A user's manual for

CREATE YOUR OWN FORCE FIELD

(CRYOFF)

a force matching code designed for the

Adaptive Force Matching

(AFM)

Omololu Akin-ojo, Jicun Li, Zhonghua Ma,
T. Ryan Rogers, Ying Yuan
and
Feng Wang

I. CRYOFF Introduction

"CRYOFF" stands for CReate Your Own Force Field, and is a program developed in the Wang group to created force fields based on adaptive force matching (AFM) procedure.

A partial list of journal articles that describe the AFM procedure is listed below, the users are strongly encouraged to read these articles before performing any AFM fitting.

1. Akin-Ojo, O.; Song, Y.; Wang, F., Developing ab initio quality force fields from condensed phase quantum-mechanics/molecular-mechanics calculations through the adaptive force matching method. *Journal of Chemical Physics* **2008,** *129* (6), 11
2. Pinnick, E. R.; Erramilli, S.; Wang, F., Predicting the melting temperature of ice-Ih with only electronic structure information as input. *Journal of Chemical Physics* **2012,** *137* (1), 5.
3. Wang, F.; Akin-Ojo, O.; Pinnick, E.; Song, Y., Approaching post-Hartree-Fock quality potential energy surfaces with simple pair-wise expressions: parameterising point-charge-based force fields for liquid water using the adaptive force matching method. *Molecular Simulation* **2011,** *37* (7), 591-605.
4. Li, J. C.; Wang, F., Pairwise-additive force fields for selected aqueous monovalent ions from adaptive force matching. *Journal of Chemical Physics* **2015,** *143* (19), 12.

II. Compiling CRYOFF

Current Version of CRYOFF: `cry2.7.3.f90` (updated 05/14/2020)

The source code for the CRYOFF program

can be compiled with a Fortran 90 compiler with the LAPACK DGELSD routine.

The code also can be compiled with MPI for parallel execution.

To compile for MPI, the -DMPI flag needs to be provided.

The code has been tested with the intel Fortran compiler (ifort) openmpi and iMPI.

For compilation on the University of Arkansas system with the module environment for intel compatible x86 CPUs.

The author would execute

```
module purge

module load  mkl/18.0.1  impi/18.0.1  intel/18.0.1

mpiifort -fpp -DMPI cry2.7.0.f90 -O3 -lmkl_lapack95_lp64
-lmkl_intel_lp64 -lmkl_intel_thread -lmkl_core -liomp5 -lpthread
-lmpi -o cryoff_parallel.x
```

without MPI

```
ifort -fpp cry2.7.0 -O3 -lmkl_lapack95_lp64             -
lmkl_intel_lp64 -lmkl_intel_thread -lmkl_core -lpthread o
cryoff_serial.x
```

The exact compiler option is system specific, for linking with the Intel MKL library, the best source of information is the MKL link advisor.
https://software.intel.com/en-us/articles/intel-mkl-link-line-advisor

III. Execution of CRYOFF

The execution of cryoff requires two files. One is the .ff file defining fitting protocol.
The other is the .ref file to provide reference forces.
The extensions for the input file must be adhered.

To execute cryoff, type

```
cryoff testinput[.ff]
```

The .ff is force field protocol file. This file contains molecular definition, interaction types and all the system control options.

The .ref reference file contains the coordinates and forces of atoms to be fitted. The name for the reference file is provided in the .ff protocol file.

The CRYOFF produces output in an .off file. The .off file name is also supplied in the .ff file.

The force evaluations will run in parallel with MPI. If a parallel Lapack library is linked in, the `DGELSD` routine for SVD can also run in parallel. This is controlled by the environment variable OMP_NUM_THREADS. The DGELSD will only be called by MPI thread 0 and use OMP_NUM_THREADS number of CPUs. It is recommended to set OMP_NUM_THREADS to the number of CPUs on the head node.

IV. The .ff input file

The .ff file contains four categories of information: the molecular definition for atom types and intramolecular term, the intermolecular terms, charge-constraints and the system control options. In CRYOFF, the intermolecular terms are synonym for non-bonded interactions. Thus, intramolecular short-range non-bonded and coulombic interactions are modeled with these intermolecular terms.

