

# My PhD thesis

Maximilien Alexandre Ambroise

June 8, 2021

This is a fancy front page: contains supervisors, university name, location, exam date ...

Licensing

(empty)

# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>I</b>	<b>Theory: The Basics</b>	<b>7</b>
<b>2</b>	<b>Ground Work</b>	<b>7</b>
2.1	The Schrödinger Equation . . . . .	7
2.2	Basis Sets . . . . .	7
2.3	Electron Integrals . . . . .	7
<b>3</b>	<b>Hartree Fock</b>	<b>7</b>
<b>4</b>	<b>Post-Hartree Fock Ground State</b>	<b>7</b>
4.1	Configuration Interaction . . . . .	7
4.2	Perturbation Theory . . . . .	7
4.3	Coupled Cluster . . . . .	7
<b>5</b>	<b>Post-Hartree Fock Excited State</b>	<b>7</b>
5.1	Configuration Interaction . . . . .	7
5.2	Coupled Cluster Linear Response . . . . .	7
5.3	Equation-of-Motion Coupled Cluster . . . . .	7
5.4	Algebraic Diagrammatic Construction . . . . .	7
<b>6</b>	<b>Density Fitting</b>	<b>7</b>
<b>II</b>	<b>Theory: Reduced-Cost QC</b>	<b>8</b>
<b>7</b>	<b>Density Fitting</b>	<b>9</b>
<b>8</b>	<b>The ABCs of LMOs and CMOs: Orbital Representations</b>	<b>9</b>
8.1	LMOs by Reducing a Functional . . . . .	9
8.2	Projected Atomic Orbitals . . . . .	10
<b>9</b>	<b>Cholesky Molecular Orbitals</b>	<b>11</b>
<b>10</b>	<b>Local Density Fitting</b>	<b>12</b>
<b>11</b>	<b>Ground State Local</b>	<b>12</b>
11.1	Hartree Fock . . . . .	12
11.2	MP . . . . .	12
11.3	Coupled Cluster . . . . .	12
<b>12</b>	<b>Excited state local</b>	<b>12</b>
12.1	CCLR . . . . .	12
12.2	ADC . . . . .	12



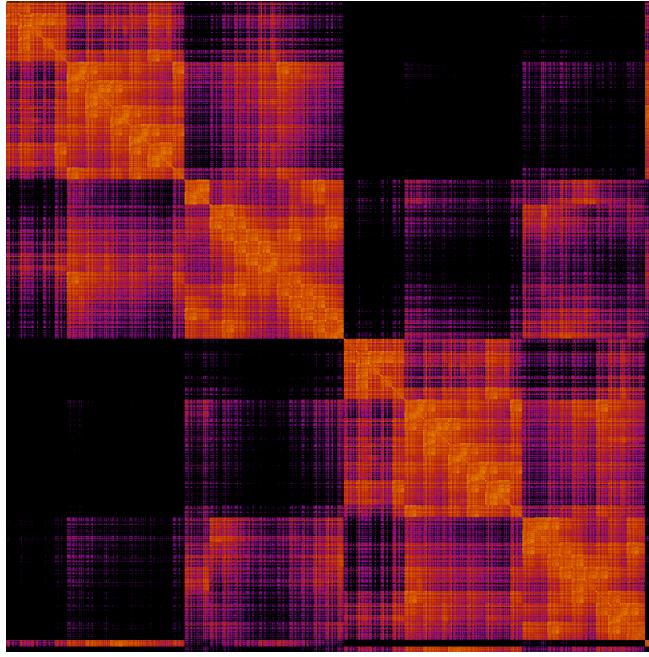


Figure 1: Adenine-guanine fock matrix Hartree FOck cc-pVTZ

## 1 Introduction

This is the introduction. Talk about sparsity. Show sparse matrix. How does sparsity arise, what can we do with it, where else does it emerge? State of arts: adc now, adc in the future

