

Hands-On Session: DP-GEN

Yongbin Zhuang

Xiamen University

Machine Learning workshop, Xiamen 2019.09.05

Other tutors: Linfeng Zhang, Yuzhi Zhang, Yunpei Liu

First of all

Update deepmd-tutorial file to Huaweicloud

```
scp deepmd_tutorial_file_path username@ip:~/
```

Unzip the file:

```
tar -zxvf deepmd_tutorial_file
```

Using dpgen with docker

Fist of all, start up the container by command:

```
docker-compose up -d
```

Let's start the DP-GEN outside the container by:

```
docker-compose exec dpgen bash -c "cd /public/data/deepmd-tutorial/cp2k_dpgen && dpgen run param.json machine_final_all.json"
```

Or, directly using the prepared bash script

```
bash exec.sh
```

Use the following command to check the status:

```
nvidia-smi
```

Start Docker

Support the execution of command **outside** or **inside** the container

Fist of all, start up the container by command:

```
docker-compose up -d
```

Let's inside one of container by:

```
docker-compose exec dpgen bash
```

If you success, you will see something like this:

```
root@797b03b3cb31:/# |
```

Try dpgen command like: You will see the help message of dpgen

```
dpgen -h
```

Start Docker

Support the execution of command **outside** or **inside** the container

Fist of all, start up the container by command:

```
docker-compose up -d
```

Let's execute command outside the container by:

```
docker-compose exec dpgen bash -c "dpgen -h"
```

You will see the help message of dpgen when outside the container

To Shut down the container, use command:

```
docker-compose down
```

Machine Learning Potential training

Training Process:

1. first principles data + Machine learning tool
2. test
3. add more data + Machine learning

Problem:

Most data are redundant, if in the field of training set.

One solution:

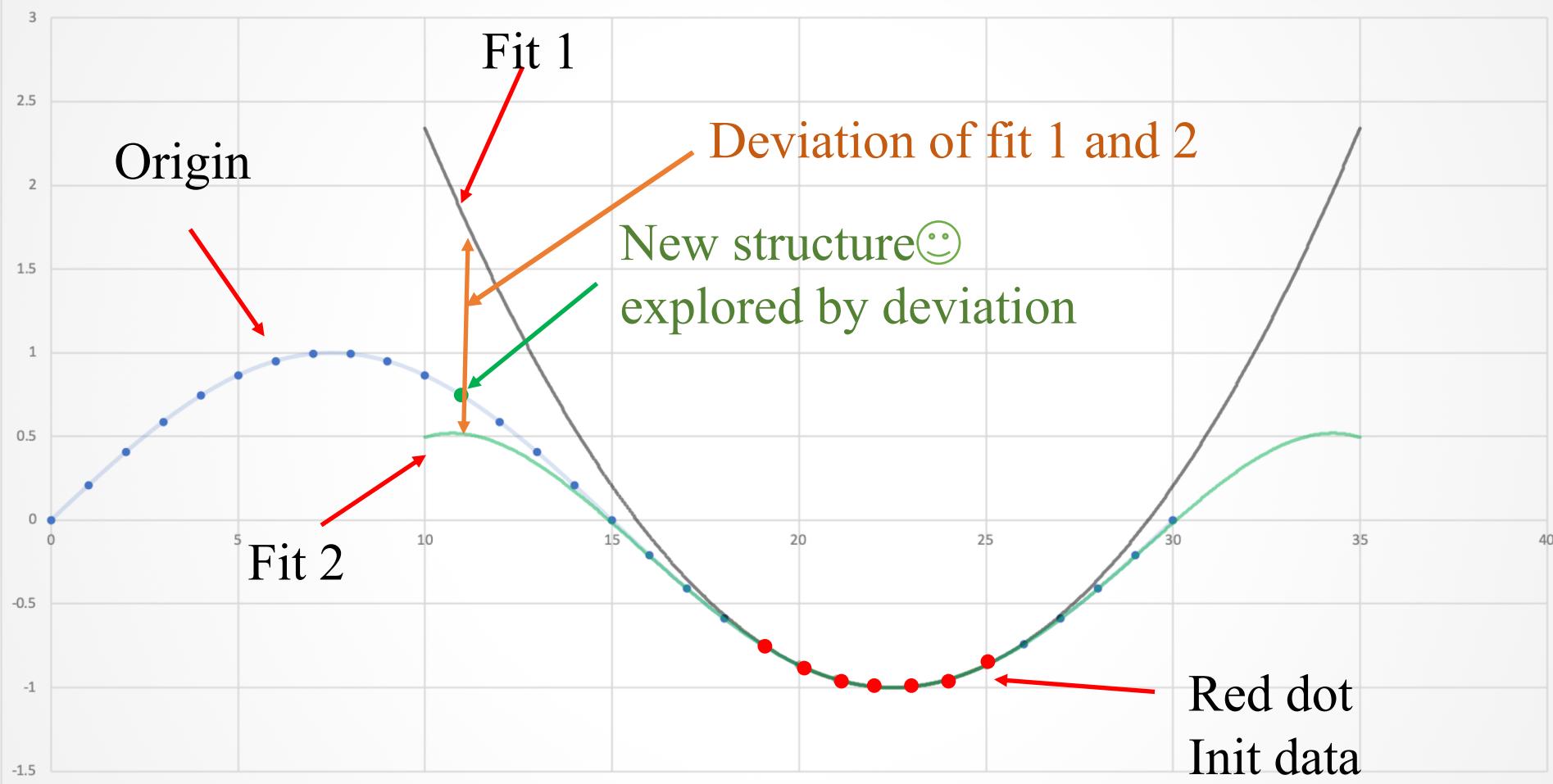
Active learning: search new data point efficiently

