PROGRAM: GLOpSAnn
Author: Juan Rodriguez-Carvajal (ILL)
Created: October 2008. Last updated (JRC): April 2020

Documentation for GLOpSAnn Program

The program **GLOpSAnn** was conceived as an example of using **CrysFML**. It has been in the repository for a long time without being publicized. We are now distributing executables within the FullProf Suite. An old version of the source code is public and may be found in the **CrysFML** repository:

http://forge.epn-campus.eu/projects/crysfml/repository/show/Program_Examples/Structures_GlobalOptimization

The input of the program is a CFL file containing the operating instructions, information about the crystallographic parameters and the data files. The program uses Simulated Annealing (SAnn) for global optimization. The program tries to find structural parameters that minimizes a particular (or composite) cost function (CF) that may be formed by differences between observed and calculated intensities, distances, valences, etc. The available cost functions at present are the following:

```
"F2obs-F2cal"     Cost(F2obs-F2cal)    : Optimization of C0 =Sum|F2obs-F2cal|/Sum|F2obs|
"Fobs-Fcal"       Cost(Fobs-Fcal)      : Optimization of C1 =Sum|Fobs-Fcal|/Sum|Fobs|
"dis-restr"       Cost(dis-restr)      : Optimization of C2 =Sum{w(dobs-dcal)^2}, w=1/var(d)
"Ang-restr"       Cost(Ang-restr)      : Optimization of C3 =Sum{w(Ang_obs-Ang_cal)^2}, w=1/var(Ang)
"Tor-restr"       Cost(Tor-restr)      : Optimization of C4 =Sum{w(Tor_obs-Tor_cal)^2}, w=1/var(Tor)
"bond-valence"    Cost(bond-valence)   : Optimization of C5 =Sum{|q-BVS|/tot_Atoms}
"bvs_coulomb"     C5 + Cost(Coulomb)   : Optimization of C6 =Sum{qi qj/dij}
"FoFc-Powder"     Cost(FoFc-Powder)    : Optimization of C7 =Sum|Gobs-Sum(Fcal)|/Sum|Gobs|
"Coordination"    Cost(Coordination)   : Optimization of C8 =Sum|Coord-Efcn|/Sum|Coord|
"Anti_Bump"       Cost(Anti_Bump)      : Optimization of C9 =Sum{(dmin/d)**power}
"Powder_Profile"  Cost(Powder_Profile) : Optimization of C10=Sum{|yiobs-yicalc|/|yiobs|}
"Powder_WProfile" Cost(Powder_WProfile): Optimization of C11=Sum{w{yiobs-yicalc}^2/(N-P)} Equivalent to Chi2
```

The names between quotes are used for identification in the input control file. Different CFs may be combined in the optimization with a weight that has to be provided by the user.

We provide here only some hints for running the program with a minimalistic description of the input file. The structure of the CFL file is free format and very simple. For running **GLOpSAnn** some instructions should be given in the CFL file. We will take an example and we will describe the need of the different parts. The example we are taking for description corresponds to an optimization in which we provide a list of integrated (overlapped) neutron powder intensities. The full file is the following:

```
Title  PbSO4 (experimental Jvi=-11 |F|) Neutrons
!        a           b           c      alpha     beta    gamma
Cell   8.485454   5.402319   6.964360  90.000   90.000  90.000
!     Space Group
Spgr   P n m a
!                    x         y         z         B         occ     Spin Charge
Atom   Pb    PB    0.18797   0.25000   0.16754   1.35290    0.50000    0.0    2.0      #color 0 0 1 1
Atom   S     S     0.06467   0.25000   0.68300   0.89361    0.50000    0.0    6.0      #color 1 1 0 1
Atom   O1    O     0.90712   0.25000   0.59675   0.57221    0.50000    0.0   -2.0      #color 0 1 1 1
Atom   O2    O     0.18635   0.25000   0.54278   0.99996    0.50000    0.0   -2.0      #color 1 0 1 1
Atom   O3    O     0.08021   0.02965   0.81211   1.07399    1.00000    0.0   -2.0      #color 1 0 0 1

! Codes for refinement
Vary xyz 0   1   0   1
!
HKL-OBS   pb_neu.int
MIN-DSPACING    1.5
```

```
WAVE  1.912
RADIATION NEUTRONS
FST_CMD  conn S O 0.0 1.8 ; poly S

OPTIMIZE    bond-valence 0.15  FoFc-Powder 0.85
LOCAL_OPTIMIZATION

! Simulated Annealing conditions
SIM_ANN

!       Name of the cost function
CostNam  FoFc_Pow+BVS

!         T_ini      anneal      num_temps
TemParM     3.0        0.95        80

!         Nalgor  Nconf nm_cycl   num_therm   accept%
Algor_T     0        6   120         0         10.0

!         Value of Seed (if SeedVAL = 0, random seed)
SeedVAL     0
!
Threshold   25.0
!         Treatment of initial configuration
InitCON   RAN
```

The reserved keywords (case insensitive) are in **bold face** and **red**. We will describe each one of them for this particular case.

