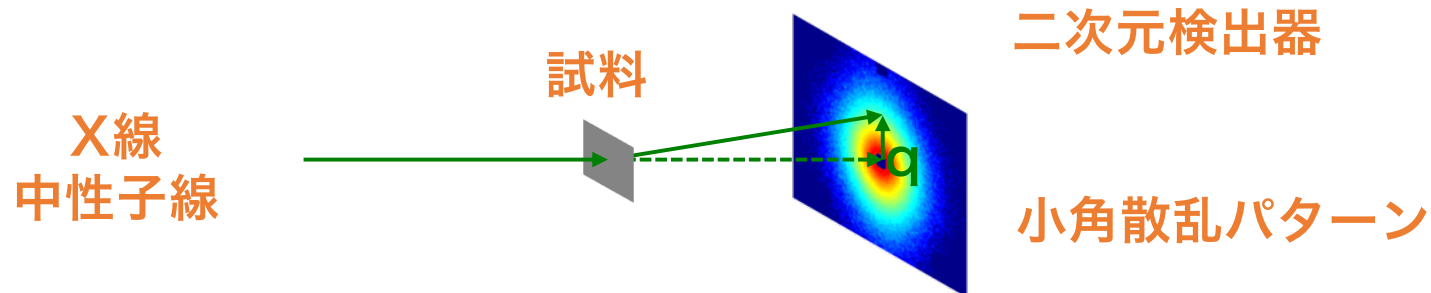
//

(誰も使っていなかった)

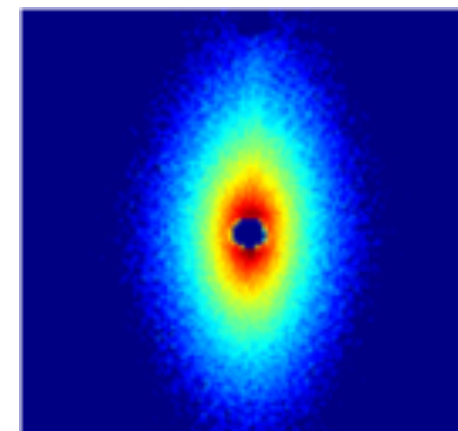


確率密度を推定する統計的手法
画像処理のスムージングと等価
数学的には畳み込み

X線・中性子小角散乱



Small-angle neutron scattering (SANS)
2D data collected by 2D detector
Scattering patterns are nonlinear
High dynamic range



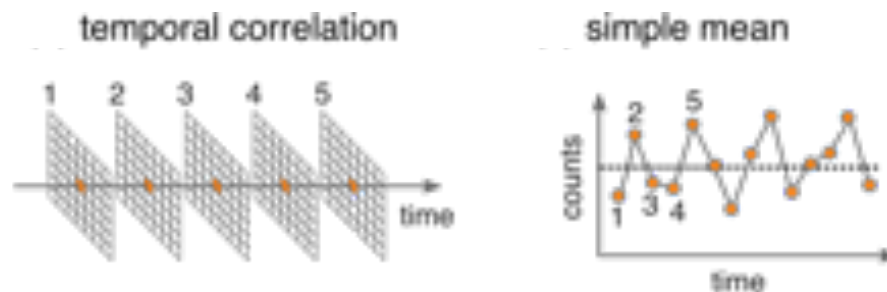
SANS pattern

問題点

限られた施設、ビームタイム
高い統計精度でのデータを得るには長時間測定が必須

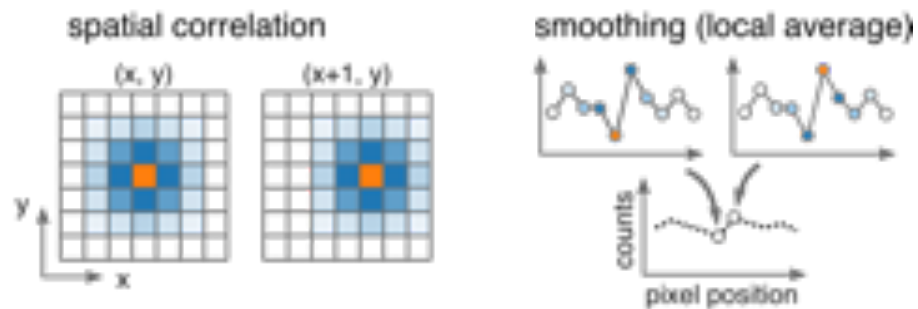
小角散乱データにおける二つの相関

時間的相関



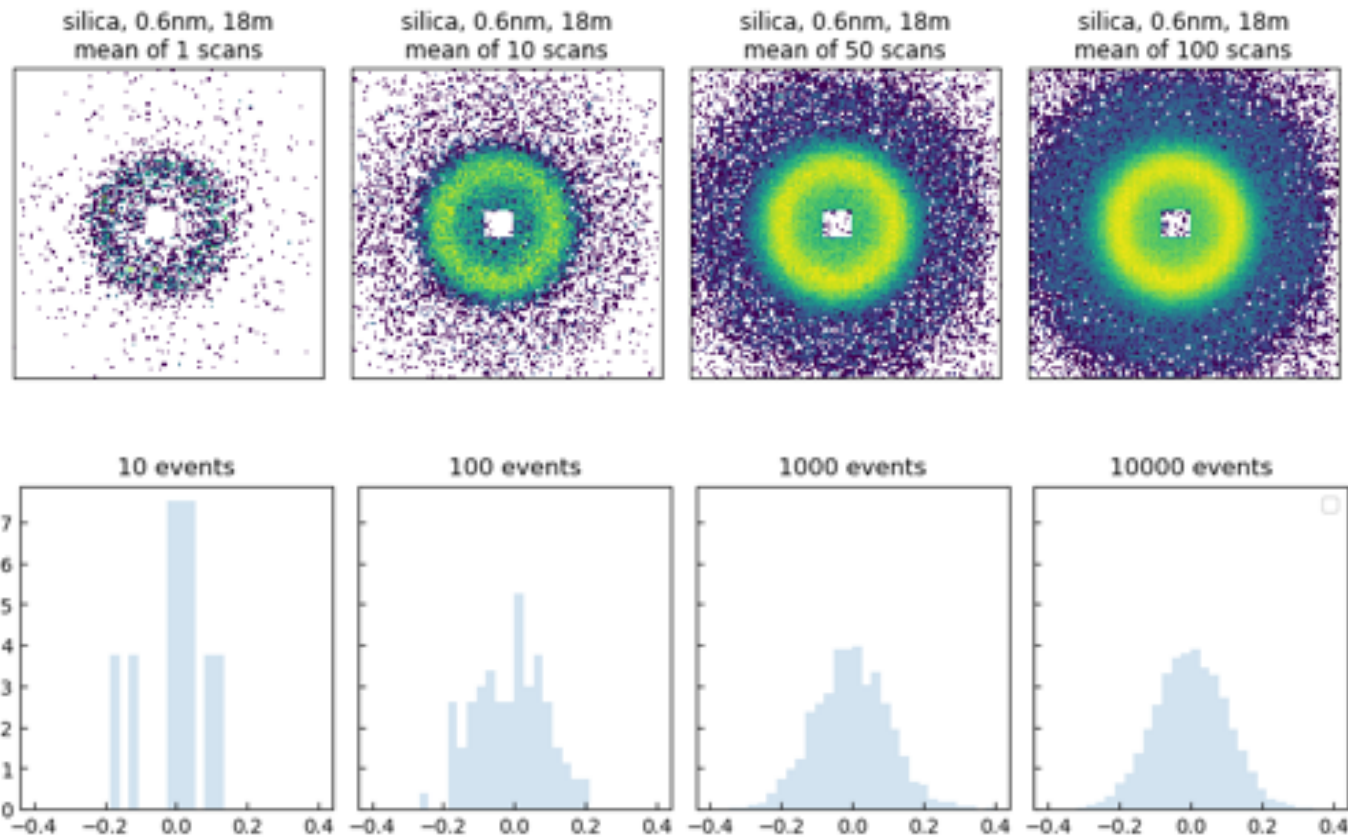
Counts per unit measurement time for a given pixel (or a bin) are always comparable.

空間的相関



If counts at a pixel (or a bin) is large/small, same for neighborig pixels (bins).

統計精度を上げるには：長時間測定



The simplest and the most used strategy to get nice data.

- Longer you measure, nicer the result looks.
- Time-consuming if a sample is a weak scatterer.