The .ff contain "cards" denoted by a case-insensitive name in brackets, e.g. `[CARd_name]` Only the first few characters of the card name is significant. This will be noted by capitalization.

Comment line are allowed in CRYOFF .ff files. Currently, all of the following characters are acceptable to set off a comment:

`;` `#` `!`

CRYOFF will ignore from that symbol through the end of the line.

Unknown keywords will be ignored by the program and will not lead to an error termination.

Different options are to be separated by space. For keywords with = (eg. step=0.1) no space should be put before or after the =.

Unless otherwise stated, the sequence of cards in the .ff file does not matter.

It is advised to have a maximum 80 character limit for each line.

1. System Options

`[FILe]   file1.ref   file2.off`

Names for the .ref file and the .off file. If extension is provided, the program will recognize the extension and set the files accordingly.
Without extensions, it is assumed the .ref file being the first argument followed by the .off file.
The program will not run if the .off already exists.

`[KEYwords]   fit_type   NORMalization=weight_option   PRInt` …

`fit_type`	Options are `INTEr` or `intra.` Without the inter keyword, cryoff will default to intramolecular fit. An intramolecular fit ignores the net forces (NETF) and molecular torque (TORQ) lines in the ref file and fit atomic forces only. With the inter keyword, only NETF and TORQ are fit. Optionally if only some molecules are to be fitted. It is possible to request a partial fit with INTEr=WAT1, METH or INTEr=WAT1.NETF The first example fits NETF and TORQ for molecule type WAT1 and METH. The second example fits only NETF of WAT1.

No selective control can be done with atomic force fits. Selective control of molecular fitting can be accomplished by using the solvation factor in the .ref file. The solvation factor based approach is recommended.
Selection of atoms by passing arguments to INTEr is a depreciated feature and may not work as expected.

NORMalization=weight_option

This an optional keyword controlling with weight of each gradient being fit. Each equation is being multiplied by the weight *w*.
The objective functions being optimized is

$$\sum_i (F_{fit}^i - F_{ref}^i)^2 (s_i \cdot w^i)^2 \qquad (1)$$

In Eq. 1 w^i is the weight for each atom, F could stand for either force or torque or charge constraint

Weight_option only affects the weight for force and torque. The weight for charge constraint is controlled only in the [CTSR] section.

The available options are

norm=RMS or norm=(RMS,2.0)
norm=REL or norm=(REL,1.5)

The WRMSF (Weighted Root Mean Square Forces) is computed as

$$\mathrm{W}RMSF = \sqrt{\frac{\sum_i \left(s_i^2 \left| F_i^{ref} \right|^2 \right)}{N}},$$

with N being the number of force components with a finite weight. (weight greater than zero tolerance (10^{-10}))
With norm=RMS, the weight is $w^i = 1/WRMSF$.
If the numbers of NETF and TORQ are not the same, the weight for each equations is further multiplied by the square root of the number of lines with fitted NETF over the number of lines with fitted TORQ to ensure the component will more equations will contribute more to the objective function.

With Norm=(RMS,2.0), the weight factor parameter 2.0 will lead to the weight for net force equations, w^i, to be multiplied by $\sqrt{2}$. The 2.0 factor has no effect when fitting atomic forces. The weight factor provides additional control over the relative weight for force and torque. However, the same goal can be accomplished by adjusting solvation factor in the .ref file, which is the recommended method since it provide more clarity.

With RELative, the weight is $w^i = \sqrt{\dfrac{1}{\max({F^i_{ref}}^2, WRMSF^2)}}$,

With norm=(REL,1.5), the weight is $w^i = \sqrt{\dfrac{1}{\max({F^i_{ref}}^2, (1.5 \cdot WRMSF)^2)}}$.

The use of relative weight might improve stability when the reference force computation is performed on out-of-equilibrium conformations. The atoms or molecules with larger forces are weighted down.

The default is Norm=RMS.

`PRInt` OPTIONAL. The print keywork can take options such as PRI=(mat,fit,debug) to control different levels of extra printing of the output. Mat will cause the Amatrix of the SVD to be printed in a separate file. Fit will cause more detailed information regarding to the fit to be printed in a separate file. Debug will cause more debugging information to be printed in the .off file.

