

---

# **Modular Quantum Chemistry Project**

**Cheol Ho Choi**

**Dec 17, 2021**



# INTRODUCTION

<b>1</b>	<b>What is MQCP?</b>	<b>3</b>
<b>2</b>	<b>What does MQCP try to accomplish?</b>	<b>5</b>
<b>3</b>	<b>How to accomplish it?</b>	<b>7</b>
<b>4</b>	<b>What is Module?</b>	<b>9</b>
<b>5</b>	<b>Why MQCP is so exciting?</b>	<b>11</b>
<b>6</b>	<b>How to create Modules?</b>	<b>13</b>
6.1	Open Quantum Project . . . . .	13
6.2	Data IO Library . . . . .	13
<b>7</b>	<b>How can I use/contribute?</b>	<b>15</b>
<b>8</b>	<b>Who are we?</b>	<b>17</b>
<b>9</b>	<b>For Contributors</b>	<b>19</b>
9.1	A Step-by-step Instruction of Module Uploading . . . . .	19
9.2	Preparing application . . . . .	19
9.3	Uploading new applications . . . . .	20
<b>10</b>	<b>The List of MQCP Datatypes</b>	<b>29</b>
10.1	Input file ( <code>inp</code> ) . . . . .	29
10.2	Simulation metadata ( <code>inf</code> ) . . . . .	30
10.3	Atomic information ( <code>xyz</code> ) . . . . .	31
10.4	Atomic basis set ( <code>bas</code> ) . . . . .	32
10.5	One-electron integrals ( <code>hst</code> ) . . . . .	34
10.6	Wavefunction ( <code>den</code> ) . . . . .	34
10.7	Shells data ( <code>shl</code> ) . . . . .	34
10.8	Log output ( <code>log</code> ) . . . . .	35
<b>11</b>	<b>The List of MQCP Modules</b>	<b>37</b>
11.1	<code>read_molecule</code> . . . . .	37
11.2	<code>apply_basis</code> . . . . .	37
11.3	<code>hsandt</code> . . . . .	38
11.4	<code>guess_huckel</code> . . . . .	38
11.5	<code>guess_hcore</code> . . . . .	39
11.6	<code>rhf_energy</code> . . . . .	39
11.7	<code>rhf_gradient</code> . . . . .	40

11.8	dft_rhf_energy	40
11.9	dft_rhf_gradient	41
11.10	dft_uroh_energy	41
11.11	dft_uroh_gradient	42
11.12	polm_band	42
<b>12</b>	<b>Workflows using Modules</b>	<b>45</b>
12.1	DFT Energy	45
12.2	HF Energy	46
12.3	B3LYP Energy Basis Effect	46
12.4	HF vs. B3LYP Computations	47
12.5	GAMESS computations	48
12.6	MRSF/TDDFT methods	49
12.7	Study of effects of inclusion polarization functions in the basis set	49
12.8	Gradient and Hessian	50
12.9	Donor-Acceptor interactions	51
12.10	Gradient in different approximations	52
12.11	Ground and Excited states	53
12.12	Intermediate matrix generation	54
12.13	SCF procedure	55
<b>13</b>	<b>Understanding MRSF-TDDFT Theory</b>	<b>57</b>
13.1	TDDFT	57
13.2	SF-TDDFT	57
13.3	MRSF-TDDFT	58
13.4	TDDFT vs. MRSF-TDDFT	58
<b>14</b>	<b>Understanding REKS Theory</b>	<b>67</b>
14.1	The KEYWORDS of REKS Theory in GAMESS	67
14.2	Backgrounds	71
<b>15</b>	<b>Band Structure</b>	<b>73</b>
<b>16</b>	<b>Dyson's Orbitals</b>	<b>75</b>
16.1	1. Calculation of Ionization Potentials and Dyson's Orbitals of based on EKT.	75
16.2	2. Calculation of Electron Affinities and Dyson's Orbitals based on EKT.	75
<b>17</b>	<b>Minimum Energy Path by geodesic interpolation</b>	<b>77</b>
<b>18</b>	<b>Nucleus Independent Chemical Shift (NICS) calculation using Dalton Package</b>	<b>79</b>
<b>19</b>	<b>Maximum Overlap Method (MOM) calculation</b>	<b>83</b>
<b>20</b>	<b>LR-TDDFT calculation</b>	<b>85</b>
<b>21</b>	<b>Delta-CR-EOMCC(2,3) calculation</b>	<b>87</b>
<b>22</b>	<b>CASSCF calculation</b>	<b>89</b>
<b>23</b>	<b>XMCQDPT2 calculation</b>	<b>91</b>
<b>24</b>	<b>MRMP2 calculation</b>	<b>93</b>
<b>25</b>	<b>Conical Intersection optimization</b>	<b>95</b>
<b>26</b>	<b>Non-adiabatic molecular dynamics (NAMD) simulaion</b>	<b>97</b>

<b>27 Singlet Ground State Optimization</b>	<b>99</b>
<b>28 Triplet Ground State Optimization</b>	<b>101</b>
<b>29 Singlet Ground State Optimization Imposing C2V Symmetry</b>	<b>103</b>
<b>30 Triplet Ground State Optimization Imposing C2V Symmetry</b>	<b>105</b>
<b>31 Ground State Optimization</b>	<b>107</b>
<b>32 Three State Conical Intersection (TCI) Optimization</b>	<b>109</b>
<b>33 MRSF-TDDFT Examples using Ethylene</b>	<b>111</b>
<b>34 Ethylene S<sub>0</sub> optimization</b>	<b>113</b>
<b>35 Ethylene tw-pyr S<sub>1</sub>/S<sub>0</sub> Conical Intersecrion (CI)</b>	<b>115</b>
<b>36 Ethylene S<sub>0</sub> minimum vibrations</b>	<b>117</b>
<b>37 Ethylene {CI (S<sub>1</sub>/S<sub>0</sub>) - S<sub>0</sub><sup>min</sup>} MEP using geodesic interpolation</b>	<b>119</b>



This site is dedicated to the development and application of modular computational platform for the general science and technology community. As the old-fashioned approach to developing and maintaining computational programs becomes obsolete, an emerging concept of software modularity offers an elegant and timely solution of the looming problems by providing an open development ecosystem where new computational approaches can be rapidly created from modules uploaded at the web repository by the interested users and developers. The first implementation of such an environment is already available in the form of a web-based platform at [MQCP](#).



## WHAT IS MQCP?

MQCP stands for Modular Quantum Chemistry Project.



## WHAT DOES MQCP TRY TO ACCOMPLISH?

MQCP is initiated to promote modularization of scientific software for

- improving interoperability
- facilitating collaborations
- reducing steep learning curve
- reducing maintenance overheads
- adapting to the future heterogeneous HW
- most of all, eliminating redundant code developments over and over again...



## HOW TO ACCOMPLISH IT?

Contrasting to the old-style packaging approach, where various methods and theories are integrated into a single software, MQCP proposes multiple software of modules.



## WHAT IS MODULE?

The modules are standalone executables in linux environment as shown below. Each module needs at least one input (blue square below) and output (red square) ports, which are basically file IO ports for data in/out.

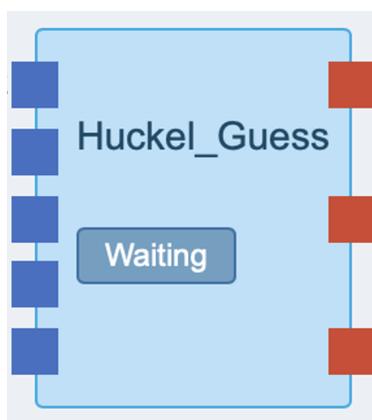


Fig. 1: An example of a module: extended Huckel method for initial orbital guess



## WHY MQCP IS SO EXCITING?

The real power of MQCP lies at the flexibility of creating **workflows** by connecting Modules. More specifically, the output of **Module A** can be directly connected to the input of **Module B** as shown below. The only restriction to this is they (input and output) should have the same datatypes for a given connection. The possible combinations grow infinitely as the number of Modules increases!

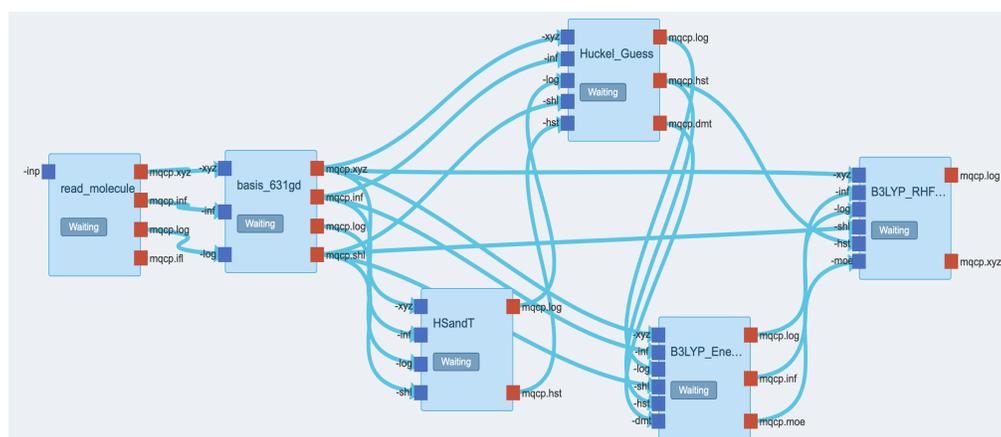


Fig. 1: An example of a workflow: R-B3LYP calculation



## HOW TO CREATE MODULES?

Two projects are undergoing.

### 6.1 Open Quantum Project

Demonstrative quantum mechanical software with the concept of modularization is being developed, which shall be utilized as a template for additional module developments. See more [here](#).

The functionalities of OQP are

- RHF, ROHF and UHF
- KS-DFT
- MRSF-TDDFT
- REKS
- Energy and Gradient

### 6.2 Data IO Library

In order to allow the interconnections of modules in the form of workflow, clear and exact definitions of data are essential. A library for this is under development.



## HOW CAN I USE/CONTRIBUTE?

Please sign up on [MQCP website](#). You can start using uploaded Modules or uploading your Modules right away!



## **WHO ARE WE?**

We are quantum chemistry group at Kyungpook National University, developing new quantum chemistry theories of MRSF-TDDFT and REKS.



## FOR CONTRIBUTORS

### 9.1 A Step-by-step Instruction of Module Uploading

All examples will be shown on `sample_module.x` in English.

### 9.2 Preparing application

Before module uploading you shall create statically-linked binary file. On the Linux operating system you might check that you application is statically-linked by the following command:

```
$ ldd your_executable
```

So, for `sample_module.x` it should provide the next output:

```
$ ldd sample_module.x
not a dynamic executable
```

In the case if you see something like that:

```
$ ldd sample_module.x
linux-vdso.so.1 (0x00007fff8d28c000)
libstdc++.so.6 => /lib64/libstdc++.so.6 (0x00007f9eb8de1000)
libpthread.so.0 => /lib64/libpthread.so.0 (0x00007f9eb8dbf000)
librt.so.1 => /lib64/librt.so.1 (0x00007f9eb8db4000)
libm.so.6 => /lib64/libm.so.6 (0x00007f9eb8c6e000)
libdl.so.2 => /lib64/libdl.so.2 (0x00007f9eb8c67000)
libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00007f9eb8c4c000)
libc.so.6 => /lib64/libc.so.6 (0x00007f9eb8a7f000)
/lib64/ld-linux-x86-64.so.2 (0x00007f9eb8ffe000)
```

It means that your application use dynamic libraries. You shall re-link it with `-static` flag.

If you already have statically-linked application, you shall archive it. Youn need to zip-archive it before uploading to the Edison platform:

```
$ zip sample_module.x.zip sample_module.x
adding: sample_module.x (deflated 58%)
```

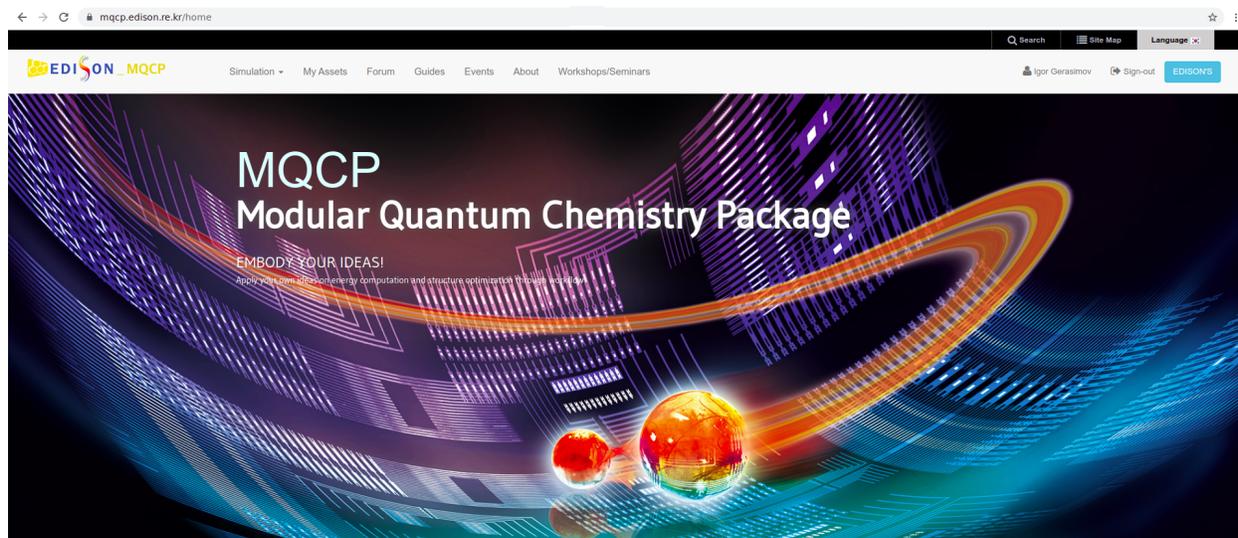
It will be used in one of the stages below.

**Warning:** On the Edison cluster CentOS 6 is used over x86\_64 arch, so that static linking on modern systems leads to the error FATAL: kernel too old. Please use older version of glibc during build process.

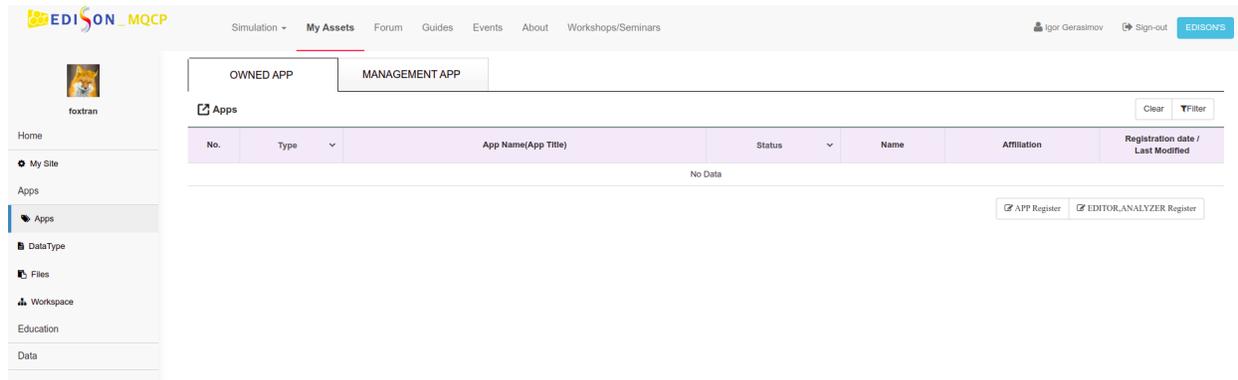
### 9.3 Uploading new applications

Firstly, you need to log-in to EDISON platform, and, then, go to MQCP Project.

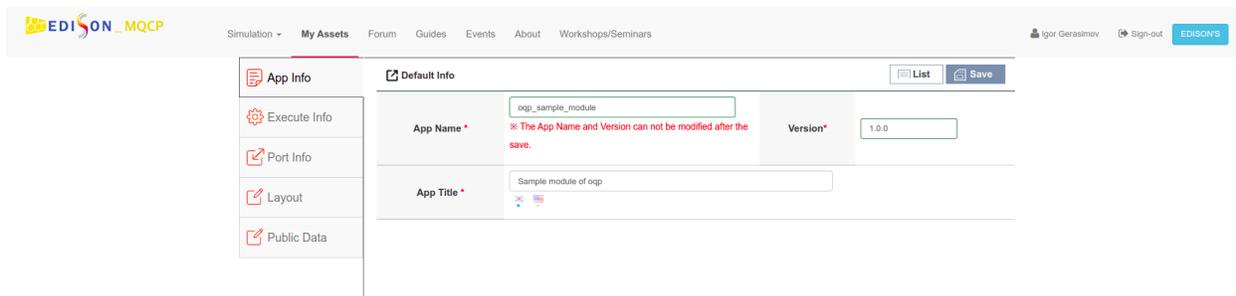
You shall see the following screen:



Navigate to My Assets tab to get to the following page:

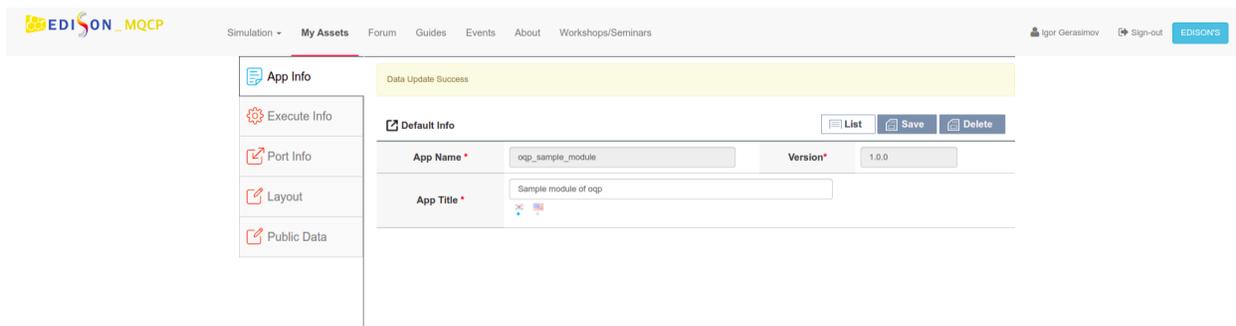


Click APP Register for creating a new application and fill all the required fields denoted by red star. On the next image, oqp\_sample\_module will be created as an example.

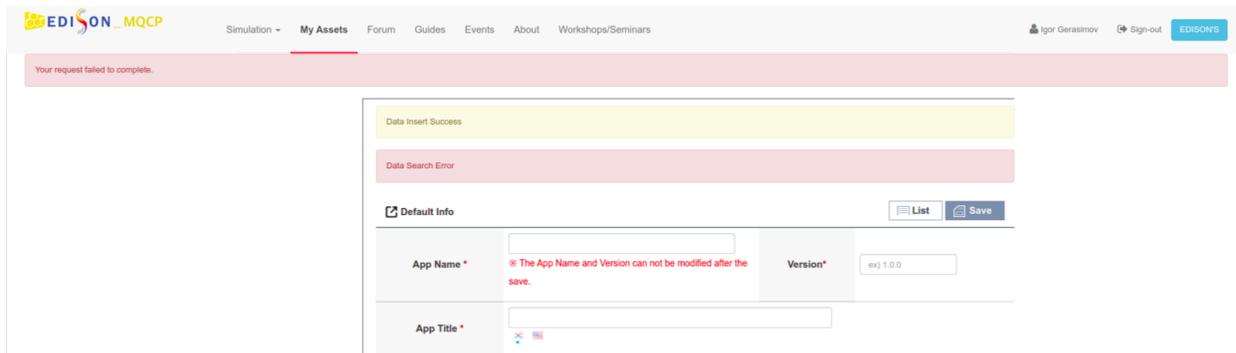


Click on save. We advise to do it after editing every section.

After clicking on save, normally you get the Data update success message:



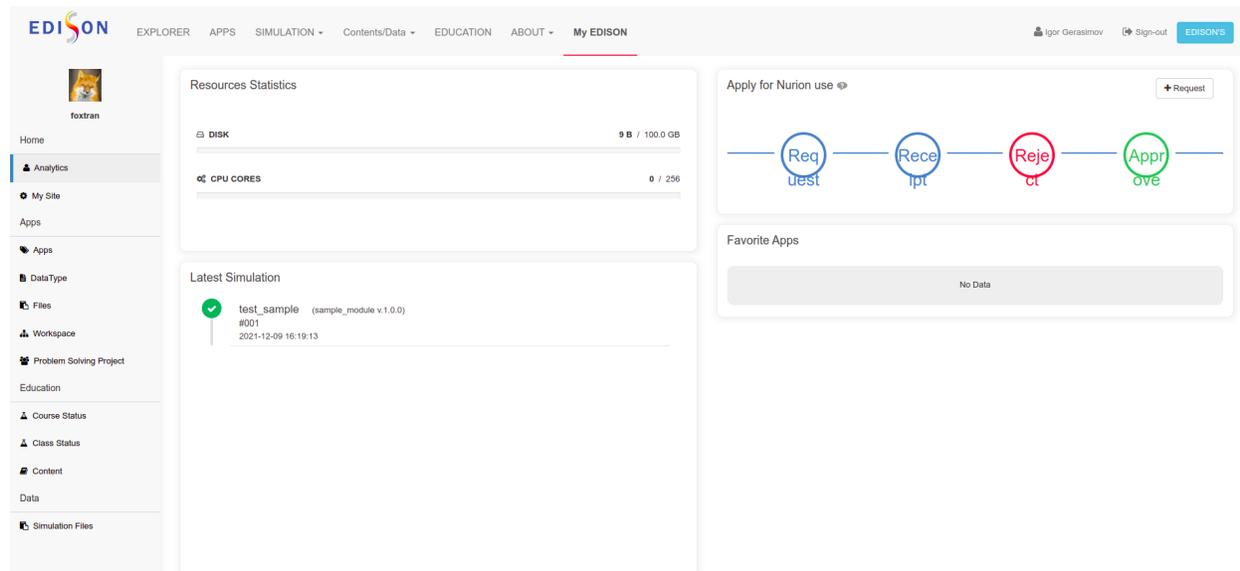
Sometimes you can see the following errors:



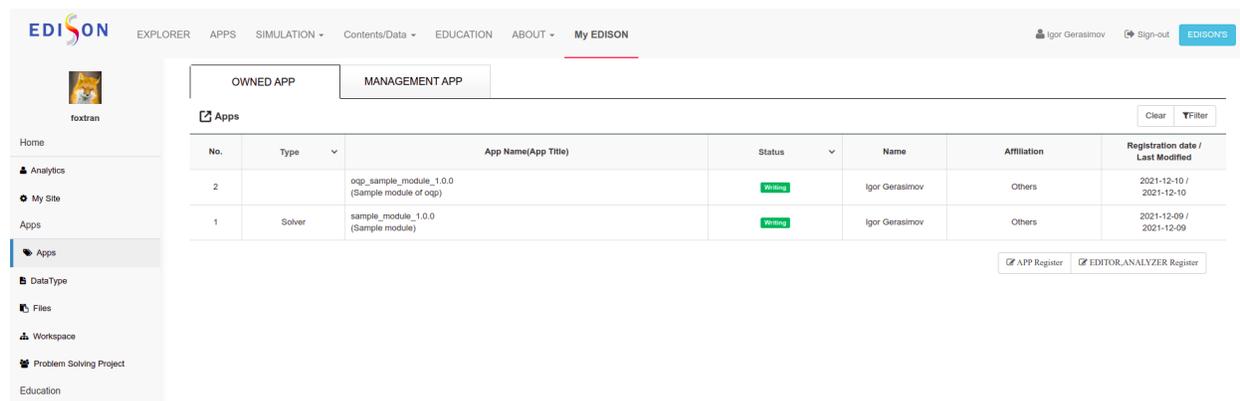
As it said, your data is saved, but you can not continue to modify it here. For continuing, go to [EDISON platform](#):



And click on My EDISON tab. You will see your analytics of your account:



Then, click on Apps tab in the left-side panel. Here, you will see your new application:



Choose your application again and continue to edit it. If it was not opened properly, try again. When the data is loaded proceed to **Execute Info**.

