Multiconfigurational approach in MOLCAS, cask matured

Valera Veryazov

Valera.Veryazov@teokem.lu.se

Division of Theoretical Chemistry
Lund University
Sweden

Overview

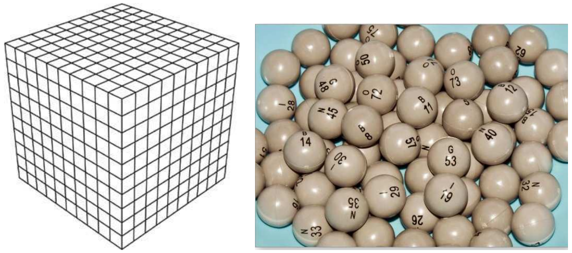
- Multiconfiguration theory in a few slides..
- Is it really difficult to select the active space?
- Molcas software
- Some state-of-the-art calculations

Quantum Chemistry, Introduction

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

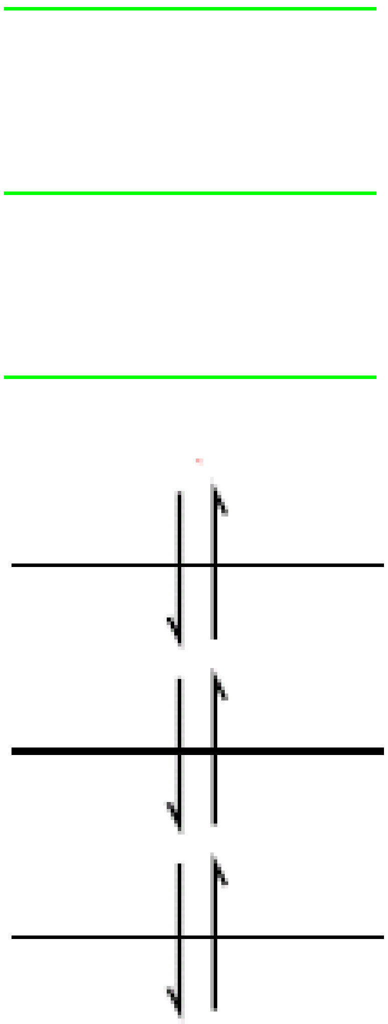
Independent particles: $\Psi = \phi(r_1) * \phi(r_2) * ... * \phi(r_n)$

Wave-function: $\Psi(r_1, r_2, ...r_n)$



a cubic box 10x10x10
1 particle= 1000 positions
26 balls=26 000 positions
26 electrons=10^{26*3}
The universe would be more simple without correlation!

Hartree-Fock theory



Hartree product $\phi(r_1) * \phi(r_2) * ... * \phi(r_n)$
$\Psi(r_1, r_2) = -\Psi(r_2, r_1)$
gives us the Hartree-Fock theory
with wavefunction in the form of
Slater determinant

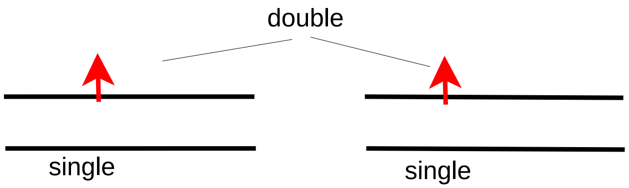
We solve HF equations for
a single configuration.

The idea fails for dissociation of H_2:

- singlet (short distance)
- triplet (long distance)

Welcome to the Zoo!

- approximate methods
  - Perturbation theory: MP2
  - Density functional theory
    - long list of functionals.....
- exact methods
  - hand made selection of excitations
  - S, SD, SDT, ... schemes
  - Coupled Cluster theory
  - The correct answer: full CI (take all of them)



Full CI

if Φ_K is antisymmetric N-electron function
Full CI : $\Psi = \sum_K C_K \Phi_K$

Number of orbitals:	n
Number of electrons	N
Spin quantum number	S

Weyl's formula for the number of configuration state functions

$$K_{n,N,S} = \frac{2S+1}{n+1} \begin{pmatrix} n+1 \\ N/2-S \end{pmatrix} \begin{pmatrix} n+1 \\ N/2+S+1 \end{pmatrix}$$

