

# Molecular systems: status and perspectives of ab initio calculations

E. Giner

Laboratoire de Chimie Theorique, CNRS, Paris, France

November 12, 2018, CEA Saclay

# The mathematical problem

- Chemistry  $\Leftrightarrow$  mainly about **electronic structure**
- **Born-Oppenheimer** approximation  
 $\Rightarrow$  focus on  $\Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\})$

# The mathematical problem

- Chemistry  $\Leftrightarrow$  mainly about **electronic structure**
- **Born-Oppenheimer** approximation  
 $\Rightarrow$  focus on  $\Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\})$

$$\hat{H} \Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\}) = E_{elec} \Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\})$$

$$\hat{H} = \sum_{i=1}^{N_e} -\frac{1}{2} \Delta_i + \underbrace{\sum_{J=1}^{N_{nucl}} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|}}_{v_{ne}=\text{electron-nuclei}} + \underbrace{\sum_{i>j} \frac{1}{r_{ij}}}_{W_{ee}=\text{electron-electron}}$$

# The mathematical problem

- Chemistry  $\Leftrightarrow$  mainly about **electronic structure**
- Born-Oppenheimer** approximation  
 $\Rightarrow$  focus on  $\Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\})$

$$\hat{H} \Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\}) = E_{elec} \Psi_{elec}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}, \{\mathbf{R}_J, Z_J\})$$

$$\hat{H} = \sum_{i=1}^{N_e} -\frac{1}{2} \Delta_i + \underbrace{\sum_{J=1}^{N_{nucl}} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|}}_{v_{ne}=\text{electron-nuclei}} + \underbrace{\sum_{i>j} \frac{1}{r_{ij}}}_{W_{ee}=\text{electron-electron}}$$

- Difference between QC calculations  $\approx \{\mathbf{R}_J, Z_J\}$

The very basics of ab initio quantum chemistry  
An heterogeneous problem  
Weak correlation: SR approaches  
Strong correlation effects

# The charm of having nuclei

Globally  $v_{ne}$  is larger  $W_{ee}$

# The charm of having nuclei

Globally  $v_{ne}$  is larger  $W_{ee}$

Electrons are **bound** near the nuclei

- **Large energy splitting** between levels
  - The shell model makes sense  $\Leftrightarrow$  **mean-field approaches**
  - Most of the molecules **exist at HF level**
  - Non mean-field part of  $W_{ee}$  **is small** (0.1 % of the Energy)
- Use hydrogen-like **atom-centered basis** (AOs basis set)
  - In practice use **polynoms**  $\times$  **gaussians**

$$\chi_i^J(\mathbf{r}) = (x - X_J)^{a_x} (y - Y_J)^{a_y} (z - Z_J)^{a_z} e^{-\alpha_i(\mathbf{r} - \mathbf{R}_J)^2}$$

- All Hamiltonian **integrals are analytical** in that basis
- **"Easy" to refine the basis set** (e.g degree of polynoms)

# The charm of having nuclei

Globally  $v_{ne}$  is larger  $W_{ee}$

Electrons are **bound** near the nuclei

- **Large energy splitting** between levels
  - The shell model makes sense  $\Leftrightarrow$  **mean-field approaches**
  - Most of the molecules **exist at HF level**
  - Non mean-field part of  $W_{ee}$  **is small** (0.1 % of the Energy)
- Use hydrogen-like **atom-centered basis** (AOs basis set)
  - In practice use **polynoms**  $\times$  **gaussians**

$$\chi_i^J(\mathbf{r}) = (x - X_J)^{a_x} (y - Y_J)^{a_y} (z - Z_J)^{a_z} e^{-\alpha_i(r - R_J)^2}$$

- All Hamiltonian **integrals are analytical** in that basis
- **"Easy" to refine the basis set** (e.g degree of polynoms)

# The charm of having nuclei

Globally  $v_{ne}$  is larger  $W_{ee}$

Electrons are **bound** near the nuclei

- **Large energy splitting** between levels
  - The shell model makes sense  $\Leftrightarrow$  **mean-field approaches**
  - Most of the molecules **exist at HF level**
  - Non mean-field part of  $W_{ee}$  **is small** (0.1 % of the Energy)
- Use hydrogen-like **atom-centered basis** (AOs basis set)
  - In practice use **polynoms**  $\times$  **gaussians**

$$\chi_i^J(\mathbf{r}) = (x - X_J)^{a_x} (y - Y_J)^{a_y} (z - Z_J)^{a_z} e^{-\alpha_i(\mathbf{r}-\mathbf{R}_J)^2}$$

- All Hamiltonian **integrals are analytical** in that basis
- **"Easy" to refine the basis set** (e.g degree of polynoms)



