

# グリッド環境に対応した量子化学シミュレーション・システムの設計と開発

## 1. 2002 年度の具体的な研究計画

### 1-1. 対論的 Hartree-Fock (HF) 解の群論的分類とポスト HF 法

本研究を始めるにあたり、相対論的効果も含めた多電子系の HF を考察し、可能な波動関数の群論的分類を行う事を目的とした。さらに、平均場近似(HF 法)をこえる取扱いの可能性について検討する事を目的とした。

### 1-2. 相対論的 DFT の定式化と解の分類

ポスト QED HF-SCF 法で考慮される多体相関を一体密度行列を用いる密度行列汎関数法(DMFT)により取り入れる為の定式化を目的とする。さらに QED DMFT 解の群論的分類について考察する。

### 1-3. 全電子 DFT プログラムの各種数値計算手法の調査研究

DFT 法の発展の歴史の開明と其中での GSO-X プログラムの位置付けを明確にする。

### 1-4. DFT と post HF との比較検討

DFT 法とポスト HF 法の計算結果を比較検討し、電子相関効果の取り込みの視点から DFT 法の有効性とその限界を明らかにする。

### 1-5. GSO に基づく DFT (GDFT) 第一原理プログラム、GSO-X の開発。

GDFT の逐次のプログラムオリジナルコードを開発し、通常の汎用 ab initio DFT プログラム(Gaussian 等)と比較・数値的信頼性の検証を行う。さらにその GDFT コードの並列化を 2002 年度の研究計画としている。

## 2. 2002 年度の進捗状況と研究成果

### 2-1. 研究計画 1-1 に関する研究成果

多電子系における相対論効果を完全に含む量子電磁力学(Quantum Electrodynamics (QED))に立脚して、可能な相対論的ハミルトニアンを導出し、四成分スピノル型波動関数によるエネルギー期待値を求めた。さらに、Thouless の定理に基づく相対論的 Slater 行列式を求め、HF 自己無撞着場(self consistent field; SCF)方程式を導出した。次に一般化 QED HF-SCF 解の群論的分類には通常ポアンカレ群(4 成分並進およびローレンツ群)、C(荷電共役)、P(パリティ)、T(時間反転)を使用するが、分子系では点群(O)と P および T の直積群  $G=O \times P \times T$  で十分である事を証明した。

ポスト QED-HF 法の開発を目的として、最も安定な Broken-symmetry (BS) QED HF-SCF 解の密度行列を対角化する事により得られた 4 成分型自然軌道(QED-NO)とその

占有数(n)より、QED-NO 軌道空間を不活性軌道(inactive)、活性軌道(active)、空軌道(virtual)に分類した。active 軌道に関する部分空間を最適化する手法として、CAS (complete active space) SCF 法を採用し、QED CASSCF 法を定式化した。さらに、CASSCF 解に基づき動的相関を求める手法として QED CASPT2 および QED multireference(MR)coupled-cluster(CC)法を考案した。

一方、QED HF-SCF 解そのものを出発点として、多体摂動論法(Moller-Plesset;MP)および CC 法を用いるポスト HF 法を考案し、QED MP および QED CC 法を定式化した。図 1 に QED HF およびポスト QED HF 法をまとめた。

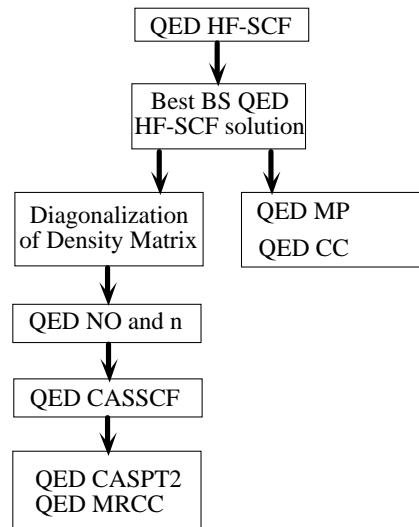


図 1 QED HF およびポスト QED HF 法の分類

## 2-2. 研究計画 1-2 に関する研究成果

量子電磁力学(QED)レベルの密度汎関数理論(DFT)を定式化する為に、4 成分型スピノル(ディラック場)を用いて定義される一体密度行列  $Q(x)$  を考察した。 $Q(x)$  は  $4 \times 4$  の 16 成分より構成されるが、スカラー(S)、ベクトル(V)、2 ランク反対称テンソル(T)、軸性ベクトル(A)、擬スカラー(P)成分(それぞれ 1,4,6,4,1 成分を持つ)に分解可能である。

$$Q(x)_{4 \times 4} = Q^S I + Q_\mu^V \gamma^\mu + Q_{\mu\nu}^T \sigma^{\mu\nu} + Q_\mu^A \gamma_5 \gamma^\mu + Q^P i\gamma_5 \quad (1)$$

ここで  $Q^x$  の定義は教科書の定義に従っている。次に、非相対論の場合と同様に全エネルギーを  $Q(x)$  の汎関数として定義する事により、Dirac-Kohn-Sham 方程式を導出した。

