# 1 Introduction

## 1.1 On Notation

I would like this section to serve as a reference the notation used throughout the document. (May be later removed or replaced by a list of acronyms.)

Let the labels i,j,k... correspond to occupied orbitals, a,b,c... correspond to virtual orbitals, and p,q,r... correspond to any orbital. Active orbitals are denoted with a prime.

## 1.2 Overview of the Problem

# 2 Theory

## 2.1 Computational Methods

### 2.1.1 Coupled Cluster

Consider the Slater determinant for a doubly occupied 1s orbtal:

$$|\Phi_{1s}\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_{1s}(x_1) & \bar{\phi}_{1s}(x_1) \\ \phi_{1s}(x_2) & \bar{\phi}_{1s}(x_2) \end{vmatrix} \tag{1}$$

Spin states are orthonomral, and the probability amplitudes for some arbitriary point $(x_1, x_2)$ can be written

$$\mathcal{P}(x_1, x_2) = |\phi_{1s}(x_1)|^2 \, |\bar{\phi}_{1s}(x_2)|^2 \tag{2}$$

That is, if the two electrons have opposite spin, the position of one electron does not depend on the other[?] i.e. the motion of these electrons are uncorrelated. These effects typically account for less than one percent of the total energy, and yet, are vital in the prediction of accurate thermo-kinetic properties. For example, the total energy of a water molecule is on the order of 184,000 $\mathrm{kJ\,mol^{-1}}$, while the heat of formation of water is only -285 $\mathrm{kJ\,mol^{-1}}$, 0.15% of the total energy. In the literature, The difference between the exact energy and the HF energy is refereed to as the correlation energy

$$E_{corr} = E_{exact} - E_{HF} \tag{3}$$

The above description of correlation energy is incomplete, as it suggests that HF would be exact for some excited state with all elections spin aligned. This of course is not the case, and the energy discrepancy arises from the reality that a wave function cannot be perfectly represented by a single Slater determinant. To combat this, post HF methods are formulated with respect to a basis of Slater determinants accessible by the HF reference state. Among the most successful approaches for solving the electron correlation problem the are coupled cluster (CC) theories. The couple cluster wave function is compactly represented by the exponential ansatz

$$|\Psi_0\rangle = e^{\hat{T}} |\Psi_{HF}\rangle \tag{4}$$

Where the cluster operator, $\hat{T}$, is defined to be the sum of $n$-electron excitation operators.

$$T = T_1 + T_2 + ... \tag{5}$$

$$= \sum_{a,i} t_i^a \{a_i a_a^\dagger\} + \sum_{i,j,a,b} t_{ij}^{ab} \{a_i a_j a_a^\dagger a_b^\dagger\} \tag{6}$$

Where $a_i$ and $a_a^\dagger$ are the second-quantized electron creation and anhilation operators, respectively. The multi-determinant nature of CC can be seen via expansion of (4)

$$e^{\hat{T}} |\Psi_{HF}\rangle = (1 + T_1 + T_2 + (T_1)^2 + T_1 T_2 + (T_2)^2) |\Psi_{HF}\rangle ... \tag{7}$$

$$= |\Psi_{HF}\rangle + \sum t_a^i |\Psi_i^a\rangle + \sum t_{ab}^{ij} |\Psi_{ij}^{ab}\rangle \dots \tag{8}$$

Where the orbital subscript and superscript correspond to orbitals whose occupation differs from the HF reference. The product of low-order excitations introduce high-order effects in the truncated result, e.g. simultaneous double excitations $(T_2)^2$ introduce corrections from quadruply excited determinants. This expansion is finite but grows as $N!$ with respect to the number of orbitals, and in practice must be truncated. The CC energy equation holds for any level of truncation:

$$E_{cc} = E_{HF} + \underbrace{\sum f_{ai} t_i^a}_{T_1} + \underbrace{\sum \langle ij|ab\rangle \, t_i^a t_j^b}_{(T_1)^2} + \underbrace{\sum \langle ij|ab\rangle \, t_{ij}^{ab}}_{T_2} \tag{9}$$

Naively, one might assume that because the CC energy depends directly on $T_1$ and $T_2$, truncating T at the singles plus doubles level will suffice. However higher order excitations indeed contribute to the energy, indirectly, though $T_1$ and $T_2$. An example is most easily shown though Goldstone diagrams, which are included in fig. 1 for completeness

$T_3 : \langle ij|ab\rangle\, t_{ijk}^{abc} \rightarrow T_1 \qquad T_4 : \langle kl|cd\rangle\, t_{ijkl}^{abcd} \rightarrow T_2$

Figure 1: Higher order excitations contribute to the energy indirectly by altering $T1$ and $T2$

To improve performance, corrections from higher-order t-amplitudes can be calculated via perturbation theory. This mixing of coupled cluster and perturbation theory is used extensively throughout quantum chemistry, and CC singles and doubles with purtubative tipples (CCSD(T)) is currently considered the 'gold-standard' to which other methods can be compared.

