Tutorial on Running DFTB via DFTB+ and Gaussian 09

Benjamin Rudshteyn, Kenneth A. Jung, Wendu Ding, Christian F.A. Negre, and Victor S. Batista*

Department of Chemistry, Yale University, P.O. Box 208107, New Haven, CT 06520-8107, and

Energy Sciences Institute, Yale University, P.O. Box 27394, West Haven, CT 06516-7394

E-mail:
victor.batista@yale.edu

Initial Version: January 5, 2015
New Version: January 3, 2017

Current "ab initio" methods were limited to very inaccurate calculations for very small molecules.
M.J.S. Dewar, *A Semiempirical Life*, 1992
via
Lewars, Errol G. *Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics*. Springer Science & Business Media, 2010.

Introduction 2

Chapter 1. Parameters 2

Chapter 2. DFTB+ 3

Chapter 3. Gaussian 09 7

Introduction

This is a tutorial for running DFTB with external parameters (chapter 1) via the DFTB+ (chapter 2) and the Gaussian 09 (chapter 3) programs. We focus on using the supercomputer Omega at Yale University that utilizes Torque/PBS, but in principle, the program can be run on any Linux supercomputer. Such clusters should have gfortran loaded by default, but otherwise, you may need it to compile the .f90 files given in Table 1.

Then you would run (for an example .f90 file called example.f90)

gfortran example.f90 -o example.o.

where the executable ends up being example.o in the example directory.

Chapter 1. Parameters

With either program, Slater-Koster parameter files (.skf) prepared by researchers specialized in the field of DFTB are required. They may be downloaded at http://www.dftb.org/parameters/download/ or requested from the relevant corresponding author.

For the website, the following group credentials are required:

Username:

victor.batista@yale.edu

Password:

s-n!Q*tB

Please note as the website says:

"Please note, that unless explicitly indicated, the sets are usually not compatible to each other. Therefore, the parameter files from two sets should not be mixed (although your program may not even notice the incompatibility of the files used)."

The parameter files usually contain README files that indicate their compatibility with other sets.

Once you go the website, click on the parameter set you want and click "Download" to get the .tar.xz file that can be opened with the **tar -xzvf** command. Don't forget to cite the relevant papers in your own work!

The parameter files for H_2O would be as follows:

1. O-O.skf
2. O-H.skf
3. H-O.skf
4. H-H.skf

The so-called *homonuclear* parameter files describe the interactions of an atom with itself and identical atoms while the *heteronuclear* parameter files describe the interactions of an atom with other nonidentical atoms. Note that both O-H.skf and H-O.skf are required.

For the sake of this tutorial, we will use the mio-1-1 set at http://www.dftb.org/parameters/download/mio/mio-1-1/. Please transfer the .tar file to your

Desktop and execute the following commands.

- cd ~/Desktop
- tar -xvzf mio-1-1.tar.gz

Chapter 2. DFTB+

Official Reference Manual: http://www.dftb-plus.info/fileadmin/DFTB-Plus/public/dftb/current/manual.pdf
Additional Documentation: http://www.dftb-plus.info/documentation/

Installing DFTB+

- Make sure, that cookies and JavaScript are enabled in your browser.
- Go to http://www.dftb-plus.info/download/.
- Use these login credentials for the group:
  Username: victor.batista@yale.edu
  Password: !ROltgW9
- Download the option under "**DFTB+ 1.3** (current stable release)"
- Click download again.
- Select this option: **executable x86_64-linux**
  which is an "Executable for x86_64 (64 bit) architecture with Linux operating system." You can also download the source code (see http://www.dftb-plus.info/fileadmin/DFTB-Plus/public/recipes/html/installation/index.html for how to compile) and precalculated test examples.
- Copy the download to your home directory on a Linux cluster called "DFTB_version_1_3".
- On the cluster, use the following command to untar the download:
  - tar -xvzf dftbplus-1.3.0.x86_64-linux.tar.xz

Running DFTB+

Download the associated DFTBscriptfiles_noexecutable.zip file and open it with the **unzip** command or double click on it in the Finder window (if you're on a Mac). The directory DFTBscriptfiles_noexecutable contains scripts used to prepare input files and submit jobs. Their use is described in Table 1. Copy it to the "DFTB_version_1_3" directory. The folder should contain the folders/files:

- dftbplus-1.3.0.x86_64-linux
- mio-1-1
- DFTBscriptfiles_noexecutable

Table 1. Files contained in DFTBscriptfiles_noexecutable in no particular order relevant for this portion of the tutorial.

