

[共同研究成果]

量子化学計算パッケージMOLPROの
電子励起状態計算への応用と並列化

山崎 馨・河野 裕彦

東北大学大学院理学研究科化学専攻

E-mail: hirohiko-kono@m.tohoku.ac.jp

要旨

我々は、東北大学サイバーサイエンスセンターとの共同研究として量子化学計算パッケージMOLPROを並列コンピュータ上で並列化し、多配置電子状態理論に基づく電子励起状態計算を高速に行える環境を整備した。ピレンの垂直励起エネルギーを2次の多参照摂動法であるMulti-state complete active space self-consistent field second-order perturbation theory (MS-CASPT2法)に基づいて計算し、その並列化による計算速度の変化と精度を調べた。MS-CASPT2法による垂直励起エネルギーは実験値と定量的な一致を示した。また、並列化によりMS-CASPT2法のルーチンは、16コアを用いてノード内並列計算を行った時に最高の計算速度を示した。並列コンピュータ上のMOLPROは、並列計算機向け線形代数パッケージScaLAPACKや最適化されたMPIの組み込みなどの追加チューニングによって、一層のパフォーマンス向上が期待される。

1. 序論

量子化学計算とは、量子力学に基づいて分子の電子分布やエネルギーを求め、これらから分子構造（結合長や結合角）など分子を特徴付ける諸量を計算する手法の総称である。量子化学計算の問題点としては、計算対象とする分子のサイズの大規模化や計算精度の向上をはかろうとすると、CPU時間やメモリ量などの必要な計算リソースが指数関数的に増大することがあげられる[1]。特に、光などの外場との相互作用によって分子内の電子が高いエネルギーを持った電子励起状態の計算ではこれがさらに顕著である。電子励起状態の計算は、効率の良い太陽電池の開発や光合成機構など光と分子の相互作用を扱う必要がある研究では不可欠である。この為、大規模系や高精度が必要とされる系において電子励起状態計算を効率良く行うことで、光と分子の相互作用が関与する現象の理論解析を迅速かつ正確に行うことが出来る。

量子化学計算パッケージには、サイバーサイエンスセンターに導入されている *Gaussian 09* [2]をはじめとして、多種多様なものが開発提供されている。その中でも、Wernerらによって開発されたMOLPRO [3]は電子励起状態の高精度計算を高速に実行できることで知られている[4, 5]。このMOLPROを、1ノードあたり32コアのCPUと512GBのメモリを搭載するサイバーサイエンスセンターの並列コンピュータにおいて並列化して運用することで、研究室のワークステーションでは不可能な大規模かつ高精度の電子励起状態計算を迅速に行えるようになることが期待される。そこで我々は、サイバーサイエンスセンターとの共同研究としてMOLPROの並列化版を並列

コンピュータ上でビルド・インストールし、電子励起状態計算を高速に行える環境を整備した。

本稿では、まず量子化学計算、特に電子励起状態の高精度計算の基礎理論である多配置電子状態理論について概説した後に、電子励起状態計算の特徴を実例によって紹介し、最後にMOLPROの並列化とそのパフォーマンスについて議論する。

2. 量子化学計算と多配置電子状態理論の基礎

量子化学において分子のエネルギーや電子の分布を計算する際には、分子を構成する v 個の原子核は空間上に固定されており (Born-Oppenheimer 近似), その周りを N 個の電子が一定のエネルギー E_n を持って束縛運動しているとする描像をとる. 分子内の電子は量子力学に従うため, 分子内の電子がどのように分布しているかを表す電子状態のエネルギー E_n は離散的 (飛び飛び) の値をとる. この電子状態を記述する波動関数 $|\Psi_n\rangle$ は Schrödinger 方程式

$$\hat{H} |\Psi_n\rangle = E_n |\Psi_n\rangle \tag{1}$$

に従う. ここで \hat{H} は, 電子が持つ全エネルギーを表すハミルトニアン演算子 (関数) であり,

$$\hat{H} = -\sum_i^N \frac{1}{2} \Delta_i - \sum_{i,\alpha}^{N,v} \frac{Z_\alpha}{r_{i\alpha}} + \sum_{i>j} \frac{1}{r_{ij}} \tag{2}$$

と書ける. この(2)式の第 1 項は電子のもつ運動エネルギー, 第 2 項は原子核と電子の間に働くクーロンポテンシャル, 第 3 項は電子の間に働くクーロンポテンシャルを表す. Z_α は α 番目の原子核が持つ電荷, $r_{i\alpha}$ は α 番目の原子核と i 番目の電子との距離, r_{ij} は i 番目のと j 番目の電子との距離を表す. なおここでは $e = \hbar = m_e = 1$ なる原子単位系を用いた.

