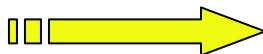# Basics of Quantum Chemistry
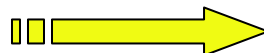
Quantum mechanical (QM) calculations:
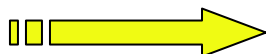
when electrons
are involved

→ spectroscopic properties
transition structures
excited states

when a high degree
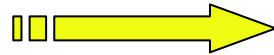of accuracy is required

→ energies and structures
of reactive sites

to avoid extensive
parameterizations
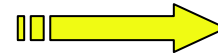
→ non standard systems
(*e.g.* radicals)

Which are the problems ?

Electron
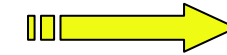correlation → Post-HF methods
Density functional theory

electronic → Linear scaling
methods

Many degrees
of freedom

nuclear
(solute) → Hybrid methods
(QM/QM, QM/MM)
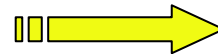
nuclear
(solvent) → Continuum, cluster,
molecular dynamics

Ab Initio Theory

Means "from first principles;" this implies that no (few) assumptions are made, and that the method is "pure" from a theoretical standpoint (in practice, may not be really true)

Even *ab initio* is based on <u>approximate</u> solutions to the Schrödinger equation:

$$\hat{H} \Psi = E \Psi$$

Exact solutions only for extremely simple systems!

Schrödinger equation: $\hat{H}\Psi = E\Psi$

$\hat{H}$ = Hamiltonion operator

$$\hat{H} = \boxed{-\frac{h^2}{8\pi^2}\sum_{A}^{nuclei}\frac{1}{M_A}\nabla_A^2} \quad \boxed{-\frac{h^2}{8\pi^2 m}\sum_{a}^{electrons}\nabla_a^2} \quad -e^2\sum_{A}^{nuclei}\sum_{a}^{electrons}\frac{Z_A}{r_{Aa}}$$

kinetic energy (nuc.) kinetic energy (elect.)

$$+\, e^2\sum_{A\,>\,B}^{nuclei}\frac{Z_A Z_B}{r_{AB}} \quad +\, e^2\sum_{a\,>\,b}^{electrons}\frac{1}{r_{ab}}$$

3 Coulombic energy terms;
one attractive, 2 repulsive

Generally one of the types:

Hartree-Fock
    considers each electron to experience a combined
    effect of all of other electrons

Correlated methods (several)
    considers individual electron interactions
    (to some extent)

Density Functional Theory (DFT)

CPU time scales as n^3 to n^4 where n = number of basis functions

Hartree-Fock Method

Main approximations introduced to allow "solution" of many-electrons Schrödinger equation

Born-Oppenheimer approximation: electrons act independently of nuclei

Hartree-Fock approximation: electrons experience the average "field" generated by all other electrons as a group, not individually

LCAO: molecular orbitals are constructed as linear combinations of atom-centered orbitals

Schrödinger equation after Born-Oppenheimer approximation

$$\hat{H} = \boxed{- \frac{h^2}{8\pi^2} \sum_{A}^{\text{nuclei}} \frac{1}{M_A} \nabla_A^2} \quad \boxed{- \frac{h^2}{8\pi^2 m} \sum_{a}^{\text{electrons}} \nabla_a^2} \quad - e^2 \sum_{A}^{\text{nuclei}} \sum_{a}^{\text{electrons}} \frac{Z_A}{r_{Aa}}$$

0

kinetic energy (nuc.)   kinetic energy (elect.)

$$+ e^2 \sum_{A > B}^{\text{nuclei}} \frac{Z_A Z_B}{r_{AB}} \quad + e^2 \sum_{a > b}^{\text{electrons}} \frac{1}{r_{ab}}$$

constant

2 Coulombic energy terms;
one attractive, 2 repulsive,
plus a constant for nuclei

What is density ?

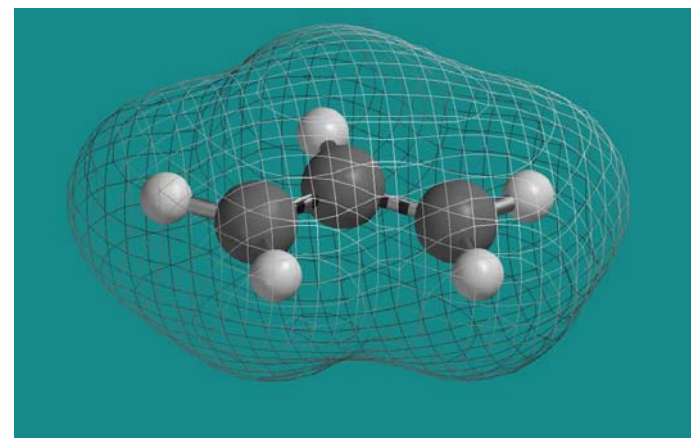- Density provides information about how something is distributed in space

- For a chemical system, electron density tells us where the electrons are likely to exist

Allyl Cation:

$$\rho \propto \Psi^* \Psi$$

Density Functional Theory (DFT)

within the KS formalism

$$E[\rho] = T_{kin} + E_{N-e}[\rho] + E_J[\rho] + \boxed{E_{XC}[\rho]} \quad \text{approximate}$$

$$E_{xc}[\rho] = \int F_{xc}(\rho, \nabla\rho)dr \quad F_{xc} = \text{exchange correlation functional}$$

T_{kin} = kinetic energy term

E_J = electron-electron repulsion term

E_{N-e} = potential energy

$E_{XC} = E_X$ (quantum mechanical exchange energy) $+ E_C$ (dynamic correlation energy)

Hybrid $E_X = aE_X^{HF} + (1-a)E_X^{DFT}$

Behavior of HF and DFT

- By definition, HF has no electron correlation

    As we will see shortly, this implies more serious errors
    for stretched/distorted bonds, i.e. disfavors overcoordination

- Pure DFT overestimates correlation

    Preference for overcoordination

- Hence success of hybrid functionals which add HF exchange
    to DFT, e.g. B3LYP…

- Hartree-Fock alone is *not* very useful – barriers are usually
    overestimated by more than DFT underestimates

Density functional theory: computational efficiency and (often)
satisfactory accuracy w.r.t. other ab inito methods

Basis Sets

"Combinations of mathematical functions used to represent atomic orbitals"