Functions of DP-GEN

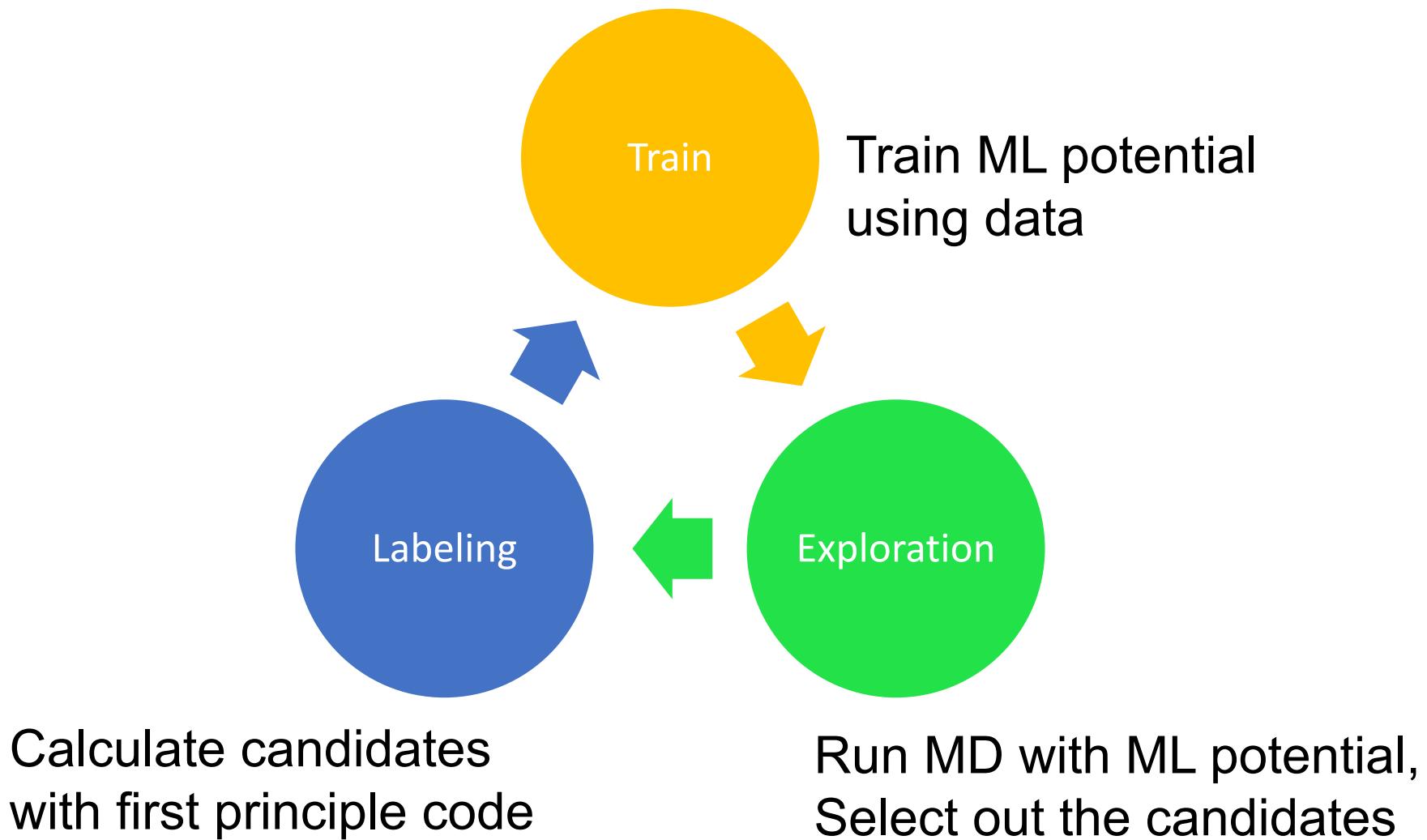
DP-GEN contains three functions:

