

(continued from previous page)

```

  2 2 2 2 2 0 1 1 0
  2 2 2 2 2 0 2 0 0
  2 2 2 2 2 1 0 0 1
  2 2 2 2 2 1 0 1 0
  2 2 2 2 2 1 1 0 0
  2 2 2 2 2 2 0 0 0
end
DoDipoleVelocity true
DecomposeFosc true
rel
dosoc true
XESSOC true
XASMOs 59
DoDTensor false
DoVelocity true
end
end

%method
SpecialGridAtoms 26 #Increase the radial integration accuracy on Fe
SpecialGridIntAcc 7 #Requested radial integration accuracy values
end

* xyz -2 5
Fe      0      0      0
Cl      2.40    0      0
Cl      -2.40   0      0
Cl      0      2.40   0
Cl      0      -2.40  0
Cl      0      0      2.40
Cl      0      0      -2.40
*
```

In ORCA 6 all these type of inputs have been unified across the wavefunction based modules (e.g. CASSCF MRCI, RASCI, ROCIS and LFT) so in a similar fashion to *X-ray Spectroscopy* one can also specify the respective RAS excitation space as following:

```

%casscf
nel 12
norb 9
nroots 1000, 1000
mult 7, 5
maxiter 1
refs ras(12:4 1/5/ 0 0) end
DoDipoleVelocity true
DecomposeFosc true
rel
  DoSOc true
  Xessoc true
  xasmos 59
  DoDTensor false
  DoVelocity true
end
end
```

In the above inputs one notes that the exact knowledge of the states to saturate the excitations space is not required. One only need to specify a large number (e.g. 1000) and the program will automatically detect the required 19 septet and 270 quintet states:

```

BLOCK 1 WEIGHT= 0.5000
Multiplicity          ...      7
#(Configurations)     ...      4
#(CSFs)               ...      4
#(Roots)              ... 1000
WARNING (ORCA_CASSCF): NRoots > NCSFs. Adjusting to maximum number of roots.
↳Please check the output carefully!
#(Roots)              ...      4

...

BLOCK 2 WEIGHT= 0.5000
Multiplicity          ...      5
#(Configurations)     ...     89
#(CSFs)               ...    105
#(Roots)              ... 1000
WARNING (ORCA_CASSCF): NRoots > NCSFs. Adjusting to maximum number of roots.
↳Please check the output carefully!
#(Roots)              ...    105

```

By now running the above input for the 4 septet and the 81 quintet states the following output is generated

```

-----
↳-----
      SOC CORRECTED      EMISSION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-----
↳-----
      Transition          Energy      Energy  Wavelength fosc(D2)      D2      |DX|
      |DY|      |DZ|      (eV)      (cm-1)      (nm)      (au**2)      (au)
      (au)      (au)
-----
↳-----
421-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000000870    0.000000
↳0.000000 0.000000 0.000007
423-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000001088    0.000000
↳0.000006 0.000006 0.000000
424-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000000898    0.000000
↳0.000002 0.000003 0.000006
422-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000000000    0.000000
↳0.000000 0.000000 0.000000
426-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000001048    0.000000
↳0.000005 0.000005 0.000003
425-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000000664    0.000000
↳0.000005 0.000004 0.000001
427-7.0A -> 420-5.0A  7084.189573 57137848.2    0.2    0.000000653    0.000000
↳0.000004 0.000004 0.000000
421-7.0A -> 419-5.0A  7084.189691 57137849.2    0.2    0.000001233    0.000000
↳0.000007 0.000005 0.000003
423-7.0A -> 419-5.0A  7084.189691 57137849.2    0.2    0.000000354    0.000000
↳0.000002 0.000002 0.000004
422-7.0A -> 419-5.0A  7084.189691 57137849.2    0.2    0.000001427    0.000000
↳0.000004 0.000008 0.000000
424-7.0A -> 419-5.0A  7084.189691 57137849.2    0.2    0.000000467    0.000000
↳0.000005 0.000002 0.000001
426-7.0A -> 419-5.0A  7084.189691 57137849.2    0.2    0.000000216    0.000000
↳0.000003 0.000002 0.000001
425-7.0A -> 419-5.0A  7084.189691 57137849.2    0.2    0.000001334    0.000000
↳0.000001 0.000002 0.000009
421-7.0A -> 418-5.0A  7084.189691 57137849.2    0.2    0.000000757    0.000000
↳0.000000 0.000005 0.000004
...

```

Finally by processing the .out file with orca_mapspc:

```
orca_mapspc fecl6_xes.out XESSOC -x07000 -x17200 -w4.0 -eV -n10000
```

and by plotting the resulted XES spectrum one will get the respective RASCI-XES spectrum presented in Fig. 3.23

This will result in Fig. 3.23.

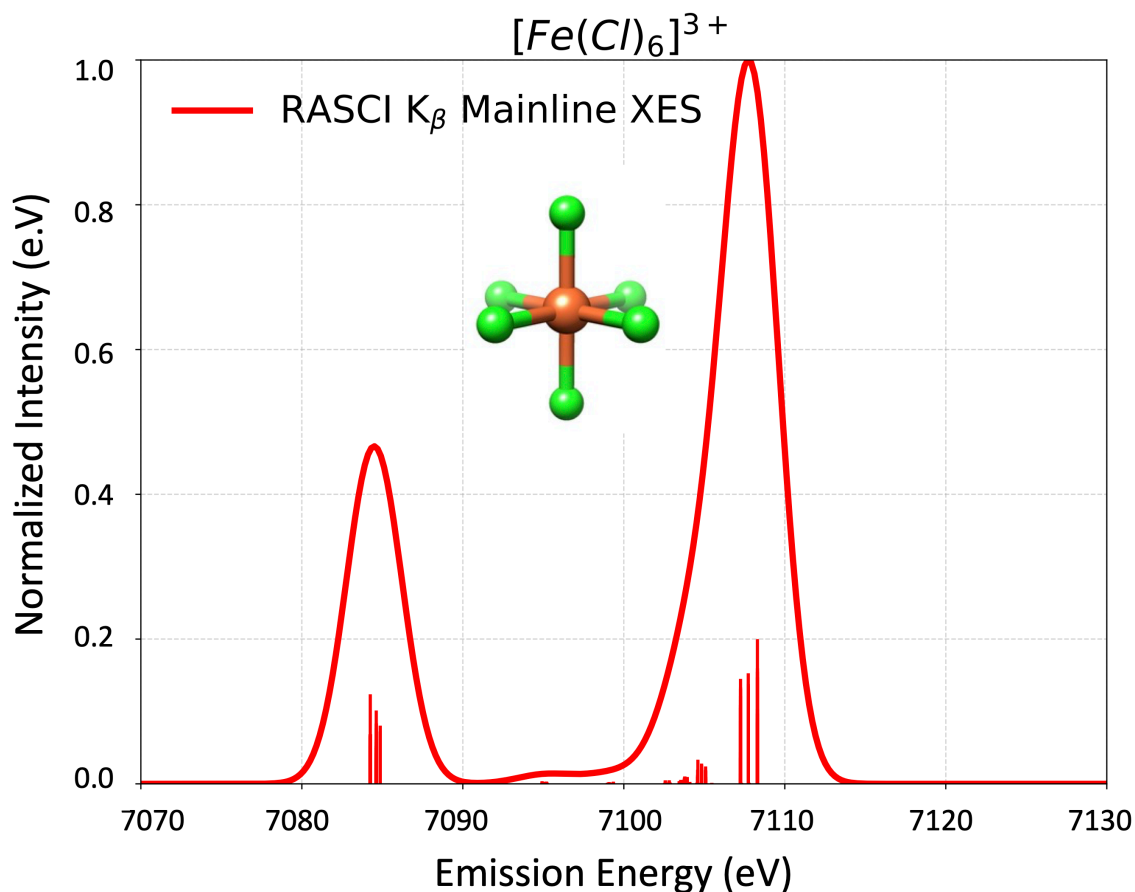


Fig. 3.23: Calculated RASCI K_{β} Mainline XES spectrum of $[Fe(Cl)_6]^{3+}$

Since Orca 6.0 the computed transition moments in the presence of SOC can be taken beyond the dipole approximation by using the OPS tool, check section [Section 5.14](#) for details.