# The charm of having nuclei

Globally  $v_{ne}$  is larger  $W_{ee}$

Electrons are **bound** near the nuclei

- **Large energy splitting** between levels
  - The shell model makes sense  $\Leftrightarrow$  **mean-field approaches**
  - Most of the molecules **exist at HF level**
  - Non mean-field part of  $W_{ee}$  **is small** (0.1 % of the Energy)
- Use hydrogen-like **atom-centered basis** (AOs basis set)
  - In practice use **polynoms**  $\times$  **gaussians**

$$\chi_i^J(\mathbf{r}) = (x - X_J)^{a_x} (y - Y_J)^{a_y} (z - Z_J)^{a_z} e^{-\alpha_i(\mathbf{r} - \mathbf{R}_J)^2}$$

- All Hamiltonian **integrals are analytical** in that basis
- **"Easy" to refine the basis set** (e.g degree of polynoms)

# The charm of having nuclei

Globally  $v_{ne}$  is larger  $W_{ee}$

Electrons are **bound** near the nuclei

- **Large energy splitting** between levels
  - The shell model makes sense  $\Leftrightarrow$  **mean-field approaches**
  - Most of the molecules **exist at HF level**
  - Non mean-field part of  $W_{ee}$  **is small** (0.1 % of the Energy)
- Use hydrogen-like **atom-centered basis** (AOs basis set)
  - In practice use **polynoms**  $\times$  **gaussians**

$$\chi_i^J(\mathbf{r}) = (x - X_J)^{a_x} (y - Y_J)^{a_y} (z - Z_J)^{a_z} e^{-\alpha_i(\mathbf{r} - \mathbf{R}_J)^2}$$

- All Hamiltonian **integrals are analytical** in that basis
- **"Easy" to refine the basis set** (e.g degree of polynoms)

The very basics of ab initio quantum chemistry  
An heterogeneous problem  
Weak correlation: SR approaches  
Strong correlation effects

## Then why QC is tedious ? the accuracy !

Ability of **unraveling chemistry**  $\Leftrightarrow$  **Accuracy of  $\Delta E$**

## Then why QC is tedious ? the accuracy !

Ability of **unraveling chemistry**  $\Leftrightarrow$  **Accuracy of  $\Delta E$**

### The famous **chemical accuracy** in quantum chemistry

- "**Chemical accuracy**"  $\approx 1$  kcal/mol  $\approx 1.6 \cdot 10^{-3}$  a.u.  $\approx 0.04$  eV
  - $\approx$  **accuracy of thermochemistry experiments**
  - $\approx 0.2$  of Boltzman probability ratio at 298 K
- Typical quantity of interest: **atomization energy (AE)**

$$AE(AB) = E(AB) - (E(A) + E(B))$$

- $10^0 < AE < 10^2$  kcal/mol
- Typical **error at HF level**:  $\approx 10^1$  or even  $\approx 10^2$  kcal/mol

## Then why QC is tedious ? the accuracy !

Ability of **unraveling chemistry**  $\Leftrightarrow$  **Accuracy of  $\Delta E$**

### The famous **chemical accuracy** in quantum chemistry

- "**Chemical accuracy**"  $\approx 1$  kcal/mol  $\approx 1.6 \cdot 10^{-3}$  a.u.  $\approx 0.04$  eV
  - $\approx$  **accuracy of thermochemistry experiments**
  - $\approx 0.2$  of Boltzman probability ratio at 298 K
- Typical quantity of interest: **atomization energy (AE)**

$$AE(AB) = E(AB) - (E(A) + E(B))$$

- $10^0 < \mathbf{AE} < 10^2$  kcal/mol
- Typical **error at HF level**:  $\approx 10^1$  or even  $\approx 10^2$  kcal/mol

## Then why QC is tedious ? the accuracy !

Ability of **unraveling chemistry**  $\Leftrightarrow$  **Accuracy of  $\Delta E$**

### The famous **chemical accuracy** in quantum chemistry

- "**Chemical accuracy**"  $\approx 1$  kcal/mol  $\approx 1.6 \cdot 10^{-3}$  a.u.  $\approx 0.04$  eV
  - $\approx$  **accuracy of thermochemistry experiments**
  - $\approx 0.2$  of Boltzman probability ratio at 298 K
- Typical quantity of interest: **atomization energy (AE)**

$$AE(AB) = E(AB) - (E(A) + E(B))$$

- $10^0 < \mathbf{AE} < 10^2$  kcal/mol
- Typical **error at HF level**:  $\approx 10^1$  or even  $\approx 10^2$  kcal/mol

What do we miss ? **The correlation effects !**

# What tools for correlation ?

## Two types of approaches

- **Density Functiona Theory (DFT)**

- Uses the 1-body density and its derivatives: **very cheap !**
- Until the last years, clearly the most used tool
- Initially: trying to know  $n^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$  with  $n^{(1)}(\mathbf{r})$
- Often fitted with experiments
- BUT: Problems with non-local correlation effects ...
- BUT: Hard to **systematically improve the quality**
- Alternative: mixing WFT and DFT ?

- **Wave Function Theory (WFT)**

- Uses the  $N$ -body wave function: **not cheap ...**
- Until recently, much less used than DFT ...
- **Systematically improvable**
- **CCSD(T) very reliable for most of chemical situation**
- BUT: CPU time
- BUT: WFT is smoothly taking over DFT ...

