

jh250034

# データ科学計算フレームワーク ODAT-SE の高度化と高性能化

工藤周平（電気通信大学）

## 概要

ODAT-SE (旧称: 2DMAT) はデータ駆動科学のためのフレームワークであり、計算モデルによるシミュレーションと並列探索手法 (PAMC 法など) を組み合わせることで、実験データをよく説明する内部パラメータの推定を効率化する。本研究は ODAT-SE の並列性能向上や、自動チューニングによる探索手法の改善、2次元物質解析や核融合計算を含むアプリとの連携強化によって、HPC 環境を利用したデータ駆動科学計算の拡大を目指す。本報告書では、ODAT-SE の並列性能向上について新規通信アルゴリズムの概要と性能測定結果、2次元物質解析アプリについて GPU 向け実装の最適化手法とその性能測定結果、また核融合計算を含むアプリとの連携強化について協同トムソン散乱スペクトルのデータ解析への応用について説明する。また ODAT-SE 本体への改良への貢献活動についても説明する。

## 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

- 北海道大学 情報基盤センター
- 東京大学 情報基盤センター

### 1.2 課題分野

- 大規模計算科学課題分野

### 1.3 参加研究者の役割分担

- 工藤周平：研究統括 テーマ A 主担当
- 星建夫：研究副統括 テーマ B 主担当
- 小林達哉：核融合関連アプリ知識協力
- 長井元輝：テーマ C GPU 向け高速化担当
- 中野陽斗：テーマ B 担当
- 山本有作：テーマ B 担当

## 2 研究の目的と意義

近年、実験データ解析にデータ駆動科学の手法を適用することが重要となっている。とくに、物質構造などの観測が難しい内部パラメータを推定するために、計算モデルに基づく計算機シミュレーション (順問題計算) による結果と、実験で得られたデータとを突き合わせ、逆問題計算を行うことが多くの実験データ解析に共通する問題となっている。(図 1)

Open Data Analysis Tool for Science and Engineering (ODAT-SE, 旧称 2DMAT) は副代表 (星) らが開発に参加している、データ駆動科学のためのフレームワークである。ODAT-SE はユーザーの持つ実験データとシミュレーションアプリを容易に結合でき、多数の統計的探索手法による逆問題計算を可能とする。

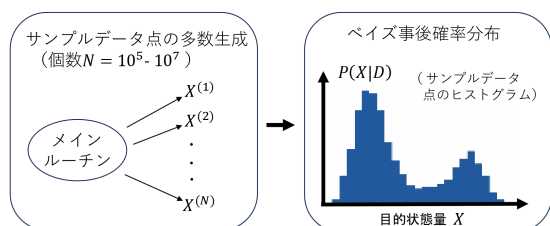


図1 ODAT-SE/PAMC 法によるベイズ推定をもとにした逆問題計算によるパラメータ推定手法の概念図

また ODAT-SE は高い並列性を持つ MCMC 法の一つである Population Annealing Monte Carlo (PAMC) 法を提供しており、幅広いアプリへ HPC 資源を活用することが可能である。ODAT-SE はすでに、2次元物質解析の分野において多くの成果を上げており、今後の発展や異分野への進展が望まれている。とくに副代表は、2024年12月より、核融合分野でのムーンショットプロジェクト（ムーンショット目標10プロジェクト「超次元状態エンジニアリングによる未来予測型デジタルシステム」、PM：星）を行っており、当分野への活用が検討されている。

そこで本研究の目的は、この ODAT-SE について、

(1) HPC 環境における性能検証・改善手法の検討、

(2) 探索手法の数理的・実装的の改良、

(3) ODAT-SE 向けアプリの展開およびアプリの高性能・高機能化、

を行うことで、より多くのアプリ・実験に対しデータ駆動科学の手法を展開し、HPC 資源を活用した実験データ解析の新たな可能性を拓くものである。なお、本研究は HPC 手法発展が主目的であり、実問題でより大きな計算資源を必要とする問題は、並行する富岳プロジェクトを活用している (hp240304 など)。