Here, you should provide some information:

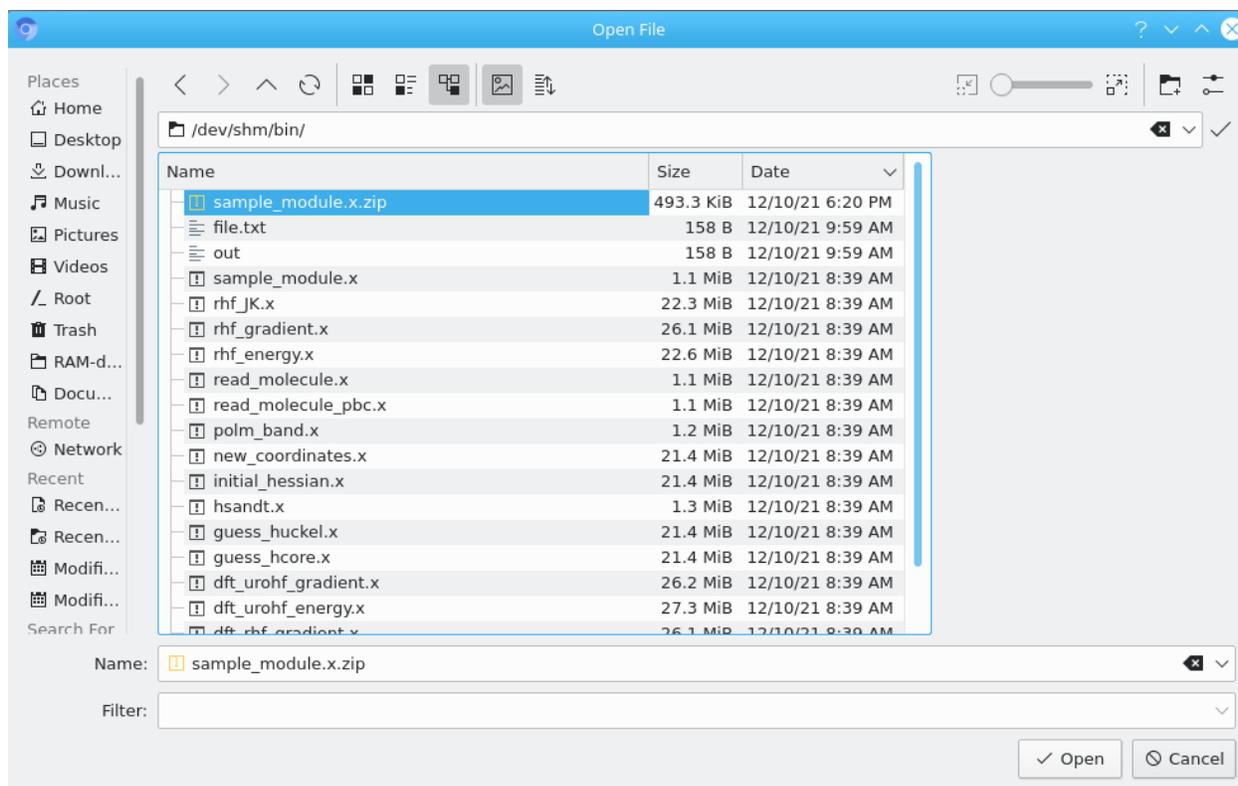
- File Name. Name of executable file. In example, it is `sample_module.x`
- Open Level. Since we are using binaries, choose `OPEN_BINARY`
- App Type. Choose one between `Solver` and `Converter`
- Run Type. Choose type of parallelization.
- New Cluster. Choose queue. It is needed for providing default queue.
- Upload Case. Clean.

It should look so:

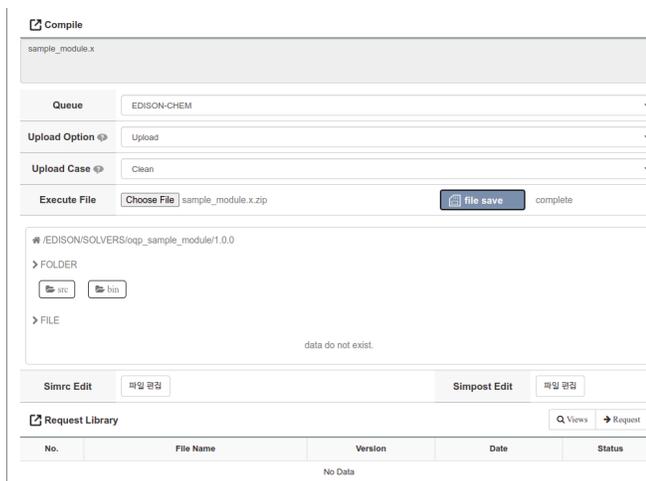
The screenshot shows a web interface for configuring an application. On the left is a sidebar with navigation options: App Info, Execute Info (selected), Port Info, Layout, and Public Data. The main content area is titled 'Execute Information' and contains several sections:

- Execute Information:** Fields for File Name (sample\_module.x), Open Level (OPEN\_BINARY), App Type (Solver), Run Type (Sequential), New Cluster (EDISON), Queue (CHEM), and Cluster (EDISON-CHEM).
- Compile:** A message 'not found directory'.
- Queue:** A dropdown menu set to EDISON-CHEM.
- Upload Option:** A dropdown menu set to Upload.
- Upload Case:** A dropdown menu set to Clean.
- Execute File:** A 'Choose File' button and a 'file save' button.
- Request Library:** A table with columns for No., File Name, Version, Date, and Status. The table is currently empty, showing 'No Data'.

Now it is the time to upload the archived application binary we prepared before. In this example we *produced* `module_sample.x.zip` as an example. You shall to upload your own archive.



In the end of this section you should see this:



In the next tab (Port Info) provide input and output parameters.

The screenshot shows a web interface for configuring an application. On the left is a sidebar with navigation options: App Info, Execute Info, Port Info, Layout, and Public Data. The main content area is titled 'Command Line' and contains a text field with 'sample\_module.x'. Below this are three sections: 'Input Port', 'Log Port', and 'Output Port', each with a table and an 'Add' button. The tables are currently empty, showing 'No Data'.

No.	Port Name	Data Type	edison-table-list-header-port-type	Sample File	Required	Default	Delete
No Data							

No.	Port Name	Data Type	edison-table-list-header-port-type	File Path	Default	Delete
No Data						

No.	Port Name	Data Type	edison-table-list-header-port-type	File Path	Default	Delete
No Data						

The `Command Line` field is the most important here, as it defines how your application should be executed on the cluster. For the sample application, let's just add one input parameter. Click on `Input Port Add` and provide an option here:

A modal dialog box titled 'Port Name' is shown. It has a close button (X) in the top right corner. Below the title is a text input field containing '-i|' and a blue 'Add' button with a checkmark icon.

Then click `Add` to find the proper type of a file corresponding to this option, which is used in your application.

SEARCH

**Data Type Editor**

Data Type Name	Version
mqcp_counter	V1.0.0
mqcp_dmt	V1.0.0
mqcp_fcm	V1.0.0
mqcp_Geom_Cvged	V1.0.0
mqcp_hst	V1.0.0
mqcp_lfl	V1.0.0
mqcp_inf	V1.0.0
mqcp_inp	V1.0.0
mqcp_log	V1.0.0
mqcp_moe	V1.0.0
mqcp_shl	V1.0.0
mqcp_xyz	V1.0.0

Preview

- SAMPLE FILE  
[Alanine\\_dipeptide.inp](#)
- DESCRIPTION  
The input data type of mqcp
- EDITOR  
TEXT\_EDITOR
- ANALYZER  
OSPTextViewer

Clicking on Select button to confirm you choice of input parameter. The command line of your application will be updated accordingly:

- App Info
- Execute Info
- Port Info
- Layout
- Public Data

**Command Line**

sample\_module.x -i mqcp\_inp

**Input Port**

No.	Port Name	Data Type	edison-table-list-header-port-type	Sample File	Required	Default	Delete
1	-i	mqcp_inp	file	<input type="button" value="Add"/>	Y		<input type="button" value="Delete"/>

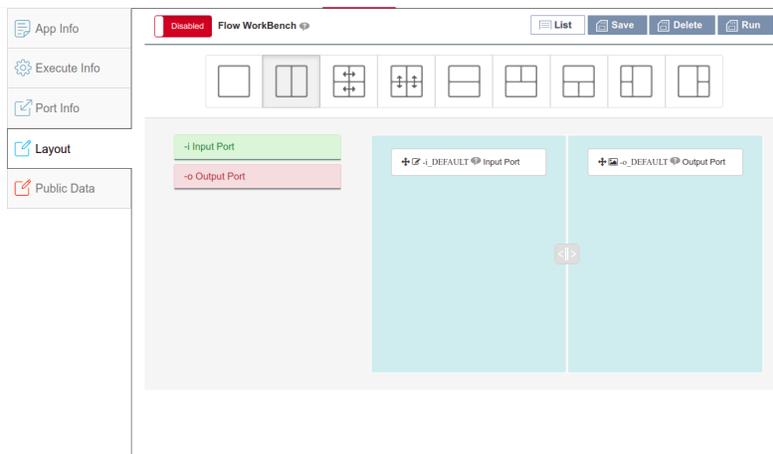
**Log Port**

No.	Port Name	Data Type	edison-table-list-header-port-type	File Path	Default	Delete
No Data						

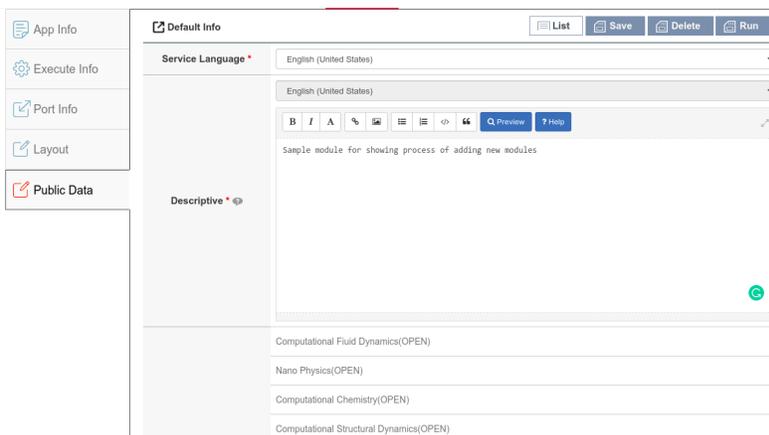
**Output Port**

No.	Port Name	Data Type	edison-table-list-header-port-type	File Path	Default	Delete
1	-o	mqcp_log	file	resultmqcp.log		<input type="button" value="Delete"/>

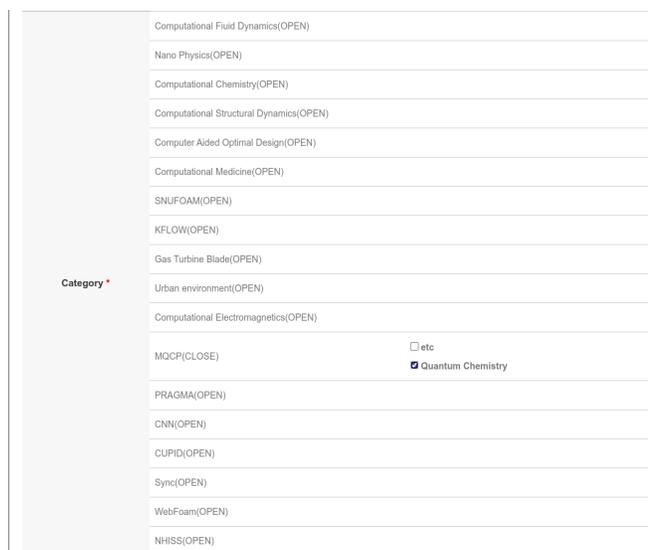
In the next tab, Layout, you shall choose the desired view of your application.



In the last tab, Public Data, you need to provide some description of your application for other users of Edison platform.



For category, you shall choose MQCP -> Quantum Chemistry



For manual, you shall provide website (as it is in screenshot) or file.

Category \*

- Computational Fluid Dynamics(OPEN)
- Nano Physics(OPEN)
- Computational Chemistry(OPEN)
- Computational Structural Dynamics(OPEN)
- Computer Aided Optimal Design(OPEN)
- Computational Medicine(OPEN)
- SNUFOAM(OPEN)
- KFLOW(OPEN)
- Gas Turbine Blade(OPEN)
- Urban environment(OPEN)
- Computational Electromagnetics(OPEN)
- MQCP(CLOSE)  etc
- Quantum Chemistry
- PRAGMA(OPEN)
- CNN(OPEN)
- CUPID(OPEN)
- Sync(OPEN)
- WebFoam(OPEN)
- NHIS(OPEN)

It was the last step when you shall click on save.

Now, you can see your new application in My Assets tab. You can edit it by clicking on application for changes.

The screenshot shows the EDISON\_MQCP web interface. The top navigation bar includes 'Simulation', 'My Assets', 'Forum', 'Guides', 'Events', 'About', and 'Workshops/Seminars'. The user is logged in as 'Igor Gerasimov'. The left sidebar shows navigation options: Home, My Site, Apps, DataType, Files, Workspace, Education, and Data. The main content area is titled 'OWNED APP' and 'MANAGEMENT APP'. Below this, there is a table of applications.

No.	Type	App Name(App Title)	Status	Name	Affiliation	Registration date / Last Modified
2	Solver	oqp_sample_module_1.0.0 (Sample module of oqp)	Working	Igor Gerasimov	Others	2021-12-10 / 2021-12-10
1	Solver	sample_module_1.0.0 (Sample module)	Working	Igor Gerasimov	Others	2021-12-09 / 2021-12-09

At the bottom of the table, there are two checkboxes:  APP Register and  EDITOR,ANALYZER Register.

## THE LIST OF MQCP DATATYPES

Datatypes specify the data format of inputs and outputs of modules, which are files in nature.

### 10.1 Input file (inp)

This datatype is used to set up runtime parameters for modules using human-readable format. It has a following format:

```
TITLE
MOLECULE DESCRIPTION SECTION
PARAMETERS SECTION
```

The title is a single line describing the calculation. It's not used by modules.

In the MOLECULE DESCRIPTION SECTION the number of atoms, the atomic types and their coordinates are given. It resembles the chemical .xyz file format. First the number of atoms is given and then every atom is specified by its nucleic charge (Q) and Cartesian coordinates (X, Y and Z) in Angstrom:

```
NATOM=<number of atoms>
Q   X   Y   Z
...
Q   X   Y   Z
```

In the parameters section various options for modules are specified in the format OPTION=VALUE. Only one option per line is allowed.

Here is an example of input file:

```
C3H8 molecule
natom=11
 6.0  -0.2814116433  -0.0914064155   0.0569146050
 6.0  -0.0108420173   1.3451955038  -0.4302553427
 6.0  -1.5164903822  -0.7099985842  -0.6138764961
 1.0   0.5852250728  -0.7142140575  -0.1449089075
 1.0  -0.3980187758  -0.0846805453   1.1288285918
 1.0   0.2402967369   1.3696657263  -1.4851529300
 1.0  -0.9033547396   1.9429289671  -0.2719580118
 1.0   0.8176724909   1.7961170180   0.1140239055
 1.0  -1.9541584238  -1.5030395028  -0.0171310512
 1.0  -1.2780973834  -1.1336105793  -1.5855662126
 1.0  -2.2852454650   0.0377688194  -0.7774286105
charge=0
scftype=rhf
```

## 10.2 Simulation metadata (inf)

This datatype contains common parameters of simulation which is shared among modules. It can be read and written as an unformatted Fortran 90 structure.

The corresponding Fortran 90 type is defined as follows:

```

type inf
  integer      :: natom      !< The number of atom
  integer      :: charge     !< Molecular charge
  integer      :: hamilton   !< The method of calculations,
                             !< 1= EHT, 2=CNDO/2, 3=INDO, 10=HF, 20=DFT
  integer      :: scftype   !< Reference wavefunction,
                             !< 1= RHF 2= UHF 3= ROHF
  integer      :: nelec      !< The number of electron
  integer      :: nelec_A   !< The number of alpha electron
  integer      :: nelec_B   !< The number of beta electron
  integer      :: mult      !< Spin multiplicity
  integer      :: nvelec    !< The number of valence electron
  integer      :: nocc      !< The number of occupied orbitals
                             !< nOCC = nelec/2 for RHF
                             !< nOCC = nelec_A for ROHF/UHF with mult=3
                             !< nOCC = nelec/2 for ROHF/UHF with mult=1
  integer      :: maxit     !< The maximum number of iterations
  real(REAL64) :: conv      !< Convergency criteria of SCF
  integer      :: idamp     !< Damp the density if idamp=1
  integer      :: mulliken  !< Print out Mulliken population charge if
  ↪mulliken=1
  integer      :: nbasis    !< The number of basis set functions
  integer      :: n2basis   !< n2basis : nbasis*(nbasis+1)/2
  integer      :: l1       !< Used for compatibility, L1 = nbasis
  integer      :: l2       !< Used for compatibility, L2 = nbasis *
  ↪(nbasis+1)/2 = n2basis
  integer      :: l3       !< Used for compatibility, L3 = nbasis * nbasis
  real(REAL64) :: energy    !< Total energy
  real(REAL64) :: enuc     !< Nuclear repulsion energy
  integer      :: nalpha   !< Number of alpha electrons
  integer      :: nbeta    !< Number of beta electrons
  character(len=20) :: &
    basis_name = ''       !< The basis set name for ab initio method
  integer      :: npfunc   !< The number of p functions
  integer      :: mem      !< Available memory size in Mega Byte
  integer      :: runtime  !< Run type
                             !< 1 - energy, 2 - gradient, 3 - geometry
  ↪optimization
  integer      :: geomit   !< Maximum number of geometry optimization
  ↪iteration
  integer      :: guess    !< Not used now
  real(REAL64) :: xdamp    !< Not used now
  real(REAL64) :: psinrm  !< Wavefunction normalization
  real(REAL64) :: ehf1    !< One-electron energy
  real(REAL64) :: vee     !< Two-electron energy
  real(REAL64) :: nenergy !< Nuclear repulsion energy

```

(continues on next page)

(continued from previous page)

```

real(REAL64)      :: etot      !< Total energy
real(REAL64)      :: vne       !< Nucleus-electron potential energy
real(REAL64)      :: vnn       !< Nucleus-nucleus potential energy
real(REAL64)      :: vtot      !< Total potential energy
real(REAL64)      :: tkin      !< Total kinetic energy
real(REAL64)      :: virial    !< Virial ratio (v/t)
real(REAL64)      :: olde      !< For geometry optimizations, energy of the
↳previous step
real(REAL64)      :: depre     !< For geometry optimizations, predicted energy
↳change
integer(I2B)      :: nolds     !< For geometry optimizations, number of steps
↳performed
integer           :: acell     !< Not used now
real(REAL64)      :: ebot      !< Bottom bound of energy range in band
↳calculations
real(REAL64)      :: etop      !< Upper bound of energy range in band calculations
real(REAL64)      :: &
    dfttyp(20)     !< Parameters of XC functional
character(len=1024) :: &
    XC_functional_name = '' !< Name of the XC functional
integer           :: &
    periodic_dim
end type

```

**Todo:** Replace with generic config format

## 10.3 Atomic information (xyz)

Describes molecular structure.

The xyz file contains the total number of atoms as well as the atomic information.

The formatting of the mqcp.xyz file format is as follows:

```

Natom=<N:=number of atoms>
ATOM(1)
ATOM(2)
...
ATOM(N)

```

The ATOM type is defined as follows

```

TYPE atom
  INTEGER(INT32) :: &
!   The number of atomic basis set
    basis_n, &
!   The highest principal quantum number, n
    n_max, &
!   The real high principal quantum number, n:w
    nc_atm, &

```

(continues on next page)

```

!      The number of core electrons
      izcore
REAL(REAL64) :: &
!      Atomic number or nuclear charge
      zn, &
!      Number of valence electrons
      ezn, &
!      Cartesian coordinates
      coord(3), &
!      The highest zeta
      zet_atm, &
!      Atomic mass
      mass, &
!      Gradient
      De(3)
CHARACTER(LEN=2) :: &
!      Atomic symbol
      Symbol
END TYPE atom

```

## 10.4 Atomic basis set (bas)

Provides basis set library.

This type of data is used to pass basis set library to the MQCP program. The library format is basically GAMESS(US)-style with minor adjustments. The basis set file contains several basis set entries for different elements separated by blank line. Symbols !, \$, #, and & are recognized as in-line comment symbols and the rest of the line will be ignored. Comments can appear anywhere in the file.

```

! this is a comment
ATOM BASIS ENTRY
<blank line>
ATOM BASIS ENTRY
<blank line>
ATOM BASIS ENTRY
<blank line>

```

**Note:** The last blank line is *mandatory*.

Atom basis set entry starts from the line containing the element name or its chemical symbol (e.g. CARBON or C). Next, one or more shell entries appears. They must not be separated by blank lines. The atom basis entry format:

```

<Element name | symbol>
SHELL ENTRY
SHELL ENTRY
...
SHELL ENTRY

```

The shell entry contains information of the shell type, the contraction degree, and all exponents and contraction coefficients for every primitive Gaussian in the shell. Shell type denotes the angular momentum and is given as a letter S,

P, D, F, G, H, or I, corresponding to angular momentum 0, 1, 2, 3, 4, 5, and 6 respectively.

**Warning:** In contrast to GAMESS(US) basis set format, Pople's SP (or L) shell types are not supported. S and P components of these shells should be present as separate entries.

After the header, representing the shell type and contraction degree, the parameters of primitive Gaussians are given one at a line. Each primitive gaussian statement contains integer sequence number (which is ignored and is present only for compatibility), and two double precision numbers corresponding to the exponential and contraction coefficients respectively. Hereby, the shell entry format is as follows:

```
<shell type := {S,P,D,F,G,H,I}>      <N, contraction degree>
1      exponent 1      contraction coefficient 1
...
k      exponent k      contraction coefficient k
...
N      exponent N      contraction coefficient N
```

Here is an example of the 6-31G++ basis set entry for the carbon:

```
CARBON
S   6
1      0.3047524880E+04      0.1834737132E-02
2      0.4573695180E+03      0.1403732281E-01
3      0.1039486850E+03      0.6884262226E-01
4      0.2921015530E+02      0.2321844432E+00
5      0.9286662960E+01      0.4679413484E+00
6      0.3163926960E+01      0.3623119853E+00
S   3
1      0.7868272350E+01      -0.1193324198E+00
2      0.1881288540E+01      -0.1608541517E+00
3      0.5442492580E+00      0.1143456438E+01
S   1
1      0.1687144782E+00      0.1000000000E+01
S   1
1      0.4380000000E-01      0.1000000000E+01
P   3
1      0.7868272350E+01      0.6899906659E-01
2      0.1881288540E+01      0.3164239610E+00
3      0.5442492580E+00      0.7443082909E+00
P   1
1      0.1687144782E+00      0.1000000000E+01
P   1
1      0.4380000000E-01      0.1000000000E+01
```

The set pre-formatted basis sets is available [here](#). All provided basis sets can be also downloaded from the [Basis Set Exchange](#) website.

## 10.5 One-electron integrals (hst)

In the `hst` data type the matrices of one-electron integrals are stored. They are namely:

- Core hamiltonian matrix ( $H_{\text{core}}$ )
- Matrix of overlap integrals ( $S$ )
- Matrix of electronic kinetic energy integrals ( $T$ )

The core hamiltonian matrix comprises both electron-nuclei Coulomb attraction integrals ( $V^{\text{en}}$ ) and electronic kinetic energy integrals and is virtually a sum of  $T$  and  $V^{\text{en}}$  matrices:

$$H_{\text{core}} = V^{\text{en}} + T$$

$H_{\text{core}}$ ,  $S$ , and  $T$  are symmetric ( $N \times N$ ) matrices. They are represented in the `packed format` where only an upper (U) triangle of a matrix is stored. Matrix dimensions  $N$  is equal to the number of basis set function for the given molecule. In the output, these three matrices are written as a plain text sequentially one after another using Fortran 90 unformatted output. The order, in which matrices are written is precisely  $H_{\text{core}}$ , then  $S$ , and finally  $T$ .

## 10.6 Wavefunction (den)

This datatype contains information about the molecular orbitals and electronic density for  $\alpha$  and optionally  $\beta$  spin. It includes the following components in order:

- $\alpha$ -spin density matrix ( $D_{\alpha}$ );
- matrix  $Q$  of orthonormal molecular orbitals constructed from the atomic basis set;
- $\alpha$ -spin molecular orbitals;
- eigenvalues of the  $\alpha$ -spin molecular orbitals;
- (optional)  $\beta$ -spin density matrix;
- (optional)  $\beta$ -spin molecular orbitals

$D_{\alpha}$  and  $D_{\beta}$  density are symmetric  $N \times N$  matrices and are stored in a `packed format` (upper triangle). Here,  $N$  denotes the number of basis set functions.  $\alpha$  and  $\beta$  molecular orbitals, as well as matrix  $Q$  are  $N \times N$  matrices which are stored in the `full square format`.  $\alpha$ -orbital energies is a vector of dimension  $N$ . All the data is written using the Fortran 90 unformatted I/O.

## 10.7 Shells data (shl)

This datatype contains a structured shell data for a given molecule. It can be read and written as an unformatted Fortran 90 structure.

The corresponding Fortran 90 type is defined as follows:

```

TYPE shell_structure
    REAL(REAL64), DIMENSION(MXGTOT) :: &
!     Exponential coefficients for shells
    EX, &
!     Contraction coefficients for shells
    CS, CP, CD, CF, CG, CH, CI

```

(continues on next page)

(continued from previous page)

```

INTEGER(INT32), DIMENSION(MXSH) :: &
!   Location of the first exponent and the first
!   contraction coefficient contained in a particular shell
KSTART, &
!   Atomic center indices
KATOM, &
!   Shell types, 1,2,3,4,5,6,7 for S,P,D,F,G,H,I respectively
KTYPE, &
!   Degrees of shell contraction
KNG, &
!   Indices of first AO of the shell in the basis set
KLOC, &
!   Starting and ending indices of the shell
!
!           S   P   D   F   G   H   I   L
!   KMIN   1   2   5  11  21  36  57   1
!   KMAX   1   4  10  20  35  56  84   4
KMIN, KMAX

!   Normalization constant
REAL(REAL64) :: PNRM(84)

!   AO symbol
CHARACTER(LEN=8) :: BFLAB(MXAO)
END TYPE shell_structure

```

## 10.8 Log output (log)

This datatype is a simple text-based log file which contain status report of all modules. It has no specific format.



