

jh220010

FMO プログラム ABINIT-MP の高速化と超大規模系への対応

望月 祐志 (立教大学)

概要

本課題は、jh210036-NAH から継続しており、「富岳」や「不老」 Type I 等の A64FX スーパーコンピュータ向けの ABINIT-MP プログラムの改良を主として研究開発活動を行っています。関連テーマとして、GPU (NVIDIA V100) を搭載する「不老」 Type II をプラットフォームにタンパク質構造推定アプリケーション AlphaFold2 の整備と量子計算シミュレータ cuQuantum の利活用を進めています。ABINIT-MP の高速化では、2020 年の版に比して 1.5~2 倍を達成しました (jh210036-NAH 終了時は 1.2~1.5 倍)。また、大規模系への対応では、液滴モデルでタンパク質から遠方の水分子群をクラスタリングする前処理スクリプトを開発し、新型コロナウイルスのスパイクタンパク質を水和状態で扱えるようになりました。AlphaFold2 に関しては、タンパク質複合体の構造予測が改良された v2.3.1 を 2023 年 4 月 5 日付で「不老」 Type II のライブラリとして公開しました。cuQuantum の利用では、4 成分の相対論的な量子化学計算によるスピン-軌道分裂エネルギーの高精度算定の量子シミュレーションに供しました。18 量子ビットの場合、CPU に比しての GPU による加速は 42.7 倍となりました。

1. 共同研究に関する情報

(1) 共同利用・共同研究を実施している拠点名

名古屋大学 情報基盤センター

(2) 課題分野 (該当するものを残す)

大規模計算科学課題分野

(3) 共同研究分野

超大規模数値計算系応用分野

(4) 参加研究者の役割分担

望月が全体の進捗を管理すると共に、テスト計算を手掛けています。また、学会発表などの情報発信も手掛けています。中野氏 (国立医薬品食品衛生研究所) と坂倉氏 (計算科学振興財団) は、望月と連携して ABINIT-MP プログラムの改良を進めています。渡邊氏 (HPC システムズ) は性能評価と版管理を担当しています。森脇氏 (東京大学) は AlphaFold2 の

整備、杉崎氏 (2023 年 4 月より慶應義塾大学)

は cuQuantum を用いた量子シミュレーションの実行を各々担っています。片桐氏 (名古屋大学) と大島氏 (2022 年 10 月から九州大学) は「不老」の責任者として計算機科学の立場からアドバイスと支援を行っています。土居氏 (立教大学) は、Python で大規模 FMO ジョブ実行のための支援ツールを作成しています。研究員 (奥脇氏) と大学院生 (満田氏、森下氏) は教員の指導を受けながら本課題にコミットしています。

2. 研究の目的と意義

本 jh220010 課題は jh210036-NAH 課題からの継続であり、(I) ABINIT-MP プログラムの A64FX スパコン向けの高速度化と大規模系への対応を「不老」 Type I をプラットフォームとして行っています。ABINIT-MP は、国産の FMO プログラムとして 20 年以上の歴史を有

し、「富岳」を頂点とする HPCI 拠点でライブラリプログラムとして提供されていることもあり、生物物理や医薬品関係の研究室・企業で一定の利用者を有してきています。従って、その改良・更新は公益性を持つものです。

2022 年度の活動では、新たに 2 つの副軸を立てました。(II)として、アミノ酸の FASTA 形式配列からタンパク質の立体構造を高精度の予測する AlphaFold2 (文献: Jumper et al., Nature 596 (2021) 583) の「不老」 Type II での整備と公開を加えました。GPU (NVIDIA V100) と SSD を備える、このシステムは最適なプラットフォームです。また、(III)として、量子コンピュータでの量子計算を GPU を使って高速にシミュレートできる cuQuantum (<https://developer.nvidia.com/cuquantum-sdk>) の同環境での整備と利活用も図りました。(II)は、(I)と分野的に重なるために共通の公益性があります。(III)は、ノイズの影響の無いシミュレータとして、将来のエラー補正付きの量子コンピュータ (FTQC) の実機に向けた方法論・ソフトウェア整備に対する利便性を提供するものです。

3. 当拠点の公募型研究として実施した意義

まず、(I)に関連して記します。2020 年度の試行運用段階の「富岳」での新型コロナウイルスの特別プロジェクトに参画し、多数の応用成果が得られたのですが、当時の ABINIT-MP Ver. 1 Rev. 22 の「限界」を認識しました。そのため、2021 年度の jh230036-NAH から、高性能計算分野の専門家である片桐氏、大島氏との共同作業を始めています。片桐らが管理・運営する「不老」 Type I は「富岳」に比べて混雑が少なく、CPU クロックも一定であるため、コード改良のテスト環境としても好適なプラットフォームです。