# Quantum chemistry from outside: why so many **acronyms**

(Some) **Acronyms for Wave Function Theory ...**

- HF, MP2, CEPA- $n$ , CISD(SC)<sup>2</sup>, CCSD(T), BCCD(T), EOM-CCSD(T), PNO-CCSD(T), DLPNO-CCSD(T), ...
- CASCI, CASSCF, MCSCF, MRMP2, XMCQDPT, CASPT2, MS-CASPT2, NEVPT2, SC-NEVPT2, PC-NEVPT2, QD-NEVPT2, JMMRPT2, ...
- CIPSI, HBCI, MPS, DMRG, FCIQMC, *i*FCIQMC, ACI, SORCI, DDCI, FOBOCI, ...
- SS-MRCC, SU-MRCC, VU-MRCC, JM-MRCC, Mk-MRCCSDT, *ic*-MRCC, ...
- F12-MP2, F12-CCSD(T), F12-NEVPT2, F12-DLPNO-NEVPT2, RS-DFT, ...



## An heterogeneous problem

Weak correlation: SR approaches

Strong correlation effects

# acronyms

## Density Functional Theory acronyms ...



# Molecular simulations: why so many acronyms ?

Mainly two answers ...

- 1 Quantum chemists have a **bigger ego issue**
  - The researcher ego is quite uniformly distributed in science
  - Why bigger ego than mathematicians ?
- 2 The theoretical chemistry problem is very **heterogeneous**

# Molecular simulations: why so many acronyms ?

Mainly two answers ...

- ① Quantum chemists have a **bigger ego issue**
  - The researcher ego is quite uniformly distributed in science
  - Why bigger ego than mathematicians ?
- ② The theoretical chemistry problem is very **heterogeneous**

# Molecular simulations: why so many acronyms ?

## Mainly two answers ...

- ① Quantum chemists have a **bigger ego issue**
  - The researcher ego is quite uniformly distributed in science
  - Why bigger ego than mathematicians ?
- ② The theoretical chemistry problem is very **heterogeneous**

## The main objectives of molecular simulation

- Predict and/or interpret molecular experiments
- Basically infinite possibilities at human scale !

**Molecular simulation is as diverse as chemistry can be !**

# An overview of the heterogeneity of theoretical chemistry

## Two main variables to define a chemistry experiment

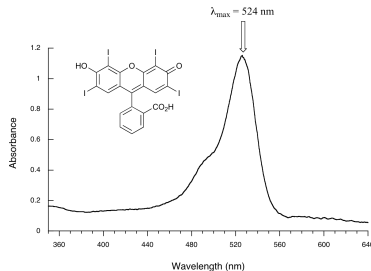
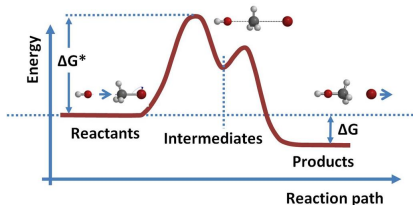
- Types of **molecular properties**  
⇒ Different objects to compute
- Types of **molecular systems**  
⇒ Different size of systems  
⇒ Different level of  $e$ - $e$  correlation

**Different implications for the theoretician**

# A few examples of molecular properties

## Different chemical problematics

- **Formation of molecules**
  - Gd state energy
  - Energy derivatives with **R**
- **UV/visible spectroscopy**
  - Excited states
  - Oscillation strength
- **Magnetic spectroscopy**
  - Open shell systems
  - Energy derivatives with **B**
- **Any combination ...**



# A few examples of molecular systems

## Different types of systems

- **Size of the system**

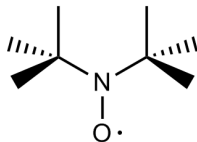
- from  $10^0$  to  $10^2$  electrons

- **Elements in the systems**

- *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
- *d* or *f* atoms (Fe, Dy, ...):  
"Hard" atoms

- **Electronic structure**

- Are there unpaired electrons ?
- Is HF a good representation ?



## A few examples of molecular systems

### Different types of systems

- **Size of the system**

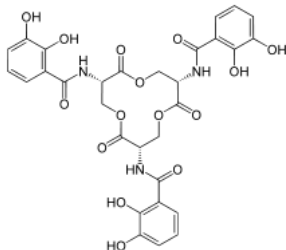
- from  $10^0$  to  $10^2$  electrons

- **Elements in the systems**

- *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
- *d* or *f* atoms (Fe, Dy, ...):  
"Hard" atoms

- **Electronic structure**

- Are there unpaired electrons ?
- Is HF a good representation ?





# A few examples of molecular systems

## Different types of systems

- **Size of the system**

- from  $10^0$  to  $10^2$  electrons

- **Elements in the systems**

- *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
- *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms

- **Electronic structure**

- Are there unpaired electrons ?
- Is HF a good representation ?

