

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 8 日現在

機関番号：63903

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25730079

研究課題名(和文)量子化学大規模並列計算アルゴリズムの開発とその実装

研究課題名(英文)Development and implementation of algorithms for massively parallel quantum chemistry calculations

研究代表者

石村 和也 (ISHIMURA, Kazuya)

分子科学研究所・理論・計算分子科学研究領域・特任研究員

研究者番号：80390681

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：大規模並列量子化学計算アルゴリズムを開発・実装し、作成したプログラムSMASHをオープンソース(Apache 2.0)ライセンスで公開した。京コンピュータ10万CPUコアでも高い並列化効率と実行性能を示し、比較的高精度での巨大分子の電子状態及び構造計算が実用的になった。頻繁に用いられる1,2電子積分などの計算ルーチンをライブラリ化し他のプログラムへの組み込みを容易にした。応用面では、ナノサイズ金クラスターの精密な構造計算、他のプログラムと組み合わせた第一原理分子動力学計算による化学反応の活性化エネルギー計算を行った。

研究成果の概要(英文)：Algorithms for massively parallel quantum chemistry calculations were developed and implemented, and the developed program SMASH was released under the open source (Apache 2.0) license. It shows high parallel efficiency and CPU performance on 100,000 CPU cores of the K Computer. It is practical to calculate electronic structures and geometries of large molecules with relatively high accuracy. Frequently used routines such as one- and two-electron integral calculations are modularized to reuse easily. As applications, the structures of nano-sized gold clusters were calculated accurately, and the activation energy calculation of a chemical reaction using first-principle molecular dynamics was performed by combining with another program.

研究分野：計算化学

キーワード：量子化学 スーパーコンピュータ 並列計算 計算化学 巨大分子系 オープンソースライセンス

### 1. 研究開始当初の背景

分子の電子状態、エネルギー、構造を求める量子化学計算は、触媒や電池の材料設計、創薬などに幅広く用いられている。取り扱われる分子サイズは年々大きくなり、大きな置換基を付けることによる電子分布の変化や分散力などいわゆる弱い相互作用の多用により、新たな機能を持った物質・反応の設計が期待されている。また、複数の計算手法及びプログラムを組み合わせて、巨大な分子や動的な効果の計算も実用的な段階になっており、使いやすいプログラム及びライブラリが求められている。

他方で、利用できる計算機システムのノード数とコア数は、スーパーコンピュータ、研究室レベルの PC クラスタでも増加し続け、並列計算が不可欠になっている。特に、ノード間に MPI、ノード内に OpenMP を使うハイブリッド並列化が主流になり、京コンピュータでもこの方法が標準となっている。CPU 内部の構造も複雑になり、ハードウェアを理解した上でのアルゴリズム・プログラム開発が重要となっている。

### 2. 研究の目的

京コンピュータなどスーパーコンピュータを使って、これまでは困難であったナノサイズ分子の高精度電子状態計算を行うため、新たな MPI/OpenMP ハイブリッド並列計算アルゴリズムを開発し、実装する。エネルギー計算だけでなく、エネルギー微分計算や電荷解析計算、分子軌道の可視化なども実装し、分子構造最適化や様々な解析も実行できるようにする。

開発したプログラムコードをオープンソース (Apache 2.0) ライセンスで公開し、多くの研究者が自由に大規模並列計算を実行できるようにする。開発面では、どの計算手法でも頻繁に用いられる原子軌道 1, 2 電子積分計算ルーチンなどをライブラリ化する。一部の計算ルーチンの呼び出しや他のプログラムとの組み合わせを容易にし、量子化学分野における開発コストを削減する。

### 3. 研究の方法

(1) いわゆる弱い相互作用を取り扱える分子軌道計算方法の中で最も計算コストの少ない 2 次の摂動 (MP2) エネルギーの MPI/OpenMP ハイブリッド並列計算アルゴリズムを開発した (図 1)。均等な計算負荷及びデータ分散、総通信量はプロセス数に関わらずほぼ一定、BLAS Level3 を用いた効率的な演算、これらすべてを満たしている。