All the possible approximations can be requested with the following commands:

```
#Non-Relativistic/Relativistic treatment

%casscf
DoDipoleLength      true
DoDipoleVelocity    true
DecomposeFosc       true
DoFullSemiclassical true
end
```

Note

The recommended process for generating XES spectra

- within the electric dipole approximation (EDA) is via DoDipoleLength or DoDipoleVelocity

- beyond EDA, employing the Full Field Matter Interaction Operator FFMIO via the `DoFullSemiclassical`

Orca_mapspc can process all these files. The list of the relevant keywords in the case of RASCI-XES protocol is

```
XESSOC #This will process the SOC corrected beyond the electric dipole_
↪approximation (EDA)
XESSOCFFMIO #This will process the SOC corrected FFMIO X-ray Emission spectra
```

A more complete list can be found in (*orca_mapspc*)

Note

This is in principle the recommended protocol for computing X-ray related spectra XAS, XES, RXES, RIXS:

- As it provides more direct and computationally efficient protocol for computing
- Metal K, L and M edge XAS spectra
- K_α : Fe $2p \rightarrow$ Fe $1s$
- $K_{\beta_{1,3}}$: Fe $3p \rightarrow$ Fe $1s$ (Mainline XES)
- $K_{\beta_{2,5}} / K_{\beta'}$: (Cl $2p +$ Fe $3d \rightarrow$ Fe $1s$ (VtC-XES)
- as well the respective resonant XES (RIXS) spectra.
- Combined with NEVPT2 the protocol can actually treat quite challenging cases in terms of electron correlation.
- We should mention that within the limits of their applicability XES and RIXS spectra can also be computed on the basis of Static Ground State DFT (SGS-DFT), see discussion in *Static Ground State DFT (SGS-DFT)* and ROCIS family of methods, see discussion in *Resonant Inelastic Scattering Spectroscopy*.
- A general discussion regarding X-ray spectroscopy experiments and how to approach them on the basis of modern computational spectroscopy methodologies is provided in reference [458]

3.14 Approximate Full CI Calculations in Subspace: ICE-CI

3.14.1 Introduction

In many circumstances, one would like to generate a wavefunction that is as close as possible to the full-CI result, but Full CI itself is out of the question for computational reasons. Situations in which that may be desirable include a) one wants to generate highly accurate energies for small molecules or b) one wants to sort out a number of low-lying states or c) one wants to run *CASSCF calculations* with larger active spaces than the about fourteen orbitals that have been the state of the art for a long time.

ORCA features a method that has been termed Iterative-Configuration Expansion Configuration Interaction (ICE-CI).^[443, 444] It is based on much older ideas brought forward by Jean-Paul Malrieu and his co-workers in the framework of the CIPSI (an abbreviation for a method with a rather bulky name *Configuration Interaction by Perturbation with multiconfigurational zeroth-order wave functions Selected by Iterative process*) in the early 1970s.

The goal of the ICE-CI is to provide compact wavefunction(s) (e.g. one or several states) close to the full-CI limit at a small fraction of the computational cost. However, ICE-CI itself is not designed to deal with hundreds of atoms or thousands of basis functions. Thus, unlike, say DLPNO-CCSD(T) which is a high accuracy method for treating large systems, ICE-CI is either a highly robust high accuracy method for very small systems or a “building block” for large systems. By itself it can treat a few dozen electrons and orbitals – e.g. *much* more than full CI – but it cannot

do wonders. Its scope is similar to the density matrix renormalization group (DMRG) or Quantum Monte Carlo Full CI (QMCFCI) procedures.

ICE-CI should be viewed as a multireference approach. It is self-adaptive and robust, even in the presence of near or perfect degeneracies. It yields orthogonal states (when applied to several states) and spin eigenfunctions. It also yields a density and a spin density.

3.14.2 The ICE-CI and CIPSI Algorithms

The general idea of ICE-CI is straightforward: Consider a many-particle state that has at least a sizeable contribution from a given configuration \mathbf{n}^0 (this is a set of occupation numbers for the active orbitals that are $n_p^0 = 0, 1$ or 2 (p = any active orbital)). By nature of the non-relativistic Hamiltonian only configurations that differ by at most two orbital occupations from \mathbf{n}^0 will interact with it. We can use perturbation theory to select the subset of singles and doubles that interact most strongly with \mathbf{n}^0 and then solve the variational problem. We can then analyze the CI vector for configurations that make a dominant contribution to the ground state. Say, we single out the configurations with $C_I^2 > T_{\text{gen}}$. This defines the “generator” set of configurations. The other configurations are called “variational” configurations. They are treated to infinite order by the variational principle, but are not important enough to bring in their single and double excitations. In the next iteration, we perform singles and doubles relative to these general configurations and select according to their interaction with the dominant part of the previous CI vector (truncated to the generators). This procedure can be repeated until no new important configurations are found and the total energy converges (See Fig. 3.24).

Flowchart

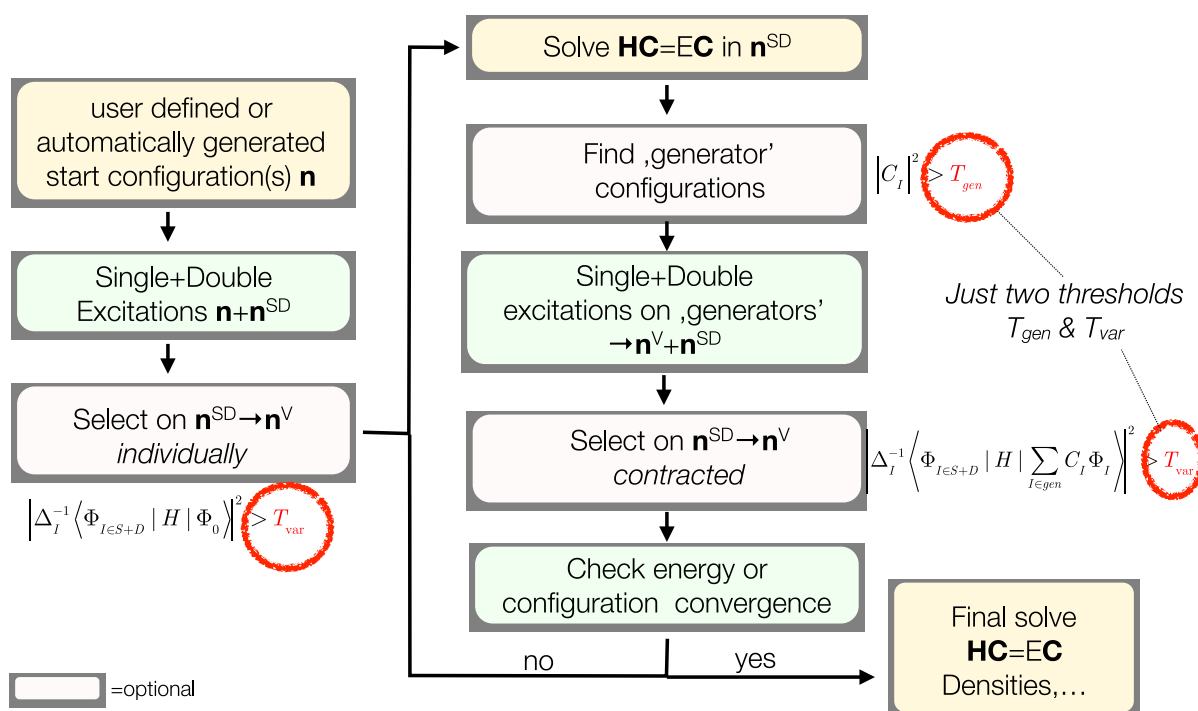


Fig. 3.24: Flowchart of the ICE-CI procedure.

The described procedure is very similar to Malrieu’s three level CIPSI procedure. One major technical difference, is that ICE is centered around *configurations* and *configuration state functions* rather than *determinants*. A configuration is a set of occupation numbers 0, 1 or 2 that describes how the electrons are distributed among the available spatial orbitals. A *configuration state function* (CSF) is created by coupling the unpaired spins in a given configuration to a

given total spin S . In general there are several, if not many ways to construct a linearly independent set of CSFs. CSFs on the other hand can be expanded in terms of Slater determinants, but there are more Slater determinants to a given configuration than CSFs. For example for a CAS(14,18) calculation one has about 10^9 determinants, but only about 3×10^8 CSFs and 3×10^7 configurations. In the configuration based ICE (CFG-ICE) all logic happens at the level of configurations. That is, it is the relationship between two configurations that determines whether and if yes, by which integrals the CSFs or determinants of two given configurations interact. Since the configuration space is so much more compact than the determinant space substantial computational benefit can be realized by organizing the calculation around the concept of a configuration. In general, in CFG-ICE all CSFs that belong to a given configuration are included and all selection quantities are summed over all CSFs of a given configuration before it is decided whether this CSFs is included or not. In the configuration state functions based ICE (CSF-ICE) the logic of generation and selection occurs at the level of individual CSFs and therefore we get rid of the requirement to carry around all the CSFs for a given configuration. This provides substantial gains in the case of molecules containing a large number of transition metal atoms, where each atom contains a high-spin center. In such cases only a few CSFs of a the dominant CFG play a dominant role and other show negligible contribution to the wavefunction. Finally, in some cases the original determinant based CIPSI procedure could be preferred. Such cases can be handled by the determinant based ICE termed DET-ICE. The three variants of ICE therefore cover all the possible types of multi-reference systems that one encounters in quantum chemistry.