S=0 (singlet), n=8

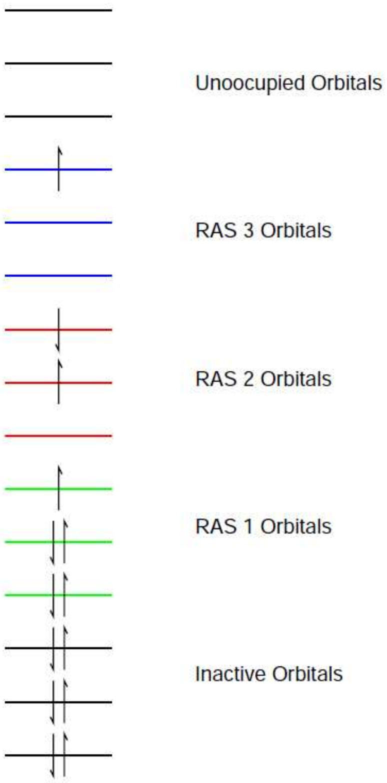
N	2	4	6	8	10	12	14	16
K	36	336	1176	1764	1176	336	36	1

It grows!

```
 2                                      3
 4                                     20
 6                                    175
 8                                   1764
10                                  19404
12                                 226512
14                                2760615
16                               34763300
18                              449141836
20                             5924217936
22                            79483257308
24                          1081724803600
26                         14901311070000
28                        207426250094400
30                       2913690606794775
32                      41255439318353696
34                     588272005095043456
36                    8441132926294525952
38                  121805548126430068736
40                 1766594752418699542528
```

Multiconfigurational approach

CASSCF =
Full CI on a limited set of orbitals
(active orbitals)
RASSCF =
Full CI on a limited set of orbitals plus SD..
for limited set of orbitals



CASPT2 - PT2 theory on top of CASSCF
RASPT2 - PT2 theory on top of RASSCF

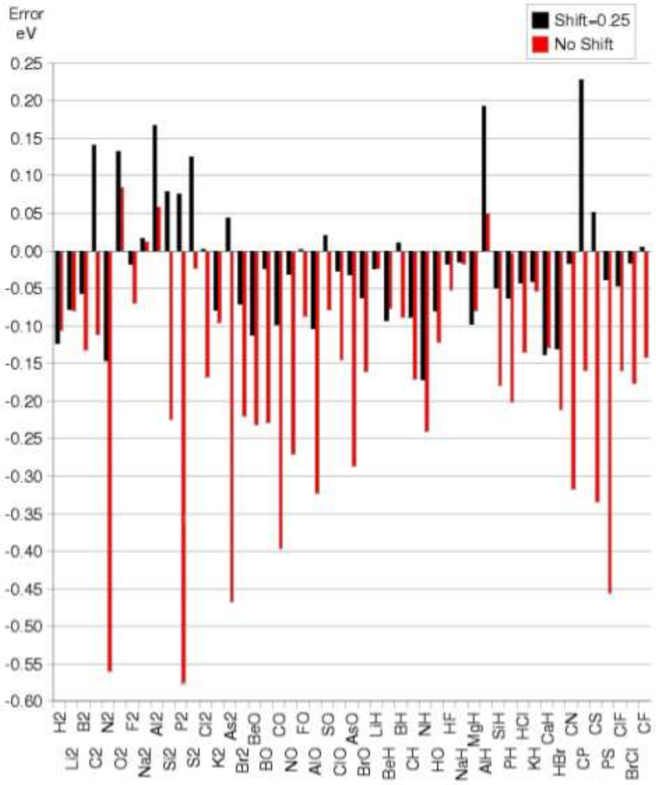
Multistate CASPT2
RAS State Interaction

Is it accurate?

Dissociation energies for two-atomic molecules.

Note, PT2 hamiltonian strictly speaking allows some parametrization.

Here the results for two variants the standard one 'red' and with IPEA shift 'black'



Why do we [still] use DFT?

- CASSCF is not a 'black box' method
- The results really depends on the active space!
- It is very easy to go beyond the size limits
- PT2 theory can not improve problems with CAS
- popular myths:
    - the selection of the active space is an ART
    - small active space will do the trick
    - active space can cure itself
    - intruder states in CASPT2
- not easy to implement

