

Manual for the mini-workshop

Information

SSH/STM4 softwares are installed in D:\ drive, you may find shortcut in D:\. To use STM4:

- double click on go.bat/STM4.bat icon.

To login the HPC server, using SSH/Secured Shell client and following information:

- Host: 202.117.249.10
- user name: USPEX
- Password: 123abc

W0. Brief tutorial on data and their analysis.

The subdirectory **resultsX** contains many files and folders:

The most important files are:

- **Individuals** - gives details of all produced structures (energies, unit cell volumes, space groups, variation operators that were used to produce this structure, k-points mesh that was used to compute its final energy, degrees of order, *etc.*).
- **BESTIndividuals** gives this information for best structures of each generation.
- **extended_convex_hull** - only for variable-composition calculations, where it gives all thermodynamically stable compositions, their enthalpies (per atom).
- **gatheredPOSCARS, fitness.dat** - relaxed structures, their fitnesses (may or may not be identical to enthalpies, depending on what you ask the code to optimize!).
- **BESTgatheredPOSCARS, BESTfitness** - the same data for the best structure in each generation.
- **OUTPUT.txt** - summarizes input variables, structures produced by USPEX, and their characteristics.
- **Parameters.txt** - this is a copy of the **INPUT.txt** file used in this calculation, for your reference.
- **hardness.dat** - an estimate of hardness of all structures.
- graphical files (***.TIFF**) for rapid analysis of results.
- **goodStructures_POSCARS** is a file, made available in v.9 of USPEX. It's extremely convenient for analysis, as it reports all the different structures in order of decreasing stability, starting from the most stable structure and ending with the least stable one. Analyzing this file is particularly convenient for those people who do not have STM4.
- **symmetrized_structures.cif** - all structures that are symmetrized using Stokes program.

The "less" important files/folders are:

- **origin** - shows which structures originated from which parents and through which variation operators.

- **non_optimized_structures** - gives all produced structures before relaxation.
- **enthalpies_complete** - here enthalpies for all structures in each stage of relaxation are given.
- **quasiEntropy.dat** - shows the diversity of structures in each generation.
- **gatheredPOSCARS_order** gives of the same information as **gatheredPOSCARS** , and in addition for each atom it gives the value of the order parameter.
- a lot of other files.
- Folder **AuxiliaryFiles** gives some of the old files - **VOLUMES**, **compositions_nospace.dat**, **enthalpies_nospace.dat** , etc.

W1. High-pressure prediction with fixed composition.

Step1: Preparing the calculation

- In your own folder your_name, make a new directory work_1(chosen by your own);
- `cd ~/your_name/work_1;`
- copy the USPEX code to your directory: `cp -r ~/USPEX/Code/* ./;`
- copy the case-specific files: `cp -r ~/USPEX/Test/T1_Physics_GULP/* ./;`

Step2: Browse the INPUT.txt file and files in folder Specific to double check everything is OK

Note: In the INPUT.txt file, a line with a ":" contains one parameter; lines between two "%" -lines is a data block for complex parameters.

- The parameters in the section "TYPE OF RUN AND SYSTEM" are "must-have" parameters, which determines the case-specific information.
- The parameters in the section "POPULATION", "SURVIVAL OF THE FITTEST AND SELECTION" and "VARIATION OPERATORS" are the key parameters for the evolutionary algorithm. Tune with care to achieve better results.
- The parameters in the sections "CONSTRAINTS" and "CELL" are useful for enhancement of calculation efficiency.
- The section "DETAILS OF AB INITIO CALCULATIONS" determines how the calculation will be done. Keep in mind the calculation workflow and choose proper settings.

Important: The words right after ":" or "%" are the keywords and must stay as it is (or as from the manual), because USPEX look for the words to read the value. The words must also stay at least one space away from their neighbors.