# Part I

## Theory: The Basics

### 2 Ground Work

#### 2.1 The Schrödinger Equation

#### 2.2 Basis Sets

#### 2.3 Electron Integrals

### 3 Hartree Fock

### 4 Post-Hartree Fock Ground State

#### 4.1 Configuration Interaction

#### 4.2 Perturbation Theory

#### 4.3 Coupled Cluster

### 5 Post-Hartree Fock Excited State

#### 5.1 Configuration Interaction

#### 5.2 Coupled Cluster Linear Response

#### 5.3 Equation-of-Motion Coupled Cluster

#### 5.4 Algebraic Diagrammatic Construction

### 6 Density Fitting

## Part II

# Theory: Reduced-Cost QC

While computational chemistry has manifested itself as a popular and widely used tool, its inherently steep scaling limits its applicability to cheaper methods like DFT, onto small or middle sized molecules for post-Hartree Fock methods. Even Hartree Fock with  $O(N^3)$  expensive when comparing to the rest of the world of computer science (examples?). Very early on, with the development of CI methods, (Pulay) effort has been put into reducing the prefactor and computational complexity for QC methods. Example: One of the most time-consuming steps in Post-Hartree Fock is the formation of the molecular orbitals, e.g. formation of the OVOV block of the MO integral tensor as encountered in CC and MP2:

$$(ia \mid jb) = \sum_b^{vir} C_{\sigma b} \sum_a^{vir} C_{\nu a} \sum_j^{occ} C_{\lambda j} \sum_i^{occ} C_{\mu i} (\mu\nu \mid \lambda\sigma) \quad (1)$$

By efficiently refactoring the sums, the MO-AO integral transformation scales as  $O(N^4)$ . There are two reasons for the large cost: first, and quite obviously, a rank-4 tensor scales very fast, hence  $O(N^4)$ , also becomes a bottle neck for tensor contractions as many indices. Secondly, for large basis sets with triple-zeta quality or higher, or basis sets with diffuse functions, the virtual orbital space is very large, and can be multiple times the size of the occupied space. Attempts to reduce scaling can be grouped into two groups: screening-based methods and domain-based methods. Screening-based methods recast existing equations into the AO basis and use the sparsity and fast decay between AOs to establish highly efficient screening algorithms to lower the scaling of integral transformation. Domain-based methods stay in the MO basis, but a localized one, and attempt to assign domains of virtual molecular orbitals to a single LMO or a pair of LMOs, to obtain a more compact representation. Other attempts at mitigating the cost of MO-AO transformation is to exploit the rank sparsity of the AO ERI tensor. Density fitting and Cholesky decompositions can refactor the ERI tensor into a product of two 3-dim tensors. Tensor Hypercontraction goes even further and decomposes into 4 2-dim tensor. Density does not inherently lower scaling of methods, but rather reduces the prefactor associated with integral transformation. In special cases, decomposition techniques allow a refactoring of the working equations into lower scaling. Examples include the coulomb part of the Fock-build ( $O_3$  to  $O_2$ ) and SOS-MP2, SOS-CC2 or SOS-ADC(2) ( $O_5$  to  $O_4$ ). Density fitting and local approximations can be combined, to yield the best of both world in what is known as local density fitting. All of the above methods have their fair share of problems, some more than others. We will first address principles of density fitting, before looking at possible orbital representations, and how they can be used for reduced scaling. Also go to local density fitting, and finally how the methods are implemented for ground state (HF,MP2,CCSD) and excited state computations (CI, CCLR, ADC).

## 7 Density Fitting

About density fitting. Principles

## 8 The ABCs of LMOs and CMOs: Orbital Representations

Occupied and virtual molecular orbitals can generally be represented in two ways: canonical molecular orbitals (CMOs) and local molecular orbitals (LMOs). CMOs are the eigenvectors of the Fock matrix obtained by solving the eigenvalue problem

