

2019 年度 卒業論文

分子の擬対称性の認証

神戸大学工学部情報知能工学科

梁有希

指導教員 天能精一郎 教授

2020 年 2 月 14 日



分子の擬対称性の認証

梁有希

要旨

物質構造の単位として原子や分子が存在し、これら微視的な結合単位が集まって巨視的物質を構成している。この分子の電子状態や性質について論ずる際に用いられるのが点群対称性である。

小分子の幾何学的構造には高い点群対称性を有するものが多く存在しており、量子化学計算の高速化に利用されてきた。一方、タンパク質などの大規模分子においては、厳密な点群対称性は一般的に存在せず、対称性は考慮されないことがほとんどである。しかし、大規模分子においても、大まかな視点で見れば近似的な対称性が存在している場合が多く、この近似的な対称性を擬対称性と定義すると、擬対称性は電子状態に反映されると考えられる。

本研究は、対称性がないとされてきた分子について分子構造から擬対称性を認証し、量子化学計算なしに定性的な軌道エネルギーの議論をすることを目的とした。慣性モーメントテンソルを用いた点群対称性判定プログラムに、機械学習の決定木を使って擬対称性を認証する判定条件を導入し、分子構造の擬対称性を認証する擬対称性認証 (Quasi-Symmetry Recognition, QSR) プログラムを作成した。また、小分子と銀クラスターについて QSR で判定した擬対称性ごとに軌道エネルギー準位を並べたところ、小分子では擬対称性においても厳密対称性と同様に軌道エネルギーの縮重が見られた。一方、銀クラスターでは軌道エネルギーの縮重は見られなかったものの擬対称性に分けることで軌道エネルギーの分子サイズ依存性が明確になった。

本研究が、従来のケモインフォマティクスに代わり、分子構造情報のみにもとづく簡便な電子状態解析の新たな手法への第一歩となることを期待している。

目次

第1章 序論	1
1.1 背景	1
1.2 目的	3
第2章 理論	4
2.1 群	4
2.1.1 群の定義	4
2.1.2 物体の対称要素	4
2.1.3 対称操作と対称要素	4
2.2 分子の対称による分類	5
2.2.1 C_1 , C_i , C_s 群	5
2.2.2 C_n , C_{nv} , C_{nh} 群	5
2.2.3 S_n 群	6
2.2.4 D_n , D_{nh} , D_{nd} 群	6
2.2.5 立方群	7
2.3 一般的な分子の点群を決定するためのフローチャート	8
2.4 QSR	9
第3章 実装	10
3.1 プログラムの説明	10
3.1.1 回転対称軸の特定	10
3.1.2 回転対称軸の次数 n の特定	10
3.1.3 C_2 軸の決定	10
3.1.4 対称平面	11
3.2 構造データに基づく分子の群論的分類	11
3.3 閾値	13
3.3.1 決定木	13
3.3.2 説明変数と目的変数	14
3.3.3 最適な閾値の推定	15

第 4 章 結果と考察	16
4.1 擬対称性の判定	16
4.1.1 CC14	16
4.1.2 Peptide-ethane	17
4.1.3 AcOH-AcOH	17
4.2 閾値が大きい場合	17
4.2.1 AA1	18
4.2.2 Pyridine-pyridine pipi	18
4.2.3 Benzene-ethane	19
4.3 閾値の推定	19
4.4 擬対称性が C_1 である分子	25
4.5 軌道エネルギーとの関係	26
4.5.1 小分子の軌道エネルギー	26
4.5.2 銀クラスタの軌道エネルギー	31
 第 5 章 結論	 35
謝辞	36
参考文献	37

第1章 序論

1.1 背景

物質構造の単位として原子や分子が存在し、これら微視的な結合単位が集まって巨視的物質を構成している。その簡単な例として、酸素分子と炭素分子が化学反応によって酸化炭素の分子となる機構に関し、最初に化学者の考えた説明は、各分子から何かカギかホックのようなものが出ており、互いに引っかかりを起こして新しい繋がりを生ずるという素朴なモデルであった。やや進んで、原子価線で結びつくという系統的取扱いが Lewis などにより提唱され、原子価飽和の現象も説明された[1]。また、岩塩などのような見事な結晶が成長する機構も、異極性結合として量子化学誕生以前にすでに立派な定量的理論が作られていた。次いで量子力学の発展にともない、原子や分子やまた固体物質の発光機構やエネルギー準位の定量的理論が創り上げられ、また古典的には説明できなかった等極性結合も、量子力学的共鳴という見地から Heitler と London の理論によって理解されるようになった[2]。他方、金属結合という新しい凝集機構も説明されていった。

ところで、これら原子、分子あるいは結晶内の電子準位を与える理論や、上述の多岐にわたる化学結合の理論が成功した根底の理由は、モデルすなわち素朴なカギ、ホックないしは原子価線、さらにその後量子力学の理論に置き換えられた波動関数の数量的空間依存性にあるのではなく、これら物理形象が対称性をそなえており、理論がそれを正しく反映することができたという点にある。対称性を数学的に取り扱う具として、群論の役割がクローズアップされてきて、群論の応用が実り多い結果をもたらした理由もまたここにある。

群論はこれまで様々な形で物理学に应用されてきた。例えば、古典結晶物理学である。巨視的な物理量を表す各種テンソルの対称性の議論などが当てはまる。そして上でも述べた量子力学である。群論は量子力学の問題を取り扱う新しい利器として Wigner により提唱され、原子スペクトルを解析することに始まった[3]。次いで Bethe により、分子や結晶の電子状態の問題に群論が適用された[4]。今日までに原子、分子、結晶について多くの精緻な研究がなされるに至ったのには、群論が大きな影響を与えている[5]。

分子などの幾何学的構造を不変に保つ場合、その分子は幾何学的な対称性を有する。

そして、その構造を不変に保つ操作のことを対称操作という。対称操作の集合は群をなし、これを点群という。点群は、分子の対称性や結晶の局所的対称性を記述するものであり、分子の電子状態や振動取扱う際に、また結晶の局所的性質の対称性を論ずる際に重要である。

この点群が量子化学に用いられた例を挙げる。一つ目は、分子積分である。原子や分子の電子状態を求めるには Schrödinger 方程式を解く必要がある。

$$\hat{H}\psi = E\psi \quad (1.1)$$

Schrödinger 方程式が厳密に解けるのは非常に単純な系に限られており、一般的な分子について解析的に解くことはできない。この Schrödinger 方程式の近似解を求める際に出てくるのが分子積分である。分子積分の中で最もよく現れる積分は重なり積分である。原子 A と B をそれぞれ中心に有する原子軌道関数を $\chi_A(\mathbf{r})$, $\chi_B(\mathbf{r})$ とすれば、重なり積分 S_{AB} は、

$$S_{AB} = \int \chi_A^*(\mathbf{r})\chi_B(\mathbf{r})d\tau \quad (1.2)$$

で定義される。二つの軌道関数が同一の場合($\chi_A = \chi_B$)、完全に一致するので、 $S_{AB} = 1$ となる。反対に、 χ_A と χ_B が完全に重なっていないときは $S_{AB} = 0$ となる。

重なり積分は軌道関数の既約表現に依存するので、 χ_A と χ_B が異なる既約表現の場合は、 $S_{AB} = 0$ となる。このように数値計算をしなくても既約表現から一部の分子積分が 0 であることがわかるので、分子計算を省略し、分子積分の高速化に寄与してきた。

二つ目は Woodward-Hoffmann 則である。Woodward-Hoffmann 則は”反応の前後において反応に関与する電子の所属する分子軌道の対称性は保存される”と主張する。これによって様々なペリ環状反応が起こりうるかどうか、またその立体特異性が説明される。例えば Diels-Alder 反応である。Diels-Alder 反応とは、Fig.1.1 のように共役ジエンにアルケンが付加して六員環構造を生じる環化付加反応である。



Fig.1.1 Diels-Alder reaction.

この反応は、六員環遷移構造(TS)をもち、六電子が環状に非局在化した芳香族類似の安定化を受けている。反応は協奏的かつ立体特異的に起こり、反応の進みやすさと立体化学は、反応物の共役二重結合または電子対の数(TS で非局在化する電子数)に関係し、分子軌道の対称性で説明できる。これは反応の進行過程で軌道の相互作用が同位相で起こる必要があることに基づいている。これは対称性が反応機構の解析に用いられた例と言える。

1.2 目的

小分子の幾何学的構造には高い点群対称性を有するものが多く存在している。点群対称性を有していれば選択則に基づいて不要な分子積分を計算する必要がない。量子化学では分子積分計算の高速化において対称性による選択則が利用されてきた。一方、タンパク質などの大規模分子においては、厳密な点群対称性は一般的に存在せず、対称性は考慮されないことがほとんどである。しかし、大規模分子においても、大まかな視点で見れば近似的な対称性が存在している場合が多く、その対称性は電子状態に反映される。この近似的な対称性を擬対称性と定義する。特に、擬対称性によって軌道エネルギー準位が予想できる可能性がある。その場合は、反応性の高い領域や逆に関与しない領域を量子化学計算することなく特定することが可能になる。特定がうまくいけば、反応の理論解析の高速化につながられる。この近似的な対称性を求める手段としてデータサイエンス的な認証技術は相性がよく、大規模計算への重要な貢献につながると期待できる。

本研究では、まず点群対称性を判定するプログラムを python 言語で作成する。次にそのプログラムを応用して擬対称性を特定するプログラムを作成する。そして、小分子の擬対称性を判定して、機械学習のデータセットを作成し、 C_1 分子の擬対称性を判定するのに適切な閾値を求める。

第2章 理論

2.1 群

本節では群の定義の紹介と対称操作、対称要素の説明を行う。

2.1.1 群の定義

g_1, g_2, \dots, g_n が集合 G の元であって、次の 4 つの公理が満たされているとする。

- (1) G は積に関して閉じている。すなわち、任意の 2 つの元 g_i と g_j の積 $g_i g_j$ が G に属する。
- (2) $g_k(g_i g_j) = (g_k g_i)g_j$ (結合則)
- (3) $eg = ge = g$ となる g の要素 e が一意的に存在する。(単位元の存在)
- (4) $gg^{-1} = g^{-1}g = g$ となる g の要素 g^{-1} が一意的に存在する。(逆元の存在)

以上の 4 公理を満たす集合 G のことを”群”と呼ぶ。集合 G の元である g の個数を群の”位数”と呼ぶ[5]。

2.1.2 物体の対称要素

ある行為を行なったあとで物体がもともとと同じに見えるとき、その行為を”対称操作”という。代表的な対称操作には、回転、鏡映、および反転がある。それぞれの対称操作にはそれに対応する”対称要素”が存在する。これは点や線や面があって、これらの点、線、または面に対して対称操作を行う[6]。

2.1.3 対称操作と対称要素

少なくとも一個の共通の点を不変に保つような、色々な操作に対応する対称要素に従って物体を分類することで”点群”を帰属できる。この種の対称操作は 5 種類ある。それぞれの対称要素に対する対称操作を以下に示す。

・恒等要素(E):

何もしない。

・鏡面(σ):面での鏡映。

鏡面(σ)における鏡映操作は 3 種類ある。

もし面が主軸に平行であれば、この面は”垂直”であるといい σ_v で表す。2本の C_2 軸の間の角を二等分する垂直な鏡面を”二等分鏡面”といい、 σ_d で表す。対称面が主軸に垂直であるときは、この面は”水平”であるといい、 σ_h で表す。

・対称中心、反転中心(i):

中心に関する全ての原子の反転。

・回転(C):

軸回りの回転。軸回りの回転 C とは、 n 回対称軸 C_n 回りの n 回回転のことであり、 $\frac{360^\circ}{n}$ だけ回転させることである。

・回映(S):

軸回りの回転をしたのち回転軸に垂直な面内での鏡映を行う。

2.2 分子の対称による分類

本節では、分子の対称による分類の紹介をする。

2.2.1 C_1, C_i, C_s 群

・ C_1 群: E 以外に対称要素を持たない分子は C_1 群に属する。

$$C_1 = \{E\} \quad (2.1)$$

例 CB_3ClFI

・ C_s 群: E 以外に対称面 σ のみを有する分子は C_s 群に属する。

$$C_s = \{E, \sigma\} \quad (2.2)$$

例 SOCl_2

・ C_i 群: E 以外に反転 i のみの要素を持つ分子は C_i 群に属する。

$$C_i = \{E, i\} \quad (2.3)$$

例 $(\text{CH}(\text{OH})\text{COOH})_2$

2.2.2 C_n, C_{nv}, C_{nh} 群

全て $n \geq 2$ とする。

・ C_n 群: E 以外に C_n 軸を1本のみ有する分子は C_n 群に属する。

例 H_2O_2 (C_2 対称性)

