Coupled Cluster Type Excited State Calculations in ORCA.

Here I will briefly describe some example calculations in the ORCA program that can be used to access vertical excitation energies. The main focus is on the equation of motion coupled cluster and similarity equation of motion coupled cluster calculations. Detailed descriptions of the methods exist in the ORCA manual. The EOM and STEOM calculations can also be done using ACES2 and this allows comparison of results between different packages. All of the calculations I discuss can be found in the eom_excite and steom_excite directories under ~nooijen/Orca_examples_2017 on chem400 computers.

A. EOMCC calculations.

ORCA provides EOMCC calculations for excited states (EOM-CCSD), for ionized states (IP-EOM-CCSD), i.e. removal of one electron from usually closed shell ground state), and finally electron attached states (EA-EOM-CCSD), i.e. addition of an additional electron. We will look at only one molecular example here (see the labs on ACES2 for a fuller exploration of possibilities): namely formaldehyde using the ma-def2-tzvp basis set, which includes some diffuse character in the basis.

The input and output is straightforward. Here is the input for EE EOM-CCSD:

```plaintext
!RHF EOM-CCSD TightSCF ma-def2-TZVP
%maxcore 1200

* xyz 0 1
H 0.000000 0.934473 -0.588078
H 0.000000 -0.934473 -0.588078
C 0.000000 0.000000 0.000000
O 0.000000 0.000000 1.221104
*

%mdci
nroots 5
CCSD2 false
dorootwise true
dotdm true
doleft false
followcis true
printlevel 3
end
```

We are asking for the 5 (lowest) singlet states (nroots), that have the largest overlap with initial CIS eigenstates (followcis true), and only calculate right hand eigenstates. The excitation energies in such a calculation are correct EOM-CCSD eigenvalues but the oscillator strenghts (intensities of absorption spectrum) are approximate. To get true EOM-CCSD oscillator strenghts turn on doleft, which implies a more expensive calculation.
The relevant output of the calculation can be found looking for EOM-CCSD RESULTS. Here is what I get:

**EOM-CCSD RESULTS (RHS)**

----------------------

IROOT= 1: 0.144762 au  3.939 eV  31771.5 cm**-1

<table>
<thead>
<tr>
<th>Amplitude</th>
<th>Excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.394491</td>
<td>7 -&gt; 9</td>
</tr>
<tr>
<td>0.515586</td>
<td>7 -&gt; 12</td>
</tr>
<tr>
<td>-0.130841</td>
<td>7 -&gt; 15</td>
</tr>
<tr>
<td>-0.119684</td>
<td>7 -&gt; 20</td>
</tr>
</tbody>
</table>

Ground state amplitude: -0.000000

Percentage singles character= 92.29

IROOT= 2: 0.266800 au  7.260 eV  58555.8 cm**-1

<table>
<thead>
<tr>
<th>Amplitude</th>
<th>Excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.626051</td>
<td>7 -&gt; 8</td>
</tr>
<tr>
<td>-0.128984</td>
<td>7 -&gt; 11</td>
</tr>
<tr>
<td>0.131423</td>
<td>7 -&gt; 13</td>
</tr>
<tr>
<td>0.125503</td>
<td>7 -&gt; 16</td>
</tr>
<tr>
<td>0.117587</td>
<td>7 -&gt; 17</td>
</tr>
</tbody>
</table>

Ground state amplitude: -0.000000

Percentage singles character= 91.84

IROOT= 3: 0.297166 au  8.086 eV  65220.4 cm**-1

<table>
<thead>
<tr>
<th>Amplitude</th>
<th>Excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.144486</td>
<td>7 -&gt; 8</td>
</tr>
<tr>
<td>-0.622198</td>
<td>7 -&gt; 11</td>
</tr>
<tr>
<td>0.166691</td>
<td>7 -&gt; 17</td>
</tr>
<tr>
<td>-0.146855</td>
<td>7 -&gt; 19</td>
</tr>
</tbody>
</table>