「不老」 Type II が(II)の AlphaFold2、(III)の cuQuantum の実行環境として好ましいのは既述の通りですし、なにより大島氏は

GPU 計算の専門家でもあります。

本課題の構成メンバーは、量子化学・計算化学 (望月、中野氏、坂倉氏、杉崎氏、土居氏)、バイオインフォマティクス (森脇氏)、計算機科学 (片桐氏、大島氏、渡邊氏) のように複数の分野に跨がっており、学際性を重視する JHPCN の活動目的にも符合しています。

4. 前年度までに得られた研究成果の概要

jh210036-NAH で第一に取り組んだのは課題名のように ABINIT-MP の高速化と超大規模系への対応です。高速化については、FMO 計算で常用される 2 次摂動レベル (FMO-MP2) のジョブの高速化で従前の Ver. 1 Rev. 22 に比較して 1.2~1.5 倍を得ました (対象系と基底関数に依存)。プログラムの改良のポイントは、片桐氏らの指針を基にした 2 電子積分ルーチン群の SIMD 化が主で、通信周りやプリントアウト量の制御などの小改良も行った結果です (主に坂倉氏、中野氏)。

構造ゆらぎを含めた相互作用解析で必要となる、分子動力学 (MD) 由来のタンパク質の液滴モデル群の処理ではフラグメントの総数は 1 万を超えることが普通です。しかし、Ver. 1 Rev. 22 では、GUI 用の配列群を保持していたために、「富岳」ではフラグメント数で 5.5 千個の扱いが限界でした。そのため、これら整理・削除して 1.1 万フラグメントまでの動作を確保しました (中野氏)。これにより、3 次摂動の FMO-MP3 も可能となりました。

SIMD 化と GUI 配列の整理を反映させた版を Ver. 2 Rev. 4 として 2021 年 9 月にリリースしました。この Ver. 2 Rev. 4 を HPCI 拠点でライブラリプログラムとして整備する活動も行いました。なお、GUI を重視する利用者の利便を図るため、Ver. 1 Rev. 22 も併存させています。詳しくは、既報 (望月ら, J. Comp. Chem. Jpn. 20 (2021) 132) を参照してください。

```

do npq=1,ngij
  if (abs(dkabm(npq)) > tv) then
    do nrs=1,ngkl
      if (abs(dkabm(npq)*dkcdm(nrs)) > tv) then
        ze = 1.0_8/(zetam(npq)+etam(nrs))
        a0 = dkabm(npq)*dkcdm(nrs)*sqrt(ze)
        rz = etam(nrs)*ze
        rho = zetam(npq)*rz
        tt = ((qm(1,nrs)-pm(1,npq))*(qm(1,nrs)-pm(1,npq)) &
              +(qm(2,nrs)-pm(2,npq))*(qm(2,nrs)-pm(2,npq)) &
              +(qm(3,nrs)-pm(3,npq))*(qm(3,nrs)-pm(3,npq)))*rho
        if (tt <= 36.0_8) then ! Tf = 2*m+36 (for the case of m=0)
          ts = 0.5_8+tt*fmt_inv_step_size
          delta = ts*fmt_step_size-tt
          ssss(0) = (((fmt_table(3,ts)*inv6*delta &
                    +fmt_table(2,ts)*inv2)*delta &
                    +fmt_table(1,ts))*delta &
                    +fmt_table(0,ts))*a0
        else
          ssss(0) = sqrt(pi_over4/tt)*a0
        end if
        sint(1) = sint(1)+ssss(0)
      end if
    end do
  end if
end do

do npq=1,ngij
  if (abs(dkabm(npq)) <= tv) cycle
  do nrs=1,ngkl
    if (abs(dkabm(npq)*dkcdm(nrs)) <= tv) cycle
    ix = ix + 1
    ze = 1.0_8/(zetam(npq)+etam(nrs))
    xa0(ix) = dkabm(npq)*dkcdm(nrs)*sqrt(ze)
    rz = etam(nrs)*ze
    rho = zetam(npq)*rz
    xtt(ix) = ((qm(1,nrs)-pm(1,npq))*(qm(1,nrs)-pm(1,npq)) &
              +(qm(2,nrs)-pm(2,npq))*(qm(2,nrs)-pm(2,npq)) &
              +(qm(3,nrs)-pm(3,npq))*(qm(3,nrs)-pm(3,npq)))*rho
  enddo
enddo
sint(1) = 0.0_8
do npqrs=1,ix
  tt = xtt(npqrs)
  if (tt <= 36.0_8) then ! Tf = 2*m+36 (for the case of m=0)
    ts = 0.5_8+tt*fmt_inv_step_size
    delta = ts*fmt_step_size-tt
    ssss(0) = (((fmt_table(3,ts)*inv6*delta &
                +fmt_table(2,ts)*inv2)*delta &
                +fmt_table(1,ts))*delta &
                +fmt_table(0,ts))*xa0(npqrs)
  else
    ssss(0) = sqrt(pi_over4/tt)*xa0(npqrs)
  end if
  sint(1) = sint(1)+ssss(0)
enddo

