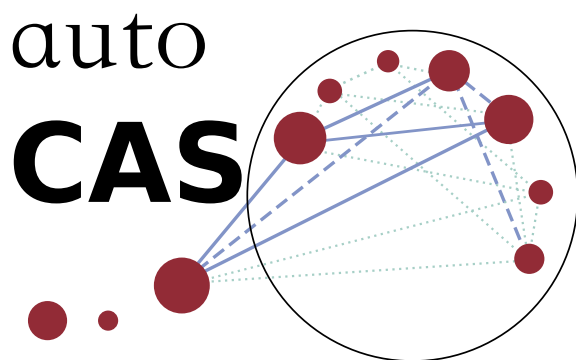THE SCINE AUTOCAS DEVELOPERS:
CHRISTOPHER J. STEIN, MAXIMILIAN MÖRCHEN,
AND MARKUS REIHER

auto
CAS

Version 1.1.1 – Mai 2022

MANUAL

ETH ZÜRICH

Contents

4

1

Introduction

Multi-configurational calculations are indispensable for many areas of computational chemistry. Whenever a molecule has several close-lying frontier orbitals, a common situation for polyaromatic molecules, transition metal complexes, and transition states in general, the approximation of the wave function with a single Slater determinant is inadequate and multi-configurational approaches are required. To provide a well-defined electronic-structure model, the most common multi-configurational methods require the selection of a subset of orbitals and electrons, the so-called *active* orbitals and electrons, for which all configurations are included in the final wave function. Unfortunately, this selection is mainly based on empirical guidelines and experience. This unfortunate fact makes this most general class of wave function calculations a niche for experienced experts.

With the AUTOCAS program, we base the active orbital selection on rigorously defined orbital entropy measures that can be obtained from qualitatively correct but only partially converged density matrix renormalization group (DMRG) calculations. The DMRG approach is chosen to be able to respect large orbital spaces and the entropy measures are a diagnostic that, by contrast to natural occupation numbers, show a large spread that allows for a better discrimination of the various electron-correlation situations. Based on these entropy concepts, AUTOCAS follows an automated selection protocol[1] that fully automates the active orbital selection procedure up to the point where multi-configurational calculations turn into black-box methods.

In this manual, we describe the installation of the software, describe in detail two example calculations as an introduction into the program and describe the most import functions and options.[2] We further describe the most common scenarios in which AUTOCAS can be used and how the necessary template files can be manually modified for a specific project.

A prospect on features in future releases and references for further reading are added at the end of this manual.

[1] See C. J. Stein, M. Reiher, *J. Chem. Theory Comput.* **2016**, *12*, 1760.

[2] Throughout this manual, the most import information is displayed in the main text, whereas useful additional information is given as a sidenote like this.

2

Obtaining the Software

THE AUTOCAS GRAPHICAL USER INTERFACE is distributed as binary software in the framework of the SCINE project (www.scine.ethz.ch). To obtain the software you need to enter your name and e-mail address on our website (www.scine.ethz.ch/download/autocas), choose your operating system and agree to the end user license agreement (EULA). The download will start immediately after this registration.

Note that you should start AUTOCAS from your terminal on macOS because if you start it with a double click on the icon in the *Applications* folder the operating system might not allow to start OpenMolcas calculations through AUTOCAS.

2.1   *System Requirements for DMRG Calculations*

The DMRG calculations are usually memory bound. In QCMaquis, the calculation of the entanglement entropies requires the transformation of the matrix-product-state wave function from a compact spin-adapted form to a larger, non spin-adapted representation. This step requires a substantial amount of memory and we recommend to run these calculations on machines with at least 64 GB of memory (or even 256 GB for active spaces with more than 50 orbitals).

3

Installation

Because autoCAS is distributed as a ready-to-use binary file, no installation is required. Since autoCAS interfaces the electronic structure program OpenMolcas and our DMRG program QCMaquis, you have to make sure that these two programs are successfully installed. Information on these installations can be found on the QCMaquis website (www.reiher.ethz.ch/software/maquis) and the QCMaquis manual that you will find on those web pages.

3.1   Notes for macOS

If you decide to start autoCAS from another folder you have to set the QT_PLUGIN_PATH accordingly in the terminal where you want to run autoCAS:

```
export QT_PLUGIN_PATH=path/to/your/folder/autoCAS.app/Contents/PlugIns
```

After you unpacked the .tar archive for macOS you have a folder called *autoCAS.app*.[1] In order for your operating system to find the required plugins you should move this folder to */Applications*.

[1] Note that this is really a folder.

3.2   OpenMolcas / QCMaquis Test

We refer to the main folder of the autoCAS program as the *autoCAS home* folder. In this folder you will find a folder called *tests/interface*. There, you will find a small sample input (ethane.inp) that you should run in order to confirm that your OpenMolcas / QCMaquis installation is ready for use with autoCAS. If the calculation finishes without errors you are ready to start autoCAS. Otherwise you might want to consider contacting us or the QCMaquis developers for support.

In order to run calculations with autoCAS, you need to select a *run file*[2] that specifies modules to be loaded, sets the paths for Open-Molcas and QCMaquis, contains settings concerning memory and the number of cores and enables autoCAS to set a project name.

[2] This file has the file extension *.env* and does not need to be executable. autoCAS will automatically modify these settings.

In principle, this script should be very similar to your usual Open-Molcas run script. The following example should allow you to set up such a file for your installation:

```bash
#!/bin/bash

# set MOLCAS
export myMOLCAS=/Users/YourName/devel/openMolcas/molcas/build
source $myMOLCAS/qcmaquis/bin/qcmaquis.sh

export QCMaquis_CPUS=4

input=$1
project={{project_name}}
scratch=/scratch/${USER}

#mkdir -p ${scratch}/${project}

CurrDir=$PWD

export MOLCAS=$myMOLCAS
export MOLCAS_PROJECT=$project
export MOLCAS_MEM=12000
export MOLCAS_WORKDIR=${scratch}

pymolcas -b 1 -f $CurrDir/$input
```

[3] These two lines are set during your installation of Open-Molcas with the QCMaquis interface. Please just copy them here.

[4] As QCMaquis is parallelized we recommend to set the number of CPUs to the maximum number of cores available on your system.

[5] autoCAS expects the OpenMolcas input file as the first argument.

[6] Do not change this line. autoCAS will fill in the correct project name and requires exactly this name and format.

This file needs to be prepared only once and can be reused for all future AUTOCAS calculations with the specified installation of OpenMolcas and QCMaquis. We recommend to copy this *run.env* file to your *project folder* (e.g. where you keep the *.xyz* file) before you start a calculation.

You are now ready to start your first multi-configurational calculation with AUTOCAS.

4
Example Calculations

To give you a first introduction to the AUTOCAS program, we present here two step-by-step example calculations that guide you from the initial set up of your multi-configurational calculation to the final calculation including the calculation of dynamical correlation.

4.1   *Example 1: Step-by-Step – Butadiene*

Prepare the calculation by setting up a *project folder* and copy the *run.ev* file that you created in the last chapter and the *butadiene.xyz* structure file from the *autoCAS home/tests* folder.