## THE LIST OF MQCP MODULES

Modules on MQCP need to specify inputs and outputs, which correspond to the arguments of functions or subroutines. Technically, inputs and outputs are the reading/writing of files within modules.

### 11.1 read\_molecule

This module should be called prior to calling any other module. It prepares the calculation by filling up the MQCP internal data structures: simulation metadata and molecular structure. The human-readable input file provided by the user must contain molecular coordinates and all relevant input parameters.

This module has one input and three output ports.

Inputs:

1. *inp* - MQCP input file

Outputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *log* - log file

### 11.2 apply\_basis

The purpose of this module is to create molecular basis set from the file with molecule structure (atomic coordinates and elements) and basis set library.

This module has four input and four output ports.

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *bas* - basis set library
4. *log* - log file

Outputs:

1. *xyz* - updated molecular structure
2. *inf* - updated MQCP simulation metadata

3. *shl* - molecular basis set
4. *log* - updated log file

### 11.3 hsandt

This module computes the following classes of one-electron integrals:

- Atomic basis set overlap integrals ( $S$ )
- Electron-nuclei Coulomb attraction integrals ( $V^{\text{en}}$ )
- Electronic kinetic energy integrals ( $T$ )

These integrals are stored as matrices and saved into the *hst format* file. Matrix elements are computed as follows.

Overlap integrals:

$$S_{ij} = \phi_i | \phi_j$$

Electronic kinetic energy integrals matrix elements:

$$T_{ij} = -\frac{1}{2} \phi_i | \nabla^2 | \phi_j$$

Electron-nuclei Coulomb attraction matrix elements:

$$V_{ij}^{\text{en}} = \sum_{n \in \text{nuc}} \phi_i | \frac{Q_n}{|r - R_n|} | \phi_j$$

Here,  $\phi_i$  denotes  $i$ -th basis set function,  $Q_n$  and  $R_n$  - are charge and coordinates of  $n$ -th nuclei,  $r$  - is electronic coordinate.

This module has four input and two output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *log* - log file

Outputs:

1. *hst* - one-electron integral matrices
2. *log* - updated log file

### 11.4 guess\_huckel

This module computes the initial guess to the density matrix using the semiempirical extended Huckel method. This method uses Huzinaga's MINI basis set. The resulting wavefunction is projected on the molecular basis set provided as an input. This method is applicable to the elements from hydrogen up to radon.

This module has five input and three output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices
5. *log* - log file

Outputs:

1. *dmt* - density matrix
2. *log* - updated log file
3. *hst* - one-electron integral matrices

## 11.5 guess\_hcore

This module computes the initial guess to the density matrix by diagonalization of the one-electron Hamiltonian. This method is applicable to the all elements.

This module has five input and three output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices
5. *log* - log file

Outputs:

1. *dmt* - density matrix
2. *log* - updated log file
3. *hst* - one-electron integral matrices

## 11.6 rhf\_energy

This module computes the electronic energy and molecular orbitals using closed-shell variant of the Hartree-Fock method. The module uses direct calculation of the two-electron integrals tensor and applies Schwartz inequality to screen out small values of integrals. The self-consistent field equations of the Hartree-Fock method are solved with the help of DIIS method. The module needs an initial guess density matrix for running.

This module has six input and three output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices

5. *dmt* - initial guess density matrix

6. *log* - log file

Outputs:

1. *moe* - resulting molecular orbitals and energy

2. *inf* - updated MQCP simulation metadata

3. *log* - updated log file

## 11.7 rhf\_gradient

This module computes the electronic energy gradient using closed-shell variant of the Hartree-Fock method. It takes the wavefunction of the converged Hartree-Fock calculations as an input, which can be obtained using *rhf\_energy module*.

This module has six input and two output ports:

Inputs:

1. *xyz* - molecular structure

2. *inf* - MQCP simulation metadata

3. *shl* - molecular basis set

4. *hst* - one-electron integral matrices

5. *moe* - the converged Hartree-Fock wavefunction

6. *log* - log file

Outputs:

1. *xyz* - updated molecular data which contains atomic gradients

2. *log* - updated log file

## 11.8 dft\_rhf\_energy

This module computes the electronic energy and molecular orbitals using closed-shell variant of the density functional theory (DFT) method. The module uses direct calculation of the two-electron integrals tensor and applies Schwartz inequality to screen out small values of integrals. The self-consistent field equations are solved with the help of DIIS method. The module needs an initial guess density matrix for running. Grid-DFT contribution is computed with the help of the LibXC library for DFT functionals.

This module has six input and three output ports:

Inputs:

1. *xyz* - molecular structure

2. *inf* - MQCP simulation metadata

3. *shl* - molecular basis set

4. *hst* - one-electron integral matrices

5. *dmt* - initial guess density matrix

6. *log* - log file

Outputs:

1. *moe* - resulting molecular orbitals and energy
2. *inf* - updated MQCP simulation metadata
3. *log* - updated log file

## 11.9 dft\_rhf\_gradient

This module computes the electronic energy gradients using closed-shell variant of the density functional theory (DFT) method. It takes the density of the converged Kohn-Sham DFT calculation, which can be obtained using *dft\_rhf\_energy module*. Grid-DFT derivative contribution is computed with the help of the LibXC library for DFT functionals.

This module has six input and three output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices
5. *moe* - the converged DFT density
6. *log* - log file

Outputs:

1. *xyz* - updated molecular data which contains atomic gradients
2. *log* - updated log file

## 11.10 dft\_urohf\_energy

This module computes the electronic energy and molecular orbitals using unrestricted or restricted open-shell density functional theory (DFT) method. The module uses direct calculation of the two-electron integrals tensor and applies Schwartz inequality to screen out small values of integrals. The self-consistent field equations are solved with the help of DIIS method. The module needs an initial guess density matrix for running. Grid-DFT contribution is computed with the help of the LibXC library for DFT functionals.

This module has six input and three output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices
5. *dmt* - initial guess density matrix
6. *log* - log file

Outputs:

1. *moe* - resulting molecular orbitals and energy
2. *inf* - updated MQCP simulation metadata
3. *log* - updated log file

### 11.11 dft\_urohf\_gradient

This module computes the electronic energy gradients using unrestricted or restricted open-shell density functional theory (DFT) method. It takes the density of the converged Kohn-Sham DFT calculation, which can be obtained using *dft\_urohf\_energy module*. Grid-DFT derivative contribution is computed with the help of the LibXC library for DFT functionals.

This module has six input and three output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices
5. *moe* - the converged DFT density
6. *log* - log file

Outputs:

1. *xyz* - updated molecular data which contains atomic gradients
2. *log* - updated log file

### 11.12 polm\_band

This module runs band structure computations. Bands calculations make sense if the molecular structure is an ordered oligomer with insignificant deviations from periodic regularity.

To carry out such calculations, restrictions are introduced on the order of atoms in the initial structure:

1. periodically repeated units are given one by one from the one end of the oligomer to the other;
2. the order of atoms in each unit must be identical;
3. terminal groups of atoms (normally terminal H atoms) must be numbered last.

The keyword ATOM\_CELL must be given in the MQCP input file to define the number of atom in the unit cell. Optional keywords EN\_LOWER and EN\_UPPER can be specified in the input file to define desired energy range (in eV). The input example is the following:

```
The input example of [C2H2]15 oligomer.
natom=62
6.0      0.305450260      -17.455528492      0.000000000
1.0      1.375447159      -17.458106142      0.000000000
6.0      -0.386647810      -16.250082413      0.000000000
1.0      -1.456644709      -16.247504763      0.000000000
6.0      0.311250090      -15.047984853      0.000000000
```

(continues on next page)

(continued from previous page)

```
...
6.0      0.386647810      16.250082413      0.000000000
1.0      1.456644709      16.247504763      0.000000000
6.0     -0.305450260      17.455528492      0.000000000
1.0     -1.375447159      17.458106142      0.000000000
1.0      0.231780490      18.380884162      0.000000000
1.0     -0.231780490     -18.380884162      0.000000000
charge=0
Hamilton= dft
Basis=6-31G(d)
charge=0
scftype=rhf
runtype=energy
atoms_cell=4
en_lower=-12
en_upper=5
$libxc functional=b3lyp5 $end
```

This module has six input and one output ports:

Inputs:

1. *xyz* - molecular structure
2. *inf* - MQCP simulation metadata
3. *shl* - molecular basis set
4. *hst* - one-electron integral matrices
5. *moe* - the converged orbitals
6. *log* - log file

Outputs:

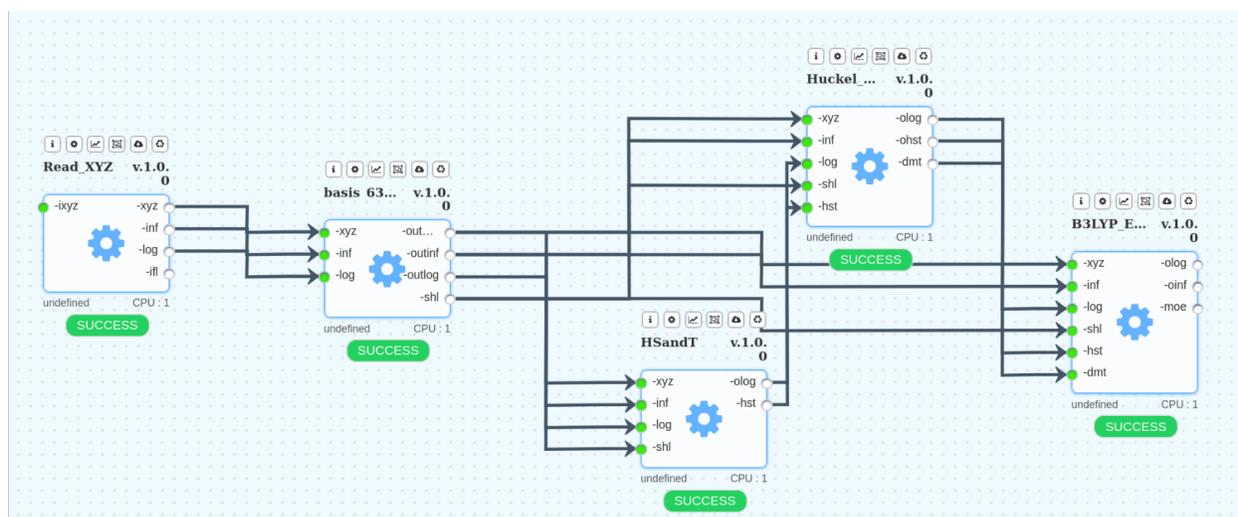
1. *log* - updated log file



## WORKFLOWS USING MODULES

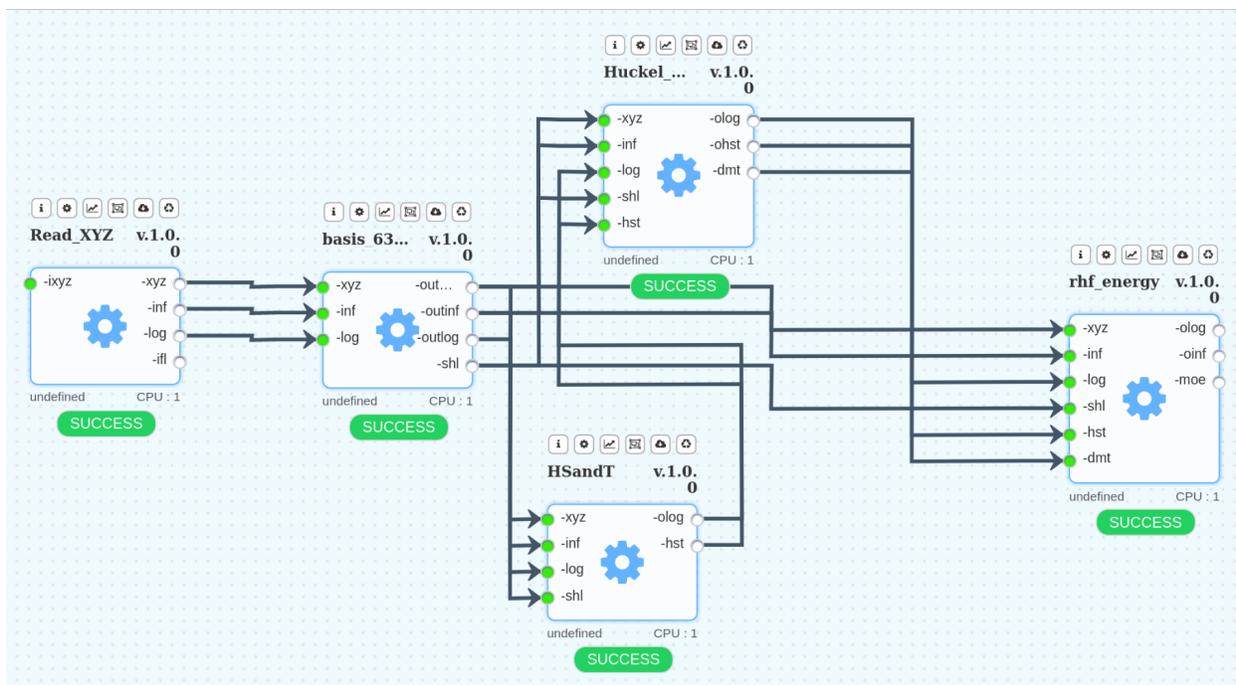
### 12.1 DFT Energy

In terms of practical application, the most popular are calculations of the total electronic energy of a molecule at B3LYP/6-31G(d) level of theory. To carry out such calculations, the following scheme is proposed. This **E\_B3LYP\_631GD** workflow assumes simple XYZ input of atomic coordinates.



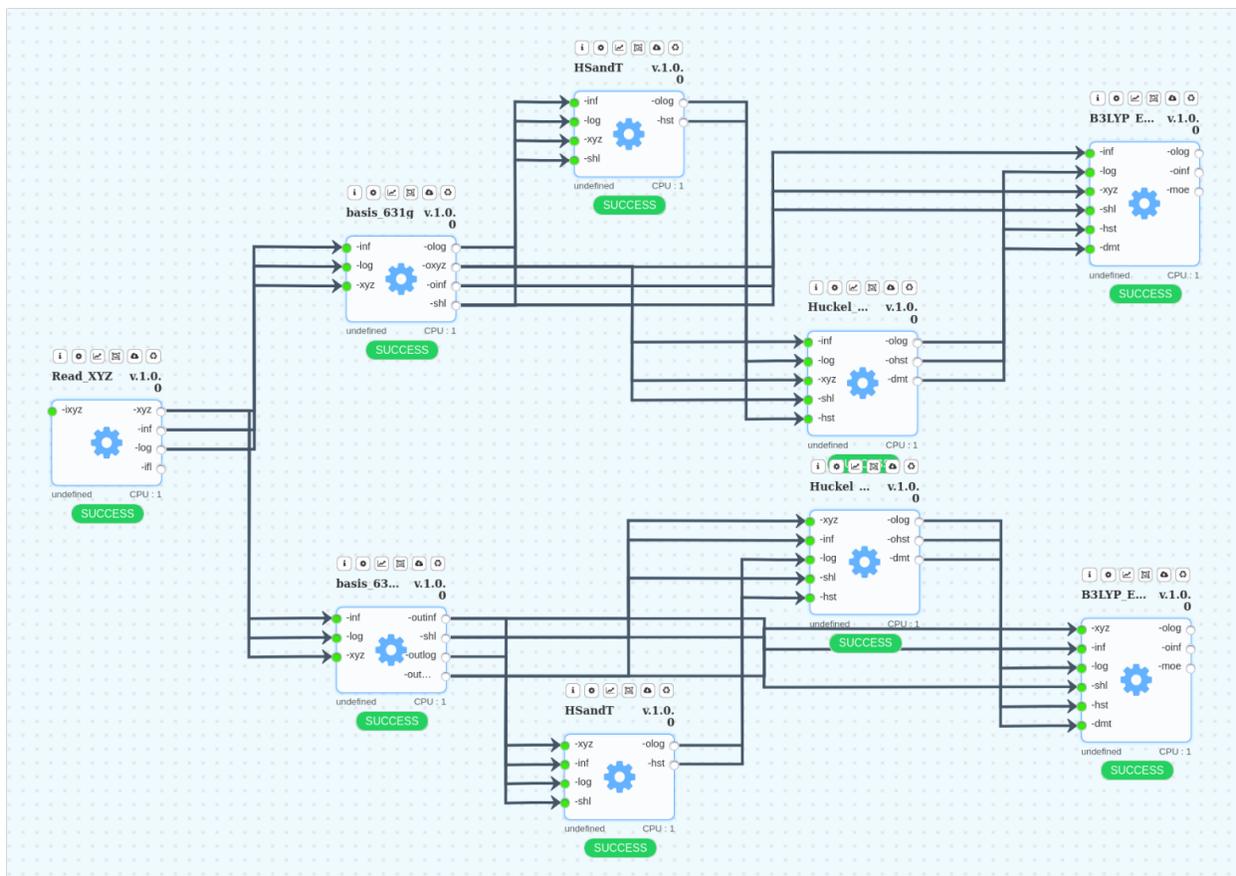
## 12.2 HF Energy

For educational purposes and for comparing the results, the following workflow is useful. This **HF\_631GD** workflow performs computations similar to the above, but at the ab-initio RHF/6-31G(d) level of theory.



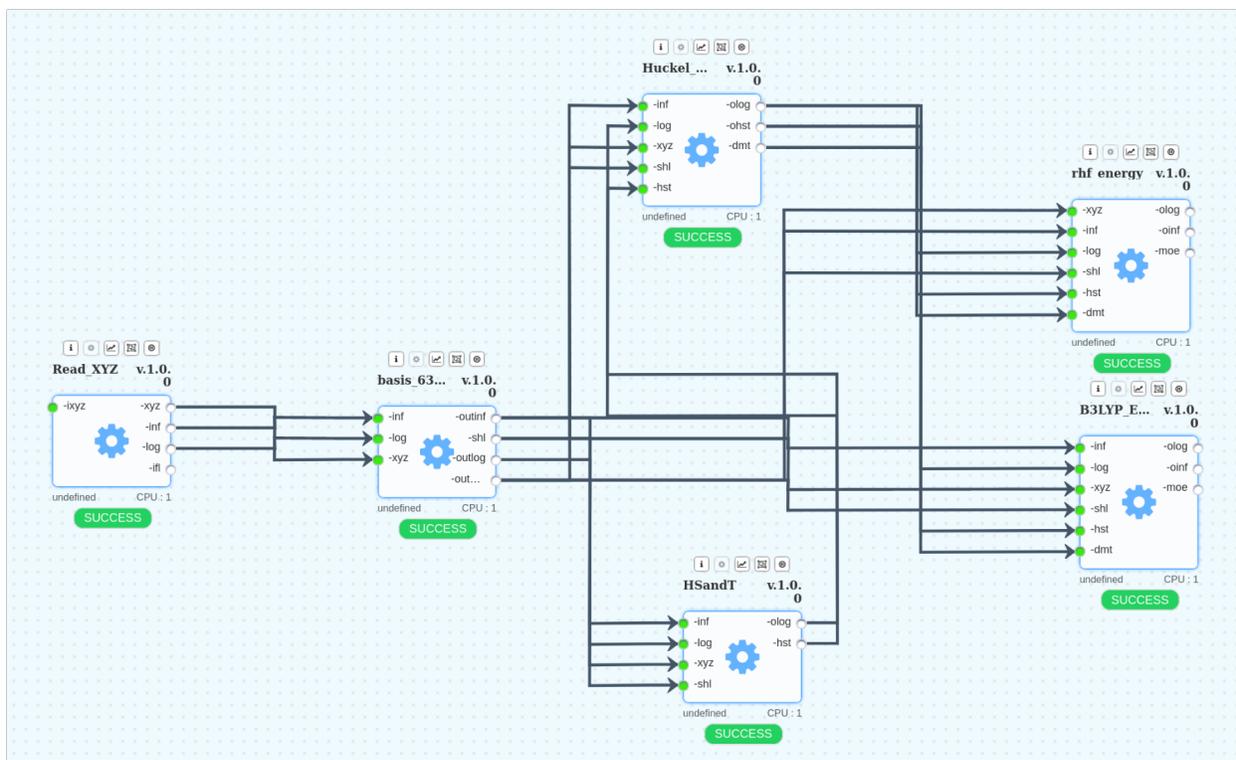
## 12.3 B3LYP Energy Basis Effect

Simultaneous energy calculations with B3LYP method but with two different basis sets (6-31G and 6-31G(d)) are available with the **E\_B3LYP\_631G\_631GD** workflow. Both of them are shown below. A simple XYZ input is needed. The results of simultaneous calculations with two different approach are provided in one output file.



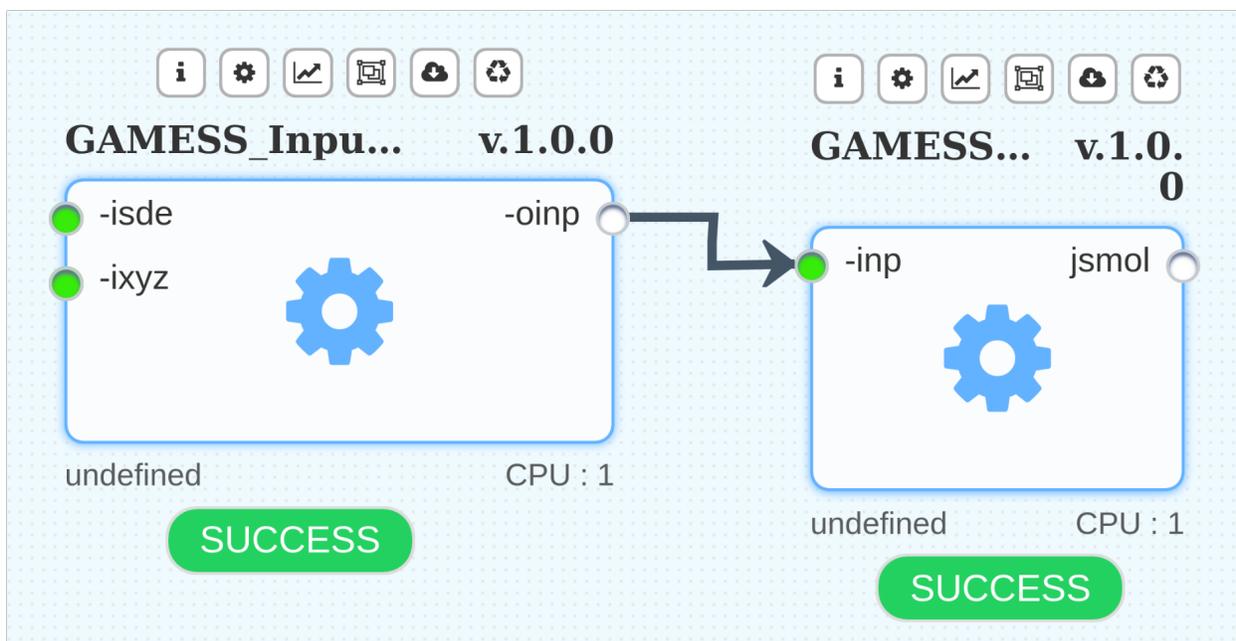
## 12.4 HF vs. B3LYP Computations

Simultaneous energy calculations by RHF and DFT(B3LYP) methods with the 6-31G(d) basis set are available by means of `E_HF_B3LYP_631GD` workflow. A simple XYZ input is needed. The results of simultaneous calculations with two different approach are provided in one output file.



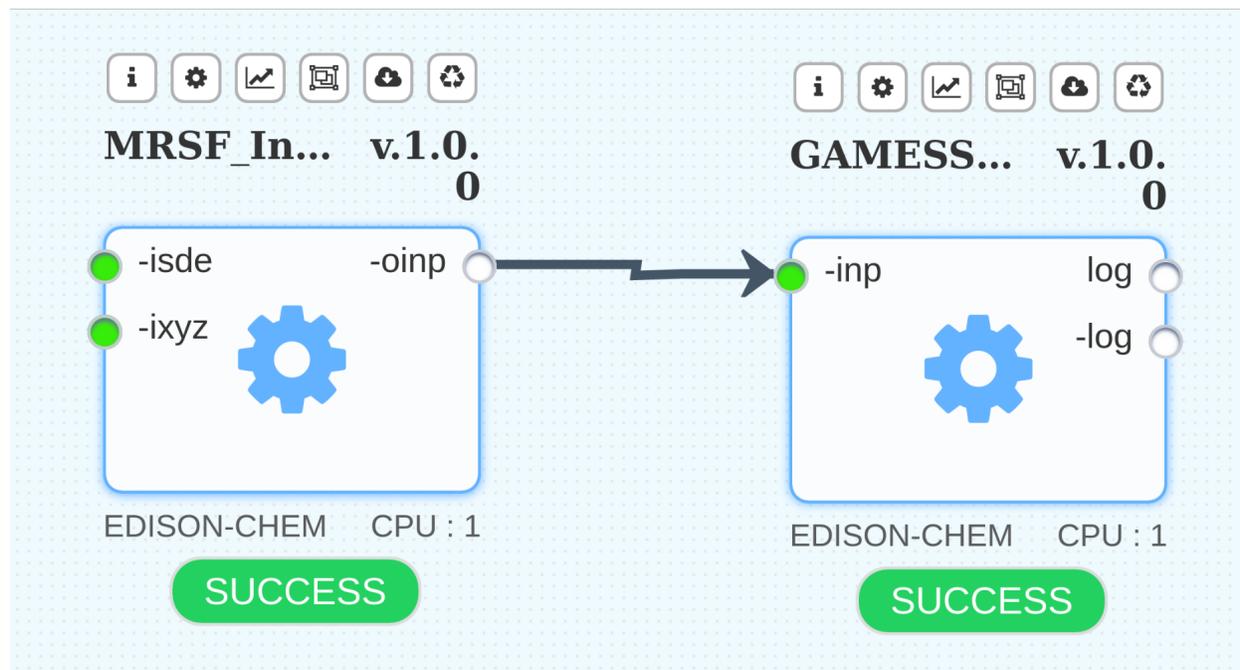
## 12.5 GAMESS computations

A simple but powerful **GAMESS\_KMOL** workflow allows one to use conventional GAMESS as a built-in module in calculations. The input file can be made both by generating in the graphical mode and uploaded by the user. A simple standard GAMESS input is provided. The user also needs to provide the coordinates of the calculated molecule in XYZ format.



