

研究報告書

「機械学習に基づく効率的な界面物性探索法の開発」

研究タイプ: 通常型

研究期間: 2015 年 12 月～2019 年 3 月

研究者: 烏山 昌幸

1. 研究のねらい

情報科学と材料科学の融合により新規材料の発見を加速する材料情報学という分野が広く注目を集め始めている。材料情報学は、これまで専門家が経験と勘により、試行錯誤的に行ってきた材料探索を統計科学的な推論により自動化する試みである。人的なボトルネックを解消し、統計的に客観的な意思決定が実現されることで探索効率を劇的に向上できる可能性がある。しかし一方、計算の簡単さからこれまでのほとんどの研究では、解析の対象が完全結晶に限定されてきた。現実には、界面や粒界と呼ばれる結晶粒の作る境界面での物理的な性質の変化が材料のパフォーマンスに大きく影響するため、これを無視することはできない。例えば、普及が進む電気自動車業界から強い注目を集める全固体電池では、電解質材料内の粒界で生じる抵抗がボトルネックとなり得る。本課題では主に結晶粒界をターゲットとして、局所的な乱れを含む材料の情報学的な解析方法論を構築する。解析を困難にする粒界特有の事情として以下の3点に着目する。

- 1) 粒界近傍は完全結晶のような均一な構造を持たないため、系の中の平均的な性質を観察するだけでは不十分であり、粒界面近傍での局所的な変化を解析する必要がある。
- 2) 結晶の回転角や並進移動などの自由度により、あり得る構造を大量に考えることができってしまうため、効率的に重要な候補構造を絞らなければならない。
- 3) 粒界の理論計算を考える場合、周期境界条件により、しばしば非常に大きな構造モデル（ユニットセル）を考える必要がある。そのように大きなモデルは通常、原子数が大きくなり、計算の実行がより困難になる。

これらの問題を解決する汎用的な方法論を構築し、粒界解析のスピードを大幅に向上することで、より広大な材料空間を効率的に探索する材料情報学の枠組みを構築することが本課題の目標となる。また、材料科学での情報学的アプローチのさらなる普及を目指し、粒界に限らず材料情報学に機械学習の新たな方法論を導入する研究も同時に視野に入れていく。

2. 研究成果

(1) 概要

ここでは主要な研究成果を、テーマ A. 粒界原子エネルギー予測 [1], テーマ B. 粒界構造探索 [2], テーマ C. その他の材料情報学研究 [3], に分けて述べる。テーマ A は、粒界近傍の原子エネルギーを機械学習でモデル化する研究であり、原子環境が局所的に乱れる粒界近傍でエネルギーの局所的な振る舞いを観察する高速かつスケーラブルな方法論を構築した。テーマ B では、高い自由度を持つ粒界構造を探索する最適化法を開発した。多様な回転角間で転移学習による知識共有を行った上で、計算コストを陽に考慮するコスト考慮型探索を行う点が特徴的な点である。テーマ C は粒界以外の材料情報学研究であり、プロトン伝導

体の拡散経路を機械学習で推定する新たな枠組みを開発したものについて述べる。

(2) 詳細

テーマ A. 粒界原子エネルギー予測 [1]:

粒界近傍で局所的にどうエネルギーが変化するかは、非常に重要な課題である。密度汎関数法(DFT)に基づく原子エネルギーの計算法に(Shiuhara, et. al., Phys. Rev. B, 2011)がある。この方法は、ベーダー分割と呼ばれる定義に基づき、電子密度を分割し局所的な原子エネルギーを定めることができ、粒界近傍のエネルギー分布を観察できる。ただし、DFTに基づくため計算量が大きく、膨大な候補のスクリーニングには適用できず、また、セルサイズの大きな粒界が候補に含まれる場合、計算はより困難になってしまう。

そこで、機械学習の LASSO 回帰モデルを拡張し、原子エネルギーを高速に予測する手法を開発した。図1に処理の全体像を示す。入力となる粒界構造は計算の軽いポテンシャルで緩和したものを用意し、訓練データとなる DFT 原子エネルギーは、小さな粒界モデルを用いて用意することで DFT 計算にかかる時間的コストを最小限に抑えた。各原子周辺の局所的な環境を動径分布関数や SOAP(Smooth Overlap of Atomic Positions)記述子(De, et. al., Phys. Chem. Chem. Phys. 2016)を用いて抽出し、それらの線形結合により DFT の結果を再現する。さらに、原子エネルギーと系全体のエネルギーの両方を fitting に用いることで更なる精度改善を図った。