- ・ C_{nv} 群： C_n 軸 1 本と、 σ_v を n 個有する分子は C_{nv} 群に属する。

例 NH_3 (C_{3v} 対称性)

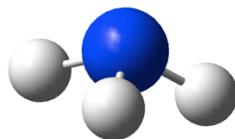


Fig.2.1 Structure of NH_3

- ・ C_{nh} 群： C_n 軸 1 本と σ_h を 1 つ有する分子は C_{nh} 群に属する。

例 H_3BO_3 (C_{3h} 対称性)

2.2.3 S_n 群

- ・ S_n 群： 回映 S_n を持つ。 n が 4 以上の偶数の場合にこの軌道を用いて表される。

$n = 2$ の場合は C_i と一致する。

例 $\text{C}(\text{C}_6\text{H}_5)$ (S_4 対称性)

2.2.4 D_n , D_{nh} , D_{nd} 群

全て $n \geq 2$ とする。

- ・ D_n 群： C_n 軸 1 本とこの C_n 軸に垂直な C_2 軸を n 本有する分子は D_n 群に属する。

例 $\text{C}_{10}\text{H}_{16}$ (D_2 対称性)

- ・ D_{nh} 群： D_n の要素を有し、かつ主軸 (C_n 軸) に垂直な鏡面 (σ_h) を有する分子は D_{nh} 群に属する。

例 C_2H_4 (D_{2h} 対称性)

- ・ D_{nd} 群： D_n の要素を有し、かつ全ての隣接した C_2 軸の間の角を 2 等分する垂直な n 個の鏡面 (σ_d) を有する分子は D_{nd} 群に属する。

例 C_2H_6 (D_{2d} 対称性)

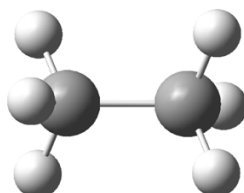


Fig.2.2 Structure of C_2H_6

2.2.5 立方群

・ T 群： 正四面体を不変に保つ回転操作から成る分子は T 群に属する。その位数は 12 である。

・ T_h 群： T の要素を有し、かつ反転 i を有する分子は T_h 群に属する。

・ T_d 群： T の要素を有し、かつ鏡面 d を有する分子は T_d 群に属する。

例 CH_4

・ O 群： 立方体、あるいは正八面体を不変に保つ回転操作から成る分子は O 群に属する。その位数は 24 である。

・ O_h 群： O の要素を有し、かつ反転 i を有する分子は O_h 群に属する。

例 C_8H_8

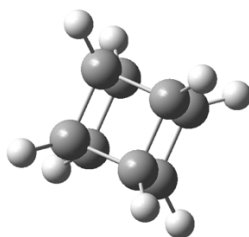


Fig.2.3 Structure of C_8H_8

・ I 群： 正十二面体、あるいは、正二十面体を不変に保つ回転操作から成る分子は I 群に属する。

・ I_h 群： I の要素を有し、かつ反転 i を有する分子は I_h 群に属する。

例 C_{60}

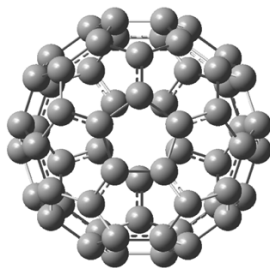


Fig.2.4 Structure of C_{60}

2.3 一般的な分子の点群を決定するためのフローチャート

一般的な分子の点群を決定するためのフローチャートは Fig.2.5 のとおりである [7]。

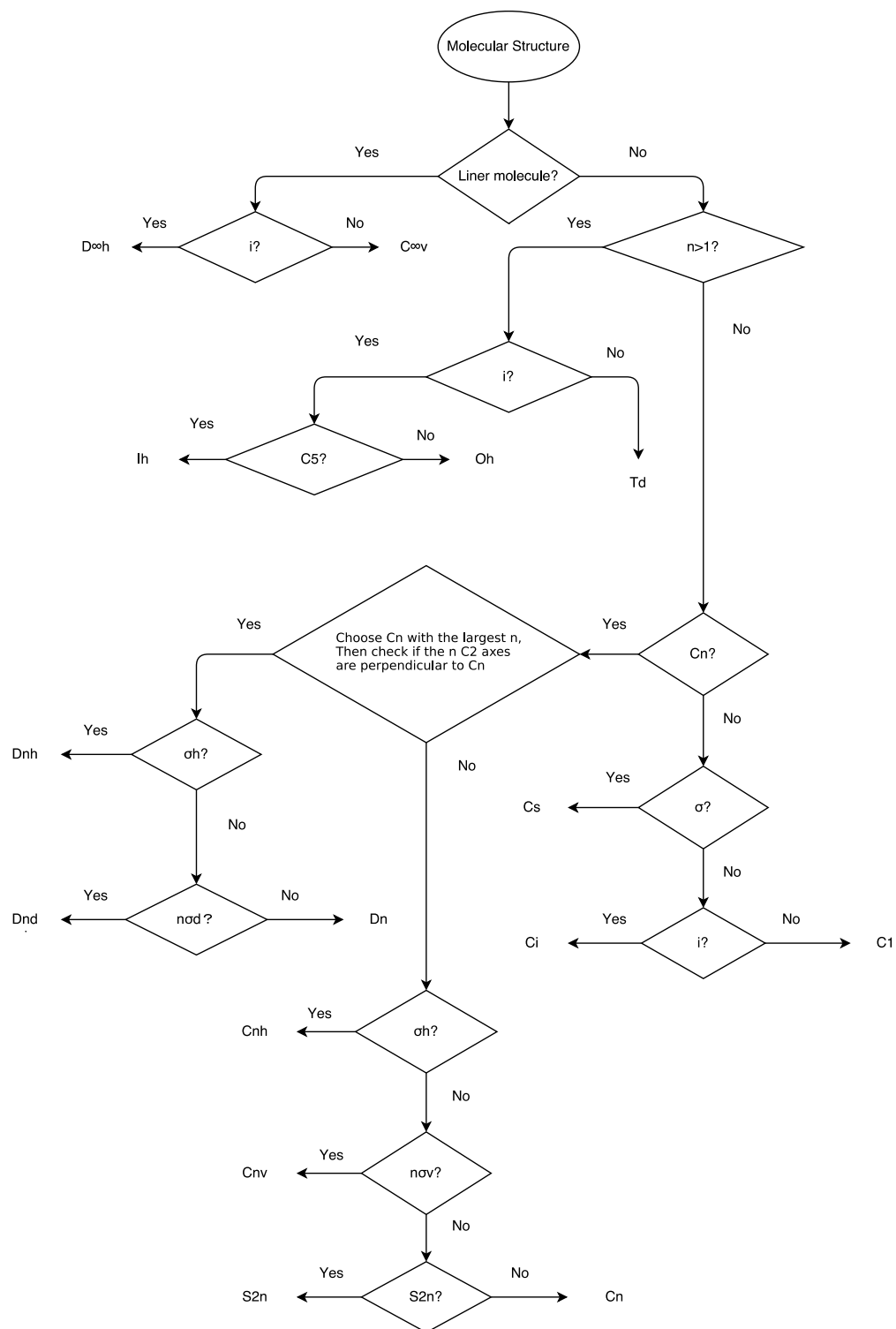


Fig.2.5 Flowchart for determining a point group of a molecule.

2.4 QSR

点群対称性の判定を応用して、分子構造からその分子の擬対称性を判定するプログラムを作成した。このプログラムを擬対称性認識(Quasi Symmetry Recognition, QSR)と名付けた。プログラムのアルゴリズムは第 3 章で説明する。

第 3 章 実装

3.1 プログラムの説明

本節では、QSR で点群対称性を判定するアルゴリズムの説明を行う。

3.1.1 回転対称軸の特定

- (1) 慣性モーメントテンソルを計算し、慣性テンソルを対角化して慣性主軸を決定する。
- (2) 座標軸を慣性主軸に合わせる。その際、特定の軸を新しい座標系の Z 軸にする。

3.1.2 回転対称軸の次数 n の決定

第 3.1.1 節で得られた回転軸の次数 n の決定の仕方は以下のとおりである。

- (1) 同一の Z 座標値を有する
- (2) 同一の原子種である
- (3) Z 軸からその原子までの距離が等しい

これら三つの基準を満たす原子群を一つのクラスにまとめる。これを円形セットと呼ぶ。次にそのクラス内において、そのメンバ数を m とするなら、Z 軸からどれかの原子までをベクトルとみなしてそれを $\frac{360^\circ}{n}$ ずつ回転していく。回転後のベクトルが全てどれかの原子までのベクトルと一致するなら、これは m 次のクラスである。そうでないときは順次 m から 1 ずつ差し引いて同様の処理を行う。

3.1.3 C_2 軸の決定

まず、第 2.3.1 節で説明した慣性主軸を Z 軸とした座標系における XY 平面で円形セットを探索する。ここで C_2 軸は、セット内の任意の原子とセット内の他の 1 つの原子の中間点を通過する必要がある。もし C_2 軸が存在しなかった場合、次に XY 平面にない円形セットから 1 つの原子を選択する。なお、 C_2 軸は、XY 平面への参照原子の投影と、参照セットに対向するセット内の原子の投影によって形成される角度を 2 等分する必要がある。

3.1.4 対称平面

・ σ_h 面の特定

全原子について同一の X,Y 座標値を有する原子群において、Z 座標値については反対符号であることを確認する。

・ σ_v 面の特定

回転対称軸に垂直な正多角形の原子のグループを考える。正多角形の中心からどれか一つの頂点までの切断線ベクトルを求め、その切断線に平行な平面で分子が対称であることを確認する [7]。

3.2 構造データに基づく分子の群論的分類

本節では、QSR のフローチャートを紹介する。QSR ではまず慣性主軸と慣性モーメントの計算を行い、その固有値の縮退度から球型分子、対称頂点分子、直線分子に分類し、それぞれ別のプログラムで点群判定を行う。QSR の全体のフローチャートは Fig.3.1 のとおりである。

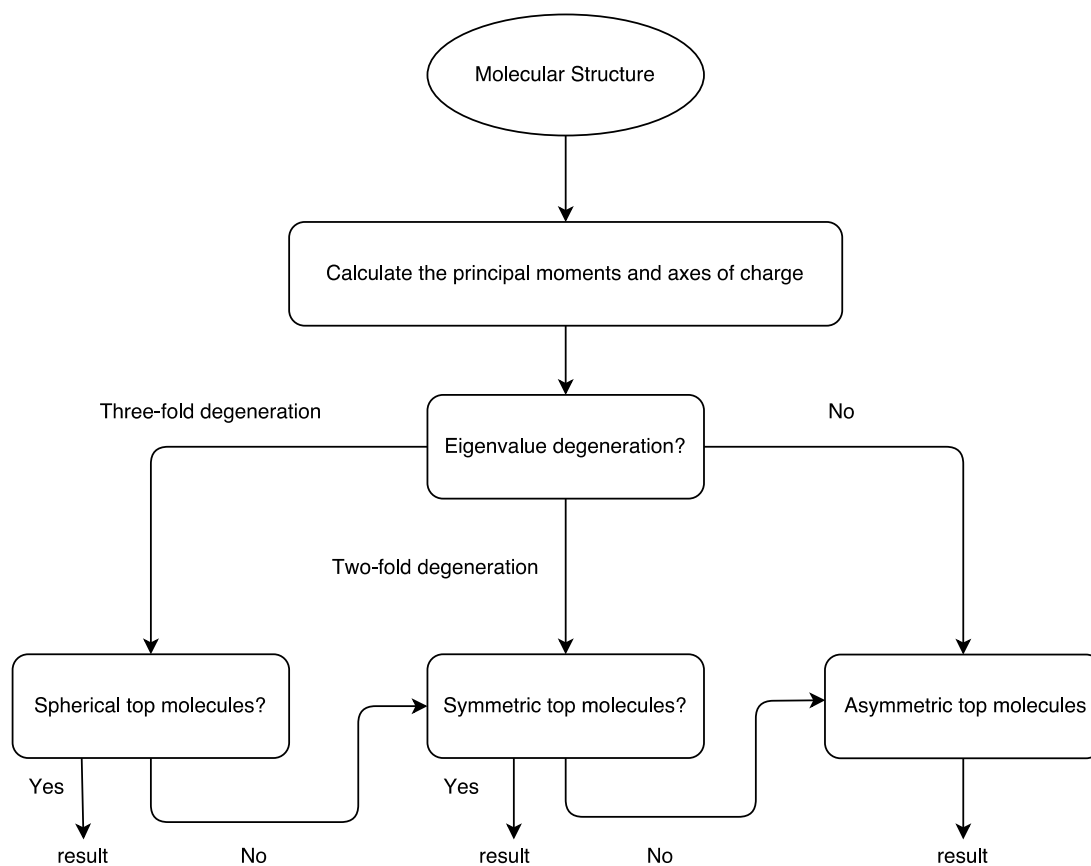


Fig.3.1 Flowchart for determining a point group of a molecule implemented in QSR.

