

Intermolecular interactions

János Ángyán

Équipe Modélisation Quantique et Cristallographique, LCM3B

– 2008 –

Part I

Supermolecule approach

Overview of the supermolecule approach

- interaction energy is given by a very small difference of large numbers

$$\Delta E(R) = E^{AB}(R) - E^A - E^B$$

- advantages
 - straightforward to "implement"
 - the whole PES can be treated on an equal footing
 - antisymmetry of the wave function is guaranteed
 - no problem of multipole or PT convergence
- drawbacks
 - small difference of large numbers
 - systematic errors are much larger than ΔE itself
 - limited to size-consistent methods
 - basis set superposition error
 - difficult to interpret

Choice of the quantum chemical method

- semiempirical methods: inappropriate
- DFT: problems with dispersion (see later)
- ab initio
 - Hartree-Fock: misses dispersion, but reasonable for polar systems, H-bonds
 - Configuration Interaction: inappropriate due to size inconsistency (excepted full CI)
 - MBPT: at low orders (MP2) polarizabilities are of only UCHF quality
 - coupled cluster: CCSD(T) is considered as the best compromise
 - explicitly correlated (r_{12}) methods: costly but good for small systems
 - Quantum Monte Carlo: able to provide reliable benchmarks values

Quality of charge density and response properties

- should provide the best possible monomer properties
- attention: a correlated method may "hide" non-correlated monomer properties
- MP2 dimer calculations:
 - monomer charge densities are of HF quality
 - monomer polarizabilities are uncoupled HF quality
- density functional charge densities are usually good

Asymptotic behavior of the Hartree-Fock interaction energy

$$\begin{aligned}
 E_{\text{int}}^{\text{HF}} &\sim \frac{1}{R^3} \mu_A^{\text{HF}} \mu_B^{\text{HF}} \\
 &+ \frac{1}{R^4} (\mu_A^{\text{HF}} \theta_B^{\text{HF}} + \theta_A^{\text{HF}} \mu_B^{\text{HF}}) \\
 &+ \frac{1}{R^5} (\theta_A^{\text{HF}} \theta_B^{\text{HF}} + \mu_A^{\text{HF}} \Omega_B^{\text{HF}} + \Omega_A^{\text{HF}} \mu_B^{\text{HF}}) \\
 &+ \frac{1}{R^6} [\alpha_A^{\text{CHF}} (\mu_B^{\text{HF}})^2 + \mu_A^{\text{HF}} \chi_B^{\text{HF}} + \theta_A^{\text{HF}} \Omega_B^{\text{HF}} + A \leftrightarrow B] \\
 &+ \dots \\
 &+ \frac{1}{R^9} [\beta_A^{\text{CHF}} (\mu_B^{\text{HF}})^3 + (\mu_A^{\text{HF}})^3 \beta_B^{\text{CHF}} + \alpha_A^{\text{CHF}} \mu_A^{\text{HF}} \alpha_B^{\text{CHF}} \mu_B^{\text{HF}} + \dots] \\
 &+ \dots
 \end{aligned}$$

Asymptotic behavior of the MBPT2 interaction energy

$$\begin{aligned}
 E_{\text{int}}^{\text{MP2}} \sim & \frac{1}{R^3} (\mu_A^{\text{MP2}} \mu_B^{\text{HF}} + \mu_A^{\text{HF}} \mu_B^{\text{MP2}} + \mu_A^{\text{MP2}} \mu_B^{\text{MP2}}) \\
 & + \frac{1}{R^4} (\mu_A^{\text{MP2}} \theta_B^{\text{HF}} + \theta_A^{\text{MP2}} \mu_B^{\text{HF}} + A \leftrightarrow B) \\
 & + \frac{1}{R^5} (\theta_A^{\text{MP2}} \theta_B^{\text{HF}} + \mu_A^{\text{MP2}} \Omega_B^{\text{HF}} + \Omega_A^{\text{MP2}} \mu_B^{\text{HF}} + A \leftrightarrow B) \\
 & + \frac{1}{R^6} [\alpha_A^{\text{MP2}} (\mu_B^{\text{HF}})^2 + 2 \alpha_A^{\text{CHF}} \mu_B^{\text{MP2}} \mu_B^{\text{HF}} + \mu_A^{\text{MP2}} \chi_B^{\text{HF}} \\
 & \quad + \mu_A^{\text{HF}} \chi_B^{\text{MP2}} + \theta_A^{\text{MP2}} \Omega_B^{\text{HF}} + \theta_A^{\text{HF}} \Omega_B^{\text{MP2}} + A \leftrightarrow B \\
 & \quad + \frac{3}{\pi} \int_0^\infty \alpha_A^{\text{UCHF}}(i\omega) \alpha_B^{\text{UCHF}}(i\omega) d\omega] \\
 & + O(R^{-7})
 \end{aligned}$$

Asymptotic behavior of the CCSD interaction energy

$$\begin{aligned}
 E_{\text{int}}^{\text{CCSD}} \sim & \frac{1}{R^3} (\mu_A^{\text{CCSD}} \mu_B^{\text{HF}} + \mu_A^{\text{HF}} \mu_B^{\text{CCSD}} + \mu_A^{\text{CCSD}} \mu_B^{\text{CCSD}}) \\
 & + \frac{1}{R^4} (\mu_A^{\text{CCSD}} \theta_B^{\text{HF}} + \theta_A^{\text{CCSD}} \mu_B^{\text{HF}} + A \leftrightarrow B) \\
 & + \frac{1}{R^5} (\theta_A^{\text{CCSD}} \theta_B^{\text{HF}} + \mu_A^{\text{CCSD}} \Omega_B^{\text{HF}} + \Omega_A^{\text{CCSD}} \mu_B^{\text{HF}} + A \leftrightarrow B) \\
 & + \frac{1}{R^6} [\alpha_A^{\text{CCSD}} (\mu_B^{\text{HF}})^2 + 2 \alpha_A^{\text{CHF}} \mu_B^{\text{CCSD}} \mu_B^{\text{HF}} + \mu_A^{\text{CCSD}} \chi_B^{\text{HF}} \\
 & \quad + \mu_A^{\text{HF}} \chi_B^{\text{CCSD}} + \theta_A^{\text{CCSD}} \Omega_B^{\text{HF}} + \theta_A^{\text{HF}} \Omega_B^{\text{CCSD}} + A \leftrightarrow B \\
 & \quad + \frac{3}{\pi} \int_0^\infty \alpha_A^{\text{CHF}}(i\omega) \alpha_B^{\text{CHF}}(i\omega) d\omega + ???] \\
 & + O(R^{-7})
 \end{aligned}$$

Choice of the basis set

Requirements:

- First-order energy: correct multipoles and tails of monomers $\rightarrow$ diffuse functions
- Second-order energy: good (static and dynamic) polarizabilities $\rightarrow$ polarization functions
 - dipole polarizabilities: d functions
 - quadrupole polarizabilities: f functions
 - ...
- balanced basis set: valence and polarization part of similar quality
 - good valence - poor polarization : underestimates induction
 - poor valence - good polarization : overestimates induction

Precise calculations require very large basis sets, (multiple zeta quality + f,g,h orbitals).
High angular momentum functions can be replaced by *bond functions*.

Role of bond functions

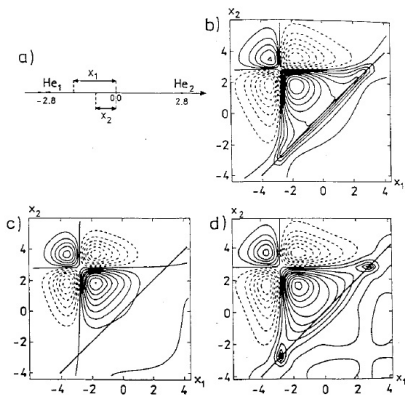


Figure 1. Plots of the dispersion function for the equilibrium He dimer from ref 61. The electron coordinates, x_1 and x_2 are defined in part a. The accurate dispersion function (calculated with Gaussian geminals) is shown in plot b. If the dispersion term is calculated with a monomer-centered basis set, one obtains plot c. If every electron is allowed to use the complete dimer-centered basis set (DCBS) as well as the bond functions, one obtains plot d.

Since the basis functions stick on the atoms, one has a different basis set (of unequal quality) in the function of the intermolecular separation/orientation.

$$\Delta E = E_{\text{AUB}}(\mathbf{R}, \mathbf{\Omega}) - E_{\text{AUB}}(\mathbf{R} = \infty, \mathbf{\Omega}) = E_{\text{AUB}}(\mathbf{R}, \mathbf{\Omega}) - E_{\text{A}} - E_{\text{B}}$$

At small separations the partner basis functions improve the monomer description (nothing to do with the physical interaction) and lead to an extra stabilization of the complex.

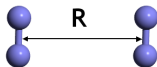
Boys-Bernardi counterpoise correction:

At each geometry evaluate monomer energies in the dimer basis

$$\Delta E = E_{\text{AUB}}(\mathbf{R}, \mathbf{\Omega}) - E_{\text{A}}(\text{A} \cup \text{B} | \mathbf{R}, \mathbf{\Omega}) - E_{\text{B}}(\text{A} \cup \text{B} | \mathbf{R}, \mathbf{\Omega})$$

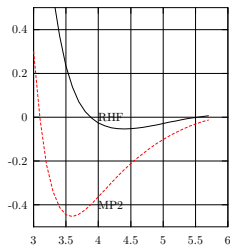
Problems: e.g. monomer properties in the dimer basis have lower symmetry

Nitrogen dimer

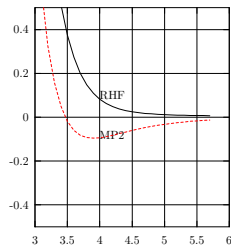


6-311+G(d)

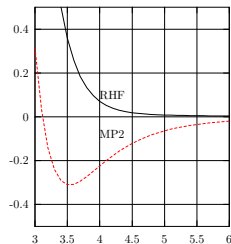
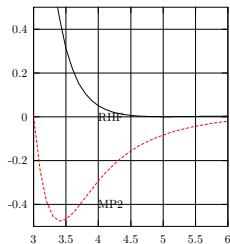
without BSSE correction



with BSSE correction



AUG-cc-pVTZ



Basis incompleteness correction

It is more judicious to consider the CP correction as a *basis set incompleteness correction*. If the monomer basis were complete, the difference

$$E_A(A \cup B | \mathbf{R}, \mathbf{\Omega}) - E_A$$

would vanish. For incomplete basis set, this error (per monomer) represents an extra stabilisation, that should be subtracted from the *total energy*:

$$E_{AUB}^{corr}(\mathbf{R}, \mathbf{\Omega}) = E_{AUB}(\mathbf{R}, \mathbf{\Omega}) - (E_A(A \cup B | \mathbf{R}, \mathbf{\Omega}) - E_A) - (E_B(A \cup B | \mathbf{R}, \mathbf{\Omega}) - E_B)$$

New total energy functional, that can be used to calculate BSSE-corrected dimer properties, and one can even define BSSE-free effective Hamiltonian. (Mayer, 2003)
The interaction energy is defined with respect to the “true” monomer energies:

$$\Delta E = E_{AUB}^{corr}(\mathbf{R}, \mathbf{\Omega}) - E_A - E_B$$

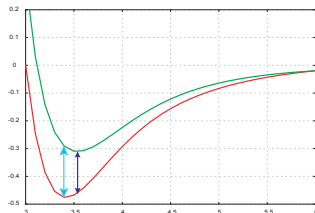
Counterpoise correction "overshoots"?

- Argument: the CP method overcorrects BSSE, since it uses the occupied orbitals of the partner to improve the description of the monomers
- Proposal to "remedy": use only the virtual orbitals of the partner in the CP procedure (VCP)
- Example: He_2 at $R = 5.6\text{a.u.}$ energy in K

basis	no CP	VCP	FCP	PT
1	-10.29	-7.66	-7.44	-7.33
2	-39.67	-12.09	-8.81	-8.81
3	-8.48	-8.31	-8.28	-8.30
4	-9.82	-9.29	-9.19	-9.19

- Error: in the dimer calculation monomer A does not use the occupied orbitals of B. The fact that the orbitals of B are occupied, is at the origin of the overlap repulsion. If we do not allow A to explore the occupied space of B, we underestimate the physical repulsion.
- Conclusion: we need all the ghost orbitals!

