

ベイズ最適化の基礎と材料工学への応用

松 井 孝 太* 金 森 研 太**
 豊 浦 和 明*** 竹 内 一 郎****

1. はじめに

「設計」や「計画」は、あらゆる科学的営みと切り離すことができない重要な問題の1つである。例えば、新たな科学的知見を得るために実験計画を立てる、あるタスクをより効率的に解決するためのロボットを設計する、疾患に有効な薬剤を設計する、天然資源を利用するための探鉱計画を立てる、というように、設計や計画が問題の根幹となっている例は枚挙に暇がない。近年、設計・計画問題に対するデータ駆動型アプローチ、また、それを実現するための情報技術である機械学習が注目されている。データ駆動型アプローチは、データ分析の方法だけでなくデータ収集の方法も考察の対象とする。特に後者の特性は、ものづくりなどデータ収集コストが非常に高い分野においては重要な意味を持つ。本稿では、データ収集と分析を統合的に扱うデータ駆動型アプローチの1つであるベイズ最適化⁽¹⁾に注目し、その基本的なアイデアを解説し、材料工学分野での研究を紹介する。

2. ベイズ最適化の基本概念

ある実験条件を入力 x とし、その条件の下での実験結果を観測 y としたとき、 x と y の間の関係 f は実験そのものであり、数学的に陽に書き下すことができない。このような f をブラックボックス関数と呼び、ベイズ最適化は図1のようなブラックボックス関数を含むシステムを最適化することを目的とした機械学習の方法である。

いま、我々の目的はブラックボックス関数 f を最大にする入力 $\mathbf{x}^*$ を見つけることとしよう：

$$\mathbf{x}^* = \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad (1)$$

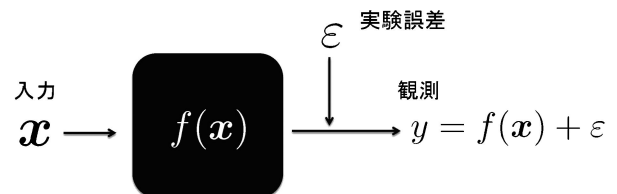


図1 ブラックボックス関数を含む入出力関係。

ここで、 $\mathcal{X}$ は、例えば可能な実験条件の集合を表す。 $\mathcal{X}$ が有限集合のとき、すなわち、 $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ のように可能な実験条件がリストアップされているような場合、この最適化問題をプール型と呼ぶ。一方で、実験条件の指定も最適化と同時にを行うとき、すなわち、 $\mathcal{X} = \mathbb{R}^d$ のように各パラメータが値を取りうる空間のみが与えられている場合をクエリー型と呼ぶ。以下では、最もシンプルな問題設定であるプール型の最適化について解説する。

ベイズ最適化の基本的な思想は、可能な限り少ない実験回数でブラックボックス関数 f の性質を理解し、これを最適化する、という点に集約される。そのため、ベイズ最適化は以下のプロセスの繰り返しとして実現される。

1. f に関する事前知識あるいは信念を統計モデル P として表現する。
2. P に基づいて次に実験すべき条件 $\mathbf{x}_{\text{next}}$ を $\mathcal{X}$ から選択する。
3. 実験条件 $\mathbf{x}_{\text{next}}$ において実験を行い、 y_{next} を観測する。 $(\mathbf{x}_{\text{next}}, y_{\text{next}})$ に基づいて P を更新する。

(1) ガウス過程による f のモデル化

まず1について、目的関数 f に対して統計モデルを導入する動機は、 f がブラックボックスであってもその予測分布を

* 理化学研究所革新知能統合研究センター；特別研究員

** 名古屋工業大学；大学院生

*** 京都大学工学研究科；准教授，理化学研究所革新知能統合研究センター；客員研究員

**** 名古屋工業大学工学研究科；教授，理化学研究所革新知能統合研究センター；チームリーダー，物質材料研究機構；招聘研究員（〒466-8555 名古屋市昭和区御器所町）

Introduction to Bayesian Optimization and Its Application to Material Engineering; Kota Matsui*, Kenta Kanamori**, Kazuaki Toyoura*** and Ichiro Takeuchi**** (*RIKEN Center of Advanced Intelligence Project, Tokyo. **Nagoya Institute of Technology, Nagoya. ***Kyoto University, Kyoto. ****Nagoya Institute of Technology, Nagoya. RIKEN Center of Advanced Intelligence Project, Tokyo. NIMS, Tsukuba)

