

グラフニューラルネットワークと生成モデルを用いた 非晶質系動力学予測システム開発

芝 隼人 (兵庫県立大学)

概要

近年、深層学習が飛躍的発展したことにより、複雑な計算を経て得られるシミュレーション結果を高精度に予測することができるようになった。本課題では、課題代表者らがこれまで開発し最高予測精度を示したグラフニューラルネットワーク“BOnd TArgeting Network(BOTAN)”を用いて、ガラス物理学の専門家と高性能計算に長い経験を持つ研究者を揃えた体制を活用する形で、ガラスの動力学予測への適用拡大を図ることを目的とした。今年度は特に、これまでにガラスの機械学習・深層学習について中心的に研究を展開してきた研究者との共同研究により、BOTAN 最新モデルの予測性能を含めたさまざまな側面を定量的に比較するベンチマーキングを実施した。これらモデルの現状についてのレビューを執筆・公開したとともに、それぞれのモデルの予測特性をデータセットにまとめ新たに公開している。物理特性を考慮することによるモデルの性能向上そのものは一巡したと思われる現状において、今回の学習性能ベンチマーキングはガラスの機械学習の利活用の重要なガイドラインとなる。

1 共同研究に関する情報

1.1 共同研究を実施した拠点名

- 東京大学 情報基盤センター
- mdx

1.2 課題分野

- 大規模計算科学課題分野

1.3 共同研究分野 (HPCI 資源利用課題のみ)

- 超大規模数値計算系応用分野

1.4 参加研究者の役割分担

- 芝 隼人 (兵庫県立大学)
研究取りまとめ・全般の実行
- 下川辺 隆史 (東京大学)
データ駆動計算手法開発・高性能 GPU 利

用支援

- 川崎 猛史 (名古屋大学)
シミュレーション・物理モデリング
- 華井 雅俊 (東京大学)
深層学習モデル開発
- 別所 秀将 (名古屋大学)
シミュレーション・物理モデリング
- 西垣内 大祐 (名古屋大学)
シミュレーション・物理モデリング

2 研究の目的と意義

近年、深層学習モデルは、数値シミュレーションのアルゴリズムの代用としての役割を果たす場面が見られるようになってきた。例えば、量子力学レベルで電子の振る舞いを考慮に

入れたシミュレーション結果を学習させたモデルを、古典分子動力学シミュレーションの力場を推論に使用する機械学習分子動力学は、高精度かつ大規模な原子分子集合体のシミュレーションに道を拓いた、最も顕著な成功例の一つに数えられる。

その一方で、物理系の時間発展を扱うシミュレーションに対する深層学習・人工知能の利用は、これからの課題として控えている研究対象と言える。科学技術シミュレーションでは多くの場合、支配方程式が時間および空間座標を変数にもつ微分方程式で与えられ、それを離散化、逐次的に多数回の積分を行っていくことで数値的に解かれる。これには非常に大きな演算量のみならず時間を要することになる。半導体集積回路の微細化限界によるムーアの法則の終焉が控える中、シミュレーション結果をデータとして有効に扱った省電力アルゴリズム開発に、大きな期待が持たれる。

本課題では、物理シミュレーションの中でも、最も長時間にわたるダイナミックレンジでの現象を示す「ガラス（非晶質固体）」の分子動力学予測を取り上げた。ガラスとは「液体を急激に冷やして固まった物質」を指しており、我々の日常的な時間スケールで見れば固体のように見えるが、極めて長時間ではゆっくりと構造が移り変わり（構造緩和）する物質である。さらに、典型的な結晶固体とは異なって原子分子の配置が乱雑であり、構造に一見したところ特徴はなく液体の性質を残している、と言える。ガラスに普遍的に見られる特徴として、ある乱雑な構造を持ったガラスを長時間で見ると、動きやすい領域と動きにくい領域とはモザイク状にコントラストが広がっていく現象が挙げられる。この物理的起源が完全に解明されるには至っていないが、長年の研究を通じて、初期の（乱雑）構造が運動しやすい領域の空間的分布が決定づ

