

(continued from previous page)

5	C	dxy	0.0	0.1	0.5	0.0	0.0	0.0
6	H	s	7.5	0.0	7.5	2.5	0.0	2.5
7	H	s	7.5	0.0	7.5	2.5	0.0	2.5
8	H	s	7.5	0.0	0.0	10.0	0.0	9.9
9	H	s	7.5	0.0	7.5	2.5	0.0	2.5
10	H	s	7.5	0.0	7.5	2.5	0.0	2.5
11	H	s	7.5	0.0	0.0	10.0	0.0	9.9
			18	19	20	21	22	23
			-0.49833	-0.33937	-0.33937	0.13472	0.13472	0.18198
			2.00000	2.00000	2.00000	0.00000	0.00000	0.00000
			-----	-----	-----	-----	-----	-----
0	C	s	0.1	0.0	0.0	0.0	0.0	2.2
0	C	pz	0.0	8.1	24.4	7.8	23.4	0.0
0	C	px	0.1	0.0	0.0	0.0	0.0	0.6
0	C	py	10.4	0.0	0.0	0.0	0.0	1.7
0	C	dxz	0.0	0.4	0.2	0.7	0.7	0.0
0	C	dyz	0.0	0.2	0.0	0.7	0.0	0.0
0	C	dx2y2	0.0	0.0	0.0	0.0	0.0	0.2
0	C	dxy	1.0	0.0	0.0	0.0	0.0	0.5
1	C	s	0.1	0.0	0.0	0.0	0.0	2.2
1	C	pz	0.0	8.1	24.4	7.8	23.4	0.0
1	C	px	0.1	0.0	0.0	0.0	0.0	0.6
1	C	py	10.4	0.0	0.0	0.0	0.0	1.7
1	C	dxz	0.0	0.4	0.2	0.7	0.7	0.0
1	C	dyz	0.0	0.2	0.0	0.7	0.0	0.0
1	C	dx2y2	0.0	0.0	0.0	0.0	0.0	0.2
1	C	dxy	1.0	0.0	0.0	0.0	0.0	0.5
2	C	s	0.0	0.0	0.0	0.0	0.0	2.2
2	C	pz	0.0	32.5	0.0	31.2	0.0	0.0
2	C	px	0.0	0.0	0.0	0.0	0.0	2.2
2	C	py	11.6	0.0	0.0	0.0	0.0	0.0
2	C	dxz	0.0	0.1	0.0	0.3	0.0	0.0
2	C	dyz	0.0	0.0	0.8	0.0	1.8	0.0
2	C	dx2y2	0.0	0.0	0.0	0.0	0.0	0.7
2	C	dxy	0.4	0.0	0.0	0.0	0.0	0.0
3	C	s	0.1	0.0	0.0	0.0	0.0	2.2
3	C	pz	0.0	8.1	24.4	7.8	23.4	0.0
3	C	px	0.1	0.0	0.0	0.0	0.0	0.6
3	C	py	10.4	0.0	0.0	0.0	0.0	1.7
3	C	dxz	0.0	0.4	0.2	0.7	0.7	0.0
3	C	dyz	0.0	0.2	0.0	0.7	0.0	0.0
3	C	dx2y2	0.0	0.0	0.0	0.0	0.0	0.2
3	C	dxy	1.0	0.0	0.0	0.0	0.0	0.5
4	C	s	0.1	0.0	0.0	0.0	0.0	2.2
4	C	pz	0.0	8.1	24.4	7.8	23.4	0.0
4	C	px	0.1	0.0	0.0	0.0	0.0	0.6
4	C	py	10.4	0.0	0.0	0.0	0.0	1.7
4	C	dxz	0.0	0.4	0.2	0.7	0.7	0.0
4	C	dyz	0.0	0.2	0.0	0.7	0.0	0.0
4	C	dx2y2	0.0	0.0	0.0	0.0	0.0	0.2
4	C	dxy	1.0	0.0	0.0	0.0	0.0	0.5
5	C	s	0.0	0.0	0.0	0.0	0.0	2.2
5	C	pz	0.0	32.5	0.0	31.2	0.0	0.0
5	C	px	0.0	0.0	0.0	0.0	0.0	2.2
5	C	py	11.6	0.0	0.0	0.0	0.0	0.0
5	C	dxz	0.0	0.1	0.0	0.3	0.0	0.0
5	C	dyz	0.0	0.0	0.8	0.0	1.8	0.0
5	C	dx2y2	0.0	0.0	0.0	0.0	0.0	0.7
5	C	dxy	0.4	0.0	0.0	0.0	0.0	0.0
6	H	s	7.4	0.0	0.0	0.0	0.0	11.5

(continues on next page)

(continued from previous page)

7	H	s	7.4	0.0	0.0	0.0	0.0	11.5
8	H	s	0.0	0.0	0.0	0.0	0.0	11.5
9	H	s	7.4	0.0	0.0	0.0	0.0	11.5
10	H	s	7.4	0.0	0.0	0.0	0.0	11.5
11	H	s	0.0	0.0	0.0	0.0	0.0	11.5

Since the molecule is aligned in the x-y plane, the desired orbitals must have 'pz' character. We see that the occupied  $\pi$ -orbitals number 16, 19, 20 and the unoccupied ones start with 21 and 22. However, the sixth high-lying  $\pi^*$ -orbital cannot easily be found. The task is much easier using natural orbitals from a correlated calculation. Let us run a simple selected RI-MP2 calculation and look at the natural orbitals.

```
! RHF SV(P) RI-MP2 def2-svp/C
! moread
%moinp "Test-CASSCF-Benzene-1.gbw"

%mp2
# TNat is an alias:
# - selected MP2 wave function (1e-5)
# - compute natural orbitals with the unrelaxed density.
# alternative: "natorbs true" with "density unrelaxed".
# ==> produces a gbw file with the suffix ".mp2nat"
tnat 1e-5
end

# ...etc, input of coordinates
```

The calculation prints the occupation numbers:

```
N[ 16] = 1.95798013
N[ 17] = 1.95798013
N[ 18] = 1.95330649
N[ 19] = 1.91852498
N[ 20] = 1.91852498
N[ 21] = 0.06544306
N[ 22] = 0.06544306
N[ 23] = 0.02820679
N[ 24] = 0.02405988
```

From these occupation number it becomes evident that there are several natural orbitals which are not quite doubly occupied MOs. The most promising orbitals are right at the HOMO-LOMO gap. These are typically the  $\pi$  and  $\pi^*$  orbitals. In a CASSCF(6,6), the orbitals 18-23 would be selected as active.

### Tip

The first active orbital always starts with '(number of electrons - number of active electrons )/2'.

Let us see what these orbitals are before starting the full CASSCF optimization:

```
! SV(P)
! moread noiter
%moinp "Test-CASSCF-Benzene-2.mp2nat"

# Enable reduced orbital population per MO
%Output
Print[P_ReducedOrbPopMO_L] 1
End
...
```

Leading to:

```
*****
* LOEWDIN POPULATION ANALYSIS *
*****
```

```
**** WARNING: LOEWDIN FINDS 41.4430078 ELECTRONS INSTEAD OF 42 ****
**** SKIPPING LOEWDIN ANALYSIS ****
```

```
-----
LOEWDIN REDUCED ORBITAL POPULATIONS PER MO
-----
```

```
THRESHOLD FOR PRINTING IS 0.1%
```

```
...
```