BSSE-free geometry optimization



- BSSE correction at the minimum of the uncorrected surface is larger than in the true minimum.
- Find the minimum of the corrected dimer energy!

Gradient optimizations using "composite" energy expression (and gradients).

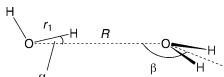


Fig. 1 Structure of the water dimer.

Basis set	PES	R	r_1	α	β	ΔE
6-31G**	CP-corr	2.991	0.967	2.7	129.5	-21.21
	Standard	2.913	0.967	9.8	99.1	-19.41
cc-pVDZ	CP-corr	3.048	0.969	-0.1	138.2	-17.61
	Standard	2.912	0.970	5.7	104.5	-15.02
cc-pVTZ	CP-corr	2.973	0.965	3.6	129.6	-18.70
	Standard	2.908	0.966	5.1	115.6	-18.28
cc-pVQZ	CP-corr	2.935	0.964	4.6	127.2	-19.71
	Standard	2.901	0.965	4.8	121.0	-19.62
aug-cc-pVDZ	CP-corr	2.977	0.972	5.7	122.3	-18.62
	Standard	2.921	0.973	6.4	119.6	-18.45
aug-cc-pVTZ	CP-corr	2.933	0.968	5.5	124.7	-20.21
	Standard	2.905	0.969	4.9	124.3	-19.75
aug-cc-pVQZ	CP-corr	2.917	0.966	5.6	124.8	-20.59
	Standard	2.903	0.966	5.1	124.7	-20.54
exp		2.910		5 ± 10	123 ± 10	-22.78 ± 2.93

Local correlation methods

- Take into account excitations from localized orbitals to virtual orbitals in the same space domain
- Local-MP2 (Pulay), Local-CCSD(T) (Werner et al.)
- Exclusion of "charge-transfer" excitation decreases dramatically the BSSE

CBS: complete basis set limit

- Energy extrapolation in a series of calculations by systematically improving the basis set, e.g.

aug-cc-pVXZ

X=D,T,Q,5,6

- Extrapolation formulae:

$$\begin{array}{lll} E_X^{HF} = E_{CBS}^{HF} + Ae^{-\alpha X} & E_X^{corr} = E_{CBS}^{corr} + BX^{-3} & \text{Helgaker} \\ E_X^{HF} = E_{CBS}^{HF} + AX^{-\alpha} & E_X^{corr} = E_{CBS}^{corr} + BX^{-\beta} & \text{Truhlar} \end{array}$$

- Impractical for costly CCSD(T) calculations
- MP2 has similar basis set dependence

$$\Delta E_{CBS}^{CCSD(T)} = \Delta E_{CBS}^{MP2} + (\Delta E_{medium}^{CCSD(T)} - \Delta E_{medium}^{MP2})$$

- *in principle at the CBS limit there is no BSSE!*

Size consistency

- Energy of non-interacting dimer ($R \rightarrow \infty$) should be equal to the sum of monomer energies
- The energy of an N -particle system should be proportional to the number of particles (N) even when $N \rightarrow \infty$
- example of N noninteracting H_2 minimal basis molecules (Szabó-Ostlund) at the DCI level.

- Exact correlation energy

$$E_{corr}(N) = N((\Delta - (\Delta^2 + K_{12}^2)^{1/2})$$

where

$$\Delta = \langle \Psi_{11}^{22} | \hat{H} - E_0 | \Psi_{11}^{22} \rangle \quad K_{12} = (12|21)$$

- DCI correlation energy of N non-interacting H_2

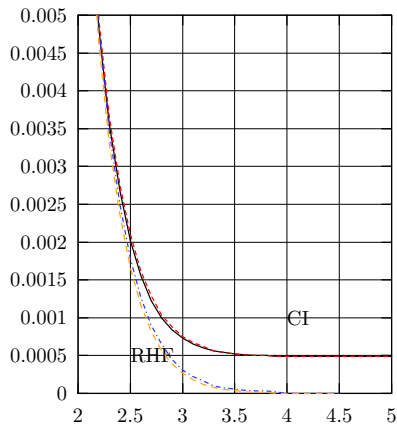
$$E_{corr}^{DCI}(N) = \Delta - (\Delta^2 + NK_{12}^2)^{1/2})$$

- the correlation energy converges to zero with $N \rightarrow \infty$:

$$\lim_{N \rightarrow \infty} \frac{E_{corr}^{DCI}(N)}{N} = 0$$

Size consistency problem

H₂ dimer
minimal basis



- CISD interaction energy does not tend to zero!

- CISD monomer wf:

$$\sigma_g^{A^2} - \lambda \sigma_u^{A^2}$$

- product wf of the dimer:

$$(\sigma_g^{A^2} - \lambda \sigma_u^{A^2})(\sigma_g^{B^2} - \lambda \sigma_u^{B^2})$$

- contains up to *quadruple* excitations:

$$\lambda^2 \sigma_u^{A^2} \sigma_u^{B^2}$$

- unbalanced description of monomer and dimer

Size-consistency of MBPT methods

N-mer of non-interacting H₂ molecules

Second order energy

$$E^{(2)} = \sum_{i=1}^N \frac{|\langle \Psi | \hat{W} | \Psi_{1_i \bar{1}_i}^{2_i \bar{2}_i} \rangle|^2}{2(\varepsilon_1 - \varepsilon_2)} = \frac{N K_{12}^2}{2(\varepsilon_1 - \varepsilon_2)}$$

is N times the monomer MP2 energy.

Third order energy correction

$$E^{(3)} = \frac{N K_{12}^2 (J_{11} + J_{22} - 4J_{12} + 2K_{12})}{4(\varepsilon_1 - \varepsilon_2)^2}$$

The general result is true: the MP_n energy correction is size-consistent.

Other size-consistent methods

The size consistency can be qualitatively discussed by invoking the one- and two-particle excitation operators

$$\hat{T}_1 = \sum_{ar} T_a^r a_r^+ a_a \qquad \hat{T}_2 = \sum_{abrs} T_{ab}^{rs} a_r^+ a_s^+ a_b a_a$$

- SDCl wave function of a complex

$$(1 + \hat{T}_1^A + \hat{T}_2^A)|\Psi^A\rangle(1 + \hat{T}_1^B + \hat{T}_2^B)|\Psi^B\rangle = |\Psi^{AB}(SDCl)\rangle + (\hat{T}_1^A \hat{T}_2^B + \dots)|\Psi^A \Psi^B\rangle$$

which is not of SCDI form

- coupled cluster wave functions (CCSD) of a complex

$$\exp(\hat{T}_1^A + \hat{T}_2^A)|\Psi^A\rangle \exp(\hat{T}_1^B + \hat{T}_2^B)|\Psi^B\rangle = \exp(\hat{T}_1^A + \hat{T}_1^B + \hat{T}_2^A + \hat{T}_2^B)|\Psi^A \Psi^B\rangle$$

is still of the CCSD form.

Interaction energy partitioning schemes

- Kitaura-Morokuma scheme
- CSOV: Constrained Space Orbital Variation
- valid for HF, MCSCF and DFT methods, respects antisymmetry
- energy components:

$$\Delta E = E_{HL} + E_{IND} + E_{CT}$$

variational space

- Heitler-London

$$E_{HL} = E_{ES} + E_{EX-REP}$$

- induction uses "own" virtual orbitals

$$E_{IND} = E_{IND,A} + E_{IND,B}$$

- charge transfer uses the full virtual space

$$E_{CT} = E_{CT,A} + E_{CT,B}$$

Some state-of-the-art examples

Table 3 Interaction energies (in kcal mol⁻¹) for model complexes (set S22). Deformation energy of monomers is not included

No.	Complex (symmetry)	$\Delta E^{\text{MP2}}_1^a$	$\Delta E^{\text{MP2}}_2^a$	$\Delta E^{\text{MP2}}_{\text{CBS}}^a$	$E^{\text{int}}_{\text{CCSD(T)/CBS}}^b$	Geometry ^c
Hydrogen bonded complexes (7)						
1	(NH ₃) ₂ (C _{2h})	-3.02 (QZ)	-3.10 (SZ)	-3.20	-3.17 (qz)	CCSD(T)/QZ
2	(H ₂ O) ₂ (C _s)	-4.75 (QZ)	-4.89 (SZ)	-5.03	-5.02 (qz)	CCSD(T)/QZ
3	Formic acid dimer (C _{2h})	-17.88 (QZ)	-18.23 (SZ)	-18.60	-18.61 (tz)	CCSD(T)/TZ
4	Formamide dimer (C _{2h})	-15.19 (QZ)	-15.52 (SZ)	-15.86	-15.96 (tz)	CCSD(T)/TZ
5	Uracil dimer (C _{2h})	-19.90 (TZ)	-20.28 (QZ)	-20.61	-20.65 (tz-fd)	MP2/TZ-CP
6	2-pyridoxine · 2-aminopyridine (C ₁)	-15.91 (TZ)	-16.77 (QZ)	-17.37	-16.71 (tz-fd)	MP2/TZ-CP
7	Adenine · thymine WC (C ₁)	-14.92 (TZ)	-15.89 (QZ)	-16.54	-16.37 (dz)	MP2/TZ-CP
Complexes with predominant dispersion contribution (8)						
8	(CH ₄) ₂ (D _{3d})	-0.42 (QZ)	-0.46 (SZ)	-0.51	-0.53 (qz)	CCSD(T)/TZ
9	(C ₂ H ₄) ₂ (D _{2d})	-1.43 (QZ)	-1.57 (SZ)	-1.62	-1.51 (qz)	CCSD(T)/QZ
10	Benzene · CH ₄ (C ₃)	-1.66 (QZ)	-1.75 (SZ)	-1.86	-1.50 (tz-fd)	MP2/TZ-CP
11	Benzene dimer (C _{2h})	-4.70 (aT)	-4.85 (aQ)	-4.95	-2.73 (adz)	MP2/TZ-CP
12	Pyrazine dimer (C _s)	-6.56 (aT)	-6.76 (aQ)	-6.90	-4.42 (tz-fd)	MP2/TZ-CP
13	Uracil dimer (C ₂)	-10.63 (TZ)	-10.99 (QZ)	-11.39	-10.12 (tz-fd)	MP2/TZ-CP
14	Indole · benzene (C ₁)	-6.44 (TZ)	-7.42 (QZ)	-8.12	-5.22 (dz)	MP2/TZ-CP
15	Adenine · thymine stack (C ₁)	-12.30 (TZ)	-13.83 (QZ)	-14.93	-12.23 (dz)	MP2/TZ-CP
Mixed complexes (7)						
16	Ethene · ethine (C _{2v})	-1.57 (QZ)	-1.62 (SZ)	-1.69	-1.53 (tz)	CCSD(T)/QZ
17	Benzene · H ₂ O (C _s)	-3.28 (QZ)	-3.43 (SZ)	-3.61	-3.28 (tz-fd)	MP2/TZ-CP
18	Benzene · NH ₃ (C _s)	-2.44 (QZ)	-2.57 (SZ)	-2.72	-2.35 (tz-fd)	MP2/TZ-CP
19	Benzene · HCN (C _s)	-4.92 (aT)	-5.06 (aQ)	-5.16	-4.46 (tz-fd)	MP2/TZ-CP
20	Benzene dimer (C _{2v})	-3.46 (aT)	-3.55 (aQ)	-3.62	-2.74 (adz)	MP2/TZ-CP
21	Indole · benzene T-shape (C ₁)	-6.16 (TZ)	-6.65 (QZ)	-7.03	-5.73 (dz)	MP2/TZ-CP
22	Phenol dimer (C ₁)	-6.71 (TZ)	-7.33 (QZ)	-7.76	-7.05 (tz-fd)	MP2/TZ-CP