計算の対象とする分子中で N 個の電子がまず独立に運動しているとする. この時, 各電子の確率的な空間分布を特徴付ける関数を分子軌道 ϕ_i ($i = 1, 2, \dots, N, \dots$) といい, それぞれ軌道エネルギー ε_i ($\varepsilon_1 \leq \varepsilon_2 \leq \dots$) を持つ. この時, ϕ_i は規格直交化されるように選ぶことが一般的である. 分子内の N 電子がどの分子軌道に収まっているかを電子配置と呼ぶ. この際, 最も低いエネルギーを持つ電子配置 ψ_0 は, エネルギーの低い分子軌道 ϕ_1 から順に N 個の電子を詰めていくことで構築できる. この基底電子配置以外にも, 基底配置から N 個の電子のうち j 個 ($1 \leq j \leq N$) を空軌道 ϕ_a, ϕ_b, \dots ($N+1 \leq a \leq b \leq \dots$) に移し替えた, よりエネルギーの高い j 電子励起配置 ψ_μ ($\mu = 1, 2, \dots$) も考えられる. また, 量子力学的な要請から, 電子配置 ψ_μ は 2 つの電子の自由度 (位置とスピン) を入れ替えたときに符号が反転するように決められる (Pauli の禁制原理). さて, 分子の電子状態 $|\Psi_n\rangle$ は, 取り得る電子配置をすべて考慮し尽くすことで厳密解を得ることが理論的にはできる[6]. つまり $|\Psi_n\rangle$ の厳密解は,

$$|\Psi_n\rangle = \sum_{\mu=0}^{\infty} c_{n\mu} \psi_\mu \tag{3}$$

と形式的に表される．ここで $c_{n\mu}^2$ は、 μ 番目の電子配置 ψ_μ を見いだしうる確率を意味するから、

$$\sum_{\mu=0}^{\infty} c_{n\mu}^2 = 1 \quad (4)$$

を満たす必要がある．この時、波動関数 $|\Psi_n\rangle$ は規格化されているという．また、この様に考慮しうる電子配置を考え尽くしたときにその電子状態におけるエネルギー E_n も厳密解をとりうる．

しかし、 E_n の数値的厳密解は、特殊な方法を用いて原子やごく小さな分子に対してのみ現時点では計算可能であるので[7]、次のような近似が通常行われる．まず、(3) 式の電子状態 $|\Psi_n\rangle$ が **有限個の**電子配置を用いてほぼ正確に近似できると仮定し、そのために必要な M 個の電子配置 $\psi_0, \psi_1, \dots, \psi_{M-1}$ の寄与のみを第 1 近似として考慮する(多配置展開)．この時、(4) 式の規格化条件は採用した N 個の配置で満足されると仮定する．次に、考慮する分子の電子配置 ψ_μ を構成する分子軌道が、分子を構成する原子が持つ電子の確率的空間分布(原子軌道)の重ね合わせで書けると近似する (LCAO-MO 近似)．さらに、計算を簡略化する為に、Gauss 型関数の線形結合を用いて原子軌道を近似的に表現することにする¹．この原子軌道を表現する Gauss 型関数群のことを Gauss 型基底関数(GTO)と呼ぶ．実際の計算では、 $|\Psi_n\rangle$ の規格直交性を保ったままエネルギーを最小化するように、この GTO の係数と(3) 式の $c_{n\mu}$ を反復法による逐次近似を用いて最適化する[8]．このようにして分子のエネルギーを計算する手法を Multi-configuration-self-consistent-field (MCSCF)法と呼ぶ．また、MCSCF 法とこれを第 1 近似としてより高精度な電子状態やそのエネルギーを求める手法を総称して多配置電子状態理論という．なお、MOLPRO はこの多配置電子状態理論による計算を得意とする量子化学計算パッケージである．また、考慮する電子配置が 1 つである場合 MCSCF 法は単に self-consistent-field (SCF) 法と呼ばれる．SCF 法やこれを出発点としてより精度の高い電子状態や分子のエネルギーを求めていく理論を総称して単配置電子状態理論と呼ぶ．