Ground state amplitude: 0.000000

Percentage singles character= 92.27

IROOT= 4: 0.352627 au  9.595 eV  77392.7 cm**-1

<table>
<thead>
<tr>
<th>Amplitude</th>
<th>Excitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.409207</td>
<td>6 -&gt; 9</td>
</tr>
<tr>
<td>0.451505</td>
<td>6 -&gt; 12</td>
</tr>
<tr>
<td>-0.173430</td>
<td>7 -&gt; 14</td>
</tr>
<tr>
<td>0.129666</td>
<td>7 -&gt; 18</td>
</tr>
<tr>
<td>0.120947</td>
<td>6 -&gt; 12</td>
</tr>
</tbody>
</table>

Ground state amplitude: 0.041420

Percentage singles character= 86.87

You obtain the excitation energy for each state, and the character of the excitation, expressed in terms of excitations from occupied to virtual orbitals. In addition the percent singles character is indicated. This can be used to gauge the accuracy of an excitation (%singles > 85-90 %) to trust accuracy.
The use of IP and EA calculations is equally straightforward. The corresponding input for an **IP-EOM-CCSD** calculation is

```
!RHF IP-EOM-CCSD TightSCF ma-def2-TZVP
%maxcore 12000

* xyz 0 1
H   0.000000  0.934473  -0.588078
H   0.000000  -0.934473  -0.588078
C   0.000000  0.000000   0.000000
O   0.000000  0.000000   1.221104
*

%mdci
nroots 3
DoEOM true
CCSD2 false
docosxeom false
doaox3e false
dorootwise true
printlevel 3
end
```

with relevant output:

```
---------------
EOM-CCSD RESULTS (RHS)
---------------

IROOT= 1: 0.396824 au   10.798 eV   87092.7 cm**-1
   Amplitude  Excitation
      0.691327    7 -> x
Percentage singles character=   93.36

IROOT= 2: 0.531161 au   14.454 eV   116576.4 cm**-1
   Amplitude  Excitation
      0.694772    6 -> x
Percentage singles character=   93.31

IROOT= 3: 0.588428 au   16.012 eV   129145.1 cm**-1
   Amplitude  Excitation
      0.690403    5 -> x
Percentage singles character=   92.37
```

Here 7 -> x means one ionizes from orbital 7.

Likewise for an **EA-EOM-CCSD** calculation:
!RHF EA-EOM-CCSD TightSCF ma-def2-TZVP
%maxcore 12000

* xyz 0 1
H  0.000000  0.934473  -0.588078
H  0.000000  -0.934473  -0.588078
C  0.000000  0.000000   0.000000
O  0.000000  0.000000   1.221104
*

%mdci
nroots 3
DoEOM true
CCSD2 false
docsxeneom false
doaox3e false
dorootwise false
dotdm false
doleft false
followcis true
nrootsperbatch 5
printlevel 3
doloop

with output

----------------------
EOM-CCSD RESULTS (RHS)
----------------------

IROOT= 1:  0.039614 au  1.078 eV  8694.3 cm**-1
   Amplitude    Excitation     0.704833    x ->  8
Percentage singles character=     99.16

IROOT= 2:  0.040616 au  1.105 eV  8914.2 cm**-1
   Amplitude    Excitation     0.649450    x ->  9
     -0.254223    x ->  12
Percentage singles character=     96.32

IROOT= 3:  0.067465 au  1.836 eV  14806.9 cm**-1
   Amplitude    Excitation     -0.705738    x ->  10
Percentage singles character=     99.46

Here x -> 8 means adding an electron to orbital 8, is the dominant contribution to the state.

B. EOM-CCSD2 calculations.
The most expensive step, in particular for IP-EOM-CCSD calculations, is the CCSD calculation of the ground state. A reasonable first approximation is to replace the CCSD ground state amplitudes by their MP2 counterparts. This can be accomplished for any of the calculations by switching `ccsd2 true` in the `%mdci` keyword list. Such calculations are included in the `eom_excite` directory.