✓ Minimal
   H: 1s;   C, N, O: 1s, 2s, 2px, 2py, 2pz

✓ Slater type orbitals (STO)
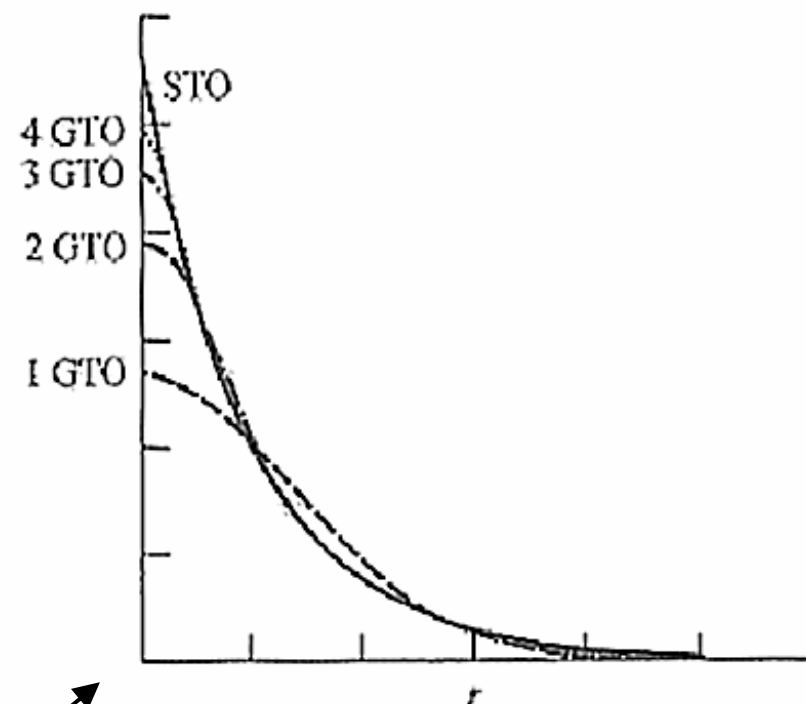   too difficult to solve analytically when combined

✓ Gaussian type orbitals (GTO)
   simpler to manipulate mathematically; combinations of Gaussian function
   [$c\, x^n\, y^m\, z^l \exp(-\alpha r^2)$] can approximate STO's

Basis functions: STO vs GTO

STO: more accurate near nuclei, but computationally expensive

GTO: less accurate near nuclei, but computationally cheaper

One uses then several GTO's grouped into « contracted gaussianss » (CG). Each CG is a linear combination of « primitives » (PG)

♦ Optimal <u>combinations</u> of Gaussian functions have been developed and are generally employed in *ab initio* work.

Split Basis Sets

Minimal (small) basis sets such as STO-3G do not adequately describe non-spherical (anisotropic) electron distribution in molecules

♦ "Split" valence basis sets (3-21G; 6-31G, etc.) were developed to overcome this problem: "each split valence atomic orbital is composed of a variable proportion of two (or more) functions of different size or radial extent"

for a "looser" bond:
($b > a$)

a + b =

for a "tighter" bond:
($a > b$)

a + b =
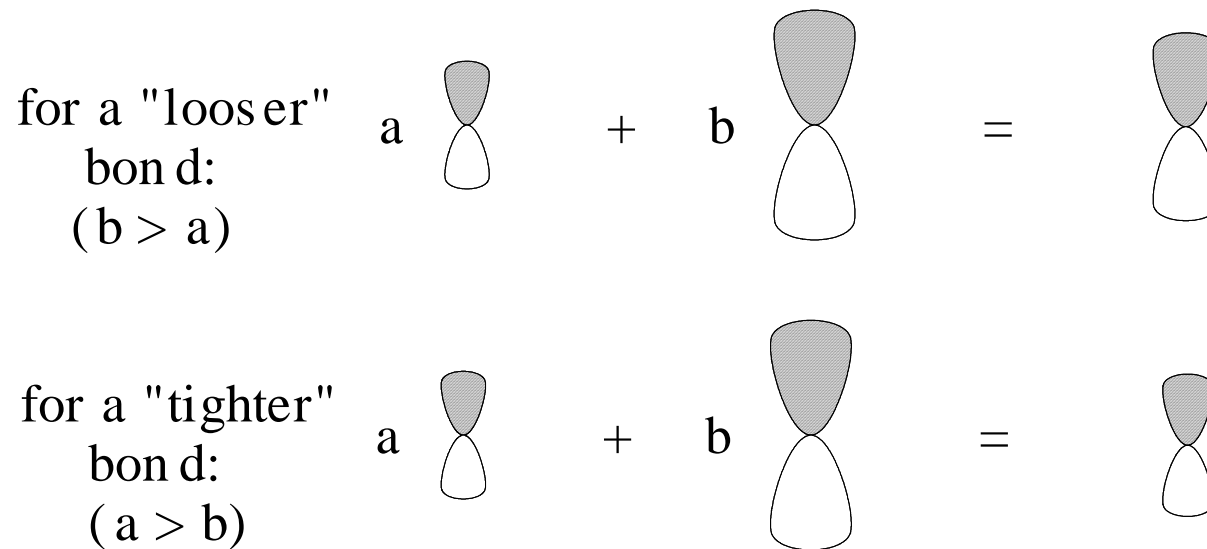
3-21G → commonly used simple split basis set

♦ good for HF geometry calculations on 1st row elements, not good for heavier elements or for accurate energies

   ♦ 3 primitive Gaussian functions for inner core (subvalence) electrons
   ♦ 2 Gaussians for contracted (small) valence orbitals
   ♦ 1 Gaussian for extended (large) valence orbitals

   6-31G, 6-311G (the latter has <u>two</u> different sizes of extended Gaussian functions for valence orbitals)

Polarization functions → 6-31G(d), or 6-31G(d,p) (a.k.a. 6-31G* and 6-31G**)

♦ adds 'd' function to 'heavy' atoms, 'p' function to H

a ⊗ + b ⊗ = ⊗
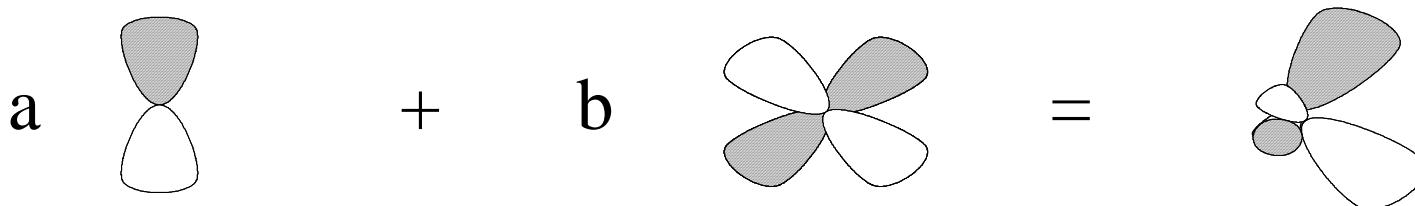
Diffuse functions

→ 6-31+G

♦ adds an additional, <u>larger p</u> function to heavy (non-hydrogen or helium) atoms; indicated by + before G

→ 6-31++G

♦ adds an additional, <u>larger p</u> function to heavy (non-hydrogen or helium) atoms and an additional <u>larger s</u> function to light elements (hydrogen and helium)

Diffuse functions are useful to describe anions, molecules with lone pairs, excited states, TS

<u>Example</u> :   ⊗ + d ⊗ = ⊗   (p diffuse orbital added)
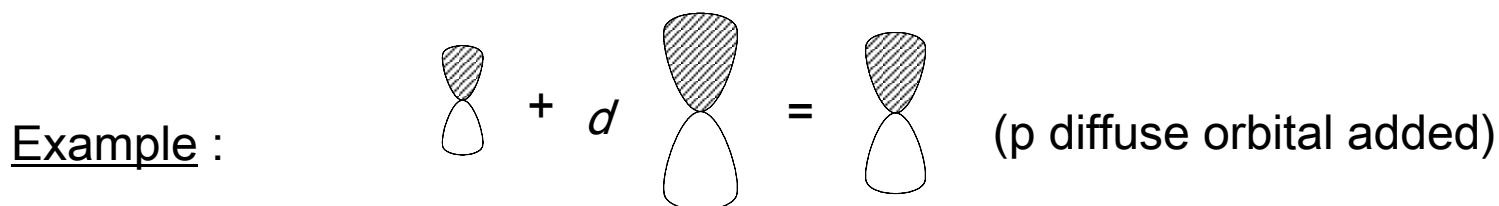
Basis Sets: Common combinations