球型分子の判定部分は Fig.3.2 のスキームを用いた。

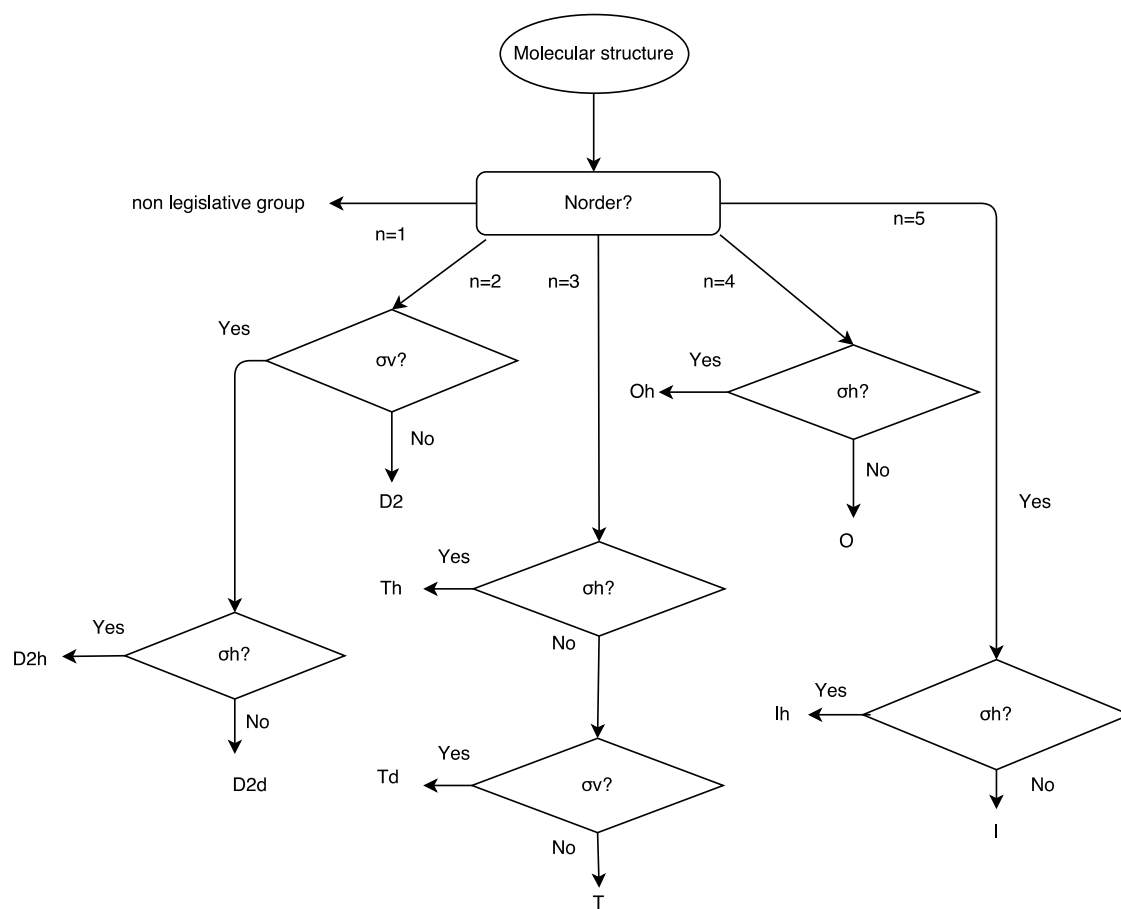


Fig. 3.2 Flowchart for determining a point group of spherical top molecules.

次に、対称頂点分子の判定部分は Fig.3.3 のスキームを用いた。

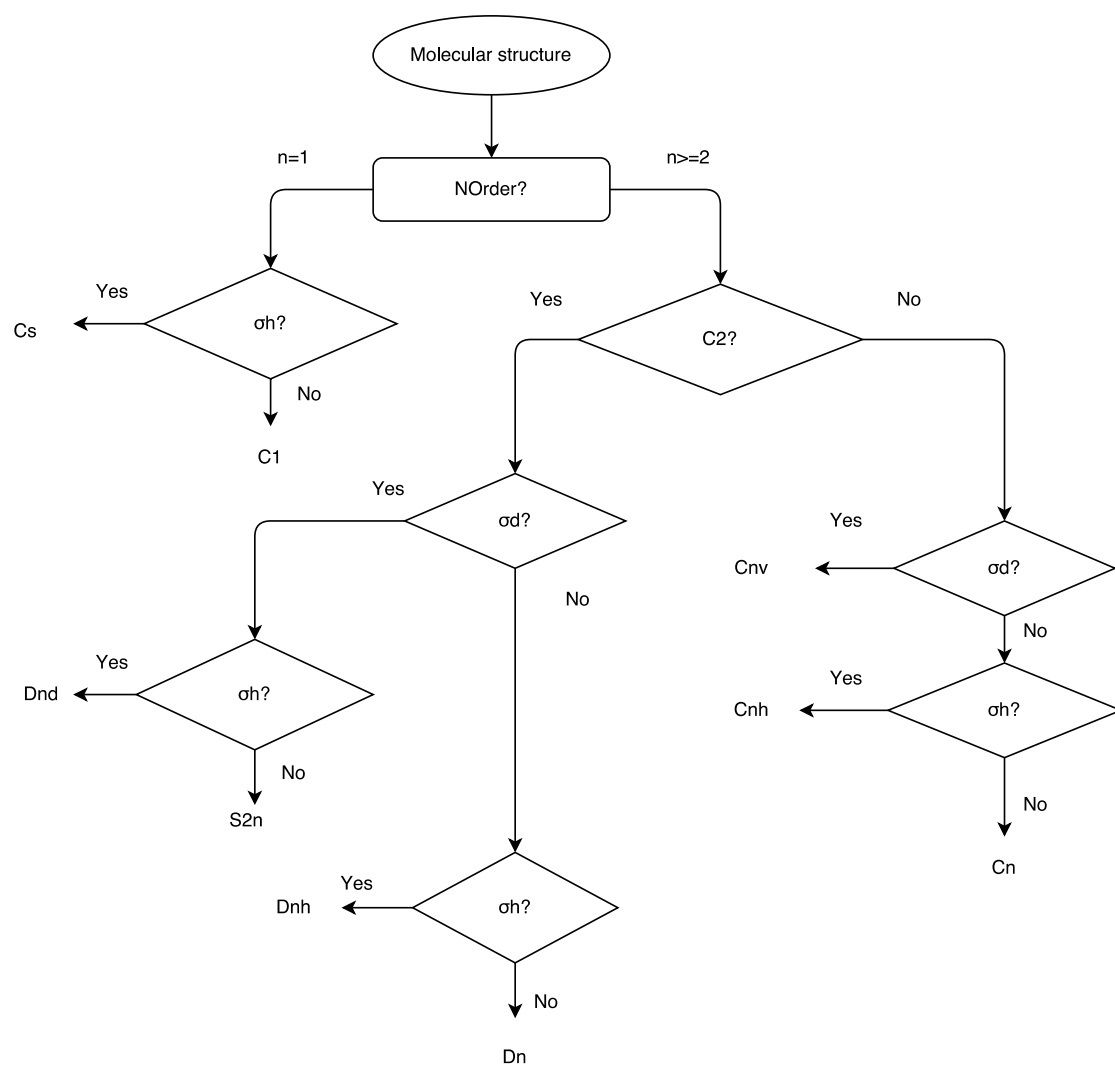


Fig.3.3 Flowchart for determining a point group of symmetric top molecules.

3.3 閾値

本節では、擬対称性の判定のために工夫した閾値について説明する。QSR の判定部分の閾値を段階的に緩和することで、擬対称性の判定が可能かどうか検討した。

3.3.1 決定木

本研究では最適な閾値の推定に決定木(Decision tree)を用いた。決定木は木構造を用いて分類や回帰を行う機械学習の手法の一つである。決定木には対象を分類する問題を解く分類木と対象の数値を推定する問題を解く回帰木がある。今回は回帰木を使用した。

また、より良い決定木モデルを作成するために、bagging と Random Forest を用いた。bagging は、bootstrap 法により弱学習器を構築し、それから最終的な学習器を構築する方法である[8]。bagging のフローチャートを Fig.3.5 に示す。

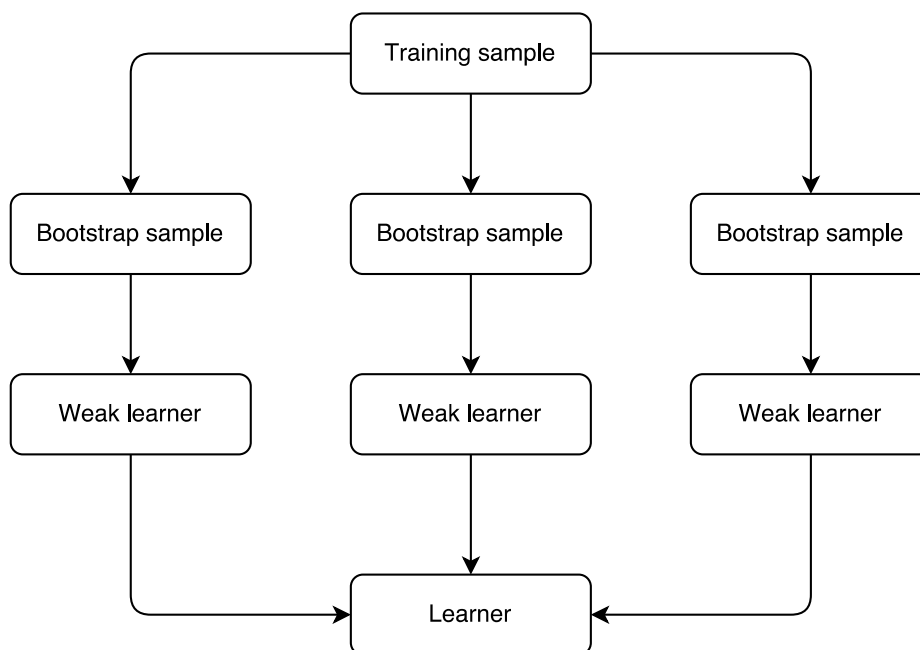


Fig.3.5 Flowchart of bagging

今回決定木を作成した方法は以下のとおりである。

- (1) まず、弱学習器を Random Forest として bagging する。
- (2) その際に各復元抽出サンプル(bootstrap sample)に対して構築された弱学習器の中から、テストデータに対して最もスコアが高い Random Forest モデルを抽出する。
- (3) さらにその Random Forest の中に弱学習器として構築された決定木モデルの中から、テストデータに対して最もスコアが高い決定木モデルを抽出する。

3.3.2 説明変数と目的変数

対称頂点分子について、最適な閾値の範囲を推定した。説明変数 x は閾値の値であり。 $10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 0.05, 0.1, 0.15, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0$ の 17 通りの値を用いた。目的変数 y は、 C_1 群のとき $y = 0$ 、 $C_s, C_n, C_{nv}, C_{nh}, D_n, D_{nh}, D_{nd}$ 群($n = 2, 3, 4$)のとき $y = 1$ 、 $C_5, C_{5v}, C_{5h}, D_5, D_{5h}, D_{5d}$ 群のとき $y = 2$ とした。

3.3.3 最適な閾値の推定

作成した決定木モデルをもとに、閾値の範囲を推定する際にはジニ不純度(Gini impurity)を基準とした。ジニ不純度は、決定木アルゴリズムにおける不純度を算出するための指標として用いられており、ジニ不純度が低いほどノード分けが上手くなされているといえる。

あるノード t 内でクラス i に属するトレーニングサンプルの個数を n_i とすると、クラス i に属するトレーニングサンプルの割合 $p(i|t)$ は、

$$p(i|t) = \frac{n_i}{n} \quad (3.1)$$

である。

ジニ不純度 $I_G(t)$ は $p(i|t)$ を用いて以下の式で表される。

$$I_G(t) = 1 - \sum_{i=1}^c p(i|t)^2 \quad (3.2)$$

第 3.3.2 節における目的変数が 1 のクラスかつジニ不純度が最も低いノードの条件分けを最適な閾値の範囲とした。

第 4 章 結果と考察

4.1 擬対称性の判定

プログラム QSR を使い、その閾値を緩和することで擬対称性を求めることが可能になった。各対称性から代表的な分子を例として Table 4.1 にまとめた。閾値 10^{-7} での判定結果をここでは”厳密対称性”と定義する。

Table 4.1 Quasi symmetry of molecules in Section 4.1

Molecule	Exact symmetry	Quasi-symmetry
CC14	C_1	C_2
Pyridine-ethane	C_1	C_s
AcOH-AcOH	C_1	C_{2h}

4.1.1 CC14

CC14 の擬対称性は C_2 と判定された。Fig. 4.1(a) を見る限り厳密対称性でも C_2 と判定されそうに見えるが、Fig. 4.1(b) を見るとずれているのがわかる。このずれによって厳密対称性は C_1 と判定されてしまうが、閾値を緩和することで C_2 と擬対称性を判定することが可能である。

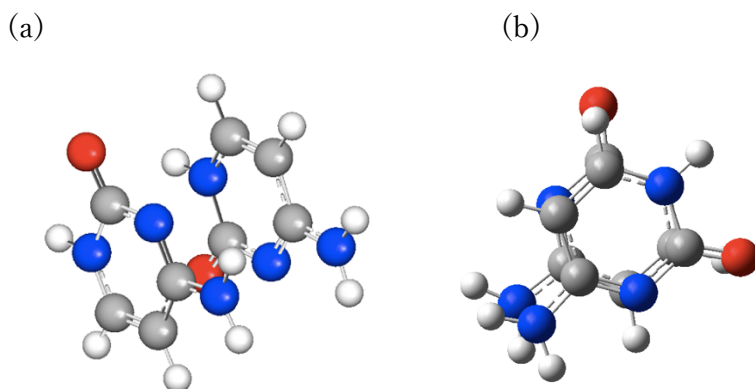


Fig. 4.1 Structure of CC14: (a) side and (b) top views.