## 12.6 MRSF/TDDFT methods

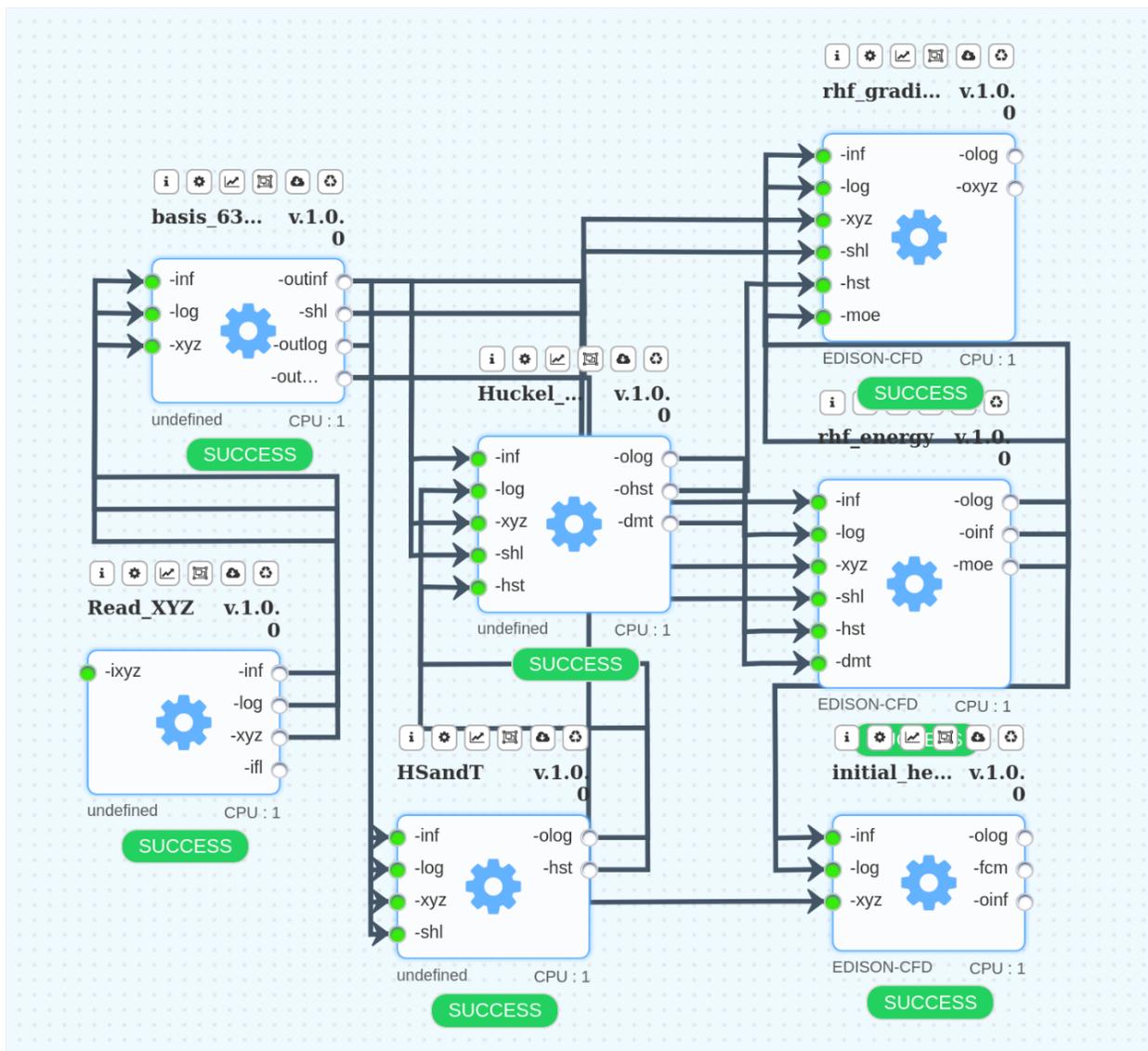
Computations with the Mixed-Reference Spin-Flip Time-Dependant Density Functional as well as Linear Response Time-Dependant Density Functional methods are available via the built-in GAMESS module. The **MRSF\_TDDFT** workflow below allows one to prepare an input file in a user-friendly way. The input file could be generated in the graphical mode, uploaded by user or the attendant sample input can be used.



## 12.7 Study of effects of inclusion polarization functions in the basis set

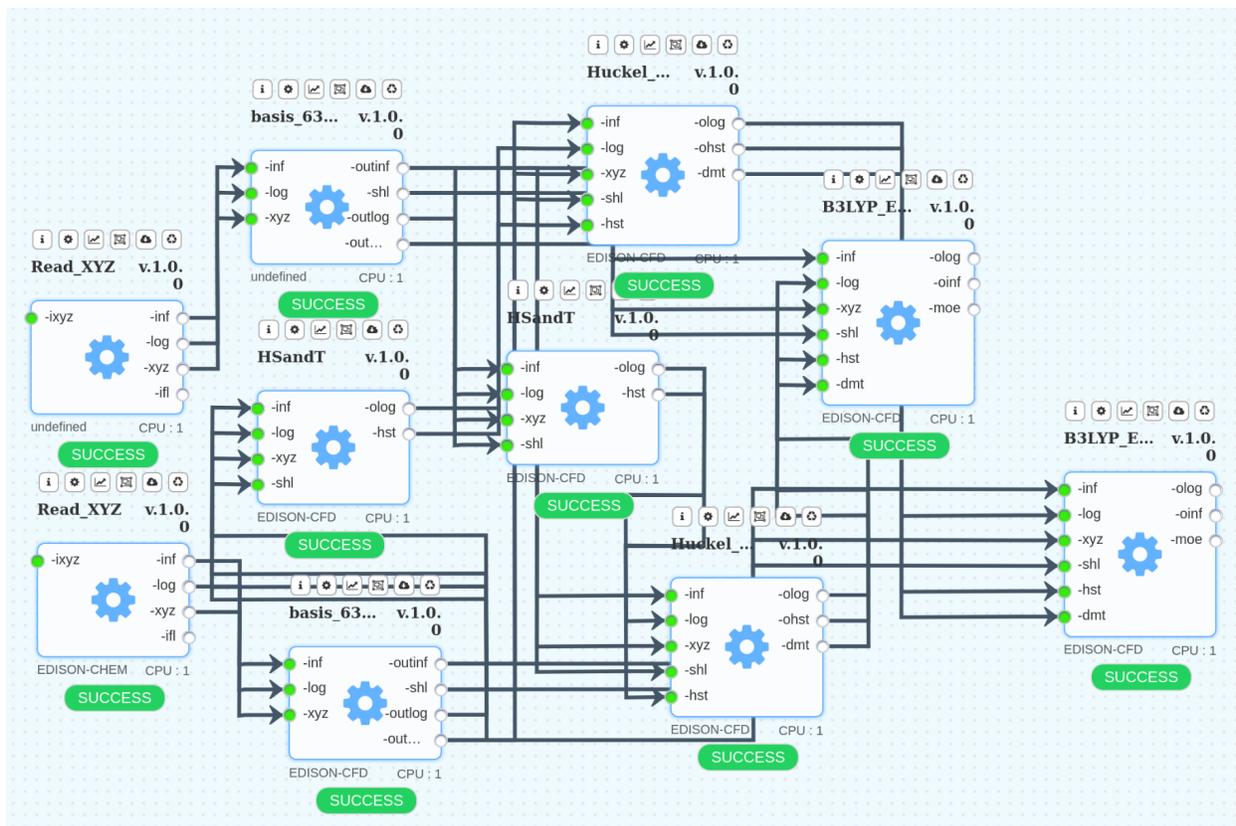
**Polarization\_Function** workflow is shown in the figure and it allows to perform simultaneous computations for the same molecule at the DFT (B3LYP) and HF method with two basis sets, one of which, 6-31G, is without polarization functions and the other, 6-31G(d) includes one polarization function for all atoms except for hydrogen atoms. The only geometry of the molecule (.xyz file format) should be provided as an input and four output files are generated as the result.





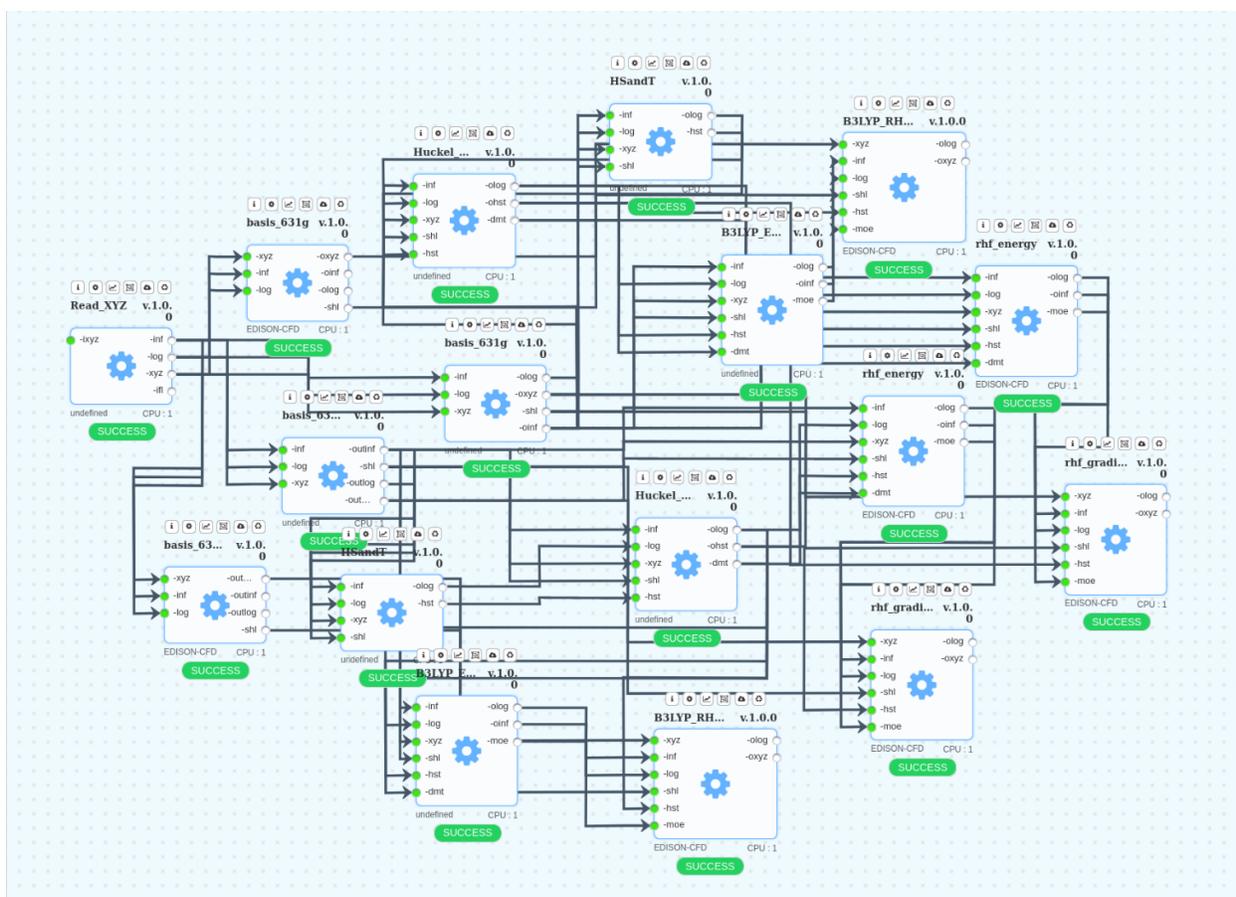
## 12.9 Donor-Acceptor interactions

The value of the donor-acceptor interaction can be estimated by comparing the total energy of the donor-acceptor complex and the total energy of its fragments. **DAinteraction** workflow allows simultaneous B3LYP/6-31G(d) energy calculations of two different fragments of a donor-acceptor complex. Both fragments must have a closed electronic shell. Geometries of both fragments (.xyz file format) should be provided as an input and two output files are generated as the result.



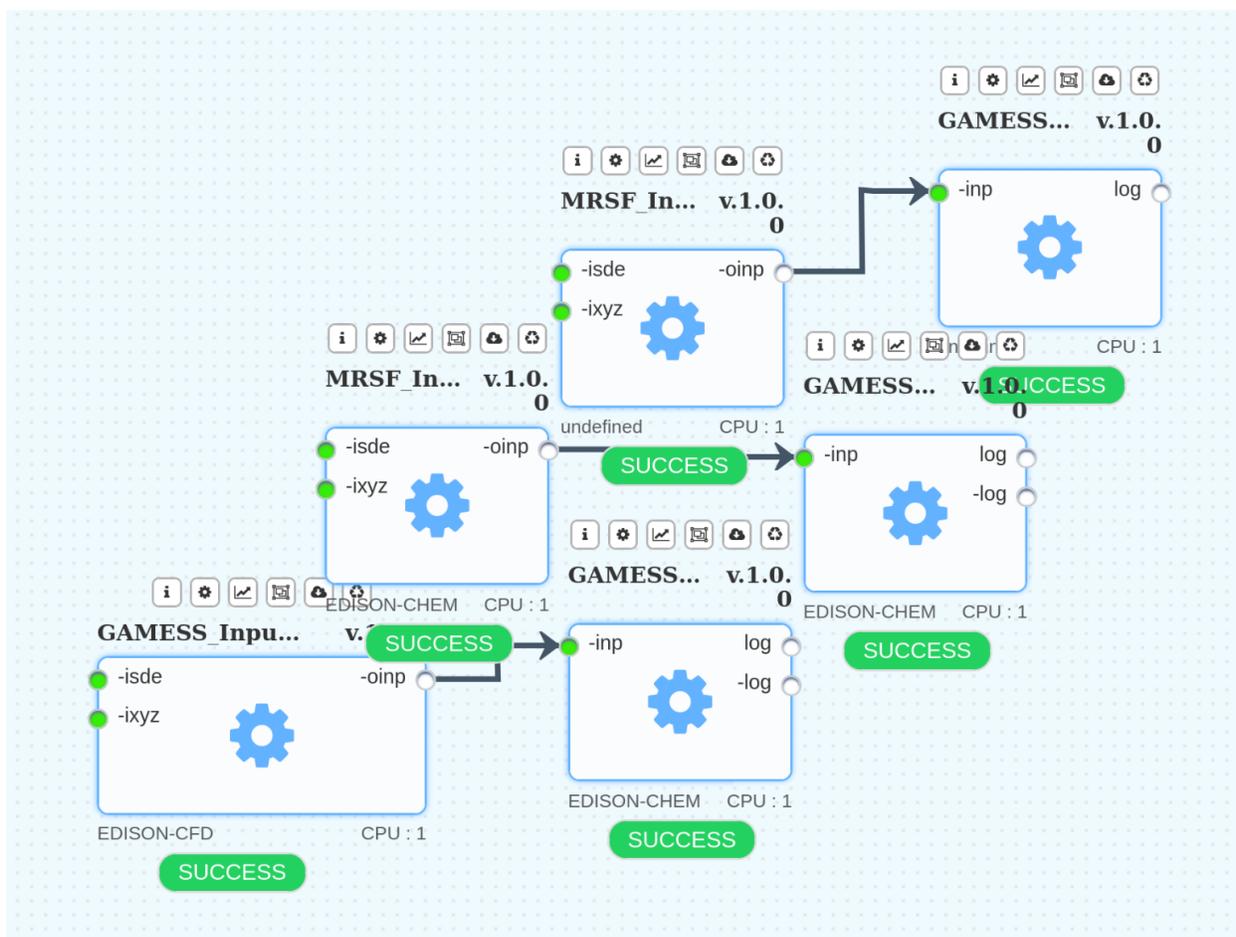
## 12.10 Gradient in different approximations

**Gradients** workflow allows simultaneous computations of Gradient in different approximations. This performs parallel calculations of the gradient matrix for the same molecule in the DFT(B3LYP) and HF methods with different sets of basis functions(6-31G and 6-31G(d)). The result of the calculations is 4 output files.



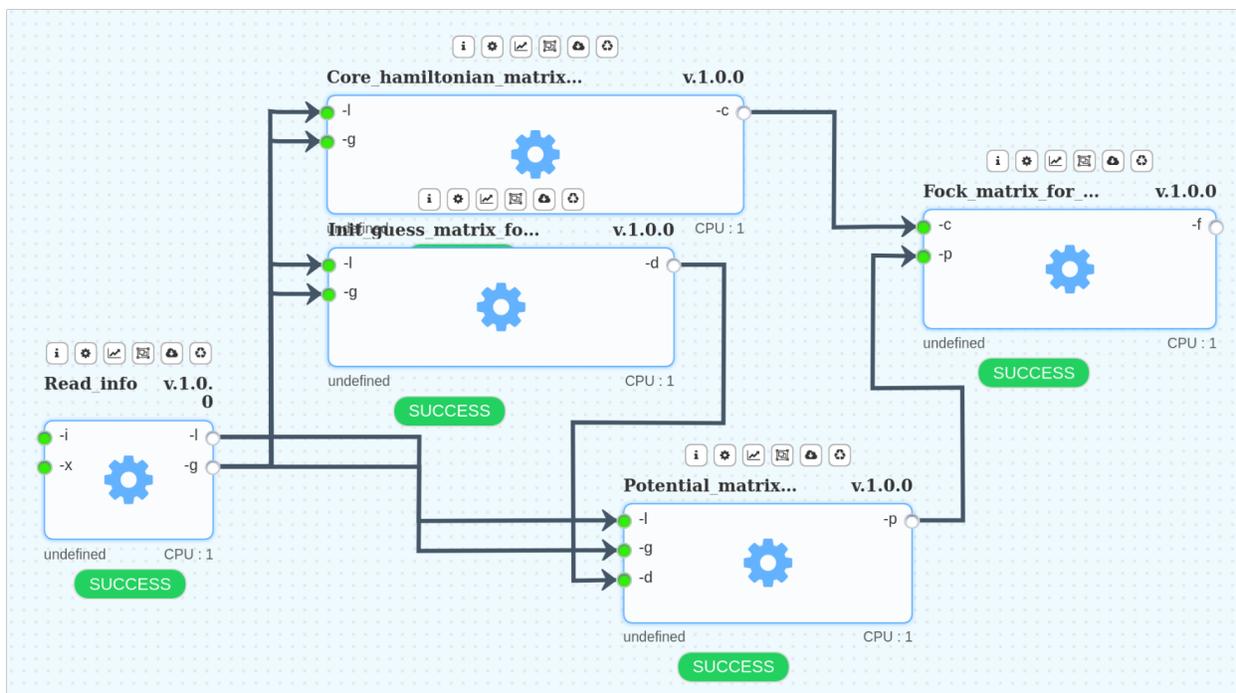
## 12.11 Ground and Excited states

Investigation of the excited states of a molecule usually requires a number of formally independent calculations for the same initial geometry of the molecule. Cascade **GroundState\_2MRSE** workflow allows such calculations to be carried out simultaneously. It is assumed that the same initial geometry is used for standard (single point properties or geometry optimization) calculations and for calculations (single point properties or geometry optimization) of the lowest and the second excited states.



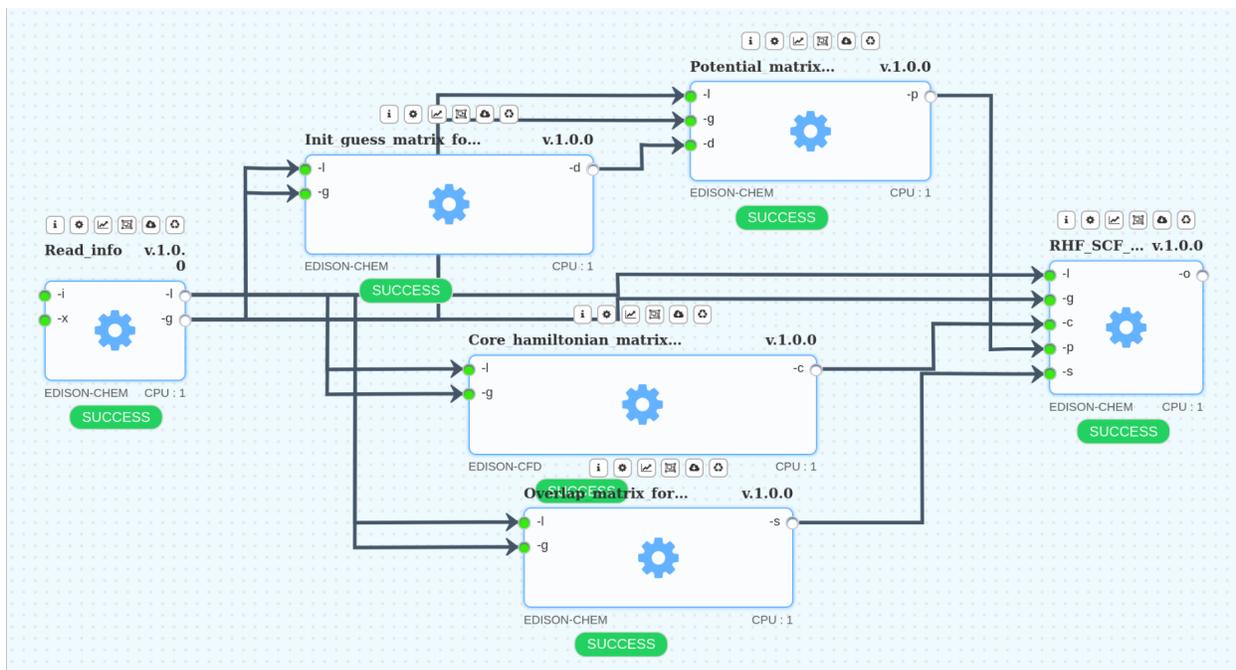
## 12.12 Intermediate matrix generation

For both developers' and educational purposes, it is extremely useful to be able to get the matrices used in standard energy calculations of the electronic structure. The **Fock\_matrix\_Workflow** workflow designed to derive the Fock matrix and such its constituents, the matrix of one-electron integrals and the matrix of potential energy. Calculations are available for a wide range of basis sets that can be selected graphically (-i port). The coordinates of the original matrix must be provided in XYZ format (-x port). The results are available in a .dat format file.



## 12.13 SCF procedure

As a base stone of the further development of modular variability, the **RHF\_SCF\_Procedure\_Workflow** workflow is proposed. It does ab-initio RHF SCF procedure. Calculations are available for a wide range of basis sets that can be selected graphically (`-i` port). The coordinates of the original matrix must be provided in XYZ format (`-x` port). The results are available in a `.dat` format file.





## UNDERSTANDING MRSF-TDDFT THEORY

### 13.1 TDDFT

Perhaps, the linear-response (LR)-TDDFT is the most practical formulation of TDDFT, which can be used if the external perturbation is relatively small in the sense that it does not completely destroy the ground-state of a given system. As a result, any variation of the system in the form of responses will depend only on the ground-state wave-function. For example, it simply means that any excited states can be obtained as derived quantities (response states) of ground state, which eliminates the needs for new additional methods to obtain excited states. It should be emphasized that not only excited states but also ground states can be obtained as response states by spin-flip techniques.

Although it has become one of the most popular quantum theories for excited states, there are a number of well-known failures of the popular LR-TDDFT method:

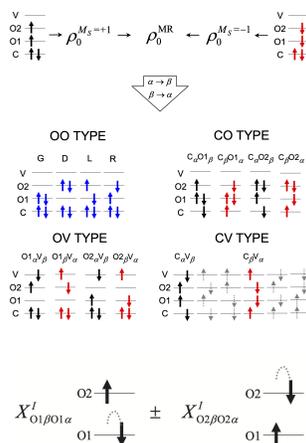
- failure to capture non-local properties for long-range charge transfer excited states,
- failure to capture double excitation characters in excited states,
- poor description of static correlation of the closed-shell reference state undergoing bond breaking
- and lack of coupling between the ground and excited states for conical intersections (CI) and avoided crossings.

### 13.2 SF-TDDFT

All these drawbacks can be efficiently corrected by the spin-flip (SF)-TDDFT method. It considers an open-shell high-spin triplet state (such as ROHF) as a reference instead of the closed-shell reference of LR-TDDFT. However, the conventional formulation of SF-TDDFT selects only one  $M_S = +1$  component of the triplet reference (See the figure below), which leads to a considerable **spin contamination** in the excited states. For example, the state with  $S^2 = 1.00$  in the Be Atom example below represents neither singlet nor triplet state.