It should be noted that although the procedure contains a perturbative element, the final energy is strongly dominated by the variational energy and hence, for all intents and purposes, the ICE-CI procedure is variational (but not rigorously size consistent – size consistency errors are on the same order of magnitude as the error in absolute energy).

3.14.3 A Simple Example Calculation

Let us look at a simple calculation on the water molecule:

```
#
# Check the ICECI implementation
#

! SV

%ice  nel 10          # number of active electrons
      norb 13         # number of active orbitals
      nroots 1        # number of requested roots
      integrals exact  # exact 4-index transformation
                        # can be set to RI to avoid bottlenecks

      icetype  CFGs    # The configuration based ICE-CI
                     CSFs # The CSF based ICE-CI
                     DETs # The determinnat based ICE-CI
                        # please choose one of these three

      Tgen 1e-04       # value for Tgen. Default is 1e-4
      Tvar 1e-11       # value for Tvar. Default is 1e-11 (1e-7*Tgen)

      etol 1e-06       # energy convergence tolerance
end

* int 0 1
O 0 0 0 0.0   0.000 0.000
H 1 0 0 1.0   0.000 0.000
H 1 2 0 1.0 104.060 0.000
*
```

Let us look at the output:

```
-----
ORCA Iterative Configuration Expansion
- a configuration driven CIPSI type approach -
```

(continues on next page)

(continued from previous page)

```

-----
(some startup information)

Making an initial 'Aufbau' configuration      ... done
Performing S+D excitations from      1   configs ... done (  0.0 sec) NCFG=581
Performing perturbative selection        ... done (  0.0 sec)
  # of configurations before selection      ... 581
  # of configurations after selection       ... 191
  'rest' energy (probably not very physical) ... -3.736391e-10

          *****
          *   ICECI MACROITERATION   1   *
          *****

  # of active configurations =      191
  Now calling CI solver (269 CSFs)
(...)
CI SOLUTION :
STATE   0 MULT= 1: E=      -76.0463127108 Eh W=   1.0000 DE= 0.000 eV      0.0 cm**-1
      0.95752 : 2222200000000
  Selecting new configurations            ... done (  0.0 sec)
  # of selected configurations            ...   191
  # of generator configurations           ...    69
  Performing single and double excitations relative to generators ... done (  0.0
↪sec)
  # of configurations after S+D            ... 13174
  Selecting from the generated configurations ... done (  0.1 sec)
  # of configurations after Selection      ... 3827
  Root   0:      -76.046312711      -0.000000063      -76.046312773
(...)
          *****
          *   ICECI MACROITERATION   3   *
          *****

  # of active configurations =      3866
  Now calling CI solver (9606 CSFs)

CI SOLUTION :
STATE   0 MULT= 1: E=      -76.0539542296 Eh W=   1.0000 DE= 0.000 eV      0.0 cm**-1
      0.95097 : 2222200000000
(...)

          ***** ICECI IS CONVERGED *****
(one final CI)

          *****
          **   ICECI Problem solved in   2.6 sec   **
          *****

FINAL CIPSI ENERGIES
  Final CIPSI Energy Root   0:      -76.053954291 EH

```

From the output the individual steps in the calculation are readily appreciated. The program keeps cycling between variational solution of the CI problem, generation of new configurations and perturbative selection until convergence of the energy is achieved. Normally, this occurs rapidly and rarely requires more than five iterations. The result will be close to the Full CI result.

Let us look at a H₂O/cc-pVDZ calculation in a bit more detail (See Fig. 3.25). The calculation starts out with a single Hartree-Fock configuration. The first iteration of ICE-CI creates the singles and doubles and altogether 544 configurations are selected. These singles and doubles bring in about half of the correlation energy. Already

the second iteration, which leads to 73000 selected CSFs provides a result close to the full CI. At this point up to quadruple excitations from the Hartree-Fock reference have been included. It is well known that such quadruple excitations are important for the correct behavior of the CI procedure (near size consistency will come from the part of the quadruple excitations that are products of doubles). However, only a very small fraction of quadruples will be necessary for achieving the desired accuracy. In the first iteration the procedure is already converged and provides 99.8% of the correlation energy, using 0.5% of the CSFs in the full CI space and at less than 0.2% the calculation time required for solving the full CI problem. Hence, it is clear that near exact results can be obtained while realizing spectacular savings.

| H ₂ O, cc-pVDZ, 1s frozen, 8 electrons, 23 orbitals T _{Gen} =10 ⁻⁴ , T _{var} =10 ⁻¹¹ | | | | |
|--|--|----------------|------------------|--------------------------|
| | Energy | N(CFG) | N(CSF) | time(sec) |
| Hartree-Fock | -76.0260966 | 1 | 1 | 2 |
| | 1977 CFG's generated, 544 selected | | | <0.1 |
| Iter=1 | -76.2292278 | 544 | 805 S+D | 0.3 |
| | 98717 CFG's generated, 22471 selected | | | 1.1 |
| Iter=2 | -76.2412460 | 22471 | 73296 T+Q | 17.1 |
| | 10677 CFG's generated, 23241 selected | | | 2.2 |
| Iter=3 | -76.2412769 | 23241 | 75705 P+H | 20.1 |
| | 10677 CFG's generated, 23241 selected | | | 3.3 |
| Iter=4 | -76.2412769 | 23241 | 75705 S+O | 20.1 |
| | 99.75% of EC | 0.57% | 0.4% | 83.5 (1 final CI) |
| | | | | 0.18% of FCI |
| Full-CI | -76.2418240 | 4065963 | 18818646 | 44529.0 |

Fig. 3.25: An ICE-CI calculation on the water molecule in the cc-pVDZ basis (1s frozen)

3.14.4 Accuracy

The accuracy of the procedure is controlled by two parameters T_{gen} and T_{var} . Since we have found that $T_{\text{var}} = 10^{-7} T_{\text{gen}}$ always provides converged results, this choice is the default. However, T_{var} can be set manually. It can be reduced considerably in order to speed up the calculations at the expense of some accuracy. Our default values are $T_{\text{gen}} = 10^{-4}$ and $T_{\text{var}} = 10^{-11}$. This provides results within about 1 mEh of the full CI results (roughly speaking, a bit better than CCSDT for genuine closed-shell systems).

During the development of ICE-CI systematic test calculations have been performed using a reference set of 21 full CI energies on small molecules. The convergence pattern of the mean absolute error is shown in Fig. 3.26. It is evident from the figure that the convergence of ICE-CI towards the FCI result is very smooth and that high accuracy can be obtained. In fact, the default settings lead to an accuracy of <1 mEh deviation to the full-CI result. μEh accuracy can be achieved by further tightening. The achieved accuracy relative to accurate coupled-cluster results shows that the accuracy of even CCSDTQ can be surpassed by ICE-CI. The achievable accuracy is only limited by the value of T_{gen} and much less so by the value of T_{var} . Hence, it is advisable to use a value for T_{var} that is essentially converged and control the accuracy of the procedure by T_{gen} .