分子内の電子が取り得る最低エネルギー状態である電子基底状態における分子の安定構造の付近では SCF 法によって計算された近似的な電子状態が良い第 1 近似であることが多い．しかし、電子励起状態においては電子状態を近似するためには安定構造の領域でも多数の電子配置が必要な場合が多く、高精度な計算の出発点として MCSCF 法が頻繁に採用される．

3. 電子励起状態計算の実例

それでは、多配置電子状態理論の電子励起状態計算への適用例を紹介しよう．本稿では、電子基底状態 (S_0) から一重項電子励起状態 S_1 と S_2 それぞれに分子の構造を保ったまま変化させるのに必要なエネルギー (垂直励起エネルギー) を有機分子ピレン

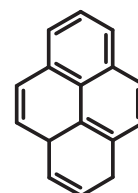


図 1: ピレンの分子構造

¹ 原子や水素分子・窒素分子などの小さな分子では、GTO の代わりに、空間グリッドを用いた数値基底関数を用いることも可能である．この応用例としては文献[10]を参照のこと．

[9](図 1)において計算した. なお, ピレンはバイオテクノロジーにおける試薬の中で光を出すパーツなどとして用いられる蛍光分子である.

今回は, 幾つかの電子状態のエネルギーの加重平均を最小化する多状態の MCSCF 法である State-averaged-CASSCF 法(SA-CASSCF 法)[11,12]と, これを出発点として, 多配置展開に漏れた電子配置の寄与を多項式展開してその2次の項まで取り込む2次の多参照摂動法である MS-CASPT2 法[13]を使った. MCSCF 法において考慮する電子配置としては, 最安定電子配置とこの配置において電子が収容されている分子軌道(被占有軌道)のうちエネルギーが高い順に4個と, 電子が収容されていない分子軌道(空軌道)のうちエネルギーが低い順に4個の計8軌道の中に含まれる8つの電子が取り得る全ての励起配置を考慮した ([8e,8o]-SA-CASSCF 法).

また, 単配置電子状態理論に基づく簡便な電子励起状態計算手法である configuration interaction singles 法 (CIS 法)と CIS(D)法, および電子密度と分子のエネルギーを結びつける汎関数に B3LYP を用いた時間依存密度汎関数法 (TD-B3LYP 法)を比較として用いた. CIS 法と TD-B3LYP 法では, 最安定な電子配置における被占有軌道から空軌道へと1つの電子が移動した(遷移した)電子配置である 1 電子励起配置のみを考慮に入れている. また, CIS(D)法は, 2 電子励起配置の寄与を部分的に取り込むことによって CIS 法で得られた電子状態とそのエネルギーの記述の改善を図っている. なお, ある程度の高精度計算にも対応した中型の GTO である cc-pVDZ と呼ばれる基底関数をエネルギー計算に統一的に用いた.

SA-CASSCF 法と MS-CASPT2 法の計算には MOLPRO 2010.1 [3]を, その他の手法の計算には Gaussian 09 [2]を用いた.

表 1 に, ピレンの S_0 状態から S_1 と S_2 状態への垂直励起エネルギーを上述の計算手法を用いて計算した結果を示す. なおこれまでの実験から, S_1 と S_2 状態は B_{2u} と B_{1u} と呼ばれる対称性をそれぞれ持つことが分かっている[14].

単配置電子状態理論に基づいて 1 電子励起配置のみを考慮した手法である CIS 法と TD-B3LYP 法は, S_1 状態と S_2 状態の順番

を間違えており, 定性的にも一致しない結果となっている. しかし, 2 電子励起配置の寄与を部分的に考慮した CIS(D)法と, 多配置電子状態理論である SA-CASSCF 法は, S_1 状態と S_2 状態の順番を正しく計算できており定性的に正しい計算であるといえる. さらに, SA-CASSCF 法を出発点とする 2 次の多参照摂動法である MS-CASPT2 法は, 0.1eV 以下の誤差で実験値 [15]と定量的に一致している. このことから, 多参照電子状態理論は正しく使用すれば定性的または定量的に実験で得られた垂直励起エネルギーを再現できることが期待できる手法の1つであるといえる.