- Init: prepare the initial data
- Run: start the active learning process
- Test: do the auto test for ML potential

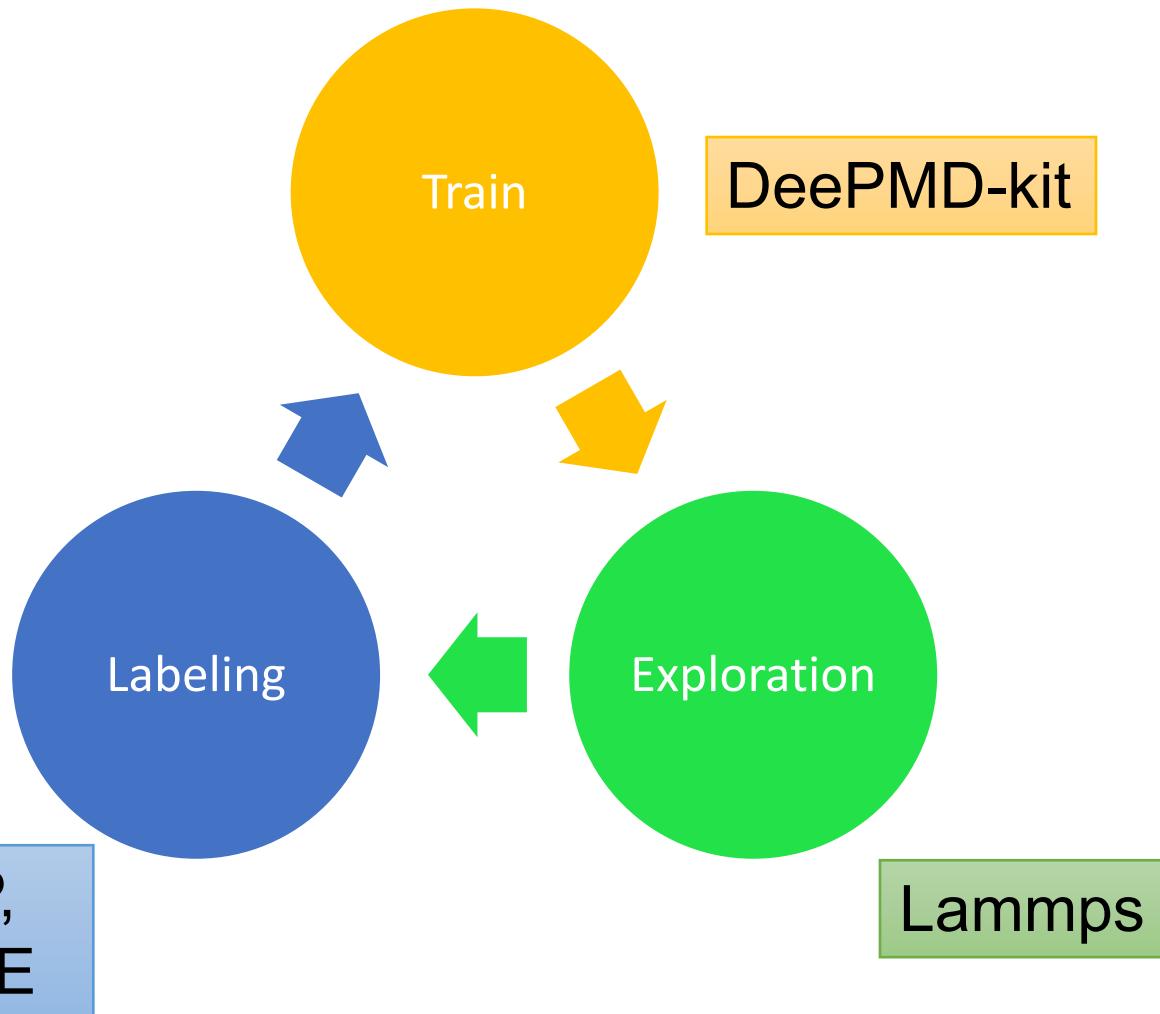
Illustration of the principle of active learning



