

「理論・実験・計算科学とデータ科学が連携・融合した先進的マテリアルズインフォマティクスのための
基盤技術の構築」研究領域 領域活動・評価報告書
－2019年度終了研究課題－

研究総括 常行 真司

1. 研究領域の概要

計測・分析技術の進歩、コンビナトリアル合成などのハイスループット実験手法の発展、スーパーコンピュータに代表される計算機能力の飛躍的向上、第一原理計算などの強力な計算科学から得られる高精度な知見などにより、物質・材料科学における原理解明が進むとともに関連するデータが短時間で大量に得られるようになってきました。また、大量かつ複雑なデータからそれらの持つ意味や新しい知識を引き出す情報科学技術の進展もめざましいものがあります。

本研究領域では、これら実験科学、理論科学、計算科学、データ科学の連携・融合によって、それぞれの手法の強みを活かしつつ相互に得られた知見を活用しながら新物質・材料設計に挑む先進的マテリアルズインフォマティクスの基盤構築と、それを牽引する将来の世界レベルの若手研究リーダーの輩出を目指します。

具体的には、

- 1) 社会的・産業的に要求される機能を実現する新物質・材料の発見の促進、設計指針の構築
- 2) 大規模・複雑データから構造・物性相関や物理法則を帰納的に解明する手法の開発とそれを用いた新材料の探索・設計
- 3) 未知物質の物性を高精度に予測し、合成・評価の実験計画に資する候補物質を高速・大量にスクリーニングする手法の構築
- 4) 多種多様な物質データを包括的に整理・記述・可視化する新しい物理的概念や方法論の構築
- 5) データ科学と物質・材料科学の連携・融合に資する物性データ取得・蓄積・管理手法の開発、データベースの整備、各種計算・解析ツールの構築

などの研究を対象とします。

研究推進にあたっては、情報科学研究者と物質・材料科学研究者等が連携し互いに触発しながらシナジー効果を得る体制を整え、エネルギー、医療、素材、化学など多くの産業応用に資する物質・材料の設計を劇的に加速しうる先駆的・革新的な研究を推進し、物質・材料科学にパラダイムシフトを起こすことを目指します。

2. 事後評価対象の研究課題・研究者名

件数： 13件

※研究課題名、研究者名は別紙一覧表参照

3. 事前評価の選考方針

選考の基本的な考えは下記の通り。

- 1) 選考は、「理論・実験・計算科学とデータ科学が連携・融合した先進的マテリアルズインフォマティクスのための基盤技術の構築」領域に設けた選考委員11名の協力を得て、研究総括が行う。
- 2) 選考方法は、書類選考、面接選考及び総合選考とする。
- 3) 選考に当たっては、さきがけ共通の選考基準

(URL: <http://www.jst.go.jp/pr/info/info1128/index.html> の他、以下の点を重視した。

本研究領域は、これまで物質・材料科学を牽引してきた物理学・化学・材料工学など単独の学術分野だけでは達成しえない目標を多く含み、情報学・統計学・数理科学などの分野との連携・融合が欠かせないため、提案者自身と異なる知識・技能をもつ専門家と、提案内容について予め議論を深めてある内容であることを重視した。

4. 事前評価の選考の経緯

一応募課題につき領域アドバイザー11名が書類審査し、書類選考会議において面接選考の対象者を選考した。続いて、面接選考および総合選考により、採用候補課題を選定した。

選考	書類選考	面接選考	採択数
対象数	76件	26件	12件

備考:

1) 加えて、以下を今年度の事後評価対象とする。

・中村 壮伸 研究者

ライフイベントにより研究を一時中断したため。

5. 研究実施期間

2016年12月～2020年3月

6. 領域の活動状況

領域会議: 7回

研究総括の研究実施場所訪問: 17回

7. 事後評価の手続き

研究者の研究報告書を基に、評価会(研究報告会、領域会議等)での発表・質疑応答、領域アドバイザーの意見などを参考に、下記の流れで研究総括が評価を行った。

(事後評価の流れ)

2019年11月 評価会開催

2019年12月 研究総括による事後評価

2019年12月 被評価者への結果通知

8. 事後評価項目

(1) 研究課題等の研究目的の達成状況

(2) 研究実施体制及び研究費執行状況

(3) 研究成果の科学技術及び社会・経済への波及効果(今後の見込みを含む)

※該当する成果がある場合には「世界レベルの若手研究リーダーの輩出の観点から、本さきがけ研究が、研究者としての飛躍につながったか(今後の期待を含む)」を加味して評価を行った。

9. 評価結果

総論

マテリアルズインフォマティクスという言葉が一般にはあまり知られていない頃、手探りの状態で始まった本さきがけ領域であったが、その後のわずか一年でも、社会情勢や研究環境は急激に変化した。今年度研究期間終了となる2期生には、応募時点で機械学習の技術を使いこなしていた研究者もおり、またそうでない場合も1期生のサポートなどを受け、比較的スムーズに研究を開始することができた。領域内外で数々の共同研究が生まれたほか、外部向けの公開シンポジウムや、海外研究者との情報交換を図る国際シンポジウム、また様々な学会での招待講演、学会誌への解説記事執筆などを通じ、マテリアルズインフォマティクスのアウトリーチにも、2期生は大きく貢献した。

その中で溝口照康研究者は、界面構造解析へのベイズ最適化の応用に始まり、次々と新しいデータ科学的手法を導入して研究の幅を広げ、機械学習を用いた界面実験データ解析でも顕著な成果を挙げた。また積極的にプレスリリースを行い、多数の論文や解説記事を上梓し、66件もの招待講演・基調講演を行って、マテリアルズインフォマティクスの普及にも大きく貢献した。

1. 中村 壮伸 研究者 「トポロジカルデータ解析に基づくアモルファス構造の包括的記述と特徴抽出」

近年応用数学の分野で誕生したトポロジカルデータ解析(パーシステントホモロジー)と呼ばれるデータ解析

手法によって、シリカのようなランダムネットワーク構造と金属ガラスのようなランダム充填構造を統一的に取り扱えることを発見した。さらにアモルファス固体の降伏転移の記述にもパーシステントホモロジーが適用できることを見出した。これはパーシステントホモロジーが、単に構造の記述だけでなく、物性値とも深くかかわっていることを示す、重要な知見である。

本研究では、長い歴史のあるアモルファス構造の研究に、数学的観点から新しい切り口を見出しており、その手法は今後様々なアモルファス物性の記述に展開できる可能性がある点が評価できる。

2. 池野 豪一 研究者「機械学習と第一原理計算による新規スペクトル解析技術の確立」

相対論的多電子状態理論によるX線吸収スペクトルの高精度第一原理計算に必要な、電子間相互作用を表す二電子積分の高速計算アルゴリズムを開発し、期間中に約千件の遷移金属L2,3端X線吸収スペクトルを計算した。また理論計算と実測データの一致度を調べる目的で、圧縮センシングを利用したピーク分解法を開発した。当初計画では、計算されたデータを用いて、結晶構造から理論スペクトルを予測する機械学習モデルを構築するはずであったが、期間中に計算できたデータ量では不足していたため、学習データを米国のMaterials Projectが提供するO-K端X線吸収スペクトル15万件に切り替えて予測モデルの構築を試みた。その結果、多層ニューラルネットワークモデルを用いて、8割程度の高い確率でスペクトルの概形を、その原子配列から直接予測することが可能であることが示された。

本研究では、高速計算によって機械学習に必要なX線吸収スペクトルのデータベースを構築することが重要な目的の一つとなっており、そのための手法開発で成果があがった。また十分なデータ量のデータベース構築までにはいたらなかったのは残念であるが、米国で行われたように十分な計算資源と人的資源が投入できるならば、当初目的に沿った予測モデルができる可能性を示すことができた点は評価できる。今後の展開については、実験研究者からのニーズを見極めて検討すべきであろう。

3. 緒明 佑哉 研究者「はく離挙動を制御する指針の確立によるナノシート材料の機能設計」

電荷を帯びた宿主層と層間ゲストイオンで構成された層状遷移金属酸化物をはく離し、ナノシート材料を生成するための指針確立を目的とした研究である。当研究者は、「剛直な層状物質」としての層状無機化合物の層間にゲスト有機分子を導入することで「柔軟な層状物質」を得て、これを有機溶剤中へ分散させてナノシートを得る新しいはく離手法を開拓してきた。この手法は、相関ゲストと分散媒の組み合わせを変えることではく離挙動が変わることから、小規模実験データの機械学習とスパースモデリングにより予測モデルを構築し、効率的なナノシート生成を実現することに成功した。さらに同様の手法を用いて、リチウムイオン二次電池負極の有機電極活物質の探索を行い、これまでにない有望な候補物質を発見することに成功したほか、電極触媒の探索にも同手法を応用しつつある。

本研究で用いられたデータは、すべて自前の実験で得られたものであり、データ量は非常にわずかである。にもかかわらずそれらをうまく使い、さらに実験研究者の経験も加味することで、実際に役だつ結果を得ることができた。データ科学的手法により材料創製のプロセスを最適化するプロセスインフォマティクスとして、高く評価できる成果である。

4. 大塚 朋廣 研究者「機能性固体微細材料のマイクロレベル電子物性解析基盤技術の構築」

本研究者はこれまで半導体量子ドット等を用いて微細材料中の電子状態を局所的($\sim 10\text{nm}$)に、高精度($\sim 10\mu\text{eV}$)かつ低擾乱($\sim 100\text{aA}$)に測定できる独自のマイクロプローブを開発してきた。本研究では熱雑音を低減した低温下で高精度にかつデータ科学的処理に使える大量のデータを短時間で取得するために、低温冷凍機内で高周波を活用した高周波反射測定(\sim 数 100MHz)を行う測定系を構築、改良した。この測定結果とマイクロな理論、シミュレーションの比較から、量子ドットの局所電子・スピン状態を解明した。また高周波技術を活用したマイクロプローブの高速化によって、1秒間に ~ 1 億データもの大量のデータ取得が可能となったことから、ベイズ確率を活用した解析手法を開発し、さらにそれを測定と同時にリアルタイムで実行するため、FPGAによる信号解析回路を開発した。さらにマイクロプローブ試料の静電ポテンシャル最適化にも機械学習を適用した。

本研究における数々の実験手法、解析手法、FPGAの利用手法等の開発は、量子情報処理に向けた将来の大規模量子デバイスの実現に資する基盤技術として非常に重要で波及効果が大いと思われる、材料研究ではないにせよ、広い意味でのマテリアルズインフォマティクスの成果として高く評価される。今後、プレスリリースや解説記事等も活用し、成果の普及に努めてほしい。

5. 熊谷 悠 研究者「半導体材料開発のための計算データベース構築」

マテリアルズインフォマティクスによる材料開発に不可欠な、第一原理電子状態計算結果のデータベース作成を行った研究である。世界的に有名な同種のデータベース (Materials Project, Aflow, OQMD) は、取り扱う物性が主に結晶構造、全エネルギー、一般化勾配近似を用いた電子構造に限られており、高精度なバンドギャップ値や光吸収係数、点欠陥特性などの半導体物性に関する情報が欠落している。本研究ではデータベースを作成するための計算ワークフローを自動化するプログラムを開発し、バンドギャップを有する非金属材料を主な対象として、これらの物性値を含む信頼性の高いデータベースを作成した。

計算による材料物性データベースの構築は、Materials Genome Initiative により圧倒的に米国が先行しているが、本研究で作成されたデータベースは、データの質などいくつかの点でそれらを凌駕する貴重なものであり、半導体や誘電体の材料開発において強力な研究基盤となるポテンシャルを有している。本データベースは現在のところ未公開であるが、公開に向けた戦略を練り、研究者コミュニティに向けた発信を行うべきである。遠からず、このデータベースを用いた材料開発の成果が出ることを期待したい。

6. 坂上 貴洋 研究者「高分子物質のトポロジカル構造解析による新規物性の探索と設計」

濃厚高分子溶液の大規模シミュレーションで得られる数値データから、トポロジカルな拘束を表現する情報を抽出し、絡み合い状態の高分子系の構造的特徴を明らかにすることを目的とし、そのための新たな解析手法として、「動的絡み合い解析法」と「パーシステントホモロジー法」を構築、提案した。これらは、従来用いられてきた管模型とレプテーション模型では取り扱いの困難な環状鎖系の絡み合い構造を、鎖間の協同運動や新たに導入したトポロジカル貫通の概念によって解析する手法である。

本研究で提案された解析手法は、高分子物理の研究手法として興味深いもので、今後、高分子研究に新しい観点をもたらす成果として評価できる。一方で、現時点では構造の解析手法にとどまっており、今後、物性値との関係を明確にしたり、環状系と鎖状系の違いの理解に繋げたりといった方向での発展が望まれる。

7. 志賀 元紀 研究者「物質・材料の微細構造計測におけるインフォマティクス技術の開拓」

微細構造計測データの解析コストの大幅な削減を目指し、自動的かつ高精度に計測データを解析するための統計的機械学習法の開発を推進した。STEM-EELS のスペクトルイメージングを入力として、代表的な成分のスペクトル波形と成分の空間分布を自動的に同定する技術を開発したほか、同様の手法を利用して、TEM/SEM-EDS 等の計測時間を大幅に短縮できることも示した。領域内の共同研究を活発に進め、実験とシミュレーションによって得られたアモルファス構造から、中距離秩序を同定する手法や、混成されたスペクトルのピークを高速に解析する手法などを開発した。開発されたソフトウェアの多くはフリーウェアとして公開されている。

本研究者は情報科学の研究者であるが、材料研究者と積極的にコミュニケーションをとり、さまざまな共同研究を通じて多彩なデータ科学的手法を応用し、マテリアルズインフォマティクスの研究を先導したことが、高く評価できる。データ科学に馴染みの薄かった本領域の材料研究者にとっては良い相談相手となり、本研究領域全体への貢献が大きかったことも評価したい。

8. 辻 直人 研究者「強相関電子系に対する機械学習を用いた高精度量子多体計算の新たな数理アプローチの開拓」

強相関電子系は超伝導や磁性などで興味深い性質を示し、現代物性物理学の手法は研究テーマの一つとなっているが、そのシミュレーションに用いられる量子モンテカルロ法には、サンプリング数とともに数値誤差が増大する負符号問題と呼ばれる困難があることが知られている。この負符号問題の程度は、波動関数を展開する基底の選び方に依存することから、本研究では負符号を抑えて計算効率を最大化するための基底を推定するモデルを構築した。具体的には、推定モデルとして対応するフェルミハバードモデルと(ハードコア)ボースハバードモデルの自由エネルギー差をとった ΔF 指標と、フェルミハバードモデルとボースハバードモデルの間の Kullback-Leibler (KL) 情報量を相補的に用いることで、選んだ基底が最適なものであるかどうかを簡単に判定できることを示した。また機械学習の手法を用いて、量子モンテカルロ法の確率の重みを計算するコストを削減するアルゴリズムを開発した。これにより、強相関電子系の量子モンテカルロ法の計算コストを大幅に削減することに成功した。

本研究は、物性物理学の難問を解決するために不可欠な計算物理学的手法に、データ科学が貢献できることを示した一例となった。期間中は手法開発にとどまったが、引き続き開発した手法を応用し、物性物理学の課題解決につなげて重要な課題解決につなげてほしい。

9. 原 祐 研究者「円錐交差データベースに基づく蛍光分子自動設計法の開発」

有機 EL や蛍光プローブなどへの応用が期待される分子の蛍光に着目し、量子化学計算に基づいて分子設計を行う方法の開発を目指した研究である。そのためにまず、本研究者が拡張を進めてきた円錐交差探索手法を用い、芳香族化合物を中心に様々な有機分子の円錐交差構造を探索し、データベースを作成した。また SMILES 記法を用いた参照分子リスト生成手法と分子評価の方法を組み合わせ、蛍光分子を予測した。並行して、より高速で汎用性の高い円錐交差探索手法の開発や、GRRM を用いた反応経路探索結果を、SMILES 記法を用いて解析する手法を開発した。

本研究では、励起状態に着目した分子設計という難問に取り組み、そこに至る複数の課題に解決策を提示して、一定の道筋をつけることができたことを評価したい。期間中には、見出した分子を実際に作るまでできなかったのが残念であるが、今後実験家との共同研究や会社への委託により、ぜひ実証まで行ってほしい。

10. Jalem Randy 研究者「材料シミュレーションとインフォマティクスを用いたデータ駆動型リチウムイオン導電性セラミックスの探索」

第一原理に基づく材料シミュレーションと機械学習を組み合わせることによって、全固体電池で用いる固体電解質として優れた材料を系統的・効率的に探索することを目的とした研究である。結晶物性の一般的な記述子として、各原子を点とみなしたボロノイ分割から得られる多面体データ(頂点の数や辺の数、辺の長さや多面体の面積や体積など)が利用できることを示した。イオン伝導度の高い材料の探索に関しては、A、M、X、および Z をイオン置換させるようなタボライト型 $AMXO_4Z$ 化合物の組成空間をスクリーニングし、ベイズ最適化を用いることで効率化が図れることを示した。このほか全固体電池の正極—固体電解質界面である $LiCoO_2$ 正極— Li_3PS_4 固体電解質界面を、第一原理計算によって調べ、Li イオン移動の機構を明らかにしている。

多成分からなる固体電解質という複雑な実用材料の理論探索に様々な手法を用いて挑戦し、材料開発へのデータ科学的アプローチの良い実例となった。材料の設計指針までは得られていないものの、いくつかの材料を具体的に提案するところまでできたことは、大きな成果である。

11. 本郷 研太 研究者「ベイズ事後分布を探索重みに活用した物質構造探索の予測性向上」

ベイズ推定を用いて、膨大な化合物空間の中から所望の物性を持つ新規化合物(埋蔵分子)を効率的に探索することを目的とした研究である。この手法ではベイズの定理に基づき、尤度を、PubChem 構造データを用いた第一原理電子状態計算データから学習した物性予測の回帰モデル、化合物構造の事前分布を SMILES 形式と自然言語モデルの n-gram モデルに基づき構築し、それらの積をベイズ反転させることで、所望の物性を持つ化合物構造を見出す事後分布を構築する。具体的には、HOMO-LUMO ギャップと内部エネルギーを対象物性量として化合物探索に取り組み、データベースに収録されていない新規構造を持った化合物候補を見出した。また有機太陽電池のドナー材料探索に取り組み、実際の材料作成まで至っていないが、仮想分子ライブラリーを構築した。

中心的アイデアであるベイズ構造探索法は、吉田亮氏(統計数理研)らとの共同研究によるものであり、本研究はその具体化と実証を目指したと考えられる。そのために自然言語モデルの導入や多様なデータベースを利用した様々な機械学習をテストし、手法の実証につながる可能性のある探索結果が複数得られたことは、評価できる内容である。一方で、研究期間内に、実際に有用な化合物の合成までは到達できなかった。今後の展開に期待したい。

12. 溝口 照康 研究者「情報科学手法を利用した界面の構造機能相関の解明」

材料物性にとって重要な界面の構造解析を、機械学習をはじめとしたデータ科学的手法により加速することを目指した研究である。そのために、仮想スクリーニングやクリギング(ベイズ最適化)、転移学習といった機械学習法をシミュレーションに組み合わせることで界面シミュレーションを加速する手法、階層型クラスタリングおよびニューラルネットワークを利用することで界面計測データ解析を加速する手法を開発した。

界面構造は複雑で自由度が高く、実験的にも理論的にも取り扱いが困難である。その解析に、様々なデータ科学的手法を利用できることを示したことは、界面研究への大きな貢献であると同時に、マテリアルズインフォマティクスの有用性を端的に示す研究例としても、高く評価できる。また、数多くの著作や招待講演によって、マテリアルズインフォマティクスの普及やコミュニティの啓発にも大きく貢献した。

13. 森 寛敏 研究者「特定混合比で発現する特異物性を利用した新材料創成のための第一原理分子シミュレーションと機械学習の連携」

本研究では、分子の第一原理計算結果と実験データベースの機械学習を用いて、従来困難とされてきた混合冷媒（機能的混合溶液）の熱力学物性の迅速予測に挑戦した。冷媒の候補となり得る7万超の有機分子について第一原理計算を行い、分子の電子的特徴量をデータベース化した後、作成した特徴量と、アメリカ化学工学会データベースに登録されている、純物質の標準沸点（1166化合物）・標準融点（1190化合物）の実験データを正解情報とした、教師あり機械学習（ガウス過程回帰）を実施し、分子の単位表面電荷分布と、体積や表面積、Coulomb-Matrix など、種々の分子構造に由来する情報を特徴量とした場合、標準沸点・標準融点が高精度で予測できることを明らかにした。また同様に、二成分混合物に関して共沸の可能性を判定する識別器を作成したところ、9割弱の正解率で共沸可能性の判定が可能になった。

一見して機械学習からは遠い分野でテーマを絞って研究を進め、明確かつ有用な結論を導き、化学工学への突破口を示した、素晴らしい成果である。産業的なインパクトは非常に大きいと考えられる。

10. 評価者

研究総括 常行 真司 東京大学 大学院理学系研究科 教授

領域アドバイザー（五十音順。所属、役職は2020年3月末現在）

青柳 岳司 産業技術総合研究所 機能材料コンピューテーショナルデザイン研究センター
総括研究主幹

伊藤 聡 物質・材料研究機構 統合型材料開発・情報基盤部門 情報統合型物質・
材料研究拠点 拠点長／材料データプラットフォームセンター センター長

射場 英紀 トヨタ自動車(株) 先進材料技術部 電池材料技術・研究部
チーフプロフェッショナルエンジニア

小谷 元子 東北大学 材料科学高等研究所 主任研究者・教授

佐藤 寛子 情報・システム研究機構 准教授／チューリッヒ大学 研究員

田中 功 京都大学 大学院工学研究科 教授

知京 豊裕 物質・材料研究機構 統合型材料開発・情報基盤部門 情報統合型物質・
材料研究拠点 副拠点長

津田 宏治 東京大学大学院 新領域創成科学研究科 教授

寺倉 清之 産業技術総合研究所 機能材料コンピューテーショナルデザイン研究センター
エグゼクティブアドバイザー

中井 浩巳 早稲田大学 理工学術院 教授

樋口 知之 中央大学 理工学部 教授

（参考）

件数はいずれも、2020年3月末現在。

(1) 外部発表件数

	国内	国際	計
論文	11	138	149
口頭	275	137	412
その他	10	7	17
合計	296	282	578

(2) 特許出願件数

国内	国際	計
1	1	2

(3) 受賞等

・緒明 佑哉

1. 文部科学大臣表彰若手科学者賞、「結晶からつくる有機および無機高分子材料に関する研究」、2018年4月
2. Polymer Journal 論文賞-日本ゼオン賞、「Emergence of temperature-dependent and reversible color-changing properties by the stabilization of layered polydiacetylene through intercalation」、2019年5月.

・大塚 朋廣

1. 日本物理学会 若手奨励賞、「局所的な電子状態の高速な制御、検出手法の研究」、2017年3月
2. 理化学研究所 研究奨励賞、「固体ナノ構造中局所電子状態の精密高速観測と制御」、2017年3月
3. 理化学研究所 CEMS Award、「固体ナノ構造中局所電子状態の精密高速観測と制御」、2017年4月
4. 新世代研究所 ATI 研究奨励賞、「超高速量子プローブによるナノデバイス中の電子輸送ゆらぎの研究」、2017年5月
5. 矢崎科学技術振興記念財団学術奨励賞、「超高速量子ドットプローブを用いた固体微細デバイス中の局所電子状態のダイナミック計測」、2018年3月
6. 文部科学大臣表彰若手科学者賞、「固体微細構造中局所電子状態の精密高速観測と制御の研究」、2018年4月
7. 東北大学通研研究交流会 優秀発表賞、「将来の情報処理・通信に向けた固体ナノ構造を用いた量子デバイス」、2018年9月

・熊谷 悠

1. 第41回応用物理学会解説論文賞、「Design and exploration of semiconductors from first principles: A review of recent advances」、2019年9月
2. 第39回本多記念研究奨励賞、「第一原理計算に基づいた半導体物性理論の構築」、2018年5月.
3. 第15回村上奨励賞、「半導体物性に関する理論研究」、2018年5月

・溝口 照康

1. 風戸研究奨励会・風戸賞、「ELNES 理論計算に関する基礎的研究と物質研究への応用」、2017年3月
2. 日本金属学会・功績賞、「原子レベル計測、第一原理計算および情報科学手法を用いた物質のナノ構造解析」、2017年3月.

・森 寛敏

1. 新化学技術推進協会(JACI) 第6回化学技術研究奨励賞、「水素吸蔵性合金ナノ粒子内の有効空隙最大化を目指した第一原理水素吸蔵空間可視化法の開発とその応用」、2017年5月
2. MTMS'18 Best Poster Prize、「Designing ionic liquids for efficient CO2 capture: A materials informatics study」、2018年9月.
3. 化学工学会秋季年会 最優秀講演賞、「CO2 吸収特性に優れたイオン液体探索のための電子状態インフォマティクス」、2018年9月
4. 伊藤科学振興会 研究奨励賞、「浸透圧調整分子の溶液化学: 第一原理有効フラグメントポテンシャル-分子動力学計算によるアプローチ」、2018年10月
5. GREEN Chemistry Prize、「Exploring ionic liquids for efficient CO2 capture and storage by electronic structure calculation and machine learning」、2018年11月
6. 化学工学会(東京大会) 優秀賞、「電子状態インフォマティクスの手法による沸点・融点の推算」、2019年3月

(4)招待講演

国際 132件
国内 51件

別紙

「理論・実験・計算科学とデータ科学が連携・融合した先進的マテリアルズインフォマティクスのための基盤技術の構築」領域 事後評価実施 研究課題名および研究者氏名

研究者氏名 (参加形態)	研究課題名 (研究実施場所)	現職(2020年3月末現在) (応募時所属)	研究費 (百万円)
中村 壮伸 (兼任)	トポロジカルデータ解析に基づくアモルファス構造の包括的記述と特徴抽出 (産業技術総合研究所 機能材料コンピュータシヨナルデザイン研究センター)	産業技術総合研究所 機能材料コンピュータシヨナルデザイン研究センター 研究員 (東北大学原子分子材料科学高等研究機構)	34
池野 豪一 (兼任)	機械学習と第一原理計算による新規スペクトル解析技術の確立 (大阪府立大学 大学院工学研究科)	大阪府立大学 大学院工学研究科 准教授 (大阪府立大学 21世紀科学研究機構)	32
緒明 佑哉 (兼任)	はく離挙動を制御する指針の確立によるナノシート材料の機能設計 (慶應義塾大学 理工学部)	慶應義塾大学 理工学部 准教授 (同上)	37
大塚 朋廣 (兼任)	機能性固体微細材料のマイクロレベル電子物性解析基盤技術の構築 (東北大学 電気通信研究所)	東北大学 電気通信研究所 准教授 (理化学研究所 創発物性科学研究センター)	37
熊谷 悠 (兼任)	半導体材料開発のための計算データベース構築 (東京工業大学 科学技術創成研究院)	東京工業大学 科学技術創成研究院 准教授 (東京工業大学 元素戦略研究センター)	39
坂上 貴洋 (兼任)	高分子物質のトポロジカル構造解析による新規物性の探索と設計 (青山学院大学 理工学部)	青山学院大学 理工学部 准教授 (九州大学大学院 理学研究院)	11
志賀 元紀 (兼任)	物質・材料の微細構造計測におけるインフォマティクス技術の開拓 (岐阜大学 工学部)	岐阜大学 工学部 准教授 (同上)	32
辻 直人 (兼任)	強相関電子系に対する機械学習を用いた高精度量子多体計算の新たな数理アプローチの開拓 (理化学研究所 創発物性科学研究センター)	理化学研究所 創発物性科学研究センター 研究員 (同上)	27
原淵 祐 (兼任)	円錐公差データベースに基づく蛍光分子自動設計法の開発 (北海道大学大学院 理学研究院)	北海道大学大学院 理学研究院 助教 (同上)	36
Jalem Randy (兼任)	材料シミュレーションとインフォマティクスを用いたデータ駆動型リチウムイオン導電性セラミックスの探索 (物質・材料研究機構 エネルギー・環境材料研究拠点)	物質・材料研究機構 エネルギー・環境材料研究拠点 主任研究員 (物質・材料研究機構 ナノ材料科学環境拠点)	32
本郷 研太 (兼任)	ベイズ事後分布を探索重みに活用した物質構造探索の予測性向上 (北陸先端科学技術大学院大学 情報社会基盤研究センター)	北陸先端科学技術大学院大学 情報社会基盤研究センター 准教授 (北陸先端科学技術大学院大学 先端科学技術研究所)	36

溝口 照康 (兼任)	情報科学手法を利用した界面の構造 機能相関の解明 (東京大学 生産技術研究所)	東京大学 生産技術研究所 教授 (同上)	36
森 寛敏 (兼任)	特定混合比で発現する特異物性を利用した新材料創成のための第一原理 分子シミュレーションと機械学習の連 携 (中央大学 理工学部)	中央大学 理工学部 教授 (お茶の水女子大学 基幹研究院)	38

研究報告書

「トポロジカルデータ解析に基づくアモルファス構造の包括的記述と特徴抽出」

研究タイプ: 通常型

研究期間: 2015年12月～2019年5月

研究者: 中村壮伸

1. 研究のねらい

アモルファス構造においては、結晶や液体などのような解析的な表現に適した適切な参照系が存在しない。それゆえシミュレーションによる計算結果を理論的に説明・統合する方法論がない。データ科学が質的に新しいものの見方を提供するとすれば、アモルファス構造のような理論が不得手とする対象においてこそ期待できるものであり、本研究はそのための基盤となる新技術創出に資するものである。