C. bt-PNO-EOM-CCSD calculations.

Replacing CCSD amplitudes by MP2 amplitudes as discussed in section B, is not always satisfactory. A cost-effective alternative is to use approximate the ground state CCSD amplitudes by the DLPNO (or PNO) analog. This means the ground state is calculated efficiently. The rest of the calculations proceed in the canonical basis. Still this can lead to significant speed-ups with almost no loss of accuracy compared to canonical EOMCC calculations. Here is an input file for bt-pno-eom-ccsd:

```plaintext
!RHF bt-PNO-EOM-CCSD TightSCF ma-def2-TZVP def2-TZVP/C def2/J
%maxcore 1200
* xyz 0 1
    H   0.000000  0.934473 -0.588078
    H   0.000000 -0.934473 -0.588078
    C   0.000000  0.000000  0.000000
    O   0.000000  0.000000  1.221104
*

%mdci
nroots 5
CCSD2 false
dorootwise true
dotdm true
doleft true
followcis true
printcis true
printlevel 3
end
```

The PNO CCSD requires an auxiliary basis set (/C) and one can use additional /J basis set to speed up HF calculations. See the notes on auxiliary basis sets in the ORCA manual. The rest of the input, and most of the output is exactly the same as for conventional EOM calculations. The same technique can be used with IP and EA PNO-EOM-CCSD calculations. For examples see the `eom_excite` directory.

Similarity Transformed EOMCC calculations (STEOM)

A STEOM calculation proceeds in a number of steps (see also ORCA manual): Coupled Cluster calculation of ground state, solution of IP-EOM and EA-EOM equations to get all ‘relevant’ roots, and a final diagonalization of a dressed Hamiltonian matrix over the CIS excited
determinant space. This last step is very fast and this is a major appeal of the method. The prior steps are used to get modified one- and two-electron matrix elements, building in correlation effects using the CCSD, IP-EOM and EA-EOM solutions, using similarity transforms of the hamiltonian.

A major simplification in ORCA compared to ACES2, is the automatic choice of ‘active' space, and the optimization of occupied and virtual orbital spaces for the purpose of excited state calculations. This means the number of nroots to solve for in IP-EOM and EA-EOM is done automatically, based on a single threshold.

The following input file is a basic canonical STEOM calculation for formaldehyde using the ma-def2-TZVP basis set.

```
!STEOM—CCSD TightSCF ma—def2—TZVP
%
maxcore 12000

* xyz 0 1
H 0.000000 0.934473 -0.588078
H 0.000000 -0.934473 -0.588078
C 0.000000 0.000000 0.000000
O 0.000000 0.000000 1.221104
*

%mdci
NRoots 8       #no of roots
printlevel 3   #printing level
OTHRESH 0.005  # Threshold for occupied orbital for the active space selection, different from the current defaults.
VTHRESH 0.005  # Threshold for virtual orbitals for the active space selection
IPSTHRS 70     # Thresh hold for discarding doubly excited IP states
EASTHRS 70     # Thresh hold for discarding doubly excited EA states
NDAV 400       # Davidsons dimension
Maxiter 2000   # maxiter, probably last two is just overkilling
DTol 1e—5      # Davidsons tolerance
DoROOTWISE true # use the root wise solver
FollowCIS true # track CIS root
end
```

The most relevant parameters in the input that the user may want to take into consideration are: a) the number of excitation energies one wants to calculate (nroots). In STEOM it typically does not cost more time to calculate more roots. However, using the automatic active space selection scheme, the number of roots calculated in the IP-EOM and EA-EOM segments may be affected.
I can list the output of the calculation. The first aspect is the IP-EOM calculation. Here are the relevant output sections:

-----------------------------------------------
RHF EOM-CCSD CALCULATION
-----------------------------------------------

EOM type ...
Eigenvectors ...
Solver ...
K(R) Formation ...
Convergence check ...
Convergence threshold ...
Root homing ...
Preconditioning update ...
Reduced space size (times number of roots) ...
Number of roots in the CIS initial guess ...
Number of roots to be optimized ...
Solving roots ...
Number of amplitudes to be optimized ...