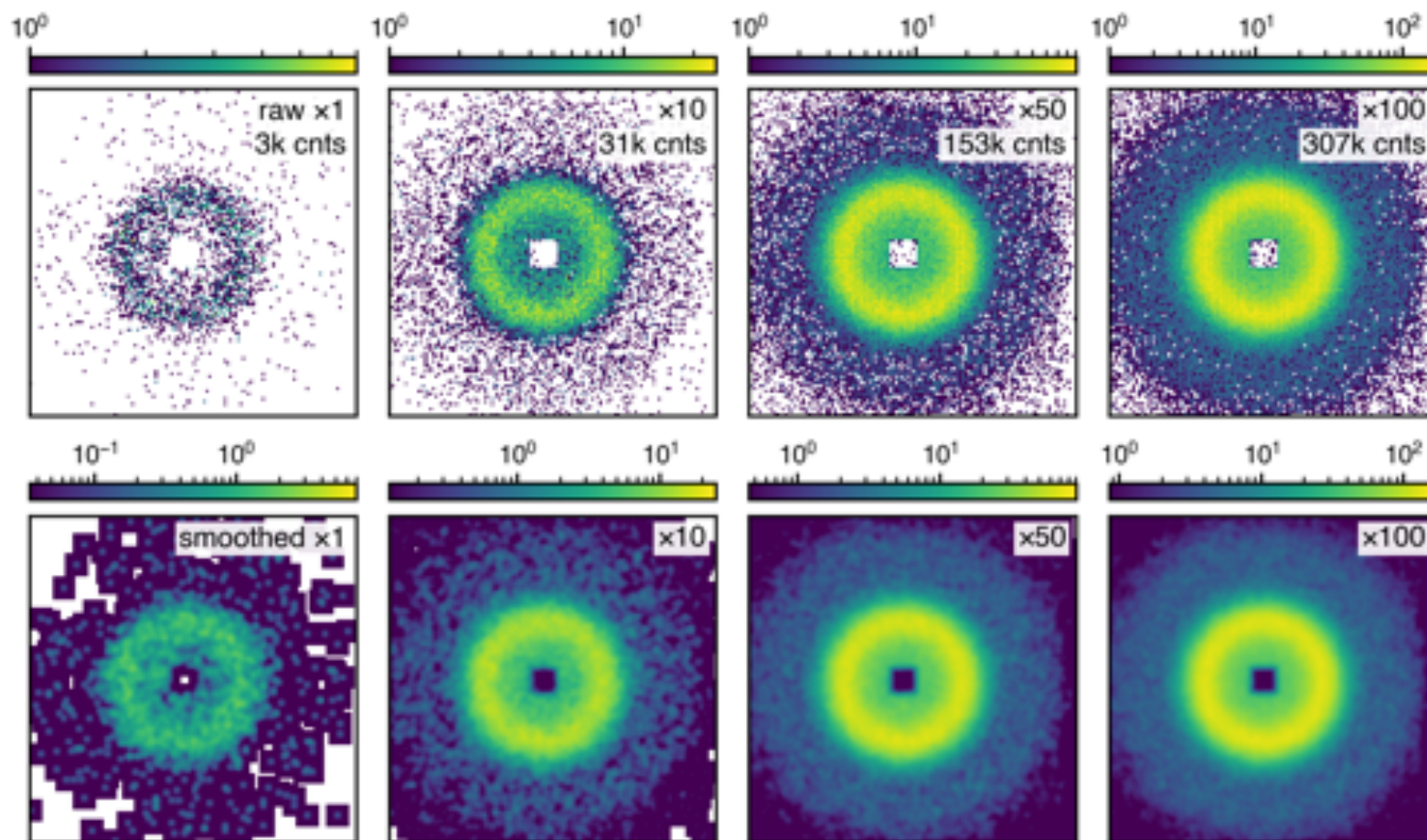
カーネル密度推定による効果

sample: silica particle, $r=60\text{nm}$, 36 vol.%,

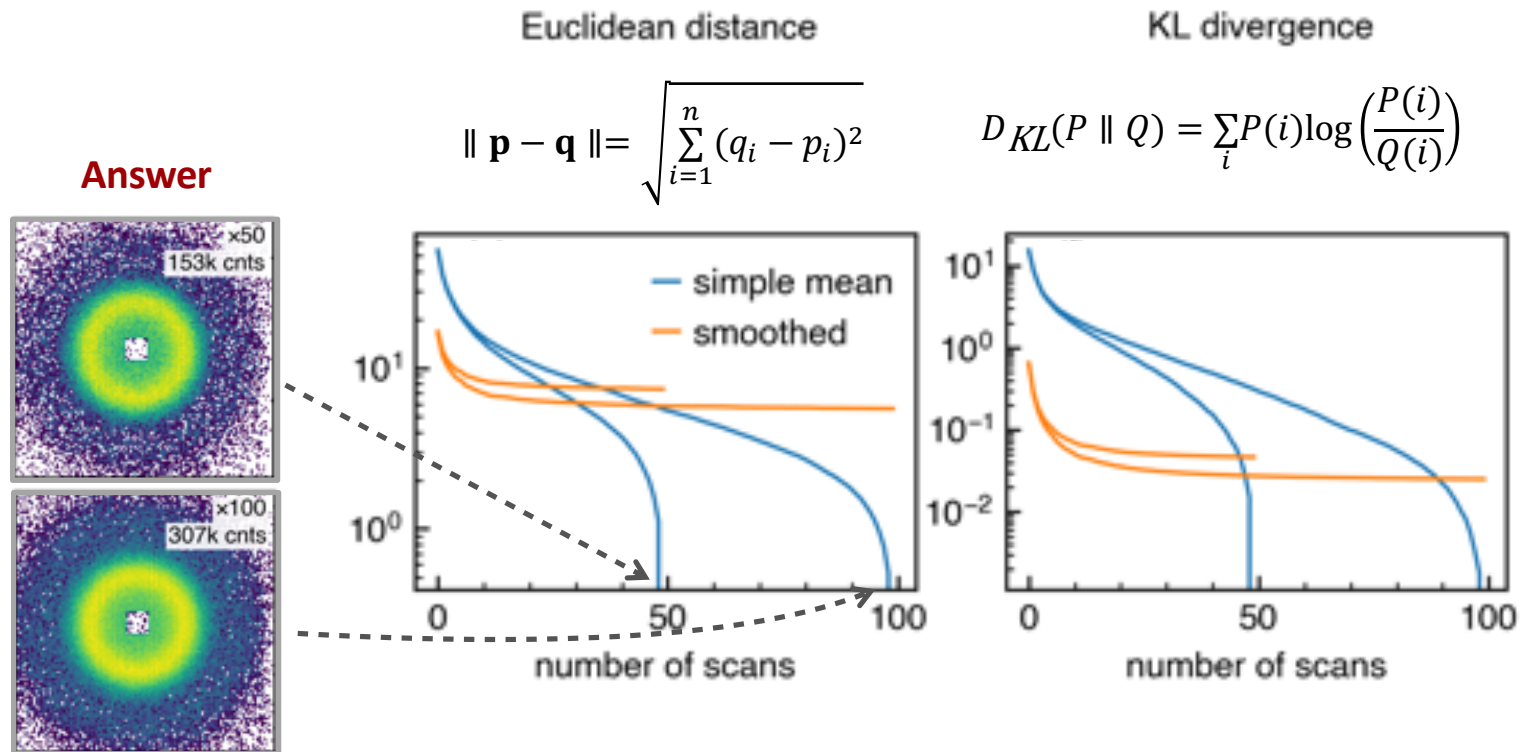
instrument: SANS-I@SINQ (PSI) , $\lambda=0.6\text{ nm}$, detector distance 18 m

100 short scans (total counts for 100 scans = 300k, a typical rule of thumb)

KDE : Gaussian kernel ($\sigma = 1\text{ pixel} < \text{resolution}$)