	18	19	20	21	22	23
	-0.49831	-0.33935	-0.33935	0.13474	0.13474	0.18198
	1.95331	1.91853	1.91853	0.06544	0.06544	0.02821
	-----	-----	-----	-----	-----	-----
0 C pz	16.5	8.1	24.4	7.6	22.9	16.2
0 C dxz	0.0	0.4	0.2	1.4	0.7	0.1
0 C dyz	0.1	0.2	0.0	0.7	0.0	0.3
1 C pz	16.5	8.1	24.4	7.6	22.9	16.2
1 C dxz	0.0	0.4	0.2	1.4	0.7	0.1
1 C dyz	0.1	0.2	0.0	0.7	0.0	0.3
2 C pz	16.5	32.5	0.0	30.5	0.0	16.2
2 C dxz	0.1	0.1	0.0	0.1	0.0	0.4
2 C dyz	0.0	0.0	0.8	0.0	2.8	0.0
3 C pz	16.5	8.1	24.4	7.6	22.9	16.2
3 C dxz	0.0	0.4	0.2	1.4	0.7	0.1
3 C dyz	0.1	0.2	0.0	0.7	0.0	0.3
4 C pz	16.5	8.1	24.4	7.6	22.9	16.2
4 C dxz	0.0	0.4	0.2	1.4	0.7	0.1
4 C dyz	0.1	0.2	0.0	0.7	0.0	0.3
5 C pz	16.5	32.5	0.0	30.5	0.0	16.2
5 C dxz	0.1	0.1	0.0	0.1	0.0	0.4
5 C dyz	0.0	0.0	0.8	0.0	2.8	0.0

This shows us that these six orbitals are precisely the  $\pi/\pi^*$ -orbitals that we wanted to have active (you can also plot them to get even more insight). If the orbitals wouldn't be in the correct position to enter as active (18-23), orbitals can be swapped with the `rotate` keyword described [elsewhere](#).

Now that we know that the desired orbitals are in the correct order, we can optimize the orbitals with CASSCF:

```
! SV(P)
! moread
%moinp "Test-CASSCF-Benzene-2.mrci.nat"

%casscf
# define CAS(6,6)
nel 6
norb 6

# ground state singlet
mult 1
nroots 1

switchstep nr # for illustration purpose
end
```

To highlight the feature `SwitchStep` of the CASSCF program, we employ the Newton-Raphson method (NR) after a certain convergence has been reached (`SwitchStep NR` statement). In general, it is recommended to use the default convergence settings! The output of the CASSCF program is:

```
-----
CAS-SCF ITERATIONS
```

(continues on next page)

(continued from previous page)

```
MACRO-ITERATION    1:
--- Inactive Energy E0 = -224.09725414 Eh
CI-ITERATION      0:
-230.588253032    0.000000000000 (   0.00)
CI-PROBLEM SOLVED
DENSITIES MADE

<<<<<<<<<<<<<<<<INITIAL CI STATE CHECK>>>>>>>>>>>>>>>>

BLOCK   1 MULT= 1 NROOTS= 1
ROOT    0:  E=   -230.5882530315 Eh
0.89482 [   0]: 222000
0.02897 [  14]: 211110
0.01982 [  29]: 202020
0.01977 [   4]: 220200
0.01177 [  65]: 112011
0.01169 [  50]: 121101

<<<<<<<<<<<<<<<<INITIAL CI STATE CHECK>>>>>>>>>>>>>>>>

E(CAS)=  -230.588253032 Eh DE=    0.000000e+00
--- Energy gap subspaces: Ext-Act = 0.195   Act-Int = 0.127
--- current l-shift: Up(Ext-Act) = 1.40   Dn(Act-Int) = 1.47
N(occ)=   1.96393 1.90933 1.90956 0.09190 0.09208 0.03319
||g|| =    1.046979e-01 Max(G)=   -4.638985e-02 Rot=53,19
--- Orbital Update [SuperCI(PT)]
--- Canonicalize Internal Space
--- Canonicalize External Space
--- SX_PT (Skipped TA=0 IT=0): ||X|| =          0.063973050 Max(X) (83,23) =   -0.
↳ 035491133
--- SFit(Active Orbitals)

MACRO-ITERATION    2:
--- Inactive Energy E0 = -224.09299157 Eh
CI-ITERATION      0:
-230.590141151    0.000000000000 (   0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)=  -230.590141151 Eh DE=   -1.888119e-03
--- Energy gap subspaces: Ext-Act = 0.202   Act-Int = 0.126
--- current l-shift: Up(Ext-Act) = 0.90   Dn(Act-Int) = 0.97
N(occ)=   1.96182 1.90357 1.90364 0.09771 0.09777 0.03549
||g|| =    2.971340e-02 Max(G)=   -8.643429e-03 Rot=52,20
--- Orbital Update [SuperCI(PT)]
--- Canonicalize Internal Space
--- Canonicalize External Space
--- SX_PT (Skipped TA=0 IT=0): ||X|| =          0.009811159 Max(X) (67,21) =   -0.
↳ 003665750
--- SFit(Active Orbitals)

MACRO-ITERATION    3:
===>>> Convergence to 3.0e-02 achieved - switching to Step=NR
--- Inactive Energy E0 = -224.07872151 Eh
CI-ITERATION      0:
-230.590260496    0.000000000000 (   0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)=  -230.590260496 Eh DE=   -1.193453e-04
```

(continues on next page)

(continued from previous page)

```

--- Energy gap subspaces: Ext-Act = 0.203   Act-Int = 0.125
--- current l-shift: Up(Ext-Act) = 0.73   Dn(Act-Int) = 0.81
N(occ)=  1.96145 1.90275 1.90278 0.09856 0.09857 0.03589
||g|| =    8.761362e-03 Max(G)=    4.388664e-03 Rot=43,19
--- Orbital Update [      NR]
AUGHESS-ITER   0: E=   -0.000016434 <r|r>= 2.70127912e-05
AUGHESS-ITER   1: E=   -0.000021148 <r|r>= 2.91399830e-06
AUGHESS-ITER   2: E=   -0.000021780 <r|r>= 4.01336069e-07 => CONVERGED
DE(predicted)=  -0.000010890 First Element= 0.999987718
<X(rot)|X(rot)>=    0.000024564
--- SFit(Active Orbitals)

MACRO-ITERATION   4:
--- Inactive Energy E0 = -224.07787812 Eh
CI-ITERATION      0:
-230.590271490   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)= -230.590271490 Eh DE=   -1.099363e-05
--- Energy gap subspaces: Ext-Act = 0.202   Act-Int = 0.125
--- current l-shift: Up(Ext-Act) = 0.40   Dn(Act-Int) = 0.47
N(occ)=  1.96135 1.90267 1.90267 0.09866 0.09866 0.03599
||g|| =    6.216730e-04 Max(G)=    1.417079e-04 Rot=66,13
---- THE CAS-SCF GRADIENT HAS CONVERGED ----
--- FINALIZING ORBITALS ---
---- DOING ONE FINAL ITERATION FOR PRINTING ----
--- Forming Natural Orbitals
--- Canonicalize Internal Space
--- Canonicalize External Space

MACRO-ITERATION   5:
--- Inactive Energy E0 = -224.07787811 Eh
--- All densities will be recomputed
CI-ITERATION      0:
-230.590271485   0.000000000000 (    0.00)
CI-PROBLEM SOLVED
DENSITIES MADE
E(CAS)= -230.590271485 Eh DE=    5.179942e-09
--- Energy gap subspaces: Ext-Act = -0.242   Act-Int = -0.002
--- current l-shift: Up(Ext-Act) = 0.84   Dn(Act-Int) = 0.60
N(occ)=  1.96135 1.90267 1.90267 0.09866 0.09866 0.03599
||g|| =    6.216710e-04 Max(G)=    1.544017e-04 Rot=29,12
-----
CASSCF RESULTS
-----

Final CASSCF energy      : -230.590271485 Eh   -6274.6803 eV

```

First of all, we see how the program cycles between CI-vector optimization and orbital optimization steps (so-called unfolded two-step procedure). After 3 iterations, the program switches to the Newton-Raphson solver which then converges very rapidly. Orbital optimization with the Newton-Raphson solver is limited to smaller sized molecules, as the program produces lengthy integrals and Hessian files. In the majority of situations the default converger (SuperCI(PT)) is the preferred choice.<sup>[430]</sup> In difficult situations, one should consider the *TRAH optimizer*.