# A few examples of molecular systems

## Different types of systems

- **Size of the system**
  - from  $10^0$  to  $10^2$  electrons
- **Elements in the systems**
  - *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
  - *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms
- **Electronic structure**
  - Are there unpaired electrons ?
  - Is HF a good representation ?

H																	He	
Li	Be											B	C	N	O	F	Ne	
Na	Mg											Al	Si	P	S	Cl	Ar	
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr	
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe	
Cs	Ba			Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra			Rf	Db	Sg	Bh	Hs	Mt	Os	111	112	Uut	Uuq	Uup	Uuh	Uus	Uuo

\* Lanthanide Series

\* Actinide Series

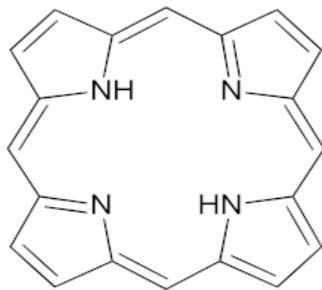
La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr

Alkali metals	Alkaline earth metals	Lanthanide	Actinide	Transition metals	Poor metals	Metalloids	Other Nonmetals	Halogens	Noble Gases
---------------	-----------------------	------------	----------	-------------------	-------------	------------	-----------------	----------	-------------

## A few examples of molecular systems

### Different types of systems

- **Size of the system**
  - from  $10^0$  to  $10^2$  electrons
- **Elements in the systems**
  - *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
  - *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms
- **Electronic structure**
  - Are there unpaired electrons ?
  - Is HF a good representation ?



# A few examples of molecular systems

## Different types of systems

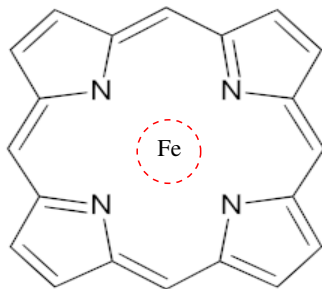
- **Size of the system**
  - from  $10^0$  to  $10^2$  electrons
- **Elements in the systems**
  - *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
  - *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms
- **Electronic structure**
  - Are there unpaired electrons ?
  - Is HF a good representation ?

1	2																	18
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18			
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	
109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	

## A few examples of molecular systems

### Different types of systems

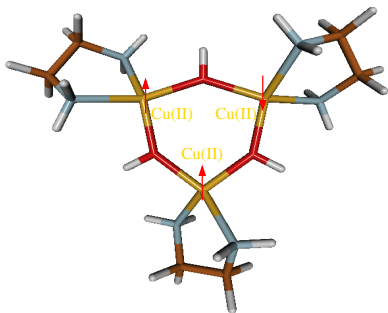
- **Size of the system**
  - from  $10^0$  to  $10^2$  electrons
- **Elements in the systems**
  - *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
  - *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms
- **Electronic structure**
  - Are there unpaired electrons ?
  - Is HF a good representation ?



## A few examples of molecular systems

### Different types of systems

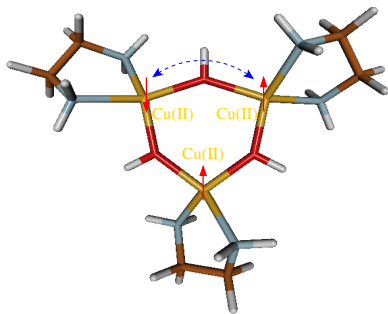
- **Size of the system**
  - from  $10^0$  to  $10^2$  electrons
- **Elements in the systems**
  - *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
  - *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms
- **Electronic structure**
  - Are there unpaired electrons ?
  - Is HF a good representation ?



## A few examples of molecular systems

### Different types of systems

- **Size of the system**
  - from  $10^0$  to  $10^2$  electrons
- **Elements in the systems**
  - *s* or *p* atoms (H, C, N,...):  
"Easy" atoms
  - *d* or *f* atoms (Fe, Dy, ..):  
"Hard" atoms
- **Electronic structure**
  - Are there unpaired electrons ?
  - Is HF a good representation ?



## Weak correlation effects: the basics

### Predominance of the HF Slater determinant

$$|\Psi^{\text{FCI}}\rangle = |\text{HF}\rangle + \underbrace{\sum_k c_k |k\rangle}_{c_k \ll 1}$$

- Correlation **does not change de HF picture**

$$E^{\text{corr}} = \sum_k c_k \underbrace{\langle \text{HF} | H | k \rangle}_{\text{small quantities}}$$

- Total  $E^{\text{corr}}$  as a **sum of many small contributions**

### Typical systems

- Closed-shell molecules near **their equilibrium geometries**
- Open-shell systems in a **high-spin state**



# Perturbative analysis of weak correlation effects

$$E^{\text{corr}} = \sum_{iajb} c_{ij}^{ab} V_{ij}^{ab} = \sum_{iajb} e_{ij}^{ab}$$

$$c_{ij}^{ab} \approx \frac{V_{ij}^{ab}}{\epsilon_i - \epsilon_a + \epsilon_j - \epsilon_b}$$

$$V_{ij}^{ab} = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_i(\mathbf{r}_1) \phi_a(\mathbf{r}_1) \frac{1}{r_{12}} \phi_j(\mathbf{r}_2) \phi_b(\mathbf{r}_2)$$