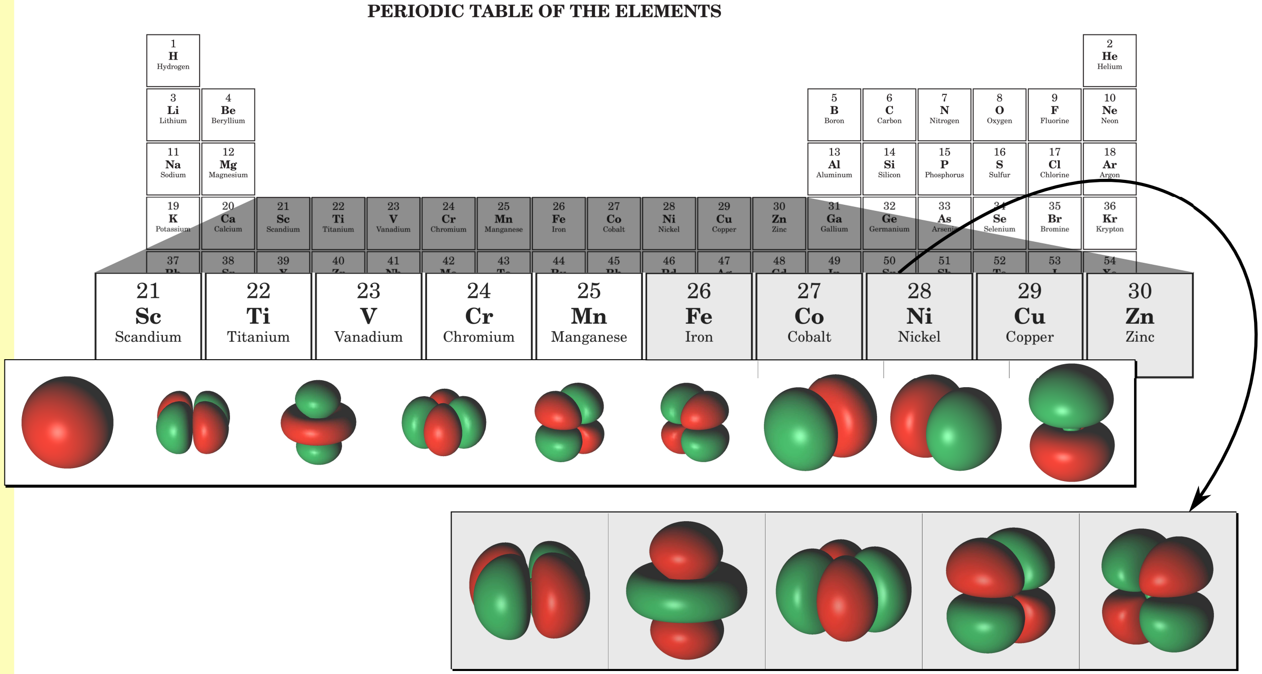
Selection of active space

"Sometimes trivial, sometimes more difficult, sometimes impossible" (Björn Roos)

- depends on the problem, and can not be automatized
- for the ground and excited states:
  - π orbitals for aromatic molecules
  - for transition metals and lanthanides/actinides based on periodic table
- for dissociation: must include reactant and product

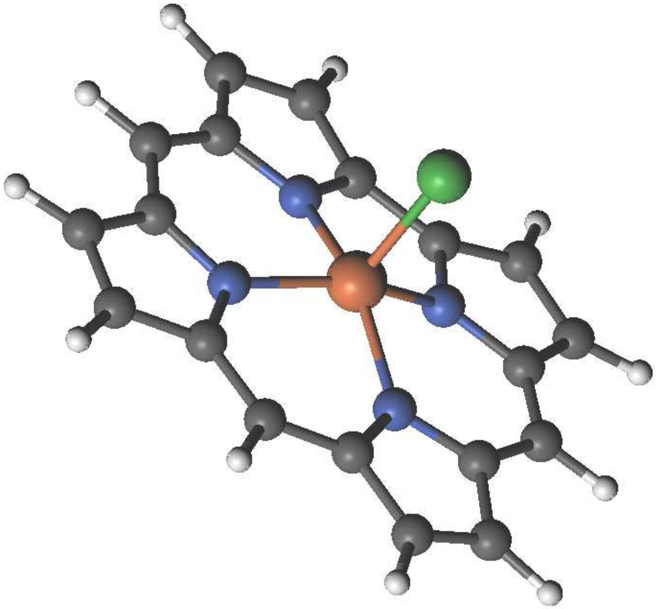
V.Veryazov, P-Å.Malmqvist, B.O.Roos, IJQC 111 (2011) 3329-3338