けているのではないかと、という説がある。すなわち、乱雑で一見特徴を抽出することが不可能と思われる原子分子の配置情報そのものに、どの原子分子（たち）が将来動きやすいのかの情報が埋め込まれているという考えが根強く支持されているのである。

上述の理由により、ガラス物質の理解そして活用の際に、「（乱雑な）構造からの分子動力学をどのように予測できるのか」ことは一番大きな問いであり続けてきた。これに対して、隣接配向秩序をもとに構成された構造秩序パラメータの適用や、準安定粒子配置に対する基準振動解析による低周波振動モードの計算など、構造と分子動力学との関係を抽出する試みが多く行われてきたが、決定的な方法は長く見出されていない状況と言える [例えば次のレビュー論文を参照：H. Tanaka *et al.*, *Nat. Rev. Phys.* **1**, 333 (2019)]。この状況の中で、近年、機械学習・深層学習手法による分子動力学分布の予測がこの問題に対して非常に有効であることが認知され、この10年ほどでさまざまな手法が提示されている。とりわけ、有限温度環境下においてガラスを構成しているそれぞれの粒子がどの程度、構造緩和に寄与する長時間運動を示すのかを定量的に予測するタスクにおいて、グラフニューラルネットワーク（GNN）を用いた教師付き学習が他の（学習モデルを含む）手法を凌駕して最高の性能をもつことが Google DeepMind のグループによる研究 [V. Bapst *et al.*, *Nature Physics* **16**, 448 (2020)] により示され、大きな関心を呼んだ。

3 当拠点公募型研究として実施した意義

本課題は東大情報基盤センターの芝と下川辺により開発された GNN モデル、“BOnd TArgeting Network” (BOTAN) を発展的に展開

させることを意図して実施したものである。BOTAN を含む GNN モデルや、物理的忠実性の高い MLP モデル、解釈可能モデルなどさまざまな提案されている中、種類の学習モデルの開発には主に物理系・材料科学系の研究者が参入しているものの、特に国外ではデータ科学・高性能計算の専門家が本格的にテーマに参入している状況ではないと言える。GNN による学習には多くの GPU 時間が要求され、学習環境（計算機環境）に則した最新モデルの開発・利活用が重要である。

今回課題では BOTAN の中心的開発者の芝、GPU コンピューティングの専門家である下川辺、ガラス物理を長年専門とする川崎、そしてグラフニューラルネットワークや生成モデルに最新の知見を持った華井らをメンバーとした学際的体制によって、高い次元で課題に取り組むことができたと考える。

4 前年度までに得られた研究成果の概要

本課題は芝が研究代表者として実施する 2023 年度の新規課題であるが、研究内容の一部は 2022 年度 JHPCN 課題「時空間発展するシミュレーションを予測する代理モデルの開発」（課題番号 jh220052、代表：下川辺隆史・東京大学情報基盤センター）から継承している。この課題の実施期間中、芝と下川辺は、前述の Google DeepMind が 2020 年論文において提唱した GNN モデルを、粒子間の長時間相対運動を更なる学習対象として取り入れた GNN モデル “BOnd TArgeting Network (BOTAN)” を提案、粒子運動予測の精度を大幅に改良することに成功した。[H. Shiba, M. Hanai, T. Shimokawabe, and T. Suzumura, *J. Chem. Phys.* **158**, 184501/1-11 (2023)]。