The main source of spin contamination comes from the **red missing** responses (the red configurations below) leading to spin incompleteness. A fundamental solution for this problem is to include **red missing** configurations in the response space of SF-TDDFT.

Yet another important but not much appreciated source of spin contamination is from the **mismatched** contributions of **L** and **R** of **OO TYPE** (the blues). This is because they are coming from different orbitals' spin-flip transition as shown below. The former and latter comes from spin-flip of O1 and O2 orbitals, respectively. Sometimes, the **mismatched** contributions introduce a major spin contamination.



### 13.3 MRSF-TDDFT

The **red missing** configurations can be added into response space by the  $M_S = -1$  component of ROHF. A *hypothetical* single reference by combining  $M_S = +1$  and  $M_S = -1$  components of ROHF triplet can be constructed by a spinor-like transformation. See more [here](#). The resulting mixed-reference SF-TDDFT (MRSF-TDDFT) eliminates the spin contamination of SF-TDDFT, allowing automatic identification of the electronic states as singlets and triplets. It should be emphasized that MRSF-TDDFT produces not only **excited** but also **ground** electronic states. Therefore, open shell singlet such as diradicals, which cannot be studied by the Kohn-Sham DFT, can be naturally described by MRSF-TDDFT.

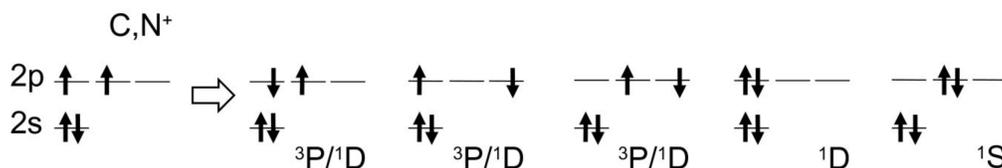


### 13.4 TDDFT vs. MRSF-TDDFT

There are multiple advantages of MRSF-TDDFT over TDDFT. Here, we list just some of them.

### 13.4.1 Missing States of C and N+ Atoms

In the case of C and N+ atoms, MRSF-TDDFT produces 8 states of  $^3P, ^1D, ^1S$  as



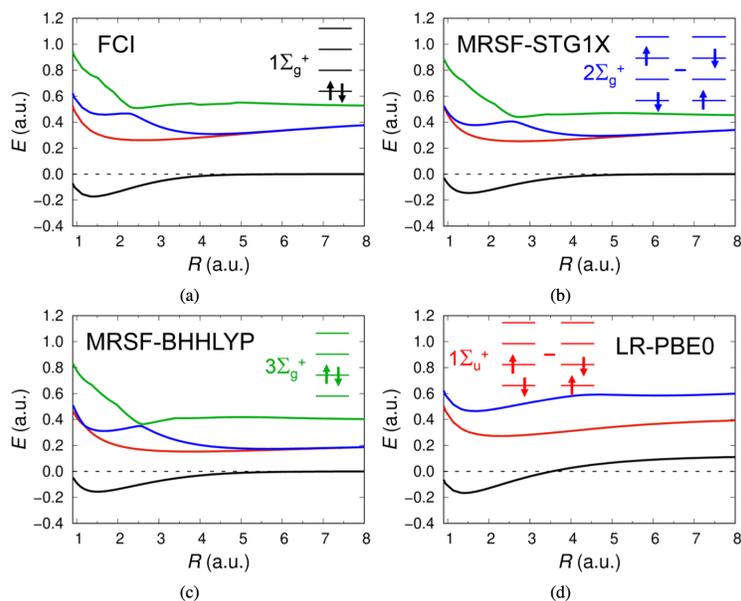
On the other hand, only 4 states can be represented by TDDFT as



This demonstrates that the conventional TDDFT misses many electronic states. Especially, the doubly excited configuration of  $^1S$  is completely missing in TDDFT.

### 13.4.2 Missing States of H<sub>2</sub>

The  $3\Sigma_g^+$  state (*green*), entirely composed of doubly excited configuration is missing in TDDFT (LR) but presents in MRSF-TDDFT as shown below in the case of  $H_2$  dissociation.

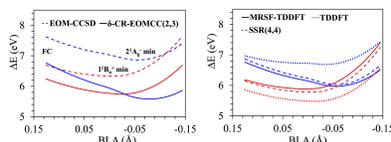


The doubly excited configuration makes greater contribution to the ground electronic state in the much stretched region, eventually leading to flattening the dissociation curve at the correct dissociation limit shown by the dashed line. On the other hand, the corresponding ground state of LR-PBE0 curve (TDDFT) has to be obtained by DFT, which does not have the correct asymptotic behaviour. See more [here](#).

### 13.4.3 Curve Crossing of Butadiene

In linear all-trans polyenes, internal conversion (IC) between the  $1^1B_u^+$  (optically bright at the Franck-Condon (FC) geometry, the red curve below) and the  $2^1A_g^-$  (optically dark, the blues) states has long been argued. However, it hasn't been proven until recently.

One of the difficulties arising when describing the  $1^1B_u^+$  and the  $2^1A_g^-$  states is their radically different nature. The former state comprises a one-electron HOMO  $\rightarrow$  LUMO transition and it displays pronounced ionic characteristics, while the latter is dominated by HOMO  $\rightarrow$  LUMO double excitations, requiring a balanced theory with both dynamic and nondynamic electron correlation.



As seen above, while both MRSF-TDDFT as well as REKS(4,4) produces the right CI as seen in  $\delta$ -CR-EOMCC(2,3), the TDDFT cannot. See more [here](#).

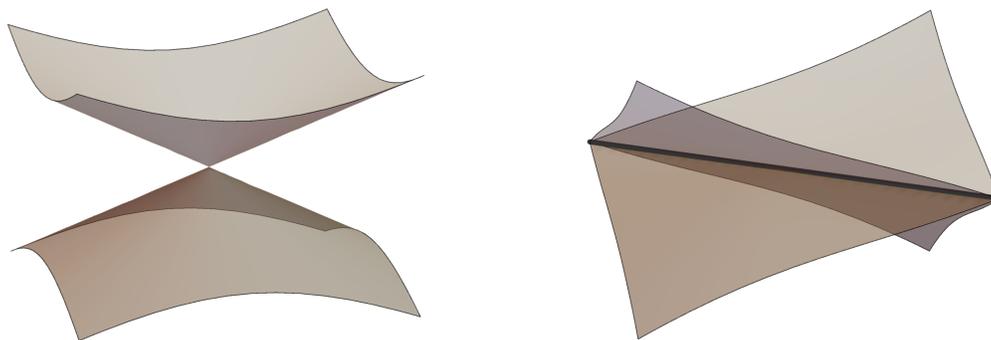
### 13.4.4 Conical Intersection Topology

Conical Intersection (CI) is a molecular geometry at which two (or more) adiabatic electronic states become degenerate. As the intersecting electronic states at a CI are coupled by the non-vanishing non-adiabatic coupling, CIs provide efficient funnels for the state to state population transfer mediated by the nuclear motion. The degeneracy of the intersecting states at a CI is lifted along two directions in the space of internal molecular coordinates  $\mathbf{Q}$ , which are defined by the gradient difference and derivative coupling vectors (**GDV** and **DCV**, respectively) given by

$$\mathbf{x}_1 = \frac{1}{2}(\nabla_{\mathbf{Q}}E_{S_1} - \nabla_{\mathbf{Q}}E_{S_0}) \quad (1a)$$

$$\mathbf{x}_2 = \langle \Psi_{S_1} | \nabla_{\mathbf{Q}} | \Psi_{S_0} \rangle \quad (1b)$$

for the case of a crossing between the ground ( $S_0$ ) and the lowest excited  $S_1$  singlet states. The degeneracy is lifted linearly along the **GDV** and **DCV** directions, which lends the potential energy surfaces (PESs) of the intersecting states the topology of a double cone below, hence the name. The remaining  $3N-8$  internal coordinates leave the degeneracy intact, thus defining the crossing seam (or the intersection space) of the CI.



The popular TDDFT fails to yield the correct dimensionality of the CI seam and predicts a linear crossing (3N-7) as shown in the right figure above. The problem is simply rooted in the absence of coupling between the ground (described by **reference** DFT) and first excited states (described by **response** of TDDFT).

On the other hand, both ground and first excited states of MRSF-TDDFT are obtained by the same **response**, producing the correct double cone topology. See more [here](#).

### 13.4.5 The Reference and Response Triplets

MRSF-TDDFT currently calculates singlet, triplet and quintet states using the ROHF triplet reference. As a result, one may be confused by the two different triplets of reference and responses. Triplet has three states with  $M_s = +1, 0, -1$ . Since they are energetically degenerated, nearly all quantum chemistry program calculates the simplest possible state of  $M_s = +1$  as a representative triplet state of ROHF. This particular triplet is the high-spin triplet. On the other hand, the triplet states generated by MRSF-TDDFT are  $M_s = 0$  triplet. Both *reference* as well as lowest *response* triplets are supposed to describe the same lowest triplet states with just different  $M_s$ . However, their energies are not degenerated. This is because the *reference* triplet is variationally obtained, while the *response* triplet is calculated by linear response theory. When you use MRSF-TDDFT, it is always better not to utilize the *reference* triplet in your analysis.

### 13.4.6 The Be Atom Case

This simple example can explain a great deal of MRSF-TDDFT calculations. The input example of MRSF-TDDFT with BHHLYP/6-31G basis set for GAMESS is

```
$CONTRL SCFTYP=ROHF RUNTYP=ENERGY DFTTYP=BHHLYP TDDFT=MRSF MULT=3 $END
$TDDFT  NSTATE=5 IROOT=1 MULT=1 $END
$BASIS  GBASIS=N31 NGAUSS=6 $END
$SCF    DIRSCF=.T. $END
$GUESS  GUESS=HCORE $END
$DATA
Be
C1
BERYLLIUM  4.0    0.0    0.0    0.0
$END
```

The important keywords are

- TDDFT=MRSF : Specifying MRSF-TDDFT method
- NSTATE=5 : Total number of excited state requested
- IROOT=1 : The target state. For example, the particular state for further geometry optimization.
- MULT=3 and =1 : MULT specifies spin multiplicity. Singlet and triplet are 1 and 3, respectively. As you can see, there are 2 different places of MULT. The one in \$CONTRL group specifies the spin multiplicity of reference wavefunction. In this case, the ROHF reference state. The MULT in \$TDDFT group specifies the spin multiplicity of response states, which are the ones generated by MRSF-TDDFT theory.

This input will prints out the SCF procedures of

ITER	EX	TOTAL ENERGY	E CHANGE	DENSITY CHANGE	DIIS ERROR	INTEGRALS
↪	SKIPPED					
1	0	-14.2683005477	-14.2683005477	1.320748431	0.0000000000	213
↪	0					
2	1	-14.4986108029	-0.2303102552	0.182168555	0.0000000000	213
↪	0					

(continues on next page)

(continued from previous page)

3	2	-14.5065330145	-0.0079222116	0.012352910	0.0000000000	213	↵
↵	0						
4	3	-14.5065504739	-0.0000174594	0.000562434	0.0000000000	213	↵
↵	0						
CONVERGED TO SWOFF, SO DFT CALCULATION IS NOW SWITCHED ON.							
5	4	-14.5662398400	-0.0596893661	0.035506584	0.0000000000	213	↵
↵	0						
6	5	-14.5666215197	-0.0003816797	0.004443966	0.0000000000	213	↵
↵	0						
7	6	-14.5666251825	-0.0000036628	0.000788830	0.0000000000	213	↵
↵	0						
8	7	-14.5666253040	-0.0000001215	0.000145036	0.0000000000	213	↵
↵	0						
DFT CODE IS SWITCHING BACK TO THE FINE GRID							
9	8	-14.5666031723	0.0000221317	0.000694477	0.0000000000	213	↵
↵	0						
10	9	-14.5666032709	-0.0000000986	0.000133009	0.0000000000	213	↵
↵	0						
11	10	-14.5666032742	-0.0000000033	0.000027388	0.0000000000	213	↵
↵	0						
12	11	-14.5666032743	-0.0000000001	0.000006101	0.0000000000	213	↵
↵	0						
13	12	-14.5666032744	-0.0000000000	0.000003515	0.0000000000	213	↵
↵	0						
14	13	-14.5666032744	-0.0000000000	0.000003071	0.0000000000	213	↵
↵	0						
-----							
ENERGY CONVERGED							
-----							
TIME TO FORM FOCK OPERATORS=		0.0	SECONDS (	0.0	SEC/ITER)		
FOCK TIME ON FIRST ITERATION=		0.0,	LAST ITERATION=	0.0			
TIME TO SOLVE SCF EQUATIONS=		0.0	SECONDS (	0.0	SEC/ITER)		
THE CONVERGED ORBITALS WILL UNDERGO GUEST/SAUNDERS							
CANONICALIZATION FOR SPIN-FLIP TDDFT.							
FINAL RO-BHLYP ENERGY IS		-14.5666032744	AFTER	14	ITERATIONS		

As you can see above, the final ROHF-DFT with BHLYP functional energy is -14.5666032744 Hartree, which corresponds to  $1s^2 2s^1 2p^1$  electron configuration. (The ground state electron configuration of Be atom is  $1s^2 2s^2$ ) There are 3  $2p$  orbitals, where one can put an electron. Since ROHF choose one of three possible  $2p$  orbitals, it inevitably breaks the symmetry of them and makes one of them more stable than the other two, which is exactly seen in the final orbitals below.

			1	2	3	4	5
			-4.3406	-0.1895	-0.0433	0.0141	0.0141
			A	A	A	A	A
1	BE	1 S	0.996731	-0.228170	0.000000	0.000000	0.000000
2	BE	1 S	0.027657	0.364242	0.000000	0.000000	0.000000
3	BE	1 X	0.000000	0.000000	0.067753	0.021479	0.409712
4	BE	1 Y	0.000000	0.000000	0.108319	0.404382	-0.030851

(continues on next page)

(continued from previous page)

5	BE 1 Z	0.000000	0.000000	0.548382	-0.082535	-0.044529
6	BE 1 S	-0.010506	0.689151	0.000000	0.000000	0.000000
7	BE 1 X	0.000000	0.000000	0.063955	0.035090	0.669317
8	BE 1 Y	0.000000	0.000000	0.102248	0.660611	-0.050398
9	BE 1 Z	0.000000	0.000000	0.517623	-0.134822	-0.072739
		6	7	8	9	
		0.3356	0.3384	0.3384	0.3614	
		A	A	A	A	
1	BE 1 S	0.000000	0.000000	0.000000	0.007075	
2	BE 1 S	0.000000	0.000000	0.000000	2.004865	
3	BE 1 X	-0.147030	0.017292	1.271128	0.000000	
4	BE 1 Y	-0.235062	1.255760	-0.046935	0.000000	
5	BE 1 Z	-1.190244	-0.250139	-0.147753	0.000000	
6	BE 1 S	0.000000	0.000000	0.000000	-1.925538	
7	BE 1 X	0.148722	-0.015718	-1.155390	0.000000	
8	BE 1 Y	0.237768	-1.141422	0.042661	0.000000	
9	BE 1 Z	1.203939	0.227360	0.134299	0.000000	

Even though, the three orbitals (3, 4 and 5 above) are degenerate, the orbital energy of 3 is -0.0433, while those of 4 and 5 are 0.0141 as a result of the ROHF calculation above.

The major transition contributions of each states by MRSF-TDDFT is seen from

```

-----
SPIN-ADAPTED SPIN-FLIP EXCITATIONS
-----

STATE # 1 ENERGY = -2.296236 EV
SYMMETRY OF STATE = A
S-SQUARED = 0.0000
DRF      COEF      OCC      VIR
-----
3 -0.992120      3 -> 2
5 -0.124389      2 -> 3

STATE # 2 ENERGY = 2.393067 EV
SYMMETRY OF STATE = A
S-SQUARED = 0.0000
DRF      COEF      OCC      VIR
-----
8 0.989883      3 -> 4
11 -0.122346     3 -> 5
17 -0.071605     3 -> 7

```

As seen above, first 2 **response states** are listed as STATE # 1 and 2 with the coefficients of major contributions. For example, the major transition of STATE # 1 is 3 (OCC) → 2 (VIR) spin flip transition, which means that the  $\alpha$  spin electron in orbital # 3 goes to  $\beta$  beta spin electron in orbital # 2 with the coefficient of -0.992120. As a result of this transition, the orbital # 2 now becomes doubly occupied with  $\alpha$  and  $\beta$  electrons. One can also see the state energy of -2.296236 eV of STATE # 1, which is the relative energy with respect to the ROHF **reference**. The program also lists the summary as

SUMMARY OF MRSF-DFT RESULTS							
STATE	ENERGY	EXCITATION	<S^2>	TRANSITION DIPOLE, A.U.			
← OSCILLATOR	HARTREE	EV		X	Y	Z	←
← STRENGTH							
1 NEGATIVE ROOT(S) FOUND.							
1 A	-14.6509883896	-2.296	0.0000	0.0000	0.0000	0.0000	0.0000
0 A	-14.5666032744	0.000		(REFERENCE STATE)			
2 A	-14.4786596677	2.393	0.0000	0.1468	-2.0547	0.3876	0.5048
3 A	-14.4786596263	2.393	0.0000	2.0757	0.0963	-0.2754	0.5048
4 A	-14.4703992515	2.618	0.0000	-0.2223	-0.3554	-1.7999	0.4112
5 A	-14.3795404853	5.090	0.0000	-0.0000	-0.0000	0.0000	0.0000
TRANSITION	EXCITATION	TRANSITION DIPOLE, A.U.			OSCILLATOR		
	EV	X	Y	Z	DIP	STRENGTH	
1 -> 2	4.689	0.1468	-2.0547	0.3876	2.0961	0.5048	
1 -> 3	4.689	2.0757	0.0963	-0.2754	2.0961	0.5048	
1 -> 4	4.914	-0.2223	-0.3554	-1.7999	1.8481	0.4112	
1 -> 5	7.386	-0.0000	-0.0000	0.0000	0.0000	0.0000	
2 -> 3	0.000	-0.0000	0.0000	-0.0000	0.0000	0.0000	
2 -> 4	0.225	0.0000	-0.0000	0.0000	0.0000	0.0000	
2 -> 5	2.697	-0.1368	-0.2186	-1.1066	1.1363	0.0853	
3 -> 4	0.225	0.0000	-0.0000	-0.0000	0.0000	0.0000	
3 -> 5	2.697	0.1126	0.1799	0.9107	0.9351	0.0578	
4 -> 5	2.472	-0.8435	-1.1528	0.3320	1.4665	0.1303	
SELECTING EXCITED STATE IROOT= 1 AT E= -14.6509883896							

STATE number 0 is the ROHF **reference**. The **response states** are from STATE number 1. In the case of above, the STATE 1 has the negative energy of -2.296 as compared to the **reference ROHF energy**. The first **response state**, STATE 1 corresponds to the ground singlet state, which is typically lower than the first triplet state. Therefore, the negative energy is not unusual. When you report the relative vertical excitation energy (VEE), you should use the lowest singlet state (STATE 1) as your reference state, not the **reference ROHF (STATE 0)**. The program also reports the corresponding  $S^2$ , transition dipole and oscillator strengths of each **response states**. The  $S^2$  indicates the spin state (0 = singlet, 2 = triplet states), while the oscillator strengths indicates the strength of absorption. As you can see, the **SELECTING EXCITED STATE IROOT= 1** indicates that the STATE 1, which is the ground singlet state is chosen for further calculations such as geometry optimizations.

The result of singlets and triplets by various methods are summarized in the table below. The MR-SF-TDDFT corresponds to the current example. The  $^1S$  is the STATE 1 above, which serves the reference energy of all the other states. The  $^3P_z, ^3P_x, ^3P_y$  should be degenerated. However, they are not exactly degenerated, since the orbital optimization step by **reference ROHF** breaks the symmetry among them by selectively choosing one particular orbital out of three  $p$  orbitals. As a result, MRSF-TDDFT produces 2.900 and 2.667 eV as compared to the  $^1S$  ground singlet state. Likewise, MRSF-TDDFT produces 4.913 and 4.690 eV VEEs of  $^1P_z, ^1P_x, ^1P_y$  singlet states. The  $S^2$  values are presented in the parentheses. As compared to MRSF-TDDFT, SF-TDDFT is missing one degenerate state of  $^1P_{x,y}$ . In fact, SF-TDDFT mixes it with  $^3P_{x,y}$  producing a half-half mixture of singlet and triplet states with the averaged singlet (2.667 eV of MRSF-TDDFT) and triplet (4.690 eV of MRSF-TDDFT) energy of 3.688 eV. This half-half mixture can be easily seen from its  $S^2$  value of 1.00 in the parentheses. The  $S^2 = 1.00$  represents neither singlet nor triplet state.

TABLE II. Ground state total energies (hartree) and excitation energies (eV) for the Be atom using a 6-31G basis set. The value of  $\langle S^2 \rangle$  is given in parentheses, in units of  $\hbar^2$ .

State	SF-TDDFT	MR-SF-TDDFT(0) <sup>a</sup>	MR-SF-TDDFT	MR-SF-CIS <sup>b</sup>
$^1S$	-14.650 997 (0.0003)	-14.650 997 (0.0)	-14.650 988 (0.0)	-14.583 980 (0.0)
$^3P_z$	2.877 (1.98)	2.899 (2.00)	2.900 (2.00)	2.107 (2.00)
$^3P_{x,y}$	3.688 (1.00)	3.688 (2.00)	2.667 (2.00)	2.107 (2.00)
$^1P_z$	4.935 (0.02)	4.913 (0.00)	4.913 (0.00)	6.042 (0.00)
$^1P_{x,y}$		3.688 (0.00)	4.690 (0.00)	5.952 (0.00)

<sup>a</sup>With the neglect of the pairing strengths.

<sup>b</sup>Using 100% HF exchange within the MR-SF-TDDFT formalism [i.e.,  $c_H = 1$  in Eq. (2.7)].



## UNDERSTANDING REKS THEORY

### 14.1 The KEYWORDS of REKS Theory in GAMESS

To start a REKS/SSR calculation the pre-optimized egeenvectors (orbitals) need to be supplied. The user is strongly advised not to use the default GAMESS orbital guess. The initial orbitals can be optimized, e.g., during an RHF (or RKS) single-point calculation. The standard mechanism of reading the guess orbitals (GUESS=MOREAD) is used.

The REKS/SSR calculation is invoked by setting SCFTYP=REKS in the \$CONTRL section. The REKS/SSR calculation is controlled by the following keywords, which should occur in the \$REKS section of the input; the \$REKS section is terminated by the \$END keyword.

```
rexType   = 0.....SA-REKS(2,2); default
            1.....SSR(2,2) = 2SI-2SA-REKS(2,2)
            2.....SSR(3,2) = 3SI-2SA-REKS(2,2)

rexTarget = 0.....the averaged state of SA-REKS; default
            1.....either PPS (of SA) or S0 (of SSR)
            2.....either OSS or S1

WPPS      = float.....weighting factor of the PPS state in the SA-REKS(2,2) ensemble (0
↳<= WPPS <= 0.5); default = 0.5
```

These keywords control the state to be calculated. If WPPS is set to 1.0 exactly, then a single state (not SA-REKS) REKS calculation will be performed. The analytical gradient, the (relaxed) density matrix, and (optional) the IPs or EAs are calculated for the target state selected by the rexTarget keyword.

The following keywords control the REKS SCF cycles.

```
rexShift  = float.....orbital level shift used in SA-REKS calculation; default = 0.0

rexDIIS   = yes/no....to do or not to do DIIS with REKS; default = yes
```

This keyword can be used to localize (or delocalize) the REKS active orbitals on a specific atom (group of atoms). It is useful for experts, when they know what they are doing.

```
rexLdL    = 0.....do nothing; default
            1.....localize the REKS active orbitals by minimizing the square of their
↳absolute overlap
            2.....delocalize the REKS active orbitals
```

The relaxed density is calculated by default, when the analytic gradient calculation is invoked; RUNTYP=GRADIENT or RUNTYP=OPTIMIZE. If RUNTYP=ENERGY, the relaxed density matrix of the target state can be calculated by:

rlxDen = yes/no....to do/**or not** to do the relaxed density matrix **for** the target state

Calculation of the ionization energies and the respective Dyson orbitals is controlled by the following keywords:

rexEKT = yes/no....to compute/**not** to compute the IPs **from the** Extended Koopmans'  $\downarrow$   
 $\hookrightarrow$ Theorem; default = no

The following keyword can be used in combination with rexEKT to calculate the electron affinities:

EKTEA = yes/no....to compute/**not** to compute the EAs **from the** Extended Koopmans'  $\downarrow$   
 $\hookrightarrow$ Theorem; default = no

These keywords are used to control the CP-REKS computation. CP-REKS obtains the orbital response (in the form of the Z-vector), which is needed for the analytic gradient, relaxed density matrix, and the EKS energies.

rexCG = yes/no....to use conjugate gradient/direct matrix inversion **for** solving CP-  
 $\hookrightarrow$ REKS equations; default = yes

rxCGit = N....max number of CG iterations; default = 100

rxCGth = float....convergence criterion **for** CG; default = 1.0E-6

The default CP-REKS convergence criterion ( $1.0 \cdot 10^{-6}$ ) provides for the analytic gradients, which are exact up to fifth or sixth digit. If more accurate gradients are required, the criterion can be decreased to  $1.0 \cdot 10^{-7}$ . Finer criteria may require more CP-REKS iterations, however hardly affect the accuracy of the computed gradient.