原子軌道 2 電子積分 ( $\mu\nu|\lambda\sigma$ ) の 4 つのインデックスを分子軌道 2 電子積分 ( $ai|bj$ ) に変換する部分が計算の大半を占めている (ギリシャ文字: 分子軌道、 $i, j$ : 占有分子軌道、 $a, b$ : 非占有分子軌道)。前半 2 つの変換では原子軌道を、後半 2 つでは分子軌道を各 MPI プロセスに割り振ることで、総通信量 (原子数の

約 4 乗に比例して増加) がほぼ一定になっている。

膨大な中間データ (原子数の約 4 乗に比例して増加) はメモリ上に分散保存している。一度に保存しきれない場合は占有軌道ペア (図 1 の最外ループ  $ij$ -block) を分割する。一方、原子軌道 2 電子積分を分割回数分重複して計算することになるが、計算全体中の割合は 2-3 割程度であり、余分にかかるコストはそれほど大きくはない。

```
do ij-block
  do  $\mu\lambda$  (原子軌道、MPI 並列)
  !SOMP parallel do schedule(dynamic,1)
  do  $\nu\sigma$ 
    原子軌道 2 積分 ( $\mu\nu|\lambda\sigma$ ) 計算
    第 1 変換 ( $\mu i|\lambda\sigma$ )
  enddo
  第 2 変換 ( $(\mu i|\lambda j)$ , dgemm)
enddo
do ij (分子軌道、MPI 並列)
  MPI_sendrecv ( $\mu i|\lambda j$ )
  第 3 変換 ( $(\mu i|bj)$ , dgemm)
  第 4 変換 ( $(ai|bj)$ , dgemm)
  MP2 エネルギー計算
enddo
enddo
MPI_reduce (MP2 エネルギー)
```

図 1 MP2 エネルギー並列計算アルゴリズム

(2) Hartree-Fock 法、密度汎関数法 (DFT) のエネルギー 1 次微分計算では、エネルギー計算同様原子軌道 2 電子積分の計算のコストが最も大きい。Gauss 基底関数の微分は、軌道角運動量と係数の異なる Gauss 関数になるためである。4 重ループの最外ループを OpenMP で、第 3 ループを MPI ランクで分散し、ループ内の計算が終了した後に reduction 節でノード内、`mpi_allreduce` ルーチンでノード間の 1

```
!SOMP parallel do schedule(dynamic,1)
!SOMP reduction(+:Gradient)
do  $\mu = nbasis, 1, -1$ 
  do  $\nu = 1, \mu$ 
     $\mu\nu = \mu(\mu+1)/2 + \nu$ 
     $\lambda_{start} = \text{mod}(\mu\nu + \text{mpi\_rank}, \text{nproc}) + 1$ 
    do  $\lambda = \lambda_{start}, \mu, \text{nproc}$ 
      do  $\sigma = 1, \lambda$ 
        原子軌道 2 電子積分項の微分計算
        +Gradient 配列への足し込み
      enddo
    enddo
  enddo
enddo
mpi_allreduce (Gradient)
```

図 2 原子軌道 2 電子積分項の微分計算並列アルゴリズム

次微分値を足し合わせるアルゴリズムを開発した(図2)。OpenMPでは動的に仕事を割り振り、さらに仕事の多いインデックスから少ない順にループを回している。MPIではIF文を使わずに非常に少ないコストで割り振ることで、並列化効率を高めている。1電子積分項、DFT計算の汎関数の数値積分の微分計算についても、同様の方法でMPI/OpenMPハイブリッド並列化を行った。