As we are mainly interested in the low-energy part of the spectrum, it can be useful to define an effective hamiltonian, $\bar{H}$, as a similarity transform under $e^T$,

$$\bar{H} = e^{-T} H e^{T} \tag{10}$$

$$\bar{H}\Psi = E\Psi \tag{11}$$

Such transformations leave eigenvalues unchanged. The coupled cluster equations which must be solved to obtain the t-amplitudes are generated via left-projection of excited states onto 11.

$$\begin{aligned} \langle \Phi | \bar{H} | \Phi_0 \rangle &= E_{CC} \\ \langle \Phi_i^a | \bar{H} | \Phi_0 \rangle &= 0 \\ \langle \Phi_{ij}^{ab} | \bar{H} | \Phi_0 \rangle &= 0 \\ &\vdots \end{aligned} \tag{12}$$

Coupled cluster is size consistent for any level of truncation. That is, for two systems A and B,

$$E_{AB} = E_A + E_B \tag{13}$$

The main drawback of CC is their prohibitive computational complexity. CCSD, and CCSD(T) formally scale as $\mathcal{O}(n_{occ}^2 n_{vrt}^4)$ and $\mathcal{O}(n_{occ}^3 n_{vrt}^4)$, respectively. Ideally, linear or better scaling is desired. Local-orbital methods work well for finite, non-conjugated, ground state

systems while scaling near-linearly. These methods will be described in detail in section 2.2 There is currently no sub-linear method for excited states, although efficient algorithms do exist.

### 2.1.2 Equation-of-Motion Coupled Cluster

Originally described in the context of the equivalent time dependent linear-response coupled cluster theory, equation-of-motion coupled cluster (EOM-CC) is best motivated by simultaneously considering two Schrödinger equations: one each for a reference wavefunction and an excited state.

$$H\Psi_0 = E_0\Psi_0 \tag{14}$$

$$H\Psi_k = E_k\Psi_k \tag{15}$$

The excited state is related to the ground state through the action of some operator, $\hat{R}_k$.

$$|\Psi_k\rangle = \hat{R}_k\,|\Psi_0\rangle \tag{16}$$

Left multiplying eq. (14) by $R_k$ and subsequent subtraction from eq. (15) yields an expression for the energy difference:

$$[H, \hat{R}_k]\,|\Psi_0\rangle = \Delta E_k\,|\Psi_0\rangle \tag{17}$$

The simplest way to broach the traditional CC framework is to now chose the reference state to be a CC wavefunction, $|\Psi_{CC}\rangle = e^T\,|\Phi\rangle$, as in **??**, and $\hat{R}$ to be an excitation operator:

$$\begin{aligned}\hat{R}_k &= \underbrace{\sum r_i^a\{a^\dagger i\}}_{R_1} + \underbrace{\sum r_{ij}^{ab}\{a^\dagger i b^\dagger j\}}_{R_2} + \ldots \\ &= R_1 + R_2 + \ldots\end{aligned} \tag{18}$$

Using that the excitation operators $\hat{R}$ and $\hat{T}$ commute, eq. (17) can be re-written in terms of the transformed Hamiltonian and the SCF eigenstate:

$$[H, \hat{R}_k]\,|\Psi\rangle = \Delta E_k\;|\Psi\rangle \tag{19}$$

$$[H, \hat{R}_k]e^T\,|\Phi\rangle = \Delta E_k\,e^T\,|\Phi\rangle$$

$$[e^{-T}He^T,\;\hat{R}_k]\,|\Phi\rangle = \Delta E_k\,|\Phi\rangle$$

$$[\bar{H}, \hat{R}_k]\,|\Phi\rangle = \Delta E\hat{R}_k\,|\Phi\rangle \tag{20}$$

As $\bar{H}$ is non-hermitian, each $R_k$ has a left hand eigenvector, $\hat{L}_k$, which acts as a de-exciation operator. The $\hat{L}_k$'s are given by the left eigenvalue problem and intermediate normalization conditions,

$$\langle\Phi|\,\hat{L}_k\bar{H} = \langle\Phi|\,\hat{L}_k\Delta E \tag{21}$$

$$\langle\Phi|L_kR_l|\Phi\rangle = \delta_{k,l} \tag{22}$$

Unlike $\hat{T}$, the set of $\hat{R}_k$ need not be particle conserving. As a result, different sectors of Fock-space can be made accessible by making different choices for the set of $\hat{R}_k$. Of particular significance are single-electron attachments (EA), ionization potentials (IP), and excitations (EE), which can be selected by choosing $R$ to generate N-1, N+1, and (singly excited) N particle eigenstates:

$$R_{IP} = \sum r_i\{i\} + \sum r^a_{ij}\{a^\dagger ij\} + ... \tag{23}$$

$$R_{EA} = \sum r^a\{a^\dagger\} + \sum r^{ab}_i\{a^\dagger ib^\dagger\} + ... \tag{24}$$

$$R_{EE} = \underbrace{\sum r^a_i\{a\dagger i\}}_{\text{primary space (P)}} + \underbrace{\sum r^{a,b}_{i,j}\{a^\dagger ib^\dagger j\}}_{\text{complementary space(Q)}} + ... \tag{25}$$

These energies lie within the one hole (IP) one particle (EA) and one particle one hole (EE) sectors of fock space. These equations are decoupled, and can therefore be treated separately. In practice, EE amplitudes are based on preceding IP/EA calculations. To obtain accurate energies, contributions from states with one excitation level above the energies of interest must be included (i.e. 1p2h, 2p1h, 2p2h for IP, EA, and EE, respectively). This so-called complementary (Q) space scales as $N^6$ with the basis and is considerably larger than the primary space (P) defined by the first term in $R_k$.