以上の目的を基に、2025 年度においては以下の実現を計画した：

■**テーマ A ODAT-SE の大規模並列化の実現:** ODAT-SE の大規模並列環境への対応は発展途上であり、現状、高並列環境においては、(Python 言語由来の) 意図しないファイル IO の発生や、ジョブ記述の煩雑さが問題となっている。これらを解決し、真に HPC に適したフレームワークとすることが (A) の課題である。

■**テーマ B ODAT-SE の PAMC 法のさらなる高度化と検証:** 2024 年度の成果を発展させ、自動チューニング技術による PAMC 法のパラメータ (レプリカ数、不確かさ強度パラメータ変化率、リサンプリング操作発動条件、などの) 最適化手法の改良・性能検証・自動化を行う。また、Hamiltonian Monte Carlo (HMC) 法などの近年の優れた手法を取り入れ、性能検証を行う。

■**テーマ C 二次元物質シミュレーションアプリ sim-trhepd-rheed (STR) のさらなる高速化と、新たなアプリの検討:** STR は大規模先端観測技術である「全反射高速陽電子回折 (TRHEPD)」のためのシミュレーションアプリであり、その高度化は貴重な実験データの解析のために重要である。2025 年度は STR のさらなる高速化に取り組む。またこの技術を基に、核融合プラズマ解析などのアプリに対する ODAT-SE の適用可能性を検討する。

### 3 当拠点公募型研究として実施した意義

本研究は、データ駆動科学フレームワーク ODAT-SE を軸に、計算・実験・データ科学の三分野にまたがる知見を集約するものである。研究代表者 (工藤) は大規模並列計算において成果を残しており、ODAT-SE とアプリの連携高度化や、ODAT-SE 自体の高性能環境への改

良について貢献が期待される。また副代表者は ODAT-SE の開発に携わっているだけでなく、物質科学の専門家として、ODAT-SE アプリの検証や実験家との連携を進めており、さらに近年は、ODAT-SE をベースに核融合科学計算へのデータ駆動科学手法の展開を進めている。このように、本研究は極めて幅広い適用可能性を持つ ODAT-SE を軸に、他分野にまたがる研究に貢献することを目的とする。このため本研究においては、単一のアプリや環境の利用ではなく、多様なアプリを対象に多数の環境を利用できる学際大規模情報基盤共同利用・共同研究拠点 公募型共同研究が適する。

#### 4 前年度までに得られた研究成果の概要

**■STR アプリ高スループット化実現** 本テーマでは研究代表らが提案した高速 STR 計算手法を基に、ODAT-SE・STR アプリ間の連携強化による高速化手法について検討した。とくに、PAMC 法の高並列性を活用し、複数解候補に対する順問題計算のバッチ型計算を行うことで、ODAT-SE・アプリ間通信のオーバーヘッドやアプリ起動コストを削減した。また、高並列性を活かした STR アプリの GPU 実装に取り組んでおり、十分な個数の順問題計算を並列に実行する場合、CPU では一計算あたり 0.12 秒かかる計算を、GPU 上で 3.3 ミリ秒で完了することが可能となっている。これは従来実装と比べてトータルで約 1 万倍の高速化に相当する。(図 2)

**■自動チューニング技術による ODAT-SE PAMC 法の改良とその検証** PAMC 法のパラメータ (レプリカ数、不確かさ強度パラメータ変化率、リサンプリング操作発動条件、などの) 最適化手法の手法開発とテスト計算を行ってきた。レプリカ数については、Gellman-Rubin