表 1: 各種の方法を用いたピレンの垂直励起エネルギー. UB3LYP/6-31+G(d)法で構造を最適化し, 基底関数に cc-pVDZ を用いて励起エネルギーを計算した.

Method	$S_1 (B_{2u}) / \text{eV}$	$S_2 (B_{1u}) / \text{eV}$
CIS	4.614	4.450
TD-B3LYP	3.745	3.679
CIS(D)	3.807	4.514
[8e,8o]-SA-CASSCF	4.779	5.448
[8e,8o]-MS-CASPT2	3.257	3.846
Exp. [15]	3.375	3.857

4. MOLPRO の並列化

並列化には, shared memory 型の並列化ライブラリである Grobal Array Tool kit 4.3.2 (GA)[16] を用いた. GA の通信ライブラリには GA に同梱されている TCGMSG を利用した. また, 数値計算ライブラリ MKL に含まれる BLAS と LAPACK を GA と MOLPRO に組み込み, 計算の更なる高速化を図った. ソフトウェアのコンパイルの際には, Intel Compiler version 11.1.069 を用いた. 結果として, MOLPRO を並列コンピュータ上において 1 ノード 32 並列・メモリ容量 460 GB で運用することに成功した.

以下, 本共同研究においてビルドされた並列化版 MOLPRO における MS-CASPT2 法のルーチンの並列化パフォーマンスを示す. 第 3 章で説明した MS-CASPT2 法によるピレンの垂直励起エネルギー計算をテストジョブとして用いた. 使用するメモリ量は 1 コアあたり 1800 MW (14.1 GB) に設定した.

並列化前(並列数 1)と並列化後にノード内並列計算を行った場合(並列数 2・4・8・16・32) の計算所要時間(実時間)の変化を図 2 に,

速度向上率 (n 並列時の計算 実時間

÷ 並列化前の計算実時間)を図 3 に示

す. 並列化前には 16418.129 秒を

本計算に要した. 並列数を増やすに

つれ計算実時間は減少し, 16 並列時に

最短となった. この時の所要時間は

5248.29 秒, 速度向上率は 2.82 を記録

した. しかし, 32 並列時には所要時間

が 16 並列時よりも増加し, 並列化率

も悪化している. これはコア間の

ノード内通信の増加による速度低下

が並列化による並列化による高速化

を上回ってしまったためであると

推論される.