Now, you can start the AUTOCAS program. Upon start, the main screen will look like this:
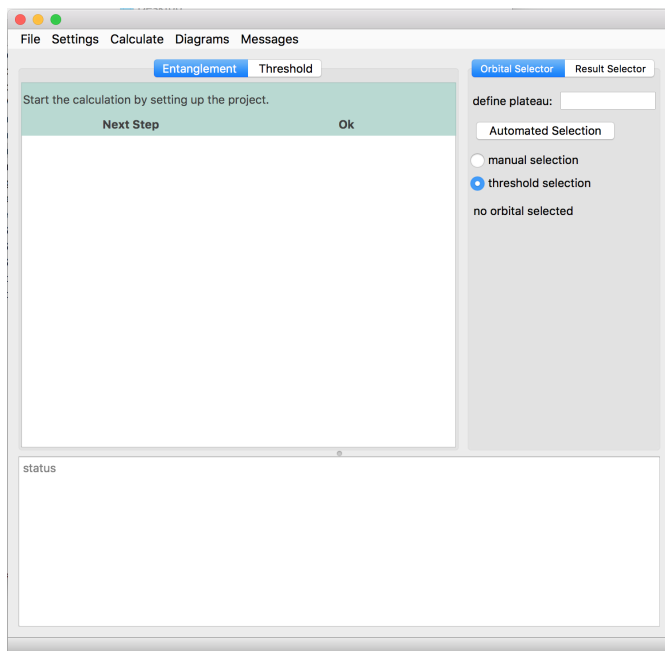

Figure 4.1: AUTOCAS main screen. The popup messages in the main screen give status information, warnings, or, as shown here, guide to the next step of the calculation.

The main screen is described in more detail in the following chap-

ter. With a click on *Next Step* in the popup message located in the center window you open the *Project Designer* with which you will set up your calculation.
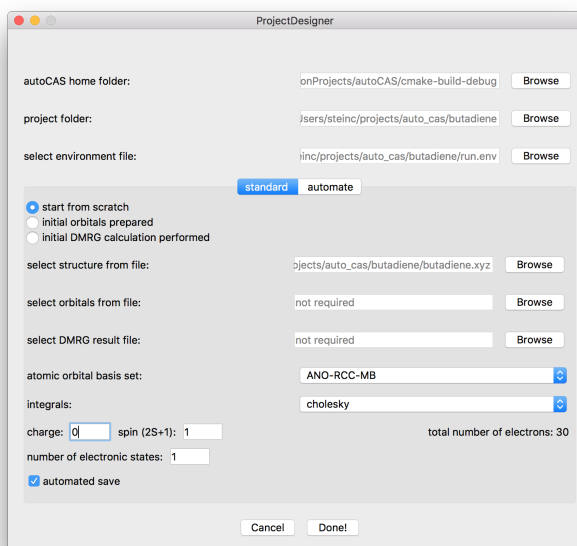

Figure 4.2: AUTOCAS Project Designer. All settings for the calculation are defined here.

We briefly discuss the settings chosen for this example:

1. **autoCAS home folder:** This is the main folder of the AUTOCAS installation. It contains all template files required by AUTOCAS.[1] The path is set automatically and hardly ever needs to be changed.

   [1] The JSON formatted template files will be discussed in the 'Customize template files' chapter.

2. **project folder:** Allows you to select the project folder for the current calculation. Here, we select the folder containing the *run.env* and *butadiene.xyz* files we set up before starting AUTOCAS.

3. **environment file:** This is the *run.env* file. A correctly setup environment file is necessary for OpenMolcas and its QCMaquis interface to work.

4. **start selection:** We choose here to start from scratch meaning that we have calculated nothing yet and start only from the *.xyz* structure file.[2]

   [2] The other options will be discussed in the 'Possible scenarios' chapter.

5. **structure file:** Here, we select the *butadiene.xyz* file.

6. **basis set:** Select the default *ANO-RCC-MB* atomic orbital basis set. For most calculations it is convenient to start with a minimal basis set and project the pre-optimized orbitals on a larger basis with the OpenMolcas EXPBAS module.[3]

   [3] See the OpenMolcas pages for a documentation of this module.

…

7. **integrals:** We choose Cholesky decomposition of the integrals.

8. **charge and spin:** AUTOCAS determines the number of electrons in the system from the number of protons as evident from the structure file and the charge specified here. The default spin is always the lowest spin possible but can be adjusted manually.

9. **electronic states:** As we are only interested in the ground state, we keep the default value of one electronic state to be calculated.[4]

10. **automated save:** We recommend to always keep that box checked for automated saving of intermediate results and therefore to allow for a simple restart option.

Note that we did not change any settings on the *automate* tab and therefore will run a calculation where we start each individual step manually. A fully automated calculation will be demonstrated in the next example.

After you clicked the *Done* button you will be asked to give a name to your calculation.[5] Your calculation is now set up and you can generate an initial orbital basis by clicking on *Next step* on the popup message appearing after closing the Project Designer.

[4] Please always select the number of electronic states at this point. It is not possible to change this setting later.

[5] AUTOCAS usually suggests the name of your structure file. The project name and therefore the name of your scratch directory will be named according to the name you give here.
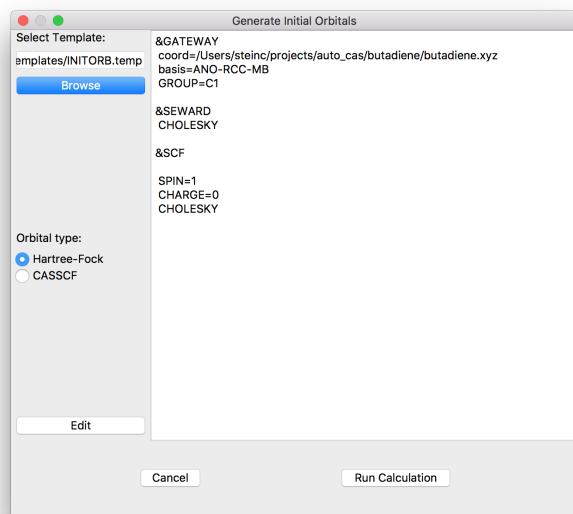
Figure 4.3: Initial orbital generation screen. For Hartree–Fock orbitals you do not need to change anything. If you want to start your calculations from previously prepared CASSCF orbitals (e.g. from another structure), you need to manually insert the path to the hdf5 orbital file and the number of active electrons here.

Since we want to start our initial DMRG calculation from Hartree–Fock orbitals in this example we do not need to specify anything in this step and can click on *Run Calculation* right away. AUTOCAS will then setup a folder structure, copy the *run.env* file with the correct project name and make it executable, write the OpenMolcas input

file and start the calculation. Two popup messages will appear on successful completion of the orbital generation calculation: a status message that informs you about the completion of the last step[6] and a message guiding you to the next step as seen before.

A click on *Next step* opens the initial orbital selection screen.
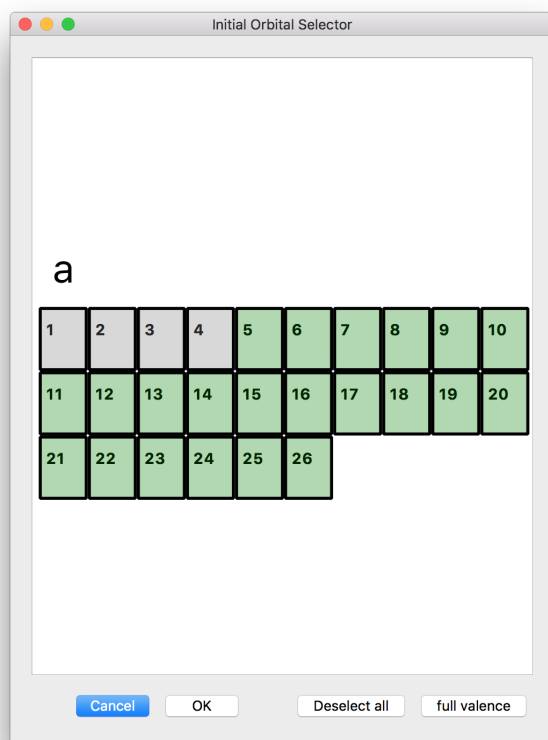
Figure 4.4: Initial orbital selection screen. Each orbital is represented by a numbered card and ordered according to its irreducible representation.