アモルファス構造の汎用的な記述子としては動径分布関数などがすでに存在する。この量は理論・実験研究との親和性が高く構造記述の中心的物理量であった。しかしながら、多くのアモルファス材料では中距離秩序構造と呼称される多体構造が構造の特徴づけに中心的役割を果たすことが示唆され、動径分布関数では不十分であるとされている。近年発明されたパーシステントホモロジーという計算幾何学的手法は点データに対して数学的に定義された「穴」を検出し特徴付ける理論である。そこで、本研究において多体構造を記述しうるパーシステントホモロジーがアモルファス構造の汎用的な記述子となりうるかどうかを問うため、典型的な系に対して適用研究を行う。

2. 研究成果

(1) 概要

研究テーマ A 「パーシステントホモロジーによる多様なガラス構造の統一的記述」

代表的なアモルファス構造としてシリカのようなランダムネットワーク構造と金属ガラスのようなランダム充填構造がある。従来はこれらの構造は別々の特徴づけがなされ統一的な取り扱いがなされていなかったためデータ科学の適用が困難であった。我々はパーシステントホモロジーを用いることでこれらの構造を統一的に取り扱うことが可能であることを発見した。

研究テーマ B 「パーシステントホモロジーによるアモルファス構造の降伏転移の記述」

アモルファス固体に対する降伏転移を支配する新たな秩序変数の提案を行った。アモルファス固体の降伏転移は近年レプリカ解析を用いて定義される秩序変数で転移を議論されてきたが、我々はこの転移を配置データに対してパーシステントホモロジーを適用し、そこで得られるノイズに対する頑健な構造の頻度の変化として捉えることが可能であることを見出した。

研究テーマ C 「データ駆動で得られた変数に対する自由エネルギー地形の新しい定義」

パーシステンスダイアグラムに代表されるデータ駆動で得られる変数と物性を繋ぐための枠組みを統計力学と微分幾何学に基づいて構築した。構造物性相関を議論する際に物性値が記述子の選択の任意性に依存しない条件を微分同相変換に対する不変性として定式化できることを発見し、そのような記述子を反応座標とした自由エネルギー地形の新しい定義を与えた。

(2) 詳細

研究テーマ A「パーシステントホモロジーによる多様なガラス構造の統一的記述」

ランダムな構造を持つ固体であるガラスの代表的な構造は2種類ある。一つは金属ガラスや分子性液体からなるガラスである。これらの物質は方向性のない相互作用を持ち、また最近接原子の配位数が可変であるという特徴を持ち、この性質からランダム充填構造と呼ばれる構造を取る。もう一つはシリカガラスなどのような方向性の強い共有結合で相互作用をする物質からなるガラスで、ランダムネットワーク構造と呼ばれる構造を取る。これらの材料はともにガラス転移を経て運動が凍結するという特徴をもつものの、構造としては全く異なるため異なる特徴づけが伝統的になされてきた。具体的にはランダム充填構造に対しては動径分布関数を考え、対応する液体構造における第二ピークの分裂を特徴づけに使い、ランダムネットワーク構造に対しては構造因子の最小の波数のピークの発達を特徴づけに使っていた。我々は今世紀に発明されたパーシステントホモロジーという数学的手法を用いることによりこれらを統一的に扱うことが可能であることを見出した。パーシステントホモロジーを用いることにより、ランダム充填構造の特徴は空隙として、ランダムネットワーク構造の特徴はリングとして、それぞれ2次元、1次元の穴として統一的に扱うことができることがわかった。また、それらの特徴として現れる穴はパーシステンスダイアグラムの対角線から離れた領域に現れる、ノイズに対して頑健な穴であることも明らかになった。これにより物理的な固さに現れる安定性を、パーシステンスダイアグラムのもつ擾乱に対する安定な穴の性質として解釈することが示唆された。

研究テーマ B「パーシステントホモロジーによるアモルファス構造の降伏転移の記述」

研究テーマ A で示唆された材料構造の安定性とパーシステンスダイアグラムが表現する擾乱に対する頑健性の関係に関する仮説の検証を推し進めた。パーシステントホモロジーでは擾乱に対する穴の頑健性はパーシステンスダイアグラムに現れる対応する点の対角線からの距離のよって評価することができることが知られている。我々はガラス転移の研究で標準的に用いられているモデルである Kob-Andersen-Lennard-Jones binary-mixture モデルを用いて剪断変形に対する降伏現象の冷却速度依存性を調査した。この系の降伏現象に対して、スピングラスなどで用いられるオーバーラップと呼ばれる量を用いてある種の相転移として扱う方法が近年提案されている。我々はオーバーラップを用いず配置データだけから相転移を検出する方法が好ましいと考え、降伏前後における構造の変化を頑健な穴の変化として定量化し、両者の関連性を問うた。その結果、パーシステンスダイアグラムに現れるロバストな穴の数が降伏点を境に急激に減少していることを見出した。したがって、パーシステンスダイアグラムに現れる対角線から距離の離れた点の数が降伏現象を相転移として記述する際のある種の秩序変数とみなせる可能性が示唆された。この結果は研究テーマ A と合わせてパーシステンスダイアグラムで抽出される構造やその頻度が熱統計力学的に意味のある量になっている例であると考えられる。

研究テーマ C 「データ駆動で得られた変数に対する自由エネルギー地形の新しい定義」

パーシステンスダイアグラムや主成分分析で得られた変数のようにデータ駆動で抽出され

る記述子はデータ科学による構造物性相関を議論する際に確かに良い記述を与えうる変数であることが多い。しかしながら変数の出自が物理的に担保されているわけではなく手続きとして定義が与えられているだけであるため熱統計力学的な現象論的な理解とは相性が悪い。実際、それらの変数を用いる場合には形式的定義のみから自由エネルギー地形を導入する必要がある。先行研究によれば、自由エネルギー地形には変数の計量の選択に依存した任意性があることが指摘されていた。パーシステンスダイアグラムや主成分分析のような物理的な出自が明らかでない勝手に選んだ変数の場合には変数の計量が物理的に妥当であるかどうかを検証する必要がある。従来の定式化ではそのような検証は理論の中に含まれておらず、そのため任意性の問題については未解決のままであった。我々は、自由エネルギー地形が軽量の選択に依存しないように定義しなければならないこと物理的要請として認めること(一般共変性原理)によって、一意な定義を与えることに成功した。この定義によれば自由エネルギー地形は従来用いられていた頻度分布の計算だけではなく、拡散行列の計算も必要であることを示した。これにより、データ駆動で得られた変数から物理的な要請をみたとすう自由エネルギー地形を計算することが可能になり、理論的な保証に基づいて構造物性相関を問うことが可能になった。特にパーシステンスダイアグラムから操作的に自由エネルギー地形を計算する手続きも明らかになった。

3. 今後の展開

パーシステントホモロジーを用いることにより多様なアモルファス構造の特徴を統一的に扱うことが可能になり、データ科学との融合が可能になった。実際、我々の研究以降にパーシステントホモロジーと機械学習と組み合わせることによりガラス転移点の推定に関する研究などが行われている。

アモルファス構造のレオロジーにおいて示されたノイズに対する頑健な穴が物性を記述する構造の担い手になるというシナリオは他の物理系においても成り立つことが予想される。今後多様な例題で同様のシナリオの成立を確認する必要があり、おそらく肯定的な結果が得られると考えられる。

パーシステンスダイアグラムを反応座標とする自由エネルギー計算の手法開発が可能になる。これは本研究計画の目的であるトポロジカルデータ解析による構造物性相関の汎用的手法の提案のゴールであるため是非とも完成させたい。

4. 自己評価

1. 研究目的の達成状況

汎用性の高い構造物性相関の評価枠組みを作るという目的に関しては自由エネルギー地形の理論を作ることによって達成できたが、トポロジカルデータ解析に限らない用途の手法開発を含んだためやや当初の目的からそれてしまった。また典型例の網羅的な解析を行う計画に関しては先行研究の情報収集に思ったよりも時間がかかってしまい十分に達成ができなかった。

2. 研究の進め方

研究体制に関しては異動に伴い研究補助者の雇用が途切れてしまったことなどで不安定な体制で行わざるを得なかった。そのため補助者への給与を目的とした研究費執行に支障

をきたしかなりの残額が生じてしまった。研究開始時に代替手段も考慮に入れた計画を寝るべきであった。

3.研究成果の波及効果

トポロジカルデータ解析の研究者人口はこの数年で圧倒的に増え、我々の研究をきっかけに材料科学への適用を考えている企業などの研究の相談を受けたことも多数あった。このようにニーズがあることを認識し、さらに発展させ社会に還元させる必要があることがわかった。

4.研究者自身の評価

データ科学の中でもトポロジカルデータ解析は予測のツールではなく記述のツールであったため、定量的な予言に関わる部分は理論物理をベースに行った。したがって、機械学習などのような予測のツールとしてのデータ科学との連携は不十分であった。これは研究計画には含まれていなかったが領域会議等を通じて、やるべき問題と認識しつつも期間内に行うことができなかった問題である。今後、本さきがけで知り得た技術や研究者と連携をとり、予測ツールとの連動も行っていきたい。

5. 主な研究成果リスト

(1)論文(原著論文)発表

- | |
|---|
| 1. Hiraoka, Y., Nakamura, T., Hirata, A., Escolar, E. G., Matsue, K., & Nishiura, Y. (2016). Hierarchical structures of amorphous solids characterized by persistent homology. <i>Proceedings of the National Academy of Sciences</i> , 113(26), 7035–7040. |
| 2. Shirai, T., & Nakamura, T. (2019). Microscopic Description of Yielding in Glass Based on Persistent Homology. <i>Journal of the Physical Society of Japan</i> , 88(7), 074801. |
| 3. Nakamura T. (2019) The Unique and Physical Definition of the Free-Energy Landscape
arXiv:1803.09034 |

(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)

ガラスの「形」を数学的に解明

https://www.wpi-aimr.tohoku.ac.jp/jp/news/press/2016/20160614_000643.html

研究報告書

「機械学習と第一原理計算による新規スペクトル解析技術の確立」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 池野 豪一

1. 研究のねらい

現代の材料開発においては、物質の特性や化学反応を支配する特徴的な原子構造(格子欠陥・界面・表面など)を解明し、制御することが求められる。内殻電子分光法は物質中の電子が光子・電子と相互作用することにより、量子的なエネルギー準位間を遷移する過程を観測する手法であり、物質中の特定元素周辺の局所的な原子構造・電子状態を反映したスペクトルが得られる。最近では測定装置や技術の発達により、透過型電子顕微鏡を用いた1原子カラム毎のスペクトル測定や、化学反応の追跡するための時間分解分光測定の発展など、分光測定から得られる情報は飛躍的に増加している。

一方で、「得られたスペクトルを如何に解析するか」については現在の大きな課題となっている。最も簡便な方法は参照物質の実験スペクトルと比較する指紋照合法であるが、その比較方法は人間の「目視」による極めて原始的かつ定性的なものである。より信頼性の高い解析手法としては、第一原理計算を用いた理論指紋照合法がある。この方法では、候補となる原子配列に対して電子状態計算から理論スペクトルを求め、実験結果と照合することで局所原子配列・電子状態を明らかにする。高精度な解析が可能となる反面、個々のスペクトルの理論計算に必要なコストは高く、最新の分光実験で得られる大量のスペクトルデータ全てに対して第一原理計算を実施するには膨大な時間を要することになる。これらの分光実験で測定されるスペクトルデータが持つ情報を最大限活用し、材料開発に応用するためには、スペクトル形状と原子構造の関連を定量的かつ高速に予測することができる新しい手法の開発が求められる。

本研究のねらいは、第一原理計算と情報科学的学習を組み合わせ、大量のスペクトルデータを効率的かつ定量的に解析し、局所原子構造・電子状態に関する情報を抽出することができる新しい解析技術を確立することにある。電子状態計算を経由せずに、原子構造からスペクトル形状を直接予測できる回帰モデルを作成することにより数百～数千の構造に対するスペクトルを短時間で計算し、候補構造を絞り込むことができる。これにより、測定される大量のスペクトルデータが持つ情報を最大限活用することが可能となり、材料開発・設計につながる新しい知見が得られると期待される。

2. 研究成果

(1) 概要

本研究目標を達成するにあたり、以下の3つの研究テーマを設定した。

- (1) 遷移金属 $L_{2,3}$ 端 X線吸収スペクトルデータベースの作成
- (2) 圧縮センシングによるピーク分離法および実験・理論スペクトルの定量照合法の確立
- (3) 機械学習によるスペクトル予測モデルの構築

研究テーマ(1)について、遷移金属 $L_{2,3}$ 端は、空間的に局在化した 2p,3d 軌道間の電子遷移を観測するものであり、応用上重要な遷移金属化合物の物性に関与する電子準位の情報が得られる。しかし、その計算にはこれらの軌道を占有する電子間に働く強い相互作用をあらわに取り扱う必要がある。計算には、研究者らが開発を進めてきた相対論多電子計算を用いる。高精度なスペクトルの予測モデルを作成するには、数万を越えるデータが必要であるため、計算の高速化を実現するため、幾つかのアルゴリズムの開発と実装を行った。これら新規に開発したプログラムを利用した自動計算システムを構築し、価数・配位子の種類・配位数・対称性が異なる遷移金属化合物のスペクトルデータの収集を進めている。研究期間内では、機械学習に用いるに十分なデータ数を収集することができなかつたため、今後も継続してデータベース作成を進める。

研究テーマ(2)に関しては、圧縮センシングアルゴリズムの一つである ESPRIT 法を利用したピーク分解法を開発した。この方法では、スペクトルをローレンツ関数の和として分解し、スペクトルをピーク数(J)、ピーク位置(s_j)、ピーク幅(γ_j)で表現し直すことにより、その「一致度」数値的に評価することを試みた。

研究テーマ(3)について、スペクトルの機械学習法の検討を行った。ここで機械学習の目的変数はスペクトルそのものであり、これは入射光のエネルギーに対する非負値の連続関数という特徴を持つ。この様な目的変数に対する回帰モデル作成は殆ど行われておらず、回帰手法そのものの確立が必要となる。本研究では、既存の電子状態計算データベース(Materials Project)で提供されている X 線吸収スペクトルデータベースを利用して、様々な回帰手法の検討を行った。その結果、多層ニューラルネットワークモデルを用いて、8 割程度の高い確立でスペクトルの概形を、その原子配列から直接予測することが可能であることを示した。

(2) 詳細

研究テーマ(1) 遷移金属 $L_{2,3}$ 端 X 線吸収スペクトルデータベースの作成

相対論多電子計算法の速度向上を目指して電子間相互作用を表す二電子積分の高速計算アルゴリズムの開発を行った。新たに開発した方法では、電子密度の多重極展開を行い、二電子積分をフーリエ空間で実行することにより、一次元振動の和として求める。また、動径関数をテーマ(2)で開発した ESPRIT 法を用いて指数関数の和としてフィッティングすることで、振動積分を解析的に、かつ高精度に計算できる手法を開発した。

また、異常高原子価イオンを含む化合物に対しては内殻電子の電子励起と同時に配位子から励起中心の遷移金属イオンへ電荷移動がスペクトル形状に大きく影響を与える。この配位子からの電荷移動は、配置間相互作用計算において新たな電子配置を付け加えることで計算が可能となるが、大規模なハミルトニアン行列の固有値問題を解く必要がある。本研究では、行列固有値の反復解法を用いて遷移の始状態の波動関数のみを求め、シフト線形方程式に基づいて系の Green 関数を高速に計算できるアルゴリズム(修正 RSCG 法)を開発した。これにより、多様な系に対して効率的かつ高精度なスペクトルを求めることが可能となった。

これらの手法を実装した上で、クラスター計算機上に自動計算システムを構築し、スペクトルデータベースの作成を試みた。しかし、研究期間内では、各遷移金属元素に対して得られ

ているスペクトルは千程度と、機械学習に用いるに十分なデータ数には至っていない。研究期間後もスペクトルデータの収集を継続的に実施し、機械学習モデルの作成を行う予定である。

研究テーマ(2) 圧縮センシングによるピーク分離法および実験・理論スペクトルの定量照合法の確立

圧縮センシングアルゴリズムの一つである ESPRIT 法を実装した。本アルゴリズムは、フィッティングパラメータの数、初期値に関する事前情報を必要とせず、等間隔メッシュ上で観測されるデータのみから指数関数の和として関数を復元できるという特徴がある。スペクトルデータをフーリエ変換し、指数関数の和でフィッティングした後、逆フーリエ変換を施すことでローレンツ関数の和としてピーク分離ができることを示した。当初の計画では、ピーク分離結果を元に、そこからの実験・理論スペクトルの変換関数を作成する計画であった。しかし、検討を進めた結果、分離後のピーク情報からスペクトルの一致度を判定することが困難であり、適切な変換関数の作成には至っていない。ただし、本手法はノイズのあるデータに対しても使用できること、ピーク位置に関する事前知識も不要という特徴から、実験スペクトルのピーク分離等、別の用途に活用できると考えられる。

研究テーマ(3) 機械学習によるスペクトル予測モデルの構築

Materials Project が提供する多重散乱理論(FEFCO コード)に基づいて得られた O-K 端 X 線吸収スペクトルを収集し、結晶構造から理論スペクトルを直接予測する機械学習モデルの構築を試みた。まず、カーネル Ridge 回帰を使ってスペクトルの形状の直接予測を試みた。入力には物質の化学組成と、動径分布関数やクーロン行列といった原子構造から求められる記述子を用いた。回帰を行う前に、各エネルギーで離散化したスペクトルデータに対して非負値分解を行い、スペクトルを少数の独立成分の線形和として表現し、その重み係数を回帰の対象とすることで、計算量の削減を図った。約 2500 のデータに対してこの手法を適用したところ、テストデータに対する予測精度を表す決定係数(R^2)の平均は 0.43 となり、十分な汎化性能のあるモデルの構築はできなかった。

そこで、学習に用いるデータ数を増やし、約 15 万のスペクトルデータに対する機械学習を試みた。データ数が 5 万を越えたあたりから非負値分解の計算コストが激増し、先の手法を適用することは困難となった。そこで、多層ニューラルネットワークを用いた予測モデルの作成を行った(図 1 参照)。この方法では、データベースからランダムに少数のスペクトルを選択し、ニューラルネットワークの重みを学習させる過程を繰り返すことで、予測精度を逐次更新することができる。ネットワークの層の数やバッチサイズ、エポック数などのハイパーパラメータについては、ベイズ最適化を用いて決定した。その結果、約 4 分の 3 のスペクトルに対して主なピークの位置や強度など、スペクトルの主な特徴を捉えた予測結果を得ることができた。一方、残りの約 4 分の 1 については、予測結果とオリジナルの理論スペクトルとの間に大きな誤差が見られた。データを検証したところ、実験的には不安定な仮想物質に対する予測精度が低いという傾向が見られたが、学習データに類似の構造からのスペクトルを含めることで、精度の改善が期待できる。

更に高精度なスペクトル予測モデルの構築を目指して、crystal graph convolution neural network (CGCNN) を用いた予測モデルの作成を行った。この手法では、結晶構造をグラフ構造として取り扱う。頂点に原子の特徴量、辺に原子間の結合を表す特徴量を保持している。さらに畳み込みニューラルネットワークを用いることで、目的変数に応じた特徴量ベクトルの最適化を行うことができる。前述のニューラルネットワークと同様のデータセットを用いて学習を行ったところ、同程度の精度のスペクトル予測モデルが得られた。

以上の結果から第一原理計算なしに、原子構造から直接スペクトル形状を捉える予測モデルを作成することが可能であるという確信を得た。

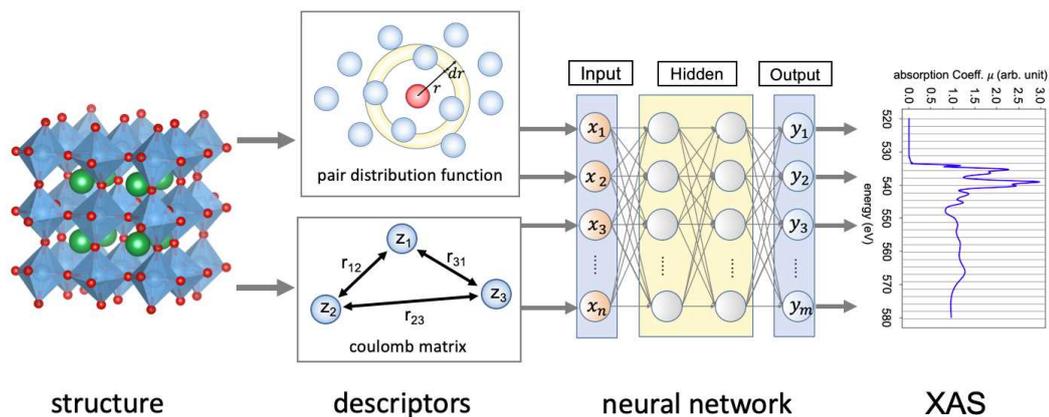


図 1 ニューラルネットワークを用いた XAS 予測モデルの概念図。電子状態計算なしに、原子構造から直接スペクトル形状を予測する。

3. 今後の展開

本研究を通して、原子配列から電子状態計算なしでスペクトル形状を予測することが可能であるという知見を得た。今回対象としたのでは O-K 端のみであるが、他の元素の任意の吸収端においても、十分な学習データを集めることができれば同様の手法ができると考えられる。したがって、今後の研究においては如何にして効率よくデータを収集するかが課題となる。本研究で実施していた遷移金属 $L_{2,3}$ 端 X 線吸収スペクトルは複雑な物理現象を扱う必要がある。継続的な研究テーマとしては、このデータベース作成を継続し、機械学習モデルの作成を行うことが挙げられる。また、遷移金属 $L_{2,3}$ 端の理論・実験スペクトルデータベースは存在しないことから、将来的にデータベースの公開も視野に入れる。

また、放射光分野において、実験スペクトルのデータベース作成に向けて議論が進められている。実験スペクトルの数は計算で得られるものよりもデータの数やバリエーションは下がるものと考えられる。これらのデータから如何に効率良く機械学習モデルを作成するかは、多岐に渡る科学分野において内殻電子分光法を活用する上で大事な研究テーマである。

4. 自己評価

本研究の鍵となるのは、電子状態計算なしに、原子配列から直接 X 線吸収スペクトルを予測する機械学習モデルの構築である。本研究では、このようなモデルの作成が可能であることを実証できたことから、研究目的の根幹をなす部分については達成できたものとする。

研究成果の欄に挙げた 3 つの研究テーマそれぞれに関しての達成度合いを見ると、まずテーマ(1)のデータベース作成については時間的要因により達成には更なる時間を要する。テーマ(2) に関しては、圧縮センシングを用いた事前知識を必要としないピークの分離方法については開発できた。しかし、この結果を元にスペクトルの照合を行うという目的に対しては、未達成であり、今後の研究課題とする。テーマ(3) については、前述の通りおおむね達成できたものとする。

研究期間全体に渡り、学生または研究補助者 1 名と協力する体制で進めてきた。研究費は前半(1、2 年目)にはデータベース作成のための計算機クラスターの導入、後半(3、4 年目)には機械学習に利用する GPU 搭載の計算機の導入に利用された。研究の進捗に応じて適切な、経費の執行を行ったものと判断する。

現代の自然科学研究において、X 線吸収分光法は物質の微視的状态を知るための重要な実験手法として認識され、多岐に渡る分野で活用されている。本研究で得た機械学習モデル作成に関する知見を用いて様々な元素・吸収端に対して信頼できる予測モデルを作成することができれば、スペクトルから局所原子配列に関する情報を迅速に抽出することが可能となり、物質評価のプロセスを効率化することができる。近年では、学術用途のみならず、二次電池や触媒などの、民間企業で開発されている先端材料の物質評価に X 線吸収分光法が利用されるようになってきている。本研究の成果に基づいて、これらのデータ解析が機械的に実行できるようになれば、新たな材料開発指針を得ることができ、開発サイクルの改善につながるものと期待される。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Hidekazu Ikeno and Teruyasu Mizoguchi, "Basics and applications of ELNES calculations", <i>Microscopy</i> 66 (2017) 305-327.
2. Masato Urasaki, Hidekazu Ikeno, " <i>Ab-Initio</i> Multiplet Calculations Using Iterative Algorithms for X-ray Absorption Spectra at Transition Metal $L_{2,3}$ -edges", <i>Materials Transactions</i> 59 (2018) 1057-1061.
3. Hidekazu Ikeno, "Spherical Bessel transform via exponential sum approximation of spherical Bessel function" <i>Journal of Computational Physics</i> 355 (2018) 426-435.
4. Hidekazu Ikeno, "mxpfit: A library for finding optimal multi-exponential approximations", <i>Computer Physics Communications</i> 230 (2018) 135-144.
5. Ikuya Yamada, Akihiko Takamatsu, Kaisei Asai, Takuto Shirakawa, Hideo Ohzuku, Akihiro Seno, Tasuku Uchimura, Hiroshi Fujii, Shogo Kawaguchi, Kouhei Wada, Hidekazu Ikeno and Shunsuke Yagi, "Systematic Study of Descriptors for Oxygen Evolution Reaction Catalysis in Perovskite Oxides" <i>The Journal of Physical Chemistry C</i> 122 (2018) 27885-27892.

(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)

1. 池野豪一、「遷移金属 $L_{2,3}$ 端 X 吸収分光と多電子系第一原理計算による正極材料の化学状態解析」、立命館 SR 公開シンポジウム「軟 X 線分光を用いた二次電池研究の最前線」、2016 年 11 月 11 日、(立命館大学)。[招待講演]
2. Hidekazu Ikeno, “Ab-initio Multiplet Calculations for X-ray Absorption Spectroscopy: Application to Cathode Materials”, Frontiers in Materials Science (FMS2017), Sep. 4-6, 2017 (Greifswald, Germany). [招待講演]
3. H. Ikeno and N. Miyamoto, “Machine Learning Approach for the Analysis of X-ray Absorption Spectra”, PRESTO International Symposium on Materials Informatics, Feb. 9-11, 2019 (University of Tokyo, Japan).
4. 池野豪一、宮本直幸、「X 線吸収スペクトルの機械学習」、第 22 回 XAFS 討論会、2019 年 9 月 2-4 日、(京都大学。[招待講演]
5. Hidekazu Ikeno, DFT-CI calculations for RIXS and RIXS-MCD”, Frontier of theoretical approach in x-ray spectroscopy (PF 研究会「X 線分光理論の新展開: 構造・電子状態解析から磁性研究まで」), Oct. 3-4, (Photon Factory, Tsukuba, Japan). [招待講演]

研究報告書

「はく離挙動を制御する指針の確立によるナノシート材料の機能設計」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 緒明 佑哉

1. 研究のねらい

本研究は、設計された構造と機能を有するナノシート材料を得るために、層状化合物の構成単位である層を化学的にはがす「はく離」の挙動を制御するための包括的な指針を、実験・計算・データの連携によって見出すことを目指した。ナノシート材料は、高比表面積、2次元異方的な形態、柔軟性、量子サイズ効果など、特異な構造・形状に由来した様々な特性から、近年、「2次元材料」として期待されている。しかし、一般に、層状物質を前駆体として、層をはがす(＝はく離)プロセスによってナノシート材料の合成が行われているものの、このような「壊す」プロセスの制御性は高くない。すなわち、ナノシート材料の問題点として、収率が低いこと、サイズ制御が容易ではないこと、分散性・表面修飾の制御が容易ではないことなどが課題として挙げられる。特に、遷移金属酸化物などの層状物質では、はく離手法に限られており、特異な機能や様々な応用が期待されているものの、構造制御されたナノシートを効率的に得ることが容易ではなかった。我々の研究グループでは、「剛直な層状物質」としての層状無機化合物の層間にゲスト有機分子を導入することで「柔軟な層状物質」を得て、これを有機溶剤中へ分散させてナノシートを得る新しいはく離手法を開拓してきた。本手法では、層間ゲストと分散媒の組み合わせを変えることで、収率やサイズを含めたはく離挙動の制御を行える可能性がある。そこで、本研究提案では、マテリアルズインフォマティクス(MI)を活用し、ナノシート材料の予測に基づいた収率の向上やサイズの制御を第1のねらいとした。

本はく離手法は我々の研究グループ独自の系であり、MIを適用するための十分な量のデータを文献値などから取得することは難しい。すなわち、自前で取得した小規模な実験データにMIを適用する必要がある。一般に、化学や素材に関連した研究開発では、特に実験値を主体としている場合などは、必ずしも大規模なデータがあるとは限らない。そこで、本研究提案では、ナノシート材料に関する自前の小規模実験データへどのようにMIを適用するかを検討し、より一般に適用できる方法論として確立することとその実践例を提示することが第2のねらいである。

2. 研究成果

(1) 概要

電荷を帯びた宿主層と層間ゲストイオンで構成された層状遷移金属酸化物を対象とし、MIを活用したナノシート材料の設計指針の確立を行った。我々がこれまで開拓してきた手法により、層間にゲストとして有機分子を導入することで柔軟な層状有機無機複合体を得て、これを有機溶剤中へ分散させてナノシートを作製した。このとき、ゲストと分散媒の種類を約100通り変化させてナノシートを合成し、収率の実測値を取得した。この収率を目的変数とし、層間ゲストと分散媒に関連した少数の本質的なパラメータ(記述子)で収率を予測することを目指し

た。機械学習によるスパースモデリングと我々実験者・研究者の考察を融合し、少数の記述子を用いた単純な収率予測モデルを構築した。この収率予測モデルを使い、未知の宿主層化合物、未知のゲスト分散媒の組み合わせにおいて、高収率および低収率を実現する条件を探索した。このとき抽出された高収率および低収率となる条件で実験を行うと、予測に近い収率でナノシートを最小実験数で得ることができた。このように、自前の少数実験データに対してスパースモデリングと研究者の考察を融合した予測モデルの構築により、ナノシートのサイズやその分布の制御も可能となった。これらの結果は、はく離によるナノシート材料の合成において、設計・予測に基づく制御に近づいたことを意味している。