4.1.2 Pyridine-ethane

Pyridine-ethane の擬対称性は C_s と判定された。

(a)

(b)

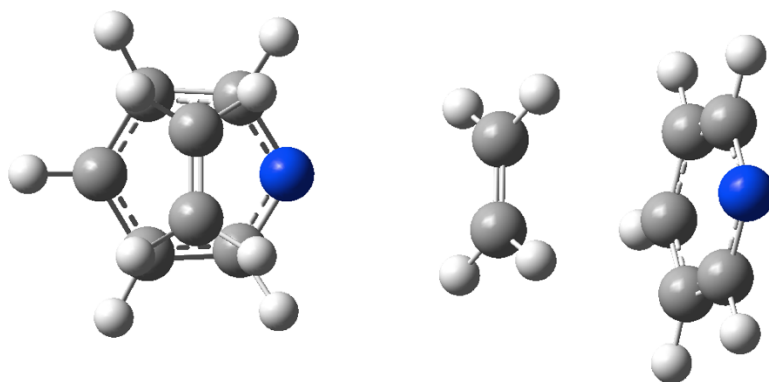


Fig.4.2 Structure of pyridine-ethane: (a) top and (b) side views.

4.1.3 AcOH-AcOH

AcOH-AcOH の擬対称性は C_{2h} と判定された。

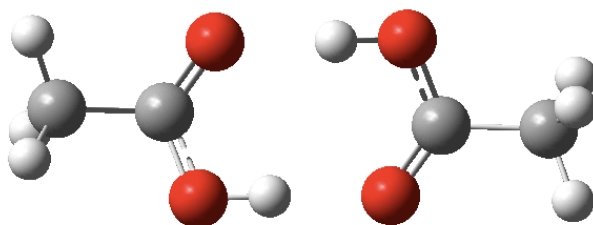


Fig.4.3 Structure of AcOH-AcOH

4.1.2、4.1.3 の分子は例 1 よりも高い対称性に見えるが、厳密対称性では C_1 と判定された。

4.2 閾値が大きい場合

閾値を緩和することで擬対称性の判定が可能になったが、閾値をあまりに大きくすると不適切な判定結果が出てしまう。各対称性から代表的な分子を例として Table4.2 にまとめた。

Table 4.2 Quasi symmetry of molecules in Section 4.2

Molecule	Exact symmetry	Quasi-symmetry
AA1	C_1	C_s
Pyridine-pyridine pipi	C_1	C_{2h}
Benzene-ethane	C_1	C_{2v}

4.2.1 AA1

AA1 の擬対称性は C_s と判定された。さらに閾値を大きくした場合 C_5 という判定が出たが、Fig.4.4 からわかるとおり、見た目には 5 回回転軸に相当する回転軸は存在しない。

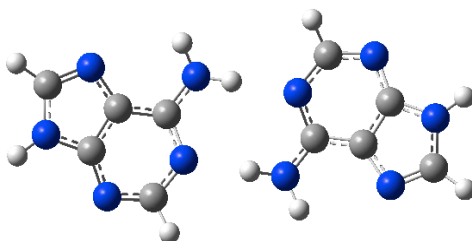


Fig.4.4 Structure of AA1

4.2.2 Pyridine-pyridine pipi

Pyridine-pyridine pipi の擬対称性は C_{2h} と判定された。さらに閾値を大きくした場合 C_{5h} という判定ができるが Fig.4.5 からわかるとおり、こちらも見た目には 5 回回転軸に相当する回転軸は存在しない。

(a)

(b)

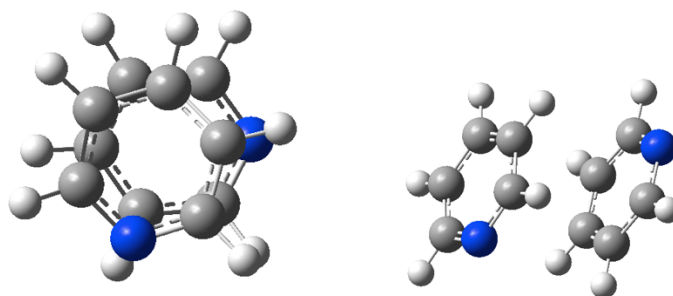


Fig.4.5 Structure of pyridine-pyridine pipi: (a) top and (b) side views.

4.2.3 Benzene-ethane

Benzene-ethane の擬対称性は C_{2v} と判定された。さらに閾値を大きくした場合 C_{5v} という判定が出たが Fig.4.6 からわかるとおり、見た目には 5 回回転軸に相当する回転軸は存在しない。

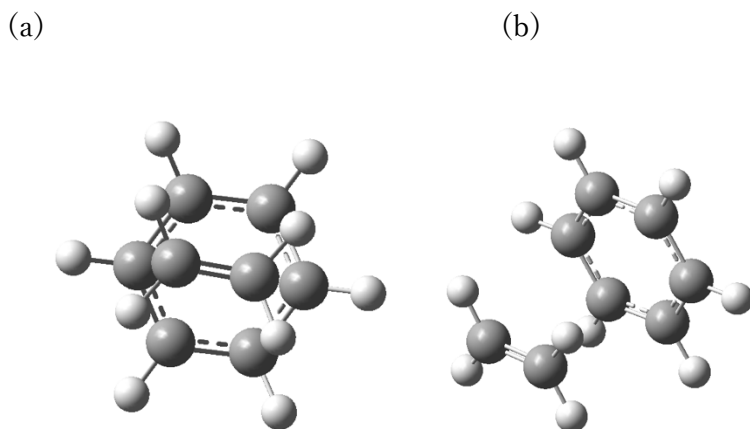


Fig.4.6 Structure of benzene-ethane: (a) top and (b) side views.

4.3 閾値の推定

機械学習で生成した決定木モデルの中からテストスコアが最高のモデルをもとに閾値の推定を行う。3.3.2 で説明した目的変数について $y = 0$ の場合 $\text{class} = C_1$ 、 $y = 1$ の場合 $\text{class} = \text{good}$ 、 $y = 2$ の場合 $\text{class} = C_5, D_5$ として決定木を表示した。作成した決定木のうち C_{2v}, D_{2h}, C_s を取り上げて考察する。

・ C_{2v} の場合

擬対称性が C_{2v} と判定された分子に対する深さ 2 の決定木を Fig.4.7 に示す。

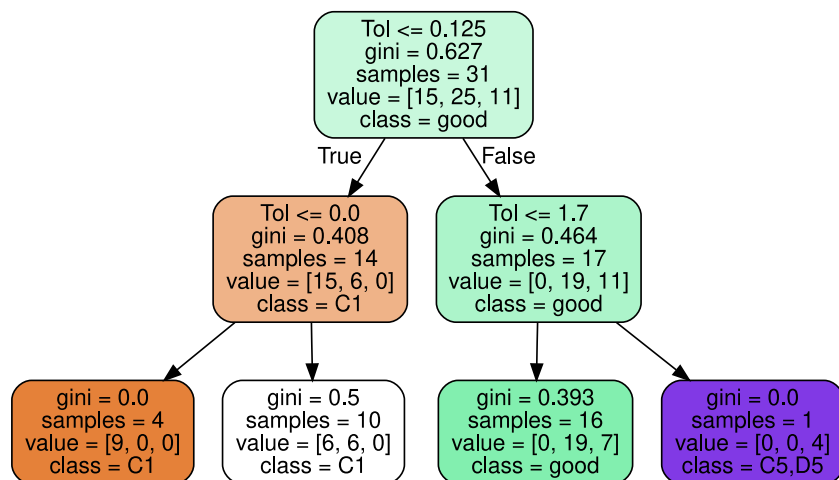


Fig.4.7 Decision tree of threshold for C_{2v} quasi-symmetry (depth=2)

*gini*とはジニ不純度のことであり、値が0に近いほどそのクラスに属する分子の割合が多い。つまり、class=good かつ *gini*が低いものが適切な閾値であると考えられる。ここで *Tol*は閾値のことであり、また、 $Tol \leq 0.0$ の、0.0 は 0.0001 のことである。Fig.4.9を見ると $Tol \leq 0.125$ では C_1 と判定された分子が多く、 $Tol > 1.7$ では 5 回回転と判定されてしまう分子が多く閾値が大きい。

さらに範囲を絞るため決定木の深さを4にしたものが Fig.4.8 である。

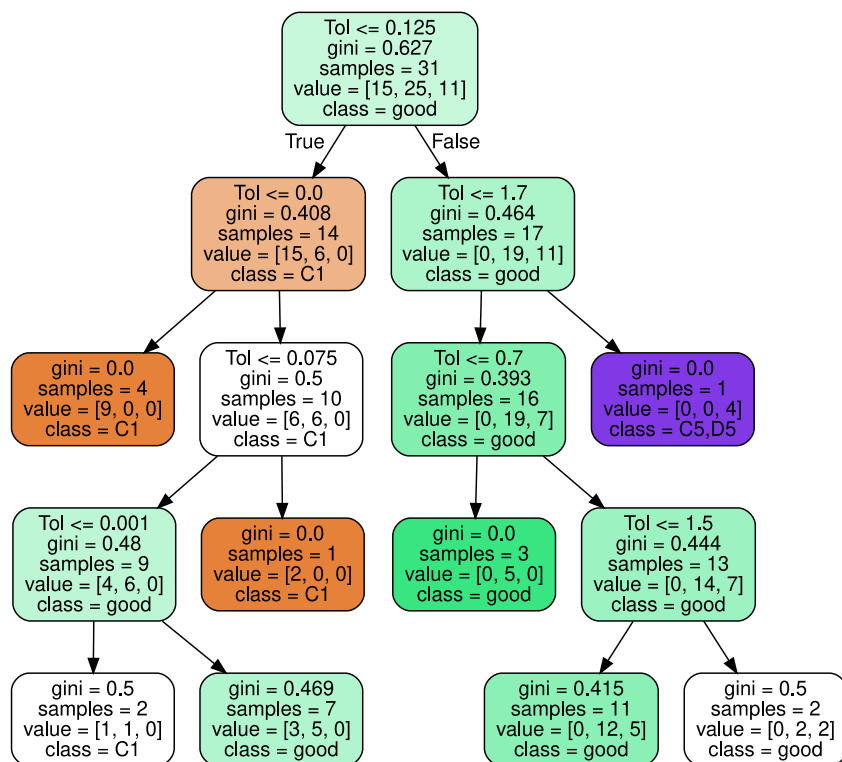


Fig.4.8 Decision tree of threshold for C_{2v} quasi-symmetry (depth=4)