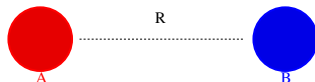
**Weak correlation effects**  $\Leftrightarrow c_{ij}^{ab} \ll 1$ :

- $V_{ij}^{ab}$  is small
- and/or  $\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$  is large

Two main regimes: **long-range** and **short-range** correlation

## Weak long-range correlation effects: $V_{ij}^{ab}$ is small

$$V_{ij}^{ab} = \int d\mathbf{r}_1 d\mathbf{r}_2 \underbrace{\phi_i(\mathbf{r}_1)\phi_a(\mathbf{r}_1)}_A \frac{1}{r_{12}} \underbrace{\phi_j(\mathbf{r}_2)\phi_b(\mathbf{r}_2)}_B$$

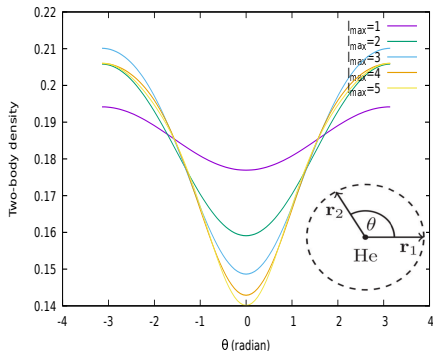


- $r_{12} \propto R \Leftrightarrow V_{ij}^{ab} \propto \frac{1}{R^3}$
- $E^{\text{MP2}} \propto -\left(V_{ij}^{ab}\right)^2 \propto -\frac{1}{R^6} \Leftrightarrow$  **attractive dispersion forces**
- Scales as the **pairs of atoms** and depends on  $\{\mathbf{R}_J\}$ 
  - Hydrogen bonds (DNA)
  - Molecular-surface interaction (solid state)

## WFT in a finite basis set and Kato's cusp condition

### Problems of the basis-set

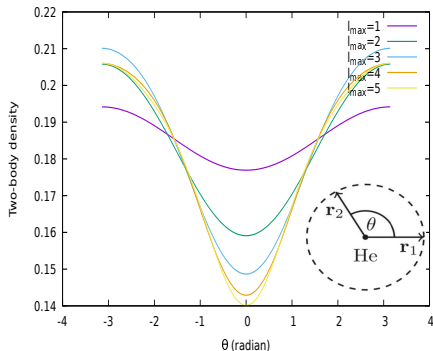
- No cusp in a **finite basis-set**
- Wrong near  $r_{12} = 0$



## WFT in a finite basis set and Kato's cusp condition

### Problems of the basis-set

- No cusp in a **finite basis-set**
- Wrong near  $r_{12} = 0$



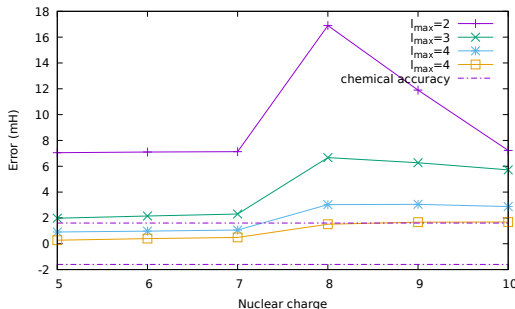
Affects local properties of  $\Psi$   
Is it really a problem for  $\Delta E$ ?

# Illustration of the impact of short-range correlation effects

## FCI Atomic ionization potentials

$$IP(A) = E(A) - E(A^+)$$

- Only **short-range** effects
- Slow convergence**  
 $\propto \frac{1}{(l_{max})^3}$
- Require **large basis-set**
- Impact **energy differences**
- Favours the **less correlated**



Impacts also molecular properties

# Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor** (F12) to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems  
DLPNO-F12-MP2, DLPNO-F12-CCSD(T)

- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...

## Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor (F12)** to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems  
DLPNO-F12-MP2, DLPNO-F12-CCSD(T)

- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...

## Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor (F12)** to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems  
DLPNO-F12-MP2, DLPNO-F12-CCSD(T)

- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...



## Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor** (F12) to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems  
DLPNO-F12-MP2, DLPNO-F12-CCSD(T)
- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...

## Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor** (F12) to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems

DLPNO-F12-MP2, DLPNO-F12-CCSD(T)

- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...

## Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor** (F12) to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems  
DLPNO-F12-MP2, DLPNO-F12-CCSD(T)

- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...

## Challenges in WFT for weak correlation effects ?

Nowadays challenge : push CCSD(T) to the limit !