```

図 1. [ss|ss]型の2電子積分生成ルーチンでのループ分割の例

jh210036-NAH では、立教大学での以前の試み (Saitou et al., ChemBio-Informatics J. 18 (2018) 58) を元に深層学習の応用研究も行いました (土居氏)。これは、タンパク質内のアミノ酸残基間の相互作用エネルギーを可視化したヒートマップで可視化した画像を用いて、TensorFlow によって α ヘリックス構造と β シート構造を判定させるもので、「不老」 Type II 上で GPU による加速は 11 倍でした。GPU 利用を経験したことが、jh220010 課題での AlphaFold2 と cuQuantum の使用の「呼び水」になった面はあります。

5. 今年度の研究成果の詳細

(I) の ABINIT-MP の改良では、常用される 2 次摂動レベルの FMO-MP2 ジョブの高速化で、Ver. 1 Rev. 22 を従前として 1.5~2 倍となりました。2021 年度の jh210036-NAH の終了時では 1.2~1.5 倍でしたので、2 着実な加速を得たこととなります。高速化に併せて大規模系への対応と機能の向上を図っており、次の公開版 (Ver. 2 Rev. 8) を 2023 年 6 月に

リリースする予定で、暫定作業版にてテストを続けています (研究業績-学術論文の項目の [1])。この Ver. 2 Rev. 8 ですが、これまで同様に「富岳」や「不老」などの HPCI 拠点にてライブラリプログラムとして公開していきます (2023 年夏季以降)。

次に、ABINIT-MP の高速化の具体的な内容について記します。Ver. 2 Rev. 4 のリリース後、2 電子積分ルーチン群の追加の改良として、レジスタスピルの低減を期待してループの分割を行いました。図 1 は、片桐研究室の院生の満田氏が s 軌道 4 つの組 [ss|ss] の積分ルーチンで試行した分割例です (満田氏は、ABINIT-MP の高速化の検討の第 84 回情報処理学会の発表で学生奨励賞を受賞)。こうしたループ分割を [ss|ss] も含め、坂倉氏が系統的に実施しました。軌道タイプの組み合わせに依りますが、SIMD とループ分割の併用でチューニング前に比べて 20%~30% の加速を得ました。また、jh230036-NAH の報告書中で触れた Fock 行列の構築ルーチンでは $(1/2)^n$ ($n=1, 2, 3$) の因子を乗じ添字の同値性

表 1. Trp-Cage の FMO-MP2 ジョブのタイミングと加速

Ver. / Rev.	Date	Sec.	Acc.
6-31G*			
V1 R22	2020/6/3	469.6	1.00
V2 R4	2021/9/16	413.9	1.13
V2 R5	2021/12/9	344.2	1.36
V2 R5'	2022/7/13	294.2	1.60
V2 R5''	2022/9/8	282.7	1.66
V2 R6	2022/11/1	240.3	1.95
cc-pVDZ			
V1 R22	2020/6/3	1059.9	1.00
V2 R4	2021/9/16	876.1	1.21
V2 R5	2021/12/9	706.7	1.50
V2 R5'	2022/7/13	622.0	1.70
V2 R5''	2022/9/8	522.1	2.03
V2 R6	2022/11/1	463.8	2.29

判断の if 分岐を排除する改良も施し、当該処理を従前に比べて 30%加速させました（坂倉氏）。上記の積分ルーチンのループ分割、Fock 行列構築の if 分岐削除をした作業版を Ver. 2 Rev. 5 としました。この版に対し、モノマー段階での自己無撞着電荷 (SCC) のアンダーソン外挿をこれまでの Fock 行列ベースから密度行列ベースに変更するオプションを追加した版（中野氏）を Ver. 2 Rev. 5' としました。これとは別に、やはりモノマー SCC 段階のコスト低減を意図して 2 電子積分を In-core モードで処理できるバッファリング機能を追加して、Ver. 2 Rev. 5'' を作成しました（坂倉氏）。積分バッファリングは 4 つの軌道の角運動量の和で制御され、メモリが潤沢に利用できれば [dd|dd] までの全量の保持も可能ですが、デフォルトでは d 関数が 1 つまでとしています。そして、これら Rev. 5 の派生作業版 2 つをまとめたのが Rev. 6 です（渡邊氏、坂倉氏、中野氏）。