The **run** function of DP-GEN

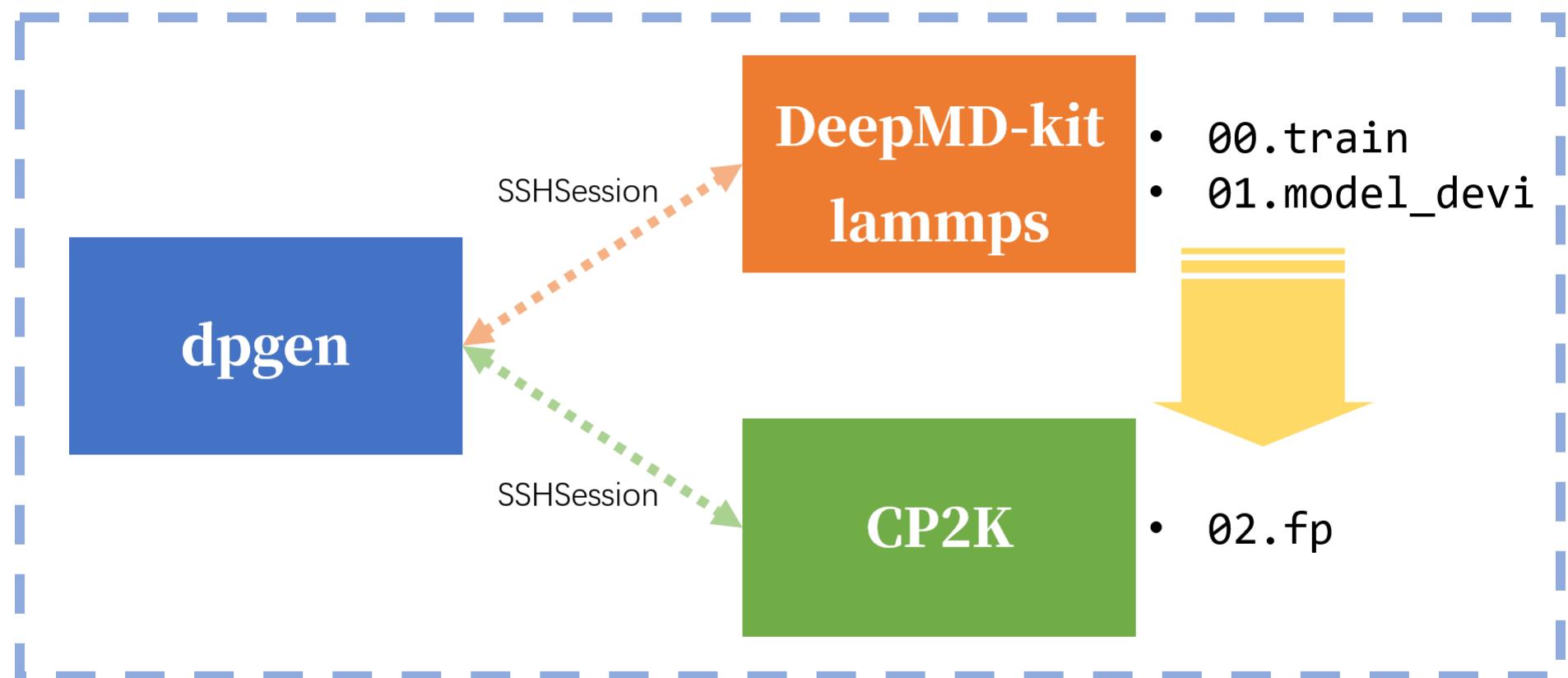


The software used by DP-GEN



About Docker

N Docker Containers = N virtual Machines



Three Containers with different installed software

General structure of feed in files

Param.json

1. Basic parameter

2. param for DeePMD-kit

3. param for Lammmps MD

4. param for FP code

Machine.json

Manage computer
resources and setting for

1. DeePMD-kit

2. Lammmps

3. Fp code (cp2k)

Param.json: basic setting

1. Include the types of elements and the mass of elements,
2. The order of properties should be consistent with that of elements

```
"type_map": [  
    "H",  
    "C"  
],  
"mass_map": [  
    1,  
    12  
,
```

Param.json: training

1. Train an ensemble of models (default 4) by DeePMD-kit, with different initial parameter from random seed.
2. Obtain the graph. 00*.pb in 00.train folder.
3. The number of batch size multiplied with atom number should larger than 32.

Param.json: model_devi.jobs

1. Specify the basic setting for MD, especially temp
2. The index of iteration is the index of devi.jobs
3. The initial structure for MD is from sys_configs

```
"model_devi_jobs": [  
    {  
        "sys_idx": [  
            0  
        ],  
        "temps": [  
            100  
        ],  
        "press": [  
            1  
        ],  
        "trj_freq": 10,  
        "nsteps": 500,  
        "ensemble": "nvt",  
        "_idx": "00"  
    },
```

Param.json: force deviation

```
"model_devi_f_trust_lo": 0.05,  
"model_devi_f_trust_hi": 0.15,  
"model_devi_f_max": 0.25,
```