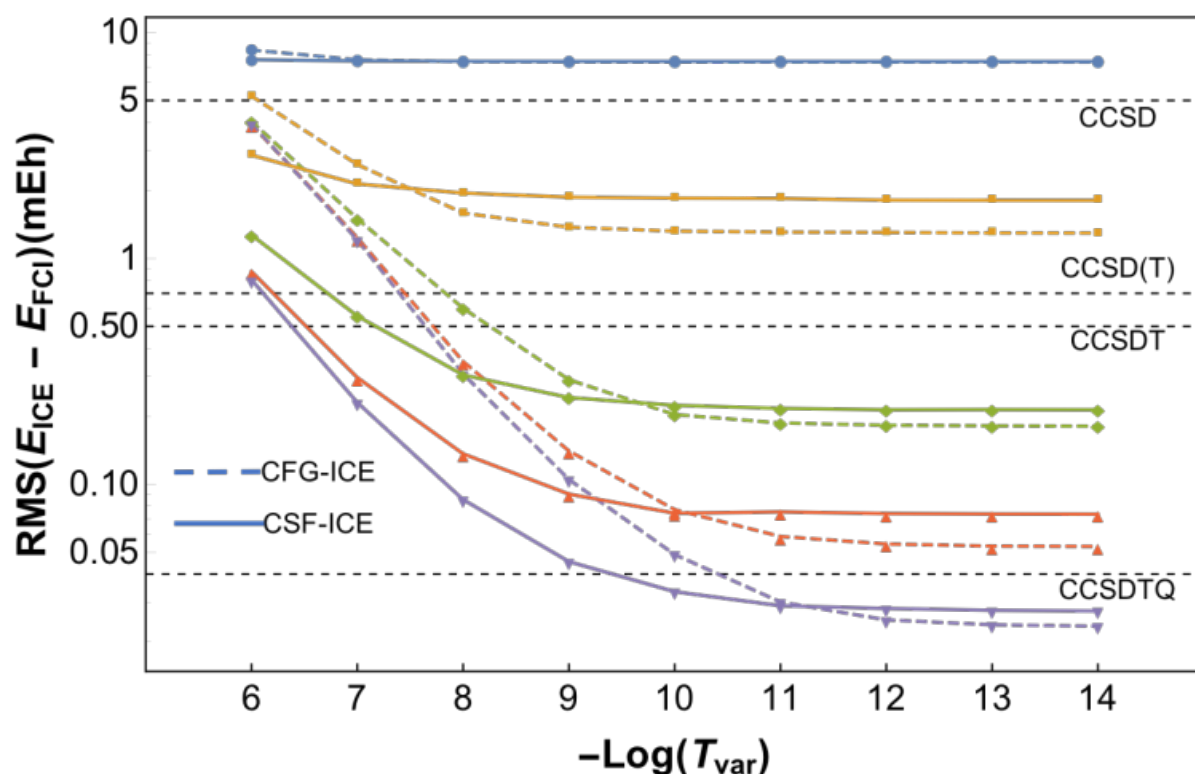


Fig. 3.26: Convergence of the ICE-CI procedure towards the full CI results for a test set of 21 full CI energy. Shown is the RMS error relative to the Full CI results. The corresponding errors for various coupled-cluster variants is shown by broken horizontal lines.

3.14.5 Scaling behavior

ICE-CI will break the factorial scaling of the full CI problem and scale polynomially. The actual order of the polynomial scaling is system dependent and accuracy dependent. In order to provide some impression, consider some calculations on linear polyene chains.

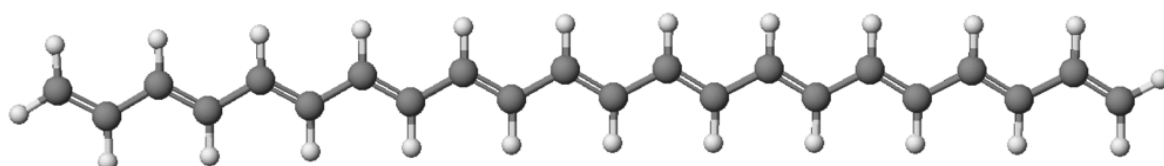


Fig. 3.27: Polyene chains used for scaling calculations.

The results are displayed in the Fig. 3.28. It is evident from Fig. 3.28 that ICE-CI breaks the factorial scaling of the full CI problem. In fact, for a thresholds of $T_{\text{gen}} = 10^{-4}$, 10^{-3} and 10^{-2} the observed scalings are approximately $O(N^8)$, $O(N^7)$ and $O(N^6)$ respectively. These numbers will obviously be very system dependent but should serve as a rough guide. The calculations become quickly much more expensive if T_{gen} is tightened. A rule of thumb is that each order of magnitude tightening of T_{gen} increases the computation time by a factor of 10. The above calculations have been performed on a simple desktop computer and it was already possible to solve a CAS(30,30) problem in less than one day of elapsed time using the default thresholds. Large active spaces will require either loosening of the thresholds or large, more powerful machines.

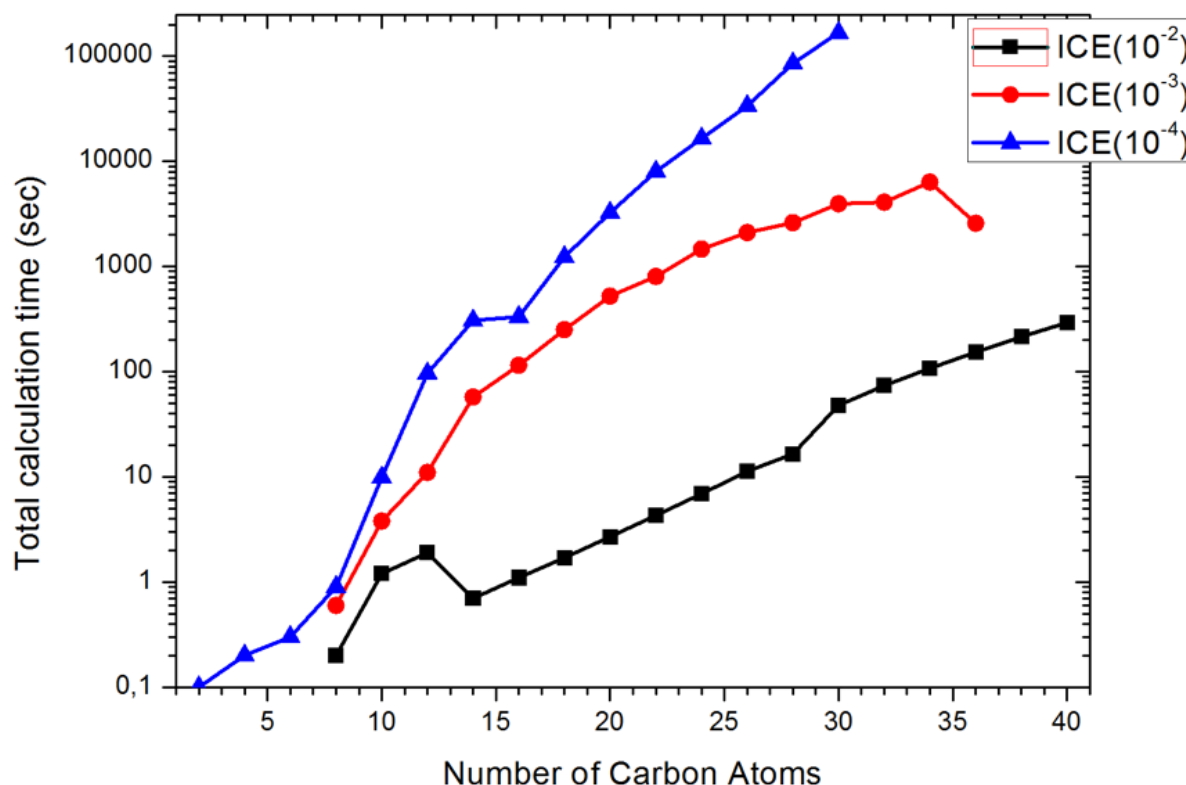


Fig. 3.28: Scaling behavior of ICE-CI for linear polyene chains (Full π -electron active space) as a functions of system size for different generator thresholds.

3.14.6 Accuracy of the Wavefunction

The accuracy of the many particle wavefunction is not straightforward to check. A reasonable measure, however, is how well it converges towards the exact result for one-electron expectation values. Since every expectation value can be written in terms of natural orbitals of the one-particle density as:

$$\langle \hat{O} \rangle = \langle \Psi | \sum_o \hat{o}(\mathbf{x}_i) | \Psi \rangle = \sum_{pq} D_{pq} \langle \psi_p | \hat{o} | \psi_q \rangle = \sum_p n_p \langle \tilde{\psi}_p | \hat{o} | \tilde{\psi}_p \rangle$$

where $\hat{o}(\mathbf{x}_i)$ is an arbitrary one-particle operator, D_{pq} is the density matrix of the ICE-CI wavefunction, $\tilde{\psi}_p$ are the natural orbitals of the ICE-CI wavefunction and n_p are their occupations numbers. It is reasonable to take the deviation of the natural orbital occupation numbers as a measure for wavefunction convergence.

For example, we treat the $\text{H}_2\text{O}/\text{cc-pVDZ}$ problem again. From the results in Fig. 3.29 it becomes evident that the ICE-CI wavefunction is fairly accurate. At the default threshold the occupation numbers agree to within 10^{-3} with the full CI reference numbers, which means that expectation values will be of similar accuracy. Interestingly, the largest errors occur in the region of the HOMO-LUMO gap, where apparently all approximate wavefunction approaches tend to depopulate the high lying orbitals too much and put too much electron density in the low lying empty orbitals. From comparison, it is seen, that the CCSD natural occupation numbers for this problem are significantly less accurate. Hence, this is evidence that the ICE-CI wavefunction is properly converging to the right result.

