

2E09

大規模分子の超並列計算のための MPI/OpenMP ハイブリッド並列 RI-MP2 解析的エネルギー勾配計算アルゴリズム (理研 AICS) ○河東田 道夫, 中嶋 隆人

MPI/OpenMP hybrid parallel RI-MP2 analytical energy gradient algorithm for massively parallel computation of large molecules (RIKEN AICS) ○Michio Katouda, Takahito Nakajima

【序】近年のスーパーコンピュータは、マルチコア CPU を複数搭載したノードを高速なインターコネクタで接続したマルチコア超並列クラスタシステム構成が主流となっており、現在ではスーパーコンピュータ「京」のような数万ノード・数十万 CPU コアで構成された数 10 ペタフロップス級の演算性能を持つスーパーコンピュータも出現している。発表者らの研究チームでは、「京」の性能を最大限に引き出し大規模分子の計算を可能とする次世代分子科学計算ソフトウェア NTCChem[1]の開発を行っている。ナノマテリアルや生体分子といった巨大分子およびその複合系の安定構造探索や化学反応経路追跡を行うためには、ファンデルワールス力などの弱い相互作用を正しい振る舞いで記述することが可能で、大規模な分子系を高精度に取り扱うことのできる電子相関理論に基づいた解析的エネルギー勾配計算手法およびソフトウェアが必要となる。本研究では、Møller–Plesset 2 次摂動(MP2)法に対する解析的エネルギー勾配計算を「京」などのマルチコア超並列クラスタシステムで高速・高並列に行うことを目的として、RI-MP2 解析的エネルギー勾配法の MPI/OpenMP ハイブリッド並列アルゴリズムおよびソフトウェアの開発を行った。

【方法】RI-MP2 解析的エネルギー勾配法は、従来法と比べて計算コストを大幅に削減することが可能なため、大規模分子の計算に適した手法である。発表者らの以前の研究では、RI-MP2 エネルギー計算の MPI/OpenMP ハイブリッド並列アルゴリズムの開発を行った[2]。本アルゴリズムは、「京」で推奨されている MPI/OpenMP ハイブリッドプログラミングモデルを採用し、さらに出来る限り多くの MPI プロセス数を用いて計算を行うため、仮想軌道を MPI 並列分割軸と用いた設計となっており、数千 MPI プロセスを用いた計算でも高い並列性能を実現することが可能である。今回開発した RI-MP2 解析的エネルギー勾配計算アルゴリズムでも、同様の設計指針を採用した。図 1 に今回開発したアルゴリズムの概略を 1 粒子・2 粒子密度行列計算部分(ステップ 2 および 3)を中心に示す。図に示されている 4 つの計算ステップの処理について、最外のループに対して MPI プロセス並列を行なっている。また、ステップ 2 の処理では、他のグループが開発したアルゴリズムでは占有軌道のループを MPI 並列分割軸として並列化を行なっているが、本アルゴリズムでは仮想軌道のループが最外になるようにし、仮想軌道を MPI 並列分割軸として用いて並列化を行っている。MPI 並列化された最外ループの内側で行われる処理については OpenMP でスレッド並列化を行っている。RI-MP2 計算で計算コストが $O(N^6)$ でボトルネックとなる 4 中心積分生成計算や 1 粒子・2 粒子密度行列の計算(図 1 内の総和 Σ の計算)は行列-行列積計算の形で処理を行うことが可能である。これらの処理を実装する際には、最適化 BLAS ライブラリに実装されているスレッド並列版の DGEMM を用いることで高いスレッド並列性能および実効性能を達成することが可能である。

【結果】今回開発したアルゴリズムを分子科学計算ソフトウェア NTCChem に実装し、「京」で並列性能ベンチマークテストを行った。表 1 にバッキーキャッチャー $C_{60}@C_{60}H_{28}$ の RI-MP2/cc-pVTZ エネル

ギー勾配計算の計算時間および高速化率を示す。「京」4096CPU コアを用いた計算の高速化率を4096倍とした場合、「京」32768CPU コアを用いた計算では19398倍の加速を達成しており、CPU コア数の増加に対して良好な並列性能を達成することが確認された。また、「京」32768CPU コアを用いた場合の計算時間は30分であり、「京」での構造最適化計算も十分に実現可能な計算時間である。本研究で開発されたソフトウェアを「京」で用いることにより、150原子4000原子軌道規模の分子のRI-MP2構造最適化計算がルーチンワークとして実行可能となり、ナノマテリアル設計や生体分子機能解明への応用が期待される。

[Step 1]
 Evaluate $V_{PQ}^{1/2} = (P|Q)^{-1/2}$ for all P, Q
 Evaluate $B_{ia}^Q = \sum_P V_{PQ}^{1/2} \sum_V C_{va} \sum_\mu C_{\mu i} (\mu\nu|P)$ for all i, P and $a \in A$
 Store B_{ia}^Q to distributed memory