By looking at the line "3 3 3 3" after "abinitioCode" in the INPUT.txt, one know that GULP code was used to do all five stages of calculation. Thus in Specific folder one must prepare five input files for GULP, which are `ginput_{1..5}` and `goptions_{1..5}`.

Note: About pressure. System pressure is an internal parameter for energy calculation code. For GULP calculation, it is set in `ginput_{1..5}` files with default unit GPa. **New:** In Release >10.0, USPES

takes a new parameter in INPUT.txt: ExternalPressure (in unit of GPa):

20 : ExternalPressure

Step 3: Submit calculation

- Open the file `job: gedit job;`
- modify the last string "USPEXrun" in the line "#PBS -N USPEXrun", replace it with `W1_(your_initial)`, save it;
- submit the job by issuing command: `qsub job;`
- you may monitor your jobs by `qstat`, your job will disappear from the `qstat` lines when it is finished;
- **Calculation time: >12 hrs !!**

Step 4: check the results

- `cd results1;`
- `tail OUTPUT.txt` to see last line with "Job Finished at";
- check the graph files by `eog` command `eog Energy_vs_N.tif;`
- check text file by `less` or `gedit` command.

W2. High-pressure structure prediction with variable composition.

Step1: Preparing the calculation

- In your own folder `your_name`, make a new directory `work_2`(chosen by your own);
- `cd ~/your_name/work_2;`
- copy the USPEX code to your directory: `cp -r ~/USPEX/Code/* ./;`
- copy the case-specific files: `cp -r ~/USPEX/Test/T7_varcomp_GULP/* ./;`

Step 2: check the INPUT.txt and Specific folder

In the INPUT.txt file, focus on (they are different from fixed-composition calculation):

- 301 : `calculationType`
- % `numSpecices`
- `firstGeneMax`
- `minAt` and `maxAt`
- `fracTrans`
- `Latticevalues`

Step 3: Submit calculation

- Open the file job: `gedit job`;
- modify the last string "USPEXrun" in the line "#PBS -N USPEXrun", replacing it with `W2_(your_initial)`, and save it;
- submit the job by issuing command: `qsub job`;
- you may monitor your jobs by `qstat`, your job will disappear from the `qstat` lines when it is finished;
- **Calculation Time: 4:06 hrs**

Step 4: check the results

- `cd results1`;
- `tail OUTPUT.txt` to see last line with "Job Finished at ...";
- check the graph files by `eog` command, in this case, particularly, `eog extendedConvexHull.tif`;
- check text file by `less` or `gedit` command.

Tutorial: Explanation on convex hull of enthalpy.

W3. Fixed-composition optimization of hardness.

Step1: Preparing the calculation

- In your own folder `your_name`, make a new directory `work_3`(chosen by your own);
- `cd ~/your_name/work_3`;
- copy the USPEX code to your directory: `cp -r ~/USPEX/Code/* ./`;
- copy the case-specific files : `cp -r ~/USPEX/Test/T6_hardness_GULP/* ./`;

Step 2: check the INPUT.txt and Specific folder

In the INPUT.txt file, focus on (they are different from energy calculation):

- `optType`
- `% valences`
- `mutationRate`
- `DisplaceInLatmutation`

Step 3: Submit calculation

- Open the file `job: gedit job;`
- modify the last string "USPEXrun" in the line "#PBS -N USPEXrun", replacing it with `W3_(your_initial)`, and save it;
- submit the job by issuing command: `qsub job;`
- you may monitor your jobs by `qstat`, your job will disappear from the `qstat` lines when it is finished;
- **Calculation time: 28 mins**

Step 4: check the results

- `cd results1;`
- `tail OUTPUT.txt` to see last line with "Job Finished at";
- check `hardness.txt` for the hardness for all structures;
- check the graph files by `eog` command, in this case, particularly, `eog Fitness_vs_N.tif;`
- check text file by `less` or `gedit` command.

Tutorial: Explanation on Minimization/Maximization of fitness function/hardness.

W4. Zero-pressure variable-composition optimization of the dielectric constant with GULP.