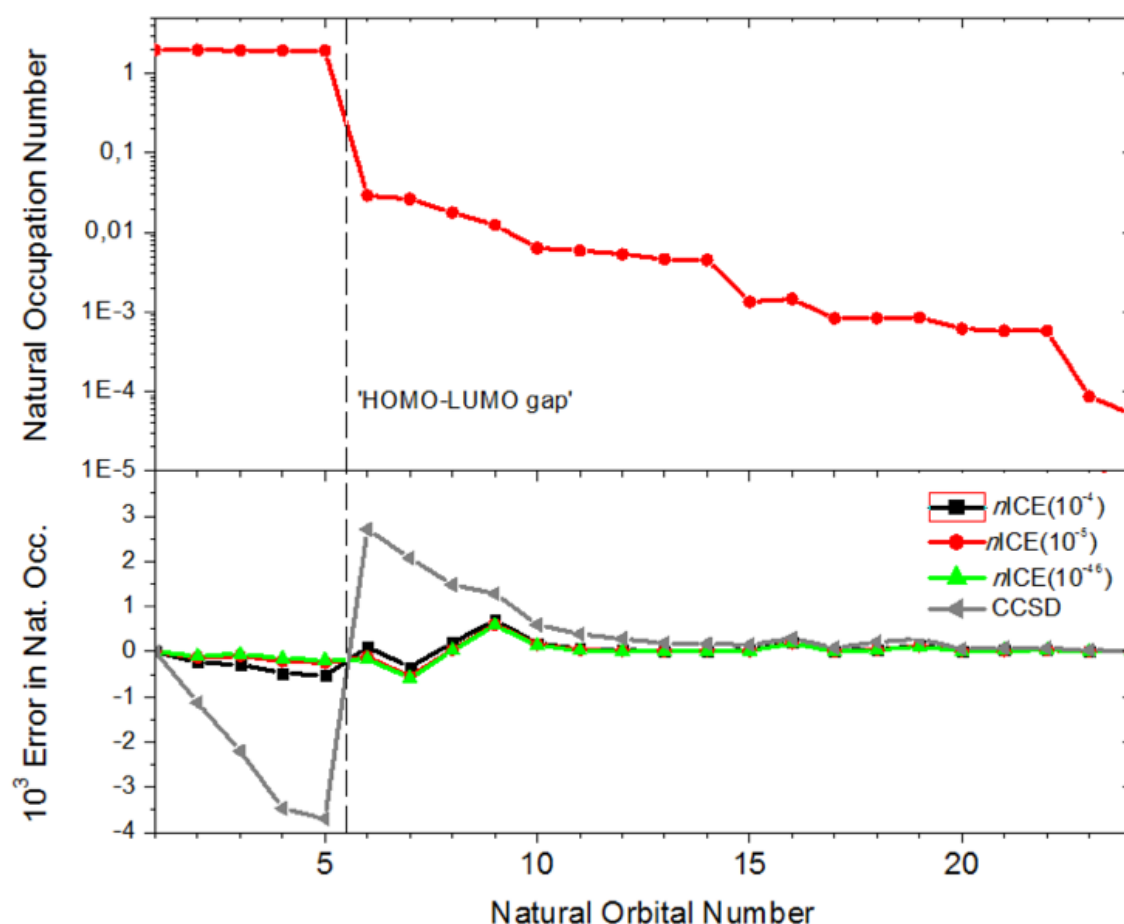


Fig. 3.29: Convergence of the ICE-CI natural orbital occupation numbers. The upper panel is showing the Full CI occupation numbers, the lower panel the deviation of the ICE-CI values from these exact values. For comparison, the CCSD natural orbital occupation numbers are also provided.

3.14.7 Potential Energy Surfaces

You can use ICE-CI to scan entire potential energy surfaces. In general, the non-parallelity error along a potential energy surface is very small. Thus, ICE-CI yields consistent quality throughout the surface.

For example, let us look at the potential energy surface of the N_2 molecule (Fig. 3.30) – a common test case for quantum chemical methods. There are not too many methods that would dissociate the triple bond of N_2 correctly – ICE-CI is one of them. The potential energy surface is entirely smooth and also correctly behaves in the dissociation limit. Near the minimum it is very close to high-level coupled-cluster methods that, however, all fail badly as the bond is stretched.

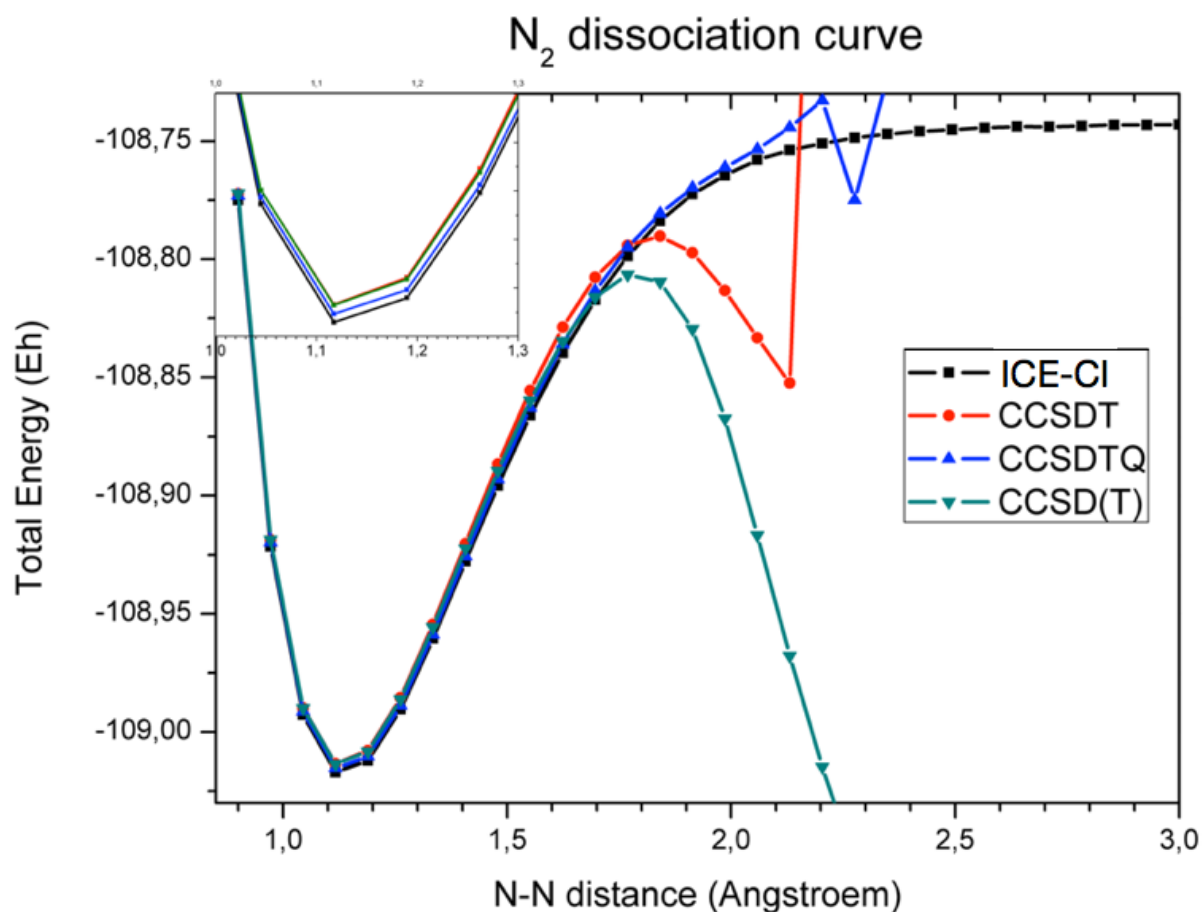


Fig. 3.30: Potential energy surface of the N₂ molecule in the SV basis. For comparison higher level coupled-cluster results are also shown.

It is interesting to observe the variations of the ICE-CI wavefunction along the dissociation potential energy surface. As an example, we look at the dissociation curve of H₂O where both O-H bonds are simultaneously stretched (Fig. 3.31). It is seen that the ICE-CI method is extremely parallel to the full CI curve at all distances. Hence, the description of the bond remains consistent, even when Hartree-Fock becomes a bad approximation. The agreement is particularly good if MP2 natural orbitals are used in the ICE-CI procedure. With the default value of $T_{\text{gen}} = 10^{-4}$ and MP2 natural orbitals the error is consistently below 0.2 mEh. For tighter thresholds, the error is below 0.05 mEh. By contrast, the CCSD(T) method shows relatively large deviations from the full CI results and also behaves very non-parallel as a function of O-H distance.