In the orbital selection screen each of the molecular orbitals that were generated in the last step are represented by a card with a number ordered according to their irreducible representation. Please click on *full valence* to select all valence orbitals of this molecule.[7] This option is available for Hartree–Fock orbitals and for CASSCF orbitals (only in C_1 symmetry). Confirm the selection of the orbitals for the initial DMRG calculation by clicking on *Ok*. AUTOCAS will now generate a *dmrg* subfolder and copy a modified version of the initial orbitals to this folder.

A *status message* informs you that we selected an initial active space with 22 electrons in 22 orbitals. Please click on *Next step* again to open the initial DMRG calculation preparation screen.

Since we do not have to adjust any parameters in this calculation,

[7] Press the key 'F' to see which orbitals are identified as valence orbitals prior to the selection.

please click on *Run calculation* to start the initial DMRG calculation.[8]

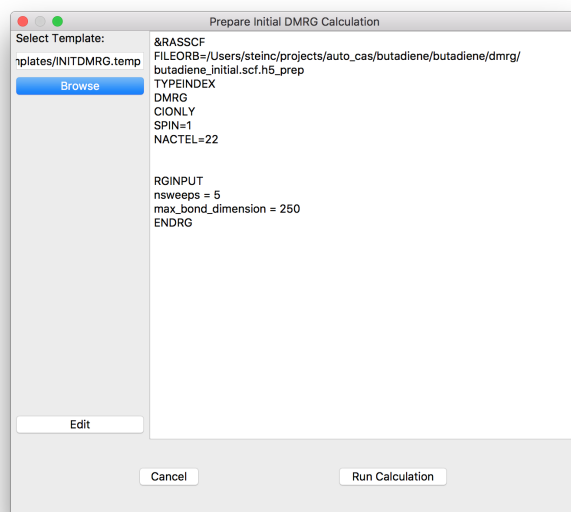AUTOCAS will start the initial DMRG calculation through the OpenMolcas / QCMaquis interface now. Note that this is usually by far the computationally most demanding part of the calculation. Initial DMRG calculations with more than 80 active orbitals are likely to crash because of huge memory demands that will not be available on most systems. [9]

Since this will take several minutes, you might try a right click on the entanglement screen (the white space in the center where the popup messages appear). A short menu opens that allows you to display information about the general settings of your current AUTOCAS session and the settings of the current calculation.

After the DMRG calculation has finished, the entanglement information is automatically loaded and displayed as mutual information diagram and a threshold diagram. [10] You can switch between both diagrams via the *Entanglement / Threshold* tabs located at the top center of the main screen.

With a click on *Automated selection* you evoke the automated selection of the final active orbital space based on the entanglement entropies that were calculated from the initial DMRG calculation. The initial suggestion of selecting all 22 orbitals is rejected by the program because of the low multi-conifgurational character of the wave function as measured by the $Z_{s(1)}$ diagnostic.[11] You confirm this with a click on *Ok* in the corresponding warning dialog and the program selects a final active space consisting of four electrons in

[8] Sometimes you might want to make sure that the entanglement information is converged with respect to the DMRG settings. See S. Keller, M. Reiher, *Chimia* **2014**, *68*, 200 or arXiv:1401.5497. In that case please raise the number of sweeps and the maximum bond dimension here.

[9] Consider to setup your calculation on your personal laptop or workstation and carry out this expensive step on a compute cluster. To do so, copy your scratch folder and your project folder to the cluster, start autoCAS and read in the *save.cas* file. You can then continue the calculation from this point.

[10] See C. J. Stein, M. Reiher, *J. Chem. Theory Comput.* **2016**, *12*, 1760. for a detailed description of the selection protocol and the two types of diagrams.

[11] See C. J. Stein, M. Reiher, *Mol. Phys.* **2017**, *115*, 2110. for more information on the definition of $Z_{s(1)}$.
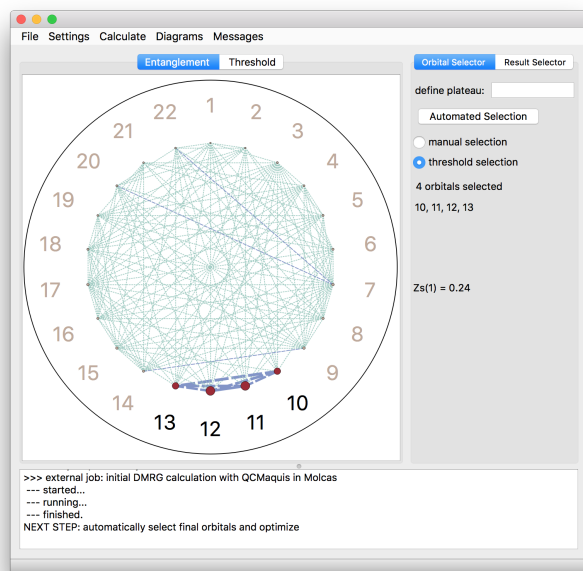
four active orbitals. This is the full π-space of butadiene and certainly a reasonable choice for an active space for this molecule.

The final step of the calculation, the orbital optimization and CASPT2 calculation with the final active space, is requested with a click on *Next step* in the popup message or alternatively with a click on *Run final calculation* in the *Calculate* menu.

You can now choose whether you want to just optimize the automatically selected active orbitals or add dynamical correlation by a subsequent CASPT2 calculation.[12] In case of a CASPT2 calculation feel free to adjust the value of the IPEA or imaginary shift according to your preference with a click on *Edit*.[13]

The final calculation is again started with a click on *Run calculation*. The status messages in the log view at the bottom of the main screen and another popup message inform you upon successful completion of the final calculation. You can export the threshold and mutual information diagrams as a publication-ready pdf file with a click on the *Diagrams* menu entry.

Now you might want to play around with the manual selection or the threshold selection.[14]

Congratulations! You finished your very first multi-configurational calculation with AUTOCAS.

[12] See C. J. Stein, V. von Burg, M. Reiher *JCTC* **2016**, *12*, 3764. for an analysis of the suitability of automatically selected active spaces with CASPT2 calculations.

[13] We generally recommend to not use any IPEA shift in CASPT2.

[14] Hovering over the circles corresponding to the orbitals in the mutual information diagram reveals additional information on the single-orbital entropies and occupation numbers.
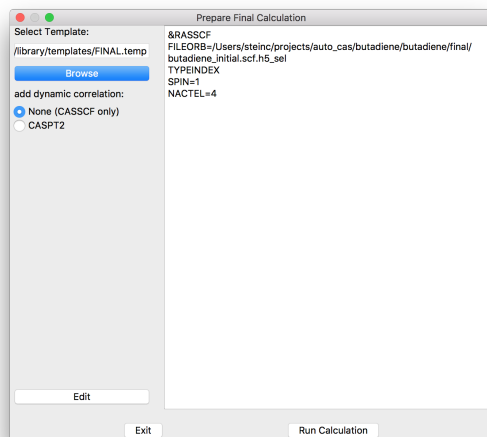
Figure 4.7: Final calculation preparation screen. You can choose to optimize the orbitals only (CASSCF) or carry out a CASPT2 calculation.

4.2   Example 2: Fully Automated – the Two Lowest Electronic States of Stretched Ethylene
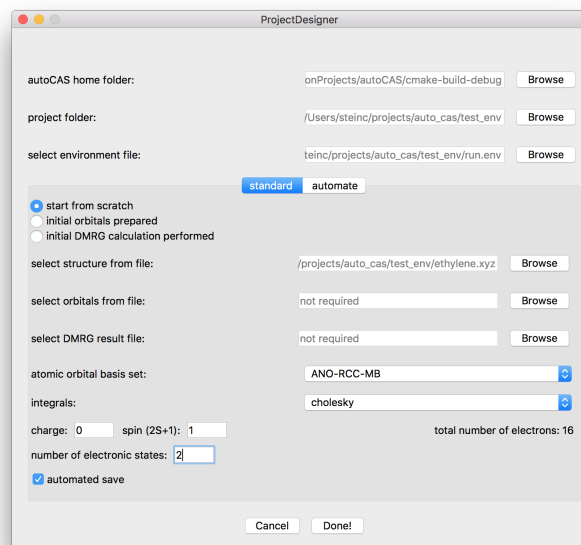
Figure 4.8: Project Designer screen with the *standard* tab. The settings we select here are almost the same as in the previous example but we want to calculate the wave function for the two lowest electronic states.