それでは, 更なる MOLPRO の

高速化のためにはどのような方法が

採れるであろうか. 先述したように,

MOLPRO の速度向上率が飽和して

しまう要因はノード内通信の増加

である. MOLPRO の並列化で用いた

GA は内部の通信ライブラリとして

同梱の TCGMSG を用いている. この

ライブラリはサイバーサイエンスセ

ンターの並列コンピュータに最適化

されていないので並列化版 MOLPRO

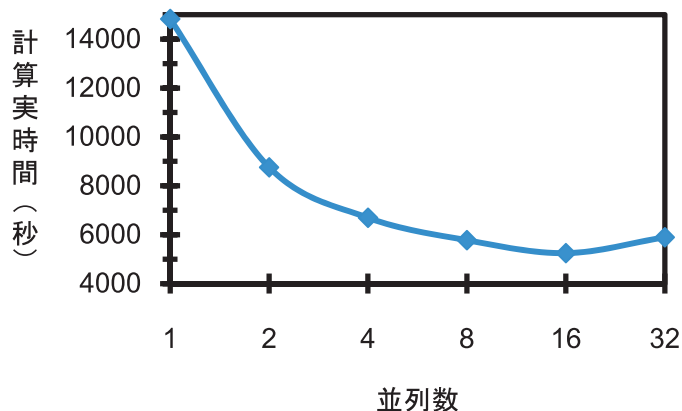


図 2: MS-CASPT2 法によるピレンの垂直励起エネルギー計算の計算実時間

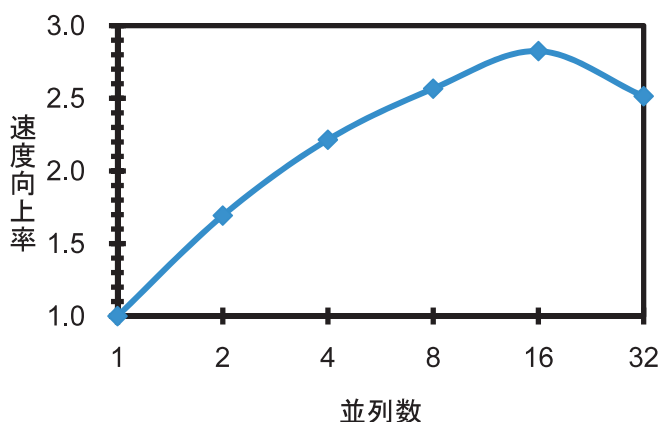


図 3: MS-CASPT2 法によるピレンの垂直励起エネルギー計算の速度向上率

の計算速度低下の要因になっていることが推論される。そのため、TCGMSG をサイバーサイエンスセンターに設置されている並列コンピュータ向けに最適化された MPI に置き換えることが有効であろう。

また量子化学計算の大部分は行列の演算によって行われているため、並列計算機向けの線形代数パッケージである ScaLAPACK の MOLPRO と GA への組み込みも計算速度向上のための有効な手段であると考えられる。

5. 結論

本稿では多配置電子状態理論とその電子励起状態計算の特徴を概説し、量子化学計算パッケージ MOLPRO の並列化について紹介した。サイバーサイエンスセンターの並列コンピュータが持つ潤沢な計算リソースと並列化された MOLPRO を組み合わせることによって、TDDFT 法などの簡便な電子励起状態計算手法では困難な高精度な電子励起状態計算が可能になった。並列コンピュータ上の MOLPRO には計算速度向上の余地が残されているので、更なるチューニングを行うことが望まれる。

謝辞

本稿に掲載した結果の一部はサイバーサイエンスセンターの並列コンピュータを用いて得られた。Intel Compiler version 11.1.069 の導入に当たっては、小野敏氏(共同利用支援係)にご尽力いただいた。また、MOLPRO の並列化とその運用に当たっては、小野敏氏と山下毅氏(同上)から多大な技術的支援をいただいた。これらの支援・協力をこの場を借りて感謝する。

参考文献

- [1] F. Jensen, *Introduction to Computational Chemistry*, 2nd edition, **2006**, Wiley.
- [2] M. J. Frisch *et al.*, *Gaussian 09*, Revision A.2, **2009**, Gaussian, Inc., Wallingford CT.
- [3] H.-J. Werner *et al.*, MOLPRO, version 2010.1, a package of *ab initio* programs, see <http://www.molpro.net>.
- [4] B. O. Roos, *Theory and Applications of Computational Chemistry: The First Forty Years*, **2005**, Elsevier, 725-764.
- [5] 永瀬茂・平尾公彦 著 『分子理論の展開』, **2002**, 岩波書店
- [6] P. O. Löwdin, *Phys. Rev.*, **1955**, 97, 1474-1489.
- [7] 例えば H. Nakashima *et al.*, *J. Chem. Phys.* **2007**, 127, 224104.
- [8] 金田行雄・笹井理生 監修, 笹井理生 編 『分子システムの計算科学—電子と原子が織り成す多体系のシミュレーション』 **2010**, 共立出版
- [9] A. A. Martí *et al.*, *Acc. Chem. Res.* **2007**, 40, 402-409.

- [10] T. Kato, H. Kono, *J. Chem. Phys.* **2008**, *128*, 184102.
- [11] H.-J. Werner *et al.*, *J. Chem. Phys.*, **1985**, *82*, 5053.
- [12] P. J. Knowles *et al.*, *Chem. Phys. Lett.* **1985**, *115*, 259.
- [13] H.-J. Werner *et al.*, *Mol. Phys.*, **1996**, *89*, 645-661.
- [14] J. Zhang *et al.*, *J. Phys. Chem. A*, **2010** *114* (42), 11117-11124.
- [15] E. A. Mangle *et al.*, *J. Phys. Chem.*, **1986**, *90* (5), 802-807.
- [16] J. Nieplocha *et al.*, *Int. J. High Perf. Comp. App.*, **2006**, *20*, 203.