従って最も一般的な場合には 16 成分に関する汎関数を考える必要があり、このレベルの理論を QED 密度行列汎関数(DMFT)法と呼ぶ事にした(図 2 参照)。一方、(1)式において、4 カレント成分のみ考えた場合には、QED カレント密度汎関数理論(CDFT)を得る。さらに、カレント部分を軌道カレント密度(I)とスピン密度(S)のカレントに分割して、後者のみを考慮すれば相対論的(R)スピン密度汎関数法(RSDFT)を、前者のみの場合には相対論的密度汎関数理論(RDFT)を得る。以上の定式化に成功した。

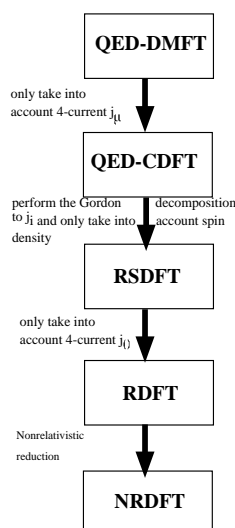


図 2 様々な相対論的/非相対論的 DFT の関係

一方、Dirac-Kohn-Sham 方程式より得られた解は QED HF-SCF 解の分類と同様に、点群(O)、パリティ(P)、時間反転(T)の直積群  $G=Q \times P \times T$  により分類可能である事を示した。

### 2-3. 相対論的極限への移行について

相対論的な 4 成分スピノル表現(Dirac 場)を 2 成分スピノル は非相対論的極限に近づくにつれて 2 成分スピノルに大別できる。

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \quad (2)$$

ここで、 $\phi$ と $\chi$ はそれぞれ大きい成分と小さい成分の二成分スピノルである。従って、その極限では $\chi$ 成分が無視できるとすれば、(1)式で  $Q^S$  と  $Q_0^V$  成分は一致し、通常非相対論的(NR)理論における電子密度に還元される。一方、軸性ベクトル成分  $Q_u^A$  から NR 理論におけるスピン密度が出てくる事が示される。一方、(1)式の他の密度行列成分は NR の極限で消える事がわかる。

従って、全ての相対論的補正(スピン軌道相互作用項など)を無視して、2 成分スピノル $\phi$ のみを考慮すれば、図 2 に示すように非相対論的密度汎関数理論(NRDFT)が得られる。

### 2-4. 全電子 DFT プログラムの各種数値計算手法の調査研究結果と、本プロジェクトで開発する DFT 計算コードの位置付け

DFT の利点は、電子相関の効果を波動関数展開ではなく、交換相関項であらわせる点にある。そして、その第一原理プログラムの最大の特徴も交換相関項にある。これは、交換相関項が、密度やその勾配さらには Hessian の非整数乗であらわさる為、厳密には基底で

展開できない事に起因する。歴史的に全電子第一原理 DFT プログラムは様々な数値計算手法を用いたアプローチがあり、現在も多くの系統がある。いずれも計算精度や描像に応じた最適の系を持ち、現在も利用されているものである。本プロジェクトで DFT プログラムを開発する為に、目的として第一原理計算 MO プログラム AMOSS や MD プログラム Cosgene との連成を念頭に起き、各種数値計算手法の調査比較検討を行った。今回開発したプログラムの位置付けを明確にする為、本節では調査検討結果を簡潔に報告する。

#### 第一世代：マフィンティン近似+多重散乱法

全電子の DFT プログラムは K.H.Johnson によって 1966 年に開発された。この方法の特徴は、ポテンシャルを、(i)原子核からある一定の範囲内では球対称、(ii)その範囲外では一定とする、「マフィンティン近似」を採用している事にある。この為、局在電子と遍歴電子が明確に区別できる固体電子系には適しており現在も非球対称補正を加えた Full potential linear augmented plane wave (FLAPW) 法として Viena 大学の K. Schwartz 教授らのグループの Wien2K 等、比較的高精度なバンド計算として利用されているが、化学結合や解離には不向きである。

#### 第二世代：離散変分(DV)法

D. E. Ellis, G.S. Painter らによって 1970 年に考案された手法で、現在も京都大学足立研究室の DVSCAT 等、材料科学の分野で広く用いられている。この方法では一電子・二電子積分ならびに DFT 特有の交換相関項を、多中心の数値積分を擬乱数を用いて計算する。生成する積分点は数百から数千点で、比較的計算量が少ない為、現在でもナノサイズのクラスターの単点(分子構造を固定した)計算には有効で、実際研究結果が発表されている。しかしながら、化学反応のように分子構造を変化させる場合、数値積分の精度によるノイズが生じる為、そのままでは QM/MM には適さない。

#### 第三世代：線形原子結合軌道(LCAO)+fitting function 法

量子化学で用いられている LCAO 基底への DFT の導入は、H. Sambe と R. H. Felton らによって 1975 年に初めて行われた。この方法では交換相関項のみならず、クーロン項も、AO 基底での fitting により AO の三中心積分で計算される。原子核を動かしても DV のような数値積分精度によるノイズが発生しない。よって一般に DV よりも高精度である。又、この後述べる第四世代の数値計算法にくらべて計算コストが比較的安くなる。しかしながら、展開する AO 基底関数が貧弱だと電子相関項も貧弱となる欠点を持つ。国内で最も著明な DFT プログラムである柏木・佐藤グループによる ProteinDF や海外でも DGauss 等は、この数値的手法を用いている。