### 3.13.7 Guess: Atomic Valence Active Space (AVAS)

Very good starting orbitals that are targeted to a specific user-given active space can be generated with the Atomic Valence Active Space (AVAS) procedure. [449, 450] The general idea is that the user provides a set of atomic orbitals (AO) of a minimal basis set that are sufficient to qualitatively represent the final CASSCF active orbitals. Typical examples are

- $p_z$  orbitals of a  $\pi$  system chromophore in a molecule
- five valence (or 10 double-shell) d orbitals of a transition-metal (TM) atom in a molecule
- seven valence (or 14 double-shell) f orbitals of a lanthanide or actinide atom in a molecule

Then, by the help of linear algebra (singular-value decomposition) AVAS rotates the starting molecular orbitals (MOs) such that they have maximum overlap with the target AOs. With those rotated MOs that have a sufficiently large singular value ( $> 0.4$  (default)) are considered as active orbitals. In that manner, AVAS automatically determines an active space, i.e. the number of active orbitals and electrons, that is now specified by the target AOs.

#### Warning

AVAS overwrites the number of active orbitals and electrons in the %casscf block!

As a first example, we now consider  $\text{CuCl}_4^-$  in a minimal active space

```
! cc-pvtz TightSCF Def2/JK PModel NoIter
! AVAS (Valence-D)

%maxcore 3000

%paras
  cucl = 2.291
end

* int -1 1
Cu 0 0 0 0.0 0.0 0.0
Cl 1 0 0 {cucl} 0.0 0.0
Cl 1 2 0 {cucl} 90.0 0.0
Cl 1 3 2 {cucl} 90.0 180.
Cl 1 4 3 {cucl} 90.0 180.
*
```

The keyword `! AVAS (Valence-D)` seeks for all transition-metal atoms in the molecule and inserts a single minimal d basis function for each TM atom. All five component  $M_L$  of the basis function are then considered. The AVAS procedure prints singular / eigen values for the occupied and virtual orbital space and easily finds the desired minimal active space CAS(9,5).

```
-----
INITIAL GUESS: Atomic Valence Active space (AVAS)
-----
```

```
AVAS threshold      : 0.400000
AVAS minimal basis set : MINAO
AVAS list           : Shell | 3 2 0> at atom 0 (system 0)
\\:\\: Shell | 3 2 1> at atom 0 (system 0)
\\:\\: Shell | 3 2 -1> at atom 0 (system 0)
\\:\\: Shell | 3 2 2> at atom 0 (system 0)
\\:\\: Shell | 3 2 -2> at atom 0 (system 0)

AVAS P matrix eig. val ( Occupied) : 0.966698
AVAS P matrix eig. val ( Occupied) : 0.974913
AVAS P matrix eig. val ( Occupied) : 0.977443
```

(continues on next page)

(continued from previous page)

```

AVAS P matrix eig. val ( Occupied) : 0.977443
AVAS P matrix eig. val ( Occupied) : 0.985233
AVAS P matrix eig. val ( Virtual) : (0.032829)
AVAS P matrix eig. val ( Virtual) : (0.024687)
AVAS P matrix eig. val ( Virtual) : (0.022047)
AVAS P matrix eig. val ( Virtual) : (0.022047)
AVAS P matrix eig. val ( Virtual) : (0.014546)

```

```

AVAS electrons      : 9
AVAS orbitals       : 5

```

The five initial active orbitals after being processed by AVAS indeed look like the desired Cu d-orbitals.

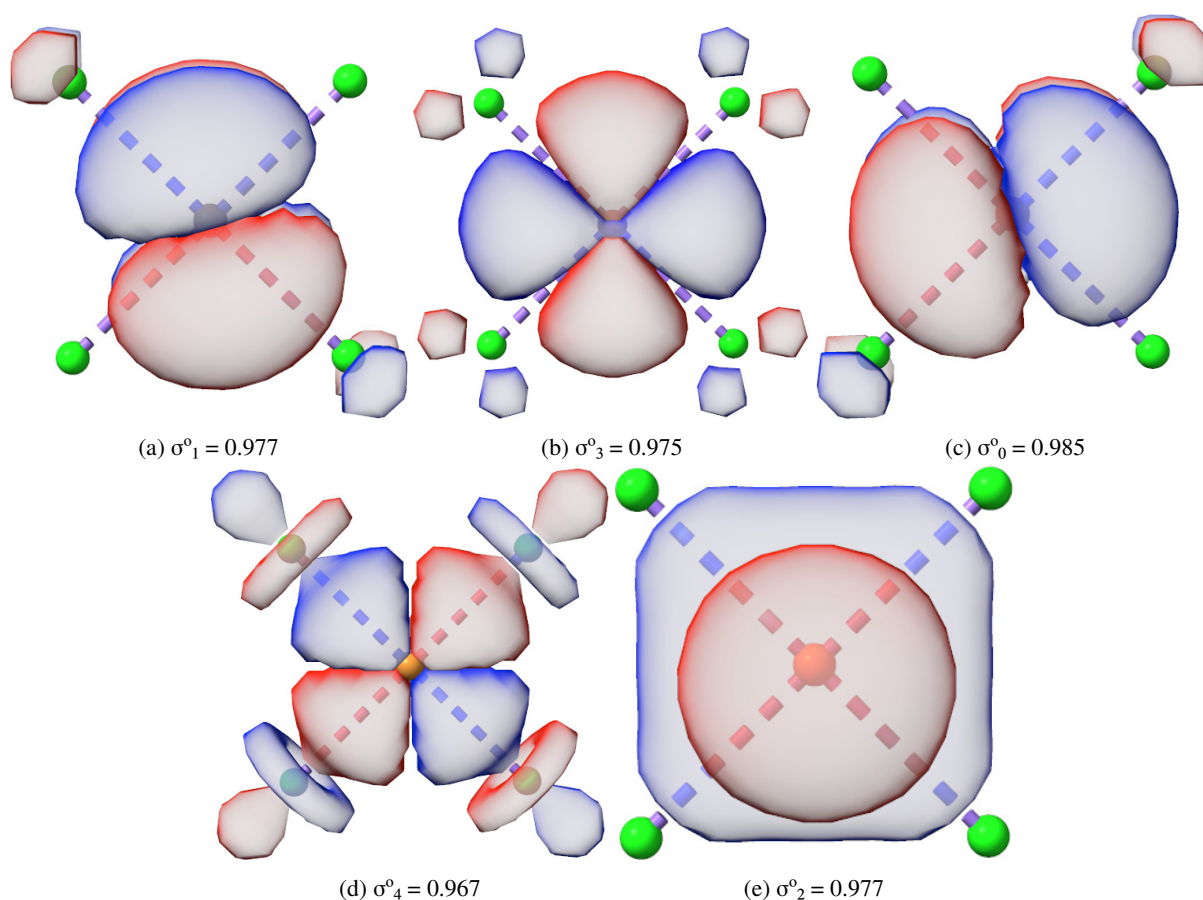


Fig. 3.14: Initial Minimal AS orbitals of  $\text{CuCl}_4^-$  generated by AVAS.

The same calculation can be started also by using the `%scf avas ... end end` block.

```

%scf
avas
  system
    shell  3,  3,  3,  3,  3
    l      2,  2,  2,  2,  2
    m_l    0,  1, -1,  2, -2
    center 0,  0,  0,  0,  0
  end
end
end

```

Here, it is also possible to use target basis functions at different atoms (`center`) and to select only a subset of

functions in a shell ( $m_l$ ). Note that if not all functions of a shell (3p, 5d, 7f) are selected, the molecule should be oriented manually to accomplish the desired basis function overlap.