PBC Optional. Request calculations with periodic boundary conditions. CRYOFF will get the box vectors from the `box=(...)` string in the 2nd line of the .ref file.
With PBC, coulombic interaction will be treated with ewald summation.
In addition, the following keywords will also be required to control ewald.
The rcou and rvdw keywords have no effect without PBC. When PBC is not used, all pairwise interactions are evaluated regardless of distance.

`Etol=`<*real*>	The Ewald tolerance. Default 10e-6
`Rcou=`<*real*>	Coulombic cutoff in Å. Default rcou=10
`rvdw=`<*real*>	van der Waals cutoff in Å. Default rvdw=10

This keyword is currently not working.

`MAXFrame=`<*int*> Optional. Control the number of frames to be read from the .ref file. The default behavior is to use all frames available in the .ref file.

USEFrame=(initial,final)
Optional, only the range of frames between the initial and final (inclusive) will be fitted.

KCV=4 Optional, request the computation of the kfold cross validation (CV) values.

KCV=(Random,6)	KCV=4 will sequentially separate the data into 4 folds to perform the CV computation. KCV=(Random,6) will randomly separate the data into 6 folds to perform the CV computation.

`[OPTimization] method step=0.1 INItialize=FRACtion CONV=1e-5 MAXITeration=1000`

method	Current choices are SIMplex, POWell, and DE (differential evolution) Without these keywords, the default is linear fit only.
method.Read	The read suffix as in SIM.R will instruct the program to read the set up for the non-linear optimizations such as the initial vertices for the simplex. The non-linear optimization set up should appear immediately follow the [OPT] line. With one set of all non-linear parameters for each line. The best way to construct this section is from the parameters printed out in the .off file.
INItialize	optional. Take options READ, FRACtion or UNIT. Control the generation of the initial simplex or trial moves for the Powell's quadratically convergent algorithm. READ will instruct CRYOFF to read trial move parameters, which should be placed after the [OPT] line, one vector per line. UNIT will construct moves that is of uniform length determined by the step directive. (see below), FRAC will construct moves that proportional to the parameter. The default is FRAC.
Step=0.1	optional, step size in the generation of initial simplex or trial moves. (The default is 0.1 for FRAC and 1.0 for UNIT)
CONV=1e-5	Set the convergence of the optimizations. CRYOFF use absolute difference to determine convergence rather than fractional difference.
MAXITeration=100	Set the maximum number of iterations allowed before optimization terminates. The default is 100.

2. Molecular Definitions

Example: (Underlined Capitalization means required keyword.)

```
[MOLeculetype] METQM
 [atoms] 7
  1    Hm_QM              Hm_QM
  2    Hm_QM              Hm_QM
  3    Hm_QM              Hm_QM
  4    Hm_QM              Hm_QM
  5    Cm_QM              Cm_QM
  6    NetF        NetF
  7    Torq        NetF
 [Bonds]  1
  HAR   4    FIX     1.090588    761.636067
    1  5
    2  5
    3  5
    4  5
 [Angles] 1
  HAR   6    FIX    109.709999     70.391957
    1   5   2
    1   5   3
    1   5   4
    2   5   3
    2   5   4
    3   5   4
```

```
---------------------------------------------------------------------------------------------------------
[moleculetype]   Mol_name


[atoms]   N
     1    vdwname₁      Coulname₁
     2    vdwname₂       Coulname₂
     .        .        .
      i*    vdwnameᵢ      Coulnameᵢ nv v1 r1 v2 r2 … vnv rnv

     .        .        .
    (N-1)  NetF
     N    Torq
```

`Mol_name`	Name of the molecule. Must match a molecule name in the .ref file.
`N`	Number of lines in the .ref file for the `[moleculetype]` molecule. NetF and Torq are special entries for the molecular net force and torque. If provided, these count toward the number of lines.
i	The first column is the atom number. If the atom number is tagged with an asterisk. The atom will be a virtual site. (The first atom of the molecule can not be a virtual site.) The virtual site can not have any force acting on them. However, the interaction with the virtual site will cause forces to be distributed to the atomic sites.

```
nv c1 v1 c2 v2 … cnv vnv
```

specifies how the position of the virtual site is related to the atomic sites for redistribution of forces.

nv is the number of atomic sites the virtual site coordinate is to be derived from. **c1**, **c2** … are coefficients, v**1**, v**2** … are numbers corresponding to the vectors of the atomic sites. The location of the virtual site is thus

$$\sum_{i=1}^{nv} (c_i) * (\vec{v}_i)$$

For the BLYPSP-4F model, the specification is thus

4*   M   M   3 0.6 1 0.2 2 0.2 3

Assuming site 1 is oxygen, sites 2 and 3 are hydrogens.