Equations (18) and (23) to (25) depict EOM-CC as a CI procedure using a transformed CC Hamiltonian: $\hat{R}$ is a CI-like construction of the excitation manifold, and the eigenvalues $\Delta E_k$ are exactly the excitation energies of the excited states $\Psi_k$. The issue of size-extensiveness should then be addressed. The root of the size-extensivity problem in truncated CI are the $\Delta EC_n$ terms. In EOM, the analogous $\Delta ER_n$ terms are explicitly set to zero as a result of the constraints set by the CC equations

*I would like to show this algebraicly, need to write it out* (26)

. While not rigorously size-extensive, EOM-CC*will* correctly report extensive properties for states which are within the truncation level of T (e.g. single-electron excitations for EOM-CCSD). The exception is states involving single electron charge-transfer, where an electron on subsystem A is ionized and attached to subsystem B. Such states are described by the 1h+2h1p determinants on A and the 1p+2p1h determinants on B. As a result contributions from overall 3h3p (triply excited) determinants are required. Excited state properties can be defined as energy derivatives and made to scale properly, while ionization potentials and electron attachments are size-intensive, regardless of truncation in $R$ or $T$

*easy to show diagrammatically but is there a good argument for it outside of diagrammatic analsys?* (27)

.

Here will be some EOM-CC results.

### 2.1.3 Partitioned EOM-CCSD

Diagonalization of the transformed Hamiltonian is without doubt the most intensive step. More precisely, diagonalization is dominated by operations involving the $N_{bas}^6$ QQ matrix block (table 1). Typically, direct (Davidson) diagonalization is used, and the slow step involves matrix-vector multiplication of the form $QQ \cdot \hat{v}$. While diagonal QQ elements are too large to be ignored, off-diagonal QQ elements correspond to e.g. simultaneous double excitations in EE-EOM, and are therefore small. Partitioned EOM-CC exploits this by replacing the QQ block with the diagonal terms of the Fock matrix (fig. 2). The complexity reduction due to this approximation is on the order of $\mathcal{O}(Q^3) \rightarrow \mathcal{O}(Q)$.

| Method | Primary space (P) | Secondary space (Q) | Approximate matrix rank |
|---|---|---|---|
| IP-EOM | 1h | 2h1p | $N \cdot N^3$ |
| EA-EOM | 1p | 2p1h | $N \cdot N^3$ |
| EE-EOM | 1p1h | 2p2h | $N^2 \cdot N^4$ |

Table 1: Primary and secondary spaces of partitioned EOM

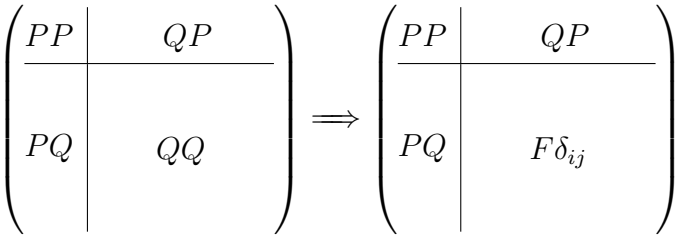



Figure 2: Partitioned EOM-CC reduces computational complexity by replacing the QQ block with the diagonal part of the Fock matrix

Partitioned EOM has proved to give accurate predictions of second-order properties which are dominated by contributions from single excitations. Nooijen et. al.

$$(Chemical Physics Letters 266 (1997) 456 - 464) \tag{28}$$

have reported NMR chemical shifts obtained via partitioned EOM-CC give results comparable to unpartitioned EOM while being much more computationally attractive.

I would like more examples of p-EOM-cc in the final draft. Examples are somewhat elusive. I am not sure if partitioned EOM deserves its own section and I may just append it to the previous.

### 2.1.4 STEOM-CCSD

While EOM-CCSD uses a linear parametrization of the excitation manifold, another approach is to take an exponential parametrization, which results in a doubly-transformed Hamiltonian $G$,

$$G = \hat{\bar{H}} = \{e^{\hat{S}}\}^{-1} \bar{H} \{e^{\hat{S}}\} \tag{29}$$

$$= \{e^{\hat{S}}\}^{-1} e^{-T} H e^{T} \{e^{\hat{S}}\} \tag{30}$$

$$= g_0 + \sum_{p,q} g_q^p \{p^\dagger q\} + \sum_{p,q,r,s} g_{rs}^{pq} \{p^\dagger r q^\dagger s\} + \dots \tag{31}$$

The new operator $\hat{S}$ is defined with respect to a user-chosen active space, and it is required that the STEOM eigenvectors are well represented within this space. Scaling with respect to the active space is modest and as such relatively large active spaces can generally be chosen. In the case that the required active space is prohibitively large, a rotation of the virtuals may

help reduce complexity, as the CC equations are invariant with respect to orbital rotations.

In the following, $i, j, k, l$ and $a, b, c, d$ denote occupied and virtual orbitals, $p, q, r, s$ denote general orbitals, as usual. Explicitly active particle/hole states will be indicated by a cap ($\cap$) superscript, and explicitly inactive orbitals with cup ($\cup$) superscript, e.g. $i^{\cap}$ is active and $a^{\cup}$ inactive. This notation is reasonably accessible if thought of in terms of light switches.
$\hat{S}$ consists of two parts; $\hat{S} = S^+ + S^-$, which are further broken down into single and double quasi-particle parts.

$$\hat{S}^+ = \hat{S}_1^+ + \hat{S}_2^+ = \sum_{a^{\cup}, b^{\cap}} S_{b^{\cap}}^{a^{\cup}} \{a^{\cup\dagger} b^{\cap}\} + \frac{1}{2} \sum_{a,b,c^{\cap},i} S_{i,c^{\cap}}^{a,b} \{a^\dagger i b^\dagger c^{\cap}\} \tag{32}$$

$$\hat{S}^- = \hat{S}_1^- + \hat{S}_2^- = \sum_{i^{\cup}, j^{\cap}} S_{i^{\cup}}^{j^{\cap}\dagger} \{j^{\cap\dagger} i^{\cup}\} + \frac{1}{2} \sum_{i,j,a,k^{\cap}} S_{ij}^{ak^{\cap}} \{a^\dagger i k^{\cap\dagger} j\} \tag{33}$$