The previous example demonstrated how AUTOCAS replaces the experience based manual based active orbital space selection by a rigorous protocol based on orbital entropy measures. In the next example we show how this allows one to fully automate multi-configurational calculations, hence turning this kind of calculations

from error prone and tedious set-up procedures into a black-box approach.

Prepare the calculation by setting up a *project folder* that contains the *run.env* file and the *ethylene.xyz* structure file from the *autoCAS home→tests* folder. Keep in mind that this structure has a stretched CC bond distance compared to the equilibrium structure.[15]

You can now start the AUTOCAS program and switch to the *Project Designer* with a click on *Next step* in the blue popup message. Enter all data as shown in the screenshot.

Apart from the number of electronic states that we set to '2' here, the setup is similar to the first example for the *standard* tab. Since we want to run an automated calculation we need to specify more settings in the *automate* tab:

> [15] The final active space should therefore include not only the two π_{CC}-orbitals but also the bonding and antibonding $\sigma_{CC}^{(*)}$ orbitals.

> Figure 4.9: The *automate* tab of the Project Designer screen. Note that the current version of AUTOCAS will run the calculation in C_1 symmetry and select the full valence space as active orbital space for the initial DMRG calculation, which will require significant memory resources.

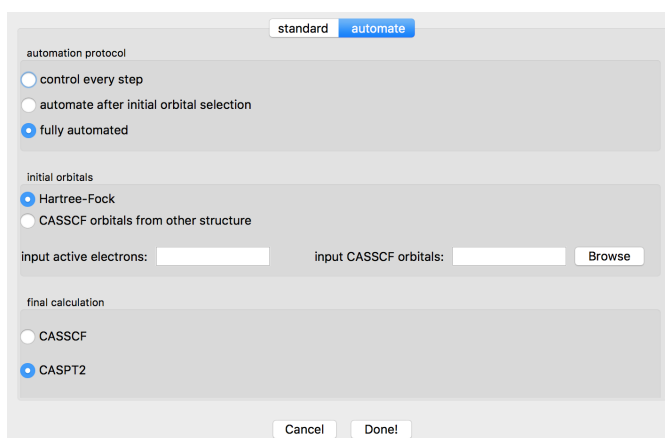We select a *fully automated* calculation now. [16] As in the previous example we want to start the automated active space selection from Hartree–Fock orbitals. In many cases it is a good idea to start from previously prepared CASSCF orbitals that were either optimized for a similar structure (e.g. several structures along a reaction pathway)[17] or from a previous iteration of the AUTOCAS automated selection procedure.[18]

You finish your setup with a click on *Done* and confirm the suggested name for the calculation in the next dialog. In contrast to the previous example the popup message in the main screen suggests to start the automated procedure as next step and you confirm with a click on *Next step*. All individual steps of the calculation are now carried out automatically. After the calculations have finished you may analyze the mutual information diagrams for both electronic states by selecting the individual results with a double click on the respective items in the *Result Selector* tab on the right hand side of the main screen.

> [16] The option *automate after initial orbital selection* can be chosen if an initial orbital basis was already prepared.

> [17] See C. J. Stein, M. Reiher, *Chimia* **2017**, *71*, 170 or arXiv:1702.00450 for more information on the automated selection along reaction coordinates.

> [18] For larger molecules with strong multi-configurational character the Hartree–Fock orbitals might not be a suitable basis. In such cases converged results can be obtained by applying the automated selection procedure repeatedly on iteratively improved molecular orbitals.
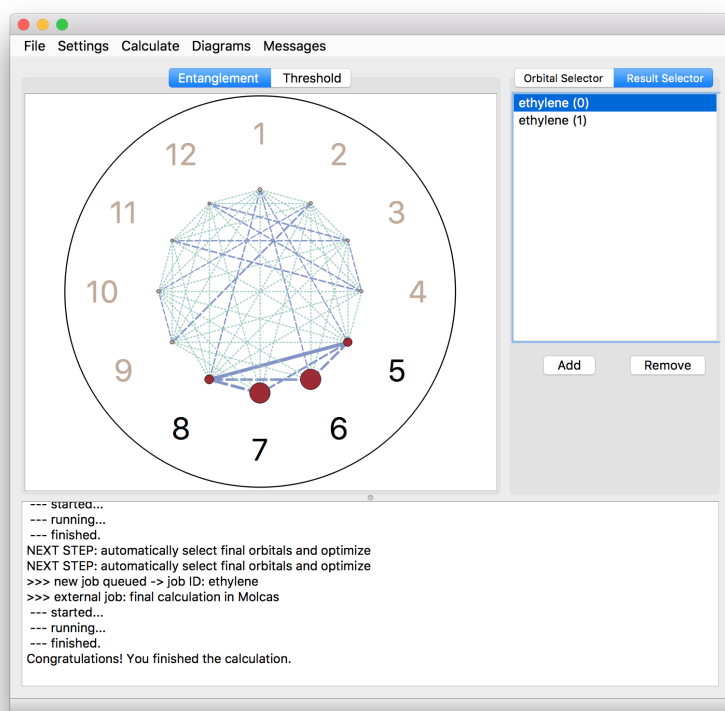
Figure 4.10: Main screen after the calculation finished. Select the individual results in the *Result Selector*.

5
Overview: Most Important Functionalities

We briefly describe the settings of the most import screens in AUTO-CAS.

5.1 Main Screen

The main screen contains all functions that allow you to select the active orbitals; most importantly the *Automated Selection* button.

We strongly advise to run any final calculation with the automatically selected orbitals. If you have reasons to mistrust the automated selection, we suggest to choose another orbital subset by manually selecting a plateau in the threshold diagram. The manual selection is a feature that allows you to test other orbital subsets and their influence on the multi-configurational diagnostic $Z_{s(1)}$.

1. **menu** bar: Contains the *File*[1], *Settings*, *Calculate*, *Diagrams* and *Messages* menu. The *Settings* menu allows you to specify certain parameters of the calculation such as charge and spin or the structure file. You can further select QCMaquis result files to display the entanglement information. *Calculate* allows you to carry out the individual steps of a multi-configurational calculation. You can export the mutual information and threshold diagrams in pdf format via the *Diagrams* menu and the *Messages* menu allows you to switch on/off the popup messages.

   [1] The *Create autoCAS project* function in the File menu will mark the beginning of your calculation in most cases.

2. **diagram** tab: Here you can select between the threshold and mutual information diagrams.[2]

   [2] This is of course only possible once a DMRG result file has been loaded.

3. **diagram** window: In this window, the mutual information and threshold diagrams are displayed. When the *entanglement* tab is active you can collect some information about the current settings with a right click. Resting your cursor on the circles corresponding to the orbitals will give you further information.

4. **orbital/result** selector: If the *Orbital Selector* tab is active, it will collect information about the currently selected orbitals, the multi-

1. menu bar

4. orbital/result selector

2. diagram tab

5. plateau set

3. diagram window

6. automated selection

Figure 5.1: AUTOCAS main screen.

7. status messenger

8. selection switch

9. selected orbitals info

10. multi-configurational diagnostic

configurational diagnostic $Z_{s(1)}$ and the tools for the orbital selection. The *Result Selector* tab allows you to switch between different calculations (e.g. different electronic states).

5. **plateau** set: This sets the plateau threshold for the automated selection protocol.[3]

6. **automated selection**: Evokes the automated active space selection based on orbital entanglement measures.

7. **status** messenger: Supplies you with information on the status of your calculation and possible errors.