- **Treat bigger systems**

- The correlation effects are essentially **local**
- Use **localized molecular orbitals** to treat dispersion forces  
Pair Natural Orbitals (PNO) and Local Domains (DL-PNO)

- **Smaller basis-set error**

- The wave function should satisfy Kato's **cusp condition**
- Use **explicit correlation factor** (F12) to reduce basis-set error

$$\Psi_{f12}(\mathbf{r}_1, \mathbf{r}_2) = e^{-\gamma r_{12}} \Psi(\mathbf{r}_1, \mathbf{r}_2)$$

F12-MP2, F12-CCSD, F12-CCSD(T)

- **Combine these two approaches** to reach big systems  
DLPNO-F12-MP2, DLPNO-F12-CCSD(T)

- **Compute Energy derivatives**

Geometry optimization, molecular properties etc ...

## Chemical examples of strong correlation

### Low-spin open-shell

- **Covalent bond** breakings:



- ***d* and *f*** elements:

- Single center: e.g. Iron-porphyrin
- Multi center: e.g. Tris-OH

- **On-site repulsion *U***:

- Tends to break pairs
- Tends to localize electrons
- Short-range  $\leftrightarrow$  long-range

- Correlation effects **in *d* elements**:

- A lot of *e* on a site
- Very polarizable !

# Chemical examples of strong correlation

## Low-spin open-shell

- **Covalent bond** breakings:



- ***d* and *f* elements:**

- Single center: e.g. Iron-porphyrin
- Multi center: e.g. Tris-OH

- **On-site repulsion *U*:**

- Tends to break pairs
- Tends to localize electrons
- Short-range ↔ long-range

- Correlation effects **in *d* elements:**

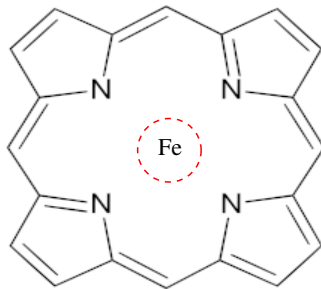
- A lot of *e* on a site
- Very polarizable !

The periodic table shows elements color-coded by groups: Alkali metals (red), Alkaline earth metals (orange), Lanthanoids (light blue), Actinoids (dark blue), Transition metals (green), Poor metals (yellow), Metalloids (light green), Nonmetals (dark green), Halogens (purple), and Noble gases (pink). The *d* block (transition metals) and *f* block (lanthanoids and actinoids) are highlighted with red dashed boxes, indicating the focus of the lecture on strong correlation effects in these elements.

## Chemical examples of strong correlation

### Low-spin open-shell

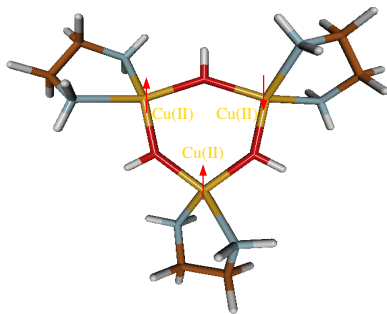
- **Covalent bond** breakings:  
 $AB \rightarrow \dot{A} \dots \dot{B}$
- ***d* and *f*** elements:
  - Single center: e.g. Iron-porphyrin
  - Multi center: e.g. Tris-OH
- **On-site repulsion *U***:
  - Tends to break pairs
  - Tends to localize electrons
  - Short-range  $\leftrightarrow$  long-range
- Correlation effects **in *d* elements**:
  - A lot of *e* on a site
  - Very polarizable !



## Chemical examples of strong correlation

### Low-spin open-shell

- **Covalent bond** breakings:  
 $AB \rightarrow \dot{A} \dots \dot{B}$
- ***d* and *f*** elements:
  - Single center: e.g. Iron-porphyrin
  - Multi center: e.g. Tris-OH
- **On-site repulsion *U***:
  - Tends to break pairs
  - Tends to localize electrons
  - Short-range  $\leftrightarrow$  long-range
- Correlation effects **in *d* elements**:
  - A lot of *e* on a site
  - Very polarizable !

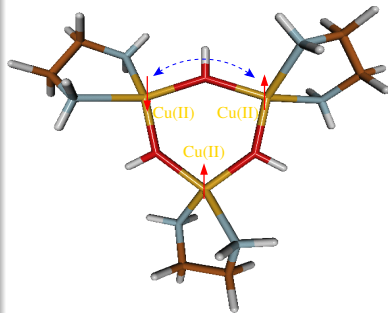




## Chemical examples of strong correlation

### Low-spin open-shell

- **Covalent bond** breakings:  
 $AB \rightarrow \dot{A} \dots \dot{B}$
- ***d* and *f*** elements:
  - Single center: e.g. Iron-porphyrin
  - Multi center: e.g. Tris-OH
- **On-site repulsion *U***:
  - Tends to break pairs
  - Tends to localize electrons
  - Short-range  $\leftrightarrow$  long-range
- Correlation effects **in *d* elements**:
  - A lot of *e* on a site
  - Very polarizable !