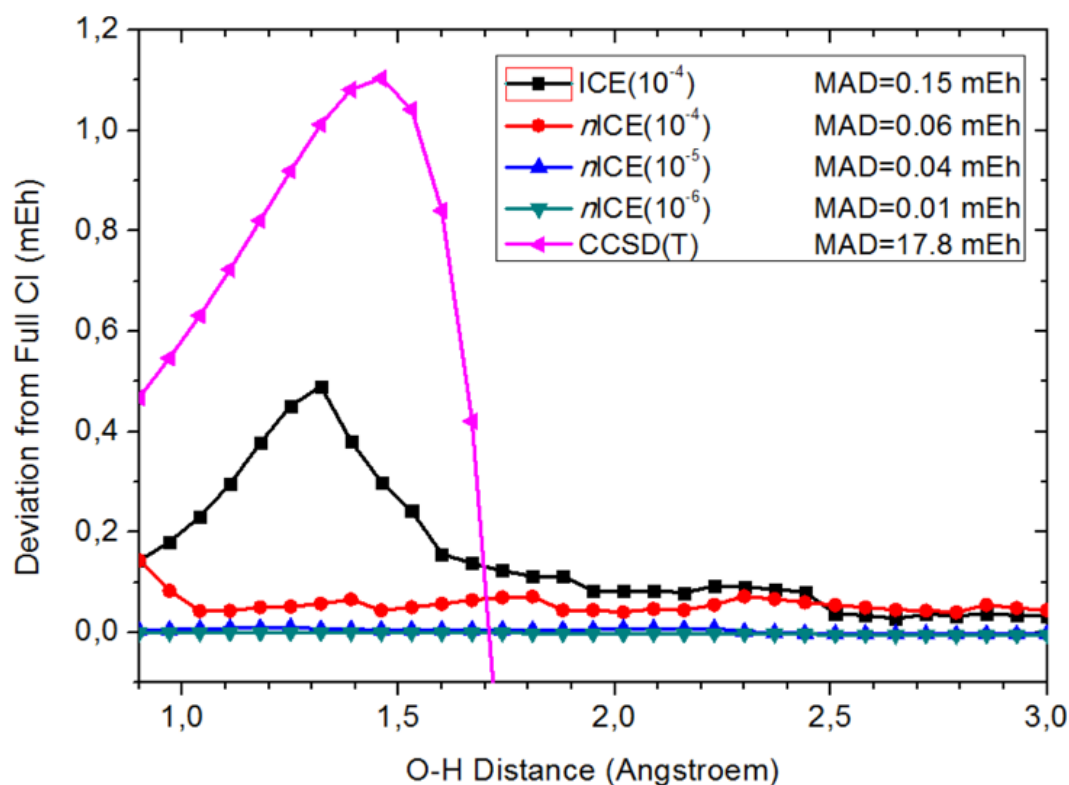


Fig. 3.31: Non-parallelity error of ICE-CI for the H_2O molecule in the SV basis. Shown is the deviation from the full CI value as a function of O-H distance (both bonds stretched). For comparison, the CCSD(T) curve is also shown

It is instructive to analyze the ICE-CI wavefunction along the dissociation pathway (Fig. 3.32). It becomes apparent that the wavefunction stays compact along the entire surface, even in the dissociation limit, where the weight of the Hartree-Fock wavefunction drops to less than 25%. Even in this drastic limit, the ICE-CI wavefunction consists of only about 60000 CSFs, which is very similar to the size of the wavefunction at equilibrium geometry. As the wavefunction becomes more multiconfigurational, the number of generator configurations goes slightly up from the equilibrium value of 77 to a maximum of 118 and finally 112 at dissociation. It is also interesting to note that along the entire dissociation pathway no configuration with more than 8 open shells is generated, which means that no more than quadruple excitations are contained in the ICE-CI wavefunction. The number of iterations required in the ICE-CI procedure also stays constant along the surface at 4 iterations, which impressively shows that a dominant configuration is not necessary for a successful ICE-CI calculation.

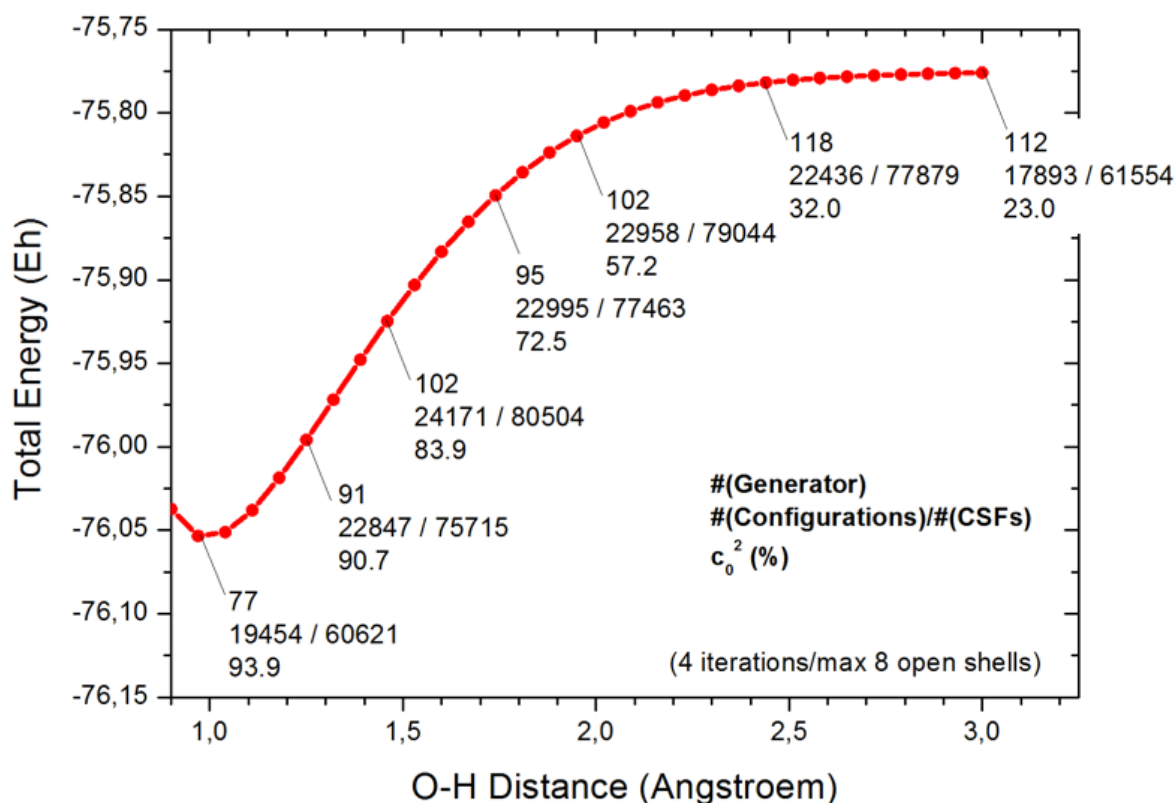


Fig. 3.32: Analysis of the ICE-CI wavefunction along the O-H dissociation pathway.

3.14.8 Excited States

ICE-CI can be used to obtain some insight into excited states starting from no knowledge at all. Of course, the best way to start an excited state calculation is to have some idea which configurations are important for the low-lying states of the system. If this is not the case, an automated procedure is used. The program will first generate an “Aufbau” configuration using the orbitals that are provided on input. Starting from this Aufbau configuration, single excitations at the configuration level are performed and the Hamiltonian is diagonalized for the required number of roots. These roots are then analyzed for the leading configurations and the regular ICE-CI procedure is started from those configurations. For example, look at a calculation on the CN radical. In this case, we know the relevant orbitals and leading configurations for the lowest four roots (a doublet Σ ground state, a doublet Π excited state and a doublet Σ excited state) and hence can provide them in the input file as shown below.

```
#
! cc-pVDZ VeryTightSCF

%maxcore 4096

%casscf nel 7
      norb 4
      nroots 4
      mult 2
      end

%ice nel 9
     norb 26
     nroots 4
     cimode 3
     tvar 1e-11
     tgen 1e-4
```

(continues on next page)

(continued from previous page)

```

      refs { 2 2 2 2 1 }
            { 2 2 2 1 2 }
            { 2 2 1 2 2 }
            { 2 1 2 2 2 }
      end
    end

* xyz 0 2
C 0 0 0
N 0 0 1.07
*
```

The result is shown below. The excitation energies are reasonable but not highly accurate due to the limitations of the basis set (experimentally the doublet Π state is at 1.32 eV and the doublet Σ state at 3.22 eV). There is a very slight symmetry breaking in the doublet Π state that arises from the selection procedure. It should be noted that the state averaged CASSCF excitation energies are 0.25 eV and 3.18 eV.

```

STATE 0 MULT= 2: E= -92.4542949092 Eh W= 0.2500 DE= 0.000 eV 0.0 cm**-1
0.67408 : 22221000000000000000000000000000
STATE 1 MULT= 2: E= -92.3777338429 Eh W= 0.2500 DE= 2.083 eV 16803.2 cm**-1
0.65997 : 22212000000000000000000000000000
0.12092 : 22122000000000000000000000000000
0.11545 : 22221000000000000000000000000000
STATE 2 MULT= 2: E= -92.3776916150 Eh W= 0.2500 DE= 2.084 eV 16812.5 cm**-1
0.69860 : 21222000000000000000000000000000
0.16565 : 22122000000000000000000000000000
STATE 3 MULT= 2: E= -92.3412510872 Eh W= 0.2500 DE= 3.076 eV 24810.3 cm**-1
0.51251 : 22122000000000000000000000000000
0.16825 : 22212000000000000000000000000000
0.11151 : 21222000000000000000000000000000
```

Below, it is described how to do ICE-CI calculations on excited states if the dominant configurations are not known.