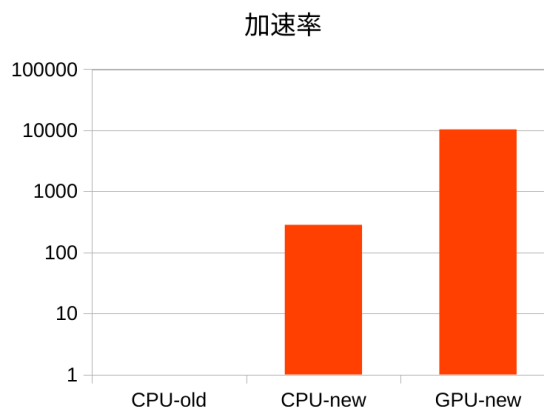


図 2 STR 高速化の結果. 右: CPU, GPU 環境における波数 23 の順問題計算 1 つあたりの計算時間の加速率 (対数グラフ). CPU は北大 Grand Chariot 環境 (1 CPU), GPU は東大 Wisteria-a (1GPU). CPU-old は推定時間.

型診断条件を PAMC に対して考案し、テスト計算で有用性を示した。不確かさ強度パラメータ変化率については、変化率が大きすぎると、正しい分布が得られない事例がみつき、対処方法の立案とテスト計算を行なった。リサンプリング操作発動条件については、少数レプリカが確率分布上の重みを寡占する事例がみつき、対処方法の立案とテスト計算を行なった。

#### 5 今年度の研究成果の詳細

本年度は、逆問題型データ解析ソフトウェア ODAT-SE を物質科学および核融合科学を含む幅広い科学分野に適用可能とすることを目的として、プログラム内部構造の改良を行った。具体的には、テーマ A, B に関連し、実験を模擬する順問題 (フォワード) モジュールの定義部分と、そのモジュールを前提として動作する探索アルゴリズム (並列ベイズ最適化や並列モンテカルロ法など) をコード上で完全に分離し、順問題モジュールの仕様を明確に策定した。この

改良により、順問題モジュールを ODAT-SE 本体から独立して開発することが可能となり、多様な実験系に対する柔軟かつ効率的なデータ解析が実現された。また、ODAT-SE の普及活動の一環として、2025 年 5 月 8 日に核融合科学研究所においてハイブリッド形式で「データ解析フレームワーク ODAT-SE ミーティング」を開催した (<https://indico.nifs.ac.jp/e/odatse-meeting-202505/>)。

またテーマ A, B に関連し、ODAT-SE における PAMC 法のリサンプリング計算における通信最適化による並列性能向上に取り組んだ。PAMC 法は多数の MCMC 計算 (レプリカ) が独立に実行される極めて並列性の高い構造を持つが、リサンプリング計算に同期やデータ通信が必要となり、並列計算におけるボトルネックとなる可能性がある。リサンプリング計算ではレプリカの重要度に応じて、乱数を用いて、レプリカを削除したり複製したりするため、データ移動パターンが不規則であり、単なる隣接通信や全体全通信では表せない特殊なデータ移動が発生する。そのため現状の ODAT-SE では全データを 1 ノードに集め、非並列に計算を行うためスケラビリティが悪い。そこで、1 次元最適輸送をもとにした準最適通信計画アルゴリズムを設計し、性能をテストした。本アルゴリズムは、1 次元ネットワーク構造を仮定し、レプリカが通信前に複製済みであることを想定して、データ量×ホップ数を最小とする通信計画を立てた上で、実際の通信においては複製前のレプリカを送信する。また送信するレプリカは複製数の大きいものを優先することで通信量を減らす。このアルゴリズムは、2 段階の最適化を行っているため全体最適ではなく、またより複雑な構造のネットワークにも対応していないが、従来実装と比べて通信量を大幅に減らすことができる。実際、Himmelblau 関数

の最小値を探索する小規模なテスト問題を北海道大学 情報基盤センター Grand Charriot2 の 64 MPI プロセス上で動作させたところ、従来 14.8 秒かかっていた通信時間を 8.63 秒にまで削減することに成功した。本アルゴリズムは次元の大きい問題に対して大規模計算を行う場合により効果的となると考えられる。