The s-amplitudes are built using operators which create electrons in occupied orbitals ($i^\dagger$) and destroy electrons in unoccupied orbitals($a$). Under second quantization, the action of $i^\dagger$ and $a$ destroy the particle-hole quasi-particles produced by excitation operators. As the reference state, by definition, contains no quasi-particles, reference state will always vanish under action of S. Thus non-zero $S$ amplitudes must describe the coupling between excited determinants. This can be readily shown by examining how the construction of $S$ constraints elements of $\hat{G}$.

$$g_i^a = g_i^{ab} = 0 \tag{34}$$

$$S^+ \rightarrow g_{a^{\cap}}^{b^{\cup}} = g_{ic^{\cap}}^{ab} = 0 \tag{35}$$

$$S^- \rightarrow g_{i^{\cup}}^{i^{\cap}} = g_{jk}^{ai^{\cap}} = 0 \tag{36}$$

Equation (34) is equivalent to the CCSD equations, while the constraints set in eq. (35) and eq. (36) define the $S^+$ and $S^-$ operators, respectively. Projecting G against states which would typically survive the action of, for example, $g_{jk}^{ai^{\cap}}$, allows for physical interpretation of these constraints:

$$g_{jk}^{ai^{\cap}} \equiv \langle \Phi_{jk}^a | G | \Phi_{i^{\cap}} \rangle = 0 \tag{37}$$

$g_{jk}^{ai^{\cap}}$ describes coupling between active 1h states and 2h1p states. This interaction is required to vanish, and thus the amplitude equations explicitly decouple excitation blocks from higher

excitations. The $G$ amplitudes are typically obtained by solving the IP/EA-EOMCC equations, followed by a normalization procedure described by an eigenvalue problem in $S$

By definition, $G$ is approximately upper triangular with respect to excitation blocks. Defining the primary space as the reference determinant plus single excitations,

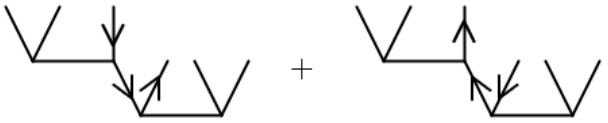
$$G = \begin{array}{c} \\ \langle 0| \\ \langle S| \\ \langle D| \\ \langle T| \end{array} \begin{array}{c} \begin{array}{cccc} |0\rangle & |S\rangle & |D\rangle & |T\rangle \end{array} \\ \begin{pmatrix} E_{CC} & X & X & 0 \\ 0 & X & X & X \\ 0 & \sim & X & X \\ \sim & \sim & \sim & X \end{pmatrix} \end{array} \rightarrow \begin{pmatrix} \mathbf{PP} & \mathbf{PQ} \\ \approx \mathbf{0} & \mathbf{QQ} \end{pmatrix} \tag{38}$$

Where $\sim$ indicates small matrix elements, and X indicates elements of typical magnitude. Solving the $G$ amplitude equations is equivalent to diagonalizing G over the excitation manifold, as is the case in EOM-CC.

The main draw of STEOM is two-fold: because the QP block is transformed to zero, each excitation block can be diagonalized separately. Excitation energies can be extracted directly from $PP$, while traditional EE-EOM requires diagonalization of the QQ space. Furthermore, EE-STEOM-CCSD is comparable to EE-EOM-CCSD(T) in accuracy because expansion of the exponential and projection over the primary space yeilds an implicit tripples correction

$$\{S_2^+ S_2^-\} |\phi_a^i\rangle + S_2 |\Phi_{ij}^{ab}\rangle \tag{39}$$

The resulting terms must be triply excited states, as the operation of $S_2$ increases net excitation by one level. Diagrams of the second term are given below to further show this result:

$S_2 |\Phi_{ij}^{ab}\rangle =$  $+$ 

The connectedness of eq. (39) guarantees size-extensive results for charge-transfer states without explicit treatment of the QQ space.

## 2.2 Efficient Electronic Structure

### 2.2.1 Molecular Symmetry

Electronic structure calculations are performed at a fixed molecular geometry, and thus it is useful to define orbitals that transform in a well defined fashion with the Hamiltonian. Molecular orbitals transform as irreducible representations within the molecule's point group, a closed set of symmetry operations under which the molecule remains unchanged. A visual example, the pi system of benzene, is comprised of four irreducible representations within the $D_{6h}$ point group. $\Gamma_\pi = \Gamma_{A_{2u}} + \Gamma_{E_{1g}} + \Gamma_{E_{2u}} + \Gamma_{B_{2g}} \in D_{6h}$

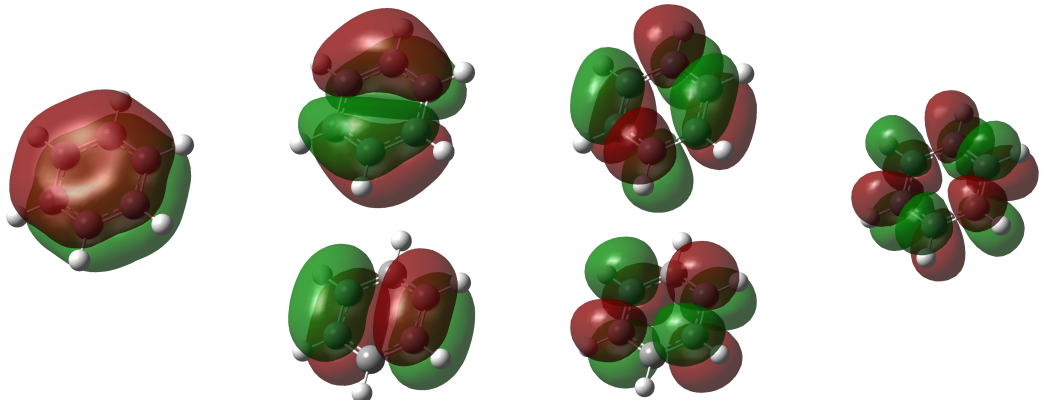



Figure 3: Canonical $\pi$ orbitals of benzene

Rather than describe a specific arrangement of atoms, point groups are an abstraction of the number of ways an object be symmetric, much the same way the number two is an abstraction of any pair of objects. In the context of coupled cluster, integrals and t-amplitudes are heavily constrained by symmetry, e.g. a t-amplitude is zero unless the product of its orbital's irreducible representations is $A_1$.