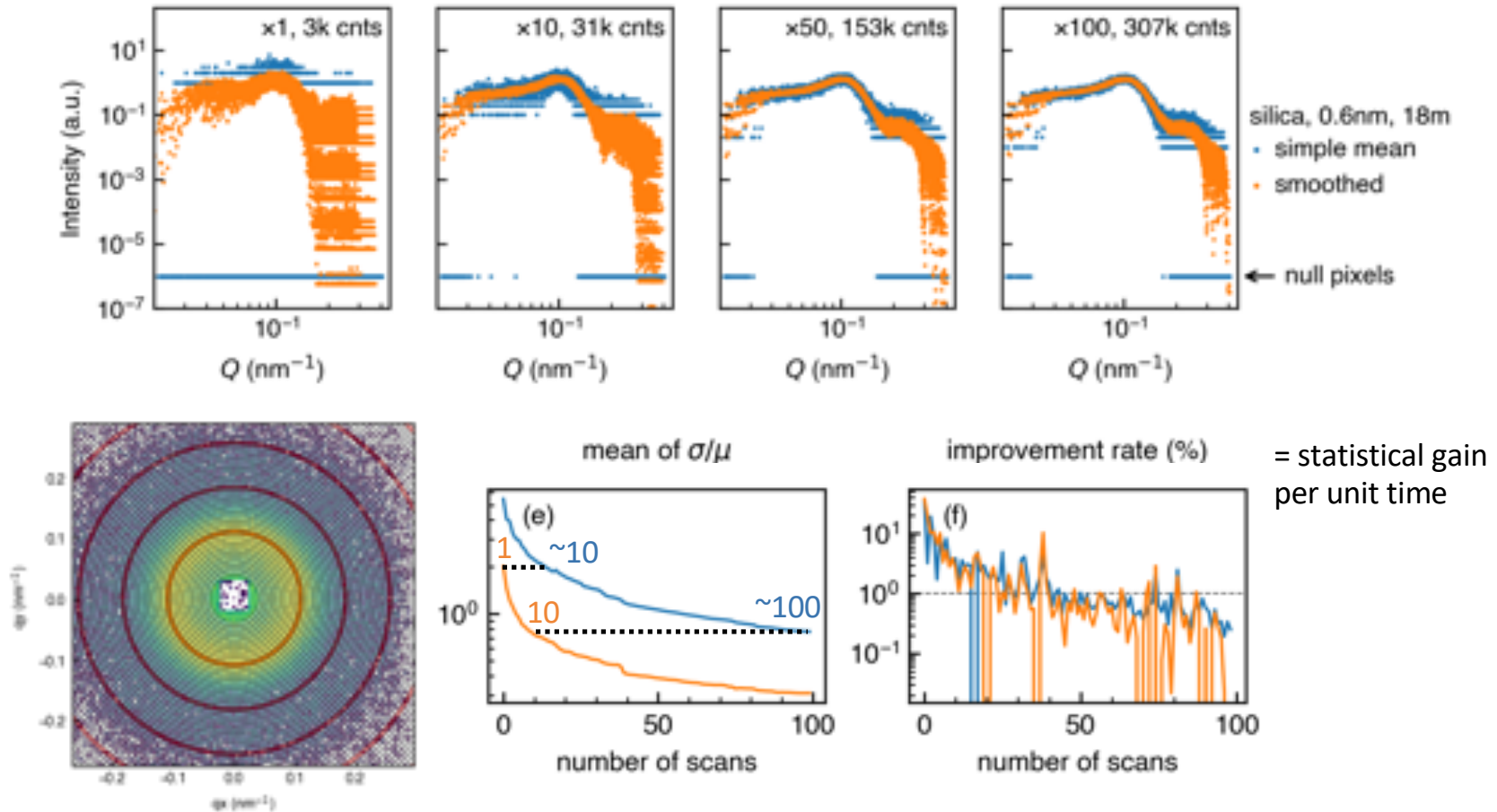
カーネル密度推定による効果：長時間測定との比較



- Smoothing makes SANS data better in terms of both metrics.
- “How better?” is a difficult question if a metric needs “*an answer*”.

カーネル密度推定による効果：分散

128x128 = 16384 pnts



- n scans with smoothing $\sim 10n$ scans without smoothing
- Smoothing can accelerate SAS measurement $\times 10$ faster to get data with same variance

Case 2 : 適応型実験デザイン法 (Adaptive design of experiments)

適応型実験デザイン法では、計測・解析・機械学習を融合することで合理的かつ効率的な実験計画が自動的に策定され、実験者に依存しない計測が可能になる

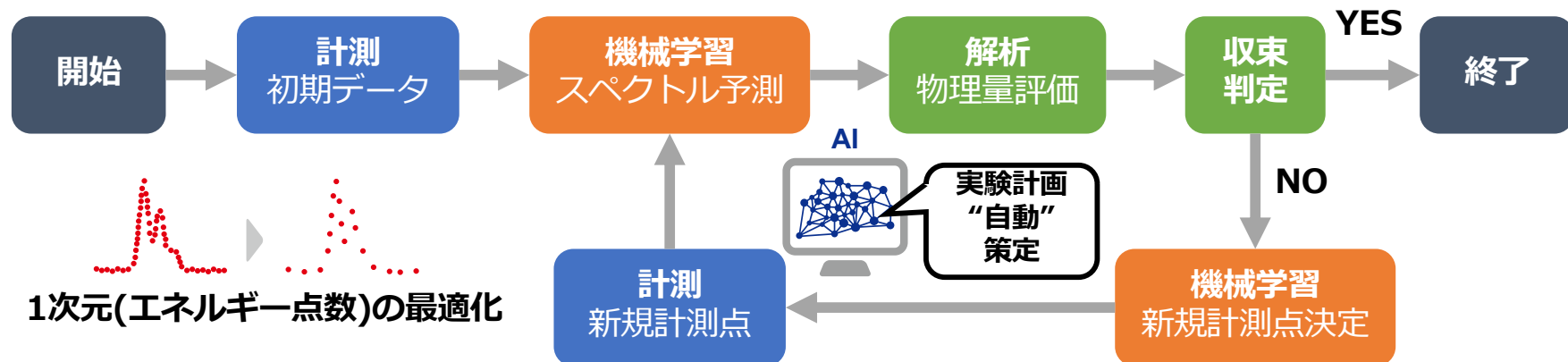
従来の実験デザイン法



適応型実験デザイン法 (Ver. 1)

T. Ueno *et al.*, npj Comput. Mater. 4, 4 (2018)

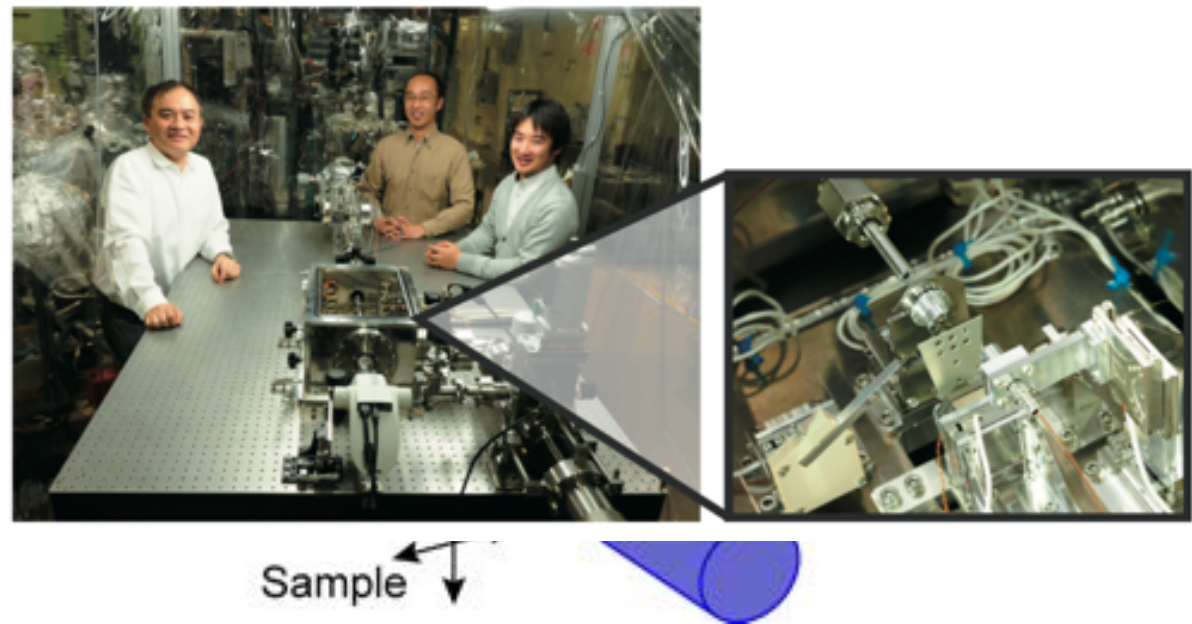
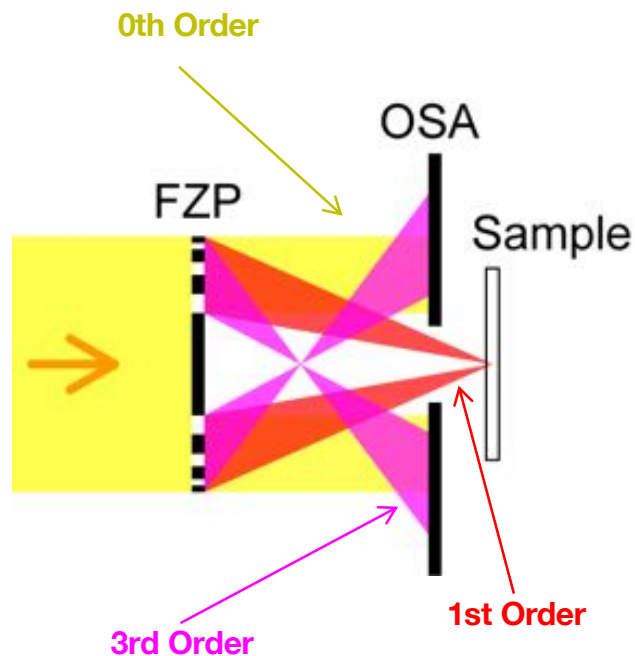
- 計測・機械学習・解析を融合
- 機械学習による計測点の自動決定
- 実験終了基準はスペクトルから算出する物理量