$$\epsilon = \max_i \sqrt{\langle \|f_i - \langle f_i \rangle\|^2 \rangle}$$

step	max_devi_e	min_devi_e	avg_devi_e	max_devi_f	min_devi_f
0	3.094068e+00	7.583044e-01	1.244792e+00	2.033439e-02	1.111021e-02
10	3.114559e+00	7.641919e-01	1.252722e+00	6.620141e-02	3.026825e-02
20	3.138356e+00	7.704017e-01	1.261834e+00	4.945633e-02	2.293204e-02
30	3.154803e+00	7.738079e-01	1.267152e+00	2.693371e-02	1.194895e-02
40	3.166201e+00	7.740074e-01	1.271595e+00	6.375671e-02	1.946995e-02
50	3.155127e+00	7.748555e-01	1.266299e+00	3.537826e-02	1.286718e-02
60	3.133582e+00	7.738328e-01	1.256576e+00	3.413587e-02	2.526567e-02
70	3.106401e+00	7.669162e-01	1.244871e+00	6.479719e-02	1.785133e-02
80	3.084239e+00	7.592567e-01	1.234835e+00	3.676270e-02	2.189086e-02
90	3.064068e+00	7.516044e-01	1.224792e+00	2.033439e-02	1.111021e-02
100	3.044895e+00	7.439521e-01	1.214722e+00	6.620141e-02	3.026825e-02
110	3.025722e+00	7.363098e-01	1.204652e+00	4.945633e-02	2.293204e-02
120	3.006559e+00	7.286675e-01	1.194582e+00	2.693371e-02	1.194895e-02
130	2.987396e+00	7.210252e-01	1.184512e+00	6.375671e-02	1.946995e-02
140	2.968233e+00	7.133829e-01	1.174442e+00	3.537826e-02	1.286718e-02
150	2.949070e+00	7.057406e-01	1.164372e+00	3.413587e-02	2.526567e-02
160	2.929907e+00	6.980983e-01	1.154302e+00	6.479719e-02	1.785133e-02
170	2.910744e+00	6.904560e-01	1.144232e+00	3.676270e-02	2.189086e-02
180	2.891581e+00	6.828137e-01	1.134162e+00	2.033439e-02	1.111021e-02
190	2.872418e+00	6.751714e-01	1.124092e+00	6.620141e-02	3.026825e-02
200	2.853255e+00	6.675291e-01	1.114022e+00	4.945633e-02	2.293204e-02
210	2.834092e+00	6.608868e-01	1.103952e+00	2.693371e-02	1.194895e-02
220	2.814929e+00	6.532445e-01	1.093882e+00	6.375671e-02	1.946995e-02
230	2.795766e+00	6.456022e-01	1.083812e+00	3.537826e-02	1.286718e-02
240	2.776603e+00	6.379599e-01	1.073742e+00	3.413587e-02	2.526567e-02
250	2.757440e+00	6.303176e-01	1.063672e+00	6.479719e-02	1.785133e-02
260	2.738277e+00	6.226753e-01	1.053602e+00	3.676270e-02	2.189086e-02
270	2.719114e+00	6.150330e-01	1.043532e+00	2.033439e-02	1.111021e-02
280	2.699951e+00	6.073907e-01	1.033462e+00	6.620141e-02	3.026825e-02
290	2.680788e+00	6.007484e-01	1.023392e+00	4.945633e-02	2.293204e-02
300	2.661625e+00	5.931061e-01	1.013322e+00	2.693371e-02	1.194895e-02
310	2.642462e+00	5.854638e-01	1.003252e+00	6.375671e-02	1.946995e-02