Step1: Preparing the calculation

- In your own folder `your_name`, make a new directory `work_4`(chosen by your own);
- `cd ~/your_name/work_4;`
- copy all the files from Test folder: `cp -r ~/USPEX/Test/T22_diel_GULP/* ./;`

Step 2: check the INPUT.txt and Specific folder

In the INPUT.txt file, focus on (they are different from energy calculation):

```
301 : calculationType (dimension: 0-3; molecule: 0/1; varcomp: 0/1)
diel_sus : optType (1=enthalpy, 2=volume, 3=hardness, 4=struc_order,
5=aver_dist)
```

```
% numSpecices
1 0 1
0 2 3
% EndNumSpecices
%%%%%%%%%
```

```
% Here come the atomic numbers of the atoms involved
% atomType
Ca Al 0
% EndAtomType
```

abinitioCode (which code from CommandExecutable shall be used for calculation?)

```
3 3 3 3 3
```

```
ENDabinit
```

The first four steps will do structure relaxation, and the last step will calculate dielectric constants.

goptions_4:

```
hideshells unit unfix opti conjugate comp
```

goptions_5:

```
hideshells unit unfix opti prop conjugate comp
```

Step 3: Submit calculation

Open the file job: gedit job; modify the last string "USPEXrun" in the line "#PBS -N USPEXrun", replacing it with W3_(your_initial), and save it; submit the job by issuing command: qsub job; you may monitor your jobs by qstat, your job will disappear from the qstat lines when it is finished; **Calculation time: ~3 hrs**

Step 4: check the results

cd results1; tail OUTPUT.txt to see last line with "Job Finished at"; check the graph files by eog command, in this case, particularly, eog extendedConvexHull.tif; check text file by less or gedit command.

W5. Molecular crystal structure prediction.

Step1: Preparing the calculation

- In your own folder your_name, make a new directory work_5(chosen by your own);
- cd ~/your_name/work_5;
- copy the USPEX code to your directory: cp -r ~/USPEX/Code/* ./;
- copy the case-specific files: cp -r ~/USPEX/Test/T11_molecule_DMACRYS/* ./;
- MOL_1 file

The **MOL_x** file describes the structure of a molecule to be used as a whole entity, and defines which degrees of freedom will be frozen during structure relaxation. This file and its format differ from SIESTA's Z_Matrix file (**MOL_x**) gives Cartesian coordinates of the atoms, whereas Z_Matrix file

defines atomic positions from bond lengths, bond angles and torsion angles). Z_Matrix file is created using the information given in the **MOL_x** file, {\it i.e.} bond lengths and all necessary angles are calculated from the Cartesian coordinates. Which lengths and angles are important and should be used for Z_Matrix to define atomic positions - this is exactly what columns 5-7 specify. Let's look at the **MOL_1** file for methane CH₄:

```
CH4
Number of atoms: 5
C   0.00000  0.00000  0.00000  0 0 0  1 1
H   0.63512  0.63512 -0.63512  1 0 0  2 1
H  -0.63512 -0.63512 -0.63512  1 2 0  3 1
H   0.63512 -0.63512  0.63512  1 2 3  4 1
H  -0.63512  0.63512  0.63512  1 2 3  5 1
```

The first atom is Carbon, its coordinates are defined without reference to other atoms (`0 0 0').

The second atom is Hydrogen, its coordinates (in molecular coordinate frame) in Z_matrix will be set only by its distance from the first atom, (i.e. Carbon described above), but no angles (`1 0 0').

The third atom is Hydrogen, its coordinates will be set by its distance from the first atom, and the bond angle 1-2-1, but no torsion angle - hence we write `1 2 0'.

The fourth atom is Hydrogen, its coordinates will be set by its distance from the first atom, bond angle 4-1-2, and torsion angle 4-1-2-3 - hence, we write `1 2 3' and so forth... until we reach the final, 5th atom, which is Hydrogen, defined by its distance from the 1st atom (Carbon), valence angle 5-1-2 and torsion angle 5-1-2-3 - hence `1-2-3'.