The REKS output:

Generally the REKS output is similar to an SCF output. The differences are the following:

During the SCF iterations the following information is reported:

```

=====
ITER EX      NR/2          TOTAL ENERGY  OFFDIAG FOCK      E CHANGE  DENSITY CHANGE  $\downarrow$ 
 $\hookrightarrow$ DIIS ERROR  VIR. SHIFT
=====
  1  0  0.602257551    -78.2672692323  0.001142168  -78.2672692323  0.000943561  $\downarrow$ 
 $\hookrightarrow$ 0.000000000  0.300000000
  2  1  0.602059530    -78.2672958088  0.000462067  -0.0000265766  0.000535339  $\downarrow$ 
 $\hookrightarrow$ 0.000000000  0.300000000
  3  2  0.602003714    -78.2673013806  0.000316234  -0.0000055717  0.000371568  $\downarrow$ 
 $\hookrightarrow$ 0.000000000  0.300000000
  ...
 23 22 0.602001670    -78.2673044126  0.000001411  -0.0000000000  0.000000922  $\downarrow$ 
 $\hookrightarrow$ 0.000000000  0.300000000
 24 23 0.602001743    -78.2673044127  0.000001155  -0.0000000000  0.000000725  $\downarrow$ 
 $\hookrightarrow$ 0.000000000  0.300000000

-----
          DENSITY CONVERGED
-----

TIME TO FORM FOCK OPERATORS=      0.6 SECONDS (      0.0 SEC/ITER)
TIME TO SOLVE SCF EQUATIONS=      0.0 SECONDS (      0.0 SEC/ITER)

```

(continues on next page)

(continued from previous page)

```
FINAL SA-REKS(2,2)-BHLYP ENERGY IS      -78.2673044127 AFTER  24 ITERATIONS
SA-REKS(2,2)-BHLYP: FON( 8)=1.20400349 FON( 9)=0.79599651
```

```
=====
SA: final printout
=====
```

```
Triplet:..... -78.338296401100
Doubly excited singlet (DES):..... -78.193048697566
Open Shell Singlet (OSS):..... -78.198436057138
Perfectly Paired Singlet (PPS):..... -78.336172768187
```

```
Lagrangian Wrs:..... -0.009374829317
```

```
SA: the averaged state is reported
```

```
SA: final energy =      -78.267304412662
```

```
SA: transition dipole between the OSS and PPS states
```

```
      Dipole = ( -0.09468334, -0.20302895, -0.70896273); |D| = 0.74351452 Debye
```

Most of the reported numbers are self-explanatory. For exceptionally smart, the FINAL SA-REKS(2,2)-BHLYP ENERGY is the final energy of the averaged state. The FONs are the fractional occupation numbers of the active orbitals; there are 2 such orbitals. The Triplet, DES, OSS, PPS energies are the energies of the triplet state, doubly excited singlet state, open-shell singlet state, the perfectly spin-paired singlet state calculated using the optimized SA-REKS orbitals and the FONs (for DES and PPS). In the triplet state, the active orbitals are occupied with the spin-up (i.e.,) electrons. The transition dipole is the transition dipole between the PPS and the OSS states. The Lagrangian Wrs is the Lagrangian matrix element between the active orbitals. No, it is non-zero in the open-shell SCF methods; may eventually vanish due to the symmetry.

The SSR calculation reports some more information. For example, for the same molecule (make an educated guess) the following lines appear in the output file after the Lagrangian line:

```
=====
3SI-2SA-REKS(2,2)-BHLYP states:
      E_k          C_{PPS}          C_{OSS}          C_{DES}
-----
SSR state 0      -78.336199986385      -0.99989739      -0.01420509      -0.00185074
SSR state 1      -78.214569352312          0.01194721      -0.75564298      -0.65487476
SSR state 2      -78.176888216893          0.00790406      -0.65482968          0.75573515
=====
```

```
SSR: the ground state is reported
```

```
SSR: unrelaxed occupation numbers: FON( 8)=1.20579413 FON( 9)=0.79420587
```

```
SSR: final energy =      -78.336199986385
```

```
SSR: transition dipole between the S1 and S0 states
```

(continues on next page)

(continued from previous page)

```
Dipole = ( -0.08125385, -0.17450835, -0.60891702); |D| = 0.63861983 Debye
```

This was an SSR(3,2) calculation. The energies of the three states are given in the table format. Typically, only the **|S0|** and **|S1|** energies are to be used. In each line of the table, the numbers after the energy value give the coefficients of the respective configurations (PPS, OSS, DES) in the respective SSR state.

After solving the SCF and the SSR secular equations, only the unrelaxed density matrix and the unrelaxed FONs are available. The relaxed density matrix (and the relaxed FONs) requires solving the CP-REKS equations and it becomes available after the CP-REKS section of the output. An example:

```
==> Solving CP-REKS equations
==> for SSR32 state S0
==> Memory required for CP-REKS is 515110 words
Solving CP-REKS equations by CG method
convergence threshold 0.000001000000 within maximum of 100 iterations
====> Search for Z-vector guess in the input file.
====> Z-vector found in the input file. Read as a guess.
CG: iteration 0, residue 0.003997135930
CG: iteration 1, residue 0.003303471349
CG: iteration 2, residue 0.000965987510
CG: iteration 3, residue 0.000293042567
CG: iteration 4, residue 0.000138157639
CG: iteration 5, residue 0.000033103894
CG: iteration 6, residue 0.000012019543
CG: iteration 7, residue 0.000003573194
CG: iteration 8, residue 0.000001443686
CG: iteration 9, residue 0.000000376722
CG: Residue ( 0.000000376722) is less than the threshold ( 0.000001000000)

==> CP-REKS done
```

The relaxed occupation numbers are printed only when the EKT calculation is requested. Example:

```
SSR: relaxed occupation numbers: FON( 8)=1.16274571 FON( 9)=0.81971186
Extended Koopmans' Theorem for Ionization Energies
==> EKT: there are 9 Dyson's orbitals with non-zero strengths
```

```
-----
EKT orbitals, energies, and pole strengths
-----
```

			1	2	3	4	5	
ENERGY			-10.632402	-10.613955	-0.801197	-0.672061	-0.526661	
STRENGTH			1.000000	1.000000	0.999954	0.999945	0.999161	
			A	A	A	A	A	
1	C	1	S	0.994153	-0.004312	-0.106364	-0.161461	-0.024532
2	C	1	S	0.038580	-0.000067	0.205788	0.321172	0.050538
3	C	1	X	-0.000146	-0.000017	-0.033281	-0.033301	0.051425
4	C	1	Y	-0.000057	-0.000031	-0.010352	0.021109	0.076615
5	C	1	Z	-0.000016	0.000037	-0.053451	0.089955	0.089309
6	C	1	S	-0.009647	0.002203	0.194890	0.352609	0.063799
7	C	1	X	0.000739	-0.000270	-0.012144	-0.009870	0.032084
8	C	1	Y	0.000072	-0.000098	-0.003239	0.005316	0.043842

(continues on next page)

(continued from previous page)

9	C	1	Z	0.000313	-0.001517	-0.010751	0.040195	0.039508
10	C	1	XX	-0.006140	-0.000095	-0.000146	0.007445	0.002667
11	C	1	YY	-0.006184	-0.000132	-0.007308	-0.009427	-0.002450
12	C	1	ZZ	-0.006283	-0.000291	0.010781	-0.001851	0.000039

The EKT output is similar to the usual MO output, except that the Dyson orbital norms are reported as the STRENGTH(s) of the respective ionizations. In the states with fractionally occupied orbitals, the norms become fractional. After the CP-REKS calculation, the subsequent density matrix analysis (e.g., the Mulliken charges) employs the relaxed density matrix.

## 14.2 Backgrounds

The Manual for the rex computations is provided [here](#).



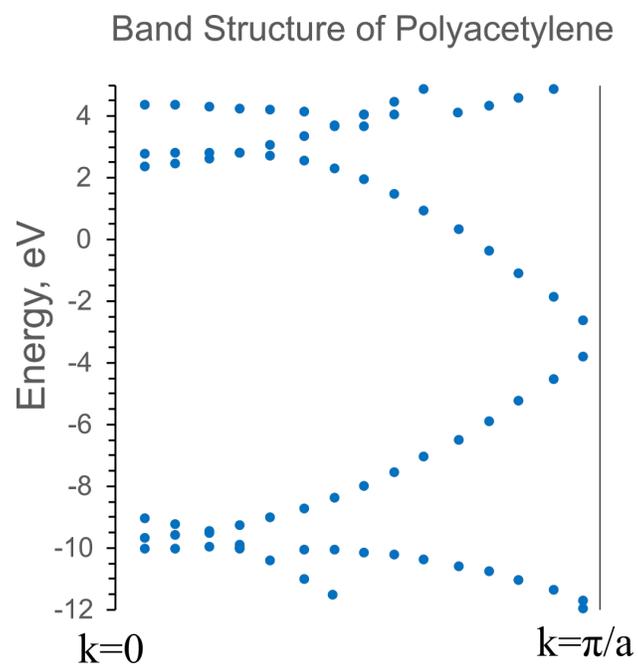
## BAND STRUCTURE

Band structure computations are available on the top of standard workflow SCF/post-SCF molecular computations with the BANDS module. Bands calculations make sense if the molecular structure is an ordered oligomer with insignificant deviations from periodic regularity.

To carry out such calculations, restrictions are introduced on the order of atoms in the initial structure: 1) periodically repeated units are given one by one from the one end of the oligomer to the other; 2) the order of atoms in each unit must be identical; 3) terminal groups of atoms (normally terminal H atoms) must be numbered last. The keyword ATOM\_CELL must be given in the mqcp input file to define the number of atom in the unit cell. Optional keywords EN\_LOWER and EN\_UPPER can be specified in the input file to define desired energy range (in eV). The input example is the following:

```
The input example of [C2H2]15 oligomer.
natom=62
6.0      0.305450260      -17.455528492      0.000000000
1.0      1.375447159      -17.458106142      0.000000000
6.0     -0.386647810      -16.250082413      0.000000000
1.0     -1.456644709      -16.247504763      0.000000000
6.0      0.311250090      -15.047984853      0.000000000
...
6.0      0.386647810      16.250082413      0.000000000
1.0      1.456644709      16.247504763      0.000000000
6.0     -0.305450260      17.455528492      0.000000000
1.0     -1.375447159      17.458106142      0.000000000
1.0      0.231780490      18.380884162      0.000000000
1.0     -0.231780490      -18.380884162      0.000000000
charge=0
Hamilton= dft
Basis=6-31G(d)
charge=0
scftype=rhf
runtype=energy
atoms_cell=4
en_lower=-12
en_upper=5
$libxc functional=b3lyp5 $end
```

As a result, a table of the dependence of the energy of orbitals on the wave vector is given in the output file. The figure below shows a graphical representation of the dispersion of orbital energies in the Brillouin zone of polyacetylene as the result of calculations.



## DYSON'S ORBITALS

Quasiparticle states calculations are available with IPDO and EADO modules. The input file is a standard MRSF input with additional options, MREKT=.T. for IPDO and MRDAE=.T. for EADO, in TDDFT section. The particular computation examples are the following.

### 16.1 1. Calculation of Ionization Potentials and Dyson's Orbitals of based on EKT.

**IonizationPotentials** presents ionization energies of first 6th Dyson's orbitals and symmetries of orbital electron distributions calculated for pyridine and thiophene using the functional BH&HLYP with the basis sets 6-311G(d,p).

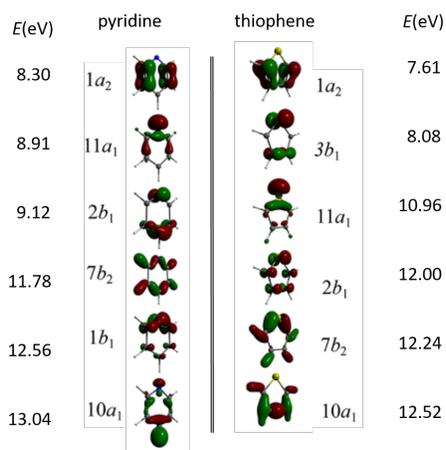
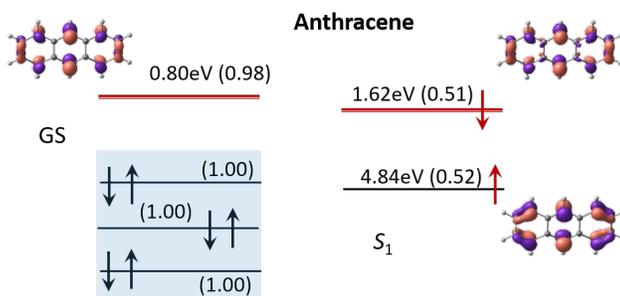
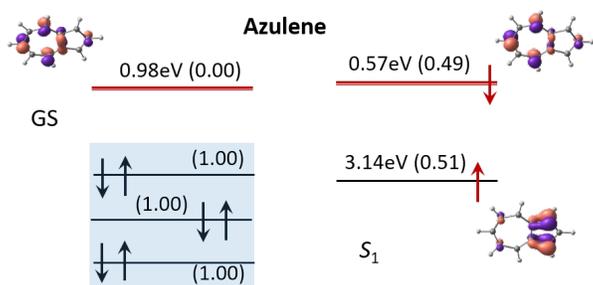


Fig. 1: First six Dyson's HOMOs and the corresponding ionization potentials for the pyridine and thiophene molecules

### 16.2 2. Calculation of Electron Affinities and Dyson's Orbitals based on EKT.

**ElectronAffinities** presents energies of electron affinities and orbital electron distributions for Ground (GS) and first excited states  $S_1$  of azulene and anthracene calculated using BH&HLYP with the basis sets 6-311G(d,p). Electron occupation corresponds to (1-norm) in parentheses where double is close to 1 and single is about a half.

Electronic affinities of azulene and anthracene molecules



MINIMUM ENERGY PATH BY GEODESIC INTERPOLATION

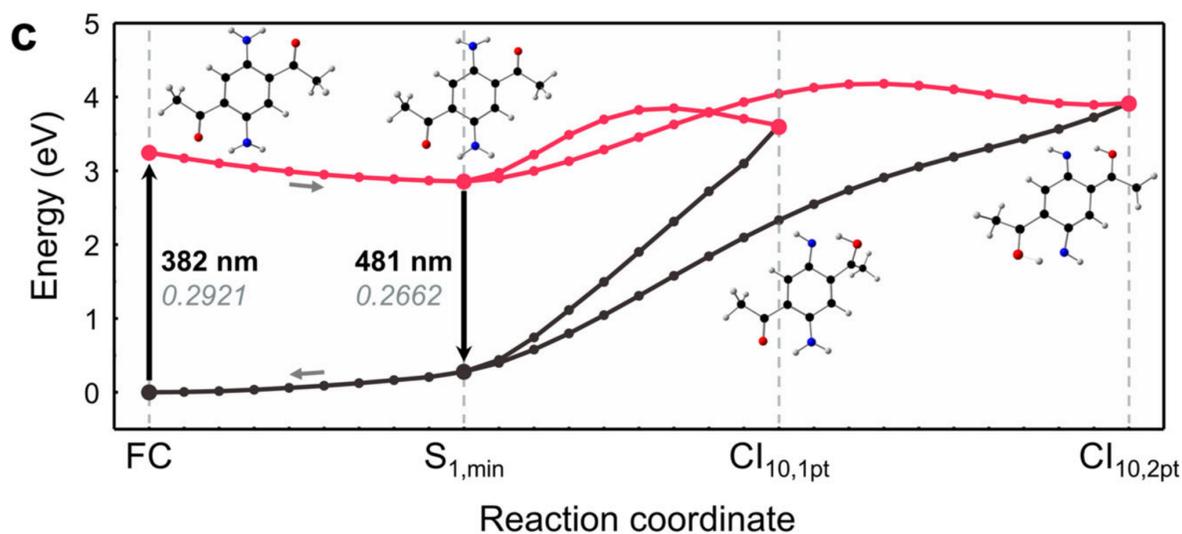
```
26
slfc
C 3.5074247314 0.3939945603 0.9561861839
C 3.7430749260 1.7842751404 0.8448560724
C 2.2261337252 -0.0579432126 0.6663989400
C 1.1840656346 0.7763155739 0.2723400367
C 1.4197493849 2.1665706794 0.1609385874
C 2.7010691216 2.6184919661 0.4506994956
N 4.4929415572 -0.5096377488 1.2756518870
N 0.4342322172 3.0701092111 -0.1585939230
C -0.1373319954 0.1920310365 -0.0536972344
O -1.0495946023 0.8670208928 -0.4881352761
C -0.3615458522 -1.2884341119 0.1482942157
C 5.0646069048 2.3685208639 1.1706703299
O 5.9766508781 1.6937349040 1.6058692066
C 5.2892619238 3.8486101124 0.9671948773
H 2.0490335062 -1.1175514715 0.7382152340
H 2.8781390989 3.6780910695 0.3788567937
H 5.3197918007 -0.1225164745 1.6879206682
H 4.1838944160 -1.3832384608 1.6534377841
H -0.3927688414 2.6829179836 -0.5705394985
H 0.7432748350 3.9437964192 -0.5362096520
H -1.3948773182 -1.5070508453 -0.0898719622
H -0.1567258677 -1.5836908173 1.1740794058
H 0.2849248363 -1.8744653080 -0.5007717299
H 6.3226530998 4.0671439452 1.2050330686
H 4.6432258826 4.4360613603 1.6155557834
H 5.0839659969 4.1428327319 -0.0588392940
26
slmin
C 3.4761900485 0.4269445477 1.0005782683
C 3.7810792620 1.8283788032 0.8640811736
C 2.1990570653 -0.0627847739 0.6962965902
C 1.1464985855 0.7318890499 0.2517597757
C 1.4516080827 2.1332233599 0.1146131197
C 2.7285626812 2.6230763437 0.4194472611
N 4.4351827379 -0.4070644472 1.4237551748
N 0.4928983086 2.9670874961 -0.3093993543
C -0.1623213857 0.1965618224 -0.0522583917
O -1.1020148873 0.9007180314 -0.4528085932
```

(continues on next page)

(continued from previous page)

C	-0.3988330787	-1.2844624423	0.1254910682
C	5.0895838269	2.3638539859	1.1690337052
O	6.0295679290	1.6596430324	1.5687918298
C	5.3255065354	3.8452995076	0.9934910447
H	2.0427184108	-1.1214526583	0.8203282149
H	2.8847328238	3.6818009713	0.2955286689
H	5.3400891076	0.0019046668	1.6309033353
H	4.2571999553	-1.3856935040	1.5299325718
H	-0.4120348220	2.5582687450	-0.5168667968
H	0.6706893292	3.9457379869	-0.4158557266
H	-1.4321566221	-1.4910584825	-0.1245708569
H	-0.2139476838	-1.6039518689	1.1499509230
H	0.2447709419	-1.8771143289	-0.5232477595
H	6.3585566810	4.0522547842	1.2444582748
H	4.6811311911	4.4366253746	1.6427089996
H	5.1409549758	4.1663039968	-0.0306025204

- This is example of xyz file to generate geodesic MEP used in Nat Commun 12, 5409 (2021).
- To generate geodesic MEP, one has to create xyz file that contains two geometries (starting / end points)



## NUCLEUS INDEPENDENT CHEMICAL SHIFT (NICS) CALCULATION USING DALTON PACKAGE

```
### dal input
**DALTON INPUT
.RUN PROPERTIES
**WAVE FUNCTION
.HF
.MCSCF
*CONFIGURATION INPUT
.SYMMETRY
1
.SPIN MULTIPLICITY
1
.INACTIVE
50
.CAS SPACE
2
.ELECTRONS
2
*OPTIMIZATION
.STATE
2
**PROPERTIES
.SHIELD
**END OF DALTON INPUT

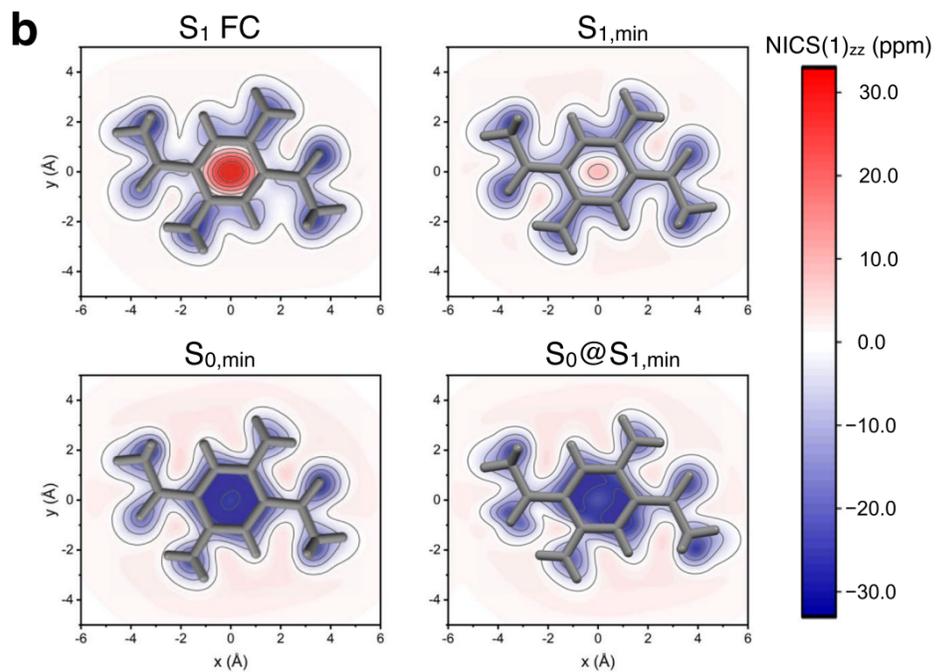
### mol input
ATOMBASIS
p-DAPA
comment
Atomtypes=5 NoSymmetry
Charge=8.0 Atoms=2 Basis=6-31G*
O      6.842672938      -1.322637495      -0.162989887
O      -6.842806547      1.322426516      0.164581259
Charge=7.0 Atoms=2 Basis=6-31G*
N      -2.873078816      4.437821733      -0.186528694
N      2.873188598      -4.437720436      0.186773595
Charge=6.0 Atoms=10 Basis=6-31G*
C      -1.509342993      2.231853428      -0.035812371
```

(continues on next page)

(continued from previous page)

C	-2.647694531	-0.186634704	-0.027091463
C	1.114090031	2.323111837	-0.001476878
C	2.647736711	0.186558340	0.027630884
C	1.509371237	-2.231894884	0.036199806
C	-1.114091660	-2.323119957	0.001798048
C	5.430726067	0.482124954	-0.014101475
C	6.560085910	3.101394554	0.113080808
C	-5.430801807	-0.482151274	0.014160237
C	-6.560086713	-3.101092723	-0.116013825
Charge=1.0 Atoms=12 Basis=6-31G*			
H	1.982920439	4.162690998	-0.032902303
H	-1.982852555	-4.162706982	0.033189390
H	-4.716232184	4.292201888	0.219892747
H	-2.008141452	5.995991523	0.445999286
H	4.716506059	-4.291926296	-0.219004079
H	2.008137723	-5.995983626	-0.445424854
H	8.596643166	2.912195785	0.167509737
H	5.922869022	4.110825239	1.784525572
H	6.031880619	4.215477089	-1.531034553
H	-8.596561221	-2.911839727	-0.171084349
H	-6.033129431	-4.217774292	1.526944615
H	-5.921352775	-4.108991587	-1.787906959
Charge=0.0 Atoms=5 Basis=pointcharge			
X	0.000021090	-0.000038182	0.000000000
X	0.000021090	-0.000038182	0.944862994
X	0.000021090	-0.000038182	1.889725989
X	0.000021090	-0.000038182	3.779451977
X	0.000021090	-0.000038182	5.669177966

- This is NICS calculation on top of CASSCF(2,2)/6-31G\* wavefunction used in Nat Commun 12, 5409 (2021).
- Note that Dalton requires two inputs (input.dal / input.mol).
- We calculate NICS values using ghost atoms (X) that are put above the center of benzene ring (0, 0.5, 1.0, 2.0, 3.0 Angstrom).