Selection, based at periodic table



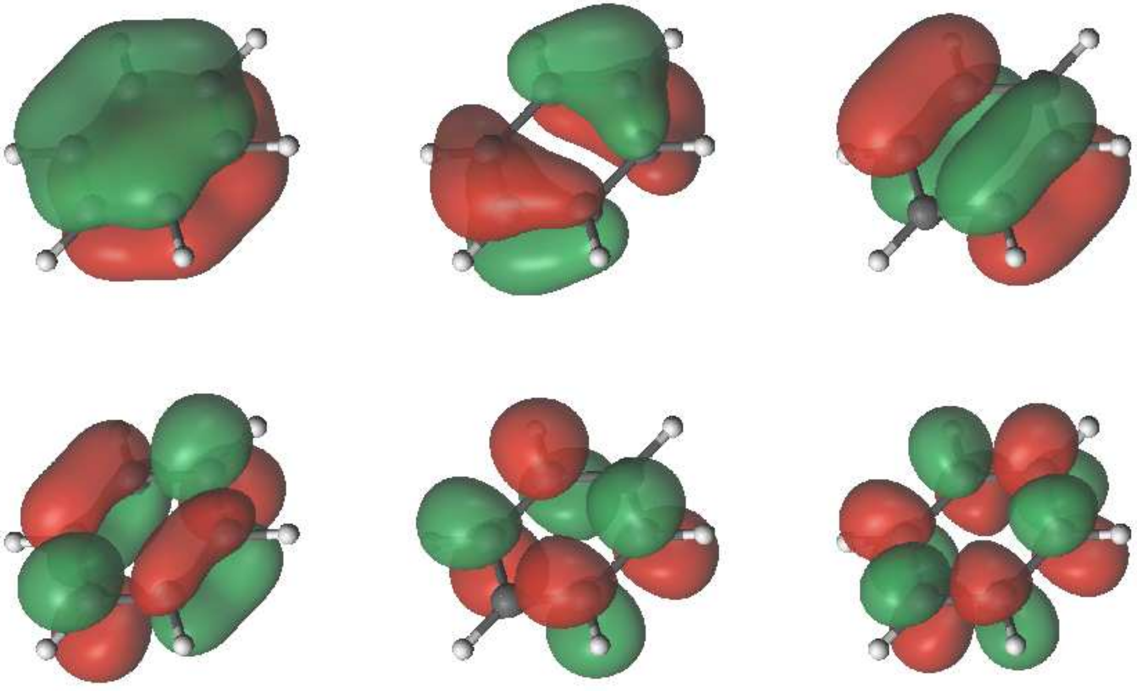
way too many...

So, we would like to include: π-bonds, $d-$shell (double?), etc.



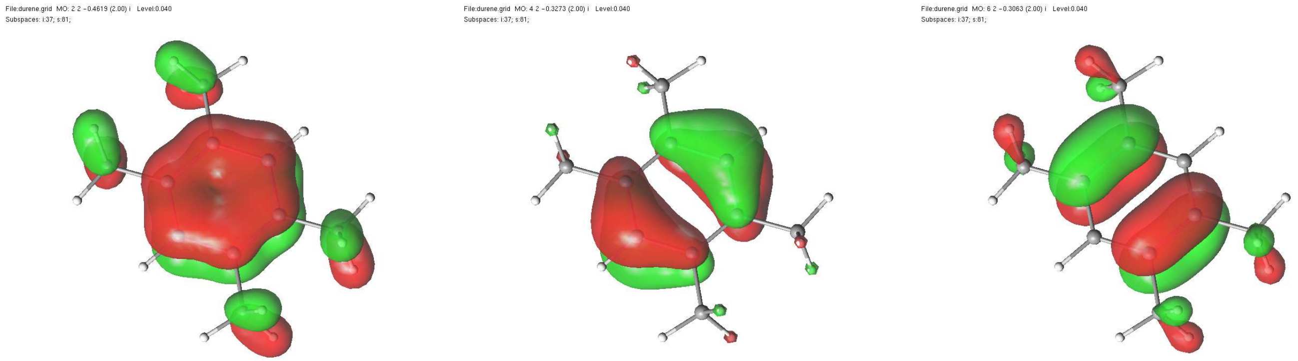
"around HOMO-LUMO"

benzene molecule: 6 $\pi-$ orbitals: 3 occupied and 3 virtual



HOMO-4, HOMO-2, HOMO, LUMO, LUMO+1, LUMO+2

from benzene to durene



We can easily recognize π orbitals,
but they are not located around HOMO-LUMO!

For virtuals: **LUMO+14, LUMO+15, LUMO+43**

Implementation of CASSCF/CASPT2

- CPU (BLAS) demanding
- extremely high demand for memory and I/O
- difficult to parallelize
- So, in many codes - very limited CAS/CASPT2
  (which makes it useless)

Molcas: early 90’s to present day

- started by Björn Roos in early 90’s
- non-commercial, but licensed software
- open for developers and contributors
- main nodes: Lund and Uppsala
- about 50 developers
- used at more than 200 academic groups
- `http://www.molcas.org`

Molcas software

Software for Quantum Chemistry

- for ground and excited states,
- for molecular structure, chemical and photochemical reactions,
- for solution chemistry, chemical bonding, and much more,
- for all elements of the periodic table.
- from DFT to Coupled Cluster
- Cholesky decomposition and resolution of identity
- analytical or numerical gradients
- runs on any platform

Molcas software, cont.