Filename	Language	Usage
optimizedftb.sh	Bash	This file will perform the preparation and submit the calculation.
xyztogen-prog.f90	Fortran 90	Will convert the input *.xyz* file into a *.gen* file.
dftb_in-make-opt.f90	Fortran 90	Will prepare the input *dftb_in.hsd* file for DFTB+.
torque.sh	Bash/Torque	Will be used to generate a submission script for a job on Omega-type systems.
lsf.sh	Bash/LSF	Will be used to generate a submission script for a job on Grace-type systems.
converge.sh	Bash	Monitor the convergence of the job
dftb_in.hsd	DFTB input	Template for generating the DFTB input file.
ethane.xyz	XYZ file	Example file that comes with the tutorial.

Follow these commands:

- **cd ~/DFTB_version_1_3/DFTBscriptfiles_noexecutable**
- **gfortran xyztogen-prog.f90 -o xyztogen-prog.o**
- **gfortran dftb_in-make-opt_all.f90 -o dftb_in-make-opt_all.o**
- **cd ~/scratch/**
- **mkdir testDFTB**
- **cd testDFTB**
- **cp ~/DFTB_version_1_3/DFTBscriptfiles_noexecutable/ethane.xyz .**

Now you are ready to submit a calculation with the example ethane molecule. For your convenience, we have written a bash script to take care of the tedious tasks involved in setting the calculation up. In order to run such a calculation, simply execute the following command (the script is well commented if you are familiar with bash, so check it out).

~/DFTB_version_1_3/DFTBscriptfiles_noexecutable/optimizedftb.sh ethane.xyz molecular

The directory should now contain the following files (br287 would be replaced with your netid).

- torque_ethane.sh
- lsf_ethane.sh
- H-H.skf -> /home/fas/batista/br287/DFTB_version_1_3/mio-1-1/H-H.skf
- H-C.skf -> /home/fas/batista/br287/DFTB_version_1_3/mio-1-1/H-C.skf
- dftb_in.hsd

- coords.xyz
- coords.gen
- C-H.skf -> /home/fas/batista/br287/DFTB_version_1_3/mio-1-1/C-H.skf
- C-C.skf -> /home/fas/batista/br287/DFTB_version_1_3/mio-1-1/C-C.skf

torque_ethane.sh is the submission script to submit the calculation on PBS systems like Omega and lsf_ethane.sh is the submission script to submit the calculation on LSF systems like Grace.
The .skf files contain the parameters for the unique elements in the system.
The .hsd and .gen files are the inputs needed for DFTB.
coords.xyz is a copy of the first argument .xyz ($1) file used in the above command.

The second argument ($2) used in the command above is for a molecular (0D) calculation. Omit it for a periodic calculation (you need still need to edit the lattice vectors in the .gen file and the k-points in the .hsd file).

Afterwards, just submit the calculation with Omega

- qsub torque_ethane.sh

or with Grace

- bsub < lsf_ethane.sh

It should take a minute. Then you will have the following new files in your directory with a different jobid in the Yale files).

- timing.txt
- results.tag
- geom.out.xyz
- geom.out.gen
- ethane.out
- ethane.5350620.rocks.omega.hpc.yale.internal.out
- ethane.5350620.rocks.omega.hpc.yale.internal.err
- detailed.xml
- detailed.out
- charges.bin
- band.out
- autotest.tag

You can monitor the convergence of the optimization with a script as follows
~/DFTB_version_1_3/DFTBscriptfiles_noexecutable/converge.sh ethane.out
To exit the view of the numbers, which should be decreasing, simply hit “q”.

At the bottom of the successful output of ethane.out you will see the following. The first line is the total energy and the last line indicates a successful geometry optimization. The **geom.out.xyz** contains the optimized geometry. The .err file would indicate if there was a problem. The timing.txt file contains the information about the start and end time of the calculation (also the

total time is in the .err file).

Total Energy: -5.7015287391 H
Total Mermin free energy: -5.7015287391 H
Maximal force component: 0.926539E-06
>> Charges saved for restart in charges.bin

Geometry converged

Please refer to the DFTB+ manual for other options. If you need to use a different parameter set, modify the following line (18) in the optimizedftb.sh script appropriately:
skffolder=$HOME/DFTB_version_1_3/mio-1-1

Additional Notes

One thing that needs attention in dftb_in.hsd is this line:
MovedAtoms=1:-1

This selects which atoms in the system will be moved during optimization. If the number is negative, it means start counting from the last atom in the system: -1 means the last atom, -2 means the penultimate atom, etc. So, if certain atoms need to be constrained to their initial position, they need to be excluded. For example, if the first 10 atoms need to be constrained, then the line becomes:
MovedAtoms=11:-1