The final column corresponds to optimization flag of the torsion angle. Ideally, this flag for should be 1 the first three atoms, and 0 for the others. If there exist any other flexible torsion angle, put 1 there.

The 8th column corresponds to the atom number, which is required by DMACRYS.

- MOL_1 (VASP)

T12_molecule_VASP

```
CH4
Number of atoms: 5
C   0.00000  0.00000  0.00000  0 0 0  1
H   0.63512  0.63512 -0.63512  1 0 0  1
H  -0.63512 -0.63512 -0.63512  1 2 0  1
H   0.63512 -0.63512  0.63512  1 2 3  0
H  -0.63512  0.63512  0.63512  1 2 3  0
```

Step 2: check the INPUT.txt and Specific folder

In the INPUT.txt file, focus on (they are different from bulk calculation):

```
310 : calculationType
*****
*****
*           VARIATION OPERATORS           *
*****
*****
```

```
0.50 : fracGene (fraction of generation produced by heredity)
0.20 : fracRand (fraction of generation produced randomly from space groups)
0.20 : fracAtomsMut (fraction of the generation produced by softmutation)
0.10 : fracRotMut (fraction of the generation produced by rotational
mutation, only for molecules)
```

Step 3: Submit calculation

- Open the file job: `gedit job`;
- modify the last string "USPEXrun" in the line "#PBS -N USPEXrun", replacing it with W5_(your_initial), and save it;
- submit the job by issuing command: `qsub job`;
- you may monitor your jobs by `qstat`, your job will disappear from the `qstat` lines when it is finished;
- **Calculation Time: ~6 hrs**

Step 4: check the results

- `cd results1`;
- `tail OUTPUT.txt` to see last line with "Job Finished at";
- check the graph files by `eog` command `eog Energy_vs_N.tif`;
- check text file by `less` or `gedit` command.

W6. Evolutionary metadynamics with GULP.

T16_metadynamics_GULP

1. Introduction

Evolutionary metadynamics is a very powerful method for finding the global minimum, and many low-energy metastable structures that are accessible from the starting structure. It even can be used for finding possible phase transition pathways. Evolutionary metadynamics is a hybrid of Martonak's metadynamics and Oganov-Glass evolutionary approach. To do an efficient search, the starting structure has to be high-quality and is given in the file POSCAR_1. Currently, evolutionary metadynamics is enabled only with the VASP and GULP codes at the moment.

2. Prepare files

To switch to the evolutionary metadynamics mode, you have to:

I. Specify

META : calculationMethod

1 : calculationType

II. Create file **POSCAR_1** in your folder.

III. Specify the population size (in this case, this is the number of softmutations at each metastep):

30 : populationSize

IV. Specify the following metadynamics-only options:

variable **ExternalPressure**

meaning: the pressure at which you want to perform the calculation, in kbars.

Default: no default

% ExternalPressure

2000.00

% EndPressure

GaussianWidth: the width of each of the Gaussians added to the energy surface to accelerate phase transitions. A good rule of thumb is to choose it close to $0.1-0.15L$, where L is the minimum length of the unit cell, in Angstroms.

Default: $0.15L$ (Å)

format:

0.60 : GaussianWidth

GaussianHeight: the height of each of the Gaussians added to the energy surface to accelerate phase transitions. A good rule of thumb is to choose it close to $L(\Delta h)^2 G$, where L is the average length of the unit cell in Angstroms, Δh is the Gaussian width in Angstroms (see below), and G is the shear modulus in kbars.

Default: $1000*(0.15L)^2*L = 22.5L^3$ (Å³kbar)

4000 : GaussianHeight

variable: **FullRelax**

meaning: metadynamics as such only relaxes structures at fixed cell. For analysis, you need to perform complete structure relaxation (with cell relaxation).