#### 第四世代：LCAO+多中心 Fuzzy cell 型数値積分

現在、Gaussian98、GAMESS をはじめ、多くの汎用 DFT プログラムで用いられている方法で、多中心の Lebedev 求積法に、Becke や Stratmann らの Fuzzy セル法を組み合わせ、DFT の交換相関項を計算する。その他の項は Hartree-Fock(HF)計算と同様の Gauss 型関数の積分を行う。通常一原子当たり数千点から数万点と計算コストはかかるが、その分一般に精度がよく、積分点を増やせば増やすほど系統的に精度はあがるという利点がある。これは特に QM/MM に必要な Gradient の計算に有利になる。又、積分領域に分割可能な数値積分である為、並列化に適している。さらに、交換相関項以外は HF と共通であ

る為、HF との Hybrid 法に適している。

以上のような調査結果と、AMOSS との連成による HF との Hybrid 法、さらに Cosgene との連成による DFT ベースの QM/MM がコンピューティンググリッド技術グループの一つの目標である事から、我々のオリジナルコードとしては(i) HF 法との Hybrid にも (ii) 化学反応ダイナミクスにも適した第四世代の LCAO+多中心 Fuzzy cell 型数値積分を用いる事とした。

## 2-5. DFT と post HF との比較検討

DFT ベースの QM/MM が化学反応に有用であるには、DFT が他の高精度計算法と同等のパフォーマンスを示す必要がある。特に重要なのは遷移状態近傍の記述である。この遷移状態近傍での電子状態の分類から、化学反応は、Woodward-Hoffmann 則のように軌道対称性の側面から許容反応と禁制反応に区別される他、遷移状態近傍がラジカル性を持つかどうかを加えた、(1)軌道対称性許容+非ラジカル(allowed-nonradical;AN)、(2) 軌道対称性許容+ラジカル (allowed-radical;AR)、(3) 軌道対称性禁制+非ラジカル (forbidden-nonradical;FN)、(4) 軌道対称性禁制+ラジカル(forbidden-radical;FR)の四つに大別される。典型的な例であるブタジエンとエチレンの Diels-Alder 反応で言えば、通常の協奏的反応が AN 型の反応にあたり、anti および Gauche-out 型の遷移状態は AR 型に該当する。我々はその遷移状態の活性化障壁エネルギーについて、種々の post HF 法と DFT 法による比較検討を行った。

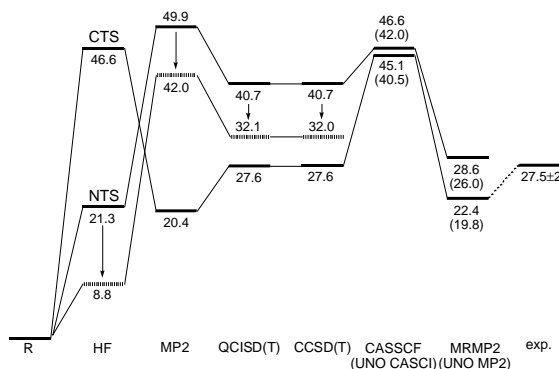


図3 post HF 法によるブタジエンとエチレンの Diels-Alder 反応の活性化エネルギー値は kcal/mol で ( ) 内は UNO CASCI と UNO MP2 の値を示す。点線ラインはスピン射影操作の後の値。

図3は、post HF の計算結果であり、MP2、quadratic CI、CCSD、CASSCF、Multireference MP2 (MRMP2) といった高精度計算ではいずれも協奏的反応(CTS)遷移状態が安定となっている。一方、Hybrid 型も含め、DFT の各種汎関数での同様な計算をおこなったところ、図4に示すように B3LYP、BLYP で CTS 型が安定となっている。特に、MRMP2 といった多配置波動関数ベースの方法を BLYP で再現できている等、密度勾配補正をおこなった DFT や、HF を部分的に Hybrid した B3LYP 法の有効性を示す結果となった。

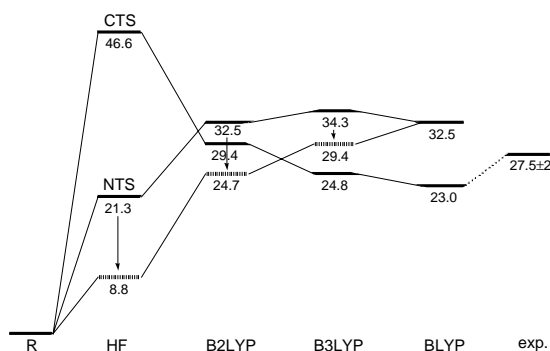


図4 DFTによるブタジエンとエチレンのDiels-Alder反応の活性化エネルギー(kcal/mol)