Without the "molecular" argument for the bash script, xyztogen-prog will make gen format coordinate files for periodic systems. Using the following gen file as an example:

```
3
S O
H
1  1  0.000 0.000 0.000
2  2  0.759 0.000 0.504
3  2  0.759 0.000 -0.504
0.0 0.0 0.0
10.48992200 0.00000000 0.00000000
0.00000000 15.18400000 0.00000000
0.00000000 0.00000000 20.00000000
```

It starts with the number of atoms in the system. The *S* in the first line indicates it is a periodic system (supercell) in Cartesian coordinates. If non-periodic system is in interest, then a letter *C* should be used. The second line lists the all the atom types in the system, and their order provide them assignments, which are used in the coordinate section. From the 3rd line to the 5th line is the coordinate section. If it is periodic system, then starting from the 6th line is the lattice vector information. The 6th line is the coordinate origin in Å. The 7th to the 9th line is the lattice vectors in the orthogonal basis.

The DFTB executable is also available to be directly available in your path if you load the appropriate module on either Grace or Omega:
module load Apps/DFTB/1.3

Chapter 3. Gaussian 09

Official Reference Manual: http://www.gaussian.com/g_tech/g_ur/l_keywords09.htm
DFTB information: http://www.gaussian.com/g_tech/g_ur/k_dftb.htm
Useful Presentation: http://www.dftb.org/fileadmin/DFTB/public/presentations/ACS-2006-SanFrancisco/presentations/342_Zheng.pdf (in the Dropbox too)
Thanks to the Master: Lufeng Zou, Ph.D. at help@gaussian.com.

Gaussian can run DFTB either with so-called analytical parameters (no external files needed) for a limited number of common elements (H, C, O N, some metals) as described on http://www.gaussian.com/g_tech/g_ur/k_dftb.htm or with external parameters from the DFTB+ website discussed above.

Analytical Parameters

Say you wanted to run the calculation on alanine on Omega. To use DFTB-A (the analytical form), you would use an input file such as that on the previous page. Note that analytical frequencies (simply the keyword **freq**) are allowed here. Essentially everything else is just like in the usual DFT calculations you might do. Note that if you have `geom=connectivity` information at the end of the file, then the `@GAUSS_EXEDIR:dftba.prm` should come after (with a blank line in between). You can see my original file at **alanine_dftba.com** in the Dropbox with the parameter set dftb_3ob-3-1.

```
%chk=alanine_dftba.chk
%mem=59GB
%nprocshared=12
#p opt freq dftba

www.gaussian.com/g_tech/g_ur/k_dftb.htm

0 1
 N                  -3.28848772   -0.46599040    1.09104316
 H                  -2.55349183   -0.63894718    1.74668559
 C                  -4.00268244    0.76221574    1.46831704
 H                  -4.78912804    0.94727950    0.76677964
 C                  -4.60062616    0.59597288    2.87772352
 C                  -3.02104903    1.94879731    1.46383921
 H                  -3.81418056    0.41090913    3.57926091
 H                  -5.12048218    1.48997327    3.15233784
 H                  -5.28267015   -0.22847015    2.88083474
 O                  -1.76378278    1.75416337    1.73730016
 O                  -3.44883506    3.14609267    1.18659449
 H                  -3.92591200   -1.23649791    1.09395084
 H                  -2.69923841    3.73038639    1.05129592

@GAUSS_EXEDIR:dftba.prm
```

See if you got the same result I did:

SCF Done: E(RDFT-SCTB) = -16.8187991095 A.U. after 7 cycles

Prepare the External Parameters

Say you wanted to run the same calculation on alanine on Grace with the mio-1-1 parameters. They have to be modified. Therefore follow these commands:

- **cd ~/DFTB_version_1_3/mio-1-1**
- **mkdir atomicnumbers**
- **cp *skf atomicnumbers**
- **cd atomicnumbers**
- **~/DFTB_version_1_3/DFTBscriptfiles_noexecutable/appendtoskffile_NEW.sh**

The format of the .skf file is described in
http://www.dftb.org/fileadmin/DFTB/public/misc/slakoformat.pdf (also in the Dropbox)
The bash script modifies the .skf files according to the following rules for the first line of the .skf file (it must contain only 4 numbers):

a. "gridDist being the distance between the grid points of the integral table" – this is the original first number in the .skf file
b. "nGridPoints being the number of points in the table" - "actual number of lines" between the lines between the header and the next instance of "Spline". The script calculate this number so you do not have to. It is unclear what this number originally is in the .skf
c. The atomic number of the first element e.g. for H-C.skf, this would be 1. For S-S.skf, this would be 16. This information is derived from the neighboring file periodictable_forsearch.txt.
d. The atomic number of the second element e.g. for H-C.skf, this would be 6. For S-S.skf, this would be 16. This information is derived from the neighboring file periodictable_forsearch.txt.