# Chemical examples of strong correlation

## Low-spin open-shell

- **Covalent bond** breakings:



- ***d* and *f*** elements:

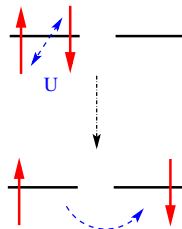
- Single center: e.g. Iron-porphyrin
- Multi center: e.g. Tris-OH

- **On-site repulsion *U***:

- Tends to break pairs
- Tends to localize electrons
- Short-range  $\leftrightarrow$  long-range

- Correlation effects **in *d* elements**:

- A lot of *e* on a site
- Very polarizable !



# Chemical examples of strong correlation

## Low-spin open-shell

- **Covalent bond** breakings:



- ***d* and *f*** elements:

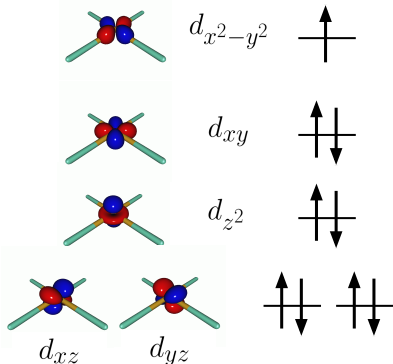
- Single center: e.g. Iron-porphyrin
- Multi center: e.g. Tris-OH

- **On-site repulsion *U***:

- Tends to break pairs
- Tends to localize electrons
- Short-range  $\leftrightarrow$  long-range

- Correlation effects **in *d* elements**:

- A lot of *e* on a site
- Very polarizable !



# Chemical examples of strong correlation

## Low-spin open-shell

- **Covalent bond** breakings:



- ***d* and *f*** elements:

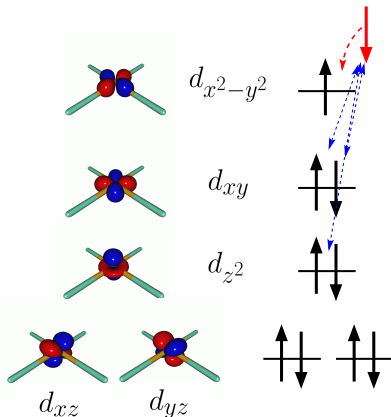
- Single center: e.g. Iron-porphyrin
- Multi center: e.g. Tris-OH

- **On-site repulsion *U***:

- Tends to break pairs
- Tends to localize electrons
- Short-range  $\leftrightarrow$  long-range

- Correlation effects **in *d* elements**:

- A lot of *e* on a site
- Very polarizable !



## Qualitative description of strongly correlated systems

- **Unpaired electrons and low-spin**
  - Covalent bond breakings:  $AB \rightarrow \dot{A} \cdots \dot{B}$
  - Magnetic systems:
- **Rapidly large expansion for  $|\Psi^{(0)}\rangle$  !**

$$|\Psi^{(0)}\rangle = \sum_{I=1}^{10^3-10^6} c_I |I\rangle$$

- **The ratios  $\frac{c_I}{c_J}$  drive most of the physical properties**
- **Between the  $|I\rangle$  and  $|J\rangle$** 
  - Large interactions
  - Energetic degeneracies
  - $\frac{\langle J|H|I\rangle}{\Delta E_{IJ}} \gg 1$
- **Non perturbative**

## Quantitative description: the physics beyond $|\Psi^{(0)}\rangle$

$$|\Psi\rangle = |\Psi^{(0)}\rangle + \sum_i c_i |\phi_i\rangle$$

- **In general**  $|c_i| \ll 1 \Leftrightarrow$  **Perturbative**
- **Standard weak correlation** (cusp, dispersion forces)
  - Weak differential correlation effects
- **Differential correlation effects**
  - The  $|I\rangle$  are different
  - Correlation effects depend on  $|I\rangle$
- **Change  $|\Psi^{(0)}\rangle$** 
  - Affects the  $\langle J|H|I\rangle$  and  $\Delta E_{IJ}$
  - Renormalization of  $H$
- **Size consistency**
  - Able to break bonds
  - Correct scaling of the energy with  $N$

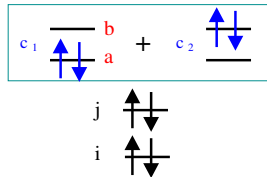
# The Complete Active Space based approach

## CAS-CI( $n, m$ ) approach

- All determinants within  $n$   $e$  and  $m$  orbitals
- Variational energy  $\Leftrightarrow$  no divergences

$$|\Psi^{(0)}\rangle = \sum_I c_I |I\rangle$$

$$E_{\Psi^{(0)}} = \min_{\{c_I\}} \langle \Psi^{(0)} | H | \Psi^{(0)} \rangle$$



## CAS-SCF approach

- Add the orbital optimization

$$|\Psi^{(0)}\rangle = e^{\hat{\kappa}} \sum_I c_I |I\rangle$$

$$E_{\Psi^{(0)}} = \min_{\{c_I, \hat{\kappa}\}} \langle \Psi^{(0)} | H | \Psi^{(0)} \rangle$$