## 2-6. 列版 GSO-DFT コードの開発

2002年度は、GDFTのプログラムコードGSO-X (X=HF,LSDA,GGA,Hybrid。原理的にはMP、CCにも拡張可能)をまず逐次的なレベルで開発し、Gaussian98等の結果との比較を行い、さらに、NEC開発のab initio分子軌道プログラム、AMOSSとの連成を念頭に置いて、GDFT交換相関項の数値積分コードの並列化を、MPI/LAMを用いて行った。

### 2-6-1. GDFT 逐次コードの開発とその数値計算法の概略

我々の選択した第四世代DFTプログラム、GSO-Xの場合、DFTの交換相関項の数値積分は、

- (1) 各原子毎の一中心積分を動径×球面に分割して行い、
- (2) 積分点の位置を決定し、
- (3) 他の原子中心とのセル重率により各点の積分重率を計算し、
- (4) 交換相関項を数値積分する、

という方式を取る。(1)の動径方向に関しては一次元のEuler-Maclaurinの公式を用いて計算し球面上ではLebedevの求積法を用いる。球面積分に関しては動径の大きさに応じて積分点を剪定する、Pruned積分点を用いるが、開発コードとしては、基本はPopleらのSG-1と呼ばれるPruned方式を採用しているが、自由度を持たせた非一意なPruned方式を可能としている。多中心積分の場合、(3)にあげた異なる二つの原子間の積分領域をどうするかが問題となるわけであるが、その最もシンプルな解決法は、固体物理でいう所のWigner-Seitz cell、もしくは数学でいうところのボロノイ多面体を用いる事である。すなわち、二原子の midpoint に垂直な平面で各原子の積分領域を分割し、その積分領域のみでその原子中心の積分点を生成するというものである。しかしながらこれは二原子間の中間領域に存在する電子が、両方の原子の影響を受けながらふるまう、という化学結合の描像にそぐわない方法である。そこで、Beckeによって提案されたFuzzy cell法を用いる。すなわち、上記WS cellのような各中心積分について自原子の積分領域では積分重率を1にし別の原子の積分領域では積分重率を0にするという階段関数型の積分重率の代わりに、1から0まで徐々に減衰するような積分重率を(2)の原子間積分重率として採用する。我々の開発したGSO-Xのループの構造を図5に示す。

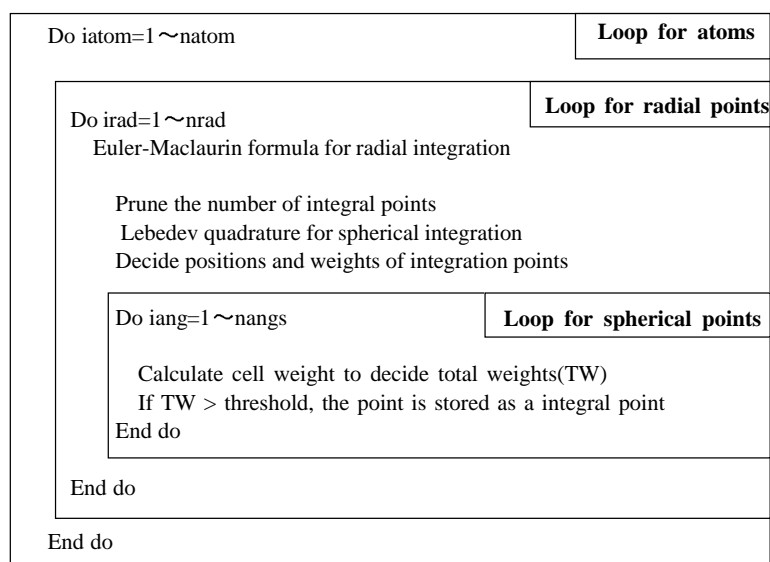


図 5 逐次的数値積分ルーチン

### 2-6-2. GSO-X による数値計算の結果

前節で述べた逐次的数値積分ルーチンを用いて、逐次型プログラムを構築したわけであるが、本節ではこれをもちいた計算例を示す。

我々の開発した GSO-X プログラムの最大の特徴は、傾いたスピンを記述できる事である。そのような状況は、(i) スピン傾斜を伴う化学反応や (ii) スピンフラストレーションやスピン縮重を起こす金属多核錯体で見られる。

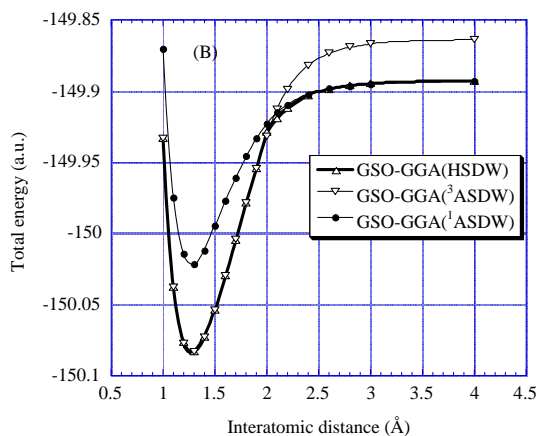


図 6 GDFT(GSO-GGA)法で計算した酸素分子のポテンシャル曲線

まず、前者の例として酸素分子の解離について述べる。酸素分子は最も基本的かつポピュラーな分子の一つでありながら、その結合解離の DFT による研究は皆無であった。それは、結合状態が分子全体で三重項であり、解離状態は三重項酸素原子二つである事に起因する。GSO-X による計算結果を図 6 に示す。HF の場合で知られているように、三重項酸素分子のような開殻分子は ASDW が安定となる(³ASDW=Axial spin density wave)。ところがその電子配置は、原子間距離の増加とともに三重項酸素原子と一重項酸素原子に乖離し、エネルギーは二つの三重項酸素原子よりも高くなる。一方、乖離極限で三重項酸素原子