Keywords: bayesian optimization, experimental design, blackbox function, gaussian process, acquisition function

2018年9月10日受理[doi:10.2320/materia.58.12]

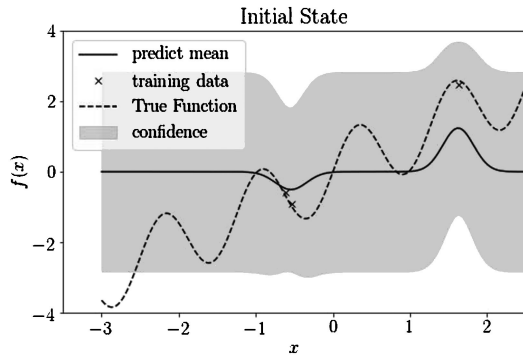


図2 ガウス過程によるブラックボックス関数 f のモデリング。破線が真の関数、実線がガウス過程による予測関数の平均、灰帯は予測関数の不確実性を表す。

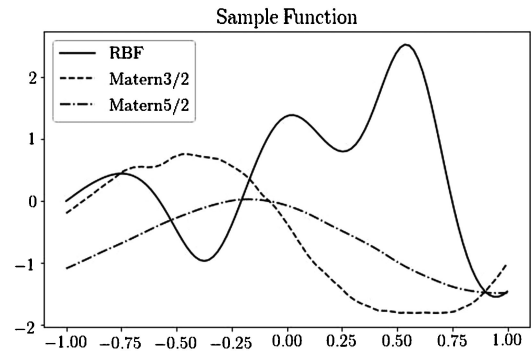


図3 各カーネルを用いたガウス過程からのサンプル。Matern3/2, Matern5/2, RBF の順により滑らかな関数がサンプリングされている。

求めることができる点にある。予測分布を考察することで、 f の予測値のみならず、その不確実性もモデル化することができる。ベイズ最適化の枠組みでは、 f の統計モデルとしてガウス過程(Gaussian process)⁽²⁾が用いられることが多い。ガウス過程とは、一言で言えば関数に対する正規分布であり、平均関数 $\mu(\mathbf{x})$ と共分散関数(カーネル関数) $k(\mathbf{x}, \mathbf{x}')$ によって特徴付けられる。図2にガウス過程による関数モデリングの例を示す。

(a) 平均関数と共分散関数の選択

ガウス過程モデルの事前分布の平均関数は、実応用では定数関数 $\mu(\mathbf{x}) \equiv \mu_0$ (特に $\mu_0 = 0$)とすることが多い。ベイズ最適化の進行とともに観測データに基づいて推定していく方針がよくとられる。

一方で、共分散関数は、目的関数の事前知識(滑らかさや振幅など)をモデルに反映させる、データ点に対して類似度を定義する(よく似た実験設定ならば出力も近い値をとることを期待する)といった重要な役割を持つ。本稿では、ベイズ最適化の文脈で良く用いられるRadial Basis Function (RBF)カーネルとMatern3/2カーネル、Matern5/2カーネルの3種類を紹介する。以下では $r = \|\mathbf{x} - \mathbf{x}'\|$ とする。

• RBF カーネル

$$k(\mathbf{x} - \mathbf{x}') := \exp\left(-\frac{r^2}{2\ell^2}\right) \quad (2)$$

• Matern3/2 カーネル

$$k(\mathbf{x} - \mathbf{x}') := \left(1 + \frac{\sqrt{3}r}{\ell}\right) \exp\left(-\frac{\sqrt{3}r}{\ell}\right) \quad (3)$$

• Matern5/2 カーネル

$$k(\mathbf{x} - \mathbf{x}') := \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right) \quad (4)$$

ここで、 ℓ はバンド幅を制御するパラメータである。それぞれのカーネルを用いたガウス過程からサンプリングされた関数の例を図3に示す。カーネルの選び方によって、どの程度の滑らかさを関数に仮定するかをコントロールできることが分かる。一般に、ブラックボックス関数は滑らかでないことが想定されるため、より滑らかでない関数を表現できる