5 今年度の研究成果の詳細

本課題で設定していた目標のうち、「学習モデルの高性能化」の検討については進展があった。2022 年 11 月にフランス・パリにて CNRS・高等師範学校が主催して開催されたワークショップ “Machine learning glassy dynamics” に開催し、ガラス物理学分野を牽引する研究者らと数多くの若手が集まり、ガラスへの機械学習の適用について議論を実施した（なお、2024 年 5 月現在、このワークショップの招待講演が YouTube のビデオチャンネルにて閲覧可能である。URL:<https://www.youtube.com/playlist?list=PLLusiyVUCDJnprBuih5yDgToIzpsIdCi4>）。ワークショップに参加したメンバーのうち、特に関連研究において深いコミットメントをしている一部の研究者が組織され、分野の現状と将来の方向性を示すためのレビュー記事を執筆することになった。

このレビュー記事の作成においては、(i) ガラスにおいて重要な特徴的構造の抽出、動力学的予測などへの機械学習利用の現状のまとめ、および、アルゴリズムやサンプリングなど、将来の方向性についての議論、を行うことを目標としたほか、(ii) 将来の更なる研究に向けた、各種の機械学習モデルの比較検討を共同で実施する、ことを最初に合意した上で作業を開始した。今回課題代表者が主に (ii) に関与した。この点について、次の成果が得られた。

1. 3次元 Kob-Andersen ガラス形成液体に対して、分野で提案された主要な機械学習モデルの持つ予測性能、転移学習性能などをベンチマークして取りまとめた。その際、以前の論文で Google DeepMind のグループが提示したデー

タセットにはいくつかの不便な点があるため、今回の学習性能ベンチマークにおいては、BOTAN 論文とともに 芝・下川辺らが公開に供したシミュレーションデータを基準データセットとした。予測結果は新たなデータセットとして取りまとめ、一般の研究者の検証や利用に役立ててもらえるように Zenodo リポジトリ上に公開した（研究業績一覧 [12] を参照）。

2. 前項の作業の過程で、「固有構造」(Inherent Structure) と通称される状態を学習の入力として使用することの重要性が浮き彫りになってきた。通常、有限温度でシミュレーションされているガラスにおいては、熱ゆらぎの影響があるため各粒子の運動中心の周辺で熱運動を行う。しかし、ガラスの遅い緩和に伴う構造遷移は、座標についての多変数関数であるエネルギー地形における準安定状態を徐々に降っていくものであると考えられ、機械学習・深層学習モデルにとっては、むしろ固有構造の方が特徴を「捉えやすい」入力である可能性が高い。参考まで、BOTAN、およびその前身モデルである Google DeepMind が提示した GNN モデルの予測性能を図 1 に示した。レビュー論文のプレプリントとは [1] とは別に、国内会議 [8] のために用意した学習性能のプロットである。固有構造を入力に取ることで、以前の GNN モデルの予測性能（赤線）も、BOTAN と同水準に向上していることが分かる。

なお、熱的な粒子配置の状態から（最急降下法などによる準安定構造へのクエンチを行わずに）、直接粒子の運動度を予測するタスクにおいては、BOTAN は他の（教師つき学習の）モデルを凌駕してトップの予測性能に位置している。距離変化を学習

対象としてグラフの辺上にアサインすることで、隣接粒子間で生じる弾性歪みの度合いを学習する BOTAN 特有の機構が、学習モデルの内部で背後に隠された固有構造を推定する機能をもたらしっていると推測される。

3. 従来、ガラスの運動の予測については 3 次元のガラス形成液体に対する予測が実施されてきたが、今回、2 次元 Kob-Andersen 液体のシミュレーション結果の提供を Gerhard Jung 博士（モンペリエ大学 [当時]）から受け、各機械学習モデルによる検証を行った。2 次元の場合にも BOTAN が良好な予測性能を発揮することが示された。
4. 転移学習に対するテストとして、ある温度で訓練されたモデルが、他の温度のガラス形成液体の動力学にどの程度有効であるかをベンチマークした。この結果、BOTAN を含む多層パーセプトロンを構成要素とした学習手法に、一定の優位性が認められた。