6-31G(d)   Common 'moderate' basis set

6-31G(d,p)

6-31+G(d,p)   Good compromise

6-31++G(d,p)

Many other basis sets are in use, and basis sets can be modified/customized/optimized easily

Alphabet of Basis Sets

After > 30 years, only a handful of basis sets still used:

- STO-3G – The last MBS standing…
- "Pople-style" – $m\text{-}n_1...n_X G$ X-zeta

       m = # prim in core n_i = # prim in i^{th} valence AO

       3-21G – Pathologically good geometries for closed-
                 shell molecules w/HF (cancellation of errors)

       6-31G, 6-31G*, 6-31G**, 6-31+G, 6-31++G

       * = polarization on non-H   ** = polarization on all

       + = diffuse on non-H   ++ = diffuse on all

- cc-pvXz, aug-cc-pvXz – X-zeta - "correlation-consistent"
       best, but tend to be larger than Pople sets

Physical Interpretation

• Could just say more functions = more complete, but this gives no insight...

n-zeta:

c_{small} ◯ + c_{large} ⬭ → ◯
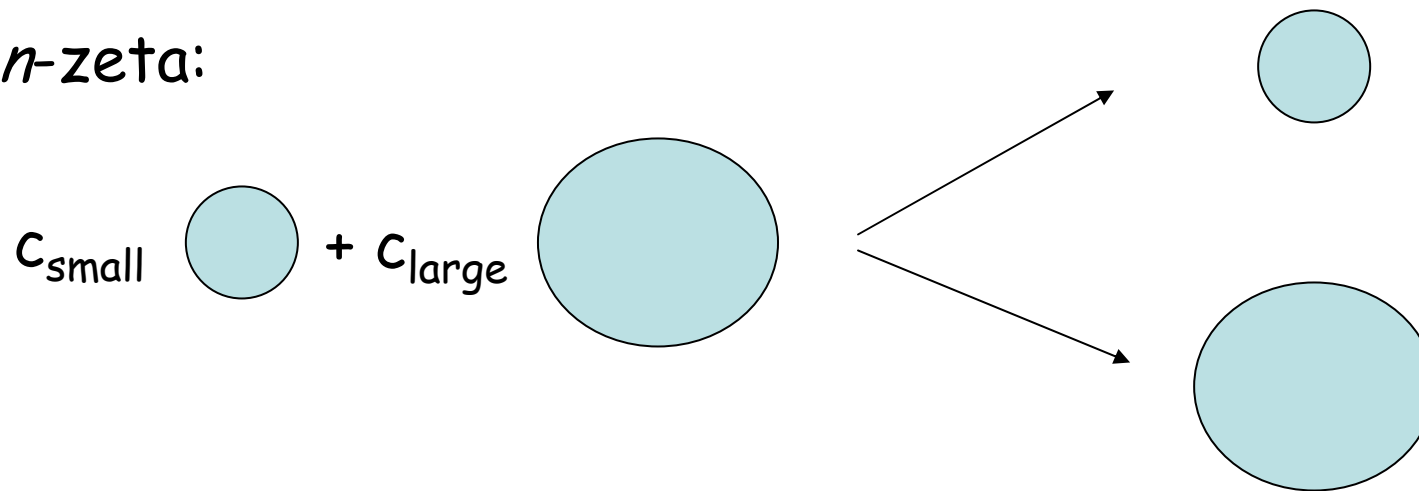
Allows orbitals to "breathe,"
i.e. to change their radial extent

Physical Interpretation II

Polarization functions:

c_s ⬭ $+c_p$ ⬭
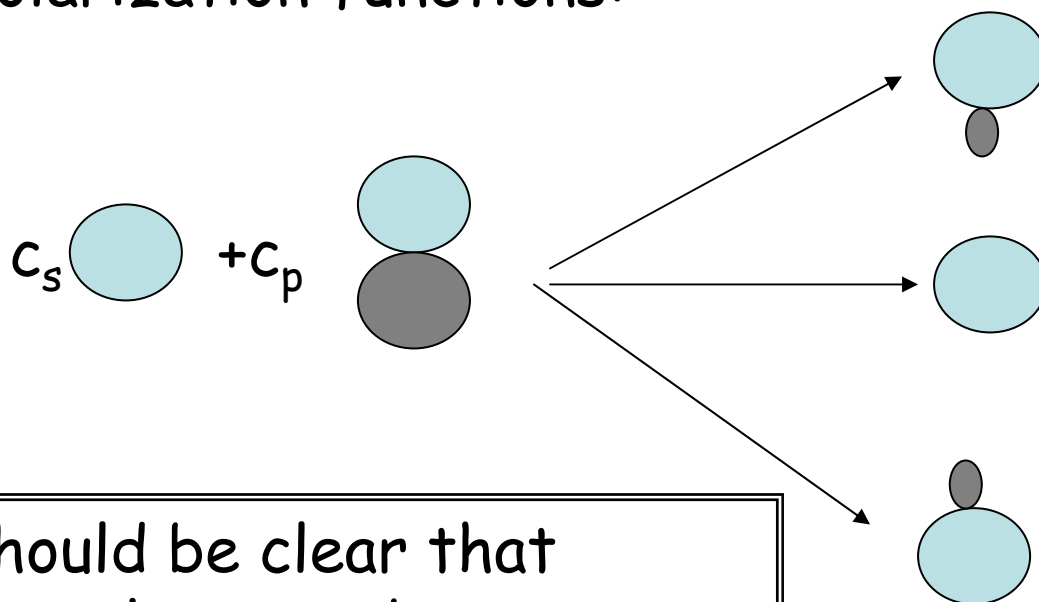
It should be clear that extra valence and polarization functions will be most important when bonds are stretched or atoms are overcoordinated
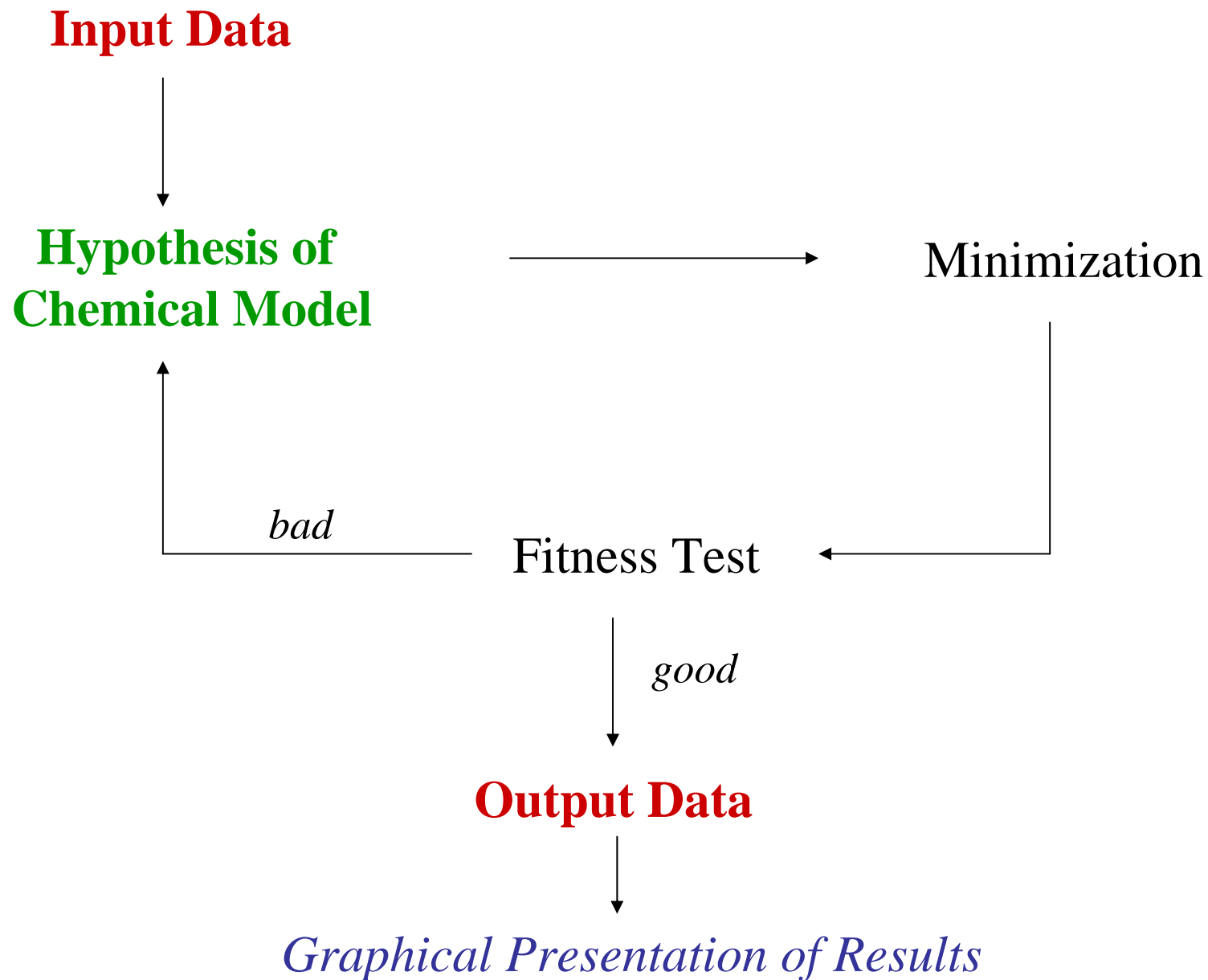
Example is for H atom; in general, polarization functions allow orbitals to "bend"