320	2.623299e+00	5.778215e-01	9.93182e+00	3.537826e-02	1.286718e-02
330	2.604136e+00	5.701792e-01	9.83112e+00	3.413587e-02	2.526567e-02
340	2.584973e+00	5.625369e-01	9.73042e+00	6.479719e-02	1.785133e-02
350	2.565810e+00	5.548946e-01	9.63072e+00	3.676270e-02	2.189086e-02
360	2.546647e+00	5.472523e-01	9.53002e+00	2.033439e-02	1.111021e-02
370	2.527484e+00	5.396100e-01	9.43032e+00	6.620141e-02	3.026825e-02
380	2.508321e+00	5.319677e-01	9.33062e+00	4.945633e-02	2.293204e-02
390	2.489158e+00	5.243254e-01	9.23092e+00	2.693371e-02	1.194895e-02
400	2.470000e+00	5.166831e-01	9.13022e+00	6.375671e-02	1.946995e-02
410	2.450837e+00	5.090408e-01	9.03052e+00	3.537826e-02	1.286718e-02
420	2.431674e+00	5.013985e-01	8.93082e+00	3.413587e-02	2.526567e-02
430	2.412511e+00	4.937562e-01	8.83112e+00	6.479719e-02	1.785133e-02
440	2.393348e+00	4.861139e-01	8.73142e+00	3.676270e-02	2.189086e-02
450	2.374185e+00	4.784716e-01	8.63172e+00	2.033439e-02	1.111021e-02
460	2.355022e+00	4.708293e-01	8.53202e+00	6.620141e-02	3.026825e-02
470	2.335859e+00	4.631870e-01	8.43232e+00	4.945633e-02	2.293204e-02
480	2.316696e+00	4.555447e-01	8.33262e+00	2.693371e-02	1.194895e-02
490	2.297533e+00	4.479024e-01	8.23292e+00	6.375671e-02	1.946995e-02
500	2.278370e+00	4.402599e-01	8.13322e+00	3.537826e-02	1.286718e-02
510	2.259207e+00	4.326176e-01	8.03352e+00	3.413587e-02	2.526567e-02
520	2.240044e+00	4.249753e-01	7.93382e+00	6.479719e-02	1.785133e-02
530	2.220881e+00	4.173330e-01	7.83412e+00	3.676270e-02	2.189086e-02
540	2.201718e+00	4.096907e-01	7.73442e+00	2.033439e-02	1.111021e-02
550	2.182555e+00	4.020484e-01	7.63472e+00	6.620141e-02	3.026825e-02
560	2.163392e+00	3.944061e-01	7.53502e+00	4.945633e-02	2.293204e-02
570	2.144229e+00	3.867638e-01	7.43532e+00	2.693371e-02	1.194895e-02
580	2.125066e+00	3.791215e-01	7.33562e+00	6.375671e-02	1.946995e-02
590	2.105903e+00	3.714792e-01	7.23592e+00	3.537826e-02	1.286718e-02
600	2.086740e+00	3.638369e-01	7.13622e+00	3.413587e-02	2.526567e-02
610	2.067577e+00	3.561946e-01	7.03652e+00	6.479719e-02	1.785133e-02
620	2.048414e+00	3.485523e-01	6.93682e+00	3.676270e-02	2.189086e-02
630	2.029251e+00	3.409100e-01	6.83712e+00	2.033439e-02	1.111021e-02
640	2.010088e+00	3.332677e-01	6.73742e+00	6.620141e-02	3.026825e-02
650	1.990925e+00	3.256254e-01	6.63772e+00	4.945633e-02	2.293204e-02