成果となるレビュー論文の原稿は現在査読中であるが、すでに投稿初版原稿をプレプリントとして公開している [arXiv:2311.14752 (2023)] ので、本報告書においては詳細なデータなどは割愛させていただいた。また、このレビュー論文原稿では生成モデルを用いたガラスのサンプリング手法開発の現状や、GNN モデルや Transformer など最新モデルを使用した分野の発展方向の提示を行っているのので、合わせてご参照いただきたい。

6 今年度の進捗状況と今後の展望

本課題では、外力剪断のもとにあるガラス物質のシミュレーションに対する予測も研究

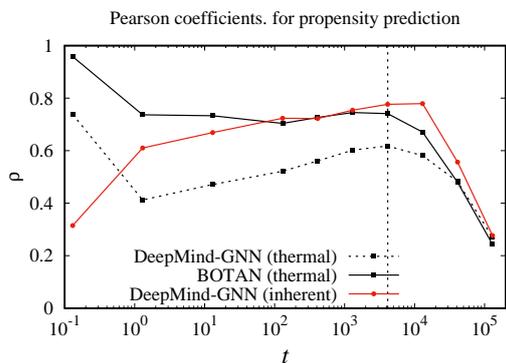


図1 図1：3次元 Kob-Andersen ガラス形成液体のシミュレーション結果から、教師つき学習によってガラス液体を構成する粒子が示す運動度（動力学）を学習した GNN モデル（Google DeepMind 提案の GNN, および BOTAN）の予測性能をピアソン係数 (ρ) で評価し、プロットした。横軸 t は、どの予測対象としている時間幅である。縦の点線はガラスが構造緩和し初期の構造が失われる典型時間であるところの「 α -緩和時間」を示す。

課題として取り上げた。周期境界条件を剪断系に拡張的に実装した Lees-Edwards 境界条件のもとで、逐次、最急降下により得られた固有構造に対してゼロ温度で最急降下をかけるタイプのシミュレーションを実装し、3次元 Kob-Andersen ガラス形成液体に対して適用、データセットを拡張するところまでは完了している。しかし、前述の研究が進んでいく中で、固有構造による予測に対して BOTAN がこれまでの GNN に対して持っている有効性は限られることをはじめとして、想定していない内容が判明してきた。

来年度、GNN モデルを使用あるいは開発する際、任意のガラスの配置に対して固有構造を計算する手続きを組み込んだワークフローも作成した上で、改めて剪断系に見られる特有の動力学的性質の予測ができるようにモデルの改良を進めていくことで、剪断系に対して有効な

GNN モデルを具体的な形で世に問えるように努めたい。

また、今年度課題では当初「生成モデルによる動力学サンプリング手法の開発」も目標に掲げていたが、生成モデルを適用する研究の状況変化が激しいこともあり、現時点では具体的計算に至っていない。拡散モデルの他にも、確率分布の発展を扱うための最新の生成モデルが提案が続いており、来年度には最新の知見を導入したモデルの導入にこぎ着けたいと考えている。

なお、2023年11月には JHPCN および代表者が所属する兵庫県立大学データ計算科学連携センターの共催という形で、研究会「液体・ガラスへのデータ駆動アプローチ - グラフニューラルネットワークとその周辺 -」を行った。非平衡物理、材料計算科学の分野の研究者による講演、およびデータ科学や高性能計算分野の研究者など合計 11 件を実施し、グラフニューラルネットワークが科学技術シミュレーションにおいて使用されている最新の状況を見ることができ、今後の研究には大いに役立つこととなった。会場利用料の補助をいただいた JHPCN に深く感謝する。