以上のような小規模実験データに基づく研究者主導の MI、すなわち機械学習と研究者の考察を組み合わせた手法を、より一般的な方法論として確立していくために、適用事例の追加を試みた。我々の研究グループで扱っていた別な材料系で、高分子材料の電気化学特性の向上への応用を試みた。具体的には、リチウムイオン二次電池の有機負極活物質の探索を行った。20 種類弱の低分子有機化合物の充放電測定を行い、充放電容量の実測値を取得した。この充放電容量を説明する少数のパラメータをスパースモデリングと考察の融合によって抽出し、充放電容量の予測モデルを構築した。得られた予測モデルに基づいて、電気化学特性が未知の低分子有機化合物の容量予測を行い、最少実験数で高容量を示す化合物を発見することができた。さらに、この化合物を重合・ナノ構造制御することで得られた高分子材料は、先行研究と比較して高い容量と耐久性を両立した有機負極活物質であることがわかった。

(2) 詳細

研究テーマ A 「はく離挙動の制御によるナノシートの収率向上とサイズ制御」

本研究では、電荷を帯びた宿主層とゲストイオンの静電相互作用で構成された層状遷移金属酸化物を、ナノシート作製の対象物質とした。はく離を行う温度や時間といったパラメータがはく離挙動に与える影響は、研究者の経験や勘に基づいて推測できるため、後で実験的に最適化を行うこととした。まず、宿主層を層状酸化チタンに決め、異なる 8 種類のゲストと 13 種類の分散媒の組み合わせ合計 104 通りの条件ではく離実験を行い、ナノシートの収率測定を行った(図1)。このとき、極端に収率が低い条件や生成したナノシートの凝集が進んでしまう条件は、収率の値および動的光散乱(DLS)による粒度分布測定の結果によるスクリーニングを行って除外した。このように、ゲスト分散媒約 60 通りの組み合わせによる収率の値を目的変数として訓練データに取り入れた。次に、例えば分散媒の分子量や分極率など、収率の値を説明すると推定したゲスト分散媒および双方に関係するパラメータ 35 個を説明変数として挙げた。以上の訓練データセットに対し、機械学習によるスパースモデリングと考察を融合し、少数の記述子を用いて収率を予測する単純なモデルの構築を目指した。スパースモデリングのための機械学習として、罰則項付きの線形重回帰分析の Minimax concave plus (MCP) および全状態探索を活用し、5 個の記述子による収率予測モデルを構築した。さらに、我々の研究者としての化学的な考察によって記述子を絞り込むことや、訓練データ内のデータの偏りを排除することで、最終的に 2 個の記述子による収率予測モデルを構築することができた。本収率予測モデルの予測精度は決して高くはないが、実験科学者が未知の条件に関して予測

を行うことで実験数を削減するには十分な精度であった。実際に、本収率予測モデルを用い、未知な条件におけるテストを行った。ホスト層を訓練データから変更して層状酸化マンガンを用い、未知なゲスト分散媒の組み合わせ 156 通りの条件において予測収率を計算した。これら 156 通りの条件から、高収率上位および下位 10 条件の合計 20 条件のみではく離実験を行うと、平均収率はそれぞれ 24.0 % および 7.09 % であった。他のホスト層に変えた場合であっても、同様に予測収率の上位と下位で有意な差を確認できた。さらに、同様の手法でサイズおよびサイズ分布の制御が可能であった。以上より、目的としていたナノシート材料の設計指針の一部が確立できたと言える。研究テーマ A およびそこから派生した成果は、5. 主な研究成果リスト(3), (4), (5)の論文として公表している。

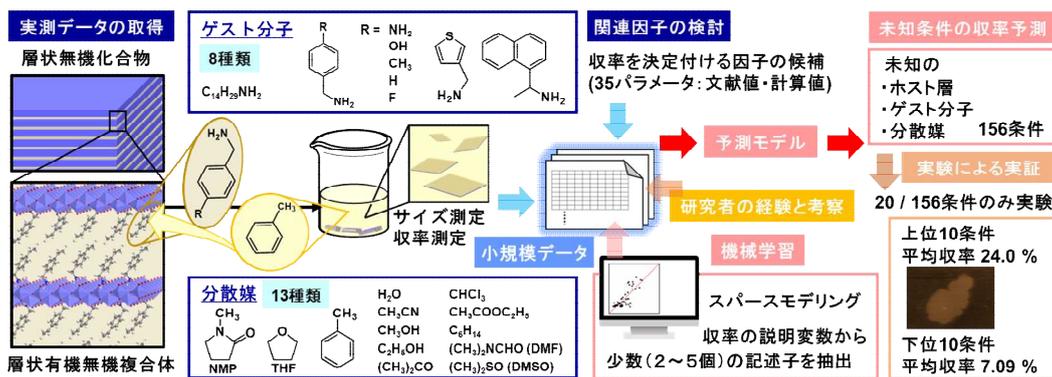


図1. 小規模実験データへの実験主導型 MI の適用によるナノシート収率の向上に関する研究成果の概要. 同様の手法はサイズやサイズ分布制御にも有効である.

研究テーマB 「小規模実験データを活用した実験主導型MIの確立に向けた取り組み」

研究テーマAでは、実験科学者自前の小規模実験データに基づき、スパースモデリングと研究者の考察を融合することで簡易な予測モデルを構築し、これを活用することで未知なる系の予測を行い、最少実験数で目的の材料合成を達成していた。本手法の具体的な手順は、まず、自前の実験によって実験データを取得し、目的とする構造や物性を説明するパラメータを挙げる。次に、機械学習を活用したスパースモデリングと考察を経て、少数の記述子による予測モデルをくみ上げる。さらに、この予測モデルに基づいて、未知の系で目的とする構造や物性に関する予測を行った上で、目的を達成できると予測された条件に絞って最少実験数で実験を行う。本手法は、より多様な系へ適用可能であると考え、いくつかの機能材料の構造制御や特性向上への応用を試みた。ここでは、リチウムイオン二次電池の有機高分子負極活物質の探索への応用を紹介する(図2)。現在、リチウムイオン二次電池の電極活物質は無機材料が主流である。資源の枯渇などの問題から、有機材料の電極活物質に関する研究も行われている。有機電極活物質については、正極に関する分子設計の指針はある程度知られているものの、負極に関する設計指針は明確ではなかった。そこで本研究では、これまで経験的に言われてきた共役骨格とカルボキシ基をあわせ持つ化合物を中心に 16 種類弱の市販の低分子有機化合物を選び、充放電測定を行って容量の実測値を取得した。この充放電容量を説明するパラメータとして、対象としている低分子化合物の軌道順位や電解液との親和性などの 24 個の説明変数を準備した。研究テーマAと同様に、スパースモデリングと考察に基づき、3 個の記述子を使った容量予測モデルを構築した。この予測式を用い、負極と

しての性能が未知の 11 個の市販の低分子化合物の容量の予測値を計算した。その中で、予測容量の高かった 3 個の化合物について実験を行うと、 100 mA g^{-1} の高電流密度での充放電測定で、2 個の化合物 (benzo[1,2-b:4,5-b']dithiophene (BdiTp) および trans,trans-1,4-diphenyl-1,3-butadiene) がそれぞれ 227 mA h g^{-1} および 310 mA h g^{-1} と、高容量を示すことがわかった。さらに、この BdiTp は重合によって高分子量化すると、容量が向上することも予測モデルからわかった。実際に、BdiTp 重合体を電解重合によって合成しナノ構造制御を行うと、電流密度 20 mA g^{-1} の充放電で 933 mA h g^{-1} の高容量と、1000 回の充放電を繰り返しても容量劣化が少ない高耐久性が両立できた。このような高容量と高耐久性の両立は、報告されている先行研究の中で最高性能に匹敵するものであった。本結果は、小規模実験データを活用した実験主導型 MI が有効な手法であり、様々な材料系へ適用できる可能性を示唆している。以上の研究テーマ B およびそこから派生した成果は、5. 主な研究成果リスト(1), (2)の論文として公表されている。

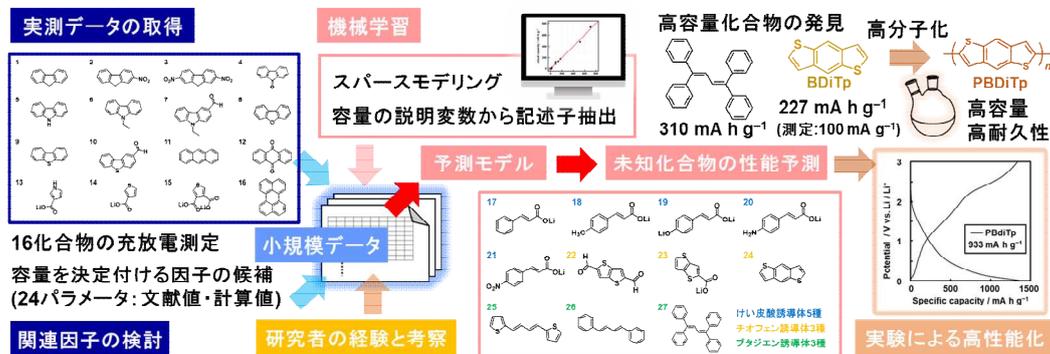


図2. 小規模実験データへの実験主導型 MI の適用によるリチウムイオン二次電池有機高分子負極活物質の新規高分子材料探索に関する研究成果の概要.

3. 今後の展開

本研究で確立したナノシートの収率向上およびサイズ制御に関する予測モデルは、様々な層状物質へ適用可能であると考えている。今後、得られた設計指針を、様々な層状物質や特にこれまで収率が低いことやサイズ制御が困難であった層状物質へ適用する。これらの取り組みにより、ナノシート材料を資源・環境・エネルギー分野へ応用することをより積極的に進める。

本研究は、小規模実験データに基づいた機械学習によるスパースモデリングと研究者の考察を融合した「実験主導 MI」が、多様な実験研究へ適用可能であることを示唆している(図3)。一般的な MI は、大規模なデータを必要とすること、プロセスがブラックボックス化しやすきこと、結果の解釈が容易ではないことが課題であった。しかし、我々の実験主導 MI では、研究

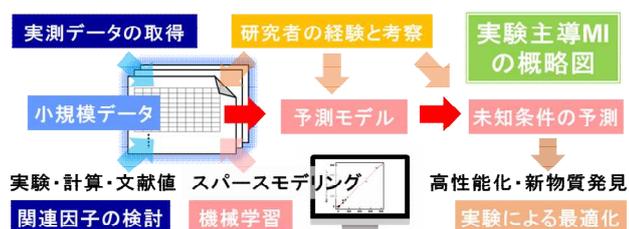


図3. 小規模データに機械学習と研究者の考察を融合させた実験主導 MI の概略図.

者の経験や考察を常に十分に活用することで、データが小規模であることを補完しつつ、結果やプロセスを解釈しながら単純なモデルを構築するため、実験科学者に即している点が特徴として挙げられる。本手法についてデータ科学的な妥当性を検証しながら、他の材料の合成プロセスの最適化や特性の向上・制御に応用することを積極的に進める。以上より、これまで小規模なデータのため適用が容易ではなかった系や、文献値やデータベースの値が利用しにくい系でも MI が適用できるようになり、研究開発の加速・高効率化が可能となる。

4. 自己評価

本研究課題は、「マテリアルズインフォマティクスを活用したナノシート材料の設計指針の構築」が第1の目的であり、さらにその過程で確立してきた「小規模データを活用した実験主導型 MI」を他の機能材料の探索に応用することが第2の目的である。いずれも研究期間内に成果は挙がり、付随した成果を含めて多くの論文を出すことができた(5. 主な研究成果リスト参照)。特に重要な成果は、論文のみならず JST から2件のプレスリリースも行き、合計6件の新聞報道などにも至った。さらに、MIに関連した解説記事の執筆依頼や招待講演の依頼を、化学および材料関係の学会、企業、関連団体等から受けるようになった。以上より、研究目的をほぼ達成でき、その社会への波及効果も大きかったと考えている。

本さがけ参画当初は、データ科学や機械学習の知見は皆無であったが、研究開始期に本さがけ1期生の物質・材料研究機構・袖山慶太郎主任研究員や北海道大学・小林正人講師の手ほどきが理解を進めるきっかけとなった。さらに、3期生の筑波大学・五十嵐康彦准教授との連携を通じ、データ科学的手法の活用やその解釈を大きく前進させることができたと考えている。また、領域会議での総括・アドバイザーの先生方や研究者との議論の中からも多くのヒントを得ることができた。このように領域内での連携や交流は、研究期間内のみならず今後の研究活動に対しても大きな意義があったと感じている。研究経費は、必要な大型機器を研究開始期に迅速にそろえられたことで効率的に実験を進めてデータを取得でき、研究期間の後半ではデータ科学的な取り組みに力を入れることができた。

最終的には、研究期間の後半で「小規模データを活用した実験主導型 MI」という形での発展性が見え始めた点は、予想外の収穫であった。実験科学者の自前実験データや産業界の研究開発部門にあるデータは必ずしも大規模でないが、本研究の方法論を応用することで、MIを使った研究開発の効率化が可能であると考えられる。また、機械学習のみに頼らず、研究者の経験・勘・考察を積極的に導入することが、MIを真に効果的に活用するために重要であることが見えつつあることも、重要な知見であると考えている。今後、本研究を出発点とし、MIに関する新しい学理を構築できるよう深化させること、より多くの適用例を見出す中から MIでしかできない材料や機能を実現すること、産業界との交流を深めることに重点を置いてより一層の発展に取り組んでいきたいと考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. K. Watanabe, H. Imai, *Y. Oaki, "Paper-based device of designer soft layered polymer composites for measurement of weak friction force" *Journal of Materials Chemistry C* **2019**, *8*, 1265.(裏表紙イラストに選出)
2. H. Numazawa, Y. Igarashi, K. Sato, H. Imai, *Y. Oaki, "Experiment-Oriented Materials Informatics for Efficient Exploration of Design Strategy and New Compounds for High-Performance Organic Anode" *Advanced Theory and Simulations* **2019**, *2*, 1900130 (10 pages).(表紙裏イラストに選出、2018-2019 Top Downloaded Papers に選出)
3. S. Yano, K. Sato, J. Suzuki, H. Imai, *Y. Oaki, "Amorphous 2D materials containing a conjugated polymer network" *Communications Chemistry* **2019**, *2*, 97 (Nature Research Chemistry Community にハイライトされる).
4. G. Nakada, Y. Igarashi, H. Imai, *Y. Oaki, "Materials-Informatics-Assisted High-Yield Synthesis of 2D Nanomaterials through Exfoliation," *Advanced Theory Simulations* **2019**, *2*, 1800180 (11 pages). (表紙裏イラストに選出)
5. G. Nakada, H. Imai, *Y. Oaki, "Few-layered titanate nanosheets with large lateral size and surface functionalization: potential for the controlled exfoliation of inorganic-organic layered composites", *Chemical Communications* **2018**, *52*, 244. (表紙裏イラストに選出)

(2) 特許出願

研究期間累積件数: 1件

発明者: 緒明佑哉, 佐藤宏亮, 矢野翔一郎, 鈴木惇平

発明の名称: グラファイト又はグラフェン様の層状構造を有する重合体

出願人: 学校法人慶應義塾

出願日: 2018/5/2

出願番号: PCT/JP2018/017519

(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)

1. 【受賞】平成30年度科学技術分野の文部科学大臣表彰 若手科学者 (2018年4月).
2. 【プレスリリース】JST プレスリリース「マテリアルズインフォマティクスを活用しリチウム電池負極用の有機材料で世界最高水準の性能を達成～少ない実験データに経験知と機械学習を融合して～ (2019年9月6日).
3. 【新聞報道】「MI 活用し有機負極材料 慶大・東大 最高水準の容量・耐久性」, 化学工業日報 (2019年9月6日).
4. 【プレスリリース】JST プレスリリース「マテリアルズインフォマティクスを活用してナノシート材料の高効率合成が初めて可能に (2019年1月10日).
5. 【新聞報道】「材料開発にAI革命」, 日本経済新聞(日曜版) (2019年2月17日).

研究報告書

「機能性固体微細材料のマイクロレベル電子物性解析基盤技術の構築」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 大塚 朋廣

1. 研究のねらい

近年、エネルギー、環境問題に伴う省エネルギー化の推進や、情報化社会の進展による情報処理能力の強化のため、スピントロニクスや量子エレクトロニクス等の革新デバイス、革新エレクトロニクスを創出して、電子デバイスの超低消費電力化、情報処理能力の向上といった性能の飛躍的進歩を実現する社会的重要性が増している。この革新デバイス、革新エレクトロニクスにおいては、従来にも増して固体微細材料中の局所電子状態が本質的に重要な役割を果たすため、固体微細材料中でのマイクロな電子物性を直接的に測定し、その物理現象を解明することが重要となる。

固体微細材料の評価を行う上で、電子物性に直接アクセスできる電子輸送測定は有効なツールの一つであり、さらに固体微細材料内部に生じる局所電子状態に直接的にアクセスできるマイクロなプローブが有用である。本研究者はこれまで半導体量子ドット等を用いた局所電子状態プローブを開発してきた。この手法では測定対象からプローブ量子ドット内人工量子準位への電子のトンネル等を解析することにより、測定対象内の局所電子状態を調べることができる。そこで本研究では固体微細材料中の局所電子状態に直接アクセスできる独自のマイクロプローブによる分析手法を、固体微細材料内電子状態のシミュレーション等の理論、測定データを効率的に活用するデータ科学と融合させ、機能性固体微細材料のマイクロレベルからの電子物性解析基盤技術の構築を行う。これにより固体微細材料内での電子物性分析精密化、マイクロな物理現象の解明等を進め、機能性固体微細材料、デバイスの開発を効率的に進めるための基盤技術を開発する。

2. 研究成果

(1) 概要

固体微細材料中局所電子状態についてのマイクロプローブによる測定手法を、電子状態のシミュレーション等の理論、データを効率的に活用するデータ科学手法と融合させて、機能性固体材料のマイクロレベルからの電子物性を調べる基盤技術の構築を行った。本研究では大きく分けて2つの研究項目として、「マイクロプローブ測定と電子状態シミュレーションの融合による機能性固体微細材料の基礎物理解明手法の開発」、「マイクロプローブとデータ科学の融合による新ダイナミクス測定手法、及び機能性固体微細材料のダイナミクス解析基盤技術の開発」を実施した。

第一の研究項目では固体微細材料として半導体量子ドットと電極が結合した固体微細構造を測定ターゲットとして、半導体量子ドットを用いたマイクロプローブを活用して、局所電子・スピン状態を単一電子のレベルで測定した。そしてトンネル過程を取り入れたマイクロな理論、シミュレーションの結果と比較することにより、その物理過程を解明した。

第二の研究項目ではマイクロプローブからのデータ解析にデータ科学手法を取り入れた。ベイズ確率を活用した解析手法について開発し、高精度な状態推定を行った。またマイクロプローブ試料等の微細構造試料制御についても、機械学習を用いた手法について研究を行い、固体微細構造試料において重要となる微小電極電圧による静電ポテンシャル形成の最適化に機械学習を適用した。

(2) 詳細

研究テーマ A: マイクロプローブ測定と電子状態シミュレーションの融合による機能性固体微細材料の基礎物理解明手法の開発

固体微細材料を用いて有用なデバイスを創製していくためには、その内部の局所電子状態をマイクロに観測し、その物理現象を解明することが重要である。本研究者はこれまで半導体量子ドット等を用いて微細材料中の電子状態を局所的($\sim 10\text{nm}$)に、高精度($\sim 10\ \mu\text{eV}$)かつ低擾乱($\sim 100\text{aA}$)に測定できる独自のマイクロプローブを開発してきた。本研究ではこのマイクロプローブ測定を局所電子状態のマイクロな理論、シミュレーションと組み合わせて固体微細材料中の物理現象を解明した。まず半導体量子ドットを用いたマイクロプローブ試料について、半導体微細加工技術を用いて作製した。またマイクロプローブの測定について、熱雑音を低減した低温下で高精度にかつデータ科学的処理に使える大量のデータを短時間で取得するために、低温冷凍機内で高周波を活用した高周波反射測定(\sim 数 100MHz)を行う測定系を構築、改良した。これらの測定セットアップを用いて、半導体量子ドットと電極が結合した固体微細構造における、局所電子・スピン状態を測定した。この結果、局所電子の電荷・スピン状態の時間変化を単一電子のレベルで測定することに成功した。

この測定結果について、トンネル過程を取り入れたマイクロな理論、シミュレーションを実施して比較を行った。この結果、単一電子トンネル過程における電荷とスピン状態の時間変化の差異、高次のトンネル過程によるスピン状態変化等の新現象について、その物理過程をマイクロに解明した。これらの成果は、固体微細材料中の局所電子状態解明の基礎物理で重要であるとともに、スピン量子ビット等における状態初期化や、スピン状態操作等に活用できるため、単一電子状態を活用したスピントロニクスや量子エレクトロニクス応用等に向けても有用である。またここで実証したマイクロプローブによる局所的な電子物性測定手法は、従来からの半導体微細材料だけでなく、近年注目を集めている原子層材料やトポロジカル材料等、他の固体微細材料測定にも応用可能である。

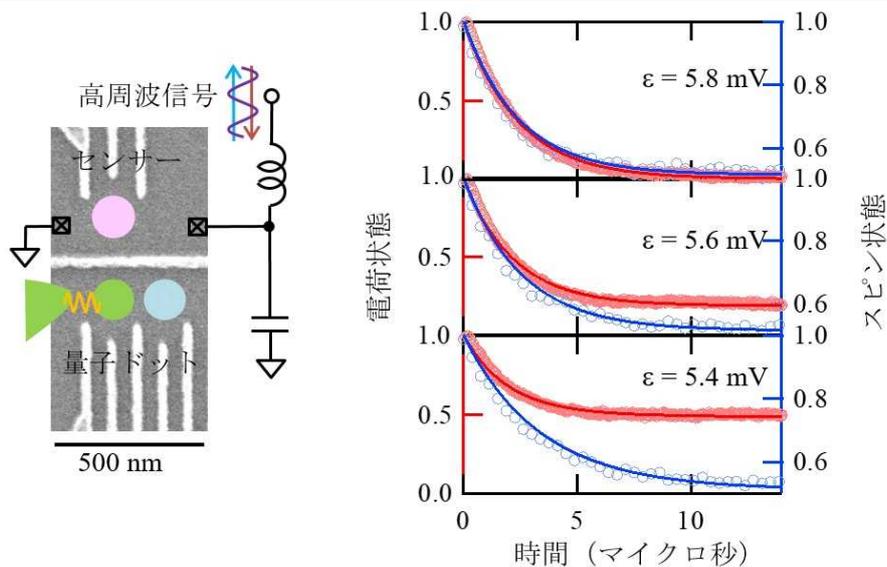


図1 ミクロプローブの電子顕微鏡写真と、ミクロプローブを用いた固体微細材料中の単一電子電荷・スピン状態の観測。

研究テーマ B: ミクロプローブとデータ科学の融合による新ダイナミクス測定手法、及び機能性固体微細材料のダイナミクス解析基盤技術の開発

高周波技術を活用したミクロプローブの高速化によって、短時間で大量のデータ取得が可能となり(1秒間に~1億データ)、その解析にデータ科学手法を活用できる状況が整ってきた。本研究ではこのミクロプローブからのデータ解析にデータ科学手法を活用する研究を行った。ミクロプローブ測定では測定対象の電子状態を、プローブ量子ドット内の単一電荷の有無から解析する。この単一電荷の有無について、実験では避けることのできないノイズのある実測定データから推定を行うのであるが、本研究ではこの推定にベイズ確率を活用した手法を開発した。この新手法では従来の解析では十分に活用されていなかった実時間のすべての測定データを活用することができ、より高精度な状態推定を実現できる。この動作をシミュレーションおよび実験データを用いて調べた。またこのベイズ手法による解析をミクロプローブ測定と同時にリアルタイムで実行するために、高速なロジックゲート処理を実現できる FPGA による信号解析回路を開発した。

また機械学習を用いたミクロプローブ試料制御についても研究を実施した。ミクロプローブ試料等の半導体微細構造試料では、複数の微小電極電圧により形成される半導体微細構造内の静電ポテンシャルが電子機能を創成する。この静電ポテンシャルの最適化は、これまで実験家の経験と感により実現されてきたが、本研究ではこの過程に機械学習を適用する研究を行った。シミュレーションにより微小電極電圧に対する半導体量子ドット電荷状態のデータセットを用意し、これに対してランダムフォレスト、畳み込みニューラルネットワーク等の機械学習手法を用いて学習を行った。この学習器を利用して、電極電圧の評価、自動調整の手法を研究した。またミクロプローブ測定の手法を、半導体量子ドット中の電子スピンを用いた量子ビットの実験研究等にも活用した。これらの成果は、ミクロプローブを活用して固体微細材料の局所電子状態ダイナミクスの解明を進め、デバイスの高速化等を推進していく上で有用である。さらに

機械学習によるデバイス最適化等は、量子情報処理に向けた将来の大規模量子デバイス等の実現においても重要となる。

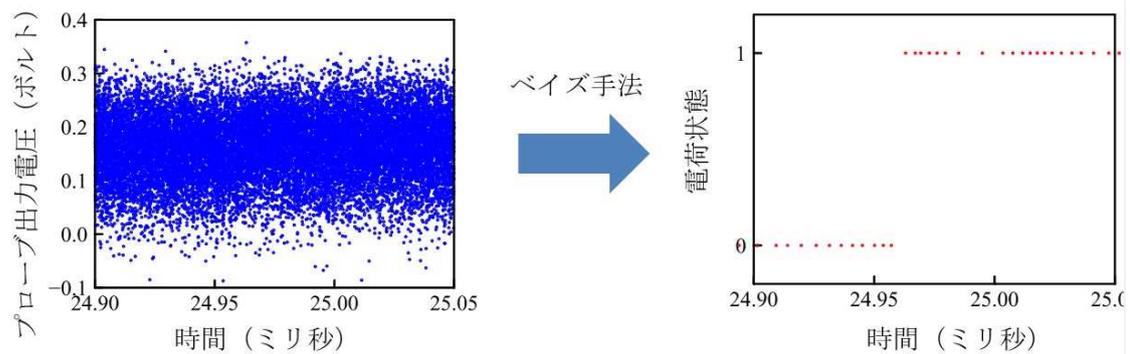


図2 大きなノイズを含むマイクロプローブからのデータ列と、ベイズ手法を活用したマイクロプローブデータ解析。

3. 今後の展開

本研究によりマイクロプローブによる分析手法を、固体微細材料内電子状態のシミュレーション等の理論、マイクロプローブによる測定データを効率的に活用するデータ科学と融合させ、機能性固体微細材料のマイクロレベルからの電子物性解析基盤技術を実現した。今後の展開としては、この電子物性解析技術を従来の半導体微細材料だけでなく、他の系にも適用していくことがあげられる。具体的なターゲットとしては、近年注目を集めている原子層材料やトポロジカル材料等を考えている。また機械学習を用いた固体微細構造試料調整の効率化、大規模化も今後重要性がさらに増してくると考えられる。半導体量子ドット等を用いた量子ビット集積による実用的な量子コンピュータ実現のためには、大規模に集積化した量子デバイスが必要であり、このデバイス調整過程は従来の人力による手法では実現できない可能性がある。そこで機械学習による調整の効率化、大規模化が有用となると考えられる。

4. 自己評価

「マイクロプローブ測定と電子状態シミュレーションの融合による機能性固体微細材料の基礎物理解明手法の開発」、「マイクロプローブとデータ科学の融合による新ダイナミクス測定手法、及び機能性固体微細材料のダイナミクス解析基盤技術の開発」、いずれの項目においても成果を出すことができたと考えている。特にさきがけ参画により、領域会議での議論や共同研究等を通して、データ科学手法を学び、研究に取り入れることができた点が、これまで実験のみを主体としてきた本研究者にとって重要であった。一方、課題としては開発したマイクロレベル電子物性解析基盤技術を、従来の半導体微細材料にとどまらず、幅広い系に適用していくことがあげられる。これにより近年注目を集めている原子層材料やトポロジカル材料等について、新しい観点からの実験データを提示し、研究の進展に貢献できると考えられる。また多体系や大規模系への拡張も今後の課題である。

将来的には、固体微細材料における局所電子状態を活用して、スピントロニクス、量子エレクトロニクス等の新エレクトロニクスデバイスの創製に貢献し、省エネルギー化の推進や、情報化社会に向けた情報処理能力の強化に貢献していきたいと考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. T. Otsuka, T. Nakajima, M. R. Delbecq, P. Stano, S. Amaha, J. Yoneda, K. Takeda, G. Allison, S. Li, A. Noiri, T. Ito, D. Loss, A. Ludwig, A. D. Wieck, and S. Tarucha, “Difference in charge and spin dynamics in a quantum dot–lead coupled system”, *Physical Review B* 99, 085402 (2019).
2. T. Nakajima, A. Noiri, J. Yoneda, M. R. Delbecq, P. Stano, T. Otsuka, K. Takeda, S. Amaha, G. Allison, K. Kawasaki, A. Ludwig, A. D. Wieck, D. Loss, and S. Tarucha, “Quantum non–demolition measurement of an electron spin qubit”, *Nature Nanotechnology* 14, 555 (2019).
3. T. Ito*, T. Otsuka*, T. Nakajima, M. R. Delbecq, S. Amaha, J. Yoneda, K. Takeda, A. Noiri, G. Allison, A. Ludwig, A. D. Wieck, and S. Tarucha, “Four single–spin Rabi oscillations in a quadruple quantum dot”, *Applied Physics Letters* 113, 093102 (2018), (*equal contribution).
4. J. Yoneda, K. Takeda, T. Otsuka, T. Nakajima, M. R. Delbecq, G. Allison, T. Honda, T. Kodera, S. Oda, Y. Hoshi, N. Usami, K. M. Itoh, and S. Tarucha, “A quantum–dot spin qubit with coherence limited by charge noise and fidelity higher than 99.9%”, *Nature Nanotechnology* 13, 102 (2018).
5. T. Otsuka, T. Nakajima, M. R. Delbecq, S. Amaha, J. Yoneda, K. Takeda, G. Allison, P. Stano, A. Noiri, T. Ito, D. Loss, A. Ludwig, A. D. Wieck, and S. Tarucha, “Higher–order spin and charge dynamics in a quantum dot–lead hybrid system”, *Scientific Reports* 7, 12201 (2017).