660	1.971762e+00	3.180831e-01	6.53802e+00	2.693371e-02	1.194895e-02
670	1.952599e+00	3.104408e-01	6.43832e+00	6.375671e-02	1.946995e-02
680	1.933436e+00	3.027985e-01	6.33862e+00	3.537826e-02	1.286718e-02
690	1.914273e+00	2.951562e-01	6.23892e+00	3.413587e-02	2.526567e-02
700	1.895110e+00	2.875139e-01	6.13922e+00	6.479719e-02	1.785133e-02
710	1.875947e+00	2.808716e-01	6.03952e+00	3.676270e-02	2.189086e-02
720	1.856784e+00	2.732293e-01	5.93982e+00	2.033439e-02	1.111021e-02
730	1.837621e+00	2.655870e-01	5.83912e+00	6.620141e-02	3.026825e-02
740	1.818458e+00	2.579447e-01	5.73942e+00	4.945633e-02	2.293204e-02
750	1.799295e+00	2.503024e-01	5.63972e+00	2.693371e-02	1.194895e-02
760	1.780132e+00	2.426601e-01	5.53902e+00	6.375671e-02	1.946995e-02
770	1.760969e+00	2.350178e-01	5.43932e+00	3.537826e-02	1.286718e-02
780	1.741806e+00	2.273755e-01	5.33962e+00	3.413587e-02	2.526567e-02
790	1.722643e+00	2.207332e-01	5.23992e+00	6.479719e-02	1.785133e-02
800	1.703480e+00	2.130909e-01	5.13922e+00	3.676270e-02	2.189086e-02
810	1.684317e+00	2.054486e-01	5.03952e+00	2.033439e-02	1.111021e-02
820	1.665154e+00	1.978063e-01	4.93982e+00	6.620141e-02	3.026825e-02
830	1.645991e+00	1.901640e-01	4.83912e+00	4.945633e-02	2.293204e-02
840	1.626828e+00	1.825217e-01	4.73942e+00	2.693371e-02	1.194895e-02
850	1.607665e+00	1.748794e-01	4.63972e+00	6.375671e-02	1.946995e-02
860	1.588502e+00	1.672371e-01	4.53902e+00	3.537826e-02	1.286718e-02
870	1.569339e+00	1.595948e-01	4.43932e+00	3.413587e-02	2.526567e-02
880	1.550176e+00	1.519525e-01	4.33962e+00	6.479719e-02	1.785133e-02
890	1.531013e+00	1.443102e-01	4.23992e+00	3.676270e-02	2.189086e-02
900	1.511850e+00	1.366679e-01	4.13922e+00	2.033439e-02	1.111021e-02
910	1.492687e+00	1.290256e-01	4.03952e+00	6.620141e-02	3.026825e-02
920	1.473524e+00	1.213833e-01	3.93982e+00	4.945633e-02	2.293204e-02
930	1.454361e+00	1.137440e-01	3.83912e+00	2.693371e-02	1.194895e-02
940	1.435198e+00	1.061017e-01	3.73942e+00	6.375671e-02	1.946995e-02
950	1.416035e+00	984594e-02	3.63972e+00	3.537826e-02	1.286718e-02
960	1.396872e+00	908171e-02	3.53902e+00	3.413587e-02	2.526567e-02
970	1.377709e+00	831748e-02	3.43932e+00	6.479719e-02	1.785133e-02
980	1.358546e+00	755325e-02	3.33962e+00	3.676270e-02	2.189086e-02
990	1.339383e+00	678902e-02	3.23992e+00	2.033439e-02	1.111021e-02
1000	1.320220e+00	602479e-02	3.13922e+00	6.620141e-02	3.026825e-02