Most important point: multiconfigurational theory
CASSCF/RASSCF and CASPT2/RASPT2

How big and how fast?

- no hard limit in active space, but memory, disk, etc. running out at
  - about 16 on 16 orbitals (no symmetry) for CASSCF
  - about 40 on 40 orbitals for RASSCF
- Number of atoms about 120
- basis set size about 2000

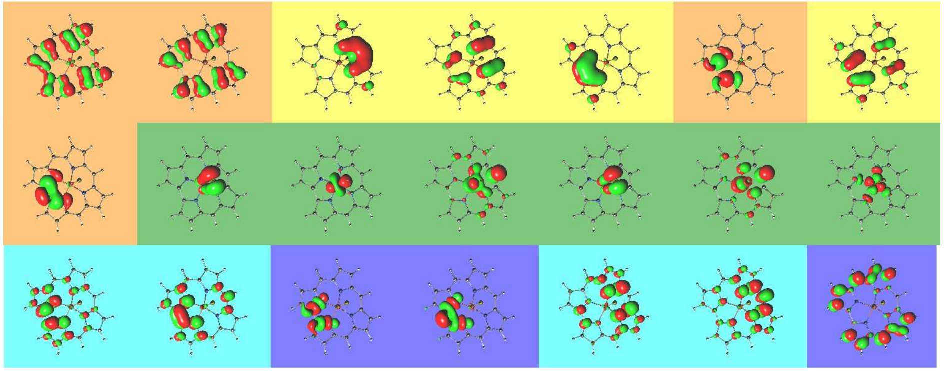
Swiss army knife for CAS in MOLCAS

Tools to be used in C[R]AS calculations:

- GV - simple graphical interface for active space selection
- expbas - a tool to transform density, computed with different basis
- localization - orbitals localization
- cmocorr - a tool to compare initial and final active spaces
- RASSCF(SD) probe - a technique to identify potentially active orbitals

Visual inspection of orbitals

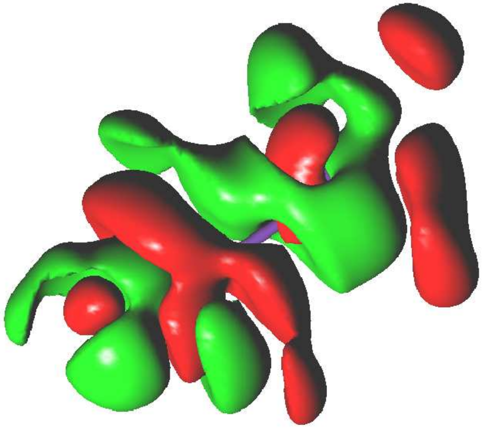
- run preliminary calculation
- visualize orbitals, using GV code



- identify the active space
- use orbital file with selected active space,
  as an input for CASSCF/RASSCF calculation

Basis set

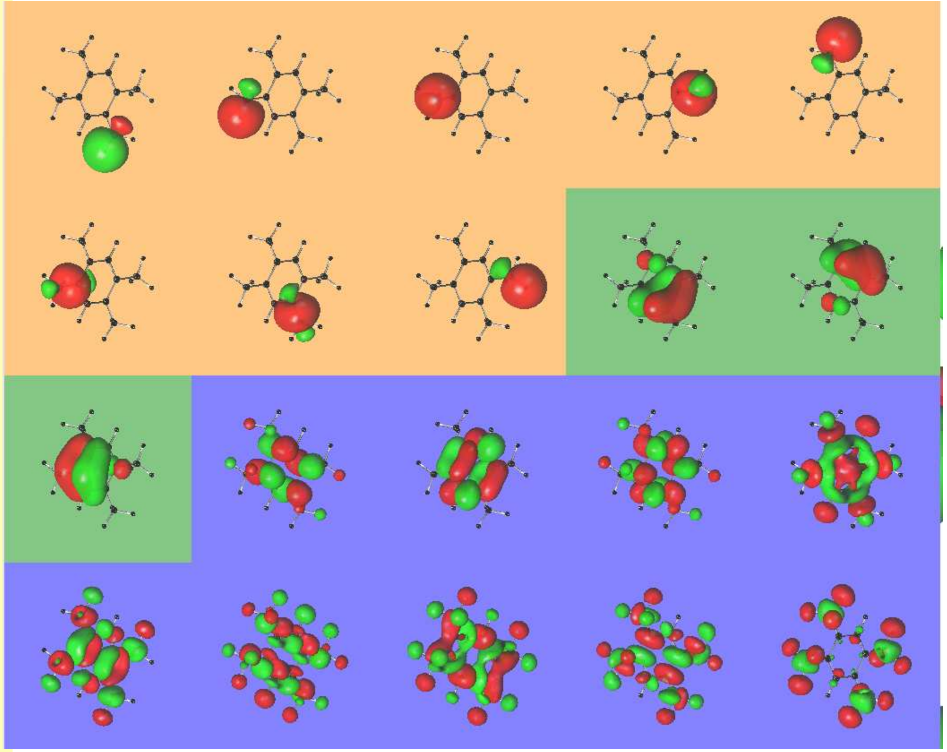
The basis set must be large in ’real calculations’.
However, it is very difficult to recognize an orbital,
computed with extended basis set.