7 研究業績一覧（発表予定も含む）

学術論文（査読あり）

今年度成果となる論文について現在投稿中であるものの、査読が継続している状態である。

国際会議プロシーディングス（査読あり）

該当なし

国際会議発表（査読なし）

- [1] Hayato Shiba, Masatoshi Hanai, Toyotaro Suzumura, Takeshi Kawasaki,

- and Takashi Shimokawabe, “Deep learning glasses: how well neural network models can capture the dynamics heterogeneity?”, 9th International Discussion Meeting on Relaxations in Complex Systems (9IDMRCS), 幕張メッセ (千葉市美浜区), 2023年8月13日 [招待講演]
- [2] Hayato Shiba, Masatoshi Hanai, Toyotaro Suzumura, and Takashi Shimokawabe, “Enhancing the predictive power of static structure in glassy systems by machine learning relative motion”, 7th International Soft Matter Conference, 大阪国際会議場 (大阪市北区), 2023年9月6日 [一般講演]
- [3] Hayato Shiba, Masatoshi Hanai, Toyotaro Suzumura, Takashi Shimokawabe, “Enhancing the predictive power of static structure in glassy systems by machine learning relative motion”, 34th IUPAP Conference on Computational Physics (CCP2023), 神戸国際会議場 (神戸市中央区), 2023年8月6日 [ポスター講演]
- [4] Hayato Shiba, Masatoshi Hanai, Toyotaro Suzumura, Takashi Shimokawabe, “Enhancing the predictive power of static structure in glassy systems by machine learning relative motion”, 28th International Conference on Statistical Physics (STATPHYS28), 東京大学本郷キャンパス (東京都文京区), 2023年8月8日 [ポスター講演]
- 国内会議発表 (査読なし)**
- [5] 芝 隼人, 「ガラスの深層学習 – シミュレーションとデータサイエンスの新たな出会い、そして可能性」, KOBE HPC サマースクール (初級) 2023, 神戸大学 (神戸市中央区), 2023年9月19日 [招待講演]
- [6] 芝 隼人, 「ガラスの長時間分子動力学シミュレーションの深層学習」, 第30回 AT 研究会オープンアカデミックセッション (ATOS30), 山梨大学 甲府東キャンパス (山梨県甲府市), 2023年10月27日
- [7] 芝 隼人, 「ガラス動力学の深層学習 – 現在とその後?」液体・ガラスへのデータ駆動アプローチ - グラフニューラルネットワークとその周辺 -, ニチイ学館 神戸ポートアイランドセンター, 神戸市中央区, 2023年11月28日 [一般講演]
- [8] 芝 隼人・下川辺 隆史, 「ガラス動力学の深層学習における準安定構造の使用の影響」第37回分子シミュレーション討論会, 福井県民ホール AOSSA (福井県福井市), 2023年12月5日 [ポスター講演]
- [9] Hayato Shiba, “Deep learning of simulated glassy dynamics”, 45th ASE Seminar: International Workshop on ”Integration of Simulation/Data/Learning and Beyond”, 東京大学情報基盤センター (千葉県柏市、オンライン講演での参加), 2023年11月29日
- [10] 芝 隼人, 「ガラス動力学の深層学習 – 乱雑構造の中に隠れた未来を掘り起こす」, 令和5年名古屋大学宇宙地球環境研究所研究集会 STE シミュレーション研究会: 計算科学とデータ科学の融合に向けて, 神戸大学六甲台キャンパス (神戸市灘区), 2023年12月20日 [招待講演]

- [11] 芝 隼人, 「ガラスとグラフニューラルネットワーク - 究極の長時間ダイナミクスへのデータ駆動科学による挑戦」, Supercomputing Japan 2024, タワーホール船堀 (東京都江東区), 2024 年 3 月 12 日 [招待講演]

公開したライブラリ等

- [12] GlassBench, Zenodo Dataset [doi: 10.5281/zenodo.10118191]
(Author: G.Jung(+), Data collectors: R. M. Alkemade(+), D. Coslovich(+), F. P. Landes(+), F. S. Pezzicoli(+), L. Filion(+), L. Berthier(+), G. Biroli(+), H. Shiba) , 2023 年 11 月 21 日公開

その他 (特許, プレス発表, 著書等)

該当なし