Hobza et al. Phys. Chem. Chem. Phys., 2006, 8, 1985

Ammonia dimer

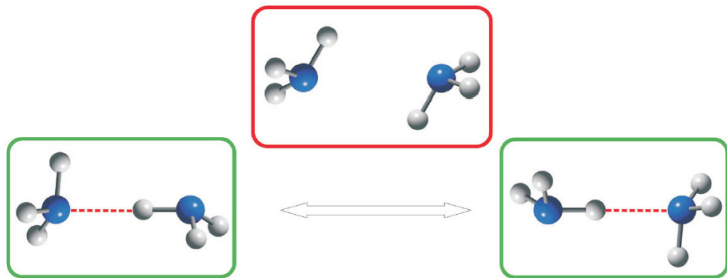


FIG. 3
Ammonia dimer structure as it appears after vibrational averaging

Water dimer

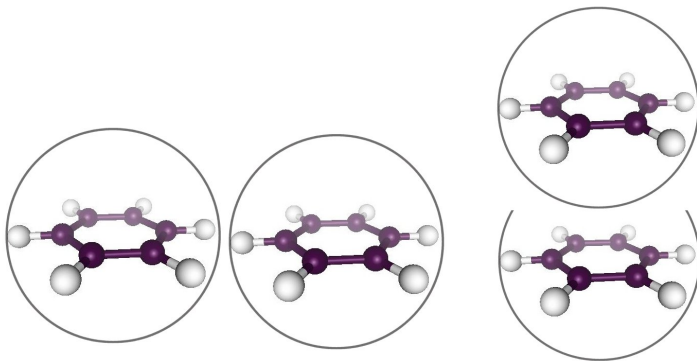
	SAPT	CCSD(T)/MP2	CCSD(T)/MP2-R12	Extrapol.	Exp.
α	4.6	4.3	4.8	5.5	5 ± 10
θ	124	128.2	122.5	124.4	123 ± 10
R_{OO}	5.58	5.527	5.472	5.503	5.50 ± 0.01
E_{min}	-4.86	-4.92	-4.94	-5.02	-5.44 ± 0.7

Part II

Distributed multipoles and polarizabilities

Why distribute multipoles?

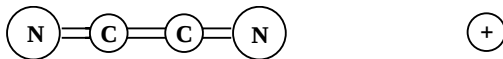
Shape convergence problem: Molecular dimensions may become comparable or larger than intermolecular separations, invalidating the multipole expansion.



By reducing the domains of multipole expansion, the convergence of the series is ensured for all possible intermolecular contacts.

Further difficulties with induction energy

Uncertainty in the choice of the origin may influence strongly induction energies:



Field	=	0.027	0.050	0.111	a.u
Dipole	=	0.036	0.066	0.145	D
- U_{ind}	=	0.126	0.397	2.013	kcal/mole

The field of the point charge varies strongly along the molecular axis. Depending of the (arbitrary) choice of the origin, the induction energy can vary considerably. Which is the “true” induction energy?

Properties of distributed polarizabilities

- Exact reconstitution of molecular multipole polarizabilities

$$\alpha_{LM,L'M'} = \sum_{\ell \leq L} \sum_{\ell' \leq L'} W_{L-\ell,M-m}^a \alpha_{\ell m,\ell' m'}^{aa'} W_{L'-\ell',M'-m'}^{a'}$$

with the multipole translation function:

$$W_{L-\ell,M-m}^a = \left[\begin{pmatrix} L+M \\ \ell+m \end{pmatrix} \begin{pmatrix} L-M \\ \ell-m \end{pmatrix} \right]^{\frac{1}{2}} R_{L-\ell,M-m}(\mathbf{r}^a - \mathbf{R}_A)$$

- Sum rule: charge conservation

$$\sum_a \alpha_{00,\ell m}^{ab} = 0$$

How to obtain distributed multipole operators?

Partitioning schemes to define atoms (or functional groups)

- $\mathbb{L}^2$ (basis set) partitioning in AO basis

$$\hat{Q}_{\ell m}^a = \sum_{\mu\nu\lambda} S_{\lambda\mu}^{-1} f_{\mu\nu}^a \left(\sum_{\ell' \leq \ell} W_{\ell-\ell', m-m'}^a \langle \chi_\mu | \hat{Q}_{\ell' m'}^{\mu\nu} | \chi_\nu \rangle \right) \chi_\lambda^+ \tilde{\chi}_\nu^-$$

- $\mathbb{R}^3$ (real-space) partitioning

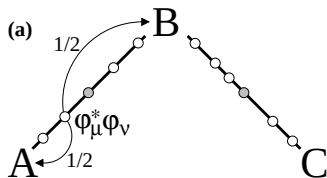
$$\hat{Q}_{\ell m}^a = \sum_{ij} \langle \phi_i | \hat{Q}_{\ell m}^a | \phi_j \rangle_A \phi_i^+ \phi_j^-$$

with the atomic multipole integral defined over an atomic domain, Ω_a :

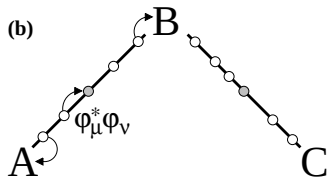
$$\langle \phi_i | \hat{Q}_{\ell m}^a | \phi_j \rangle_A = \int_{\Omega_a} d\mathbf{r} \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) R_{\ell m}^a(\mathbf{r})$$

Basis set partitioning schemes

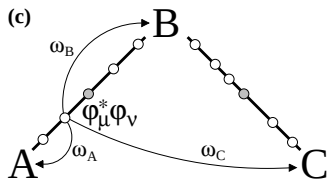
- Mulliken, Sokalski's CAMM



- Stone



- Claverie



Distributed multipole schemes

weight	order	rms	$\Delta\theta$	(std.dev.)	$\Delta\rho$	(std.dev.)
SOKA	HEXA	2.042	0.900	(0.744)	-0.060	(1.839)
STON	HEXA	1.121	0.490	(0.402)	-0.040	(0.981)
CLAV	HEXA	0.829	0.360	(0.290)	-0.040	(0.700)
SOKA	OCTU	3.718	1.620	(1.218)	0.020	(2.976)
STON	OCTU	2.190	0.920	(0.735)	0.030	(1.697)
CLAV	OCTU	1.753	0.730	(0.628)	0.040	(1.350)
SOKA	QUAD	16.641	8.360	(7.237)	2.060	(12.852)
STON	QUAD	11.306	5.480	(4.865)	1.410	(9.644)
CLAV	QUAD	10.086	4.800	(4.386)	1.350	(8.957)
SOKA	DIPOL	28.599	13.520	(14.679)	-5.180	(24.532)
STON	DIPOL	19.182	8.370	(7.622)	-3.560	(15.676)
CLAV	DIPOL	16.562	7.150	(6.356)	-2.770	(13.209)
SOKA	CHRG	25.254	10.010	(4.928)	1.590	(20.409)
STON	CHRG	28.896	10.220	(7.281)	8.530	(13.211)
CLAV	CHRG	34.747	11.070	(8.089)	12.550	(13.815)

Precise regeneration of electrostatic potential requires high-order distributed multipoles

Basis set dependence / partitioning schemes

Charge-flow polarizability in CO : α_{qq}^{CO}

	6-31G**	[5s4p2d]	[8s6p3d]
Mulliken	-15.015	-143.176	-900.053
Stone	-21.371	-205.894	-1815.740
Gauss-Hermite	-3.760	-60.152	-636.540
Smoothed Gauss-Hermite	-1.125	-2.399	-9.286
Analytical	-1.355	-1.415	-1.422
Topological	-0.816	-0.864	-0.902

TPEP: Topologically Partitioned Electric Properties

- atomic domains from Bader's AIM (Atoms in Molecules) theory
- response functions calculated at TDHF, TDMP2 level
- all local and non-local contributions up to a given multipolar order
- possibility of simplification by “relocalizing” two-center contributions

Ángyán, Jansen, Loos, Hättig and Hess; *Chem. Phys. Lett.* **219** (1994) 267.

Hättig, Jansen, Hess and Ángyán; *Can. J. Chem.* **74** (1996) 976.

Jansen, Hättig, Hess and Ángyán; *Mol. Phys.* **88** (1996) 69.

Hättig, Jansen, Hess and Ángyán; *Mol. Phys.* **91** (1997) 145.

TPEP: Charge-flow polarizabilities

	6-31G**	[5s4p2d]	[10s6p3d]
CO			
α_{qq}^{CO}	-0.816	-0.864	-0.902
$\overline{\alpha}$	9.256	12.204	12.299
H₂O			
α_{qq}^{OH}	-0.358	-0.369	-0.387
α_{qq}^{HH}	-0.043	-0.041	-0.044
$\overline{\alpha}$	5.020	8.492	8.482
urea			
α_{qq}^{OC}	-0.350	-0.385	-0.392
α_{qq}^{NC}	-0.203	-0.242	-0.251
α_{qq}^{NN}	-0.106	-0.124	-0.125
$\overline{\alpha}$	23.274	32.543	32.555

Unlike basis set partitioning methods , Bader's real space partitioning leads to stable charge flow polarizabilities with respect to the basis set.

Distributed polarizability matrix of H₂O

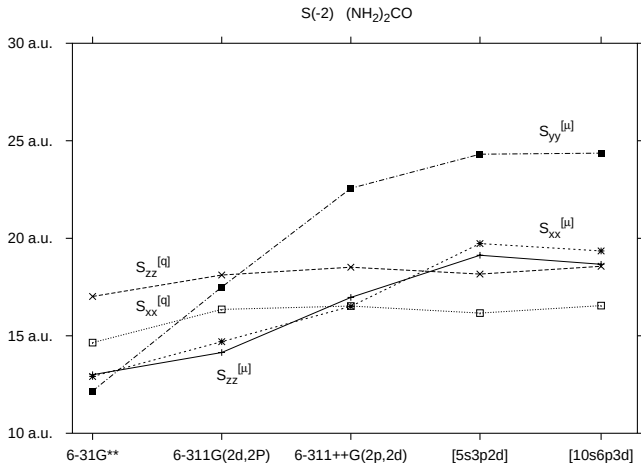
Illustration of the general structure by selected elements up to dipole-dipole. (Other elements are smaller than 0.1)

	O_q	H_q	H_q	O_{xx}	O_{yy}	O_{zz}	H_{xx}	H_{yy}	H_{zz}	H_{xx}	H_{yy}	H_{zz}
O_q	0.738	-0.369	-0.369			-0.220						
H_q	-0.369	0.418	-0.049			0.110						0.194
H_q	-0.369	-0.049	0.418						0.194			
O_{xx}				4.067								
O_{yy}					6.906							
O_{zz}						5.156						
H_{xx}							0.247					
H_{yy}								0.124				
H_{zz}									0.228			
H_{xx}										0.247		
H_{yy}											0.124	
H_{zz}												0.228

Exact reconstruction of the molecular dipole polarizability tensor:

$$\begin{pmatrix} 9.167 & 0.000 & 0.000 \\ 0.000 & 7.821 & 0.000 \\ 0.000 & 0.000 & 8.488 \end{pmatrix}$$

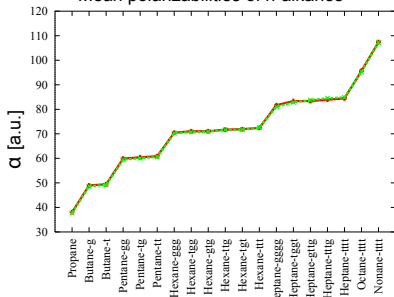
TPEP: Basis set dependence – urea



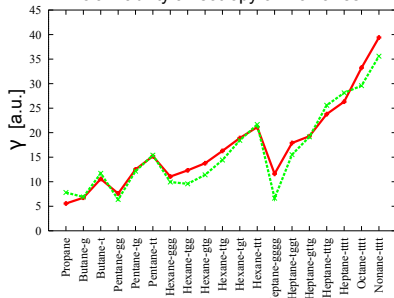
Charge-flow components depend weakly on the basis set.