- make a calculation with minimal basis set
- identify orbitals
- use expbas tool to transform them to larger basis set
- (bonus) better starting point

Localization of orbitals

Localization is mostly used to get charges, bond orders.
But it can also be used to recognize the orbitals.



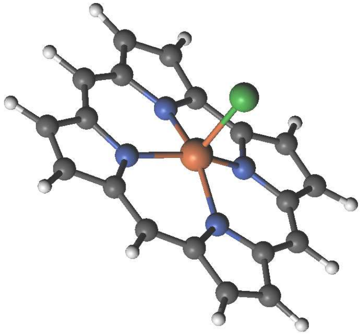
RAS probing

The idea:
give a freedom, and see how it was used...

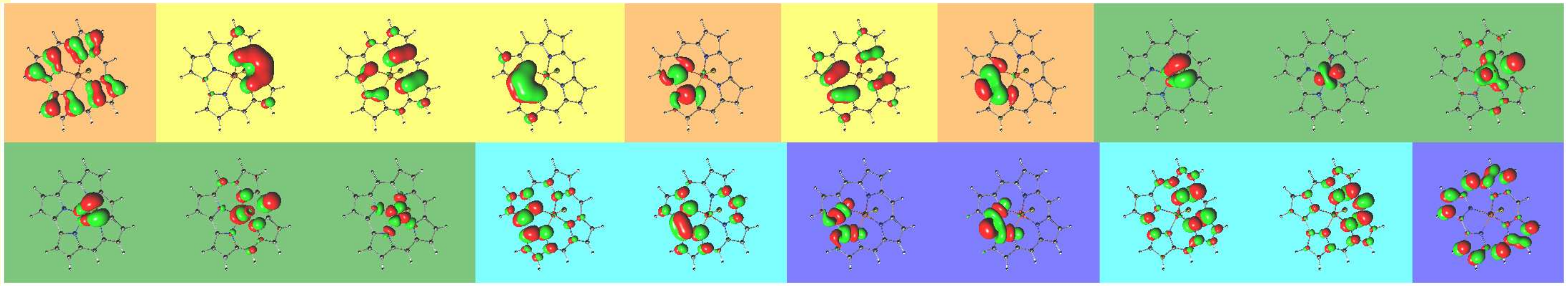
- add orbitals to the active space for RASSCF calculation
- check the occupation numbers
- keep those in between 0 and 2
- make another active space

A success story..

Chloroiron corrole.
A study published in 2008 suggests active space /0/14/0.



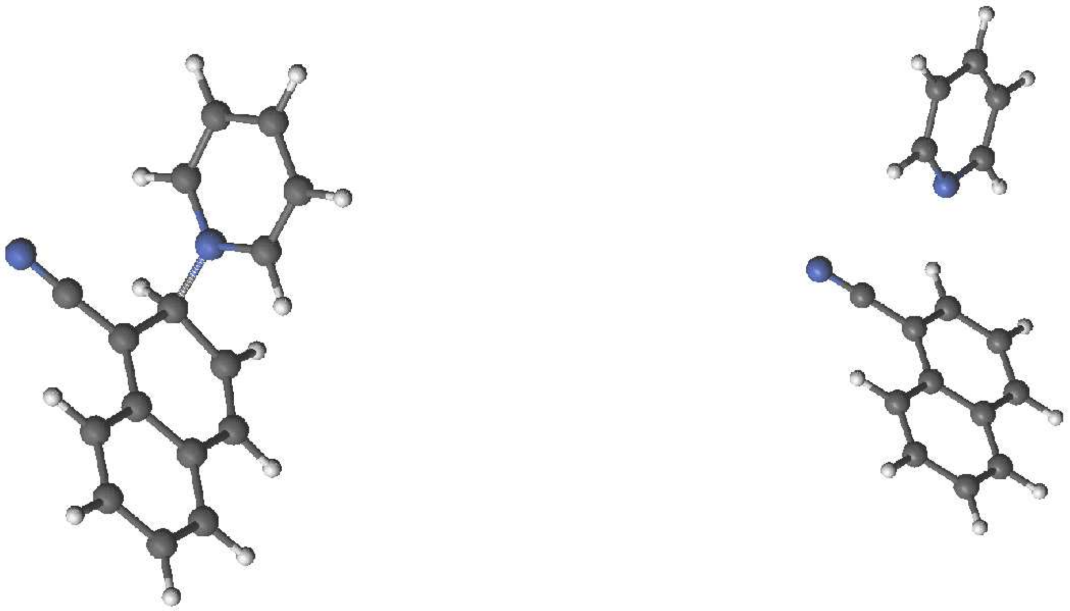
initial RASSCF calculation was made for /6/6/6/ set.
the set can be reduced to /4/6/4/,
which gives results similar to /0/14/0/