I used the following details from Lufeng to write the script:
"The "actual number of lines" are the lines between the header and the "Spline". For example, in "C-H.skf", there are 602 lines before the "Spline", minus the first headline, it'll be 602-1=601 lines, instead of the "600" given in the skf file. So you'll change "**600**" to "**601**".

For parameter files of the same atom, there are two headlines. For example, in "H-H.skf", there are 553 lines before "Spline", so 553-2=551 lines. In addition, there should be only four parameters in total, so please remove the additional "2", i.e., replace the line:

2.000000000000E-02, **550**, **2** 1 1

with

2.000000000000E-02, **551**, 1 1

Then you should be able to run the DFTB job without problems."

Note: You can modify the script to delete the resulting .backup files by adding the command **rm *.backup**. Alternatively, you can undo the changes it makes with the following loop:
for f in *backup ; do mv $f $(basename $f .backup) ; done

Preparing the Input File

Now you have to tell your Gaussian input file to read these parameters. The file should be organized in the following way (don't use `dftb=ReadParameters` instead of `dftb=Read`. We have not tested `dftb=chkparameters`, but it may allow you to omit the list of parameters in the second calculation. Note here, freq=numer must be used, because of some technical details described at the bottom of http://www.gaussian.com/g_tech/g_ur/k_dftb.htm. You cannot use `opt freq` and skip the second calculation entirely, because the freq section needs to be fed the parameters. The `/N` in the parameters list indicates that the parameters will not be printed into the output file (they can be omitted). You can see my original file at **alanine_dftb_lufeng.com** in the Dropbox with the parameter set dftb_3ob-3-1.

```
%chk=alanine_dftb_mio.chk
%mem=59GB
%nprocshared=12
#p opt dftb=read

www.gaussian.com/g_tech/g_ur/k_dftb.htm

0 1
 N                  -3.28848772   -0.46599040    1.09104316
 H                  -2.55349183   -0.63894718    1.74668559
 C                  -4.00268244    0.76221574    1.46831704
 H                  -4.78912804    0.94727950    0.76677964
 C                  -4.60062616    0.59597288    2.87772352
 C                  -3.02104903    1.94879731    1.46383921
 H                  -3.81418056    0.41090913    3.57926091
 H                  -5.12048218    1.48997327    3.15233784
 H                  -5.28267015   -0.22847015    2.88083474
 O                  -1.76378278    1.75416337    1.73730016
 O                  -3.44883506    3.14609267    1.18659449
 H                  -3.92591200   -1.23649791    1.09395084
 H                  -2.69923841    3.73038639    1.05129592

@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-O.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-O.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-O.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-O.skf /N

--Link1--
```

```
%chk=alanine_dftb_mio.chk
%mem=59GB
%nprocshared=12
#p guess=read geom=check freq=numer dftb=read

www.gaussian.com/g_tech/g_ur/k_dftb.htm

0 1

@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/N-O.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/H-O.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/C-O.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-N.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-H.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-C.skf /N
@/home/fas/batista/br287/DFTB_version_1_3/mio-1-1/atomicnumbers/O-O.skf /N
```

You should get an energy very close to the analytical version.

Note from Lufeng if you are using super-new parameters: "You're using a newer parameter file that treats the 4d electrons explicitly for Cd (with filled 4d10 shell). However, the DFTB code in Gaussian expects the older convention where the filled d shell is not included, thus giving the error.

Therefore, this new parameter file cannot be used here. Is it designed for DFTB**3**? Note DFTB3 and DFTB are two different methods and Gaussian only provide DFTB method. However, it is possible to run the calculation if you can find an order set of DFTB parameter files, that includes Cd without explicit 4d electrons. You might have to change the parameter files for other elements so that these parameter files are compatible."

Quickly Prepare the Parameters List
Use such a loop (replace the elements C O N H with the ones you care about:

```
for f in C O N H
do
      for g in C O N H
      do
            skf=$f-$g.skf
            echo @$HOME/DFTB_version_1_3/mio-1-1/atomicnumbers/$skf /N
      done

done
```

The output from this command can be appended to your inputfile.

Can I Run ONIOM with DFT-B?
YES!!!!!!!!!! Use such a route line (if uff is the low layer):

```
#p opt oniom(dftb=read:uff) geom=connectivity
```