	In fact the virtual site location is read in from the ref file and used as is. The CRYOFF code only use the virtual site definition to redistributed forces. Thus it is important that the virtual site positions in the ref file is correct.
`vdwname`$_i$	Atom name (a character string) of the ith atom in the molecule. It is imperative that these atom names match those found in the first column of the .ref file!
`Coulname`$_i$	Optional. The code allows a different atom type to be used for Coulombic Interactions. This way different atoms will have different charges but share the same vdw types (vdwname) and vice versa. If this field is omitted. The coulombic type will be the same as vdw type. Although this field is optional, the input file has to be consistent. If the coulname is provided for the first atom of the molecule. It must also be provided for all the rest of the atoms. If it is not provided to the first atom, it should not be provided for any other atoms.

```
-------------------------------------------------------------------------------------------------------
[bonds]   num
   type     N        fit_flag    P1   P2 ... PN
     int_atom1     int_atom2
     int_atom1     int_atom2
        .             .
        .             .
        .             .
```

`num`	Number of bond types in this molecule.
`type`	Bond type. A list of all supported bond types is given in Table 1.
`N`	Number of atom pairs with this bond `type`. There must be *N* lines of atom pairs listed immediately afterwards that is interacting according to `type`. int_atom$_1$ and int_atom$_2$
`fit_flag`	Only options are `FIT` or `FIX`. FIT will instruct cryoff to optimize this parameter. FIX will disable the optimization of the parameter. The parameter will be fixed at the value provided. When the parameter is fit, the actual parameter provided will be the initial guess for non-linear optimizations. The value provided will not affect the least square determination of the linear parameters.

`P`$_i$	Parameters of the FF term denoted by `type` previously on the same line. The number of parameters for each `type` can be found in Table 1 below, and varies for each interaction type.
`int_atom`$_i$	Atom number of one of the atoms interacting according to `type`. The `int_atom`$_i$ number is one of the "`i`" listed in the `[atoms]` sub-card.

Table 1. List of all available FF terms for bonded interactions. Use the "CRYOFF symbol" column of this table to call upon its related functional form in an `[bonds]` sub-card of your .ff file.

Bond Interaction description	CRYOFF symbol	Functional form (All distances are in unit of Å, energy in kcal/mol, and angle in degrees)	Nonlinear parameters
harmonic, 2-atom	HAR	$U = \frac{P_2}{2} \cdot (R - P_1)^2$	None
quartic, 2-atom	QUA	$U = \frac{P_2}{2} \cdot (R - P_1)^2 + \frac{P_3}{3} \cdot (R - P_1)^3 + \frac{P_4}{4} \cdot (R - P_1)^4$	None

---

```
[angles]   num
    type       N       fit_flag    P1   P2 ...  PN
       int_atom1     int_atom2     int_atom3
           .             .              .
           .             .              .
           .             .              .
```

`num`	Number of angle types in this molecule.
`type`	Angle type. A list of supported angle types is given in Table 2.
`N`	Number of lines list triplets of atoms interacting according to this `type`. There must be *N* lines of atoms listed immediately afterwards that is interacting according to `type`. For the int_atom$_1$-int_atom$_2$-int_atom$_3$ angle, int_atom$_2$ is the center atom.
`fit_flag`	Only options are `FIT` or `FIX`. FIT will instruct cryoff to optimize this parameter. FIX will disable the optimization of the parameter. The parameter will be fixed at the value provided. When the parameter is fit, the actual parameter provided will be the initial guess for non-linear optimizations. The value provided will not affect the least square determination of the linear parameters.

