

## Quantum chemistry and wavefunction based methods for electron correlation

Joachim Sauer Institut für Chemie, Humboldt-Universität



HUMBOLDT-UNIVERSITÄT ZU BERLIN

The Royal Swedish Academy of Sciences has awarded **The 1998 Nobel Prize in Chemistry** in the area of quantum chemistry to

Walter Kohn, University of California at Santa Barbara, USA and John A. Pople, Northwestern Univ., Evanston, Ill., USA (British citizen).

#### **Citation:**

"to Walter Kohn for his development of the **density-functional theory** and to John Pople for his development of **computational methods in quantum chemistry**."







### **Surface Reaction**



### **DFT problems with barriers**



### Zeolite catalysts: active sites (transition metal ions, protons) in a "surface-only" silica matrix



### Zeolite catalysis: Methanol-to-hydrocarbons



### Methylation of ethene, propene, trans-2-butene in H-MFI



\*S. Svelle, P.A. Ronning, S. Kolboe, *J. Catal.* **2004**, *224*, 115. S. Svelle, P.O. Ronning, U. Olsbye, S. Kolboe, *J. Catal.* **2005**, *234*, 385.

### Methylation of alkenes in H-MFI - pbc DFT



Svelle, Tuma, Rozanska, Kerber, Sauer, JACS 131 (2009) 816

### Pragmatic solution: DFT + Dispersion (DFT+D)

Many predecessors with DFT, HF+Disp, e.g. Ahlrichs and Scoles Many more sophisticated schemes

$$E_{total} = E_{DFT} + E_{disp}$$

$$F_{damp}(R) = \frac{1}{1 + e^{-\alpha(R/R_0 - 1)}}$$

$$F_{damp}(R) = \frac{1}{1 + e^{-\alpha(R/R_0 - 1)}}$$

 $s_6 = 1.4/1.3/0.7$  for BLYP/BP86/PBE

Grimme, J. Comput. Chem., 2004, 25, 1463; 2006, 27, 1787.

Parameters are as good for solids (condensed systems) as for molecules Note, however, Mg<sup>2+</sup> very different from Mg, whereas O<sup>2-</sup> similar to O Implementation of Ewald sum for 1/r<sup>6</sup> for periodic systems

Kerber, Sierka, Sauer, J. Comput. Chem. 29 (2008) 2088.

### Methylation of alkenes in H-MFI - pbc DFT+D



Svelle, Tuma, Rozanska, Kerber, Sauer, JACS 131 (2009) 816

### Methylation of alkenes in H-MFI - pbc DFT+D



Svelle, Tuma, Rozanska, Kerber, Sauer, JACS 131 (2009) 816

### **Errors on energy barriers (Truhlar et al.)**

### kcal/mol

	heavy atom	transfer (12)	NS <sup>d</sup> (16)		unimol and assoc <sup>e</sup> (10)		hydrogen transfer (38)	
methods	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE
PBE1PBE	-6.62	6.62	-1.87	2.05	-0.58	2.16	-4.22	4.22
B3LYP	-8.49	8.49	-3.25	3.25	-1.42	2.02	-4.13	4.23
MP2	11.76	11.76	0.74	0.74	4.71	5.44	3.69	4.14
TPSSh	-11.51	11.51	-5.78	5.78	-2.94	3.23	-5.97	5.97
BLYP	-14.66	14.66	-8.40	8.40	-3.38	3.51	-7.52	7.52
TPSS	-14.65	14.65	-7.75	7.75	-3.84	4.04	-7.71	7.71
PBE	-14.93	14.93	-6.97	6.97	-2.94	3.35	-9.32	9.32

#### TABLE 5: Mean Errors for HTBH38 and NHTBH38 Database<sup>a,b,c</sup>

Nucleophilic substitution

Reference data:

CCSD(T) with extrapolation to complete basis set limit("W1 theory")