This describes the input for IP-EOM: search for 4 roots in particular. (Look for IP-EOM in output)

The final results of the IP-EOM are in the usual spot:

-------------------------
EOM-CCSD RESULTS (RHS)
-------------------------

IROOT= 1: 0.396827 au  10.798 eV  87093.4 cm**-1
   Amplitude   Excitation
      -0.691316    7 -> x
Percentage singles character= 93.36

IROOT= 2: 0.531160 au  14.454 eV  116576.2 cm**-1
   Amplitude   Excitation
      0.694773    6 -> x
Percentage singles character= 93.31

IROOT= 3: 0.588416 au  16.012 eV  129142.4 cm**-1
   Amplitude   Excitation
      0.686718    5 -> x
Percentage singles character= 92.37

IROOT= 4: 0.636787 au  17.328 eV  139758.6 cm**-1
   Amplitude   Excitation
      0.685386    4 -> x
Percentage singles character= 91.34
IP STATE= 0:  percentage singles  93.362
IP STATE= 1:  percentage singles  93.315
IP STATE= 2:  percentage singles  92.367
IP STATE= 3:  percentage singles  91.342
Final no active IP roots:  4

The EA-EOM calculation follows the IP-EOM calculation. The input section (search for EA-EOM in output file):

---------------------------------------------------
RHF EOM-CCSD CALCULATION
---------------------------------------------------

<table>
<thead>
<tr>
<th>EOM type</th>
<th>... EA-EOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvectors</td>
<td>... RHS</td>
</tr>
<tr>
<td>Solver</td>
<td>... Davidson</td>
</tr>
<tr>
<td>K(R) Formation</td>
<td>... MO</td>
</tr>
<tr>
<td>Convergence check</td>
<td>... for each root separately</td>
</tr>
<tr>
<td>Convergence threshold</td>
<td>... 1.0E-05</td>
</tr>
<tr>
<td>Root homing</td>
<td>... on</td>
</tr>
<tr>
<td>Preconditioning update</td>
<td>... CIS</td>
</tr>
<tr>
<td>Reduced space size (times number of roots)</td>
<td>... 400</td>
</tr>
<tr>
<td>Number of roots in the CIS initial guess</td>
<td>... 6</td>
</tr>
<tr>
<td>Number of roots to be optimized</td>
<td>... 6</td>
</tr>
<tr>
<td>Solving roots</td>
<td>... separately</td>
</tr>
<tr>
<td>Number of amplitudes to be optimized</td>
<td>... 32930</td>
</tr>
</tbody>
</table>

And the corresponding output

-------------------------------
EOM-CCSD RESULTS (RHS)
-------------------------------

IROOT= 1:  0.039597 au  1.077 eV  8690.5 cm**-1
   Amplitude Excitation
     -0.682394  x ->  8
     0.111950   x -> 14
Percentage singles character=  99.15

IROOT= 2:  0.040595 au  1.105 eV  8909.6 cm**-1
   Amplitude Excitation
     -0.625944  x ->  9
     -0.281125  x -> 12
     0.135340   x -> 15
Percentage singles character=  96.31

IROOT= 3:  0.067486 au  1.836 eV  14811.6 cm**-1
Amplitude    Excitation
-0.686422     x -> 10
-0.154377     x -> 18
Percentage singles character= 99.47

IROOT= 4: 0.067151 au 1.827 eV 14738.0 cm**-1
Amplitude    Excitation
-0.665325     x -> 11
 0.194984     x -> 17
Percentage singles character= 99.40

Etc.