表 1 に、20 残基の人造タンパク質 Trp-Cage (PDB-ID=1L2Y) の FMO-MP2 ジョブを、「不老」 Type I で実行したタイミングと加速をまと

めました（望月自身でテスト・評価）。ノード数を 10 で、OpenMP を 24 スレッドで 1 ノードに 2 プロセスを割り当て、MP2 部分は全段 DGEMM 処理としました。表中では、リリースないし取りまとめた日付も示してある。Ver. 1 Rev. 22 を基準にした加速では、Ver. 2 Rev. 4 に対して Rev. 5 系では追加改良の効果が確認できません。6-31G*基底よりも s 関数の短縮長が長い cc-pVDZ 基底の場合、積分周りの改良による加速がより効いています。特に、Rev. 6 では 2 倍を超えた高速化を達成しています。

他のタンパク質も「不老」 Type I 上で改良の効果をチェックしましたが、1.5 倍～2 倍の加速が得られています（望月）。例えば、新型コロナウイルスのメインプロテアーゼ (PDB-ID=6LU7) では、Rev. 6 で、Ver. 1 Rev. 22 比では 6-31G*で 1.54 倍、cc-pVDZ で 1.74 倍でした。留意点として、Ver. 2 Rev. 5'' でのモノマー SCC でのアンダーソン外挿の密度ベースでのオプションで大反復の回数が増える場合があります。そこで、2023 年 4 月のテスト版ではエネルギー収束が設定値に達すると元の Fock ベースに切り替えるオプションを追加しました（中野氏）。これにより、安定した改善が得られるようになっていきます (Rev. 8 でのリリースに反映予定)。

次に、超大型分子系への取り組みについて記します。2021 年度から、望月が代表となって「富岳」の感染症対策の一般課題と並走していることもあり、2022 年度の後半から新型コロナウイルスのスパイクタンパク質の新しい構造 (クローズ型が PDB-ID=6XLU、オープン型が PDB-ID=6XMO) の分子動力学 (MD) シミュレーション由来の多サンプル構造での超大规模 FMO 計算に挑戦しており、2023 年度も継続中です (hp210026→hp220025→hp230017 & hp220352)。2020 年度の計算 (K. Akisawa et al., RSC Adv. 11 (2021) 3272) と異なり、図 2 に図示するように今回は液滴モデル

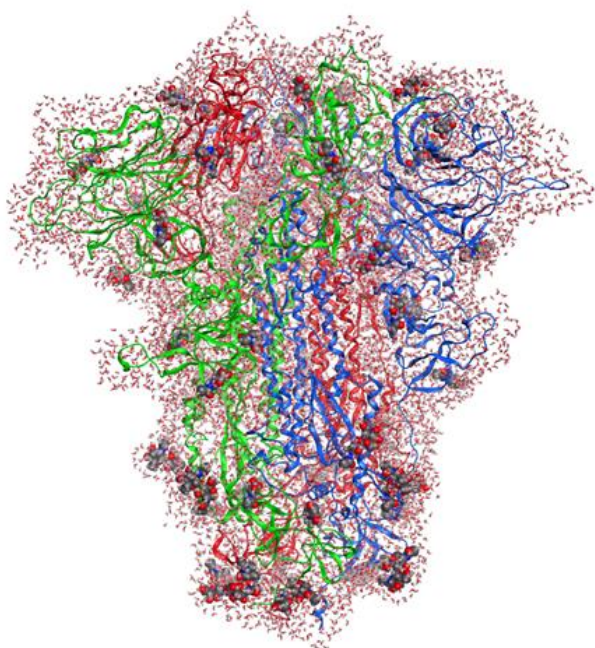


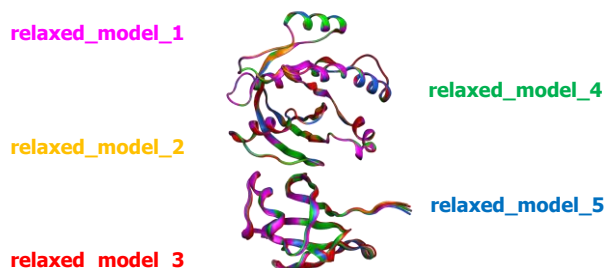
図 2. スパイクタンパク質の液滴モデル (PDB-ID=6XLU)