### **Divide and Conquer - Models and Methodsn**

**DFT (PBE/plane waves)**, VASP 20.2x20.5x13.5 Å; 96 +18 atoms Periodic Boundary Conditions





### Hybrid high level : low level method

Hybrid MP2(cluster):PBE+D(pbc) + $\Delta$ CCSD(T) method  $E_{hybrid}$  (S,C) =  $E_{DFT+D}$ (S) +  $[E_{MP2}(C) - E_{DFT+D}(C)]$ High level correction

Step 0: PBE+D optimization, periodic boundary conditions (pbc) Frequency calculation for stationary points, ZPVE
Step 1: Hybrid MP2(cluster):PBE+D(pbc) optimization
Step 2: Basis set extrapolation to CBS limit, single point
Step 3: CCSD(T)-MP2, small cluster model
Tuma, Sauer, CPL 2004, 387, 388; PCCP, 2006, 8, 3955

Kerber, Sierka, Sauer, *J. Comput. Chem.* 2008, *29*, 2088 Reuter/Scheffler, PRL 98 (2007) 176103 (CO/Cu(111))

Stoll, JPC A 113 (2009) 11483 (Be, Mg crystals)

Energy as functional of orbitals or electron density



### Self-interaction correction

Coulomb and exchange terms cancel - self-interaction correction (SIC)

Kohn-Sham
$$E_J$$
+ $E_{XC}$  (density) $\frac{1}{2} \iint \rho(x_1) \frac{1}{r_{12}} \rho(x_2) dx_1 dx_2 + \int dx \rho(x) V_{XC} [\rho, \nabla \rho]$  $\frac{1}{2} \langle ii | ii \rangle$  $\frac{1}{2} \langle ii | ii \rangle$  $+ \langle i | V_{XC} [\rho, \nabla \rho] | i \rangle \neq 0$ 

Coulomb and exchange do not cancel - self-interaction error

\* 
$$\iint i^*(x_1)j^*(x_2)\frac{1}{r_{12}}i(x_1)j(x_2)dx_1dx_2 = \langle ij|ij \rangle$$

### Energy as functional of orbitals or electron density



### Cluster anions $(V_2O_5)_n^-$



### CCSD(T) calculations confirm most stable structures



Orbital functional for electron correlation: 2nd order Møller-Plesset perturbation theory



Covers dispersion Functionals including this term and Fock exchange: double hybrid

 $E^{DFT} = E_T + E_N + E_J + E_{XC} [\rho] + a_x E_X^{F30} + a_{MP2} E_C^{MP2}$ 



### Wave function expansion

$$\Phi = \sum_{n} C_{n} \Phi_{n} = \Phi_{o} + \sum_{i,a} C_{i}^{a} \Phi_{i}^{a} + \sum_{\substack{i < j, \\ a < b}} C_{ij}^{ab} \Phi_{ij}^{ab} + \sum_{\substack{i < j < k, \\ a < b < c}} C_{ijk}^{abc} \Phi_{ijk}^{abc} + \dots$$

Result depends on

- configuration space "full configuration interaction"
- size of one-particle space (M, given by the basis set) how many configurations of a given substitution level ?

$$\frac{M!}{N!(M-N)!} = 4 \cdot 10^{11} \text{ for CH}_4 \text{ DZP basis M=70, N=10}$$





### Electron correlation - "Post Hartree Fock"

$$\Phi = \Phi_o + \sum_{i,a} C_i^a \Phi_i^a + \sum_{\substack{i < j, \\ a < b}} C_{ij}^{ab} \Phi_{ij}^{ab} + \sum_{\substack{i < j < k, \\ a < b < c}} C_{ijk}^{abc} \Phi_{ijk}^{abc} + \dots$$

$$\langle ab | | rs \rangle = \langle ab | rs \rangle - \langle ab | sr \rangle \qquad \int \psi_a^*(x_1) \psi_b^*(x_2) \frac{1}{r_{12}} \psi_r(x_1) \psi_s(x_2) dx_1 dx_2 = \langle ab | rs \rangle$$

$$E_{corr} = E - E^{HF} = \sum_{\substack{r < s \\ a < b}} C^{rs}_{ab} \left\langle \Psi_0 \left| \hat{H} \right| \Psi^{rs}_{ab} \right\rangle = \sum_{\substack{r < s \\ a < b}} C^{rs}_{ab} \left\langle ab \right| \left| rs \right\rangle$$

