EASYPARM

User manual

Version 2.0 – November 2024

Abdelazim M. A. Abdelgawwad

abdelazim.abdelgawwad@uv.es

Institut de Ciència Molecular, Universitat de València, P.O. Box 22085, València 46071, Spain

Table of Contents

1. Introduction

easyPARM is a computational tool with two primary functions. First, it derives bond-stretching and bond-angle bending force field parameters for metal-containing molecular systems using the Seminario method.[1] This ensures accurate parameter generation tailored for the AMBER software suite [2], based on the Hessian matrix produced by frequency calculation in Gaussian [3] or Orca software.[4]

Second, easyPARM enables charge restraints on specific atoms using the Restrained Electrostatic Potential (REsP) fitting approach.[5] This helps achieve a more accurate electrostatic potential around the molecule while keeping charge distribution physically meaningful.

With these capabilities, easyPARM is an efficient tool for accurately modeling metal-containing systems, ensuring both structural and electrostatic precision.

1.1 Key Features

- Generates parameters for metal-containing systems using a combination of the general AMBER force field (GAFF) or the AMBER force field and the missing parameters for bond stretching and angle bending involving metals and linked atoms with Seminario method
- Capable of generating parameters for entire structures in non-metal systems using either GAFF or AMBER
- Outputs essential files (frcmod, pdb, mol2, and lib) for initiating simulations or preparing metal-containing systems

1.2 Citation

Please cite the following references if you use the easyPARM code:

- Abdelgawwad AMA, Francés-Monerris A. easyPARM: Automated, Versatile, and Reliable Force Field Parameters for Metal-Containing Molecules with Unique Labeling of Coordinating Atoms. ChemRxiv. 2024; doi:10.26434/chemrxiv-2024-f8wp4
- Github (https://github.com/Abdelazim-Abdelgawwad/easyPARM.git)

2. Installation

2.1 Prerequisites

To successfully install and run easyPARM, ensure that the following prerequisites are met:

- Python 3 or higher: easyPARM requires Python version 3 or above to function properly. You can download the latest version of Python from the official Python website.
- periodictable package: This package provides access to the periodic table of elements, which is essential for the metal-specific calculations in easyPARM. To install it, use the following command in your terminal or command prompt:

```
pip install periodictable
```

- scipy package: SciPy is a scientific computing library that easyPARM utilizes for various numerical calculations, including those related to force fields and molecular mechanics. To install SciPy, use the following command:

```
pip install scipy
```

Additional package requirements may arise during execution. If prompted, install them via pip or contact us for assistance.

Once the necessary dependencies are installed, you'll be ready to run easyPARM to generate the required force field parameters.

2.2 Running easyPARM

Choose one of the following methods:

1. **Run the script directly**: Navigate to the directory containing the easyPARM script and execute the following command.

```
./easyPARM.sh
```

2. **Set up an alias for easy access**: To run easyPARM from any location in your terminal, you can create an alias by adding the following line to your .bashrc file:

```
alias easyPARM='/full/path/to/easyPARM.sh'
```

3. Usage Guide

3.1 Required Input Files

1. **Hessian Matrix Files**: The Hessian matrix is required for subsequent calculations and can be provided in one of the following formats:
   a. **Checkpoint File (.chk):** Contains the Hessian matrix from a Gaussian calculation.
   b. **Formatted Checkpoint file (.fchk):** Contains the Hessian matrix from a Gaussian calculation.
   c. **Gaussian ouput (.log):** Contains the Hessian matrix from a single-point frequency calculation. Ensure that the keyword `iop(7/33=1)` is included in the input.
   d. **Orca output (.hess):** Contains the Hessian matrix from a single-point frequency calculation in Orca.
2. **Optimized Structure (XYZ Format):** Contains the optimized molecular geometry.
3. **Charge Output:** This file provides information on charge distribution (ESP charges) and can be in one of the following formats:
   a. Gaussian Charge Output (RESP charges)
   b. Orca Charge Output (CHELPG charges)

3.2 Step-by-Step Tutorial

1- AMBER Configuration

The code will prompt you to confirm whether AMBER is already loaded or if you need to specify its path. You will see the following prompt:

```
================================
```

```
 Amber Configuration Menu
==============================
Select your option:
1- Use currently loaded Amber
2- Specify Amber installation path
Enter your choice:
```

- If you can use the module load amber command, select option 1.
- If you prefer to manually provide the AMBER installation path, select option 2.

Option 1:

```
Enter your choice: 1
Amber is assumed to be already loaded. Skipping sourcing.
```

Option 2:

```
Enter your choice: 2
Please provide the path for Amber:
```

2- System Charge.