走査型透過X線顕微鏡 STXM

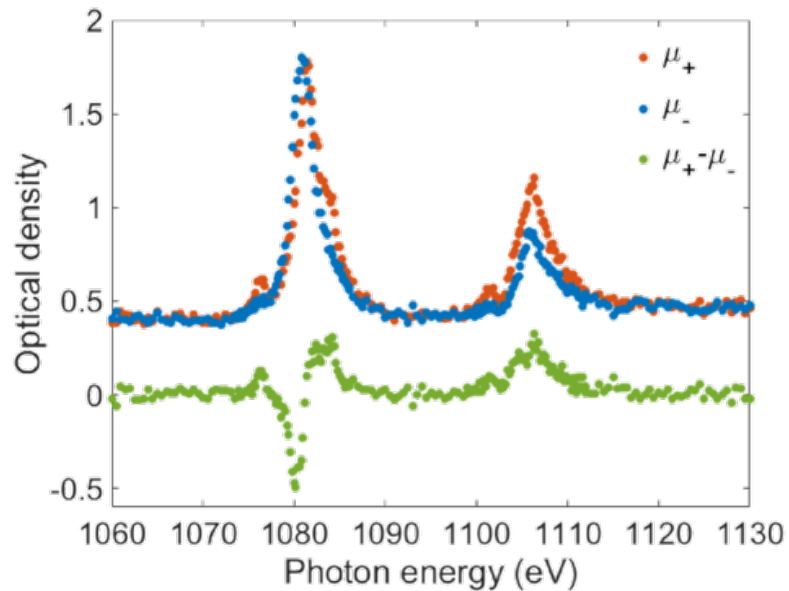
STXM : Scanning Transmission X-ray Microscopy

- Fresnel zone plate (FZP) focuses x-rays and Order sorting aperture (OSA) select first order diffraction.
- Scanning the sample position and detecting the intensity of transmitted x-rays to obtain a microscopic image
- Spatial resolution : 10 ~ 100 nm

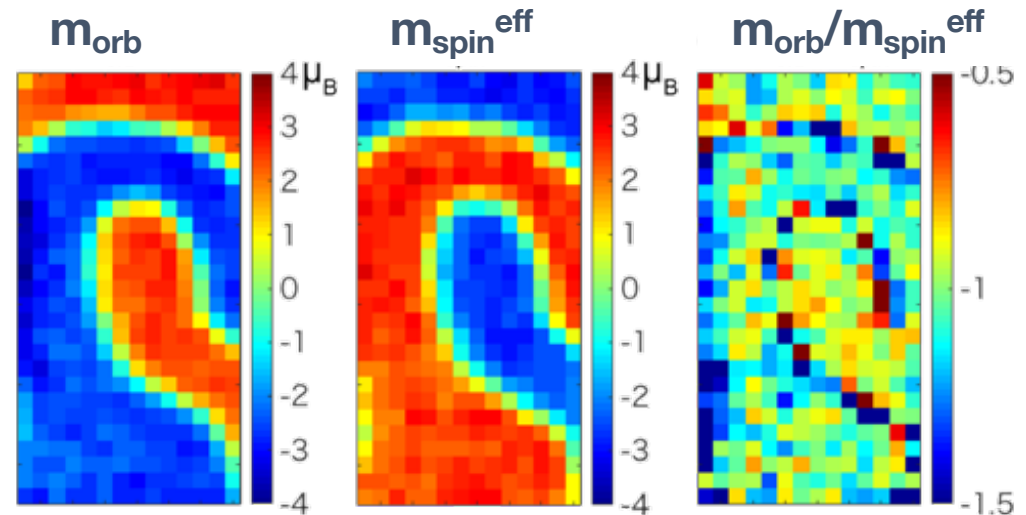


STXM を用いたナノスペクトロスコピー

単一ピクセル (30 nm) 領域での
XAS & XMCD スペクトル



磁気モーメントマップ



T. Ueno et al., AIP Adv. 7, 056804 (2017)

- Evaluation of magnetic moments ... Good S/N spectra
- Enormous numbers of spectra (10,000 for 100 x 100 pixels)
- Large area with high spatial resolution



High-throughput measurement is necessary

適応型実験デザイン法の実証

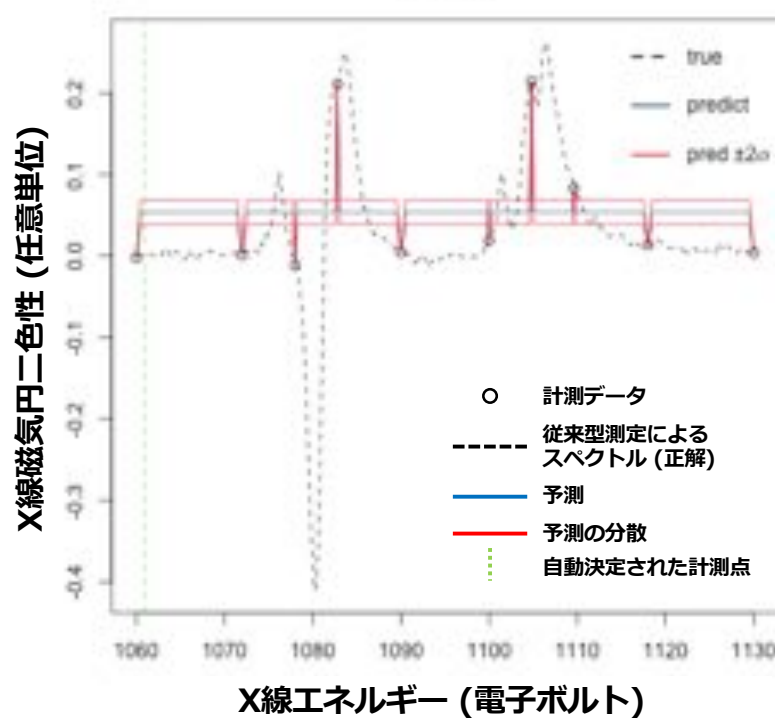
ARTICLE OPEN

Adaptive design of an X-ray magnetic circular dichroism spectroscopy experiment with Gaussian process modelling

Tetsuro Ueno^{1,2,3}, Hideitsu Hino⁴, Ai Hashimoto², Yasuo Takeichi², Masahiro Sawada⁵ and Kanta Ono^{2,3,6}