Matern カーネルが用いられることが多い。

(2) 探索方針の考察

次に2について、「次にどの設定で実験をするべきか(図2の横軸上のどの点で関数値を評価すべきか)」を決定する際の基本的なアイデアは、ある点 x において実験を行い、観測値 y を得ることでどの程度最適解 $\mathbf{x}^*$ に関する情報が得られるかを考察することである。上述のように目的関数 f に対して統計モデルを導入することで、この問題に対して様々な指標に基づいてアプローチすることが可能となる。本節では、代表的な3つのアプローチを紹介する。以下では、 m 点 $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathcal{X}$ が既に観測されているとし、現在の最良の目的関数値を $\tau = \max_{i=1, \dots, m} f(\mathbf{x}_i)$ とおく。

(a) 改善確率に基づく探索

このアプローチでは、候補点 $\mathbf{x}$ が現在の最良関数値 τ を改善する確率(probability of improvement)を評価する。

$$\text{PI}(\mathbf{x}) := \Pr(f(\mathbf{x}) > \tau) \quad (5)$$

PIが最大となる $\mathbf{x} \in \mathcal{X}$ を次の評価点とすることで、目的関数値を改善する確率が高い点を次々と評価していく探索方針が得られる。

(b) 期待改善度に基づく探索

このアプローチでは、候補点 $\mathbf{x}$ が現在の最良関数値 τ を改善する期待値(expected improvement)を評価する。

$$\text{EI}(\mathbf{x}) := \mathbb{E}[\max\{f(\mathbf{x}) - \tau, 0\}] \quad (6)$$

ここで、 $\mathbb{E}[X]$ は X の期待値を、 $\max\{a, b\}$ は a と b のうち値の大きい方を返す関数である。EIが最大となる $\mathbf{x} \in \mathcal{X}$ を次の評価点とすることで、目的関数の期待改善度が高い点を次々と評価していく探索方針が得られる。

(c) 上側不確実性に基づく楽観的探索

このアプローチは、図2における青実線の上側の青帯の高さ(upper confidence bound)を評価する。

$$\text{UCB}(\mathbf{x}) := \mu(\mathbf{x}) + \beta\sigma(\mathbf{x}) \quad (7)$$

ここで、 $\sigma(\mathbf{x})$ は共分散関数 k に基づいて定まる $\mathbf{x}$ の分散、 β は探索と活用のトレードオフ(分散 $\sigma(\mathbf{x})$ を重視するか、平均 $\mu(\mathbf{x})$ を重視するか)をコントロールするという役割を持つ。

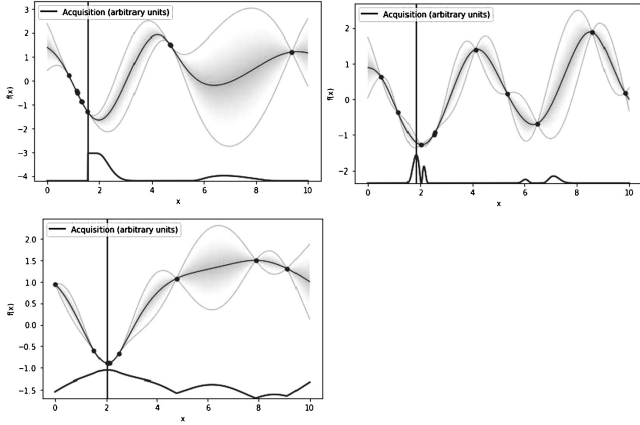


図4 (左上), (右上), (左下)がそれぞれPI, EI, UCBの獲得関数(下部の実線)を表す. いずれも非凸関数となっていることが分かる.

本節で紹介したPI, EI, UCBは獲得関数(acquisition function)と呼ばれ, ベイズ最適化において本質的な役割を果たす. 一方で, 「どの獲得関数を用いれば, より少ない探索回数で f の最適解が得られるか」は当然ながら問題に依存する. 問題の構造を良く理解した上で, 適切な獲得関数を選択(場合によっては新たに設計)することが重要である.