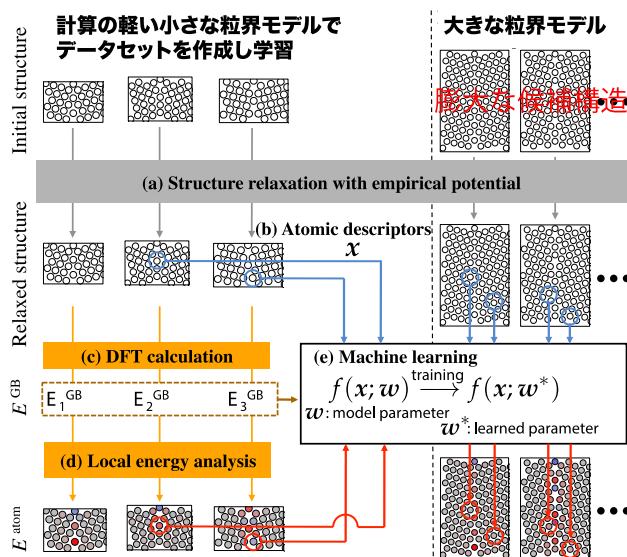


図 1 機械学習による原子エネルギー予測の処理フロー。

このように学習されたモデルが、訓練データより大きな粒界に対して正確な予測が可能なのか FCC-A1 の傾角対称粒界を用いて検証した。図2に示す通り、構築した機械学習モデルは非常に高い精度で原子エネルギーを予測することができた。粒界の周期性の指標である Σ 値が 20 以下のものでも学習し、 Σ 60 程度までの粒界を予測した場合には、予測値と DFT の値の相関係数は約 0.98 と非常に高い値を達成した。この結果は、粒界における局所構造同士

の類似性によって解釈できる。粒界はセルサイズが大きく違っても、面近傍に現れる基本的な局所構造は少数の構造ユニットで記述できることが多いことは知られていた。そのため、一見すると、完全な外挿補間に思われるような、小さな粒界から推定したモデルによる大きな粒界の予測でもかなりの精度が達成されることがわかった。論文では、記述子の空間内で、原子環境同士の距離分布を観察し、これを検証した。また、いくつかの記述子を比較検証し、DFT 原子エネルギー予測には SOAP が非常に有効であることも確認した。

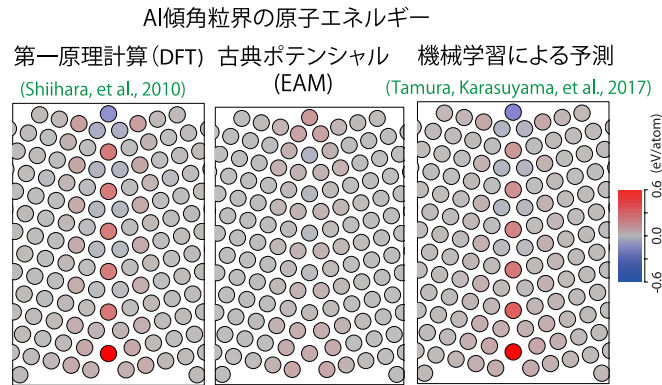


図 2 機械学習による原子エネルギー予測の例

テーマ B. 粒界構造探索 [2]:

安定構造の決定は結晶構造の性質を知るうえで最も基本的な情報の一つであるが、粒界においては回転角や軸、並進移動など、完全結晶では考えなかった自由度を考慮する必要がある。よく用いられる調査法に、粒界の回転角を変えながら、それぞれの角度における並進移動の最安定エネルギーを探し、プロットする方法がある。このようなプロットを作ることで、どのような角度に特異的に安定な構造が現れるのか調べることができる。しかし、多様な回転角それぞれの安定構造を探すため、計算量が非常に多く、単一の組成について調べるだけでも、 naïve な全計算では 10 万回を超える安定構造計算が必要になることも多い。

本課題では、機械学習におけるマルチタスク学習 (転移学習) とコスト考慮型戦略を導入した探索方法論を構築した (図3)。提案法では、多様な角度の探索タスクそれぞれで情報を共有することで、探索を加速する。情報の共有は構造の類似性を記述子により判断して行う。つまり、構造が似ているほど、同様のエネルギー曲面を持つ可能性が高いと確率的に判断する。特に恩恵が大きいのが、セルサイズの小さい粒界で得た曲面を、セルサイズの大きい粒界と共有できる場合である。提案法では、コスト考慮型の戦略を導入し、コスト効率の高いものから順番に探索をしていく。この戦略では探索の初期段階ではコストの低い粒界が選択され、エネルギー曲面の情報が蓄積された後で、よりコストの高い大きな粒界の計算へとアルゴリズムが自動的に進んでいくことで全体の効率を最大化する。