3.14.9 Tips and Tricks

ICE-CI can be used very fruitfully together with, say, MP2 natural orbitals. This usually results in results that are closer to full CI results and at the same lead to more compact wavefunctions (it may be called nICE). The use of MP2 natural orbitals is requested by choosing `UseMP2nat true` inside the `%ice` block. Alternatively, improved virtual orbitals can be used (requested by `UseIVOs true`). A comparison is shown in Scheme *Comparison of MP2 natural orbitals and improved virtual orbitals for the ICE-CI procedure (H2O molecule, cc-pVDZ basis, equilibrium geometry)*. It is evident that the calculations based on the MP2 natural orbitals show an error relative to full CI that is almost a factor of two smaller than the corresponding result with canonical orbitals while at the same time the wavefunction is more compact by more than 30%. Hence, the use of MP2 natural orbitals appears to be a very good idea in conjunction with the ICE-CI procedure. This also holds when MP2 itself is a bad approximation (for example in the dissociation limit of the H2O molecule as shown above). On the other hand, the IVOs behave very similar to canonical orbitals and hence, seem to offer fewer advantages.

| | #(Iterations) | #(CSF) | Energy/Eh | Error/ μ Eh |
|----------------------------|---------------|------------------------|---|-----------------|
| Canonical orbitals: | 4 | 85220 | -76.241 306 7 | 518 |
| Improved virtual orbitals: | 4 | 85064 | -76.241 316 6 | 508 |
| MP2 natural orbitals: | 4 | 59447 (0.3% of FCI) | -76.241 434 0 (99.82% of E _{CI}) | 390 |
| Full CI result | | 18818646 | -76.241 824 0 | |

Fig. 3.33: Comparison of MP2 natural orbitals and improved virtual orbitals for the ICE-CI procedure (H₂O molecule, cc-pVDZ basis, equilibrium geometry)

If ICE-CI is used in conjunction with MP2 natural orbitals, there also is the possibility of letting the program automatically choose the active space (this is called auto-ICE). The general idea is simple – we base the active space on the MP2 natural orbitals and their occupation numbers. All orbitals between occupation number say 1.98 down to 0.02 will be included in the active space. A relevant input is shown below.

```
! cc-pVDZ aug-cc-pV6Z/C Auto-ICE
%ice nmin 1.99 nmax 0.01 end

%paras R= 1.0 end

* int 0 1
O 0 0 0 0 0 0
H 1 0 0 { R } 0 0
H 1 2 0 { R } 104 0
*
```

If we scan along the H₂O dissociation surface one can see that despite changing active spaces, the dissociation curves are smooth and remain fairly parallel to the full CI dissociation curve. Depending on the tightness of the thresholds the active space may change from a small 6 electrons in 5 orbitals to a larger 8 electrons in 7 or 8 orbitals upon dissociation. This is the expected behavior as the σ -antibonding orbital becomes more stable along the bond stretching coordinate. Hence, these results are encouraging in as far as in many situations the program will be able to select a sensible active space without extended input from the user.

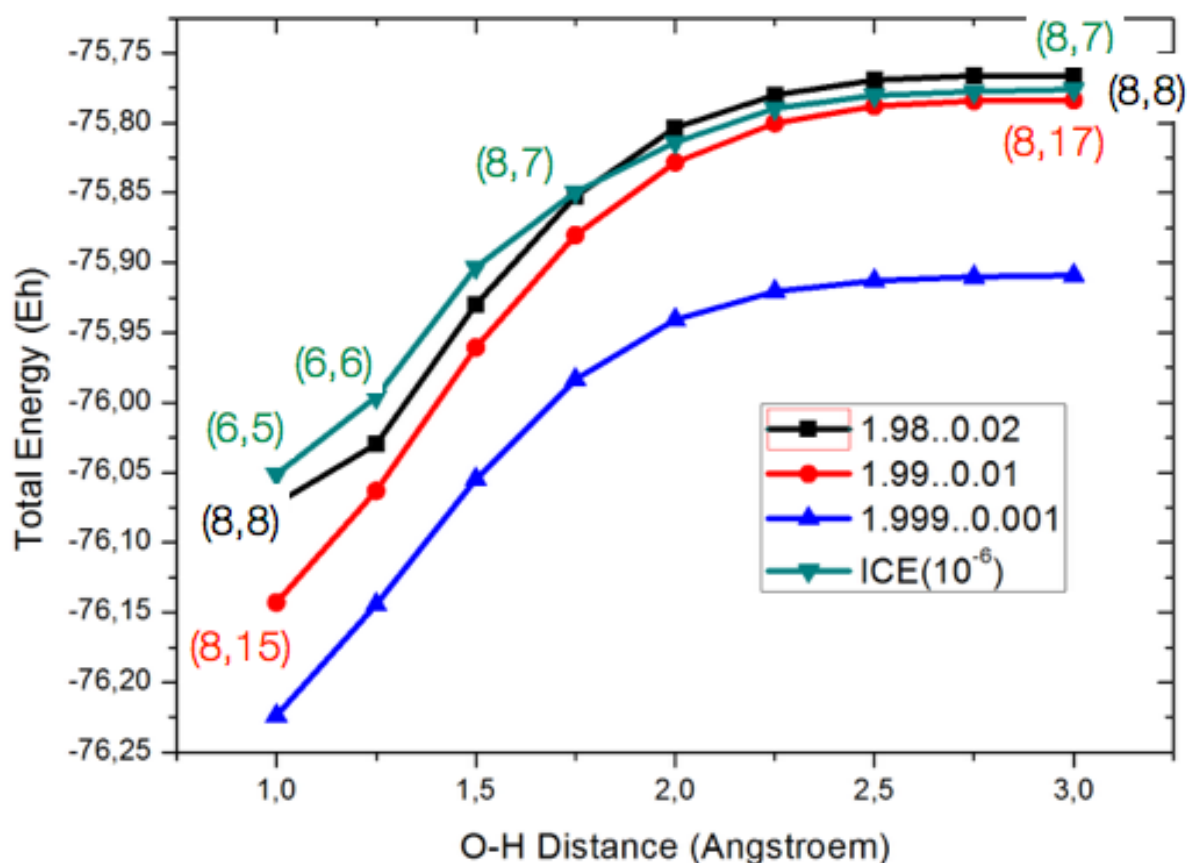


Fig. 3.34: Automatic active space selection along the H₂O dissociation surface. The reference curve (blue triangles) is the ICE-CI method for the full orbital space with the default parameters.

Another place, where automatic selection comes in conveniently is in the calculation of excited states. If there are no user supplied configurations, what happens is that the program will first choose an Aufbau “reference” configuration and then perform all single excitations relative to this configuration. The program will then diagonalize the Hamiltonian over the this set of configurations to create 0th order approximations for the chosen number of roots of interest and then initiate the ICE-CI procedure starting from the leading configurations of these states. Here is an example for the benzene molecule:

```
! RHF def2-SVPD def2-SVP/C Auto-ICE
%cclib "/user/me/orca/cclib/orcacc"
%ice nroots 5
    nmin 1.98
    nmax 0.02
    integrals ri
end

* int 0 1
C 0 0 0 0.000000 0.000 0.000
C 1 0 0 1.389437 0.000 0.000
C 2 1 0 1.389437 120.000 0.000
C 3 2 1 1.389437 120.000 0.000
C 4 3 2 1.389437 120.000 0.000
C 5 4 3 1.389437 120.000 0.000
H 1 2 3 1.082921 120.000 180.000
H 2 1 3 1.082921 120.000 180.000
H 3 2 1 1.082921 120.000 180.000
H 4 3 2 1.082921 120.000 180.000
H 5 4 3 1.082921 120.000 180.000
```

(continues on next page)

(continued from previous page)

```
H 6 5 4 1.082921 120.000 180.000
*
```

(The %cclib statement is explained below and is not mandatory here). The Auto-ICE procedure comes up with as many as 24 electrons in 19 orbitals, which already is a fairly heavy calculation. The procedure converges in five iterations and provides indeed the correct states: the ground state, the $^1B_{2u}$ state at 6.4 eV, the $^1B_{1u}$ state at 8.9 eV and a degenerate $^1E_{1u}$ state at 10.0 eV. These excitation energies are still in error by about 2 eV relative to experiment, which is mainly due to missing dynamic correlation. However, the correct states and their sequence has been found.

The ICE-CI can be used to find the ground state if the actual ground state is not known. To this end, one simply has to turn off the selection steps. This makes the calculations more expensive, but they will converge to the lowest state. In the example below (again, the H_2O molecule) we start from a random quintuply excited configuration – the ICE-CI still finds the ground state after four iterations:

```
%ice
  nel      8
  norb     23
  nroots    1
  tvar 1e-11
  tgen 1e-04
  etol 1e-06

# selection
SelStart false
SelIter  false

# algorithm details
useivos    false
integrals  exact
cimaxdim    5 #Davidson expansion space = MaxDim * NRoots
cimode      3

# spatial sym (only supported for ICEType = CFGs)
irrep 0
# startup (optional)
refs { 2 1 0 1 0 2 1 0 1 }
      end
end
```

However, if one wants to converge to an excited state, one should turn on the selection. In the example below (once more the water molecule) one can converge to the second excited singlet state by judicious choice of the start configuration:

```
%ice
  nel 8
  norb 23
  nroots 1
  tvar 1e-11
  tgen 1e-04
  etol 1e-06

# selection
SelStart true
SelIter true

# algorithm details
useivos false
integrals exact
cimaxdim 5
cimode 3
```

(continues on next page)

(continued from previous page)

```

# spatial sym (only supported for ICEType = CFGs)
irrep 0
{ #} startup (optional)
refs
  { 2 2 1 2 0 1 }
end
end

```

3.14.10 Large-scale approximate CASSCF: ICE-SCF

ICE-CI can be used in place of the default `CISStep` in a *CASSCF* framework. In this way, much *larger CASSCF calculations* than previously possible can be envisioned. In using the ICE-CI in this way, the active orbitals should be chosen as natural orbitals in order to ensure a proper canonicalization. In general, ICE-CI results will not be invariant with respect to the choice of orbitals. However, in practice we have not found this to be problematic. We refer to this as ICE-SCF.

The use is simple: In the `%casscf` block choose `CISStep ICE`. Optionally, starting configurations (CFGs) or configuration state functions (CSFs) can be defined in the `refs` sub-block. All other parameters are accessible through the 'CI' sub-block in `%casscf`. The program also supports *FIC-NEVPT2 calculations with the ICE* as reference wavefunction.[459]

```

%casscf
...
# mandatory
CISStep ICE

# optional starting references for ICE
# using the refs sub-block
# refs {2 2 ...} end

# optional refined settings
# with keywords as in the %ice block
ci
  ICEType CFGs      # select ICE expansion type:
                    # 'CFGs' (default), 'DETS' or 'CSFs'
  TGen 1e-4         # controls accuracy (default = 1e-4)
  TVar 1e-11        # default = 1e-7 * TGen
  MaxIter 100       # number of allowed cycles (default = 64)

  # Scratch space dimension (number of singles/double excitations per CSF)
  # If scratch is too small, the program may crash.
  CIBufferLength 10240 # ~82kB (factor 8/103 to get the memory in kB):
end
end

```

The entire remaining input is the one for standard CASSCF calculations. In this way one can do CASSCF calculations with very large active space in reasonable turnaround times. We have not observed convergence problems that are worse than in the standard CASSCF procedure. The results in Fig. 3.35 show that the deviations from regular CASSCF energies are very small. The largest deviation observed for C₂H₄ is on the order of 0.2 mEh, which appears acceptable. Note that the *CASSCF tutorial* also covers larger examples and excitations energies computed with the ICE-CI as CI solver.

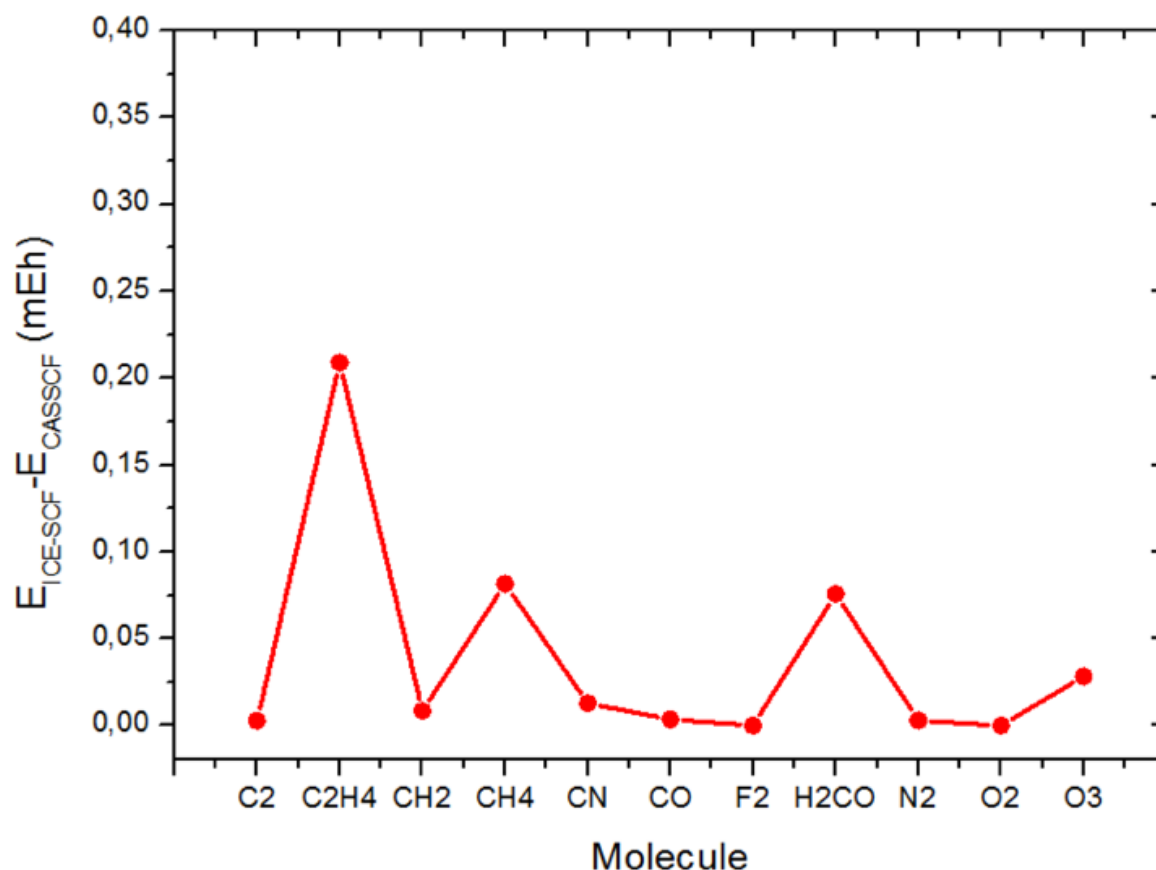


Fig. 3.35: Deviations of ICE-SCF from CASSCF energies for a selection of molecules (standard truncation parameters $T_{\text{gen}} = 10^{-4}$ and $T_{\text{var}} = 10^{-11}$)

Since CASSCF is fully variational, it is possible to optimize geometries with that procedure. It is our experience so far, that the ICE-SCF geometries are virtually indistinguishable from CASSCF geometries (an example is shown in Fig. 3.36).

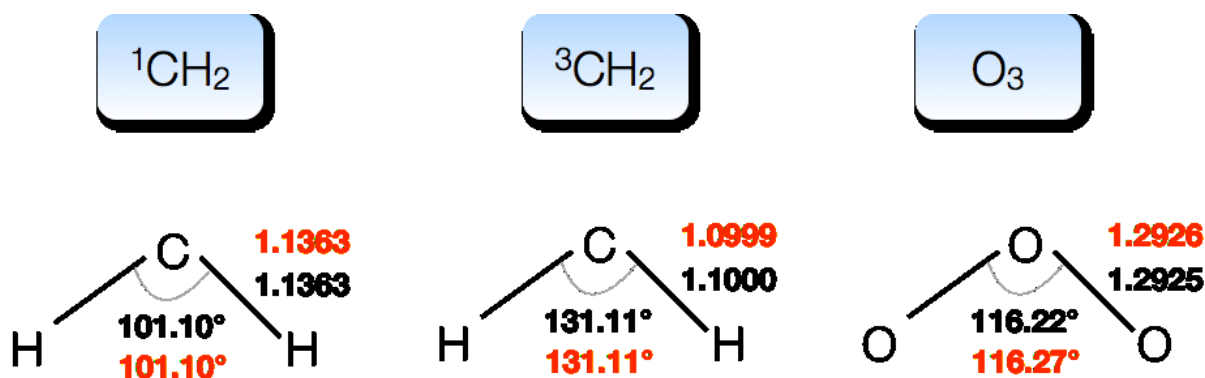


Fig. 3.36: CASSCF and ICE-SCF optimized geometries for methylene and ozone (cc-pVDZ basis set, default parameters).