を一重項にカップリングした配置は、原子間距離が近い領域では最安定スピン配置ではなくなる(ASDW)。軌道毎のスピン傾斜を考慮した解(HSDW=Helical spin density wave)は、この二つのポテンシャル曲線が交差する点の近傍で安定となり、二曲線を連続的に繋ぐ。なお、ASDW については Gaussian98 で同様の計算が可能であり、実際同じ基底関数で計算したところ、この図の ASDW の二曲線と重なり、精度的にプログラムの問題がない事が検証された。一方、GDFT 特有の解については Gaussian98 等汎用プログラムでは計算不可能な為、エネルギー値の比較はできないが、Mulliken のスピン密度解析(三次元に拡張版)を用いて、スピン密度を三次元でプロットしたものを確認したところ、図 7 に示すような三次元スピン構造の連続的変化という妥当な計算結果を得ている。

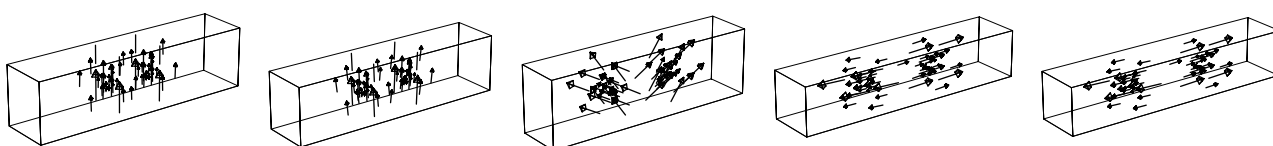


図 7 GDFT で計算した酸素分子の解離の三次元スピン構造

同様の計算結果を鉄と酸素の結合解離でも確認しており、図 8 にその三次元スピン構造の変化を示す。

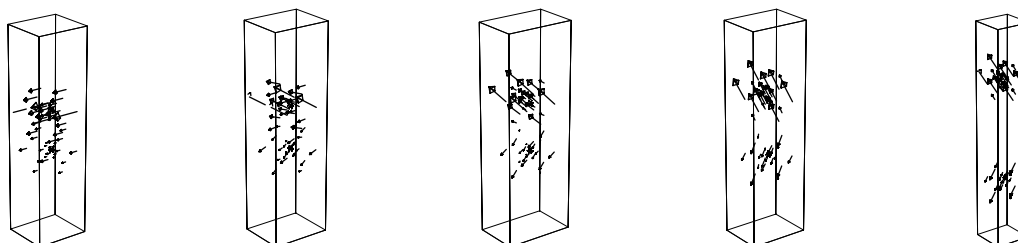


図 8 GDFT で計算した鉄酸素の解離の三次元スピン構造

これら酸素分子の解離や金属と酸素の結合解離自体も生体酵素反応でしばしば見られる重要な化学反応のカテゴリであるが、その酵素活性中心側はというと金属多核錯体、すなわち(ii)の場合である事が多い。ここでは例として Cubane 構造を持つ  $Mn(II)_4O_4$  の三次元スピン構造を図 9 に示す。(A)が GDFT 特有の解であり、(B)は Gaussian98 等でも計算可能な解である。エネルギーの相対的安定性を比較した結果、(A)の三次元スピン構造がより安定となった。実際の系では、Jahn-Teller 歪みや混合原子価になった場合等を考えると常にこのような三次元スピン構造が最安定である保証はないが、少なくとも第一原理計算法としては、Cubane 構造を持つ金属多核錯体の取り扱いには GDFT を用いる必要がある事が明らかである。