$0.125 < Tol \leq 0.7$ のとき、 $gini = 0.0$ であり最適な閾値の範囲であると考えられる。

・ C_{2h} の場合

擬対称性が C_{2h} と判定された分子に対する深さ 2 の決定木を Fig.4.9 に示す。

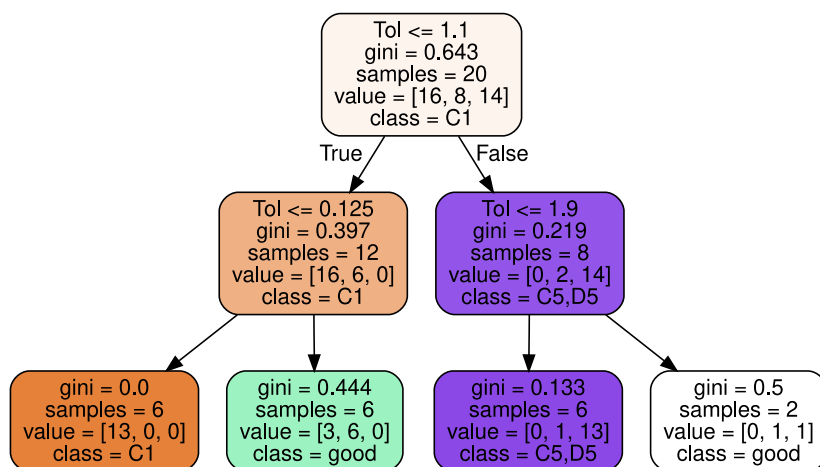


Fig.4.9 Decision tree of threshold for C_{2h} quasi-symmetry (depth=2)

Fig.4.9 を見ると $Tol \leq 0.125$ では C_1 と判定された分子が多く、 $Tol > 1.1$ では 5 回回転と判定されてしまう分子が多く閾値が大きい。

さらに範囲を絞るため決定木の深さを 4 にしたものが Fig.4.10 である。

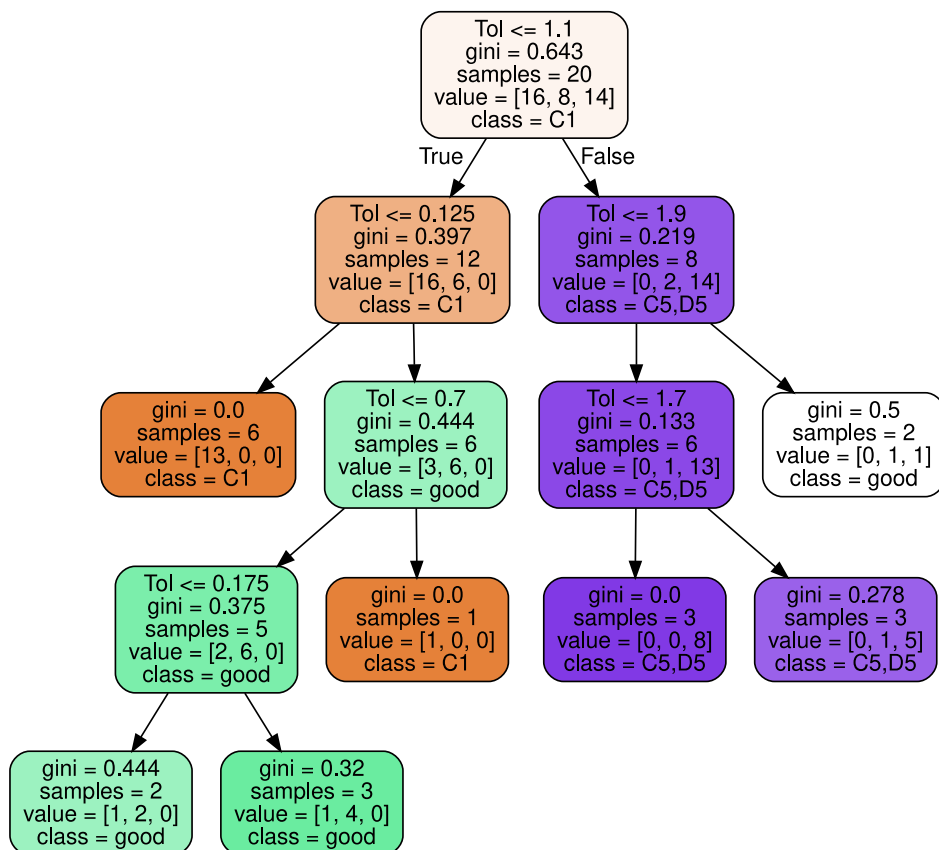


Fig.4.10 Decision tree of threshold for C_{2h} quasi-symmetry (depth=4)

$0.175 < Tol \leq 0.7$ のとき、 $gini = 0.32$ であり最適な閾値の範囲であると考えられる。

・ C_s の場合

擬対称性が C_s と判定された分子に対する深さ 2 の決定木を Fig.4.11 に示す。

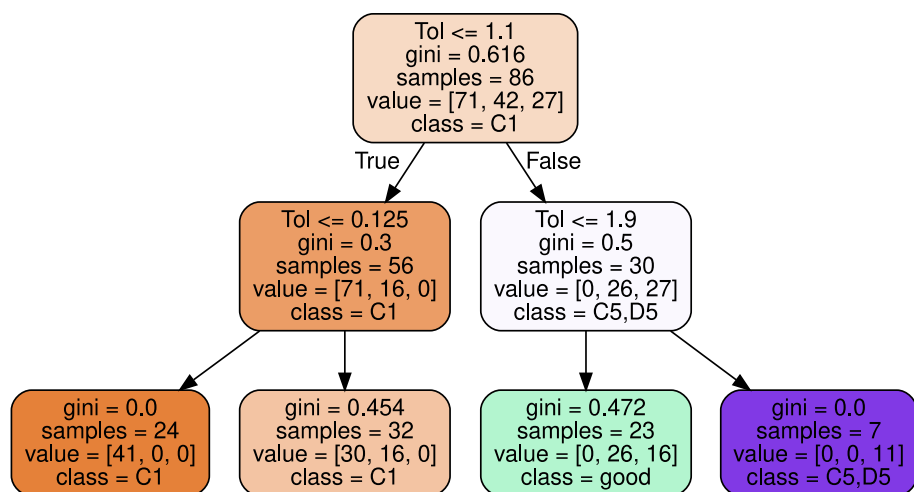


Fig.4.11 Decision tree of threshold for C_s quasi-symmetry (depth=2)

擬対称性 C_s の分子は擬対称性が認証される閾値の値が比較的大きい。

さらに範囲を絞るため決定木の深さを4にしたものが Fig.4.12 である。

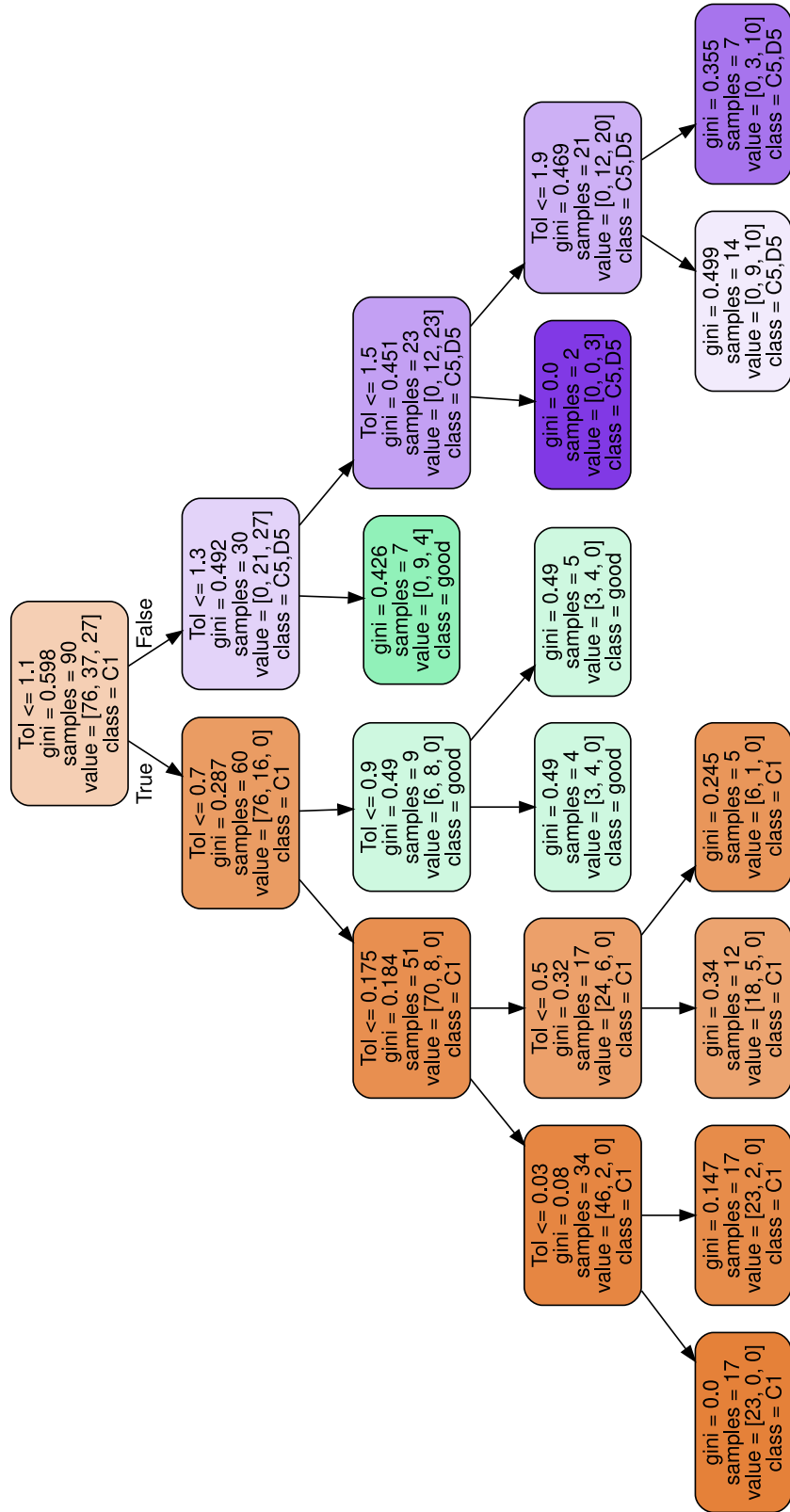


Fig.4.12 Decision tree of threshold for C_s quasi-symmetry (depth=4)

Fig.4.12 を見ると $class = good$ のもので比較的 $gini$ が低いのは、 $1.1 < Tol \leq 1.3$ の場合と $0.7 < Tol \leq 0.9$ の場合であり、それぞれ $gini = 0.426$ と $gini = 0.49$ である。

擬対称性が C_{2v} , C_{2h} を有する分子群について正しい擬対称性が判定される閾値の範囲の幅は広い。これは、正しい擬対称性が判定され始める閾値の値は個々により異なることが原因である。一方、 C_5 や D_5 といった判定が出始める閾値の値は概ね等しいという結果が得られた。どの分子群においても $Tol \leq 0.9$ では 5 回回転対称と判定される分子はなかった。

以上のことから、閾値の下限は個体差により明確に定められないが、上限は 0.9 が好ましいと考えられる。

4.4 擬対称性が C_1 である分子

調べた分子の中には 4.3 で調べた閾値の範囲の中で擬対称性が C_1 のままであるものがあつた。このような分子は分子構造を見ても全く対称性がない。Fig.4.13 と Fig.4.14 に擬対称性が C_1 の分子構造を示す。

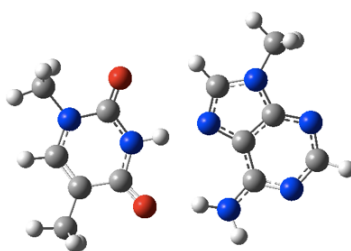


Fig.4.13 Structure of mAmTH

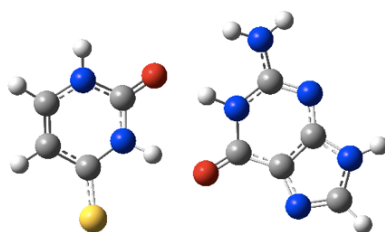


Fig.4.14 Structure of G4thioUwobble

C_{2v} や C_{2h} といった高い擬対称性だけでなく擬対称性が C_1 であるという結果も軌道エネルギーについて議論する際に重要である。また、分子計算をする際の分子のスクリーニングに用いることが可能である。

4.5 軌道エネルギーとの関係

4.5.1 小分子の軌道エネルギー

小分子について得られた対称性ごとに軌道エネルギーを表示した図をFig4.15~Fig4.20に示す。図について空白を挟んで右側が厳密対称性を有する分子、左側が擬対称性を有する分子である。