で表面の糖鎖も含めて扱っています(糖はボール表示で、3本のタンパク鎖は青赤緑で区別)。タンパク質の部分のアミノ酸残基数は3.3千、水とイオンを含めた総フラグメント数は1.8万に達します。残念ながら、Ver. 2 Rev. 4で行った作業配列の整理が未だ十分でないためか、Ver. 2 Rev. 6でも、この1.8万フラグメントのモデルをFMO-MP2計算することはできませんでした。そこで発想を変え、タンパク質との相互作用がほぼ無い遠方の水分子群はクラスターとしてまとめ、実効的にフラグメント数を低減する Python スクリプトを土居氏が新たに開発しました。このクラスタリングによって、フラグメント数は1万程にまで低減されて計算が可能となりました。並列計算の観点から見ると、水分子群が1つのフラグメントにまとめられると酸素原子の数が4,5個となり、アミノ酸残基部分と比べて重原子数の差が小さくなって粒度バランスの改善が期待できます。1構造サンプルでのタイミングでは、「富岳」の8ラックで48スレッド/3072プロセスの設定でFMO-MP2/6-31G*ジョブが2時間程で完了します。

2023年4月から、6XLUの液滴モデルの応用計算を「富岳」で進めています。この際、多サンプル構造の計算ジョブの投入と膨大な結果データの回収、相互作用エネルギー解析などのために様々な Python スクリプトが開発されて使われています(土居氏、奥脇氏)。

本 jh220010 課題とは直接は関係しないのですが、Ver. 2 Rev. 8では既述の高速化と超大規模系への対応だけでなく、計算機能も Ver. 2 Rev. 4に比べて強化されます。相互作用エネルギー解析では、分散力の寄与を相関補正から分離するオプション、静電エネルギー

Superimposed image of 5 structures



Comparison with PDB structure

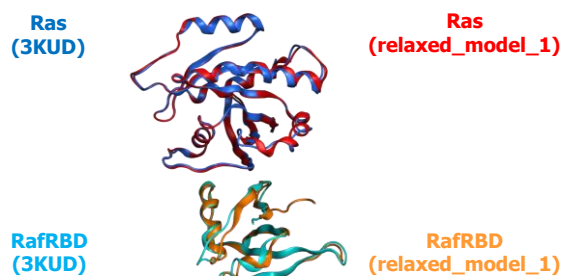


図 3. AlphaFold2による構造予測の例 (PDB-ID=3KUD)

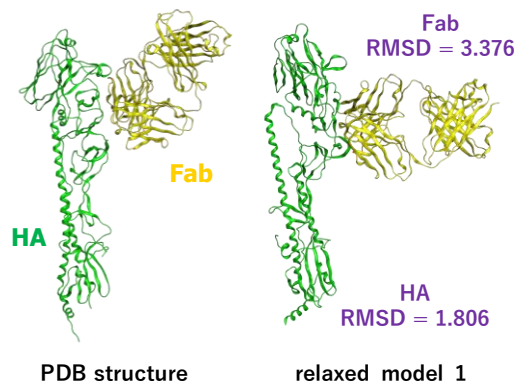


図 4. AlphaFold2による構造予測の例 (PDB-ID=1E08)

ギーを環境静電ポテンシャルを再現する点電荷 (RESP) から算定するオプションが追加になります。また、対象系の興味領域で励起エネルギーとイオン化ポテンシャルを高い精度で計算することも可能になります (詳しくは学術論文の[1]を参照)。

次に、(II)の AlphaFold2 (以下 AF2 と略記) についてです。AF2 は、Google 系の英 Deep Mind 社による公開ソフトウェアで、アミノ酸残基の FASTA 形式のシーケンスの入力だけで立体構造を高精度に予測できます。国内では、森脇氏が最も早い段階から導入した「伝道師」の一人 (<https://bit.ly/3Lnnagg>) となっています。大島氏とのコラボレーションで「不老」 Type II への導入が 2021 年 11 月から v2.2.1 を皮切りに始まりました。AF2 では深層学習の一種である Transformer (呼称は Evoformer) が処理のポイントの 1 つですが、その前段階で TB オーダーのデータベース探索による Multiple Sequence Alignment 情報の作成が必要です。大島氏によれば、このステップが SSD の利用によって数十倍に高速化されます (<https://bit.ly/3oT8WMK>)。大島氏と森脇氏の連携によって、AF2 の更新が 2022 年度も続けられ、最新の公開版は v2.3.1 です (研究業績-公開したライブラリの項目の[2])。

AF2 が導入された初期段階から、望月は FMO 応用計算での利用を意識して使用テストを行ってきました。図 3 は、AF2 による予測精度が高いとされる Ras-GDP と RafRBD のタンパク質複合体 (PDB-ID=3KUD) の v2.2.1 による予測構造です。構造が 5 つ出力され、モデル 1 で PDB 構造と比較すると、平均自乗変位 (RMSD) は Ras 部で 0.73、RafRBD 部で 0.74 と良好です。この系の処理時間は Type II の 1 ノードで 1.5 時間でした。ただ、FMO 計算で相互作用エネルギーを評価したところ、主鎖に比べて側鎖の位置が「微妙」なためか、荷電性残基間では 60 kcal/mol 程度の差が元