(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)

< 学会発表 >

1. T. Otsuka, T. Nakajima, M. R. Delbecq, P. Stano, S. Amaha, J. Yoneda, K. Takeda, G. Allison, S. Li, A. Noiri, T. Ito, D. Loss, A. Ludwig, A. D. Wieck, and S. Tarucha, “Charge and Spin Dynamics in a Quantum Dot–Lead Hybrid System”, *International Symposium on Frontiers of Quantum Transport in Nano Science*, Kashiwa, Japan, Nov. 9, 2018 (invited).
2. T. Otsuka, S. Nagayasu, T. Nakajima, M. R. Delbecq, J. Yoneda, K. Takeda, A. Noiri, S. Li, T. Ito, S. Tarucha, “Speed up of quantum dot charge sensing utilizing Bayesian estimation”, *International Conference on the Physics of Semiconductors*, Montpellier, France, Jul. 30, 2018 (oral).
3. 大塚朋廣、永安修也、中島峻、Matthieu R. Delbecq、米田淳、武田健太、野入亮人、Sen Li、伊藤匠、樽茶清悟、「固体微細構造マイクロプローブ計測へのデータ科学手法応用」、日本表面真空学会関東支部講演大会、東京、2019年4月13日(招待講演)

<受賞>

4. 文部科学大臣表彰若手科学者賞 2018年4月17日
5. 日本物理学会若手奨励賞 2017年3月18日

研究報告書

「半導体材料開発のための計算データベース構築」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 熊谷 悠

1. 研究のねらい

近年の計算機性能の向上と計算手法の改善により、実験結果に依存しない第一原理に基づいた新材料探索が現実のものとなってきた。その先駆的な仕事の1つとして、我々は2016年に、600程度の異なる組成を持つ亜鉛窒化物に対して系統的な第一原理計算を行い、内21物質が電子材料として有望な物質であることを提案した(Hinuma *et al.*, Nature commun., 2016)。そして最も有望な新規物質である CaZn_2N_2 について、実験グループと協働することで合成に成功している。一方、21物質の内、8物質が材料データベースですでに報告されている物質であった。さらに、内4物質に関しては、過去に合成報告がされていたにもかかわらず、詳細な理論・実験解析が行われていなかったために、半導体材料として見過ごされてきた物質である。このように既報物質の中にも半導体材料として見過ごされている革新的物質が数多く存在し、第一原理計算を系統的に行うことで、それらの既報物質の中から、人類に役立つ新材料を見出すことができると期待される。

しかしながら、世界中で開発されている Materials Project(<https://materialsproject.org>)や Aflow(<http://afloplib.org>)、OQMD(<http://oqmd.org>)などの計算材料データベースでは、取り扱う物性が主に結晶構造、全エネルギー、一般化勾配近似を用いた電子構造に限られており、高精度なバンドギャップ値や光吸収係数、点欠陥特性などの半導体物性に関する情報が欠落している。そこで本研究では、バンドギャップを有する非金属材料を主な対象とし、多様な半導体物性値を有する計算材料データベースを構築することを目的とした。また得られた大規模データに対して機械学習を行うことで、高精度かつ高速に物性を予測するための回帰式を提案した。

本研究で目指すデータベースは、半導体材料のみならず、触媒材料やイオン伝導体などの非金属材料全般において有用なものになると思われる。そこでプロジェクト終了後には、様々な応用的観点からの解析を行うとともに計算データの公開を行う。

2. 研究成果

(1) 概要

本研究ではまず、計算材料データベースを構築する際に必要な第一原理計算を自動実行するための pydefect コード、vise コード及び自動計算コードの3つのプログラム開発を行なった。本研究で採用した自動計算システムでは、有向非巡回グラフを用いて逐次的に行う処理を表現し、それぞれの処理をさらに小さな単位の処理を組み合わせることで表現する。このように個別に開発・テストされたパーツを様々な組み合わせることで、手動では不可能な数の系統的な第一原理計算が可能となる。

次に、これらの計算プログラムを用いて計算材料データベースの開発を行なった。まず実験の結晶構造の実験値の再現性が良いとされる PBEsol 汎関数を用いて 32,000 物質の構造

最適化を行い、それらのバンド図・状態密度の計算を行った。そして、バンドギャップを有すると判定された 12,000 物質に関しては、摂動的 HSE06 法を用いてより高精度なバンド構造の計算を、数千物質に関しては光吸収スペクトル、動的安定性の計算を行った。

また、数千種類の酸化物を対象に、誘電定数や酸素空孔形成エネルギーの計算を行い、それぞれの傾向を議論した。さらにランダムフォレストによる機械学習を行い、誘電定数や酸素空孔形成エネルギー共に、比較的良好な精度で予測できることを検証した。

(2) 詳細

計算材料データベースを構築するためには、数百万回以上の第一原理計算を実行する必要がある、この数の計算を手動で行うことは、人的コストの面で不可能である。そこで、第一原理計算入力ファイルの自動生成、計算過程でエラーが生じた場合の自動的なハンドリング、得られた計算結果の自動解析とデータベース化、といったことが重要となる。本研究では、(i) 点欠陥計算に必要な入力ファイルの生成と出力結果の解析を行う pydefect コード、(ii) 第一原理計算パッケージ VASP コード (Kresse and Hafner, 1993) の入力ファイルの生成と得られた出力結果の解析を行うための vise コード、(iii) 自動計算ワークフロープログラム Fireworks (Jain *et al.*, 2015) および atomate (Mathew *et al.*, 2017) を独自に拡張した自動計算コードの 3 つのプログラムを開発した。これらのプログラムは python3 を用いて構築されており、それぞれ、およそ 1 万行程度で

構成されている。尚、これらのプログラムの一部に関しては本プロジェクト終了後に公開する予定である。

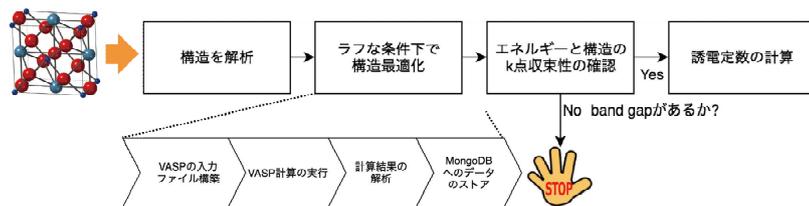


図 1. 本研究で用いた第一原理計算ワークフローモデルの例。

本研究で用いた、Fireworks における計算ワークフローモデルの例を図1に示す。誘電定数の例で示すように、逐次的に行う処理を、有向非巡回グラフを用いて表し、またそれぞれの処理に関しても、さらに小さな単位の処理を組み合わせることで表現する。このように様々な組み合わせさせたパーツからフローを構築することで、手動では不可能な数の系統的な第一原理計算が可能となる。また本ワークフローモデルでは、重複する物質の第一原理計算の回避や mongoDB を用いた容易なデータマネージメントなどの利点がある。これら 3 つの計算プログラムを用いて得られた計算材料データベースの詳細について以下に示す。

研究テーマ A「計算材料データベース作成」

本研究では、図2に示すような流れで、既報物質を対象に系統的な第一原理計算を行った。本

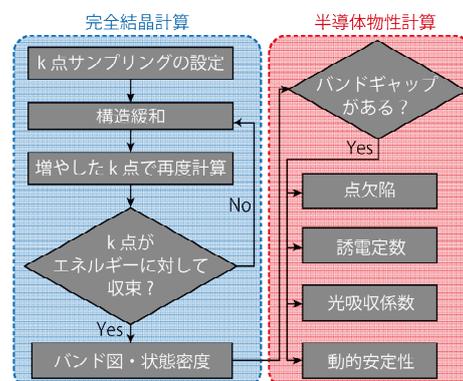


図 2. 本研究での計算材料データベース構築のためのワークフロー。

研究では、実験の結晶構造の実験値の再現性が良いとされる PBEsol 汎関数を用いた。まず 32,000 物質のバンド図・状態密度の計算を行い、バンドギャップを有すると判定された 12,000 物質に関しては、我々が提案した摂動的 HSE06 法(Hinuma et al., Phys. Rev. B, 2017)を用いて高精度なバンド構造の計算を行った。また、数千物質に関しては光吸収スペクトル、動的安定性の計算を行った。

研究テーマ B「酸化物の誘電定数および酸素空孔の大規模計算とその機械学習」

本研究では、酸化物の誘電定数および酸素空孔の大規模計算を実行した。誘電定数に関しては、高誘電性を有するワイドギャップ半導体の探索に、点欠陥に関しては酸化物半導体や酸素イオン伝導体の探索に重要な指標となる。

ここで対象とした酸化物は、(i)対称性が P1 でない、(ii)アニオンが酸素のみで構成されている、(iii)40原子以下の原子数でユニットセルが構成されている、(iv)Materials Project の計算結果より磁性を有していない、などの条件を満たす数千種類の酸化物である。

本研究により得られた電子系および格子系誘電定数とバンドギャップの関係性を図3上に示す。このように、電子系誘電定数の最大値はバンドギャップと反比例の関係にあることがわかる。一方、格子系誘電定数は、バンドギャップとの強い相関を示さず、また電子系と比べて大きな値になる傾向がある。このことから、ワイドギャップ高誘電物質の探索には、格子系誘電定数に注目することが重要であると言える。また、機械学習の手法であるランダムフォレストを用いた回帰を行った結果(図3下)、電子系誘電定数は極めて高い精度で予測されている。一方、格子系誘電定数は、電子系には劣るものの、高誘電物質探索に有用なレベルでの予測が可能となっている。

次に、図4上に酸素分子を基準とした酸素空孔の形成エネルギーを示す。これから、酸素空孔形成エネルギー4.64 eVを平均として、0.2 eV から 7.5 eV と幅広く分布している様子がわかる。また誘電定数と

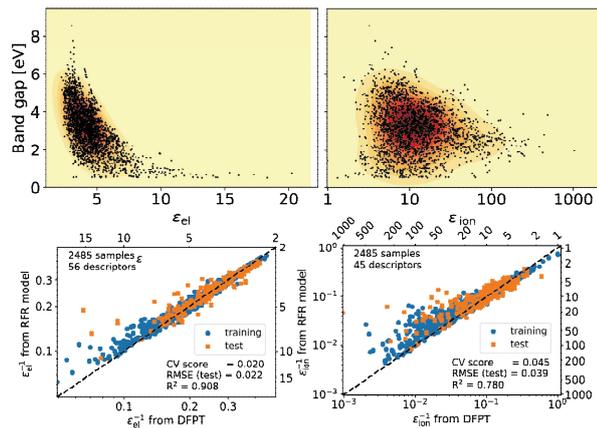


図 3. (上) 誘電定数のバンドギャップ依存性と (下)機械学習による予測。

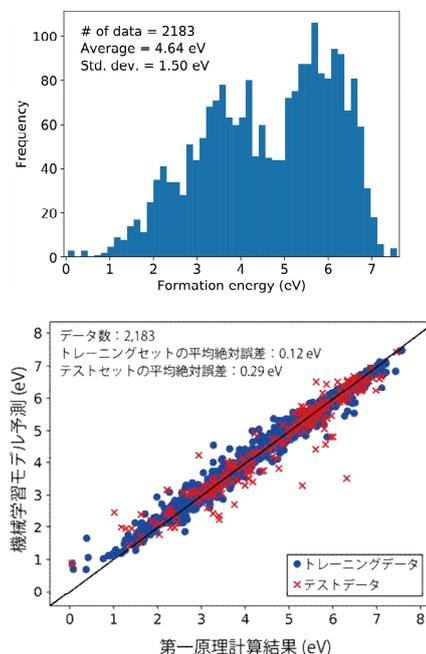


図 4. (上) 酸素空孔形成エネルギーの分布図と(下)機械学習による酸素空孔形成エネルギーの予測。

同様、機械学習を用いることで、0.3eV 以下の誤差で酸素空孔形成エネルギーを予測することができることが可能となった。

3. 今後の展開

本さがけプロジェクト終了後は、開発した自動計算プログラムを、新物質・新材料探索へと展開していく。本プロジェクトでは既存の物質のみを対象として研究を行ってきたが、今後は、プロトタイプ構造に基づいた系統的・網羅的計算や、第一原理計算と機械学習を1つのワークフロー中に融合することで、新物質・新材料探索の手法開発を行う。また、予測された物質に関しては実験家と協働することで、新材料探索の開発速度を向上させる仕組みの構築をめざしていく。

4. 自己評価

本研究における成果で最も重要な成果は、上述した3つのプログラムである。公開されている既存プログラムを最大限に活用しつつ、必要な部分のみプログラム開発を行うことで、少人数でありながら非常に有用な自動計算システムを構築できたと自負している。

また研究成果に関しても、32,000 物質のバンド図・状態密度、12,000 物質の高精度バンドギャップ、数千以上の点欠陥、誘電定数、光吸収係数の計算結果を得ることができており、個人プロジェクトとしては十分な成果と言える。また、物質探索に関しても、一定の成果が上がっており、ある程度当初の目標を達成できたものと自負している。

研究の進め方については、自動計算プログラムの作成にかなりの時間がかかってしまったことが惜まれる。しかしながらプログラム上の問題は研究初期段階に見通すことは極めて困難であることからある程度は止むを得ないものと考えている。また、研究費の執行に関しては、プログラムが整った最終年度に大きく用いることでメリハリのある使い方ができた。

研究成果については、特許に関わる検討が終了し次第、論文発表とプログラムの公開により社会への還元を行う。そして、得られた基盤技術に関しては、日本国内での情報共有を行い、上述した新物質・新材料探索を含めてさらなる研究へと発展させていきたい。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. T. Gake, Y. Kumagai (共同第一著者), C Freysoldt and F. Oba
“Finite-size corrections for defect-involving vertical transitions in supercell calculations”
Phys. Rev. B, **101** (2020) 020102(R).
2. N. Tsunoda, Y. Kumagai, and F. Oba
“Stabilization of small polarons in BaTiO₃ by local distortions”
Phys. Rev. Mater., **3** (2019) 114602.
3. N. Tsunoda, Y. Kumagai(共同第一著者), M. Araki, and F. Oba
“One-dimensionally extended oxygen vacancy states in perovskite oxides”
Phys. Rev. B, **99** (2019) 060103(R)-1-5.
doi.org/10.1103/PhysRevB.99.060103
4. N. Tsunoda, Y. Kumagai, A. Takahashi, and F. Oba

<p>“Electrically benign defect behavior in ZnSnN₂ revealed from first principles” <i>Phys. Rev. Applied</i>, 10 (2018) 011001–1–6. doi.org/10.1103/PhysRevApplied.10.011001</p>
<p>5. <u>Y. Kumagai</u>, N. Tsunoda (共同第一著者), and F. Oba “Point defects and <i>p</i>-type doping in ScN from first principles” <i>Phys. Rev. Applied</i>, 9 (2018) 034019–1–10. doi.org/10.1103/PhysRevApplied.9.034019</p>
<p>6. <u>Y. Kumagai</u>, K. Harada, H. Akamatsu, K. Matsuzaki, and F. Oba “Carrier-Induced Band-Gap Variation and Point Defects in Zn₃N₂ from First Principles” <i>Phys. Rev. Applied</i>, 8 (2017) 014015–1–12, <i>Editors’ suggestion</i>. doi.org/10.1103/PhysRevApplied.8.014015</p>
<p>7. <u>Y. Kumagai</u>, K. T. Butler, A. Walsh, and F. Oba “Theory of ionization potentials of nonmetallic solids” <i>Phys. Rev. B</i>, 95 (2017) 125309–1–10. doi.org/10.1103/PhysRevB.95.125309</p>

(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

“酸化物物性の系統的計算”

熊谷 悠

第 66 回応用物理学会春季学術講演会、東京、3/9 – 3/12 (2019).

“半導体材料開発のための計算材料データベース”

熊谷 悠

日本物理学会第 73 回年次大会、東京、3/25 (2018).

“First-principles calculations on point defects in semiconductors and insulators”

Yu Kumagai

The 20th Asian Workshop on First-Principles Electronic Structure Calculations,
Nanjing, China, Oct. 30 – Nov. 1 (2017).

受賞

第 39 回本多記念研究奨励賞 (平成30年5月)

第 15 回村上奨励賞(平成30年9月)

研究報告書

「高分子物質のトポロジカル構造解析による新規物性の探索と設計」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 坂上 貴洋

1. 研究のねらい

身の周りには、ゴムや、ガラス、繊維、プラスチックからゲルなどの食品、化粧品、更には生体まで高分子性の物質が溢れている。高分子が、そうでない低分子性物質と比べて材料として注目されるのには様々な理由があるが、高分子性物質の材料としての最も顕著な特性は「絡み合い」に起因する。絡み合いは、「紐はすり抜けることが出来ない」ということに由来する高分子系に内在する相互作用であり、材料形成の基礎となる高分子濃厚系の流動物性に本質的に重要であるばかりでなく、結晶化過程、高分子ガラスや繊維の強度にも関与する。従来、このトポロジカルな拘束は理論的には「管模型」により表現され、その動的性質は「レプテーション模型」により記述されてきたが、近年、これらの理論では記述できない現象が様々な系において報告されてきている。これらの理解は高分子科学の基礎的問題として重要であり、同時にこれからの新規材料開発においても大きな可能性を秘めている。シミュレーションにより大量のデータ採取が可能となりつつある現在、そこから本質的な情報を抽出し、今後の材料設計の指針となるような基礎理論構築が強く望まれる。

本研究では、大規模計算機シミュレーションから得られる大量の高次元データから、トポロジカルな拘束を表現する情報を抽出し、理論研究との融合により、新規な絡み合い状態の高分子系の構造的特徴を明らかにする。また、これらの構造特性と付随する力学特性、流動物性を結びつけ、従来の枠を超えて「絡み合い」の概念を深めていくことを目指す。これにより様々な新規絡み合い構造について確固たる物理的描像を獲得し、材料科学や生命科学分野の実験的研究との連携も視野に入れつつ、絡み合い構造の制御という視点から新規高分子性材料開発の道を切り拓く。

2. 研究成果

(1) 概要

本研究では、濃厚高分子溶液の大規模シミュレーションから得られる数値データから、系の絡み合いについての情報を引き出すための新たな解析手法として、「動的絡み合い解析法」と「パーシステントホモロジー法」を構築、提案した。

「動的絡み合い法」は、系内の協働運動から絡み合いについての時空間構造を抽出する手段を構築し、この解析方法を従来からよく知られる線形鎖の濃厚絡み合い溶液と、従来の絡み合い描像では記述することのできない「新規絡み合い構造」のプロトタイプである環状高分子鎖の濃厚溶液に適用した。

また、「パーシステントホモロジー法」では、環状鎖の系に特有の鎖間の形態状態「貫通形態」(threading configuration)に対する明確な定義を与え、パーシステントホモロジーの手法を適用、拡張することにより貫通形態を定量化する手法を確立した。

環状鎖の系では、分子鎖内の結び目、異なる分子鎖間の絡み目についてのトポロジカルな拘束が顕著に現れる結果、ダイナミクスはもちろんのこと、鎖の形態も非自明なものとなり、近年、単なる理論上のモデルではなく、その興味深い物性について、化学、材料科学、生物学を含む多様な分野から関心が持たれている。本研究成果の適用により、これらの系を定量的に解析していく上で有用な指針を与えることが期待される。

(2) 詳細

研究テーマ1「高分子濃厚系の動的絡み合い解析法の確立」

・はじめに

濃厚高分子溶液における絡み合い相互作用の記述は、高分子物理学における難問として知られ、同時にその重要性から多くの研究者の関心を惹きつけてきた。現在、これについての標準的理解は「管模型」によるところが大きい。管模型では、絡み合い相互作用という極めて複雑な多体問題を、絡み合いを引き起こす周りの高分子(障害物)により形成された管状の領域中に閉じ込められた一本鎖の問題に置き換える。絡み合い高分子系のダイナミクスやレオロジーの議論においては、個々の高分子鎖の運動モードとして、管状の領域をくねくねと這うような運動(レプテーション)が想定される。このような多体問題を絡み合いの「平均場中」での一体問題としてみなすという見方は広く受け入れられており、数値シミュレーションや実験データから絡み合いに起因するダイナミクスを調べる場合にも、一本鎖の運動(例えば、鎖の重心や、鎖中のモノマーの拡散など)に着目した解析が行われることが多い。また、絡み合いの空間構造を可視化、定量化する手法として、原始経路解析(primitive path analysis)という手法が知られるが、この解析法のアイデアは管模型に由来する。ところが、上述(上記1. 研究のねらい参照)のように、環状鎖系を代表とする新規絡み合い系においては、従来のような管模型、レプテーション模型による記述自体が非自明であり、絡み合いの概念そのものに対しても挑戦的な問いかけがなされているといつてよい。ここでは、絡み合いという現象にたいして、新たな見方、解析手法の提案が望まれる。

・動的絡み合い解析法の導入

本研究では、従来とは異なるアプローチとして、異なる鎖間の協働運動の解析を通して絡み合い構造を定量化する手法～動的絡み合い解析法～を提案し、実際に線形鎖と環状鎖(リンク無し)の濃厚溶液に適用した。この解析法の背後にある考え方は極めて単純であり、ある意味で原始経路解析と相補的なものと考えられる。すなわち、原始経路解析法は、管模型における管の空間構造の解析に特化しているのに対し、動的絡み合い解析法では、絡み合いにより引き起こされる鎖間の協働運動を定量的に解析し、そこから絡み合い構造の時空間構造を抽出することができる。

図1に、環状鎖濃厚溶液に動的絡み合い解析法を適用することにより得られた協働運動の様子をベクトル場として可視化したものを示す。このような分子鎖間の協働的な運動は、これまであまり議論されることはなかったものの、鎖の慣性半径と同程度(もしくはそれ以上)の大きなスケールで起こっており、モノマーの緩和時間のような微視的な時間スケールに比べてずっと長い時間に渡り持続することが観測される。線形鎖の濃厚系においても一見すると類似

の協同運動が見られるが、それらのデータを定量的に解析することにより、線形鎖の系と環状鎖の系とでは、絡み合いの時空間構造が大きく異なることが明らかになった。また、これらの結果に対し、平均場近似を適用することにより、現象論的な理論モデルとの比較を試みた。その結果、線形鎖の系についての結果は、従来のレプテーション理論と矛盾しないことがわかった。その一方、環状鎖の系では、絡み合っている鎖の本数が時間スケールに応じて変化することが示唆されるなど、新規絡み合い構造の包括的理解に資する新たな知見が得られた。

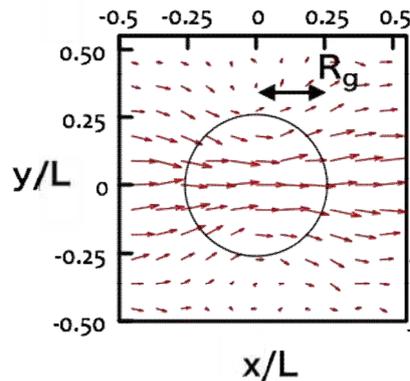


図1: 分子動力学シミュレーションから得られるデータに動的絡み合い解析法を適用することにより、異なる鎖間の運動の協同性を定量化することができる。上図は環状鎖の濃厚溶液(リンク無し)で $\Delta t \sim 10^3 \tau_0$ の時間スケール (τ_0 : LJ time unit) において見られる協同運動の様子をベクトル場として可視化したもの。

研究テーマ2「環状鎖系におけるトポロジカルな拘束の定量化に向けたパーシステントホモロジー法の拡張」

・はじめに

トポロジ的絡み合い(リンク)のない環状鎖の濃厚溶液では、トポロジー不変の拘束条件が顕著となり、線形鎖の系とは大きく異なる振る舞いを示す。一般に、環状高分子鎖の物理においては未開拓な部分が多く、その振る舞いや物性について不明な点も多い。ここでの困難の主要因として、鎖間の「関係」を記述する手段が未熟であることがあげられる。例えば、互いにリンクしていない二つの環状鎖を取り上げよう(図2)。この図に示した二つの例では、どちらもトポロジーは同じ(リンク無し)であるが、その位置関係は定性的に異なり、右図では、一方の環状鎖が他方を「貫通」(threading)している。このような、貫通現象は、系の性質(個々の鎖の形態統計やダイナミクス)に影響することが期待されるが、現在、これを定量化、記述する手法は皆無であると言ってよい。

本研究では、上記の問題に対しパーシステントホモロジーの手法を用いた解析法を提案した。

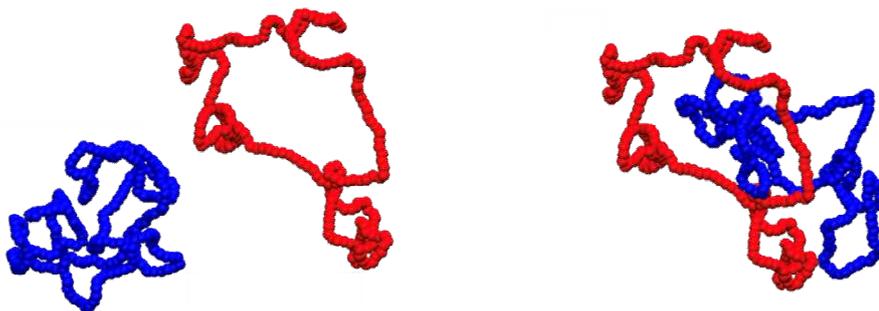


図2: 絡み合い(リンク)無しの二つの環状鎖の配位関係。左図と異なり、右図では、青色の環状鎖が赤色の環状鎖の中を「貫通」している。

・「ホモロジカル貫通」(H-threading)の概念導入

そもそも、3次元空間で複雑なコンフォメーションをとっている環状鎖に対して「貫通」事象を定義することは自明ではない。そこで、それぞれの環状鎖から得られるパーシステントダイアグラム(PD)に集合演算を施し、環状鎖間のホモロジー的な「関係」を表現する新たなPDを作成する手法を考案した。また、これらのPDの解析から、様々な空間スケールで見られる「貫通」様の事象を抜き出し、これを「ホモロジカル貫通」(H-threading)として定義した。

この解析法を数値シミュレーションより得られたリンク無し環状鎖の濃厚溶液のデータに適用することにより、以下の結果を得た。(i) リンク無し環状鎖の濃厚溶液においては、個々の環状鎖は3次的にクラumpingした複雑な形態をとり、様々なスケールにおいて分子内にループ状構造が見られる。(ii) これらの分子鎖内ループのうち、ほとんど全てのループが近傍に存在する他の鎖により貫通されている。(iii)系内に稀に大きく開いたループが存在し、その穴のサイズ分布は分子量でスケールされる特徴的な分布を示す。また、本手法により、注目した環状鎖が他のどの鎖を貫通しているのか(能動貫通)、もしくは、他のどの鎖により貫通されているのか(受動貫通)といった量を統計的に解析することも可能である。今後、これらの知見を系の実際の物性に結びつけるには、ここで定量化した貫通事象の時間発展を調べる必要があると考えている。

3. 今後の展開

現在までのところ、動的絡み合い解析法により、短～中時間領域における絡み合い構造の解析を行うことに成功している。今後、絡み合いの緩和時間に相当するより長時間領域の解析を行うことにより、さらなる知見が得られると期待される。特に、リンク無し環状鎖系においては、長時間領域でのダイナミクスについて十分な理解がなく、このような基礎的問題の解決において本質的に重要な貢献ができる可能性がある。また、提案した解析手法を、様々な系(例えば、剪断流下の環状鎖濃厚系や、線形鎖と環状鎖の混合系など)に適用することも興味深い。

パーシステントホモロジーの活用による環状鎖間の関係を定量化する試みにおいては、本研究において定義したホモロジカル貫入構造の時間発展の様子を調べる。これにより、鎖同士がリンクできないというトポロジカルな拘束が系のダイナミクスやレオロジーにどのような影

響を及ぼすかを明らかにする。また、ここで提案した方法論自体について、パーシステントホモロジーの新たな活用法の一つとして、その可能性を追求していきたい。

4. 自己評価

本研究では、「大規模計算機シミュレーションから得られる大量の高次元データから、トポロジカルな拘束を表現する情報を抽出し、理論研究との融合により、新規な絡み合い状態の高分子系の構造的特徴を明らかにする」という研究目的に対し、二つの新しい解析法を提案した。研究の進め方については紆余曲折があり、敢えて当初の予定とは異なる方面へ進めることもあったものの、「従来の枠を超えて『絡み合い』の概念を深める」という大きな目的に対して、重要な成果を得ることが出来たと考えている。研究実施体制については、研究が進展した最終年度に研究補助員を二名雇用し、これにより効率的なデータ解析を行うことができた。

高分子系の「絡み合い」という実用的に重要な(しかしまだまだ未解明な部分の多い)問題に対して、新たな解析法を提案出来たことの意義は大きいと考えている。絡み合い構造の制御という考えは、今後の新規物性の開拓において大きな可能性を秘めており、そこにおいて、本研究成果が本質的に重要な役割を果たすことを期待する。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

- | |
|--|
| 1. S. Put, T. Sakaue, C. Vanderzande, Active dynamics and spatially coherent motion in chromosomes subject to enzymatic force dipoles, Phys. Rev. E 99, 032421 (2019). |
| 2. T. Sakaue, Statistical physics of ring polymers based on topological volume concept, Reactive and Functional Polymers 134, 150–155 (2019). |
| 3. T. Sakaue, Compressing a confined DNA: from nano-channel to nano-cavity, J. Phys.: Cond. Matt. 30, 244004 (2018). |
| 4. T. Sakaue, Topological free volume and quasi-glassy dynamics in the melt of ring polymers, Soft Matter 14, 7507–7515 (2018). |
| 5. Y. Hayase, T. Sakaue, H. Nakanishi, Compressive response and helix formation of a semiflexible polymer confined in a nanochannel, Phys. Rev. E 95, 042143:1–6 (2017). |

(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)

1. T. Sakaue, “On topological glass in melt of non-concatenated ring polymers” Knots and Polymers: Aspects of topological entanglement in DNA, proteins and graph-shaped polymers, 2017/8/10, Tokyo, Japan.
2. T. Sakaue “Some Topics on Dynamics of Chromosomal Loci” The Arctic Biophysics Meeting on Epigenetics and Chromosome Dynamics, 2019/1/21, Kiruna, Sweden.
3. T. Sakaue “Statistical Physics of Topologically Constrained Polymers”

Polymer meets Topology, 2019/1/30, Tokyo, Japan.