(3) 構造最適化の計算時間短縮においては、エネルギー1次微分計算の高速化だけではなく、最適化サイクル数の削減も重要となる。そこで、最もシンプルなCartesian座標以外に、分子の結合、角度、二面角の情報を使うredundant座標(P. Pulay, G. Fogarasi, J. Chem. Phys., 96 (1992) 2856)を導入し、強い共有結合でつながっていない分子間の結合を追加するなどの工夫も行った。構造最適化計算では、通常準Newton法と2次微分の更新にはBFGS法を用いて安定な構造を求める。初期2次微分にCartesian座標では単位行列を用いるが、redundant座標では分子力場パラメータを用いることで、大幅に値を改善する。Hessianの固有値を補正し確実にエネルギーが下がる方向に原子座標を動かすRational Function Optimization(RFO)法(J. Simons, P. Jørgensen, H. Taylor, J. Ozment, J. Phys. Chem., 87 (1983) 2745)も実装した。

(4) エネルギー計算のみならずエネルギー微分計算でも原子軌道1電子及び2電子積分計算を行うことから、これらの計算ルーチンをライブラリ化して、容易に呼び出せるように実装した。アルゴリズムはすでに開発したもの(K. Ishimura, S. Nagase, Theoret. Chem. Account, 130 (2011) 317)を改良して、必要なデータ(座標、基底、積分値)はすべてFortranの引数で受け渡して容易に呼び出せるようにし(図3)、さらに他の研究者も簡単に使えるようにした。

```

subroutine int2elec(twoeri, exijkl, coijkl,
xyzijkl, nprimijkl, nangijkl, nbfijkl,
maxdim, mxprsh, threshex)
twoeri  2電子積分計算値 (Output)
exijkl  primitive関数の指数 (以下 Input)
coijkl  primitive関数の係数
xyzijkl xyz座標
nprimijkl primitive関数の数
nangijkl 軌道角運動量(s=0, p=1, d=2,...)
nbfijkl  基底関数の数(s=1, p=3, d=5or6,...)
maxdim  twoeriの最大次元数
mxprsh  最大primitive関数の数
threshex exp(-x2)計算の閾値

```

図3 原子軌道2電子積分ルーチンのインターフェイス

(5) プログラムの公開にあたって、マニュアル、サンプルインプット、ウェブサイトの準備をした。インプット形式を極力シンプルにして、サンプルインプットを見るだけで実行したい計算のインプットファイルが作成できるようにした。さらに、計算して得られた分子軌道を可視化できるように、vtkファイルを作成するプログラムも作成した。

#### 4. 研究成果

(1) MP2エネルギーのテスト計算(Taxol(C<sub>47</sub>H<sub>51</sub>NO<sub>14</sub>), 6-31G(d)(970基底))をXeon E5-2697 v3(2.6GHz, 28CPUコア/ノード, 128GBメモリ/ノード)1ノードで行った。ij-block分割数を変えて行ったところ、表1のように分割数が増えるごとに2割程度の計算時間増加だけで、使用メモリ量を1/2, 1/3へと削減することができた。

次に、京コンピュータ(2.0GHz, 8CPUコア/ノード, 16GBメモリ/ノード)を用いてC<sub>150</sub>H<sub>30</sub>(6-31G(d), 2160基底)の並列計算を行った(表2)。1万CPUコアレベルでも高い並列加速率を示し、18432CPUコアでは、46.9秒と非常に短い時間で計算を完了した。

このアルゴリズムにより、弱い相互作用の高精度並列計算がPCクラスからスーパーコンピュータまでで効率よく実行できるようになった。

表1 MP2エネルギー計算時間及びメモリ使用量((C<sub>47</sub>H<sub>51</sub>NO<sub>14</sub>), 6-31G(d))

ij-block 分割数	1	2	3
MP2 計算時間(sec)	765.7	943.4	1121.5
使用メモリ量(GB)	101.0	50.5	33.7

表2 MP2エネルギー計算時間及び並列加速率((C<sub>150</sub>H<sub>30</sub>), 6-31G(d))

