九州大学情報基盤研究開発センター

民間利用成果報告書

提出日： 2020 年　10 月　30 日

利用課題名		有機二光子吸収材料計算に向けた機械学習モデルの構築				
課題責任者	企業名	パナソニック株式会社 印				
	フリガナ 氏名	四橋　聡史 印	部署名	テクノロジー本部	職位	主幹研究員
	連絡先	住　所：〒571 － 8508 大阪府門真市大字門真 1006 番地 E-mail： yotsuhashi.s@jp.panasonic.com		TEL 080-2278-4543 FAX		
利用期間		2020 年　4 月　1 日　～　2020 年　9 月 30 日				
利用成果公開延期希望の有無		■ 即時公開　・　□ 公開延期（成果公開予定：　　年　　月）				

※利用成果報告書は原則公開ですが、課題終了後最大 2 年間公開を延期することが可能です。

- 本様式の変更はできません。
- 補足資料を付加することは可能です。

受付番号	2	受付日 2 年 11 月 6 日	受付印	

■利用計画全体の概略（申込書と同じ内容を記述してください）

1） 利用目的

次世代光ディスクへのさらなる大容量化、電力消費量低減に向けてディスク内部に３次元に情報を記録する体積記録ディスクが注目されており、その材料候補の一つとして非線形光学応答を示す有機分子が提案されている。非線形光学応答を示すには一光子ではなく二光子で吸収励起される必要がある。この様な二光子吸収励起現象を詳細に検討し、定量的に信頼できる精度で予測するためには、高い近似レベル（TD-DFT）の二光子吸収断面積計算が有効である。しかしながら、TD-DFT に基づく二光子吸収断面積計算では、複雑かつ多くの励起状態を考慮した計算が必要であり、その計算負荷は非常に高い。そのため実験結果の解析に用いることは可能であるものの、計算主導による大量のスクリーニングを実現するには計算負荷が課題となる。そこで本研究では、TD-DFT に基づく二光子吸収断面積計算結果をより負荷の低い計算（TD-DFTB: TD-DFT の約 1/1000、DFT: TD-DFT の約 1/100）から予測するための機械学習モデルの構築を行う。そのために必要な学習データを用意するために、ITO 専有ノードを活用し約 4.5 万件の分子セットを対象に TD-DFT および TD-DFTB 計算を網羅的に実施する。実施した計算結果を用いて機械学習モデルを構築し、全体の計算時間を 1/10 以下に削減することを目標とする。

2） 利用意義

二光子吸収断面積という複雑で計算負荷の高い物理量に対して、負荷の低い TD-DFTB あるいは DFT の計算結果からこれを導く機械学習モデルを構築することにより、所望の特性を持つ材料の探索を飛躍的に広げることが可能となる。

化学物質の数は登録されているものだけで約２億種類と言われている。未知の材料の可能性も含めるとあらにその数は膨大なものになる。これらの中から所望の特性を持つ材料をすべて計算で求めるのは大変計算コストのかかる仕事となり、現時点での計算コストと計算リソースでは現実的な方法ではない。本研究の成果により、数万件の低コスト計算から得られる機械学習モデルで計算負荷の削減を実現できると、この膨大な化学物質の空間からこれまでにない材料を探索できる可能性が大いに広がる。

3） 必要性

機械学習モデルの構築には一度は数万件の高負荷な計算（TD-DFT）を実行する必要がある。このデータを計画的なスケジュールで半年の短期間でそろえるために、常時使用可能な大規模な計算リソースが必要となる。

■成果の概要

1） 本課題で得られた具体的な成果

TD-DFT レベルの二光子吸収断面積を予測する機械学習モデルを構築するために必要なデータを収集するため、まず、分子構造をドナー、リンカー、アクセプターの３つのグループのフラグメントのＳＭＩＬＥＳの組み合わせからなる数万件の二光子吸収材料分子を作成した。これらの分子に対し、目的変数である二光子吸収断面積の値を得るために TD-DFT 計算を行い、また説明変数を得るためには計算コストが低い励起状態計算(TD-DFTB)を実施し、得られたデータから機械学習モデルを構築した。さらに、より低コストで計算できる基底状態の量を説明変数として利用できる可能性を模索するために、基底状態計算(xTB)も行った。得られた機械学習モデルを比較検討した結果、説明変数を基底状態計算で得られる量にした場合でも高精度な予測が可能であることが明らかになった。この予測器を活用することで、二光子吸収断面積を計算するために必要なコスト(計算時間)を従来の TD-DFT 計算と比べて約 1/1000 に削減することに成功した。

2） 社会・経済への波及効果の見通し

昨今のビッグデータ時代の到来に伴い、急増する大容量データを長期的にかつ安全に保存するためのストレージ開発が急務となっており、二光子吸収材料を用いた高記録密度容量の光ディスク・アーカイバに対する期待が高まっている。本研究では、実験はもちろん計算でも評価に高コストを要した二光子吸収特性を低コストで予測できる機械学習モデルを構築し、膨大な候補材料分子に対する特性評価を可能にした。本成果により、計算及びインフォマティクスの主導によって、今後の二光子吸収材料開発を加速することができる。実際に本アプローチによって開発した材料を用いた次世代光ディスク・アーカイバが商品化された暁には、通信インフラを支えるストレージの提供という形で社会に貢献できるため、大きな波及効果も期待できる。

3） その他

特になし。

■利用アンケートにご協力ください

1) 利用に関して有益であった事項

ハードウェアの面において一定数の計算機を占有利用させて頂けた点は、機械学習モデルの教師データとして必要な膨大な分子数のデータベース作成を限られた期間内で行う上で非常に大きなメリットであった。特にTDDFT等の比較的計算コストが高い手法を用いた量子力学計算において、高エネルギーな励起状態を介した二光子吸収特性を理論計算・実用予測しようとすると使用メモリ量などの大きな計算機性能が要求される。また機械学習モデルを構築する上でどの程度の高励起状態数までを考慮すべきか、計算精度と計算コストの間の実用的なバランスをとるには試行錯誤も必要であるが、そのためにも大きな計算機資源が要求される。 これらの問題を解決する上で、高性能な計算機を常に複数利用できた状況は、並列で進む各分子の計算の実行状況の管理においても非常に有益であった。 またソフトウェアの面において量子力学ソフトウェアのインストールに必要な環境構築を行う上で、モジュール環境やコンパイラも充実していた点は非常にメリットが大きく、gaussian等の既存インストールされた量子力学ソフトウェアの充実具合にも非常に助けられた。ストレージ容量についても膨大な数の分子についての計算結果を一時的に保存しておく上で充分なサイズが用意されており、今回の膨大な数の分子の中で高い二光子吸収特性を示す候補分子を機械学習モデルによって予測する上で、滞りなく業務を進める事ができ非常に有益であった。

2) 利用に関して生じた問題点など

現状、特になし。

3) ユーザーサポートとして必要と考えられることについて

現状、特になし。

4) 施設利用に係る感想・改善を希望することについて

現状、特になし。

5) 本事業で得られた成果や公表する予定の成果があれば以降に記述をお願いします

現状、特になし。