4. 坂上貴洋 「環状高分子鎖濃厚系におけるトポロジーの効果」
高分子学会講演会（高分子開発における MI・AI・計算科学からのアプローチ）東京
2019年2月15日
5. 坂上貴洋 “Structure and dynamics of chromatin: perspective from polymer physics”
第57回日本生物物理学会年会 宮崎 2019年9月24日

研究報告書

「物質・材料の微細構造計測におけるインフォマティクス技術の開拓」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 志賀 元紀

1. 研究のねらい

物質・材料科学分野において、第一原理に基づく数値計算の信頼性・精度が向上するにともなう、様々な材料に関する計算・計測結果が大規模に蓄積されるようになってきた。一方、SPring-8などの大型計測施設や電子顕微鏡などの各研究機関の計測機器の高性能化・自動化が目覚ましく、以前よりも精密な微細構造の計測が可能となってきた。しかしながら、この発展の結果として1回の計測におけるデータサイズが肥大化し続けており、データ解析のコストが新規に合成された物質・材料の評価プロセスにおけるボトルネックとなる。本研究課題では、微細構造計測データの解析コストの大幅な削減を目指し、自動的かつ高精度に計測データを解析するための統計的機械学習法の開発を推進する。

数多くある物質・材料構造計測データの中でも、特に、走査透過型電子顕微鏡(STEM: Scanning Transmission Electron Microscopy)におけるエネルギー損失スペクトル計測(EELS: Electron Energy Loss Spectroscopy)のイメージング計測データ(STEM-EELS スペクトルイメージデータ)の解析法を開発する。この計測技術は、試料平面の各位置に電子線を照射して各位置でのスペクトルを計測するものであり、平面上の計測点が横方向500点×縦方向500点さらに分光スペクトルのチャンネル数2,000という標準的な計測データサイズでもデータ点数は約5億と非常に大きい。この計測スペクトルのピーク位置やピーク形状は物質の電子状態と対応しており、スペクトルイメージは計測物質構造の同定に用いられるものの、EELS スペクトルはピーク形状が単純な関数でモデル化できなかったり、複雑な形状のバックグラウンド成分を含んだりするため、単純なピーク関数によるフィッティングを用いる解析が困難である。そこで、ピーク形状を仮定しないモデルに基づく機械学習法を開発し、解析精度の向上を図る。本研究課題で対象とする物質・材料には、規則的な構造である結晶だけでなく、不規則な構造であり解析が困難な非晶質も含めており、新規構造や構造秩序の同定を目指す。また、さきがけ領域内外の関連研究者と物質・材料科学の実問題解決のための共同研究に取り組み、インフォマティクスを基盤とする新しいデータ解析技術を開拓する。

2. 研究成果

(1) 概要

本研究で得られた主な成果を、研究テーマ A: スペクトルイメージ解析のための統計的機械学習法の開発、研究テーマ B: 非晶質の構造解析、研究テーマ C: その他の物質・材料解析のための機械学習法の開発 に分けて述べる。テーマ A では、スペクトルイメージングを入力として、計測領域中に含まれる化学成分の空間分布と成分スペクトルを出力する統計的機械学習法を開発した。テーマ B では、非晶質構造の構造モデルを用いた仮想計測データを用いて中距離秩序の評価を行なった。テーマ C では、さきがけ領域内の共同研究成果であり、理

論計算の効率化のための機械学習法を構築した。以下に、テーマごとの研究成果を詳細に述べる。

(2) 詳細

研究テーマ A「スペクトルイメージ解析のための統計的機械学習法の開発」

網羅計測されたスペクトルイメージングを入力として、代表的な成分のスペクトル波形と成分の空間分布を自動的に同定する技術を構築した。STEM-EELS のスペクトルイメージング解析において同定される成分は電子状態のことであり、この解析結果によって元素の空間配置(2次元配置)や配位数に関する情報を自動解析できる。開発手法の特徴は、空間的な不自然な重なりを除去できること、また、スパース罰則項により成分数を学習できることであり、いくつかの実データ解析の結果とともに論文[1]において発表した。発表と同時に、アカデミックや企業の研究者が方法を容易に使えるように Github にて開発コードを公開した [https://github.com/MotokiShiga/stem-nmf]。ところで、上述の手法はガウス雑音モデルを仮定しており、低カウントのスペクトルの雑音モデルに適していない。この問題点に着目してポアソン分布を雑音モデルとする手法を開発し、論文[4]において発表を行った。この手法は、低カウントのスペクトルから高精度な解析が可能であるため、TEM/SEM-EDS 等の計測時間を大幅に短縮できることを示せた。開発法を Python のパッケージ malspy (Machine Learning for Spectral Imaging)にまとめて一般公開しており [https://github.com/MotokiShiga/malspy]。開発パッケージは、ToF-SIM や TEM-EDS などの異なる種類のスペクトルイメージングに対しても高精度で解析できることを確認しており、その汎用性が示された[1,3]。今後もこのパッケージを更新し、新しい手法の追加や機能の強化を行う予定である。以上の成果から、本研究テーマに関して、概ね当初の計画通りの目標を達成できたと考えている。

研究テーマ B「非晶質の構造解析」

研究計画当初は、非晶質構造の解析のために、長時間電子線照射した際に現れる EELS スペクトルの高エネルギー帯域の振動成分 EXELFS (Extended Electron Energy Loss Fine Structure) の解析に取り組む予定であった。しかしながら、長時間計測によって試料ダメージが発生し、解析できる良質な計測データを得ることが困難であった。こうしたことから、実計測を今回あきらめて、理論計算(シミュレーション)を用いて生成した構造モデルの解析に取り組んだ。

大規模な構造モデルの作成のために、まず、古典分子動力学法によって、液体状態の温度からガラスになる温度まで急冷するシミュレーションを行い、その後、さきがけ同領域の共同研究者の小原真司博士(物質・材料研究機構)から提供された X 線回折および中性子線回折データを用いる逆モンテカルロ法によって構造モデルを精錬し、実験データを再現するモデルを構築した。構築した構造モデルに対して、電子回折イメージングの仮想計測実験を行った。この計測を選んだ理由は、実計測に必要な電子線照射時間が EXELFS よりも短くなり、また、逆空間の構造を高解像度な回折像で得られるため、原子配置の詳細な情報を得られるからである。ところで、非晶質の重要な解析の1つは、一見ランダムに見える原子配置や化学結合から、秩序を同定することである。仮想計測データから特定の回折パターンを取り出す解析を

行い、化学結合リングの強い相関構造(中距離秩序)を同定する手順を考案した。様々な合成条件の非晶質構造を解析・比較した研究成果を学会発表[5]および論文[5]において発表した。仮想計測データの解析によって新しい知見を得られたので、実計測から中距離秩序を評価することが今後の課題である。課題達成を目指して、今後も実験研究者との共同研究を継続してゆきたい。

研究テーマ C「その他の物質・材料解析のための機械学習法の開発」

スペクトル解析において、テーマ A と異なる手法も構築した。例えば、スペクトルからガウス関数あるいはローレンツ関数で近似されるピークを高速に列挙する手法を開発し論文[Sudou, Shiga, et al., *Journal of the Korean Astronomical Society*, 50 (6), 157-165, 2017]にて発表を行った。この手法をさらに発展させ、3D ラマンスペクトルにおいて混成された計測値を解きほぐし、正しいピーク位置を推定する機械学習法を開発し、論文[2]にて発表した。

物質・材料分野の理論計算を効率化するための機械学習法の開発にも取り組んだ。例えば、さきがけ同領域の烏山昌幸准教授(名古屋工業大学)や塚田祐貴准教授(名古屋大学)とともに、実験データと合致する構造パラメータを効率的に探索する機械学習法を開発して論文[6]にて発表した。また、さきがけ同領域の世古淳人准教授(京都大学)や烏山昌幸准教授らとプロトン伝導経路の効率探索に関する共同研究に取り組んだ [Kanamori, et al., *Physical Review B*, 97, 125124, 2018]。

3. 今後の展開

今後に取り組むべき課題は、開発した機械学習法の流用、あるいは、異なる微細構造計測データ解析のために新たな機械学習法を開発し、非晶質の原子配置や構造パラメータを同定することである。実計測のみでの検証が困難な可能性が高いので、本研究期間中に構築した構造モデルを用いた仮想実験によって開発法の性能を十分に検証することが必要である。実験研究者や理論研究者との連携により、解析例を増やして、構造特徴と物性・材料特性の関係に関する知見を蓄積することで、材料設計につなげるアプローチを模索してゆきたい。

4. 自己評価

研究目的の達成状況:

近年、機械学習法によって微細構造計測データを解析する試みは行われていたものの、それほど広く浸透している状況ではなかった。しかしながら、昨今では、検出器や計測技術の進展が目覚ましく、大規模なデータを網羅的に記録してソフトウェアで処理する流れに移行しつつあり、機械学習の重要性が広く認知されるようになってきた。こうした状況において、本研究課題の目的である、物質材料計測において広く用いられるスペクトルイメージングの統計的機械学習法の開発・整備を達成できたことは非常に重要なものと考えている。当初の計画では、開発法の対象データを STEM-EELS と設定していたものの、開発法は他のスペクトル解析にも有効でありその汎用的を示すことができた。また、実計測時間を短縮するため、低カウントのスペクトルを対象とした手法の有効性も示すことができた。したがって、当初の第一目標を達成できたと考えている。

一方で、本研究で取り組む予定であったスペクトルの周波数解析の課題、既存の解析法で困難な物質構造を同定する課題に関しては、今のところ十分な研究成果を得られていない。達成できなかった原因の一つには、物質・材料科学分野の計算・実験の知識・技術が不足していた点が考えられ、個人研究の限界を痛感した。この不足点を補強するために、さきがけ研究者とのネットワークを構築し、具体的な共同研究を開始している。さきがけ研究を通じて構築したネットワーク、領域会議・自主勉強会で得た知見をさらに深めて、目標達成のための取り組みを継続したい。

研究の進め方(研究実施体制及び研究費執行状況):

研究を進めるにあたって物質・材料科学との連携が必要不可欠であったので、さきがけ領域内で密に議論できる共同研究者のネットワークを新たに構築し、必要な知識や技術を交換した。また、機械学習法の開発や大規模な分子動力学計算を行う必要があったので、CPU コア数・メモリが多い計算サーバーや GPU 搭載計算サーバーを本研究費によって購入して計算環境を構築した。

研究成果の科学技術及び社会・経済への波及効果:

微細構造計測データの定量的かつ高精度な自動解析法は材料探索・開発のスピードを加速させるために必要不可欠である。本研究課題では、スペクトルイメージング計測データを中心として、この技術の発展に貢献した。また、本研究成果をアカデミックや企業の方が自由に使えるように一部の開発コードを公開し、手法の普及に努めた。その結果として、他の研究者による研究論文や企業での材料評価等において、公開コードの利用を確認でき、こうした試みの重要性を実感できた。ところで、微細構造計測データにはスペクトル以外にも数多くの種類のものが存在する。その中には、複雑かつ膨大なデータサイズのため解析法が十分に整備されていないものがある。機械学習を中心とするインフォマティクス技術を用いて、こうした計測データの解析法を新たに構築し、幅広く使えるようにソフトウェアを整備することが重要な課題である。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. [Motoki Shiga](#), Kazuyoshi Tatsumi, Shunsuke Muto, Koji Tsuda, Yuta Yamamoto, Toshiyuki Mori, Takayoshi Tanji, “Sparse Modeling of EELS and EDX Spectral Imaging Data by Nonnegative Matrix Factorization”, *Ultramicroscopy*, 170, 43–59, 2016.
2. Hongxin Wang, Han Zhang, Bo Da, [Motoki Shiga](#), Hideaki Kitazawa, Daisuke Fujita, “Informatics-Aided Raman Microscopy for Nanometric 3D Stress Characterization”, *Journal of Physical Chemistry C*, 122 (13), 7187–7193, 2018.
3. 石倉航, 高橋一真, 山嵜崇之, 青木弾, 福島和彦, 志賀元紀, 青柳里果, “多変量解析を利用した TOF-SIMS イメージデータフュージョンとスパースモデリングおよび機械学習による TOF-SIMS スペクトル解析”, *Journal of Surface Analysis*, 25 (2), 103–114, 2018.
4. [Motoki Shiga](#), Shunsuke Muto, “Non-negative matrix factorization and its extensions for

spectral image data analysis”, *e-Journal of Surface Science and Nanotechnology*, 17, 148–154, 2019.

5. Yohei Onodera, Shinji Kohara, Shuta Tahara, Atsunobu Masuno, Hiroyuki Inoue, Motoki Shiga, Akihiko Hirata, Koichi Tsuchiya, Yasuaki Hiraoka, Ippei Obayashi, Koji Ohara, Akitoshi Mizuno, Osami Sakata, “Understanding diffraction patterns of glassy, liquid and amorphous materials via persistent homology analyses”, *Journal of the Ceramic Society of Japan*, Vol. 127, No.12, 853–863, 2019.

6. Yuhki Tsukada, Shion Takeno, Masayuki Karasuyama, Hitoshi Fukuoka, Motoki Shiga, Toshiyuki Koyama, “Estimation of material parameters based on precipitate shape: efficient identification of low-error region with Gaussian process modeling”, *Scientific Reports*, 9, 15794, 2019.

(2)特許出願

研究期間累積件数:0件

(2)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)

著作物

1. Motoki Shiga, Shunsuke Muto, “High Spatial Resolution Hyperspectral Imaging with Machine-Learning Techniques”, *Nanoinformatics* (Editor Prof. Isao Tanaka), Ch. 9 (p.179–203), Springer, 2018.

主な学会発表

1. 志賀元紀, “スペクトラムイメージ解析における機械学習”, 情報統合型物質・材料開発イニシアティブ チュートリアルセミナー 第6回「計測インフォマティクス」, 科学技術振興機構 東京本部別館, 東京都千代田区, 2017年11月1日.(チュートリアル講演)

2. 志賀元紀, “スペクトラムイメージ解析のための統計的機械学習法”, 第31回日本放射光学会年会・放射光科学合同シンポジウム, 企画講演『情報科学を活用した放射光科学の展開』, つくば国際会議場, 茨城県つくば市, 2018年1月8日.(招待講演)

3. Motoki Shiga, “Statistical Machine Learning for Spectrum Image Data Analysis”, The 19th KIM-JIM Symposium –Recent Advances in Artificial Intelligence and Simulations in Materials Science and Engineering–, Daejeon Convention Center, Daejeon, Korea, October 25, 2018. (Invited Talk)

4. Motoki Shiga, Motoki Fukaya, Shunsuke Muto, “Statistical Hypothesis Testing of the Number of Chemical Components in Spectrum Image Data”, The 19th International Microscopy Congress (IMC19), International Convention Center, Sydney, Australia, September 9–14, 2018. (Mini-oral and digital poster)

5. 志賀元紀, 平田秋彦, 小原真司, 小野寺陽平, “オングストロームビーム電子回折イメージングを用いた非晶質構造の網羅的解析”, NIMS 先端計測シンポジウム 2019, 物質・材料研究機構, 東京, 2019年3月7日.(特別講演)

研究報告書

「強相関電子系に対する機械学習を用いた高精度量子多体計算の新たな数理アプローチの開拓」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 辻直人

1. 研究のねらい

本研究では、高精度量子多体計算に機械学習の手法を取り入れて計算効率を最適化する新たな数理手法を開発し、強相関電子系の物性予測・物質設計の汎用的フレームワークを確立する。それによりこれまで解析することができなかった物質群に対する高精度計算の道を拓き、また計算効率を最適化する高次元パラメータの学習モデルによる表現を得ることで強相関電子系の本質に迫るような新知見を得ることを目指す。

高精度量子多体計算の中でも動的平均場理論は電子相関効果を取り入れられる有力な手法であり、有限サイズクラスターの問題に帰着させることができる。これは高精度な量子モンテカルロ法によって解かれるが、負符号問題という本質的な困難を抱えており、現状では軌道数やクラスターサイズに厳しい制限がある。負符号の出やすさ、つまり計算効率はクラスターの1粒子基底の選び方に鋭敏に依存し、問題ごとに最適な基底が存在する。これまでは人間が知恵を絞って計算効率を向上させる基底を探す努力がなされてきた。最適な基底は一般に平面波基底とも局在基底とも異なり、人間が直観的に理解しにくいものとなっている。計算効率を最大化する基底を探す問題は一種の最適化問題とみなすことができる。

そこで、機械学習の手法を利用して、低エネルギー有効モデルのパラメータが与えられたときに計算効率が最もよくなる基底を推定する学習モデルを構築する。量子モンテカルロ法の計算効率の指標として、符号の期待値を用いる。この量は、物理量の期待値と比較して高速で計算することができるため、様々な基底に対して符号の期待値を計算し、大量の訓練データを蓄積させることができる。そこから符号の期待値の分布を観測データとして得ることができる。この分布になるべく近くなるような確率分布を生成する学習モデルを設計する。

構築された学習モデルの基底状態を求めることで、量子モンテカルロ法の計算効率を最大化する基底が得られる。実際、計算効率がどの程度改善されたかは量子モンテカルロ法自身によって自己判定することができる。これを様々な学習モデルの候補について試すことで、最も効率よく最適化基底を選び出せる学習モデルを導く。それを動的平均場理論に組み込むことで、軌道数やクラスターサイズの制限を緩和した汎用的な高精度量子多体計算手法を確立することをねらう。

2. 研究成果

(1) 概要

強相関電子系のシミュレーションに使われる連続時間量子モンテカルロ法において、計算効率を左右する負符号問題の度合いは展開に用いる1粒子基底の取り方に大きく依存する。本研究では、負符号を抑えて計算効率を最大化するような基底を推定するモデルを構築し

た。具体的には、推定モデルとして対応するフェルミハバードモデルと(ハードコア)ボースハバードモデルの自由エネルギー差をとった ΔF 指標というものを提案した。量子モンテカルロ法の負符号問題は、電子がフェルミオンであることを反映してフェルミオンの反交換関係に起因する。 ΔF 指標はちょうどフェルミオンの反交換関係をボソンの交換関係に置き直した際の影響の度合いを定量化したものになっている。また、 ΔF 指標だけでは不十分な場合、フェルミハバードモデルとボースハバードモデルの間の Kullback-Leibler (KL) 情報量を相補的に用いることで、最適基底の推定精度が上がる。 ΔF 指標と KL 情報量を様々なサイト数・軌道数のクラスターモデルに適用し、選んだ基底が最適かどうかをすばやく判定することができることを示した。また、ベイズ最適化を用いて量子モンテカルロ法の計算効率を最大化するような基底の探索を行った。

量子モンテカルロ法には、負符号問題以外にも、軌道の数が多くなると確率の重みを計算するコストが指数関数的に増大する困難がある。強結合展開型の量子モンテカルロ法において、モンテカルロが生成する大量の情報を機械学習して、効率よくサンプリングを行う手法を開発した。モンテカルロの確率分布を学習するモデルとして、虚時間軸上を相互作用する古典粒子が配置された学習モデルを考えた。多軌道アンダーソン模型に対してテストした結果、フント結合が大きくない場合に精度よくモンテカルロの確率分布を推定できることがわかった。得られた学習モデルを使ってモンテカルロのアップデートを行うことで、計算コストを大幅に削減させることができた。また、学習モデルによるアップデートを繰り返した後、真の確率分布を計算してアップデートを行うことで、バイアスをかけることなく数値的に厳密にモンテカルロの自己相関時間を短くすることができる。これによって、動的平均場理論のソルバーとして広く使われる強結合展開型の量子モンテカルロ法の計算効率が格段に向上した。

(2) 詳細

研究テーマ A「量子モンテカルロ法の最適基底選択問題」

固体電子系において、電子相関効果に起因して超伝導や磁性など様々な物理的性質をもった量子相が現れる。しかし、電子相関効果を正しく考慮したマテリアルシミュレーションは容易ではなく(例えば密度汎関数理論はモット絶縁体を金属と予測してしまう)、平衡系・非平衡系を含めて様々なアプローチがとられている(発表論文[1-5])。その中でも、動的平均場理論は局所的な電子相関効果を厳密に取り込む手法として第一原理計算と組み合わせることで強相関物質に対して広く用いられ、成功を収めている。

動的平均場理論は、電子系の格子モデルを 1 サイトの不純物ないし有限サイズのクラスターモデルに置き換えて自己無撞着に解く手法である。クラスターモデルを解くための汎用的かつ高精度な量子多体計算手法として、連続時間量子モンテカルロ法がある。量子モンテカルロ法は一般に、確率の重みが負になってしまう負符号問題という困難を抱えていて、動的平均場理論を使ったマテリアルシミュレーションのボトルネックになっている。負符号問題の度合いは量子モンテカルロ法の展開に使う 1 粒子基底の選び方に大きく依存することが知られている。3 サイトクラスターモデルの場合に、符号の期待値が基底のパラメーター空間上でどのように分布しているかを図 1 左に示した。

本研究では、量子モンテカルロ法の計算効率を最大化する基底を見つけるための推定モデ

ルを構築した(学会発表[2-4])。負符号問題は、電子がもつフェルミ統計性に起因する。このフェルミ統計性を、負符号が発生しないボース統計性に置き換えたとき、その「差」が負符号の影響の大きさを表していると考えられる。そこで、フェルミオンのクラスターモデルと対応するボソンのクラスターモデルの間の自由エネルギー差(ΔF 指標)を推定モデルに用いた。3 サイトクラスターに対して、平面波基底・局在基底・中間基底をとったときの ΔF 指標の振る舞いを図1右に示した。この結果は、実際の量子モンテカルロ法の計算効率を決める符号の期待値の振る舞いと定量的に合致する。特に中間基底は ΔF 指標の値が厳密に1になっており、この基底を選ぶことで負符号を厳密にキャンセルすることができる。このように、推定モデルを使うことで選んだ基底が最適かどうかをすばやく判定することができるようになった。 ΔF 指標と相補的に、フェルミオンのクラスターモデルとボソンのクラスターモデルの間の Kullback-Leibler (KL) 情報量も用いた。また、ベイズ最適化によって、計算効率を最大化する基底の探索を行った。

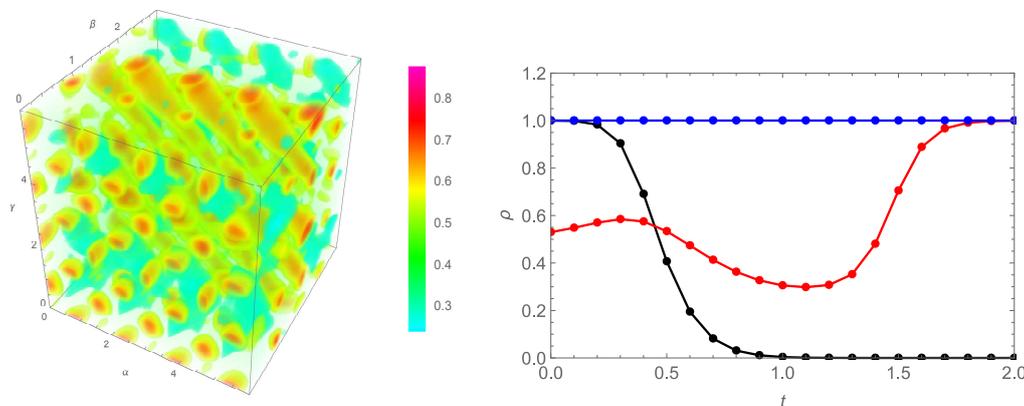


図1. (左) 3 サイトクラスターに対して、量子モンテカルロ法の計算効率を表す符号の期待値を基底のパラメータ空間でプロットしたもの。(右) 3 サイトクラスターにおいて、平面波基底(赤)・局在基底(黒)・中間基底(青)に対する推定モデルによる計算効率の予測値。横軸は電子のホッピングパラメーター。その他のパラメーターはクーロン相互作用 $U=5$ 、逆温度 $\beta=10$ 。

研究テーマ B「機械学習を用いた強結合展開型の連続時間量子モンテカルロ法」

連続時間量子モンテカルロ法の中でも多軌道の強相関系に威力を発揮するのが、強結合展開型のアルゴリズムである。この方法は、局所的な相互作用項を厳密に取り入れる代わりに、非局所的な運動項を摂動展開によって扱う。1 サイトクラスターの場合には負符号問題が抑制される特徴があるが、多軌道系において軌道の数が増えると、確率の重みを計算するコストが指数関数的に増大してしまうという困難がある。

そこで、機械学習の手法を用いて、強結合展開型量子モンテカルロ法において確率の重みを計算するコストを緩和するアルゴリズムを開発した(学会発表[1,3-5])。強結合展開型では、確率の重みは行列式の部分と局所的なトレースの部分に分離する。行列式の部分は、虚時間軸上で粒子が相互作用する古典的なモデルを学習モデルとして用いた。局所トレース部分は、相互作用の形が密度・密度型の場合にセグメント表示というものができるとを参考にし、各相互作用点において状態が変化していくとして状態ごとにエネルギーが与えられた古

典モデルを学習モデルとして採用した。量子モンテカルロ法を走らせながら訓練データを蓄積し、得られたデータから線形回帰によって学習モデルに含まれるパラメータを最適化した。2軌道のアンダーソンモデルに対して機械学習を行った結果を図2に示した。フント結合が小さいと相互作用は密度・密度型に近づくので、精度良く学習できていることがわかる。フント結合が大きくなると、一般に学習精度は落ちる。

機械学習の結果を用いて、量子モンテカルロ法の確率重みを計算することに応用した。まず、学習モデルを使ってモンテカルロのアップデートを繰り返す。学習モデルは古典モデルになっていて計算コストが非常に低い。そのため高速でアップデートを繰り返すことができる。一定回数のアップデートを行った後、真の確率重みを計算し、重みの比率からアップデートを採用するかどうかを判定する。この方法の利点は、真の確率重みを使うため近似やバイアスが入ることなく数値的に厳密なアルゴリズムになっていること、学習モデルによるアップデートを繰り返すことでモンテカルロサンプル間の自己相関時間を劇的に減らすことができることが挙げられる。自己相関時間が減少するとサンプルの独立性が高まり、少ない数のサンプルでも統計誤差を抑えることができるようになる。自己相関時間が十分長いシステムの場合、計算コストを $O(N^3+N*2^{6N_{orb}})$ から $O(N^2+N)$ に落とすことができた (N は強結合展開の次数、 N_{orb} は軌道数)。以上の結果によって、強相関係の材料シミュレーションに用いる強結合展開型の量子モンテカルロ法の計算コストが大幅に緩和された。

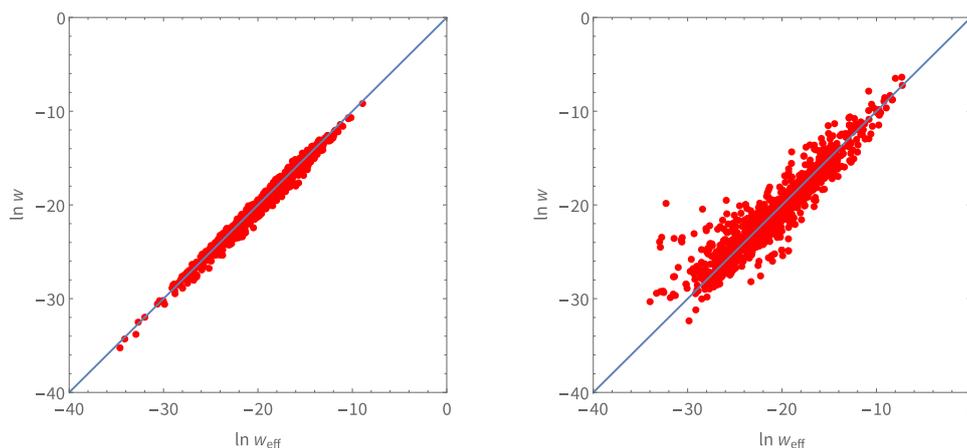


図2. 2軌道アンダーソンモデルに対して、強結合展開型量子モンテカルロ法の確率の重みの分布を機械学習した結果。縦軸が真の確率重みの値、横軸が学習モデルによる予測値である。(左) フント結合 $J=0$ の場合。 R^2 値は0.977。(右) $J=1$ の場合。 R^2 値は0.841。その他のパラメータは、クーロン相互作用 $U=4$ 、化学ポテンシャル $\mu=2$ 、逆温度 $\beta=10$ 、バンド幅 $W=4$ 。

3. 今後の展開

本研究で開発した機械学習を用いた強結合展開型量子モンテカルロ法に残る課題として、学習モデルを訓練する計算コストをいかに削減するかが課題である。一度学習してしまえば計算コストは軽減されるが、学習するための訓練データを蓄積するのに結局量子モンテカルロ法をまわさなければならない。訓練を行う際には確率重みだけを計算すればよく、計算コストの高い物理量の計算は必要ないが、それでも軌道の数が多くなると訓練プロセスで多大な

時間がかかってしまう。できれば、システムごとに毎回学習しなくても最初からある程度学習パラメーターが予測できていることが望ましい。例えば確率重みの局所トレース部分は、密度・密度型の相互作用に近似すれば学習パラメーターの値がわかる。この方向で今後研究を展開していくことが考えられる。

また、フント結合の値が大きいときに学習精度が落ちてしまうという問題も残っている。これは、相互作用の形が密度・密度型から離れてしまうために、密度・密度型の相互作用をよく表すように作られた学習モデルでは確率重みをうまく表現できなくなるのが原因と考えられる。多くの物質群ではフント結合の値はそれほど大きくないので問題にならない場合が多いが、大きなフント結合をもつ強相関物質(Sr_2RuO_4 など)もあるのでそれらにも適用できるようにしたい。密度・密度型ではない、スピンフリップ項やペアホッピング項を精度よく表現できるように学習モデルを改善していくことが今後の方向性として考えられる。

4. 自己評価

本研究の目的であった、量子モンテカルロ法の計算効率を最大化する展開基底を推定する課題については、精度のよい推定モデルを構築することができた。これによって、選択した展開基底が最適のものであるかどうかを素早く判定することが可能になった。さらに、なぜ平面波基底や局在基底などの人間にとって理解しやすい基底ではなく中間的な基底が最適であるかについて、理由を理解することができた。この部分に関しては本研究の目的を達成しており、評価できる。さきがけ領域会議でいただいたコメントやアドバイスの中で、量子モンテカルロ法を走らせている最中に得られる大量のデータを機械学習に利用できないかという提案があり、当初の想定にはなかった方向に研究を発展させることができた。その結果、機械学習を用いた強結合展開型の量子モンテカルロ法の開発が進み、確率重みを計算するコストを大幅に削減する方法が確立した。この成果は、今後動的平均場理論と組み合わせることで強相関電子系のマテリアルシミュレーションに大いに役立つことが期待される。本研究結果はすでに各種学会において発表済みであるが、論文執筆が当初の予定より遅れており、この部分はマイナスの評価である。強相関電子系や超伝導体のマテリアルシミュレーションを使った非平衡ダイナミクスの研究を並行して進めており、Annual Review of Condensed Matter Physics 誌から招待されて総説論文が出版されたり、Nature Communications 誌に論文が受理された等の成果があった。研究実施体制については、基本的に個人研究の形で研究を進めてきたが、量子モンテカルロ法の専門家であるフリップール大学の Philipp Werner 教授と国際共同研究を行うことができた。研究費の執行状況はおおむね計画通り進み、クラスター計算機を導入して量子モンテカルロ法の研究に用いることができた。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. L. Schwarz, B. Fauseweh, N. Tsuji, N. Cheng, N. Bittner, H. Krull, M. Berciu, G. S. Uhrig, A. P. Schnyder, S. Kaiser, D. Manske, "Classification and characterization of nonequilibrium Higgs modes in unconventional superconductors", Nature Communications, 11, 287 (2020).