AVAS can be also used very conveniently in the same fashion for double-d shell calculations with transition-metal complexes (`! AVAS (Double-D)`). For each 3d transition-metal center in a molecule all 3d and 4d target functions are considered. Similarly, double-shell active spaces can be also set up for 4d and 5d transition-metal complexes.

There is also a similar keyword for lanthanides and actinides. `! AVAS (Valence-F)` attempts to set up an active space with 7 f functions for each lanthanide or actinide atom in a molecule. There is also the possibility to run double-f shell calculations using the `! AVAS (Double-F)` keyword.

To avoid this issue for  $\pi$  active space calculation, all three 2p target AOs are considered first but they are weighted by the three component of the principle axis of inertia with the largest moment. [450] For those inertia moment calculations, masses are ignored and only the centers of the desired target p AO are considered.

For a CAS(10,9)  $\pi$ -active space calculation on tryptophan, the AVAS input read

```
! cc-pVTZ PModel NoIter

%scf
  avas
    tol 0.4
    system
      center 0, 1, 2, 3, 4, 5, 10, 11, 12
      type pz, pz, pz, pz, pz, pz, pz, pz, pz
    end
  end
end

* xyz 0 1
C      0.4512549872      2.3796411953      0.0577773122
C      0.1094760583      1.0035547288     -0.1566676092
C      1.7801675822      2.8072137170      0.2571892289
C      2.7806901872      1.8262977582      0.2356692574
C      2.4656511421      0.4546661378      0.0230301052
C      1.1452272475      0.0339609344     -0.1736410480
H      2.0259509453      3.8615102004      0.4187388370
H      3.8237760997      2.1214062744      0.3833595222
H      3.2752609035     -0.2812140701      0.0109117373
H      0.9203743659     -1.0244858148     -0.3388820373
C     -1.3215206968      0.9316755285     -0.3177965838
C     -1.7965156128      2.2386249398     -0.2022300378
N     -0.7296902726      3.0958334808      0.0227806512
H     -0.8107596679      4.0971334562      0.1485860796
H     -2.8167763088      2.6080109542     -0.2688439980
C     -2.1029028025     -0.3291635000     -0.5688909937
C     -3.4238543678     -0.4065989881      0.2267575199
H     -1.4745479852     -1.1909157350     -0.2884954137
H     -2.3461010681     -0.4478113906     -1.6421457333
C     -3.9423325138     -1.8379287141      0.1258142785
N     -4.3742952299      0.5836598444     -0.2812451173
H     -3.2051519657     -0.1892488262      1.2846794690
O     -3.2924970778     -2.6708957465      0.9924074621
O     -4.8043368378     -2.2232843366     -0.6488988164
H     -3.6480373076     -3.5631013900      0.8277551234
H     -5.2270970579      0.5578136152      0.2816027849
H     -4.6658127461      0.2911757460     -1.2180819802
*
```

and leads to the following output

```
-----
INITIAL GUESS: Atomic Valence Active space (AVAS)
```

(continues on next page)



(continued from previous page)

```

-----
AVAS threshold      : 0.400000
AVAS minimal basis set : AUTO
AVAS list           : Shell | 2 1  0> at atom 0 (system 0, type pz)
                    : Shell | 2 1  1> at atom 0 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 0 (system 0, type pz)
                    : Shell | 2 1  0> at atom 1 (system 0, type pz)
                    : Shell | 2 1  1> at atom 1 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 1 (system 0, type pz)
                    : Shell | 2 1  0> at atom 2 (system 0, type pz)
                    : Shell | 2 1  1> at atom 2 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 2 (system 0, type pz)
                    : Shell | 2 1  0> at atom 3 (system 0, type pz)
                    : Shell | 2 1  1> at atom 3 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 3 (system 0, type pz)
                    : Shell | 2 1  0> at atom 4 (system 0, type pz)
                    : Shell | 2 1  1> at atom 4 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 4 (system 0, type pz)
                    : Shell | 2 1  0> at atom 5 (system 0, type pz)
                    : Shell | 2 1  1> at atom 5 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 5 (system 0, type pz)
                    : Shell | 2 1  0> at atom 10 (system 0, type pz)
                    : Shell | 2 1  1> at atom 10 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 10 (system 0, type pz)
                    : Shell | 2 1  0> at atom 11 (system 0, type pz)
                    : Shell | 2 1  1> at atom 11 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 11 (system 0, type pz)
                    : Shell | 2 1  0> at atom 12 (system 0, type pz)
                    : Shell | 2 1  1> at atom 12 (system 0, type pz)
                    : Shell | 2 1 -1> at atom 12 (system 0, type pz)

AVAS P matrix eig. val ( Occupied) : (0.000004)
AVAS P matrix eig. val ( Occupied) : (0.000014)
AVAS P matrix eig. val ( Occupied) : (0.000292)
AVAS P matrix eig. val ( Occupied) : (0.040014)
AVAS P matrix eig. val ( Occupied) : 0.978162
AVAS P matrix eig. val ( Occupied) : 0.986637
AVAS P matrix eig. val ( Occupied) : 0.993225
AVAS P matrix eig. val ( Occupied) : 0.994300
AVAS P matrix eig. val ( Occupied) : 0.996447
AVAS P matrix eig. val ( Virtual) : 0.999996
AVAS P matrix eig. val ( Virtual) : 0.999986
AVAS P matrix eig. val ( Virtual) : 0.999708
AVAS P matrix eig. val ( Virtual) : 0.959986
AVAS P matrix eig. val ( Virtual) : (0.021838)
AVAS P matrix eig. val ( Virtual) : (0.013363)
AVAS P matrix eig. val ( Virtual) : (0.006775)
AVAS P matrix eig. val ( Virtual) : (0.005700)
AVAS P matrix eig. val ( Virtual) : (0.003553)

AVAS electrons      : 10
AVAS orbitals       : 9

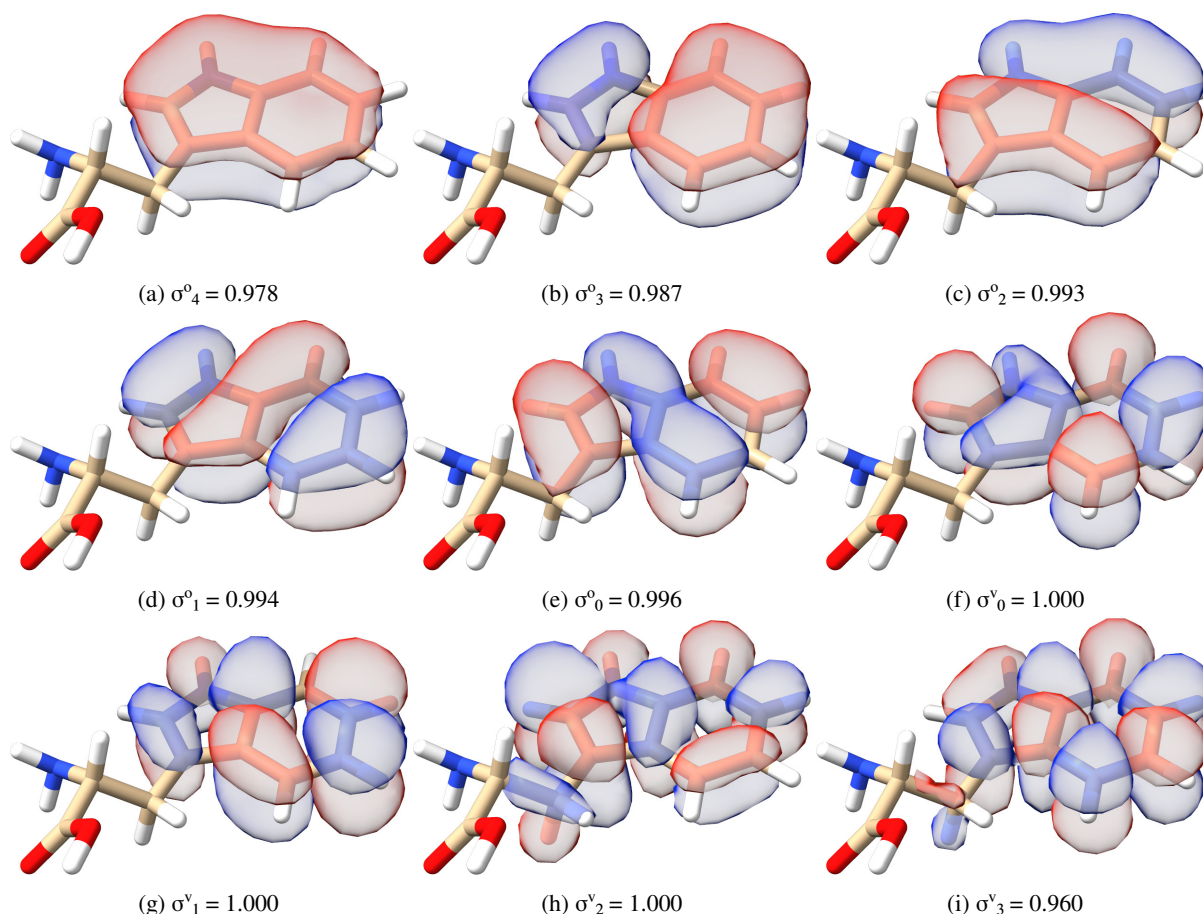
```

```

-----
INITIAL GUESS DONE ( 0.3 sec)
-----