$$\Gamma_a \otimes \Gamma_i \otimes \cdots = A_1 + \ldots \tag{40}$$

This fact is leveraged by organizing MOs into blocks of symmetry, where only products of orbitals in the same symmetric block are considered, the resulting matrix of t-amplitudes

are is block diagonal:

$$\hat{T}_n = \begin{array}{c|cccc} & A_1 & \Gamma_2 & \cdots & \Gamma_n \\ \hline A_1 & X_{11} & 0 & \cdots & 0 \\ \Gamma_2 & 0 & X_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \Gamma_n & 0 & 0 & \cdots & X_{nn} \end{array}$$

This partitioning reduces the computational cost by a constant factor of $h^2$, where h is the number of irreducible representations. However, because the resulting MO's are de-localized over the the entire molecule, computational effort increases exponentially. Currently, methods involving localized orbitals are of particular interest

## 2.3 Local Correlation Theory

### 2.3.1 Overview

The harsh scaling of CC methods largely stems from the delocalized character of the canonical orbitals [**?**, **?**]. As electrons are added to the system, the number of correlated electron pairs increases quadratically, while the number of strongly-correlated electron pairs grows linearly. The computational complexity is dominated by weakly interacting pairs, whose interaction falls with $R^{-6}$. These interactions cannot be ignored, however calculations involving highly localized orbitals with an approximate treatment of long-range forces have shown to give accurate results while scaling like DFT. Two methods of particular interest are domain-based local pair natural orbital (DLPNO), and cluster-in-molecule (CIM). DLPNO solves the canonical equations under a local approximation, while CIM divides the system into fragments and indirectly assembles the energy from many parallel calculations. Both methods are commonly implemented in conjunction with perturbation theory. Currently, work is ongoing to combine these methods. The correlation energy can be written

$$E^{cor} = \sum_{ijab} \left( 2 \langle ij|ab\rangle - \langle ij|ba\rangle \right) \tau_{ij}^{ab} \tag{41}$$

individual t-amplitudes $\tau_{ij}^{ab}$ can be calculated at either the CC or MP2 level:

$$\tau_{ij}^{ab} = \begin{cases} t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a, & (CCSD) \\ t_{ij}^{ab}, & (MP) \end{cases} \tag{42}$$

eq. (41) is invariant under orbital rotations, and is valid in both the canonical and localized representation. Efficient partitioning of this summation is a convenient basis of discussion for local methods.

The first local correlation method was develoepd by Pulay et. al. in 1983, wherein MPPT calculations were performed using orthogonal localized molecular orbitals (LMO's) and nonorthongonal projected atomic orbitals (PAO) for the occupied and virtual space, respectively [1]. The local schema demonstrated that $> 98\%$ of the correlation energy can be recovered via the use of a unique, local, domain in addition to an approximate treatment of long-range effects [2, 3, 4]. This approach was further refined by Werner and Schütz, who generalized the method for use in coupled-cluster type calculations [5, 6, 7, 8].

### 2.3.2 DLPNO-CC

This method was further refined by constructing approximate natural orbitals specific to each electron pair. CI calculations involving two-electron systems rapidly converge when utilizing NO's ordered by decreasing occupation number [**?**]. LPNO domains grow slowly with respect to basis size, and help prevent the local domains from becoming prohibitively large. [**?**, **?**]. DLPNO carries out eq. (41) over the virtuals $a, b$. The resulting equation expresses the correlation energy as a sum over all electron pairs:

$$E^{cor} = \sum_{ij} E_{ij}^{corr} = \sum_{ij} \sum_{ab} \Big( 2 \langle ij|ab\rangle - \langle ij|ba\rangle \Big) \tau_{ij}^{ab} \tag{43}$$

The number of pairs scales quadratically with the system size, but the contributions from a large percentage of pairs are negligible. Each electron pair is associated with a virtual space of constant size. The pairs are pre-screened, and weakly interacting pairs may be discarded. Surviving amplitudes are coupled and can be determined with CC or MP2, depending on their expected interaction strength. For large systems, this pre-screening reduces the number of t-amplitudes to be calculated by a factor of roughly $10^5$, while retaining $> 99.5\%$ of the correlation energy [**?**]. As a result, LPNO scales linearly for systems of containing less than 80 atoms. For larger systems, the improved domain-based LPNO technique is applicable. At the time of its publication, Neese et. al. showed successful DLPNO treatment for systems typically outside the reach of canonical CC or even LPNO-CC (Figure 4). In their implementation, less than 10% of all electron pairs were given full-accuracy treatment. The

canonical CC equations are modified under DLPNO, and, if not properly implemented, the associated integral transformations can up the majority of the DLPNO runtime. Existing, optimized CC code is incompatible with DLPNO. As a result, new DLPNO implementations require a considerable effort.