TPEP: Distributed polarizability model for n-alkanes

Mean polarizabilities of n-alkanes



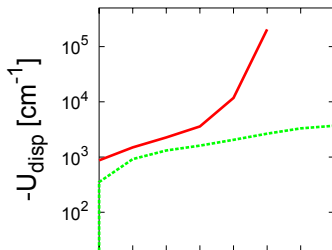
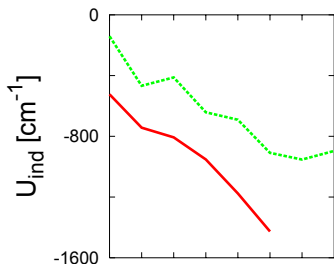
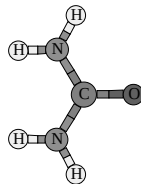
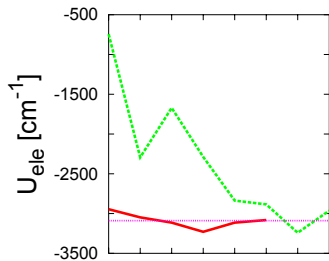
Polarizability anisotropy of n-alkanes



- Calculate distributed group polarizabilities (CH_3 , CH_2) in local frame
- Translate two-center terms to group centers
- Retain charge-flow terms
- Average group polarizabilities on a set of C_2 - C_6 n-alkanes
- Use transferable group parameters to derive polarizabilities of higher n-alkanes and/or different conformations
- **good transferability**

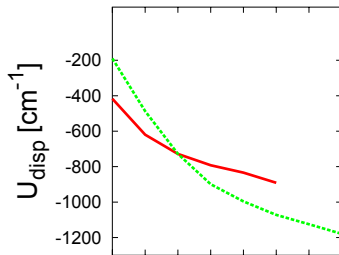
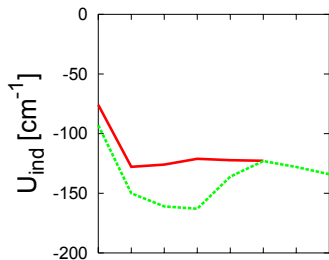
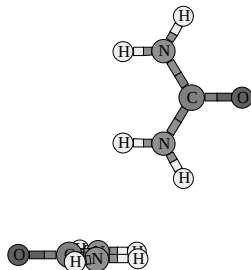
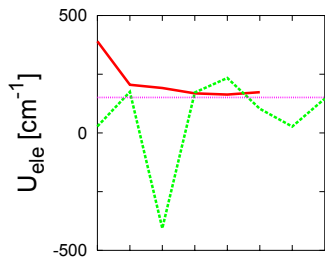
Stone, Hättig, Jansen and Ángyán; *Mol. Phys.* **89** (1996) 595.

TPEP: Convergence of energy components – urea



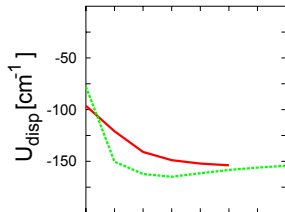
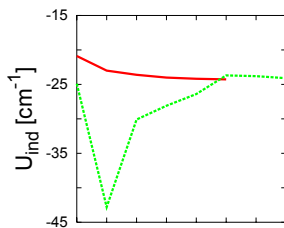
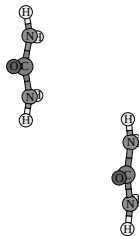
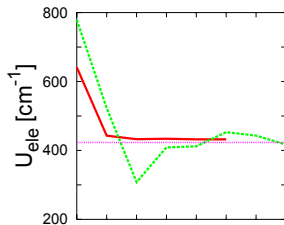
mono-centered vs. distributed multipole interaction energy components

TPEP: Convergence of energy components – urea



TPEP: Convergence of energy components – urea

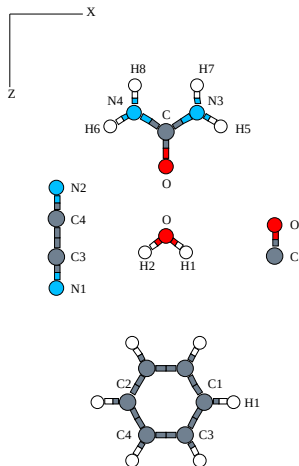
Convergence of energy components



mono-centered vs. distributed multipole interaction energy components

Atomic charge-charge polarizabilities

A – B	α_{qq}^{AB}	A – B	α_{qq}^{AB}
CO		H2O	
O – C	-0.864	O – O	0.739
NCCN		H1 – H1	0.411
N1 – N1	1.796	H1 – O	-0.369
C3 – C3	1.735	(NH ₂) ₂ CO	
C3 – N1	-1.256	C – C	0.912
C4 – C3	-0.437	O – C	-0.385
N2 – N1	-0.499	O – O	1.034
C3 – N2	-0.042	N3 – O	-0.211
C ₆ H ₆		N3 – C	-0.242
C1 – C1	2.389	N3 – N3	1.405
C2 – C1	-0.316	N4 – N3	-0.124
C3 – C1	-0.800	H5 – N3	-0.375
C4 – C1	0.103	H5 – N4	-0.022
H1 – H1	0.868	H7 – N3	-0.402
C1 – H1	-0.421	H5 – H5	0.545
C1 – H2	-0.055	H7 – H7	0.583
C1 – H3	-0.102		



Charge-flow polarizabilities are significant between bonded atoms, but certain charge-flows between nonbonded atoms can be large also (cf. benzene, dicyane).

Charge-flow polarizabilities: water dimer

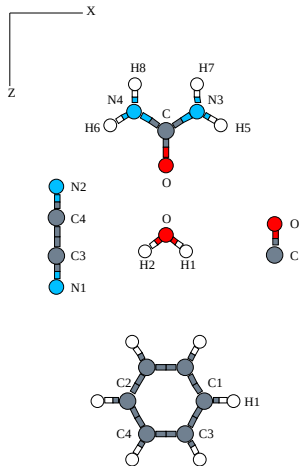


- Isolated water molecule: $\alpha_{qq}^{OH} = -0.369$
- α_{qq}^{OH} values are influenced by intermolecular forces
- Intermolecular charge-flows are weak!
($\alpha_{qq}^{O\dots H} = -0.003$, $\alpha_{qq}^{O\dots O} = -0.041$)

in het Panhuis, Popelier, Munn and Ángyán; *J. Chem. Phys.* **114** (2001) 7951.

Dipole-dipole polarizabilities

Atom	$\alpha\beta$	-2	Atom	-2
CO			(NH ₂) ₂ CO	
C	<i>xx</i>	5.816	C	<i>xx</i> -0.174
C	<i>zz</i>	4.975	C	<i>yy</i> 1.167
O	<i>xx</i>	5.251	C	<i>zz</i> 0.062
O	<i>zz</i>	5.532	O	<i>xx</i> 6.613
H ₂ O			O	<i>yy</i> 6.300
O	<i>xx</i>	5.280	O	<i>zz</i> 6.502
O	<i>yy</i>	7.240	N3	<i>xx</i> 4.679
O	<i>zz</i>	6.156	N3	<i>xz</i> -0.463
H1	<i>xx</i>	1.018	N3	<i>yy</i> 7.580
H1	<i>yy</i>	0.290	N3	<i>zz</i> 4.157
H1	<i>zz</i>	0.713	H5	<i>xx</i> 1.239
H1	<i>xz</i>	0.515	H5	<i>xz</i> 0.236
NCCN			H5	<i>yy</i> 0.390
N1	<i>xx</i>	7.423	H5	<i>zz</i> 0.696
N1	<i>zz</i>	6.140	H7	<i>xx</i> 0.726
C3	<i>xx</i>	3.133	H7	<i>xz</i> -0.201
C3	<i>zz</i>	-0.312	H7	<i>yy</i> 0.451
			H7	<i>zz</i> 1.429
C ₆ H ₆			C ₆ H ₆	
C1	<i>xx</i>	2.334	H1	<i>xx</i> 2.179
C1	<i>zz</i>	0.970	H1	<i>zz</i> 1.370
C1	<i>yy</i>	6.340	H1	<i>yy</i> 1.280



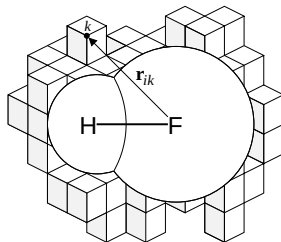
TPEP: Conclusions

- The topological partitioning scheme is successful for defining atomic and/or group polarizabilities
- Both static and frequency-dependent polarizabilities can be obtained in this way
- Compact and accurate model should involve at least explicit charge-flow polarizabilities and one-centre dipole-dipole terms
- United atom model is necessary to avoid divergences due to close contacts (e.g. H-bonds)
- Classical dipole-induced-dipole models should be handled with care in the vicinity of equilibrium intermolecular separations (screened interaction may be necessary).
- Accurate models are very complex!

OPEP: Optimally Partitioned Electrostatic Properties

- **Aim:** reduce the number of parameters without losing accuracy
- Generate high-quality database of electrostatic/induction/dispersion energies on a grid
- Define a convenient set of distributed multipole/polarizability/dispersion coefficient parameters
- Obtain the best parameters by fit to the database on the grid

OPEP: Determining distributed multipoles



Evaluation of the quantum mechanical molecular electrostatic potential on a grid of N_p points:

$$\mathcal{V}(\mathbf{r}_k) = \sum_A Z_A T(\mathbf{r}_k, \mathbf{r}_A) - \sum_{\mu} \sum_{\nu} P_{\mu\nu} \int \varphi_{\mu}^*(\mathbf{r}') \varphi_{\nu}(\mathbf{r}') T(\mathbf{r}, \mathbf{r}') d\mathbf{r}'$$

Approximation of the expansion of the potential by a truncated multipole series:

$$f(\{Q^i\}) = \sum_k^{N_p} \left[\mathcal{V}(\mathbf{r}_k) - \sum_i^{N_s} \sum_{\ell m} T_{00, \ell m}(\mathbf{r}_k, \mathbf{r}_i) Q_{\ell m}^i \right]^2 \text{ is minimal}$$

"Mechanism" of atomic charge fit

- Atomic charge distributions are usually multipolar
- Higher atomic multipoles are mimicked by fractional charges on neighboring atoms
- Effective atomic charge is the sum of an "intrinsic" charge + "multipole-mimicking" charges due to the environment
- Chemical meaning of these charges is often lost
- Strong conformation-dependence is due to "multipole-mimicking" charge contributions
- Transferability, in general, lost or deteriorated
- Analogous "mechanism" holds for polarizability fit

Chipot, Ángyán, Ferenczy and Scheraga; *J. Phys. Chem.* **97** (1993) 6628.

Ángyán and Chipot; *Int. J. Quantum Chem.* **52** (1994) 17.

OPEP: General least-squares formulation

The reference quantum mechanical electrostatic potential computed on a grid of N_p points is approximated by:

$$\tilde{v}_i = \sum_k T_{ik} Q_k \quad \forall i = 1, \dots, N_p$$

The optimal series of distributed multipoles is obtained by minimizing:

$$f(\{Q_k\}) = \sum_i^{N_p} \left(v_i - \tilde{v}_i \right)^2$$

The solution, $\mathbf{Q} \equiv (Q_1, \dots, Q_{N_c})$, that minimizes $f(\{Q_k\})$ also satisfies the so-called Normal equations of the least-squares problem:

$$\mathbf{T}^T \mathbf{T} \mathbf{Q} = \mathbf{T}^T \mathbf{V}$$

$\mathbf{T}$ is an $N_p \times N_c$ matrix containing the T_{ik} elements. $\mathbf{V}$ is an N_p vector containing the v_i elements.

OPEP: General least-squares formulation (contd.)