The code needs the total charge of the system

```
Please provide the total charge:
```

3- Providing the Optimized Structure for Seminario Method

To proceed with the Seminario method, you will need to provide the optimized XYZ geometry file. The code will prompt you to enter the name of this file:

```
Please provide the optimized XYZ geometry file:
```

4- Charge Method

The code supports two types of charge calculation methods. You can choose the method that best fits your needs and system requirements. During execution, the code will prompt you to make a selection from the following menu:

```
==================================
 Charge Method Selection Menu
==================================
Select the charge calculation method:
1- Gaussian (RESP charges)
2- Orca (CHELPG charges)
Enter your choice:
```

Option 1: Gaussian (RESP charges)

If you select this option, the code will proceed to generate the .mol2 file using RESP charges from the Gaussian output. (Refer to the example input in the Appendix for further details).

- **mol2 File Generation**

After specifying the total charge, the code will prompt you to choose the input format for generating the .mol2 file. You will be presented with the following options:

```
==============================
```

```
 Charge Output Selection Menu
================================
Please select the input format:
1- Gaussian Output (.log file)
2- Resp (.gesp file)
3- PDB Structure (.pdb file)
Enter your choice:
```

Next, you'll need to select the charge calculation method for the system. The recommended method is **RESP**. The menu will appear as:

```
================================
 Charge Method Selection Menu
================================
Please select the charge method (recommended: RESP):
1- RESP (resp)
2- Mulliken (mul)
3- ESP (esp)
4- AM1-BCC (bcc)
Enter your choice:
```

You will also be asked to specify the atom type for the system. The options are **GAFF** or **AMBER** or **GAFF2**:

```
================================
 Atom Type Selection Menu
================================
Please select the atom type:
1- Amber Force Field (AMBER)
2- General Amber Force Field (GAFF)
3- General Amber Force Field (GAFF2)
Enter your choice:
```

Once you have selected the atom type, the code will prompt you to provide the output file containing the charge information. Accepted file formats include .log and .resp. You will see the following prompt:

```
Please provide the charge output file (e.g., .log, .resp):
```

Option 2: Orca (CHELPG charges)

If you select this option, the code will proceed to generate the .mol2 file using **CHELPG** charges from the Orca output. (Refer to the example input in the Appendix for further details).

- **mol2 File Generation**

After specifying the total charge, the code will prompt you to choose the atom type for the system. The options are **GAFF** or **AMBER** or **GAFF2**:

```
==================================
 Atom Type Selection Menu
==================================
Please select the atom type:
1- Amber Force Field (AMBER)
```

2- General Amber Force Field (GAFF)
3- General Amber Force Field (GAFF2)
Enter your choice:

Once you have selected the atom type, the code will prompt you to provide the output file containing the charge information. Accepted file formats include .log and .out. You will see the following prompt:

Please provide the charge output file (e.g., .log, .out):

5- **Seminario Method Setup**

To run the Seminario method, you must provide one of the required input files. The system will guide you through selecting the appropriate format based on your input.

```
================================
 Input Format Selection Menu
================================
Please select the format you will provide:
1- Orca Output
2- Gaussian Output
3- Gaussian Checkpoint
4- Gaussian Formatted Checkpoint
Enter your choice:
```

1. Handling an Orca file:

If you select the Orca output option, two files are required to properly run the Seminario method:

- Orca Output File (.log): You will be prompted to provide the Orca output file. However, if you selected the CHELPG charge method, this file is not required.
- Orca Hessian File (.hess): You will also need to provide the Orca Hessian file.

Please provide the Orca output file (.log):
Please provide the Orca hessian file (.hess):

2. Handling a Gaussian Output

If you select the Gaussian output option, you will need to provide one file to proceed:

- Gaussian Output File (.log or .out): The system will prompt you for the Gaussian output file.

Please provide the Gaussian output file (.log or .out):

3. Handling a Checkpoint file:

If you choose to provide a Gaussian checkpoint (.chk) file, the system will check for Gaussian availability. You'll be asked whether Gaussian is already installed on your system or if you need to manually provide the path to Gaussian's formchk utility.

Select your option:
1- Gaussian is already loaded (formchk is available)
2- Provide the Gaussian path
Enter your choice (1 or 2):

4. Handling a Formatted Checkpoint file:

If you provide a formatted checkpoint (.fchk) file, the system will automatically proceed to the next step, with no additional input required.

6- Output Generation

After successfully completing all the preceding steps and ensuring that all required inputs are correctly provided, the software will generate the necessary parameter files for further use. Upon completion, you will receive a set of output files that may include:

```
==================================
                Output Files
==================================
mol2      : COMPLEX.mol2
Frcmod  : COMPLEX.frcmod
pdb       : COMPLEX.pdb
lib       : COMPLEX.lib
```

7- Atomic Charge Restraint (Optional)

The program provides the option to restrain the atomic charge on specific atoms. Upon execution, the following prompt will appear:

Would you like to restrain the charge on specific atoms? (Yes or No):

- Selecting No will terminate the operation.
- Selecting Yes will allow you to provide details for one or more atoms

7.1 Specify the Number of Atoms

If you choose to proceed, the following prompt will appear:

How many atoms do you want to restrain?

At this point, you will be required to specify the exact number of atoms for which the charge is to be restrained.