You only need double substitution to calculate the exact correlation energy, but to determine the doubles coefficients exactly, you need all substitutions

### Many body perturbation theory (MBPT)

$$\begin{aligned} H_{0} &= \sum_{\zeta} F(\zeta) \\ \Phi_{0} &= - \text{Determinant consisting of canonical HF orbitals} \\ E^{(0)} &= \sum_{i} \mathcal{E}_{i} \\ \langle \Phi_{0} | \hat{H} | \Phi_{a}^{r} \rangle &= 0 \quad (\text{Brillouin Theorem}) \\ \langle \Phi_{0} | \hat{V} | \Phi_{ab}^{rs} \rangle &= \langle \Phi_{0} | \hat{H} | \Phi_{ab}^{rs} \rangle &= \langle ab \| rs \rangle &= \sum_{\mu\nu\kappa\lambda} c_{\mu\alpha} c_{\nu b} c_{\kappa r} c_{\lambda s} \langle \mu \nu \| \kappa \lambda \rangle \end{aligned}$$

$$\begin{aligned} \mathcal{E}_{corr} &= E^{(2)} = -\sum_{a < b} \sum_{r < s} \frac{\langle ab \| rs \rangle}{\mathcal{E}_{r} + \mathcal{E}_{s} - \mathcal{E}_{a} - \mathcal{E}_{b}} \langle ab \| rs \rangle &= E^{MP2} \end{aligned}$$

$$\begin{aligned} \mathcal{U}_{abb}^{rs} \rangle = \langle MP2 \\ \mathcal{U}_{abb}^{rs} \rangle \end{aligned}$$

## Energy (Property) SCF MP3 Limiting value MP4 MP2

### Typical oscillating behaviour of results obtained with the MP method

### **Coupled Cluster Expansion**

$$\Psi = e^{\mathsf{T}} \Phi_0 = (1 + \mathsf{T} + \frac{1}{2}\mathsf{T}^2 + \frac{1}{3!}\mathsf{T}^3 + \dots)\Phi_0$$

$$e^{T} = 1 + T_1 + (T_2 + \frac{1}{2}T_1^2) + (T_3 + T_2T_1 + \frac{1}{3!}T_1^3) + (T_4 + \frac{1}{2}T_2^2 + T_2^2)$$

$$\begin{split} & \mathsf{T}_{1} \Phi_{0} = \sum_{i} \ \sum_{m} t_{i}^{m} \Phi_{i}^{m} & t \text{ - Amplitudes} \\ & \mathsf{T}_{1}^{2} \Phi_{0} = \sum_{i,j} \ \sum_{m,n} t_{i}^{m} t_{j}^{n} \Phi_{ij}^{mn} & \text{disconnected (unlinked) doubles} \\ & \mathsf{T}_{2} \Phi_{0} = \sum_{i,j} \ \sum_{m,n} t_{ij}^{mn} \Phi_{ij}^{mn} & \text{connected (linked) doubles} \end{split}$$

The presence of disconnected (double) substitution ensures size-consistency

# Coupled Cluster Energy $E = \left\langle \Phi_0 \left| H \right| 1 + T_1 + (T_2 + \frac{1}{2}T_1^2) \Phi_0 \right\rangle$ (if HF orbitals are used) $E_{corr} = \sum_{i < i} \sum_{m < n} (t_{ij}^{mn} + t_i^m t_j^n - t_i^n t_j^m) \left\langle ij \right| |mn\rangle$