```

and initial active orbitals.

Fig. 3.15: Initial  $\pi$  AS orbitals of tryptophan generated by AVAS.

It is also possible to specify the number of active electrons `nel` and orbitals `norb` directly. For such a calculation, the AVAS singular value decomposition threshold `tol` is ignored. In the following calculation, the strongly occupied orbital from the previous CAS(10,9) ( $\sigma^0_4$  in Fig. 3.15) calculation is omitted.

```
%scf
avas
  system
  norb 8
  nel 8
  center 0, 1, 2, 3, 4, 5, 10, 11, 12
  type pz, pz, pz, pz, pz, pz, pz, pz, pz
end
end
end
```

It is also possible to do the AVAS start MO generation for several `systems` independently and then re-orthonormalize all MOs at the end similar to [450]. This becomes interesting for generating starting orbitals for multiple  $\pi$  chromophores like the bridged bithiophene biradical

```
%scf
avas
  system
  norb 4
  nel 3
  center 0, 1, 2, 3, 4 # C / S atoms system 1
  type pz, pz, pz, pz, pz
end
system
```

(continues on next page)

(continued from previous page)

```

norb 4
nel 5
center 38, 39, 40, 41, 43 # C / S atoms system 2
type pz, pz, pz, pz, pz
end
end
end

```

or the FeTPP molecule.

```

! SVP NoIter
! PModel

%scf
avas
system
center 0
type d
end
system
center 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, ↵
↵19, 20, 29, 30, 31, 32
type pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, pz, ↵
↵pz, pz, pz, pz, pz, pz
end
end
end
end

*xyz 0 1
Fe 0.0000 0.0000 0.0000
N 1.9764 0.0000 0.0000
N 0.0000 0.0000 1.9884
N -1.9764 0.0000 0.0000
N 0.0000 0.0000 -1.9884
C 2.8182 0.0000 -1.0903
C 2.8182 0.0000 1.0903
C 1.0918 0.0000 2.8249
C -1.0918 0.0000 2.8249
C -2.8182 0.0000 1.0903
C -2.8182 0.0000 -1.0903
C -1.0918 0.0000 -2.8249
C 1.0918 0.0000 -2.8249
C 4.1961 0.0000 -0.6773
C 4.1961 0.0000 0.6773
C 0.6825 0.0000 4.1912
C -0.6825 0.0000 4.1912
C -4.1961 0.0000 0.6773
C -4.1961 0.0000 -0.6773
C -0.6825 0.0000 -4.1912
C 0.6825 0.0000 -4.1912
H 5.0441 0.0000 -1.3538
H 5.0441 0.0000 1.3538
H 1.3558 0.0000 5.0416
H -1.3558 0.0000 5.0416
H -5.0441 0.0000 1.3538
H -5.0441 0.0000 -1.3538
H -1.3558 0.0000 -5.0416
H 1.3558 0.0000 -5.0416
C 2.4150 0.0000 2.4083
C -2.4150 0.0000 2.4083
C -2.4150 0.0000 -2.4083

```

(continues on next page)

(continued from previous page)

C	2.4150	0.0000	-2.4083
H	3.1855	0.0000	3.1752
H	-3.1855	0.0000	3.1752
H	-3.1855	0.0000	-3.1752
H	3.1855	0.0000	-3.1752
*			

For those systems, all AVAS starting MOs have the desired  $\pi$  or  $d$  character as illustrated for the “active frontier orbitals” in Fig. 3.16.