Provided that the rank of $\mathbf{T}$ is equal to N_c , the matrix product $\mathbf{T}^T\mathbf{T}$ can, in principle, be inverted, so that the optimal, least-squares solution is:

$$\mathbf{Q} = (\mathbf{T}^T\mathbf{T})^{-1} \mathbf{T}^T\mathbf{V}$$

Molecular moments are regenerated by translating local multipoles:

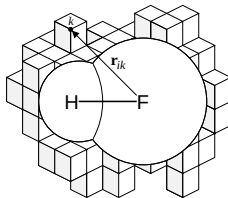
$$Q_{\ell m}(\mathbf{r}) = \sum_i \sum_{j=0}^{\ell} \sum_{k=-j}^j Q_{jk}^i W_{\ell-j, m-k}^i$$

Ángyán, J. G.; Chipot, C., A comprehensive approach to molecular charge density models: From distributed multipoles to fitted atomic charges, *Int. J. Quantum Chem.* **1994**, 52, 17–37.

OPEP: Electrostatic potential database

- Electrostatic potential from QM calculation
 - penetration effects should be minimized by choosing distant grids, cpu-intensive
- Exact multipolar potential calculated from the overlap multipole moments (in Gaussian basis set)
 - free from penetration effects
- Electrostatic potential from high-order TPEP atomic multipoles
 - slight inaccuracies

OPEP: Determining distributed polarizabilities



Induction energy, $\mathcal{U}_{\text{ind},k}$, resulting from the polarization of the molecule by a charge q_k , located at point $\mathbf{r}_k$:

$$\mathcal{U}_{\text{ind},k} = \mathcal{E}_{\text{total},k} - \mathcal{E}^0 - q_k \mathcal{V}(\mathbf{r}_k) \quad \forall k = 1, \dots, N_p$$

Approximation of the induction energy by a selection of distributed polarizabilities, introducing α :

$$\tilde{\mathcal{U}}_{\text{ind},k} = -\frac{1}{2} \sum_{i,j} \sum_{\ell=0}^{N_{\ell}(i)} \sum_{m=-\ell}^{\ell} \sum_{\ell'=0}^{N_{\ell'}(j)} \sum_{m'=-\ell'}^{\ell'} q_k T_{\ell m}^{ki} \alpha_{\ell m, \ell' m'}^{ij} T_{\ell' m'}^{jk} q_k \quad \forall k = 1, \dots, N_p$$

Require that the functional:

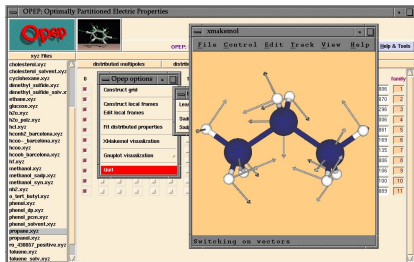
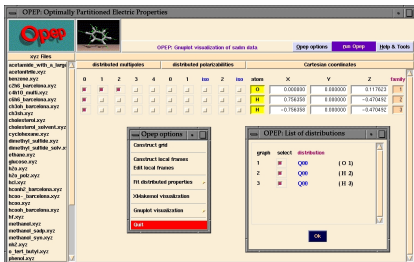
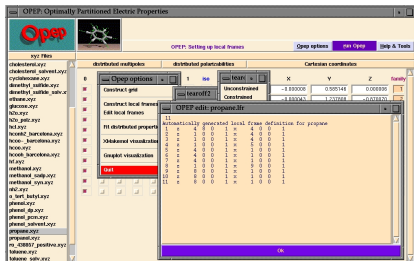
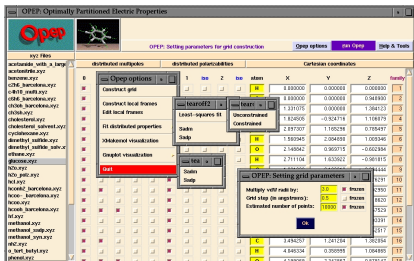
$$f(\{\alpha_{\ell m, \ell' m'}^{ij}\}) = \sum_k \left[\mathcal{U}_{\text{ind},k} - \tilde{\mathcal{U}}_{\text{ind},k} \right]^2 \quad \text{is minimal}$$

Celebi, Ángyán, Dehez, Millot and Chipot; *J. Chem. Phys.* **112** (2000) 2709.

Dehez, Soetens, Chipot, Ángyán and Millot; *J. Phys. Chem. A* **104** (2000) 1293.

Dehez, Chipot, Millot and Ángyán; *Chem. Phys. Lett.* **338** (2001) 180.

OPEP: A tool...



Ángyán, Chipot, Dehez, Hättig, Jansen and Millot *J. Comp. Chem.* **24** (2003) 997.

OPEP: Fitted polarizability models for H₂O

Polarizability	model A	model B	model C	model D	model E
$\alpha_{00,00}^{OH_1}$	-3.232	-1.673		0.708	-0.622
$\alpha_{11s,11s}^{OO}$		9.082	7.472	7.381	5.474
$\alpha_{10,10}^{OO}$		4.517	5.486	6.439	3.959
$\alpha_{11c,11c}^{OO}$		2.574	4.554	5.852	3.307
$\alpha_{11s,11s}^{H_1H_1}$			0.899	0.958	1.078
$\alpha_{10,10}^{H_1H_1}$			1.474	1.843	1.149
$\alpha_{11c,11c}^{H_1H_1}$			2.008	2.721	1.348
$\alpha_{10,11c}^{H_1H_1}$			0.689	1.149	0.281
$\alpha_{22s,22s}^{OO}$					22.677
$\alpha_{21s,21s}^{OO}$					48.276
$\alpha_{20,20}^{OO}$					30.351
$\alpha_{21c,21c}^{OO}$					24.599
$\alpha_{22c,22c}^{OO}$					58.807
$\alpha_{20,22c}^{OO}$					50.072
RMSD (mH)	0.450	0.071	0.030	0.029	0.004
Δ (%)	31.21	7.63	5.11	4.72	0.67

$$RMSD = \left[\frac{1}{N_p} \sum_k \left(\mathcal{U}_{ind,k} - \tilde{\mathcal{U}}_{ind,k} \right)^2 \right]^{1/2}$$

$$\Delta = \frac{1}{N_p} \sum_k \left| \frac{\mathcal{U}_{ind,k} - \tilde{\mathcal{U}}_{ind,k}}{\mathcal{U}_{ind,k}} \right|$$

- Model A: charge flows only;
- Model B: charge flows + dipole polarizability on the oxygen atom;
- Model C: dipole polarizabilities only;
- Model D: charge flows + dipole polarizabilities;
- Model E: same as Model D + quadrupole polarizability on the oxygen atom

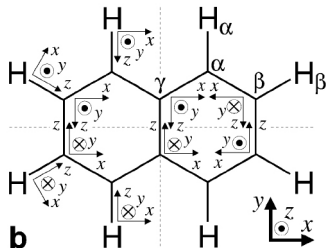
OPEP: Reconstituted multipole polarizabilities – H₂O

Polarizability	model A	model B	model C	model D	model E	TDMP2
$\alpha_{11s,11s}$	0.000	9.082	9.269	9.297	7.630	7.877
$\alpha_{10,10}$	7.984	8.649	8.433	8.376	7.793	7.836
$\alpha_{11c,11c}$	13.206	9.408	8.569	8.403	8.544	8.697
$\alpha_{11s,21s}$	0.000	3.496	0.108	−0.108	−1.212	−0.824
$\alpha_{10,20}$	2.014	3.050	−0.832	−0.849	−1.134	−0.462
$\alpha_{10,22c}$	−12.711	−6.577	−3.412	−2.904	−3.838	−4.550
$\alpha_{11c,21c}$	−20.337	−9.533	−7.841	−7.363	−8.185	−6.955
$\alpha_{22s,22s}$	0.000	0.000	11.017	11.738	35.890	25.637
$\alpha_{21s,21s}$	0.000	1.346	5.370	5.636	54.200	23.062
$\alpha_{20,20}$	0.508	1.156	11.601	12.257	41.145	23.602
$\alpha_{21c,21c}$	31.318	16.588	38.767	47.020	55.880	34.585
$\alpha_{22c,22c}$	20.234	10.471	24.608	28.927	79.226	30.031
$\alpha_{22c,20}$	−3.207	−1.659	−8.141	−8.444	42.389	−0.953

Model A: charge flows only; Model B: charge flows + dipole polarizability on the oxygen atom; Model C: dipole polarizabilities only; Model D: charge flows + dipole polarizabilities; Model E: same as Model D + quadrupole polarizability on the oxygen atom

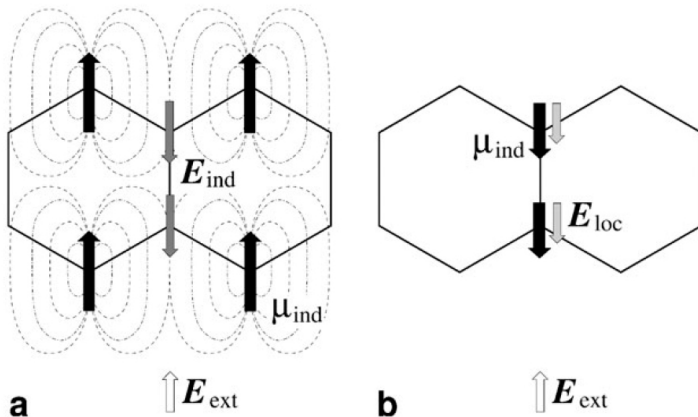
OPEP: naphthalene: atomic dipole polarizabilities

	distributed		molecular	MP2
$\alpha_{10,10}^{C\alpha}$	24.098	$\alpha_{10,10}$	74.66	70.9
$\alpha_{11c,11c}^{C\alpha}$	38.469	$\alpha_{11c,11c}$	175.65	162.0
$\alpha_{11s,11s}^{C\alpha}$	18.599	$\alpha_{11s,11s}$	121.34	119.5
$\alpha_{10,11c}^{C\alpha}$	3.169			
$\alpha_{10,10}^{C\beta}$	12.398			
$\alpha_{11c,11c}^{C\beta}$	10.722			
$\alpha_{11s,11s}^{C\beta}$	8.309			
$\alpha_{10,11c}^{C\beta}$	-1.960			
$\alpha_{10,10}^{C\gamma}$	-12.316			
$\alpha_{11c,11c}^{C\gamma}$	-10.563			
$\alpha_{11s,11s}^{C\gamma}$	-16.488			
RMSD/ 10^{-3} a.u.	0.025			
% $\Delta\epsilon$	1.948			



Negative dipole polarizabilities on C_{γ} .

Negative atomic polarizabilities?



OPEP uses a *implicitly* interacting sites within the molecule/fragment.

OPEP: Induction energy database

- Finite perturbation calculation

$$\mathcal{U}_{\text{ind},k} = \mathcal{E}_{\text{total},k} - \mathcal{E}^0 - q_k \mathcal{V}(\mathbf{r}_k) \quad \forall k = 1, \dots, N_p$$

- nonlinear effects
 - penetration effects
 - costly
- Second order perturbation expression

$$\mathcal{U}_{\text{ind},k} = -q_k^2 \sum_{i \neq 0} \frac{\langle \psi_0 | \hat{V}^k | \psi_i \rangle \langle \psi_i | \hat{V}^k | \psi_0 \rangle}{E_i^A - E_0^A}$$

- strictly linear response
 - penetration effects
 - can be evaluated at different levels of approximation: Luque et al. (UCHF)
 - could be done at TDHF or TDMP2 or higher

OPEP: Induction energy database (contd)

- From high-order TPEP model

$$\mathcal{U}_{\text{ind},k} = -\frac{1}{2} \sum_{i,j} \sum_{\ell m} \sum_{\ell' m'} q_k T_{\ell m}^{ki} \alpha_{\ell m, \ell' m'}^{ij} T_{\ell' m'}^{jk} q_k$$

- strictly linear
- no penetration
- quality depends on the length of development; cheap

OPEP: Conclusions

- The electrostatic potential constitutes a fingerprint of the molecule, and, thus, probably represents an appropriate quantity from which atomic charges and multipoles can be derived.
- Point charges obtained from a **global** fitting procedure are inherently not transferable, and strongly conformational dependent.
- Chemically counter-intuitive potential derived charges, together with their poor transferability, can be ascribed to the incompleteness of the model to describe subtle and complex contributions of local multipole moments of higher order.
- Conformational dependence can be alleviated by adding higher order multipoles (typically dipoles) in the fitting procedure.
- Unlike in TPEP, in OPEP fully flexible models to be constructed, from simple charge flows to quadrupole contributions and beyond.