T. Ueno, K. Ono et al., npj Comput. Mater. 4, 4 (2018)

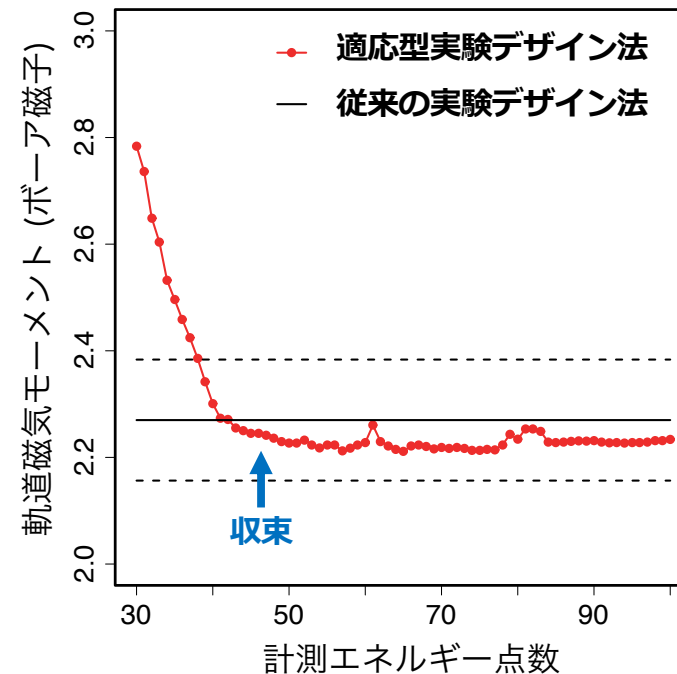
適応型実験デザイン法による計測の様子 (ガウス過程回帰) trial = 1



ガウス過程回帰

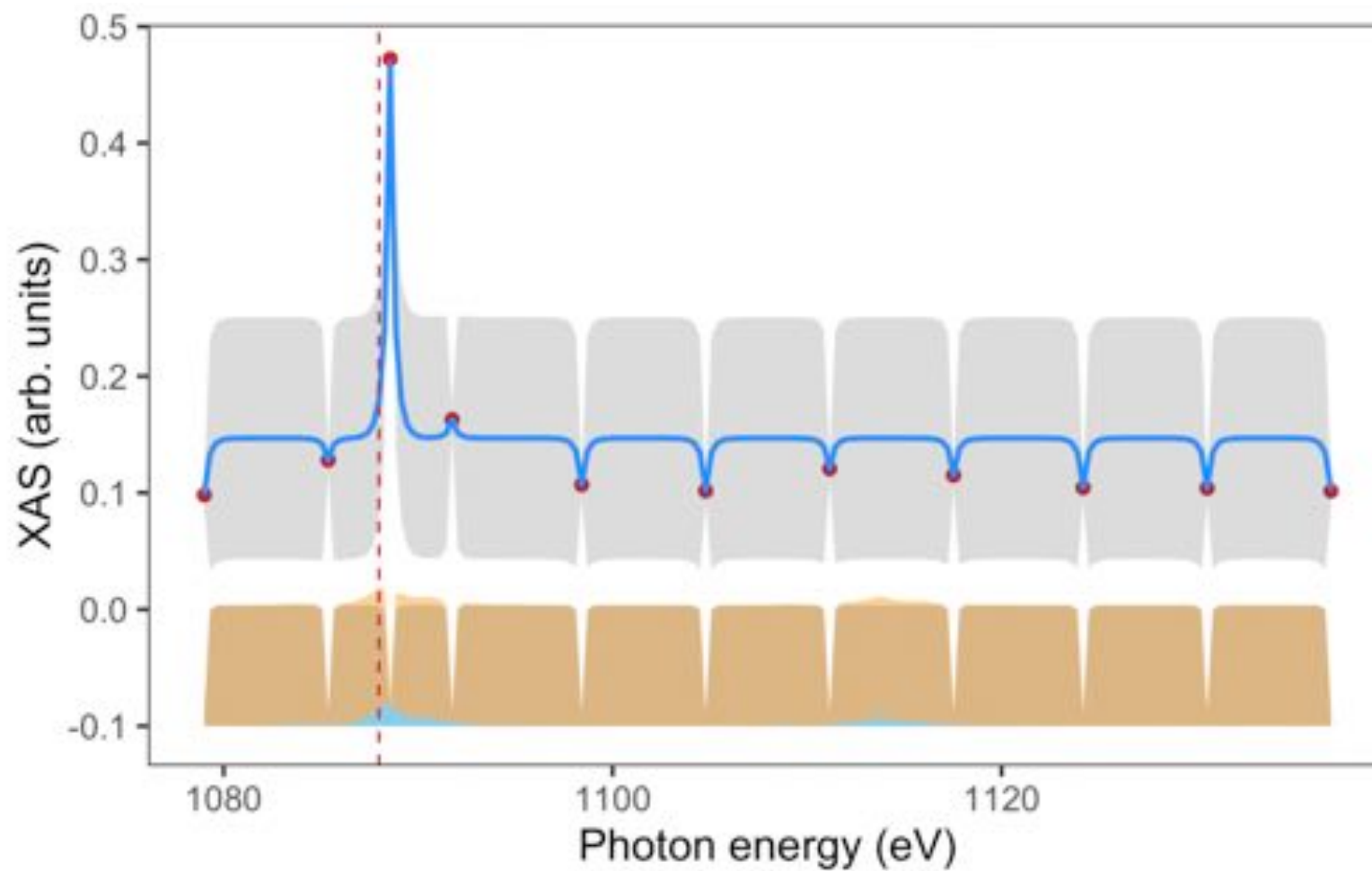
- ✓ 少数の計測データから複雑なスペクトル形状を予測可能
- ✓ 次にどこを計測すれば予測精度が上がるかがわかる (分散がわかる)

計測点数 vs 物理量 (軌道磁気モーメント)

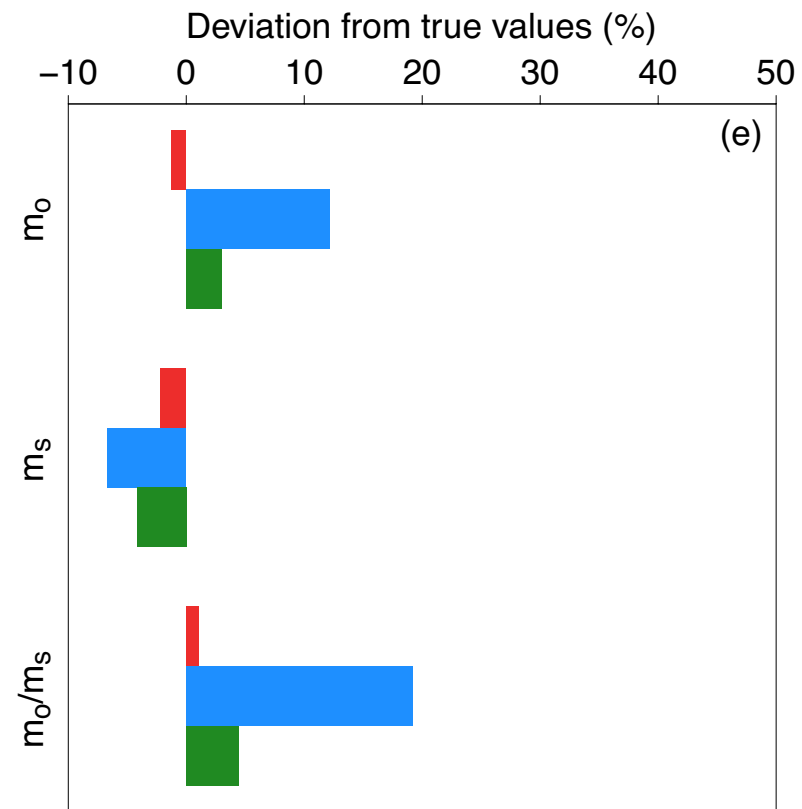
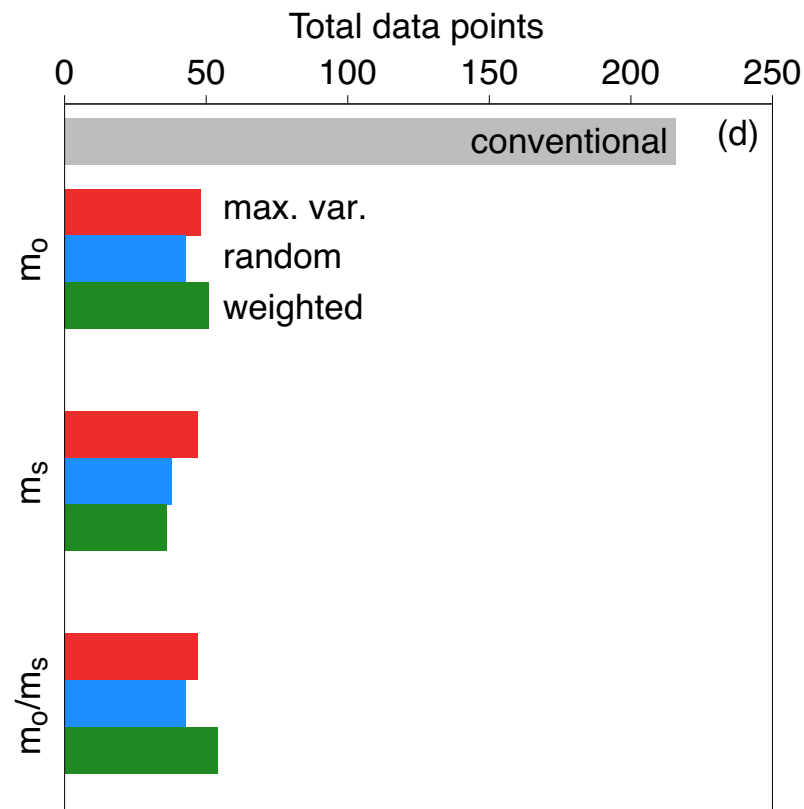


従来の実験デザイン法(200点以上の計測)と同等の精度を40~50点の計測で達成
→ 5倍の効率化

適応型実験デザイン法を用いたX線スペクトル計測



計測ポイントの数と誤差



従来の計測法に比べて 1/5 の計測点
5 倍の高速化が可能になった

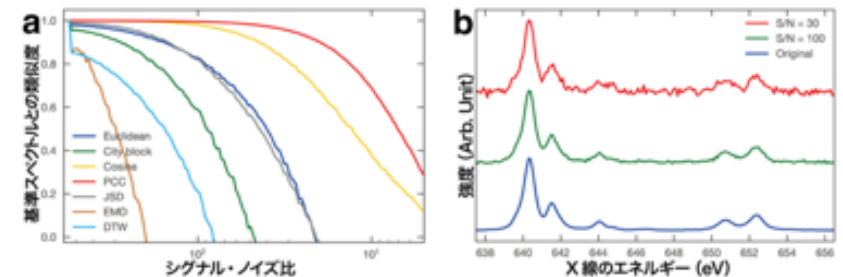
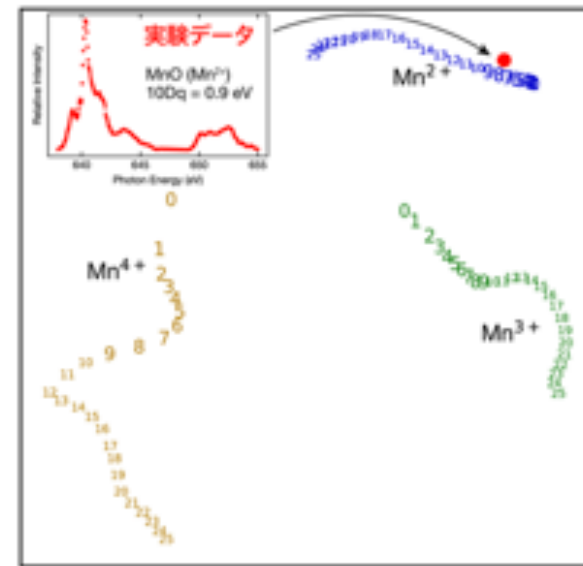
十分な精度で物理量を求めることが可能

