

機械学習と第一原理計算の融合による半導体格子欠陥の原子構造・特性の解明

名古屋大学 大学院工学研究科 物質科学専攻

横井 達矢

1. テーマ設定の背景

半導体材料は、既存技術だけでなく、超高速通信や量子計算、再生可能エネルギーといった次世代の通信・エネルギー技術でも不可欠である。他方、実材料中には材料作製に応じて、点欠陥や転位、粒界といった局所的な原子配列の乱れ『格子欠陥』が必ず存在し、材料特性に多大な影響を及ぼす。その影響には有益・有害の両面があり、またバルク材にはない新奇特性の起源となる可能性もある。よって、次世代技術の高度な要求を満たす半導体材料の創成には、格子欠陥の原子・電子レベルの知見とそれに根差した材料設計が必須である。

先行研究では、量子力学にもとづく密度汎関数理論 (density functional theory, DFT) 計算を用い、格子欠陥の原子配列や形成エネルギー、電子構造が理論解析されてきた。しかし格子欠陥の原子配列は完全結晶に比べて対称性が低く、また格子欠陥の種類によっては多数の原子が関与する原子配列の乱れを伴う。従って、数百～数千原子に及ぶ大規模計算セルがしばしば必要となる。また平衡状態の原子配列や格子欠陥特性の評価には、多数の構造緩和計算や長時間の分子動力学 (molecular dynamics, MD) 計算が必要となる。通常のDFT計算の計算時間は、電子数の3乗に比例するため、上記のような解析は困難である場合が多い。よって多くのDFT計算は、孤立した点欠陥や低指数の界面など単純な系に限定されており、実際の半導体中の格子欠陥に関する系統的理解には程遠い。

上記に対し、筆者らはDFT計算データを人工ニューラルネットワーク (artificial neural network, ANN) に学習させたANN原子間ポテンシャルを構築し、構造緩和やMD計算などの分

子シミュレーションに統合した。これにより、DFT計算を直接行うことなく、高精度で系のエネルギーや原子にかかる力的高速予測が可能か検証した。そして本手法をSiやGe、化合物半導体に適用し、格子欠陥の原子配列や安定性、特性を予測した。そして最終的にANN原子間ポテンシャルの予測とDFT計算を比較し、予測精度を検証した。

2. 素形材分野との関連性

本研究は、実際の半導体材料に必ず存在する格子欠陥に注目し、原子配列と特性に関する原子・電子レベルの系統的理解を実現するため、ANNに基づく理論計算手法を構築した。この内容は、既存材料だけでなく新規材料にも不可欠な内容であり、材料科学および格子欠陥物理を大きく発展させる可能性をもつ。

3. 研究開発の成果

本稿では、Si粒界を対象として構築したANN原子間ポテンシャルの研究結果を示す。図1は学習データの誤差を示している。学習データには、完全結晶だけでなく、点欠陥や表面、低 Σ 値の対称傾角粒界を含めた。各データ点は個々の学習データに対応しており、対角線に近いほどDFT計算との誤差が小さいことを意味する。この図より、系のポテンシャルエネルギーおよび原子にかかる力ともに、データ点は対角線付近に分布している。よって、粒界の学習データに対しても学習が十分進行している。全学習データに対する平均絶対誤差 (mean absolute error, MAE) は、エネルギーが5.7 meV/atom、原子にかかる力が111 meV/Åとなった。

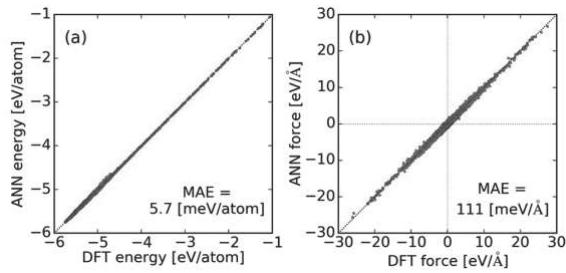


図1 学習データに対する誤差: (a) 系のエネルギー、(b) 原子にかかる力。

また図2(a)では、粒界の学習データについてANN原子間ポテンシャルと従来の経験的原子間ポテンシャルのMAEを比較している。この結果より、経験的原子間ポテンシャルの中で最もMAEが小さい Tersoffポテンシャルでも、ANN原子間ポテンシャルに比べてMAEが約3倍大きいことが分かる。図2(b)は各データ点の誤差を示しているが、Tersoffポテンシャルでは対角線から大きくずれた点が複数存在しており、高精度予測は困難であるといえる。

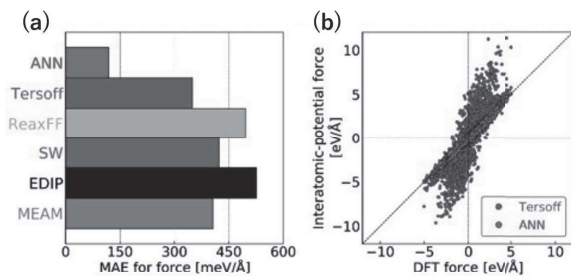


図2 $\Sigma 5(210)$ 粒界の学習データに対する誤差: (a) ポテンシャル間のMAEの比較、(b) ANN/Tersoffポテンシャルの比較。

図3は学習データとは異なる回転軸や界面をもつ粒界に対するエネルギーの誤差を示している。この結果より、 $\Sigma 11(113)$ 粒界を除いて、ほぼ全てのデータ点が対角線付近に分布している。またMAEは1.7-3.1 meV/atomと、学習データに近い値であった。したがって今回構築したANN原子間ポテンシャルは、学習データにない粒界に対しても高い予測能力を維持することが示された。なお、ANN原子間ポテンシャルが $\Sigma 11(113)$ 粒界のエネルギーを全体的に過小評価する理由は、学習データの種類が不十分であり、十分内挿できていない原子環境があったためと考えら

れる。この点は、今後一度学習させたANN原子間ポテンシャルを用いて、カバーできていない原子環境を効率的に探索することで解決できると考えられる。

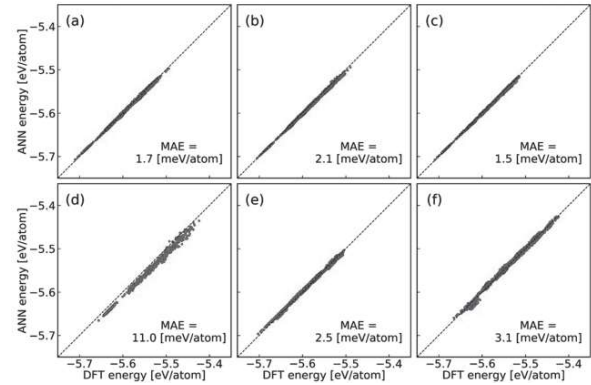


図3 テストデータの粒界に対するエネルギーの誤差: (a) $\Sigma 13(320)/[001]$ 、(b) $\Sigma 25(430)/[001]$ 、(c) $\Sigma 41(540)/[001]$ 、(d) $\Sigma 11(113)/[\bar{1}10]$ 、(e) $\Sigma 21(15\bar{4})/[111]$ 、(f) $\Sigma 5(001)/[001]$ ねじり粒界。

4. 訴求点

本研究では機械学習と理論計算を融合させ、半導体における格子欠陥の原子構造と形成エネルギーを高精度・高速で予測する手法を構築した。この手法は本稿で述べたSiだけでなく、多元系の化合物半導体や、さらには半導体以外の物質も同様に適用できる。よって波及効果の大きい内容であるといえる。

参考文献

- 1) T. Yokoi et al., Phys. Rev. Mater. 4 (2020) 014605.