の PDB 構造での値と出ていました。AF2 の構造を「そのまま」で FMO 計算にかけるのは適当ではなく、MD による追加の構造緩和が必要と思われる。

図 4 は、v2.2.1 が「破綻」した例で、インフルエンザウイルスのヘマグルチニン (HA) と Fab 抗体の複合体 (PDB-ID=1E08) の予測構造を PDB のものと比較したものです。HA と Fab 界面の位置は大きくズレ、さらに重鎖と軽鎖から成る Fab 部分の構造の対応もよくありません。AF2 の v2.3.1 では、複合体の構造予測が改善されているようですので、2023 年度の活動で再度評価するつもりです (望月)。

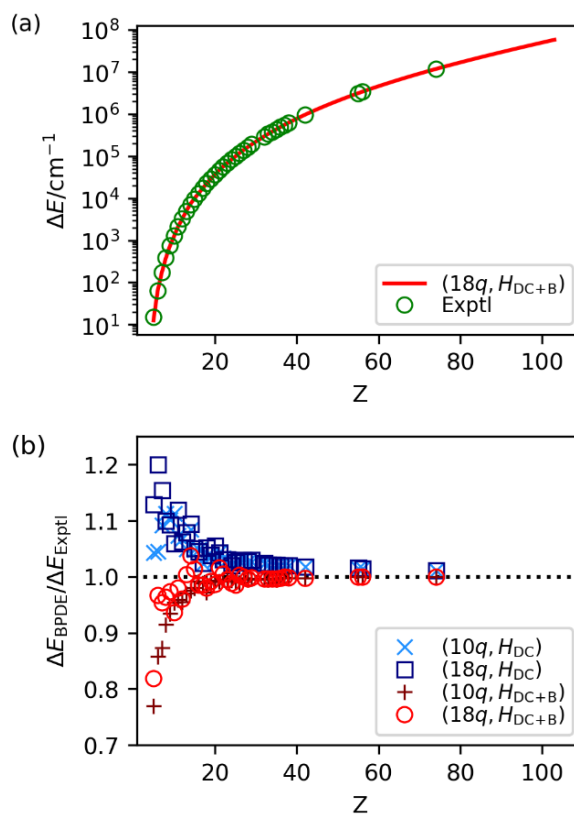


図 5. (a) スピン軌道分裂の計算値. (b) 量子ビット数と Breit 項の影響評価.

表 2. スピン軌道分裂の量子計算のタイミングと加速

#Qubits	WS	Flow		Acc.
	CPU	CPU	& GPU	
8	628	731	177	3.55
10	2197	2267	588	3.74
16	73452		4830	15.2
18	387328		9081	42.7

最後に、(III)の cuQuantum の利活用についてです。量子コンピュータの応用では、量子化学分野が有望とされています。2022 年度は、状態間のエネルギー差をベイズ推定の援用によって直接算定できる杉崎氏の手法 (Sugisaki et al., Phys. Chem. Chem. Phys. 23 (2021) 20152) を 4 成分の相対論計算に拡張し、B 原子と等電子系イオンの $2p_{1/2}$ 状態と $2p_{3/2}$ 状態のスピ軌道分裂エネルギーを評価しました。当該内容は主担当者の杉崎氏が arXiv に公開した上で、査読付きの雑誌に投稿中です (研究業績-その他の項目の [2])。量子化学計算のスキームと結果の詳細は、arXiv 論文を参照いただくとして、要点を以下にまとめます。

周期表を原子番号 5 番の B 原子から 103 番の Lr^{98+} イオンまで系統的に扱うため、クーロンの伝達遅延に関する Breit 補正も含めた Dirac の 4 成分ハミルトニアンを用いました。小成分の基底は、大成分から動力学均衡の関係から自動生成しました。Dirac-Hartree-Fock 計算では電子の 1 個少ない $(1s)^2(2s)^2$ で解き、空間対称性を保持した上で、相関計算の活性軌道空間として量子ビット (qubit) 数で 8 から 18 までを設定しました。18 量子ビットでは、 $\{1s, 2s, 2p, 3s, 3p\}$ の軌道がカバーされます。第二量子化ハミルトニアンの Jordan-Wigner 変換には、cuQuantum がサポートしている業界標準の OpenFermion と Cirq を用いました。