`P_i` Parameters of the FF term denoted by `type` previously on the same line. The number of parameters for each `type` can be found in Table 2 below, and varies for each interaction type.

`int_atom_i` Atom number of one of the atoms interacting according to `type`. The `int_atom_i` number is one of the "`i`" listed in the `[atoms]` sub-card.

Table 2. List of all available FF terms for intramolecular angle interactions. Use the "CRYOFF symbol" column of this table to call upon its related functional form in an `[angles]` sub-card of your .ff file.

Angle Interaction description	CRYOFF symbol	Functional form (All distances are in unit of Å, energy in kcal/mol, and angle in degrees)	Nonlinear parameters
harmonic, 3-atom	HAR	$U = \frac{P_2}{2} \cdot (\theta - P_1)^2$	None
quartic, 3-atom	QUA	$U = \frac{P_2}{2} \cdot (\theta - P_1)^2 + \frac{P_3}{3} \cdot (\theta - P_1)^3 + \frac{P_4}{4} \cdot (\theta - P_1)^4$	None

```
-----------------------------------------------------------------------------------------------------------
[dihedrals]   num
   type      N      fit_flag    P₁   P₂ ... Pɴ
      int_atom₁     int_atom₂     int_atom₃    int_atom₄
          .             .             .             .
          .             .             .             .
          .             .             .             .
```

`num` Number of dihedral types in this molecule.

`type` dihedral type. A list of supported dihedral types is given in Table 3.

`N` Number of lines list of atoms interacting according to this `type`.

There must be *N* lines of atoms listed immediately afterwards that is interacting according to `type`.
int_atom$_1$-int_atom$_2$-int_atom$_3$-int_atom4

`fit_flag` Only options are `FIT` or `FIX`.
FIT will instruct cryoff to optimize this parameter. FIX will disable the optimization of the parameter. The parameter will be fixed at the value

provided. When the parameter is fit, the actual parameter provided will be the initial guess for non-linear optimizations. The value provided will not affect the least square determination of the linear parameters.

`P`$_i$	Parameters of the FF term denoted by `type` previously on the same line. The number of parameters for each `type` can be found in Table 3 below, and varies for each interaction type.
`int_atom`$_i$	Atom number of one of the atoms interacting according to `type`. The `int_atom`$_i$ number is one of the "`i`" listed in the `[atoms]` sub-card.

Table 3. List of all available FF terms for intramolecular dihedral angle interactions. Use the "CRYOFF symbol" column of this table to call upon its related functional form in an `[dihedrals]` sub-card of your .ff file. Note that the CRYOFF symbols for dihedral interactions have 4 characters, whereas most other symbols are only 3 characters.

Dihedral Interaction description	CRYOFF symbol	Functional form (All distances are in unit of Å, energy in kcal/mol, and angle in degrees)	Nonlinear parameters
Harmonic dihedral 4-atom	HARD	$U = \frac{P_2}{2} \cdot (\phi - P_1)^2$	none
Normal Cosine dihedral, 4-atom	NCOS	$U = P_1 \cdot [1 + \cos(P_2 \cdot \phi - P_3)]$	P_2, P_3
Cosine dihedral, 4-atom	COS	$U = P_2 \cdot [1 + \cos(P_3 \cdot \phi - P_1)]$	P_3

```
[EXClusions]   N
    int_atom1    int_atom2
       .            .
       .            .
       .            .
```

`N`	Number of intramolecular non-bonded interactions to be excluded for this molecule
`int_atom`$_1$	

int_atom$_2$	The atoms (index) of the pair of intramolecular non-bonded interaction to exclude, without listing them in the exclusion list, the atoms will interact with each other via Coulombic and van der Waals even if they are covalently bonded to each other .

```
[FUDge]    N    fudge_vdw       fudge_Cou
    int_atom1     int_atom2
       .             .
       .             .
       .             .
```

N	Number of intramolecular non-bonded interactions to scaled by the fudge factors.
fudge_vdw	Scaling factor for the van der Waals interaction
fudge_Cou	Scaling factor for the Coulombic interaction
int_atom$_1$	
int_atom$_2$	The atoms (index) of the pair of intramolecular non-bonded interaction to scale the short-range (van der Waals) interaction by fudge_vdw and the coulombic interaction by fudge_Cou.