2. R. Shimano, N. Tsuji, “Higgs mode in superconductors”, Annu. Rev. Condens. Matter Phys. (2020), in press. (Online publication: doi: 10.1146/annurev-conmatphys-031119-050813)
3. S. Sayyad, N. Tsuji, A. Vaezi, M. Capone, M. Eckstein, H. Aoki, “Momentum-dependent relaxation dynamics of the doped repulsive Hubbard model”, Phys. Rev. B 99, 165132 (2019).
4. R. Iwazaki, N. Tsuji, S. Hoshino, “Nature of the superconducting fluctuations in photoexcited systems”, Phys. Rev. B 100, 104521 (2019).
5. K. Katsumi, N. Tsuji, Y. I. Hamada, R. Matsunaga, J. Schneeloch, R. D. Zhong, G. D. Gu, H. Aoki, Y. Gallais, R. Shimano, “Higgs Mode in the d -Wave Superconductor $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$ Driven by an Intense Terahertz Pulse”, Phys. Rev. Lett. 120, 117001 (2018).

(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)

1. 辻直人、Philipp Werner、「機械学習を用いた強結合展開型の連続時間量子モンテカルロ法」、日本物理学会秋季大会、岐阜大学、2019年9月。
2. 辻直人、「動的平均場理論に対する量子モンテカルロ法の最適基底選択問題」、日本物理学会秋季大会、同志社大学、2018年9月。
3. Naoto Tsuji, “Self-learning continuous-time quantum Monte Carlo and optimization of the single-particle basis”, PRESTO International Symposium on Materials Informatics, Tokyo, February 2019.
4. 辻直人、「非平衡動的平均場理論の最近の進展」、招待講演、第2回動的平均場近似計算に関する情報交流会、埼玉大学、2018年6月。
5. Naoto Tsuji, “Machine learning quantum Monte Carlo simulations for strongly correlated electron materials”, Materials Research Meeting 2019 (MRM2019), Yokohama, December 2019.

研究報告書

「円錐交差データベースに基づく蛍光分子自動設計法の開発」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 原 潤 祐

1. 研究のねらい

一般的な分子設計のサイクルは三段階からなる。ある対象分子に対し、実験・計算から分子の機能や化学反応の機構を調べる評価の段階、得られた情報を収集・整理する蓄積の段階、そして、得られた過去の知見から新規分子を提案する予想の段階である。近年、計算機性能の向上と計算手法の発展に伴い、分子の安定構造や励起エネルギーを高速に計算することが可能となった。また、量子化学計算に基づく化学反応の解析では、GRRM(Global Reaction Route Mapping)のような反応経路自動探索プログラムの発展に伴い、分子に対する事前知識によらない化学反応の機構解析が可能となりつつある。しかし、新規分子の予想については、現在も熟練の研究者が行う場合がほとんどである。そこで、本研究では、新規分子の予想の自動化に着目し、量子化学計算に基づいて分子設計を行う方法の開発を目指す。

本研究では、有機 EL や蛍光プローブなどへの応用が期待される分子の蛍光に着目する。一般の光反応では、光励起した分子は、円錐交差を通じた内部転換過程、シーム交差を通じた項間交差過程、蛍光過程、りん光過程によって最終的には基底状態へと失活する。この時、円錐交差構造とそれに到達するための反応経路を網羅的に調べることで、励起分子が内部転換によって無輻射失活するか否かを議論することができる。しかし、円錐交差領域の情報は実験的にも理論的にも求めることが難しく、交差構造の情報を蓄積することは困難であった。これは、分子の交差領域滞在時間が光反応全体に比べて非常に短いため実験的に観測することが容易ではなく、理論的には円錐交差構造が分子の安定構造とはかけ離れているために推定が難しいことに起因する。これに対して提案者は、円錐交差構造を網羅的・自動的に探索する手法の開発に取り組んできた。現在では、研究室規模の計算機を用いた場合でも30原子程度の分子に対して、交差探索計算を適用出来るようになってきている。

本研究では、多様な分子に対して円錐交差探索を適用し、円錐交差データベースを構築する。また、得られた円錐交差の情報から、個々の光反応機構を解析する。さらに、円錐交差データベース、量子化学計算、情報学手法を組み合わせることで、新規分子構造を提案する手法の開発を目指す。

2. 研究成果

(1) 概要

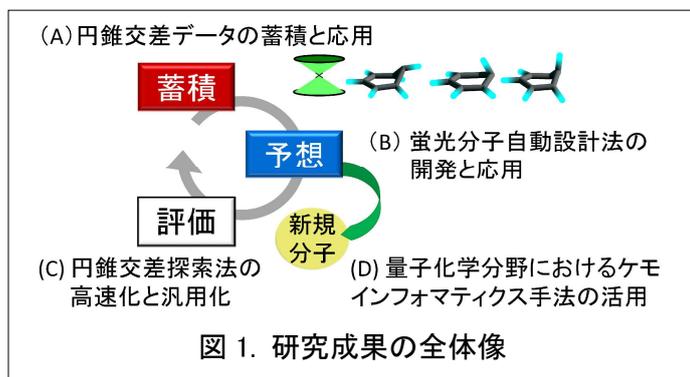
初めに、円錐交差データの蓄積と、データベース構築に取り組んだ。基底状態の分子に対する構造データベースは存在するが、光反応で重要となる円錐交差領域の情報は求めることが難しく、データの蓄積は困難であった。本研究では、提案者が拡張を進めてきた円錐交差探索手法を用い、芳香族化合物を中心に、様々な有機分子の円錐交差構造を探索し、情報(交差構造、エネルギー)を利用可能な形で集約した。データベース構築においては、探索計

算の結果を簡便に登録し、情報を取り出しやすい形で格納する方法の研究も進めた。また、得られた円錐交差の情報を、様々な光反応の機構解析に応用した(図1A)。

蛍光分子自動設計法を開発では、新規分子構造予想の自動化に取り組んだ。新規分子構造の予想は、参照分子リスト生成手法と分子評価の方法を組み合わせる方針で進めた。参照分子リストでは、ケモインフォマティクス分野で用いられる SMILES 記法を用い、与えられた分子骨格に対して置換を行うことで分子リストを得る方法を導入した。これにより、多くの未知分子構造を得ることが出来た。得られた分子構造リストに対して、円錐交差探索法に基づく励起状態計算を用いて蛍光特性の評価を行った。その結果、1000分子を超える参照分子リストの中から、計算に基づいて蛍光を発しやすいと予想される分子を選ぶことが出来た。また、データベースに基づき円錐交差エネルギーを推定し、参照分子リストから有力候補分子を予想する方法についても開発を進めた(図1B)。

円錐交差データの蓄積を加速するために、交差探索手法の汎用化と高速化に取り組んだ。通常の TDDFT 法に基づく交差探索法を導入することで、溶媒効果を考慮した分子特性評価を可能にした。続く研究として、これまで

30原子程度の分子にしか適用できなかった探索手法を70原子以上の大きな分子に対しても適用できるように拡張した(図1C)。また、反応経路探索結果の解析におけるケモインフォマティクス手法の利用を進め、実際のデータ解析で活用した(図1D)。



(2) 詳細

研究テーマ A 「円錐交差データベースの構築と円錐交差データに基づく光反応解析」

円錐交差のデータ蓄積を行うにあたって、GRRM プログラムを用いた円錐交差探索計算の実行と計算結果の収集を出来る限り自動的に行う必要があった。研究初期には、探索計算の入力ファイル作成と計算結果の自動集計を行う枠組みを構築することで、芳香族化合物を中心として様々な分子に対する円錐交差探索計算を適用した。計算結果の集計においては、データ格納の形式を統一することにより、分子構造を入力とする計算結果の検索も可能となった。また、データとして登録された円錐交差構造を整理し、情報学手法を適用するために必要な情報を簡便に取り出せる方法を導入した。これにより、円錐交差構造に基づいた情報学的な解析を行うことが可能になり、円錐交差構造の類似性と交差エネルギーに関する研究を行うことができた。円錐交差計算では、スピンフリップ(SF)-時間依存密度汎関数法(TDDFT)を用いた交差探索を230種の分子に対して適用した。さらに、本研究(テーマ C)で開発した Energy Shift (ES) - TDDFT 法を適用することにより、100種以上の分子の円錐交差構造を最適化した。全体として、当初の目標とした分子数を達成した。

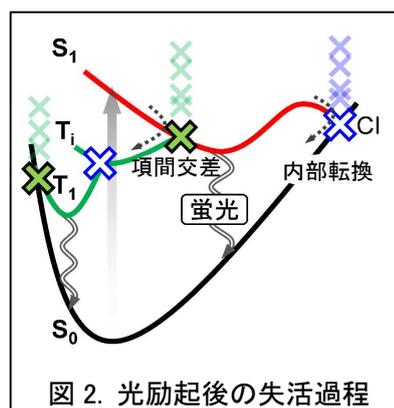
円錐交差データを用いた光反応の機構解析にも取り組んだ。インドール、イソインドール、キノリン、イソキノリンに対して、円錐交差、円錐交差に至る反応経路の障壁、励起状態安定

構造、振動子強度、三重項状態とのシーム交差、スピン軌道相互作用を計算することで、蛍光量子収率、及び、三重項状態への項間交差量子収率が異なる機構を説明した(論文4)。また、分光実験との共同研究によって、桂皮酸誘導体の失活過程、2'-ヒドロキシカルコンの超高速過程に関して機構を解析した。続く研究では、大環状芳香族化合物における蛍光量子収率骨格依存性の機構を円錐交差と励起状態の安定性に基づき解析した。この研究を通じて、円錐交差探索に基づく励起状態解析が実際の蛍光分子設計においても有効であることが示された(論文1)。

研究テーマB「蛍光分子自動設計法の開発と応用」

自動設計法では、参照分子リストの生成を行い、量子化学計算と情報学手法に基づき有力分子を選ぶ方法について研究を進めた。参照分子リストでは、与えられた分子骨格に対して置換を行うことで分子リストを得る方法を導入した。また、研究者が予め用意した分子部分構造を用いて、3次元空間で分子部品を立体的に組み上げることで分子リストを生成する方法についても手法の開発に取り組んだ。これらの方法を組み合わせることで、多くの未知分子構造を得た。

蛍光過程では、図2に示すように、内部転換、項間交差、蛍光の3つの過程が競合する。そのため、これらの過程を量子化学計算に基づき評価することが重要となる。本研究では、得られた分子構造リストに対して、高速な円錐交差探索を用いた絞り込みを行い、TDDFTに基づく交差構造を最適化することで順位付けする方法を適用した。この方法を上述の参照分子リストと組み合わせることで、量子化学計算に基づき未知分子構造の中から有力分子を提案することが可能になった。作成した方法を用いることにより、1000分子を超える未知分子リストの中から、



から、内部転換、項間交差過程、蛍光過程の競合を考慮した計算結果に基づき、蛍光過程が有利となる分子を予測した(学会1)。また、データベースの情報に基づき交差エネルギーを推定する方法についても研究を進め、新規分子リストへと適用した。続く研究として、同さきがけ研究領域に所属する相澤直矢博士と協力し、蛍光分子設計技術の実際の分子開発への応用にも着手した。

研究テーマC「円錐交差データベース構築の加速を目指した探索手法の高速化と汎用化」

円錐交差データの蓄積における探索計算のコストが、さきがけ期間を通じて1つの課題であった。そのため、本研究中にも、より高速で汎用性の高い円錐交差探索手法の開発を並行して進めた。従来、GRRMプログラムでは、ポテンシャル交差構造で極小となるペナルティ関数を用いる交差探索手法が用いられていた。本研究では、ペナルティ関数に基づく交差最適化の収束が一般的に良くないことに着目し、勾配射影法を導入することで交差最適化の収束性を高めることで探索を効率化した(論文5)。更に、同さきがけ研究領域に所属した畑中美穂准教授と共同し、通常のTDDFT法に基づく円錐交差探索法の開発に取り組んだ。畑中美穂准教授が開発したEnergy Shift法を有機分子の交差計算に活用し、通常のTDDFTに基づ

いて円錐交差計算が行える手法を開発した(論文3)。この手法の開発により、研究開始時に問題点として考えていた、実際の蛍光量子収率の評価において溶媒効果を考慮する必要がある点と、全ての失活経路(内部転換・項間交差・蛍光)を同等の計算レベルで評価する必要がある点に対しても取り扱うことが可能となった。本手法は、テーマBの蛍光分子評価の計算で利用した。また、半経験的な方法を取り入れた高速な円錐交差探索法の開発にも取り組んだ。大環状芳香族化合物の課題では、密度汎関数強束縛(DFTB)を用いた反応経路探索を用い、66原子分子の円錐交差を探索した(論文1)。更に、経験的分子軌道法を用いた円錐交差探索法の高速化にも取り組み、70原子以上を含む大きな分子に対しても、交差探索が適用可能であるという結果を得た(学会1)。

研究テーマD「反応経路探索結果の解析におけるケモインフォマティクス手法の活用」

近年の反応経路探索の発展に伴い、未知の分子に対する反応経路計算が可能となってきたが、その一方で、探索計算の中で多くの反応経路が得られるため、どのように得られた結果を解析するかが課題となってきている。例えば、 γ -ketohydroperoxid (KHP)を反応物とする反応経路を単成分-人工力誘起反応(SC-AFIR)法によって探索した場合、KHPからKHP以外の分子構造に到達する反応の遷移状態が401個、KHPのコンフォメーション変化の遷移状態が99

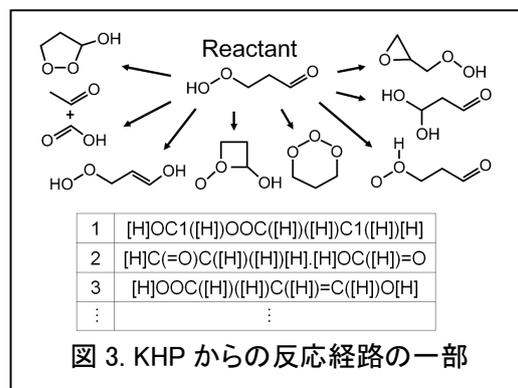


図3. KHPからの反応経路の一部

個得られる。この結果について何種類の反応経路が得られたのかを研究者が解析するのは容易ではない。これに対して提案者は、SMILES記法を用いることでGRRMプログラムによって得られた反応経路を結合パターンの変化によって分類し、解析した(論文2)。また、円錐交差データベースの解析においても、分子リストの可視化手法を導入することでデータ解析を効率化した。

3. 今後の展開:

本研究では、芳香族化合物を中心に多数の分子に対して円錐交差構造を計算した。今後、更に多くの分子に対して交差情報を蓄積する。また、既に計算されたデータには、これまで光反応の機構が理論的に議論されていない分子も含まれるため、交差に基づく光反応の解析にも応用を進める。本研究では、円錐交差構造、基底状態安定構造、励起状態安定構造のデータを登録する形で開発を進めてきたが、実際の光反応では遷移状態やスピン軌道相互作用の情報も重要となる。これらの情報も同時に管理できるように拡張することで、一般の光機能性分子の設計へと応用の幅が広がると考えている。また、計算効率と汎用性が向上した円錐交差探索法は、実際に応用が期待される大きさの光機能性分子に対しても適用できると考えており、今後、更なる応用が期待される。

蛍光分子設計法に関しては、実際に未知の分子を予測したが、その分子が合成可能かと、安定に存在できるかについては現在考慮できていない。また、光励起後に分解してしまう反応経路については交差探索によって暗に含まれるが、基底状態へ失活後に分解するかについては評

価していない。一方で、実際の分子設計においては、これらの点は重要な要素であるため、予測した分子の安定性の問題を解決することで、より手法の有用性を高めることが出来ると考えている。今後、実際の実験研究との共同によって、予想した新規分子を合成して蛍光特性評価を行うことで方法の有用性を検証していく。また、実験研究者と議論においては、本研究の中で開発を進めた計算結果可視化手法が有用であると考えており、一般の反応解析にも利用可能な形に拡張を進める。

4. 自己評価:

本研究では、円錐交差データベースの構築と、円錐交差の情報に基づく蛍光分子自動設計法の開発についてそれぞれ研究を進めた。円錐交差データベースに関しては、当初予定した分子の数を超えて、分子の円錐交差構造を計算し情報を蓄積することが出来た。円錐交差データの蓄積を通じて、自動探索によって得られる計算結果をどのようにまとめ、簡便にとりだせる形として集約するかという観点で試行錯誤し、探索データ蓄積方法の観点でも進展があったと考えている。また、円錐交差探索手法の高速化・汎用化を進めたことで、当初想定していたよりも多様な分子に対する円錐交差構造の最適化を適用することが出来た。円錐交差計算については、当初問題として考えていた溶媒効果の影響に関して、溶媒中の分子に対する機能を評価することが可能となった。本手法は、今後の光化学研究における活用も期待できる。蛍光分子自動設計については、分子構造を生成する手法を導入し、与えられた構造に対して、量子化学計算および機械学習を用いて順位付けを行うことで、有力な分子構造を選ぶことが出来た。このことは、円錐交差構造の情報を未知分子の予測へと応用した点で意義があると考えている。

本研究で用いた円錐交差の自動探索計算は、1つの分子に対して多くの交差構造を網羅探索するため、相応の計算コストが求められる。特に、本課題では、実在しない分子や、全く蛍光を発しない分子に関してもデータを得る必要があったため、一般の量子化学の研究では除外されるよう分子に対する計算を数多く実行する必要があった。これに対して本課題では、計算機システムの導入に加え、計算機センターを利用することで計算を行った。また、これらの計算資源を活用することで、予測された分子の評価も行うことが出来た。提案者自身がプログラム開発を行うことで、実験研究者や情報学研究者と議論しながら柔軟に研究を進められたと考えている。

本研究を通じて蓄積した円錐交差データは、これまで計算することさえ容易ではなかった円錐交差構造を多数の分子に対して探索しており、将来、蛍光分子だけでなく一般の光機能性分子の設計に利用できると期待される。同時に、本研究を通じて導入された効率的・汎用的な円錐交差探索法は、今後の光反応の理論解析をリードする手法になると期待される。蛍光分子設計については、現在、実験研究者との連携を進めており、予測した分子を実際に合成することで交差まで考慮した分子設計の有用性を示し、社会貢献に近づけると考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. K. Ikemoto, T. Tokuhira, A. Uetani, Y. Harabuchi, S. Sato, S. Maeda, H. Isobe, Fluorescence Enhancement of Aromatic Macrocycles by Lowering Excited Singlet State Energies, *J. Org. Chem.*, 85, 1, 150–157 (2020).
2. S. Maeda, Y. Harabuchi, On Benchmarking of Automated Methods for Performing Exhaustive Reaction Path Search, *J. Chem. Theory Comput.*, 15, 4, 2111 (2019).
3. Y. Harabuchi, M. Hatanaka, S. Maeda, Exploring approximate geometries of minimum energy conical intersections by TDDFT calculations, *Chem. Phys. Lett.*, 2, 100007 (2019).
4. Y. Harabuchi, K. Saita, S. Maeda, Exploring radiative and nonradiative decay paths in indole, isoindole, quinoline, and isoquinoline, *Photochem. Photobiol. Sci.*, 17, 315 (2018).
5. Y. Harabuchi, T. Taketsugu, S. Maeda, Combined Gradient Projection / Single Component Artificial Force Induced Reaction (GP/SC-AFIR) Method for an Efficient Search of Minimum Energy Conical Intersection (MECI) Geometries, *Chem. Phys. Lett.*, 674, 141 (2017).

(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)

主要な学会発表:

1. 原 渕 祐、W. Thiel、前田 理、半経験的分子軌道法を用いた S_0/S_1 -円錐交差構造の網羅探索: 量子化学計算に基づく蛍光分子設計に向けて、1E2-33、日本化学会 第100春季年会 (2020)、東京理科大学 野田キャンパス, 東京, 2020/3/22.
2. Y. Harabuchi and S. Maeda, Systematic Exploration of Conical Intersection Geometries Between the Ground and First Excited Electronic States Based on Time Dependent Density Functional Theory: Application to Photoreactions, IC067, APATCC 2019, Sydney, Australia, 2019/10/3. (invited)
3. 原 渕 祐, 前田 理、SMILES 記法による反応経路の分類: γ -ketohydroperoxide の反応経路探索, P18、シンポジウム「化学反応経路探索のニューフロンティア2019」、名古屋大学 野依記念学術交流館, 愛知, 2019/9/16.
4. Y. Harabuchi, Automated search for internal conversion and intersystem crossing pathways: Application to photoreactions, The 77th Okazaki Conference Series: International Symposium on Ultrafast Dynamics in Molecular and Material Sciences, 分子科学研究所、愛知県, 2017/3/6. (invited)
5. Y. Harabuchi, Automated Search for Minimum Energy Conical Intersection and Seam of Crossing Geometries: Application to Photoreactions, CJK-WTCC-III, KAIST, South Korea, 2017/1/11. (invited)

研究報告書

「材料シミュレーションとインフォマティクスを用いたデータ駆動型リチウムイオン導電性セラミックスの探索」

研究タイプ: 通常型

研究期間: 2016年10月～2020年03月

研究者: JALEM Randy

1. 研究のねらい

本研究では、材料シミュレーションと機械学習を組み合わせることによって、全個体電池用の固体電解質として優れた材料を系統的・効率的に探索する目的とした。

2. 研究成果

(1) 概要

現在のリチウムイオン電池は可燃性の有機電解液を用いているため安全性の問題がある。不燃性のセラミックスを固体電解質として用いた全固体電池はより高い安全性と一層の大型化、そして高エネルギー密度化を実現する革新的な電池として期待されている。このような全固体電池は、電気自動車の電源として期待される。本研究では、材料シミュレーションと機械学習を組み合わせることによって、固体電解質として優れた材料を系統的・効率的に探索する目的とした。

(2) 詳細

研究テーマA「異なる結晶構造に対してマテリアルズ・インフォマティクスの手法を適用することができるジェネラルな記述子の開発」

ICSD と呼ばれる結晶構造データベースには9000件という膨大な構造プロトタイプが登録されている。そのデータベースに固体電解質として優れた材料を機械学習で系統的・効率的に探索するために、ジェネラルな材料の説明変数が必要である。そこで、考えたアイデアとしてはポロノイ分割という構造表現方法である。結晶構造中の各原子を点とみなして、ポロノイ分割をし、得られたポロノイ分割図形の頂点の数や辺の数、辺の長さや多面体の面積や体積などを、ヒストグラムにして、これを機械学習で用いる説明変数とした。結晶構造がバラバラのLiを含有する酸化物1000サンプルについて、凝集エネルギー、密度、バンドギャップ、分解エネルギーなどのデータを抽出し、そしてGBR法で、ポロノイ分割図の記述子を用いて機械学習を行った結果、高い予測性能が得られることがわかった。今後、本手法を使ってさらに大きな材料空間で優れた特性を有する材料の探索に活用する予定である。

研究テーマB「DFT計算とベイズ最適化を組み合わせた全固体電池用高速イオン伝導体材料の効率的探索」

イオン伝導のような遷移状態が関与する材料特性に対して、高精度な密度汎関数理論(DFT)計算による評価は、計算コストが非常に高く適用例は限定的である。これの解決策は、DFT計算に加えてベイズ最適化(BO)を適用することである(DFT+BO)。これは、機械学習

による予測とその不確実性情報の両方を効果的に利用することを目的としており、優れたイオン伝導を示す可能性の高い候補化合物に計算リソースを効果的に割り当てることができる。DFT+BO のアプローチを検証するために、A、M、X、および Z をイオン置換させるようなタボライト型 $AMXO_4Z$ 化合物の組成空間をスクリーニングした。タボライト型構造はイオン伝導性が高い既知材料があることから選択した。なお、DFT 計算ではイオン伝導度ではなくイオンの移動エネルギー(E_b)を評価とした。 E_b が低いほどイオン伝導性は高くなる。また組成 A は Li、Na が該当し、それぞれがキャリアイオンになると想定している。318 種類のタボライト型 Li および Na 含有化合物に対して今回開発した方法を適用したところ、最適化合物を 90% 確率で発見するための計算量は、網羅的全材料を評価する計算量に対する 30% まで圧縮できることが確認された。更に、全固体電池で要求される 0.3 eV 未満の移動エネルギーを有する化合物の再現性能は、ランダム探索よりも 2 倍有効であることが示された。このアプローチは、高速イオン伝導体の大規模な材料スクリーニングにおける計算上のボトルネックを解決できるものと期待される。

研究テーマ C「第一原理計算を用いた複雑な固体材料界面の電子・イオン状態の調査」

本研究では、全固体電池の正極—固体電解質界面である $LiCoO_2$ 正極— Li_3PS_4 固体電解質界面を調査し、界面構造サンプリングを行いました。その結果、界面を挟んで Co^{3+} イオンと P^{5+} イオンおよび O^{2-} イオンと S^{2-} イオンの交換(相互拡散)がエネルギー的に安定であることがわかった。さらに PO_4^{3-} というユニットが PS_4^{3-} ユニットよりも安定であることも示された。これらはこれまでの実験観測と一致しており、本計算手法の予言性の高さを保証するものである。さらにヘテロ固固界面付近の各 Li イオンサイトの Li の化学ポテンシャル、界面を横切る Li イオンのポテンシャルエネルギー面などを計算した結果、充電初期に界面付近の電解質から電子と Li イオンがまず移動することがわかった。前者は固体電解質 Li_3PS_4 の酸化に相当し、後者は電解質界面に Li イオンの動的な欠乏層が形成されることを意味し、いずれも実験観測を説明するものとなっており Li イオン伝導の界面抵抗と強く相関していることが明らかになった。本研究で得られた界面のマイクロな描像は、酸化物電解質や界面イオン伝導の改善に有効とされる酸化物緩衝層の効果も説明可能なものであり、より一般的な理論を与えるものとなっている。このような知見は界面制御の指針獲得、最適設計を促進し、固体デバイスの実用化・高度化に貢献することが期待される。

3. 今後の展開

実験の方と連携して、材料スクリーニングから得られた逆ペロブスカイト型 $Li_aX_bZ_c$ 系の新規材料の合成を行う。

4. 自己評価

本研究では、全個体電池用の固体電解質として優れた材料を系統的・効率的に探索する目的とした。それを実現するために、材料の探索方法を開発したが、その開発した方法の実際の有効性の評価はまだです。実験のフィードバックと合成結果からによって、開発した材料スクリーニング方法の探索基準と材料空間の定義のしかたを改善する可能性がある。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. Bo Gao, Randy Jalem, Yanming Ma, Yoshitaka Tateyama, "Li+ Transport Mechanism at the Heterogeneous Cathode/Solid Electrolyte Interface in an All-Solid-State Battery via the First-Principles Structure Prediction Scheme", *Chemistry of Material* 2020, 32, 85–96.
2. Yoshitaka Tateyama, Bo Gao, Randy Jalem, Jun Haruyama, "Theoretical picture of the positive electrode-solid electrolyte interface in all-solid-state battery from electrochemistry and semiconductor physics viewpoints", *Current Opinion in Electrochemistry* 2019, 17, 149–157.
3. Randy Jalem, Masanobu Nakayama, Yusuke Noda, Tam Le, Ichiro Takeuchi, Yoshitaka Tateyama, Hisatsugu Yamasaki, A General Representation Scheme for Crystalline Solids based on Voronoi-Tessellation Real Feature Values and Atomic Property Data, *Sci. Tech. Adv. Mater.* 2018, 19, 231–242.
4. Randy Jalem, Kenta Kanamori, Ichiro Takeuchi, Masanobu Nakayama, Hisatsugu Yamasaki, Toshiya Saito, Bayesian-Driven First-Principles Calculations for Accelerating Exploration of Fast Ion Conductors for Rechargeable Battery Application, *Sci. Rep.* 2018, 8, 5845.
5. Hiromasa Shiiba, Nobuyuki Zettsu, Miho Yamashita, Hitoshi Onodera, Randy Jalem, Masanobu Nakayama, Katsuya Teshima, Molecular Dynamics Studies on the Lithium Ion Conduction Behaviors Depending on Tilted Grain Boundaries with Various Symmetries in Garnet-Type Li₇La₃Zr₂O₁₂, *J. Phys. Chem. C* 2018, 122, 21755–21762.
6. Randy Jalem, Combining DFT Calculations and Bayesian Optimization for the Efficient Search of Fast Ion Conducting Materials for All-Solid-State Battery Applications (DFT 計算とベイズ最適化を組み合わせた全個体電池用高速イオン伝導体材料の効率的探索), *Japan Engineering & Technology Intelligence (JETI)* 2018, 66, 33–37.
7. Masanobu Nakayama, Kenta Kanamori, Koki Nakano, Randy Jalem, Ichiro Takeuchi, Hisatsugu Yamasaki, Data-driven Materials Exploration for Li-ion Conductive Ceramics by Exhaustive and Informatics-aided Computations, *Chem. Record* 2018, 18, 1–9.
8. Nataly Carolina Rosero-Navarro, Ryunosuke Kajiura, Randy Jalem, Yoshitaka Tateyama, Akira Miura, and Kiyoharu Tadanaga, "Significant reduction in the interfacial resistance of garnet-type solid electrolyte and lithium metal by thick amorphous lithium silicate layer", *ACS Appl. Energy Mater.* 2020, 3, 6, 5533–5541.
9. Bo Gao, Randy Jalem, and Yoshitaka Tateyama, "Surface-Dependent Stability of the Interface between Garnet Li₇La₃Zr₂O₁₂ and the Li Metal in the All-Solid-State Battery from First-Principles Calculations", *ACS Appl. Mater. Interfaces* 2020, 12, 14, 16350–16358.
10. Masanobu Nakayama, Kenta Kanamori, Koki Nakano, Randy Jalem, Ichiro Takeuchi, Hisatsugu Yamasaki, "Data-driven Materials Exploration for Li-ion Conductive Ceramics by Exhaustive and Informatics-aided Computations", *Chem. Record* 2018, 18, 1–9.