[Step 2]
 do $a \in A = \text{MOD}(\text{Myrank}, N_{\text{proc}}) + 1$ (MPI parallel)
 Read B_{ia}^Q for all a, P from distributed memory
 do $i_{\text{rank}} = 0, N_{\text{proc}} - 1$
 $B = \text{MOD}(\text{Myrank} + i_{\text{rank}} \cdot N_{\text{proc}}, N_{\text{proc}}) + 1$
 Following procedures in this loop are performed
 for all i, j and $a \in A$ and $b \in B$
 Send B_{ia}^P to Myrank-1
 Receive B_{jb}^P from Myrank+1
 Evaluate 4-center MO integrals $(ia|jb) = \sum_P B_{ia}^P B_{jb}^P$
 Evaluate $T_{ij}^{ab} = \frac{(ia|jb)}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$ and $t_{ij}^{ab} = 2T_{ij}^{ab} - T_{ji}^{ab}$
 Evaluate MP2 correlation energy $E^{(2)} = t_{ij}^{ab} (ia|jb)$
 Evaluate 1-particle density matrix (DM) $P_{ij}^{(2)} = -\sum_k t_{ik}^{ab} T_{jk}^{ab}$
 Evaluate 2-particle DM $Y_{ia}^Q = \sum_j t_{ij}^{ab} B_{jb}^Q$
 end do
 Evaluate 1-particle DM $P_{bc}^{(2)} = \sum_j (2T_{ij}^{ab} - T_{ji}^{ab}) T_{ij}^{ac}$ for all i, b, c
 Evaluate 2-particle DM $\Gamma_{ia}^P = \sum_Q Y_{ia}^Q V_{QP}^{-1/2}$ for all i, P
 Store Γ_{ia}^P for all i, P to distributed memory
 Evaluate 2-particle DM $X_{PQ} = \sum_i \Gamma_{ia}^P B_{ia}^Q$ for all P, Q
 end do
 Allreduce $E^{(2)}, P_{ij}^{(2)}, P_{ab}^{(2)}$, and X_{PQ}

Evaluate 2-particle DM $\gamma_{PQ} = \sum_R X_{PR} V_{RQ}^{-1/2}$ for all P, Q
 Evaluate non-separable part of MP2 energy gradient $\frac{dE_{\text{MP2}}}{dx} = -2 \sum_{PQ} \gamma_{PQ} (P|Q)^{(x)}$

[Step 3]
 do $i_{\text{rank}} = 0, N_{\text{proc}} - 1$
 Send Γ_{ia}^P for all $i, P \in \text{Myrank} \cdot i_{\text{rank}}$, and $a \in \text{Myrank}$ to Myrank- i_{rank}
 Receive Γ_{ia}^P for all $i, P \in \text{Myrank}$, and $a \in \text{Myrank}$ from Myrank+ i_{rank}
 end do
 do $P \in \text{Myrank}$ (MPI parallel)
 Evaluate MP2 Lagrangian $L^3 = \sum_i \Gamma_{ia}^P \sum_V C_{va} \sum_\mu C_{\mu i} (\mu\nu|P)$ for all a, q
 Evaluate MP2 Lagrangian $L_{iq}^4 = \sum_a \Gamma_{ia}^P \sum_V C_{va} \sum_\mu C_{\mu a} (\mu\nu|P)$ for all i, q
 Evaluate part of MP2 energy weighted DM $W_{ij}^{(2)}[I] = 2 \sum_a \Gamma_{ia}^P (ja|P)$ for all i, j
 Evaluate 2-particle DM $\Gamma_{\mu\nu}^P = \sum_a C_{va} \sum_i C_{\mu i} \Gamma_{ia}^P$ for all μ, ν
 Evaluate non-separable part of MP2 energy gradient $\frac{dE_{\text{MP2}}}{dx} = 4 \sum_{\mu\nu} \Gamma_{\mu\nu}^P (\mu\nu|P)^{(x)}$
 end do
 Allreduce L_{aq}^3, L_{iq}^4 , and $W_{ij}^{(2)}[I]$

[Step 4]
 Evaluate rest part of MP2 Lagrangian L_{aq}
 Solve coupled perturbed Hatree-Fock equation to get $P_{\mu\nu}^{(2)}$
 Evaluate MP2 energy weighted DM $W_{\mu\nu}^{(2)}$
 Evaluate separable part of MP2 energy gradient

図 1. MPI/OpenMP ハイブリッド並列 RI-MP2 解析的エネルギー勾配計算アルゴリズムの概略

表 1. 「京」での MPI/OpenMP ハイブリッド並列 RI-MP2 エネルギー勾配計算の計算時間と高速化率 (C₆₀@C₆₀H₂₈ RI-MP2/cc-pVTZ, 3992 原子軌道, 374 相関軌道, 3618 仮想軌道, 10560 補助基底)

ノード数	CPU コア数	計算時間[秒]	高速化率
512	4096	8560	4096
1024	8192	4611	7604
2048	16384	2842	12336
4096	32768	1807	19398

[参考文献] [1] NTChem2013, http://labs.aics.riken.jp/nakajimat_top/ntchem_j.html

[2] M. Katouda and T. Nakajima, J. Chem. Theory Comput. **9**, 5373 (2013).