Title — This keyword is followed by whatever comment

Cell — Followed by 6 real numbers corresponding to cell parameters (in Å) and angles (in degrees)

SpGR — Followed by the symbol or the number of the space group

Atom — Followed by two alphanumeric symbols (***label*** and ***species*** for determining the scattering factor) and up to 7 real numbers that have the following meaning (the order is important): ***x***, ***y***, ***z*** (fractional coordinates or the atom), ***Biso*** (isotropic temperature factor), ***occ*** occupation factor (proportional to multiplicity of the site/multiplicity of the general position), ***spin*** (magnetic moment, in Bohr magnetons) and formal charge. Information after the symbol **#** may be passed to the program FullProf Studio (or VESTA)

Vary — This keyword symbols and numbers. Example: **Vary** ***xyz 0 1 0 1*** means that we vary the coordinates of all the atoms between 0 and 1 (inside the unit cell), the second 0 may be substituted by a step value (depending on the algorithm used for optimizing) and the last 1 means that the values satisfy periodic boundary conditions (if in a particular point one variable goes outside the limits, e.g. 1.3214 takes the value 0.3214).
Related instructions: **FIX** and **EQUAL**.
Examples: `fix x_Fe y_O4;` `Equal y_Fe y_P 0.25;`
`fix xyz La1 La2 La3 Mn1 Mn2 Mn3 Mn4 Sr`

Hkl-Obs — Followed by the name of the file containing integrated intensities. Only needed in the case experimental diffraction data are provided.

PROFILE-OBS — Followed by the name of the file containing profile intensities. The profile intensities file is that generated by **FullProf** with extension *.spr*. Only needed in the case experimental diffraction data are provided.

Min-Dspacing — Followed by the value d_{min} for which reflections with d_{hkl} below d_{min} are not used in the optimization procedure. Only needed if reflection intensities are provided.

Wave — Followed by the value of the radiation wavelength. Only needed if diffraction intensities are provided.

Radiation — Followed by type of radiation used: *Neutrons*, *X-rays* or *Electrons*. For electrons the only cost function options are Fobs-Fcalc or F2obs-F2calc for intensities obtained from precession images.

FST_CMD — Followed by commands for visualization programs **FullProf Studio** or **VESTA**. Example: **FST_CMD** *conn S O 0.0 1.8 ; conn Pb O 0.0 2.6*.

OPTIMIZE — Followed by the names and weight of the cost functions to be optimized
E.g.: *bond-valence 0.15 FoFc-Powder 0.85*

Ref_Within — If this keyword is present, followed by a real number (*ref_within delta*), the program performs an optimization (local or global) of coordinates within a parallelepipedic box around the initial atom positions of size *delta* along each direction *x*, *y*, *z*. The size is given in angstroms. E.g.: *ref_within 0.2*

Local_Ref — If this keyword is present, followed by an integer, the program performs a local optimization instead of a global optimization E.g. *Local_Ref n_points*, where *n_points* is the number of random starting points. This is an alternative to simulated annealing (see below) when we want to obtain the local minimum closest to the initial configuration. This option should be used together with **Ref_within** and the *n_points* starting points of the algorithm are generated within the box defined in **Ref_Within**.

Local_Optimization — If this keyword is present after **Optimize** the program performs a local optimization if the value of the cost function is just below a threshold value provided by the user (see below)

SIM_ANN — This keyword starts a series of lines constituting the information needed for making a simulated annealing optimization.

CostNam — This keyword is followed by the username of the cost function.
E.g. `FoFc_Pow+BVS`

TemParM This keyword is followed by three real numbers, *T_ini*, *anneal*, *num_temps* representing the initial temperature, the annealing factor and the maximum number of temperatures for the simulation.