3. Definition of Intermolecular Interactions

```
[COUlomb]    N
     A.name1      A.name2      fit_flag     q.prod
        .            .            .            .
        .            .            .            .
        .            .            .            .
```

N	Number of coulombic interactions to be evaluated between all atom pairs in the system. (With CRYOFF, the pairs for which Coulombic interactions are to be evaluated has to be defined explicitly, since no atomic charges are defined in the .ff file. Only charge products of the pairs are provided/fitted.) Intramolecular coulombic interactions should also be defined here.
A.name$_i$	Name of atom in pair interacting according to Coulomb's law. Each `A.name`$_i$ is one of the atom names listed in the `[atoms]` cards in one of molecules.

`fit_flag`	Only options are `FIT` or `FIX`. FIT will instruct cryoff to optimize this parameter. With FIX, the parameter. will be fixed at the value provided.
`q.prod`	Product of the charges Parameter. Product of charges of `A.name`$_1$ and `A.name`$_2$
Suggestion:	There is no need to include charge products between atoms not being fit. (solvation factor zero.) However, charge products between atoms being fit and not being fit should be included. For inter fitting, there is no need to provide intra coulombic terms, since the charge products fitted from these terms won't affect the objective function.

```
--------------------------------------------------------------------------------------------------------
[<vdW_type>]    N
    A.name1      A.name2     fit_flag       P1    P2 ... PN_[min:max:step]
       .            .            .          .     .      .
       .            .            .          .     .      .
       .            .            .          .     .      .
```

<vdW_type>	van der Waals (non-coulombic short-range non-bonded) interaction. Available options listed in Table 4.
`N`	Number of pairs of the *<vdW_interaction>* type to be provided in the ff file.
`A.name`$_i$	Name of the atoms of the *<vdW_type>* interaction pair. Each A.namei is one of the atom names listed in the [atoms] cards in one of molecules.
`fit_flag`	Only options are `FIT` or `FIX`. FIT will instruct cryoff to optimize this parameter. With FIX, the parameter. will be fixed at the value provided.
`P`$_i$	Parameters of the intermolecular, *<vdW_type>* function. The definition of parameters for each *<vdW_type>* can be found in the Table 4. If `P`$_i$ is a non-linear parameter, refer to the format explained for the next entries, `min`, `max`, and `step`.
`_[min:max[:step]]`	For each nonlinear parameter,(see Table 4) a range is to be provided to the optimization algorithm. This is to be provided as "`_[min:max[:step]]`" No space allowed between the underscore and the parameter value(initial

guess). The step is optional. If omitted, the step specification provided in the [KEYword] card will be used. Note that the step size provided here overrides the step specification in the OPT keyword.

Omitting the "`_[min:max[:step]`" string will lead to the nonlinear parameter being fixed. In this case, the fit_flag will only lead to the linear parameters to be optimized. CRYOFF will determine nonlinear parameter according to Table 4.

As a simple example of how the van der Waals interactions in the .ff file are interpreted by CRYOFF, consider the one power law described in the sample .ff file below:

```
[POWer]          1
  Cm_QM    Ow_QM      FIX    -1095.567     -6_[-7,-5]
```

According to Table 4, the `[Power]` functional form is:

$$U_{power} = P_1 \cdot R^{P_2}$$

therefore, the interaction in the .ff file, is:

$$U_{C_m O_w} = -1095.567 \cdot {R_{C_m O_w}}^{-6}$$

With the FIX flag, the range specification for the non-linear parameter is ignored.

Table 4. Functional forms of all possible van der Waals FF terms (intermolecular interactions) that can be parameterized by CRYOFF. The "CRYOFF symbols" are the minimal symbols used in the .ff file to denote their functional forms; more transparent labels are suggested (e.g. "Buckingham" is preferred instead of merely "Buc"). Some functional forms include nonlinear parameters, which may require special treatment in the .ff file.