Exciplex formation

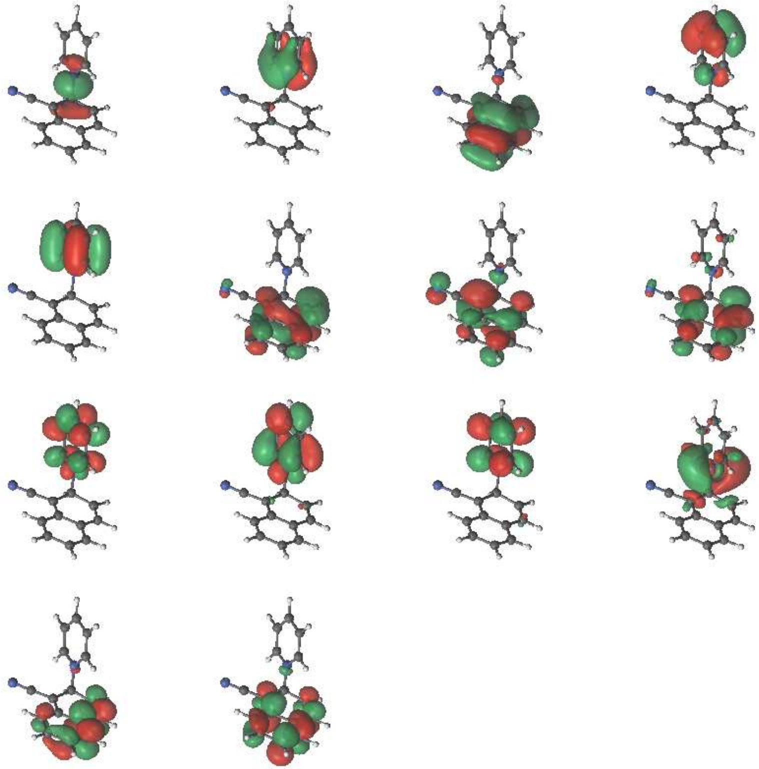
Pyridine and cyanonaphtalene can form an exciplex with distance 1.45 Å between the molecules.

experiment shows repulsive interaction for the ground state and a weak minimum for excited states



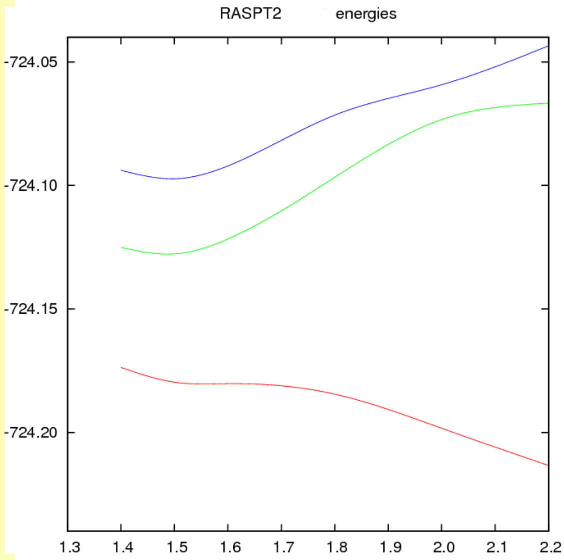
23 orbitals.