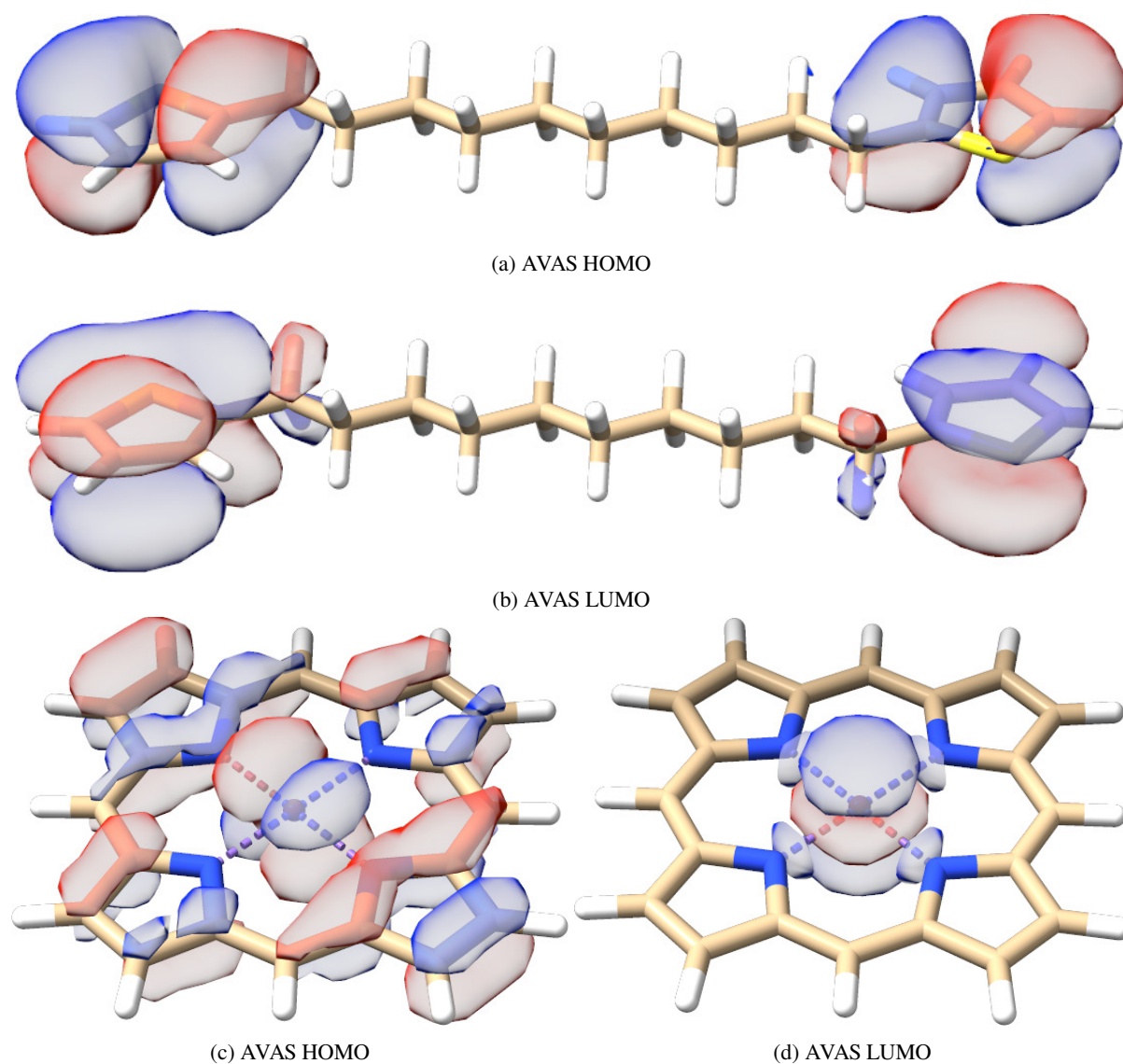


Fig. 3.16: Initial HOMO and LUMO AVAS orbitals of a bridged bithiophene biradical and FeTPP.

### 3.13.8 Example: Symmetry

The CASSCF program can make some use of symmetry. Thus, it is possible to do the CI calculations separated by irreducible representations. This allows one to calculate electronic states in a more controlled fashion.

Let us look at a simple example:  $C_2H_4$ . We first generate symmetry adapted MP2 natural orbitals. Since we opt for initial guess orbitals, the computationally cheaper unrelaxed density suffices:

```
! def2-TZVP def2-TZVP/C UseSym RI-MP2 conv # conventional is faster for small_
↪molecules

%mp2
density unrelaxed # unrelaxed is sufficient for guess orbitals.
natorbs true
end

* int 0 1
C 0 0 0 0 0 0
C 1 0 0 1.35 0 0
H 1 2 0 1.1 120 0
H 1 2 3 1.1 120 180
H 2 1 3 1.1 120 0
H 2 1 3 1.1 120 180
*
```

The program does the following. It first identifies the group correctly as  $D_{2h}$  and sets up its irreducible representations. The process detects symmetry within SymThresh ( $10^{-4}$ ) and purifies the geometry thereafter:

```
-----
SYMMETRY DETECTION
-----

The point group will now be determined using a tolerance of 1.0000e-04.
Splitting atom subsets according to nuclear charge, mass and basis set.
Splitting atom subsets according to distance from the molecule's center.
Identifying relative distance patterns of the atoms.
Splitting atom subsets according to atoms' relative distance patterns.
Bring atoms of each subset into input order.
The molecule is planar.
The molecule has a center of inversion.
Analyzing the first atom subset for its symmetry.
The atoms in the selected subset form a 4-gon with alternating side lengths.
Testing point group D2h.
Success!
This point group has been found:      D2h
Largest non-degenerate subgroup:     D2h

Symmetry-perfected Cartesians (point group D2h):

Atom          Symmetry-perfected Cartesians (x, y, z; au)
0      -1.275565140397    0.000000000000    0.000000000000
1       1.275565140397    0.000000000000    0.000000000000
2      -2.314914514054    1.800205921988    0.000000000000
3      -2.314914514054   -1.800205921988    0.000000000000
4       2.314914514054    1.800205921988    0.000000000000
5       2.314914514054   -1.800205921988    0.000000000000

-----
SYMMETRY-PERFECTED CARTESIAN COORDINATES (A.U.)
-----

Warning (ORCA_SYM): Coordinates were not cleaned so far!
```

(continues on next page)

(continued from previous page)

-----  
SYMMETRY REDUCTION  
-----

ORCA supports only abelian point groups.  
 It is now checked, if the determined point group is supported:  
 Point Group ( D2h ) is ... supported

(Re)building abelian point group:  
 Creating Character Table ... done  
 Making direct product table ... done  
 Constructing symmetry operations ... done  
 Creating atom transfer table ... done  
 Creating asymmetric unit ... done

-----  
ASYMMETRIC UNIT IN D2h  
-----

#	AT	MASS	COORDS (A.U.)		BAS
0	C	12.0110	-1.27556514	0.00000000	0
2	H	1.0080	-2.31491451	1.80020592	0

-----  
SYMMETRY ADAPTED BASIS  
-----

The coefficients for the symmetry adapted linear combinations (SALCS)  
 of basis functions will now be computed:

Number of basis functions ... 86  
 Preparing memory ... done  
 Constructing Gamma(red) ... done  
 Reducing Gamma(red) ... done  
 Constructing SALCs ... done  
 Checking SALC integrity ... nothing suspicious  
 Normalizing SALCs ... done

Storing the symmetry object:  
 Symmetry file ... Test-SYM-CAS-C2H4-1.sym.tmp  
 Writing symmetry information ... done

It then performs the SCF calculation and keeps the symmetry in the molecular orbitals.

NO	OCC	E (Eh)	E (eV)	Irrep
0	2.0000	-11.236728	-305.7669	1-Ag
1	2.0000	-11.235157	-305.7242	1-B3u
2	2.0000	-1.027144	-27.9500	2-Ag
3	2.0000	-0.784021	-21.3343	2-B3u
4	2.0000	-0.641566	-17.4579	1-B2u
5	2.0000	-0.575842	-15.6694	3-Ag
6	2.0000	-0.508313	-13.8319	1-B1g
7	2.0000	-0.373406	-10.1609	1-B1u
8	0.0000	0.139580	3.7982	1-B2g
9	0.0000	0.171982	4.6799	4-Ag
10	0.0000	0.195186	5.3113	3-B3u
11	0.0000	0.196786	5.3548	2-B2u
12	0.0000	0.242832	6.6078	2-B1g
13	0.0000	0.300191	8.1686	5-Ag
14	0.0000	0.326339	8.8801	4-B3u
...	etc			

The MP2 module does not take any advantage of this information but produces natural orbitals that are symmetry adapted:

```

N[ 0] (B3u) = 2.00000360
N[ 1] ( Ag) = 2.00000219
N[ 2] ( Ag) = 1.98056435
N[ 3] (B3u) = 1.97195041
N[ 4] (B2u) = 1.96746753
N[ 5] (B1g) = 1.96578954
N[ 6] ( Ag) = 1.95864726
N[ 7] (B1u) = 1.93107098
N[ 8] (B2g) = 0.04702701
N[ 9] (B3u) = 0.02071784
N[10] (B2u) = 0.01727252
N[11] ( Ag) = 0.01651489
N[12] (B1g) = 0.01602695
N[13] (B3u) = 0.01443373
N[14] (B1u) = 0.01164204
N[15] ( Ag) = 0.01008617
N[16] (B2u) = 0.00999302
N[17] ( Ag) = 0.00840326
N[18] (B3g) = 0.00795053
N[19] (B3u) = 0.00532044
N[20] ( Au) = 0.00450556
etc.

```

From this information and visual inspection you will know what orbitals you will have in the active space:

These natural orbitals can then be fed into the CASSCF calculation. We perform a simple calculation in which we keep the ground state singlet ( $A_{1g}$  symmetry, irrep=0) and the first excited triplet state ( $B_{3u}$  symmetry, irrep=7). In general the ordering of irreps follows standard conventions and in case of doubt you will find the relevant number for each irrep in the output.