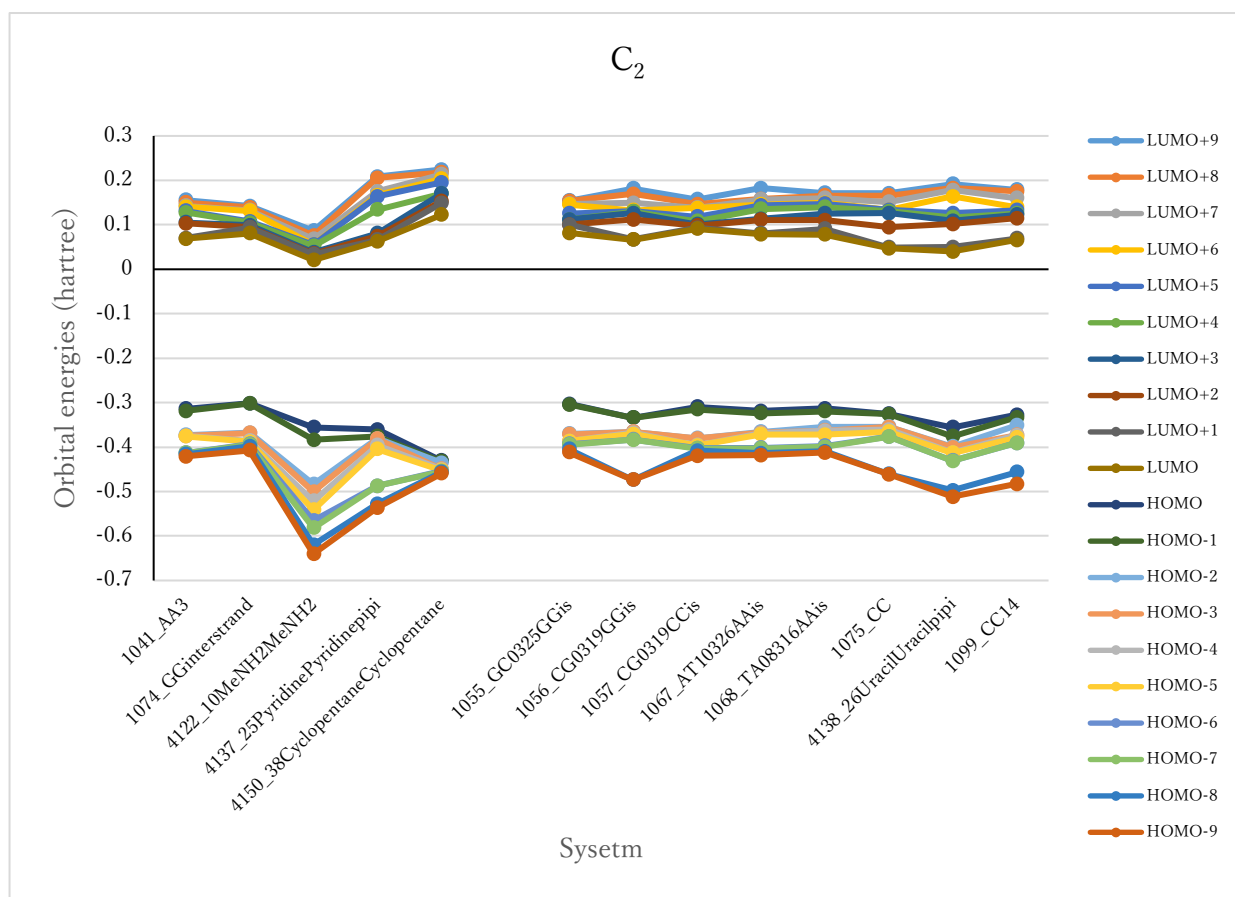


Fig.4.15 Valence orbital energy levels of molecules assigned as C_2 quasi-symmetry

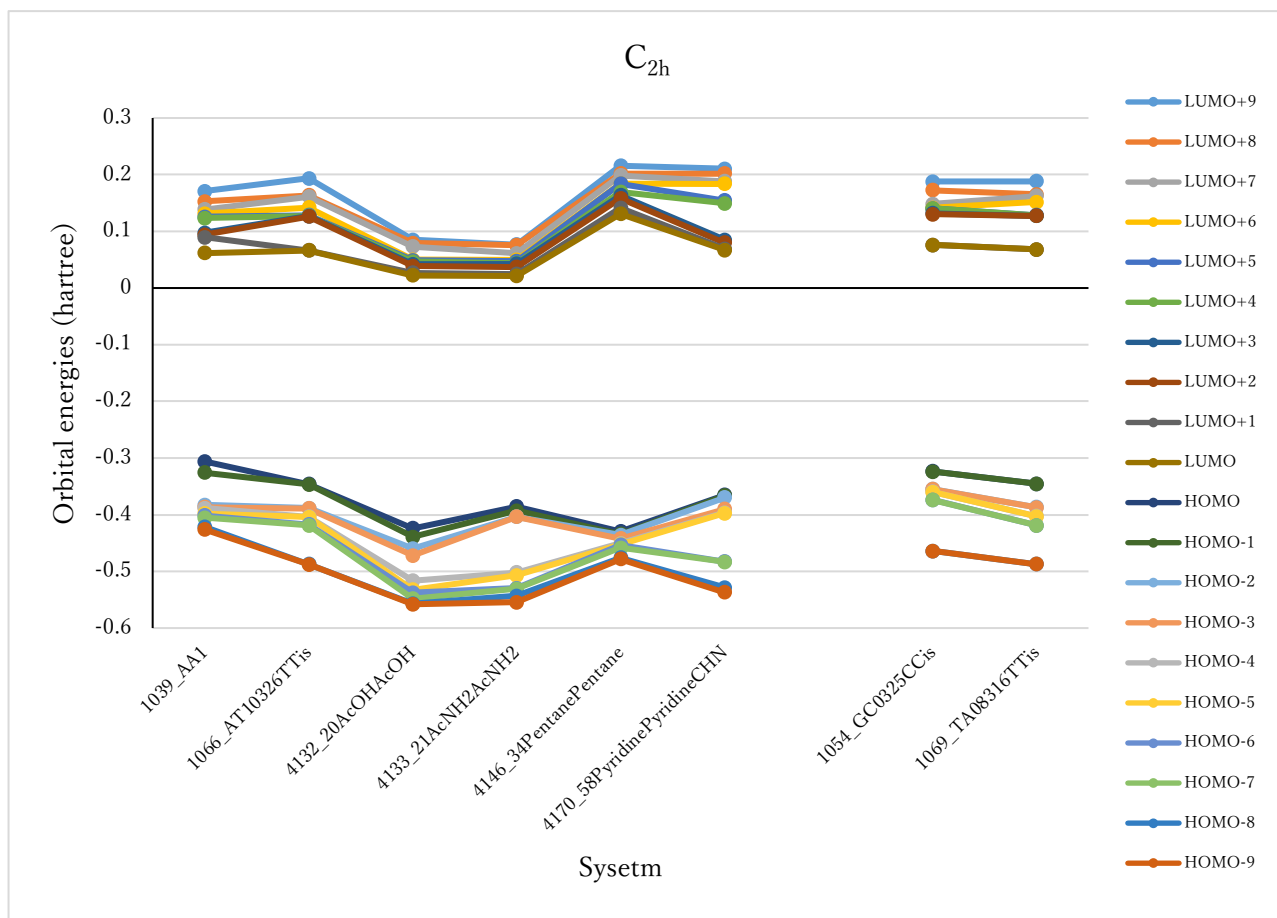


Fig 4.16 Valence orbital energy levels of molecules assigned as C_{2h} quasi-symmetry

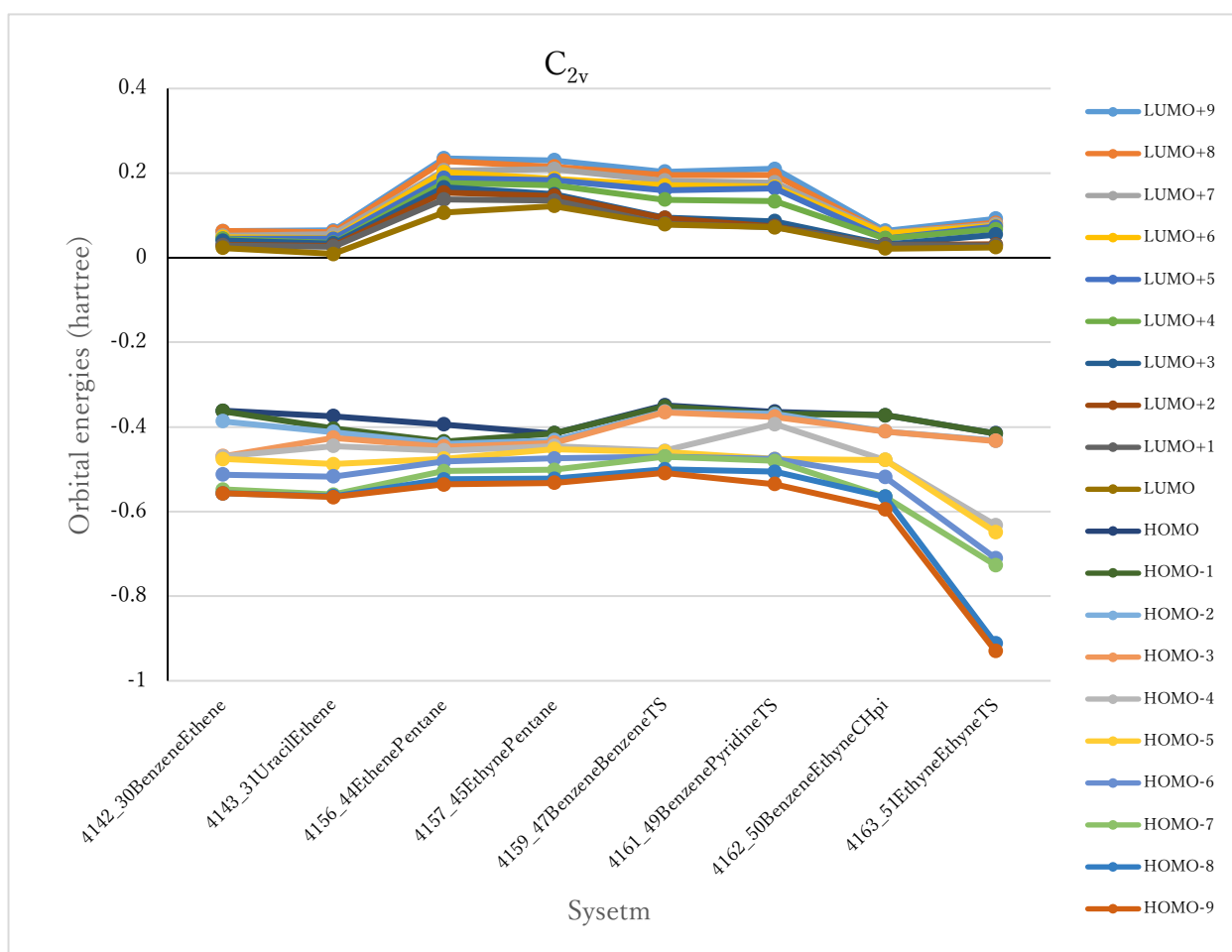


Fig 4.17 Valence orbital energy levels of molecules assigned as C_{2v} quasi-symmetry