MAXIMUM OVERLAP METHOD (MOM) CALCULATION

```

$CONTRL SCFTYP=ROHF RUNTYP=ENERGY DFTTYP=BP86 ICHARG=1
TDDFT=MRSF MAXIT=200 MULT=3 ISPHER=1 $END
$TDDFT NSTATE=8 IROOT=1 MULT=1 $END
$SCF DIRSCF=.T. diis=.t. soscf=.f. damp=.t. shift=.t.
swdiis=1e-4 mom=.t. $END
$BASIS GBASIS=SPK-DZP $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$GUESS guess=moread norb=586 $END
$DFT DFTTYP=BP86 swoff=1e-6 sgl=.t. NRAD0=50 NLEB0=110 $END
$DATA
MeCbl 2.00
C1
N 7.0 -0.0624430 -0.1132311 1.7483714
Co 27.0 -0.0061051 0.0164836 -0.4207477
C 6.0 -3.2784169 -0.1968389 -0.4357658
C 6.0 -0.0480683 0.1521549 -2.4156871
H 1.0 0.3963554 1.1102324 -2.7139713
H 1.0 0.5126009 -0.6847181 -2.8508229
H 1.0 -1.0991660 0.1044608 -2.7278236
H 1.0 -4.3662141 -0.2693960 -0.4385652
C 6.0 -2.5657964 -1.3936154 -0.4756937
C 6.0 -3.2295671 -2.7497398 -0.5028914
C 6.0 -2.7304937 1.0827493 -0.4201847
C 6.0 -3.5612378 2.3438803 -0.4288008
N 7.0 -1.2167798 -1.4938324 -0.4945363
N 7.0 -1.4047733 1.3581454 -0.4000123
C 6.0 -0.8316992 -2.8219770 -0.5260048
C 6.0 -2.0444940 -3.7267350 -0.5348367
C 6.0 -1.1959306 2.7224982 -0.4943585
C 6.0 -2.5151505 3.4527244 -0.6185229
C 6.0 0.4657791 -3.2769236 -0.5319828
H 1.0 0.6330652 -4.3534559 -0.5386831
C 6.0 0.0311018 3.3424730 -0.4891331
H 1.0 0.0596883 4.4279750 -0.5769618
C 6.0 1.5975802 -2.4173960 -0.5539798
C 6.0 3.0481635 -2.8466095 -0.6310661
C 6.0 1.2614677 2.6434512 -0.3505205
C 6.0 2.6358480 3.2684188 -0.2365601
N 7.0 1.4890249 -1.1132326 -0.5421504
C 6.0 2.8023916 -0.4595798 -0.7444275

```

(continues on next page)

(continued from previous page)

```

N 7.0  1.3213202  1.3394639  -0.2641629
C 6.0  2.7024681  0.8684759  -0.0062929
C 6.0  0.0790660  -1.2325912  2.5477945
H 1.0  0.2458472  -2.2172693  2.1261106
C 6.0  -0.0266652  -0.8707181  3.8706666
H 1.0  0.0256305  -1.4424043  4.7896892
N 7.0  -0.2341697  0.4910496  3.8646645
H 1.0  -0.3556976  1.0818008  4.6823436
C 6.0  -0.2494666  0.9121245  2.5755304
H 1.0  -0.3961902  1.9478540  2.2904586
C 6.0  3.8036821  -1.5364699  -0.3063101
C 6.0  3.5758765  2.0568291  -0.4362581
H 1.0  2.9104427  -0.2457789  -1.8249529
H 1.0  2.7981622  0.6753905  1.0784427
H 1.0  3.8415539  1.9573456  -1.4995035
H 1.0  4.5028482  2.1339770  0.1439416
H 1.0  -3.8704695  -2.8826991  0.3807287
H 1.0  2.7862550  4.0687539  -0.9735636
H 1.0  -2.5903115  3.9200573  -1.6117427
H 1.0  2.7552589  3.7222912  0.7614094
H 1.0  -4.3186871  2.3171427  -1.2235215
H 1.0  3.9829582  -1.4613543  0.7769532
H 1.0  4.7676256  -1.4579825  -0.8226923
H 1.0  3.2737546  -3.2072263  -1.6488382
H 1.0  3.2736080  -3.6702306  0.0595988
H 1.0  -2.5958611  4.2599996  0.1211719
H 1.0  -2.0432622  -4.3609201  -1.4323997
H 1.0  -4.1031145  2.4414718  0.5246480
H 1.0  -2.0293491  -4.4021608  0.3318621
H 1.0  -3.8822994  -2.8415614  -1.3831333
$END
--- OPEN SHELL ORBITALS --- GENERATED AT Tue Mar  9 13:41:44 2021
Ref orbital at 1.85
E(RO-BP86)= -2604.7250403089, E(NUC)= 3915.3398553345, 100 ITERS
$VEC
~~~
$END

```

- This is example input of using MOM. This method can be used when you want to use specific reference orbital given in \$VEC.

## LR-TDDFT CALCULATION

```
$CONTRL SCFTYP=RHF RUNTYP=ENERGY DFTTYP=B3LYP ICHARG=0
TDDFT=EXCITE MAXIT=200 MULT=1 ISPHER=1 $END
$TDDFT NSTATE=6 MULT=1 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
Butadiene TDDFT
Cnh 2
C      6.0      -0.410990219      -1.798958603      0.000000000
C      6.0      -0.559475119      -0.470573512      0.000000000
H      1.0      -1.263114858      -2.463113554      0.000000000
H      1.0      -1.554995711      -0.040726039      0.000000000
H      1.0      -0.571968115      2.252448386      0.000000000
$END
```

- This is LR-TDDFT/B3LYP input for butadiene used in J. Phys. Chem. Lett. 2021, 12, 9720-9729



## DELTA-CR-EOMCC(2,3) CALCULATION

```
$CONTRL SCFTYP=RHF RUNTYP=ENERGY ICHARG=0 CCTYP=CR-EOML
MAXIT=200 MULT=1 ISPHER=1 NUMGRD=.T. $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=1000 $END
$EOMINP nstate(1)=2,0,0,1 MTRIP=4 $END
$DATA
Buta delta-CR-EOMCC(2,3)
Cnh 2
C 6.0 -0.410990219 -1.798958603 0.000000000
C 6.0 -0.559475119 -0.470573512 0.000000000
H 1.0 -1.263114858 -2.463113554 0.000000000
H 1.0 -1.554995711 -0.040726039 0.000000000
H 1.0 0.571968115 -2.252448386 0.000000000
$END
```

- This is delta-CR-EOMCC(2,3) input for butadiene used in J. Phys. Chem. Lett. 2021, 12, 9720-9729



## CASSCF CALCULATION

```
$CONTRL SCFTYP=MCSCF RUNTYP=ENERGY ICHARG=0
MAXIT=200 MULT=1 ISPHER=1 MLEVEL=2 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=1000 $END
$GUESS GUESS=MOREAD NORB=17 NORDER=1 IORDER(13)=14,13,15,16,17 $END
$MCSCF CISTEP=ALDET FOCAS=.F. SOSCF=.T. MAXIT=200 $END
$DET GROUP=C2H STSYM=BU
NCORE=13 NACT=4 NELS=4 NSTATE=10 WSTATE(1)=1,1 $END
$DATA
Butadiene CASSCF(4.4)
Cnh 2

C      6.0      -0.410990219      -1.798958603      0.000000000
C      6.0      -0.559475119      -0.470573512      0.000000000
H      1.0      -1.263114858      -2.463113554      0.000000000
H      1.0      -1.554995711      -0.040726039      0.000000000
H      1.0      -0.571968115      2.252448386      0.000000000
$END
--- NATURAL ORBITALS OF MCSCF --- GENERATED AT 21:17:43 13-MAY-2021
Orbital at FC geometry
E(MCSCF)=      -154.9954698060, E(NUC)=      103.8314056120
$VEC
~~~
$END
```

- This is CASSCF(4,4) input for butadiene used in J. Phys. Chem. Lett. 2021, 12, 9720-9729
- We calculate 2 state average CASSCF of Bu state of butadiene in  $C_{2h}$  symmetry.
- Note that we reordered initial orbital.



## XMCQDPT2 CALCULATION

```
$CONTRL SCFTYP=MCSCF RUNTYP=ENERGY ICHARG=0
MAXIT=200 MULT=1 ISPHER=1 MLEVEL=2 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=1000 $END
$GUESS GUESS=MOREAD NORB=17 NORDER=1 IORDER(13)=14,13,15,16,17 $END
$MCSCF CISTEP=ALDET FOCAS=.F. SOSCF=.T. MAXIT=200 $END
$DET GROUP=C2H STSYM=BU
NCORE=13 NACT=4 NELS=4 NSTATE=10 WSTATE(1)=1,1 $END
$MRMP MRPT=MCQDPT $END
$MCQDPT STSYM=BU NSTATE=10 KSTATE(1)=1,1 WSTATE(1)=1,1
XZERO=.T. $END
$DATA
Butadiene XMCQDPT2(4.4)
Cnh 2
C      6.0      -0.410990219      -1.798958603      0.000000000
C      6.0      -0.559475119      -0.470573512      0.000000000
H      1.0      -1.263114858      -2.463113554      0.000000000
H      1.0      -1.554995711      -0.040726039      0.000000000
H      1.0      -0.571968115      2.252448386      0.000000000
$END
--- NATURAL ORBITALS OF MCSCF --- GENERATED AT 21:17:43 13-MAY-2021
Orbital at FC geometry
E(MCSCF)=      -154.9954698060, E(NUC)=      103.8314056120
$VEC
~~~
$END
```

- This is XMCQDPT2(4,4) input for butadiene used in J. Phys. Chem. Lett. 2021, 12, 97209729
- Note that we reordered initial orbital.

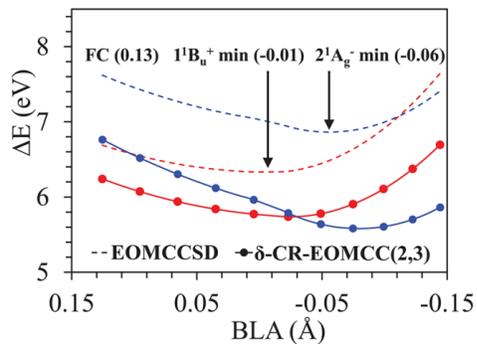


## MRMP2 CALCULATION

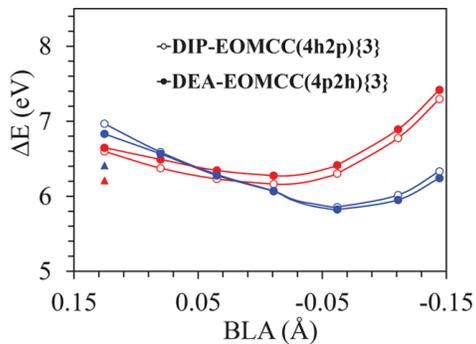
```
$CONTRL SCFTYP=MCSCF RUNTYP=ENERGY ICHARG=0
MAXIT=200 MULT=1 ISPHER=1 MLEVEL=2 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=1000 $END
$GUESS GUESS=MOREAD NORB=17 NORDER=1 IORDER(13)=14,13,15,16,17 $END
$MCSCF CISTEP=ALDET FOCAS=.F. SOSCF=.T. MAXIT=200 $END
$DET GROUP=C2H STSYM=BU
NCORE=13 NACT=4 NELS=4 NSTATE=10 WSTATE(1)=1,1 $END
$MRMP MRPT=MCQDPT $END
$MCQDPT STSYM=BU NSTATE=10 KSTATE(1)=1,0 WSTATE(1)=1,0
XZERO=.F. $END
$DATA
Butadiene MRMP2
Cnh 2

C      6.0      -0.410990219      -1.798958603      0.000000000
C      6.0      -0.559475119      -0.470573512      0.000000000
H      1.0      -1.263114858      -2.463113554      0.000000000
H      1.0      -1.554995711      -0.040726039      0.000000000
H      1.0      -0.571968115      2.252448386      0.000000000
$END
--- NATURAL ORBITALS OF MCSCF --- GENERATED AT 21:17:43 13-MAY-2021
Orbital at FC geometry
E(MCSCF)=      -154.9954698060, E(NUC)=      103.8314056120
$VEC
~~~
$END
```

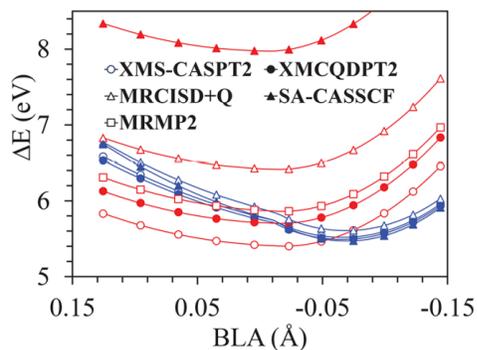
- This is MRMP2 input for butadiene used in J. Phys. Chem. Lett. 2021, 12, 9720-9729
- Note that we reordered initial orbital.



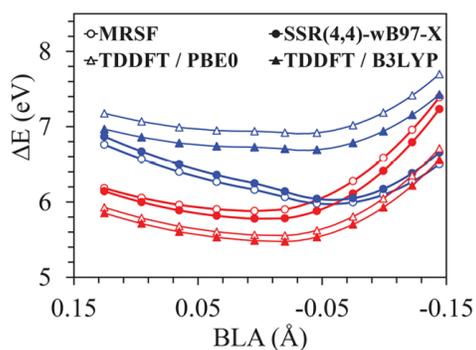
(a)



(b)



(c)

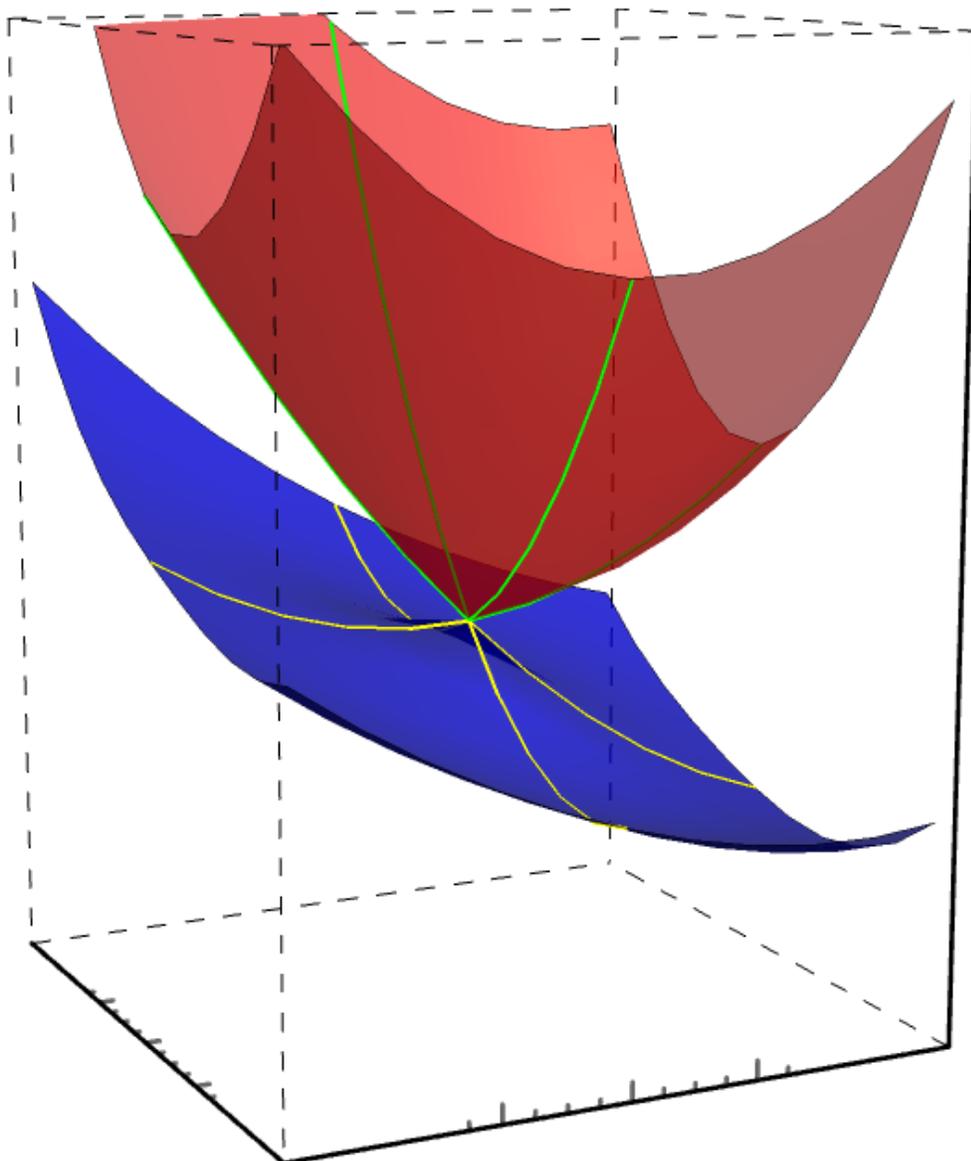


(d)

## CONICAL INTERSECTION OPTIMIZATION

```
$CONTRL SCFTYP=ROHF RUNTYP=CONICAL DFTTYP=BHLLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$CONICL IXROOT(1)=1,2 $END
$TDDFT NSTATE=3 IROOT=2 MULT=1 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
$STATPT NSTEP=150 $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
Thymine CI10 optimization
C1
C      6.0  -6.3694490194  3.3066040959  0.6701853348
C      6.0  -5.1383739860  2.5156794815  0.3450907496
C      6.0  -5.2128492716  1.0973250281  0.5420397967
C      6.0  -7.4147564086  1.5675884803 -0.5727641378
C      6.0  -3.9774834523  3.2436330141 -0.2019998520
N      7.0  -6.1240564597  1.0691991511 -0.6096436795
N      7.0  -7.5125831049  2.6034163480  0.3503600316
O      8.0  -8.3186768725  1.1989581009 -1.2681645527
O      8.0  -6.3549614051  4.3990557850  1.1739453250
H      1.0  -5.9018654502  0.5164108991 -1.4202506441
H      1.0  -8.4041193810  3.0517186642  0.4596019320
H      1.0  -3.3162568968  2.5573719510 -0.7163454661
H      1.0  -3.4282807909  3.6411117522  0.6581871107
H      1.0  -4.2491229046  4.0903025662 -0.8242064803
H      1.0  -5.6839825963  0.7285646824  1.4439495321
$END
```

- This is input for optimization of conical intersection (CI) used in J.Chem.Phys.Lett.(2021), 12, 4339.
- We optimized  $CI_{0/1}$  (ground state / first excited state) of thymine molecule with MRSF-TDDFT/BH&HLYP/6-31G\* method. (IXROOT(1)=1,2)

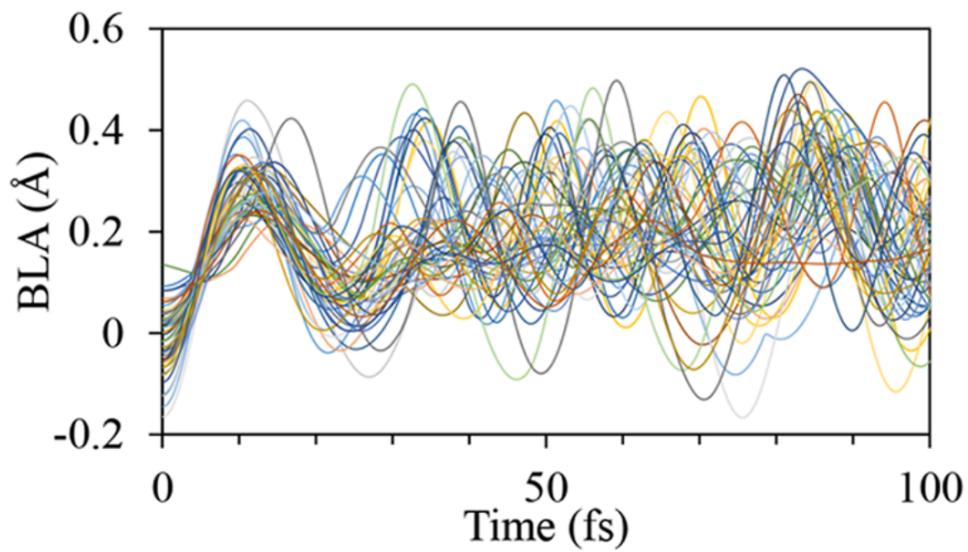
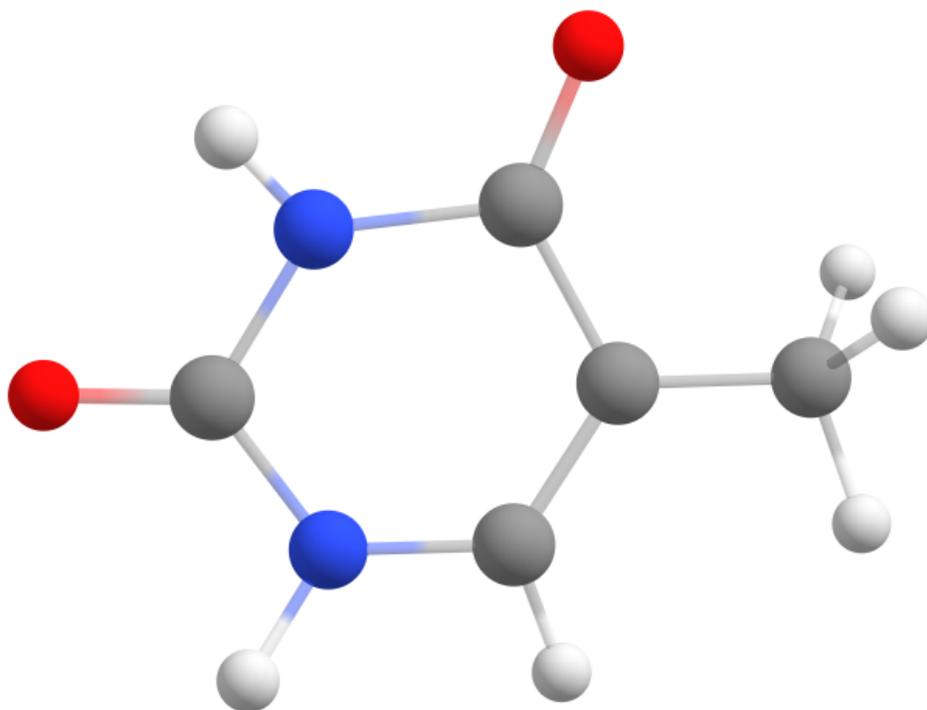


## NON-ADIABATIC MOLECULAR DYNAMICS (NAMD) SIMULATION

```
$CONTRL SCFTYP=ROHF RUNTYP=MD DFTTYP=BHLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 UNITS=BOHR $END
$TDDFT NSTATE=3 IROOT=3 MULT=1 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
Thymine NAMD simulation
C1
C 6.0 -11.8505071 6.5206747 -0.0300030
C 6.0 -9.5107081 5.0382484 0.0296647
C 6.0 -9.7175136 2.4272968 0.1217397
C 6.0 -14.2404466 2.2839256 -0.0366132
C 6.0 -7.0401877 6.4071584 -0.3197621
N 7.0 -11.9358493 1.2290798 0.2578589
N 7.0 -14.1275838 4.9013071 0.0834579
O 8.0 -16.1954657 1.2417354 -0.2958192
O 8.0 -12.1015559 8.7906497 0.1551644
H 1.0 -12.0940252 -0.7267109 0.7216047
H 1.0 -15.6585341 5.3990041 0.2736094
H 1.0 -5.4516499 4.9448097 -0.0114726
H 1.0 -6.9028034 7.5913218 1.4849747
H 1.0 -6.8090855 8.1187178 -2.0100609
H 1.0 -8.2001318 1.1305770 -0.1053907
$END
$MD READ=.F. MBT=.T. MBR=.T.
TTOTAL=0 DT=5e-16 NSTEPS=4000 MDINT=VVERLET
NVTNH=0 BATH(1)=300.0 RSTEMP=.F. JEVERY=1 KEVERY=1
THRSHE=10 NAMD=.T. $END
```

This is input for NAMD simulation used in J.Chem.Phys.Lett.(2021), 12, 4339.

We simulate thymine molecule with MRSF-TDDFT/BH&HLYP/6-31G\* method. We excited thymine to bright  $S_2$  state (IROOT=3) and propagate it until 2 ps with 0.5 fs timestep. (DT = 0.5 fs, NSTEPS=4000 → total 2 ps) \* Note that we block the hopping which has greater energy gap between electronic state than 10 kcal/mol (THRSHE=10) \* Note that one can geometry and velocity obtained from Wigner sampling with READ=.T. and TVELQM(1)= 3N values of velocity in atomic unit in MD group.



## SINGLET GROUND STATE OPTIMIZATION

```
$CONTRL SCFTYP=ROHF RUNTYP=OPTIMIZE DFTTYP=BHLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$TDDFT  NSTATE=5 IROOT=1 MULT=1 $END
$SCF    DIRSCF=.T. DIIS=.T. $END
$BASIS  GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
O-Benzynes
C1
H        1.0    2.5147391212  -0.0000000000  -0.1133955920
C        6.0    1.4425296188  -0.0000000000  -0.1119841121
C        6.0    0.6976318813  -0.0000000000   1.0655138930
H        1.0    1.2156127180  -0.0000000000   2.0075731367
C        6.0   -0.6976162877  -0.0000000000   1.0655155622
H        1.0   -1.2156221862   0.0000000000   2.0075631596
C        6.0   -1.4425256576   0.0000000000  -0.1119742488
H        1.0   -2.5147353494  -0.0000000000  -0.1134027300
C        6.0   -0.6195526066  -0.0000000000  -1.2085042416
C        6.0    0.6195387483   0.0000000000  -1.2085068269
$END
```

Above is the input for Singlet Ground state optimization used in J. Chem. Theory Comput. 2021, 17, 2, 848-859.

We simulated o-benzynes diradical with MRSF-TDDFT/BH&HLYP/cc-pVTZ method. One ground state singlet and four lowest singlet states are calculated (NSTATE = 5).