(2)特許出願

研究期間累積件数:1件

1.

発 明 者: ハレム ランディ

発明の名称 : リチウムリッチアンチペロブスカイト化合物、リチウムイオン二次電池用固体電解質およびリチウムイオン固体二次電池

出 願 人: ハレム ランディ

出 願 番 号 : 2020-033117

研究報告書

「ベイズ事後分布を探索重みに活用した物質構造探索の予測性向上」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 本郷 研太

1. 研究のねらい

「物性の第一原理計算」と「機械学習などのデータ科学」との融合展開「マテリアルズ・インフォマティクス(MI)」において、その究極の目的は、「所望の物性を実現する構造を予測する事」である。本研究は、新規 MI 手法を開発して、「膨大な化合物空間」の中から所望の物性を持つ新規化合物、すなわち、「埋蔵分子」を効率的に探索する。そのための手法として、本研究ではベイズ推定に基づくアプローチを採用する。当該手法は、ベイズの定理に基づき、尤度を物性予測の回帰モデル、化合物構造の事前分布を自然言語モデルに基づき構築することで、それらの積をベイズ反転させることで、所望の物性を持つ化合物構造を見出す事後分布を構築する。すなわち、回帰モデルを逆に辿り、「化合物らしさ」の事前情報を統計的重み付けすることで、所望物性を持つ化合物を効率的に探索することが可能となる。このベイズ構造探索法は、所望物性を持つ化合物を探索する仮想スクリーニングの枠組みを超えた新しい化合物探索手法であり、広大な化合物空間から効率的・選択的に、所望物性を持つ化合物候補を提案することができる。本研究は、特に、回帰モデルと事前分布の構築で、物質・材料研究としては大規模なビッグデータを利用して、探索精度の向上を実現する。

本研究では、ベイズ構造探索法の確立、その適用事例として、有機太陽電池用途の新規化合物探索を実施する。ドラッグデザインやケモ・バイオインフォマティクスなどの関連分野を含め、逆問題としての物質探索の先行研究は多くない。特に、任意の物性量レンジに存在する化合物を探索する実用段階には至っていない。これらまでの予備的研究から、順問題回帰モデルの物性予測性能が、最終的な物質構造探索の信頼性に大きな影響を与えることがわかっている。回帰モデルの性能向上に向けては、モデル選択のみならず、パラメータ学習のために大量の学習データが必要となる。本研究課題で、ベイズ物質構造探索の実働事例を確立すれば、マテリアルズ・インフォマティクス分野の様々な問題に対する有効なアプローチとして、広く普及するものと期待される。

2. 研究成果

(1) 概要

本研究では、(1) 学習データ生成のためのハイスループット第一原理計算、(2) 埋蔵化合物探索を目的とするベイズ構造探索法の研究基盤確立、及び、(3) ベイズ構造探索法の実材料開発に取り組んだ。

(1) ハイスループット第一原理計算

各種化合物データベースから既知化合物情報を入手し、第一原理計算による物性算定を行い、データベース化を進めている。本研究課題で得られた化合物の部分セットに対して、実験と計算結果を比較して、計算精度の検証を行った[論文発表 2/4/5]。これらの部分データに対

して解析を行った結果、物質探索に資する記述子の解明[論文発表 2/4]や高熱伝導率を持つポリマー結晶を発見できた[論文発表 4]。

(2) ベイズ構造探索法の開発

本研究では、具体的な問題設定として、HOMO-LUMO ギャップと内部エネルギーを対象物性量とする化合物探索に取り組み、ベイズ構造探索法の基盤を確立することができた。本研究では、尤度関数として線形回帰モデルを採用し、PubChem データベースからランダムに選んだ 10,000/6,674 個の化合物物性を学習/テストデータとした。対象物性の計算には Gaussian09 を用いた密度汎関数法(DFT/B3LYP)により算定した。事前分布としては、自然言語モデルに基づく化合物構造生成法を開発し、PubChem データベース収録の 50,000 化合物構造データを用いて学習を行った。化合物類似度を指標として、得られた仮想化合物とデータベース登録既知化合物を照合した結果、データベースには登録されていない新規構造を持つ化合物候補がベイズ構造探索法によって提案されていることがわかった[論文発表 1]。

(3) ベイズ構造探索法による実材料開発

本研究では、ベイズ構造探索による実材料開発の適用事例として、有機太陽電池のドナー材料探索に取り組んだ。実際の材料作成まで至っていないが、仮想分子ライブラリーを構築している。本研究課題に関連して、ベイズ構造探索法を活用した高熱伝導率ポリマー材料開発で、当該法の有用性を確認できた[論文発表 3]。

(2) 詳細

研究課題(1) 学習データ生成のためのハイスループット第一原理計算

研究テーマ A: 既存化合物構造に対する第一原理計算の実施

化合物の公開データベースとして、PubChem、Polymer Genome、Materials Project などのデータベースから既知化合物の情報を入手し、第一原理計算を実施して、独自のデータベースを構築している。本研究課題は、後段の MI 研究展開への学習データ提供以外にも、それ自体に利用価値があり、次に示すスピンアウト的研究成果が得られた。発表論文 5: 有機化合物の部分集合データを活用して、分子理論と組み合わせることで、液体プロセスなどの濡れ性制御に必要なハマカ一定数を評価し、実験との比較検証を行った。得られた計算結果は、十分な精度を持ち、ハマカ一定数の計算科学的算定アプローチが確立できた。発表論文 2: 層状化合物系の第一原理計算と呼応するマーデルング・ポテンシャルの相関関係を見出すことができた。当該ポテンシャルは、第一原理計算よりも計算コストの低い物理量であり、新規層状化合物の探索スクリーニングにおける良い記述子として活用できると期待される。論文発表 4: Polymer Genome Project に登録されているポリマー結晶構造を利用して、熱伝導率算定のベンチマーク計算を行った。その結果、計算科学の枠組みで、低温でポリエチレン結晶を超える熱伝導率を持つポリマー結晶を発見することができた。また、熱伝導率と相関する調和近似の範疇で算定可能な記述子を発見することができた。当該記述子を利用すれば、計算コストの非常に大きな熱伝導率計算を実施せずに、大量のポリマー結晶の仮想スクリーニングを実施することが可能となる。

研究課題(2) 埋蔵化合物探索を目的とするベイズ構造探索法の研究基盤確立

ベイズ構造探索法の構築では、ベイズの定理より、尤度関数 $P(Y|S)$ と事前分布 $P(S)$ の積(に比例する)として、所望の物性 Y を持つ化合物構造 S の事後分布 $P(S|Y)$ を算定することができ、この事後分布に従ってサンプリングを行うことで、化合物候補を得ることができる。本課題では、次の小項目に分けて、研究を進めた。

研究テーマ B: 事前分布 $P(S)$ の構築

化合物 S を SMILES 形式で記述することで、化合物を文字列として扱うことが可能となる。その結果、自然言語処理の言語生成モデルとして知られている n -gram モデルを適用することが可能となる。この n -gram モデルを化合物構造に対する文法制限を考慮して拡張することで、化合物生成器を構築できる。拡張 n -gram モデルの学習では、元素列の出現確率を既知化合物データベースから頻度として学習している。本研究では、具体的には、PubChem データベースからランダムサンプリングした数千~万程度の化合物を学習データとして活用した。

研究テーマ C: 順問題学習の回帰モデル $P(Y|S)$ の構築

回帰モデルとしては、線形・非線形回帰の複数モデル(ガウス過程回帰、サポートベクトル回帰、ランダムフォレスト、勾配ブースティング等々)を検証した。また、記述子としては、各種フィンガープリント(PubChem、MACS 等々)を検証した。学習データとしては、数万程度の既知化合物の第一原理計算の結果を利用している。

研究テーマ D: ベイズ逆問題予測に基づく候補分子構造生成の実装

研究テーマ B と C の尤度関数 $P(Y|S)$ と事前分布 $P(S)$ を用いて、事後分布 $P(S|Y)$ を構築し、この分布に従って、モンテカルロサンプリングを実施する。得られた候補化合物については、第一原理計算を行い、性能検証を行った。

研究テーマ E: 実装のパッケージ化

研究テーマ B/C/D のパッケージ化までは完了していないが、本研究課題に関連する一部の構造生成プログラムに関しては研究成果として独立に原著論文出版に向けて、Python パッケージ化を進めている(arXiv:1911.08071)。

研究課題(3) ベイズ構造探索法の実材料開発

研究テーマ F: 有機太陽電池のドナー材料探索/埋蔵分子の発掘

研究課題(2)で確立した研究基盤を利用して、有機太陽電池のドナー材料探索に取り組んだ。回帰モデルと事前分布の構築には、研究課題(1)の物性・構造データや HCEP(Harvard Clean Energy Project)データベースの物性・構造データを活用した。得られた候補化合物については、仮想化合物データベースに登録している。後述の今後の展開に示すように、実験研究者からの協力の下、得られた候補化合物から実際に有機太陽電池を作製し、エネルギー変換効率の測定に取り組む。なお、ベイズ構造探索法の実材料開発への適用事例として、関連研究として、高熱伝導率ポリマー開発を行い、実証研究を行った[論文発表 3]。

3. 今後の展開

本研究では、有機太陽電池ドナー材料の仮想化合物データベース構築まで到達することができた。候補化合物をドナー材料として実際に合成・作製することで、ベイズ構造探索法の実証研究事例として、本研究課題を完成させる。当該材料の実際の合成・作製では、学内実験研究者との共同体制を確立している。また、当該探索法の汎用性を実証するために、研究

課題(1)で得られた第一原理計算データベースを活用して、新規課題に取り組む。具体的には、高熱伝導率ポリマー結晶の探索を行う。現在、学習データを生成するための第一原理計算を継続している。また、候補ポリマーユニットからポリマー結晶構造を決定するためには、結晶構造探索手法の研究基盤確立が必要となるが、粒子群最適化法に基づく結晶構造探索法を利用する(既に別系統の化合物に適用し、新たな結晶相の発見につながっている)。ベイズ構造探索法と結晶構造探索法を組合せ、新規ポリマー結晶探索研究を展開する。なお、実際のポリマー合成・結晶作製についても、学内実験研究者との共同体制を確立している。

4. 自己評価

本研究は、(1)ハイスループット第一原理計算、(2)ベイズ構造探索法の開発、(3)太陽電池のドナー材料開発へのベイズ構造探索法の適用を目的として、課題設定を行った。課題(1)と課題(2)については、当初の計画通り、目的を達成することができた。本研究提案手法は、既知化合物群からの仮想スクリーニングで発見できない新たな化合物を研究対象に据えることを可能とする新しいMI研究であり、当該研究分野における重要な研究成果であると考えている。また、その研究基盤を利用するための第一原理計算データを大量に確保することができた。この点では、研究費を利用して導入した計算サーバ(データ保存のストレージやバックアップシステムを含む)や、データ整理を目的に雇用した研究補助員、及び、学生が大いに貢献している。得られた計算データを素早く整理したことで、得られたデータセットの部分集合の解析を行う余裕が生まれ、その結果として当初の計画では想定していなかった研究成果も得られており、ベイズ構造探索手法に基づく新たな研究展開への道を拓くことが可能になったと考えている。課題(3)は、本研究提案手法の有効性に関する実証研究であるが、候補化合物の提案とそのデータベース化に留まり、研究期間内での新規材料の発見には至らなかった。特に、未知構造の化合物を対象とする場合、その合成可能性を考慮した研究課題設定が必要であった。合成可能性を検証する枠組みの整備と、それに基づく事前確率の再構築(すなわち、合成可能性を考慮した候補化合物生成)が今後の課題と考えている。課題(3)に関連して、来年度以降の研究費確保の見通しは立っており、当該課題を継続実施し、新規材料の発見を目指す。実際に新材料の発見に繋がれば、データ駆動に基づく新たな材料開発が大きく前進するものと期待される。MI 材料開発アプローチの導入は、これまで絨毯爆撃的かつ経験的な実験計画に基づき行われてきた材料開発を刷新するもので、材料開発に伴うコストと時間の削減に大きく寄与するものと期待される。

5. 主な研究成果リスト

(1)論文(原著論文)発表

1. Hisaki Ikebata, Kenta Hongo, Tetsu Isomura, Ryo Maezono, Ryo Yoshida, "Bayesian molecular design with a chemical language model", *Journal of Computer-Aided Molecular Design*, 2017, 31, 379-391.
2. Daichi Kato, Kenta Hongo, Ryo Maezono, Masanobu Higashi, Hironobu Kunioku, Masayoshi Yabuuchi, Hajime Suzuki, Hiroyuki Okajima, Chengchao Zhong, Kousuke Nakano, Ryu Abe, Hiroshi Kageyama, "Valence Band Engineering of Layered Bismuth

	Oxyhalides toward Stable Visible-Light Water Splitting: Madelung Site Potential Analysis”, Journal of the American Chemical Society, 2017, 139, 18725–18731.
3.	Stephen Wu, Yukiko Kondo, Masa-aki Kakimoto, Bin Yang, Hironao Yamada, Isao Kuwajima, Guillaume Lambard, Kenta Hongo, Yibin Xu, Junichiro Shiomi, Christoph Schick, Junko Morikawa, Ryo Yoshida, “Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm”, npj Computational Materials, 2019, 5, 66:1–11.
4.	Keishu Utimula, Tom Ichibha, Ryo Maezono, Kenta Hongo, “Ab Initio Search of Polymer Crystals with High Thermal Conductivity”, Chemistry of Materials, 2019, 31, 4649–4656, (selected in Virtual Issue on Machine-Learning Discoveries in Materials Science, Chemistry of Materials, 2019, 31, 8243–8247.)
5.	Hideyuki, Takagishi, Takashi Masuda, Tatsuya Shimoda, Ryo Maezono, Kenta Hongo, “Method for the Calculation of the Hamaker constants of Organic Materials by the Lifshitz Macroscopic Approach With DFT”, Journal of Physical Chemistry A, 2019, 123, 8726–8733.

(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)

1.	本郷研太、「第一原理計算とベイズ統計に基づく新物質探索」、マテリアルズ・インフォマティクスによる材料開発と活用集(技術情報協会) 第3章第3節、pp.62–69。(著書・執筆分担)
2.	Kenta Hongo, “Recent Advances in Materials Simulations and Informatics”, International Congress on Pure & Applied Chemistry (ICPAC) 2018, 2018/03/10, Siem Reap, Cambodia (Invited).
3.	Kenta Hongo, “Computational materials design from ab initio simulations to ab initio materials informatics”, The 10th International Conference of the Asian Consortium on Computational Materials Science (ACCMS–10), 2019/07/26, City University of Hong Kong, China (Invited).
4.	Kenta Hongo, “Data-driven approach to computational materials design”, 20th International Union of Materials Research Societies International Conference in Asia (IUMRS–ICA), 2019/09/24, Perth Convention and Exhibition Centre, Perth, Australia (Invited).
5.	Kenta Hongo, “Data-driven approach to molecular design”, The 5th International Conference on Molecular Simulation 2019 (ICMS 2019), 2019/11/04, Jeju, Republic of Korea (Invited).

研究報告書

「情報科学手法を利用した界面の構造機能相関の解明」

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 溝口照康

1. 研究のねらい

界面における原子構造は単結晶の構造とは異なるため、単結晶では得られない物性が得られる。たとえば、過電流から電気機器を保護するためのバリスタ素子には酸化亜鉛が使用されているが、バリスタ特性は多結晶酸化亜鉛のみで得られる。つまり、バリスタ特性は界面があつて初めて得られる機能である。界面は新機能を発現するだけでなく、悪影響を及ぼすこともある。例えば、チタン酸ランタンリチウムの単結晶は優れたリチウムイオン伝導特性を示し、固体電解質として期待されてきたが、界面において大幅に伝導特性が低下することが知られている。このような界面における特異な特性は、結晶内部と異なる界面固有の原子構造に起因している。そのため、界面構造を解析することが材料研究において最も重要なテーマとされてきた。

界面の機能は界面固有の構造に起因しているため、界面の構造解析がこれまで盛んに行われてきた。特に、高い空間分解能を有する透過型電子顕微鏡(TEM)や電子線エネルギー損失分光(EELS)のような実験的手法と、第一原理計算や分子動力学計算(MD)のようなシミュレーション法が界面の構造解析にひろく用いられてきた。

一方で、界面構造を決定することは容易ではない。これは界面の有する自由度が原因である。自由度を制限した単元素金属の対応格子理論(CSL)に基づく対称傾角粒界でも、1種類の粒界あたり数百～数万個の候補構造が存在しており、そのシミュレーションには膨大な計算時間を要する。さらに、界面から取得されるTEMやEELS計測データも膨大かつ複雑なため、その解析は容易ではない。

本研究のねらいは、機械学習をはじめとした情報科学手法を利用することで、界面構造解析を加速させることにある。

具体的には、(1)仮想スクリーニングやクリギング、転移学習といった機械学習法をシミュレーションに組み合わせることで界面シミュレーションを加速すること、(2)階層型クラスタリングおよびニューラルネットワークを利用することで界面計測データ解析を加速させる。(1)と(2)により界面構造データベースを作成し、さらに界面構造の物性データを計算する。得られた構造と物性の相関関係を機械学習により解析し、界面における構造と機能の相関性を明らかにすることが本研究のねらいである。

本研究により界面構造相関が明らかになれば、界面の構造から物性を予測することが可能となり、界面設計の実現につながると期待される。

2. 研究成果

(1)概要

本研究は、機械学習を併用して界面構造解析を効率化し、さらに界面における構造と機能との相関性(構造機能相関)を明らかにすることを目的としている。主な研究内容は、1)既存

手法のさらなる高速化, 2) 多元系化合物への利用, 3) 計測データ解析の効率化と, 4) 界面における構造機能相関の解明である. 本申請研究により, これらの研究項目に対して成果を得た. たとえば, 転移学習をクリギング法に組み合わせることにより, 従来のクリギングと比較して3倍(クリギング無しの場合と比較して3600倍)の高速化に成功した. さらに, 酸化物界面の高速決定にも成功し, 多元系化合物においても機械学習法が有効であることを実証した. また, 階層型クラスタリングと決定木を利用することで, スペクトルデータを高速に解釈・予測する手法を開発した. さらに, ニューラルネットワークを利用することでスペクトルデータから結合距離や結合角度, 電子構造などの物質情報を直接決定するための手法も開発した. 界面における構造機能相関に関して, 界面の空孔偏析挙動と界面構造との相関性を調べた. 界面近傍の各原子サイトから取得される幾何学的な情報と空孔偏析挙動との相関性を明らかにした. 同研究を通し, 汎用的な予測モデルを構築するための指針や, これまで議論されてこなかった界面構造の基礎的な知見を明らかにすることが出来た.

(2) 詳細

研究成果

研究テーマ1「転移学習を利用したクリギングによる界面構造決定の加速」

界面構造を決定するうえで「仮想スクリーニング」と「クリギング」という機械学習の手法を利用し, 構造決定のプロセスを加速させてきた. 二つの手法を比較すると, 仮想スクリーニングでは予測モデルを使用するために数百~数千個の候補構造を実際に計算する必要があるが, 一度予測モデルを構築することができれば, 従来の手法の数万倍程度高い効率で界面構造を決定することができる. 一方で, クリギングは予測モデル構築の手間がないものの, その効率は数百倍程度であり仮想スクリーニングには劣る.

そのようなクリギングをより加速させることを目的として, クリギングと転移学習を組み合わせた手法を開発した. 転移学習は, ある課題で学習した結果を, 他の類似した課題に再利用することで, より高精度(もしくは高効率)化する手法である.

本手法では, モデル CSL 粒界の Σ 粒界に対し, Σ 3粒界をクリギングした際に得られた予測モデルを, そのまま Σ 5に利用して再学習する. そこで得られた新たな学習モデルを次の Σ 7粒界に利用する. そのようにして, 各粒界において学習モデルを Σ 値順に転移することにより, クリギングの試行回数を減らすことに成功した. 通常のクリギングよりも3倍程度, 従来の手法よりも3600倍程高速化することに成功した.

同成果は, 転移学習を物質解析に利用した重要な成果であり, 図1に示すようなイメージ図を作成してプレスリリースを行い, 国内外で報道された.

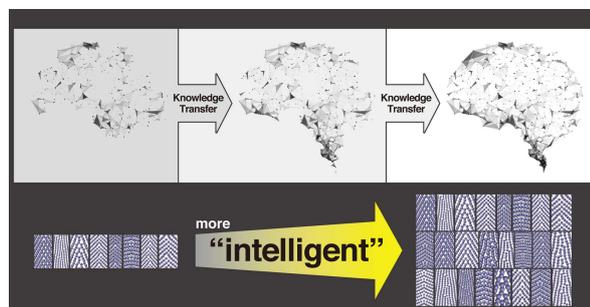


図1 転位学習により多くの界面構造を決定することに成功

研究テーマ2「機械学習を利用した多元系化合物粒界構造決定」

「仮想スクリーニング」及び「クリギング」の手法開発においては, 単金属のような単純な物

質系を対象としてきた。一方で、工業的に使用されている材料は多元系化合物が主である。多元系化合物の粒界構造決定のためにクリギングを使用した。クリギングは2つの結晶粒を準備するのみで界面構造決定することが可能なため、界面終端面の組み合わせ数が多くなる多元系にも有効である。本研究では、MgO, NiO のような岩塩型構造に加え、CeO₂ や TiO₂ のような複雑な結晶構造を有する化合物の界面構造決定を行った。それらの粒界では2万から4万程度の候補構造が存在している。

クリギングを実施した結果、すべての化合物において十数回から数十回の試行計算によって最安定構造をえることに成功した。単金属だけでなく、多元系化合物においてもクリギングが有効な理由を明らかにするために、界面エネルギーの剛体変位に対する依存性を解析した。その結果、金属と多元系化合物の両方において、剛体変位に対して粒界エネルギーが連続的に変化し、不連続的な変化を生じないことが明らかとなった。本研究を通し、物質を構成する元素の種類によらず、その界面エネルギー分布が連続的であればクリギング法が有効であると結論付けることができた。

研究テーマ3「機械学習を利用した計測データ解析」

界面の解析においてシミュレーションに加えて高分解能電子顕微鏡を用いた像やスペクトル計測も行われている。特に、像だけではわからない結合距離のわずかな変化や、結合角度の変化、さらに電子構造の変化はスペクトルから取得される。一方で、そのようなスペクトルデータを解釈し、結合距離や、角度、電子構造の情報に変換するためには、シミュレーションを用いたデータ解析が不可欠である。さらに、スペクトルに現れるわずかな変化を正確に解釈するためには専門的な経験が必要となる。そのような研究者駆動型のスペクトル解析における諸問題を解決するために機械学習を活用した。

具体的には、教師無し階層型クラスタリングと、教師あり決定木を用いた。今回開発した手法を模式的に図2に示す。(a)スペクトルデータを、(b)階層型クラスタリングにより分類し、(c)さらにその分類結果を教師として物質情報(結合距離、角度、電子構造)の決定木を作る。デンドログラム(b)と決定木(c)という2つの「木」を利用することで、

スペクトルの解釈に加え、物質情報からスペクトルを予測することも可能であることを示した。

さらに、スペクトルを解釈するだけでなく、スペクトルから直接物質情報や物性情報を取得するための手法も開発した。同手法はスペクトルを input、構造情報や物性情報を output とし、ニューラルネットワークを形成するというものである。実際に SiO₂ 多型から取得された 1000 本以上の O-K 端というスペクトルから、結合距離、結合角度、電子構造(イオン性、共有結合

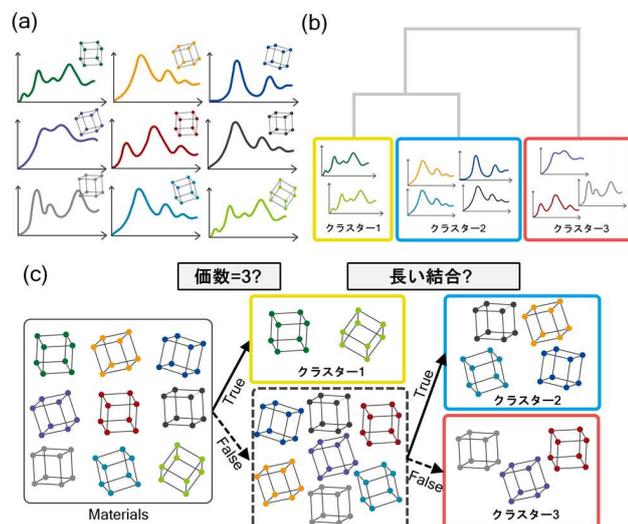


図 2(a) スペクトルデータベース, (b)階層型クラスタリングによるスペクトルの分類, (c)スペクトルの分類結果を教師とした構造情報の決定木

性), さらにスペクトルの遷移エネルギーを予測するためのモデルを構築した。その結果, すべての物質情報を高い精度で予測することに成功した。今回開発した手法により, 界面の構造ではなく, 物性を定量的に可視化することが可能になると期待できる。

これらデータ駆動型スペクトル解析法は JST からプレスリリースされ, 国内外で広く報道された。

研究テーマ4「界面における構造機能相関の理解」

機械学習法を利用することで界面構造のデータベースを効率的に作成することができる。さらに, その界面構造に対して網羅的に界面物性を計算することで界面物性のデータベースを作成することもできる。構築した界面構造と界面物性のデータを解析することで界面における構造機能相関を理解することが可能になる。