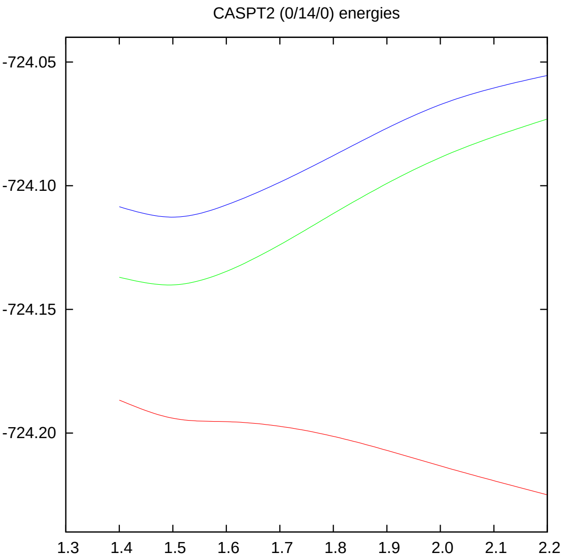
Complex pyridine-CN-napth



Two variants of RAS spaces:
pyridine-like (for non interacting molecules) /9/4/8/
and complex (for short distances) /11/3/9/ *vs.* CAS /0/14/0/.

RAS vs. CAS

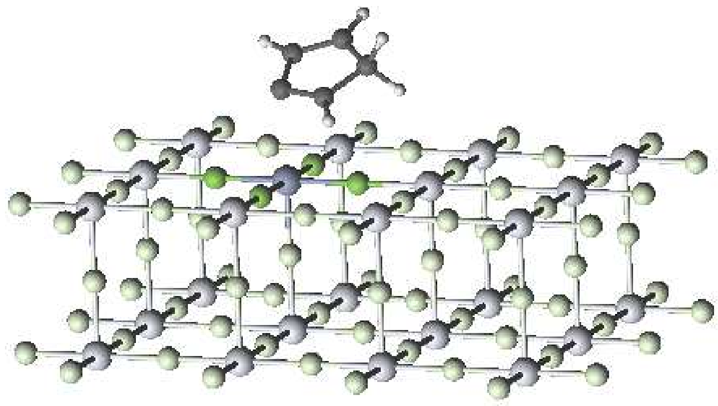




The similar results with very different cost!
in both cases - a minimum in the excited state has been found

Material Science applications

Many applications in material science are related to bond creation/breaking. Precise treatment is often needed.

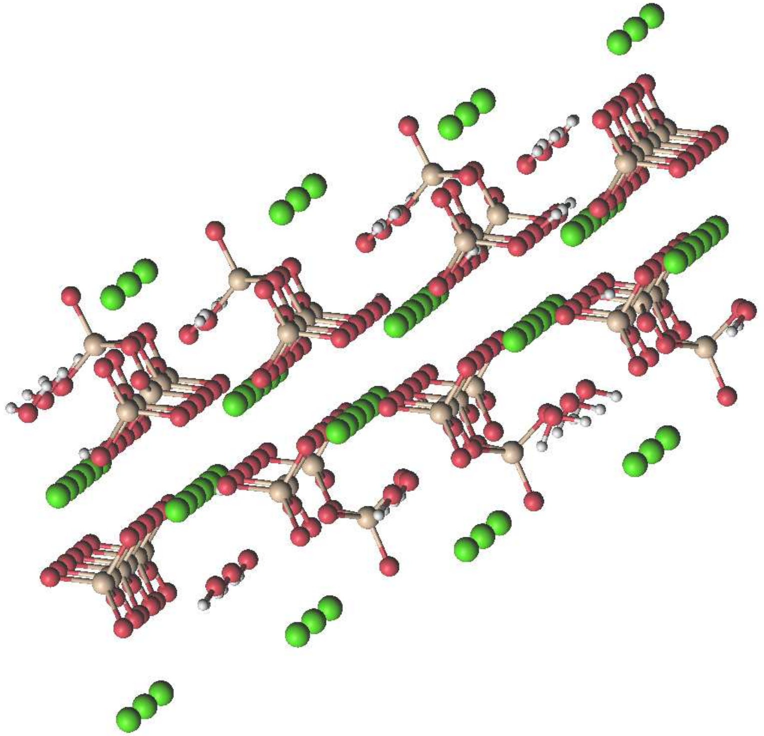


To describe molecules of a surface we are developing

- new schemes to decrease the number of virtual orbitals
- new schemes for geometry optimization
- basis sets to mimic embedding
- parallelization and general speed up

Crystal growth problem

A study of formation of C-S-H nanoparticles:
the basic blocks of concrete (sponsored by BASF).
Interaction of surface with water and $Si(OH)_4$



todo list – work in progress

- better adjustment for modern hardware
- better parallelized
- new geometry optimizer for numerical gradients
- analytical CASPT2 gradients
- better GUI
- cluster embedding
- free student edition

Thanks to Molcas team,
especially Per-Åke Malmqvist, Roland Lindh,
Per-Olof Widmark, Victor Vysotskiy, Steven Vancoillie