Param.json: list of candidates

```
iter.000000/01.model_devi/task.000.000007 300
iter.000000/01.model_devi/task.000.000005 370
iter.000000/01.model_devi/task.000.000004 380
iter.000000/01.model_devi/task.000.000007 290
iter.000000/01.model_devi/task.000.000004 110
iter.000000/01.model_devi/task.000.000007 210
iter.000000/01.model_devi/task.000.000003 320
iter.000000/01.model_devi/task.000.000004 200
iter.000000/01.model_devi/task.000.000000 70
iter.000000/01.model_devi/task.000.000005 120
iter.000000/01.model_devi/task.000.000001 410
iter.000000/01.model_devi/task.000.000008 240
iter.000000/01.model_devi/task.000.000008 440
iter.000000/01.model_devi/task.000.000008 300
iter.000000/01.model_devi/task.000.000005 130
iter.000000/01.model_devi/task.000.000000 130
```

In the folder of **02.fp/ candidate.shuffled.000.out**

Param.json: setting for fp task

```
"fp_style": "cp2k",
"shuffle_poscar": false,
"fp_task_max": 20, ← Task number for fp
"fp_task_min": 5,
"fp_pp_path": ".",
"fp_pp_files": [],
"fp_params": {
  "cutoff": "400",
  "rel_cutoff": "50",
  "functional": "PBE",
  "pair_potential_path": "dftd3.dat",
  "pair_ref_functional": "PBE",
  "basis_path": "BASIS_MOLOPT",
  "pp_path": "GTH_POTENTIALS",
  "element_list": ["H", "C"],
  "basis_list": ["DZVP-MOLOPT-GTH", "DZVP-MOLOPT-GTH"],
  "pp_list": ["GTH-PBE-q1", "GTH-PBE-q4"]}
```

setting for cp2k

Record for dpgen

Index of Iteration	Stage in Iteration	Meaning
0	0	Make train
0	1	Run train
0	2	Post train
0	3	Make model devi
0	4	Run model devi
0	5	Post model devi
0	6	Make fp
0	7	Run fp
0	8	Post fp

Machine.json

- manage resources for
 - DeePMD-kit,
 - Lammmps
 - cp2k.
- Support job manager: Slurm, PBS, Local, cloud machine(amazon etc.) ..
- Automatic and general
- Robust for abnormal jobs

Acknowledgements

Yunpei Liu(Xiamen Univ.): Set up docker and test DP-GEN with me

Linfeng Zhang(Princeton Univ.): main developer, technical support

Han Wang(IAPCM): main developer, technical support

Yuzhi Zhang(Peking Univ.): main developer, technical support

Thank you
for
your attention