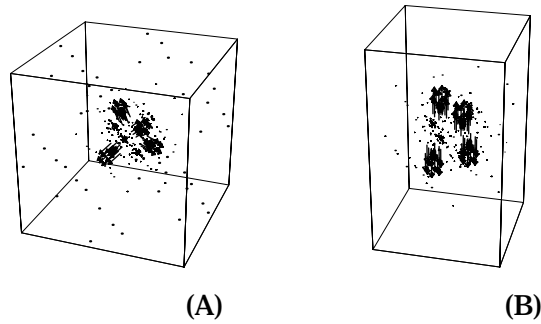


図 9 GDFT で計算した  $\text{Mn(II)}_4\text{O}_4$  の三次元スピン構造 (A) GDFT 解、(B)UDFT 解。

### 2-6-3. GSO-X の並列コードの開発

2-6-1 節で述べたように逐次的 DFT の数値積分は原子 × 動径 × 球面で行われる。従って、第一の並列化の方法として、球面のループを各プロセスに分割し、並列化するスキームが考えられる(球面分割型並列化)。これを用いれば、原子ループでも動径ループでも、自分の担当でないならスキップできる為、その分数値計算時間が短くなる。しかしながら、これも前述のとおり数値積分の際、各球面の積分点の剪定を行っており、各球面は等価ではないため球面分割型では各プロセスの負担が均等にならず、その分、並列化効率は落ちる。しかし、プロセス数に対し、原子数が増えればふえる程、この不均等がならされてこの欠点は補正されると予想される。一方、上記逐次型の計算でセル重率の計算がコストが高い事を考慮すれば、`ipts` をトータルの積分点の `index` として、セル重率の前で各プロセスに分割する方法が考えられる。この方法では、各プロセスの担当の積分点が空間に均一にばらまかれる為均一分解型並列化と呼ぶ。この並列化では、数値積分ルーチンのループで、スキップできる計算過程は球面分割型に比べ減るが、各プロセスの計算負担がほぼ等しくなる為数値積分の並列効率が高くなる。これら二つの並列コードのループ構造を、図 10、図 11 に記す。

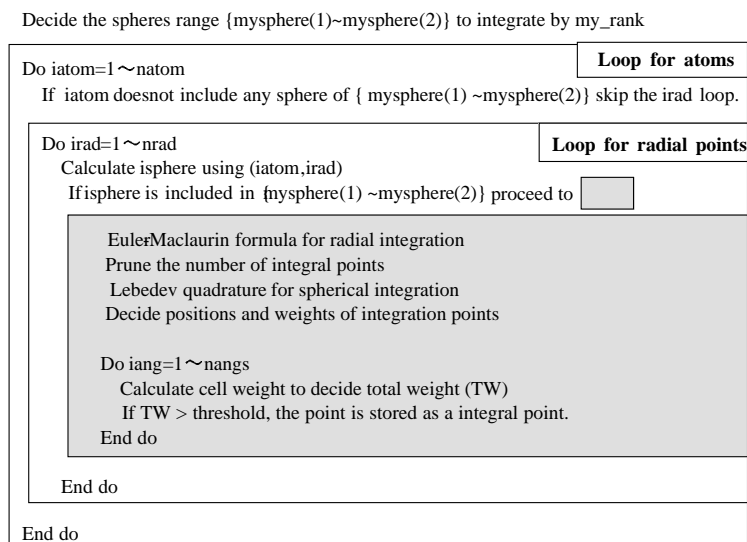


図 10 球面分割型並列化のループ構造

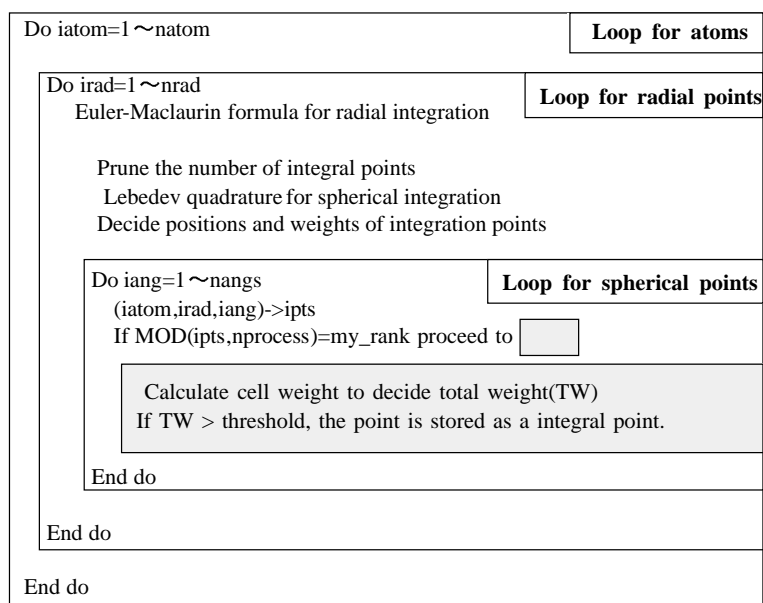


図 11 均一分解型並列化のループ構造

又、球面分割型と均一分解型の違いを直感的に理解する為、6 原子分子( $D_{6h}$ )をこれら二つのスキームで 8 台で並列化した時、図 12 (A)に示す全積分点の各プロセスの負担する積分点を図 12 (B)、(C)に示す。ただし、プロットの見やすさの為、実際の計算に用いる積分点よりは点の取り方を粗くしている。