図 5 の (a) に、18 量子ビットで Breit 補正まで含めて計算したスピ軌道分裂値をプロットしました。緑丸が実験値ですが、原子番号が大きくなってもきわめて良好な対応を示していることが見てとれます。個別の値では、B 原子の計算値 $12.5287 \pm 0.9813 \text{ cm}^{-1}$ に対して実験値は 15.287 cm^{-1} です。原子番号 74 の W^{69+} では、各々 $11800183 \pm 44 \text{ cm}^{-1}$ と 11802000 cm^{-1} となります。全域での実験値との平均二乗誤差は 605.3 cm^{-1} となりました。

図 5-(b) の Breit 項の寄与の評価では、18 量子ビットの場合、原子番号が大きくなると明らかな改善が見られます。

杉崎氏所有のワークステーションと「不老」Type II での計算時間を表 2 にまとめました。加速はワークステーションと「不老」の CPU+GPU の場合との比で示しています (Xeon 間で大差ないため)。量子ビット数が増えると GPU 加速は顕著になり、18 量子ビットでの加速は 42.7 倍に達しました。原子番号を 5 から 103 までカバーする計算をワークステーションで行えば 1 年程かかるころでしたが、GPU を備えたスーパーコンピュータである「不老」Type II では僅か 1 週間程で済みました。これは、量子シミュレータ cuQuantum の威力を例示する好例と言えます。

6. 進捗状況の自己評価と今後の展望

(I) の ABINIT-MP の A64FX 向けの改良に関しては、FMO-MP2 ジョブに関して Ver. 1 Rev. 22 比の高速化で 2 倍に至ったのは一区切りです。また、2020 年から懸案であったスパイクタンパク質の水和モデルの扱いを可能にした点も良かったと思います。しかし、「富岳」や「不老」Type I の A64FX 機での理論性能比では未だ 4%程の効率に留まっている残念な状態です (測定は渡邊氏)。こうしたことから、総合的な達成度は 80%と判断します。2023 年度、jh230001 課題として jh220010 での活動が継続し、併せて片桐氏が代表で関係する jh230005 課題とも連動しますので、理論性能比の向上も注意して高速化を進めたいと考えています。

高速化改良の具体的な方向性ですが、2 電子積分ルーチン群の再構成が 1 つの軸になりそうです。実は、ABINIT-MP から FMO 固有の処理を省きつつハートリー・フォック (HF) 計算のみを切り出して作成したミニアプリで GPU 化の試みを既に試みています。これは、東京大学情報基盤センターの「OakForest-

PAGS II 導入に向けた GPU 移行サポート」の中で行われたものです。NVIDIA の成瀬彰氏は、OpenACC を使いつつ、CUDA コアを満たすようにバッチ処理化するために積分ルーチン群を {s, p, d} 軌道の組み合わせに応じてまとめ、Fock 行列の構築もメモリ転送でのロスを減らすために積分ルーチンにマージしました。Wisteria/Aquarius でのテストでは、GPU による加速は 6-8 倍となりました (望月)。試金石的として十分な結果と思われます。

ポスト「富岳」の FS では、「演算加速器」が検討されていますので、ABINIT-MP の GPU 対応も視野に入れた活動が今後必要で、そこから得られる「ヒント」や「知見」を A64FX での性能改善に活かすこともできそうです。

jh23001 課題では、HPCI センターでの Ver. 2 Rev. 8 への更新を含めたライブラリ整備を積極的に進めます。その中では、NEC の SX-Aurora TSUBASA を有する東北大学の A0BA-A へベクトル化対応版を整備することになりますので、これまでの名古屋大学に加えて東北大学も拠点になり、関係するメンバーも追加になっています (学術論文の [1] には、ベクトル化に関する参考記述あり)。

(II) の AlphaFold2 関係では、更新版 v2.3.1 の整備・公開を行いましたので、その点は問題ありません。ただ、実証応用が不足しましたので、達成度は 80% です。jh23001 課題では AlphaFold2 の更新と共に、生体分子シミュレーションの専門家も加わっていることから、タンパク質のモデリングでの活用も行う予定です。

(III) の cuQuantum に関しては、スーパーコンピュータでの量子シミュレーションで GPU 加速のメリットを明示した点で達成度は 90% と思います (国内では初の試み)。jh23001 課題では、良質な基底関数 (それだけ軌道数が増加) を使って遷移金属原子を含む小型分子の計算を行う計画を立てていますが、内殻軌道を凍結した上で仮想軌道を擬

似自然軌道を使ってコンパクト化する工夫が要りそうです。いずれにせよ、量子ビットの数が多くなるため、cuQuantum の強みが出るはずで

7. 研究業績

(1) 学術論文 (査読あり)

[1] “FMO プログラム ABINIT-MP の整備状況 2022”, 望月祐志*, 中野達也, 坂倉耕太, 渡邊啓正, 佐藤伸哉, 奥脇弘次, 秋澤和輝, 土居英男, 大島聡史, 片桐孝洋, J. Comp. Chem. Jpn., 21 (2022) 106-110.