For example, here (using !LargePrint):

```

-----
CHARACTER TABLE OF GROUP D2h
-----
GAMMA  O1  O2  O3  O4  O5  O6  O7  O8
Ag :   1.0  1.0  1.0  1.0  1.0  1.0  1.0  1.0
B1g:   1.0  1.0 -1.0 -1.0  1.0  1.0 -1.0 -1.0
B2g:   1.0 -1.0  1.0 -1.0  1.0 -1.0  1.0 -1.0
B3g:   1.0 -1.0 -1.0  1.0  1.0 -1.0 -1.0  1.0
Au :   1.0  1.0  1.0  1.0 -1.0 -1.0 -1.0 -1.0
B1u:   1.0  1.0 -1.0 -1.0 -1.0 -1.0  1.0  1.0
B2u:   1.0 -1.0  1.0 -1.0 -1.0  1.0 -1.0  1.0
B3u:   1.0 -1.0 -1.0  1.0 -1.0  1.0  1.0 -1.0

```

```

-----
DIRECT PRODUCT TABLE OF GROUP D2h
-----
**      Ag  B1g B2g B3g Au  B1u B2u B3u

Ag      Ag  B1g B2g B3g Au  B1u B2u B3u
B1g     B1g Ag  B3g B2g B1u Au  B3u B2u
B2g     B2g B3g Ag  B1g B2u B3u Au  B1u
B3g     B3g B2g B1g Ag  B3u B2u B1u Au
Au      Au  B1u B2u B3u Ag  B1g B2g B3g
B1u     B1u Au  B3u B2u B1g Ag  B3g B2g
B2u     B2u B3u Au  B1u B2g B3g Ag  B1g
B3u     B3u B2u B1u Au  B3g B2g B1g Ag

```

We use the following input for CASSCF, where we tightened the integral cut-offs and the convergence criteria using !VeryTightSCF.

```

! def2-TZVP  Conv UseSym
! moread
%moinp "Test-SYM-CAS-C2H4-1.mp2nat"
%casscf
# define CAS(4,4)
nel          4
norb         4

# This is only here to show that NR can also be used from
# the start with orbstep
orbstep      nr
switchstep   nr

# The lowest singlet and triplet states. The new feature
# is the array "irrep" that lets you give the irrep for
# a given block. Thus, now you can have several blocks of
# the same multiplicity but different spatial symmetry
irrep        0,7
mult         1,3
nroots       1,1
end

* int 0 1
C 0 0 0 0 0 0 0
C 1 0 0 1.35 0 0
H 1 2 0 1.1 120 0
H 1 2 3 1.1 120 180
H 2 1 3 1.1 120 0
H 2 1 3 1.1 120 180
*
```

And gives:

```

-----
SCF SETTINGS
-----
Hamiltonian:
  Ab initio Hamiltonian  Method          .... Hartree-Fock (GTOs)

General Settings:
  Integral files         IntName          .... Test-SYM-CAS-C2H4-1
  Hartree-Fock type      HFTyp            .... CASSCF
  Total Charge           Charge           .... 0
  Multiplicity           Mult             .... 1
  Number of Electrons    NEL              .... 16
  Basis Dimension        Dim              .... 86
  Nuclear Repulsion      ENuc             .... 32.9609050695 Eh

Symmetry handling       UseSym            .... ON
Point group              .... D2h
Used point group         .... D2h
Number of irreps         .... 8
  Irrep  Ag has 19 symmetry adapted basis functions (ofs= 0)
  Irrep  B1g has 12 symmetry adapted basis functions (ofs= 19)
  Irrep  B2g has 8 symmetry adapted basis functions (ofs= 31)
  Irrep  B3g has 4 symmetry adapted basis functions (ofs= 39)
  Irrep  Au has 4 symmetry adapted basis functions (ofs= 43)
  Irrep  B1u has 8 symmetry adapted basis functions (ofs= 47)
  Irrep  B2u has 12 symmetry adapted basis functions (ofs= 55)
  Irrep  B3u has 19 symmetry adapted basis functions (ofs= 67)
```

And further in the CASSCF program:



```

Symmetry handling      UseSym      ... ON
Point group           ... D2h
Used point group       ... D2h
Number of irreps      ... 8
  Irrep   Ag has    19 SALCs (ofs=  0) #(closed)=  2 #(active)=  1
  Irrep  B1g has    12 SALCs (ofs= 19) #(closed)=  1 #(active)=  0
  Irrep  B2g has     8 SALCs (ofs= 31) #(closed)=  0 #(active)=  1
  Irrep  B3g has     4 SALCs (ofs= 39) #(closed)=  0 #(active)=  0
  Irrep   Au has     4 SALCs (ofs= 43) #(closed)=  0 #(active)=  0
  Irrep  B1u has     8 SALCs (ofs= 47) #(closed)=  0 #(active)=  1
  Irrep  B2u has    12 SALCs (ofs= 55) #(closed)=  1 #(active)=  0
  Irrep  B3u has    19 SALCs (ofs= 67) #(closed)=  2 #(active)=  1
Symmetries of active orbitals:
MO =  6  IRREP= 0 (Ag)
MO =  7  IRREP= 5 (B1u)
MO =  8  IRREP= 2 (B2g)
MO =  9  IRREP= 7 (B3u)

Setting up the integral package      ... done
Building the CAS space               ... done (7 configurations for Mult=1,
↪ Irrep=0)
Building the CAS space               ... done (4 configurations for Mult=3,
↪ Irrep=7)

```

Note that the irrep occupations and active space irreps will be frozen to what they are upon entering the CASSCF program. This helps to setup the CI problem.

After which it smoothly converges to give:

6:	1.986258	-0.753012	-20.4905	3-Ag
7:	1.457849	-0.291201	-7.9240	1-B1u
8:	0.541977	0.100890	2.7454	1-B2g
9:	0.013915	0.964186	26.2368	3-B3u

As well as:

```

-----
SA-CASSCF TRANSITION ENERGIES
-----

```

```

LOWEST ROOT =      -78.110314788 Eh -2125.490 eV

```

STATE	ROOT	MULT	IRREP	DE/a.u.	DE/eV	DE/cm** <sup>-1</sup>
1:	0	3	B3u	0.163741	4.456	35937.1

### 3.13.9 Example: Breaking Chemical Bonds

Let us turn to the breaking of chemical bonds. As a first example we study the dissociation of the H<sub>2</sub> molecule. Scanning a bond, we have two potential setups for the calculation: a) scan from the inside to the outside or b) from the outside to inside. Of course both setups yield identical results, but they differ in practical aspects i.e. convergence properties. In general, **scanning from the outside to the inside is the recommended procedure**. Using the default guess (PModel), starting orbitals are much easier identified than at shorter distances, where the antibonding orbitals are probably 'impure' and hence would require some additional preparation. To ensure a smooth potential energy surface, in all subsequent geometry steps, ORCA reads the converged CASSCF orbitals from the previous geometry step. In the following, `TightSCF` is used to tighten the convergence settings of CASSCF.

```
!Def2-SVP TightSCF
```

```
%casscf
# define CAS(2,2)
```

(continues on next page)

(continued from previous page)

```
nel          2
norb         2
# singlet ground state
mult         1
nroots       1
end

# Scanning from the outside to the inside
%paras
R [4.1 3.8 3.5 3.2 2.9 2.6 2.4 2.2
   2 1.7 1.5 1.3 1.1 1 0.9 0.8
   0.75 0.7 0.65 0.6]
end

* xyz 0 1
H   0.0   0.0   0.0
H   0.0   0.0   {R}
end
```

The resulting potential energy surface (PES) is depicted in [Fig. 3.17](#) together with PESs obtained from RHF and broken-symmetry UHF calculations (input below).

```
! RHF Def2-SVP TightSCF

# etc...
```

And

```
! UHF Def2-SVP TightSCF

%scf
  FlipSpin 1
  FinalMs 0.0
end

# etc...
```

#### **Note**

The `FlipSpin` option does not work together with the parameter scan. Only the first structure will undergo a spin flip. Therefore, at the current status, a separate input file (including the coordinates or with a corresponding coordinate file) has to be provided for each structure that is scanned along the PES.