Optimization

To optimize basis set parameters, one relies on the principle that HF theory is *variational*; that is, it converges to the "true" value of the energy and will never go below that value.

The constants and exponents that describe the Gaussian functions are varied sequentially until the ***lowest energy*** is obtained.

Input Data

↓

Hypothesis of Chemical Model ——————→ Minimization

↑ *bad*

Fitness Test ←——————

↓ *good*

Output Data

↓

Graphical Presentation of Results

Direct SCF

- Setup
- Form Fock Matrix
- Evaluate Integrals
- Density converged?
  - no
  - yes
- Populations,...
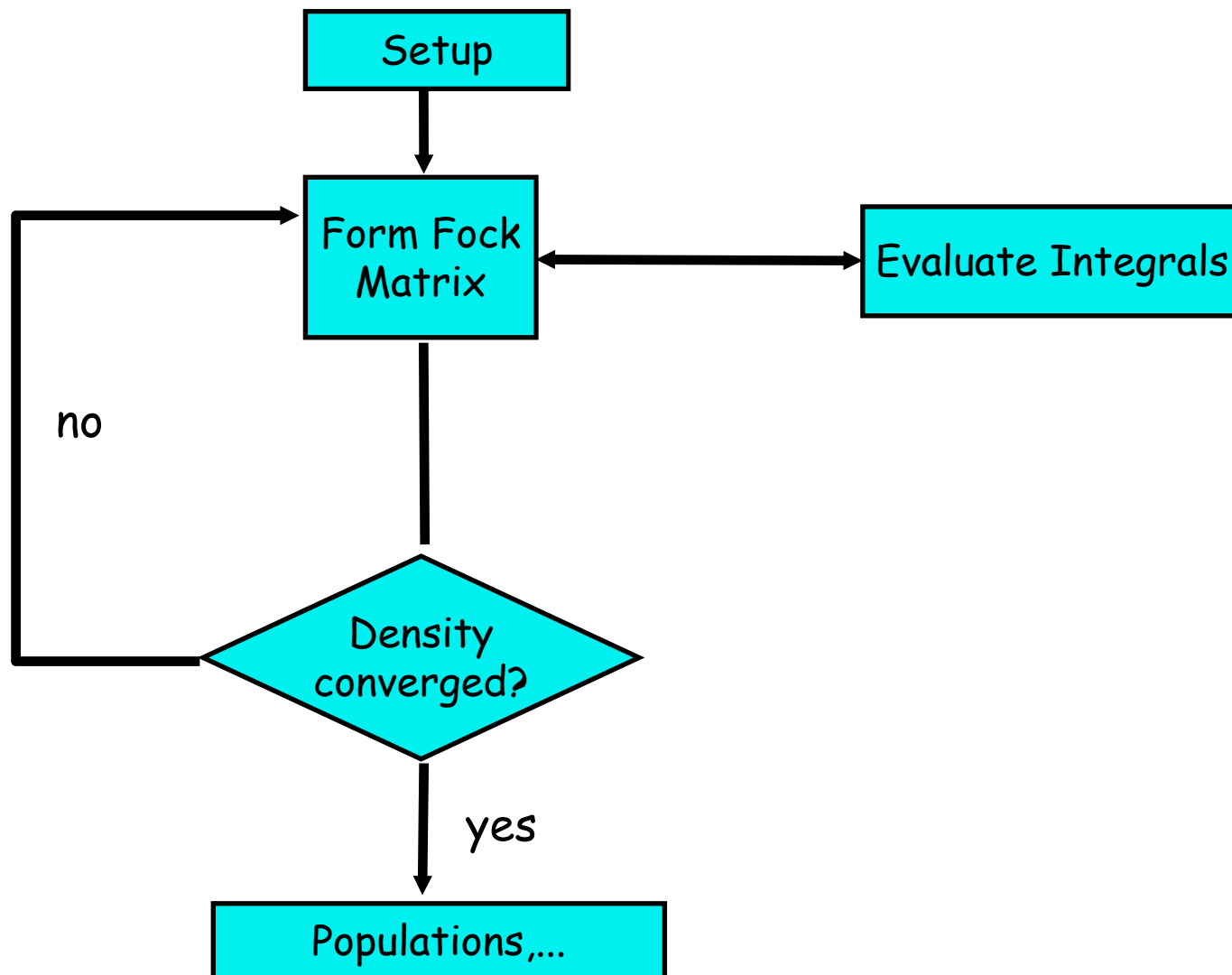
PES Topography
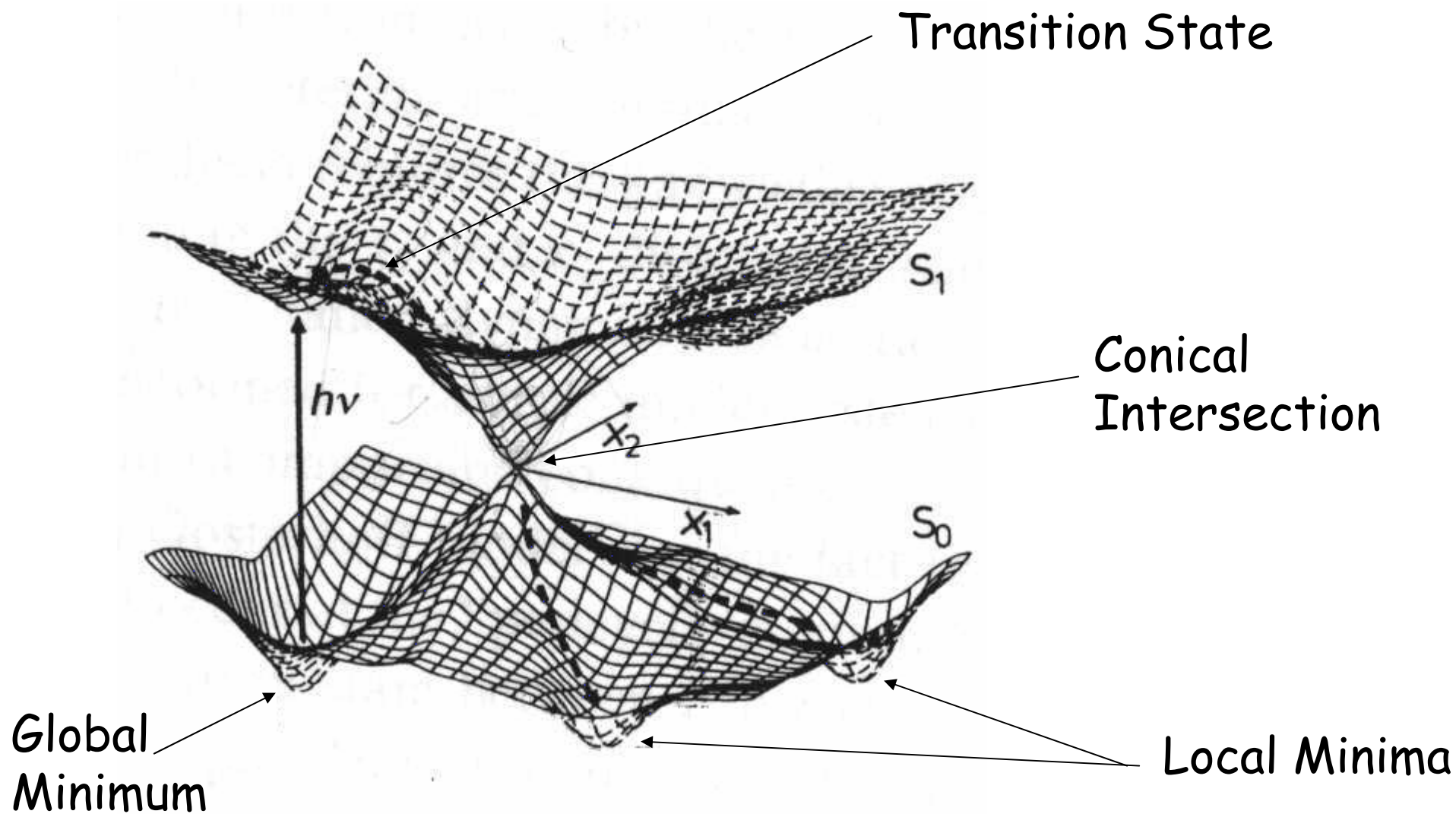
Summary of choices:
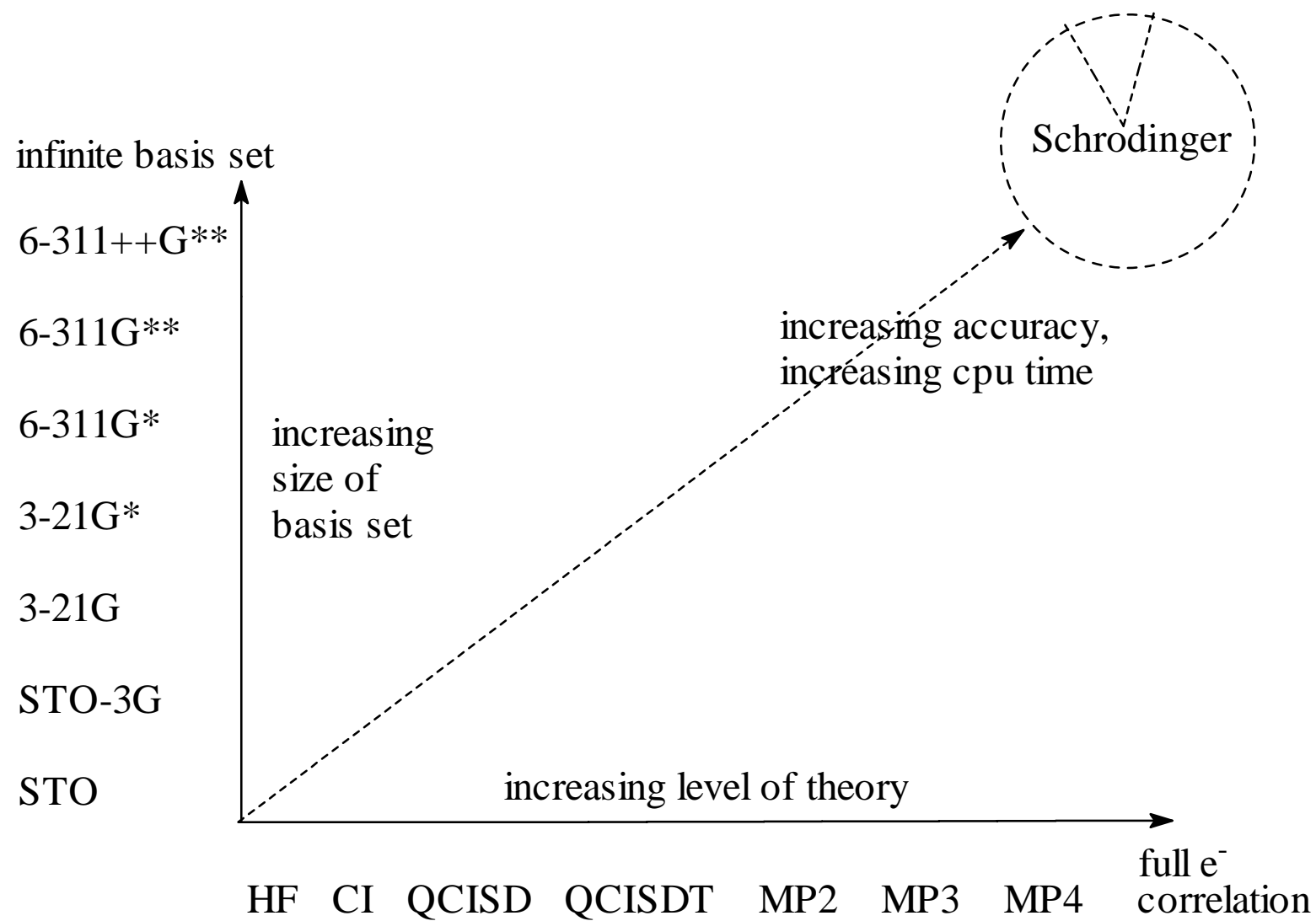
Continuum Solvent Models (e.g. PCM)