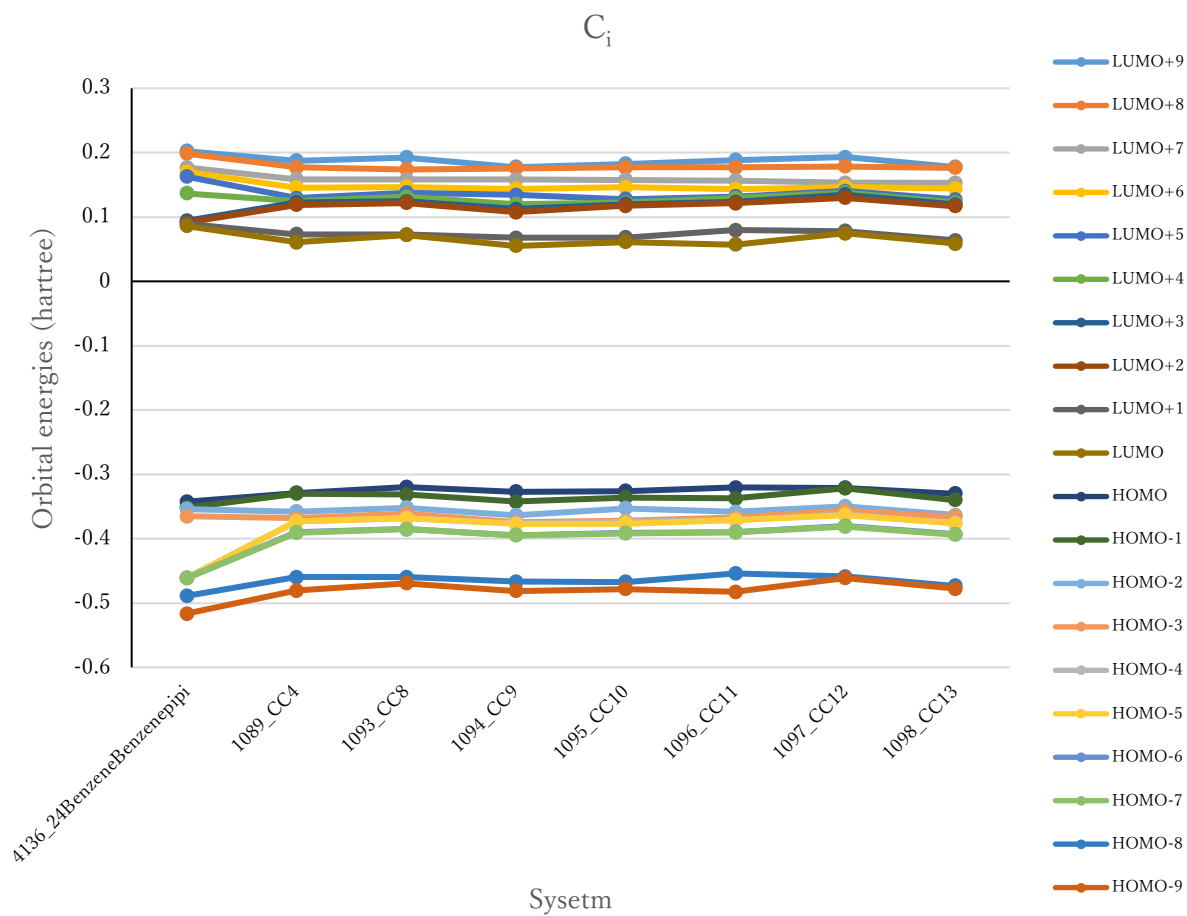


Fig 4.18 Valence orbital energy levels of molecules assigned as C_i quasi-symmetry

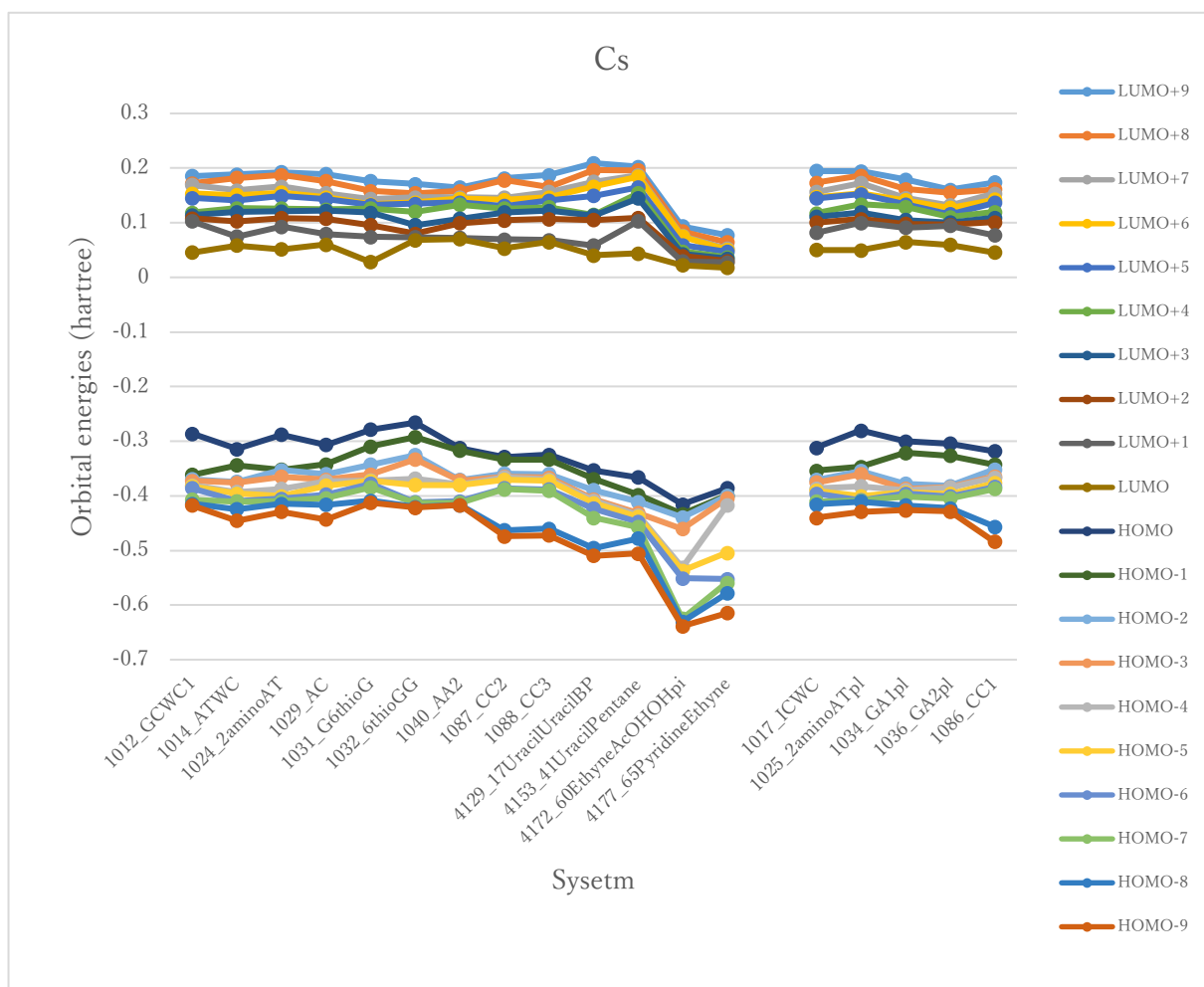


Fig 4.19 Valence orbital energy levels of molecules assigned as C_s quasi-symmetry

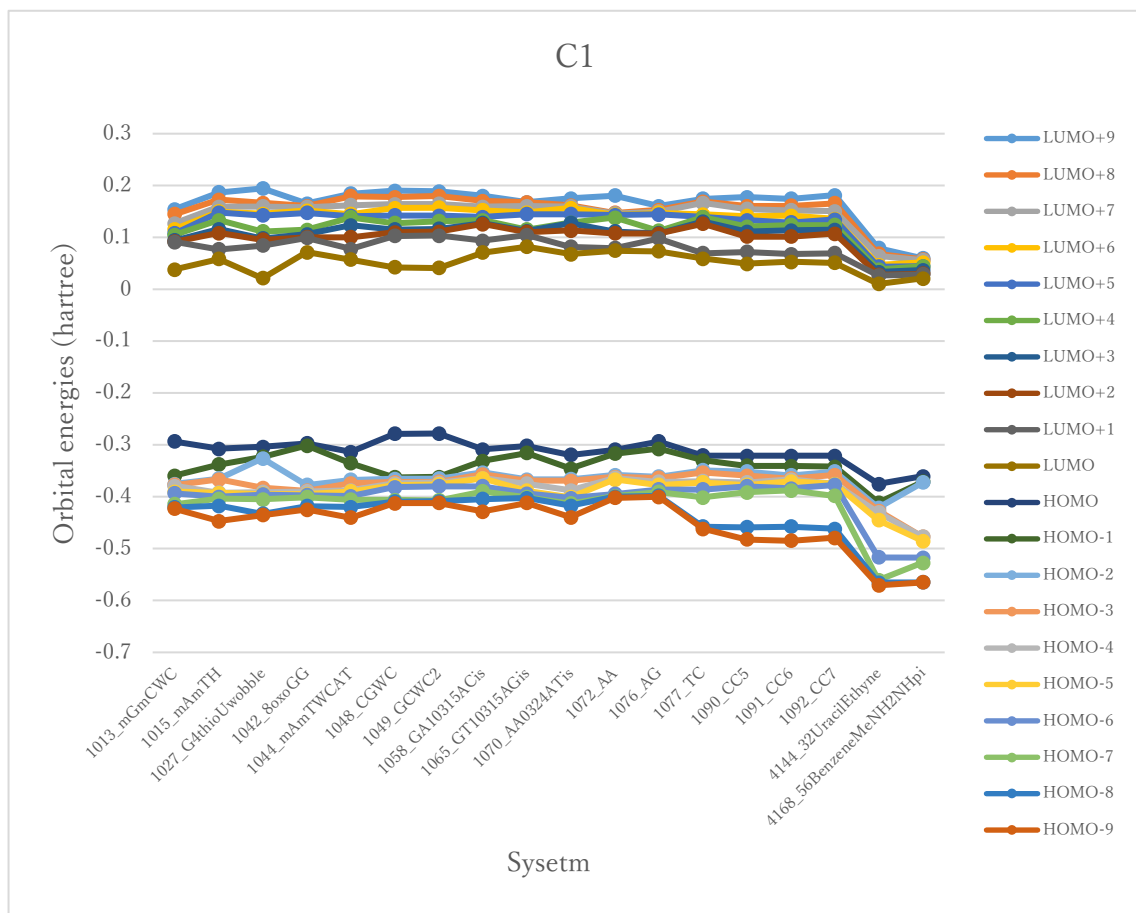


Fig 4.20 Valence orbital energy levels of molecules assigned as C_1 quasi-symmetry

C_2 , C_{2h} , C_{2v} は HOMO と HOMO-1、および LUMO と LUMO+1 がそれぞれ、縮重あるいはかなりエネルギーが近い擬縮重をしている。それに対して C_i や特に C_1 , C_s では縮重はほとんど見られない。また、 C_{2v} は LUMO+5 と LUMO+6 など HOMO と HOMO-1、および LUMO と LUMO+1 以外の軌道にも縮重が見られる。擬対称性が高い分子の軌道エネルギーほど縮重が見られることから、厳密対称性の場合と同様に軌道エネルギー計算なしに軌道エネルギー準位を予想することが可能になる。

4.5.2 銀クラスターの軌道エネルギー

銀クラスターについて得られた対称性ごとに軌道エネルギーを表示した図を Fig.4.21~Fig.4.23 に示す。

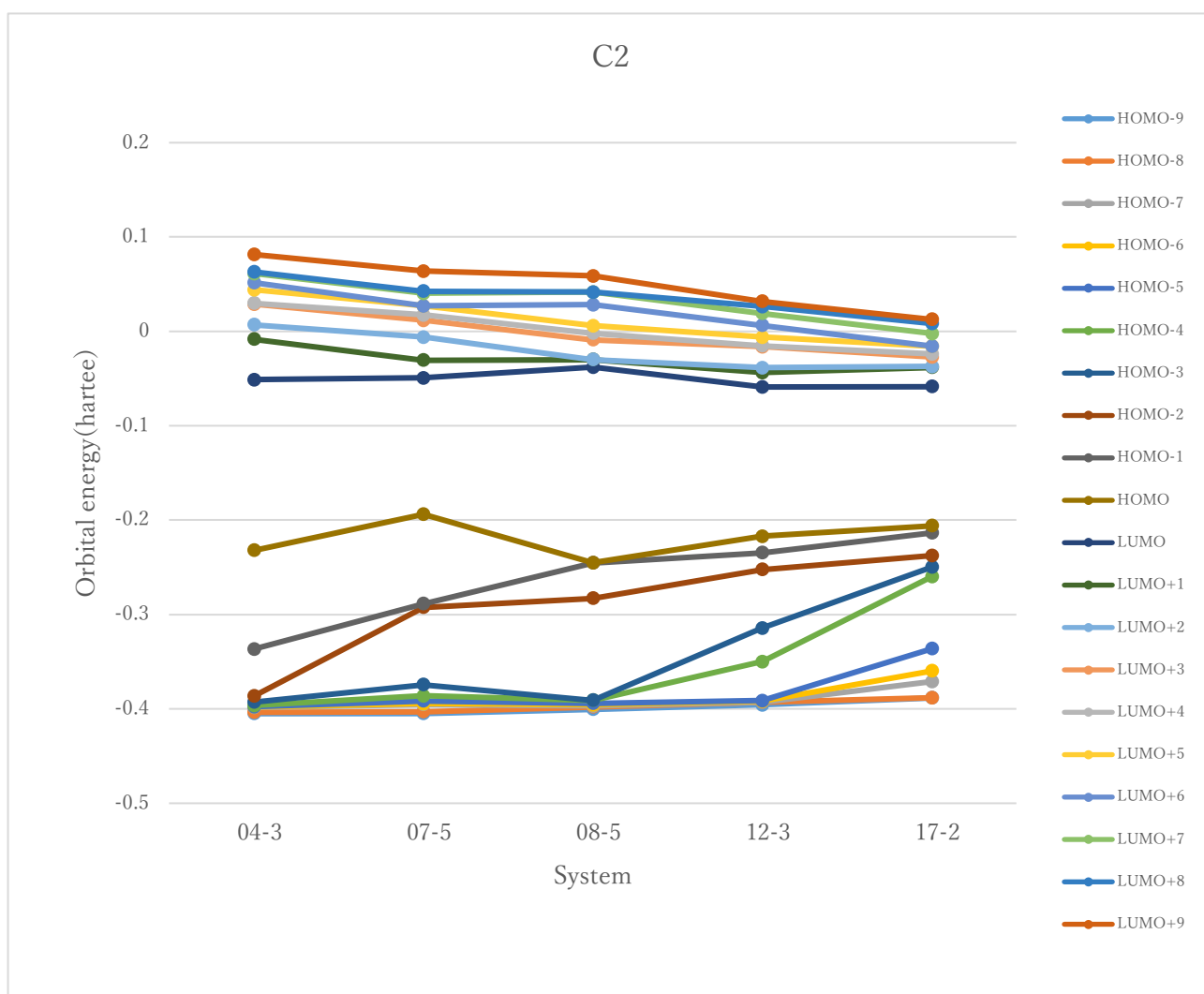


Fig 4.21 Valence orbital energy levels of Ag clusters assigned as C_2 quasi-symmetry