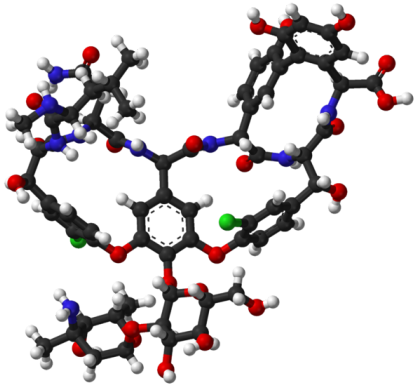


Figure 4: Chemical structure of vancomycin: 176 atoms, 3593 basis functions (def2-TZVP)

### 2.3.3 Cluster in Molecule

In contrast to the 'direct' nature of DLPNO, the 'fragment' based CIM schema divides the system into a number of localized clusters. Each occupied orbital is assigned a subset of orbitals, $P_i$, which, by some measure, interact strongly with the central orbital, $i$. The correlation energy is taken to be a sum over the clusters, and thus eq. (41) is carried out over a single occupied orbital index:

$$E^{cor} = \sum_{i}^{N_{occ}} E_i^{corr} = \sum_{i} \sum_{jab} \Big( 2 \langle ij|ab\rangle - \langle ij|ba\rangle \Big) \tau_{ij}^{ab} \tag{44}$$

There are three important advantages of CIM that should be kept in mind; parallelization, the multi-level extension and compatibility with existing code.

**Parallelization** The CC equations only need to be solved for a series of clusters, instead of for the system as a whole. As these calculations can be carried out in parallel, the overall scaling of CIM depends only on the largest cluster size. In addition, the size of the clusters, is independent on the overall system-size [9]. The memory and disk requirements are also constant

**Multi-level CIM** Multi-level CIM allows different clusters to be treated with varying levels of theory. Chemically important motifs can be treated with a high-accuracy method, while atoms far from a reaction center can be treated with a more efficient low-accuracy method. The mixed-accuracy energy is still the sum over all clusters:

$$E^{corr} = \sum_{i \in P} E_i^{CCSD} + \sum_{j \in Q} E_j^{MP2} \tag{45}$$

Where $P$ and $Q$ are clusters which are to be treated with e.g. CCSD and MP2, respectively. This concept of mixed-accuracy techniques will be later discussed in context of our ground-state method.

**Compatibility** Once the clusters are defined, it is possible to use existing canonical CC code with CIM. This is ease of implementation is the exception rather than the norm. As it stands, a CIM implementation using CCSDTQ code could be developed with a moderate development overhead.

Despite its advantages, CIM is not without faults. The efficiency of the CIM technique is strongly dependant on the size of the clusters, and also the manner in which the clusters are constructed. Canonical CC calculations are carried out on each individual cluster, and if generous thresholds are chosen to ensure accurate results, the resulting clusters may be too large to reasonably treat. Going too far in the other direction, clusters may be too small to yield accurate energies. It is clear the construction of clusters must be done with some level of care. As it is our hope that our work contributes positively to this problem, and as such the major steps of CIM cluster construction are outlined below:

1. Localized occupied orbitals are generated via some procedure. Foster-Boys was initially used, but scales as $\mathcal{O}(N^3)$. Pipek-Mezey or natural localized orbitals have also been implemented [8].

2. Each occupied LMO $i$ is assigned to an atom with the greatest Mulliken atomic charge. This atom is the 'central atom' of the LMO.

3. For a given LMO $i$, LMOs which interact strongly with the central orbital are added to the MO domain. Originally, distance thresholds were used, but methods which do

not utilize real-space cut-offs have also been developed. In such cases quantities such as differential overlap integrals are used.

4. If any domains are completely enveloped by larger domains, they are combined. Thus a single $P_i$ may consist of multiple LMO's and the overall number of clusters is reduced.

5. For each LMO in a given domain, an AO domain is constructed in a similar manner to the MO domains.

6. The occupied LMOs of a cluster $P_i$ are taken to be the LMOs projected onto the final AO domain of the cluster.

7. Virtual LMOs are construed via projection of the atom centred basis functions, $|\chi_k\rangle$, into the virtual space:

$$|\tilde{\chi}_\mu\rangle = \left(1 - \sum_{i=1}^{N_{occ}} |\phi_i\rangle \langle\phi_i|\right) |\chi_\mu\rangle \tag{46}$$

$$= \sum_{k}^{K} |\chi_k\rangle V_{k\mu} \tag{47}$$

For a given AO domain, virtual MOs cam be written as a linear combination of basis functions within the domain:

$$|\phi_a\rangle = \sum \chi_s X_{sa} \tag{48}$$

The coefficients of the transformation matrix $X$ can be used to select virtual MO's which are well represented by the basis functions in a given AO domain. These virtuals are assigned to clusters via this coefficient and subsequently localized. The result is each cluster is given a unique set of virtual, localized MOs, which can be well represented by its AO domain.

Further improvements to this technique involve the use of quasi-canonical MO's, so the MP2 corrections for each cluster can be obtained directly rather than iteratively [10]. A 2016 paper by Li et al. highlights some illustrative uses of CIM. CIM was shown to accurately predict conformational energy differences, reaction barriers, and binding energies of several large systems ($C_{18}H_{38}$, $(H_2O)_{144}$, etc), while scaling linearly [9].

I can also talk about the DLPNO-CIM combination, though that may be enough for its own section. I am going to leave it like this for now.