"Real" solvent $\longrightarrow$ **distribution function $g(r)$** $\longrightarrow$

$\longrightarrow$ **homogeneous distribution**
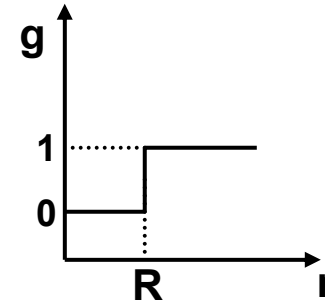
♦ **sharp reduction of degrees of freedom**

♦ **implicit average on accessible solvent conformations: make use of $\varepsilon(T)$**

♦ **both electrostatic and non-electrostatic contributions can be included**

A sharp boudary ("cavity") is defined between solute and solvent

solvation charges
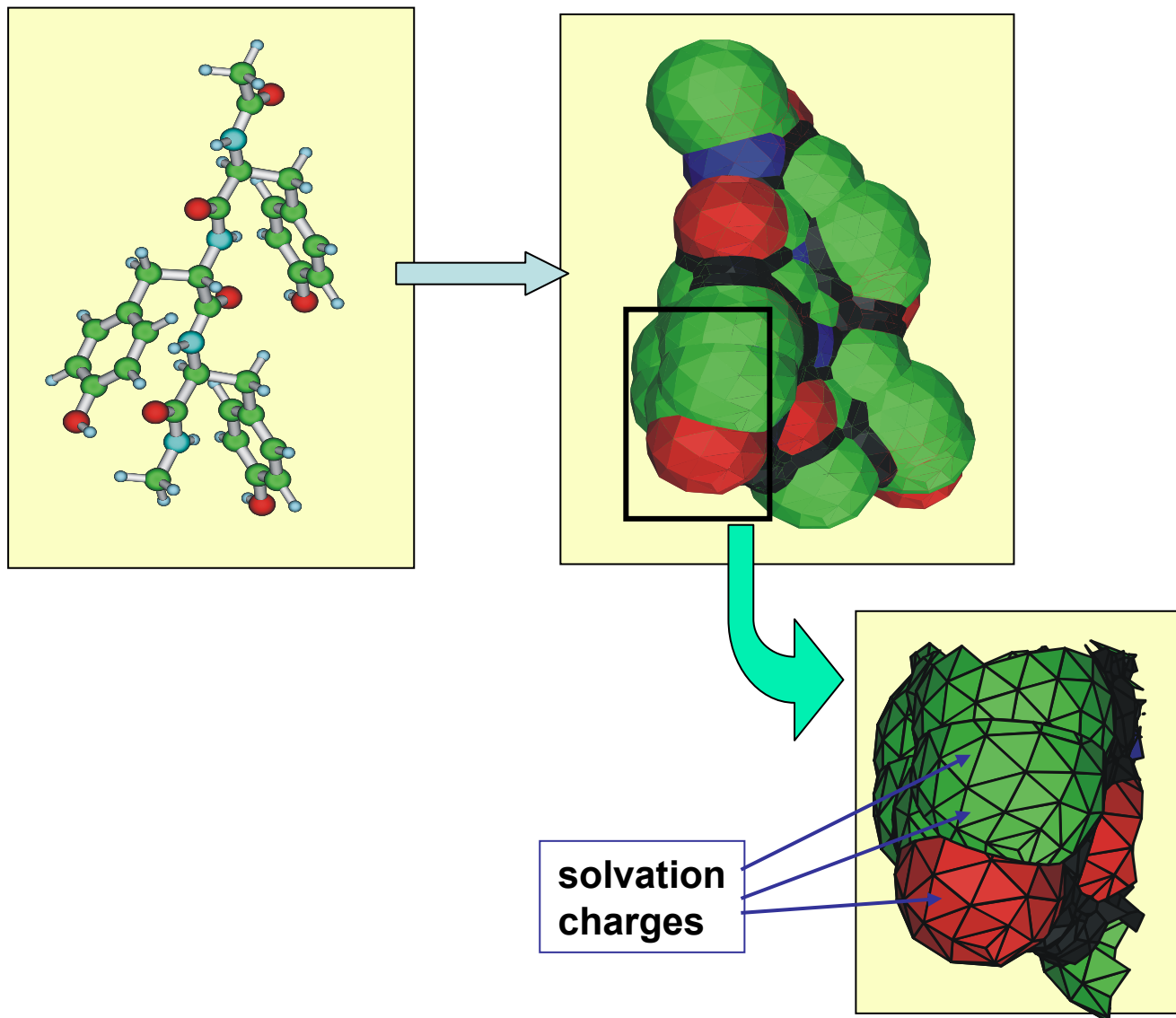
The molecular (free) energy in solution comes from
three distinct contributions:

♦ electrostatic solute-solvent interactions → also included in the molecular Hamiltonian (affecting electronic properties)
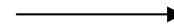
♦ cavitation energy (entropic)

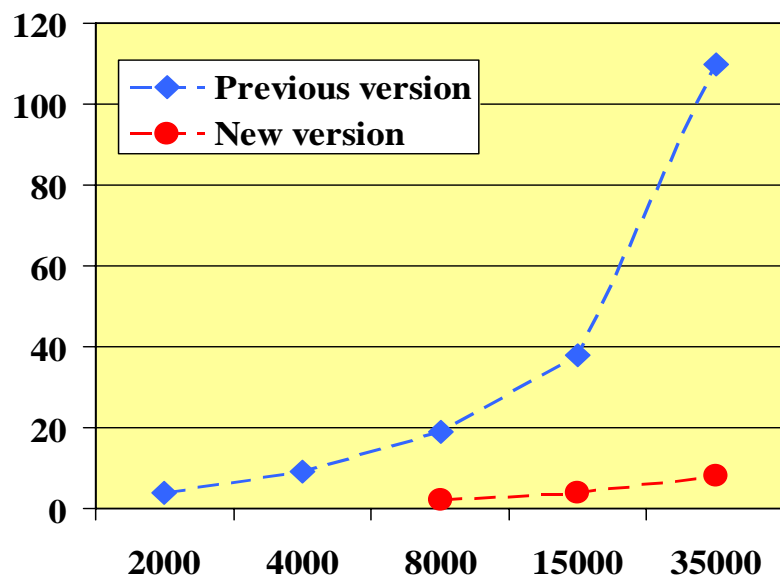♦ non-electrostatic (dispersion-repulsion) interactions