Equations for determination of "amplitudes" (coefficients)

$$\begin{split} \sum_{im} t_i^m \left\langle \Phi_l^p \left| \boldsymbol{H} \right| \Phi_i^m \right\rangle + \sum_{ij,mn} (t_{ij}^{mn} + t_i^m t_j^n - t_i^n t_j^m) \left\langle \Phi_l^p \left| \boldsymbol{H} \right| \Phi_{ij}^{mn} \right\rangle + \\ \sum_{ij} (t_{ij}^{mn} t_k^o + \dots t_i^m t_j^n t_k^o - \dots) \left\langle \Phi_l^p \left| \boldsymbol{H} \right| \Phi_{ijk}^{mno} \right\rangle = \mathsf{E} t_l^p \end{split}$$

### Iterative solution

### CC-SD equation (HF-Orbitale)

$$\begin{split} \left\langle \Phi_{i}^{m} \left| H \right| (T_{1} + T_{2} + \frac{1}{2}T_{1}^{2} + T_{1}T_{2} + \frac{1}{6}T_{1}^{3})\Phi_{0} \right\rangle &= E \cdot t_{i}^{m} \\ \left\langle \Phi_{ij}^{mn} \left| H \right| (1 + T_{1} + T_{2} + \frac{1}{2}T_{1}^{2} + T_{1}T_{2} + \frac{1}{6}T_{1}^{3} + \frac{1}{2}T_{2}^{2} + \frac{1}{2}T_{1}^{2}T_{2} + \frac{1}{24}T_{1}^{4})\Phi_{0} \right\rangle \\ &= E \cdot (t_{ij}^{mn} + t_{i}^{m}t_{j}^{n} - t_{i}^{n}t_{j}^{m}) \end{split}$$

CCSD(T) - (T)riple-contribution according to MP4 equation, but T1 and T2 amplitudes are taken from CCSD equations Wavefunction-based methods for electron correlation

I. According to the way, the expansion coefficients are determined variational: Configuration Interaction (CI methods) Size-consistency problem CIS (CI with singles) for excited states perturbation theory: Møller-Plesset MP2, MP4-SDTQ system of equations: Coupled Cluster (iterative) CCSD; CCSD(T)

### Wavefunction-based methods for electron correlation

II. Substitutions included in wavefunction expansion all in given one-particle basis: full CI up to quadruple: MP4-SDTQ CCSD(T) up to connected triple substitution: (an expensive standard method) up to doubles:  $CI-SD \rightarrow CPF$ doubles only: MP2 (the simplest wavefunction-based standard method)) CI-S singles only: (simplest method for excited states)

**TABLE 1.** Statistical Data (in kcal mol<sup>-1</sup>) for 21 Reaction Energies (Small Molecules) Calculated with Different Quantum Chemical Methods with the def2-QZVPP Basis Set<sup>a</sup>

	deviat	deviation from CCSD(T) (exp.) reference				
	MAD	MD	MAX			
CCSD(T)	(1.76)	(-0.82)	(6.73)			
CCSD	3.41 (4.78)	-3.02 (-3.84)	14.13 (17.23)			
QCISD	3.07 (4.44)	-2.71 (-3.53)	12.11 (16.27)			
CEPA/1	1.08 (2.53)	-0.35 (-1.17)	2.99 (7.14)			
NCPF/1	1.03 (2.42)	-0.26 (-1.08)	2.44 (6.84)			
SCS-MP3	2.22 (3.01)	-1.21 (-2.03)	7.06 (11.91)			
SCS-MP2	3.41 (2.78)	-0.34 (-1.17)	9.58 (15.36)			
MP2	5.54 (4.99)	-2.76 (-3.58)	26.39 (33.12)			
B2PLYP	2.68 (2.45)	0.55 (—0.28)	6.33 (7.93)			
B3LYP	3.18 (2.91)	1.72 (0.90)	8.72 (7.26)			

Neese et al., Acc. Chem. Res. 42 (2009) 641-648

### The scaling problem

Table 7.4Formal scaling behavior, as a function of basisfunctions N, of various electronic structure methods