CPU コア数	4608	9216	18432
計算時間(sec)	152.5	83.4	46.9
並列加速率	4608.0	8425.9	14983.4

(2) DFT(B3LYP)エネルギー微分計算の並列性能を京コンピュータを用いてC<sub>150</sub>H<sub>30</sub>(cc-pVDZ, 2250基底)の測定を行った(表3)。エネルギー計算とは異なり行列対角化が無く、さらにノード内、ノード間の計算負荷分散のコストを極力少なくしたため並列化効率が非常に高く、CPUコア数と並列加速率がほぼ同じ値となっている。

表3 DFT(B3LYP)エネルギー微分計算時間及び並列加速率((C<sub>150</sub>H<sub>30</sub>), cc-pVDZ)

CPU コア数	4096	8192	16384
計算時間(sec)	101.2	50.8	25.5
並列加速率	4096.0	8159.7	16255.5

(3) redundant座標の性能を調べるため、2つの分子(Luciferin(C<sub>11</sub>H<sub>8</sub>N<sub>2</sub>O<sub>3</sub>S<sub>2</sub>), Taxol(C<sub>47</sub>H<sub>51</sub>NO<sub>14</sub>))について、初期構造をHartree-Fock/STO-3Gで最適化したものを使

い、B3LYP/cc-pVDZ での構造最適化回数を測定した。表 4 のように、Cartesian 座標に比べ、redundant 座標は最適化回数が 1/5 から 1/6 に大幅に削減することができた。初期 2 次微分が改良されたため、特に初めの数サイクルのエネルギー低下が大きく改善されている。

(2) のエネルギー微分計算と組み合わせることで、構造最適化 1 サイクルの高速化とサイクル数削減を同時に実現することができ、巨大分子の構造を短時間で求めることが可能になった。

表 4 B3LYP/cc-pVDZ 構造最適化回数 (初期構造: HF/STO-3G)

	Cartesian	Redundant
Luciferin	63	11
Taxol	203	40

(4) 開発したプログラムを SMASH (Scalable Molecular Analysis Solver for High performance computing) と名付け、2014 年 9 月に Apache2.0 ライセンスでバージョン 1.0.0 を公開した。2015 年 1 月に対応する計算機システムを増やし、単体性能を向上させたバージョン 1.1.0 を公開した。1 年半の間でダウンロード数は 596 で、そのうち海外からは約 20% であった。

複数の他のプログラムへの原子軌道 1、2 電子積分ルーチンの組み込みが始まっており、量子化学分野での効率的なプログラム開発及び性能向上に寄与できたと考えられる。

(5) SMASH プログラムを用いて、金 24 原子とその保護基からなる原子数 264、直径 2.5nm の分子  $\text{Au}_{24}(\text{SePh})_{20}$  (図 4) の計算を京コンピュータで行った。中心部分の金 4 原子クラスター 2 つの距離は、様々な DFT 汎関数を用いても実験値を再現せず、計算コストの大きい MP2 法で再現でき、架橋だけではなくクラスター間の弱い相互作用が構造において重要であることを示した。この計算では、6TByte 以上の大量のデータを各ノードに分散することで実行が可能となった。

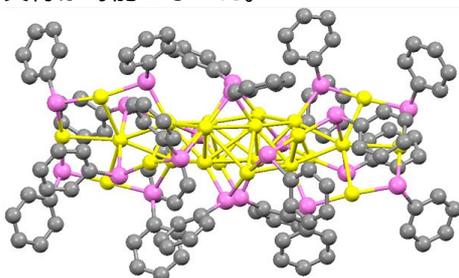


図 4  $\text{Au}_{24}(\text{SePh})_{20}$  分子図

原子数 560、5 重項開殻系の金属酸化物担体とロジウムクラスターの計算では、電子の詰まっていない軌道のエネルギーを補正し SCF 繰り返し計算途中での軌道の混ざり方を制限する Level Shift 法なども実装し、複雑