} → "classical" description
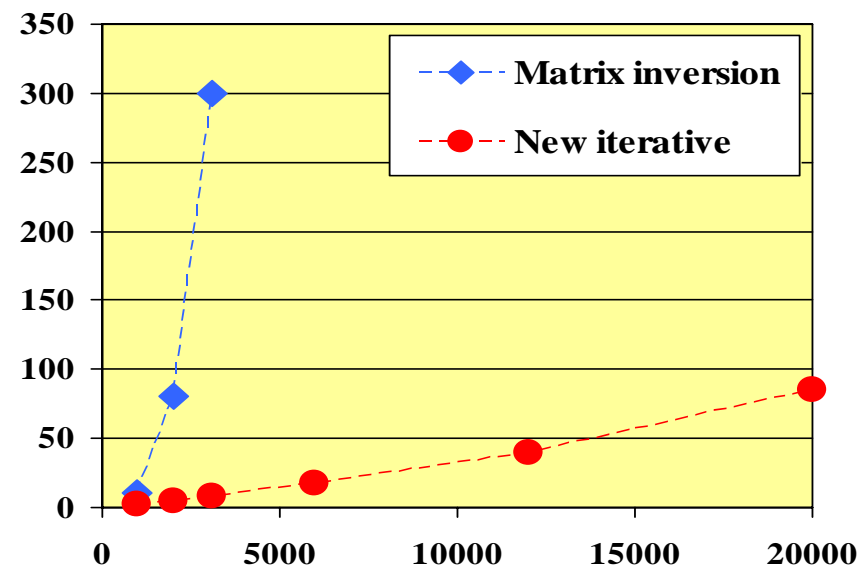
Solvation free energies computed with great accuracy:
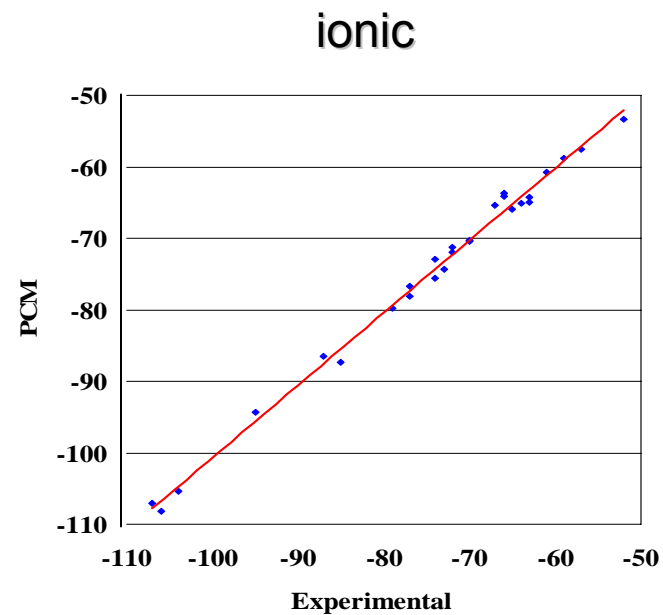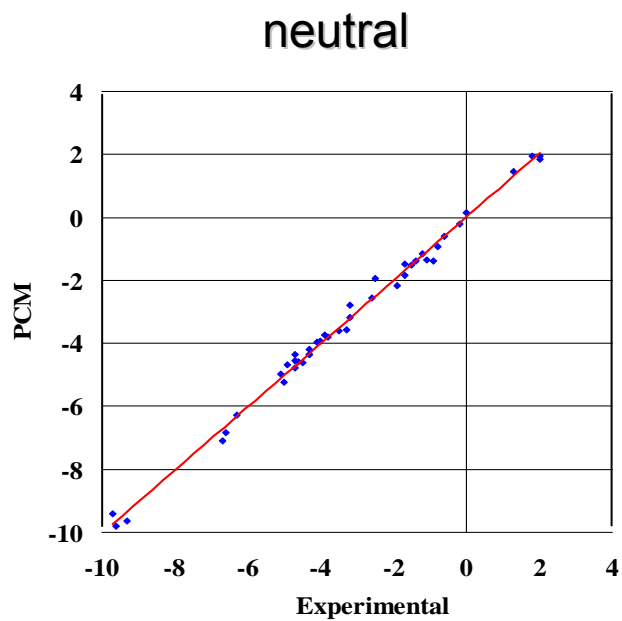
The procedure is also applicable to very large systems

CPU times (sec) to build the cavity
(w.r.t. the number of surface tesserae)

CPU times (sec) to compute the solvation charges
(w.r.t the number of surface tesserae)

Solvent as a continuum?

Sharp reduction of the degrees of freedom

Very accurate description of the solute

Possibly inadequate for the first solvation layer of polar solutes in protic solvents

Implicit averaging of accessible solvent conformations (at a given temperature)

Good reproduction of bulk effect

Solvent molecules taking part to reactions

A mixed discrete/continuum model
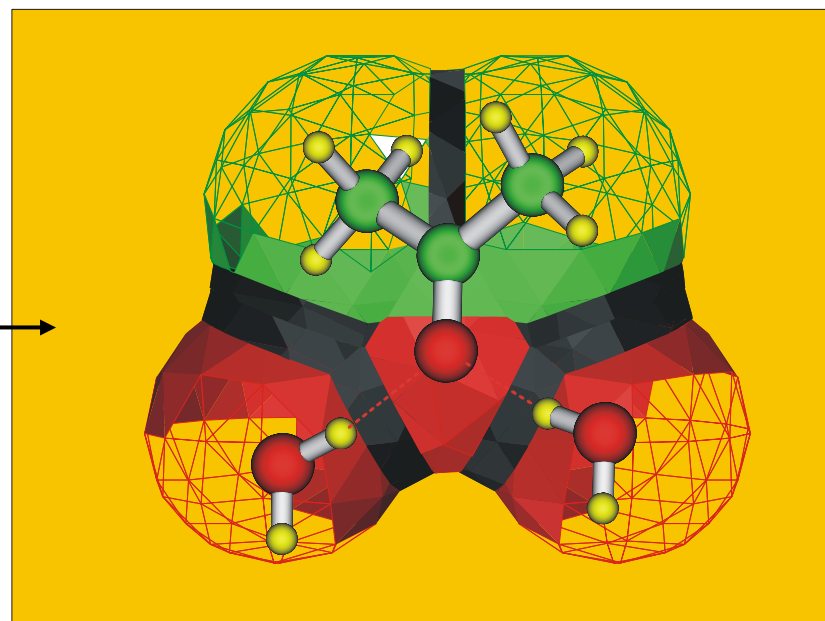
Chemical shifts calculations

The Spin Hamiltonian

$$H = \boxed{S \cdot g \cdot B + \sum_N S \cdot A \cdot I_N + S \cdot D \cdot S} \quad \textbf{EPR}$$
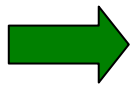
$$\boxed{\sum_N I_N \cdot (1 - \sigma_N) \cdot B + \sum_{N,M} I_N \cdot (D_{NM} + K_{NM}) \cdot I_M} \quad \textbf{NMR}$$

g, **A**, **D**, σ can be expressed as **2nd derivatives** of the **energy**
with respect to **the external field** and/or **electron/nuclear spin**