8. **selection** switch: Allows you to switch between a *manual selection* (where you can manually select the active space with a click on the circle corresponding to the orbital in the mutual information diagram) and a *threshold selection* (where you can select subsets of orbitals with high single-orbital entropy based on the threshold diagram).

[3] Although the plateau threshold is a parameter of the automated active space selection protocol there is no need to change the default threshold. If you, however, decide otherwise, you must report your plateau threshold when you publish results obtained with AUTOCAS.

9. **selected orbitals** info: Lists the number of currently selected orbitals.[4]

10. **multi-configurational diagnostic**: The $Z_{s(1)}$ diagnostic measures the multi-configurational character of your approximate DMRG wave function.[5] A value > 0.2 corresponds to significant multi-configurational character and is therefore the threshold of the orbital selection protocol.

[4] For calculations with several excited states you will also find a list of orbitals that are selected for the other states. The final active space will consist of the union of all selected orbitals for each state.
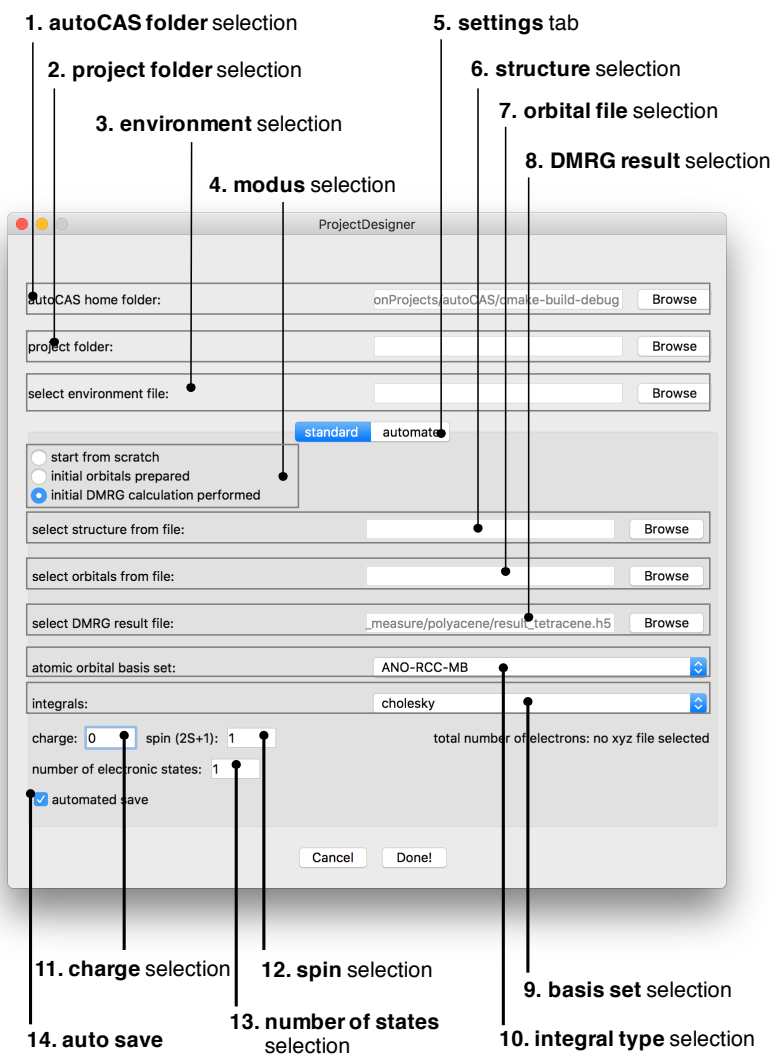
[5] See C. J. Stein, M. Reiher, *Mol. Phys.* **2017**, *115*, 2110.

5.2 Project Designer

1. autoCAS folder selection

2. project folder selection

3. environment selection

4. modus selection

5. settings tab

6. structure selection

7. orbital file selection

8. DMRG result selection

ProjectDesigner

autoCAS home folder: onProjects/autoCAS/cmake-build-debug [Browse]

project folder: [Browse]

select environment file: [Browse]

standard | automated

○ start from scratch
○ initial orbitals prepared
● initial DMRG calculation performed

select structure from file: [Browse]

select orbitals from file: [Browse]

select DMRG result file: _measure/polyacene/result_tetracene.h5 [Browse]

atomic orbital basis set: ANO-RCC-MB

integrals: cholesky

charge: 0   spin (2S+1): 1   total number of electrons: no xyz file selected

number of electronic states: 1

☑ automated save

[Cancel] [Done!]

11. charge selection

12. spin selection

13. number of states selection

14. auto save

9. basis set selection

10. integral type selection

1. **autoCAS folder** selection: Here, you can select the *autoCAS home* folder that contains the template files for the basis sets, integral types, information about each element and the OpenMolcas input template files.[6] This folder is selected automatically and there is

[6] See chapter *Customize template files* on page 29 for information about how to modify these files.

usually no need to change it.

2. **project folder** selection: Select the parent folder where you want your calculations to be started from. It is convenient to store the structure file and environment file here.

3. **environment** selection: Select the environment file that is adapted to your local OpenMolcas / QCMaquis installation.[7]

4. **modus** selection: Here you can choose if you want to start a calculation from scratch (you just have a structure file), if you want to start from already prepared orbitals, or if you want to start from an initial DMRG calculation. **Please note that when selecting the last two options you are responsible for supplying integral information etc. in your appropriately named scratch folder. Otherwise the OpenMolcas calculation will crash.**

5. **settings** tab: Allows you two switch between the *standard* settings screen and the *automate* screen discussed in the next section.

6. **structure** selection: Select your *.xyz* structure file. If an element is not yet in the *ELEMENTS* template file, see Chapter *Customize template files* on page 29 for a way to add it manually.[8]

7. **orbital file** selection: In the case you have already prepared a set of initial molecular orbitals in the OpenMolcas hdf5 orbital file format you can select them here.

8. **DMRG result** selection: In the case you have already carried out an initial DMRG calculation and you only want to analyze the entanglement. You can select the hdf5 result file here.

9. **basis set** selection: Select your atomic orbital basis set.[9]

10. **integral type** selection: Specify the atomic orbital integral type. You will mainly use *standard* for small molecules and *cholesky* for large molecules.

11. **charge** selection: Set the total charge of the calculation.

12. **spin** selection: Set the total spin multiplicity (2S+1) of your calculation. By default, AUTOCAS will always assume the lowest spin possible (singlet for even number of electrons, doublet for odd number of electrons).

13. **number of states** selection: Specify how many electronic states you want to include in your calculation.

[7] Once you selected a parent folder this will be the the folder where AUTOCAS directs you to first.
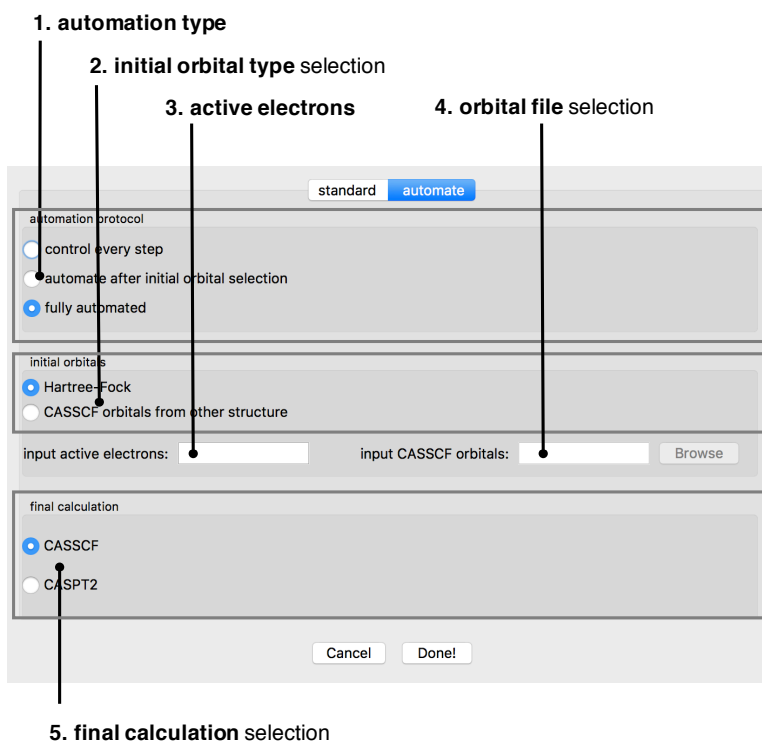
[8] AUTOCAS determines the number of electrons in the system from this file and the total charge you set.