Algor_T This keyword is followed by four integer numbers, *Nalgor*, *Nconf*, *nm_cycl* and *num_therm* representing the type of algorithm (0/1 for Corana algorithm. Other values: conventional (fixed step)), the number of configurations considered for each temperature (if *Nconf* > 1 parallel SAnn is performed), the maximum number of Monte Carlo cycles per temperature, the number of thermalization cycles before counting averages. And, finally, a real number *accept%* that tells to the program a convergence criterion: if the percentage of accepted configurations is lower than the given value, the program has converged and stop.

SeedVAL This keyword is followed an integer value. If zero the seed of the random generator is taken according to the system clock. Otherwise the same sequence of random number is conserved from run to run.

Threshold This keyword is followed a real number representing the value of the const function below which a local optimization is applied.

InitCON This keyword tells to the program how to generate initial configuration. If the keyword is followed by RAN, the initial configuration is taken at random but respecting the symmetry constraints given in the initial coordinates. For instance if an atom is placed in a special position the fixed coordinates are not moved in the algorithm. In all other cases the initial configuration is that provided by the user in the list of atoms.

Examples of OPTIMIZE instructions

The weight factors in OPTIMIZE instructions are normalized to 1.0 inside the program.

Distance restraints combined with list of observed structure factors:

```
OPTIMIZE   dis-restr 1

OPTIMIZE   Fobs-Fcal 1.0  dis-restr 1.0  ang-restr 1.0

DFIX   3.19620  0.00000  Ni  Ni_3.545
DFIX   2.90276  0.00000  Ni  Fe_1.554
DFIX   2.06756  0.00000  Ni  O1_2.455
DFIX   2.13013  0.00000  Ni  O2
DFIX   2.04270  0.00000  Ni  O4_4.554
. . .
DFIX   3.08789  0.00000  O2  O4_2.455
DFIX   2.80803  0.00000  O2  O4_4.554
. . .
AFIX    57.099    0.123  Fe  O3_1.556  O3_7.556
AFIX    55.618    0.068  P  Li  Li_3
AFIX    56.016    0.087  P  Li  Fe_1.554
AFIX    50.853    0.084  P  Fe_1.554  Li
. . .
```

The list of distances (or angles) for restraints may be provided inside the CFL file or in a separated file that is accessed by the keyword: **RESTR_FILE** followed by the name of the restraint file containing **DFIX** and/or **AFIX** instructions. If only distance restraints are provided we do not need intensity of reflections at all and all the corresponding parameters and keywords need not to be provided. A list of

angle and distance restraints in the appropriate format may be generated by using the program **Bond_STR** from the **FullProf Suite**.

```
OPTIMIZE  dis-restr 2.0  bvs_coulomb 1.0  1.0  coordination 4.0
```

Notice that the cost function **bvs_coulomb** has two weight factors corresponding to the BVS and to the Coulomb potential. The charges of ions (provided in the **Atom** instructions) are needed for this kind of cost function.

When **coordination** cost function is to be optimized the program need a coordination instruction of the form:

```
COORDINATION   n1  n2   n3  n4  . . .  nt
```

Where the integer numbers `n1 n2 …` up to the total number of atoms `nt` in the asymmetric unit are the expected coordination numbers. If some of them are zero, the restraint is not applied to the corresponding atom.

```
OPTIMIZE  dis-restr 1.0  anti_bump 1.0  Fobs-Fcalc 4.0
```

When **anti_bump** cost function is to be optimized the program need a series of Anti-Bump instructions that adopt the following form: *Damin Element1 Element2 distance n*, where *n* is the power exponent of the penalty function: $(d_{amin}/d)^n$. The instructions are written inside the CFL file as:

```
!Anti-Bump instructions

damin  La   La  3.6  2
damin  Mn   Mn  2.8  2
damin  Mn   La  3.2  2
damin  O    O   2.5  2
damin  La   O   2.2  2
damin  Mn   O   1.6  2
```

Powder profile intensities may be used for solving a crystal structure. In such a case two intensity files are needed provided with the keywords: **HKL-OBS** and **PROFILE-OBS**. The intensity files are those generated by **FullProf** after a Le Bail fit with output for Simulated Annealing (*ctrl*.int and profile intensities (*.spr).

```
OPTIMIZE   Powder_Profile 1.0

Or

OPTIMIZE   Powder_WProfile 1.0 dis-restr 1.0
```

In the FullProf distribution there is a subdirectory of the Examples directory containing files ready to run and test **GLOpSAnn**.