本研究では, 単元素金属の対称傾角粒界, 非対称傾角粒界に関して網羅的に空孔の偏析エネルギーを計算し, 界面における空孔偏析と構造との相関性を調べた。記述子として, 界面近傍における結合距離や結合本数, 原子数密度などを用い, 目的変数として各サイトにおける空孔偏析エネルギーを回帰した。各回転軸において個別の予測モデルを構築したが, 各回転軸の予測モデルを他の回転軸の物性予測に利用した際, 予測精度が悪いことが明らかとなった。つまり, 各回転軸で構築した予測モデルは汎用的でなく, 他の回転軸には利用できないことが分かった。

この結果を踏まえ, 汎用的な予測モデルを構築するために, 異なる回転軸の粒界物性を用いて予測モデルを再度構築した。同予測モデルは回転軸が異なる粒界の物性を高精度で予測することができた。また, そのような予測モデルの構築法が, 粒界物性だけでなく粒界構造自体の予測にも有効であることも分かった。以上の研究を通し, 界面物性及び界面構造の汎用的予測モデルの構築方法を明らかにした。

また, 得られた構造と機能の相関性を調べ, 機能の予測モデルを構築し, 構造記述子の重要度を調べた。その結果, 界面への空孔偏析に関しては, 界面近傍に生じる構造ひずみ(短結合および長結合)が重要な役割をはたしていることが明らかとなった。

これらの研究から, 界面における物性(主に空孔偏析)と構造との相関性を明らかにすることができた。

3. 今後の展開

本研究により, シミュレーションおよび計測を用いた界面構造解析を加速するための基盤技術を確立できたと考えている。今後の展開として考えられるのは次の2点である。1) 実用的な物質系への利用, 2) 手法確立を通して得られた知見のさらなる調査。

1) に関して, 本研究では単一物質間に形成される界面の構造決定を対象としてきた。一方で, 電子デバイスや電池においては, 異種物質が接するヘテロ界面が存在し, デバイス性能に重要な役割をはたしている。本研究で確立した手法を, そのような実用性の高いヘテロ界面への利用が期待できる。

2) に関して, 界面構造・物性の回転軸依存性や, 転移学習の学習順序依存性など, 本申請研究において, 材料学的に興味深い知見が得られている。それらについてさらに調査をすること, これまでの研究では得ることができなかった界面構造に関する重要な知見を得ることができ

ると期待できる。

4. 自己評価

本研究は、機械学習を利用した界面構造解析の加速と構造機能相関の解明を目的としており、上述のような研究成果はその目的を達成していると結論付けることが出来る。研究においては本研究費により導入した装置や計算機が有効に機能したと考えている。学術論文の投稿に加えて、学会誌への解説記事の執筆や、様々な学会・会合での講演・講師などを数多く行ってきた。また、得られた研究成果についてなるべくプレスリリースを行い、ひろくアウトリーチすることが出来たと考えている。

一方で、研究成果に興味を持っていただく機会があったものの、社会実装されているとはいえない。本研究で確立された手法の公開や、データの公開などを今後積極的に行いたいと考えている。

今回のさきがけ研究では「界面」にフォーカスした研究を実施した。一方で、多結晶材料中には転位や表面といった別の格子欠陥が多数存在している。機械学習がそのような格子欠陥の解析にも有効であると考えられるので、界面同様に機械学習を利用することで「格子欠陥インフォマティクス」の確立に尽力したいと考えている。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. H. Oda, S. Kiyohara, K. Tsuda, and T. Mizoguchi, "Transfer Learning to Accelerate Interface Structure Searches", J. Phys. Soc. Jpn (Letter), 86 (2017) 123601-1-4.
2. S. Kikuchi, H. Oda, S. Kiyohara, and T. Mizoguchi, "Bayesian optimization for efficient determination of metal oxide grain boundary structures", Physica B, 532 (2018) 24-28.
3. S. Kiyohara and T. Mizoguchi, "Searching the stable segregation configuration at the grain boundary by a Monte Carlo tree search ", J. Chem. Phys., 148 (2018) 241741-1-6
4. S. Kiyohara, T. Miyata, K. Tsuda, and T. Mizoguchi, "Data-driven approach for the prediction and interpretation of core-electron loss spectroscopy", Scientific Reports, 8 (2018) 13548-1-12
5. S. Kiyohara, M. Tsubaki, Kunyen Liao, and T. Mizoguchi, "Quantitative estimation of properties from core-loss spectrum via neural network", J. Phys.: Materials, 2 (2019) 024003-1-9
6. H. Oda, S. Kiyohara and T. Mizoguchi, "Machine learning for structure determination and investigating the structure-property relationships of interfaces", J. Phys.: Materials, 2 (2019) 034005-1-8
7. T. Mizoguchi and S. Kiyohara, "Machine learning approaches for ELNES/XANES", Microscopy, in press. [Invited review]

査読付き学術論文: 14 件

日本語解説: 8 件

(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)

主要な学会発表

“基調講演”“機械学習を活用した機能コア解析”

日本金属学会 2019 年度秋季大会, 岡山大学, 岡山, 9/12, 2019

“データ科学で理解するスペクトロスコピー”

日本化学会 2019 春季大会, 甲南大学, 神戸, March 17, 2019

“Data-driven approach for interface and spectrum”

PRESTO Inter. Symposium on Materials Informatics, Koshiba-hall, Tokyo, Feb. 9, 2019

“機械学習と原子分解能計測による結晶界面の構造解析”

地球惑星科学連合 2018 年年会, 幕張メッセ, 千葉, May 23, 2018

“原子分解能 STEM-EELS および第一原理計算による液体・気体の解析”

鉄鋼協会春季大会, 千葉工業大学, 新習志野, March 20, 2018

“機械学習を利用した粒界構造の効率的探索”

日本物理学会 秋季大会, 岩手大学, 岩手, 9/22, 2017

招待講演・基調講演:66件

受賞

溝口照康

功績賞, 日本金属学会, 2017 年 3 月

溝口照康

風戸賞, 風戸研究奨励会, 2017 年 3 月

著作物

“Nanoinformatics”, Springer (2018), e-book ISBN 978-981-10-7617-6

Chapter 8, “Atomic-Scale Nanostructures by Advanced Electron Microscopy and Informatics” 執筆担当

“Machine Learning in Chemistry: The Impact of Artificial Intelligence”,

Royal Society of Chemistry, (2020) 印刷中, (Hugh Cartwright edited)

Chapter “Machine learning for core-loss spectroscopy” 執筆担当

プレスリリース

1. 2019年3月26日 「理論計算や専門知識いらず！人工知能がスペクトルから物質の構造や機能を直接定量 ～物質開発の加速に期待～」
記事掲載：日刊工業新聞，日経産業新聞
オンライン掲載：オプトロニクス，日経オンライン，日刊工業オンライン，PhysOrg, EurekAlert, Medical Development, Engineering.com
2. 2018年9月7日 「人工知能が専門家の約2万倍の速さでスペクトルを解釈」
記事掲載：日刊工業新聞，鉄鋼新聞
オンライン掲載：オプトロニクス，日経オンライン，大学ジャーナル，Technology network, Science Daily, Lab manager, Phys Org, Scientific Computing, Asian Scientist, Spectroscopy Europe and Asia
3. 2017年12月15日 「世界初、液体中の原子1つ1つの運動を観察！ ～高性能電池や溶媒の開発、液体中の現象解明に革新～」
記事掲載：日刊工業新聞，化学工業日報，鉄鋼新聞
オンライン掲載：オプトロニクス，日経オンライン，マイナビ，Motor Cars, エキサイトニュース，Fabcross, Science Magazine, Nano werk, Mining, 日刊工業オンライン
4. 2017年12月12日 「電子顕微鏡で気体分子の挙動や特性に迫る」
オンライン掲載：オプトロニクス，日経オンライン，PhysOrg, Science Daily, Science News Line, Azo Materials, Nano werk, Technology Breaking News
5. 2017年11月15日 「人工知能が「繰り返し成長すること」で計算コストを1/3600に削減 ～界面構造を高速に決定し、高性能な物質開発を加速～」
記事掲載：日刊工業新聞，化学工業日報
オンライン掲載：マイナビ，fabcross, JPubb, Jnet21, nifty, 日経オンライン

研究報告書

特定混合比で発現する特異物性を利用した新材料創成のための 第一原理分子シミュレーションと機械学習の連携

研究タイプ: 通常型

研究期間: 2016年10月～2020年3月

研究者: 森 寛敏

1. 研究のねらい

我々の生活を支える空調技術は、エネルギー消費により実現される。投入したエネルギーに対する冷却効率の向上は、省エネルギー・環境的観点から重要な課題である。空調効率の決定因子の一つに冷媒が挙げられる。冷媒の善し悪しは、蒸発潜熱や蒸気圧の大きさに決まり、従来、それらの特性に優れたフロン類が使われてきた。だが、フロン類は塩素が分子内に含まれる場合、オゾン層破壊の原因になり、塩素を含まない場合でも、その高い安定性と赤外光吸収率により地球温暖化の原因となる。CO₂等の自然冷媒を用いる手法も開発されているが、その冷却効率は低く、CO₂の化学特性から現存する空調デバイスの冷媒を単純に入れ替えることも現実的ではない。

この背景の下、近年、各種冷媒材料を「混合」することで冷却効率を向上させようとする試みが行なわれている。系中で生じる同種・異種分子間の相互作用がモル分率に対して非線形に変化することを利用し、純冷媒で実現し得ない良特性を生む可能性を探索しているわけである。従来、混合物性は主に実験結果を基に語られ、純物性のモル分率に対する線形和からのずれ、すなわち活量で説明されてきた。だが、活量の定量的予測法は確立されていない。そのため、混合冷媒の開発は、地球規模の喫緊の課題であるにも関わらず、その開発は困難を極めている。

本研究では、第一原理分子シミュレーションと機械学習の連携によるこの状況の打破を目標に掲げた。すなわち、環境・エネルギー分野を支える機能材料「混合冷媒」の開発に際し、無限に存在しうる「成分の組み合わせ・組成」と「冷媒機能」の関係解明に挑戦した。混合冷媒の冷却効率を最大化するには、混合系でありながら温度すべり現象を示さない「共沸混合物」を探索せねばならない。本研究では、混合の化学的非線形性が分子間相互作用に由来することに着目し、各成分を構成する分子フラグメントの電子状態と、各種混合冷媒の候補となる混合物の共沸特製の関係性解明を目指した。

2. 研究成果

(1) 概要

本研究では、分子集合体としての溶液物性が、その構成要素である各分子フラグメントの一分子物性と分子間相互作用により記述できることに着目することで、従来、困難とされてきた混合冷媒（機能性混合溶液）の熱力学物性の迅速予測に挑戦した。冷媒の候補となり得る7万超の有機分子（構造を系統的に変化させたフルオロカーボン・ハイドロカーボン・ハイドロフルオロカーボン・フルオロオレフィン・ハイドロフルオロオレフィン・ハイドロオレフィン）につい

て、RI-BP86/TZVPD レベルの密度汎関数計算(分散力補正含む)により、分子体積・表面積・アスペクト比などの分子の幾何的特徴量と、単位表面積あたりの電荷密度分布、多極子モーメントなどの分子の電子的特徴量をデータベース化した。作成した特徴量と、アメリカ化学工学会データベースに登録されている、純物質の標準沸点(1166 化合物)・標準融点(1190 化合物)の実験データを正解情報とした、教師あり機械学習(ガウス過程回帰)を実施したところ、分子の単位表面電荷分布(σ -moment)と、体積や表面積、Coulomb-Matrix など、種々の分子構造に由来する情報を特徴量とした場合、標準沸点・標準融点予測の平均絶対誤差が、それぞれ 3 および 10 K になること」が分かった。また同様に、二成分混合物に関して共沸の可能性を判定する識別器を作成したところ、9 割弱の正解率で共沸可能性の判定が可能になった。今後、地球温暖化係数を加味して、作成した共沸判定器および沸点予測器を活用することで、共沸混合冷媒の迅速設計を行うことができる。

(2) 詳細

本研究では、分子集合体としての溶液物性が、その構成要素である各分子フラグメントの一分子物性と分子間相互作用により記述できることに着目することで、従来、困難とされてきた混合冷媒(機能性混合溶液)の熱力学物性の迅速予測に挑戦した。

研究テーマ①

「混合溶媒・冷媒の構成要素となる有機分子フラグメントの電子状態 DB の蓄積」

テーマ①では、二つの観点から有機分子フラグメントの電子状態データベースの構築を実施した。

まず、「最終的に二成分混合共沸冷媒の候補が予測された際、その分子を材料として昇華できるか」、という観点からデータベースの構築に着手した。具体的には、アメリカ暖房冷凍空調学会(ASHRAE)に登録され安全性が考慮された冷媒分子(217 種類)について、電子状態を系統的に評価した。即ち、RI-BP86/TZVPD レベルの密度汎関数計算(分散力補正含む)により、分子体積・表面積・アスペクト比などの分子の幾何的特徴量と、単位表面積あたりの電荷密度分布、多極子モーメントなどの分子の電子的特徴量をまとめた。

続いて、既存の ASHRAE に登録済みかどうかに関わらず、候補となり得る塩素非含有化合物(炭素数=C₂-C₈)の中から、混合冷媒の構成要素となり得る良熱力学物性を示す候補分子を迅速探索するため、構造を系統的に変化させたフルオロカーボン・ハイドロカーボン・ハイドロフルオロカーボン・フルオロオレフィン・ハイドロフルオロオレフィン・ハイドロオレフィン(合計7万種超)について、上述と同様な密度汎関数計算を実施し、これら化合物群の電子状態データベースを構築した。アメリカ化学工学会データベースに登録されている、純物質の標準沸点(1166 化合物)・標準融点(1190 化合物)の実験データを正解情報とした、教師あり機械学習(ガウス過程回帰)を実施したところ、分子の単位表面電荷分布(σ -moment)と、体積や表面積、Coulomb-Matrix など、種々の分子構造に由来する情報を特徴量とした場合、標準沸点・標準融点予測の平均絶対誤差が、それぞれ 3 および 10 K になること」が分かった。

研究テーマ②

「実験データを正解情報とした純物質の熱力学物性高精度予測方法の開発」

続くテーマ②では、テーマ①で作成した冷媒候補となる有機化合物の幾何的特徴量と電子的特徴量を活用して、純物質熱力学物性の高精度予測方法の開発に取り組んだ。アルカンやアルケンなどの単純な化合物に限れば、化学工学分野において Joback 法や量子化学計算に基づく COSMO-RS 法などの熱力学物性予測法が確立されている。だが、Joback 法は、熱力学物性を予測した化合物が未知の官能基を有する場合、適用することができない。また COSMO-RS 法については、アメリカ化学工学会データベースに登録されている有機化合物群の標準沸点(1166 化合物)・標準融点(1190 化合物)予測に網羅的に適用したベンチマークを取ったところ、実測値と予測値の平均絶対誤差が標準沸点では 17.5 K、標準融点については 46.0 K 存在することが明らかとなった。この実測値と予測値の大きな違いは、当該手法を、直接冷媒探索に用いる際には慎重を期すべきであることを意味する。そこで本研究では、COSMO-RS 計算時に利用する分子の単位表面電荷分布(σ -moment)と、体積や表面積、Coulomb-Matrix など、種々の分子構造に由来する情報を特徴量とした機械学習により、COSMO-RS 法の標準沸点・標準融点の推算精度向上を検討した。その結果、ガウス過程回帰を用いた場合、有機純物質の標準沸点・標準融点予測には、「500 分子程度の熱力学データを学習させれば、データ数に対して十分に収束した学習モデルが得られること」、「分子の幾何的特徴量と電子的特徴量のみをそれぞれ使用した標準沸点・標準融点予測モデルの精度比較より、標準沸点・標準融点を支配している要因は、幾何構造が主な部分を占めること」、「Coulomb-Matrix を特徴量として加えた場合には、異なる沸点・融点を与える幾何異性体の区別より精密にできるようになり標準沸点・標準融点予測の平均絶対誤差が、それぞれ 3 および 10 K になること」が分かった。

研究テーマ③

「実験データを正解情報とした二成分混合物の熱力学物性高精度予測方法の開発」

テーマ②で実施した、純物質の熱力学物性予測の高精度化を受け、テーマ③では、テーマ①で作成した冷媒候補となる有機化合物の幾何的特徴量と電子的特徴量を活用した、二成分混合物の共沸物性を機械学習により予測する方法の開発を行った。具体的には、混合物が共沸点を持つか否かを判定する識別器を、教師あり機械学習法により作成することとした。

特徴量は、テーマ②と同じものを用いた。正解情報としての共沸・非共沸のデータは、Dortmund Data Bank に登録されている 54074 件の実験データ(共沸混合物 707 組・非共沸混合物 53367 組)を用いた。サポートベクトルマシン(3 次 SVM)による識別学習を実施した結果、混合物の学習データ数を系統的に変えた機械学習により、データ数が(共沸:非共沸)=500:500 のとき、十分に学習曲線が収束し、9 割弱の高い識別正解率を得た。共沸が生じるか否かについて、簡便に実施できる小分子の量子化学計算から得た分子の幾何的特徴量と電子的特徴量を組み合わせることで、化学工学分野で重要な共沸の有無を、高精度かつ迅速に推定できる技術を開発した。

3. 今後の展開

本研究では、二成分混合系の共沸物性に焦点を当てた研究を実施したが、開発した技術は、共沸以外の混合熱力学物性の予測にも応用できると考えられる。具体的には、共融点の予測、表面張力などの予測に引き続き取り組みたい。

4. 自己評価

本研究では、共沸混合冷媒の設計を目指した、共沸点の予測器の開発を一つのマイルストーンとした。研究提案時に定義した、このマイルストーンについては、一定の成果を得ることができたと感じている。しかし、この成果は、社会・経済への波及効果を、この3年半で生み出すことはできず、理学的研究の枠組みに留まった点は、反省すべき点だと考えている。想定したよりも、共沸予測器の開発に時間を要したのが原因であるが、実用レベルでの共沸予測の精度向上のために、正解情報として、理論計算値ではなく、熱力学的「実験データ」そのものを精査する方向に最終的に拘った点は、本プロジェクト全体を完遂してよかった点だと考えている。今後、本研究で開発した技術を、社会実装できるよう、更なる研究に挑戦していく。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

- | |
|--|
| 1. N. Kuroki, H. Mori*, Applicability of Effective Fragment Potential Version 2 – Molecular Dynamics (EFP2-MD) Simulations for Predicting Dynamic Liquid Properties Including Supercritical Fluid Phase, <i>J. Phys. Chem. B</i> , 2019, 123, 194–200 (Inside Cover). |
| 2. N. Kuroki, H. Mori*, Applicability of Effective Fragment Potential version 2 – Molecular Dynamics (EFP2-MD) Simulations for Predicting Excess Properties of Mixed Solvents, <i>Chem. Phys. Lett</i> , 2018, 694, 82–85 (Frontiers Article). |
| 3. N. Kuroki*, S. Maruyama, H. Mori*, Theoretical Strategy for Improving CO ₂ Absorption of Mixed Ionic Liquids Focusing on the Anion Effect: A Comprehensive COSMO-RS Study, <i>Ind. Eng. Chem. Res.</i> , 2020, 59, 8848–8854 (Inside Cover). |

(2) 特許出願

該当なし

(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)

- | |
|--|
| 1. [学会発表(基調講演)] 森 寛敏、有効フラグメント法を基とした第一原理分子シミュレーションと機械学習との連携による溶液物性予測法、分離技術会年会、S1-1、名古屋、2019年、5月 |
| 2. [学会発表] H. Mori, Materials Informatics toward Green Cooling., 7th Asia-Oceania Conference on Green and Sustainable Chemistry (AOC7-GSC 2018), Singapore, Nov., 2018. |

3. [学会発表・受賞] N. Kuroki, H. Mori, Exploring ionic liquids for efficient CO₂ capture and storage by electronic structure calculation and machine learning., 7th Asia-Oceania Conference on Green and Sustainable Chemistry (AOC7-GSC 2018), Singapore, Nov., 2018., GREEN Chemistry Prize 受賞
4. [学会発表(招待講演)] H. Mori, Materials Informatics for Designing Functional Liquids, 8th International Symposium on Molecular Thermodynamics and Molecular Simulation. (MTMS 2018), IL07, Narashino, Japan, Aug., 2018.
5. [学会発表(招待講演)] H. Mori, Design on Functional Liquid base on *ab initio* Materials Informatics, International Congress on Pure & Applied Chemistry (ICPAC 2018), PCC29, Siem Reap, Cambodia, Mar., 2018.

人工知能技術＋物質構造解析＝夢の社会の実現

溝口 照康（東京大学 生産技術研究所・教授）

研究課題名：「情報科学手法を利用した界面の構造機能相関の解明」 研究期間：2016.9～2020.3

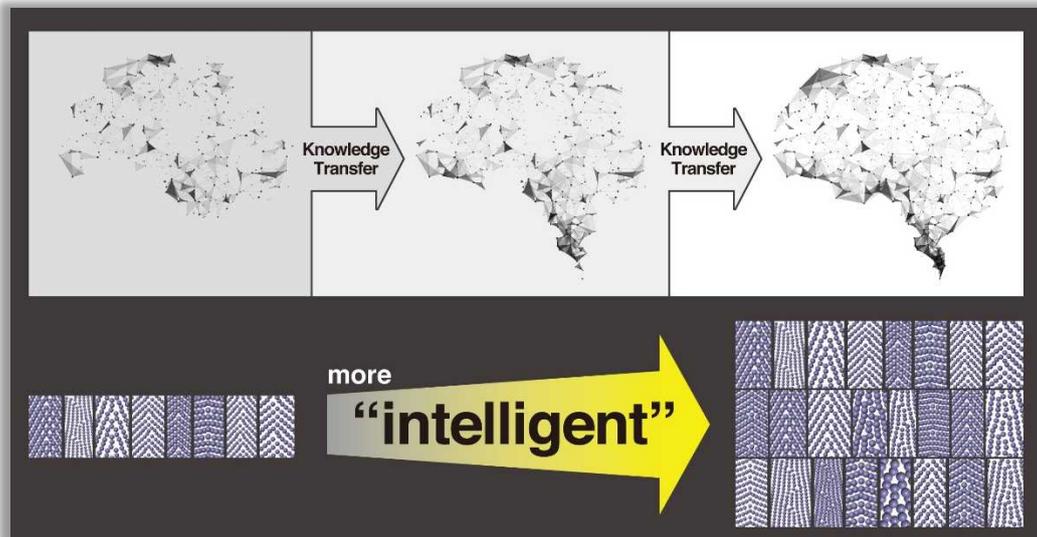


図1 人工知能が「繰り返し学習すること」で計算コストを1/3600に削減

物質内部に無数に存在する「界面」の構造は、電池や触媒など多くの機能に決定的な役割を果たしています。その複雑な構造を明らかにするためには膨大な計算量が必要でした。本研究により、転移学習という技術を使って人工知能が繰り返し学習することで、界面構造決定のための計算コストを約1/3600まで削減することに成功しました。

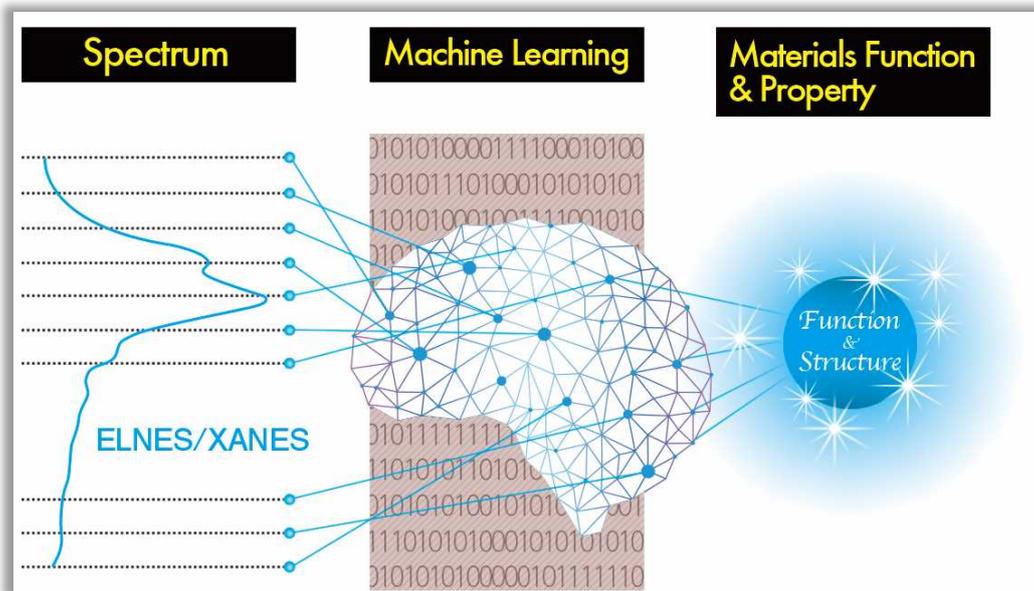


図2 理論計算や専門知識を用いずに人工知能がスペクトルから物質の構造や機能を直接定量

物質開発の現場では、その構造を明らかにするためにスペクトルの計測が行われる。一方で、スペクトルから機能まで推測するには、複雑な理論計算や専門知識が必要だった。本研究では、人工知能技術を利用し、スペクトルから直接、物質の構造と機能を決定する新手法を開発した。

過電流から電気機器を保護するためのバリスタ素子や、2019年にノーベル賞を受賞したリチウムイオン電池で使われているイオン伝導体など、世の中にある物質の多くは「界面」と呼ばれる領域を多数含有しております。界面は特異な特性をしめすため、界面の構造を解析することは材料研究において最も重要なテーマとされてきました。

一方で、複雑な界面の構造を決定するためには膨大な量の計算、もしくは複雑な実験を行う必要があり、界面の研究は困難と考えられてきました。

本研究は人工知能技術を利用して、物質の界面構造を解析する手法を加速することを目的としました。

本研究により、人工知能技術の転移学習とベイズ最適化を組み合わせ、約3600倍も高速に界面構造を決定する手法を開発しました。

さらに、人間の脳を模したニューラルネットワークを利用することで、スペクトルデータから物質情報を直接決定するための手法も開発し、界面構造の計算だけでなく、実験的な解析にも有効な手法を開発することができました。

現在工業的に使用されている物質の多くにおいて界面は重要な役割を果たしており、物質開発のスピードを上げるためには、界面の構造を決定し、その機能を理解することが不可欠です。今回開発した手法を利用することで、界面の構造をより効率的に決定することができ、物質の開発スピードが加速されることが期待されます。

物質開発のスピードが加速されれば、長期間充電不要な携帯電話や電気自動車をつくることができたり、環境負荷のほとんどない発電システムを実現できたりするかもしれません。今後は、人工知能と物質構造解析を組み合わせた研究をさらに発展させ、そんな夢の社会をつくることに貢献していきたいと思っております。

>>参考情報

➤ 論文

1. H. Oda, S. Kiyohara, K. Tsuda, and T. Mizoguchi, "Transfer Learning to Accelerate Interface Structure Searches", J. Phys. Soc. Jpn, 86 (2017) 123601-1-4.
2. S. Kikuchi, H. Oda, S. Kiyohara, and T. Mizoguchi, "Bayesian optimization for efficient determination of metal oxide grain boundary structures", Physica B, 532 (2018) 24-28.
3. S. Kiyohara and T. Mizoguchi, "Searching the stable segregation configuration at the grain boundary by a Monte Carlo tree search", J. Chem. Phys., 148 (2018) 241741-1-6
4. S. Kiyohara, T. Miyata, K. Tsuda, and T. Mizoguchi, "Data-driven approach for the prediction and interpretation of core-electron loss spectroscopy", Scientific Reports, 8 (2018) 13548-1-12
5. S. Kiyohara, M. Tsubaki, Kunyen Liao, and T.

Mizoguchi, "Quantitative estimation of properties from core-loss spectrum via neural network", J. Phys.: Materials, 2 (2019) 024003-1-9

6. H. Oda, S. Kiyohara and T. Mizoguchi, "Machine learning for structure determination and investigating the structure-property relationships of interfaces", J. Phys.: Materials, 2 (2019) 034005-1-8
7. T. Mizoguchi and S. Kiyohara, "Machine learning approaches for ELNES/XANES", Microscopy, in press. doi.org/10.1093/jmicro/dfz109 [Invited review]

➤ 本執筆

1. "Nanoinformatics", Springer (2018)
Chapter 8, "Atomic-Scale Nanostructures by Advanced Electron Microscopy and Informatics" 執筆担当
2. "Machine Learning in Chemistry: The Impact of Artificial Intelligence", Royal Society of Chemistry, (2020) 印刷中, (Hugh Cartwright edited)
Chapter "Machine learning for core-loss spectroscopy" 執筆担当

➤ 受賞

1. 功績賞, 日本金属学会, 2017年3月
2. 風戸賞, 風戸研究奨励会, 2017年3月

➤ プレスリリース

1. 2019年3月26日「理論計算や専門知識いらず！人工知能がスペクトルから物質の構造や機能を直接定量 ～物質開発の加速に期待～」
記事掲載：日刊工業新聞, 日経産業新聞等
2. 2018年9月7日「人工知能が専門家の約2万倍の速さでスペクトルを解釈」
記事掲載：日刊工業新聞, 鉄鋼新聞等
3. 2017年12月15日「世界初、液体中の原子1つ1つの運動を観察！ ～高性能電池や溶媒の開発、液体中の現象解明に革新～」
記事掲載：日刊工業新聞, 化学工業日報, 鉄鋼新聞等
4. 2017年12月12日「電子顕微鏡で気体分子の挙動や特性に迫る」
オンライン掲載：オプトロニクス, 日経オンライン, 等
5. 2017年11月15日「人工知能が「繰り返し成長すること」で計算コストを1/3600に削減 ～界面構造を高速に決定し、高性能な物質開発を加速～」
記事掲載：日刊工業新聞, 化学工業日報等