また, 獲得関数は一般に非凸関数となるため, 大域最適解を求めることが難しい(図4参照). プール型の最適化問題の場合, 全ての候補点で獲得関数の値を評価することができれば可能であるが, 候補点の数が多い場合には非現実的になってしまう. 現実的な方法としては, マルチスタートの非線形最適化手法によって良い局所最適解を見つけるといったヒューリスティクスが用いられる.

3. 回帰におけるベイズ最適化

本節では回帰問題, すなわち, 図1における y が実数値を取る場合($y \in \mathbb{R}$)におけるベイズ最適化を概説する.

回帰問題の場合, ガウス過程によるモデリングには, データ $D_n = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\}$ を観測した下で, 新たな点 x における関数値 $f(x)$ が正規分布し, その事後平均と事後分散(すなわち統計モデルの更新式)が解析的に書ける, という利点がある.

$$f(\mathbf{x}) | D_n \sim N(\mu_m(\mathbf{x}), \sigma_m(\mathbf{x})),$$

$$\mu_m(\mathbf{x}) = \mu(\mathbf{x}) + x(\mathbf{x})^\top (K + \sigma^2 I)^{-1} (y - \mu)$$

$$\sigma_m^2(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - k(\mathbf{x})^\top (K + \sigma^2 I)^{-1} k(\mathbf{x})$$

ここで, $k(\mathbf{x}) = (k(\mathbf{x}_1, \mathbf{x}), \dots, k(\mathbf{x}_m, \mathbf{x}))^\top$, $K_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{y} = (y_1, \dots, y_m)$ であり, I は単位行列, σ^2 は実験誤差 ε の分散を表す.

$\mu_m(\mathbf{x})$ と $\sigma_m(\mathbf{x})$ に基づいて2.(2)節で示したような獲得関数が定義され, その最大値を与える点が次の探索点として選択される. 図5に回帰におけるベイズ最適化の実行例を示す.

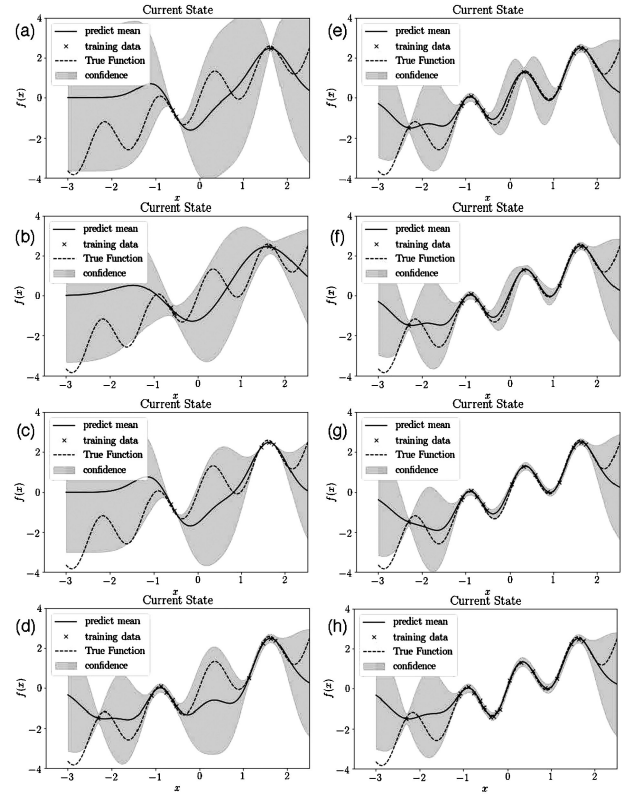


図5 回帰におけるベイズ最適化の実行過程((a)から(h)に向かって最適化が進行する). 獲得関数にはUCBを使用した. 山の高い部分にある程度調べて次の山に移るといった振る舞いを示す. 関数の極大となる点を重点的に探索する様子がみてとれる.

4. 判別におけるベイズ最適化

これまでの, 入力 $\mathbf{x}$ に対して出力 $\mathbf{y}$ が連続値である場合を考察してきた. 一方, 現実の問題では出力が連続値を取らない状況も想定される. 例えば, 関数 f の最大値ではなく, f がある閾値を超えるかどうかのみ興味がある場合を考えよう. 簡単のため, 閾値を0とし, $f(\mathbf{x}) > 0$ のとき成功 ($y=1$), $f(\mathbf{x}) \leq 0$ のとき失敗 ($y=0$) と定義する. すると問題は, 「ノイズの含まれる2値の出力から成功する確率が最も高い $\mathbf{x}$ をできるだけ少ない探索回数で列挙する」という2値判別問題とみなすことができる.