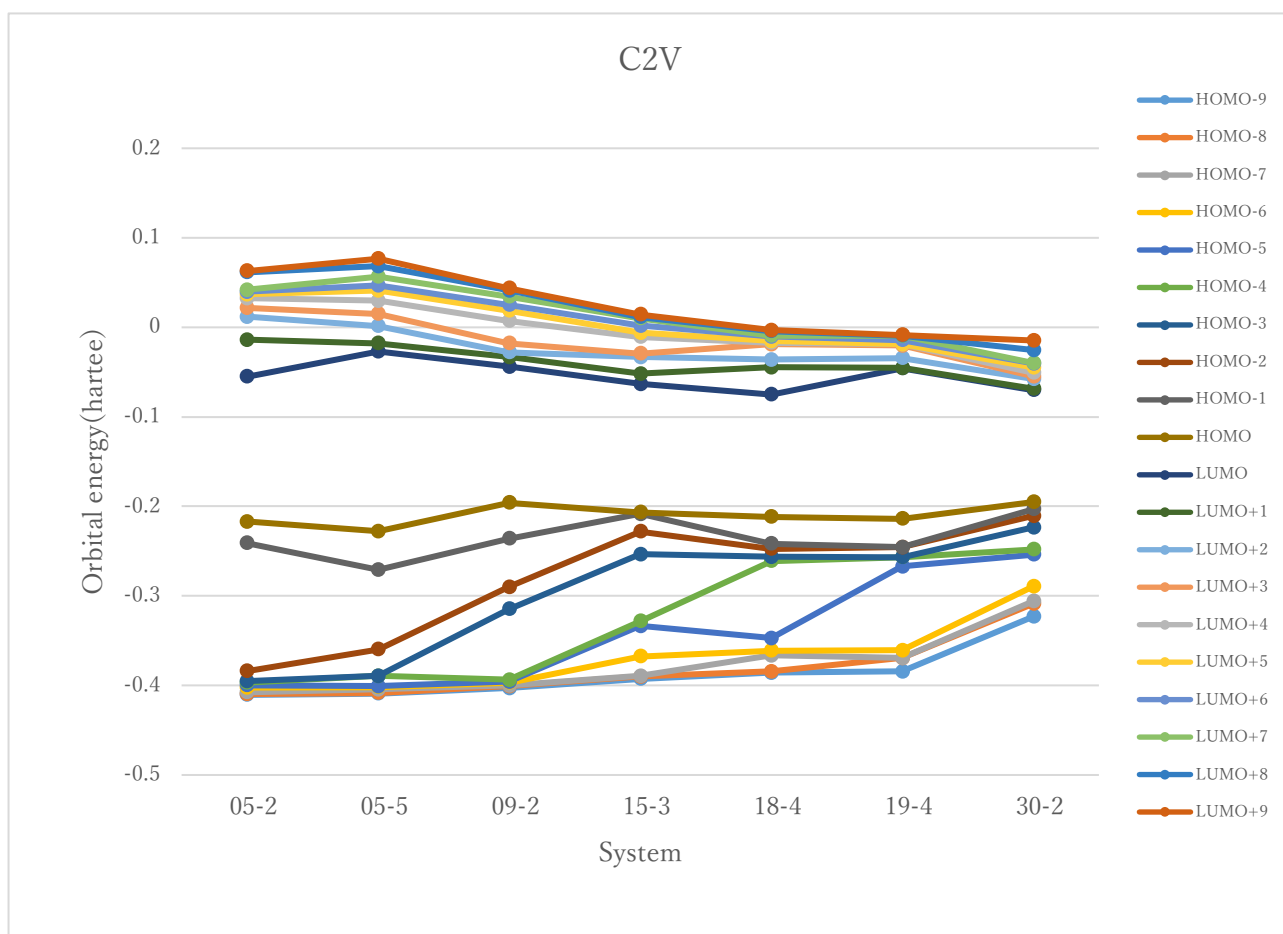


Fig 4.22 Valence orbital energy levels of Ag clusters assigned as C_{2v} quasi-symmetry

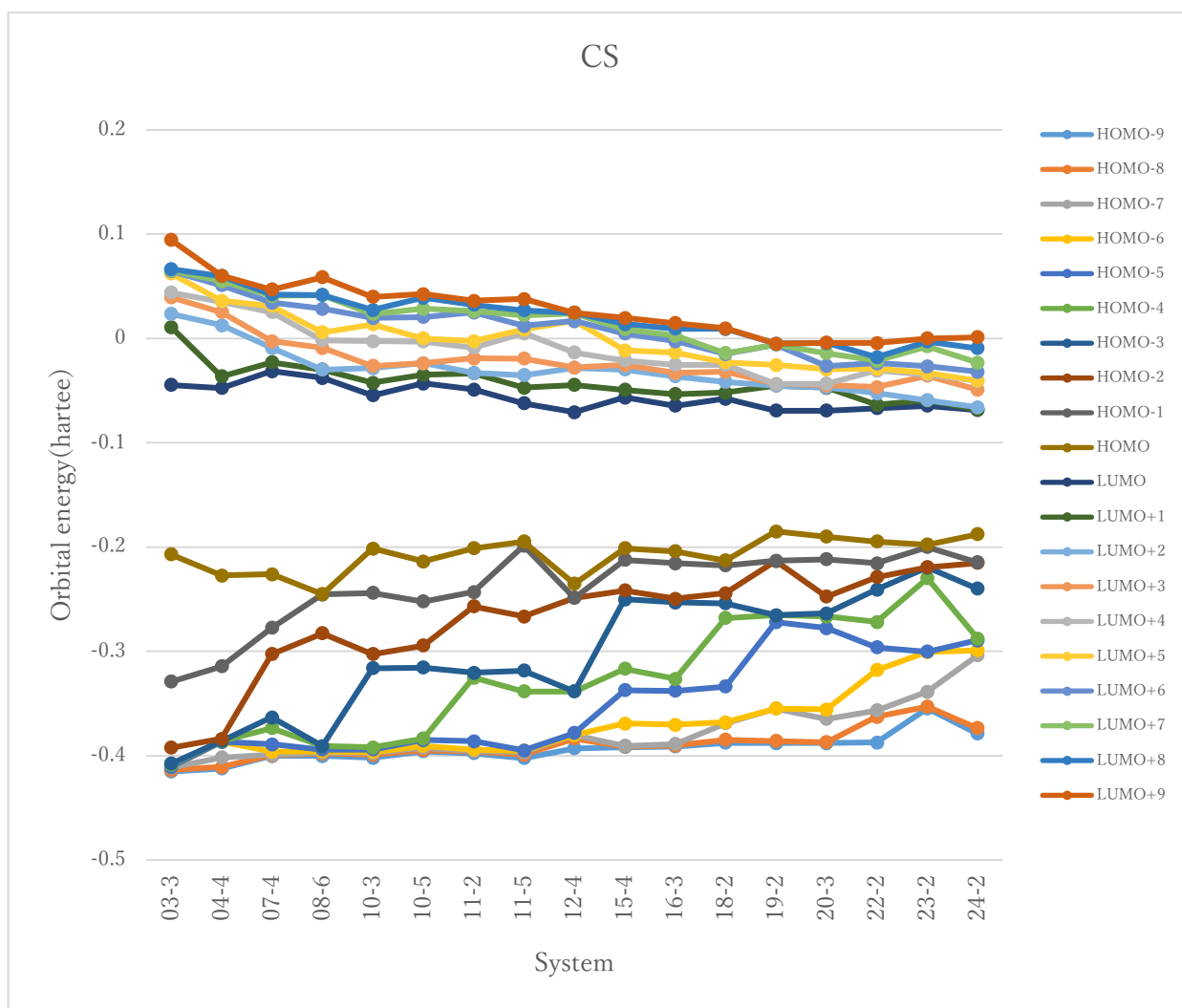


Fig 4.23 Valence orbital energy levels of Ag clusters assigned as C_s quasi-symmetry

銀クラスターでは擬対称性ごとに分類しても HOMO-LUMO 軌道での軌道エネルギーの縮重は見られなかった。また、通常は金属クラスターを分子サイズごとに並べた場合 Fig4.23 のように軌道エネルギーが上下しながら 0 に漸近していくが、 C_2 や C_{2v} のように高い対称性が見られる場合は滑らかに 0 に漸近していく。つまり、軌道エネルギーの分子サイズ依存性が明確になった。

第 5 章 結論

本研究では、点群対称性を判定するプログラムを応用して擬対称性を判定するプログラム QSR を作成した。そして、擬対称性を求めるにあたって適切な閾値を推定した。結果として、一部の小分子について擬対称性を判定することができた。また、決定木を用いて調べた閾値の範囲について下限は個体差があるため明確には定められないが、閾値を緩和する上限は 0.9 が好ましいという結果が得られた。

得られた擬対称性から三つの結果が得られた。一つ目は、擬対称性が C_1 の分子の構造には全く対称性がないため、これは分子計算をする際の分子のスクリーニングに用いることが可能である。二つ目は小分子において擬対称性ごとに軌道エネルギー準位が縮重している。三つ目は、銀クラスタ分子は擬対称性で分類すると軌道エネルギーの分子サイズに対する依存性が明確になった。

今回は小分子と銀クラスタの擬対称性を調べたが、今後は金属錯体や銀以外の金属クラスタの擬対称性について調べたい。金属錯体には厳密対称性がなく、d軌道で分子の分類を行い解析などに活かしている。金属クラスタは、構造全体の擬対称性が原因で軌道エネルギー準位に規則性がみられ、それは磁性などの物性に影響を与えている。擬対称性がわかればその分類や物性に与える影響の裏付けをすることや、より正確な分類が可能になると考えられる。また、金属錯体や金属クラスタはそれぞれ構造が類似しているため、小分子で検証した閾値よりも狭い範囲で適切な閾値の推定が可能になると考えられる。また、含有する原子種が似ているので軌道エネルギー準位との関係も見出せると考えられる。

謝辞

本研究及び本論文を作成するに際して、ご指導を頂いた天能精一郎教授ならびに常田貴夫特命教授、上島基之先生に深謝いたします。また研究室生活で多くの助言をしていただいた CS51 の皆様に感謝いたします。

参考文献

- [1] Gilbert N. Lewis, “The Atom and the Molecule”, J. Am. Chem. Soc. 38(4) (1916), pp762-785.
- [2] W. Heitler and F. London, “Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik”, Zeitschrift für Physik , 44(1927), pp455-472.
- [3] E. Wigner, “Über nicht kombinierende Terme in der neueren Quantentheorie”, Zeitschrift für Physik , 40 (1926), pp492-500.
- [4] H. Bethe, “Termaufspaltung in Kristallen.”, Ann. d. Phys., 3 (1929), pp133.
- [5] 犬井 鉄郎, 田辺 行人, 小野寺 嘉孝, “応用群論 –群論表と物理学–”, 裳華房 (1991), pp1-7.
- [6] P. W. Atkins, “アトキンス物理化学(上)第6版”, 東京化学同人(2001), pp461-468.
- [7] 矢野 敬幸, 吉村 康子, “パソコンによる分子の対称要素の発見と点群への分類”, Hitotsubashi University Repository(1993), pp380-385.
- [8] 塚本邦尊, 山田典一, 大澤文孝, “東京大学のデータサイエンティスト育成講座 python で手を動かして学ぶデータ分析”, マイナビ出版(2019), pp288.