Scaling behavior	Method(s)		
N <sup>4</sup>	HF		
N <sup>5</sup>	MP2		
N <sup>6</sup>	MP3, CISD, MP4SDQ, CCSD, QCISD		
N <sup>7</sup>	MP4, CCSD(T), QCISD(T)		
N <sup>8</sup>	MP5, CISDT, CCSDT		
N <sup>9</sup>	MP6		
N <sup>10</sup>	MP7, CISDTQ, CCSDTQ		

# Electron correlation: requires better (larger) basis sets than Hartree Fock or DFT

Polarisation functions indispensable/diffuse functions recommended

Pople school	correlation consistent			
(GAUSSIAN)	Dunning			
6-311G** 6-311+G(2df,p) 6-311+G(3df,2p)	cc-pVDZ cc-pVTZ cc-pVQZ	3s2p1d/2s1p 4s3p2d1f/3s2p1d 5s4p3d2f1g/4s3p2d1f		
+ diffuse s-function	aug - diffuse functions (one for each angular momentum)			

### Dunning (aug)-cc-pVnZ Basis Sets

	S	р	d	f	g	h	No.AO X/H
cc-pVDZ	3/ <u>2</u>	<u>2</u> /1	1/0				14/05
cc-pVTZ	4/ <u>3</u>	<u>3</u> /2	2/1	1/0			30/14
cc-pVQZ	5/ <u>4</u>	<u>4</u> /3	3/2	2/1	1/0		55/30
cc-pV5Z	6/ <u>5</u>	<u>5</u> /4	4/3	3/2	2/1	1/0	91/55
aug-cc-pVDZ	4/3	3/2	2/0				23/09
aug-cc-pVTZ	5/4	4/3	3/2	2/0			46/23
aug-cc-pVQZ	6/5	5/4	4/3	3/2	2/0		80/46
aug-cc-pV5Z	7/6	6/5	5/4	4/3	3/2	2/0	127/80

I/k refers to first row atom/h atom

How to reach the basis set limit for a given method?

- Use basis sets of increasing size
- Reaching the basis set limit faces the scaling problem
- More expensive methods generally need larger basis sets to reach limit
- Extrapolation formulas available

### Basis set extrapolation - CBS limit

(aug)-cc-pVXZ

 $E_{\operatorname{corr},X} = E_{\operatorname{corr,lim}} + AX^{-3}$ 

$$E_{\text{corr,lim}} = \frac{X^3}{X^3 - (X-1)^3} E_{\text{corr},X} - \frac{(X-1)^3}{X^3 - (X-1)^3} E_{\text{corr},X-1}$$

Two-point extrapolation A. Halkier et al.,CPL 286 (1988) 243 other and more sophisticated schemes, see refs. in T. H. Dunning, Jr., JPC A 104 (2000) 9062

Note: You need to extrapolate also the Hartree-Fock energy

### O<sub>2</sub> dissociation energy (kJ/mol)

Method	Basis set	De	$D_0$
Obsd		505	494
UCCSD(T)FC	cc-pVTZ	476	
	cc-pVQZ	490	
	CBS(T,Q)	501	
	aug-cc-pVTZ	481	
	CBS	501	
	CBS+cv	503	
G2	G2	494	484

### EMSL Computational Results DataBase - 255 Molecules, 41 Atoms http://www.emsl.pnl.gov:2080/proj/crdb/

(kJ/mol)				
	R <sub>Au-S</sub>	ΔE		
PBE/TZVPP	247	-53.8		
PBE+D <sup>b</sup> /TZVPP//PBE	247	-56.9		
B3LYP/TZVPP//B3LYP	260	-26.5		
B3LYP+D <sup>b</sup> /TZVPP	260	-30.7		
CCSD(T)//B3LYP+D <sup>a</sup>	ΔΕ	∆E/CBS(X,Y		
aug-pVDZ	-27.4	-		
aug-pVTZ	-32.8	-34.7 (D,T)		
aug-pVQZ	-35.4	-37.0 (T,Q)		
aug-pV5Z	-36.7	-37.7 (Q,5)		