The final STEOM-CCSD results are found in the section

------------------
STEOM-CCSD RESULTS
------------------

IROOT= 1: 0.142001 au 3.864 eV 31165.7 cm**-1
Amplitude    Excitation
 0.121366     4 -> 12
-0.551041     7 -> 9
-0.820264     7 -> 12

Percentage Active Character 99.71

Amplitude    Excitation in Canonical Basis
 0.110313     4 -> 12
-0.566404     7 -> 9
 0.765155     7 -> 12
-0.185996     7 -> 15
-0.181297     7 -> 20

IROOT= 2: 0.268364 au 7.303 eV 58899.1 cm**-1
Amplitude    Excitation
-0.957335     7 -> 8
 0.273797     7 -> 11

Percentage Active Character 99.85

This output follows a similar pattern as for EOM calculations, but there are only single excitations. In
the output the percentage active character is provided. Ideally this should be larger than about 97% or
so for results to be sufficiently reliable. One can lower the OTHRsh or VTHRsh parameters to
include more occupied or virtual orbitals respectively in the active space.

We also obtain absorption spectra:

-----------------------------------
ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
As for IP/EA/EE EOM-CCSD calculation one can speed up the CCSD part of the calculation, by either using the CCSD2 option, or more accurately, using the bt-PNO-STEOM-CCSD approach. To run the CCSD2 variant, simply switch on CCSD2 in the %mdci namelist. An input file for the latter calculation is

```
!bt-PNO-STEOM-CCSD TightSCF ma-def2-TZVP def2-TZVP/C def2/J NORMALPNO
%maxcore 12000
*
* xyz 0 1
H 0.000000 0.934473 -0.588078
H 0.000000 -0.934473 -0.588078
C 0.000000 0.000000  0.000000
O 0.000000 0.000000  1.221104
*

%mdci
NRoots  8        #no of roots
printlevel 3     #printing level
OTHRESH 0.005    # Threshold for occupied orbital for the active space selection, different from the current defaults.
VTHRESH 0.005    # Threshold for virtual orbitals for the active space selection
```
As in the case of bPNO-EOM-CCSD calculations one has to specify auxiliary basis set to perform the DLPNO-CCSD calculation. This type of STEOM calculation can be nearly as accurate as the canonical STEOM calculation, but for large molecules there is a substantial speed-up. The EOM parts of the calculation are still run at the canonical level, and there is little speed-up for these parts of the calculation.

A very recent development is a DLPNO scheme also for the IP and EA segments of the calculation. This type of calculation may need further scrutiny regarding the selection of the active space. Here is a preliminary input file. We can explore.

!STEOM-DLPNO-CCSD TightSCF ma-def2-TZVP def2-TZVP/C def2/J NORMALPNO
%maxcore 12000

* xyz 0 1
  H    0.000000  0.934473 -0.588078
  H    0.000000 -0.934473 -0.588078
  C    0.000000  0.000000  0.000000
  O    0.000000  0.000000  1.221104
*

%mdci
NRoots 8 #no of roots
printlevel 3 #printing level
OTRSH 0.005 # Threshold for occupied orbital for the active space selection, different from the current defaults.
VTRSH 0.01 # Threshold for virtual orbitals for the active space selection
IPSTHRS 70 # Thresh hold for discarding doubly excited IP states
EASTHRS 70 # Thresh hold for discarding doubly excited EA states
TCutPNOSingles 6.6e-10 # singlesPNO need to be rescaled to get good DLPNOEA
DoNewActSch false # new active space selection scheme
NDAV 400 # Davidsons dimension
Maxiter 2000 # maxiter, probably last two is just overkilling
DTol 1e-5 # Davidsons tolerance
DoROOTWISE true # use the root wise solver
FollowCIS true         # track CIS root
end

The output of the calculations follows the same logic as canonical calculations. You are ready to explore.