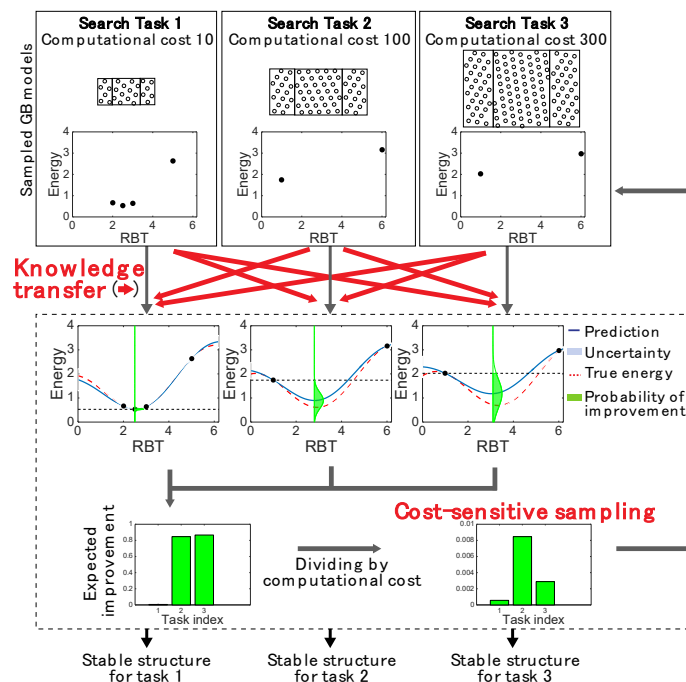


図 3 コスト考慮型マルチタスクベイズ最適化による粒界探索.

提案した探索アルゴリズムの探索効率を FCC-AI を用いて検証したものが図4である. 横軸が回転角, 縦軸が粒界エネルギーで, 縦軸の粒界エネルギーが小さいほど安定性が高いことを意味する. 回転角を完全結晶(0度)から変えて境界面を作っていくと, 粒界エネルギーが不規則に変化していく. 提案法は網羅調査の 0.3% 程度のコストしかかけていないにもかかわらず, ほぼ最安定なエネルギーを検出していることがわかる.

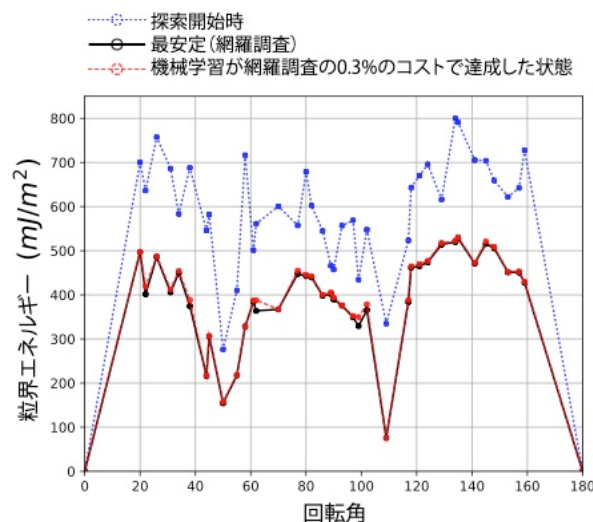


図 4 提案法の効率評価. 赤い線が, 機械学習が網羅探索の 0.3% のコストで到達した状態.

テーマ C. その他の材料情報学研究 [3]:

材料情報学の普及に貢献するため、粒界以外の材料情報学に関する研究も期間中に推進した。ここでは、そのうちの一つとして、プロトン伝導体の伝導経路解析に機械学習を活用する方法について簡単に述べる。DFT によるプロトン伝道経路解析には、まずプロトンのサイトを発見し、サイト間を nudged elastic band(NEB)法でつなぐ方法がよく用いられるが、この方法は、初期経路として、最適な経路に近いものを与える必要がある。この研究課題は、ベイズ最適化の考えを拡張し、プロトンの安定点と経路上のボトルネック点を効率的に発見し、確率的にポテンシャル障壁を評価する枠組みを開発したものである。

3. 今後の展開

本研究で構築した材料情報学ツールをより実践的な状況へ展開していく。粒界の原子エネルギー予測については、構築した機械学習モデルにより実際に多量の候補に原子エネルギーの予測値を与えスクリーニングを行う。また、多元系などより複雑な系への適用を目指す。さらに、小さな粒界から大きな粒界が予測可能であるという知見をもとに、なるべく小さな粒界によって、予測精度を最大限高めるような訓練データ集合の如何に集めるかの方法論を構築する(これは機械学習で能動学習と呼ばれる枠組みに相当する)。粒界の構造探索法についても同様に、実践的な設定への投入を目指す。特に電解質材料など、粒界解析の重要性の高いターゲットへの適用を考えていく。また、ソフトとしてパッケージ化し、フリーで誰でも利用できる形で整備することで、研究成果の波及を目指す。