Case 3-1 : 計測データ解析の自動化 -X線スペクトル-

ARTICLE OPEN

Automated estimation of materials parameter from X-ray absorption and electron energy-loss spectra with similarity measures

Yuta Suzuki^{1,2}, Hideitsu Hino³, Masato Kotsugi^{1,4} and Kanta Ono^{2,4}



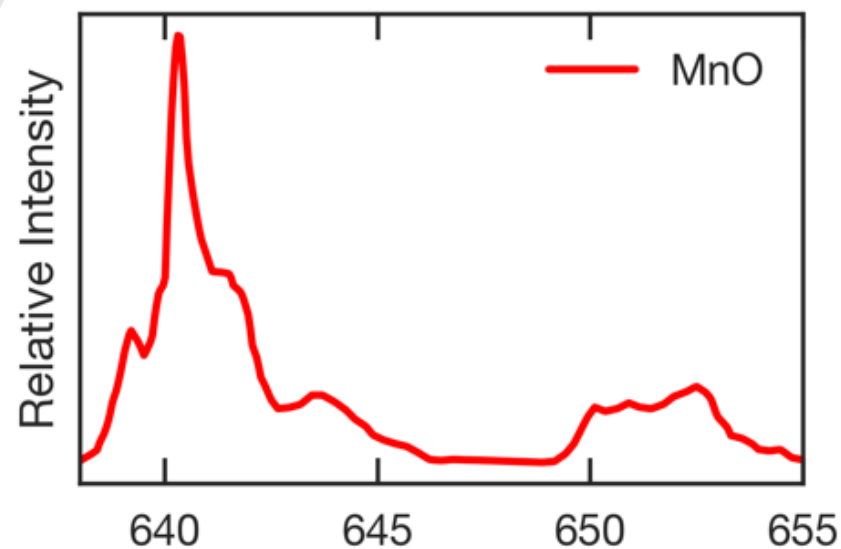
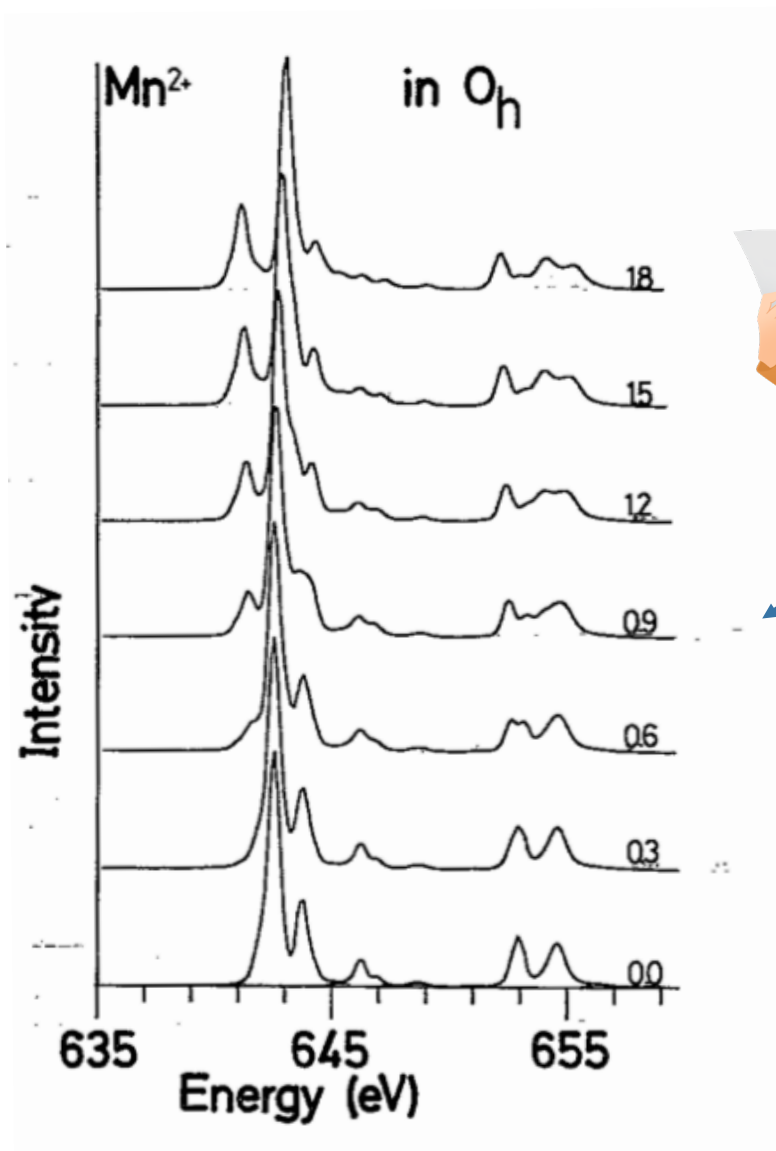
Y. Suzuki, K. Ono et al., *npj Comp. Mat.* **5**, 39 (2019).

機械学習によりX線吸収スペクトル解析の自動化に成功
データの類似度に着目し定量的なスペクトルの解析を実現

解析自動化のためのワークフロー



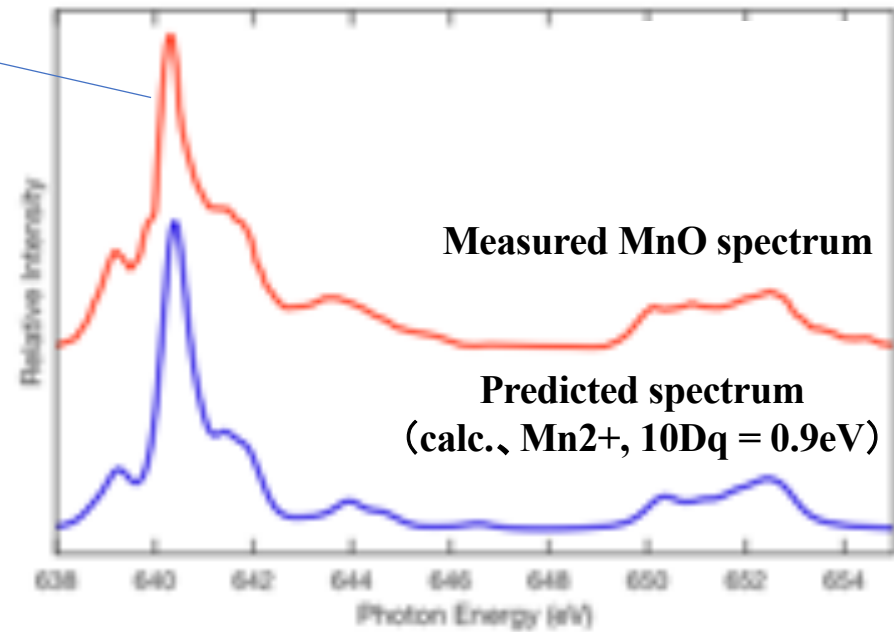
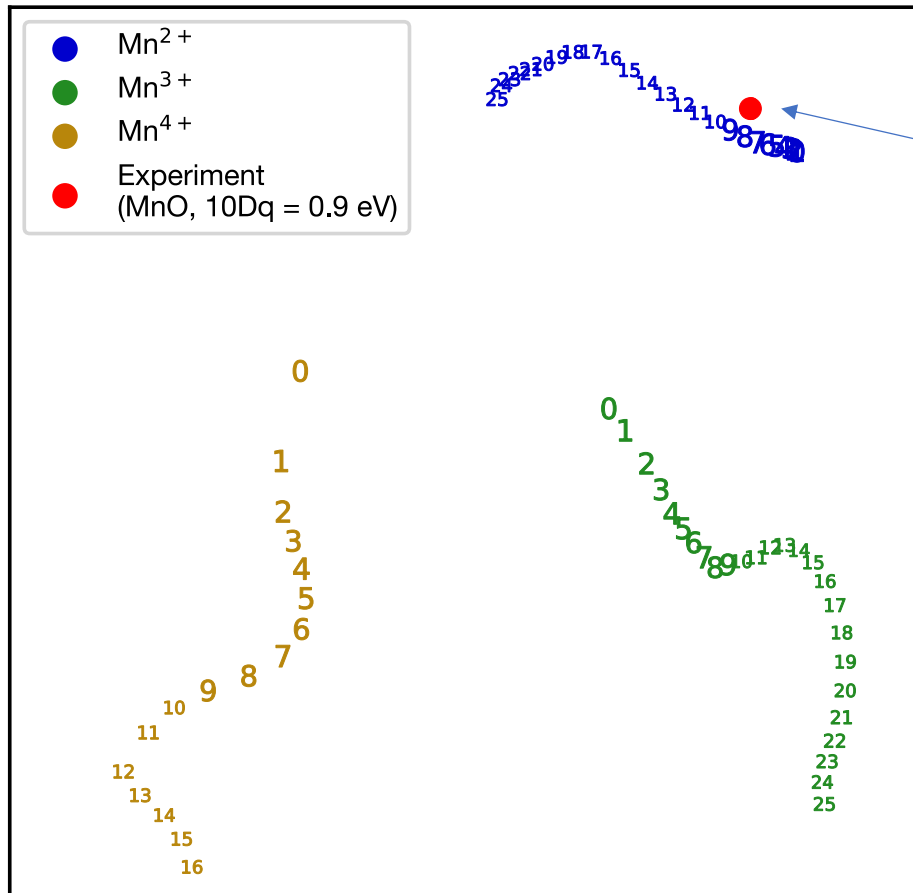
X線スペクトルからの物理量の推定