OPEP: Conclusions

- The pathological appearance of positive $\alpha_{00,00}^{ij}$ or $\alpha_{1m,1m'}^{ii}$ reflects redundancies in the model, thereby suggesting that parameters to be fitted should be selected carefully to warrant physical realism.
- Reproduction of molecular polarizabilities within “chemical accuracy” from the experimentally measured quantities requires the usage of large basis sets. Standard, Pople-type basis sets, e.g. 6-311++G(2d,2p) are clearly not enough and should be replaced by more complex ones like (13s8p3d/7s2p)/<8s5p3d/4s2p>.
- Mathematically “correct” models show a conspicuous tendency to exaggerate atomic anisotropy.
- Computation of induction energies, including correlation effects, on a grid of points is very CPU-intensive.
- Alternative ways to generate induction energy grids (TPEP, second order perturbational expression) should be further explored to ensure computational feasibility.
- Generalization to frequency dependent induction energies may allow to access dispersion coefficients.

Part III

DFT and intermolecular forces

DFT and intermolecular forces

- Thomas-Fermi-Dirac model and Gordon-Kim method
- Kohn-Sham method
- When DFT works well for nonbonded interactions
- When DFT fails:
 - charge transfer complexes
 - dispersion forces
- Remedies
 - tentative to "improve" functionals
 - empirical corrections
 - range-separated hybrid with explicit long-range correlation
- Recent applications

Thomas-Fermi-Dirac model

- Sum of kinetic, nuclear attraction, Hartree, exchange and correlation energies

$$E_{\text{TFD}}[\rho] = T_{\text{TF}}[\rho] + E_{ne}[\rho] + E_H[\rho] + E_x[\rho] + E_c[\rho]$$

- These terms are either trivial or parameterized from the uniform electron gas

- Potential energy classically

$$E_{ne}[\rho] = \int d\mathbf{r} v_{ne}(\mathbf{r}) \rho(\mathbf{r})$$

- Kinetic energy (in atomic units)

$$T_{\text{TF}}[\rho] = \frac{3(3\pi^2)^{2/3}}{10} \int d\mathbf{r} \rho^{5/3}(\mathbf{r})$$

- Hartree energy

$$E_H[\rho] = \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

- Slater exchange energy

$$E_x[\rho] = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \int d\mathbf{r} \rho^{4/3}(\mathbf{r})$$

- Minimize the energy with constraint $\int d\mathbf{r} \rho(\mathbf{r}) = N$ using the Lagrange multiplier, μ

$$\frac{\delta}{\delta \rho(\mathbf{r})} \left(E_{\text{TFD}}[\rho] - \mu \int \rho(\mathbf{r}) \right) = 0$$

TFD equation

- Differential equation for the electron density (μ is called chemical potential)

$$\mu = \frac{1}{2}(3\pi^2)^{\frac{2}{3}}\rho^{\frac{2}{3}}(\mathbf{r}) - \frac{4}{3}\rho^{\frac{1}{3}}(\mathbf{r}) + v_{ne}[\rho](\mathbf{r}) + v_H[\rho](\mathbf{r})$$

- The density is singular at $\rightarrow 0$, because the Hartree potential is singular
- Wrong decay of the density: $1/r^{\frac{3}{2}}$
- No shell structure in atoms
- No Friedel oscillations in solids
- Teller has shown that the TFD model is unable to describe chemical bonding

Gordon-Kim method

The density can be kept constant in weak intermolecular complexes

$$\rho_{AB} = \rho_A + \rho_B$$

using e.g. Hartree-Fock densities of the monomers.

Calculate interaction energy from the TFD energy functional (no optimization).

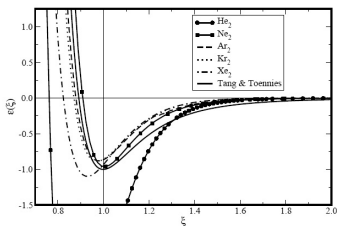
$$\begin{aligned}\Delta E = & C_F \int \{ \rho_{AB}^{5/3}(\mathbf{r}) - \rho_A^{5/3}(\mathbf{r}) - \rho_B^{5/3}(\mathbf{r}) \} d\mathbf{r} \\ & + \left\{ - \int \left(\frac{Z_A}{r_A} + \frac{Z_B}{r_B} \right) \rho_{AB}(\mathbf{r}) d\mathbf{r} + \int \frac{Z_A}{r_A} \rho_A(\mathbf{r}) d\mathbf{r} + \int \frac{Z_B}{r_B} \rho_B(\mathbf{r}) d\mathbf{r} \right\} \\ & + \left\{ \frac{Z_A Z_B}{R} + E_H[\rho_{AB}] - E_H[\rho_A] - E_H[\rho_B] \right\} \\ & - C_x \int \{ \rho_{AB}^{4/3}(\mathbf{r}) - \rho_A^{4/3}(\mathbf{r}) - \rho_B^{4/3}(\mathbf{r}) \} d\mathbf{r} + E_c[\rho_{AB}] - E_c[\rho_A] - E_c[\rho_B]\end{aligned}$$

Gordon-Kim: results

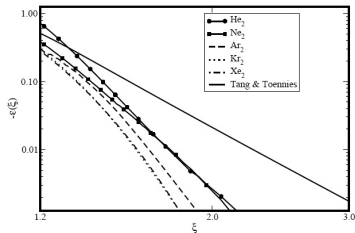
Overall agreement with experiment is quite satisfactory (table from Lee-Parr-Yang book):

System	r_0 (calc.)	r_0 (exp.)	ΔE	ΔE (exp.)
He-He	2.49	2.96	62.5	16.5
Ne-Ne	2.99	3.09	56	63
Ar-Ar	3.62	3.70	175	195
Kr-Kr	3.89	3.87-4.08	248	270-340
Xe-Xe	4.15	4.5	417	342-360

Our calculations with HF/aug-cc-pVQZ monomers

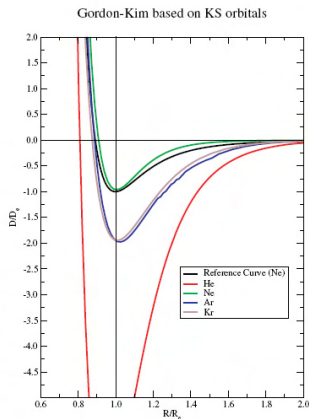
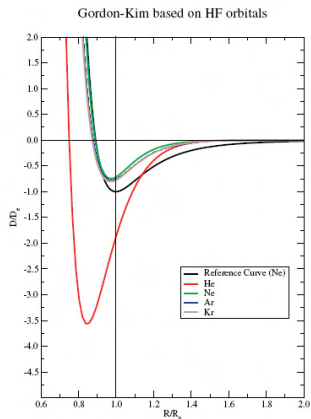


nice vdW minima (excepted He₂)



wrong asymptotic behaviour

Gordon-Kim: sensitive to the monomer densities



Kohn-Sham method

- TFD did not work because of the kinetic energy functional; Kohn and Sham suggested:

$$T_{TF}[\rho] \quad \Rightarrow \quad \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} | \Psi \rangle$$

- Search for an independent particle system, which reproduces the *exact density*:

$$E = \min_{\rho \rightarrow N} \left\{ \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} | \Psi \rangle + E_{Hxc}[\rho] + \int d\mathbf{r} v_{ne}(\mathbf{r}) \rho(\mathbf{r}) \right\}$$

- leads to an effective Schrödinger equation of a system of non-interacting electrons

$$(\hat{T} + \hat{V}_s) |\Psi\rangle = \mathcal{E} |\Psi\rangle$$

- where $v_s(\mathbf{r})$, Kohn-Sham potential ensures that ρ is the exact density

$$v_s(\mathbf{r}) = \frac{\delta E_{Hxc}[\rho]}{\delta \rho(\mathbf{r})} + v_{ne}(\mathbf{r})$$

- unfortunately, the exact $E_{Hxc}[\rho]$ is not known...

Kohn-Sham method: functionals

How to approximate the exchange-correlation functional?

$$E_{xc}[\rho] = \int d\mathbf{r} \epsilon_{xc}(\mathbf{r}) \rho(\mathbf{r})$$

- LDA: use electron gas results locally, $\epsilon_{xc}[\rho(\mathbf{r})](\mathbf{r})$
- GGA: take into account the rate of variation of ρ : $\epsilon_{xc}[\rho(\mathbf{r}), |\nabla\rho(\mathbf{r})|](\mathbf{r})$
- meta GGA: depends on higher derivatives of the density,...
- hybrid functionals: mixing a certain amount of Hartree-Fock exchange
- functionals are either derived ab initio from the electron gas and by imposing exact conditions (Perdew)...
- ...or obtained by more-or-less extensive fitting to experimental or other data (Becke)...

Example of water dimer: LDA and GGA

Table 12-4. Computed interaction energy (ΔE) and counterpoise-corrected interaction energies (ΔE_{CP}) of the water dimer [kcal/mol]. The BSSE is given in parentheses.

Basis Set	SVWN			BLYP			SLYP			BVWN		
	ΔE	ΔE_{CP}		ΔE	ΔE_{CP}		ΔE	ΔE_{CP}		ΔE	ΔE_{CP}	
6-31++G(d,p)	10.8	9.8	(1.0)	5.6	4.8	(0.8)	13.2	12.3	(0.9)	4.1	3.4	(0.7)
6-311++G(d,p)	10.5	9.5	(1.0)	5.4	4.6	(0.8)	12.9	11.9	(1.0)	4.0	3.3	(0.7)
6-311++G(3df,2p)	9.3	9.1	(0.2)	4.5	4.2	(0.3)	11.7	11.5	(0.2)	3.1	2.9	(0.2)
aug-cc-pVDZ	— ^a			4.3	4.1	(0.2)	11.6	11.6	(0.0)	2.9	2.8	(0.1)
aug-cc-pVTZ	9.0	9.1	(-0.1)	4.2	4.2	(0.0)	11.4	11.6	(-0.2)	2.9	2.8	(0.1)
aug-cc-pVQZ	9.0	9.1	(-0.1)	4.2	4.2	(0.0)	11.3	11.6	(-0.2)	2.9	2.9	(0.0)

^a No C_s -*trans* minimum structure.

Example of water dimer: hybrid functionals

Table 12-5. Computed interaction energy (ΔE) and counterpoise-corrected interaction energies (ΔE_{CP}) of the water dimer [kcal/mol]. The BSSE is given in parentheses.

Basis Set	B3LYP			B3PW91			<i>m</i> PW1PW91		
	ΔE	ΔE_{CP}		ΔE	ΔE_{CP}		ΔE	ΔE_{CP}	
6-31++G(d,p)	6.0	4.6	(1.4)	5.5	4.7	(0.8)	6.2	5.3	(0.9)
6-311++G(d,p)	5.8	5.1	(0.7)	5.3	4.5	(0.8)	5.9	5.1	(0.8)
6-311++G(3df,2p)	4.8	4.6	(0.2)	4.3	4.0	(0.3)	4.9	4.6	(0.3)
cc-pVDZ	– ^a			7.3	3.9	(3.4)	7.9	4.5	(3.4)
cc-pVTZ	6.1	4.5	(1.6)	5.3	4.0	(1.3)	5.9	4.6	(1.3)
cc-pVQZ	5.3	4.6	(0.7)	4.6	4.0	(0.6)	5.2	4.6	(0.6)
cc-pV5Z	4.8	4.6	(0.2)	4.2	4.0	(0.2)	4.8	4.6	(0.2)
aug-cc-pVDZ	4.7	4.5	(0.2)	4.2	4.0	(0.2)	4.8	4.6	(0.2)
aug-cc-pVTZ	4.6	4.6	(0.0)	4.0	4.0	(0.0)	4.6	4.6	(0.0)
aug-cc-pVQZ	4.6	4.6	(0.0)	4.0	4.0	(0.0)	4.6	4.6	(0.0)

^a No C_s -*trans* minimum structure.