4. 自己評価

研究目的の達成状況/研究の進め方:

研究開始当初は情報科学的観点と材料科学的な観点の違いから具体的な方向性を見定めるのに苦心したが、研究が進むにつれ道筋がかなりクリアに見えてくるようになった。粒界近傍での局所的な DFT エネルギーの振る舞いを予測する機械学習モデルの構築と粒界の構造探索アルゴリズムの開発、さらにはそれらの中でセルサイズの影響によるコストの違いを考慮し粒界計算のスケラビリティを向上する試みは、最終的には当初想定した枠組みに相当するものが構築できた。エネルギー予測において、小さな粒界から大きな粒界が予測できたことは、機械学習による粒界スクリーニングを考える上で非常に重要であり、これを活かして大量粒界のスクリーニングの実践や、コストを考慮した訓練データ収集といった次の展開を考えていくことが可能になった。構造探索でも、転移学習との組み合わせで、コストを考慮しつつ情報を転移する枠組みが構築できた。

一方で今後の大きな課題として残されたのは、構築したアルゴリズムの実践投入である。実用的な枠組みが構築できたものの、実際の新規材料/構造の発見に期間内に至ることはできなかった。これを達成することが、今後の材料情報学の未来に関わる重要な課題となると考えている。

研究成果の科学技術及び社会・経済への波及効果:

世界的な競争激化が進む次世代材料の開発において粒界の解析は最重要課題の一つであるにも関わらず、計算の難しさや、探索すべき候補の膨大さから取り扱いが容易ではないとされてきた。膨大な候補に対して、人的、時間的、費用的コストのかかる実験や計算をいきなり行うことは損失が非常に大きい。そのため、粒界や界面のように複雑性が高く、先見知識の効かない対象こそ、機械学習のようなコストの低いアプローチによるスクリーニングが必要不可欠となる。本研究で構築した枠組みは、粒界を如何に低コストにスクリーニングするかの一般方法論を与えており、今後の材料情報学の発展とともに広く普及していく可能性があると考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. T. Yonezu, T. Tamura*, I. Takeuchi, M. Karasuyama*, Knowledge-Transfer based Cost-effective Search for Interface Structures: A Case Study on fcc-Al [110] Tilt Grain Boundary, <i>Physical Review Materials</i> , vol. 2, no. 11, 113802, 2018
2. T. Tamura*, M. Karasuyama*, R. Kobayashi, R. Arakawa, Y. Shiihara, and I. Takeuchi, Fast and Scalable Prediction of Local Energy at Grain Boundaries: Machine-learning based Modeling of First-principles Calculations, <i>Modelling and Simulation in Materials Science and Engineering</i> , vol.25, no.7, 075003, 2017
3. K. Kanamori, K. Toyoura*, J. Honda, K. Hattori, A. Seko, M. Karasuyama, K. Shitara, M. Shiga, A. Kuwabara, and I. Takeuchi*, Exploring a potential energy surface by machine learning for characterizing atomic transport, <i>Physical Review B</i> , vol.97, no.12, 125124, 2018.
(*: corresponding author)

(2) 特許出願

研究期間累積件数:0 件

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

招待講演

- ものづくり企業に役立つ応用数理手法の研究会, 「統計的機械学習による材料データ解析: データ駆動材料発見に向けて」, 2018 年 10 月 29 日
- 応用物理学会秋季学術講演会, 「機械学習による粒界データ解析: データ駆動型材料探索に向けて」, 2018 年 9 月 18 日

- 第27回日本 MRS 年次大会シンポジウム「計算機シミュレーションによる先端材料の解析・機能創成」招待講演, 「統計的機械学習に基づく粒界データ解析: データ駆動型材料探索に向けて」, 2017 年 12 月 5 日

国際会議

- T. Tamura, M. Karasuyama, R. Kobayashi, R. Arakawa, Y. Shinhara, and I. Takeuchi, Fast and scalable prediction of local energy at grain boundaries: Machine-learning based modeling of first-principles calculations, 9th International Conference on Multiscale Materials Modeling (MMM), 2018

プレスリリース

- 名古屋工業大学「互いに協調し進化する人工知能システムで多数の複雑な界面の安定構造を同時に超高速探索する新手法を開発」 2018 年 11 月 27 日
<https://www.nitech.ac.jp/news/press/2018/7125.html>