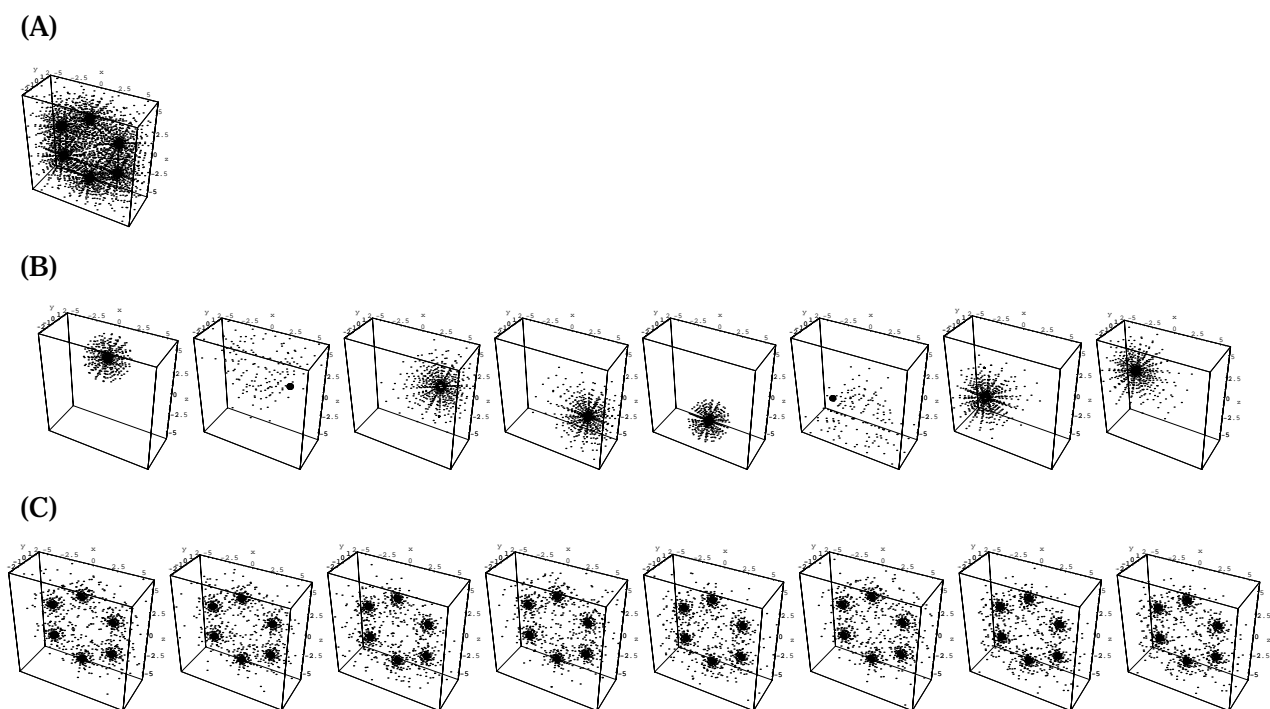


図 12 6 原子分子( $D_{6h}$ )を並列化し数値積分した際の各プロセス(計 8 プロセス)担当の積分点の空間分布  
 (A) 全体の積分点 (B) スキーム 1 (球面分割型並列化) (C) スキーム 2 (均一分解型並列化)

原子数を変えた際の球面分割型並列化の各ノードの分担積分点の数を図 13 にプロットした。10 原子の場合では負荷が最小のプロセスと最大のプロセスで 3 倍近い積分点数のばらつきがあるのに対し、20 原子、30 原子と、原子数が増えるに従って、この不均一性が弱まりならされていく様子が見て取れる。

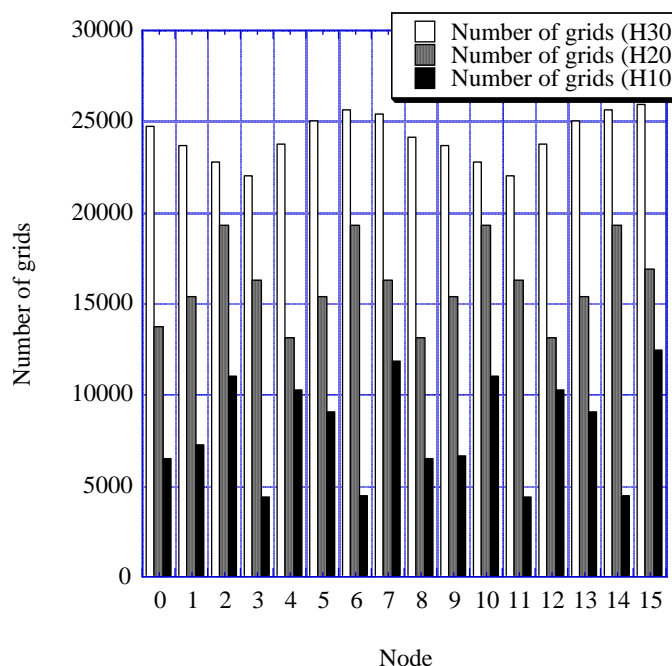


図 13 16 プロセスで球面分割型並列化した際の各ノードの分担積分点の数(10、20、30 原子)

交換相関積分とそのための積分点の並列化を評価するために、

- (1) 計算条件の読み込み。
- (2) 初期軌道の読み込み。
- (3) 密度行列の構築。
- (4) MPI を用いたマスターから各プロセスへの密度行列の送信。
- (5) 各プロセス毎の積分点生成。
- (6) 各プロセス毎の交換相関項の積分。
- (7) 各プロセスからマスターへの結果の送信。