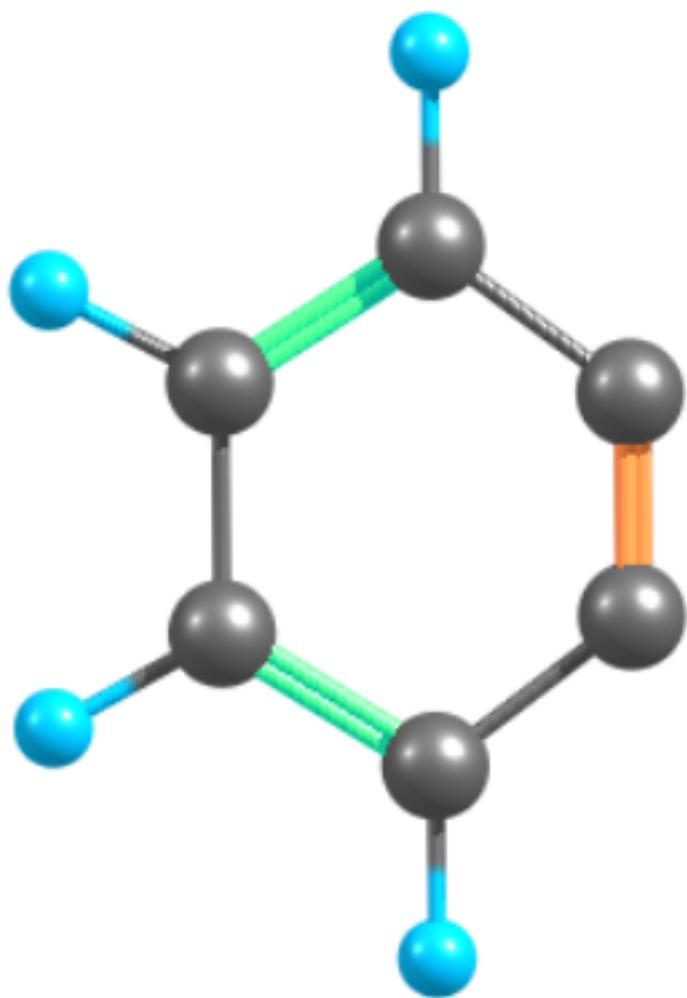
## TRIPLET GROUND STATE OPTIMIZATION

```
$CONTRL SCFTYP=ROHF RUNTYP=OPTIMIZE DFTTYP=BHLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$TDDFT  NSTATE=5 IROOT=1 MULT=3 $END
$SCF    DIRSCF=.T. DIIS=.T. $END
$BASIS  GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
O-Benzynes
C1
HYDROGEN  1.0  2.5147391212  -0.0000000000  -0.1133955920
CARBON    6.0  1.4425296188  -0.0000000000  -0.1119841121
CARBON    6.0  0.6976318813  -0.0000000000  1.0655138930
HYDROGEN  1.0  1.2156127180  -0.0000000000  2.0075731367
CARBON    6.0  -0.6976162877  -0.0000000000  1.0655155622
HYDROGEN  1.0  -1.2156221862  0.0000000000  2.0075631596
CARBON    6.0  -1.4425256576  0.0000000000  -0.1119742488
HYDROGEN  1.0  -2.5147353494  -0.0000000000  -0.1134027300
CARBON    6.0  -0.6195526066  -0.0000000000  -1.2085042416
CARBON    6.0  0.6195387483  0.0000000000  -1.2085068269
$END
```

Above is the input for Triplet Ground state optimization used in J. Chem. Theory Comput. 2021, 17, 2, 848-859.

We simulated o-benzyne diradical with MRSF-TDDFT/BH&HLYP/cc-pVTZ method. One Triplet Ground state and four triplet lowest states are calculated (NSTATE = 5).

The above 2 inputs, singlet ground state ( $S_0$ ) optimization and triplet ground state ( $T_0$ ) optimizations are used for calculating singlet-triplet (ST) gap of diradicals and diradicaloids.



**SINGLET GROUND STATE OPTIMIZATION IMPOSING C2V  
SYMMETRY**

```
$CONTRL SCFTYP=ROHF RUNTYP=OPTIMIZE DFTTYP=BHLLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$TDDFT NSTATE=5 IROOT=1 MULT=1 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
C4H6
CnV 2

CARBON 6.0 0.000000000 0.000000000 -0.036668884
CARBON 6.0 -1.203065556 0.000000000 -0.711856855
CARBON 6.0 0.000000000 0.000000000 1.456105051
HYDROGEN 1.0 -2.138811217 0.000000000 -0.189475446
HYDROGEN 1.0 -1.230113805 0.000000000 -1.784658245
HYDROGEN 1.0 0.000000000 0.921362405 1.999599679
$END
```

Above is the input for Singlet Ground state optimization by imposing C2V symmetry used in J. Chem. Theory Comput. 2021, 17, 2, 848-859. We Simulated Trimethylenemethane (TMM) with MRSF-TDDFT/BH&HLYP/cc-pVTZ method. One Ground state singlet and four lowest singlet states are calculated (NSTATE = 5). In \$TDDFT MULT=1 denote singlet.

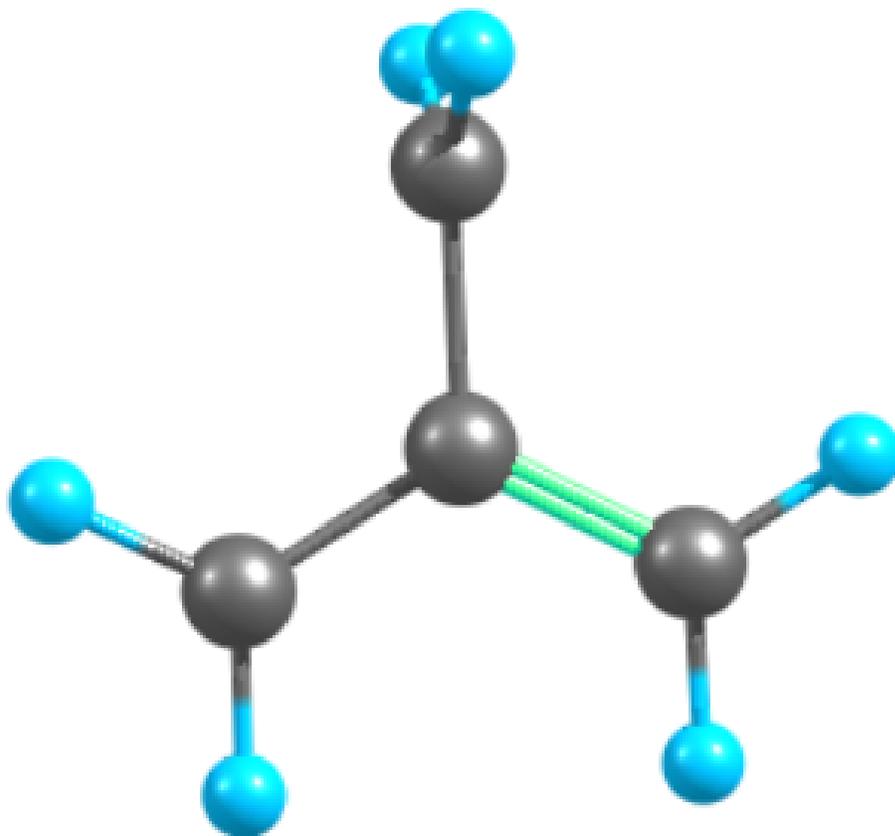


TRIPLET GROUND STATE OPTIMIZATION IMPOSING C2V  
SYMMETRY

```
$CONTRL SCFTYP=ROHF RUNTYP=OPTIMIZE DFTTYP=BHLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$TDDFT NSTATE=5 IROOT=1 MULT=1 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=CCT $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
C4H6
CnV 2

CARBON 6.0 0.000000000 0.000000000 -0.036668884
CARBON 6.0 -1.203065556 0.000000000 -0.711856855
CARBON 6.0 0.000000000 0.000000000 1.456105051
HYDROGEN 1.0 -2.138811217 0.000000000 -0.189475446
HYDROGEN 1.0 -1.230113805 0.000000000 -1.784658245
HYDROGEN 1.0 0.000000000 0.921362405 1.999599679
$END
```

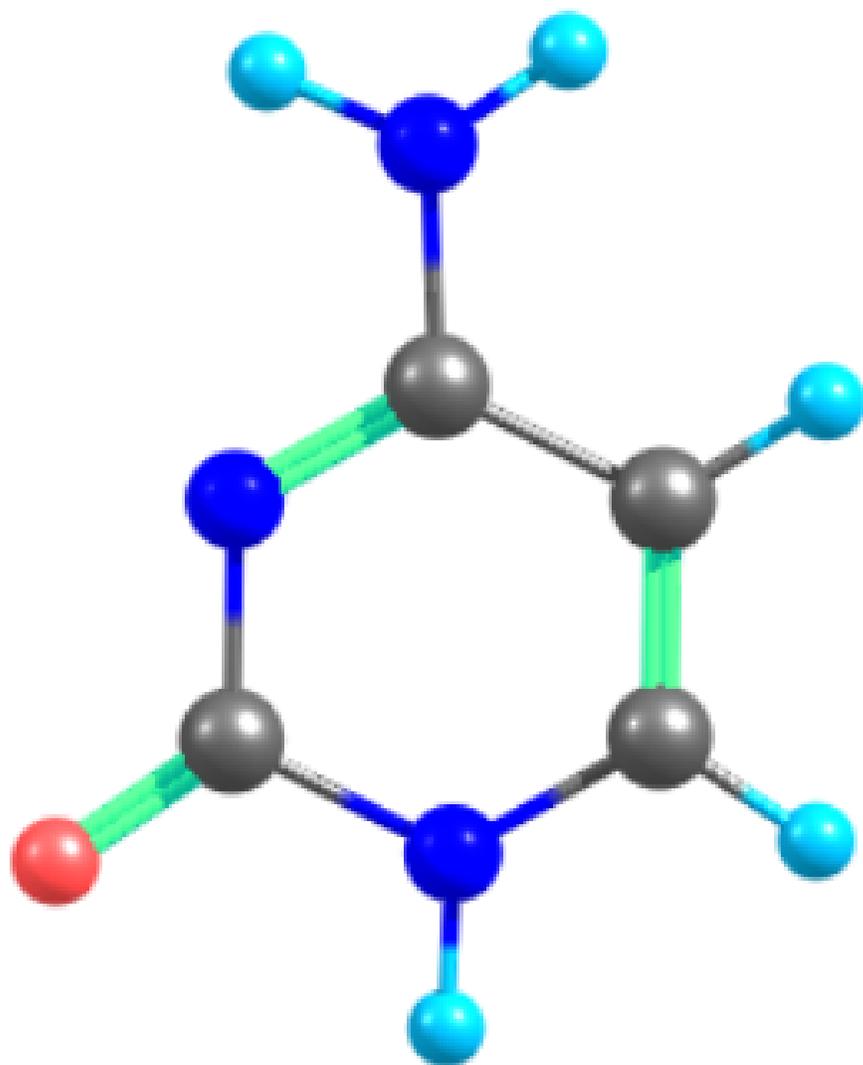
Above is the input for triplet Ground state optimization by imposing C2V symmetry used in J. Chem. Theory Comput. 2021, 17, 2, 848-859. Simulation is performed for Trimethylenemethane (TMM) with MRSF-TDDFT/BH&HLYP/cc-pVTZ method. One triplet Ground state and four lowest triplet states are calculated (NSTATE = 5). In \$TDDFT MULT=3 denote triplet.



## GROUND STATE OPTIMIZATION

```
$CONTRL SCFTYP=ROHF RUNTYP=OPTIMIZE DFTTYP=BHLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$TDDFT NSTATE=3 IROOT=1 MULT=1 $END
$SCF DIRSCF=.T. DIIS=.T. $END
$BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
Cytosine
C1
NITROGEN 7.0 -0.0086829468 0.0149277828 0.0660559171
NITROGEN 7.0 2.3458590473 -0.0091621139 0.0642813499
NITROGEN 7.0 -1.1420053387 -0.0086754893 2.0332447816
CARBON 6.0 1.1400383399 -0.0038705258 -0.6655768360
CARBON 6.0 2.3947342125 -0.0046455567 1.4100335961
CARBON 6.0 1.2551073430 0.0066010052 2.1253976227
CARBON 6.0 0.0395322993 0.0143876096 1.3718087669
OXYGEN 8.0 1.2032360931 -0.0103502365 -1.8709106973
HYDROGEN 1.0 3.1791336241 -0.0190957823 -0.4919314370
HYDROGEN 1.0 3.3724732459 -0.0122330275 1.8596591445
HYDROGEN 1.0 1.2618461171 -0.0011092818 3.1994280444
HYDROGEN 1.0 -1.1825535621 0.2440758431 2.9990286956
HYDROGEN 1.0 -1.9622049931 0.1464934116 1.4798381607
$END
```

Above is the input for Ground state optimization of Cytosine nucleobase. Simulation is performed for Cytosine Nucleobase using MRSF-TDDFT/BH&HLYP/6-31G(d) method.



## THREE STATE CONICAL INTERSECTION (TCI) OPTIMIZATION

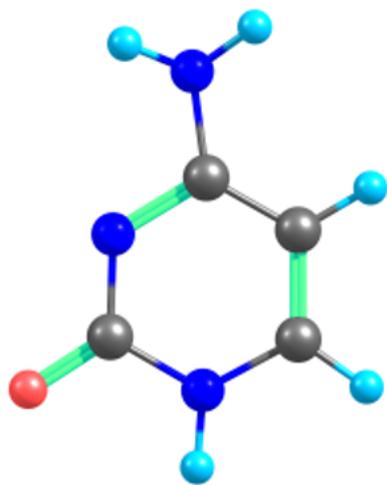
```

$CONTRL SCFTYP=ROHF RUNTYP=CONICAL DFTTYP=BHHLYP ICHARG=0
TDDFT=MRSF MAXIT=200 MULT=3 $END
$TDDFT  NSTATE=4 IROOT=4 MULT=1 TAMMD=.t. $END
$CONICL OPTTYP=PENALTT IROOT(1)=2,3,4 DEBUG=.t. SIGMA=1 $END
$SCF    DIRSCF=.T. DIIS=.T. $END
$BASIS  GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
$STATPT NSTEP=1000 hess=read opttol=1e-10 $END
$SYSTEM TIMLIM=999999100 MWORDS=500 $END
$DATA
Cytosine
C1
NITROGEN  7.0  -0.0086829468  0.0149277828  0.0660559171
NITROGEN  7.0  2.3458590473  -0.0091621139  0.0642813499
NITROGEN  7.0  -1.1420053387  -0.0086754893  2.0332447816
CARBON    6.0  1.1400383399  -0.0038705258  -0.6655768360
CARBON    6.0  2.3947342125  -0.0046455567  1.4100335961
CARBON    6.0  1.2551073430  0.0066010052  2.1253976227
CARBON    6.0  0.0395322993  0.0143876096  1.3718087669
OXYGEN    8.0  1.2032360931  -0.0103502365  -1.8709106973
HYDROGEN  1.0  3.1791336241  -0.0190957823  -0.4919314370
HYDROGEN  1.0  3.3724732459  -0.0122330275  1.8596591445
HYDROGEN  1.0  1.2618461171  -0.0011092818  3.1994280444
HYDROGEN  1.0  -1.1825535621  0.2440758431  2.9990286956
HYDROGEN  1.0  -1.9622049931  0.1464934116  1.4798381607
$END
$HESS
...
$END

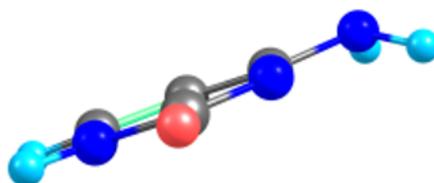
```

Above is the input for Three state Conical Intersection optimization of Cytosine nucleobase. Simulation is performed for Cytosine Nucleobase three state conical intersection using MRSF-TDDFT/BH&HLYP/6-31G(d) method. IROOT(1)=2, 3, 4 is used to locate conical intersection between  $S_1$ ,  $S_2$  and  $S_3$  excited states.

Front view



Side view



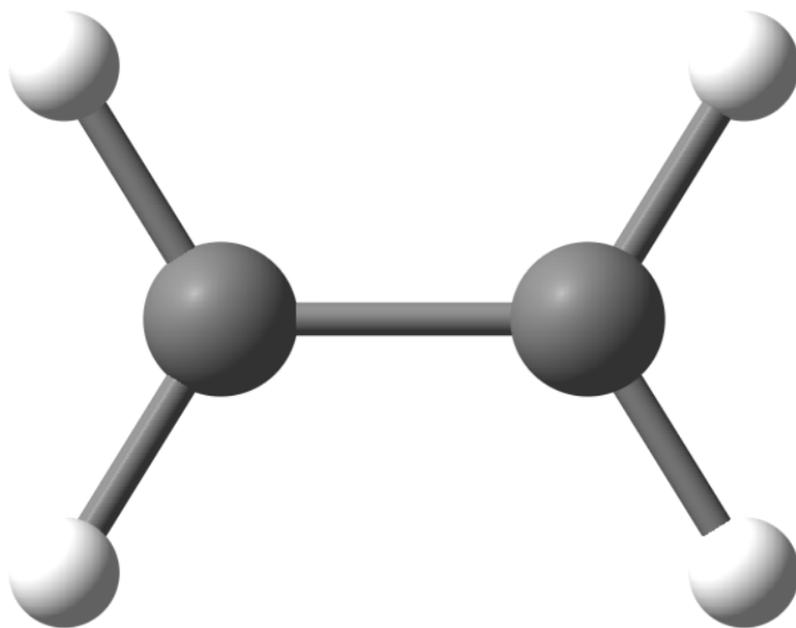
**MRSF-TDDFT EXAMPLES USING ETHYLENE**



## ETHYLENE S<sub>0</sub> OPTIMIZATION

```
$CONTRL
  SCFTYP=ROHF RUNTYP=OPTIMIZE DFTTYP=BHHLYP ICHARG=0
  TDDFT=MRSF MAXIT=200 MULT=3
$END
$TDDFT
  NSTATE=5 IROOT=1 MULT=1
$END
$SCF
  DIRSCF=.T. DIIS=.T. SOSCF=.F.
$END
$BASIS
  GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1
$END
$STATPT NSTEP=250
$END
$SYSTEM
  TIMLIM=999999 MWORDS=500
$END
$DATA
C2H4 S0 optimization
C1
C   6.0   0.180596175   -0.343800720   -0.175326809
C   6.0   1.465647227   -0.102026293   0.002214204
H   1.0   1.907395230   -0.454760198   0.920551799
H   1.0   -0.127316256   0.024318928   -1.141030591
H   1.0   -0.149587005   -1.321665903   0.137338066
H   1.0   1.822162163   0.910834885   -0.096714838
$END
```

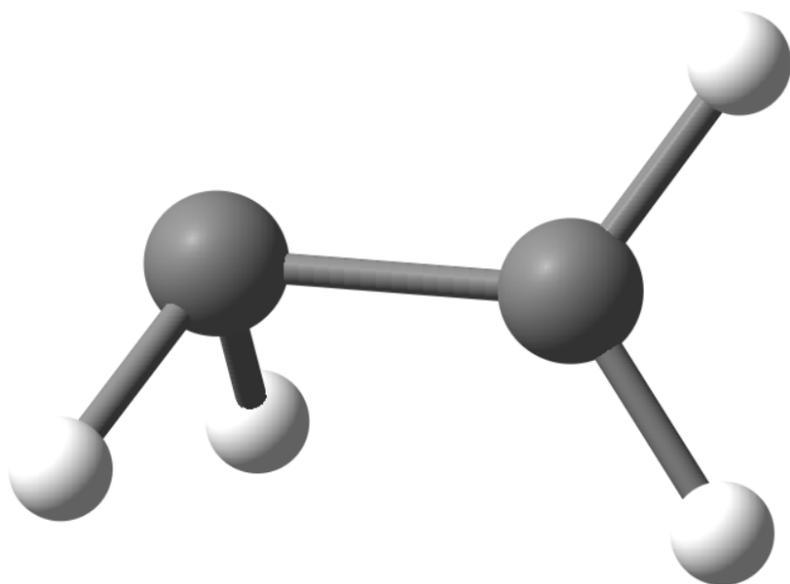
This is input for the ground state optimization of Ethylene with MRSF-TD-DFT/BHH&LYP/6-31G\*\*. Final geometry is shown on the figure below.



## ETHYLENE TW-PYR $S_1/S_0$ CONICAL INTERSECRION (CI)

```
$CONTRL
  SCFTYP=ROHF RUNTYP=CONICAL DFTTYP=BHHLYP ICHARG=0
  TDDFT=MRSF MAXIT=200 MULT=3
$END
$CONICL IXROOT(1)=1,2 $END
$TDDFT
  NSTATE=5 MULT=1
$END
$$SCF
  DIRSCF=.T. DIIS=.T. SOSCF=.F.
$END
$BASIS
  GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1
$END
$STATPT NSTEP=250
$END
$SYSTEM
  TIMLIM=999999 MWORDS=500
$END
$DATA
C2H4 S1/S0 CI
C1
C   6.0   -0.0017163764   0.0292739731  -0.0372180054
C   6.0    1.3786130141   0.0072684726  -0.0080577710
H   1.0    2.0129445494  -0.8878312901   0.0204802061
H   1.0    0.0551578071  -0.6605132877  -0.9497996163
H   1.0   -0.4811626193  -0.7365319361   0.5845879002
H   1.0    1.9748256249   0.9104670683  -0.1296487136
$END
```

This is input for the tw-pyr  $S_1/S_0$  CI optimization of Ethylene (in IXROOT, 1 corresponds to  $S_0$ , and 2 corresponds to  $S_1$  state) with MRSF-TD-DFT/BHH&LYP/6-31G\*\*. The CIs geometry is shown on the figure below.



ETHYLENE S<sub>0</sub> MINIMUM VIBRATIONS

```

$CONTRL
  SCFTYP=ROHF RUNTYP=HESSIAN DFTTYP=BHHLYP ICHARG=0
  TDDFT=MRSF MAXIT=200 MULT=3
$END
$TDDFT
  NSTATE=5 IROOT=1 MULT=1
$END
$$SCF
  DIRSCF=.T. DIIS=.T. SOSCF=.F.
$END
$BASIS
  GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1
$END
$STATPT NSTEP=250
$END
$SYSTEM
  TIMLIM=999999 MWORDS=500
$END
$DATA
C2H4 S0 vibrations
C1
CARBON      6.0   0.2737725809  -0.4418423602  -0.2974162841
CARBON      6.0   1.4258439056   0.0128270303   0.1797313338
HYDROGEN    1.0   1.9439319637  -0.4884167117   0.9807203656
HYDROGEN    1.0  -0.2442892276   0.0593372586  -1.0984148091
HYDROGEN    1.0  -0.1906628974  -1.3305950516   0.0969883247
HYDROGEN    1.0   1.8903012088   0.9015905335  -0.2145771000
$END

```

This is input for the Ethylene S<sub>0</sub> frequency calculations with MRSF-TD-DFT/BHH&LYP/6-31G\*\*

MODE	FREQ(CM**-1)	SYMMETRY	RED. MASS	IR INTENS.
1	18.554	A	1.008760	0.000000
2	6.650	A	2.121082	0.000000
3	6.102	A	2.180655	0.000000
4	0.243	A	4.668636	0.000001
5	0.065	A	4.672078	0.000000
6	0.029	A	4.672023	0.000000
7	860.005	A	1.042244	0.012138
8	1003.751	A	1.516381	0.000004

(continues on next page)

(continued from previous page)

9	1016.337	A	1.160935	2.255860
10	1116.259	A	1.007826	0.000011
11	1287.643	A	1.517699	0.000000
12	1420.859	A	1.271199	0.000000
13	1531.578	A	1.112140	0.186539
14	1752.366	A	2.858180	0.000000
15	3245.587	A	1.047561	0.299122
16	3262.542	A	1.073808	0.000004
17	3327.977	A	1.116184	0.000003
18	3353.094	A	1.118196	0.595658

No imaginary frequencies was obtained

**ETHYLENE {CI ( $S_1/S_0$ ) -  $S_0^{\text{MIN}}$ } MEP USING GEODESIC  
INTERPOLATION**

```
$CONTRL
  SCFTYP=ROHF RUNTYP=ENERGY DFTTYP=BHLLYP ICHARG=0
  TDDFT=MRSF MAXIT=200 MULT=3
$END
$TDDFT
  NSTATE=5 IROOT=1 MULT=1
$END
$SCF
  DIRSCF=.T. DIIS=.T.
$END
$BASIS
  GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1
$END
$END
$SYSTEM
  TIMLIM=999999 MWORDS=500
$END
$DATA
C2H4 CI(tw-pyr) - S0min MEP
c1
CARBON      6.0    -0.824885547    0.252330946    0.049317029
CARBON      6.0     0.555605058    0.230100380    0.078834443
HYDROGEN    1.0     1.189859013   -0.665149629    0.106935364
HYDROGEN    1.0    -0.767680817   -0.437452028   -0.863174166
HYDROGEN    1.0    -1.304227831   -0.513413659    0.671248498
HYDROGEN    1.0     1.151330123    1.133583991   -0.043161168
$END
```

Here, interpolation between two points on PES was done. For each point (frame) obtained as a result of interpolation, we carried out single point calculations in order to plot the MEP corresponding to the tw-pyr CI -  $S_0^{\text{min}}$  relaxation.

C2H4 tw-pyr Cl to S0 min MEP