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (2)$$

where the eigenvalues  $\epsilon$  are known as the molecular orbital energies of the associated CMOs. However, CMOs are not unique in the sense that there are multiple molecular representations possible which yield the same electron density  $\mathbf{P}$ . Observables such as the electron density, or the total energy, are said to be invariant under unitary transformations (Fock V (1930) Z Phys 61:26–148). The CMOs  $\mathbf{C}$  relate to other representations  $\mathbf{L}$  as

$$L_{\mu i} = U_{ii} C_{\mu i} \quad (3)$$

where  $\mathbf{U}$  is a unitary transformation matrix with  $\mathbf{U}^\dagger \mathbf{U} = \mathbb{1}$ . Typically,  $\mathbf{U}$  is chosen to generate a set of molecular orbitals which are localized on as few atoms as possible, hence local molecular orbitals. While CMOs and LMOs agree on observables, they show differences for non-observables, such as molecular orbital energy or orbital shape.

There are several reasons for choosing an LMO representation. First, as mentioned above, LMOs are used in local correlation methods, because CMOs are too delocalized, and electron correlation between LMO centres decay more rapidly. Secondly, they offer a more intuitive picture for chemists and help to interpret chemical phenomena (cite), e.g. involving lone pairs or  $\pi$  bonds. Different representations can be used to interpret different phenomena, e.g. Boys LMOs vs NTOs.

Over the years, a myriad of different schemes has been proposed on how to find appropriate transformation matrices  $\mathbf{U}$ . We will now go over some examples.

### 8.1 LMOs by Reducing a Functional

One of the most popular methods for finding LMOs consists in maximizing a localization function  $\eta(\phi)$  by successive rotation of the orbital space. The most prominent examples are Foster-Boys (FB)(1), Edmiston-Ruedenberg (ER) (2) and Pipek-Mezey (PM) (3). Their functionals can be written as

$$\zeta_{FB}(\chi) = \sum_i \langle \chi_i | \mathbf{r} | \chi i \rangle^2 \quad (4)$$

$$\zeta_{ER}(\chi) = \sum_i (\chi_i \chi_i | \chi_i \chi_i) \quad (5)$$

$$\zeta_{FB}(\chi) = \sum_i \sum_A \langle \chi_i | \mathbf{P}_A | \chi i \rangle^2 \quad (6)$$

The problem is generally solved using an iterative procedure consisting in consecutive pair-wise rotations, known as Jacobi sweeps (ALGO). These sweeps are rotated until convergence is reached, which may be slow. The methods differ within the procedure by how the rotational angle is computed, and scale differently with system size, with  $\mathcal{O}(N^3)$  for FB,  $\mathcal{O}(N^5)$  for ER and  $\mathcal{O}(N^4)$  for PM. A faster alternative to Jacobi sweeps does also exist (4).

Over the years, PM has been the more popular choice of the three: like ER and unlike FB, it conserves  $\sigma$ - $\pi$  separation (0), but it scales more favorably than ER.

Functional localization methods are most often used for rotating occupied MOs. Virtual MOs are often plagued by convergence issues and have a steep computational cost simply due to being much more numerous than occupied MOs (5). It is crucial that molecular localization should not take longer than the methods they are used for, and hence VMOs are often localized using separate methods.

## 8.2 Projected Atomic Orbitals

A set of highly localized molecular orbitals can be obtained by projecting the CMOs onto the atomic orbital basis, known as projected atomic orbitals (PAO) (0). For a set of orthonormal occupied/virtual molecular orbitals  $\{\Psi_i\}$  and  $\{\Psi_a\}$ , the projection operators are defined as (1)

$$\hat{P} = |\Psi_i\rangle \langle \Psi_i| = |\chi_\mu\rangle C_{\mu i} C_{\nu i} \langle \chi_\nu| \quad (7)$$

$$\hat{Q} = |\Psi_a\rangle \langle \Psi_a| = |\chi_\mu\rangle C_{\mu a} C_{\nu a} \langle \chi_\nu| \quad (8)$$