な電子状態のエネルギー計算の収束を可能にした。

これまでに開発した SMASH と第一原理経路積分プログラム PIMD を組み合わせた PIMD-SMASH を開発し、レプリカごとの MPI 並列、さらに SMASH の MPI/OpenMP 並列の 3 段階の並列化を行った。シクロペンタジエンとブテノン、さらに水分子を配置して Diels-Alder 反応の計算を行い、実験で示されている水の存在で遷移エネルギーが下がることを再現した。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 5 件)

(1) S. Ruiz-Barragan, K. Ishimura, M. Shiga, On the hierarchical parallelization of ab initio simulations, Chem. Phys. Lett., 査読有, 646, 2016, 130-135, DOI:10.1016/j.cplett.2016.01.017

(2) K. Ishimura, Development of massively parallel quantum chemistry program SMASH, AIP Conf. Proc., 査読有, 1702, 2015, 90053, DOI:10.1063/1.4938861

(3) M. Noda, K. Ishimura, K. Nobusada, Program Package of Photoinduced Electron Dynamics: GCEED (Grid-based Coupled Electron and Electromagnetic field Dynamics), JPS Conf. Proc., 査読有, 5, 2015, 11010, DOI:10.7566/JSPSC.5.011010

(4) J. Hasegawa, K. Yanai, K. Ishimura, Quantum Mechanical Molecular Interactions for Calculating the Excitation Energy in Molecular Environments: A First-Order Interacting Space Approach, ChemPhysChem, 査読有, 16, 2015, 305-311, DOI:10.1002/cphc.201402635

(5) M. Noda, K. Ishimura, K. Nobusada, K. Yabana, T. Boku, Massively-parallel electron dynamics calculations in real-time and real-space: Toward applications to nanostructures of more than ten-nanometers in size, J. Comput. Phys., 査読有, 265, 2014, 144-155, DOI:10.1016/j.jcp.2014.02.006

[学会発表](計 15 件)

(1) 石村和也, 超並列量子化学計算プログラム SMASH, 第 9 回分子科学討論会, 2015 年 9 月 16 日-19 日, 東京工業大学(東京都目黒区)

(2) Kazuya Ishimura, Development of Massively Parallel Quantum Chemistry

Program SMASH, 15th International Congress of Quantum Chemistry, 2015年6月8日-13日, Tsinghua University(Beijing, China)

(3) 石村和也、大規模並列量子化学計算プログラム SMASH の開発と公開、第 18 回理論化学討論会、2015年5月20日-22日、大阪大学(大阪府豊中市)

(4) Kazuya Ishimura, Development of Massively Parallel Quantum Chemistry Program SMASH, International Conference of Computational Methods in Sciences and Engineering 2015, 2015年3月20日-23日, Metropolitan Hotel Athens (Athens, Greece)

(5) 石村和也、巨大分子の構造最適化並列計算プログラムの開発、第 17 回理論化学討論会、2014年5月22日-24日、名古屋大学(愛知県名古屋市)

(6) Kazuya Ishimura, Massively Parallel Program for Quantum Chemistry Calculations, 5th Japan-Czech-Slovakia International Symposium on Theoretical Chemistry, 2013年12月2日-5日, 東大寺総合文化センター(奈良県奈良市)

(7) 石村和也、超並列量子化学計算プログラムの開発、第 16 回理論化学討論会、2013年5月15日-17日、福岡市健康づくりサポートセンター(福岡県福岡市)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

ホームページ

<http://smash-qc.sourceforge.net/>

ダウンロードページ

<https://sourceforge.net/projects/smash-qc/>

## 6. 研究組織

### (1) 研究代表者

石村 和也 (ISHIMURA, Kazuya)

分子科学研究所・理論・計算分子科学研究  
領域・特任研究員

研究者番号：80390681