## 2.4 Ground State Method

The central idea is to perform highly accurate calculations in a truncated Fock space to achieve accuracy comparable to canonical CC while using less resources. Should this scheme show promise, it can be integrated for use in DLPNO/CIM type methods. For a given reference, natural orbital occupation numbers are obtained from the one body density matrix and are used to discard orbitals with low NO occupation. First approximations of the t-amplitudes are calculated using perturbation theory. T-amplitudes completely within the truncated space are later replaced by their CC counterparts. The canonical CC equations are solved using the mixed-accuracy t-amplitudes. Once the these t-amplitudes are obtained, any existing CC code (eg. CCSDTQ) can be used with only a modest development overhead. Below, the math is outlined in detail.

One can obtain low-order approximations to the coupled cluster t-amplitudes from an MP2 correction on a Hartree-Fock reference state. As the potential is a two-body operator, contributions from singly excited states vanish. Similarly, contributions from triple and higher excitations with respect to the reference vanish due to Slater's rules (REF S&O). PT approximations to the t-amplitudes, denoted $t(\mathrm{PT})$, are analytically available following a HF reference. Let $i, j, k, l...$ and $a, b, c, d...$ denote occupied and virtual orbitals, respectively, and let $p, q, r, s...$ be general orbitals, then

$$T_2 : t_{ij}^{ab}(\mathrm{PT}) = \frac{\langle ab|ij\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \tag{49}$$

$$T_1 : t_i^a(\mathrm{PT}) = \frac{R_{ai}}{\epsilon_i - \epsilon_a} \tag{50}$$

Where $R_{ai}$ (not to be confused with the $R_K$ from EOM) is the CC residual:

$$\begin{aligned} R_{ai} = f_{ai} &+ \sum_{c,d,l} t_{il}^{cd}(PT)\Big[2\,\langle al|cd\rangle - \langle al|dc\rangle\Big] \\ &- \sum_{k,l,d} t_{kl}^{ad}(PT)\Big[2\,\langle id|kl\rangle - \langle id|lk\rangle\Big] \end{aligned} \tag{51}$$

Using the low-accuracy t-amplitudes, we generate the occupied-occupied and virtual-virtual blocks of the one-body density matrix.

$$D_{ij} = \delta_{ij} + \sum_{c} t_i^c(\mathrm{PT}) t_j^c(\mathrm{PT}) + \sum_{c,d,l} t_{jl}^{cd}(\mathrm{PT}) \Big( 2\, t_{il}^{cd}(\mathrm{PT}) - t_{il}^{dc}(\mathrm{PT}) \Big) \tag{52}$$

$$D_{ab} = \delta_{ab} - \sum_{k} t_k^a(\mathrm{PT}) t_k^b(\mathrm{PT}) - \sum_{d,k,l} t_{lk}^{bd}(\mathrm{PT}) \Big( 2\, t_{ad}^{cd}(\mathrm{PT}) - t_{lk}^{ad}(\mathrm{PT}) \Big) \tag{53}$$

Eigenvalues of (52) and (53) are natural-orbital occupation numbers, $n_k$ and $n_a$.

To truncate the Fock-space, we choose a numeric threshold $\eta$ and discard natural orbitals with $n_p < \eta$. Solving the CC equations within the truncated NO space yields a set of 'mixed-accuracy' t-amplitudes:

$$t_{pq\ldots} = \begin{cases} t_{pq\ldots}(\mathrm{PT}), & \text{if at least one } \{p, q, \ldots\} \text{ is discarded} \\ t_{pq\ldots}(\mathrm{CC}), & \text{if all } \{p, q, \ldots\} \text{ are retained} \end{cases} \tag{54}$$

Henceforth, all t-amplitudes in this section are taken to be mixed-accuracy as defined by (54). Using the CIM energy expression, the energy can be calculated in the NO basis:

$$E = \sum_{i} \Bigg( \sum_{j,a,b,\in P_i} V_{ij}^{ab} (t_{ij}^{ab} + t_i^a t_j^b + t_j^a t_i^b) \Bigg) + \sum_{i} f_{ia} t_i^a \tag{55}$$

This procedure reduces the number of orbitals involved in the CC equations. By discarding orbitals with low NO occupancy, we hope to maintain canonical CC-like accuracy for sufficiently large $\eta$. The prohibitive scaling of highly-accurate methods means that a modest reduction in orbital space will give a disproportionately large reduction in CPU time, e.g. $0.95^7 \approx 0.69$. Furthermore, the mixed-order t-amplitudes can be used with existing CC code, with reasonable development time.

## 2.5 Extrapolation

As orbitals are discarded based on their NO occupation, different values of $\eta$ may give identical results, making it inconvenient for analysis. Instead, base our analyses on the

weight of orbitals discarded in a calculation, determined by the sum of the discarded orbitals' natural occupation numbers: $\omega$

$$\omega = \sum_{p} n_p \, \forall \, \text{discarded p} \tag{56}$$

The t-amplitudes (54) and energy (55) for a given calculation can be taken as a function of the discarded weights, $E(\omega)$. The point $E(\omega = 0)$ then corresponds to the full CCSD calculation in the given basis. From a set of several truncated calculations we can perform least-squares regression:

$$E(\omega) = \sum_{i} \mathbf{W}_i \omega_i, \quad \omega_i = \begin{pmatrix} 1 \\ \omega_i \\ \omega_i^2 \\ \vdots \end{pmatrix} \tag{57}$$

To provide some cheap lunch.

### 2.6 Excited State

## 3 Results

### 3.1 Ground State

Calculations were performed at the CCSD level in the PVQZ basis using ACESII. Preliminary results show that in practise, the majority of the NO weight is represented by a small number of virtual orbitals relative to the complete basis (figure 5). On average, $\approx 25\%$ of virtuals could be discarded while maintaining a weight $> 0.99$.