[9] For most calculations it is convenient to start with a minimal basis set and project the pre-optimized orbitals on a larger basis with the OpenMolcas EXPBAS module. See the OpenMolcas pages for a documentation of that module.

14. **auto save**: This option is switched on by default and ensures that intermediate results are saved in a file called *save.cas* in your parent folder. You can interrupt your calculation any time and continue later by loading this file via *File→Open autoCAS Project*.[10]

5.3   Project Designer (Automated Tab)

Open this screen with a click on *File → Create autoCAS project*.

1. automation type

2. initial orbital type selection

3. active electrons

4. orbital file selection

5. final calculation selection

1. **automation type** : You select the degree of automation. You can either control every step (no automated procedure), start the automated calculation after you manually selected the initial orbitals, or go for a fully automated calculation.

2. **initial orbital type** selection: At the moment you can choose between an initial molecular orbital basis constructed from Hartree–Fock orbitals or from CASSCF orbitals of a previous calculation. The latter option is especially useful when you already calculated a multi-configurational wave function for a similar structure (for example when calculating several structures along a reaction path) or you have a large molecule and your initial Hartree–Fock orbital basis is likely to be a bad choice so that you iterate the AUTOCAS selection procedure more than once.

3. **active electrons** selection: If you select CASSCF orbitals as the initial orbital basis, you must specify the number of active electrons of that calculation...

4. **orbital file** selection: ... and the corresponding hdf5 orbital file.

5. **final calculation** selection: Here you can decide whether you want to only optimize the automatically selected active orbitals (CASSCF) or include dynamical correlation with a CASPT2 calculation.

5.4   Initial Orbital Generator

Open this screen with a click on *Calculate → Generate initial orbitals*.
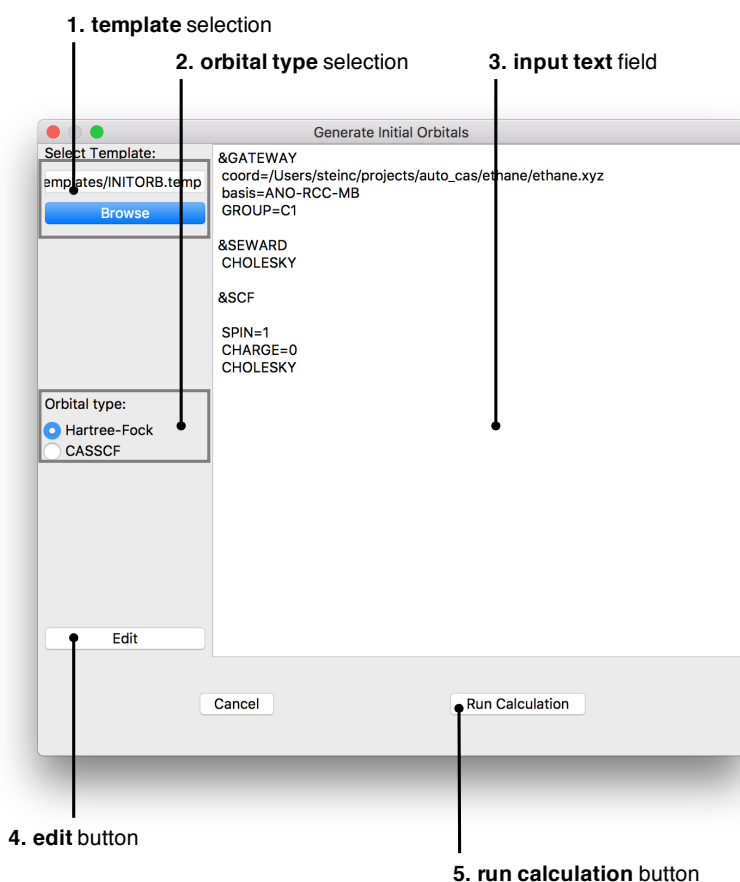
1. **template** selection: The template is automatically selected from the files stored in the *autoCAS home* folder.

2. **orbital type** selection: Here you can choose between Hartree–Fock initial orbitals or CASSCF orbitals.[11]

3. **input text** field: Displays the input for the initial orbital generation.[12]

[11] If you select CASSCF orbitals you will be asked to manually edit the input file and add the number of active electrons and the path to the orbital hdf5 file.

[12] By default: *autoCAS* does not use symmetry. If you change that manually, keep in mind that an automated selection of the *full valence* space for the initial DMRG calculation is only possible for Hartree–Fock orbitals in that case.

4. **edit** button: A click on this button allows you to manually adapt the input.

5. **run calculation** button: AUTOCAS will now create folder called *initial*, write the input and run file and start the calculation by calling OpenMolcas.

5.5 Initial Orbital Selector

Open this screen with a click on *Calculate → Select candidate orbitals*.
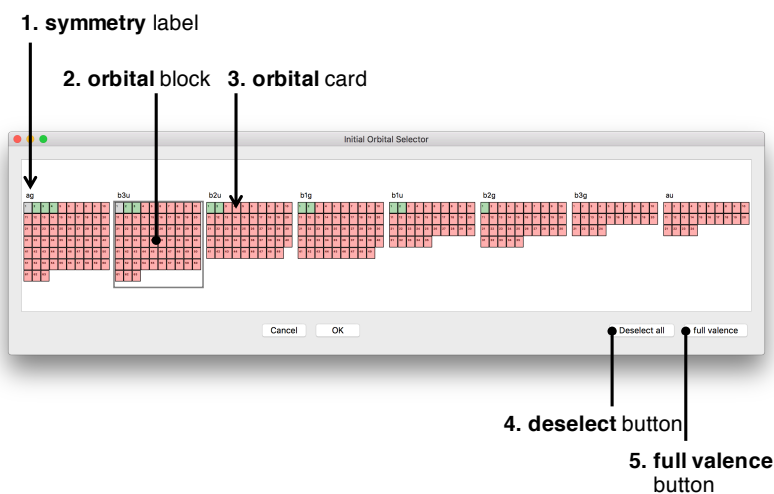
1. **symmetry** label: The symmetry label of all irreducible representations of your point group.[13]

    [13] Since most of the time you will calculate in C_1 symmetry, you will only see the 'a' label.

2. **orbital** block: Collects all orbitals of your molecular orbital basis that correspond to the same irreducible representation.

3. **orbital** card: Numbered card representing an orbital. Gray cards represent unoccupied orbitals, red cards represent virtual orbitals and green cards represent active orbitals. You can select/deselect orbitals with a click on the card.[14]

    [14] Hovering over an orbital card reveals additional information as a tool tip.

4. **deselect** button: Deselects all previously selected orbitals.

5. **full valence** button: Selects the full valence space (if that is possible for the given combination of orbital type and point group symmetry). This is reliable only for minimal atomic orbital basis sets.[15] Otherwise we recommend to select the initial orbital space with the aid of an external orbital viewer.

    [15] You can check the full valence space prior to the selection with a click on 'F'. Orbitals that belong to the valence space will be shown in blue.