出力が2値のとき, 回帰の場合のように $f(\mathbf{x})$ の事後平均や事後分散を解析的に求めることができない. 従って, 何らかの方法で $f(\mathbf{x})$ の事後分布を近似して計算する必要がある. 詳細は割愛するが, これまで観測した点における関数値 $\mathbf{f} = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_m))$ の事後分布 $p(\mathbf{f} | D_m)$ をラプラス近似によって正規分布で近似することで, 事後平均や事後分散を回帰の場合とほぼ同じ形で書き下すことができる⁽²⁾. このようにして求めた近似的な事後分布に基づいて獲得関数の設計, 統計モデルの更新を行う.

図6に判別におけるベイズ最適化の実行例を示す. 図6

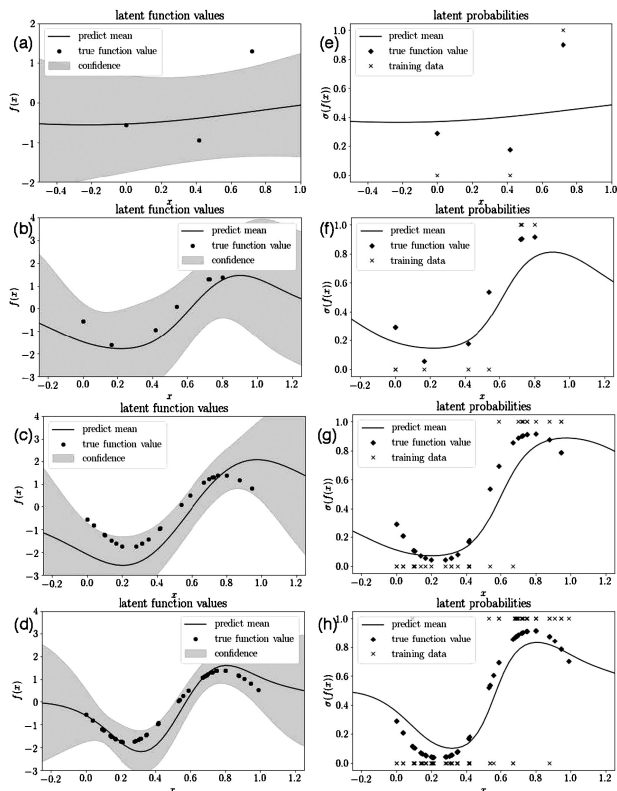


図6 判別におけるベイズ最適化の実行過程(上から下に向けて最適化が進行する). 獲得関数にはUCBを使用した. 左図が目的関数 f のモデリング, 右図が予測確率のモデリングをそれぞれ表す.

の左列は目的関数 f のガウス過程によるモデリングとそのベイズ最適化による更新過程を表している. 図6の右列は,

$$y = \begin{cases} 1 & f(\mathbf{x}) > 0 \Leftrightarrow p(y|f(\mathbf{x})) > \frac{1}{2} \\ 0 & f(\mathbf{x}) \leq 0 \Leftrightarrow p(y|f(\mathbf{x})) \leq \frac{1}{2} \end{cases}$$

なる対応によって左列のベイズ最適化プロセスを予測確率の世界に写したものである.

5. 材料工学への応用

最後に, ガウス過程モデルとベイズ最適化を材料科学の課題に利用した筆者らの最近の研究を簡単に紹介する⁽³⁾⁽⁴⁾. 電池材料などのイオン伝導性物質の伝導度を知るには, ポテンシャルエネルギー曲面内のイオン伝導経路を同定する必要がある. イオン伝導経路を同定できれば, その経路内の最安定点(エネルギー最小の点)とボトルネック点(経路内でエネルギー最大の点)を知ることができ, イオン伝導度を求めることができる. 第一原理計算などの物理シミュレーションを利用すると, コンフィギュレーション空間の各点におけるポテンシャルエネルギーを高精度に求めることができるが, ポテンシャルエネルギー関数全体を網羅的な第一原理計算で求めようとすると膨大な計算コストがかかってしまう. このた