(2) 国際会議プロシーディングス (査読あり) なし.

(3) 国際会議発表 (査読なし) なし.

(4) 国内会議発表 (査読なし)

[1] “FMO プログラム ABINIT-MP の整備状況、ならびに FMO-DPD 法の開発と応用”(依頼講演) 望月祐志*, 物性研短期研究会:「理論タンパク質物性科学の最前線:理論と実験との密な協働」, 2022/7/26.

[2] “FMO プログラム ABINIT-MP の「不老」Type I 向け改良について”(オンライン依頼講演) 望月祐志*, 第 3 回 スーパーコンピュータ「不老」ユーザ会, 2022/9/13.

[3] “A64FX スーパーコンピュータ上での ABINIT-MP の改良と大規模応用計算”(口頭) 望月祐志*, 中野達也, 坂倉耕太, 渡邊啓正, 秋澤和輝, 北原駿, 奥脇弘次, 土居英男, 山本詠士, 平野秀典, 泰岡顕治, 森脇由隆, 大島聡史, 片桐孝洋, 分子科学討論会 2022, 横浜, 2022/9/21.

[4] “FMO プログラム ABINIT-MP の整備状況 2022”(オンライン口頭) 望月祐志*, 中野達也, 坂倉耕太, 渡邊啓正, 佐藤伸哉, 奥脇弘次, 秋澤和輝, 土居英男, 大島聡史, 片桐孝洋, 日本コンピュータ化学会 2022 年秋季年会, 松本, 2022/11/26.

[4] “FMO プラグラム ABINIT-MP の現状と今後

”（オンライン口頭：依頼）望月祐志*, 分子科学研究所計算科学研究センター スーパーコンピュータワークショップ 2022, 2023/1/16.

[5] “FMO プログラム ABINIT-MP の整備状況と今後の GPU 対応”（オンライン口頭：招待）望月祐志*, 2022 年度 CMSF 勉強会「ポスト「富岳」世代の計算物質科学の展望」（第 2 回）～ GPU を使いこなすためのノウハウとポスト「富岳」検討状況の共有～, 2023/2/10.

[6] “FMO プログラム ABINIT-MP の現状と今後”（オンライン口頭：依頼）望月祐志*, 分子科学研究所計算科学研究センター スーパーコンピュータワークショップ 2022, 2023/1/16.

[7] “FMO プログラム ABINIT-MP の整備状況と今後の GPU 対応”（オンライン口頭：招待）望月祐志*, 2022 年度 CMSF 勉強会「ポスト「富岳」世代の計算物質科学の展望」（第 2 回）～ GPU を使いこなすためのノウハウとポスト「富岳」検討状況の共有～, 2023/2/10.

[8] “フラグメント分子軌道計算プログラム ABINIT-MP の高速化”（口頭）望月祐志*, 中野達也, 坂倉耕太, 渡邊啓正, 佐藤伸哉, 奥脇弘次, 土居英男, 大島聡史, 片桐孝洋, 情報処理学会第 85 回全国大会, 東京, 2023/3/3.

[9] “量子コンピュータによる微細構造分裂の直接計算法と GPU による数値シミュレーションの高速化”（口頭）杉崎研司, V. S. Prasanna, 大島聡史, 片桐孝洋, 望月祐志*, B. K. Sahoo, B. P. Das, 応用物理学会春期年会 2023, 東京 2023/3/15.

(5) 公開したライブラリなど

[1] HPCI 拠点のコンパイラ環境の変更等による ABINIT-MP のライブラリの更新依頼への対応（2022 年度）：R-CCS 「富岳」, 名古屋大学「不老」Type II, 東京大学「Wisteria」Odyssey & Aquarius（担当：望月祐志）.

[2] 「不老」Type II での AlphaFold2 のライブラリ整備に関する告知（2023/4/5 付で

v2.3.1 に更新）<<https://bit.ly/3LrCkIW>>（担当：大島聡史, 森脇由隆）.

(6) その他（特許, プレスリリース, 著書等）

[1] “ABINIT-MP プログラムによるフラグメント分子軌道(FMO)計算 1,2”（オンライン依頼講義）望月祐志*, 理研-配信講義 計算科学技術特論 B（2022）, 2022/6/9&2022/6/16 <<https://bit.ly/41VaYur>>.

[2] “Bayesian phase difference estimation algorithm for direct calculation of fine structure splitting: quantum simulation of relativistic quantum chemistry”, K. Sugisaki*, V. S. Prasanna, S. Ohshima, T. Katagiri, Y. Mochizuki, and B. P. Das, <<https://arxiv.org/abs/2212.02058>>, also submitted to a peer-review journal.