de Groot, F. & Kotani, A. Core Level Spectroscopy of Solids. (CRC Press, 2008).

de Groot, F. et.al., Phys. Rev. B 42, 5459–5468 (1990).

X線スペクトルデータの低次元への次元削減と可視化



計測スペクトルと予測したスペクトル

多次元尺度構成法 (Multi-dimensional scaling (MDS)) によるスペクトルデータの次元削減と可視化

計測データに用いられる種々の距離尺度（類似度）

- Euclid distance
- Cityblock distance
- Cosine distance
- Pearson correlation coefficient
- Jensen-Shannon Divergence
- Dynamic Time Warping
- Earth Mover's Distance

$$D_{p\text{-norm}}(\mathbf{s}, \mathbf{t}) = \left(\sum_{i=1}^N (s_i - t_i)^p \right)^{1/p}$$

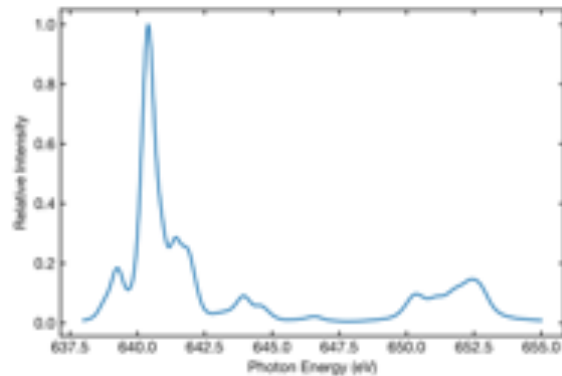
$$D_{\text{cosine}}(\mathbf{s}, \mathbf{t}) = 1 - \frac{\sum_{i=1}^n (s_i \cdot t_i)}{(\sum_{i=1}^n s_i^2)^{\frac{1}{2}} (\sum_{i=1}^n t_i^2)^{\frac{1}{2}}}$$

$$D_{\text{Pearson}}(\mathbf{s}, \mathbf{t}) = 1 - \frac{\sum_{i=1}^n (s_i - \bar{s})(t_i - \bar{t})}{(\sum_{i=1}^n (s_i - \bar{s})^2)^{\frac{1}{2}} (\sum_{i=1}^n (t_i - \bar{t})^2)^{\frac{1}{2}}}$$

$$JSD(s, t) = \frac{1}{2} D_{KL}(s, M) + \frac{1}{2} D_{KL}(t, M)$$

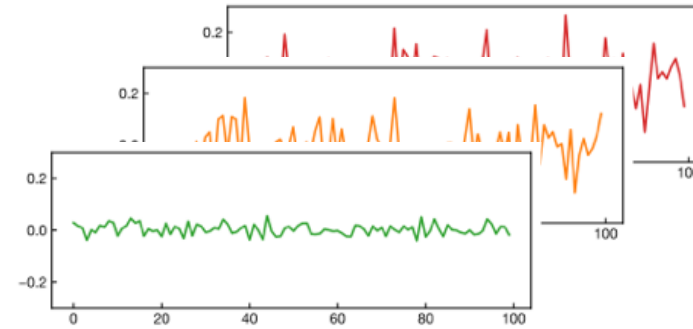
$$D_{KL}(s, t) = \sum_i s(i) \log \frac{s(i)}{t(i)} \quad M = \frac{1}{2}(s, t)$$

類似度に求められる性質：ノイズに対する冗長性

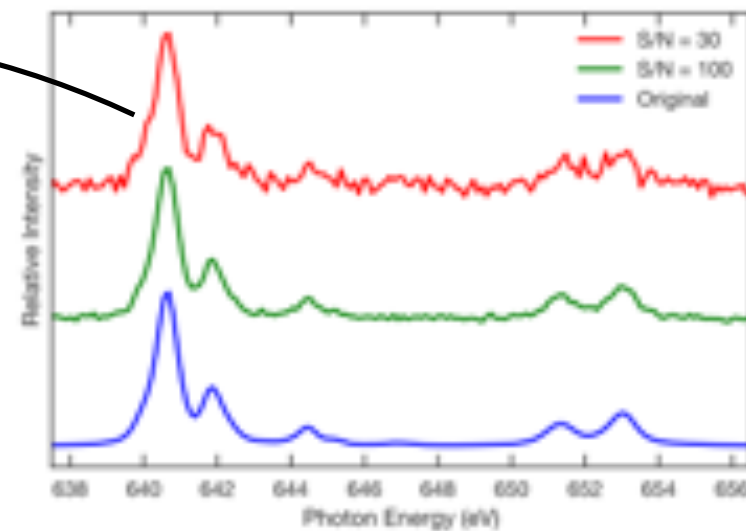
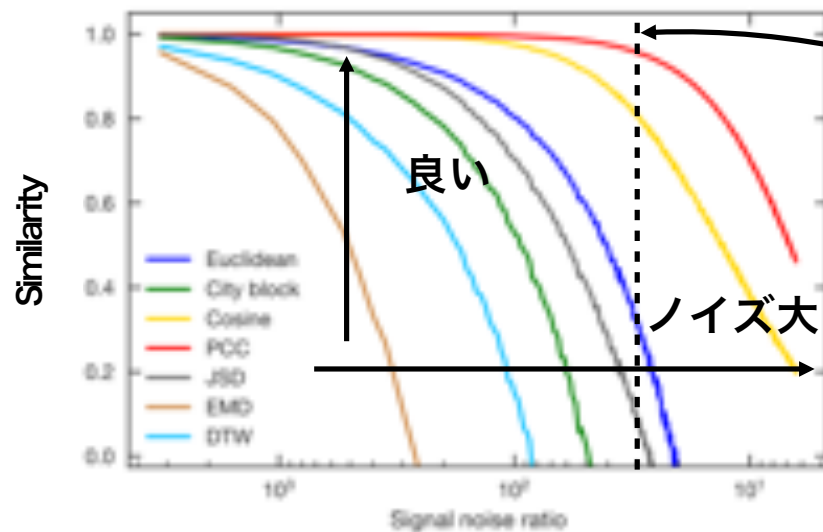


X線吸収スペクトル (シミュレーション)