からなる DFT 並列化コードを開発した。

1 台から 16 台までプロセス数を増やしていった場合の計算時間を球面分割型(Scheme1)、均一分解型(Scheme2)の並列化を用いた場合について、図 14 にプロットする。Ideal は並列化効率が 100%の場合の計算時間で、Gen は計算時間のうち、積分点生成に費やした時間である。(A)はシンプルな局所スピン密度近似である Hartree-Fock-Slater 汎関数、(B)は複雑な一般化された密度勾配近似(GGA)、Becke-Lee-Yang-Par 汎関数を用いた。この図は原子数 10 の場合の結果である。この原子数では球面分割型の並列化効率のロスがおこっている事や、並列化すべき積分点が少ないため、非並列コード部分(1)~(3)の比率が大きくなりアムアダールの法則から予想できるように並列化が完全に(100 パーセント)行われた場合の Ideal からの大幅なずれが全ての方法で見られる。

一方、原子数を 30 に増やした場合の同様のプロットを図 15 に示す。この原子数では、全体の計算時間に対する数値積分の計算時間の割合が高くなる為、その並列化効率も高く

なる。又、10 原子の場合について、球面分割型の並列化効率が均一分解型との比較において回復しているのは、図 13 に示した、負担の均一化による。このことから、巨大分子系に対しては、球面分割型も有効である事がわかる。又、汎関数の比較では、HFS では積分生成 (GEN) の計算時間の全体のそれに占める割合が約半分であったが、複雑な汎関数である BLYP では、積分全体の計算が点生成に比べひと桁近く大きい計算時間となっている。今日、通常用いられる DFT の汎関数は BLYP 等の複雑な GGA が多いので、AMOSS の連成にあたっては、生成した積分点を計算を通して格納し使用する方法よりも、積分点生成+数値積分をまとめて AMOSS の SCF ルーチンから DFT 用プロセスとして spawn する方が、Grid コンピューティング向き、かつ、本プロジェクトコンピューティンググループのプログラム設計指針に適していると言える。

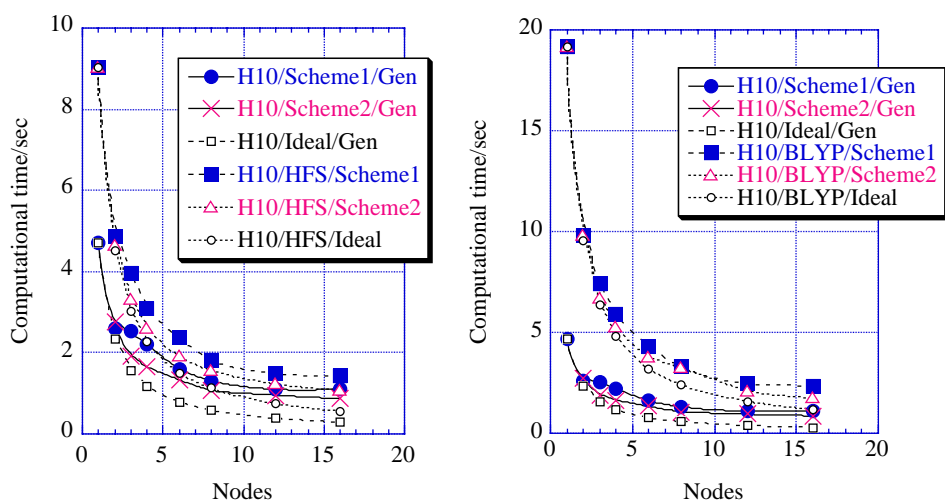


図 14 1~16 プロセスで並列化した際の計算時間(10 原子の場合)。(A)HFS, (B)BLYP 汎関数使用

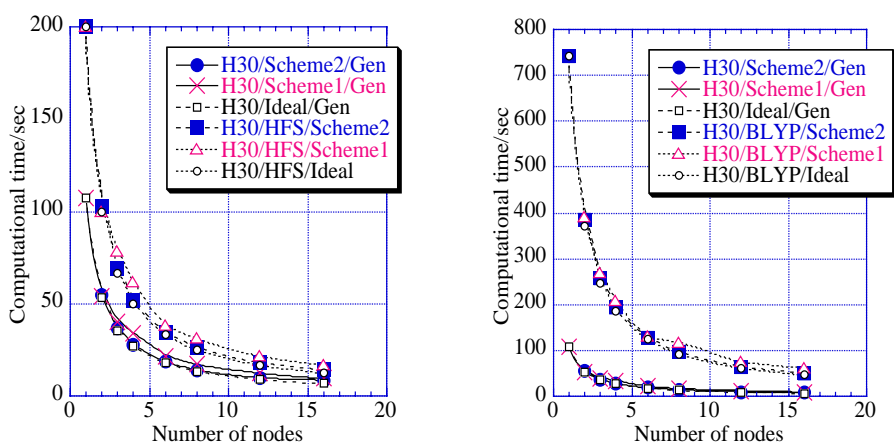


図 15 1~16 プロセスで並列化した際の計算時間(30 原子の場合)。(A)HFS, (B)BLYP 汎関数使用