5.6   Initial DMRG Calculation

Open this screen with a click on *Calculate → Run DMRG calculation*.
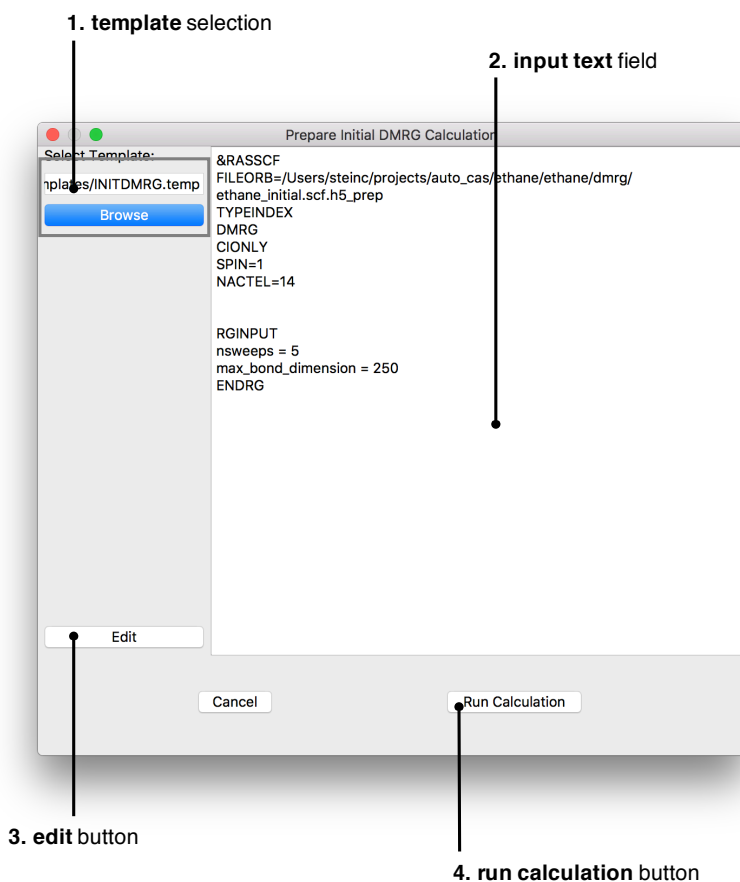
1. **template** selection: The template is automatically selected from the files stored in the *autoCAS home* folder.

2. **input text** field: Displays the input for the initial DMRG generation.

3. **edit** button: A click on this button allows you to manually adapt the input. In order to make sure that your entanglement information is converged we recommend to carry out several calculations with varying *maximum bond dimension* and *number of sweeps*.[16]

4. **run calculation** button: AUTOCAS will now create folder called *dmrg*, write the input and run file, copy the adapted initial orbital file and start the calculation by calling OpenMolcas with the QC-Maquis interface.

[16] A mutual information diagram with chaotic mutual information and overall very high single-orbital entropies for the majority of orbitals already in the ground state is usually an indication of an unconverged calculation. In this case you should rerun the DMRG calculation with a larger *maximum bond dimension* (i.e. number of renormalized block states).

5.7   Final Calculation

Open this screen with a click on *Calculate → Run final calculation*.

1. template selection

2. method selection

3. input text field

```
Prepare Final Calculation

Select Template:
/library/templates/FINAL.temp
          Browse

add dynamic correlation:
◉ None (CASSCF only)
○ CASPT2



              Edit

   Exit              Run Calculation
```

```
&RASSCF
FILEORB=/Users/steinc/projects/auto_cas/ethane/ethane/final/
ethane_initial.scf.h5_sel
TYPEINDEX
SPIN=1
NACTEL=2
```
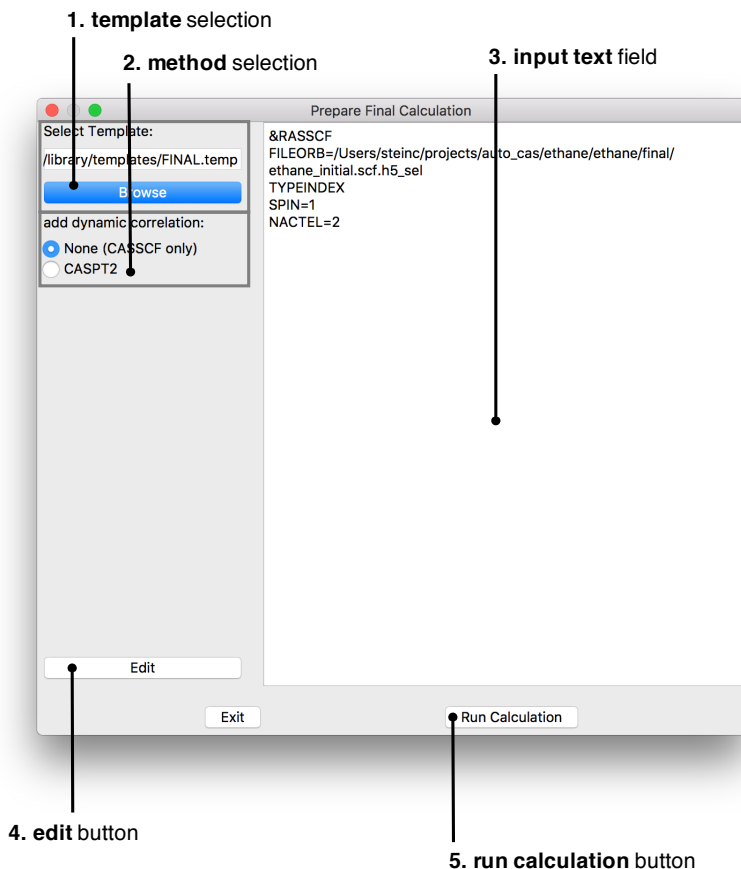
4. edit button

5. run calculation button

1. **template** selection: The template is automatically selected from the files stored in the *autoCAS home* folder.

2. **method** selection: Choose whether you want to only optimize the active orbitals with CASSCF or include dynamical correlation with CASPT2 in case your final active space is small enough for a CASSCF calculation.[17]

3. **input text** field: Displays the input for the final calculation.

4. **edit** button: A click on this button allows you to manually change the input file.

5. **run calculation** button: AUTOCAS will now create folder called *final*, write the input and run files, and start the calculation by calling OpenMolcas (and its QCMaquis interface).

[17] For active spaces with more than 14 orbitals AUTOCAS switches the orbital optimization from CASSCF to DMRG-SCF. If you do not agree, you must manually change the input file.

6
Possible Scenarios

AUTOCAS implements the automated active space selection pro-
tocol and allows the user to analyze the orbital entanglement of a
DMRG calculation in different scenarios. If you want to integrate
AUTOCAS in your workflow and require assistance or adaptations
of the program, do not hesitate to contact the developers via e-mail
(autocas@phys.chem.ethz.ch).

- **Fully automated calculation**

  This is the scenario of the second example above. You start your
  calculation with nothing but a molecular structure. This is possible
  in a fully automated manner for molecules with valence orbital
  spaces up to around 80 orbitals. Remember that the DMRG calcu-
  lations is mostly limited by the accessible memory and you should
  always run the calculation on a machine with very large memory.[1]
  We recommend to use a minimal atomic basis set and expand it
  with the OpenMolcas EXPBAS module after the final orbitals were
  optimized.[2]

- **Semi-automated calculation**

  For large molecules, the full valence space might be too large to
  be included in the active space of the initial DMRG calculation.[3]
  In this case you can select an initial space of *candidate orbitals*[4] and
  carry out the evaluation of the entanglement entropies for these or-
  bitals. Note that this manual selection of *candidate orbitals* leads to
  a procedure that is only semi-automated, the automated selection
  criteria usually substantially shrink the final active space. Further-
  more, the initial selection is not critical and you should simply
  include all orbitals that might be statically correlated. This way,
  the selection takes only a few minutes and is deterministic com-
  pared to a completely manually selected active space calculation
  that usually requires several iterations of trial and error.