またテーマ C に関連し、STR アプリ (順問題計算) の GPU 高速化に取り組んだ。STR アプリの計算は、多数 (> 10,000) の小さいサイズ (< 100) の行列 ODE 初期値問題に帰着される。これは GPU が得意とするバッチ型行列計算となるが、単にバッチ版 BLAS ライブラリを使用すると、ODE 計算内で中間行列が多数回再利用される計算構造を十分に活用できず、メモリバンド幅ネックとなる。そこで我々は STR アプリ専用計算カーネルを作成し、中間行列を GPU のレジスタやローカルキャッシュ (shared memory) に配置したまま計算可能とすることで大幅な性能効率向上を実現した。また STR アプリにおいて数値発散を防ぐために ODE の反復ごとに必要な逆行列計算を Neumann 級数をもとにした近似逆行列計算に置き換えることで行列計算ライブラリの使用を排除し、ODE 中の Global memory へのアクセスを最小限とした。さらに、最近の GPU が持つ行列積計算ユニット (tensor core) を直接利用することで GPU の最大限の性能を発揮することを目指した。この結果、最も良い条件で、Wisteria-a 上の NVIDIA A100 では (tensor core 利用時の) 理論限界性能の 87%、Miyabi 上の NVIDIA GH200 では (tensor core 利用時の) 理論限界性能の 75% を超える性能比率に到達している。(図 3) これは、Wisteria-a と Miyabi が持つ最新鋭の 72 コア CPU と比べて、それぞれ最大 17 倍と 31 倍の性能向上になっており、専用カーネル設計が優れているこ

とを示している。一方、バッチ BLAS ベースの GPU 実装ではそれぞれ CPU の最大 1.5 倍と 3.1 倍の性能向上となっていた。本結果は STR アプリを利用した TRHEPD 結晶構造解析に直接役立つだけでなく、行列積計算ユニットを活用したシミュレーションアルゴリズムの高速化の指針となる例を示すものとなっていると考える。本成果については国際学会における発表を目標に予稿論文の投稿準備を行っている。

またテーマ C に関連して、核融合分野への応用としては、核融合科学研究所の境健太郎らによるプラズマ計測実験である協同トムソン散乱 (CTS) スペクトルのデータ解析に適用した。順問題モジュールでは、CTS スペクトル  $D$  を電子温度  $T_e$ 、電子密度  $n_e$ 、電子速度  $v_e$ 、イオン温度  $T_i$ 、イオン速度  $v_i$ 、イオン価数  $Z$  の 6 変数  $X \equiv (T_e, n_e, v_e, T_i, v_i, Z)$  からなる関数としてモデル化し、尤度関数  $P(D | X)$  を定義した上で、PAMC 法を用いてベイズ事後確率分布  $P(X | D)$  を 6 次元ヒストグラムとして算出した。さらに予備的計算を通じて、尤度関数の定義に計測特性を適切に反映させることが解析精度に重要であることが明らかとなった。一般に実験スペクトルと計算スペクトルの絶対値は一致しないため規格化が必要であるが、プラズマ計測では強いイオン信号と弱い電子信号が重畳して観測されるため、強い信号に起因するノイズが弱い信号の情報抽出を困難にする問題がある。そこで、電子寄与が支配的な波長領域とイオン寄与が支配的な波長領域とに分割し、それぞれに対して独立に規格化を行う手法を導入した結果、解析結果の信頼性が大きく向上した。このことから、計測の特性を取り入れた尤度関数の設計がデータ解析の信頼性向上に極めて重要であるという知見が得られ、今後の計測データ解析における重要な指針となると考えられる。本成果については国際学会で発表を行っ