Interaction type	CRYOFF symbol	Functional form	Nonlinear parameters
Generalized Lennard Jones	GLJ	$U = \frac{P_1}{R^{P_3}} + \frac{P_2}{R^{P_4}}$	P_3 and P_4
Buckingham	BUC	$U = P_1 \cdot e^{-P_3 R} + \frac{P_2}{R^6}$	P_3
diffuse Buckingham	DBU	$U = P_1 \cdot e^{-P_3 R} + \frac{P_2 \cdot f}{R^6}, \qquad f = \frac{1}{1 + e^{-P_4 (R - P_5)}}$	P_3, P_4, and P_5
shift-truncated	STR	$U = P_1 \left(\frac{1}{R^{P_2}} - \frac{1}{{P_3}^{P_2}} + \frac{P_2 \cdot (R - P_3)}{{P_3}^{(P_2+1)}} \right)$	P_2, P_3
exponential	EXP	$U = P_1 \cdot e^{-P_2 R}$	P_2
power law	POW	$U = P_1 \cdot R^{P_2}$	P_2

power-exponential	PEX	$U = P_1 \cdot R^{P_2} \cdot e^{-P_3 R}$	P_2, P_3
diffuse power law	DPO	$U = P_1 \cdot f \cdot R^{P_2}, \qquad f = \frac{1}{1 + e^{-P_3(R-P_4)}}$	P_2, P_3, P_4
Short-range damped dispersion	SRD	$\mathrm{U} = P_1/(R^{\|P2\|} + {P_3}^{\|P2\|})$	P_2, P_3

4. Charge Constraints

```
[charge_constraints]   N
   #Cou_prods :   α(prod.#a) + β(prod.#b) + … = sum     weight
        .           .   .        .   .        .         .
        .           .   .        .   .        .         .
        .           .   .        .   .        .         .
```

`N`	Number of charge constraint equations to be read in by the program. Each line below is a charge constraint equation.
`#Cou_prods`	Number of terms in this constraint equation.
α, β	Multipliers of the charge products in the constraint equation. Typically, this is the number of atoms of a certain type in the group.
prod.#$_a$	Line number listing the charge product in the `[Coulomb]` card.
`sum`	Constraint value for the sum. Zero for neutral fragments.
`weight`	Weighting factor used by CRYOFF. A recommended value is 1000.

Example:

`If the [Coulomb] card has:`

```
[Coulomb]  2
  Cm_QM   Hw_QM     FIT        -0.326356742
  Hm_QM   Hw_QM     FIT         0.081589182
```

`A charge constraint equation specification of`

```
  2:   1(1) + 4(2) = 0.0   100
```

states that

1*Cm_QM*Hw_QM+4*Hm_QM*Hw_QM = 0.

IV. The .ref file

The .ref file contains the reference forces which you would like your FF to reproduce. The order of atoms in the .ref file must be the same as that in the molecular definition in the ff file. Each line of the .ref is space delimited.

```
---------------------------------------------------------------------------
N
Comment       box=(xx, xy, xz; yx, yy, yz; zx, zy, zz)  wgtsca=1.0
A1      xA1   yA1   zA1   Fx,A1          Fy,A1          Fz,A1          QM_wt.A1   Mol.nameA1
A2      xA2   yA2   zA2   Fx,A2          Fy,A2          Fz,A2          QM_wt.A2   Mol.nameA2
.        .     .     .     .              .              .              .          .
.        .     .     .     .              .              .              .          .
.        .     .     .     .              .              .              .          .
An      xAn   yAn   zAn   Fx,An          Fy,An          Fz,An          QM_wt.An   Mol.nameAn
NetF     0     0     0    Σn_i=1(Fx,Ai)  Σn_i=1(Fy,Ai)  Σn_i=1(Fz,Ai)  QM_wt.An   Mol.nameAn
Torq    xA    yA    zA    Σn_i=1(τx,Ai)  Σn_i=1(τy,Ai)  Σn_i=1(τz,Ai)  QM_wt.An   Mol.nameAn
---------------------------------------------------------------------------
```