[1] **A calculation with 80 active orbitals, no point-group symmetry and 500 renormalized block states requires around 250 GB of memory for the evaluation of the orbital entanglement!** The number of active Orbitals can be increased by employing the *large active space* protocol. However this protocol cannot be utilized for the final calculation.

[2] AUTOCAS determines the valence orbital space according to the orbital energies and this selection is therefore reliable only for minimal basis sets where no Rydberg orbitals with similar energies to those of the valence orbitals are present.

[3] A fully automated procedure that can also be applied to molecules with very large active spaces is currently being tested and will be part of the next release of AUTOCAS.

[4] Orbitals that are likely to be statically correlated like transition metal $3d$-orbitals or π-orbitals of aromatic systems.

- **Reaction pathway**

  If you want to determine the active space along a reaction path
  you can start at an initial structure and follow the fully automated
  procedure as described in the introductory examples. For the next
  structure along the path you may now start your evaluation from
  the set of CASSCF (or DMRG-SCF) orbitals that was optimized
  for the previous structure. In order to do so you just have to select
  CASSCF as orbital in the *Initial Orbital Generator* and add the path
  to the orbital file and the number of active electrons for this cal-
  culation. Alternatively you can select these two variables already
  in the *automated* tab of the *Project Designer* in case you want to
  carry out a fully automated calculation.[5] After you selected an ac-
  tive space for each structure, make sure that the final active space
  comprises the union of all orbitals selected along that reaction
  pathway.

  > [5] Please note that the *initial orbitals
  > prepared* option in the *Project Designer*
  > combined with the *orbital file selection* is
  > not the same! Here the set of orbitals is
  > assumed to correspond to the structure
  > currently calculated and not to the
  > orbitals of a similar structure that are to
  > be reoptimized.

- **Analyze entanglement only**

  In some cases you might have already carried out a DMRG cal-
  culation and just want to analyze the entanglement or the $Z_{s(1)}$
  multi-configurational diagnostic. In such a case you can simply
  select *Settings → Add DMRG result file* after starting AUTOCAS.[6]
  You can check the automatically selected active space, $Z_{s(1)}$ or sim-
  ply export the entanglement or threshold plot as pdf files via the
  Diagrams menu.

  > [6] You can also directly select a DMRG
  > result file in the *Project Designer* by
  > selecting the modus *initial DMRG
  > calculation performed*.

- **Transfer calculation between different computers**

  In many cases, it will be useful to set up your calculation on one
  computer (e.g. your laptop or workstation) and run the expen-
  sive DMRG calculation on a large nodes with sufficient memory.
  This is supported by AUTOCAS because you can simply load the
  current status of a calculation from the *save.cas* file in your *project
  folder*. This file is generated in case you selected the automated
  save mode (which is the default). Copying the *project folder* and the
  OpenMolcas scratch folder to a different machine will allow you to
  restart the calculation from any step of the calculation.[7]

  > [7] Since AUTOCAS stores files like run
  > files or orbital files with absolute paths
  > you have to make sure that your folder
  > structure is the same on both systems
  > or that you manually adjust to the new
  > paths in the *save.cas* file.

7
Customize Template Files

Auto CAS loads several template files to obtain information about basis sets, integral types or element specific information. In order to communicate with OpenMolcas / QCMaquis it further modifies template input files. We describe in the following how you can safely modify these files.

7.1   auto CAS *Template Files*

The three auto CAS template files (*BASIS_SETS*, *ELEMENTS* and *IN-TEGRALS*) are stored in your *autoCAS home folder* and we recommend to copy these files along with the *library* folder that contains the input template files to a different path in case you want to manipulate these files. The auto CAS template files are stored in JSON format.[1]

[1] In order to check whether your manipulated files are readable you might try https://jsonlint.com. This usually identifies formatting errors.

7.1.1   *BASIS_SETS*

You can add your favourite atomic orbital basis set here. Make sure it is available in the OpenMolcas basis set library for all elements of your calculation.

```
[
  {"name": "ANO-RCC-MB",
    "elements": ["H","He","Li","Be","B","C","N","O","F", ... ]
  },²

  ...
  {"name": "ANO-S",
    "elements": ["H","He","Li","Be","B","C","N","O","F" ... ]
  }³
]
```

[2] Add a comma after every group and element.

[3] No comma after the last group or element.

7.1.2   ELEMENTS

This file contains information about the different elements of the periodic table. AUTOCAS reads information from this file to calculate the number of protons and electrons in your molecule and the number of core and valence orbitals. In the current version of AUTOCAS, this information is stored for the first 54 elements of the periodic table. Please add the information manually for the heavier elements by manipulating this file.

```
[
  { ...
  },
  {
    "name": "C",
    "atomic number": 6,
    "number of core orbitals": 1,
    "number of valence orbitals": 4,
    "core orbitals": ["1s"],[4]
    "valence orbitals": ["2s","2p"]
  },[5]
  { ...
  }[6]
]
```

[4] The two last entries are just for clarification of the definition of the core and valence orbitals. This information is currently not processed further.

[5] Again, comma after every group and element.

[6] No comma after the last group or element.

7.1.3   INTEGRALS

The integral file only sets the keywords for the OpenMolcas input files. Most likely you will not need to modify this file.

```
[
  {"name": "cholesky",
    "keys": ["cholesky"]
  },
  {"name": "standard",
    "keys": []
  },
  {"name": "cholesky+AMFI",
    "keys": ["cholesky","AMFI"]
  }
]
```

7.2   OpenMolcas / QCMaquis Template Files

These template files are written in the so-called Jinja format. Every keyword that is replaced by a string while AUTOCAS is running is framed by two curly brackets: {{*keyword*}}. In the following we will simply list the keywords that AUTOCAS expects to be present in the respective files. Apart from these required keywords, you can manipulate the template files according to your requirements.[7]

[7] Be aware that a manipulation of these files might lead to unforeseen errors when running OpenMolcas (and the QCMaquis interface). Therefore we recommend to manipulate these files in the respective screens with a click on the *edit* button, so that changes are effective only for a given project.

7.2.1   INITORB.temp

- {{*coord_file*}}

- {{*basis*}}

- {{*cholesky*}}

- {{*uhf*}}

- {{*spin*}}

- {{*charge*}}

- {{*cirootstring*}}

7.2.2   INITDMRG.temp

- {{*orbital_file*}}

- {{*spin*}}

- {{*active_electrons*}}

- {{*cirootstring*}}

- {{*alphaorbetastring*}}

7.2.3   FINAL.temp

- {{*orbital_file*}}

- {{*spin*}}

- {{*active_electrons*}}

- {{*cirootstring*}}

In addition to these keywords three conditional statements need to be included.

- {%if dmrg %} ... {% endif %}

- {%if caspt2 %} ... {% endif %}

8
Future Releases

We are currently working on an open–source, command line based version of AUTOCAS, which will be released soon.

9
References

Please consult the following references for more details on the AU-TOCAS automated active space selection approach. We kindly ask you to cite the appropriate references in any publication of results obtained with AUTOCAS.

Initial description of the approach:

C. J. Stein, M. Reiher "Automated selection of active orbital spaces", *J. Chem. Theory Comput.* **2016**, *12*, 1760.

Automated active space selection with multi-reference perturbation theory:

C. J. Stein, V. von Burg, M. Reiher "The delicate balance of static and dynamic electron correlation", *J. Chem. Theory Comput.* **2016**, *12*, 3764.

Excited states and reaction paths:

C. J. Stein, M. Reiher "Automated identification of relevant frontier orbitals for chemical compounds and processes", *Chimia* **2017**, *71*, 170.

Multi-configurational diagnostic $Z_{s(1)}$:

C. J. Stein, M. Reiher "Measuring multi-configurational character by orbital entanglement", *Mol. Phys.* **2017**, *115*, 2110.