When FullRelax=0, no full relaxation will be performed (very fast option, but inconvenient for analysis of results).

When FullRelax=1, only the best structure of the generation will be fully relaxed (also fast, often sufficient).

When FullRelax=2, all inequivalent structures are fully relaxed (still fast, only ~ 2 times slower than =1, but brings a lot more insight. Strongly recommended for most cases).

Default: 2

2 : FullRelax

For full relaxation, in case of evolutionary metadynamics there is a slightly different format of the block abinitioCode, for example:

abinitioCode (which code shall be used for calculation? Up to now: vasp, siesta, gulp)

3 3 3 (3)

ENDabinit

In the example above, there are three stages of relaxation with fixed cell, and one stage of full

relaxation (in parentheses). Remember that in the last fixed-cell stage of relaxation, pressure tensor must be accurate – this is what drives metadynamics.

3. Submit job

In INPUT.txt, set:

```
% commandExecutable
~/bin/gulp <input >output
% EndExecutable

1      : numParallelCalcs (how many parallel calculations shall be
performed)\
0      : whichCluster (0: no-job-script, 1: local submission, 2: remote
submission)
```

Then, in your task folder vi job, it looks like:

```
#!/bin/sh
#####
#PBS -N T16_yourname
#PBS -l nodes=1:ppn=1
#PBS -l walltime=2:00:00
#PBS -q batch
#PBS -V
#PBS -S /bin/bash
#####
EXEC=~/.opt/matlab/bin/matlab
cd '$'{PBS_0_WORKDIR}
'$'EXEC <USPEX.m > log
}
then
qsub job
```

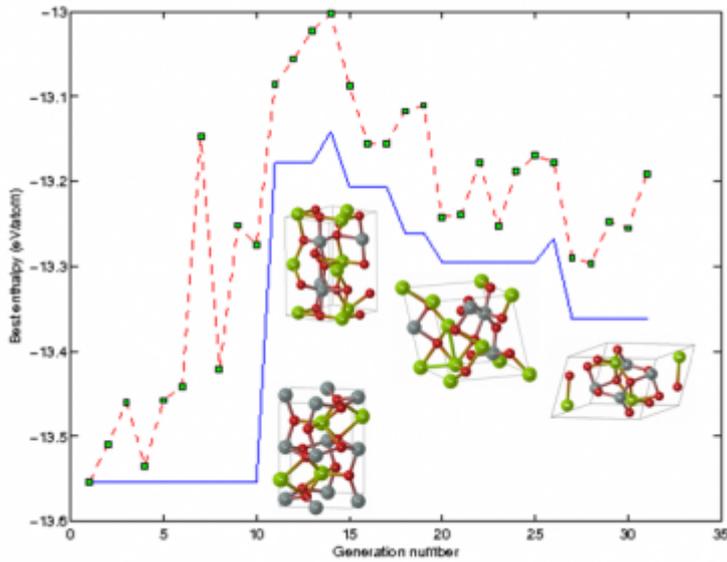
4. Download files

After run metadynamics, additional files will be found in **/results1** or **/results1/AuxiliaryFiles**, most importantly:

- **present** – pressure tensor
- **lattice.dat** – cell shape change during the simulation
- **BESTenthalpies** and **BESTenthalpies_relaxed** – enthalpies of the best structures at each generation at fixed cell and after full relaxation, respectively
- **BESTgatheredPOSCARS** and **BESTgatheredPOSCARS_relaxed** – the best structures at fixed cell and after full relaxation, respectively.
- **BestEnthalpy.tif** – Evolutionary curves of the enthalpies of the BESTgatheredPOSCARS before and

after relaxation

****5. Results and discussion****



W7. problems that participants will bring to be solved.

From:

<http://www.mse-project.org/wiki/> - **Materials Genome Project**

Permanent link:

http://www.mse-project.org/wiki/doku.php?id=uspex_training_course_in_xi_an

Last update: **2014/01/07 00:38**