Example of water dimer

Table 12-7. Selected computed properties for the water dimer (taken from Tuma, Boese, and Handy, 1999).

Property	B3LYP	B97-1	PBE1PBE	BLYP	PBE	HTCH	HTCH38
$\Delta R_{O-O}[\text{\AA}]^a$	-0.047	-0.044	-0.077	-0.009	-0.072	0.109	-0.003
$\Delta E_{CP}[\text{kcal/mol}]$	4.8	5.2	5.2	4.5	5.4	2.9	4.6
$\Delta \nu_{OH}[\text{cm}^{-1}]$	-175	-175	-199	-187	-227	-143	-181

^a Deviation from the experimental R_e of 2.952 Å.

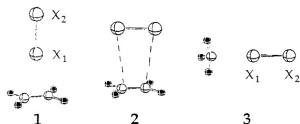
Best available estimate of the harmonic frequency shift is -121 cm⁻¹ MP4/VTZ(2df,2p) by Bleiber and Sauer, 1995.

Some general trends

- quite reasonable for minima of H-bonded complexes
- cation-molecule interactions (dominated by induction) also good
- interactions in strongly bound species are generally overestimated, weaker hydrogen bonds are often underestimated
- not necessarily uniform quality for the whole PES

Charge transfer complexes

DFT calculations on $\text{C}_2\text{H}_4 \cdots \text{X}_2$, $\text{NH}_3 \cdots \text{X}_2$, $\text{X} = \text{F}, \text{Cl}, \text{Br}, \text{and I}$).



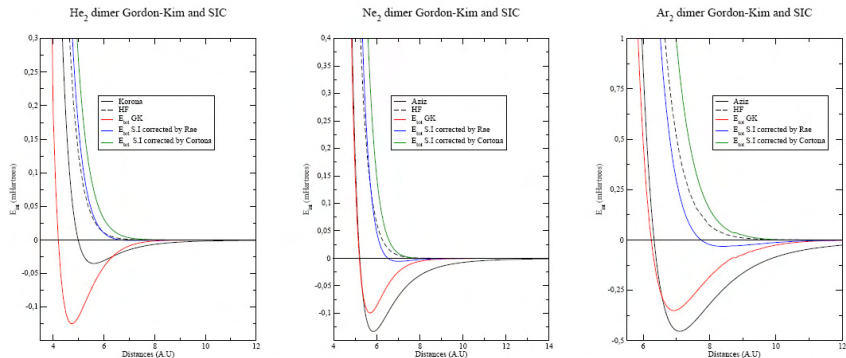
complex	LDA	BP	PW	MP2
$\text{NH}_3 \cdots \text{Cl}_2$	-20.0	-12.7	-13.2	-5.8
$\text{NH}_3 \cdots \text{F}_2$	-35.0	-25.8	-25.0	-1.3
$\text{NH}_3 \cdots \text{Br}_2$	-20.3	-12.9	-13.6	-6.4
$\text{NH}_3 \cdots \text{I}_2$	-17.9	-11.0	-10.5	-6.9
$\text{C}_2\text{H}_4 \cdots \text{Cl}_2$	-12.6	-5.2	-6.8	-1.6
$\text{C}_2\text{H}_4 \cdots \text{F}_2$	-38.4	-25.7	-25.3	-0.4
$\text{C}_2\text{H}_4 \cdots \text{Br}_2$	-14.0	-6.0	-7.8	-2.5
$\text{C}_2\text{H}_4 \cdots \text{I}_2$	-11.2	-4.3	-5.7	-2.9

Strong overestimation of intermolecular interactions

Functional failure: self-interaction error

- Spurious Hartree and exchange self interactions should cancel each other
- True for exact exchange functional, does not hold for approximate one
- Self-interaction is strong for localized electrons, weak for delocalized ones
- Qualitatively wrong electronic structure for system with both type of electrons (e.g. transition metals)
- Erroneous long-range behavior of the Kohn-Sham potential
- Insufficient binding of distant electrons
- Charge transfer is strongly overestimated

Self-interaction corrected Gordon-Kim



Self-interaction corrected potential curves are repulsive, like the Hartree-Fock curve (not surprising: the bonding here is fully due to correlation effects).

Why seems Gordon-Kim work so well?

Overlap dependence of electron gas kinetic and exchange energies

Hartree-Fock interaction potential is always repulsive

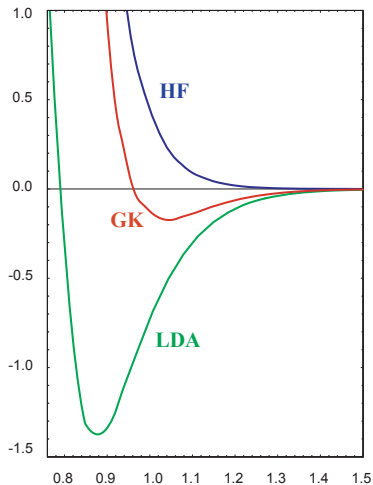
$$\Delta E_{HF} \approx (C_T^{HF} - C_X^{HF})S^2$$

Electron gas (Slater-Dirac) exchange is overestimated and deteriorates balance with kinetic energy contribution

$$\Delta E_{LDA} \approx C_T^{HF} S^2 - C_X^{LDA} S^{4/3}$$

Gordon-Kim method uses Thomas-Fermi kinetic energy functional and thus “matches the mistake with another”

$$\Delta E_{GK} \approx C_T^{GK} S^{5/3} - C_X^{GK} S^{4/3}$$



Can London dispersion be taken into account by DFT?

Controversial situation:

- LDA: strong overbinding
- Becke GGA: no binding at all
- Perdew GGA: sometimes quite reasonable binding
- Considerable effort spent to reparametrize functionals
- None of the functionals is able to describe $1/R^6$ behaviour

Prototype systems: rare gas dimers

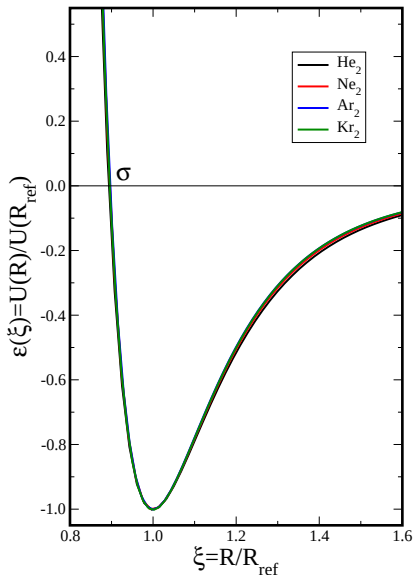
- Bonding by long-range dynamic correlation
- Experimental data

Syst.	R_{ref} [a_0]	U_{ref} [μH]
He-He	5.62	34.8
Ne-Ne	5.84	134
Ar-Ar	7.10	454
Kr-Kr	7.58	638

- Tang-Toennies potential

$$U(R) = Ae^{-bR} - \sum_{n=3}^5 f_{2n}(bR) \frac{C_{2n}}{R^{2n}}$$

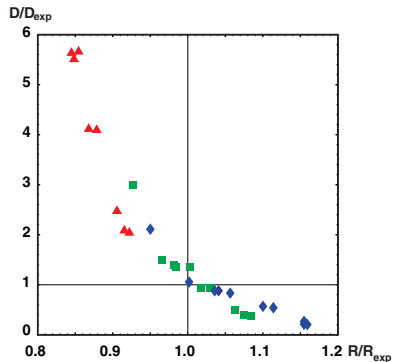
- Use *reduced potential curves*
- Allows direct comparison of different systems



Rare gas dimers

U_m^*	MAE	Std. Dev.
LDA	3.85	2.40
PBE	0.52	0.50
TPSS	0.47	0.34
CCSD(T)	0.17	0.02

r_m^*	MAE	Std. Dev.
LDA	0.14	0.04
PBE	0.04	0.03
TPSS	0.09	0.06
CCSD(T)	0.02	0.01



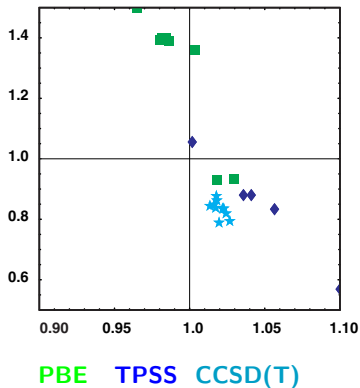
LDA PBE TPSS

Random behaviour of various functionals

Rare gas dimers

U_m^*	MAE	Std. Dev.
LDA	3.85	2.40
PBE	0.52	0.50
TPSS	0.47	0.34
CCSD(T)	0.17	0.02

r_m^*	MAE	Std. Dev.
LDA	0.14	0.04
PBE	0.04	0.03
TPSS	0.09	0.06
CCSD(T)	0.02	0.01



Wave function methods are much more reliable

Rare gas dimers: overview of functionals

I. C. Gerber and J. G. Ángyán

J. Chem. Phys. **126**, 044103 (2007)

TABLE II. Mean of absolute percentage errors (MA%E), mean of percentage of error (M%E), and standard deviation of the mean error for the ten rare gas dimers calculated by the CCSDT WFT method, as well as a series of density functionals. The methods are in the order of increasing unsigned mean error.

Method	MA %E(ϵ_m)	M %E(ϵ_m)	SDEV(ϵ_m)	MA %E(ξ_m)	M %E(ξ_m)	SDEV(ξ_m)	Ref.
CCSD(T)	16.88	-16.88	2.29	1.76	1.76	0.45	73 and 74
WL	20.34	4.61	24.13	5.89	-5.89	2.82	37 and 38
M05-2X	31.15	15.01	37.57	4.37	-3.13	4.10	72
B98	46.10	-9.45	59.68	4.20	2.30	4.64	72
PBE0	50.87	-27.24	51.60	5.17	3.58	4.94	72
TPSS	54.32	-24.61	59.96	8.50	8.35	6.59	72
MPWB1K	55.01	-22.79	64.08	5.50	4.50	5.48	72
TPSSH	55.90	-37.44	50.24	9.02	8.93	6.28	72
mPW91PW91	57.15	-39.60	50.57	15.74	15.74	7.80	72
MPW1B95	59.99	-11.37	77.77	5.45	4.50	5.41	72
MPW1K	61.17	-57.39	35.20	15.70	15.70	7.66	72
B97-1	62.17	35.35	85.53	3.98	-1.14	4.86	72
PBE	69.94	35.68	94.48	4.06	1.19	5.46	70
PBE	66.76	28.93	93.27	4.31	1.38	5.14	72
PBEPW91	68.47	33.46	95.53	4.19	0.84	5.10	72
PBE1W	70.98	33.50	99.17	4.81	1.98	5.70	72
M05	87.64	63.52	94.44	6.06	-4.50	5.42	72
PW6B95	109.45	79.14	127.98	4.39	3.26	4.48	72
PWB6K	157.46	147.70	189.06	4.87	1.16	5.89	72
PW91	285.29	281.44	311.40	5.08	-1.12	6.21	70
PW91	271.46	264.85	301.09	4.95	-0.93	6.11	72
LDA	387.03	387.03	261.48	13.27	-13.27	4.12	70
LDA	411.71	411.71	267.34	13.56	-13.56	4.04	72

Dispersion effect on the conformations of Tyr-Gly

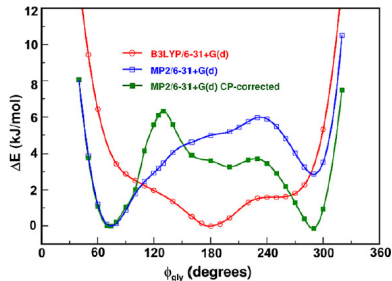
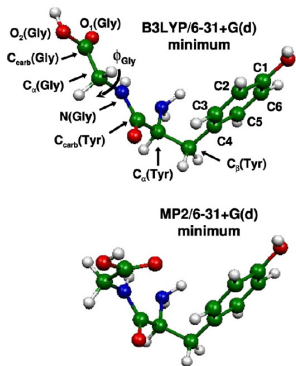


Fig. 3. B3LYP/6-31+G(d), MP2/6-31+G(d), and CP-corrected MP2/6-31+G(d) potential energy profiles for rotation around the glycine C_α-N bond.