Au-S bond distance (pm) and H<sub>2</sub>S-Au interaction energies



Н

Н

Αι



Boys Bernardi function counterpoise method



## BSSE

### BSSE Basis set superposition error BSCE Basis set convergence error

Thom H. Dunning Jr. "A road map for the calculation of molecular binding energies" J Phys Chem A 2000, 104, 9062-9080

### Notes on BSSE

- BSSE for two subsystems is an indicator of balanced basis sets
- For small basis sets BSSE correction may lead to larger deviations from experiment (or basis set limit)
- BSSE corrected results are less basis set dependent
- There may be other errors (BSCE)
- Sizeable for intermolecular complexes, but also present for chemical bonds
- BSSE elimination is not always easy. How do you correct an energy barrier for BSSE?
- There is no BSSE for plane wave basis sets

There are cases, for which the Hartree Fock determinant is not a sufficiently good starting point
 Stretched, almost dissociated bonds (H<sub>2</sub> example)
 Biradicals, polynuclear transition metal compounds
 Almost degenerate states (partially occupied d-shells)

Multireference cases CAS (Complete Actice Space) SCF MC (Multi-configuration) SCF + MR (Multireference)-CI(SD)

### The challenge of transition metal oxides

GGA functionals (PBE) are often not good enough, but hybrid functionals (B3LYP, PBE0, HSE) are one order of magnitude more expensive (with plane waves and pbc two orders of magnitude)

GGA functionals yield more delocalized electronic structures (delocalization of holes), preference for polar structures and too low energy barriers.

Partially occupied d-shells, the presence of more than one TM site and the decoupling of electron pairs in bond breaking processes requires a multi-reference treatment

- DFT limited to high-spin, low spin states require broken symmetry approximation

- in general, CCSD(T) is not good enough and not suitable as benchmark

- ab initio multi-reference methods cannot treat large enough systems reliably.

## T<sub>1</sub> diagnostics

Is CCSD(T) applicable? Is a single determinant a good reference function?

$$\Psi = \left[ 1 + T_1 + (T_2 + \frac{1}{2}T_1^2) + (T_2T_1 + \frac{1}{3!}T_1^3) + \dots \right] \Phi_0$$
  
$$T_1 \Phi_0 = \sum_i \sum_m t_i^m \Phi_i^m$$
  
$$T_1 = \left\| t_1 \right\| / \sqrt{N_e} \quad \left\| t_1 \right\| = \sum_i \sum_m (t_i^m)^2$$

### T1 > 0.02 not a clear single reference state

### H-Transfer: surface complex - gas phase model

	$H - CH_{2}$ $I$ $V - V$ $V - V$ $V - V$ $V' = V$ $V'' = V'' = V''' = V'''' = V''''' = V''''' = V''''' = V''''''''$	H $\cdot CH_2$   $ $ $ $ $0\vee \cdot \circ\vee \cdot \circ\vee \cdot \circ(1)$ - biradicaloid ken symmetry	$ \begin{array}{c} H & CH_2 \\ I & II \\ O & O \\ V & O \\ O & O \\ I & I \\ V^{III}(d^2) - triplet \end{array} $		
	$V_2O_5/SiO_2$	$OV(OCH_3)$	$OV(OCH_3)^+$		
Calculated	broken	symmetry	doublet		
E <sup>#</sup> ,B3LYP	154	147	80		
E <sup>#</sup> ,CCSD(T)	(137)	131			
Observed	$\mathbf{X}$		T1=0.07		
E <sup>#</sup> <sub>intrinsic</sub> (0 K)	137±10	estimate	52±11		
Döbler, Pritzsche, Sauer, JACS 127 (2005) 10861					

If you would like to join our group as postdoc or PhD student -

js@chemie.hu-berlin.de

(www.chemie.hu-berlin.de/ag\_sauer