# The Complete Active Space based approach

## CAS-CI( $n, m$ ) approach

- All determinants within  $n$   $e$  and  $m$  orbitals
- Variational energy  $\Leftrightarrow$  no divergences

$$|\Psi^{(0)}\rangle = \sum_I c_I |I\rangle$$

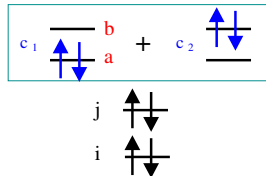
$$E_{\Psi^{(0)}} = \min_{\{c_I\}} \langle \Psi^{(0)} | H | \Psi^{(0)} \rangle$$

## CAS-SCF approach

- Add the orbital optimization

$$|\Psi^{(0)}\rangle = e^{\hat{\kappa}} \sum_I c_I |I\rangle$$

$$E_{\Psi^{(0)}} = \min_{\{c_I, \hat{\kappa}\}} \langle \Psi^{(0)} | H | \Psi^{(0)} \rangle$$





## Selected CI: Breaking the exponential wall

### Some problems of the CAS

- Exponential scaling : max is CAS(20,20) (Heroic calculation)
- Cannot take into account weak correlation effects ...

## Selected CI: Breaking the exponential wall

### Some problems of the CAS

- Exponential scaling : max is CAS(20,20) (Heroic calculation)
- Cannot take into account weak correlation effects ...

### A CAS-CI space is full of dead-wood

- **Expensive** step in CAS-CI: the **variational step**
- Treating **all**  $|I\rangle$  **at same footing** is **useless**
- Only a few percent are really needed in a variational way
- Many of them can be **treated in perturbation**

## Selected CI: Breaking the exponential wall

### Some problems of the CAS

- Exponential scaling : max is CAS(20,20) (Heroic calculation)
- Cannot take into account weak correlation effects ...

### A CAS-CI space is full of dead-wood

- **Expensive** step in CAS-CI: the **variational step**
- Treating **all**  $|I\rangle$  **at same footing** is **useless**
- Only a few percent are really needed in a variational way
- Many of them can be **treated in perturbation**

### One key question

- How to **know a priori** which are the **important ones** ?

## The CIPSI algorithm: selection based on perturbation

- 1 Given a zeroth-order wave function  $|\Psi^{(0)}\rangle$  and  $E_{\Psi^{(0)}}$
- 2 Estimate the coefficient of  $|\mu\rangle$  not in  $|\Psi^{(0)}\rangle$ :

$$c_{\mu}^{(1)} = \frac{\langle \mu | H | \Psi^{(0)} \rangle}{E_{\Psi^{(0)}} - \langle \mu | H | \mu \rangle}$$

- 3 On the fly, estimate the  $E^{(2)}$

$$E^{(2)} = \sum_{\mu} \frac{(\langle \mu | H | \Psi^{(0)} \rangle)^2}{E_{\Psi^{(0)}} - \langle \mu | H | \mu \rangle}; E_{\text{CIPSI}} = E_{\Psi^{(0)}} + E^{(2)}$$

- 4 Select the most important ones and add them to  $\Psi^{(0)}$
- 5 Check the convergence or GOTO (1)

## A brief review

### Revival of an old idea (Malrieu, 1973)

- I rediscovered it in 2011 (ironically in Malrieu's lab)
- Converge extremely fast: CAS(40,80), CAS(20,100) !
- Many times rediscovered (from 2011 at least 4 versions)
- Examples: HCl, MC-Cl, A-Cl, AdCl, ...
- We did not renamed it :)

### Alternatives

- DMRG: use matrix product states
- FCI-QMC : use a stochastic sampling of the CI equations

## Trial of conclusion for WFT

The weak correlation problems: WFT begins to compete DFT!

- Mathematically solved : CCSD(T) !!
- The main issue : CPU time
  - System size  $\Leftrightarrow$  Use local orbitals (PNO)
  - Basis set  $\Leftrightarrow$  Use explicit correlation (F12)

## Trial of conclusion for WFT

The weak correlation problems: WFT begins to compete DFT!

- Mathematically solved : CCSD(T) !!
- The main issue : CPU time
  - System size  $\Leftrightarrow$  Use local orbitals (PNO)
  - Basis set  $\Leftrightarrow$  Use explicit correlation (F12)

The strong correlation problems: a lot of work ...

- Selected CI (SCI) have pushed away the limits of CAS-CI
- Not the end of the story: the exponential wall is just farther ...
- Alternatives: MRCC, MRPT based on large SCI

## Trial of conclusion for WFT

The weak correlation problems: WFT begins to compete DFT!

- Mathematically solved : CCSD(T) !!
- The main issue : CPU time
  - System size  $\leftrightarrow$  Use local orbitals (PNO)
  - Basis set  $\leftrightarrow$  Use explicit correlation (F12)

The strong correlation problems: a lot of work ...

- Selected CI (SCI) have pushed away the limits of CAS-CI
- Not the end of the story: the exponential wall is just farther ...
- Alternatives: MRCC, MRPT based on large SCI

**The complexity of QC relies in the diversity of chemistry !**