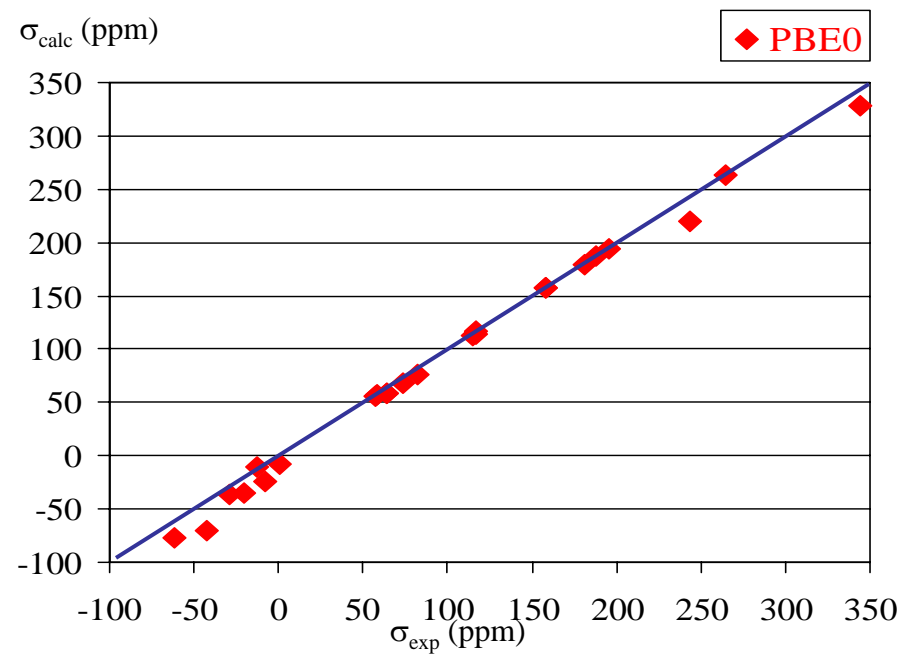$$< X >= \frac{\partial^2 E}{\partial \lambda \partial \gamma}$$

For instance
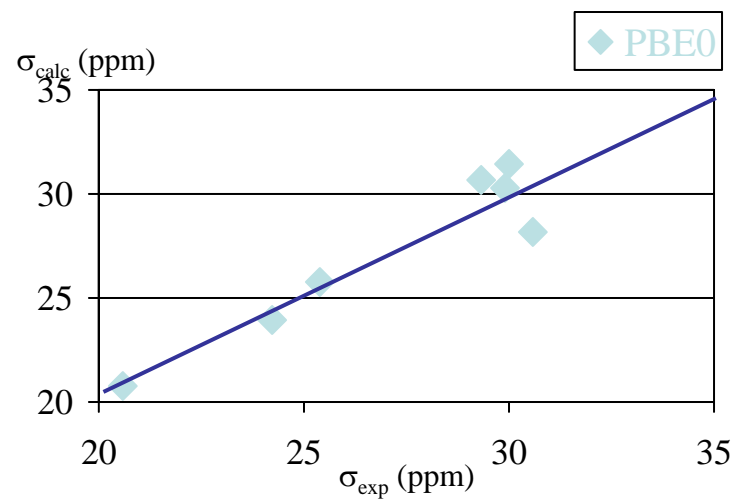$$g = \frac{\partial^2 E}{\partial B \partial s}$$

➡ Find the terms of the electronic Hamiltonian that depends on
external field and/or spin

Nuclear shielding calculated at DFT level with good accuracy for small molecules (GIAO/PBE0/6-311+G(2d,p))
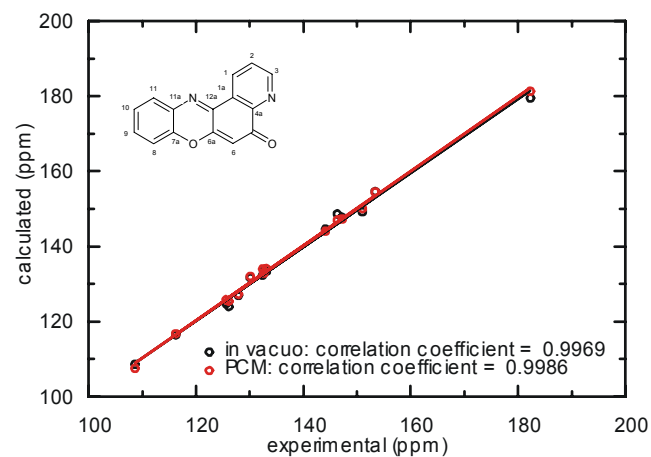
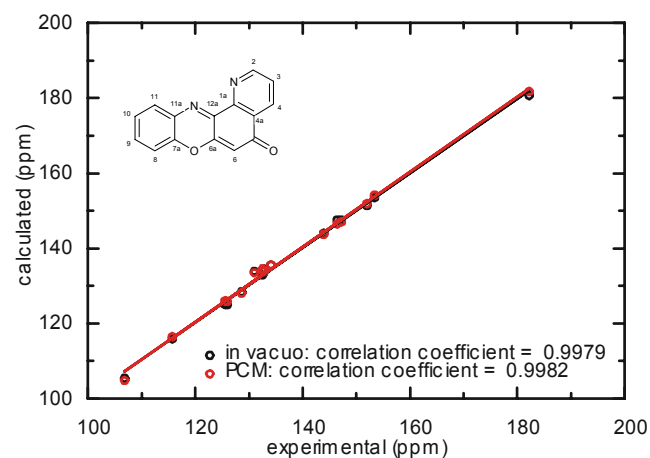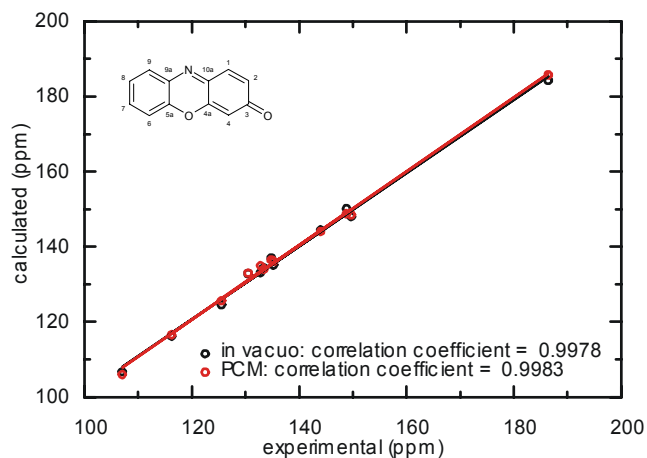**On heavy atoms (C, N) ...**

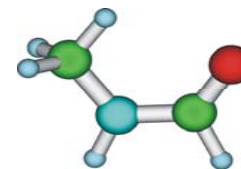**On hydrogen ...**

To evaluate a proposed signal assignment:

- E.g. ^{13}C-NMR of phenoxazones and pyridophenoxazones
  (related to the ommochromes, pigments of insects and crustaceans):

PBE0/6-311+G(d,p) // PBE0/6-31G(d);
reference compound benzene

Solvent effects on NMR spectra

^{17}O isotropic shielding in N-methyl-formamide

continuum model