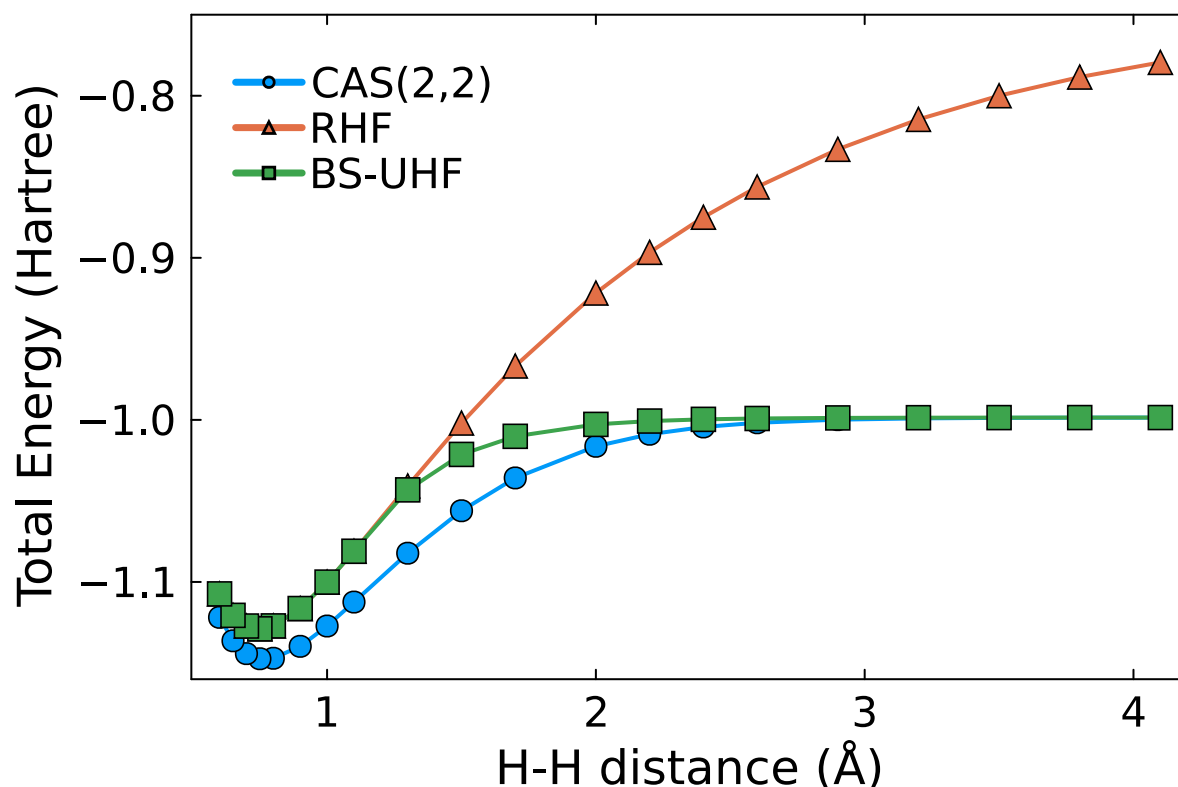


Fig. 3.17: Potential Energy Surface of the  $H_2$  molecule from RHF, UHF and CASSCF(2,2) calculations (Def2-SVP basis).

It is obvious, that the CASSCF surface is concise and yields the correct dissociation behavior. The RHF surface is roughly parallel to the CASSCF surface in the vicinity of the minimum but then starts to fail badly as the H-H bond starts to break. The broken-symmetry UHF solution is identical to RHF in the vicinity of the minimum and dissociates correctly. It is, however, of rather mediocre quality in the intermediate region where it follows the RHF surface.

A more challenging case is to dissociate the N-N bond of the  $N_2$  molecule correctly. Using CASSCF with the six p-orbitals we get a nice potential energy curve (The depth of the minimum is still too shallow compared to experiment by some 1 eV or so. A good dissociation energy requires a dynamic correlation treatment on top of CASSCF and a larger basis set).

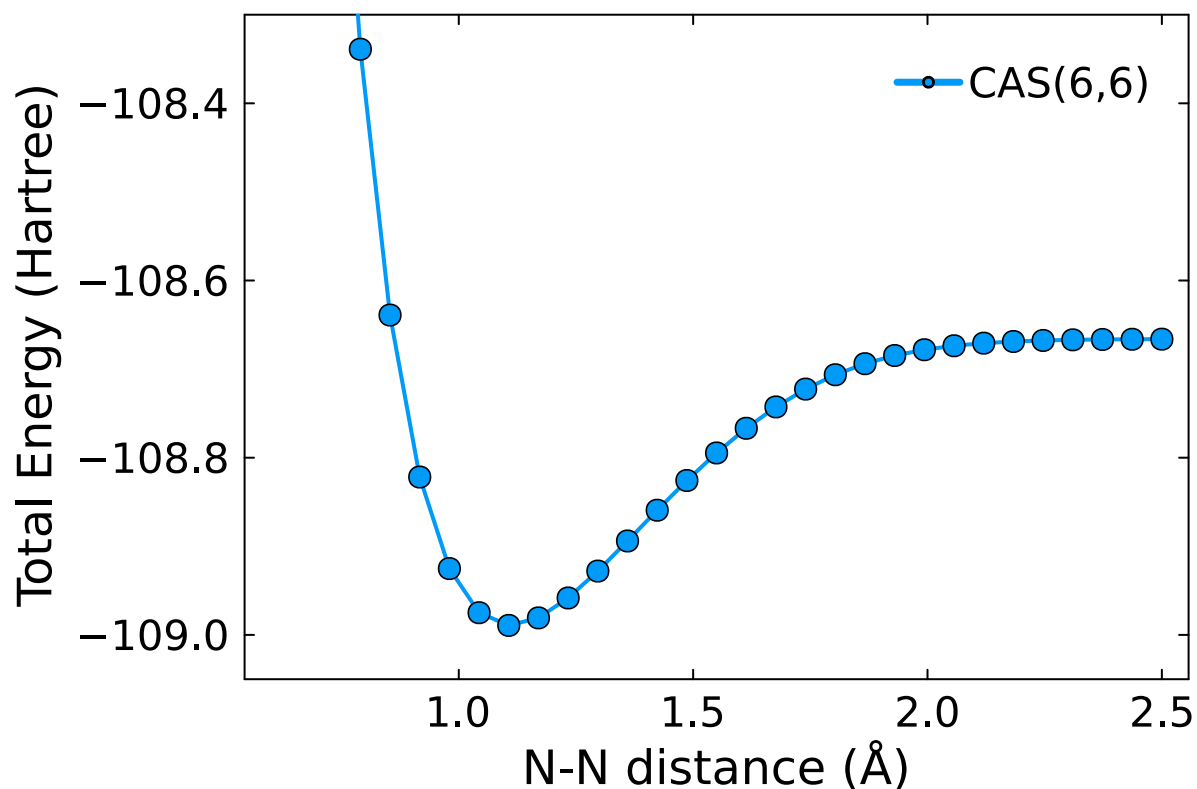


Fig. 3.18: Potential Energy Surface of the  $N_2$  molecule from CASSCF(6,6) calculations (Def2-SVP basis).

One can use the  $H_2$  example to illustrate the state-averaging feature. Since we have two active electrons we have two singlets and one triplet. Let us average the orbitals over these three states (we take equal weights for all multiplicity blocks):

```
!Def2-SVP TightSCF

%casscf
# define CAS(2,2)
nel      2
norb     2

# state-averaged triplet and singlet states
mult     3,1
nroots   1,2
end

# Scanning from the outside to the inside
%paras
R [4.1 3.8 3.5 3.2 2.9 2.6 2.4 2.2
   2 1.7 1.5 1.3 1.1 1 0.9 0.8
   0.75 0.7 0.65 0.6]
end

* xyz 0 1
H 0 0 0
H 0 0 {R}
end
```

which gives: