Some practical HINTS
on how to use MOLCAS
in quantum chemistry problems

Luis Serrano-Andres and Valera Veryazov

`version 0.98`

13 Aug. 2010

Outline

- 1.-Some comments about different programs: SCF, RASSCF, CASPT2, RASSI

- 2.-Selection of active spaces

- 3.-About structure optimizations

- 4.-Environmental effects

- 5.-Comments about hardware

- 6.-Final admonitions and request

GATEWAY/SEWARD program:

- Try the Choleski approximation (or RI)! It saves disk space and possibly time.

- Think about basis set. ANO-like basis sets have many advantages, but they are "marginal".

- Try to avoid inline basis sets, use the library.

- Remember that the quality of basis set should match quality of computational method.

- Use ANO-RCC even for atoms in the 2nd row.

- Be extremely careful when computing anions. Remember that special situations requires special basis sets.

- Use minimal or small basis set for understanding chemical problem. You always can use expbas later..

SCF program

- HF orbitals are in many cases good starting orbitals, but quite often GuessOrb can be used instead.

- Very large basis set together with HF can lead to linear dependences.

- Remember! Hartree-Fock method allows multiple solutions (even for trivial systems)

- Be reasonable selecting convergence thresholds

- UHF convergence is much poor comparing to RHF.

- Sometime you have to use Scramble keyword to break the symmetry.

SCF program (DFT):

- DFT convergence can't be better that HF convergence. Think about starting orbitals for DFT.

- Remember that DFT is a powerful method but it is still single configurational method. Don't use it beyond it's limits.

- Choose your favorite functional, and stay with your decision

- Note that MOLCAS is not the best DFT code available

Hints for using MOLCAS in quantum chemistry

1.- **Some comments about different programs:**

> **RASSCF program:**
>
> • MCSCF are multi-solution methods that heavily depend on the starting orbitals and level of calculation (roots). Try to find the most convenient for your case.
>
> • On convergence ALWAYS (ALWAYS, ALWAYS, etc) check the orbitals (gv, molden,…). MCSCF methods will lead to different solutions for active spaces of different nature. Use your chemical intuition and let the calculation guide you.
>
> • Increase LEVShift parameter in cases of slow or difficult convergence.
>
> • Sometimes RASSCF will diverge when is close to convergence. Try restarting the calculation from the previous JOBIPH file linked as JOBOLD
>
> • Analyze carefully the CI coefficients and natural occupation numbers together with the orbitals (average orbitals are fine in general for that) in order to understand the nature of the states

Hints for using MOLCAS in quantum chemistry

RASSCF program:

• Remember that *rasscf.molden contains the average orbitals, and *rasscf.molden.1, .2, etc, contain the natural orbitals of the different roots.

• Try increasing the number of SA-CASSCF roots to locate more excited states. They can be low-lying solutions at the CASPT2 level. In high symmetry cases you may also need to consider roots that have high energy at the initial steps and can become lower roots in the converged calculation. Note that the keyword CIROOT allows you to obtain more roots than you optimize the orbitals for. For example the input CIROOT; 3 6; 1 2 3; 1 1 1 will make a SA-CASSCF calculation optimizing the orbitals for the three lowest root, but give you also the energies and wave function for three more roots.
For transition metal or heavy elements molecules it might be a bad idea. Too many roots may mean unproper mixing of orbitals.

• It is NOT advisable to play games with weights for the different roots. Roots with equal weights make your calculation more clear and reproducible.

1.- **Some comments about different programs:**

RASSCF program:

• In high symmetry cases:

    ✓ Work in a symmetry point group that allows degenerate states to belong to the same irreducible representation (e.g. C_2 for linear molecules). Try C_1 too.

    ✓ Start with clean orbitals (GUESSORB or SCF normally produce them) and make use, if needed, of CLEANUP and SUPSYM, or for linear molecules: the LINEAR keyword.

    ✓ Working in a too high symmetry might prevent you from obtaining less symmetric lowest-lying localized solutions (e.g. Ni^{2+})

    ✓ To obtain good, symmetric starting orbitals (e.g. all d orbitals equal) you may make a previous calculation where you fill or empty some orbitals. Equally occupied orbitals tend to be equal. For example a calculation on a radical can be started by a closed shell calculation on the positive ion.

RASSCF program (RAS):

- use it! RASSCF is a simple way to increase an active space.

- Balance RAS1/3 and RAS2 subspaces. Try to change orbitals between these subspaces.

- Removing RAS2 space completely is not a good idea.

- Note that RAS calculations have a slower convergence, and demand more resources.

Hints for using MOLCAS in quantum chemistry

1.- **Some comments about different programs:**

CASPT2 program:

• The new IPEA = 0.25 zeroth Hamiltonian is the default from molcas-6.4.
  It particularly improves open shell cases.

• Energy differences between different states or situations are only reliable between
  calculations with the same active space size and similar reference weights in CASPT2.

• An intruder state (low reference weight in the CASPT2 state) might be informing you
  that your active space lacks an important orbital. Check the list of large perturbative
  contributions (small denominators combined with large RHS values; check the output)
  and also the occupation number of the CASPT2 orbitals.

• For weakly interacting intruder states cases try the IMAGINARY level shift parameter.
  0.05 or 0.1 au is typically enough. It is wise to compute a series: 0.0, 0.05, 0.10, 0.15…
  and check that the result is converged. Then, take the lowest value that solves your
  problem. Beware of using too large level shifts (not larger than IMAGINARY 0.20).
  Don't use the level shift to reach agreement with experiment!

1.- **Some comments about different programs:**

CASPT2 program:

• For heavy valence-Rydberg mixing cases or for closely degenerated CASPT2 states, use MS-CASPT2.

• If the MS-CASPT2 description differs a lot from the CASPT2 one, try to check the calculation by increasing the active space (introducing angular correlation if possible) until the result is converged. The "true" solution is typically between both cases (CASPT2 and MS-CASPT2). If you are suspicious about the MS-CASPT2 result, better keep the CASPT2 one. It has worked out generally well so far.

Hints for using MOLCAS in quantum chemistry

1.- **Some comments about different programs:**

RASSI program:

• Remember that the program shows first the interaction among the input states and later this description might change (because the states order and nature change) in a second part of the output. In general, ALWAYS check the changing order of states.

• For spin-orbit coupling calculations don't forget to include the CASPT2 energies as input (EJOB or HDIAG keywords) because the results depend on the energy gap. In other cases having the CASPT2 energies as input will help you to get the right oscillator strength and Einstein coefficient in the final table.

• If you have degenerate states be sure that the CASPT2 energies are degenerate. If they are not (which is common) average the energies for the degenerate set (the two components of E symmetry for example).

• Remember that the spin-orbit coupled results (e.g. TDM) depend on the number of interacting singlet and triplet states included in RASSI.

2.- **Selection of basis sets and active spaces:**

Active space selection (Check Lecture 2 notes).

• Always compare calculations with the same active space size (and nature if possible).

• Ask yourself first which is your goal. The selection of the active space depends on that.

• In ground state calculations many orbitals can have an occupation close to 2 (0) and might rotate with others in the inactive (secondary) space. It might be wise to skip them (e.g. lone pair orbitals).

• For low-lying excited states and few roots you might leave inactive quite a number of orbitals. Check with RASSCF for instance. RASSCF(SDTQ) is as good as a CASSCF calculation with the same orbital size.

• SCF orbital energies sometimes help to choose the orbitals by using the energy order criterion, but you must learn to see the problems (like lone pair orbitals having too low energies at the SCF level).

2.- **Selection of basis sets and active spaces:**

Active space selection (Check Lecture 2 notes).

• You typically will need correlating orbitals, that is, if you have a π orbital you need a π^*, the same for σ,σ^*, but not for lone pairs.

• CASSCF/RASSCF geometry optimizations are the worst case. If you miss orbitals you might end up in a totally wrong geometry (e.g. breaking a bond usually requires the bonding and antibonding orbitals in the space).

• Organic (1st row atoms) molecules usually require open shell orbitals, π, π^*, and lone pairs. If 2nd row atoms are added (S, P, Si, etc) s orbitals enter in action (s bonds are longer). CH bonds can often be left be inactive.
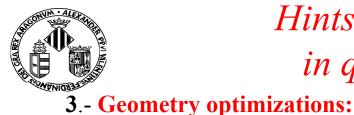

2.- **Selection of basis sets and active spaces:**

Active space selection (Check Lecture 2 notes).

• Rydberg states require additional diffuse basis sets and specific orbitals in the active space. Use basis sets centered in the charge center of the positive ion (consult the manual, page 385).

• Transition metal chemistry (1st row) sometimes requires a double d shell description in the active space.

• Lanthanides have a quite inert 4f shell that must be active together with 5d, 6s (6p). Actinides 5f, 6d, 7s.

Active space: a final advice

- **use expbas!** start from minimal basis set, decide the active space, and expand the basis to "normal". With small basis set you can clearly identify orbitals.

- **use localization!** Especially for virtual orbitals.

- **expand active space by adding RAS1/3** - give the system a freedom, and see how it reacts.

3.- **Geometry optimizations:**

- For minima on flat hypersurfaces, such in loosely bound fragments, or in slow convergence cases you might have to decrease the CUTOFF threshold in ALASKA

- Be careful with the bond angle definition if you are close to a linear bond.
  You may have to switch to the LAngle definition

- Don't forget that CASSCF does not include dynamical correlation. In some cases you better change to DFT or numerical CASPT2 optimizations or, if this is not feasible, may be preferable to run RASSCF optimizations

- Poor active spaces may lead you to symmetry broken wrong solutions (e.g. a Cs minimum for water below the true C2v one)

- Poor geometry convergence might be reduced or at least controlled by reducing the initial trust radius with the MAXSTEP keyword or/and by doing the optimization in Cartesian coordinates (CARTESIAN)

3.- **Geometry optimizations:**

• In order to obtain localized solutions it might be a good idea to feed the program with a slightly distorted geometry that helps the method to reach the non symmetric solutions. Other possibilities are to use an electric field, to add a charge far from the system or use a solvent cavity. In all cases you break symmetry and allow less symmetric situations.

• Linearly interpolated internal coordinates geometries may be a good starting point to locate a transition state. Use also the useful FindTS command. Sometimes can be wise to compute a MEP from the TS to prove that it is relevant for the studied reaction path. Try also the new Saddle approach!

• When locating a CASSCF surface crossing (MECP) ALWAYS compute CASPT2 energies at that point. The gap between the states can be large at that level. In severe cases you might have to make a scan with CASPT2 to find a better region for the crossing.

• Remember that (so far) MOLCAS does not search for true conical intersections (CIs) but minimum energy crossing points (MECP) because it lacks NACMEs. Note however that typically computed minimum energy CIs (MECIs) may not be photochemically relevant if they are not easily accessible. Barriers have to be computed. Use MEPs!!

Hints for using MOLCAS in quantum chemistry

3.- **Geometry optimizations:**

• Numerical hessians and optimizations may lead you to bad solutions when different electronic states are too close. As you move your calculation from the equilibrium geometry some of the points may belong to other state and corrupt your result. This might be the case for numerical CASPT2 crossing searchs. Use then MS-CASPT2 searchs.

• Remember that SA-RASSCF analytical gradients and SA-CASSCF analytical hessians are not implemented.

• Be careful with the change of roots and nature along a geometry optimization or MEP. For example, you start with the state in root 3 (at the CASSCF level) and reach a region of crossing root 3 – root 2. You may need to change to root 2 for your state. Not an easy solution (so far).

4.- **Solvent effects:**

• Some effects of the solvent are very specific, such as hydrogen bonds, and require to include explicit solvent molecules. Try adding a first solvent shell (optimized with molecular mechanics for instance) and then a cavity, for instance with PCM.

• Too small cavity sizes can lead you to unphysical solutions, even if they seem to match experiment.

• Remember using NonEquilibrium (final state) and RFRoot (SA-CASSCF) when required

• QM/MM is a much powerful strategy, but it requires experience and knowledge of the field

Hardware/software issues

How NOT to misuse MOLCAS?

- Do not play too much with code optimization..
  we did it for you!

- remember: Memory can be more important than CPU speed.

- do not overload MOLCASMEM

- test GA before compiling molcas in parallel

- be very careful if run MOLCAS in parallel
  think: where do you run it, what memory is used, etc.

- MOLCAS has cross node parallelization
  use multicores wisely.

Hardware/software issues

- GPUs can be used, but do not expect too much! they are designed for different codes!

- Solid state devices - very good for conventional integral codes, but average for CD/RI.

5.-Final admonitions and request:

- Not matching experiments may sometimes be good news, either because we are comparing different things or because our models do not consider all the required effect. We have to use our calculations to EXPLAIN and PREDICT the experimental data, not to simply have "good agreement". Consider that the experiment might be wrong too.

- Don't try to give "final" answers if you don't have the accuracy.

- If your calculation is not "conclusive" it might be worthless.

- Theoretical models are heavily needed. Try to give chemists chemical explanations. They don't like numbers.

5.-Final admonitions:

- (Not to be listened by experimentalists: do not feel that theory is always inferior to experiment. In many cases experimentalists need us more than we need them, because they don't know at all what they have measured).

- a word about CCSD... Don't think that there is (somewhere) a superior method, which is always correct.

Final request:

Please, help us to improve MOLCAS:

- Send us bug reports and suggestions through the web page

- Participate in the discussion forum: add questions and answers

- If you make useful tools related to MOLCAS for your personal use think in sharing them. They might be added to the package.

- Remember the developers that there is pending tools deeply needed. The list of priorities may change in that way.

- Now your turn: problems, hints, and ideas.