+



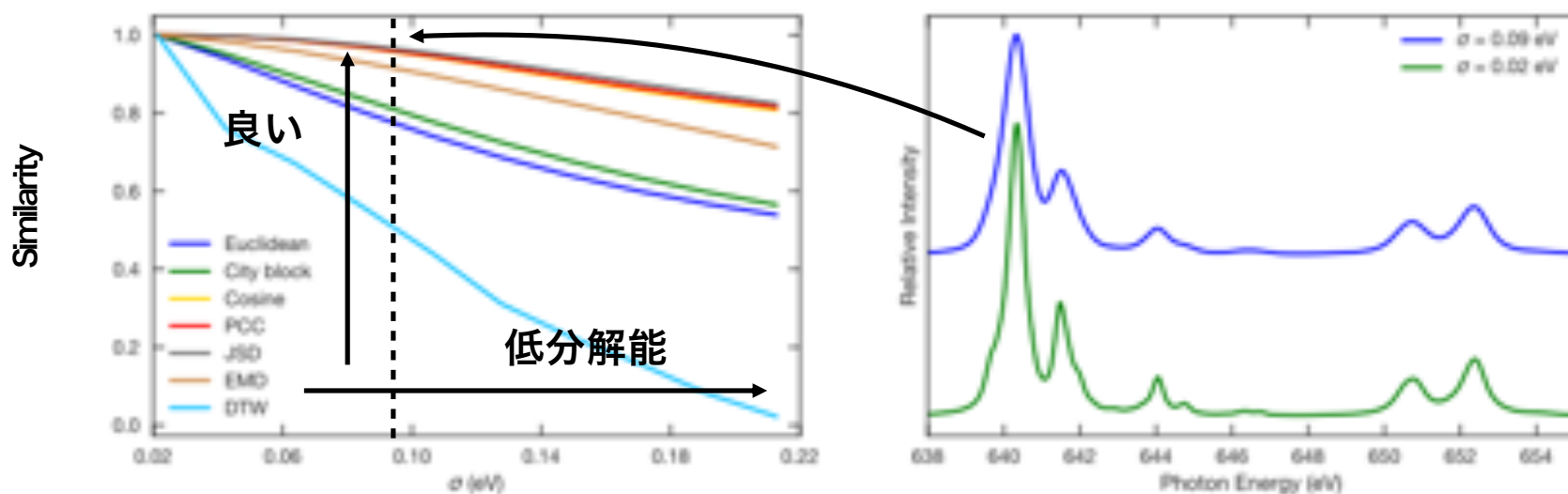
人工的にノイズを付加



- ピアソン相関係数やコサイン類似度がノイズに強い類似度になっている
- 適切な類似度 (距離尺度) を使う・作ることによってノイズに強い計測が可能に

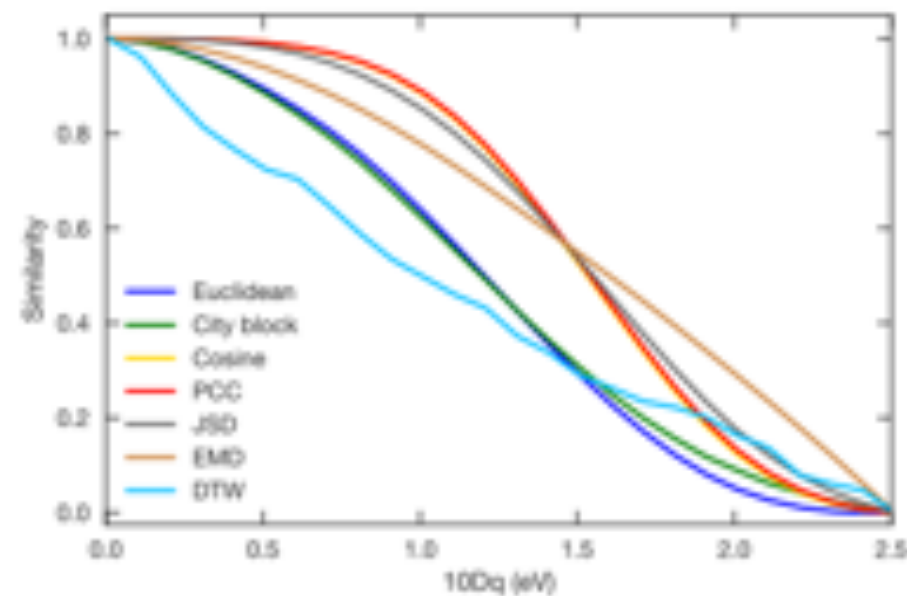
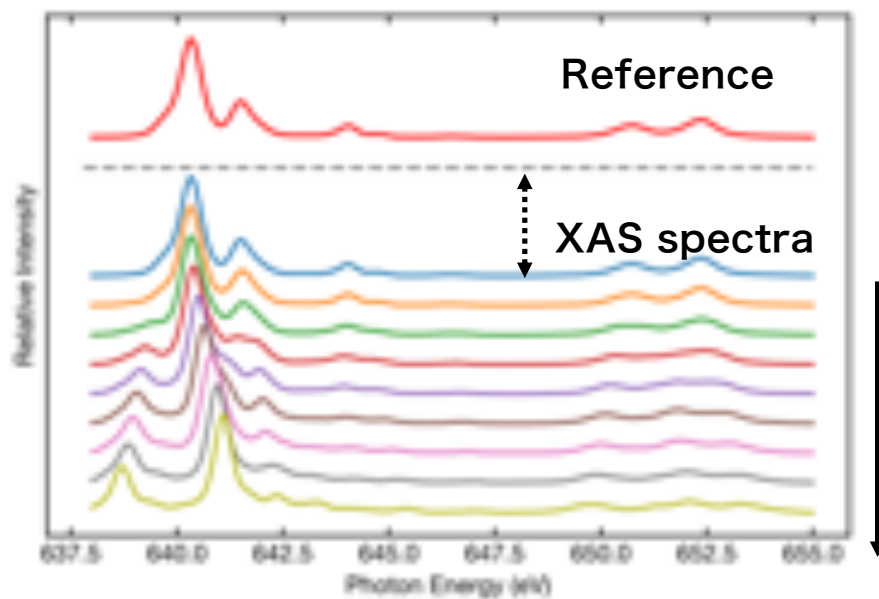
類似度に求められる性質：分解能に対する冗長性

分解能に対する類似度の冗長性



- ピアソン相関係数が分解能に強い類似度になっている
- 適切な類似度（距離尺度）を使う・作ることでノイズに強い計測が可能に

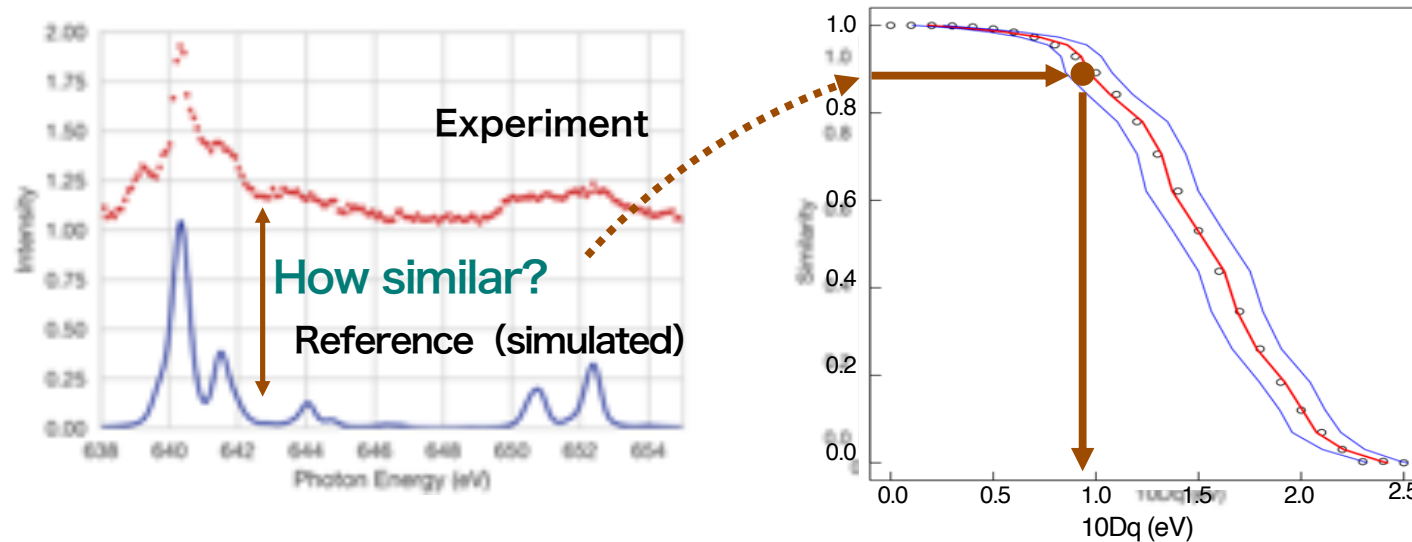
類似度に求められる性質：物理量への対応



物理量の自動推定

Estimate the physical parameter (10Dq) from experimentally obtained XAS spectrum of MnO

Y. Suzuki, KO et al., npj Comput. Mater. (2019)



- From the similarity with Pearson metric
Estimated **10Dq = 0.9 eV**
 - Identical to estimated value of professionals (0.9 eV)

We can estimate the physical parameter directly from XAS/EELS spectra

まとめ

物質・材料の計測のハイスループット化には下記の要素が重要だと考え、それぞれについて研究を行なった

- 統計的手法を用いた計測システムの効率化
 - カーネル密度推定 (KDE) を用いた小角散乱実験の高速化
- 最適な実験計画を自動で策定する方法論
 - ガウス過程を用いたX線スペクトルの適応型実験計画
- 計測データ解析の自動化
 - 次元縮約と類似度を用いたX線スペクトル解析
 - 機械学習によるX線回折パターンの解析

kanta.ono@kek.jp