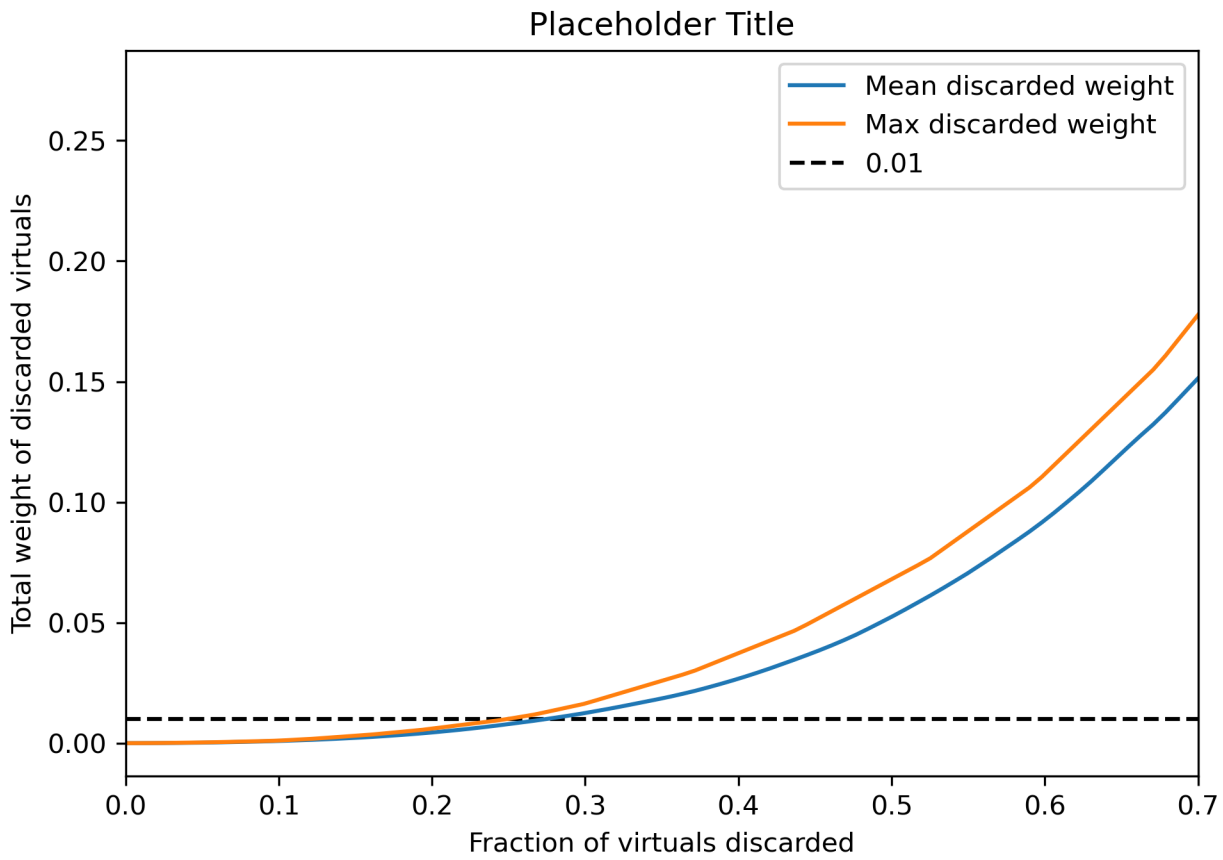



Figure 5: The majority of the NO weight belongs to a small number of virtual orbitals

### 3.2 Excited State

## References

[1] P. Pulay, "Localizability of dynamic electron correlation," *Chemical Physics Letters*, vol. 100, no. 2, pp. 151 – 154, 1983.

[2] S. Saebo and P. Pulay, "The local correlation treatment. ii. implementation and tests," *The Journal of Chemical Physics*, vol. 88, no. 3, pp. 1884–1890, 1988.

[3] S. Sæbø and P. Pulay, "Local configuration interaction: An efficient approach for larger molecules," *Chemical Physics Letters*, vol. 113, no. 1, pp. 13 – 18, 1985.

[4] S. Saebo and P. Pulay, "Local treatment of electron correlation," *Annual Review of Physical Chemistry*, vol. 44, no. 1, pp. 213–236, 1993.

[5] C. Hampel and H. Werner, "Local treatment of electron correlation in coupled cluster theory," *The Journal of Chemical Physics*, vol. 104, no. 16, pp. 6286–6297, 1996.

[6] M. Schütz and H.-J. Werner, "Low-order scaling local electron correlation methods. iv. linear scaling local coupled-cluster (lccsd)," *The Journal of Chemical Physics*, vol. 114, no. 2, pp. 661–681, 2001.

[7] J. Yang, G. K.-L. Chan, F. R. Manby, M. Schütz, and H.-J. Werner, "The orbital-specific-virtual local coupled cluster singles and doubles method," *The Journal of Chemical Physics*, vol. 136, no. 14, p. 144105, 2012.

[8] H.-J. Werner and M. Schütz, "An efficient local coupled cluster method for accurate thermochemistry of large systems," *The Journal of Chemical Physics*, vol. 135, no. 14, p. 144116, 2011.

[9] W. Li, Z. Ni, and S. Li, "Cluster-in-molecule local correlation method for post-hartree–fock calculations of large systems," *Molecular Physics*, vol. 114, no. 9, pp. 1447–1460, 2016.

[10] W. Li, Y. Guo, and S. Li, "A refined cluster-in-molecule local correlation approach for predicting the relative energies of large systems," *Phys. Chem. Chem. Phys.*, vol. 14, pp. 7854–7862, 2012.