7.2 Providing Atom-Specific Information

You will see the following prompt:

Please provide the atom number and its charge for atom 1 (e.g., 12 -0.834):

You are required to input both the atom number and its corresponding charge for each selected atom.

7.3 Inputting Gaussian or Orca Output

You will see the following prompt:

Please enter the name of output that contains all the ESP charges:

At this point, you must specify the file name containing the Electrostatic Potential (ESP) charges:
For Gaussian: Provide the **standard output file** (e.g., .log or .out) that contains the ESP charges.

For Orca: Provide the output file with the **.vpot extension**, which contains the ESP charges.

7.4 Generating the Output

After processing, a new mol2 and library file will be generated, containing the updated atomic charges based on the information you provided.

4. Appendix

Example for Generating the Required Files

4.1 Gaussian Input

When using Gaussian to generate the required files, note that the choice of the level of theory is up to the user and is not mandatory to match the examples provided. Follow the steps below:

4.1.1 Structure Optimization

First, optimize the molecular structure. Upon completion, you will obtain the optimized structure in .xyz format. Once this step is complete, proceed to the next step.

```
%mem=30GB
%nproc=24
%chk=output.chk
#p B3LYP/def2tzvp symmetry=none opt scf=(MaxCycle=60,xqc)
freq(noraman) int(ultrafinegrid) iop(7/33=1)

#Comment

charge multiplicity
coordinates
```

4.1.2 Frequency Calculation

Use the optimized structure to perform a single-point frequency calculation. Upon completion, you will obtain the following files:

- Optimized Structure (.xyz): Contains the final geometry of the optimized structure.
- Checkpoint File (.chk): Stores data from the frequency calculation, including the Hessian matrix and other important information.
- Gaussian output (.log): Contains Hessian matrix and other important information.

```
%mem=30GB
%nproc=24
%chk=output.chk
#p B3LYP/def2tzvp symmetry=none scf=(MaxCycle=60,xqc)
freq(noraman) int(ultrafinegrid)

#Comment

charge multiplicity
coordinates
```

4.1.3 Charge Calculation Using the RESP Method

Next, calculate the molecular charge using the ESP method in Gaussian. This step will produce the charge output in the .log file.

```
%mem=30GB
%nproc=24
#p B3LYP/gen pseudo=read SCF=tight Pop=(mkuff) iop(6/33=2)
iop(6/42=6) iop(6/50=1)

#Comment

charge multiplicity

coordinates

C N H 0
6-31G*
****

Ru 0
SDD
****

Ru 0
SDD

output.resp

output.resp
```

4.2 Orca Input

When using Orca to generate the required files, the choice of the level of theory is flexible and does not need to match the examples provided. Follow the steps below:

4.1.1 Structure Optimization

First, perform a structure optimization of the molecule. Upon successful completion, you will obtain the optimized structure in .xyz format. Once this step is complete, proceed to the next step.

```
!B3LYP 6-31G*  RI-SOMF(1X) defgrid2 KDIIS TIGHTSCF Opt
%pal nprocs 24
end
%basis
NEWGTO  Pt "SDD" END
NewECP Pt "SDD" end
end
%scf MaxIter 1000 end
%maxcore 5000
* xyz 1 1
Coordinates
*
```

4.2.2 Frequency Calculation and Charge Calculation Using the CHELPG Method

1. Run a Single-Point Frequency Calculation using the optimized structure.
2. Output Files: Upon completion of the calculation, you will obtain the following files:
   - Optimized Structure (.xyz): Contains the final geometry of the optimized structure.
   - Charge File (.log): Contains the calculated charges based on the CHELPG method.
   - Hessian File (.hess): Contains the Hessian matrix from the frequency calculation.
   - ESP charge (.vpot): Contains the calculated ESP charges based on the CHELPG method.

```
!B3LYP 6-31G*  RI-SOMF(1X) defgrid1 KDIIS TIGHTSCF Freq
%pal nprocs 24
end
%basis
NEWGTO  Pt "SDD" END
NewECP Pt "SDD" end
end
%scf MaxIter 1000 end
! CHELPG
%maxcore 5000
* xyz charge multiplicity
Coordinates
*
```

5. References

(1) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int J Quantum Chem* **1996**, *60* (7), 1271–1277. https://doi.org/https://doi.org/10.1002/(SICI)1097-461X(1996)60:7<1271::AID-QUA8>3.0.CO;2-W.

(2) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Jin, C.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; O'Hearn, K. A.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; D.R. Roe; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xue, Y.; York, D. M.; Zhao, S.; Kollman, P. A. Amber 20. *Amber 20* **2021**, San Francisco.

(3) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. a.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. a.; Nakatsuji, H.; Li,

X.; Caricato, M.; Marenich, a. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, a. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. a.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. a.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, a. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. G16_C01. 2016, p Gaussian 16, Revision C.01, Gaussian, Inc., Wallin.

(4) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J Chem Phys* **2020**, *152* (22), 224108. https://doi.org/10.1063/5.0004608.

(5) Comell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP Charges. *journal of American Chemical Society* **1993**, *115* (7), 9620–9631.