The projection operators are idempotent and mutually orthogonal with  $\hat{P}\hat{Q} = \mathbb{1}$ . Applying the projection operators to the set of AOs

$$\hat{P}|\chi_{\mu'}\rangle = \sum_\mu |\chi_\mu\rangle P_{\mu\nu} S_{\nu\mu'} = L_{\mu I} |\chi_\mu\rangle \quad (9)$$

$$\hat{Q}|\chi_{\mu'}\rangle = \sum_\mu |\chi_\mu\rangle Q_{\mu\nu} S_{\nu\mu'} = L_{\mu A} |\chi_\mu\rangle \quad (10)$$

yields the set of occupied and virtual PAOs  $\{\chi_I\}, \{\chi_A\}$ . Both sets span a space of  $n_{AO}$  functions each, as opposed to  $n_{occ}$  and  $n_{vir}$ . As such, just like the AO

basis, the PAO basis is redundant. CMOs are transformed to PAOs by the relationship

$$|\chi_I\rangle = (\mathbf{SC})_{Ii} |\Psi_i\rangle \quad (11)$$

$$|\chi_A\rangle = (\mathbf{SC})_{Aa} |\Psi_a\rangle \quad (12)$$

PAOs are centred on the atom on which their corresponding AO is localized. However, PAOs can still span multiple atoms. Methods which are entirely formulated in PAOs are rare (1).

PAOs also arise in the context of AO-MP2 or AO-CCSD, when rearranging the canonical equations into an AO basis, as will be shown later.

### 8.3 Cholesky Molecular Orbitals

Sparsity of the atomic density matrix is crucial for achieving low-scaling electronic structure methods. Aquilante et al. proposed (0) to define a set of occupied molecular orbitals by Cholesky decomposition of the density matrix. Analysis of the resulting Cholesky molecular orbitals (CholMOs) showed their localized character inherited from the sparsity of the density matrix.

$$\mathbf{P} = \mathbf{L}\mathbf{L}^T \quad (13)$$

Figure ... shows the sparsity of the occupied density matrix and the occupied Cholesky molecular coefficient matrix of the linear alkane H<sub>322</sub>C<sub>160</sub>. The number of CholMOs is equal to the rank of the density matrix, which is equal to the number of occupied orbitals. The CholMOs are computed by an incomplete Cholesky decomposition with full row and column pivoting (ALGO). The unitary transformation matrix is given by

$$U_{ii} = C_{\mu i} S_{\mu\nu} L_{\nu i} \quad (14)$$

The decomposition algorithm scales with  $\mathcal{O}(N^3)$  but can be made linearly scaling by using sparse matrix algebra. CholMOs have several advantages: the Cholesky decomposition is fast and non-iterative, and an initial guess for molecular orbitals is not needed.

The scheme can be extended to virtual orbitals as well, by CD of the virtual atomic density matrix  $\mathbf{Q}$ . The rank of  $\mathbf{Q}$  is equal to the number of virtual orbitals  $n_{vir}$ , therefore the prefactor of the incomplete CD increases with basis set size. Especially in the presence of diffuse functions, the rank reduction might not offer much of an advantage compared to simpler localization methods such as PAOs.

Moreover, orbitals obtained by CD are less localized than FB or ER LMOs, especially for small molecules. Low scaling is still possible using CholMOs in the context of LMO correlation methods, albeit with a larger prefactor.

CD is also used in the context of AO-MP2 to reduce the prefactor of integral transformation by using the rank sparsity of the pseudo-density matrices, as will be shown further below.

CholMOs can also be used as an initial guess for iterative localization schemes to achieve faster convergence.

#### **8.4 Natural Orbitals?**

#### **8.5 Pair Natural Orbitals**

#### **8.6 OSVs**

#### **8.7 Embedding**

#### **8.8 Natural Transition Orbitals**

Virtual Space: PAOs (Pulay), PNOs (..), OSVs, SPADEs

### **9 Local Density Fitting**

NAFs: <https://aip.scitation.org/doi/full/10.1063/1.4905005>

### **10 Ground State Local**

#### **10.1 Hartree Fock**

#### **10.2 MP**

#### **10.3 Coupled Cluster**

### **11 Excited state local**

#### **11.1 CCLR**

#### **11.2 ADC**

## **Part III**

# **Benchmarking: Timings and**