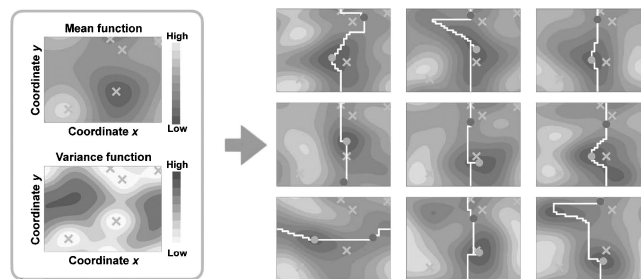


図7 2次元のガウス過程モデルとランダムサンプリングされたPE曲面における電導経路推定の概念図.

め, ガウス過程モデルとベイズ最適化の考え方を拡張し, イオン伝導経路を特徴づける部分に対して選択的に第一原理計算を行うアプローチを検討した.

2.(2)節で議論したように, ベイズ最適化では興味の対象となる物性値の予測分布を求め, 適切な獲得関数を設計する必要がある. しかしながら, ガウス過程モデルからイオン伝導経路の予測分布を直接求めることはできないため, 筆者らは1)ガウス過程モデルからのモンテカルロサンプリング, 2)動的計画法によるイオン伝導経路推定, 3)最安定点とボトルネック点に関する獲得関数を用いたベイズ最適化の3ステップのアプローチを提案した. ステップ1では, ガウス過程モデルが関数に対する確率分布であることを利用し, モンテカルロサンプリングによってポテンシャルエネルギー関数のランダムサンプルを数多く生成した. ステップ2では, ステップ1で生成したそれぞれのポテンシャルエネルギー関数に対して動的計画法を適用することで最適なイオン伝導経路を同定し, 最安定点とボトルネック点のランダムサンプルを得た. ステップ3では, ステップ2で得た最安定点とボトルネック点のランダムサンプルに基づいた獲得関数を設計し, 次に第一原理計算を適用すべきコンフィギュレーション点の選択を行った. これらの3ステップを繰り返すことで, イオン伝導経路に関する不確実性をモデリングし, モデルに基づくイオン伝導経路の推定を効率的に行うことができた.

紙面の都合上, 詳細な説明は省略するが, 図7に提案法のイメージを示す. 2次元空間のポテンシャルエネルギー(PE)曲面をガウス過程によってモデル化すると予測分布の平均関数と分散関数は図7左のように表される. また, このガウス過程モデルからランダムサンプリングによってPE曲面の候補を多数作成し, 各候補に動的計画法を適用してイオン伝導経路を求めることで, イオン伝導経路の予測分布を推定することができる. このアプローチをプロトン伝導体の伝導度推定に利用することで, 網羅的计算に比べて大幅に計算コストを削減することができた.

6. おわりに

材料科学の問題にデータ科学のアプローチを利用する試みは材料情報学(Materials Informatics)と呼ばれ注目を集めて

本研究の一部は、科学研究費(17H00758, 16H06538), JST CREST (JPMJCR1302, JPMJCR1502), 理化学研究所革新知能統合研究センター, および, JST イノベーションハブ構築支援事業・情報統合型物質・材料開発イニシアティブの補助を受けて行われた。

- (1) B. Shahriari, K. Swersky, Z. Wang, R. Adams and N. D. Freitas: Proceedings of the IEEE, **104**(2016), 148–175.
- (2) C. E. Rasmussen and C. K. I. Williams. Gaussian: Processes for Machine Learning. MIT Press, (2006).
- (3) K. Kanamori, K. Toyoura, J. Honda, K. Hattori, A. Seko, M. Karasuyama, K. Shitara, M. Shiga, A. Kuwabara and I. Takeuchi: Phys. Rev. B, **97**(2018), 125124.
- (4) K. Toyoura, D. Hirano, A. Seko, M. Shiga, A. Kuwabara, M. Karasuyama, K. Shitara and I. Takeuchi: Phys. Rev. B, **93**(2016), 054112.
- (5) J. M. Hernández-Lobato, M. Gelbart, R. Adams, M. Hoffman

竹内一郎