ており、今後論文投稿を行う予定である。

## 6 進捗状況の自己評価と今後の展望

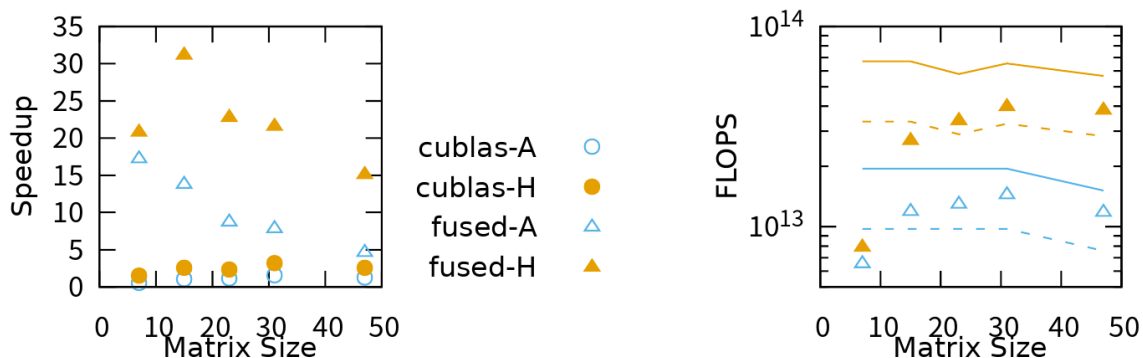
■**テーマ A** 当初想定していた、大規模並列計算環境での Python 実行における問題解決の点については、2DMAT から ODAT-SE への設計変更への対応や、環境依存問題解析の難しさから、大きな成果を挙げられなかった。一方で、ODAT-SE の機能拡張や PAMC 法の通信最適化などで代わりとなる成果を挙げられており、「真に HPC に適したフレームワークとする」という最終目標に対しては大きく貢献できたと考えられる。そのため十分な進捗があった。

特に PAMC 法の通信最適化において興味深い結果を得られたことから、今後はアルゴリズム開発や理論的なアルゴリズム解析に取り組む予定である。

■**テーマ B** 研究を担当する学生の指導計画の都合上、本年度にはまだ十分な成果を得られていないが、本テーマについては理論的研究が多くを占めており、今後精力的に研究を継続する予定である。

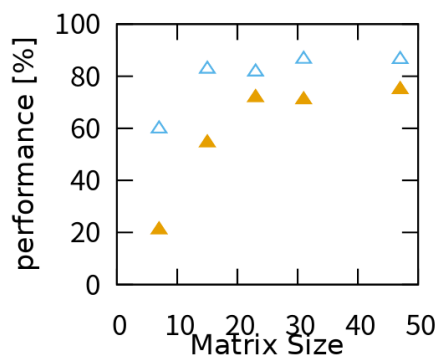
■**テーマ C** 当初の重要目標であった STR アプリの GPU 向け高速化では、最新 GPU の理論限界性能の 87% や 75% といった、極めて高度なレベルで目標を達成することができた。また本年度から追加の課題としていた核融合プラズマ解析などのアプリに対する ODAT-SE の適用可能性を検討においては、実際に CTS スペクトルのデータ解析への応用を行い、データ解析における計測特性の利用といった、実応用上重要な知見を得ることができた。また両成果について今後外部発表を行う予定である。

以上のとおり、テーマごとに達成度にばらつきがあるが、特に当初からの重要課題であったテーマ C について目覚ましい成果を上げることができたため、全体として十分な進捗があっ



(a) Relative speed to CPU

(b) Achieved performance



(c) relative performance to the ideal computation time

図3 Wisteria-a NVIDIA A100, Miyabi NVIDIA GH200 上での STR アプリ最適化結果. cublas-A, cublas-H は cuBLAS を用いた実装のそれぞれ A100, GH200 上での結果, fused-A, fused-H は同様に新規計算カーネルを用いた実装. (a) それぞれの環境での CPU 上での STR アプリ計算速度との性能比. (b) 絶対性能 (FLOPs 値). (c) 理論限界性能比.

たと評価する。またテーマ A や B についても今後の理論的研究につながる重要なステップを踏むことができたと評価する。