`N`	Number of lines with force information for the current conformation. This is similar to the .xyz file with NetF and Torq also count toward this number
`Comment`	Comment string with reserved keyword box and wgtsca
`box=(...)`	OPTIONAL. If PBC is requested, the box information is provided with the box string. "`box=(...)`" The box specification should be 9 numbers giving the (a,b,c) vectors of the unit cell. "box=`ax, ay, az, bx, by, bz, cx, cy,` and `cz"`.
`wgtsca`	optional. Default=1.0 A scaling factor for the conformation. All the solvation factors for this conformation will be multiplied by this factor before the weight is to be calculated in CRYOFF.
A_n	Atom name of the *n^th^* atom of molecule *A*. The atom names should be in the same sequence as in the .ff file except NetF and Torq, which stands for molecular force and molecular torque.
x_{An} y_{An} z_{An}	Cartesian coordinates of the site. In case of Torq, this is the center with respect to which the torque is calculated. These values are ignored for NetF.
Fx_{An} Fy_{An} Fz_{An}	Reference force for the atom or the net molecular and net molecular torque. This has to be zero for virtual sites.
$QM\_wt._{An}$	Solvation factor. This is the solvation factor in Eq. 1. Atoms not being fit has a solvation factor of 0.

`Mol.name`$_{An}$	The molecule name associated with atom *n* of molecule *A*. The name must contain the corresponding molecular name defined in the .ff file. However, a number could be added to the molecular name. For example, WAT001 or 001WAT are acceptable, assuming the molecular name is WAT.

Sample of a CRYOFF .ref file (truncated):

```
1125
QM/MM configuration from /gpfs_home/trr007/solvation/methane/qmmm-2/fm1_E3/conf001
Hm_QM    10.790001     10.840001     10.850001    -12.206840      7.586225     -6.403846     1        1METQM
Hm_QM    10.580001      9.180001     10.650001     10.934591    -30.586853      1.440173     1        1METQM
Hm_QM    10.720001     10.190001      9.160001     -5.826609     -3.981564     24.704785     1        1METQM
Hm_QM     9.220001     10.130001     10.180001     17.264311     10.622835      2.239287     1        1METQM
Cm_QM    10.330001     10.100001     10.200001    -14.154535     17.723951    -20.090967     1        1METQM
NetF      0.000000      0.000000      0.000000     -3.989082      1.364594      1.889432     1        1METQM
Torq     10.328001     10.088001     10.208001      1.165062     -1.834890      1.648742     1        1METQM
Ow_QM    10.120001      7.680001     13.310001     22.689588     18.531944    -39.502559     1      185SOLQM
Hw_QM    10.330001      7.070001     14.020001      4.217245     -6.641985      8.830513     1      185SOLQM
Hw_QM    10.810001      7.630001     12.560001    -20.663686     -8.686832     35.081140     1      185SOLQM
Mw_QM    10.300001      7.548001     13.302001      0.000000      0.000000      0.000000     1      185SOLQM
NetF      0.000000      0.000000      0.000000      6.243147      3.203127      4.409094     1      185SOLQM
Torq     10.420001      7.460001     13.296668     -8.012693     -6.162416     -8.183827     1      185SOLQM
```

……

VI. The .off output file

CRYOFF prints both to standard output and write to the .off output file. More diagnostic information regarding how the code is reading and interpreting the input file is only printed to standard output. Thus the users are strongly encouraged to carefully study both the output and the output file. While the former provide valuable information regarding if the .ff is correctly interpreted by the code, the latter provide information regarding to the fitted parameters and the quality of the fit.

The RMS values printed are normalized to per atom not per force component.

The “Force Field Potentials” output follows similar definition as input. See table 1-4 for sequence of parameters and units. Generally, the energy is in kcal/mol, the distance is in Å and angles in degree.

In the “Inter-Potential” section, MIN and MAX stands for the minimum distance in Å. This is the minimal and maximum distance being fitted in .ref file. Distances between solvation factor 0 atoms is not counted since they are not being fit.

The “Molecular-Definition” section summarizes the parameter reported previously in the Intra-Potential section. While the sequence of parameter is according to tables 1-3, the units have been converted to the same as those used for Gromacs. Energy will in kJ/mol, distance in nm and angles in degree.

The “Table-Potential” section summarizes all the coulombic and non-bonded short-range interaction for each atom pair and the print out is in the same format that is accepted by the table file generation scripts in the AFM toolkit.

The Simplex Vertices printed after each 5 iterations can be used put after the [opt] section to use with Simplex.Read.