Holroyd and van Mourik, *Insufficient description of dispersion in B3LYP and large basis set superposition errors in MP2 calculations can hide peptide conformers* Chem. Phys. Lett., 2007, 442, 42

Stabilization of branched alkanes

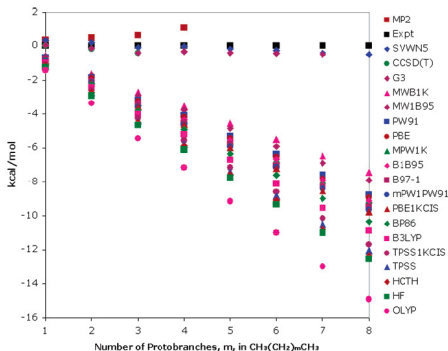


Figure 2. Deviations of various DFT functionals from experimental (0 K) protobranching stabilization energies. Negative values denote underestimation. Stabilization energies are based on eq 1. CCSD(T) and MP2 refer to CCSD(T)/aug-cc-pVTZ//MP2/6-311+G(d,p) and MP2/aug-cc-pVTZ//MP2/6-311+G(d,p), respectively, and include MP2/6-311+G(d,p) zero-point corrections. All other computations employed the 6-311+G(d,p) basis set.

Failure of functionals attributed to dispersion-like effects

Wodrich, Corminboeuf, Schleyer *Systematic Errors in Computed Alkane Energies Using B3LYP and Other Popular DFT Functionals*; Org. Lett., 2006, 8, 3631

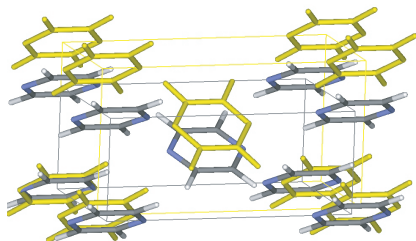
Layered structures

Layered minerals

V_2O_5	PW91	exp	Δ
a	11.55	11.51	2 %
b	3.85	3.86	1 %
c	4.85	4.37	11 %

MoS_2	PW91	exp	Δ
a	3.22	3.160	2 %
thick	3.16	3.172	1 %
dist.	4.57	2.975	35 %

Stacking in molecular crystals

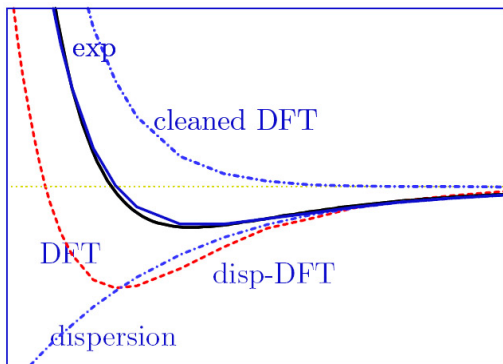


Pyrazine DFT vs. experimental structure

Strategy to treat van der Waals systems

Two-step procedure:

- remove wrong DFT minimum
- add asymptotically correct dispersion

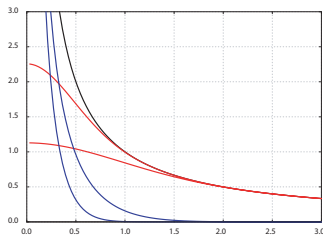


Range-separated e-e interaction

Separation of the e-e interaction

$$w_{ee}(r) = w_{ee}^{lr,\mu}(r) + w_{ee}^{sr,\mu}(r)$$

$$\frac{1}{r} = \frac{\text{erf}(\mu r)}{r} + \frac{\text{erfc}(\mu r)}{r}$$



$w_{ee}^{lr,\mu}(r)$ is not singular at $r = 0$

Generalization of Kohn-Sham theory (A. Savin)

- Treat kinetic energy and long-range e-e interaction together

$$E = \min_{\rho \rightarrow N} \left\{ \min_{\Psi_\mu \rightarrow \rho} \langle \Psi_\mu | \hat{T} + \hat{V}_{ee}^{lr,\mu} | \Psi_\mu \rangle + E_{Hxc}^{sr,\mu}[\rho] + \int d\mathbf{r} v_{ne}(\mathbf{r}) \rho(\mathbf{r}) \right\}$$

- Long-range "HK-functional"

$$F^{lr,\mu}[\rho] = \min_{\Psi_\mu \rightarrow \rho} \langle \Psi_\mu | \hat{T} + \hat{V}_{ee}^{lr,\mu} | \Psi_\mu \rangle$$

- Short-range Hartree-exchange-correlation energy functional

$$E_{Hxc}^{sr,\mu}[\rho] = F[\rho] - F^{lr,\mu}[\rho]$$

- Ψ_μ from the **long-range interacting** effective Sch. equation

$$(\hat{T} + \hat{V}_{ee}^{lr,\mu} + \hat{V}_{Hxc}^{sr,\mu}[\rho_{\Psi_\mu}] + \hat{V}_{ne}) |\Psi_\mu\rangle = \mathcal{E}_\mu |\Psi_\mu\rangle$$

RSH+MP2 approach

- Long-range dynamic correlation of electron pairs
- Double counting of correlation is avoided
- Correct $1/R^6$ asymptotic behaviour
- Simple energy correction at 2nd order

$$E^{\text{RSH+MP2}} = E^{\text{RSH}} + \sum_{i < j}^{\text{occ.}} \sum_{a < b}^{\text{virt.}} \frac{|K_{iajb}^{\text{lr}} - K_{ibja}^{\text{lr}}|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$$

- Higher-order approaches are possible (e.g. CCSD(T))

Rare gas dimers

- 10 homo- and hetero-dimers XY (X,Y=He, Ne, Ar et Kr)
- aug-cc-pVQZ basis

Percentage deviations of binding energies

	LDA	PBE	M05-2X	WL	CCSD(T)	MP2	RSH+MP2
MA%E	387	67	31	20	17	25	16
M%E	387	29	15	5	-17	-25	-16
SDEV	261	93	38	24	2	18	14

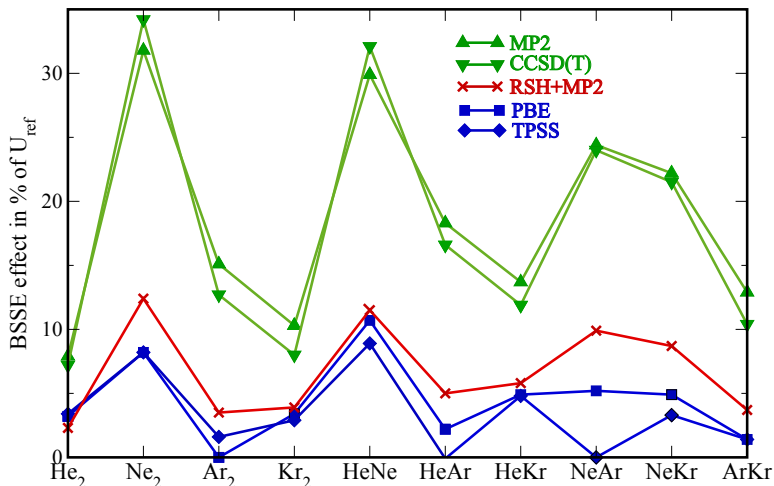
- Good MAE and ME for Wilson-Levy and M05-2X, but high SDEV
- RSH+MP2 better than standard GGA and MP2
- RSH+MP2 is of similar quality as CCSD(T)

Rare gas dimers: BSSE effects

	U_m^*		r_m^*	
	RHF+MP2	RSH+MP2	RHF+MP2	RSH+MP2
He ₂				
avtz	-0.125	-0.063	-0.007	-0.000
av5z	-0.048	-0.015	-0.004	-0.001
Ne ₂				
avtz	-0.547	-0.335	-0.049	-0.031
av5z	-0.150	-0.025	-0.011	-0.001
Ar ₂				
avtz	-0.263	-0.101	-0.020	-0.007
av5z	-0.138	-0.022	-0.005	-0.002
Kr ₂				
avtz	-0.191	-0.126	-0.013	-0.007
av5z	-0.073	-0.039	-0.003	-0.002

BSSE is considerably reduced in RSH+MP2 wrt usual MP2

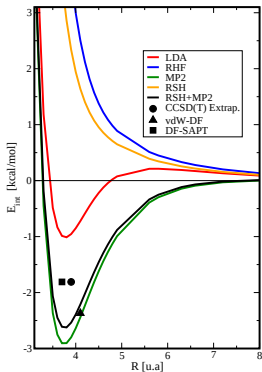
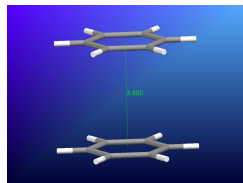
Rare gas dimers: BSSE effects



BSSE much smaller than with MP2 or CCSD(T)

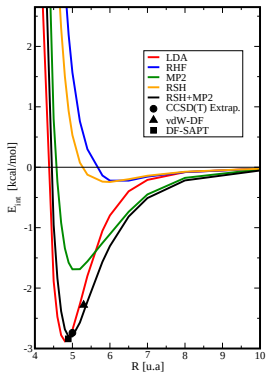
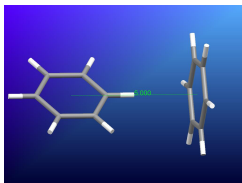
Benzene dimer

Stacked parallel



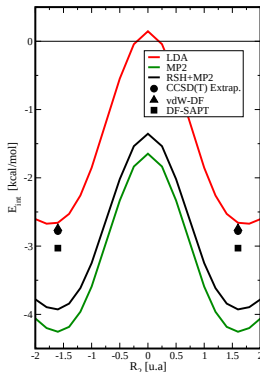
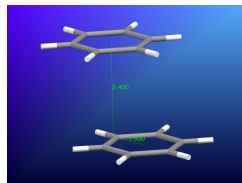
J. Ángyán (EMQC)

T-shaped



IMF

Parallel displaced



Benzene dimer

conf.	exp.	QMC	DFT-SAPT	MP2(∞)	RSH+MP2 (DZ)
T	-2.5 to -3.0	-3.0 ± 0.4	-2.8	-3.4	-2.8
PD	-2.5 to -3.0	-3.6 ± 0.4	-3.0	-4.5	-3.8
S			-1.8	-3.2	-2.6

- Good ordering of the configurations
- Stacked conformation is too strongly bound (MP2 error)
- Global improvement w.r.t. standard MP2 results

Rare gas solids: Ne, Ar, Kr

	Ne			Ar			Kr		
	E_{coh}	a	B	E_{coh}	a	B	E_{coh}	a	B
LDA	86	3.86	7.2	135	4.95	6.2	159	5.35	5.2
PBE	19	4.61	1.0	21	5.98	0.6	22	6.43	0.5
RSH+MP2	41	4.42	1.4	109	5.19	2.3	167	5.58	3.0
Exp.	27	4.35	1.1	89	5.23	2.4	123	5.61	3.6

- 1% error of the lattice constant, a (Å)
- Good agreement for the bulk modulus, B (GPa)
- Overestimated cohesion energy, E_{coh} (eV)
- Laborious convergence with essentially plane wave virtuals

Conclusions on RSH+MP2

- Good answer for the good reasons
- Long-range exchange correction removes artificial minima
- Correct $1/R^6$ asymptotic behaviour
- Reasonable vdW minima
- Favorable basis set dependence and small BSSE
- Overestimation of interaction between strongly delocalized systems